Characterisation of a novel nested encoding Laf4ir gene in the cardiovascular system

Ehteramyan, Mazdak

Awarding institution:
King’s College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to:
 Share: to copy, distribute and transmit the work

Under the following conditions:
 Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
 Non Commercial: You may not use this work for commercial purposes.
 No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Characterisation of a novel nested encoding \textit{Laf4ir} gene in the cardiovascular system

Mr Mazdak Ehteramyan
This thesis is submitted for the degree of Doctor of Philosophy (PhD)

Primary Supervisor: Dr Lingfang Zeng
Secondary Supervisor: Dr Yanhua Hu

King’s College London
Faculty of Life Science and Medicine
Cardiovascular Division
Submitted 17\textsuperscript{th} May 2019
Abstract

**Rationale:**
Recently, a novel nested intronic gene was discovered from the microarray profiling of the laminar flow-upregulated genes in mouse embryonic stem cells (ESCs). This gene is located in the intron 6 of the lymphoid transcription factor gene *Laf4/Aff3*. Therefore, this novel gene is referred as *Laf4* intron resident (*Laf4ir*). *Laf4ir* exhibits 7 exons and two transcript variants. *Laf4* and *Laf4ir* utilise opposite DNA strands for transcription. *Laf4ir* mRNA sequence contains a few potential open reading frames (ORFs) for the translation of polypeptides.

**Objective:**
In the present study, I aimed to verify whether *Laf4ir* is an encoding gene, explore the expression profile of *Laf4ir*, potential functions and its underlying mechanisms. I also aimed to investigate the role of *Laf4ir* in cardiovascular remodelling and utilise the global knockout mouse model to further understand the functions of *Laf4ir*.

**Methods and results:**
Mass spectrometry detected multiple peptides associated with two ORF polypeptides from the *Laf4ir* gene, demonstrating that *Laf4ir* is an encoding gene. *Laf4ir* ORF1 and *Laf4ir* ORF2 polypeptide expression was found in various adult organs, different stages of embryonic development, different cell types and different subcellular localisations. Immunoprecipitation and immunofluorescence staining revealed the physical association of *Laf4ir* ORF1 polypeptide with LAF4 protein suggesting the potential co-ordinated regulation of the parent and nested genes.

Overexpression of *Laf4ir* ORF2 via adenoviral gene transfer could enhance laminar flow- and VEGF-induced endothelial cell (EC) differentiation and reduce cell proliferation. This could be due to cell cycle arrest via the retention of mini-
chromosome maintenance protein 3 (MCM3) in the cytosol and as a consequence assist differentiation towards EC lineage. Overexpression of *Laf4ir* ORF2 also promoted endothelial survival under oxidative stress induced by hydrogen peroxide.

To investigate the potential contribution of *Laf4ir* in cardiovascular repair and pathology, several different cardiovascular remodelling *in-vivo* models were conducted. Results indicated the upregulation of *Laf4ir* ORF2 polypeptide in transverse aortic constriction-mediated pressure overload in the heart, and femoral artery wall following vascular injury and ischaemia. An alteration in distribution of *Laf4ir* ORF2 expression was also observed in the aorta from *ApoE−/−* mice.

To explore the functionality of *Laf4ir* further, a Cre-*loxP* *in-vivo* global knockout model was developed and utilised. A significant reduction in blood perfusion recovery was detected in *Laf4ir* heterozygous mice when compared to wild type mice at day 14-post surgery of hindlimb ischaemia. This suggests the potential contribution of *Laf4ir* in vascular repair. Phenotypic examination of *Laf4ir* knockout mice displayed a reduction in body size, an enlargement in heart size, malocclusion, abnormal arch posture and abnormal hindlimb vessel growth. Sca1+ adventitia cells isolation from transgenic mice exhibited an increase in proliferation, further supporting the inhibitory role of *Laf4ir* in proliferation.

**Conclusion:**

Overall, the novel nested *Laf4ir* gene may contribute to cardiovascular remodelling via spatiotemporal translation of different ORFs. Further detailed investigation on *Laf4ir* will undoubtedly shed new insights into cardiovascular biology and pathology. This study could also potentially encourage new discoveries for other nested mammalian genes with diverse functions.
# Table of Contents

Abstract .............................................................................................................................................. 3  
Acknowledgement .......................................................................................................................... 9  
Declaration ....................................................................................................................................... 10  
Index of tables ............................................................................................................................... 11  
Index of figures .............................................................................................................................. 12  
List of abbreviations ...................................................................................................................... 15  

Chapter 1: Introduction ....................................................................................................................... 26  
  1.1 Cardiovascular system ............................................................................................................... 26  
  1.2 Cardiovascular diseases ......................................................................................................... 28  
  1.3 Heart failure ............................................................................................................................ 28  
  1.4 Atherosclerosis ....................................................................................................................... 30  
  1.5 Endothelial dysfunction .......................................................................................................... 31  
  1.6 Endothelial cell under oxidative stress .................................................................................. 32  
  1.7 Vessel growth ......................................................................................................................... 34  
  1.8 Stem and vascular progenitor cells ....................................................................................... 37  
    1.8.1 Embryonic stem cells ....................................................................................................... 38  
    1.8.2 Adult stem cells ................................................................................................................ 40  
    1.8.3 Vascular stem/progenitor cells ....................................................................................... 41  
    1.8.4 Stem/progenitor cell proliferation ................................................................................... 47  
    1.8.5 Stem/progenitor cell differentiation ................................................................................. 49  
  1.9 Nested genes ............................................................................................................................ 58  
    1.9.1 Nested intronic genes ...................................................................................................... 59  
  1.10 LAF4 gene ............................................................................................................................. 61  
  1.11 Laf4ir novel gene ................................................................................................................... 65  
  1.12 Project design ......................................................................................................................... 73  
    1.12.1 Hypothesis ...................................................................................................................... 73  
    1.12.2 Aims ............................................................................................................................... 73  

Chapter 2: Materials and methods ..................................................................................................... 75  
  2.1 Mice ........................................................................................................................................ 75  
  2.2 Cell culture ............................................................................................................................. 75  
  2.3 ESC differentiation and shear stress ..................................................................................... 76
Chapter 3: Expression profile of Laf4ir

3.1 Introduction .................................................. 108

3.2 Aims .............................................................. 111

3.3 Results ............................................................ 112

3.3.1 Laf4ir is a novel encoding gene ......................... 112

3.3.2 Computational model of Laf4ir promoter and mRNA .... 115

3.3.3 Computational model of Laf4ir ORF1, ORF2-TV1 and ORF2-TV2 polypeptide sequences .................................. 117

3.3.4 Laf4ir expression in adult organs ......................... 123

3.3.5 Laf4ir expression during embryonic development ........... 127

3.3.6 Laf4ir expression in different cell types ................... 129

3.3.7 Laf4ir expression in bone marrow-derived cells ............. 134

3.3.8 Laf4ir sub-cellular localisation ............................ 139
Chapter 4: Potential functions of Laf4ir and its underlying mechanisms ........................................... 141

4.1 Introduction ........................................................................................................................................... 141

4.2 Aims .......................................................................................................................................................... 142

4.3 Results ...................................................................................................................................................... 143

4.3.1 Laf4ir ORF2 polypeptide enhances CD31 expression in response to VEGF and shear stress-induced stem/progenitor differentiation ................................................................. 143

4.3.2 Laf4ir ORF2 polypeptide retards proliferation in stem/progenitor cells ........................................... 146

4.3.3 Laf4ir may affect proliferation via the interaction of Laf4ir ORF2 polypeptide with MCM3 in shear stress-induced treated dESCs ................................................................. 149

4.3.4 Laf4ir ORF1 polypeptide interaction with LAF4 ............................................................................. 151

4.3.5 Laf4ir ORF2 polypeptide protects against oxidative stress induced by hydrogen peroxide in endothelial cells. .............................................................................................................. 153

4.3.6 Laf4ir ORF2 polypeptide reduces hydrogen peroxide in endothelial cells .................................... 158

Chapter 5: In-vivo model of Laf4ir in cardiovascular remodelling ............................................................ 160

5.1 Introduction .............................................................................................................................................. 160

5.2 Aims .......................................................................................................................................................... 161

5.3 Results ...................................................................................................................................................... 162

5.3.1 Upregulation of Laf4ir ORF2 and downregulation of Laf4ir ORF1 and LAF4 in a TAC model .......... 162

5.3.2 Altered expression of Laf4ir ORF2 in ApoE /- aorta ....................................................................... 164

5.3.3 Upregulation of Laf4ir ORF2 polypeptide in a hindlimb ischemia model ..................................... 169

5.3.4 Impaired vascular recovery in ischaemic-induced Laf4ir /- mice .................................................... 171

5.3.5 Upregulation of Laf4ir ORF2 polypeptide in a femoral artery injury model .................................. 174

5.3.6 Genotype of global Laf4ir knockout mice ....................................................................................... 177

5.3.7 Phenotype of global Laf4ir knockout mice ....................................................................................... 179

Chapter 6: Discussions ................................................................................................................................. 183

6.1 Co-ordination between nested Laf4ir and parent Laf4 genes .............................................................. 183

6.2 Multiple translation from single Laf4ir mRNA transcript .................................................................... 184

6.3 Laf4ir role in stem/progenitor proliferation and differentiation .......................................................... 187

6.4 Laf4ir role in endothelial cell survival against oxidative stress .......................................................... 189

6.5 Laf4ir role in cardiovascular remodelling ............................................................................................ 190

6.6 Laf4ir phenotype in global knockout mice ............................................................................................ 193

6.7 Future directions .................................................................................................................................... 194

6.8 Conclusions .......................................................................................................................................... 198

Chapter 7: Publications ............................................................................................................................... 199
Acknowledgement

I wish to express my sincere gratitude to my primary supervisor, Dr Lingfang Zeng, for his professional supervision during the past three years of my PhD study. Dr Zeng’s invaluable expertise and advice were incredibly important throughout the entire PhD journey and I will appreciate his support. I would also like to acknowledge my secondary supervisor, Dr Yanhua Hu, who performed in-vivo surgeries.

I am grateful to have the opportunity to teach and supervise a few talented postgraduate and undergraduate students: Mr Christopher Whiteman, Mr Edward Gregory and Ms Yaqing Si. I would like to thank the research associates and PhD students in our laboratory group for their advice and help: Dr Ana Moraga Yebenes, Dr Ka Hou Lao, Dr Angshumonkik Angbohang, Mr Jose Morales and Mr Yue Zhao.

I am grateful for Dr Wenguang Gu’s hands-on experience in the isolation of adventitia cells from transgenic mice aortas for cell culture and for Dr Witold N. Nowak who shared his expertise in flow cytometry in the bone-marrow derived cell study. I would like to acknowledge the support of all the other members from Professor Qingbo Xu’s group: Dr Jiacheng Deng, Dr Ioannis Kokkinopoulos, Dr Zhichao Ni, Dr Shirin Issa Bhaloo, Dr Xuechong Hong and Dr Alexandra Le Bras.

Lastly and most importantly, I must say thank you to my parents, Ms Maryam Mahdavi and Mr Masoud Ehteramyan, for their constant unconditional love and support throughout my whole life. Without their irrepressible challenges as first generation migrants to the UK from Iran, I would have not been able to accomplish any of my life ambitions. I will forever be in debt for their sacrifices and no one could ask for better parents.
Declaration

I, Mr Mazdak Ehteramyan, declare that this PhD thesis and the research to which it refers is a product of my own work. Mr Steven Lynham and Dr Xiaoping Yang obtained vital mass spectrometry data from my samples. Dr Yanhua Hu conducted hindlimb ischeamia and femoral artery injury surgery models and Prof Shah’s laboratory group conducted the trans-aortic constriction surgery model. Information derived from other sources and work carried out with the assistance of others has been appropriately cited and acknowledged.
## Index of tables

Table 1: Models of mouse endothelial cell differentiation from embryonic stem cells 50
Table 2: Phenotypic observation from *Laf4* knockout *in-vivo* model .......................... 63
Table 3: Illustration of all primers used ..................................................................................... 79
Table 4: Optimised PCR conditions for the detection of the *Laf4ir* conditional and constitute knock-out allele .............................................................................................................. 81
Table 5: PCR amplicon base pair (bp) product expected from primers X and Y .................. 81
Table 6: Primers X and Y for the detection of the *Laf4ir* conditional Knock-out allele ........ 81
Table 7: PCR amplicon product expected from primers A, B and C ........................................ 83
Table 8: Primers for the detection of the *Laf4ir* constitutive knock-out allele ..................... 83
Table 9: Illustration of all monoclonal and polyclonal primary antibodies used ..................... 89
Table 10: Antibodies for immunophenotyping bone marrow-derived cells ......................... 96
Table 11: Computational calculation of isoelectric point and molecular weight of LAF4IR polypeptides ................................................................................................................................. 110
Table 12: Mass spectrometry data identifying peptide sequences of *Laf4ir* ORF1 and ORF2 ........................................................................................................................................... 114
Table 13: Computational analysis of *Laf4ir* polypeptides .................................................... 123
Table 14: Flow cytometric data of *Laf4ir* ORF2\(^+\) cells in different bone marrow-derived cell types .................................................................................................................................................. 139
Table 15: Changes in *Laf4ir* ORF1 and ORF2 polypeptide expression in different experimental models ................................................................................................................................. 176
Table 16: Phenotypic observation from *Laf4ir* knockout *in-vivo* model ............................ 182
Index of figures

Figure 1: Schematic diagram of blood vessels. .................................................................27
Figure 2: Schematic diagram of vessel development.........................................................36
Figure 3: Representative illustration of stem cell differentiation.....................................38
Figure 4: LIF-STAT3 signalling pathway linked with self-renewal and proliferation ....39
Figure 5: Schematic illustration of vascular wall resident stem/progenitor cells...........46
Figure 6: Illustration of eukaryotic initiation of DNA replication....................................48
Figure 7: Illustration of mechanisms involved in EC differentiation from ESCs...........57
Figure 8: Schematic illustration of a nested gene.............................................................59
Figure 9: Illustration of human LAF4 mRNA expression in different tissues ..............62
Figure 10: Characterisation of ECD1 and ECD2 genes..................................................66
Figure 11: Northern blot analysis of Laf4ir in different organs.....................................68
Figure 12: The nucleotide sequences of Laf4ir transcript variant 1 and 2.....................69
Figure 13: Schematic illustration of the Laf4ir gene structure........................................70
Figure 14: Amino acid sequences of the three potential ORFs........................................71
Figure 15: Illustration of the conservation of Laf4ir gene among different species......72
Figure 16: Representative diagram of the shear stress protocol....................................77
Figure 17: PCR genotype of the Laf4ir endogenous and conditional knock-out mouse line..................................................................................................................82
Figure 18: PCR genotype of the Laf4ir endogenous, conditional and constitute knock-out mouse line........................................................................................................84
Figure 19: Diagram of a trans-aortic constriction model in mice. .................................101
Figure 20: Diagram of a femoral artery injury model in mice........................................103
Figure 21: Diagram of a hindlimb ischemia model in mice............................................105
Figure 22: pShuttle2-L151 vector transfection and immunostaining..............................112
Figure 23: Laf4ir ORF1 and ORF2 detection with other peptide sequences from mass spectrometry........................................................................................................114
Figure 24: Bioinformatics analysis of Laf4ir promoter with predicted transcription factor binding sites.................................................................116
Figure 25: Bioinformatics analysis of Laf4ir mRNA TV1 with predicted ribosome binding sites.............................................................................................................................................116
Figure 26: Predicted computational model of Laf4ir ORF1 polypeptide.................................118
Figure 27: Predicted computational model of Laf4ir ORF2-TV1 polypeptide.........................120
Figure 28: Predicted computational model of Laf4ir ORF2-TV2 polypeptide.........................122
Figure 29: Laf4ir ORF1 expression in different adult organs ...................................................125
Figure 30: Laf4ir ORF2 expression in different adult organs. ..................................................126
Figure 31: Laf4ir ORF1 and ORF2 polypeptide expressions at different stages of embryonic development. .............................................................................................................................................128
Figure 32: LAF4 protein, Laf4ir ORF1 and ORF2 polypeptide expressions in embryonic stem cells. .............................................................................................................................................131
Figure 33: Vascular progenitor cells characterised by Sca-1 expression. ..............................132
Figure 34: Laf4ir ORF1 and ORF2 polypeptide positive cells in VPCs.................................133
Figure 35: Flow cytometric analysis of Laf4ir ORF2+ bone marrow-derived cell types. .............................................................................................................................................138
Figure 36: Sub-cellular localisation of Laf4ir ORF1 and ORF2 polypeptides. .......................140
Figure 37: Overexpression of 151aa sequence upregulates CD31 expression. ....................145
Figure 38: Overexpression of 151aa sequence reduces proliferation .................................146
Figure 39: Enhanced proliferation in Laf4ir +/- vascular progenitor cells .........................148
Figure 40: Co-localisation of MCM3 protein with Laf4ir ORF2 polypeptide ......................150
Figure 41: Laf4ir ORF1 polypeptide interaction with LAF4 protein ....................................152
Figure 42: Overexpression of 151aa sequence protects endothelial cells under oxidative stress .............................................................................................................................................153
Figure 43: Overexpression of 151aa sequence reduces apoptosis in endothelial cells under oxidative stress .............................................................................................................................................155
Figure 44: Overexpression of 151aa sequence improves cellular survival in endothelial cells under oxidative stress .............................................................................................................................................157
Figure 45: Overexpression of 151aa sequence in endothelial cells reduces intracellular and extracellular H₂O₂ .............................................................................................................................................159
Figure 46: Laf4ir ORF1 and ORF2 expression in a trans-aortic constriction model. ....163
Figure 47: *Laf4ir* ORF2 polypeptide expression is endothelial cell specific in the aorta ..................................................................................................................................................................................166

Figure 48: *Laf4ir* ORF1 polypeptide and ORF2 polypeptide expression in wild type and *ApoE*⁻/⁻ aorta.................................................................................................................................................................................................................................................................167

Figure 49: *Laf4ir* ORF2 polypeptide expression in wild type and *ApoE*⁻/⁻ vascular progenitor cells..........................................................................................................................................................................................................................................................................................................................168

Figure 50: *Laf4ir* ORF1 polypeptide and ORF2 polypeptide expression in a hindlimb ischaemia model..........................................................................................................................................................................................................................................................................................................................170

Figure 51: *Laf4ir* knockout reduced foot blood perfusion in a mouse hindlimb ischaemia model..........................................................................................................................................................................................................................................................................................................................172

Figure 52: *Laf4ir* knockout reduced Sca-1⁺ and CD31⁺ cells in a mouse hindlimb ischaemia model..........................................................................................................................................................................................................................................................................................................................173

Figure 53: *Laf4ir* ORF1 polypeptide is not expressed in a femoral artery injury model. ..........................................................................................................................................................................................................................................................................................................................175

Figure 54: *Laf4ir* ORF2 polypeptide upregulation in a femoral artery injury model. ..176

Figure 55: Genotype of *Laf4ir* knockout mouse line and expression of *Laf4ir* ORF1 polypeptide, ORF2 polypeptide and LAF4 protein. ..........................................................................................................................................................................................................................................................................................................................178

Figure 56: Body size phenotypic examination of *Laf4ir* knockout mouse line. ........180

Figure 57: Heart size phenotypic examination of *Laf4ir* knockout mouse line..........181
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>18s</td>
<td>18 Svedberg unit</td>
</tr>
<tr>
<td>28s</td>
<td>28 Svedberg unit</td>
</tr>
<tr>
<td>40S</td>
<td>40 Svedberg unit</td>
</tr>
<tr>
<td>A$_{2A}$R</td>
<td>Adenosine 2A receptor</td>
</tr>
<tr>
<td>A</td>
<td>Antibody</td>
</tr>
<tr>
<td>aa</td>
<td>Amino acid</td>
</tr>
<tr>
<td>ac-LDL</td>
<td>Acetylated low-density lipoprotein</td>
</tr>
<tr>
<td>ACTB</td>
<td>Beta actin</td>
</tr>
<tr>
<td>ACTN3</td>
<td>Alpha-actinin-3</td>
</tr>
<tr>
<td>Ade3</td>
<td>Adenosine 3</td>
</tr>
<tr>
<td>Ad-L109</td>
<td>Adenovirus L109</td>
</tr>
<tr>
<td>Ad-L151</td>
<td>Adenovirus L151</td>
</tr>
<tr>
<td>AFF</td>
<td>AF4/FMR2</td>
</tr>
<tr>
<td>ALL</td>
<td>Lymphoblastic leukaemia</td>
</tr>
<tr>
<td>αSMA</td>
<td>alpha smooth muscle actin</td>
</tr>
<tr>
<td>Alt-ATXN1</td>
<td>Alternative Ataxin-1</td>
</tr>
<tr>
<td>altORF</td>
<td>Alternative opening reading frame</td>
</tr>
<tr>
<td>Ang-1</td>
<td>Angiopoietin 1</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>AP</td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td>ApoB100</td>
<td>Apolipoprotein B100</td>
</tr>
<tr>
<td>ApoE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>AT$_{1a}$R</td>
<td>Angiotensin type 1a receptor</td>
</tr>
<tr>
<td>ATXN1</td>
<td>Ataxin-1</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone morphogenetic proteins</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>BP</td>
<td>Blocking peptide</td>
</tr>
<tr>
<td>BrdU</td>
<td>Bromodeoxyuridine / 5-bromo-2'-deoxyuridine</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>Ca²⁺</td>
<td>Calcium ion</td>
</tr>
<tr>
<td>CAD</td>
<td>Caspase-activated DNase</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CD105</td>
<td>Cluster of differentiation 105</td>
</tr>
<tr>
<td>CD106</td>
<td>Cluster of differentiation 106</td>
</tr>
<tr>
<td>CD11b</td>
<td>Cluster of differentiation 11b</td>
</tr>
<tr>
<td>CD11c</td>
<td>Cluster of differentiation 11c</td>
</tr>
<tr>
<td>CD14</td>
<td>Cluster of differentiation 14</td>
</tr>
<tr>
<td>CD144</td>
<td>Cluster of differentiation 144</td>
</tr>
<tr>
<td>CD16</td>
<td>Cluster of differentiation 16</td>
</tr>
<tr>
<td>CD19</td>
<td>Cluster of differentiation 19</td>
</tr>
<tr>
<td>CD24</td>
<td>Cluster of differentiation 24</td>
</tr>
<tr>
<td>CD31</td>
<td>Cluster of differentiation 31</td>
</tr>
<tr>
<td>CD32</td>
<td>Cluster of differentiation 32</td>
</tr>
<tr>
<td>CD34</td>
<td>Cluster of differentiation 34</td>
</tr>
<tr>
<td>CD45</td>
<td>Cluster of differentiation 45</td>
</tr>
<tr>
<td>CD45</td>
<td>Cluster of differentiation 45</td>
</tr>
<tr>
<td>CD64</td>
<td>Cluster of differentiation 64</td>
</tr>
<tr>
<td>CD73</td>
<td>Cluster of differentiation 73</td>
</tr>
<tr>
<td>CD79α</td>
<td>Cluster of differentiation 79α</td>
</tr>
<tr>
<td>CD9</td>
<td>Cluster of differentiation 9</td>
</tr>
<tr>
<td>CD90</td>
<td>Cluster of differentiation 90</td>
</tr>
<tr>
<td>CD90</td>
<td>Cluster of differentiation 90</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CDS</td>
<td>Coding sequence</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>CFU-Hill</td>
<td>Colony forming unit-Hill</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary heart disease</td>
</tr>
<tr>
<td>CK1</td>
<td>Casein kinase 1</td>
</tr>
<tr>
<td>CMG</td>
<td>Cdc45/MCM2–7/GINS</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CNN1</td>
<td>Calponin 1</td>
</tr>
<tr>
<td>COUP-TFII</td>
<td>COUP transcription factor 2</td>
</tr>
<tr>
<td>cT</td>
<td>Cycle threshold</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>CVS</td>
<td>Cardiovascular system</td>
</tr>
<tr>
<td>CXCL8</td>
<td>Chemokine (C-X-C motif) ligand 8</td>
</tr>
<tr>
<td>CXCR-4</td>
<td>C-X-C chemokine receptor type 4</td>
</tr>
<tr>
<td>DAPI</td>
<td>4′-6-Diamidino-2-phenylindol</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Double distilled H₂O</td>
</tr>
<tr>
<td>DDK</td>
<td>Dbf4-dependent kinase</td>
</tr>
<tr>
<td>dESC</td>
<td>Differentiated embryonic stem cell</td>
</tr>
<tr>
<td>DF</td>
<td>Disturbed flow</td>
</tr>
<tr>
<td>DH</td>
<td>Double hexamer</td>
</tr>
<tr>
<td>DM</td>
<td>Differentiation media</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's Modified Eagle's Medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Nucleoside diphosphates</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double stranded deoxyribonucleic acid</td>
</tr>
<tr>
<td>dsRNA</td>
<td>Double stranded ribonucleic acid</td>
</tr>
<tr>
<td>dVPC</td>
<td>Differentiated vascular progenitor cell</td>
</tr>
<tr>
<td>EB</td>
<td>Embryoid body</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ECs</td>
<td>Endothelial cells</td>
</tr>
<tr>
<td>ECD1</td>
<td>Endothelial cell differentiation-related gene 1</td>
</tr>
<tr>
<td>ECD2</td>
<td>Endothelial cell differentiation-related gene 2</td>
</tr>
<tr>
<td>ECFCs</td>
<td>Endothelial colony-forming cells</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EDHF</td>
<td>Endothelium derived hyperpolarizing factor</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EDV</td>
<td>End-diastolic volume</td>
</tr>
<tr>
<td>EF</td>
<td>Ejection fraction</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>eNOS</td>
<td>Endothelial nitric oxide synthase</td>
</tr>
<tr>
<td>EPCs</td>
<td>Endothelial progenitor cells</td>
</tr>
<tr>
<td>EPO</td>
<td>Erythropoietin</td>
</tr>
<tr>
<td>ERK</td>
<td>Extracellular regulated kinase</td>
</tr>
<tr>
<td>Erk</td>
<td>Extracellular signal–regulated kinases</td>
</tr>
<tr>
<td>ESC</td>
<td>Embryonic stem cell</td>
</tr>
<tr>
<td>ESV</td>
<td>End-systolic volume</td>
</tr>
<tr>
<td>ET-1</td>
<td>Endothelin-1</td>
</tr>
<tr>
<td>ETV2</td>
<td>ETS Variant 2</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>FGF2</td>
<td>Fibroblast growth factor 2</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>FLAG</td>
<td>Flagellar</td>
</tr>
<tr>
<td>FL1-4</td>
<td>Fluorescence 1-4</td>
</tr>
<tr>
<td>Fli1</td>
<td>Friend leukemia integration 1</td>
</tr>
<tr>
<td>FLP</td>
<td>Flippase</td>
</tr>
<tr>
<td>Fox</td>
<td>Forkhead</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>FRT</td>
<td>Flippase recognition target</td>
</tr>
<tr>
<td>FSC-A</td>
<td>Forward Scatter area</td>
</tr>
<tr>
<td>FSC-H</td>
<td>Forward Scatter height</td>
</tr>
<tr>
<td>G4</td>
<td>G-quadruplex</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GC</td>
<td>Guanine-cytosine</td>
</tr>
<tr>
<td>G-CSF</td>
<td>Granulocyte-colony stimulating factor</td>
</tr>
<tr>
<td>gDNA</td>
<td>Genomic DNA</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Granulocyte-monocyte-colony stimulating factor</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>GPX</td>
<td>Glutathione peroxidase</td>
</tr>
<tr>
<td>GTEx</td>
<td>Genotype-Tissue Expression</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome wide association studies</td>
</tr>
<tr>
<td>H₂O</td>
<td>Water</td>
</tr>
<tr>
<td>H₂O₂</td>
<td>Hydrogen peroxide</td>
</tr>
<tr>
<td>H&amp;E</td>
<td>Hematoxylin and eosin</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>Trimethylation to the lysine 27 on the histone H3</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>Trimethylation to the lysine 4 on the histone H3</td>
</tr>
<tr>
<td>HA</td>
<td>Hemagglutinin</td>
</tr>
<tr>
<td>HATs</td>
<td>Histone acetyltransferases</td>
</tr>
<tr>
<td>HDAC3</td>
<td>Histone Deacetylase 3</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic kidney cell 293</td>
</tr>
<tr>
<td>HGF</td>
<td>Hepatocyte growth factor</td>
</tr>
<tr>
<td>HH14</td>
<td>Hamburger Hamilton Stage 14</td>
</tr>
<tr>
<td>HIF-1</td>
<td>Hypoxia-inducible factor-1</td>
</tr>
<tr>
<td>HLA-D</td>
<td>Human leukocyte antigen D</td>
</tr>
<tr>
<td>HMG-CoA</td>
<td>(\beta)-Hydroxy (\beta)-methylglutaryl-CoA</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>HO-1</td>
<td>Heme oxygenase-1</td>
</tr>
<tr>
<td>HOCl</td>
<td>Hypochlorous acid</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HUVECs</td>
<td>Human umbilical vein endothelial cells</td>
</tr>
<tr>
<td>IA/IE</td>
<td>I-A and I-E MHC-II molecules</td>
</tr>
<tr>
<td>ICAM-1</td>
<td>Intracellular adhesion molecule 1</td>
</tr>
<tr>
<td>ICC</td>
<td>Immunocytochemistry</td>
</tr>
<tr>
<td>ICH</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>IF</td>
<td>Immunofluorescence</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IGF</td>
<td>Insulin-like growth factor</td>
</tr>
<tr>
<td>IL-1</td>
<td>Interleukin 1</td>
</tr>
<tr>
<td>IL-6</td>
<td>Interleukin 6</td>
</tr>
<tr>
<td>INF-γ</td>
<td>Interferon gamma</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitation</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosome entry site</td>
</tr>
<tr>
<td>K⁺</td>
<td>Potassium ion</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium chloride</td>
</tr>
<tr>
<td>KOH</td>
<td>Potassium hydroxide</td>
</tr>
<tr>
<td>LAF4</td>
<td>Af4/Fmr2 family member 3</td>
</tr>
<tr>
<td>Laf4ir</td>
<td>LAF4 intron resident</td>
</tr>
<tr>
<td>LDL</td>
<td>Low-density lipoprotein</td>
</tr>
<tr>
<td>LF</td>
<td>Laminar flow</td>
</tr>
<tr>
<td>LIF</td>
<td>Leukaemia inhibitory factor</td>
</tr>
<tr>
<td>LoxP</td>
<td>Locus of Crossover in P1</td>
</tr>
<tr>
<td>LRP1</td>
<td>Low density lipoprotein receptor-related protein 1</td>
</tr>
<tr>
<td>LV</td>
<td>Left ventricular</td>
</tr>
</tbody>
</table>
Ly6C  Lymphocyte Ag 6C
Ly6G  Lymphocyte Ag 6G
MAPK  Mitogen-activated protein kinase
MS    Mass spectrometry
m/z   Mass-to-charge ratio
MCM   Minichromosome maintenance
MCP-1  Monocyte chemoattractant peptide 1
MDGA2 MAM domain containing glycosylphosphatidylinositol anchor 2
MEF   Murine embryonic fibroblasts
miRNA Micro RNA
MLL   Mixed lineage leukaemia
MMP   Matrix metalloproteinase
MOI   Multiplicity of infection
MR    Magnetic resonance
MRTFs Myocardin-related transcription factors
MSCs  Mesenchymal stem cells
MTS   3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Mw    Molecular weight
Myc   Myelocytomatosis oncogene
MYH4  Myosin Heavy Chain 4
NaAc  Sodium Acetate
NADPH Nicotinamide adenine dinucleotide phosphate
NBT   Nitro blue tetrazolium
NE    Nuclear extract
NFκB Nuclear factor kappa-light-chain-enhancer of activated B cells
NMDA  N-methyl-D-aspartate
NO    Nitric oxide synthase
NOS   Nitric oxide synthase
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOX</td>
<td>NADPH oxidase</td>
</tr>
<tr>
<td>NQO1</td>
<td>NAD(P)H Quinone Dehydrogenase 1</td>
</tr>
<tr>
<td>Nrf3</td>
<td>Nuclear factor erythroid 2-related factor 3</td>
</tr>
<tr>
<td>NRP-1</td>
<td>Neuropilin-1</td>
</tr>
<tr>
<td>NT</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>O2</td>
<td>Superoxide anion</td>
</tr>
<tr>
<td>OCCM</td>
<td>ORC/Cdc6/Cdt1/MCM2–7</td>
</tr>
<tr>
<td>OCT</td>
<td>Optimal cutting temperature compound</td>
</tr>
<tr>
<td>Oct4</td>
<td>Octamer-binding transcription factor 4</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>OH</td>
<td>Hydroxyl radical</td>
</tr>
<tr>
<td>ONOO−</td>
<td>Peroxynitrite</td>
</tr>
<tr>
<td>ORC</td>
<td>Origin recognition complex</td>
</tr>
<tr>
<td>ORF</td>
<td>Opening reading frame</td>
</tr>
<tr>
<td>OS</td>
<td>Oscillatory shear</td>
</tr>
<tr>
<td>Ox-LDL</td>
<td>Oxidised low-density lipoprotein</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCA</td>
<td>Protocatechuic acid</td>
</tr>
<tr>
<td>Pcp</td>
<td>Pupal cuticle protein</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDGF</td>
<td>Platelet-derived growth factor</td>
</tr>
<tr>
<td>PECAM-1</td>
<td>Platelet endothelial cell adhesion molecule 1</td>
</tr>
<tr>
<td>PFA</td>
<td>Paraformaldehyde</td>
</tr>
<tr>
<td>pl</td>
<td>Isoelectric point</td>
</tr>
<tr>
<td>PI</td>
<td>Propidium iodide</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphoinositide 3-kinase</td>
</tr>
<tr>
<td>PIGF</td>
<td>Placental growth factor</td>
</tr>
<tr>
<td>PKA</td>
<td>Protein kinase A</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PKB</td>
<td>Protein kinase B</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein kinase C</td>
</tr>
<tr>
<td>PLC-γ</td>
<td>Phospholipase C-γ</td>
</tr>
<tr>
<td>pre-RC</td>
<td>Pre-replication complex</td>
</tr>
<tr>
<td>PRDX</td>
<td>Peroxiredoxins</td>
</tr>
<tr>
<td>PS</td>
<td>Pulsatile shear</td>
</tr>
<tr>
<td>PTM</td>
<td>Post-translational modification</td>
</tr>
<tr>
<td>PUFAs</td>
<td>Polyunsaturated fatty acids</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>RBS</td>
<td>Ribosome binding sites</td>
</tr>
<tr>
<td>RFK</td>
<td>Replication fork</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>Ribonucleic acid sequencing</td>
</tr>
<tr>
<td>ROCK</td>
<td>Rho-associated protein kinase</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>RS</td>
<td>Replication stress</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error of the mean</td>
</tr>
<tr>
<td>S1P</td>
<td>Sphingosine-1-phosphate</td>
</tr>
<tr>
<td>Sca-1</td>
<td>Stem cells antigen-1</td>
</tr>
<tr>
<td>SCF</td>
<td>Stem cell factor</td>
</tr>
<tr>
<td>SDF1</td>
<td>Stromal cell-derived factor 1</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SHB</td>
<td>Src Homology-2 domain containing protein B</td>
</tr>
<tr>
<td>siRNA</td>
<td>small interfering ribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>SKIP</td>
<td>Ski-Related Protein</td>
</tr>
<tr>
<td>SM22</td>
<td>Smooth muscle 22</td>
</tr>
<tr>
<td>SMA</td>
<td>Smooth muscle actin</td>
</tr>
<tr>
<td>SMC</td>
<td>Smooth muscle cells</td>
</tr>
<tr>
<td>SOD</td>
<td>Superoxide dismutase</td>
</tr>
<tr>
<td>SOX-2</td>
<td>SRY (sex determining region Y)-box 2</td>
</tr>
<tr>
<td>SPC</td>
<td>Stem/progenitor cell</td>
</tr>
<tr>
<td>SPECT</td>
<td>Single photon emission computed tomography</td>
</tr>
<tr>
<td>SRF</td>
<td>Serum response factor</td>
</tr>
<tr>
<td>SS</td>
<td>Shear stress</td>
</tr>
<tr>
<td>SSC-A</td>
<td>Side scatter area</td>
</tr>
<tr>
<td>SSC-H</td>
<td>Side scatter height</td>
</tr>
<tr>
<td>SSEA-3</td>
<td>Stage-specific embryonic antigen 3</td>
</tr>
<tr>
<td>SSEA-4</td>
<td>Stage-specific embryonic antigen 4</td>
</tr>
<tr>
<td>ST</td>
<td>Static</td>
</tr>
<tr>
<td>STAT3</td>
<td>Signal transducer and activator of transcription 3</td>
</tr>
<tr>
<td>STH</td>
<td>Saitohin</td>
</tr>
<tr>
<td>TAC</td>
<td>Trans-aortic constriction</td>
</tr>
<tr>
<td>TAQ</td>
<td>Thermus aquaticicus</td>
</tr>
<tr>
<td>TCF</td>
<td>T-cell factor</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Transforming growth factor beta</td>
</tr>
<tr>
<td>TIMP</td>
<td>Tissue inhibitors of MMPs</td>
</tr>
<tr>
<td>TIS</td>
<td>Translation initiation site</td>
</tr>
<tr>
<td>TK</td>
<td>Thymidine kinase</td>
</tr>
<tr>
<td>TLR4</td>
<td>Toll-like receptor 4</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumor necrosis factor-alpha</td>
</tr>
<tr>
<td>TV1</td>
<td>Transcript variant 1</td>
</tr>
</tbody>
</table>
TV2  Transcript variant 2
TXA2  Thromboxane
UEA-1  Ulex europaeus agglutinin-1
UTRs  Untranslated regions
UV  Ultra violet
V  Volts
VCAM-1  Vascular cell adhesion molecule 1
VE-cadherin  Vascular endothelial-cadherin
VEGF  Vascular endothelial growth factor
VEGFR2  Vascular endothelial growth factor receptor 2
VPCs  Vascular progenitor cells
vWF  Von Willebrand factor
WB  Western blot
WNT  Wingless/Integrated
WNT2  Wnt Family Member 2
WT  Wild type
Wt1  Wilms' tumor-1
XOD  Xanthine oxidase
Chapter 1: Introduction

1.1 Cardiovascular system

The cardiovascular system transports blood from the heart to all the organs, tissues and ultimately cells in the body (Figure 1). The vasculature comprises of arteries, veins and capillaries. Diverse heterogeneity among the circulatory system exists. Arteries comprise of three-layered structured components: tunica adventitia, media and intima. The outermost layer is the tunica adventitia which contains connective tissue (mainly collagen), fibroblasts, nerves, resident stem/progenitor cells (SPCs), adipocytes and vasa vasorum. Smooth muscle cells (SMCs), elastic fibres and collagens make up the media layer. The inner layer is the tunica intima, which is covered by a single layer of endothelial cells (ECs) and is directly exposed to blood flow. The extracellular matrices (ECMs) form the subintimal layer and participate in the structural integrity and functional state of ECs.

