Understanding the Role of the Long Non-Coding RNA MIR503HG in Endothelial-to-Mesenchymal Transition During Vascular Remodelling

João Pedro Pinho Monteiro

Submitted for the degree of Doctor of Philosophy

University of Edinburgh

2020

Declaration

I declare that the work presented in this thesis is entirely my own work, except where stated in the text. This work has not been submitted for any other degree, and to the best of my knowledge contains no material published or written by any other person, except where stated in the text.

João Pedro Pinho Monteiro

September 2020

Acknowledgements

Because it takes more than a lab to make a PhD, I give my most sincere thanks to Maria Pinho and Julia Polivanova for their patience and unwavering support throughout these years.

Finally, I would like to dedicate this thesis to the memory of my father, Alfredo Monteiro, without whom I would not be the man I am today.

Abstract

Endothelial-to-mesenchymal transition (EndMT) is a dynamic biological process present during development and involved in a variety of pathological vascular remodelling scenarios. However, despite our growing understanding of the key cellular alterations required, the precise molecular determinants governing this phenotypical transition remain elusive. With long non-coding RNAs (lncRNA) now emerging as powerful regulators of gene expression we sought to understand their role in the process of EndMT.

To replicate EndMT *in vitro* and characterise its molecular signature, human umbilical vein endothelial cells (HUVEC) and human pulmonary artery endothelial cells (HPAEC) were exposed to a continuous co-treatment of transforming growth factor-beta 2 (TGF-β2) and interleukin 1 beta (IL-1β) for a total of 7 days. Using high-throughput RNA-sequencing analysis of these cells, a total of 103 differential expressed lncRNAs were identified. Of these, the downregulation of the lncRNA MIR503HG was found to be a prevalent feature present in multiple human primary EC types undergoing EndMT *in vitro*. Further analysis revealed that depletion of MIR503HG was sufficient to elicit a robust EndMT phenotype, with a significant increase in the expression of SNAI2, ACTA2 and COL1A1, accompanied by repression of CD31. Conversely, ectopic expression of a single MIR503HG transcript suppressed these hallmark EndMT-associated changes despite TGF-β2 and IL-1β co-treatment. Accompanying RNA-sequencing of these cells showed that the overexpression of MIR503HG alone was able to inhibit over 25% of the EndMT transcriptional profile. Crucially, these changes were found to be independent of the functional regulation of miR-503 and miR-424, found within the MIR503HG locus.

Our findings were then confirmed *in vivo* using a sugen/hypoxia-induced model of pulmonary hypertension (PH) established in endothelial lineage-tracing mice. Here, the expression of the MIR503HG mouse homolog (Gm28730) was significantly downregulated in association with an EndMT profile in the lung. Conversely, targeted up-regulation of MIR503HG in the mouse lung significantly suppressed the appearance of mesenchymal markers in CD31$^+$ cells during PH. Notably, MIR503HG availability was also found to be decreased in lung tissue sections from patients with idiopathic pulmonary arterial hypertension (IPAH) and cultured blood outgrowth ECs isolated from patients with heritable pulmonary arterial hypertension (HPAH).

Collectively, our studies identify MIR503HG as essential in maintaining EC phenotypical commitment and preventing EndMT both *in vitro* and during disease.

Lay Summary

Spanning the entirety of the human body, the vasculature is a highly complex network of blood vessels composed of several specialised cell types. Unsurprisingly, alterations to the function of these cells can have severe consequences to the structure of the vessels and ultimately drive the development of pervasive clinical conditions such as hypertension and atherosclerosis. This disease process is often referred to as vascular remodelling and it generally encompasses changes in the diameter, elasticity and responsiveness of blood vessels. During this process, the endothelial cells (ECs) forming the inner layer of the vasculature lose their identify and become highly dysfunctional. This is called endothelial-to-mesenchymal transition (EndMT). Despite our understanding of the global changes that lead to EndMT during disease, the precise mechanisms that drive this transition are relatively unknown. Novel molecules known as long non-coding RNAs (lncRNAs) have now started to show promise as regulators of cell function and may also be involved EndMT. In this thesis we describe the previously unknown role of lncRNAs in maintaining EC function and preventing EndMT.

To do this, we first replicated the EndMT process by exposing human ECs to factors known to influence vascular remodelling and drive the transition. After evaluating all RNA present in our samples, we discovered that the lncRNA MIR503HG was significantly lower in ECs undergoing EndMT. Interestingly, by specifically targeting MIR503HG in ECs to decrease its expression we were able to fully reproduce the EndMT process, while increased levels of the lncRNA prevented this.