Large arteries can be further subdivided into either elastic or muscular arteries. In elastic arteries such as the aorta and the carotid artery, there are SMCs distributed throughout the intima as well as in subintimal layers. In the media layer, multiple layers of SMCs are surrounded by elastin fibres forming SMC contractile units. This elastin structure is required to provide strong flexibility within elastic arteries to expand in response to high blood pressure. Elastic arteries are mainly positioned at close proximity to the heart. Muscular arteries such as the femoral, ulnar, and radial arteries, are smaller. Rather than mitigate systolic and diastolic pressure fluctuation, they are found in areas specific to tissues and organs. When the diameter of an artery is less than 0.3mm, it is defined as an arteriole. Although arterioles also have a three layer structure, all three layers are thinner than large arteries and constitute of 1-2 layers of smooth muscle cells in the media. Arterioles then branch into capillaries.
Figure 1: Schematic diagram of blood vessels.
Oxygenated blood is pumped from the left ventricle of the heart and circulates throughout the body. Deoxygenated blood returns to the right atrium where it is pumped to the lungs for re-oxygenation. Both artery and veins are comprised of tunica intima, tunica media and tunica external layers. Figure source: (https://healthjade.com/aortic-aneurysm)

Similarly to arteries, veins can be separated into the large veins or venules according to diameter. Large veins are also made up of three layers of tissue. However, the venous media layer contains fewer SMCs compared to same size of artery. Additionally, most veins are equipped with valves to prevent backflow. Distinct to arterioles, venules have a thinner structure and are equipped with pores. Capillaries are formed from an EC layer, supported by the basement membrane with a diameter of less than 10μm. The main function of capillaries is substance exchange from arterioles to parenchymal tissues within the microcirculation. According to structure, capillaries can be divided into three types: continuous, fenestrated and sinusoid capillaries.
Continuous capillaries provide an uninterrupted lining layer, in which ECs are connected by intercellular junctions, allowing only small molecules to pass through. Both fenestrated and sinusoidal capillaries have pores within cells or inter-cellularly, which allow big molecules and cells to pass through capillary.

1.2 Cardiovascular diseases

Cardiovascular diseases (CVDs) still account for the majority of mortalities worldwide. Significant advancements in diagnosis, treatments and management of patients have fortunately contributed in the decline of deaths linked with CVDs over the past 40 years. The annual rate of CVD-related deaths in the United States have reduced by 31% between 1998 and 2008, with similar trends observed in European Union countries; the UK has observed a decline of 68% between 1980 and 2013. Hospital admissions in the UK for all CVDs have increased by over 46 000 between 2010/2011 and 2013/2014, and CVD-related prescriptions/operations have also increased in the last decade. Pathogenesis underlying coronary heart disease (CHD) and different forms of strokes is due to atherosclerosis, a chronic inflammatory condition that starts early in life and advances with age. Typical clinical signs manifest in one of three forms: an acute coronary syndrome such as myocardial infarction, angina pectoris or sudden cardiac death. Abrupt loss of cardiac function (cardiac arrest) within a short period is usually attributed by abnormal electrical regulation, resulting in a fatal arrhythmia including ventricular fibrillation. Risk factors of CHD include sex, smoking, drug abuse, hypertension, dyslipidemia, diabetes, lack of exercise, diet and energy intake, alcohol consumption, and stress.

1.3 Heart failure

To date, more than 25 million patients have been reported to be diagnosed with heart failure. Heart failure is characterised by maladaptive cardiac function, various clinical manifestations and complex systemic perturbations. These include
hypertension, ischaemia, arrhythmia, medical/dietary noncompliance, inflammation, abnormal breathing during sleep and comorbid disorders, such as renal and pulmonary disease. Patients also have certain symptoms and signs (e.g. dyspnoea, angina, orthopnoea, oedema and weight gain), that help the process of diagnosis. Endothelial dysfunction, venous congestion, neurohormonal stimulation and cardiac hypertrophy are all indicative in the pathogenesis of heart failure. Regardless of the substantial improvements in the understanding of heart failure and advances in treatment strategies in the past four decades, management of heart failure remains unaltered. Treatments including intravenous diuretics, vasodilators, and inotropes and novel therapeutic approaches are vital to reduce the economic burden and health epidemic problems worldwide manifested by heart failure.

In order to characterise molecular mechanisms of heart failure in mice, investigators have used two main models; pressure overload by the transverse aortic constriction (TAC) and myocardial infarction by the coronary artery ligation. TAC initiates a sudden onset of severe hypertension, enhances cardiac contractility, drives left ventricular hypertrophy and in response to maladaptive hemodynamic overload, cardiac dilation and heart failure is developed. Myocardial infarction in a mouse model of coronary ligation enables researchers to investigate the pathophysiology of ischemia-reperfusion injury during cardiac remodelling. The use of mouse models of heart failure has proven crucial in better understanding the complex cardiovascular pathology.

Myocardial function, coronary circulation and haemodynamics are modulated by various different factors such as the autonomic nervous system and the production and regulation of endothelial nitric oxide (NO), prostaglandins and cytokines. Dysfunctional NO-induced vasodilation within the endothelium is known to be involved in heart failure and is referred as endothelitis. Unregulated levels of NO and development of oxidative stress are also implicated in cardiorenal syndrome. Amplification of venous arm pressure in normal volunteers has been suggested to change endothelial transcript
expression profiles and activate neurohormonal molecules and increase inflammation. To date, there have been substantial advancement in the understanding of heart failure but the precise pathogenesis of the disease is still unestablished.

1.4 Atherosclerosis

Atherosclerosis is a chronic inflammatory disorder characterised by the hardening and loss of elasticity of the artery wall with disturbed blood flow due to the development of atheromatous plaques. As the principal underlying cause of cardiovascular disease, the leading cause of death and morbidity worldwide; atherosclerosis has enormous clinical relevance.

The earliest pathophysiological stage is initiated by endothelial cell dysfunction and structural changes (such as the exposure of proteoglycans and the absence of the luminal elastin layer), which allow the build-up of low-density lipoprotein (LDL) within the sub-endothelial layer. Increased concentrations of circulating cholesterol conveyed by LDL contain apolipoprotein B100 (ApoB100) which plays an important role in the development of the disease. ApoB100 interactions with proteoglycans leads to retention of LDL molecules in the intima, where they are vulnerable to oxidative modification mediated by reactive oxygen species (ROS) or lipoxygenases or myeloperoxidase secreted from inflammatory cells. Repeated exposure to disturbed flow shear stress in atheroprone regions produces ROS by endothelial NADPH oxidase (NOX). Oxidised LDL (Ox-LDL) initiates the expression of adhesion molecules including P-selectin, VCAM-1, and E-selectin and the release of chemokines by endothelial cells. With the addition of chemokines derived from deposited platelets, an immune cell infiltration is driven: mainly by monocytes. Monocytes recruited into the sub-intima differentiate into macrophages. These macrophages undergo phagocytosis of ox-LDL via non-autoregulated scavenger receptors to form foam cells.
ECs, SMCs and foam cells release matrix metalloproteinases; propagating the development of fatty streaks. Fatty streaks are the earliest lesions in atherosclerosis composed mainly of a lipid layer of T cells, foam cells and over-proliferating SMCs beneath the endothelium. SMCs release a package of connective tissue matrix rich in collagen. Ultimately, fatty streak lesions advance into a new form of fibrous plaques due to persistent localisation of lipids, accumulation of SMCs and connective tissue. These fibrous plaques with lipid-rich cores contain largely living and/or dead foam cells. As these fibrous plaques expand and protrude into the lumen of vessels, they can eventually cause the onset of acute clinical syndromes, for example cerebral or myocardial infarction. Such major consequences arise due to a variety of alterations in the atheroma and the lumen of the blood vessel including thrombosis, haemorrhage, ulceration, or calcification.

In an atherogenic environment a variety of pro-inflammatory mediators are released from endothelial cells and leukocytes, such as interleukin 1 (IL-1), interleukin 6 (IL-6), monocyte chemoattractant peptide 1 (MCP-1), platelet-derived growth factor (PDGF) and more specifically interferon gamma (IFN-γ). Chronic inflammation can weaken the fibrous cap, leading to its rupture and the development of an occlusive thrombus. Although it is currently unclear how chronic inflammation is maintained within atherosclerotic plaques, it is clear that endothelial dysfunction plays a pivotal role in atherogenesis.

1.5 Endothelial dysfunction

The maintenance of blood flow is mediated by a thin layer of ECs known as the endothelium which lines the interior surface of blood and lymphatic vessels. ECs act as a barrier between the circulation and rest of the tissues and are involved in immune responses. The endothelium modulates vascular tone by secreting multiple vasodilators including NO, endothelium derived hyperpolarizing factor (EDHF), prostacyclin or vasoconstrictive factors such as thromboxane (TXA₂), and endothelin-1 (ET-1). NO
mediates anti-thrombotic, anti-platelet and anti-inflammatory properties and reduced permeability effect. ECs form tight and adherens junctions that regulate cell signalling and modulate vascular homeostasis. In normal conditions endothelial cells are quiescent, close in contact, non-proliferative, anti-apoptotic, and retain tight control of permeability. Laminar flow and anti-inflammatory factors including adipokine and adiponectin are shown to downregulate the expression of adhesion molecules to preserve vascular homeostasis.

In response to harmful insults, denudation of the intact endothelial monolayer causes increased permeabilisation, leukocyte accumulation, lipid deposition and pro-inflammatory events that can lead to EC dysfunction. EC dysfunction is a hallmark of cardiovascular pathogenesis, such as atherosclerosis. ECs are in contrast elongated, highly mobile, proliferative and responsive to various growth factors and cytokines. Immune cells secret pro-inflammatory mediators such as IL-2, angiotensin II, leptin, IFN-γ and tumor necrosis factor-α (TNF-α) which propagates ECs to express adhesion molecules e.g. vascular cell adhesion molecule 1 (VCAM-1) and intracellular adhesion molecule 1 (ICAM-1). Disturbed flow is shown to upregulate adhesion molecules in atheroprone regions. These processes recruits leukocytes further, enhancing EC permeability and ultimately initiated a series of complex cascade events including platelet activation, growth factor release, lipid and ECMs accumulation, and potentially neointima formation. Multiple risk factors are associated with EC dysfunction such as oxidative stress, aging, dyslipidaemia, diabetes, smoking, hypertension, diabetes, low disturbed shear stress, viral/bacterial infection, ischaemia, and stent implantation.

1.6 Endothelial cell under oxidative stress

ROS are chemically reactive molecules that contain oxygen. Multiple ROS molecules with unpaired electrons are referred as free radicals such as superoxide anion \( \text{O}_2^\cdot^- \), hydroxyl radical \( \text{OH}^\cdot^- \), and lipid radicals. Other ROS including hydrogen peroxide \( \text{H}_2\text{O}_2 \), peroxynitrite \( \text{ONOO}^- \), and hypochlorous acid \( \text{HOCl} \) are not free radicals but
are oxidative agents that contribute to oxidative stress. The main forms of ROS insults in the human body are H$_2$O$_2$, O$_2$ $^{•−}$, *OH, and ONOO$^{−}$. Oxygen is used in all higher order biological systems and a single-electron reduction of oxygen yields O$_2$ $^{•−}$. The production of O$_2$ $^{•−}$ is mainly generated by the mitochondrial electron transport chain of complex I and III, cytochrome p450s, free iron (Fe$^{2+}$), lipooxygenase, cyclooxygenase, uncoupling of nitric oxide synthase (NOS), xanthine oxidase (XOD), and NOXs. Amongst all the sources of ROS, NOXs are the major producers throughout cardiovascular components. O$_2$ $^{•−}$ is converted to H$_2$O$_2$ spontaneously or by superoxide dismutase (SOD). SOD is ubiquitously expressed and deletion of SOD isoforms in animals are not viable. Catalase and glutathione peroxidase enzymes convert harmful H$_2$O$_2$ to H$_2$O. Other nonenzyme products such as β-carotene, ascorbic acid, and tocopherols are also known to act as antioxidants. An imbalance between ROS production and defective antioxidant defence mechanisms leads to oxidative stress associated with endothelial dysfunction.

In physiological conditions, ROS mediates various important cellular functions such as mediating signalling pathways, promoting cell growth/differentiation, maintaining vascular homeostasis, regulating hormone levels, strengthening synaptic plasticity, stimulating enzymes, maintaining chemical balance, defending against invading pathogen and inducing immune responses to pathogenic environment. As signalling molecules, ROS can regulate kinase pathways including Src, Ras, JAK2, Pyk2, phosphoinositide 3-kinase (PI3K), mitogen-activated protein kinase (MAPK) and block protein phosphatases. They can activate β γ-subunit of G proteins and phospholipases and alter intracellular cation concentrations by stimulating Ca$^{2+}$ and K$^{+}$ channels. They can also modulate transcription factor functions such as hypoxia-inducible factor-1 (HIF-1), nuclear factor erythroid 2-related factor 3 (Nrf3), activator protein-1 and NFκB. Overall, it is speculated that low levels of ROS mediate important reparative processes.

ROS-mediated damage to macromolecules (e.g. nucleic acids and protein) is one key manifestation of aging. In pathological conditions, high levels of ROS damages the endothelium lining and is a key contributor to atherosclerosis, dysfunctional
angiogenesis, ischemia-reperfusion damage, myocardial infarction, reduced vascular
tone and chronic inflammation. Oxidative stress can alter intracellular reduction-
oxidation homeostasis in ECs, upregulate monocyte chemoattractant protein 1 (MCP-
1), leptin, chemokine (C-X-C motif) ligand 8 (CXCL8) and P selectin, and accelerate
monocyte recruitment to endothelial cells via toll-like receptor 4 (TLR4). ROS
enhances ox-LDL sensitivity in endothelial cells by upregulating TLR2 expression.
Oscillatory shear stress-induced ROS can activate intercellular signalling cascades
leading to adhesion molecules expression and matrix metalloproteinase (MMP)
secretion through NOX. Overall, the contribution of ROS is currently under debate for
its harmful and/or protective effects in endothelial function.

1.7 Vessel growth

During embryonic development, de novo vessel formation is driven by the
assembly of mesoderm-derived endothelial precursors (angioblasts) that differentiate
into a vascular labyrinth (vasculogenesis). Subsequent vessel sprouting
(angiogenesis) orchestrates a network of vascular remodelling to develop arteries and
veins. Incorporation of pericytes and SMCs adjacent to nascent EC tubules
provides stability and governs blood perfusion (arteriogenesis).

During adult life, vessels remain quiescent and rarely form new vessels. However, mature ECs have the ability to sense and respond to various angiogenic
stimuli. The process of “angiogenesis” involves new vessels sprouting from pre-existing
vascular compartments. Stimulated and attracted by pro-angiogenic molecules (mainly
by vascular endothelial growth factor (VEGF)), ECs transform from a quiescent inactive
state and into an active/invasive state to protrude filopodia. The protruding tip cells
spearhead new sprouts and migrate towards pro-angiogenic cues. Alongside tip cells,
proliferating stalk cells develop a lumen to provide sprout elongation. Tip cells
propagate basement membrane development and the recruitment of mural cells to
stabilise new sprouts and initiate blood flow (Figure 2C). This is mainly driven by pro-angiogenic factors such as VEGF, HGF, PDGF-B, TGF-β, ephrin-B2, NOTCH and protease inhibitors to stimulate stable, mature and functioning vessels. Prior to novel vessel formation, arterial-venous identity is required.

After the formation of these vessels, there is the need to determine their arterial-venous identity. Notch signalling is known to be a regulator of cell fate. Regulators of arterial fate include VEGF and Ephrin B2 while regulators of venous fate include Ephrin B4 and Coup-TFII. Continuation of sprouting is halted once proangiogenic signals abate and their quiescent phalanx state is resumed. This is mainly contributed by tissue inhibitors of MMPs (TIMP), cytokines such as IL-4, IL-10, IL-12, thrombospondins-1 and 2, MMP2 c-terminal lytic fragments, endostatin, angiostatin, maspin, and TGFβ. Mechanisms of vessel growth are also contributed by circulating and/or resident stem/progenitor cells which is highlighted next in this PhD thesis.
Figure 2: Schematic diagram of vessel development.

(A) Angioblasts differentiate into ECs to form tube-like structures with arterial or venous phenotypes. (B) Stages of vessel spouting: (1) tip/stalk cell identification; (2) tip cell migration and stalk cell proliferation; (3) branching organisation; (4) stalk elongation and lumen formation; and (5) perfusion and vessel maturation. (C) Progressive stages of vascular remodelling from an immature (left box) and towards a direction of stable and complete vascular plexus (right box) including adoption of a quiescent endothelial phalanx phenotype, basement membrane deposition, pericyte coverage, and branch regression. (Figure Source: Potente M, Gerhardt H, Carmeliet P. Basic and therapeutic aspects of angiogenesis. 2011. Cell Review. 146:6. 873-887 41).
1.8 Stem and vascular progenitor cells

In the past four decades, regenerative medicine in stem cells has been identified as a promising therapeutic area for various diseases, including CVD. Stem cells are a group of specialised cells capable of developing into any mature functioning cell type with the capability to proliferate, self-renew and regenerate tissues. Two specific hallmarks; self-renewal and potency, ultimately define stem cells from other cell types.

1. Self-renewal involves the maintenance of undifferentiated state through multiple cell division cycles.
2. Potency involves the capability of totipotent or pluripotent cells to differentiate to tissue/organ-specific mature cells with specialised functions induced specifically by the microenvironment.

Totipotency is the capability of one cell to divide and produce all differentiated cells in an organism (Figure 3). Totipotent stem cells can differentiate into both embryonic and extraembryonic cell types. Oocyte fertilisation by a spermatozoon forms a single totipotent cell, a zygote, which initiates a series of cell divisions to generate the morula (32-64 totipotent cells). Around 4-5 days post-fertilisation, these totipotent zygote cells develop into pluripotent stem cells to form the blastocyst. Peripheral cells (the trophoblast) of the blastocyst generate the embryonic membranes and placenta, whereas the inner cell mass develops into the foetus. Pluripotent stem cells have the potential to differentiate into any of the three germ layers (endoderm, mesoderm and ectoderm) to form the gastrula from which the complete organism develops. Multipotent stem cells can differentiate into several distinct cellular lineages of systems.
1.8.1 Embryonic stem cells

Pluripotent embryonic stem cells (ESCs), derived from the inner mass of blastocysts, have been extensively used in regenerative biology \(^{45,46}\). Cultured ESCs are immortal and can be maintained in an undifferentiated state with normal chromosomal composition. Molecular identification of ESCs is relatively well-documented. ESCs are known to express specific surface markers including alkaline phosphatase, CD9, and CD24, and a few genes associated with pluripotency such as Oct-4, Rex-1, SOX-2, Nanog, LIN28, CD90, and SSEA-3 and -4 \(^47\). The immortality feature \textit{in-vitro} maybe due to a high level of telomerase expression. Utilising ESC in research can be problematic as they can spontaneously differentiate into different cell types and lose their pluripotent state in long-term culture, it is also challenging to develop effective conditions to specifically differentiate ESCs into desired mature functioning cell types \(^48\).
Mouse ESCs are maintained in undifferentiated state by leukaemia inhibitory factor (LIF) and/or on a feeder layer of murine embryonic fibroblasts (MEF)\textsuperscript{49}. LIF is mainly produced by feeder cells, mediating its effect by activating gp130 receptors. This in turn activates transcription factor STAT3 (signal transducer and activator of transcription 3) that enhances gene expression linked to proliferative state of ESCs\textsuperscript{50}. SHP-2, a tyrosine phosphatase, interacts with the intracellular domain of the gp130 receptor and stimulates an intracellular cascade linked to self-renewal. In the absence of LIF, extracellular regulated kinase (ERK) is activated and synergistically with SHP-2, antagonise the proliferative effect of STAT3 (Figure 4). Therefore, ES cellular renewal is attenuated by the ERK and SHP-2 activated complex\textsuperscript{51}.

**Figure 4:** LIF-STAT3 signalling pathway linked with self-renewal and proliferation
Figure source (https://stemcells.nih.gov/info/2001report/appendixB.htm).
Many studies including previous work from our laboratory group have successfully driven ESCs to differentiate into specific cell type lineages, such as EC and SMCs in various conditions. The removal of LIF and supplementation of ESC culture medium with VEGF can guide cells to differentiate into EC \(^5^2\); while treatment of ESCs with PDGF cultured on collagen-IV coated flasks can lead to SMC-differentiated cell types \(^5^3\). SMCs and ECs are the essential cellular compartments of the vasculature and the use of ESC-derived SMCs and ECs can play a crucial role in further understanding physiological and pathological processes in the cardiovascular system.

### 1.8.2 Adult stem cells

Adult stem cells are rare, quiescent with a more limited self-renewal and differentiation capacity to continuously replenish all damaged/dying cells. These adult stem cells either self-renewal to main undifferentiated state or initiate a differentiation process to develop into a functioning tissue-specific cell type. In contrast to ESCs or induced pluripotent stem cells, utilisation of adult stem cells requires fewer ethical legislative hurdles and they are known to be less tumorigenic. The localisation of adult stem cells in the body is generally unclear, however some specific compartments, including hematopoietic, bone marrow, epithelial, muscular, and neural, consist of intrinsic stem cell characteristics that are better defined \(^5^4\). Hematopoietic stem cells have been extensively used in bone marrow and cord blood transplantation. Mesenchymal stem cells (MSCs) are of stromal origin and can be isolated from majority of tissues, which indicates the presence of a perivascular niche \(^5^5\). Adult-specific stem cell ‘niches’ are rare anatomic locations and therefore difficult to isolate. Some protocols enable stem/progenitor enrichment to a certain level of purity. For instance, human hematopoietic stem cells via the bone marrow are isolated and sorted for a CD34\(^+\) CD133\(^+\) CD38\(^-\) lineage population. In a normal physiological state, adult stem cells remain in a quiescent state within a niche. But during microenvironmental stimulation by various factors in response to tissue generation, maintenance and repair, adult stem cells undergo proliferation, migration and differentiation \(^5^6\).
MSCs are located in various different tissues and are able to differentiate towards mesenchymal lineage such as adipocytes, osteoblasts, myocytes and chondrocytes, but also towards other lineages including neurons, ECs and SMCs with specific micro-environmental stimulus. MSCs are relatively easy to isolate and grow in culture, compared to endothelial or other resident progenitor cells. Specific characterisation of MSCs are usually positive for CD73, CD90, CD105 and negative for CD45, CD34, CD14 or CD11b, CD79α or CD19 and HLA-D. MSCs can contribute to vascular regeneration by stimulating vasculogenesis, angiogenesis and inflammation, and thus restore function in damaged vasculature. MSC-mediated SMC differentiation is reported to be modulated by ligands such as Thrombin, TGFβ, S1P, and BMP. The activation of the intracellular RhoA-ROCK pathway followed by the formation of myocardin-related transcription factors (MRTFs) can initiate SMC gene expression via the association of SRF-CArG element. Additionally, microRNA-143/145, 1, 21 and 10a have been illustrated to play a role in MSC-derived SMC differentiation. MSC-induced EC differentiation is generally governed by VEGF, shear stress and insulin-like growth factor (IGF) and its associated signalling pathways to ultimately enhance EC expression markers. Multiple clinical trials have used MSC therapy in patients with myocardial infarction, peripheral artery diseases and other CVD-related illness. Results indicated safe usage of MSCs with differences in functional improvements.

### 1.8.3 Vascular stem/progenitor cells

For over 30 years, it has been known that ECs have substantial capacity for replication throughout the vasculature. Proposed reasons for EC turnover during vasculogenesis, angiogenesis and intima repair include three processes, (1) the ability for mature ECs to re-enter cell cycle via a dedifferentiation mechanism, (2) the mobilisation of circulating endothelial progenitor cells (EPCs) from remote locations via the circulation, and (3) the presence of lineage-specific EPCs or multi-lineage stem cells in the vessel wall.
Researchers have extensively examined a plethora of vascular progenitor cell types with the potential to develop into functioning ECs. The first isolation EPCs were identified by magnetic beads specific for CD34 and vascular endothelial growth factor receptor 2 (VEGFR2) positive cells. These cells were able to proliferate and differentiate into the mature EC-shaped cells expressing ECs markers such as CD31, E-selectin and eNOS. It was also originally thought that uptake of acetylated-LDL (ac-LDL), binding with ulex europaeus agglutinin-1 (UEA-1) and differentiation potential to cells expressing ECs markers are specific characteristics of ESPCs. Later studies identified various in vitro methods to improve the development of pure ESPCs such as colony-forming assays. This involves culturing cells on fibronectin-coated dishes with different stimulant compounds and observing adhering central rounded cells surrounded by thin flat cells. By utilising specific EC markers, these cells were referred as the colony forming unit-Hill (CFU-Hill) EPCs. It is important to note that hematopoietic monocyte and/or immune cells could mimic CFU-Hill EPCs characteristics.

Technical difficulties in recapitulating novel blood vessel growth and the low proliferative nature of EPCs in vivo, introduces huge obstacles for researchers. Later on, another population named as endothelial colony-forming cells (ECFCs) was presented. Under this method, the peripheral cells were cultured on collagen I coated plate with suitable culture medium. Under this condition for approximately three weeks, adhering cells begin to express EC markers and form a heterogeneous population with hierarchy proliferating cells in vitro. Moreover, ECFCs were able to mimic capillary-like EC characteristics both in vitro and in vivo. It has been difficult to characterise EPCs due to complications with specificity, however to date they are identified by various stem/progenitor markers such as Sca-1, CD34, c-kit, and CD133. To date, ECs are generally identified by several markers including CD31, CD144, tie-2, CD105, CD106, vWF, eNOS, VEGFR1-2 and EphrinB2.
1.8.3.1 Circulating stem/progenitor cells

The precise source of circulating EPCs is currently disputed. The bone marrow was the first site proposed to harbour stem/progenitor cells. Within the bone marrow niche, EPCs homeostasis is under the influence of various cytokines (e.g. SDF1, SCF, VEGF and Ang-1) secreted by neighbouring cells such as pericytes and ECs in a paracrine manner. When these secreted cytokines activate their corresponding receptors (c-kit, Notch, VEGFR and CXCR4), EPCs are programmed to migrate, proliferate or differentiate in response to specific signalling pathways. During pathological conditions such as ischaemia/hypoxia, secreted HIF-1 upregulates VEGF and Ang-1 expression within the bone marrow microenvironment. Stromal cells within the bone marrow are sensitive to alterations in VEGF levels and once the VEGFR/Akt-dependent signalling pathway is stimulated, eNOS is phosphorylated and activated. NO production in turn yields s-nitrosylation of MMP9 and activated MMP9 alters the quiescent state of the bone marrow niche into an active site releasing soluble SCF. SCF can mobilise c-kit+ EPCs to the circulation via the bone marrow.

Other studies have suggested the role of α4-integrin, granulocyte-colony stimulating factor (G-CSF), granulocyte monocyte-colony stimulating factor (GM-CSF), erythropoietin (EPO), HMG-CoA reductase inhibitors (statins) and exercise in the mobilisation of bone marrow-derived EPCs into the peripheral blood. The precise mechanism of mobilisation from these factors are unclear. Some of these molecules are well known to augment the PI3K/Akt-pathway and therefore may share some common signalling pathways. One may speculate the essential influence of eNOS in the increase of progenitor cell mobilisation by PI3K/Akt-dependent activation within the bone marrow stromal cells. As expected, exercise and VEGF-stimulated EPC mobilization was deteriorated in eNOS−/− mice.
1.8.3.2 Resident stem/progenitor cells

Other important sources of EPCs including the vessel wall, muscle, adipose tissue, dermis, intestine and liver are known to be involved in blood vessel formation\textsuperscript{64}. Among all these tissues, the vascular wall is the most rational and least energy/time consumption tissue to harbour EPCs as summarised in Figure 5. Local cluster of EPCs with the ability to develop into mature functioning ECs are defined as resident EPCs. The inner adventitia, (adjacent to the external elastic lamina), is speculated to reside the vascular stem cell niche. This region contains marker for stemness (Sca-1, Notch-1, CD34, c-kit, Oct-4, Stro-1, and FLK-1) \textsuperscript{71}. The resident stem/progenitor cells are committed to endothelial, smooth muscle, and macrophage lineage. Among the vessel wall, multiple populations of stem/progenitor cell types exist with potentially similar ancestry. It has been challenging to compare different subsets of vascular wall progenitor cells in different studies due to differences in species, isolation techniques, assays used, diversity in the vasculature, and the utilisation of nonspecific surface markers to locate these populations. For example, Sca-1\textsuperscript{+} (mouse) and CD34\textsuperscript{+} (human, mouse) adventitia cells consist of a combination of progenitor and mature cell types \textsuperscript{72}. Multipotent MSC-like cells are identified as both CD34\textsuperscript{+} and CD34\textsuperscript{−}, whilst CD34 is associated with adventitial EPCs. Even though Sca-1 and CD34 markers are expressed on adventitial progenitor’s \textit{in-situ}, they are reduced with culture due to expansion of MSCs. There is also co-expression of Sca-1, CD34 and c-kit in murine adventitia, indicating that their combined use could delineate subpopulations of progenitors with greater specificity and this could be further improved by the introduction of other lineage-specific markers such as VEGFR\textsubscript{2}, CD133, CD31, CD115, CX\textsubscript{3}CR1 and PDGFR-\textbeta as shown in figure 5. Overall, further investigations are needed to further characterise vascular wall progenitor cells and elucidate their specific hierarchical organisations with addition to mechanisms linked with the activation, differentiation and mobilisation in their neighbouring surroundings.
Besides the complexity of resident vascular PCs, a subpopulation of quiescent vascular resident cluster of CD31^+CD45^- cells was observed to self-expand and form colonies *in-vitro*\(^73\). Transplantation of these cells in ischaemic mice resulted in *de novo* blood vessel formation\(^73\). A distinct niche located in between the media and adventitia later is named at the vasculogenesis zone. Within this region, another resident PC subpopulation (CD34^+, VEGFR2^+, TIE2^+, CD31^- and CD45^-) was identified. This population was also able to produce de novo vessels by sprouting and form capillary sprouts\(^74\).

The three mural layers within the coronary arteries were documented to possess proliferating c-kit^+ VEGFR2^+ EPCs with the ability to connect with other supporting cells. Upon culture, followed by transplantation of these EPCs in a stenosis model of immunosuppressant dogs, vasculogenesis was seen to form human coronary arteries and there was a functional improvement of the ischemic myocardium\(^75\). It is important to note that the contribution of new blood vessel growth is not entirely contributed by resident PCs as other pro-angiogenic cells (for example, HS/PCs, fibroblast, MSCs and pericytes) and circulating bone-marrow PCs can also contribute to this phenomenon. In summary, these findings suggest that the adventitia maintains multiple types of progenitor cells that coordinates healing response to vascular injury. The progenitor cells may also contribute to vascular disease development by differentiating into proliferating smooth muscle cells or pro-inflammatory macrophages
Figure 5: Schematic illustration of vascular wall resident stem/progenitor cells

Stem/Progenitor cells (SPCs) are distributed across all of three vessel walls. EPCs mainly located within the intima and vasculogenesis zone. In the media layer, MSCs are detected along with side population of progenitor cells. Adventitia harbours most abundant populations of SPCs, including adventitia SPCs, EPs, HSCs, MSCs and perivascular SPCs. All of these populations of SPCs can be recruited under vascular signal and involve in vascular biology. SPCs, stem/progenitor cells; EPCs, endothelial progenitor cells; MSCs, mesenchymal stem cells; HSCs, hematopoietic stem cells; AMPCs, adventitial macrophage progenitor cells. (Figure Source: Psaltis PJ, Simari RD. Vascular wall progenitor cells in health and disease. Circulation Research. 2015. 116(8):1392-412).
1.8.4 Stem/progenitor cell proliferation

Active dividing eukaryotic cells undergo a series of stages referred as the cell cycle. This is known by two gap phases (G₁ and G₂) and a synthesis (S) phase, in which the genetic material is duplicated via DNA replication; and an M phase, in which mitosis partitions the genetic material and the cell divides. The G₁ phase involves metabolic alterations to prepare for cell division and in specific conditions, the cell is committed to division and transfers into the S phase. During the S phase, DNA replication results in the duplication of the genome and each chromosome contains two sister chromatids. Then at the G₂ phase, metabolic changes orchestrate cytosol molecules essential for mitosis and cytokinesis. At M phase, a nuclear division (mitosis) followed by a cell division (cytokinesis) occurs. Mitosis, although a continuous process, is conservatively separated into 5 steps: prophase, prometaphase, metaphase, anaphase and telophase. After cytokinesis whereby the cytoplasm becomes constricted, cells become quiescent at the resting G₀ phase or re-enter the cell cycle at G₁ phase.

During DNA replication, a family of around 50 proteins assemble into a replication fork to recognise DNA replication origins, load the replicative helicase on DNA, unwind DNA, produce new DNA strands, and reorganise the chromatin ⁷⁶. The process of DNA replication initiates from the G₁ phase at replication origin sites, several hours before their activation in S phase. This is a highly regulated process with specific checkpoint controls to overcome problems and safeguard the genome from damage. Defective maintenance of genomic duplication can lead to ‘replicative stress’ (RS), which is linked with various different developmental diseases and cancers ⁷⁷. The following figure 6 outlines the complex DNA replication process found in eukaryotic cells characterised by the electron microscopy over the past few decades (Figure 6):
Figure 6: Illustration of eukaryotic initiation of DNA replication. 
(A) Conserved DNA sequences contain a replication origin site for the binding of the origin recognition complex (ORC). (B) Cdc6 binds to an ORC/DNA complex, initiating the formation of a replication fork in S phase. (C) This recruits the Cdt1/minichromosome maintenance 2–7 (MCM2–7) heptamer and the loading of MCM2–7 on dsDNA resulting in the development of the ORC/Cdc6/Cdt1/MCM2–7 (OCCM) complex. (D) Cdt1 and Cdc6 dissociates from the OCCM in an ATP hydrolysis-dependent manner, which promotes the closure of the MCM2–7 ring and links with another Cdc6, resulting in the ORC/Cdc6/MCM2–7 (OCM) complex formation. (E) The OCM, vital for the loading reaction, recruits a second MCM2–7/Cdt1 heptamer forming a double hexamer (DH) that encircles dsDNA via an unknown mechanism. DH formation triggers spontaneous release of Cdc6, ORC and Cdt1 and the closure of the second MCM2–7 ring around dsDNA to form a stable complex. (F) During the S phase, preinitiation DNA replication occurs on a Dbf4-dependent kinase (DDK)-dependent phosphorylation of the DH complex and association of various different factors. This ultimately forms the important replicative helicase: the Cdc45/MCM2–7/GINS (CMG) complex, to initiate DNA-unwinding. (G) During 3′→5′ DNA-unwinding reaction, the CMG complex binds with both polymerases ε and α into a replication fork (RFK) to produce the leading and lagging strands. This DNA synthesis process is triggered at thousands of replication origins in order to completely replicate the enormous eukaryotic genome. (Figure Source: Riera A, et al. From structure to mechanism—understanding initiation of DNA replication. Genes Dev. 2017. 31(11): 1073–108876).
As stem/progenitor cells differentiate their rate of proliferation usually decreases. Most differentiated cells in adult animals are arrested in the G0 stage of the cell cycle or continuously divide in a highly regulated homeostatic fashion for natural turnover and to replace injured cells. Certain differentiated cells, however, have limited proliferate capacity including cardiac and neuronal cells.