Crucially, the presence of MIR503HG was also found to be nearly abolished in the remodelled vasculature of patients with pulmonary hypertension (PH). This was also

found to be true in an animal model of PH in association with EndMT.

Together, the data presented here identify MIR503HG as essential in preventing EndMT during disease. Thus, opening new therapeutic avenues for the treatment of vascular remodelling.

Publications

Monteiro, J. P. *et al*. Loss of MIR503HG promotes endothelial-to-mesenchymal transition during vascular remodelling. Circ. Res. **(In revision)**.

Monteiro, J. P. *et al*. Endothelial function and dysfunction in the cardiovascular system: the long non-coding road. Cardiovasc. Res. 115, 1692–1704 (2019).

Presentations

"Loss of the long non-coding RNA MIR503HG promotes endothelial-to-mesenchymal transition during vascular remodelling". Oral presentation. Cardiovascular Science Symposium 2019, University of Edinburgh, Edinburgh, UK. June 13th, 2019.

"The long non-coding RNA MIR503HG as a novel regulator of endothelial-to-mesenchymal transition in vascular remodelling". Poster presentation. Keystone Symposia - Noncoding RNAs: Form, Function, Physiology, Keystone Resort, Keystone, Colorado USA. January 13th, 2018.

"Non-Coding RNAs as Regulators of EndMT in Vascular Remodelling". Poster presentation. Non-coding RNA: Recent Insights into the Mechanisms of Action, University of Edinburgh, Edinburgh, UK. June 22nd, 2017.

Definitions/Abbreviations

AA	Arachidonic acid
ACVRL1	Activin receptor-like kinase-1
ADAM	A disintegrin and metalloproteinases
ALK	Activin receptor-like kinases
ANOVA	Analysis of variance
ANRIL	Antisense noncoding RNA in the INK4 locus
APAH	Associated pulmonary arterial hypertension
ApoE	Apolipoprotein E
BACE	Beta-secretase 1
BMEC	Brain microvascular endothelial cells
BMP	Bone morphogenetic protein
BMPR2	Bone morphogenetic protein receptor type 2
BOEC	Blood outgrowth endothelial cell
Bp	Base pair
BRG1	Brahma-like gene 1
CAD	Coronary Artery Disease
cAMP	Cyclic AMP
cDNA	Complementary DNA
cGMP	Cyclic guanosine monophosphate
ChIRP	Chromatin isolation by RNA purification
chr	Chromosome
CLIP	Cross-linking immunoprecipitation
CO2	Carbon dioxide
COSHH	Control of substances hazardous to health
Ct	Cycle threshold
CVD	Cardiovascular disease
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
ds	Double stranded
E	Embryonic day

E coli	Escherichia coli
EC	Endothelial cell
ECE	ET-1 converting enzyme
ECM	Extracellular matrix
EdU	5-ethynyl-2′-deoxyuridine
EGM-2	Endothelial cell basal media plus supplements
EHT	Endothelial-to-haematopoietic transition
EMT	Epithelial-to-mesenchymal transition
ENCODE	The Encyclopaedia of DNA Elements
EndMT	Endothelial-to-mesenchymal transition
ENG	Endoglin
eNOS	Endothelial nitric oxide synthase
ERA	Endothelin receptor antagonists
ET-1	Endothelin
EV	Extracellular vesicle
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FC	Fold change
FGF	Fibroblast growth factor
FISH	Fluorescent in situ hybridisation
FPKM	Fragments Per Kilobase of transcript per Million mapped reads
FSP1	Fibroblast-specific protein 1
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GAS5	Growth arrest specific 5
GATA6-AS	GATA6 antisense RNA
GFP	Green fluorescent protein
GR	Glucocorticoid receptor
GRE	Glucocorticoid response element
GTP	Guanosine triphosphate
GWA	Genomewide association study
HCAEC	Human Coronary Artery Endothelial Cells
HCC	Hepatocellular carcinoma
hEGF	Human erythroid growth factor

hESC	Human embryonic stem cell
Het	Heterozygous
HF	Heart failure
HIF-1α	Hypoxia inducible factor 1α
HIMEC	Human intestinal microvascular endothelial cell
HMVEC	Human dermal microvascular endothelial cell
HOTAIR	Hox antisense intergenic RNA
HOTTIP	HOXA transcript at the distal tip
HPAEC	Human pulmonary artery endothelial cell
HPAH	Heritable pulmonary arterial hypertension
HSVEC	Human saphenous vein endothelial cell
HUVEC	Human umbilical vein endothelial cell
ICAM-1	Intercellular Adhesion Molecule 1
IL-1β	Interleukin 1 beta
IP3	Inositol triphosphate
IPAH	Idiopathic pulmonary arterial hypertension
KD	Knockdown
KLF2	Kruppel-like factor 2
KO	Knock-out
LB	Luria Broth
LincRNA	Long intergenic non-coding RNA
LncRNA	Long non-coding RNA
LOXL2	Lysyl oxidase homolog 2
MALAT1	Metastasis associated lung adenocarcinoma transcript 1
MEG3	Maternally expressed gene 3
MHC	Myosin heavy chain
MI	Myocardial infarction
miRNA	MicroRNA
mM	Millimolar
MOI	Multiplicity of infection
mPAP	mean pulmonary arterial pressure
MPP	Matrix metalloproteinase
mRNA	Messenger RNA

mTOR	Mammalian target of rapamycin
ncRNA	Non-coding RNA
NF-κB	Nuclear Factor Kappa Beta
NICD	Notch intracellular domain
nM	Nanomolar
NO	Nitric oxide
NOTCH4	Notch homolog protein 4
NS	Not significant
nt	Nucleotide
Oligo	Oligonucleotide
ORF	Open reading frame
PAD	Peripheral arterial disease
PAH	Pulmonary arterial hypertension
PAP	Pulmonary arterial pressure
PCA	Principle Component analysis
PcG	Polycomb group complexes
PCR	Polymerase chain reaction
PDE-5	Phosphodiesterase type 5
PECAM1	Platelet endothelial cell adhesion molecule 1
pen/strep	Penicillin/streptomycin
PFA	Paraformaldehyde
PGI2	Prostacyclin
PH	Pulmonary hypertension
piRNA	PIWI-interacting RNA
PKC-δ	Protein kinase C-delta
PMVEC	Pulmonary microvascular endothelial cell
Pol II	Polymerase II
Pol III	Polymerase III
PRC2	Polycomb Repressive Complex 2
pre-miRNA	Premature microRNA
pri-miRNA	Primary microRNA
PSA	Prostate-Specific Antigen
PVR	Pulmonary vascular resistance

qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RAP	RNA antisense purification
RCC	Renal cell carcinoma
RIN	RNA integrity number
RIP	RNA immunoprecipitation
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNAseq	RNA sequencing
RQ	Relative Quantity
rRNA	Ribosomal RNA
RT	Reverse Transcriptase
RUNX3	Runt-related transcription factor 3
RV	Right ventricle
scRNAseq	Single-cell RNA sequencing
SEM	Standard error of the mean
sgRNA	Single guide RNA
siRNA	Small interfering RNA
SMC	Smooth muscle cell
SMILR	Smooth Muscle Induced LncRNA enhances Replication
SM-MHC	Smooth muscle myosin heavy chain
snoRNA	Small nucleolar RNA
SuHx	Sugen 5146/Hypoxia model
TBE	Tris-Borate EDTA
TF	Transcription factor
TGF-β	Transforming growth factor-β
TGF-βR2	TGF-β type II receptor
TINCR	Terminal differentiation-induced ncRNA
TNF-α	Tumour necrosis factor alpha
tRNA	Transfer RNA
TWIST1	Twist Basic Helix-Loop-Helix transcription factor 1
UBC	Ubiquitin C
UCSC	University of California, Santa Cruz

UTR	Untranslated region
VCAM-1	Vascular cell adhesion protein 1
VE-Cadherin	Vascular endothelial cadherin
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
VSMC	Vascular smooth muscle cells
WHO	World health organisation
Wnt	Wingless-type MMTV integration site family member
WT	Wild type
XCI	X-chromosome inactivation
Xist	X- inactive specific transcript
YFP	Yellow fluorescent protein
ZEB1	Zinc-finger E-box-binding homeobox 1
ZEB2	Zinc-finger E-box-binding homeobox 2
αSMA	Alpha smooth muscle actin
μg	Microgram
μl	Microliter

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Overview

The vasculature is an intricate and highly adaptable network of blood vessels spanning the entirety of the human body. Unsurprisingly, alterations to the structural and mechanistic functioning of the vascular wall can have severe consequences that underline the development of pervasive subclinical conditions such as hypertension and atherosclerosis. Vascular remodelling is the term used to describe this maladaptive process, broadly encompassing changes in vascular diameter, elasticity and responsiveness. Neointima formation and intimal hyperplasia are often present during vascular remodelling scenarios and feature in the pathogenesis of several diseases such as vein-graft restenosis [1], atherosclerosis [2] and pulmonary arterial hypertension (PAH) [3]. This is often accompanied by dysregulated cell migration, proliferation and apoptosis, along with increased inflammation and excessive deposition of extracellular matrix (ECM).