1.8.5 Stem/progenitor cell differentiation

The differentiation and production of vascular cells from stem cells is a complex multifactorial process orchestrated by various different external stimuli’s, signalling pathways, and transcriptional and post-transcriptional regulations. This section will focus on the specific differentiation processes mediated by the microenvironment including various growth factors or cytokines, extracellular matrix and mechanical forces. Table 1 on the next page illustrates up-to-date key findings of EC differentiation from ESCs using different experimental parameters, delineating its mechanisms involved via signalling pathways, and characterising expression/functional EC phenotype.
Table 1: Models of mouse endothelial cell differentiation from embryonic stem cells

<table>
<thead>
<tr>
<th>Mouse cells</th>
<th>Models of EC differentiation</th>
<th>Characterisation</th>
<th>Signalling pathways/molecules</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESC/EC</td>
<td>Collagen IV/Sca-1* cell sorting/differentiation media (DM)/VEGF</td>
<td>EC differentiation markers</td>
<td>HDAC3</td>
<td>52</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>Collagen IV/DM/Shear stress/VEGF</td>
<td>EC differentiation markers</td>
<td>Flik-1-PI3K-Akt-HDAC3-p53-p21</td>
<td>79</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>Embryoid body (EB) formation</td>
<td>EC differentiation markers</td>
<td>Wnt2</td>
<td>80</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>EB formation</td>
<td>EC/SMC differentiation markers</td>
<td>Wilms’ tumor-1 (Wt1)-Snail E-cadherin</td>
<td>81</td>
</tr>
<tr>
<td>ES/EC</td>
<td>VEGFR2* isolation/shear stress</td>
<td>EphrinB2</td>
<td>VEGF-Notch signalling pathway</td>
<td>82</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>VEGFR2* cell isolation</td>
<td>EC differentiation markers</td>
<td>VEGFR2-phospholipase Cy1</td>
<td>83</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>VEGFR2* cell isolation</td>
<td>EC differentiation markers</td>
<td>H-Ras/Erk</td>
<td>84</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>VEGFR2*/VEGF co-receptor neuropilin-1 (NRP-1*) cells</td>
<td>CD34*/CD31*</td>
<td>VEGF Semaphorin binding activities of NRP-1</td>
<td>85</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>CD144 promoter-GFP driven</td>
<td>EC differentiation markers</td>
<td>-</td>
<td>86</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>EB formation</td>
<td>CD31 or VEGFR2</td>
<td>Shb-VEGFR2/PDGF-β</td>
<td>87</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>VEGFR2*/shear stress</td>
<td>EC differentiation markers</td>
<td>VEGF-independent Flk-1 phosphorylation</td>
<td>88</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>EB formation and angiogenic growth factor mixture (EPO, IL-6, FGF2, and VEGF)</td>
<td>VEGFR2, CD31 Tie2 CD144 and Tie1</td>
<td>-</td>
<td>89</td>
</tr>
<tr>
<td>ESC and arterial EC (aEC)</td>
<td>Collagen IV-coated/DM/Hypoxia</td>
<td>Arteriogenesis and revascularization Improved cardiac function</td>
<td>HIF1α/Etv2/NOTCH1 signalling</td>
<td>90</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>FGF2, VEGF, and BMP4</td>
<td>EC differentiation markers</td>
<td>MAPK and PI3K pathway ERG and FLI1</td>
<td>91</td>
</tr>
<tr>
<td>ESC/EC</td>
<td>Collagen IV/Flk-1 sorting/VEGF H3K4me3 and H3K27me3 modifications</td>
<td>EC differentiation markers</td>
<td>ETV2, Gata2, Flk1, Sox7 and Sox18</td>
<td>92</td>
</tr>
</tbody>
</table>
1.8.5.1 Growth factors

Signalling mediated by the hedgehog family of secreted proteins is known to be pivotal in embryonic patterning, morphogenesis and vascular formation. Bone morphogenetic proteins (BMPs) function downstream of hedgehog signalling and are important for embryonic vascular development and stimulating bone formation. The WNT proteins are secreted glycoproteins that govern embryonic patterning, cell fate specification, survival, and overall organogenesis. Three Wnt signalling pathways have three distinct categories: the canonical Wnt pathway, the noncanonical planar cell polarity pathway, and the noncanonical WNT/calcium pathway. The canonical WNT signalling is more widely known to be involved in vascular development. β-catenin, downstream moderator of canonical WNT signalling, drives EC gene expression and enhances EC differentiation.

PDGF is a potent mitogen for cells of mesenchymal origin, such as SMCs. PDGF signalling, initiated by its ligands, PDGF-A, B, C and D and its corresponding receptors PDGFR-α and −β, is a mediator of SMC migration and proliferation. PDGF family of proteins have been implicated in wound healing, neointima formation and atherogenesis, and specific deletion of PGDG-B is known to be embryonically lethal due to severe failure in the cardiovascular system. As expected, in-vivo inhibition of PDGF or its receptor has been illustrated to decrease neointima hyperplasia in response to injury.

The TGF-β secreted superfamily which includes inhibins, activin, anti-müllerian hormone, BMPs, decapentaplegic and Vg-1, initiates its diverse cellular responses by stimulating various different TGF-β receptor isoforms and complex signalling pathways. TGF-β and their family members, such as TGF-β1, Nodal, activins, BMP-2, 4, 6 and 7, are expressed by vascular cells, macrophages, lymphocytes, and have been associated with the differentiation of stem/progenitor cell types to various cell lineages, including VSMCs and ECs. Signal transduction mediated by TGF-β subfamily towards the nucleus
is mainly orchestrated in a complex manner of phosphorylating different SMAD 1-8 proteins. Contradicting studies display differences in endothelial marker expression due to different models, strains and other parameters used. VEGF ligands VEGF-A, -B, -C, -D, -E and placental growth factor (PIGF), and their receptors VEGFR-1 or Flt-1, VEGFR-2 or Flk-1/KDR, and VEGFR-3, exert distinct biological functions. VEGFR-1 and VEGFR-2 are mainly expressed on vascular endothelial cells and VEGFR-3 is mainly found on lymphatic endothelial cells. VEGF-A binds to the semaphorin co-receptor, neuropilin-1 (NRP-1), in a complex which promotes VEGF binding to VEGFR2. NP1 and NP2 are transmembrane tyrosine kinase co-receptors which are vital for angiogenesis, arteriogenesis and permeability. Active/modified VEGF signal transducers, such as PI3 kinase, serine/threonine kinase Akt, protein kinase B (PKB), protein kinase C (PKC), protein kinase A (PKA), phospholipase C-γ (PLC-γ), Scr family tyrosine kinases, Ras GTPase, cyclic adenosine monophosphate (cAMP), extracellular signal-regulated kinase (ERK) and p38 mitogen-activated protein kinase, are all reported to be associated with EC differentiation, proliferation, migration, cell survival, permeability and angiogenesis. Akt/PKB pathway is essential for EC survival and PLC-γ-induced activation of ERK is pivotal in EC proliferation downstream to VEGF. Overall, VEGF signalling is fundamental in vascular development and endothelial functions.

Adhesions molecules, integrins and cadherins, can co-regulate intracellular cascade pathways of growth factor receptors, such as VEGF receptors. VEGFR activation leads to the phosphorylation of vascular endothelial-cadherin (VE-cadherin) and associated catenins, and this pathway can lead to enhancement of permeability, alteration in cytoskeletal organisation and potentially determine apoptotic fate within endothelial cells. Angiopoietin growth factors and its corresponding endothelial Tie receptors modulate vascular development and permeability, inflammation, angiogenic remodelling and tumour vascularization. Angiopoietin signalling forms complexes at
endothelial cell-cell and cell-matrix contacts associated with physiological and pathological angiogenesis processes \textsuperscript{110}.

\textbf{1.8.5.2 Shear stress}

Mechanical force in the form of laminar and disturbed shear stress is known to be involved in stem/progenitor cell differentiation towards specific cell lineages. It is proposed that mechanical force generative by blood flow in mesenchymal progenitor cell \textsuperscript{111}, EPC \textsuperscript{112} and ESC \textsuperscript{88} can induce differentiation towards EC lineage. In addition, shear stress induced-EPCs from human peripheral blood and mouse ESC-derived Flt\textsuperscript{+1} cells resulted in a larger population of cells in the S and G\textsubscript{2}-M phases of the cell cycle and thus increased cell density when compared to static conditions. Shear stress also upregulated EC markers including Flk-1, Flt-1, VE-cadherin, and PECAM-1 at both mRNA and protein level \textsuperscript{88}. Another report revealed that shear stress increased lysine acetylation of histone H3 at position 14 (K14), serine phosphorylation at position 10 (S10) and lysine methylation at position 79 (K79), thereby enhancing EC expression. This result provides another molecular basis of shear-stress stimulating EC differentiation from ESCs \textsuperscript{113}.

\textbf{1.8.5.2 Extracellular matrix}

The Extracellular matrix (ECM) provides a complex and stable platform around cells within all mammalian tissues. It provides this structural integrity by influencing cell adherence, growth, migration and apoptosis \textsuperscript{114}. The ECM also possesses a key role in stem/progenitor cell differentiation functions \textsuperscript{115}. Three major forms of ECM exist: specialised proteins (fibronectin, laminin, and fibrillin), structural proteins (collagen and elastin), and proteoglycans.
Collagens are ubiquitously found in the body, form the major basis of the ECM and consist of 12 different isoforms. Collagen types I, II and III are the most abundant and form fibrils of similar structures. Collagen type I is abundant in mainly scar tissues such as in the bone, cornea, dermis and tendon. A study illustrates the role of collagen type I in endoderm lineage differentiation in embryonic stem cells. Collagen type IV is a major component of the basal lamina (a layer adjacent to the epithelium and produced by the epithelial cells). Numerous investigations reveal the role of collagen IV in differentiation process towards EC lineage in stem/progenitor cells as highlighted in table 1. How extracellular collagens (specifically type IV) mediate EC differentiation in stem/progenitor cells is still unexplored.

Additionally, fibronectin has been shown to enhance CD34+ cells to form endothelial colonies and improve VEGF-mediated CD34+ cell migration. Later on, the molecular pathway of integrin α5 and β1 stimulation towards VEGFR3 activation and its downstream PI3 kinase/Akt signalling pathway was shown to be involved in fibronectin-mediated lymphatic endothelial cell survival and proliferation. This suggests that the components of the ECM have the capability to have a role in cellular differentiation, proliferation and survival.

**1.8.5.3 Intracellular modulators**

Previous investigations documented the link between ROS and VSMC proliferation, migration and differentiation with ROS/p38 MAPK-dependent upregulation of SRF-induced transcriptional stimulation of markers linked with VSMC differentiation. Different isoforms of NOX exist and Nox4-H2O2 has been shown to encourage VSMC differentiation from mouse ESCs. In ECs, different NADPH oxidase subunits have been identified (e.g. Nox2, p22phox, Rac1, p47phox as well as Nox1 and Nox4). Some reports have illustrated the role of NAPDH-producing ROS in EC differentiation from stem/progenitor cell types. ROS signalling is also seen to be
involved in in-vivo or in-vitro angiogenesis and vasculogenesis via the modulation of proliferation, migration and differentiation.

Both histone acetyltransferases (HATs) and histone deacetylases (HDACs) are vital modulators of chromatin infrastructure and function. HDACs are classified in four different groups depending on their sequence identity. Upon differentiation, ESCs alter complex gene-targeted and chromatin structure remodelling with the downregulation of HDACs in inhibiting differentiation of ESCs. HDAC7 and 8 have been reported to play a role in VSMC differentiation from different in-vitro and in-vivo models elucidating its mechanisms of action. HDAC3 is important for VEGF-A and shear stress-induced EC differentiation from mouse ESCs. Laminar flow stimulates HDAC3 via the Flk-1-PI3K-Akt pathway, subsequently deacetylating p53 and stimulating p21 which results in the inhibition of ESC proliferation and the enhancement of differentiation towards functional ECs. Later study confirmed that Sca1+ vascular progenitor cells isolated from differentiating ESCs could differentiate into functional ECs via activation of HDAC3.

E-twenty six (ETS) proteins are transcription factors that bind to a specific core 5′-GGA(A/T)-3′ sequence in EC associated genes. The forkhead (Fox) transcription factors, FoxC, FoxF, FoxH and Fox, have all been linked endothelial development and functions. GATA transcription factors were initially known to be involved in hematopoietic development, however more recently GATA is currently recognised as important mediators of endothelial differentiation.

Several miRNAs have been suggested to regulate vascular development, angiogenesis and endothelial functions via the fine-tuning of VEGF, Notch and Slit/Robo signalling pathways. Slit-Robo signalling governs axon repulsion within the developing neurons, however its functional implications in angiogenesis, cardiac and cancer development among other processes have also been documented. These miRNAs
include miR-126, miR-221, miR-132, miR-218, miR-23-27-24 clusters, miR-27a/b, and miR-92\textsuperscript{130}. Overall, further understanding of the precise endothelial differentiation mechanism will undoubtedly improve knowledge of vascular disease pathogenesis and enable us to regenerate vascular compartments in a clinical setting. A summary of major contributors for the progression of ESCs towards EC lineage is shown in Figure 7.
Figure 7: Illustration of mechanisms involved in EC differentiation from ESCs.

1) Upon VEGF-A and/or mechanical force environment, VEGFR2 stimulation and autophosphorylation, enhances PI3K-Akt signalling cascades and stabilises HDAC3. HDAC3 induced p53 deacetylation and p21 stimulation is associated to shear- and VEGF-induced EC differentiation. VEGFR2-PI3K signalling also enhances EC markers via Ras-ERK signalling pathways to target other genes such as Erg, Id1, Pim1, phospholipase Cγ1, Erk, wt1 and snail.

2) Notch receptor stimulation releases its intracellular domain of Notch (NICD) from the membrane and NICD forms a complex with [CBF1/RBPJ-kappa/Su (H)/Lag1] in the nucleus to activate various EC-related genes.

3) Wnt ligand activation of the canonical Wnt signalling via Frizzled receptor and LRP interaction and β-catenin and T-cell factor (TCF) protein complex in the nucleus, leads to the transcription of EC markers. The non-canonical Wnt pathway, Wnt5a-PLCγ-Ca^{2+}-PKC-Ras-ERK, also propagates ESCs towards EC lineage.

4) BMPs activates BMP receptor type 2 (BMPR2), autophosphorylation of BMPR1 and complex formation of SMADS 1, 4, 5 and 8, and consequently regulate EC gene expression. This pathway is also associated with TGFβ signalling cascades. Figure source: (Kane N et al, Pluripotent stem cell differentiation into vascular cells: A novel technology with promises for vascular regeneration. 2011. Pharmacology & Therapeutics. 129:1, 29-49) ^131.
1.9 Nested genes

Intrinsic genomic overlapping is widely shared among diverse organisms ranging from simple viruses to multi-cellular organisms. Genomic DNA sequences can encode more than one gene product, with the overlapping genes commonly situated on opposite strands. Overlapping genes can be orientated in a way where one gene is completely contained within a chromosomal region occupied by another gene. In this paradigm, the internal gene is defined as a “nested” gene. A nested gene is a gene whereby the entire coding sequence resides within the chromosomal region with start and stop codons of a larger external gene. Nested genes are different from alternatively spliced transcripts as the coding sequence for a nested gene is different from the coding sequence for its external host gene. One example is the distinct transcriptional start sites.

There are two types of nested genes: (i) intronic nested genes - genes situated within an intron of the external gene (Figure 8A) and (ii) non-intronic nested genes - genes situated completely opposite an exon of the external gene (Figure 8B). Intronic nested genes are found to be more common, specifically in eukaryotic genomes. On the other hand, non-intronic nested genes have been documented in rare cases in eukaryotes. Prokaryotic organisms exhibit a high level of genomic plasticity as around a third of its annotated genes are overlapped. Irrespective of the natural spatial arrangement, its unique biological ramifications with respect to evolutionary origins, regulations and functions remains to be established.
1.9.1 Nested intronic genes

Around 30 years ago, “Nested gene” description was initially used to define a unique gene rearrangement at the Gart locus in the Drosophila melanogaster genome. Within this locus, a gene known as adenosine 3 (ade3) was identified to exhibit 6 introns with intron 5 containing an open reading frame (ORF) of around 200 codons oriented antisense and opposite the ade3 coding strand. This ORF sequence was later found to encode a protein, now known as pupal cuticle protein (PCP). Encoding of Pcp and ade3 are from opposite DNA strands and therefore have no sequence homology between each other. This conserved Pcp nested gene among different species, is a typical characteristic example of nested intronic genes. Intronic nested genes predominately encode for functionally unrelated product(s) to their host genes.
For instance, Pcp forms the structural component of the pupal chitin-based cuticle, while ade3 encodes for proteins (phosphoribosylglycinamidase synthetase, phosphoribosylaminoimidazole synthetase, and phosphoribosylglycinamid transformylase) linked with the purine pathway. The expression profile of nested Pcp gene and host ade3 gene differ substantially: Pcp expression is specific in the epidermic during prepupal stage, whereas the ade3 gene is found during all stages of development. Nested intronic genes are usually not co-expressed with their external host genes, however in some cases this rule is not applied.

Over the past four decades, more and more reports have shown the existence of nested intronic genes within the human genome. A report found a total of 158 protein-coding intronic nested genes in the human genome via the utilisation of the NCBI Map Viewer online database. These nested genes have been showed to express sequence tags and may have protein-encoding capability. Human nested intronic genes share similar patterns of gene organisation with the Drosophila Pcp gene. Of 106 human host genes and 96 matching nested genes with Gene Ontology (GO) annotations, only 5 parallel nested gene pairs and one anti-parallel nested gene pair display similar functions. Around 63% of human nested genes are identified on the opposite DNA strand to the host gene, exhibiting antiparallel pairs. The remaining 37% nested genes are orchestrated in a parallel fashion on the same DNA strand as the host gene. Nested human genes are usually found in the large introns of host genes, in a similar manner to Pcp which resides in the largest intron of ade3. The average length of an intron harbouring a nested gene is around 10-fold higher compared to another intron from the same host gene.

Several overlapped and nested genes have been associated with human pathologies. Within the chromosomal region 13q33, a schizophrenia susceptible locus was identified to contain genes G72 and G30. The G72 and G30 are a pair of overlapped genes transcribed in opposite direction. Genetic variants at the G72/G30 locus is linked
specific psychiatric phenotypes and diagnoses. Functional studies identified the role of G72 gene product in the stimulation of N-methyl-D-aspartate (NMDA) receptors, a molecular hallmark of schizophrenia. A nested intronic gene, Saitohin (STH), was discovered to be located in the intron between exons 9 and 10 of the human tau gene. STH gene encodes its 128-aa protein and its expression was similar to tau. The tau gene is a hallmark of many neurodegenerative diseases referred as tauopathies. This study identified a genetic polymorphism that altered glutamine to arginine at amino acid position 7 (Q7R) of STH. This specific mutation was highly correlated in the homozygous state in late onset Alzheimer’s disease patient samples. To date, there is no literature indicating the role of encoding nested genes in the cardiovascular physiology or pathology. However numerous micro RNAs and antisense transcripts from nested genes have been documented to play an important role in cardiovascular physiology.

1.10 LAF4 gene

Laf4/LAF4, also known as Af4/Fmr2 family member 3, is one of the four members of the AFF (AF4/FMR2) family of genes. This AFF gene family is considered as alternative pre-mRNA splicing modulators by binding to G-quadruplex RNA-forming structures. LAF4 gene maps to the mouse chromosome 1 and human chromosome 2q11.2-q12 respectively. Human LAF4 contains high homology between the mouse and chicken, illustrating high conservation during vertebrate evolution. The following page illustrates the expression profile of human mRNA LAF4 in figure 9 (data obtained by the GTEX website: https://gtexportal.org/home/gene/AFF3).
Figure 9: Illustration of human LAF4 mRNA expression in different tissues
The global knockout model of LAF4 by the Wellcome Trust Sanger Institute in Cambridge was developed relatively recently. Several phenotypic observations from homozygote \( (\text{Aff}3^{tm1a(EUCOMM)Wtsi}/\text{Aff}3^{tm1a(EUCOMM)Wtsi}) \) and heterozygote \( (\text{Aff}3^{tm1a(EUCOMM)Wtsi}/\text{Aff}3^+ \) mice were documented. These features are listed below in Table 2.

<table>
<thead>
<tr>
<th>Affected areas</th>
<th>LAF4(^{-/-}) Female</th>
<th>LAF4(^{-/-}) Male</th>
<th>LAF4(^{-/-}) Female</th>
<th>LAF4(^{-/-}) Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behaviour/neurological</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Craniofacial</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal incisor morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal snout morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Growth/size/body</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal incisor morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal snout morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hearing/vestibular/ear</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Increased/absent threshold for auditory</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Brainstem response</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Homeostasis/metabolism</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reduced circulating fructosamine level</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Circulating serum albumin</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Limbs/digits/tail</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reduced caudal vertebrae number</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Skeleton</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal incisor morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reduced caudal vertebrae number</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal vertebral arch morphology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fusion of vertebral arches</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Vertebal transformation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2: Phenotypic observation from Laf4 knockout in-vivo model

Table illustrates phenotypic observation from the mouse genome website (http://www.informatics.jax.org/allele/MGI:4434291). Genotype ID: Homozygote (MGI:5608405) and heterozygote (MGI:5705951).

Mouse LAF4 mRNA was identified at high levels in lymphoid tissues and at a lower level in the brain and lung. LAF4 is also reported to be expressed in mouse ESCs \(^{149, 150}\). In human and mouse lymphoid cell lines, mRNA LAF4 levels were highest in
mature B cells, pre-B cells and absent in plasma cells, indicating a regulatory role during lymphoid development. Further detailed investigation found that the mouse LAF4 protein, with an amino acid sequence of 1227, is a nuclear-specific transcription factor involved in early lymphoid development.148

LAF4 mouse protein can also act on a neurodevelopmental structural protein referred as MAM Domain Containing Glycosylphosphatidylinositol Anchor 2 (MDGA2) and modulate cortical neuronal migration.151 A method using cDNA subtraction based on mirror-orientation selection, enabled investigators to identify alternative spliced forms of mouse LAF4 with diverse protein transactivation functions. Two variants were identified to be predominately expressed during embryogenesis and other forms during adult life. Transcriptional activation of LAF4 was found to be in line with initial stages of immature cortical cell type differentiation and downregulated during mature cortical neuron development stages.152 This indicates that LAF4 is involved in intellectual disability in the mouse nervous system. There is currently no literature of its implications in the CVS.

Various pathological studies illustrate that human LAF4 gene can fuse to mixed lineage leukaemia (MLL) via gene translocation to express a lymphoid nuclear protein linked with acute lymphoblastic leukaemia (ALL) in infants.153-155 A genetic study on Nievergelt syndrome patients with abnormal limb, brain and urogenital development has been found to have haploinsufficiency for the LAF4 gene.156 A microdeletion of 500kb on the human chromosome 2q11.1 encompassing LAF4 gene was reported to be the cause of the clinical manifestations and ultimately death at the age of 4 months.156 Additionally, downregulation of the LAF4 gene and its AF4/FMR2 gene family is generally associated with neurodevelopmental defects.157 Genome wide association studies (GWAS) revealed that LAF4 is associated with human high triglycerides,158 diabetic nephropathy,159 rheumatoid arthritis,160 and reproductive behaviour.161 These observations indicate that LAF4 gene may be involved in multiple physiological and
pathological processes. However, to date, there are no disorders of the CVS associated with the \textit{LAF4} gene.

\textbf{1.11 \textit{Laf4ir} novel gene}

Our laboratory group is mainly focused on further delineating the precise machinery driving stem/progenitor cells to differentiate towards EC lineage via shear stress. To further understand this underlying mechanism, our previous investigation utilised a microarray analysis [unpublished data] of gene expression profiling among mouse ESCs and differentiated ESCs subjected to 12 dynes/cm$^2$ laminar shear stress or kept at static condition.$^{79}$

Two cDNA clones attracted the most amount of interests, which were from a cDNA library constructed from the mouse aorta and vein (\textit{GenBank: AK040668.1})$^{162}$ and from the neonate heart (\textit{GenBank: AK085781})$^{163}$ mRNAs respectively. These two cDNA products were labelled as EC differentiation-related gene 1 and 2 (\textit{ECD1}, \textit{ECD2}). \textit{In situ} hybridization with chicken embryos revealed that \textit{ECD1} with an 871 bp cDNA fragment was only expressed in the peripheral blood vessels but not in the heart. \textit{ECD2} with an 2469 bp cDNA fragment was expressed in both heart and blood vessel (Figure 10A).
Figure 10: Characterisation of ECD1 and ECD2 genes.

Data is from previous work conducted in Dr Zeng’s group (unpublished data). (A) In situ hybridisation illustrated ECD1 and ECD2 expression in HH14 chicken embryos. Digoxigenin-labelled single DNA probe was incubated, followed by AP-conjugated antidigoxigenenin and NBT detection. H: Heart (B) ESCs were cultured on collagen-IV coated slides for three days in differentiation medium and then exposed to 12 dynes/cm² laminar flow for 24 hours, followed by quantitative RT-PCR analysis. Data presented were representative SEM of three independent experiments.*: p<0.05.

Blasting the ECD1 fragment to mouse genome revealed that it existed within the intron 6 of the mouse LAF4 gene, spanning two different regions. A homologous fragment was also found in the intron 7 of human LAF4 gene (AI438968 from B-cell of chronic lymphocytic leukemia, AA541378 from metastatic prostate lesion of the bone and BF751024 from normal breast tissue). The ECD1 871bp fragment might be derived from LAF4 precursor RNA or be separately transcribed in a similar fashion as the nested intronic gene Pcp and its external gene ade3 \(^{133}\).
To investigate whether \textit{LAF4} and \textit{ECD1} are the same gene, primers specific for both sequences were developed and quantitative reverse transcription polymerase chain reaction (qRT-PCR) was performed on RNA samples isolated from the differentiated mouse ESCs subjected to laminar shear stress or kept at static condition for 24hrs. Upon shear stress stimulation, expression of EC markers \textit{CD144} and \textit{CD31} were upregulated as expected. \textit{ECD1} and \textit{ECD2} were upregulated whilst \textit{LAF4 mRNA} level remained constant (illustrated in Figure 10B). This pilot experiment illustrates that both \textit{LAF4} and \textit{ECD1} are not the same gene. This also illustrates the likelihood of \textit{ECD1} and \textit{LAF4} being transcribed separately. Therefore, \textit{ECD1} was designated as \textit{LAF4} intron resident gene, abbreviated to \textit{Laf4ir}.

To investigate the potential expression of \textit{Laf4ir} in the mammalian system, a northern blot with a cDNA probe derived from the 871bp sequence fragment was conducted. 20 µg of RNA from different organs were subjected to northern blot with DIG Northern Starter kit (Sigma, UK) according to the protocol provided. The digoxigenin-labelled probe was prepared by PCR with primer set of 5’-gtg aga cac ctg cag aag ctg-3’ vs 5’-ctg cag ttt atg cac gta gttg-3’. A main band of around 2.0 kb (Figure 11) was detected in some mouse adult organs with the most abundant expression in the aorta. This primary study for the first time indicated the expression capability of \textit{Laf4ir} in different mouse organs.
Figure 11: Northern blot analysis of Laf4ir in different organs.
Data is from previous work conducted in Dr Zeng’s group (unpublished data). Detection of around 2kb band from digoxigenin-labelled cDNA probe from the 871bp fragment of Laf4ir in 20µg total RNA isolated from adult mouse organs. 18s rRNA was included as a loading control. Data presented were representative of three independent experiments.

The cloning of the 2kb ECD1 cDNA sequence from mRNA derived from laminar-flow treated differentiated mouse ESCs was achieved using 5’-RACE and 3’-RACE (rapid amplification of cDNA ends) techniques. DNA sequencing of hundreds of colonies resulted in the identification of two clones that were different in length, spanning 63 nucleotides (nt), with a size of 2059nt (Figure 12A) and 1996nt (Figure 12B), respectively.
Figure 12: The nucleotide sequences of \textit{Lafdir} transcript variant 1 and 2.

Data is from previous work conducted in Dr Zeng's group (unpublished data). The coding sequences were indicated by start and stop codons in purple and red for open reading frame 1 and 2 respectively with amino acid sequences annotated.
Blasting this 2kb sequence in the mouse genome database revealed a potential intact gene with 7kb in size and contained a total of 7 exons. This \textit{Laf4ir} sequence was established to be located specifically in the 60kb intron 6 of \textit{LAF4} (Figure 13). \textit{Laf4ir} and \textit{LAF4} uniquely utilise opposite DNA strands for transcription. Two \textit{Laf4ir} RNA transcript variants, referred as transcript variant 1 (TV1) and 2 (TV2), have been identified from mouse aorta. \textit{Laf4ir-tv1} contains 7 exons (2059 nucleotides), while \textit{Laf4ir-tv2} contains 6 exons lacking the exon 4 (1996 nucleotides) (Figure 13). Computer analysis suggests that other putative transcript variants may exist. Sequence analysis of both RNA transcripts reveals three potential opening reading frames (ORFs) for translation (Figure 13).

\textbf{Figure 13: Schematic illustration of the \textit{Laf4ir} gene structure.}
\textit{Laf4ir} is located in the intron 6 of \textit{Laf4}, comprised of 7 exons. The transcription of \textit{Laf4ir} is initiated from the opposite strand of \textit{Laf4} gene. Transcript variant 1 (\textit{Laf4ir-tv1}) contains all 7 exons with 2059 nucleotides (nt) in length, while transcript variant 2 (\textit{Laf4ir-tv2}) lacks exon 4 with 63nt fewer in length. There are three potential open reading frames with different sizes of amino acids. In this thesis, translation of \textit{Laf4ir} ORF1 and ORF2 (highlighted in red) polypeptides has been demonstrated and therefore investigated. Translation of ORF3 (highlighted in blue) polypeptide was not confirmed and therefore not investigated.
To understand the expression profile of potential polypeptides derived from *Laf4ir*, antibodies were raised against epitope of *Laf4ir* ORF1, ORF2 derived from TV1 and TV2, and ORF3 amino acid (aa) sequences by GenScript (Piscataway, NJ, USA) prior to the PhD project. The aa sequences of the three potential ORFs are illustrated in Figure 14. The antibody raised against *Laf4ir* ORF2 aa sequence is not specific for either TV1 or TV2 isoform. Therefore the utilisation of *Laf4ir* ORF2 antibody will not be able to validate specific *Laf4ir* ORF2-TV1 and ORF2-TV2 polypeptide expressions.

**Figure 14: Amino acid sequences of the three potential ORFs.**

*Laf4ir* ORF1 and 3 give rise to 45aa and 82aa, respectively. *Laf4ir* ORF2 give rise to 109aa and 151aa from TV1 and TV2, respectively. Cysteine residues are highlighted in red. The different C-terminal sequences in *Laf4ir* ORF2 in the two transcript variants are underlined. The highlighted sequence in blue is used to raise specific antibodies.

Blasting of the *Laf4ir*-tv1 cDNA mouse sequence to genomes from different species revealed around 70% homology in exon 2 and 7 from bird and lizard to human (Figure 15). No homology between mouse and frog or fish was observed (Figure 15). Importantly, the homology among other exons is quite low, suggesting that *Laf4ir* is an evolutionally late gene. The cloning of the human *Laf4ir* isoform in currently under examination. This cloning could provide interesting insights on whether *Laf4ir* also accounts for phenotypic features witnessed in patients with the microdeletion on chromosome 2q11.1<sup>156</sup>.
Figure 15: Illustration of the conservation of *Laf4ir* gene among different species.
1.12 Project design

1.12.1 Hypothesis

Based on previous investigations, the hypothesis is that *Laf4ir* gene expression can be induced during embryonic development and in adult organs/tissues in response to physiological or pathophysiological stimuli. This includes shear stress and vascular injury. The spatiotemporal translation of different polypeptides from the different open reading frames within *Laf4ir* mRNA may participate in multiple cellular processes in different cell types and tissues. Overall, *Laf4ir* may be associated with cardiovascular system remodelling.

1.12.2 Aims

The aims of this PhD thesis are highlighted accordingly to parts 1, 2 and 3 are as follows:

**Aim 1:** To verify the polypeptide coding property of *Laf4ir* from different ORFs. If *Laf4ir* is an encoding gene, the next focus would be to observe *Laf4ir* ORF1 and ORF2 polypeptide expression during embryonic development and adult life in mice and explore its cellular and sub-cellular localisation.

**Aim 2:** Initially to find out if *Laf4ir* propagates EC differentiation from stem/progenitor cell types. Elucidate the role of *Laf4ir* in proliferation and its mechanism in stem/progenitor cell types. Observe if the nested *Laf4ir* ORF1 polypeptide interact with the parent protein *Laf4*. *Laf4ir* may provide protection against oxidative stress and therefore it is critical to explore endothelial cell survival capability in response to hydrogen peroxide.
**Aim 3**: To evaluate the potential role of *Laf4ir* in cardiac hypertrophy/heart failure, mechanical and ischaemic vascular injury and atherosclerosis. To develop, utilise and examine phenotypic differences of the global knock out model of *Laf4ir*. 
Chapter 2: Materials and methods

2.1 Mice

Mice from different genetic backgrounds including C57BL/6 and ApoE<sup>−/−</sup> (8-14 weeks old; Charles River, West Sussex, UK) were used in this study. To summarise, 8 weeks old C57BL/6 mice were used for organ expression study. 10 weeks old transgenic mice were used for genotype and phenotype studies. 12 weeks old C57BL/6 mice were used for bone marrow cell expression study, TAC model, hindlimb ischaemia model, and femoral artery injury model. 14 weeks old C57BL/6 and ApoE<sup>−/−</sup> mice were used for atherosclerosis study. Animals of groups with three to five were housed at 22±2°C which was on a 12-h light/dark cycle in the King’s College London animal facility. Both types of mice were used for tissue collection and Sca1<sup>+</sup> cell isolation. Water and food pellets were available ad libitum. Mice were habituated for one week and had 30 min acclimatisation to the testing room prior to all experiments. C57BL/6 mice were used for surgeries, embryo, organ, tissue and bone marrow isolation throughout the research.

All mice were monitored according to the UK Animal Scientific Procedures Act (1986) and by the Institutional Committee for Use and Care of Laboratory Animals. Experiments were performed under my UK Home Office Personal licence number I56EDA266 and the groups UK Home Office Project Licence number 70/7266. The surgical procedures were carried out by Dr Yanhua Hu and members from Prof Shah’s laboratory group.

2.2 Cell culture

Sca1<sup>+</sup> vascular progenitor cells (VPCs) were isolated from the outgrowth of adventitial tissues of mouse arterial vessels, as previously described<sup>164, 165</sup>. Briefly, the arterial vessels were harvested from 14 weeks old wild type C57BL/6J or Apoem1Unc
mice (Charles River, Margate, Kent, UK) and cut into 2-mm rings, prior to placement on (0.04%) gelatin-coated flasks and incubated at 37°C in a humidified incubator supplemented with 5% CO₂ as described in chapter 2.8.

Stem cell culture medium was created by supplementing Dulbecco's Modified Eagle's Medium (DMEM) (ATCC, Rockville, MD, USA) with 10 ng/ml recombinant human leukemia inhibitory factor (LIF; Chemicon, Temecula, CA, USA), 10% EmbryoMax fetal bovine serum (FBS) (Millipore, UK), 0.1mmol/L 2-mercaptoethanol (ThermoFisher, UK), 100U/ml penicillin (ThermoFisher, UK), and 100µg/ml streptomycin (ThermoFisher, UK), which was added to cells and refreshed every two days until cells had reached 80% confluence. The cells were amplified and subjected to Sca-1⁺ cell purification using anti–Sca-1 immunomagnetic microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany). The purity of isolated Sca-1⁺ cells was demonstrated using flow cytometry in Figure 30 in Chapter 3.3.6. The Sca-1⁺-VPCs were maintained in stem cell culture medium and passaged every two days at a ratio of 1:4 by Trypsin-EDTA (0.25%) phenol red (ThermoFisher, UK). Up to 5 passages were used in this study.

Mouse ESCs (ES-D3 cell line, CRL-1934) were purchased from ATCC (Rockville, MA, USA) and maintained as described previously. Furthermore, these cells were treated similarly to Sca-1⁺-VPCs. Mouse EC cell line C166 (ATCC, UK) and HEK293 cells (ATCC, UK) (for adenoviral particle production) were purchased from ATCC and maintained in DMEM high glucose, GlutaMAX™ Supplement, pyruvate (ThermoFisher, UK) supplemented with 10% fetal bovine serum (FBS, ATCC), penicillin (100U/ml) and streptomycin (100µg/ml).

### 2.3 ESC differentiation and shear stress

Mouse ESCs cells were seeded on 5µg/ml of collagen IV-coated dishes in differentiation medium [DM, αMEM supplemented with 10% FBS, 0.05mmol/L β-mercaptoethanol, penicillin (100U/ml) and streptomycin (100µg/ml)] for 3 days
indicated in figure legends. Low shear stress by laminar and disturbed flow conditions in cells were induced by 150rpm rotation by a PSU-10i orbital shaker (Grant-Bio, UK). Cells on the outer region were considered to be induced by laminar flow and cells in the middle region were considered to be induced by disturbed flow (Figure 16). Static conditions included cells not treated with shear stress as a control. In previous investigations, a different flow system was used for microarray and qRT-PCR analysis compared to subsequent flow experiments used in this PhD project. This previous flow system utilised a slide chamber exposed to 12 dynes/cm² laminar flow.

Figure 16: Representative diagram of the shear stress protocol
Diagram illustrates areas of a 10cm petri dish exposed to disturbed and laminar flow in differentiated ESCs.

2.4 RNA extraction, RT-PCR and qRT-PCR

Total RNA was isolated from cells using the RNeasy Mini kit (Qiagen, UK). Briefly, after the cells had been washed with PBS to thoroughly remove the culture medium, 350µl RLT buffer was added to disrupt the cells. Lysate solution was then added into a QIAshredder spin column that was placed in a 2ml collection tube prior to centrifugation for 2min at full speed. Subsequently, 350µl ethanol was added to the homogenized lysate, and then mixed by pipetting. The sample was then transferred to an RNeasy spin
column placed in a 2 ml collection tube and centrifuged at 8000g for 15s. After the flow-through had been discarded, 700µl RW1 buffer was added to the RNeasy spin column and afterwards the sample was centrifuged at 8000g for 15s to wash the spin column membrane. Once the flow-through had been discarded, 500µl RPE buffer was added to RNeasy spin column and the sample was centrifuged at 8000g for 2min. This process was repeated twice. After the flow-through had been discarded, the RNeasy spin column was placed in a new 1.5 ml collection tube and 30µl RNase-free water was added to the RNeasy spin column. After the sample was centrifuged at 8000g for 1min, the RNA was diluted into the solution collected in the collection tube. The RNA concentration was detected using the NanoDrop 1000 Spectrophotometer (Thermo scientific, Unk). The RNA sample was then ready to be used for reverse transcription polymerase chain reaction (RT-PCR).