Within the setting of PAH, despite its complex aetiology, it is accepted that the key elements in the pathogenesis of the disease encompass a variety of structural changes to the vascular wall [4]. This includes stiffening of proximal pulmonary arteries, increased intimal and medial arterial thickness, along with the eventual development of complex neointimal lesions. Underlying many of these changes is the increased deposition of ECM and appearance of smooth muscle-like cells with a highly proliferative and migratory potential [5].

Similarly, the common use of venous bypass grafts for surgical revascularisation in patients with coronary artery disease can also be followed by a rapid decrease in lumen diameter [6]. As with the previous condition, despite their aetiological differences, the observed pathological post-grafting remodelling is characterised by the development

of neointimal lesions composed of proliferative smooth muscle cells (SMC) and ECM proteins [7,8].

Leading from the initial disruption of normal endothelial function, neointimal formation and hyperplasia is believed to be preceded by a cascade reaction involving the activation and adhesion of circulating thrombocytes and leukocytes which work to initiate medial SMC differentiation, migration and proliferation [9]. Nevertheless, while the appearance of neointimal cells is widely reported, their origin is still up for debate [10,11]. Emerging paradigms, supported by lineage tracing studies, now propose that in part these SMC-like cells are of endothelial origin [12–14]. Specifically, these studies highlight the pathological contribution offered by the endothelial cell (EC) capacity to transition between proliferative mesenchymal phenotypes [15]. With emerging evidence showing that endothelial-to-mesenchymal transition (EndMT) is present during pathological vascular remodelling, new therapeutic avenues may also arise. However, despite our growing understanding of the key cellular alterations involved, the molecular determinants governing this phenotypical transition still remain elusive. This is particularly true with the discovery of long non-coding RNAs (lncRNA) as powerful regulators of gene expression both at the transcriptional and post-transcriptional level.

In this thesis, we summarise the current body of knowledge on the role of EndMT in the context of pathological vascular remodelling, and the role played by ncRNAs in governing this phenotypical transition. Lastly, we provide evidence for the role of the lncRNA MIR503HG in preventing the initiation of EndMT both *in vitro* and during disease.

1.2 Endothelial-to-Mesenchymal Transition

Lining the totality of the vasculature, the endothelium directly interacts with nearly every system in the body to regulate most aspects of vascular development, homeostasis and pathogenesis. The EC monolayer that constitutes this barrier is in itself a distinctively versatile population, showing extraordinary physiological and morphological heterogeneity throughout different parts of the vasculature [16]. These differences reflect the variety of functions they perform and the shifts in the priority of these roles across different anatomical locations in the cardiovascular system. For instance, the angiogenic capacity of ECs plays a vital role during embryological growth, tissue development, and wound healing in damaged tissues [17]. Maintenance of vessel function, and therefore cardiovascular homeostasis, is highly dependent on the ability of the endothelium to react to external stimuli, mediating not only vasodilation and constriction but also thrombogenic, immune and inflammatory responses [18,19]. Dysregulated, these mechanisms can lead to, among others, the unrestrained vessel formation often seen in cancerous tumours [20] and the pathological remodelling of mature vessels associated with the development of pervasive conditions such as hypertension [21] and atherosclerosis [22].

Recently, several studies have started to look at the process by which ECs adopt mesenchymal phenotypes and the role that this transition plays both in the initiation and maintenance of vascular disease. This transition refers to the process whereby ECs will progressively lose their cell-specific markers, change their morphology and adopt a mesenchymal cell-like phenotype accompanied by increased motility, cytoskeletal modifications and cell-to-cell junction rearrangement [23] (Figure 1.1). First described as a key developmental process for heart valve formation, during disease these cells

are thought to acquire invasive and proliferative properties along with the ability to deposit large amounts of ECM [23].

In order to transition to a mesenchymal phenotype, ECs must first lose their cell-to-cell tight and adhesion junctions via increased cleavage of vascular endothelial (VE)-cadherin and platelet endothelial cell adhesion molecule 1 (PECAM1) proteins, accompanied by decreased gene expression [15]. Occurring in parallel with this, there are distinct changes in shape from their normal arrangement to an elongated morphology [24]. As the transition progresses, the cells will begin to express mesenchymal markers – specifically, there is an increased expression of known fibroblastic markers such as α-smooth muscle actin (α-SMA) and fibroblast-specific protein 1 (FSP1), along with various leukocyte adhesion molecules and a combination of ECM proteins, such as collagen I and III, vimentin and fibronectin [10,14,25]. The transition, and associated cellular changes, involves an intricate series of signalling pathways which regulate both its initiation and potentiation.

Endothelial Markers	EndMT Markers	Mesenchymal Markers
Downregulated	*Upregulated*	*Upregulated*
PECAM1, VE-Cadherin, vWF, Tie-1, Tie-2	SNAI1, SNAI2, ZEB1, ZEB2, TWIST1	α-SMA, SM22α, FSP1, Fibronectin, Collagen, Vimentin
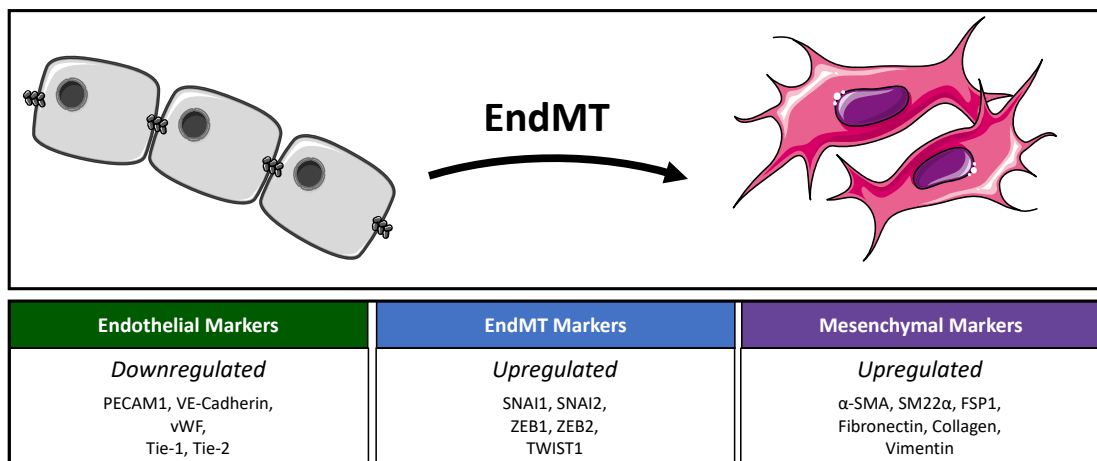
Figure 1.1: Markers of EndMT.

EndMT progression is characterised by changes in a variety of markers. Following the upregulation of EndMT-associated transcription factors, early endothelial responses include the downregulation of endothelial markers and adherens junction dismantling. As it progresses, mesenchymal markers such as cytoskeletal and ECM proteins, are strongly upregulated.

1.2.1 Signalling Pathways Regulating EndMT

EndMT responses can be induced by a variety of events leading to a rise in cellular stress, such as hypoxia [26], direct vascular injury [27] or increased inflammation [28]. These events can trigger the activation of transforming growth factor-β (TGF-β), Wnt and Notch intercellular signalling pathways, which in turn act to direct and/or mediate EndMT by interacting with several transcription factors regulating endothelial and mesenchymal gene expression [14,27,29,30].

1.2.1.1 Transcription factors

Broadly speaking, our understanding of the transcriptional factors involved in the initiation of EndMT largely build on previous foundational studies conducted on epithelial-to-mesenchymal transition (EMT). With this, five transcription factors have emerged as key regulators of both EMT and EndMT. Namely, the SNAIL family zinc finger transcription factors SNAI1 and SNAI2, the zinc-finger E-box-binding homeobox 1 (ZEB1) and 2 (ZEB2), and the twist Basic Helix-Loop-Helix transcription factor 1 (TWIST1). The transcription factors SNAI1 and SNAI2, for instance, were first shown to actively repress E-cadherin expression in transitioning epithelial cells [31,32]. During EndMT, the same transitional repressors have been found to lower the expression of the analogous VE-cadherin and PECAM1 in ECs [15]. As described by Lopez and colleagues, both proteins can bind to E-box motifs in the human VE-cadherin promoter region and thus repress the expression of the adhesion molecule [33]. Similarly, the zinc-finger E-box-binding proteins ZEB1 and ZEB2 were first described as direct transcriptional repressors of E-cadherin expression during EMT [34–36]. In the context of EndMT, while this mechanism hasn't been thoroughly explored, the two

transcription factors are often associated with loss of EC markers [37,38].