Reverse transcription was performed using the QuantiTect Reverse. Transcription Kit (Qiagen, UK) following the manufacturer’s instructions. Each RNA sample was normalised to 1µg in RNase-free water (total volume: 12 µl) and 2 µl gDNA wipeout buffer was added and incubated at 42 °C for 2min in a multiplex PCR thermal cycle machine (Techne, UK) to eliminate genomic DNA (Gdna) in the RNA sample. This 14µl sample was combined with 4 µl reverse transcription buffer, 1µl RT primer mix and 1µl quantitative reverse transcriptase enzyme. The total 20µl reaction solution was incubated in the thermal cycler at 42 °C for 15min and then at 93°C for 3min. After the reverse transcription reaction process, 80 µl RNase-free water was added to dilute each cDNA samples. 2 µl of cDNA samples (20ng) was subjected to quantitative PCR with a 10µl SYBR green dye–based PCR amplification and detection master mix (ThermoFisher, UK) and specific 1.5 µl forward and 1.5 µl reverse primers (10ug) in a Eppendorf MarsterCycler gradient S machine (Eppendorf, Enfield, CT, USA). Negative control included primers with SYBR green dye mix and no cT values were obtained in results indicating no contamination in primer mix.
2.5 Primers

The primers were designed using *NCBI primer pickup* software and synthesized by Thermofisher (UK). Below is a summary of all primers designed and utilised throughout the research (Table 3). The double delta C\(_t\) value method for gene expression quantification was used by qRT-PCR.

**Table 3: Illustration of all primers used.**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence 5' &gt; 3'</th>
<th>Location</th>
<th>GC %</th>
<th>Tm</th>
<th>Amplicon size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gapdh</td>
<td>5'-cga ctt caa cag caa ctc ctc ttc-3' 5'-ttg gtg gtc cag ggt ttc tta ctc ctt-3'</td>
<td>AY618199</td>
<td>54</td>
<td>67</td>
<td>175 bp</td>
</tr>
<tr>
<td>Lef4l</td>
<td>5'-gtt cgag tag cct tca-3' 5'-ttg aac ccc ttc ttc ctc-3'</td>
<td>Exons 3/5 and Exons 4/5</td>
<td>50</td>
<td>62</td>
<td>100 bp</td>
</tr>
<tr>
<td>Lef4</td>
<td>5'-cta cca caa cc acta cca cta c-3' 5'-cca ggt gcg tgc tat cca taa g-3'</td>
<td>NM_001290814</td>
<td>50</td>
<td>62</td>
<td>78 bp</td>
</tr>
<tr>
<td>Cd144</td>
<td>5'-gtg cct gac gac atc cga gtt-3' 5'-gac ccc tct ctc ttc ttc-3'</td>
<td>NM_009868</td>
<td>57</td>
<td>63</td>
<td>454 bp</td>
</tr>
<tr>
<td>Cd31</td>
<td>5'-gga ggt cct ggt gga cat cag-3' 5'-tgc agc gtt acg tat tca ctc-3'</td>
<td>NM_008816</td>
<td>57</td>
<td>63</td>
<td>525 bp</td>
</tr>
<tr>
<td>Cnn1</td>
<td>5'-gag tca act cag aac tgg cac-3' 5'-ttt ggg ata gag gtt agc-3'</td>
<td>MUSH1CA</td>
<td>52</td>
<td>58</td>
<td>348 bp</td>
</tr>
<tr>
<td>Sm22</td>
<td>5'-agt gga ttg tag tgc cat gtt-3' 5'-cag ttg ctt gtt gaa gtc-3'</td>
<td>MUSSM22A</td>
<td>48</td>
<td>59</td>
<td>410bp</td>
</tr>
</tbody>
</table>

2.6 DNA extraction and fragmentation

Cells were scraped and extracted with media before suspension in 2ml tubes. For the knockout mouse line, ear clips were taken from each of the new litter. Lysis buffer was added along with 1 μg/ml proteinase K. Samples were then incubated at 55°C overnight. 535μl of phenol chloroform was added and vortexed to mix. Samples were then centrifuged at 14,630rpm for five minutes. Having split into two layers, around 300μl of the top layer was removed and placed into 1.5ml tube. 40μl of 3M NaAc and 800μl 100% ethanol was then added and vortexed to mix, subsequently precipitating DNA. Samples were then centrifuged at 14,630rpm for five minutes. Supernatant was then discarded, being left with DNA pellet. This was then washed with
500µl of 70% ethanol and centrifuged for five minutes several times. Supernatant was then discarded and 40µl TE buffer added.

For DNA fragmentation analysis, extracted genomic DNA was quantified using a nano dropper and normalized to 1µg/µl with ddH2O to a final volume of 12µl. 5µl DNA loading buffer was then added. Samples were then loaded onto 1.5% agar gel using TAE buffer containing 1µg/ml ethidium bromide. Electrophoresis was run at 30V for 4 hours after which the gel was then imaged using a UV light and images obtained using a Biospectrum AC Imaging System 500. Genotyping of DNA samples from transgenic mice are described next in chapter 2.7.

2.7 Genotype of *Laf4ir* knockout mice

Genotyping for the target *Laf4ir* gene in a conditional knock-out mouse model was designed, performed and validated by the Genoway biotechnology company based in France. Ear clips were taken from new litters and their DNA isolated as described in chapter 2.6. PCR genotyping was carried out by two specific primers to validate the insertion of *loxP* sites, with the neomycin cassette flanked by FRT sites, in a multiplex PCR thermal cycle machine (Techne, UK). The PCR mix which included the TAQ DNA polymerase, nucleotides and PCR buffer were employed by Accuprime (Invitrogen, UK). The optimised PCR conditions for all primers (Table 4), the amplicon PCR products (Table 5), the primer sequences (Table 6), and illustration of the endogenous *Laf4ir* and conditional knockout-alleles (Figure 17) are listed in the following page.

Following PCR completion, a 1.5% agarose gel (Invitrogen, UK) was created in 1X TAE buffer (National Diagnostic, UK). The mixture was heated to boiling point in a microwave, then cooled down and 0.5µg/ml of ethidium bromide nucleic acid stain (NBS Biologicals, UK) was added into the solution. The gel mixture was carefully poured into a Horizon horizontal gel cast electrophoresis apparatus with the well comb inside.
Once the gel had reached room temperature and solidified, the gel was gently placed into 1X TAE buffer and the comb was carefully removed. 25 μL of amplified DNA samples together with 5 μL of 6x DNA loading dye (ThermoFisher, UK) were loaded into each well. To mark different sizes of base pairs, 0.5 μg/ml 100 bp DNA Ladder (New England, Biolabs, UK) was used in each gel. Electrophoresis was performed at 160V for 30 minutes. DNA fragments were visualized under UV light and images were obtained using a Biospectrum AC Imaging System 500.

<table>
<thead>
<tr>
<th>Reaction mix</th>
<th>Reaction conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic DNA</td>
<td>Denaturing 94°C 120s 1x</td>
</tr>
<tr>
<td>Primers</td>
<td>Denaturing 94°C 30s 30x</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Annealing 65°C 30s</td>
</tr>
<tr>
<td>PCR buffer</td>
<td>Extension 68°C 300s</td>
</tr>
<tr>
<td>TAQ polymerase</td>
<td>Completion 68°C 480s 1x</td>
</tr>
<tr>
<td>Reaction volume</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Optimised PCR conditions for the detection of the *Laf4ir* conditional and constitute knock-out allele.
Concentrations given correspond to the final concentration in the reaction tube.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>PCR amplicon size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type-allele</td>
<td>251 bp</td>
</tr>
<tr>
<td>Heterozygous conditional knock-out mice</td>
<td>251 bp + 319 bp</td>
</tr>
<tr>
<td>Homozygous conditional knock-out mice</td>
<td>319 bp</td>
</tr>
</tbody>
</table>

Table 5: PCR amplicon base pair (bp) product expected from primers X and Y

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5’-3’</th>
<th>PCR product size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Wild-type allele</td>
</tr>
<tr>
<td>X primer</td>
<td>5’-GTA GCT CAC AAA AGC CAC AGG GGC-3’</td>
<td>251 bp</td>
</tr>
<tr>
<td>Y primer</td>
<td>5’-TGT CTG TCT TCA AGA AAC TGA ACA CGG G-3’</td>
<td>251 bp</td>
</tr>
</tbody>
</table>

Table 6: Primers X and Y for the detection of the *Laf4ir* conditional Knock-out allele.
Figure 17: PCR genotype of the Laf4ir endogenous and conditional knock-out mouse line.

A) Schematic representation of the endogenous Laf4ir and the conditional Knock-alleles. LoxP sites were inserted before the promoter and after exon 3. Arrows indicated the primer binding sites. P: Promoter, E1-3: Exons 1-3, FRT: flippase recognition target.

B) PCR was conducted using genomic DNA extracted from ear clip biopsies of C57BL/6 mice. Samples containing C57BL/6 wild-type DNA (Laf4ir+/+) or without DNA (H2O) served as positive and negative controls, respectively. Sample from mice with loxP sites inserted, generated Laf4ir homozygous conditional knock-out locus (Laf4ir-loxp/loxp). PCR fragments were separated by electrophoresis on a Biospectrum AC Imaging System 500 and analysed using PROSize 2.0 analytical software.
To generate global *Laf4ir* knock-out mice, the *Laf4ir* homozygous conditional knock-out mice (*Laf4ir* 
\(^{loxP/loxP}\)) was crossed with Cre-Recombinase mice purchased from the Jackson biotechnology company in the UK. This generated the first *Laf4ir* heterozygous constitutive knock-out (*Laf4ir* \(^+-\)) and then ultimately *Laf4ir* homozygous constitutive knock-out (*Laf4ir* \(-/-\)) mice. PCR genotyping was carried out by three specific primers to validate the deletion of DNA sequence in-between *loxP* sites in a multiplex PCR thermal cycle machine (Techne, UK). The PCR amplicon products (Table 7), the primer sequences (Table 8) and the illustration of the endogenous, conditional and constitutive *Laf4ir* Knock-out alleles (Figure 18) are listed in the following below.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>PCR amplicon size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type mice</td>
<td>200 bp + 2951 bp</td>
</tr>
<tr>
<td>Heterozygous constitutive knock-out mice still</td>
<td>200 bp + 261 bp + 327 bp + 2951 bp + 3080 bp</td>
</tr>
<tr>
<td>displaying a conditional knock-out allele</td>
<td></td>
</tr>
<tr>
<td>Heterozygous knock-out mice</td>
<td>200 bp + 327 bp + 2951 bp</td>
</tr>
<tr>
<td>Homozygous knock-out mice</td>
<td>327 bp</td>
</tr>
</tbody>
</table>

**Table 7: PCR amplicon product expected from primers A, B and C.**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5′-3′</th>
<th>PCR product size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Wild-type allele</strong></td>
<td><strong>Conditional knock-out allele</strong></td>
</tr>
<tr>
<td>A primer</td>
<td>5′-CCT ATG GTG TTC TTT GCA TCC ACC AAC-3′</td>
<td>200 bp</td>
</tr>
<tr>
<td>B primer</td>
<td>5′-CCC CAC ACA TAA GCA CAC ATC CAG AAC-3′</td>
<td>200 bp</td>
</tr>
<tr>
<td>C primer</td>
<td>5′-AAA TGA AGG GAC CTT TAT CCT ACC CAC AC-3′</td>
<td>200 bp</td>
</tr>
</tbody>
</table>

**Table 8: Primers for the detection of the *Laf4ir* constitutive knock-out allele.**
Figure 18: PCR genotype of the *Laf4ir* endogenous, conditional and constitute knock-out mouse line.

**A)** Schematic representation of the endogenous, conditional and constitutive *Laf4ir* Knock-out alleles. *LoxP* sites were inserted before the promoter and after exon 3. Arrows indicated the primer binding sites. P: Promoter, E1-3: Exons 1-3, FRT: flippase recognition target.

**B)** PCR was conducted using genomic DNA extracted from ear clip biopsies of C57BL/6 *Laf4ir* wild type (*Laf4ir*+/+), *Laf4ir* heterozygous constitutive knock-out (*Laf4ir*+/−) and *Laf4ir* homozygous constitutive knock-out (*Laf4ir*−/−) locus. PCR fragments were separated by capillary electrophoresis on a Biospectrum AC Imaging System 500 and were analysed using PROSize 2.0 analytical software.
2.8 Embryo/organ/tissues/BMC/VPC isolation

**Embryos**

In total, 15 pregnant C57BL/6 mice were purchased from Charles River (Harlow, UK) and used for embryo harvesting at different embryonic stages (E8.5, 10.5, 14.5, 16.5 and 19.5). 3 mice from each known embryonic stage were sacrificed and their embryos were isolated and frozen down on dry ice. Each embryo in RIPA buffer was homogenised in lysing matrix D tubes (MP Biomedicals, UK) by a precellys 24 lysis homogeniser machine (Bertin Technologies, France) at 5,000 rpm for 3 mins three times. Each sample was then sonicated (Brandson Sonicator, UK) numerous times for a few seconds. Total protein content was isolated and subjected to western blotting as described in chapter 2.9.

**Organs**

8 weeks old C57BL/6 and 14 weeks old ApoE−/− mice were sacrificed and their organs were harvested which included the brain, heart, lungs, liver, stomach, skeletal muscle (thigh), kidneys, aorta artery, femoral artery, stomach, small intestine and spleen. Frozen samples were then homogenised as previously described, by the embryo isolation protocol and total protein content was isolated and subjected to western blotting as described in chapter 2.9.

**Bone marrow cells**

12 weeks old C57BL/6 bone marrow cells were harvested by flushing the femurs and tibias with 1xPBS. This was followed by the removal of excess tissue and/or large clots by passing cells through a 40μm cell strainer (Falcon, UK). Bone marrow cells were then incubated with an erythrocyte lysis buffer (eBioscience, UK) for 20min to remove peripheral red blood cells. All washed samples were then kept on ice for flow cytometry analysis as described in chapter 2.14.
Vascular progenitor cells

*Laf4ir* heterozygous (*Laf4ir*+/−) and wild type mice were sacrificed and their aortas were isolated in strict sterile conditions. This included both the aortic arch and root with all adipose tissue removed. The aorta was then opened and the intima and media peeled off in DMEM media with 20% EmbryoMax FBS, 100U/ml penicillin, and 100µg/ml streptomycin, exposing the distinguishable adventitia. The adventitia was then washed several times with media and then cut with scissors into small pieces (about diameter 1mm).

Small pieces of adventitia were then placed in a 0.04% gelatin-coated T25 flask to dry for 3hrs with the flask standing on its side at 37°C in a humidified incubator supplemented with 5% CO₂. The tissues were then immersed in the same maintenance medium with 1µg/ml plasmocin (InvivoGen, UK) in an upright position the same incubator. After days 4-5, maintenance medium was changed whilst after day 7, cells were passaged to increase cell number. Once sufficient cells were grown, cells were then treated with fresh maintenance stem cell medium after the first Sca-1 cell sorting.

2.9 Protein extraction and western blotting

Cells, organs, tissues and embryos were added in RIPA buffer (ThermoFisher, UK) with phosphatase and protease inhibitors (Roche, UK). The resulting cell lysate was transferred to a 1.5ml crystal clear microcentrifuge tube (Starlab, UK) and twice sonicated using a Branson Sonifier 150 at level 1, each time for 8 seconds. After a 15 min centrifugation at 15000 g speed at 4 °C, supernatant was collected and transferred to a new 1.5ml microcentrifuge tube. Quantification of protein concentration was performed using the DC Biorad protein reagent assay according to manufacturer’s instructions. Briefly, reagent A and reagent B were sequentially mixed with the sample, absorption was detected and recorded at 750 nm whilst protein concentration was measured by a spectrometer (Bio-Rad, UK) according to a resulting standard curve.
The loading buffer was prepared using 5xSDS (Thermo-Fisher, UK) supplemented with 10% 2-mercaptoethanol (Thermo-Fisher, UK). All protein samples were then normalised to 2μg/ml with SDS and distilled water. Once loading samples had been prepared, they were centrifuged briefly and incubated at 95°C for 10 min to denature proteins. NuPAGE 4-12% bis-tris gels (Thermo-Fisher, UK) were loaded into a running gel tank, and the comb was carefully removed. First 1X running buffer (Thermo-Fisher, UK) was added to the wells and each well was rinsed with a pipette to remove excess acrylamide.

After the running buffer had been added, 60μg of samples plus the 10-250 kD protein ladder (BIORAD Precision Plus ProteinTM Dual Colour Standards ladder, UK) were loaded into the wells. Any remaining wells were loaded with the 1X SDS to normalise the pH balance within the gel. The electrophoresis apparatus was then attached to an electric power supply running at 140V for 90min, after which the gel was gently removed from the plastic plates.

After running of the gel, transfer buffer consisting of 20% methanol (Thermo-Fisher, UK) and 1x NuPAGE transfer buffer (Thermo-Fisher, UK) was prepared and used to soak PVDF Blotting membrane (GE healthcare life science, United Kingdom) which was activated by 100% methanol beforehand. The PVDF membrane and filter paper was laid across the NuPAGE 4-12% bis-tris gel according to manufacturer’s instruction. Transfer procedure was run for 1 hour at 30V.

After washing with TBST (prepared from tris buffered saline (TBS) tablet (AMERCO, UK), 0.1% Tween-20 (Sigma-Aldrich, USA) and 100mL distilled water), the membrane was blocked by shaking in 5% non-fat dried milk in TBST for 1 hr at room temperature with shaking. The membrane was incubated with primary antibody diluted in 5% non-fat dried milk in TBST overnight at 4°C with shaking.
The following day, the membrane was washed in TBST three times for 5min each. The membrane was then incubated with secondary HRP-conjugated antibody diluted at 1:3000 in 5% milk in TBST for 1h, followed by washing in TBST three times for 5min each. To develop the membrane Amersham ECL Western Blotting Detection reagent (GE healthcare life science, UK) was applied for 2min at room temperature, after which the membrane was then put inside plastic film in a hypercassette and developed using Amersham hyperfilm (GE healthcare life science, United Kingdom) in the developing machine (Ecomax, Protech, UK).

Primary antibodies applied in this project are listed in Table 9. All secondary antibodies were raised in swine and were anti mouse, rat, rabbit or goat HRP-conjugated (Dako, UK). For the peptide blocking assays, the antibody was pre-incubated with blocking peptide at a ratio of 1:1 at 4°C for 24hr. Membrane was re-used by immersing in Western Blot Stripping Buffer (Thermofisher Scientific, UK) for 30 minutes followed by washing in TBST three times for 5min each.

2.10 Antibodies

The peptides of QRNRRPWSVKITSD, ADHRSTPQAGKVR and FPTRQEEVEPKQ were synthesized and used to raise *Laf4ir* anti-ORF1, ORF2 and ORF3 antibodies in mouse, rabbit and rat, respectively by GenScript (Piscataway, NJ, USA). Below is a summary of all monoclonal and polyclonal primary antibodies used in this PhD thesis project (Table 9). Experimental applications utilising the selected antibodies are as follows: Western blot (WB), Immunoprecipitation (IP), enzyme-linked immunosorbent assay (ELISA), immunocytochemistry (ICC), immunohistochemistry (IHC), fluorescence-activated cell sorting (FACS) and cell sorting.
### Table 9: Illustration of all monoclonal and polyclonal primary antibodies used

<table>
<thead>
<tr>
<th>Protein</th>
<th>Host specie</th>
<th>Supplier</th>
<th>Cat No</th>
<th>Application</th>
<th>Conjugation</th>
</tr>
</thead>
<tbody>
<tr>
<td>αSMA</td>
<td>Mouse</td>
<td>Sigma</td>
<td>C6198</td>
<td>5 μg/ml (WB)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (ICC/HHC)</td>
<td></td>
</tr>
<tr>
<td>αSMA</td>
<td>Goat</td>
<td>Abcam</td>
<td>ab21027</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td>αTubulin</td>
<td>Rabbit</td>
<td>Abcam</td>
<td>ab18251</td>
<td>1 μg/ml (WB)</td>
<td>-</td>
</tr>
<tr>
<td>BrdU</td>
<td>Sheep</td>
<td>Abcam</td>
<td>ab2285</td>
<td>0.5 μg/ml (ELISA)</td>
<td>HRP</td>
</tr>
<tr>
<td>CD31</td>
<td>Rabbit</td>
<td>Abcam</td>
<td>ab28364</td>
<td>5 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td>CD31</td>
<td>Rat</td>
<td>LifeSpan Bioscience</td>
<td>LS-C4921-50</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td>CD34</td>
<td>Rat</td>
<td>BD Pharmingen</td>
<td>553731</td>
<td>5 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td>CD144</td>
<td>Goat</td>
<td>Santa Cruz</td>
<td>Sc6458</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td>FLAG</td>
<td>Mouse</td>
<td>Sigma</td>
<td>F1804</td>
<td>5 μg/ml (WB)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MERK</td>
<td>F4049</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>FITC</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Rabbit</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-25778</td>
<td>0.1 μg/ml (WB)</td>
<td>-</td>
</tr>
<tr>
<td>Lamin B1</td>
<td>Mouse</td>
<td>Abcam</td>
<td>ab133741</td>
<td>1 μg/ml (WB)</td>
<td>-</td>
</tr>
<tr>
<td>LAF4</td>
<td>Rabbit</td>
<td>MyBioSource</td>
<td>MBS853344</td>
<td>3 μg/ml (WB)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (IP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 μg/ml (ICC/HHC)</td>
<td></td>
</tr>
<tr>
<td>MCM3</td>
<td>Goat</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-9850</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (IP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (WB)</td>
<td></td>
</tr>
<tr>
<td>ORF1</td>
<td>Mouse</td>
<td>GenScript</td>
<td>-</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 μg/ml (FACS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (WB)</td>
<td></td>
</tr>
<tr>
<td>ORF2</td>
<td>Rabbit</td>
<td>GenScript</td>
<td>-</td>
<td>2 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 μg/ml (FACS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (WB)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (IP)</td>
<td></td>
</tr>
<tr>
<td>ORF3</td>
<td>Rat</td>
<td>GenScript</td>
<td>-</td>
<td>2 μg/ml (IP)</td>
<td>-</td>
</tr>
<tr>
<td>Sca-1</td>
<td>Rat</td>
<td>Abcam</td>
<td>ab51317</td>
<td>5 μg/ml (ICC/HHC)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (FACS)</td>
<td></td>
</tr>
<tr>
<td>Sca-1</td>
<td>Mouse</td>
<td>Miltenyi Biotec</td>
<td>130-092-529</td>
<td>2 μg/ml (Cell sorting)</td>
<td>FITC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 μg/ml (FACS)</td>
<td></td>
</tr>
</tbody>
</table>
All secondary antibodies used were donkey Alexa Fluor® 488, 594 and 647-conjugated fluorophores at a concentration of 1 μg/ml (ICC/IHC/FACS) raised against different species including goat, rat, mouse and rabbit (Thermo Fisher Scientific, UK).

### 2.11 Immunoprecipitation/Mass Spectrometry

Cell monolayer was washed with 1xPBS containing 5% FBS, removed by a cell scraper (VWR, UK) and spun down at at 4°C at 1000rpm for 5 min. The supernatant was discarded, cell pellet was resuspended in 300μl IP-Buffer A (25mM Tris-Cl pH 7.5, 150mM NaCl, 1mM EDTA pH 8.0, 1.0% Triton X-100 and protease inhibitor (Roche, UK)), incubated on ice for 45 min and spun down at 10,000rpm at 4°C for 5min. Supernatant was transferred into a 1.5ml tube and the pellet was resuspended in 100μl IP-Buffer A and sonicated at 40% output for 10 sec and incubated at 4°C for 30min. Then sample was spun down at 10,000 rpm at 4°C for 5min, supernatant harvested, and the total protein concentration detected with Bradford reagent as instructed by the manufacturer’s instructions (Bio-Rad, UK). 1mg proteins were isolated and mixed with 3x volume of IP-Buffer B (20mM Tris-Cl Ph 7.5, 120mM NaCl, 2mM CaCl₂, 1mM EDTA pH 8.0, 1mM EDTA pH 8.0 and protease inhibitor (Roche, UK)). The lysate was added with 2μg normal IgG and 5μg protein-G beads (Sigma, UK), incubated on a rotator at 4°C for 1hr and spun down at 4°C at 10,000rpm for 3min. The supernatant was recovered, added with 2μg specific primary antibody (anti-ORF1, ORF2, MCM3 or LAF4) or normal IgG, and incubated on a rotator at 4°C for 2hr. Then 10μl of protein-G agarose beads was added and incubated at 4°C overnight.

The following day, sample was spun down at 4°C at 2,000rpm for 3min and washed with 1ml IP washing buffer (25mM Tris-Cl pH 7.5, 150mM NaCl, 1mM EDTA pH 8.0, 0.5% Triton X-100 and protease inhibitor) on a rotator at 4°C for 5min. Samples were subjected to western blot as described in chapter 2.9 or proteomics analysis. For proteomic analysis, sample was spun down at 4°C at 2000rpm for 3min and washing
step was repeated three more times. The pellet was resuspended in 25µl 1xSDS elution buffer (25mM Tris-Cl pH 7.5, 1mM EDTA pH 8.0, 2% SDS and 5% 2-Mercaptoethanol), incubated at 72°C for 30 min and spun at 3000rpm at room temperature for 5min. The recovered supernatant was then added with 5µl 6x DNA loading dye and 100µg total cellular protein was applied in a 4-12% SDS-PAGE running gel. The proteins between 3KDa and 25KDa were cut and subjected to in-gel trypsin digestion and subjected to mass spectrometry (MS). The mass spectrometry procedure and data analysis was conducted by Mr Steven Lynham and Dr Xiaoping Yang.

Chromatographic separation were performed using an Ultimate 3000 NanoLC system (ThermoFisherScientific, UK). Peptides were resolved by reversed phase chromatography on a 75µm*15cm C18 column using a linear gradient of water in 0.1% formic acid and 80% acetonitrile in 0.1% formic acid. The gradient was delivered to elute the peptides at a flow rate of 250 nl/min over 60 min.

The eluate was ionised by electrospray ionisation using an Orbitrap-Fusion-Lumos (ThermoFisherScientific, UK) operating under Xcalibur v4.1. The instrument was programmed to acquire using a “Universal” method by defining a 3s cycle time between a full MS scan and MS/MS fragmentation. This method takes advantage of multiple analysers on Orbitrap-Fusion-Lumos, and drives the system to use all available parallelizable time, resulting in decreasing the dependence on method parameters. The MS analysis was conducted using collision energy profiles that were chosen based on the mass-to-charge ratio (m/z) and the charge state of the peptide.

Raw mass spectrometry data were processed into peak list files using Proteome Discoverer (PD v2.2; ThermoScientific). Processed data was then searched using Mascot search algorithm (www.matrixscience.com) against the current version of the reviewed SwissProt mouse dataset downloaded from Uniprot (www.uniprot.org/uniprot/), and
Sequest search engine embedded in PD 2.2 against in-house databases built based on the specific amino acid sequences of Laf4ir ORF1, ORF2 from TV1 and OR2 from TV2.

### 2.12 Cellular fractionation

Cellular fractionation was performed according to standard methods. Briefly, the cells were washed with ice cold 1xPBS containing 5% BSA and removed by mechanical scraping with a rubber policemen (VWR, UK). The solution was spun down at 1000rpm at 4°C for 5 min and the supernatant discarded with the remaining cell pellet resuspended in 1ml ice cold 1xPBS. This solution was transferred into a 1.5ml centrifuge tube, followed by centrifugation at 10,000g at 4°C for 10 sec, and the supernatant discarded.

The cell pellet was then resuspended in 100µl hypotension buffer (10mM Tris-HCl, pH 7.5, 10mM KCl, 1mM EDTA, protease and phosphatase inhibitors) and incubated on ice for 15 min and vortexed mildly several times every 5 min. Then, 6.25µl of 10% NP-40 was added and vortexed at the highest speed for 10 sec. This solution was spun down at 13,200 rpm at 4°C for 10 sec and the supernatant harvested into a fresh tube. This solution contained the cytoplasm fraction. All samples were then subjected to western blot analysis as described in chapter 2.9.

For nuclear extract isolation, the pellet was resuspended in 1ml of cold 1xPBS and spun down at 10,000g at for 10 sec. This nuclear pellet was resuspended in 70µl of High Salt buffer C (10mM Tris-HCl, pH7.5, 420mM KCl, 1mM EDTA, 30% glycerol plus protease and phosphatase inhibitors) and incubate on ice for 45 min whilst being vortexed every 5 min. This solution was spun down at 132,000rpm at 4°C for 5 min and the supernatant recovered which contained the nuclear extract. All samples were then subjected to western blot analysis as described in chapter 2.9.
2.13 Immunofluorescence

Immunocytochemistry

ESC, differentiated ESCs, VPCs in 10cm cell culture dishes were washed with 1xPBS (ThermoFisher, UK), fixed with 4% paraformaldehyde (PFA) solution (ThermoFisher Scientific, UK) for 15 mins and washed again with 1xPBS before immunostaining. Surrounding edges of each dish was cut by a heated metal wire. Cells were blocked and permeabilised with 5% donkey serum (Sigma, UK), 0.1% Tween 20 (Invitrogen, UK) and 0.1% triton X-100 (Invitrogen, UK) in 1xPBS for 2 hr at room temperature. Incubation of cells with primary antibodies in blocking solution (5% donkey serum, 0.1% Tween 20 and 0.1% triton-100) was performed at 4°C overnight in a damp condition. Cells were also treated with only blocking solution as negative control.

The following day, cells were washed with 1xPBS three times for 5 mins each prior to secondary antibodies application for 2hr in the dark damp environment protected from light. Negative control samples consisted of only secondary antibodies. Cells were washed three times with 1x PBS and counterstained with 1 μg/ml 4′-6-Diamidino-2-phenylindol (DAPI) (ThermoFisher Scientific, UK) for 3 min, followed by washing with 1x PBS. Sections were mounted in mounting media (ThermoFisher Scientific, UK) and covered with a covering slip (Menzel-Glaser, Germany).

Routine imaging and quantification were performed on an inverted microscope (TS100, Nikon Eclipse Microscope, UK) or a confocal microscope (Leica SP5, Germany), and processed by Image J and Windows Photo Viewer software (PerkinElmer, UK). Magnification was indicated in figure legends as scale bars. For details regarding the concentrations of the antibodies used, see Table 9 in chapter 2.10.
Immunohistochemistry

Harvested heart, skeletal muscle, aorta and femoral arteries were frozen at -80 °C before OCT (VWR, UK) embedding. Cryo-sectioned tissues (10-12 mm thickness) on microscope slides (VWR, UK) were obtained using a Cryostat machine (Leica CM1950 platform). Samples on slides were then fixed with 100% cold acetone for 30 mins and then subjected to the same immunofluorescent staining protocol as described above. For details regarding the concentrations of the antibodies used, see Table 9 in chapter 2.10.

2.14 Flow cytometry analysis

Fluorescence activation cell sorting (FACS) assay was performed in ESCs, VPCs and ECs. Briefly, cells were harvested by 0.05% trypsin-EDTA (Gibco, UK) and washed in phosphate-buffered saline (1xPBS). The cells were permeabilised and fixed with 0.1% Triton™X-100 (Sigma, UK) and 1% PFA (ThermoFisher Scientific, UK) in 5% donkey serum (Sigma, UK) blocking solution for 30 min at room temperature in FACS tubes (BD Biosciences, UK). The cells were then incubated with primary antibody at room temperature for 2hrs, washed with 1xPBS several times, followed by incubation with secondary antibody at room temperature for 1hr in the dark environment protected from light. The cells were washed several times with 1xPBS. Cells were collected at a rate of 100ul/minute by a BD Accuri™ C6 Flow Cytometry machine (Beckman Coulter Inc., UK) and analyzed by flowJo software. Debris were eliminated using FSC-A vs SSC-A gating and cell doublets and clumps were eliminated by FSC-H vs FSC-A gating. Background fluorescence signal was obtained and subtracted by secondary antibody only stained cells. For the peptide blocking assays, the primary antibody was pre-incubated with blocking peptide at a ratio of 1:1 at 4°C overnight. The antibodies applied in this project are shown in table 9 in chapter 2.10.
Bone marrow cell analysis

Following bone marrow cell isolation as described in chapter 2.8, cells were subjected to FACS analysis as instructed by a relatively recent article titled; “A Protocol for the Comprehensive Flow Cytometric Analysis of Immune Cells in Normal and Inflamed Murine Non-Lymphoid Tissues” 168.

Bone marrow cells (100,000 cells per sample) were suspended in 5% donkey serum for blocking, and solutions A and B supplied by the BD Intrasure kit (BD, UK) in 1xPBS for effective fixation and permeabilisation according to manufacturer’s instruction. Additionally, 1μg/ml rat anti-mouse CD16/CD32 (MouseFc Blocker, BD, UK) was added to each sample. This is because these Fc receptors bind antibodies via their constant Fc domain rather than the antigen specific Fab domain in various leukocytes including monocytes, dendritic cells and B cells, which can lead to false positive results.

The cells were then incubated with Laf4ir rabbit anti-ORF2 antibody and all other cell surface fluorophore-conjugated antibodies as listed below in table 10 at room temperature for 2hrs. The cells were then washed with 1xPBS several times, followed by incubation with donkey anti-rabbit 488 conjugated secondary antibody at room temperature for 1hr. The cells were washed several times with 1xPBS.

These stained cells were collected and analysed using the advanced BD LSR Fortessa II flow cytometry (Becton Dickinson, UK) by Dr Witold Nowak. 100,000 cells from each sample were collected at a rate of 100ul/minute. The positive and negative control compensation protocols included anti-rat, anti-hamster and anti-mouse compensation beads (BD, UK) incubated with the cell surface fluorophore-conjugated antibodies for 30 mins at room temperature. This provides distinct positive and negative
(background fluorescence) stained populations which can be used to set compensation levels.

Data was then analysed by the Flowjo Software. For the peptide blocking assays, the primary antibody was pre-incubated with blocking peptide at a ratio of 1:1 at 4°C overnight. Debris were eliminated using FSC-A vs SSC-A gating and cell doublets and clumps were eliminated by FSC-H vs FSC-A gating. Background fluorescence signal was obtained and subtracted by cells treated with Laf4ir ORF2 antibody incubated with blocking peptide at a ratio of 1:1 for 24hr.

**Table 10: Antibodies for immunophenotyping bone marrow-derived cells**

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Clone</th>
<th>Isotype</th>
<th>Dilutions</th>
<th>Conjugate</th>
<th>Company</th>
<th>Code #</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD11b</td>
<td>M1/70</td>
<td>Rat lgG2b, κ</td>
<td>1:200</td>
<td>APCCy7</td>
<td>BD biosciences</td>
<td>557657</td>
</tr>
<tr>
<td>CD11c</td>
<td>HL3</td>
<td>Armenian Hamster lgG1, λ2</td>
<td>1:100</td>
<td>BV785</td>
<td>BD biosciences</td>
<td>563735</td>
</tr>
<tr>
<td>CD24</td>
<td>M1/69</td>
<td>Rat lgG2b, κ</td>
<td>1:1000</td>
<td>BV711</td>
<td>BD biosciences</td>
<td>563450</td>
</tr>
<tr>
<td>CD45</td>
<td>30-F11</td>
<td>Rat lgG2b, κ</td>
<td>1:800</td>
<td>BV605</td>
<td>Biolegend</td>
<td>103139</td>
</tr>
<tr>
<td>CD64</td>
<td>X54-5/7.1</td>
<td>Mouse lgG1, κ</td>
<td>1:400</td>
<td>BV421</td>
<td>Biolegend</td>
<td>139309</td>
</tr>
<tr>
<td>IA/IE</td>
<td>M5/114.15.2</td>
<td>Rat lgG2b, κ</td>
<td>1:1500</td>
<td>BV650</td>
<td>BD biosciences</td>
<td>563415</td>
</tr>
<tr>
<td>Ly6C</td>
<td>HK1.4</td>
<td>Rat lgG2c, κ</td>
<td>1:200</td>
<td>PerCPCy5.5</td>
<td>eBioscience</td>
<td>45-5932</td>
</tr>
<tr>
<td>Ly6G</td>
<td>1AB</td>
<td>Rat lgG2a, κ</td>
<td>1:200</td>
<td>AF700</td>
<td>BD biosciences</td>
<td>561236</td>
</tr>
</tbody>
</table>

**2.15 Plasmid and adenoviral construction**

The full length *Laf4ir* mRNA was cloned using commercially available 5′-RACE and 3′-RACE kits (ThermoFisher Scientific) according to manufacturer’s instructions prior to my PhD project. The cDNA sequence for the 151-aa of *Laf4ir* ORF2 was synthesized (Genscript, UK) with FLAG tag sequence inserted downstream the ATG start
codon and cloned to pShuttle2 vector, abbreviated as pShuttle2-L151, respectively. The production of pShuttle2-L151 vector was conducted by Dr Yi Li prior to my PhD project.

In this PhD project, the pShuttle2-L151 vector was transfected in mouse ESCs using Fugene 6 (Promega, UK) to explore the encoding capability of Laf4ir ORF2 from TV2. For an optimised and successful transfection of the vector, a 3:1 ratio (transfection Reagent: DNA) was used in a 24-well plate cultured in mouse ESCs. 2μg of plasmid pShuttle2-L151 vector DNA was added to the FuGENER 6 Transfection Reagent, and mixed and incubated for 15 minutes at room temperature. Then 10μl of the FuGENER 6 Transfection Reagent/ pShuttle2-L151 vector mixture was added to each well of a 24-well plate containing 490μl growth medium. The solution was mixed by pipetting and by a plate shaker for 10 seconds. The plate was returned in the incubator for 24 hrs. Cells were then subjected to immunochemistry as instructed in chapter 2.13.

Adenoviral particles not expressing any aa sequence was abbreviated as Ad-null. The pShuttle2-L151 vector was incorporated into adenoviruses to overexpress the Laf4ir ORF2 sequence from TV2 (151 aa). This newly synthesised adenovirus was abbreviated as Ad-L151. The production of the adenoviruses was created by using adenoviral X system (Clontech, UK) and amplified in HEK293 cells according to the manufacturer’s instructions. The development of Ad-null and Ad-L151 viruses were conducted by Dr Yi Li prior to my PhD project.

2.16 BrdU proliferation assay

The 3-day differentiated ESCs were infected with Ad-null or Ad-L151 virus at 10 MOI for 24hr. Then, 2,000 cells were seeded in each well of a 96 well plate for 24hrs. The dESC proliferation assay was performed using a cell proliferation BrdU (colourimetric) (Roche, UK). BrdU labelling reagent (final concentration of 10 μM) was added for 48hr. After removing the labelling medium, 200μl of Fix Denat was added in each well for 30 mins at room temperature.
The solution was removed and cells were washed with 1xPBS three times. This was followed by anti-BrdU-HRP (Abcam, UK) antibody for 2hrs. After 3 washes with 1xPBS, 100μl of substrate solution was added and incubated at room temperature. Once a change in colour was detected, 25μl of 1M H₂SO₄ was added to stop the reaction. Absorbance of HRP was measure at 370nm within 5mins of adding the stop solution by the Tecan i-control plate reader (Infinite® 200 PRO, UK). A mixture of BrdU labelling solution and its antibody in culture medium was used as blank to identify the unspecific binding of BrdU and antibody to the plate. The wells in which cells had not been incubated with BrdU but were incubated with antibody were used as background controls to evaluate the unspecific binding the antibody to dESCs. Calculated Br-dU measurement was measured between treated cells and background control cells. The relative Br-dU was defined as the ratio of A370nm of Ad-L151 group to Ad-null group with that of Ad-null group set as 1.0.

### 2.17 MTS assay

The cell survival assessment was performed with the CellTiter 96® AQueous One Solution Cell Proliferation Assay Kit (Promega, UK) with slight modification of the protocol provided. Briefly, mouse ECs were seeded in gelatin-coated 24-well plates at 5x10^4 cells/well in maintenance medium. Sixteen hr later, Ad-null or Ad-L151 were added at 10 MOI and incubated for 24hr. The cells were then incubated with serum free medium without or with 50 μM and 100 μM H₂O₂ for 24hr. Then, 60μl MTS reagent was added into each well of the 24-well assay plate in 400μl of serum-free medium at 37°C for 4 hours in a humidified, 5% CO₂ atmosphere. To stop the reaction, 75μl of 10% SDS was added to each well. Store SDS-treated plates protected from light in a humidified chamber at room temperature for 30 mins were then recorded at 490nm absorbance using a Tecan i-control plate reader (Infinite® 200 PRO, UK). A mixture of MTS and SDS in serum free medium was used as background controls. All samples were conducted in duplicates in a 96-well plate. Calculated formazan measurement was measured
between treated cells and background control cells. The relative formazan was defined as the ratio of A490nm of Ad-L151 group to Ad-null group with that of Ad-null group unchallenged with H₂O₂ set as 1.0.

2.18 Annexin V/PI staining

ECs (C166, ATCC, UK) were infected with Ad-null or Ad-L151 (10 MOI) virus for 24hr in Dulbecco’s Modified Eagle Medium (Gibco, UK), 10% FBS (ATCC, UK), 100 U/ml penicillin (ThermoFisher, UK), and 100 mg/ml streptomycin (ThermoFisher, UK). Cells were then treated with 50 μM or 100 μM H₂O₂ or H₂O in serum-free condition for 24hr. After H₂O₂ incubation period, cells were subjected to annexin V/PI staining by the annexin V/Dead Cell Apoptosis Kit (Invitrogen, UK) and FACS analysis.

Cells were harvested by trypsinisation and washed in cold 1xPBS. The cell pellets were then resuspended in 5ml 1x annexin-binding buffer in FACS tubes. The cells were re-centrifuged and the supernatant discarded and the cell pellet resuspended in 1X annexin-binding buffer to 500,000 cells/ml to have 100 μL per assay. 100 μg/mL working solution of propidium iodide (PI) nucleic acid binding dye was prepared by diluting 5 μL of the 1 mg/mL PI stock solution (Component B) in 45 μL 1X annexin-binding buffer. In every 100 μL assay, 5 μL Alexa Fluor 488 annexin V and 1 μL 100 μg/mL PI working solution was added and incubated at room temperature for 15 min.

After the incubation period, 400 μL 1X annexin-binding buffer was added, mixed gently and each tube kept on ice. The stained cells were then analysed by FACS, measuring the fluorescence emission at 530 nm (e.g., FL1) and >575 nm (e.g., FL3). In each assay, 50,000 cells were collected at a rate of 100μl/minute and analysed by the FlowJo software. Debris were eliminated using FSC-A vs SSC-A gating and cell doublets and clumps were eliminated by FSC-H vs FSC-A gating.
2.19 Hydrogen peroxide assay

500,000 ECs were infected with adeno-null or adeno-L151 virus (10 MOI) for 24hrs. Media was washed with 1xPBS and cells were exposed to different concentrations of H$_2$O$_2$ (500µM, 100µM, 50µM, 25µM and 0µM) in 1xPBS and incubated at 37 degrees for 15 mins. Then, cell lysates and media solutions were then harvested and samples were subjected to deproteinisation and neutralisation. Intracellular and extracellular H$_2$O$_2$ concentration was measured via a colorimetric hydrogen peroxide assay kit (Abcam, UK). All standards, treated samples and controls were conducted in triplicates in a 96-well plate. Standard samples included known concentrations of H$_2$O$_2$ (25, 50, 100 and 500 µM) in 1xPBS as input controls for extracellular H$_2$O$_2$ quantification. Negative control was included as 1xPBS. 46µl of each standard, treated sample and control was added in each well, with 2 µl OxiRed probe and 2 µl HRP. In total, 50 µl of each reaction mix in the 96-well plate was then incubated at room temperature for 10 min protected from light and absorbance read at 570 nm by a Tecan i-control plate reader (Infinite® 200 PRO, UK). The negative control samples were used as background controls. Calculated standard and treated samples measurement were measured against background controls.

2.20 Trans-aortic constriction model

All animal experiments in this study were performed according to protocols approved by the Institutional committee for the use and care of laboratory animals. Wild type C57bl/6 (12 weeks old) mice were purchased from Charles River (West Sussex, UK) Twelve weeks of C57bl/6 mice were subjected to minimally invasive trans-aortic constriction (TAC) under 2% isoflurane anaesthesia without sternotomy or ventilation as described previously $^{169}$ and conducted by Dr Ajay Shah’s laboratory group. This aortic constriction was performed by ligation of the transverse thoracic aorta with a 27-gauge needle using a 6-0 braided polyester suture $^{170}$. This experimental model is illustrated in figure 19. Then, sham and TAC-induced heart tissues were collected at day 3 and 7 post-
surgery. Immunofluorescence staining and western blot was performed on these tissues as indicated in chapters 2.9 and 2.13.

**Figure 19: Diagram of a trans-aortic constriction model in mice.**
Figure source: Rockman et al. Segregation of atrial-specific and inducible expression of an atrial natriuretic factor transgene in an *in vivo* murine model of cardiac hypertrophy. 1991. *PNAS.* 88: 8277-8281. 10

### 2.21 Femoral artery injury model

The femoral artery injury model was introduced into wild type C57bl/6 (Charles River, UK) and conducted by Dr Yanhua Hu as described previously 171. Mice were weighed and anaesthetised by intraperitoneal injection of ketamine (75 mg/kg; VetalarTMV, Pfizer, UK) combined with medetomidine hydrochloride (1 mg/kg; Domitor®, Orion). A pinch test was performed on the tail of mouse to confirm successful anaesthetisation. Lubricating ointment was applied to the eyes of mice to prevent corneal desiccation. Hair was removed by application of a depilatory cream from both legs to lower abdomen. The mouse was then put on an aseptic operation board and covered with an aseptic drape. The depilated area of skin was sterilized with 70% ethanol and dried with sterile cotton swabs just before surgery.
An incision into the skin over the femoral artery was made before locating femoral artery by blunt dissection of surrounding tissues. The tissues were moistened periodically using saline for irrigation. The femoral artery was isolated by gently separating the femoral nerve and vein from the femoral artery, in sequence using micro-forceps. The anterior branch of the femoral artery was looped with a 10-0 silk suture before performing an arteriotomy using a 30 gauge needle at the distal part of the bifurcation. The opening of the arteriotomy was lifted with micro-forceps before a 0.25 mm guide wire (Cross-IT 100xT, Hi-Torque, Abbot, USA) was introduced into the artery until it could be inserted no further. The wire was inserted and retracted 5 times before it was allowed to remain in the femoral artery for 3 mins. After gently and slowly removing the wire, the artery was immediately ligated. The same procedure was applied on the other femoral artery of each mouse. The wound was closed by sewing. Sham conditions involved incisions into the skin without any further invasive procedures. This experimental model is illustrated in figure 20.

Mice were then recovered from anesthesia by the antidote to ketamine, atipamezole hydrochloride (5 mg/kg; Antisedan, Orion). Mice were kept in a recovery incubator set at 37°C until re-establishment of full consciousness. All mice were monitored continuously during recovery and periodically every 24hr for 5 days. The sham and injured induced femoral arteries were harvested 3 and 7 days post-surgery and sandwiched in liver. This was to keep the integrity and structure of the small vessels for efficient cryo-sectioning and immunofluorescence staining.
**Figure 20: Diagram of a femoral artery injury model in mice.**  
(A) Common femoral artery (arrow), the superficial femoral artery (dotted arrow), and the deep femoral artery (arrow head) are uncovered and looped with sutures. (B) The size differences between the 0.25-inch wire and the arteries. (C) The wire is inserted via the deep femoral artery and retracted 5 times and held in position for 3 mins. (D) The deep femoral artery is ligated, and blood flow is resumed (arrow). Bar indicates 1 mm. Figure source: Takayama et al. A murine model of arterial restenosis: Technical aspects of femoral wire injury. 2015. *Journal of Visualized Experiments*. 97: e52561.

### 2.22 Hindlimb ischaemia model

The hindlimb ischaemia model was performed in C57BL/6J wild type and *Laf4ir* heterozygous mice and conducted by Dr Yanhua Hu as described previously. Mice were weighed, anaesthetised and prepared for surgery in the same manner as the femoral artery injury model described above. In this model, however, the right leg of the femoral artery was ligated permanently. This involved ligation of the femoral artery.
at two adjacent sites with the middle part cut. The wound was closed by sewing. Sham conditions involved incisions into the skin without any further invasive procedures on the left leg of each animal. This experimental model is illustrated in figure 21.

The lower abdomen up to the feet area of each animal’s blood flow was measured by a LDI Doppler laser scanner (Moor Instruments, Devon, UK) following the completion of surgery to confirm successful ligation. This was 30 mins post-surgery of each animal. Mice were then recovered from anaesthesia by the antidote to ketamine, atipamezole hydrochloride (5 mg/kg; Antisedan, Orion). Mice were kept in a recovery incubator at 37°C until re-establishment of full consciousness. All mice were monitored continuously during recovery and periodically every 24hr for 5 days.

At 7 and 14 days post-surgery, the blood perfusion was measured in the same manner as 30 mins post-surgery. After day 14 post-surgery, mice were sacrificed humanely. The adductor muscle tissues around both ligated right leg and non-ligated left legs of each mouse were isolated and stored at -80°C. The tissue samples were subjected to western blot analysis or immunofluorescent staining.
Figure 21: Diagram of a hindlimb ischemia model in mice. (A) 0.5cm incision introduced at the inguinal ligament of a mouse hindlimb and (B) Image of the incision. (C) Representation of the main vessels at the hind limb area. (D) Two occlusions are introduced at the proximal and distal regions of the femoral artery, followed by excision of the full segment. Figure source: Martins et al. Angiogenic Properties of Mesenchymal Stem Cells in a Mouse Model of Limb Ischemia. 2014. Methods in Molecular Biology. 1213: 147-169.

2.23 Computational modelling

Structural model analysis

To further understand the polypeptides produced by laf4ir, in-silico predictive 3D modelling was performed. 5 independent modelling platforms were utilised; template threaded models were produced by RaptorX, IntFold and I-TASSER modelling, while Phyre2 and Swiss modelling both use homology alignment methodology. Protein threading models proteins which have the same fold as proteins...
of known structures but do not account for homologous proteins with known structures. Homology modelling enables to construct an atomic-resolution model with homologous protein structures deposited in the protein data bank. Alignment of the initial models was achieved on PyMol (Schrödinger, USA) to discover structural similarities. The most accurate models from each online service were then validated by SAVES V5.0 (UCLA, USA). SAVES V5.0 runs three tests referred as Verify3D, PROVE and ERRAT. Verify3D tests the 3D position of each residue and calculates a percentage that match the verified 3D-ID. PROVE provides an overall percentage of buried atoms that may clash. ERRAT provides a score of residues that are over 95% improbability. The highest scoring models were then refined to improve their validity. Surface electrostatic calculations and cartoon images obtained were performed by the PyMol software and presented in my PhD project. To predict the secretory capacity of each protein, SecretomeP was also used.

**Post-Translational Modification**

To identify any possible sites of post-translational modification (PTM), multiple bioinformatic web services were used to analyse the sequence. These programmes analyse the sequence for possible enzymatic binding sites that could catalyse the addition of a modification. To test for phosphorylation GPS 3.0, sitePROscan and NetPhos 3.1 were used. Glycosylation was identified with NetNGlyc. SUMOplot was utilised for SUMOylation predictions, and for the addition of lipid molecules GPS-lipid was used. All tests were run at the highest settings and a threshold of above 0.75 was set (site cut-off=0.50), only the highest scoring sites were shown.

**Promoter analysis**

The sequence of the promoter region was analysed for transcription factor (TF) binding sites using TFBIND. Results above 95% identity were identified and displayed. Transcription can be regulated by the formation of G-quadruplex (G4) structures inside
promoter regions. QGRS-H predictor was used to identify any possible G4s inside the promoter.

**mRNA analysis**

To investigate the alternative translation of the 3ORFs, altORFez was used to identify possible ribosome binding sites (RBS).

### 2.24 Statistical analysis

Data were characterised as mean ± standard error of the mean (SEM). All data were analysed using GraphPad Prism 7 software with t-test or one-way analysis of variance (ANOVA) or two-way ANOVA. Figures 32B, 34, 38, 39, 44B, 49, 50B and 50D used t-test analysis. Figures 29, 30, 31, 32D, 36, 43, 45, 46, 48, 51, 52, 55, 56 and 57 used one-way ANOVA analysis. Figure 37 used two-way ANOVA analysis followed by Dunnett’s multiple comparison tests. Significance was depicted by asterisks, *: p<0.05, **: p<0.01, ***: p<0.001. A value of p < 0.05 was considered to be significant.. Image J was used for western blot quantification.
Chapter 3: Expression profile of *Laf4ir*

3.1 Introduction

As described in chapter 1.11, *Laf4ir* RNA sequence contains potential ORF sequences from transcript variant 1 and 2. ORF is a sequence of successive nucleotide triplets that are read as codons that begin with a start codon (e.g. AUG) and end with a stop codon (e.g. UAA, UGA and UAG). Millions of ORF sequences have been identified in eukaryotic genomes and only about 30,000 ORFs are known to be translated. It is now more widely speculated that we have a hidden proteome or peptideome within our genome. Currently, there is a lack of knowledge focusing on whether or not unexplored ORFs contain translational machinery with potential biological functions.

Normally, small ORFs are regarded as non-coding due to their short length, which minimises computational detection of protein-coding capability. Also, there is a traditional antagonising approach to experimentally confirm functionality of unknown protein coding genes, as it is challenging to validate annotation and curation. Issues from computational annotation arise from sequence similarities of ORFs which determine the conservation of the potential coding sequence. Therefore, the shorter the ORF protein sequence, there is a greater probability of obtaining a ‘low conservation score’. For instance, the BLAST search tool disregards the discovery of protein sequences fewer than 80 amino acids and removes those that have fewer than 20. These rules, which excludes potential small ORFs, are part of the annotation guidelines used by majority of all publically available gene sets, such as Aceview, GENCODE, RefSeq, Ensembl, VEGA and CCDS. There is currently a growing trend of studies demonstrating the translation of short ORFs and functions mediated by the novel peptides/proteins that can be broadly conserved across metazoans. Given the high number of short ORFs in the genome and the expectation that most of them
are not functional, a wide window of potential experimental discoveries of novel proteins with diverse functions arises.

The consensus of a eukaryotic mature mRNA is a monocistronic molecule with a tripartite structure: a single translated ORF or coding sequence (CDS) flanked by 5’ and 3’ untranslated regions (UTRs). Initially, a 43S preinitiation complex binds to the cap structure of the 5’end and scans the 5’UTR to find the translation initiation site (TIS). Then, the large 60S subunit joins to form a fully functional 80S ribosome with tRNA to initiate polypeptide synthesis. Consequently, it was established prior to genomic and proteomic experimental advances that each eukaryotic mature mRNA translates only one polypeptide/protein. This traditional viewpoint that ribosomes translate only one ORF (usually the longest sequence) and produce one protein/polypeptide, has been fundamentally challenged by experimental evidence for the translation of alternative ORFs (altORFs) by mass spectrometry and ribosome profiling. Currently, multiple altORFs have been discovered in mammals with diverse expression profiles and functions. These include INK4a/ARF, histone H4/OGP, XLalphas/ALEX, RPP14/HsHTD2, PrP/altPrP, ATXN1/altATXN1, AZAR/uORF5, MKKS/uMKKS1 and uMKKS2, and AT1aR/PEP7. To bypass the AUG codon of an altORF and initiate protein synthesis of the downstream reading frame, a few mechanisms (e.g. leaky scanning, reinitiation, and cellular internal ribosome entry site (IRES)) have been proposed.

Potential altORFs candidates can be predicted by in silico detection strategies which is mainly based by sequential filters with different stringencies, including a cut off size (150 or 500 nucleotides), conservation between species and the presence of a strong Kozak signal around the predicted TIS. Almost all altORFs occur at an upstream AUG codon in relation to the position of the main ORF, usually within an optimal kozak content. However, in rare cases altORFs can occur at a downstream AUG codon with optimal kozak signature in mammals. The human transcriptome
comprises of 83,886 potential altORFs with a minimum size of 40 codons while the current human proteome comprises of around 52,000 annotated proteins (RefSeq release 72). As accumulating evidence indicate the discoveries of unknown multiple proteins from the same mRNA, this unique mechanism may suggest that the *Laf4ir* gene could also contain this translational machinery.

The largest ORF sequences may indicate candidate polypeptide coding regions in the *Laf4ir* mRNA sequence. Computational analysis using ExPASy bioinformatics (https://www.expasy.org/) software for translational tool indicated that the *Laf4ir* mRNA sequence possesses three potential ORFs as indicated in Figure 12-14 in Chapter 1.11. *Laf4ir* ORF1 and ORF3 encode for 45 amino acids (aa) and 82-aa respectively in both transcript variants. *Laf4ir* ORF2 encodes polypeptides with 109-aa and 151-aa in *Laf4ir-tv1* and *Laf4ir-tv2* respectively. The 109-aa and 151-aa polypeptides of *Laf4ir* ORF2 share the same N-terminal but differ in C-terminal due to a stop codon in exon 4. Also the ExPASy bioinformatics website (https://web.expasy.org/compute_pi/), enables researchers to compute the theoretical pl (isoelectric point) and Mw (molecular weight) of a specific amino acid sequence of interest. Computational predicted results are indicated in table 11.

Table 11: Computational calculation of isoelectric point and molecular weight of LAF4IR polypeptides

<table>
<thead>
<tr>
<th>LAF4IR polypeptide</th>
<th>Amino acid sequence</th>
<th>Isoelectric point (pl)</th>
<th>Molecular weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF1 (45aa)</td>
<td>MYWGEASRQDWRSTSEQRNRRPWSVKTSRKRILESRYWNSAAV</td>
<td>10.83</td>
<td>5.54316</td>
</tr>
<tr>
<td>ORF2 from TV1 (109aa)</td>
<td>MPDLNCEVLIVIQQPISRNERIALIFASADHRSTPOAGKVRTRINSRTSHNLSTDSTQASHKAESQPKPCGLCTEHRGRQLNKIVAWCSGLLEPSFSKFRHFYPESFL</td>
<td>9.3</td>
<td>12.297</td>
</tr>
<tr>
<td>ORF2 from TV2 (151aa)</td>
<td>HKAESQPKPCGLCTEHRGRQQLNKIVAWCSGLLEPSFSKRGRGSRHQVLRHPHTQTVRWGTSANIRTVTLRITPEDVVRHPGCCPESLPS</td>
<td>10.52</td>
<td>16.87921</td>
</tr>
<tr>
<td>ORF3 (82aa)</td>
<td>MGECYFCKGSHCAENTRCGKTESSWLLSRSFALLKEGCYVAQ2PPVNDGFPTRQVEEVNPQSHATVPEMGPPIPGQ</td>
<td>5.58</td>
<td>9.05032</td>
</tr>
</tbody>
</table>

The *Laf4ir* ORF1 polypeptide is an alkaline peptide (pl= 10.83) as it contains 10 alkaline aa and 5 acidic aa. Both *Laf4ir* ORF2 polypeptides from TV1 (pl= 9.3) and TV2
(pI= 10.52) are alkaline polypeptides with 19 or 31 alkaline aa (Histidine, Lysine & Arginine) compared to 10 or 12 acidic aa (Aspartic acid & Glutamic acid) in the 109-aa or 151-aa sequences. There are three and five cysteine aa in the 109-aa and 151-aa polypeptides, respectively. Besides the intra-molecule disulphide bond(s) between two cysteine residues, both polypeptides have a spare cysteine residue to form homo- or hetero-dimers via inter-molecular disulphide bonds. The ORF3 polypeptide was not further investigated because the mass spectrometry data illustrates that the ORF3 polypeptide is not translated in dESCs.

3.2 Aims

**Hypothesis**

*Laf4ir* is an encoding gene from three potential ORFs with a diverse expression profile.

**Aims**

To demonstrate that *Laf4ir* is an encoding gene and assess the *Laf4ir* ORF1 and ORF2 polypeptide expression profile at different embryonic stages, in adult organs, and cells, and at sub-cellular level.
3.3 Results

3.3.1 *Laf4ir* is a novel encoding gene

To explore whether *Laf4ir* is an encoding nested gene like *Pcp*\(^{135}\), the cDNA sequence for the 151-aa of *Laf4ir* ORF2 was synthesized (Genscript, UK) with FLAG tag sequence inserted downstream the ATG start codon and cloned to pShuttle2 vector, designated as pShuttle2-*L151*, respectively (Figure 22). Transfection of pShuttle2-*L151* vector in ESCs and immunostaining showed FLAG positive cells (Figure 19). This result indicates the artificial protein coding ability of *Laf4ir* via ORF2.

![Figure 22: pShuttle2-L151 vector transfection and immunostaining.](image)

The *Laf4ir*-tv2 cDNA sequence for the 151-aa of *Laf4ir* ORF2 was cloned into a pShuttle2 vector, in which a FLAG tag was inserted into the N-terminal of the *Laf4ir* ORF2 (Upper panel). The constructed pShuttle2-*L151* vector was transfected in ESCs. After 24hrs, cells were subjected to immunocytochemistry with anti-FLAG FITC conjugated antibody (α-FLAG). The nuclei were counterstained with DAPI and images obtained under confocal microscope. Data presented were representative images of three independent experiments. Scale bar: 5μm.
To demonstrate the encoding capability of *Laf4ir* in a non-artificial setting, proteins were extracted from 3-day spontaneously differentiated collagen IV coated-ESCs and separated by SDS-PAGE. The protein bands between 3kDa and 25kDa were cut and subjected to high pressure liquid chromatography mass spectrometry analysis. Two and eight peptide fragments were detected from *Laf4ir* ORF1 and ORF2 aa sequences respectively. Identification details for the ten peptides are summarized in table 12. No peptide fragments were detected from ORF3, illustrating that ORF3 polypeptide is not expressed. All ten peptides were identified with a high confidence and their XCorr scores were between 0.57 and 3.43. XCorr (Cross Correlation), is a measure of the goodness of fit of the experimental peptide fragments (b- and y-ions) to the theoretical spectra created. Therefore, the higher the number, the more confident the match.

From the 8 peptide fragments of *Laf4ir* ORF2, 5 were shared by *Laf4ir*-tv1 and *Laf4ir*-tv2, while 3 peptide fragments were from the C-terminal of *Laf4ir*-tv2 (Figure 23). Eight peptides detected against ORF2 by *Laf4ir*-tv2 sequence has a coverage of 79% and five peptides detected against ORF2 by *Laf4ir*-tv1 has a coverage of 72% and two peptides detected against *Laf4ir* ORF1 has a coverage of 31%. These results indicate that both *Laf4ir* ORF1 and ORF2 can be translated from *Laf4ir* mRNA and *Laf4ir*-tv2 may be the main transcript variant. This mass spectrometry data shows that *Laf4ir* is an encoding gene with two polypeptides (ORF1 and ORF2).

In order to investigate any partners associated to *Laf4ir* ORF2 TV1 and TV2 amino acid sequence, Immunoprecipitation with *Laf4ir* anti-ORF2 antibody plus proteomics analysis was performed. The raw data was searched against the current version of the reviewed SwissProt mouse database downloaded from Uniprot ([www.uniprot.org/uniprot/](http://www.uniprot.org/uniprot/)). A total number of 428 protein partners (Master class) were identified with at least 1 peptide within the three biological replicates and thus, the identified proteins common in three replicates were used. This raw data provides an array of information regarding potential functions of *Laf4ir* in dESCs.
Table 12: Mass spectrometry data identifying peptide sequences of *Laf4ir* ORF1 and ORF2

<table>
<thead>
<tr>
<th>Peptide ID</th>
<th>Mass spectrometry sequence</th>
<th><em>Laf4ir</em> ORF-sequence</th>
<th>Theo M<em>H</em> [Da]</th>
<th>Confidence: Sequest HT</th>
<th>Xcorr: Sequest HT</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>AESQKPVCLQETEHGR</td>
<td>ORF2 - TV1 and TV2</td>
<td>1868.89</td>
<td>High</td>
<td>2.58</td>
</tr>
<tr>
<td>P2</td>
<td>GSRHQVLHPQHTVRTWGSVTANR</td>
<td>ORF2 - TV2</td>
<td>2867.50</td>
<td>High</td>
<td>2.82</td>
</tr>
<tr>
<td>P3</td>
<td>INSRTSHNLSTDQTMLASHK</td>
<td>ORF2 - TV1 and TV2</td>
<td>2110.06</td>
<td>High</td>
<td>3.43</td>
</tr>
<tr>
<td>P4</td>
<td>MDPOLNECVIQPQNER</td>
<td>ORF2 - TV1 and TV2</td>
<td>2198.09</td>
<td>High</td>
<td>2.15</td>
</tr>
<tr>
<td>P5</td>
<td>STPOAGKVNR</td>
<td>ORF2 - TV1 and TV2</td>
<td>1057.57</td>
<td>High</td>
<td>2.02</td>
</tr>
<tr>
<td>P6</td>
<td>VNRTRINSR</td>
<td>ORF2 - TV1 and TV2</td>
<td>1115.64</td>
<td>High</td>
<td>2.16</td>
</tr>
<tr>
<td>P7</td>
<td>WGSVTANRVTVR</td>
<td>ORF2 - TV2</td>
<td>1432.77</td>
<td>High</td>
<td>2.09</td>
</tr>
<tr>
<td>P8</td>
<td>HQVLRHPQHTVRTWGSVTANR</td>
<td>ORF2 - TV2</td>
<td>2567.35</td>
<td>High</td>
<td>0.57</td>
</tr>
<tr>
<td>P9</td>
<td>RPWSVK</td>
<td>ORF1</td>
<td>772.45</td>
<td>High</td>
<td>1.01</td>
</tr>
<tr>
<td>P10</td>
<td>STSEQQRNR</td>
<td>ORF1</td>
<td>977.48</td>
<td>High</td>
<td>1.53</td>
</tr>
</tbody>
</table>

**Figure 23: Laf4ir ORF1 and ORF2 detection with other peptide sequences from mass spectrometry**

Proteomics analysis detected 2 and 8 peptides from *Laf4ir* ORF1 and ORF2 sequences respectively. No peptides were detected for ORF3 sequence. The position of the peptide sequences (P: 1-10) was indicated in the amino acid sequence of all ORFs with designated colours. Proteins were isolated from dESCs and run on 4-12% SDS-PAGE gel, the protein bands corresponding to 3KDa to 25KDa were cut, in-gel digested with trypsin and then analysed by mass spectrometry. Data presented were representative images of three independent experiments.
3.3.2 Computational model of *Laf4ir* promoter and mRNA

To predict potential transcriptional and translational regulations of *Laf4ir*, computational analysis of the promoter and mRNA sequence can provide clues for potential transcription factor candidates and ribosome binding sites. The TFBIND websolver provides information regarding the potential binding sites of known transcription factor (TFs) found in the Transfac database. It was found that in total 22 TFs can bind to the promoter of *Laf4ir* with ≥ 95% identity to the sequence. Interestingly, only 2 TFs had a score of 100% identity which illustrates a strong probability that these two TFs could regulate *Laf4ir* transcription. These were NKX 25 at base 849 and CDXA at base 1268 (Figure 24 in red). The existence of G4 in the promoter was also analysed and 12 G4 structures of varying size and stability were predicted to exist in the promoter region (Figure 24 in blue).

Analysis of the mRNA sequence was also performed to determine possible regulation at the translational level. The 2 ORF’s of mRNA *laf4ir* are driven by alternative translation and to predict the possible routes by which each ORF polypeptide is synthesised, ribosome binding sites were analysed using altORFez (Figure 25). This program found the potential existence of 5 possible ribosome binding sites. These different sites may suggest how ribosomes can mediate the production of more than one polypeptide from the same mRNA.
Figure 24: Bioinformatics analysis of *Laf4ir* promoter with predicted transcription factor binding sites.
Potential TF candidates and their binding sites are shown in red, while predicted G-quadruplex structures are shown in blue.

Figure 25: Bioinformatics analysis of *Laf4ir* mRNA TV1 with predicted ribosome binding sites
Diagram displaying a total of five ribosome binding sites on the *Laf4ir* mRNA from TV1 and the 3ORFs. RBS: Ribosome binding site. #1-5: Number of each sites. Nucleotide position of *Laf4ir* TV1 mRNA: (35-97-250-466-492).
3.3.3 Computational model of Laf4ir ORF1, ORF2-TV1 and ORF2-TV2 polypeptide sequences

To assist in the prediction of potential functions of Laf4ir, computational analysis of the structural integrity of each polypeptide can be useful. Laf4ir ORF1 polypeptide was predicted to have a hairpin structure containing 2 α-helix’s (H1= 9aa, H2=7aa) with 2 residues on each helix interacting between them (Figure 26A). Its structure also contains 7 β and 4 γ turns (Figure 26A). Laf4ir ORF1 polypeptide was predicted to have 11 possible phosphorylation sites by various different kinases including casein kinase 1 (CK1), Thymidine kinase (TK), and a family of serine/threonine protein kinases (AGC) (Figure 26C). Laf4ir ORF1 polypeptide may also reserve a glycate site. This structural prediction indicates a cell cycle or cellular signalling role mediated by the potential phosphorylation of Laf4ir ORF1 polypeptide.
Figure 26: Predicted computational model of *Laf4ir* ORF1 polypeptide. 
(A) Cartoon model of peptide backbone. N’ to C’ terminal runs from Blue to Red. Images run from left to right in 90° increments. (B) Vacuum electrostatic surface model of *Laf4ir* ORF1 polypeptide, showing the negative (red) and positive (blue) charged regions available for binding. The range of electrostatic charge is shown in the bar below. (C) Table illustrates potential kinase interactions and other predicted PTM.
Structural analysis of the *Laf4ir* ORF2 polypeptide sequence derived from TV1 model revealed the predicted structure consists of 3 α-helix, (H1=6aa both H2/H3=14aa) (Figure 27A). 6aa interact between each large helix. There is one predicted parallel β-sheet containing 2 strands both 3aa in length and 1 β-α-β-motif. The rest of the structure contains 16 β-turns and 11 γ turns (Figure 27A). 16 sites for phosphorylation and 9 glycosylation sites were also predicted (Figure 27C). This indicates the potential association of *Laf4ir* ORF2-TV1 polypeptide with membrane bound components or its role in cell cycle and signal cascades via glycosylation and phosphorylation at multiple sites.
Figure 27: Predicted computational model of *Laf4ir* ORF2-TV1 polypeptide. (A) Cartoon model of peptide backbone. N’ to C’ runs from Blue to Red. Images run from left to right in 90° increments. (B) Vacuum electrostatic surface model of *Laf4ir* ORF1, showing the negative (red) and positive (blue) charged regions available for binding, range of electrostatic charge is shown in the bar below. (C) Table illustrates potential kinase interactions and other predicted PTM.
Laf4ir ORF2 polypeptide derived from TV2 was predicted to have a more complex structure due to the largest aa sequence. Laf4ir ORF2-TV2 polypeptide was found to have 11.9% α-helix containing 3 small helical regions (H1=5aa, H2=7aa, H3=6aa) (Figure 28A). The rest of the structure is highly convoluted consisting of 47 β-turns held together with 7 hydrogen bonds and 13 inverse γ-turns (Figure 28A). PTM analysis predicted a total of 24 phosphorylation sites (Figure 28C). These numerous phosphorylation sites included AGC, CK1, and MAPK. 8 glycosylation sites and 4 sites of lipid modification were also predicted. This indicates the potential association of Laf4ir ORF2-TV2 polypeptide with membrane bound components or its role in cell cycle and signal cascades via glycosylation and phosphorylation at multiple sites.

SecretomeP website analysis predicts a high probability of extracellular secretion\textsuperscript{183}. SecretomeP (SecP) score for Laf4ir ORF2-TV2 polypeptide was calculated to be 0.917 and the recommended threshold is set at 0.6 for mammalian sequences. Therefore, scores greater than 0.6 are known to have a higher probability to be secreted within the mammalian system. Table 13 summarises the predicted structures, post-translational modifications and functions of Laf4ir ORF1, ORF2-TV1 and ORF2-TV2 polypeptides.
Figure 28: Predicted computational model of Laf4ir ORF2-TV2 polypeptide. (A) Cartoon model of peptide backbone. N' to C' runs from Blue to Red. Images run from left to right in 90° increments. (B) Vacuum electrostatic surface model of ORF2-TV2 polypeptide, showing the negative (red) and positive (blue) charged regions available for binding, range of electrostatic charge is shown in the bar below. (C) Table illustrates potential kinase interactions and other predicted PTM.
Table 13: Computational analysis of *Laf4ir* polypeptides

<table>
<thead>
<tr>
<th></th>
<th>ORF1</th>
<th>ORF2-TV1</th>
<th>ORF2-TV2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predicted</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>structure</td>
<td>2 α-helixes</td>
<td>3 α-helixes</td>
<td>3 α-helixes</td>
</tr>
<tr>
<td></td>
<td>7 β turns</td>
<td>16 β turns</td>
<td>47 β turns</td>
</tr>
<tr>
<td></td>
<td>4 γ turns</td>
<td>11 γ turns</td>
<td>13 γ turns</td>
</tr>
<tr>
<td><strong>Predicted</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>post-translational</td>
<td>11 Phosphorylation sites</td>
<td>16 Phosphorylation sites</td>
<td>24 Phosphorylation sites</td>
</tr>
<tr>
<td>modifications</td>
<td>1 Glycate site</td>
<td>9 Glycosylation sites</td>
<td>8 Glycosylation sites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Sumoylation site</td>
<td>1 Glycation sites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Glycation site</td>
<td>1 Myristylation site</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Palmitoylation site</td>
<td>2 Farnesylation sites</td>
</tr>
<tr>
<td><strong>Predicted</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>functions</td>
<td>Cell cycle or cellular</td>
<td>Membrane bound or cell cycle or</td>
<td>Membrane bound or cell cycle or</td>
</tr>
<tr>
<td></td>
<td>signalling</td>
<td>cell cycle or cellular signalling</td>
<td>cell cycle or cellular signalling</td>
</tr>
</tbody>
</table>

### 3.3.4 *Laf4ir* expression in adult organs

To understand the potential functions of *Laf4ir* in different systems, it is vital to firstly explore the *Laf4ir* expression in different adult organs. This will give us clues of other unknown functions of *Laf4ir* in diverse mammalian systems. Western blot analysis detected approximately 17kDa and 16kDa bands for *Laf4ir* ORF1 and ORF2 polypeptides respectively in different organs (Figures 29-30). *Laf4ir* ORF1 polypeptide has a sequence of 45aa with a predicted molecular mass of 5.54316 kDa in Table 11 in chapter 3.1. The 17kDa molecular weight of *Laf4ir* ORF1 polypeptide suggests that there may be heavy modifications existing in this *Laf4ir* ORF1 polypeptide. No detection of around 5 kDa band might be due to the low molecular weight. Optimization of western blot procedure will be required to detect low molecular weight proteins. *Laf4ir* ORF2 polypeptide from TV1 has a sequence of 109aa with a predicted molecular mass of 12.297 kDa (Table 11), and *Laf4ir* ORF2 polypeptide from TV2 has a sequence of 151aa with a predicted molecular mass of 16.87921 kDa (Table 11). Western blot analysis of around 16kDa illustrates the native size of *Laf4ir* ORF2 polypeptide from TV2, suggesting that the *Laf4ir* TV2 may be the main mRNA transcript for the translation of *Laf4ir* ORF2 polypeptide in a wide range of organs.
To verify the specificity of the *Laf4ir* ORF1 and ORF2 antibodies used in this investigation, blocking peptides against both *Laf4ir* ORF1 and ORF2 antibodies were designed by Genscript and utilised in this experiment. Western blot analysis showed the specificity of both *Laf4ir* ORF1 and ORF2 antibodies due to the absence of any bands when incubated with blocking peptide (Figures 29-30). Therefore, it was ideal to utilise these antibodies for further experiments in this project.

*Laf4ir* ORF1 polypeptide expression was found in some organs such as the lung, liver, spleen and kidney (Figure 29). *Laf4ir* ORF1 polypeptide expression was absent in the heart, muscle, aorta, brain and stomach (Figure 29). This data suggests that the *Laf4ir* ORF1 polypeptide may have a role in pulmonology, hepatology, nephrology and the lymphatic system. *Laf4ir* ORF2 polypeptide expression, on the other hand, was more ubiquitous and found in more organs such as the heart, lung, liver, spleen, kidney, intestine, aorta, brain and stomach (Figure 30). *Laf4ir* ORF2 polypeptide expression was significantly higher in the aorta, compared to the heart (Figure 30) suggesting a greater contribution to vascular function than cardiac function. The bigger size bands indicate multiple post-translational modifications. It could also mean that the protein of interests (ORF1 and ORF2) are bound to a complex of proteins. Furthermore, sumoylation of *Laf4ir* ORF2 polypeptide from TV2 will result in the total protein size of around 29 kDa. This band size is indicated in the heart, lung, liver, spleen, kidney and brain samples (Figure 30). This data suggests a ubiquitous role of *Laf4ir* in the mammalian adult system.
Figure 29: Laf4ir ORF1 expression in different adult organs

(A) C57Bl/6J 8 week old mice organs were harvested. Total protein content of each organ was isolated and subjected to western blot analysis using Laf4ir ORF1 antibody (α-ORF1) and Laf4ir ORF1 antibody with BP (blocking peptide). The Laf4ir ORF1 polypeptide band is shown with an arrow with 17kDa size. GAPDH was used as loading control. To verify specificity Laf4ir ORF1 antibody, blocking peptide was incubated with Laf4ir ORF1 antibody for 24hr in room temperature prior to western blot analysis. (B) The total level of Laf4ir ORF1 expression in each organ was quantified in relation to lung Laf4ir ORF1/GAPDH ratio expression level set as 1.0. The data presented were representative images of 3 mice. Mean ± SEM. *: p<0.05.
Figure 30: *Laf4ir ORF2* expression in different adult organs.
(A) C57Bl/6J 8 week old mice organs were harvested. Total protein content of each organ was isolated and subjected to western blot analysis using *Laf4ir ORF2* antibody (α-ORF2) and *Laf4ir ORF2* antibody with BP (blocking peptide). The *Laf4ir ORF2* polypeptide band is shown with an arrow with 16kDa size. GAPDH was used as loading control. To verify specificity *Laf4ir ORF2* antibody, blocking peptide was incubated with *Laf4ir ORF2* antibody for 24hr in room temperature prior to western blot analysis. α=antibody. BP= Blocking peptide. (B) The total level of *Laf4ir ORF2* expression in each organ was quantified in relation to heart *Laf4ir ORF2/GAPDH* ratio expression level set as 1.0. The data presented were representative images of 3 mice. Mean ± SEM. *: p<0.05, **: p<0.01, ***: p<0.001.
3.3.5 \textit{Laf4ir} expression during embryonic development

To understand the potential role of \textit{Laf4ir} during embryonic development, it is vital to explore the \textit{Laf4ir} expression trend from early to late embryonic stages. During embryonic development, the CVS development is known to take place at embryonic days (E8.5, 10.5, 14.5, 16.5 and 19.5). The cardiogenic plate begins to form at E8-8.5 with cardiac myocyte cells developing at E10-10.5 \textsuperscript{221,222}. At E14-14.5, various vessels begin to develop such as the aorta, vena cava and coronary arteries, and at later stages E16-19 all vascular and cardiac systems are modified \textsuperscript{222}. To identify expression trend of \textit{Laf4ir} during embryogenesis, total protein content of mouse embryos at different stages of embryonic development was harvested and subjected to western blot analysis.

Results show a moderate level of expression of 17 kDa \textit{Laf4ir} ORF1 polypeptide at embryonic days E8.5 and E10.5. Between embryonic day E10.5 and E14.5 there is a significant increase in \textit{Laf4ir} ORF1 polypeptide expression (Figure 31). This increase in expression is almost absent at E16.5 (Figure 31). However, \textit{Laf4ir} ORF1 polypeptide expression then increased at the last day of embryonic development (E19.5) (Figure 31).

At E8.5 and E10.5, 16kDa \textit{Laf4ir} ORF2 polypeptide expression remains at a low level. Between embryonic day 10.5 and 14.5 there is a significant increase in \textit{Laf4ir} ORF2 polypeptide expression (Figure 31). This increase in expression is sustained up to the last day of embryonic development (E19.5) (Figure 31). Overall, this data indicates an important role of \textit{Laf4ir} ORF1 polypeptide in middle stages of embryonic development (specifically at E14.5) and \textit{Laf4ir} ORF2 polypeptide in later stages of embryonic development.
Figure 31: *Laf4ir* ORF1 and ORF2 polypeptide expressions at different stages of embryonic development.

(A) C57Bl/6J mice embryos were harvested at embryonic days (E) 8.5, 10.5, 14.5, 16.5 and 19.5. Total protein content of each embryo was isolated and subjected to western blot analysis using *Laf4ir* ORF1 and ORF2 antibodies. GAPDH was used as loading control. (B) The total level of *Laf4ir* ORF1 and ORF2 expression was quantified in relation to GAPDH expression level. The data presented were representative images of 3 mice. Mean ± SEM. *: p<0.05. **: p<0.01. ****: p<0.0001. NS: Non significant
3.3.6 *Laf4ir* expression in different cell types

To understand the contribution of *Laf4ir* in vascular regeneration, it is important to explore the *Laf4ir* expression in stem/progenitor cell types. This will give us clues in which cell types *Laf4ir* mediates its functions. *mRNA* *Laf4ir* was originally identified in ESCs and FACS analysis using antibodies specific against *Laf4ir* ORF1 and ORF2 showed *Laf4ir* ORF1 and ORF2 polypeptide expression in ESCs *in-vitro* (>90% ESCs) (Figure 32A-B).

As stem cells lose their pluripotency and self-renewal state and differentiate towards a progenitor phenotype, *Laf4ir* expression may alter due to alterations in functions. Therefore, it was important to observe the expression of *Laf4ir* ORF1 and ORF2 polypeptide expression in dESCs. I also wanted to see if the parent gene is also expressed in dESC. This will provide information if the parent gene (*LAF4*) and nested gene (*Laf4ir*) are expressed in the same cell type. Previous data (Figure 10B) illustrated the upregulation of *Laf4ir* mRNA in response to laminar flow in the previous shear stress model, and therefore I was also interested to see if expression level changes as well in this altered shear stress model. This shear stress model, illustrated in chapter 2.3 (Figure 16), enables dESCs to be exposed to both disturbed and laminar flow shear stress that is more physiologically relevant to *in-vivo* systems. Western blot analysis illustrated the expression of LAF4 protein, *Laf4ir* ORF1, and *Laf4ir* ORF2 polypeptides in dESCs (Figure 32C-D). There was a slight increase in *Laf4ir* ORF1 and *Laf4ir* ORF2 polypeptide expression among cells subjected laminar and disturbed flow, but slight decrease in LAF4 protein expression when compared to static control (Figure 32C-D).

It was important to explore the expression level of *Laf4ir* ORF1 and ORF2 polypeptides in VPCs. To verify the population of VPCs as Sca-1⁺, FACS analysis using two different Sca-1 antibodies with two different epitopes of Sca-1 protein were conducted. Results revealed around 95-99% of VPCs positive for Sca-1 (Figure 33A-B).
FACS analysis showed *Laf4ir* ORF1 polypeptide positive cells in around 70% of VPCs and *Laf4ir* ORF2 polypeptide positive cells in almost all VPCs (Figure 49 from chapter 5.3.2). The percentage of ORF2 polypeptide positive cells were significantly higher than *Laf4ir* ORF1 polypeptide positive cells. These results demonstrate the expression of *Laf4ir* polypeptides in ESC, dESC and VPC *in-vitro*.
Figure 32: LAF4 protein, Laf4ir ORF1 and ORF2 polypeptide expressions in embryonic stem cells.

(A) ESCs were cultured in maintenance media and harvested for FACS analysis. Cells were stained with Laf4ir mouse-ORF1 and rabbit anti-ORF2 antibodies and donkey anti-mouse and anti-rabbit Alexa Fluor 647 secondary antibodies. Control samples consisted of cells incubated with secondary antibody only. Blocking peptide (BP) was incubated with Laf4ir ORF1 and ORF2 antibodies for 24hrs, followed by staining to show the specificity of Laf4ir ORF2 antibody. 50,000 cells were collected at a rate of 100μl/minute. Data was quantified by FlowJo software. 

(B) Bar chart illustrates the total percentage of EPCs positive for Laf4ir ORF1 and ORF2 polypeptides. Data was presented by GraphPadPrism software. Data presented were representative of three independent experiments. Mean ± SEM. 

(C) ESCs were treated with differentiation media for 3 days and then subjected to laminar flow (LF) and disturbed flow (DF) on an orbital shaker at 150rpm for 24hr, followed by western blot analysis with antibodies indicated. Static (ST) conditions was used as control. Tubulin was used as a loading control. 

(D) The total level
of LAF4, Lof4ir ORF1 and ORF2 expression in dESCs was quantified in relation to tubulin expression level. The relative ratio of each sample was against static control which was set as 1.0. Data was presented by GraphPadPrism software. Data presented were representative of three independent experiments. Mean ± SEM.

Figure 33: Vascular progenitor cells characterised by Sca-1 expression. (A) Flow cytometry analysis conducted using mouse anti-Sca1 conjugated FITC antibody from the immunomagnetic microbeads kit (Miltenyi Biotec, Bergisch Gladbach, Germany) and rat anti-Sca-1 antibody followed by anti-rat PE-Cy7 fluorophore antibody in mouse VPCs. The epitopes for both antibodies against Sca-1 were different. 100,000 single cells were collected at a rate of 100μl/minute. Data was quantified by FlowJo software. (B) The total percentage of Sca-1+ positive cells was quantified. Data was presented by GraphPadPrism software. Data presented were representative of three independent experiments. Mean ± SEM.
**Figure 34: *Laf4ir* ORF1 and ORF2 polypeptide positive cells in VPCs.**

(A) VPCs were cultured in maintenance media and harvested for FACS analysis. Cells were stained with *Laf4ir* mouse-ORF1 and rabbit anti-ORF2 primary antibodies and donkey anti-mouse and anti-rabbit Alexa Fluor 647 secondary antibodies. Control samples consisted of cells incubated with secondary antibody only. Blocking peptide (BP) was incubated with *Laf4ir* ORF1 and ORF2 antibodies for 24hrs, followed by staining to show the specificity of *Laf4ir* ORF2 antibody. 20,000 cells were collected at a rate of 100μl/minute. Data was quantified by FlowJo software. (B) Bar chart illustrates the total percentage of VPCs positive for *Laf4ir* ORF1 and ORF2 polypeptides. Data was presented by GraphPadPrism software. Data presented were representative of three independent experiments. Mean ± SEM. **: p<0.01
3.3.7 *Laf4ir* expression in bone marrow-derived cells

It is well known that during vascular repair mechanisms, inflammatory mediators are released from an array of leukocytes. Therefore, it is important to identify and quantify the number of bone marrow-derived leukocytes and explore the level of expression of *Laf4ir*. To test this, a protocol for the preparation and FACS analysis of wild type bone marrow cells was conducted. Table 14 shows the total percentage of patent population of subtypes, cell markers used to identify these leukocytes and total percentage of *Laf4ir* ORF2*+* positive cells within each of the cell type identified. The total percentage of *Laf4ir* ORF1*+* positive cells were not analysed due to technical difficulties with *Laf4ir* ORF1 antibody. The gating and cell identification strategy for the staining protocol is shown in Figure 35 A-E and is as follows:

Debris were eliminated using FSC-A vs SSC-A gating (R1) and cell doublets and clumps were eliminated by FSC-H vs FSC-A gating (R2) and after *Laf4ir* ORF2*+* single bone marrow cells were quantified. Majority of bone marrow cells were identified to be *Laf4ir* ORF2*+* (64.2%) (Figure 35A). Next, all leukocytes were identified based on CD45 (R3). As expected, 97.8% of bone marrow cells were identified to be CD45*+* leukocytes. Neutrophils are distinguished from all the other leukocytes (R4) based on their expression of the neutrophil-specific marker, Ly6G. As expected, a good proportion of leukocytes were neutrophils (45.4%) and within the neutrophil population 75.3% were *Laf4ir* ORF2*+* (Figure 35B). Positive staining for either CD11b or CD11c on non-neutrophils is then used to distinguish the remaining myeloid leukocytes (R5) from CD11b*−* CD11c*−* lymphoid cells (R6). This specific gating does not account for the small population of lymphocytes that are CD11c*+,* but these will be identified in subsequent gating. The CD11b*−* CD11c*+* lymphoid cells can be further differentiated into MHC class II (IA/IE)*+ CD24*+* B cells and IA/IE*− CD24*−* T cells. As expected, a small population of bone marrow cells were identified as T cells (9.3%) and B cells (4.7%). It was revealed that 68.4% of T cells and 78.2% of B cells were positive for *Laf4ir* ORF2 (Figure 35C).
The myeloid leukocytes (R5) are first subdivided based on their pattern of side scatter vs IA/IE expression. The IA/IE- SCClo population (R7) contains a small population of monocytes. These can be distinguished based on their expression of CD11b and CD64: lymphocytes are CD11b- CD64-, and monocytes are CD11bhi CD64int. The monocytes can be further subdivided into Ly6Chi CD11c- inflammatory monocytes (5.9% in total) and Ly6C- CD11c+ resident monocytes (15.3% in total) (Figure 35D). It was revealed that 68.8% of inflammatory monocytes and 71.1% of resident monocytes were positive for Laf4ir ORF2 (Figure 35D).

The IA/IE+ or SCChi cells within R5 (R8) consist of dendritic cells (DCs) and eosinophils. To distinguish DCs and eosinophils, the expression of CD24 and CD64 were utilised. Within this R9 gate, 3 cell populations can be distinguished based on their expression of MHC class II and CD11b: IA/IE- CD11b+ eosinophils (0.0031%), IA/IE+ CD11b- DCs (0.6%), and IA/IE+ CD11b+ DCs (0.8%) (Figure 35E). It was revealed that 33.3% of eosinophils, 5.5% of CD11b+ DCs and 51.3% of CD11b- DCs were positive for Laf4ir ORF2 (Figure 35E). This FACS result illustrates a high abundance of Laf4ir ORF2 positive bone marrow derived immune cells with a potential wide range of immune functions.
**Figure 35: Flow cytometric analysis of *Laf4ir* ORF2⁺ bone marrow-derived cell types.**

Contour plots of windows and gating strategy used for the identification of major immune cell population in wild type 12 week old mouse bone marrow cells. Gates containing multiple cell populations were numbered (R1-R9). Gates containing a single cell population are labelled with the included cell type. These include (A) total single cells, (B) neutrophils, (C) T cells, B cells, (D) inflammatory monocytes (iMono), resident monocytes (rMono), (E) eosinophils, CD11b⁺ and CD11b⁻ dendritic cells. Each cell type were characterised by specific markers including CD45, Ly6G, CD11c, CD11b, IA/IE, CD64 and CD24. Rabbit anti-ORF2 and donkey anti-rabbit 488 antibody was utilised to quantify number of *Laf4ir* ORF2⁺ cells. Blocking peptide (BP) was incubated with *Laf4ir* ORF2 antibody for 24hrs at a 1:1 ratio concentration, followed by staining to show the specificity of *Laf4ir* ORF2 antibody. 100,000 cells were collected at a rate of 100μl/minute. Data was quantified by FlowJo software.
Table 14: Flow cytometric data of Laf4ir ORF2\(^+\) cells in different bone marrow-derived cell types

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Parent population cells (%)</th>
<th>Cell markers</th>
<th>ORF2(^+) parent population cells (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single cells</td>
<td>92.4</td>
<td>ORF2(^+)</td>
<td>63.8</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>45.4</td>
<td>CD45(^+) &gt; Ly6G(^-) &gt; ORF2(^+)</td>
<td>75.3</td>
</tr>
<tr>
<td>B cells</td>
<td>4.7</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b &gt; CD11c &gt; CD24 &gt; IA/IE &gt; ORF2(^+)</td>
<td>78.2</td>
</tr>
<tr>
<td>T cells</td>
<td>9.3</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b &gt; CD11c &gt; CD24 IA/IE &gt; ORF2(^+)</td>
<td>68.4</td>
</tr>
<tr>
<td>Inflammatory monocytes</td>
<td>5.9</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b &gt; CD11c &gt; IA/IE &gt; CD6401 &gt; CD11b &gt; CD11c &gt; ORF2(^+)</td>
<td>68.8</td>
</tr>
<tr>
<td>Resident monocytes</td>
<td>15.3</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b &gt; CD11c &gt; IA/IE &gt; CD6401 &gt; CD11b &gt; CD11c &gt; ORF2(^+)</td>
<td>71.1</td>
</tr>
<tr>
<td>Eosinophils</td>
<td>0.0031</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b &gt; CD11c &gt; IA/IE &gt; CD64 CD24 &gt; IA/IE CD11b &gt; ORF2(^+)</td>
<td>33.3</td>
</tr>
<tr>
<td>Dendritic cells CD11b'</td>
<td>0.8</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b CD11c &gt; IA/IE &gt; CD64 CD24 &gt; IA/IE CD11b &gt; ORF2(^+)</td>
<td>5.5</td>
</tr>
<tr>
<td>Dendritic cells CD11b</td>
<td>0.6</td>
<td>CD45(^+) &gt; Ly6G &gt; CD11b CD11c &gt; IA/IE &gt; CD64 CD24 &gt; IA/IE CD11b &gt; ORF2(^+)</td>
<td>51.3</td>
</tr>
</tbody>
</table>

3.3.8 Laf4ir sub-cellular localisation

To further enhance our knowledge of potential functions of Laf4ir, it is important to discover the localisation of Laf4ir ORF1 and ORF2 polypeptides at sub-cellular level. To identify sub-cellular localisation of Laf4ir ORF1 and ORF2 polypeptides, cellular fractionation was conducted in laminar flow and disturbed flow-induced and static dESCs to harvest proteins from the cytosol and nuclear extract (NE). Cellular fractionation revealed the presence of Laf4ir ORF1 polypeptide within both the cytosol and nucleus of dESCs exposed to static, laminar flow and disturbed conditions. Laminar flow and disturbed flow did seem to reduce Laf4ir ORF1 polypeptide expression compared to static conditions within the nuclear extract (Figure 36). Laf4ir ORF2 polypeptide localisation was observed to be located within the cytosol and absent in the nucleus of dESCs in static, laminar flow and disturbed conditions (Figure 36). This data indicates cytosolic localisation of Laf4ir ORF2 polypeptide and both cytosolic and nuclear localisation of Laf4ir ORF1 polypeptide in dESCs induced with shear stress.
Figure 36: Sub-cellular localisation of *Laf4ir* ORF1 and ORF2 polypeptides.

The three day spontaneously differentiated ESCs were subjected to laminar flow (LF) and disturbed flow (DF) by orbital shaker for 24hr. Static (ST) conditions was used as control. **(A)** Cellular fractionation was conducted to harvest total proteins from cytosol and nuclear extract (NE), followed by western blot analysis with antibodies as indicated. Tubulin and Lamin B1 were included as cytosolic and nuclear fraction markers, respectively. **(B)** For cytosol fraction, the total level of *Laf4ir* ORF1 and ORF2 expression was quantified in relation to tubulin level. The relative ratio of each sample was against static control which was set as 1.0. **(C)** For nuclear fraction, the total level of *Laf4ir* ORF1 expression was quantified in relation to lamin B1 level. The relative ratio of each sample was against static control which was set as 1.0. Data was presented by GraphPadPrism software. Data presented were representative of three independent experiments. Mean ± SEM.
Chapter 4: Potential functions of *Laf4ir* and its underlying mechanisms

4.1 Introduction

Previous proteomics data illustrated the identification of a total of 428 proteins in chapter 3.3.1 (Figure 23). One of the proteins associated with *Laf4ir* ORF2 polypeptide was mini-chromatin maintenance protein 3 (MCM3). MCM3 is one of the highly conserved MCM family of proteins that govern the initial stages of DNA replication.

The MCM family includes seven members: MCM2-MCM9 and MCM10. This family is regarded to function as licensing components for the S-phase of cell cycle and is expressed ubiquitously in all eukaryotes. The heterohexameric MCM expression level is generally high in different stages of cell cycle and reduced in quiescence, senescence and differentiation stages. MCM 4, 6 and 7 proteins are known to mediate 3’- 5’ DNA unwinding and DNA helicase activity. The entire MCM 2-7 complex has been shown to have *in-vitro* helicase activity where the MCM 2-5 junction performs as an ATP driven gap.

The MCM 2-7 complex is known to consist as a large part of the replicative DNA helicase and that they are in inactive form at the replication sites. Helicase activity is then initiated by the recruitment of Cdc45 and GINS proteins into a holo-helicase referred as CMG (Cdc45, MCM 2–7, GINS). More recently, a study illustrated that the MCM complex also modulates the cellular response to DNA double strand breaks during DNA replication. MCM 2, 3 and 5 proteins of the complex have been documented to play other roles during replication. For example, MCM3 is known to be acetylated by chromatin-associated acetyltransferase and phosphorylated by Checkpoint Kinase 1 (Chk1) to inhibit initiation of DNA replication. However, the functional
The intramolecular environmental contribution of chemical disulphide bonds with other molecules is critical in the maintenance of redox homeostasis in living organisms. As previously stated, Laf4ir ORF2 polypeptide sequence contains an odd number of cysteine residues. This potentially enables Laf4ir ORF2 polypeptide to form a disulphide bond with other proteins such as antioxidants. Additionally, the thiol group of the spare cysteine residue within the Laf4ir ORF2 polypeptide sequence can donate a reducing equivalent (H+ + e−) to ROS to neutralise and ultimately produce less harmful molecules. This mechanism is a hallmark of various anti-oxidants such as glutathione. In the present study, we intend to investigate whether Laf4ir ORF2 polypeptide affects cycle progression through the interaction with MCM3 and whether Laf4ir ORF2 polypeptide can function as an anti-oxidant to protect cells under oxidative stress.

As previously stated, it is well known that shear stress can drive EC differentiation from ESCs (Table 1). The discovery of Laf4ir mRNA upregulation in line with EC markers from previous preliminary study in laminar flow-induced ESCs provides information of the potential ability of Laf4ir to drive EC differentiation. Therefore, in this chapter I was also interested to investigate if Laf4ir can drive EC differentiation from stem/progenitor cell types.

4.2 Aims

Hypothesis

Laf4ir enhances EC differentiation via cell cycle arrest by retaining MCM3 in the cytosol and protects differentiated ECs against oxidative stress as an antioxidant.
Aims

To assess if *Laf4ir* stimulates stem/progenitor cell differentiation towards EC lineage. To observe if *Laf4ir* plays a role in cellular proliferation and protect EC survival from oxidative stress. To investigate the potential interaction between parent gene *LAF4* and novel nested gene *Laf4ir*.

4.3 Results

4.3.1 *Laf4ir* ORF2 polypeptide enhances CD31 expression in response to VEGF and shear stress-induced stem/progenitor differentiation

The discovery of *Laf4ir* mRNA upregulation in parallel with EC markers in laminar flow-induced ESCs prompted us to explore if *Laf4ir* has the potential to differentiate stem/progenitor cells towards EC lineage. To test this, the 151aa protein sequence derived from *Laf4ir* ORF2 in TV2 was introduced in adenoviruses, abbreviated as (*Ad-L151*). In this adenovirus, a FLAG sequence was inserted downstream the ATG codon of the 151aa ORF to form a FLAG-tagged protein, which is driven by a CMV promoter. FLAG can be used as an indicator of the exogenous *Laf4ir* ORF2 polypeptide. Adenovirus with empty pShuttle2 vector was used as control throughout this PhD project and abbreviated as (*Ad-null*).

Overexpression of LAF4IR-151aa was introduced in collagen-IV coated dESCs at 10 multiplicity of infection (MOI) for 24hr. This was then followed by laminar flow shear stress (SS) for 24hr. SS significantly increased *CD31* and *CD144* expression while significantly decreased smooth muscle protein 22 (*SM22*) and calponin h1 (*CNN1*) expression (Figure 37A). These alterations in EC and SMC marker expressions were anticipated as SS is known to drive EC differentiation. Overexpression of *Laf4ir* ORF2 significantly enhanced *CD31* expression but had no effect on *CD144*, *SM22* and *CNN1*.
expression when compared to dESCs infected with Ad-null (Figure 37A). Overexpression of Laf4ir did increase SS induced expression of CD144 and CD31 (Figure 37A).

In a similar experiment, the effect of Laf4ir ORF2 overexpression on VEGF-induced VPCs was assessed. VPCs were infected with Ad-null or Ad-L151 virus at 10 MOI for 24hr and then subjected to 10ng/ml of VEGF treatment for 3 days. Similar as SS, VEGF increased CD31 expression, which was enhanced by overexpression of Laf4ir ORF2 (Figure 37B). Different to SS, VEGF had no effect in CD144 expression but slightly increased SM22 and CNN1 expression. However, overexpression of Laf4ir ORF2 slightly increased CD144 expression in the presence of VEGF and abolished VEGF-induced SM22 and CNN1 expression (Figure 37B). Furthermore, immunofluorescent staining of VPCs infected with Ad-L151 virus showed the presence of cells with double positive staining for FLAG with CD31 and fewer FLAG and smooth muscle actin (SMA) positive cells (Figure 37C). These results suggest that Laf4ir ORF2 may favour SS or VEGF-induced stem/progenitor cell differentiation towards EC lineage However the precise mechanism is unknown.
Figure 37: Overexpression of 151aa sequence upregulates CD31 expression.

(A) The 3-day spontaneously differentiated ESCs were infected with ad-null or ad-L151 virus (10 MOI) for 24hrs in serum conditions. Then the cells were subjected to laminar flow (SS) for 24hr, followed by quantitative RT-PCR analysis. Specific primers for Cd144, Cd31, Sm22 and calponin h1 (Cnn1) were utilised to assess EC and SMC expression. Gapdh was used as loading control. (B) VPCs were infected with Ad-null or Ad-L151 virus at 10 MOI for 24hrs in serum conditions. Then the cells were subjected to 10ng/ml of VEGF treatment for 3 days, followed by quantitative RT-PCR analysis. Specific primers for Cd144, Cd31, Sm22 and Cnn1 were utilised to assess EC and SMC expression. Gapdh was used as loading control. (C) Immunofluorescence of VPCs infected with Ad-L151 virus stained for mouse anti-FLAG (red), rat anti-CD31 (green) or goat anti-SMA (green) followed by donkey anti-rat alexa-Fluor 488° and donkey anti-goat alexa-Fluor 488° and donkey anti-mouse alexa-Fluor 594° secondary antibodies. DAPI was included to counterstain nuclei. Data was quantified and presented by GraphPad. Data presented were representative of three independent experiments. Mean ± SEM. *: p<0.05. Scale bar: 10 µm
4.3.2 *Laf4ir* ORF2 polypeptide retards proliferation in stem/progenitor cells

During cellular differentiation, the exit of cell cycle is required. To investigate the role of *Laf4ir* in cellular proliferation, a BrdU (Bromodeoxyuridine / 5-bromo-2'-deoxyuridine) assay was performed. During DNA replication, BrdU (an analog of the nucleoside thymidine) can be incorporated within the newly synthesised DNA and detected by anti-BrdU antibodies and therefore specifically identifies proliferating cells. In this experiment, collagen IV-coated ESCs were cultured in differentiation media for 3 days. Cells were then infected with adeno-null or adeno-L151 virus (10 MOI) for 24hrs in serum conditions, followed by 10μM BrdU incubation for 48hrs. BrdU incorporation assay demonstrated that overexpression of LAF4IR-151aa protein significantly suppressed dESC proliferation (Figure 38).

![Graph showing BrdU incorporation assay](image)

**Figure 38: Overexpression of 151aa sequence reduces proliferation**
The 3-day spontaneously differentiated ESCs were infected with adeno-null or adeno-L151 virus (10 MOI) for 24hrs in serum conditions, followed by 10μm BrdU incubation for 48hrs. Cells were then fixed and incubated with anti-BrdU-HRP antibody. Absorption was measured at 370nm. The relative BrdU was defined as the ratio of A370nm of Ad-L151 group to Ad-null group with that of Ad-null group set as 1.0. Data was quantified and presented by GraphPad. Data presented were representative of three independent experiments. Mean ± SEM. *: p<0.05.
To further validate the suppressive role of *Laf4ir* in proliferation, various different siRNAs against mRNA *Laf4ir* sequence were developed. However, after multiple attempts of *Laf4ir* knockdown studies to profoundly diminish *Laf4ir* expression level, results were continuously unsuccessful. This could be due to the unsuitability of the siRNA fragments or the strong stability of *Laf4ir* polypeptides. In the last year of my PhD project when the *Laf4ir* knockout mice were available, VPCs were isolated from the wildtype and heterozygous *Laf4ir* +/- mice line. The occurrence of homozygous *Laf4ir*/- mice was far fewer than expected. According to Mendelian Inheritance, *Laf4ir*/- mice should take one fourth of *Laf4ir* +/- mice crossbreeding. However, the occurrence of *Laf4ir*/- mice is only about one seventh (9 out of 63). The phenomenon suggests that *Laf4ir* deficiency may cause low embryonic survival rate for the homozygous. As the homozygous mice were not developed at the time of harvest, I used the heterozygous mice as the knockdown model in this PhD thesis project. Knockdown of *Laf4ir* expression in VPCs from *Laf4ir* +/- mice was shown in Figure 39A. Proliferation rate was compared between VPCs from wild type and heterozygous mice via cell number counting with FACS. As shown in Figure 39B, VPCs from *Laf4ir*+/ mice proliferated faster than those from wild type mice, suggesting that *Laf4ir* product may suppress cell proliferation and that the *Laf4ir* deficiency attenuated this suppression effect. This result supports the findings that overexpression of *Laf4ir* ORF2 polypeptide suppressed DNA replication in stem/progenitor cells.
Figure 39: Enhanced proliferation in *Laf4ir* +/- vascular progenitor cells

(A) *Laf4ir* +/- and *Laf4ir* +/- VPCs were harvested and proteins isolated for western blot analysis with antibodies as indicated. GAPDH was used as loading control. Data was presented by Image J. Data presented were representative of three independent experiments. (B) 5,000 *Laf4ir* +/- and *Laf4ir* +/- VPCs were seeded in a 12 well plate. Cells were fed every 24hr with culture media. Total cells were harvested every 24hr for 7 days and quantified by the FACS machine. Data was quantified and presented by GraphPad. Data presented were representative of three independent experiments. Mean ± SEM. *: p<0.05.
4.3.3 *Laf4ir* may affect proliferation via the interaction of *Laf4ir* ORF2 polypeptide with MCM3 in shear stress-induced treated dESCs

To understand how *Laf4ir* plays an inhibitory role in proliferation and possibly regulates cell cycle arrest, I searched the proteomics database for *Laf4ir* ORF2 polypeptide associated proteins. Among them was MCM3, one of the major DNA replication initiation factors. To validate the proteomics data, immunoprecipitation and western blot analysis was conducted in shear stress-induced dESCs. Collagen-IV coated ESCs were treated with differentiation media for 3 days and then subjected to static or shear stress conditions. Immunoprecipitation and western blot experiment showed the association of MCM3 with *Laf4ir* ORF2 polypeptide (Figure 40A). Immunocytochemistry revealed co-staining of *Laf4ir* ORF2 polypeptide and MCM3 protein and both staining’s seemed to reside within the cytosol (Figure 40B). This result suggests that *Laf4ir* ORF2 polypeptide may interact with MCM3 protein within the cytosol and may physically retain MCM3 protein from entering the nucleus, leading to cell cycle arrest. This could potentially accommodate ESCs to undergo differentiation process.
**Figure 40: Co-localisation of MCM3 protein with Laf4ir ORF2 polypeptide**

(A) The 3-day spontaneously differentiated ESCs were subjected to laminar flow (LF) and disturbed flow (DF) on an orbital shaker at 150rpm for 24hr. Static (ST) conditions was used as control. This was followed by immunoprecipitation with anti-ORF2 and anti-MCM3 antibodies and western blot analysis with antibodies indicated. WB analysis of 60μg of cell lysate inputs (input) is shown. (B) The LF induced dESCs were subjected to immunocytochemistry staining with rabbit anti- *Laf4ir* -ORF2 and goat anti-MCM3 antibodies. Secondary antibodies included donkey anti-goat Alexa-Fluor 488 and donkey anti-rabbit Alexa-Fluor 594 antibodies. DAPI was included to counterstain nuclei. Data presented were representative of three independent experiments. Scale bar: 5μm.
4.3.4 \textit{Laf4ir} ORF1 polypeptide interaction with LAF4

As both parent LAF4 and nested \textit{Laf4ir} ORF1 and ORF2 polypeptides were found to be expressed in dESCs in chapter 3.3.6 (Figure 32C), I wondered whether they could interact with each other. LAF4 protein is known to be a nuclear-specific transcription factor \cite{148}. Previous result showed the nuclear and cytosol localisation of \textit{Laf4ir} ORF1 polypeptide and cytosol localisation of \textit{Laf4ir} ORF2 polypeptide in dESCs in chapter 3.3.8 in figure 36. I wondered whether \textit{Laf4ir} ORF1 polypeptide could interact with LAF4 in dESCs. To test this, collagen-IV coated ESCs were spontaneously differentiated in differentiation medium for 3 days and then treated with shear stress for 24hr, followed by immunoprecipitation and western blot analysis with anti-LAF4 and anti-ORF1 antibodies. As shown in Figure 41A-B, I detected the association of LAF4 protein and \textit{Laf4ir} ORF1 polypeptide in static, laminar flow and disturbed flow-induced dESCs. This suggests that the two candidates can associate with each other. The association of LAF4 protein and \textit{Laf4ir} ORF1 polypeptide was further illustrated by immunocytochemistry staining which seemed to be nuclear-specific (Figure 41B) in laminar-flow induced dESCs. These results suggest that both parent \textit{LAF4} and nested \textit{Laf4ir} genes may regulate each other via polypeptide-protein interaction in stem/progenitor cell types for specific function/s.
Figure 41: *Laf4ir* ORF1 polypeptide interaction with LAF4 protein

(A) The 3-day spontaneously differentiated ESCs were subjected to laminar flow (LF) and disturbed flow (DF) on an orbital shaker at 150rpm for 24hr. Static (ST) conditions was used as control. Immunoprecipitation was performed with anti-LAF4 and anti-ORF1 antibodies and subsequent western blot analysis with antibodies indicated. WB analysis of 60μg of cell lysate inputs (input) is shown. (B) The LF induced dESCs were subjected to immunocytochemistry staining with mouse anti-ORF1 and rabbit anti-LAF4 antibodies. Secondary antibodies included donkey anti-mouse Alexa-Fluor 488 and donkey anti-rabbit Alexa-Fluor 594 antibodies. DAPI was included to counterstain nuclei. Data presented were representative of three independent experiments. Scale bar: 5µm.
4.3.5 *Laf4ir* ORF2 polypeptide protects against oxidative stress induced by hydrogen peroxide in endothelial cells.

In this study, I used the mouse EC cell line C166 (ATCC, UK). This EC line lacks endogenous *Laf4ir* ORF2 protein and therefore is more compatible for over-expression studies (Figure 42A). Also, ESCs and VPCs in culture are non-responsive to H\textsubscript{2}O\textsubscript{2} challenge due to the presence of β-mercaptoethanol, therefore ECs were ideal to investigate the role of *Laf4ir* in oxidative stress. Over-expression of LAF4IR-151aa in ECs was observed to be successful via western blot analysis as indicated in Figure 39A. ECs overexpressing LAF4IR-151aa were then challenged with 100μM H\textsubscript{2}O\textsubscript{2} for 24hr. Images of cellular death indicated a protective role of LAF4IR-151aa in ECs in response to oxidative stress facilitated by high levels of H\textsubscript{2}O\textsubscript{2} (Figure 42B). This data may suggest another potential function of *Laf4ir* ORF2 polypeptide and its role in oxidative stress in ECs.

![Figure 42: Overexpression of 151aa sequence protects endothelial cells under oxidative stress](image)

Mouse EC line C166 was infected with *Ad-null* or *Ad-L151* viruses at 10MOI for 24hr. (A) Western blot analysis shows the expression of FLAG-tagged *Laf4ir* ORF2 polypeptide with mouse anti-FLAG and rabbit anti- *Laf4ir* ORF2 antibodies. GAPDH was included as loading control. (B) The cells were then challenged with or without 100μM H\textsubscript{2}O\textsubscript{2} in serum free medium for 24hr. Cellular death was observed under microscope and images
displayed by ImageJ. Data presented were representative of three independent experiments. Scale bar: 200μm.

To clarify the cell survival role of Laf4ir ORF2 polypeptide against oxidative stress; annexin V/PI staining, DNA fragmentation and MTS assay were performed in LAF4IR-151aa overexpressed ECs. Annexin V is a classical apoptotic marker as it is able to bind to phosphatidylserine which translocates from the cytosol towards the outer leaflet of the plasma membrane during apoptosis. PI, however, does not stain live or early apoptotic cells due to the presence of an intact plasma membrane. In late apoptotic and necrotic cells, the integrity of the plasma and nuclear membranes diminishes allowing PI to migrate through the membranes, intercalate into nucleic acids, and display red fluorescence emission.

ECs were infected with Ad-null or Ad-L151 viruses at 10MOI for 24hr and then subjected with or without 50 μM or 100 μM H₂O₂ or H₂O in serum-free condition for 24hr, followed by annexin V/PI staining and FACS analysis. Results reveal a significantly fewer number of annexin V positive Ad-L151 treated ECs compared to Ad-null treated ECs when exposed to 50 μM or 100 μM H₂O₂ for 24hr (Figure 43). Due to inadequate methodology procedure by utilising trypsinisation instead of mechanical scratching, this may affect the outcome of this result. This could explain the difference in PI⁺ cells between Ad-L151 treated ECs compared to Ad-null treated ECs when exposed to 100μM H₂O₂ for 24hr. Therefore, this result may or may not indicate an anti-apoptotic effect of Laf4ir ORF2 polypeptide during oxidative stress in ECs.
Overexpression of 151aa sequence reduces apoptosis in endothelial cells under oxidative stress

(A) Mouse EC line (C166) was infected with Ad-null or Ad-L151 viruses at 10MOI for 24hr. ECs were then treated with or without 50 μM or 100 μM H₂O₂ or H₂O in serum-free condition for 24hr, followed by FACS analysis as instructed by the annexin V/Dead Cell Apoptosis Kit manual (Invitrogen, UK). Cells were stained with Alexa Fluor 488°-tagged annexin V (AV) and red-fluorescent propidium iodide (PI) nucleic acid binding dye, provided by the kit. 50,000 cells were collected at a rate of 100ul/minute and subjected to FACS analysis. Flow cytometry analysis shows the percentage cells positive and/or negative for AV (+/-) and PI (+/-). (B) The total percentage of annexin V positive cells in each condition were analysed and presented by GraphPad Prism. Data presented were representative of three independent experiments. Mean ± SEM. *: p<0.05. **: p<0.01.
During late stages of cell death, nucleic acids are degraded by a specific nuclease referred as caspase-activated DNase (CAD). Stimulation of CAD by the caspase intrinsic and extrinsic cascade pathways results in cleavage of DNA at internucleosomal linker sites, generating fragments of around 200 base pairs known as DNA ladders. To detect fragmented DNA ladders, semi-quantitative electrophoresis methodology is used. To induce DNA fragmentation by specific concentrations of H₂O₂, mouse ECs were treated with 50 and 100μM H₂O₂ for 24hrs and the DNA extracted for electrophoresis. Result showed that 50 and 100μM H₂O₂ was not sufficient to induce DNA fragmentation (Figure 44A-Left panel). The agarose gel electrophoresis from DNA extracted from Ad-null infected ECs challenged with 200μM H₂O₂, however didshow fragmentation with a ladder-like pattern (Figure 44A-Right panel). This affect was profoundly attenuated by Laf4ir ORF2 overexpression (Figure 44A-Right panel). This data indicates an inhibitory role of Laf4ir in CAD to fragment DNA during cell death, further establishing anti-apoptotic characteristics.

The MTS ((3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)) assay is widely used to explore cellular viability. The colorimetric MTS assay protocol involves the enzymatic conversion of yellow water-soluble MTS tetrazolium to a purple insoluble formazan product by the NADPH-dependent mitochondrial succinate dehydrogenase. Viable cells with active metabolism convert MTT into formazan, whilst apoptotic or dead cells lose this ability. Therefore this colour alteration measured is directly proportional to the number of viable cells. To test the effect of Laf4ir in cellular viability of ECs exposed to oxidative stress, MTS assay was performed in Ad-null or Ad-L151 infected ECs treated with or without H₂O₂ in serum-free conditions for 24hr. Results indicate a significant increase in formazan formation in Ad-L151 infected ECs compared with Ad-null infected ECs treated with 50 μM or 100 μM H₂O₂ (Figure 44B). This data indicates the beneficial impact of Laf4ir in the metabolic state of ECs during oxidative stress.
Figure 44: Overexpression of 151aa sequence improves cellular survival in endothelial cells under oxidative stress

(A) Left panel: Mouse ECs were treated with 0, 50 and 100 μM H₂O₂ in serum-free condition for 24hr, followed by DNA extraction. 20 μg extracted genomic DNA from each sample was applied to electrophoresis for DNA fragmentation analysis. Electrophoresis was run at 30V for 4 hours and gel was imaged by a UV camera. Data presented were representative of three independent experiments. Right panel: Mouse ECs were infected with Ad-null or Ad-L151 viruses at 10MOI for 24hr, and then treated with or without 200 μM H₂O₂ in serum-free condition for 24hr, followed by DNA extraction. 20 μg extracted genomic DNA from each sample was applied to electrophoresis for DNA fragmentation analysis. Electrophoresis was run at 30V for 4 hours and gel was imaged by a UV camera. Data presented were representative of three independent experiments.

(B) Mouse ECs were infected with Ad-null or Ad-L151 viruses at 10MOI for 24hr, and then treated with 0 μM, 50 μM or 100 μM H₂O₂ in serum-free condition for 24hr, followed by MTS assay. The relative formazan developed was defined as the ratio of A490nm of Ad-L151 group to Ad-null group with that of Ad-null group without H₂O₂ exposure set as 1.0. Data was quantified and presented by Microsoft excel. Data presented were representative of three independent experiments. Mean ± SEM. *: p<0.05.
4.3.6 *Laf4ir* ORF2 polypeptide reduces hydrogen peroxide in endothelial cells

As *Laf4ir* ORF2 polypeptide provides a protective role against oxidative stress mediated by H$_2$O$_2$, and *Laf4ir* ORF2 polypeptide has a high potential to be secreted, I was next enthusiastic to see if *Laf4ir* can have an effect on intracellular and extracellular H$_2$O$_2$ concentration to reduce harmful effects of oxidative stress. To test this, ECs infected with *Ad-null* or *Ad-L151* were exposed to 500, 100, 50 or 25μM H$_2$O$_2$ challenge. Cells treated without H$_2$O$_2$ were included as control. After 15 mins, H$_2$O$_2$ concentration was measured from the media and cell lysate. In untreated cells, the endogenous H$_2$O$_2$ level was slightly reduced by the overexpression of LAF4IR-151aa. While under 25, 50, 100 and 500μM H$_2$O$_2$ challenge the intracellular H$_2$O$_2$ was greatly increased in *Ad-null* infected ECs, overexpression of LAF4IR-151aa significantly attenuated this increase when exposed to 50 and 100μM H$_2$O$_2$ challenge (Figure 45A). Results indicate a significant decline in extracellular H$_2$O$_2$ concentration from LAF4IR-151aa overexpressed ECs compared to *Ad-null* treated ECs when exposed to 25μM H$_2$O$_2$ challenge (Figure 45B). This result indicates the ability of *Laf4ir* ORF2 polypeptide to reduce intracellular and extracellular harmful concentration of H$_2$O$_2$. 


Figure 45: Overexpression of 151aa sequence in endothelial cells reduces intracellular and extracellular H$_2$O$_2$.

Mouse ECs were infected with Ad-null or Ad-L151 viruses at 10MOI for 24hr and then exposed to different concentrations of H$_2$O$_2$ (500, 100, 50, 25 or 0 μM) and incubated at 37°C for 15 mins. (A) Cell lysate was harvested and intracellular H$_2$O$_2$ concentration was quantified via the colorimetric hydrogen peroxide assay kit. Signal generated by the optical density (OD) was measured at 570 nm. The relative ratio of OD of each sample was against sample control infected with Ad-null and treated with 0 μM H$_2$O$_2$ which was set as 1.0. (B) Medium was harvested and extracellular H$_2$O$_2$ concentration was quantified via the colorimetric hydrogen peroxide assay kit. Input controls include known concentrations of H$_2$O$_2$ (500, 100, 50 or 25 μM). Negative control consisted of 1xPBS. Signal generated by the optical density (OD) was measured at 570 nm. The relative ratio of OD of each sample was against input control which was set as 1.0. Data was quantified and presented by GraphPad. Data presented were representative of three independent experiments. Mean ± SEM. *: p<0.05.
Chapter 5: *In-vivo* model of *Laf4ir* in cardiovascular remodelling

5.1 Introduction

Vascular endothelium dysfunction is a common pathophysiological hallmark of cardiovascular diseases such as heart failure and atherosclerosis. Traditional therapeutic approaches aim to alleviate further injury to protect the vasculature. Alternative strategies include stimulating vascular self-repair mechanisms, however the precise processes involved remain unclear.

The vasa vasorum constitutes a complete vascular tree-like structure, including arteries, arterioles, capillaries, venules and veins \(^{235}\). Under pathological conditions, the arterial adventitia undergoes remodelling in response to a variety of arterial injuries. Data from chapter 4 illustrates the functional aspect of *Laf4ir* in EC differentiation, proliferation and oxidative stress. Therefore, *Laf4ir* in stem/progenitor cells may act in a coordinating fashion during healing response to vascular injury. Conversely, *Laf4ir* may also contribute to vascular disease development.

*LAF4* mRNA was expressed ubiquitously in most human organs or tissues as revealed by the RNA-seq from 53 human tissue samples from the Genotype-Tissue Expression (GTEx) Project (https://www.ebi.ac.uk) including the cardiovascular system. To assess whether *Laf4ir* was also expressed and contributed to cardiovascular functions, several *in-vivo* cardiovascular remodelling models were introduced, followed by the detection of *Laf4ir* ORF1 and ORF2 polypeptides. In this chapter, I described the role of *Laf4ir* in mouse ischaemic and vascular injury models, atherosclerosis and heart failure.
As described in chapter 2.7 methodology, I recently developed a global knockout model of *Laf4ir*. This novel *Laf4ir* $^{-/-}$ mouse line was developed by the conditional knockout Cre-loxp system. Cre/loxp system consists of two vital components which are Cre recombinase and loxP sites. Cre recombinase is encoded by a *cre* gene derived from bacteriophage P1. This enzyme mediates recombination of target sequence which locates between loxP sites and is widely used in transgenic mice. In transgenic mice models, activity of Cre recombinase is under control of specific promoter. loxP sites are a pair of the sequences which flank one inserted gene.

Once loxP sites are recognized by Cre, target gene deletion occurs. In this model the Genoway biotechnology company inserted loxP sites into the promoter region and exon 3 respectively of the *Laf4ir* gene. Once genotype results validated loxP insertions in both copies of the *Laf4ir* gene, *Laf4ir* $^{loxP/loxP}$ mice were bred with CMV-Cre transgenic mice for gene deletion. The specific location of loxP sites will lead to the deletion of the flanked gene sequence by Cre recombinase. FLP-FRT recombination system, similar to Cre/loxp, consists of FLP recombinase and FRT sites. FLP protein exerts deletion once recognising FRT sites. The FLP-FRT system was used to insert a neomycin cassette for positive ESC clone selection, which was performed by GenOway. In this chapter, I examined phenotypic alterations in the global knockout of *Laf4ir* by the Cre/loxp system. Then I used the *Laf4ir* knockout mice in an ischaemia model and also isolated the aortic adventitia VPCs for *in-vitro* functional studies (described in chapter 4.3.2). I was unable to further analyse phenotypic difference in the femoral injury mouse model due to low mice stock and time restraints.

### 5.2 Aims

**Hypothesis**

*Laf4ir* is involved in cardiovascular remodelling via spatiotemporal translation of different ORFs.
Aims

To investigate the role of Laf4ir in heart failure, atherosclerosis and vessel injury repair, and to delineate further potential functions from phenotypic examinations of the global knockout in-vivo model.

5.3 Results

5.3.1 Upregulation of Laf4ir ORF2 and downregulation of Laf4ir ORF1 and LAF4 in a TAC model

To investigate if Laf4ir is associated with heart cardiac hypertrophy, a transverse aortic constriction (TAC) model was introduced in the hearts of mice. Immunofluorescent staining in sham condition at day 7 post surgery revealed some staining of Laf4ir ORF1 polypeptide and ORF2 polypeptide within the endothelium of small vessels in the left ventricle (Figure 46A). The TAC-induced heart samples at day 7-post surgery revealed a reduced level of expression of Laf4ir ORF1 polypeptide and increased level of expression of Laf4ir ORF2 polypeptide within the endothelium compared to sham controls (Figure 46A). At day 7 post-surgery, TAC-induced left ventricle showed Laf4ir ORF2 polypeptide expression located on the outer layer of endothelium in the lumen side, forming a shell on ECs (Figure 46A). There were also ORF2+/CD31− cells or particles in interstitial tissues after TAC (Figure 46A). These results suggest that Laf4ir ORF2 polypeptide may be secreted to extracellular environment in response to TAC. Western blot analysis of heart samples showed this trend of expression (Figure 46B-C).

Western blot analysis illustrated a significant reduction of Laf4ir ORF1 polypeptide expression in TAC samples at days 3 and 7 post-surgeries compared to day 7 post-surgery sham control (Figure 46B-C). Interestingly, LAF4 protein had a similar significant reduction as Laf4ir ORF1 polypeptide (Figure 46B-C). In the contrary, a
significant enhancement of Laf4ir ORF2 polypeptide expression in TAC samples at days 3 and 7 post-surgeries compared to day 7 post-surgery sham control was observed (Figure 46B-C). These results suggest that there is a translation shift from Laf4ir ORF1 to ORF2 after TAC and that the spatiotemporal translation of different ORFs may contribute to cardiac remodelling.

Figure 46: Laf4ir ORF1 and ORF2 expression in a trans-aortic constriction model.
Aortic constriction was introduced into 12-week old C57bl/6 mice and heart tissues were harvested at day 3 (TAC-D3) and 7 (TAC-D7) post-surgery. Sham control was included with similar surgery except for aortic constriction at day 7 (Sham-D7) post-surgery. (A) The left ventricle of heart samples were cryosectioned (12μm) in a horizontal manner and subjected to immunofluorescence staining. Staining included
incubation with rat anti-CD31, *Laf4ir* mouse anti-ORF1 and *Laf4ir* rabbit anti-ORF2 primary antibodies, followed by secondary antibodies consisting of donkey anti-rat 594, donkey anti-mouse 488 and donkey anti-rabbit 488, respectively. The nuclei were counterstained with DAPI (blue) and images obtained under confocal microscope. Blocking peptide (BP) was included in LAF4IR staining to show the specificity of *Laf4ir* ORF1 and ORF2 antibodies. Scale bars 25 and 50μm. (B) Total protein content of each heart was isolated and subjected to western blot analysis using LAF4, *Laf4ir* ORF1 and ORF2 antibodies. GAPDH was used as loading control. (C) Bar chart illustration of LAF4, *Laf4ir* ORF1 and ORF2 expression relative to GAPDH expression, quantified by ImageJ and presented by GraphpadPrism. The data presented were representative images of 3 mice. Mean ± SEM. **: p<0.01, ***: p<0.001.

5.3.2 Altered expression of *Laf4ir* ORF2 in *ApoE^-/-* aorta

As previous data has shown the potential role of *Laf4ir* in oxidative stress in ECs and high expression of *Laf4ir* ORF2 polypeptide in the aorta of WT mice, I wondered whether *Laf4ir* participated in atherosclerotic processes. To observe this, the expression level of *Laf4ir* ORF1 and ORF2 polypeptides in the aorta was compared between wild type and *ApoE^-/-* mice.

In WT ascending aorta, immunohistochemistry revealed non-existent expression of *Laf4ir* ORF1 polypeptide (Figure 48A) whilst *Laf4ir* ORF2 polypeptide expression was located in the intima (Figure 48A). As expected, Sca-1+ cells were detected within the adventitia. To further validate if *Laf4ir* ORF2 polypeptide was exclusively expressed in the intima, co-staining of *Laf4ir* ORF2 polypeptide with haematopoietic cell marker CD34, endothelial cell marker CD144 and smooth muscle marker SMA proteins was performed in the ascending aorta. *Laf4ir* ORF2 polypeptide was exclusively expressed in the intima, which is either CD34 or CD144 positive, and no *Laf4ir* ORF2 polypeptide expression was detected in the SMA-positive media (Figure 47). This result indicates that the *Laf4ir* ORF2 polypeptide in WT ascending aorta in mice is endothelial specific.

In *ApoE^-/-* mouse tissue, *Laf4ir* ORF2 positive cells were detectable throughout the aorta from the intima to media and the adventitia, some cells were double positive
for \textit{Laf4ir} ORF2 and Sca-1 (Figure 48B, left panel). A significant high level expression of \textit{Laf4ir} ORF2 polypeptide in the \textit{ApoE}⁻/⁻ mouse aorta was further shown when compared to heart expression by western blot analysis (Figure 48C-D). Other organs (brain and lung) showed no difference in \textit{Laf4ir} ORF2 expression in comparison to heart expression in the \textit{ApoE}⁻/⁻ mouse (Figure 48C-D). A significant high level expression of \textit{Laf4ir} ORF2 in the \textit{ApoE}⁻/⁻ mouse trachea was observed when compared to heart expression by western blot analysis (Figure 48C-D). The high specificity of \textit{Laf4ir} ORF2 antibody in western blot experiments was shown (Figure 48C). \textit{Laf4ir} ORF1 polypeptide expression was only detected in the neointima area in \textit{ApoE}⁻/⁻ but not in WT mice (Figure 48B, far right). The \textit{Laf4ir} ORF1 positive cells showed multiple nuclei with lobed appearance (Figure 48B, bottom right panel).

To observe the level of \textit{Laf4ir} ORF2 polypeptide expression in activated Sca-1 positive cells, the number of \textit{Laf4ir} ORF2 positive cells in Sca-1⁺ VPCs in culture from both WT and \textit{ApoE}⁻/⁻ were compared via FACS analysis. FACS data shows that almost all WT and \textit{ApoE}⁻/⁻ VPCs (>95%) were positive for \textit{Laf4ir} ORF2 polypeptide (Figure 49), illustrating the expression of \textit{Laf4ir} ORF2 polypeptide in Sca-1 positive cells observed \textit{in-vivo}. No differences in expression of \textit{Laf4ir} ORF2 positive cells were detected between WT and \textit{ApoE}⁻/⁻ VPCs as both conditions involve activated forms of Sca-1 cells \textit{in-vitro} (Figure 49). The high specificity of \textit{Laf4ir} ORF2 antibody in FACS experiments was demonstrated (Figure 49). These results suggest that the \textit{Laf4ir} ORF2 polypeptide may have an important role in physiological endothelial functions and that \textit{Laf4ir} ORF2 polypeptide could potentially be associated with atherosclerosis in Sca-1 positive cells.
Figure 47: *Laf4ir* ORF2 polypeptide expression is endothelial cell specific in the aorta

14-weeks old wild type mouse ascending aorta was cryo-sectioned at a longitudinal manner (12μm) and subjected to immunofluorescent staining with rabbit anti-ORF2, rat anti-CD34 or goat anti-CD144 or mouse anti-SMA antibodies. The Alexa-Fluor 488 (green) and 594 (red) secondary was used in correlation to each primary antibody as indicated by the colour of the font. The nuclei were counterstained with DAPI (DAPI) and images obtained under confocal microscope. Data presented were representative of three independent experiments. Scale bar: 75μm and 25μm.
Figure 48: Laf4ir ORF1 polypeptide and ORF2 polypeptide expression in wild type and ApoE⁻/⁻ aorta

(A) Wild type and (B) ApoE⁻/⁻ 14 weeks-old mouse ascending aorta was cryo-sectioned (12μm thickness) and subjected to immunofluorescent staining with rabbit anti-ORF2, mouse anti-ORF1 and rat anti-Sca-1 antibodies. The Alexa-Fluor 488° (green) and 594° (red) secondary was used in correlation to each primary antibody used as indicated by the colour of the font. The nuclei were counterstained with DAPI and images obtained under confocal microscope. The data presented were representative images of 6 mice. Scale bars include 250μm, 75μm, 50μm and 25μm. (C) ApoE⁻/⁻ mouse organs were isolated, proteins harvested and subjected to western blot analysis with rabbit anti-ORF2 antibody and anti-ORF2 antibody incubated with blocking peptide (BP). GAPDH was used as loading control. (D) Bar chart illustration of Laf4ir ORF2 polypeptide expression relative to GAPDH expression in ApoE⁻/⁻ mouse organs, quantified by ImageJ and presented by GraphpadPrism. The total level of Laf4ir ORF2 expression in each organ was quantified in relation to heart Laf4ir ORF2/GAPDH ratio expression level set as 1.0. The data presented were representative images of 3 mice. Mean ± SEM. **: p<0.01, ***: p<0.001.
Figure 49: Laf4ir ORF2 polypeptide expression in wild type and ApoE−/− vascular progenitor cells.

(A) VPCs from both WT and ApoE−/− were cultured in maintenance media and harvested for FACS analysis. Cells were stained with rabbit Laf4ir antibody and donkey anti-rabbit Alexa Fluor 647 secondary antibody. Blocking peptide (BP) was incubated with Laf4ir ORF2 antibody for 24hrs, followed by staining to show the specificity of Laf4ir ORF2 antibody. 20,000 cells were collected at a rate of 100μl/minute. Data was quantified by FlowJo software.

(B) Bar chart illustrates the total percentage of VPCs positive for Laf4ir ORF2 in both WT and ApoE−/− backgrounds. Data was presented by GraphPadPrism software. Data presented were representative of three independent experiments. Mean ± SEM.
5.3.3 Upregulation of *Laf4ir* ORF2 polypeptide in a hindlimb ischemia model

The discovery of the potential role of *Laf4ir* in vascular pathology of atherosclerosis, lead to the investigation of the contribution of *Laf4ir* in other vascular pathologies induced by ischaemic conditions. To explore this phenomenon, a hindlimb ischaemia model was conducted in mice.

At day 14 post surgery, surrounding injured (ischaemic) and sham (non-ischaemic) skeletal tissues were harvested and processed for western blot analysis (Figure 47A-B) and immunohistochemistry (Figure 47C). In both sham and injured muscle samples, *Laf4ir* ORF1 polypeptide expression was not present (Figure 47A). However, there was a significant increase in *Laf4ir* ORF2 polypeptide expression in ischaemic muscles compared to non-ischaemic muscles (Figure 47A-B). This was further demonstrated by immunofluorescent *Laf4ir* ORF2 polypeptide staining as a higher level of *Laf4ir* ORF2 positive cells were detected in injured skeletal tissues compared to sham control (Figure 47C). Some CD31^{+}\text{ORF2}^{+}, CD31^{+}\text{ORF2}^{-} and CD31^{+}\text{ORF2}^{+} cells were detected in sham control skeletal tissues and a greater extent of CD31^{+}\text{ORF2}^{+} cells were observed in injured ischaemic skeletal tissues (Figure 47C). A significant increase in *Laf4ir* ORF2 positive cells were detected in injured skeletal tissues compared to sham control (Figure 47D). These results indicate the responsive feature of *Laf4ir* towards ischemia via the translation of *Laf4ir* ORF2 polypeptide, potentially orchestrating angiogenesis mechanisms.
Figure 50: Laf4ir ORF1 polypeptide and ORF2 polypeptide expression in a hindlimb ischaemia model.

C57Bl/6 12 week old mice organs were subjected to hindlimb ischaemia surgery. At day 14 post surgery, non-ischaemic (sham) and ischaemic (injured) muscles were isolated. (A) Total protein content of each tissue was harvested and subjected to western blot analysis using Laf4ir ORF2 and ORF1 antibodies. GAPDH was used as loading control. (B) The total level of Laf4ir ORF2 expression in sham and injured muscles was quantified in relation to GAPDH expression level. (C) Tissues were cryosectioned (12μm thickness) and subjected to immunofluorescent staining. Primary antibodies included rabbit anti-ORF2 (α-ORF2) and rat anti-CD31 primary antibodies, followed by donkey anti-rabbit 488° and donkey anti-rat 594°, respectively. The nuclei were counterstained with DAPI and images obtained under confocal microscope. Scale bars, 25μm. (D) Bar chart illustration of Laf4ir ORF2+ positive cell quantification in injured tissue in relation to sham tissue control set as 1.0. This was quantified by ImageJ and presented by GraphpadPrism. The data presented were representative of 6 mice. Mean ± SEM. *: p<0.05.
5.3.4 Impaired vascular recovery in ischaemic-induced *Laf4ir*\(^{+/−}\) mice

The data in chapter 5.3.3 indicated the upregulation of *Laf4ir* ORF2 polypeptide in response to hindlimb ischemia. Therefore, I wanted to investigate if *Laf4ir* contributes to vascular repair mechanisms in response to ischaemic insults. To test this, a hindlimb ischaemia model was conducted in *Laf4ir*\(^{+/+}\) wild type and *Laf4ir*\(^{+/−}\) heterozygous knockout mice.

Surgical injury was introduced on the right leg and sham control was introduced on the left leg. Doppler scan for blood perfection of the lower part of the mouse body was obtained at days 0 (30 mins), 7 and 14 post-surgery for both *Laf4ir*\(^{+/+}\) and *Laf4ir*\(^{+/−}\) mice. At days 1 and 2 post-surgery, two *Laf4ir*\(^{+/−}\) mice were found to be unfit to recover and survive so they were sacrificed and not used in this investigation. This may suggest inadequate surgical procedures or *Laf4ir* may have an important role in vascular repair mechanisms.

Results showed that blood perfusion of the paw was reduced in *Laf4ir*\(^{+/−}\) mice as compared to *Laf4ir*\(^{+/+}\) mice at day 7 post-surgery (Figure 51). At day 14 post-surgery, a significant reduction in blood perfusion was observed in *Laf4ir*\(^{+/−}\) mice compared to *Laf4ir*\(^{+/+}\) mice (Figure 51). This suggests that *Laf4ir*, potentially mediated by the *Laf4ir* ORF2 polypeptide, is involved in angiogenesis or other repair mechanism processes in response to ischaemia.

To further investigate this process, surrounding muscles from both sham and injured legs of all mice were isolated after day 14 post-surgery doppler scan. The injured and uninjured muscles were then processed for immunofluorescent staining for endothelial cell marker (CD31) and vascular progenitor cell marker (Sca-1). Immunohistochemistry staining showed a fewer number of CD31\(^{+}\), Sca-1\(^{+}\) and double
positive CD31⁺Sca-1⁺ cells in injured Laf4ir⁻/⁻ tissue when compared to injured Laf4ir⁺/+ tissue (Figure 52). To measure this trend, the ratio of each CD31⁺ and Sca-1⁺ cells between injured ischaemic and uninjured non-ischaemic muscle tissues was quantified (Figure 52). Results show a reduction, but not significant difference, in the ratio of injured to uninjured tissue positive for CD31, Sca-1 and double positive for both markers from Laf4ir⁻/⁻ mouse line compared to Laf4ir⁺/+ mouse line. These results may indicate the potential role of Laf4ir, potentially by the Laf4ir ORF2 polypeptide, in vascular ischaemic injury repair.

Figure 51: Laf4ir knockout reduced foot blood perfusion in a mouse hindlimb ischaemia model.

The hindlimb ischaemia model was introduced in 2 months old Laf4ir⁺/+ and Laf4ir⁻/⁻ mice. (A) Left panel shows the representative images by a Doppler scanner of Laf4ir⁺/+ and Laf4ir⁻/⁻ groups at days 0, 7 and day 14 post-surgery. (B) Foot blood flow ratio was measured between left sham and right injured side by the doppler scanner software at days 0, 7 and day 14 post-surgery. Data was quantified and presented by GraphpadPrism. Two Laf4ir⁻/⁻ mice died at days 1 and 2 post surgery. The data presented were representative of 6 mice for Laf4ir⁺/+ and 4 mice Laf4ir⁻/⁻. Mean ± SEM. *: p<0.05.
**Figure S2: L4f4r knockout reduced Sca-1+ and CD31+ cells in a mouse hindlimb ischaemia model.**

Ischaemic (injured right leg) and non-ischaemic (uninjured left leg) muscle tissues were harvested at day 14 post-surgery and embedded in frozen OCT for cryosection (12μm) and subjected to immunofluorescent staining. **(A)** Injured ischaemic tissue from both L4f4r+/+ and L4f4r+/- mice was treated with rat anti-Sca1 and rabbit anti-CD31 primary antibodies, followed by donkey anti-rat 488 and donkey anti-rabbit 594, respectively. The nuclei were counterstained with DAPI and images obtained under confocal microscope. Scale bars, 100μm. **(B)** The total number of Sca1+ and CD31+ cells were counted by ImageJ and the ratio between ischaemic injured and non-ischaemic uninjured tissues in each L4f4r+/+ and L4f4r+/- mouse was quantified and presented by GraphpadPrism. The data presented were representative images of 6 mice for L4f4r+/+ and 4 mice L4f4r+/-; Mean ± SEM. NS: Non-significant.
5.3.5 Upregulation of *Laf4ir* ORF2 polypeptide in a femoral artery injury model

Since I discovered the role of *Laf4ir* in vascular ischaemic damage, I was curious to find out if *Laf4ir* is associated with different injury insults: in this case by mechanical means. A widely used method to access mechanical-induced endothelial injury in mice is the platinum wire induced femoral artery injury model. This model involves the physical scratching of the femoral artery endothelium. This enables researchers to investigate reendothelialisation and neointima formation linked with vascular repair and pathologies such as atherosclerosis.

In uninjured sham control day 3 post-surgery wild type mouse, *Laf4ir* ORF1 polypeptide expression in the femoral artery was found to be absent (Figure 53, left panel). Upon femoral injury at day 3 post-surgery, *Laf4ir* ORF1 polypeptide expression was also found to be absent (Figure 53, right panel). As expected, Sca-1 positive cells were found to be localised mainly in the adventitia and intima of the uninjured and injured femoral arteries (Figure 53). This result illustrates the non-existent role of *Laf4ir* ORF1 polypeptide in response to mechanical-induced endothelial injury.

In uninjured sham control day 3 post-surgery wild type mouse, *Laf4ir* ORF2 polypeptide expression was also absent in the femoral artery (Figure 54, left panel). The differences in *Laf4ir* ORF2 polypeptide expression between aorta and femoral artery may reflect the EC heterogenicity along the arterial tree. However, when the femoral artery suffered injury via platinum wire scratching, plenty of *Laf4ir* ORF2⁺ cells were detected in the adventitia region and lumen at day 3 post surgery (Figure 54, middle panel) but mainly in the adventitia at day 7 post femoral artery injury (Figure 54, right panel). In day 3 injured samples, some ECs had a shell of *Laf4ir* ORF2 polypeptides additionally (insert in Figure 54, middle panel). As expected, CD31 positive cells were
found to be localised in the intima of the uninjured and injured femoral arteries (Figure 54). These results suggest the potential involvement of \textit{Laf4ir} ORF2 polypeptide in reendothelialisation in response to EC mechanical damage. Table 15 summarises the alterations in \textit{Laf4ir} ORF1 and ORF2 expression in all experimental models utilised in my PhD project.

![Uninjured- D3 Sca-1/ORF1](image1)
![Injured- D3 Sca-1/ORF1](image2)

**Figure 53:** \textit{Laf4ir} ORF1 polypeptide is not expressed in a femoral artery injury model.

Platinum wire injury was introduced in 12-weeks old wild type mice. Injured and uninjured femoral arteries were harvested at days 3 (D3) post-surgery and embedded in liver tissue. The uninjured sham control underwent similar surgery without the inclusion of wire scratching. Cryo-sectioned tissues (12 μm thickness) were stained with mouse anti-ORF1 and rat anti-Sca-1 primary antibodies, followed by donkey anti-mouse alexa-Fluor 488 (green) and donkey anti-rat alexa-Fluor 594 (red) secondary antibodies. The nuclei were counterstained with DAPI and images obtained under confocal microscope. The data presented were representative images of 6 mice. Scale bars: 75μm.
Figure 54: *Laf4ir* ORF2 polypeptide upregulation in a femoral artery injury model.

Platinum wire injury was introduced in 12-weeks old wild type mice. Injured and uninjured femoral arteries were harvested at days 3 (D3) and 7 (D7) post-surgery and embedded in liver tissue. The uninjured sham control underwent similar surgery without the inclusion of wire scratching. Cryo-sectioned tissues (12 μM thickness) were stained with rabbit anti-ORF2 and rat anti-CD31 primary antibodies, followed by donkey anti-rabbit alexa-Fluor 488° (green) and donkey anti-rat alexa-Fluor 594° (red) secondary antibodies. The nuclei were counterstained with DAPI and images obtained under confocal microscope. The data presented were representative images of 6 mice. Scale bars: 100μm.

Table 15: Changes in *Laf4ir* ORF1 and ORF2 polypeptide expression in different experimental models

<table>
<thead>
<tr>
<th>Experimental models</th>
<th>LAF4IR ORF1</th>
<th>LAF4IR ORF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trans-aortic constriction</td>
<td>Decrease in expression</td>
<td>Increase in expression</td>
</tr>
<tr>
<td><em>ApoE</em> knockout</td>
<td>No expression</td>
<td>Change in distribution expression</td>
</tr>
<tr>
<td>Femoral artery injury</td>
<td>No expression</td>
<td>Increase in expression</td>
</tr>
<tr>
<td>Hindlimb ischaemia</td>
<td>No expression</td>
<td>Increase in expression</td>
</tr>
</tbody>
</table>
5.3.6 Genotype of global \textit{Laf4ir} knockout mice

The setup of various crossbreeding pairs of mice and the anticipating period of developing our first heterozygote and homozygote \textit{Laf4ir} knockout mice was time consuming and exciting. The first examination was to investigate the effective global knockout of \textit{Laf4ir} in our transgenic mice. To explore the genotype of newly born \textit{Laf4ir} mouse line, PCR reactions were utilised using specific primers for \textit{Laf4ir} deletion as mentioned in the methodology chapter 2.7 (Figure 55A). PCR and electrophoresis result validated the genotype profile of \textit{Laf4ir} \textsuperscript{+/-}, \textit{Laf4ir} \textsuperscript{+-}, \textit{Laf4ir} \textsuperscript{-/-} mice (Figure 55A).

It was also crucial to investigate if the deletion of the nested gene \textit{Laf4ir} would cause potential disruption of the expression of the parent gene, \textit{LAF4}. Therefore, to see if \textit{Laf4ir} deletion affected transcription and therefore translation of \textit{LAF4}, a western blot analysis was conducted using antibodies against \textit{LAF4} protein, \textit{Laf4ir} ORF1 polypeptide and ORF2 polypeptide in liver tissue samples. I used total proteins derived from the liver as previous data illustrates the expression of \textit{LAF4} protein, \textit{Laf4ir} ORF1 polypeptide and ORF2 polypeptide in the liver. Result showed that the deletion of \textit{Laf4ir} did not disrupt and affect \textit{LAF4} expression (Figure 55B). Results also showed the significant reduction in \textit{Laf4ir} ORF1 and ORF2 polypeptide expression in the liver of \textit{Laf4ir} \textsuperscript{+/-} mice compared to \textit{Laf4ir} \textsuperscript{+/-} mice, and the non-existent ORF1 and ORF2 polypeptide expression in the liver of \textit{Laf4ir} \textsuperscript{-/-} mice. This result uniquely indicates different transcriptional regulations of the parent gene \textit{LAF4} and nested gene \textit{Laf4ir}.
Figure 55: Genotype of Laf4ir knockout mouse line and expression of Laf4ir ORF1 polypeptide, ORF2 polypeptide and LAF4 protein.

(A) DNA from ear clips of 10-week old Laf4ir+/+ (+/+), Laf4ir+/- (+/-) and Laf4ir−/− (−/−) mice were extracted and subjected to PCR reaction and electrophoresis. Genotype result of all three mouse lines animals. Laf4ir wild type base pair band is 200bp and homozygous knock-out band is 327bp. (B) Liver organs from the 10-week old Laf4ir+/+, Laf4ir+/- and Laf4ir−/− mice were isolated and the protein content harvested. This was followed by western blot analysis using antibodies against LAF4, Laf4ir ORF1 and ORF2 proteins. GAPDH was included as loading control. (C) Bar chart illustration of LAF4, Laf4ir ORF1 and ORF2 expression in relation to GAPDH control from Laf4ir+/+, Laf4ir+/- and Laf4ir−/− liver organ samples. This was quantified by ImageJ and presented by GraphpadPrism. Data presented were representative of three independent experiments. Mean ± SEM. **: p<0.01.
5.3.7 Phenotype of global *Laf4ir* knockout mice

To further better understand potential functions of *Laf4ir*, phenotypic examination of this novel knockout model is vital. To date, several breeding pairs have been utilised which enabled me to have enough mice to determine phenotypic differences between wild type and heterozygote and homozygote mice. All phenotypic observations in affected areas of *Laf4ir* knockout mice have been documented in table 14. After thorough examinations, I noticed substantial differences in body sizes among *Laf4ir* *+/+,* *Laf4ir* *+/−* and *Laf4ir* *−/−* mice (Figure 56A). The total body length of 10 week old *Laf4ir* *+/+,* *Laf4ir* *+/−* and *Laf4ir* *−/−* mice were 175 mm, 170 mm and 156 mm, respectively. Then, the total body weight of 10 week old female *Laf4ir* mouse line was measured. Results showed a significant lower body weight in homozygote *Laf4ir* *−/−* mice when compared to *Laf4ir* *+/−* and *Laf4ir* *+/+* mice (Figure 56B). Homozygote *Laf4ir* *−/−* mice were also seen to have an abnormal arch posture and enlargement of incisor morphology known as malocclusion (Table 16). Overall, physical appearance examinations indicate that the *Laf4ir* gene may have a role in adult development.

Internal organ examination revealed no noticeable differences between each *Laf4ir* genotype, however an additional novel vessel around the femoral artery was observed during the hindlimb ischaemia model surgery in heterozygote *Laf4ir* *+/−* mice (Table 16). The hindlimb vessel growth was not studied in homozygous knockout mice, therefore it is not known if this phenotype is present in *Laf4ir* *−/−* mice (Table 16). During internal examination, a striking difference in heart size was observed (Figure 57A). The hearts of each animal were isolated, measured and weighed to quantify overall heart to body weight ratio. Results showed that the total length of 10 week old *Laf4ir* *+/+,* *Laf4ir* *+/−* and *Laf4ir* *−/−* mice hearts were 6.8 mm, 6.9 mm and 7 mm, respectively. Also, a significant higher heart to body weight ratio was found in homozygote *Laf4ir* *−/−* mice when compared to *Laf4ir* *+/−* and *Laf4ir* *+/+* mice (Figure 57B). Further examination is required to see if hypertrophy, hyperplasia or other factors contribute to the abnormal
growth of the heart. Overall, this observation suggests that the *Laf4ir* gene may be important in cardiac developmental processes and/or cardiac functions.

**Figure 56: Body size phenotypic examination of *Laf4ir* knockout mouse line.**

A) Photograph of 10 week old female *Laf4ir* +/+ , *Laf4ir* +/- and *Laf4ir* -/- mice. The total body length of mice was measured adjacent to a 30cm ruler. (B) Body weight of 10 week old female and male *Laf4ir* mouse line was measured from three crossbreeding pairs. Data presented were representative of *Laf4ir* +/+ (n=5), *Laf4ir* +/- (n=11) and *Laf4ir* -/- (n=3) mice. Mean ± SEM. *: p<0.05, **: p<0.01.
Figure 57: Heart size phenotypic examination of *Laf4ir* knockout mouse line

(A) Photographic images of 10 week old female *Laf4ir*/*, *Laf4ir*+/− and *Laf4ir*−/− mice of hearts were taken prior to the total heart length measurement adjacent to a 30cm ruler.  
(B) The body weight and heart weight of each female *Laf4ir* mouse line was measured and the heart to body weight ratio was quantified and presented using GraphPad Prism. The data presented were representative images of 3 mice. Mean ± SEM. *: p<0.05. **: p<0.01.
<table>
<thead>
<tr>
<th>Affected areas</th>
<th>( \text{Laf4ir}^{-/-} )</th>
<th>( \text{Laf4ir}^{+/-} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Malocclusion</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reduced body size</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal arch posture</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Enlarged heart size</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Abnormal hindlimb vessel growth</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 16: Phenotypic observation from \( \text{Laf4ir} \) knockout *in-vivo* model
Table illustrates phenotypic observation documented from the male and female adult \( \text{Laf4ir}^{-/-} \) and \( \text{Laf4ir}^{+/-} \) mice during visual external and internal examination. N/A (Not applicable).
Chapter 6: Discussions

6.1 Co-ordination between nested \textit{Laf4ir} and parent \textit{Laf4} genes

In this PhD thesis, a novel nested intronic gene (\textit{Laf4ir}) which is located in the intron 6 of the mouse transcription factor gene \textit{Laf4} was previously discovered. \textit{Laf4ir} encodes two different polypeptides, 45aa from \textit{Laf4ir} ORF1 and 109aa or 151aa from \textit{Laf4ir} ORF2 in \textit{Laf4ir-tv1} and/or \textit{Laf4ir-tv2}. The \textit{Laf4ir} ORF1 polypeptide could associate with LAF4 protein, suggesting that the nested gene and its external gene may coordinate and regulate each other. Most importantly, \textit{Laf4ir} ORF2 polypeptide expression is related to vascular remodelling.

\textit{Laf4}/LAF4 (AFF3/LAF4) is one of the four members of the \textit{AFF (AF4/FMR2)} family of genes, encoding a lymphoid transcription factor. The mRNA of \textit{LAF4} was detected in all tissues and organs including cardiovascular system. My study detected \textit{Laf4ir} ORF2 polypeptide in some mouse tissues and organs. These observations indicate that both \textit{Laf4} and \textit{Laf4ir} may be expressed in the same tissue and potentially in the same cell. Indeed, I detected the co-expression and association of LAF4 protein and \textit{Laf4ir} ORF1 polypeptide within a single ESC. \textit{Laf4} and \textit{Laf4ir} use the opposite strands for transcription, which raises a possibility that the intron 6 of \textit{Laf4} gene may not be transcribed. If the intron 6 is transcribed, the precursor RNA for \textit{Laf4} will be complementary to \textit{Laf4ir} mRNA and ultimately lead to the development of double stranded RNA. This would trigger the cellular autonomous immune system to destroy both RNAs. This would trigger the cellular autonomous immune system to destroy both RNAs. Thus, an intron-bypass transcription mechanism in genes with large introns may exist. DNA binding proteins may bring together two exons and direct RNA polymerase jump from one exon to the next bypassing the internal part, which can save energy and does not interfere with the transcription of any nested intronic genes. Alternatively, \textit{Laf4} and \textit{Laf4ir} may be transcribed at different times. Further detailed
investigation on the transcriptional regulation of *Laf4* and *Laf4ir* will help to understand the mechanisms involved in the transcription of the nested genes and their external genes and the transcription of genes with large intron.

*Laf4ir* ORF1 polypeptide and LAF4 protein were found to be associated with each other in the nucleus of dESCs. There was a similar trend of changes of *Laf4ir* ORF1 polypeptide and LAF4 protein levels in the heart in response to TAC treatment. These results suggest that *Laf4ir* ORF1 polypeptide and LAF4 protein may coordinate and modulate each other’s functions. On this aspect, *Laf4* and *Laf4ir* can be a good model to investigate this transcriptional mechanism and the interaction between a nested gene and its external gene.

### 6.2 Multiple translation from single *Laf4ir* mRNA transcript

Alternative translation of two or more peptides/proteins from an mRNA exists and more and more recent experimental data challenges the concept of one peptide/protein encoded from an mRNA. The initiation of the translation of ORFs within an mRNA molecule can be achieved by two mechanisms, the canonical G7-cap pathway and non-canonical internal ribosome entry site. A stop-reinitiation mechanism for the translation of the second ORF following the translation of the first ORF exists in a polycistronic mRNA. This newly identified nested encoding gene, *Laf4ir*, is one such representative. The two ORF polypeptides can be translated differently in a diverse range of organs or tissues. For example, *Laf4ir* ORF2 polypeptide was detected in the aorta endothelium, femoral artery vessel wall adventitia in response to vascular injury, and skeletal muscle tissues in ischemia in wild type mice but *Laf4ir* ORF1 polypeptide was not detected. These observations suggest the polypeptides from these two ORFs may exert different functions in different cells.
Furthermore, the two ORF polypeptides can be translated in the same cell such as ESCs and VPCs. Although expression of both ORFs were different in response to TAC-mediated pressure overload differently, \textit{Laf4ir} ORF1 and ORF2 polypeptides were indeed co-translated within the endothelium of coronary artery. These observations indicate that these two polypeptides may exert certain functions together.

This phenomenon is in line with various different reports indicating some alternative proteins functionally interact with their respective reference proteins. A study found an alternative translation initiation AUG codon in the +3 reading frame of human ataxin-1 (ATXN1) and that this alternative ORF encodes for an a 21-kDa polypeptide referred as Alt-ATXN1 (Alternative ATXN1) \textsuperscript{208}. This was identified by a hemagglutinin tag in-frame with Alt-ATXN1 in ATXN1 cDNA and a cell culture investigation illustrated the co-expression of both ATXN1 and Alt-ATXN1 in nuclear compartments. Interestingly, ATXN1 depletion resulted in an Alt-ATXN1 homogenous nucleoplasmic expression distribution. Also, antibodies specific for both Alt-ATXN1 and ATXN1 showed co-localisation expression in the human cerebellum \textsuperscript{208}. This result suggests that the reference ATXN1 protein regulates the subcellular localisation of the alternative Alt-ATXN1 protein. This co-regulatory subcellular distribution of alternative translational products can also be evident in both \textit{Laf4ir} ORF1 and ORF2 polypeptides.

In another study, it was documented how the upregulation of a reference protein can modulate alternative protein expression at a post-transcriptional level. An alternative protein designated as uORF5 was found upstream of the rat adenosine 2A receptor (A\textsubscript{2A}R) AUG via an anti-uORF5 antibody in rat brains and PC12 cells \textsuperscript{209}. Increased mRNA A\textsubscript{2A}R level in response to hypoxia, resulted in the upregulation of both the A\textsubscript{2A}R and uORF5 proteins. Also, activation of A\textsubscript{2A}R enhanced uORF5 protein expression via post-transcriptional regulation. Further investigation elucidated the functional role of uORF5 in gene expression and modulation of the MAPK pathway strongly associated with the A\textsubscript{2A}R signalling pathway \textsuperscript{209}. Similarly, in another study a
newly discovered alternative translated peptide named as PEP7 of the angiotensin type 1a receptor (AT$_{1a}$R) gene, was found to inhibit the beta-arrestin dependant signalling cascade of AT$_{1a}$R. A dual functional role of both alternative proteins may also be mediated by the Laf4ir gene, therefore it will be interesting to observe if Laf4ir ORF1 polypeptide can also possess cellular proliferative and survival functional roles mediated by the Laf4ir ORF2 polypeptide.

To my knowledge, the first experimental evidence of alternative proteins in the mammalian system, published in cell press in 1995, showed the presence of an alternative ORF of the INK4a tumour repressor gene capable of stimulating G$_1$ and G$_2$ cell cycle arrest. The INK4a (MTS7, CDKNS) gene encodes a protein, p16$^{INK4a}$, which prevents exit from the G$_1$ phase in the cell cycle and in this study, the alternative ORF designated as p19$^{ARF}$ was found to also mediate cell cycle arrest. In correlation to Laf4ir ORF2 polypeptide mediated inhibitory effect in proliferation found in my PhD thesis, Laf4ir ORF1 polypeptide may also underlie the dual requirement in cell cycle regulation.

The TV1 mRNA transcribed by laf4ir was predicted to have five alternative RBS by altORFez software analysis that could drive translation. There was no difference found between the two transcript variants of Laf4ir. A small initiator starter protein was predicted to be produced from position 35 of the mRNA and if produced this could initiate the binding of ribosomes to the re-initiator site at position 97 and potentially drive the expression of Laf4ir ORF1 polypeptide. Multiple studies indicate upstream alternative ORFs at the 5’UTR can inhibit expression of the downstream major ORF. However this translational mechanism is currently under debate as it speculated that when the first AUG start codon is surrounded by a weak kozak motif, the second AUG (with or without an optimal kozak sequence) can be recognised by the 40S ribosomal subunit and utilised to stimulate translation due to leaky scanning. This could be important as the start codon (AUG) of the Laf4ir ORF1 polypeptide might be flanked with an imperfect kozak sequence and regulated by elongation factors. Laf4ir
ORF1 polypeptide could be translated through the canonical mechanism or through RBS#1 and #2, while the Laf4ir ORF2 polypeptide could be translated through the stop-reinitiation mechanism. For example, under TAC conditions, some RNA binding proteins may be activated to bind to the RBS #1 and #2 sites to prevent the access of ribosome and suppress Laf4ir ORF1 translation. In contrast, some RNA binding proteins can bind to RBS #3 -5 to recruit a ribosome to initiate the translation of Laf4ir ORF2 polypeptide.

As mentioned in chapter 3.1, the human transcriptome comprises of 83,886 potential altORFs with a minimum size of 40 codons. It is therefore encouraging to judge that multiple translation in the mammalian system has been disregarded and potentially contribute substantially to the proteome. To tackle this concept would ideally be to predict and clarify expression of alternative ORFs in large scale experimental approaches. Overall, further detailed investigation will be required to elucidate the mechanisms underlying the translation selection between Laf4ir ORF1 and ORF2 polypeptides and their precise functions. Also, the Laf4ir ORF2 antibody used in this PhD thesis cannot distinguish between the two transcript variants of Laf4ir, giving the opportunity for still further differential expression in tissues. On this aspect, Laf4ir can be a good model to investigate mechanisms involved in the translation of mRNAs with two alternative ORFs from the same mRNA.

6.3 Laf4ir role in stem/progenitor proliferation and differentiation

During cell cycle, departure from the G1 phase and transition to a G0 cell cycle arrest phase is an essential step for stem cell differentiation. Previous study from our laboratory group revealed that laminar flow could upregulate p21cip1/waf1 via histone deacetylase-mediated p53 deacetylation and as a consequence encourage cell cycle arrest and enhance differentiation processes. In my PhD project, I elucidated another
mechanism for this process: the binding and interaction between MCM3 protein and \textit{Laf4ir} ORF2 polypeptide.

The MCM family of proteins are highly conserved in all eukaryotes $^{226, 247}$. MCM2-7 and MCM10 complex are generally known as DNA replication licencing factors, propagating a series of processes with other complexes to essentially bind to DNA replication origins and unwind the condensed chromatin. MCM3 protein is strongly associated with different forms of cancer cell proliferation $^{247-252}$. In my study, I found that \textit{Laf4ir} ORF2 polypeptide could interact with MCM3 in the cytosol, potentially leading to the retention of MCM3 in the cytosol. The consequence could be that nuclear MCM3 would be deficient, leading to DNA replication suppression and cell cycle arrest. Deficiency in \textit{Laf4ir} ORF2 polypeptide may allow MCM retaining in the nucleus, leading to elevated cell proliferation. This is supported by the suppressive role of \textit{Laf4ir} in dESCs overexpressing \textit{Laf4ir} ORF2 polypeptide and the opposite effect observed in VPCs from \textit{Laf4ir}$^{+/\text{+}}$ mice. Previous initial work in the role of MCM3 in cell cycle regulation, revealed how MCM3 can interact and be acetylated by MCM3AP, a novel acetyltransferase, to block the initiation of DNA replication and cell cycle progression $^{232, 253}$. More recent studies have linked the association of DExH-box helicase 9 (DHX9) and mico RNA 183 (miR\textsuperscript{-183}) with MCM3 in proliferating cancer cells $^{254, 255}$. MCM2-3 activity was seen to be directly dependant on DHX9 in osteosarcoma cellular proliferation $^{254}$. miR-183 signalling pathway is known to inhibit neuroblastoma growth, and in this study dual-luciferase reporter gene assays revealed the direct targets of miR-183 in MCM3 and MCM5 $^{255}$. In another study, Han et al demonstrated the phosphorylation and activation of MCM3 at Ser-205 position by checkpoint kinase 1 (Chk1) in the regulation of DNA replication $^{231}$. In my project, further investigation is key to establish the mechanistic regulation of \textit{Laf4ir} ORF2 polypeptide in MCM3 protein to potentially suppress stem/progenitor proliferation.
In addition to MCM3’s role in DNA replication, MCM3 can also modulate gene transcription by functioning as a co-activator or co-repressor of certain gene expressions. Hubbi ME et al reported that MCM proteins including MCM3 suppressed hypoxia inducible factor 1 (HIF-1) activity via direct interaction by an O$_2$ and hydroxylation-dependent manner $^{256}$. HIF-1 activity reduces when quiescent cells re-enter the cell cycle and this effect seems to be MCM dependent. Alvarez S et al found that MCM3 was involved in hematopoietic stem cell differentiation $^{257}$. Defected embryonic hematopoietic progenitor cellular differentiation was witnessed in an MCM3 novel knockout model leading to anaemia and reduced life expectancy $^{257}$. Overall, the *Laf4ir* gene may regulate stem cell differentiation through modulating MCM3 protein cellular location and/or interaction with transcription factors via spatiotemporal translation of the different ORFs.

### 6.4 Laf4ir role in endothelial cell survival against oxidative stress

ROS are highly active molecules, existing intracellularly and extracellularly. ROS participates in almost all physiological and pathophysiological processes depending on its level $^{258}$. At low to modest levels, ROS is beneficial to physiological processes such as cell proliferation $^{259}$, differentiation $^{260}$, migration $^{261,262}$ and cell death $^{263}$. In a damaged organ or tissue, the apoptotic cells produce ROS, which will trigger the proliferation of existing cells to repair the damage $^{264}$. However, excessive dose of ROS induces apoptosis. Cells have developed multiple antioxidant defences, one of which is glutathione $^{234}$. Glutathione exists as both reduced and oxidized states. The reduced form can neutralise ROS. Thus, the ratio of reduced to oxidized glutathione modulates cellular ROS level. The homeostasis between reduced glutathione (GSH) and oxidized glutathione (GSSG) is mediated by multiple mechanisms such as biosynthesis$^{265}$ and reduction of oxidized glutathione $^{266}$. Glutathione can form disulphide bond with the cysteine residues in proteins. Thus, some cysteine containing peptides or proteins may
be also involved in the regulation of glutathione level. The peptides from *Laf4ir* ORF2 have odd number of cysteine residue. Apart from the formation of intramolecular disulphide bonds, the thiol group of the spare cysteine residue may act in a similar way as glutathione to neutralise ROS. It can also form disulphide bond with glutathione, therefore regulating the ratio of the reduced to oxidized glutathione. In this study, over-expression of *Laf4ir* ORF2 polypeptide may protect cells from H$_2$O$_2$-induced cell apoptosis and CAD-mediated apoptosis. Further detailed investigation will be required to see if and potentially how *Laf4ir* mediates cellular protection against oxidative stress such as the potential role in glutathione, superoxide dismutase or other antioxidants, or scavenging for ROS, or potentially effecting gene expression associated with oxidative stress defence. However, *Laf4ir* ORF2 polypeptide was secreted into the extracellular environment, which neutralised and reduced the extracellular ROS levels, providing protection against the ROS from the extracellular environment.

The *in-silico* structural model of *Laf4ir* ORF2 polypeptide from TV2 suggests potential transportation out of cells by non-classical secretion and this could provide clues to the possible mechanism observed in the reduction of extracellular H$_2$O$_2$ level. The predictions also implied it may also have a high potential for protein-protein interactions. This highly positively charged protein was shown to have possible PTM sites for N-myristylation and S-farnslylation known to aid in membrane and protein-protein interactions. A binding site was also discovered on the predicted model for N-AGlyc a common PTM applied to other proteins, these predictions all suggest an involvement in dimerisation. These predications may all provide interesting insights in the anti-oxidant mechanistic role of *Laf4ir*.

### 6.5 *Laf4ir* role in cardiovascular remodelling

It has been widely reported that TAC-mediated pressure overload, low shear stress in some areas of aorta, hyperlipidemia, vascular injury can increase local ROS levels $^{267-271}$. In the heart, pressure overload stimulates cardiac hypertrophy and this
propagates the development of heart failure. During this process, mitochondrial complex I-related production of ROS was found to be enhanced with the onset of diastolic functional alterations in a rat TAC-model \(^{272}\). The upregulation of \textit{Laf}4ir ORF2 polypeptide in the TAC model implies that \textit{Laf}4ir ORF2 polypeptide may contribute to maintain the homeostasis of redox in the microenvironment.

The endothelium is highly responsive to shear stress mediated by the mechanical stimuli of blood flow in order to provide vascular tone homeostasis. Atheroprine regions are generally located near branches, exposed to disturbed, oscillatory flow with low-magnitude shear stress \(^{273}\). Conversely, it has been proposed that pulsatile high shear stress is atheroprotective with limited thrombosis, EC apoptosis, reduced permeability and inflammatory responses \(^{274}\). ECs under oscillatory disturbed flow are known to undergo a greater degree of oxidative stress than laminar pulsatile flow. An important study performed a transcriptome analysis of human umbilical vein endothelial cells (HUVECs) exposed to oscillatory shear (OS), pulsatile shear (PS) and static conditions to elucidate the dynamics of endothelial responses in a few functional pathways including oxidative stress, inflammation and cell cycle \(^{275}\). Results indicated the upregulation of ROS producing candidates (e.g. NOX4), and inflammatory markers (e.g. NF-κB, IL-8, MCP-1, Selectin E and VCAM-1), and the downregulation of antioxidants (e.g. catalase) in OS vs PS condition. This study provided evidence of the oxidative and inflammatory state of ECs exposed to OS compared to PS, which is evident in atheroprone regions. The presence of oxidative stress contributes to the oxidative modification of lipoproteins (mainly in the form of ox-LDL), endothelial dysfunction, leukocyte migration and differentiation, nucleic acid damage, VSMC proliferation, MMPs secretion, collagen degradation, platelet aggregation and other factors which accelerate atherogenic processes \(^{276}\). Perhaps, the upregulation of \textit{Laf}4ir ORF2 polypeptide in the mouse \textit{Apo}E\(^{-}\) ascending aorta implies that \textit{Laf}4ir ORF2 polypeptide potentially protects against oxidative stress by reducing H\(_2\)O\(_2\) and/or drive stem/progenitor Sca-1 positive cells towards ECs to replace dysfunctional ECs. However,
the pro-atherogenic effect *Laf4ir* ORF2 polypeptide should not be excluded. The upregulation of *Laf4ir* ORF2 polypeptide in the mouse *ApoE*−/− ascending aorta could drive atherogenic processes by stimulating inflammatory processes in bone marrow derived-leukocytes. Further detailed investigation with the cross breeding of *ApoE* −/− mice with *Laf4ir* −/− mice will provide evidence on the precise role of *Laf4ir* in atherosclerosis.

The difference in *Laf4ir* ORF2 polypeptide expression in ECs from the coronary artery, aorta and femoral artery may reflect the heterogenicity of EC origin \(^{237}\). The expression of *Laf4ir* ORF2 polypeptide in the endothelial cells of WT ascending aorta and bone marrow-derived leucocytes may serve as a purpose to modulate endothelial and inflammatory homeostasis. In the coronary artery, aorta and newly re-endothelialised endothelium of femoral artery following injury, *Laf4ir* ORF2 polypeptide seemed to form a shell for the ECs. It may stick to cell surface glyocalyx through the intermolecular disulphide bond formation between *Laf4ir* ORF2 polypeptide and glycoproteins or other interaction with glycolipids via unknown mechanisms \(^{277}\). The benefit of this shell of *Laf4ir* ORF2 polypeptides is to neutralise excessive ROS in the extracellular environment, preventing the direct damage of ROS on ECs.

During ischaemic conditions, various inflammatory mediators, growth factors and other molecules are released in the microenvironment to promote revascularisation to maintain vascular integrity and homeostasis. Angiogenesis, vasculogenesis and the contribution of circulating/resident stem/progenitor cells towards vascular components have all been heavily documented to take part in ischaemic revascularisation. In response to hindlimb ischaemia, *Laf4ir* ORF2 polypeptide was observed to be significantly upregulated. This might serve as a basis for ischaemic revascularisation. This is supported by the impaired blood perfusion phenotype observed in *Laf4ir* +/− mice compared to *Laf4ir* +/- mice in the hindlimb ischaemia model, potentially mediated by the *Laf4ir* ORF2 polypeptide. Additionally a
fewer number of Sca-1 and CD31 positive cells in \( \text{Laf4ir}^{-/-} \) compared to \( \text{Laf4ir}^{+/-} \) ischaemic muscle suggests a limited level of stem/progenitor contribution to revascularisation. This set of results indicates the role of \( \text{Laf4ir} \) in stimulating stem/progenitor differentiation toward EC lineage via cell cycle arrest to repair ischaemic vessels. The precise mechanism of \( \text{Laf4ir} \) in ischaemic revascularisation is however unestablished.

6.6 \( \text{Laf4ir} \) phenotype in global knockout mice

The \( \text{Laf4ir}^{-/-} \) mice exhibit enlarged heart at baseline and this phenotypic observation can be due to multiple factors. Perhaps \( \text{Laf4ir} \) deletion leads to cardiomyocyte hypertrophy or proliferation/differentiation, abnormal cardiac structural alterations during development and/or ageing and pathogenesis processes such as pressure-overload. \( \text{Laf4ir} \) reduces proliferation in stem/progenitor cell types, therefore the deletion of \( \text{Laf4ir} \) could lead to prenatal or postnatal cardiac hyperplasia. The computational promoter analysis found two TFs (NKX25 and CDXa) with a very high probability as potential regulators of \( \text{Laf4ir} \) gene expression. NKX25 is known to be important in embryonic cardiac development \(^{278}\), while CDXa is part of the Wnt signalling pathway important for EC differentiation \(^{279}\) and begins expression at embryonic stage 8.5 of mouse foetal development \(^{280}\). This finding may be linked to the defect in the heart size and abnormal vessel growth in \( \text{Laf4ir}^{-/-} \) and \( \text{Laf4ir}^{+/-} \) mice.

The reduced adult body size of \( \text{Laf4ir}^{-/-} \) mice indicates a potential defect in developmental process. Reduced body weight is generally a common trait in around a third of cases that use genetic knockout models in mice \(^{281}\). Body weight is highly heritable, and thousands of genes have been associated with this complex manifested appearance. Also the incidence of malocclusion in \( \text{Laf4ir}^{-/-} \) mice might be due to abnormal skeletal remodelling which could lead to misalignment of the jaw and overgrown teeth. This phenotype is in line with the abnormal incisor morphology.
observed in the parent LAF4 gene (Table 2) which suggests potential co-functionality between the Laf4ir and LAF4 genes.

6. 7 Future directions

In my PhD thesis, results have shown the expression prolife of a novel nested encoding Laf4ir gene and its role in cellular survival, proliferation and differentiation, and its implications in cardiovascular remodelling. These findings may ignite insight into the discovery of other mammalian nested genes with a diverse range of unknown functions. However, several issues are left unanswered and will need to be addressed in future studies. Extensive future investigations will utilise various different in-vitro and in-vivo studies to provide more detailed analysis of the functions and regulations of Laf4ir in cardiovascular biology and disease.

Future work will decipher the transcriptional regulation of Laf4ir. This will involve, utilising a pSiCHECK2-Laf4ir reporter to identify the shear stress response element from the Laf4ir gene promoter. Through a serial deletion of the promoter region in the reporter system, the shear response element can be defined and potential transcription factor candidates can be deducted from software analysis. The verification of the involvement of the transcription factors can be achieved through transcription factor knockdown assay. Utilising the luciferase report assay will provide evidence of inducers of Laf4ir gene expression. An array of cytokines and growth factors such as VEGF, TNF-α and INF-γ will be used as potential candidates of Laf4ir transcription and/or translation regulation. VEGFR1,2 are known mechano-transducers and perhaps this might provide a link between the shear stress responses of Laf4ir.

To investigate the translation regulation of ORFs, the Laf4ir TV1 and TV2 sequences will be cloned into a CMV-driven vector. Mutations can be introduced into different RBS sites. The wild type and mutant vectors can be transfected into ECs and
challenged with H₂O₂, followed by assessment on *Laf4ir* ORF1 and ORF2 polypeptide expression. These experiments will provide evidence to detect specific RBS responsible for *Laf4ir* ORF1 or ORF2 translation. Subsequently, biotin-labelled RNA fragment bearing the specific RBS will be synthesised and RNA-mediated protein pulldown assay will be performed using H₂O₂-treated/untreated ECs, followed by proteomics analysis. These experiments will find the potential RNA binding proteins. The association of these RNA binding proteins with ribosomal proteins will then be investigated, which will give clues on the recruitment of ribosome to *Laf4ir* mRNA at different RBS.

The functional implications of *Laf4ir* in cellular differentiation, survival, and proliferation in my PhD project are generally descriptive and lack mechanistic insight. It is unclear whether MCM3 is necessary and sufficient to mediate the biological functions of *Laf4ir*. Broader insight into the mechanistic role of *Laf4ir* ORF2 polypeptide in stem/progenitor MCM3-mediated proliferation will be addressed. This will involve exploring the cellular localisation of MCM3 among VPCs isolated from *Laf4ir* +/+ , *Laf4ir* +/- and *Laf4ir* −/−. It will be expected that more MCM3 is accumulated in the nucleus in *Laf4ir* +/- and *Laf4ir* −/− VPCs compared to *Laf4ir* +/+ VPCs. To verify the role of MCM3 retention, reconstitution of *Laf4ir* ORF2 polypeptide in *Laf4ir* +/− and *Laf4ir* −/− VPCs will be performed via adenoviral gene transfer. This will be followed by analysis of MCM3 cellular localisation and cellular proliferation assays.

It is unclear if *Laf4ir* acts as an oxygen radical scavenger or modulates antioxidant gene expression. To demonstrate the potential neutralising role of thiol groups within the *Laf4ir* ORF2 sequence, mutant *Laf4ir* ORF2 polypeptides will be created by substituting cysteine residues with serine residues. Multiple different vectors with mutated cysteine residue sites will be created. These mutant and wild type vectors will be transfected in ECs and challenged with H₂O₂, followed by H₂O₂ measurements and apoptosis analysis. This investigation will show if *Laf4ir* functions as a ROS scavenger, via the *Laf4ir* ORF2 polypeptide. It will also provide information on whether the thiol
group(s) is responsible for H$_2$O$_2$ neutralization and protection from H$_2$O$_2$-induced apoptosis and also from which cysteine residue.

To explore the mechanistic role of Laf4ir in EC differentiation, $Laf4ir^{+/+}$, $Laf4ir^{+-}$, and $Laf4ir^{-/-}$ VPCs will be subjected to VEGF-induced differentiation process. Subsequently, transcriptome RNA-sequence analysis of gene expression will be conducted. This will then be followed by the examination of EC/SMC marker analysis. This experiment will enable the identification of potential transcription factors responsible for EC differentiation. Reconstitution of Laf4ir ORF2 polypeptide into $Laf4ir^{+/+}$, $Laf4ir^{+-}$ and $Laf4ir^{-/-}$ VPCs will be used to confirm the RNA-sequencing data.

The expression of Laf4ir ORF2 polypeptide does not appear to be specific for cardiovascular tissues. This raises the question of whether it is a general factor required for cell cycle control or whether it has specific functions in the cardiovascular system. The development of loxP sites at the promoter and exon 3 of the Laf4ir gene ($Laf4ir^{loxP/loxP}$) in mice, can be used to crossbred with cell specific Cre mouse line to delete the Laf4ir gene in specific cell types to enable accurate investigation of cardiovascular cell specific functions. According to Mendelian Inheritance, $Laf4ir^{-/-}$ mice should take up 25% of the litters of the $Laf4ir^{+/+}$ X $Laf4ir^{+/+}$ crossbreeding. However, the occurrence of $Laf4ir^{-/-}$ mice was lower than expected. Therefore, the continued decrease in litter number will indicate a reduced survival rate during embryogenesis. If this phenomenon is evident in the global knockout model of Laf4ir, embryos at different stages of development will be harvested to observe embryo viability, size, anatomic features and measurements of specific abnormalities. Given the finding of an enlarged heart in the $Laf4ir^{+/+}$ and $Laf4ir^{-/-}$ mice, comprehensive phenotype studies analysing heart structure and function is needed. Future work using methods of echocardiography, computed tomography (CT), magnetic resonance (MR) and single photon emission computed tomography (SPECT) imaging in Laf4ir knockout mice can measure the end-diastolic volume (EDV), end-systolic volume (ESV), ejection fraction (EF), and left ventricular (LV)
mass to delineate the function of *Laf4ir* in cardiac biology. The potential role of *Laf4ir* in cardiac fibrosis, cardiomyocyte hypertrophy/hyperplasia or capillary density can be detected via histological and staining procedures. The cardiac remodelling in response to pressure overload can also be evaluated by utilising the *Laf4ir* / mouse in the TAC model.

A hindlimb ischemia model was developed in heterozygous mice. In the proposed study, homozygous mice will be utilised in the same procedure to determine whether there is a ‘gene dose dependent’ phenotype in revascularisation of the ischaemic limb. The *Laf4ir* / mice will also be used in a femoral artery injury model to access re-endothelialisation and potential neointima formation via immunostaining and H&E staining. The *Laf4ir* / mice will also be crossbred with *ApoE* / mice in a high fat diet to further analyse the potential role of *Laf4ir* in atherosclerosis. This would help us understand if *Laf4ir* contributes in atherogenesis or protects cells from the atherosclerotic environment. Accumulating in-vivo loss-of-function results in this mouse model will substantially highlight the functional significance of *Laf4ir* in the cardiovascular system.

Homologs were detected in the intron 7 of human *LAF4* gene to *Laf4ir* exon 2 and exon 7 respectively, suggesting that the human *LAF4IR* gene may exist. To determine whether the human *LAF4IR* isoform gene exists would require various different experiments. To clone the human isoform will require the 5′ RACE and 3′ RACE to detect the full length sequence of RNA transcript/s of human *LAF4IR*. Primers designed specifically between the exons of human *LAF4IR* will enable the detection of RNA transcript/s. RNA-sequencing analysis will assist with the confirmation of the human *LAF4IR* transcript/s. Ribosome profiling and the rabbit reticulocyte lysate system kit will enable to characterise RNA transcript/s and detect in-vitro translation from the human *Laf4ir* mRNA sequence from a specific ORF. Specific antibodies against human *Laf4ir* protein product/s from potential ORFs with the correct native protein size
detection can enable us to explore the translational paradigm of human *Laf4ir*. The use of mass spectrometry can confirm the potential translational capability of human *Laf4ir* and detect associated proteins. Knock-out/down and/or over-expression studies of human *Laf4ir* will ultimately enable us to explore the functional role of *Laf4ir* in human cardiovascular physiology and diseases.

### 6.8 Conclusions

- A novel encoding nested intronic gene, *Laf4ir*, has been discovered from the intron 6 of mouse *Laf4 gene*.
- Expression profile of *Laf4ir* ORF2 polypeptide seems to be more ubiquitous than *Laf4ir* ORF1 polypeptide in different organs and cells.
- In stem/progenitor cells, *Laf4ir* gene may regulate cell cycle progression and cell differentiation via spatiotemporal translation of different ORFs that interact with MCM3 spatially different.
- The *Laf4ir* ORF2 polypeptide may contribute to cellular survival via redox homeostasis in endothelial cells.
- In coronary artery, aorta and re-endothelialized endothelium in femoral artery, *Laf4ir* ORF2 polypeptide may be upregulated to prevent ROS-mediated EC damage.
- *Laf4ir* ORF1 polypeptide may co-ordinate with parent LAF4 protein in specific conditions.
- Overall, the novel intronic encoding mouse *Laf4ir* gene may contribute to cardiovascular remodelling through multiple mechanisms.
- Further detailed investigation on *Laf4ir*/*Laf4* gene transcription, translation and biological function analysis will undoubtedly provide new insights into cardiovascular physiological and pathological processes.
Chapter 7: Publications

7.1 Prior review articles


7.2 Abstracts for presentations

- Characterisation of a novel nested gene *Laf4ir* in the cardiovascular system.  
  **Ehteramyan M**, Yi L, Hu Y and Zeng L. Department meeting, King’s College London, December 2018. *(Oral presentation).*

- Identification of a novel shear stress responsive gene in the intron *Laf4*.  
  **Ehteramyan M**, Yi L, Hu Y and Zeng L. 2\textsuperscript{nd} International splicing conference, Lisbon Portugal, July 2018. *(Oral presentation).*

- The novel *Laf4ir* gene protects endothelial cell survival under oxidative stress  
  **Ehteramyan M**, Yi L, Hu Y and Zeng L. British Heart Foundation postgraduate symposium, King’s College London, May 2018. *(Oral presentation).*

- Identification of a novel shear stress responsive gene in the intron *Laf4*.  
  **Ehteramyan M**, Yi L, Hu Y and Zeng L. BSCR (British Society of Cardiovascular Research) Autumn meeting and Early Career Symposium, Oxford University, September 2017. *(Poster presentation).*

7.3 Awards

- Travel grant for poster presentation for PhD students organised by the BSCR Autumn Meeting and Early Career Symposium, Oxford University, 2017.
Chapter 8: References


46. Edwards RG. IVF and the history of stem cells - Embryo stem cells are poised to fulfil their considerable historical potential. *Nature*. 2001;413:349-351.


104. Whitaker GB, Limberg BJ and Rosenbaum JS. Vascular endothelial growth factor receptor-2 and neuropilin-1 form a receptor complex that is responsible for the differential signaling potency of VEGF(165) and VEGF(121). *Journal of Biological Chemistry*. 2001;276:25520-25531.


148. Ma C and Staudt LM. LAF-4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF-4, the gene fused to MLL in t(4;11) leukemias. *Blood*. 1996;87:734-745.


174. !!! INVALID CITATION !!! 173.


205. Abramowitz J, Grenet D, Birnbaumer M, Torres HN and Birnbaumer L. XL alpha s, the extra-long form of the alpha-subunit of the Gs G protein, is significantly longer than suspected, and so is its companion Alex. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101:8366-8371.