Part of a large family of helix-loop-helix (HLH) transcription factors, TWIST1 can function by both dimerising with other basic HLH member proteins and binding specific E-box motifs to regulate gene transcription in EMT and EndMT [39]. Again, this has been historically associated with repression of E-cadherin transcriptional expression, shown to promote EMT during development and in tumor metastasis [40,41]. As with SNAI1/2 , TWIST1 can also directly repress VE-cadherin transcription and is associated to EndMT-like changes in a variety of settings, including heart valve formation, vascular lesion development and fibrosis [33,42–45]. During the formation of the heart valves in mice, for example, persistent TWIST1 expression leads to increased cell proliferation and increased ECM gene expression, characteristic of early developmental EndMT [43].

Interestingly, despite being broadly described in the literature as central to EndMT, the mechanism by which these transcription factors regulate mesenchymal gene expression remains somewhat unclear and surprisingly under-researched. Nonetheless, there is clear evidence that all of these are intrinsically linked to the initiation of EndMT and associated transcriptional changes. Overexpression of SNAI1 alone is sufficient to induce an EndMT-like profile, not only repressing VE-cadherin and PECAM1, but also leading to the increased production of mesenchymal markers such as SM22α and FSP1 [40]. In a model of SMC differentiation *in vitro*, ZEB1 was shown to interact with the TGF-β effector protein SMAD3 to activate αSMA and smooth muscle myosin heavy chain (SM-MHC) promoters [46]. While not validated in ECs it's possible that similar mechanisms are recapitulated during EndMT.

Lastly, while SNAI1 and SNAI2 are often cited as key factors for EndMT induction,

overexpression of either transcription factor alone is not sufficient to fully induce transition [47,48]. Given that several of these mediators are often expressed together, it is possible that a synergistic network of factors regulating EndMT induction exists.

1.2.1.2 TGF-β Signalling

Linked to a variety of cardiovascular diseases associated with tissue remodelling [49,50], TGF-*β* signalling is well established as one of the key signalling pathways involved in the initiation of EMT and, more recently, EndMT [51,52]. This includes the thoroughly described prototypic members of the TGF-*β* Family, TGF-*β*1 and TGF-*β*2 [12,14,53].

In most cell types, members of the TGF-β family will commonly exert their cellular effects by specifically binding a transmembrane complex composed of type I and type II receptors, where upon binding type I receptors are phosphorylated by a type II receptor. Different TGF-β family proteins can bind various combinations of type I and type II receptor complexes, activating specific signaling pathways. There are seven known type I receptors, also described as activin receptor-like kinases (ALK), listed as ALK1 to ALK7. Five type II receptors have been described so far, including the activin receptor type IIA (ActRIIA), activin receptor type IIB (ActRIIB), BMP type II receptor (BMPR2), TGF-β type II receptor (TGF-βR2) and AMH type II receptor (AMHRII) [15]. Activated, the receptor complexes lead to a signaling cascade transmitted by a series of receptor-regulated (R)-SMAD proteins (i.e. SMAD1, 2, 3, 5, or 8), which can then form complexes with a common-mediator (Co)-SMAD (SMAD4), shuttle to the nucleus and participate in the transcriptional regulation of target genes [15,54,55].

In ECs, TGF-β1 and 2 will typically exert their cellular effects via the constitutively active TGF-βR2 which will transphosphorylate the broadly expressed type I receptor

AKL5 or the endothelial specific ALK1 in a common complex. Either receptor complex can then recruit and phosphorylate a variety of R-SMADS – ALK5 will induce the activation SMADS2/3, whereas ALK1 is associated with SMADS1/5/8. Notably, which receptor complex is activated has been shown to be largely context dependent, associated with a variety of factors such as TGF-β ligand concentration, relative expression levels of each receptor, and co-activated inhibitory pathways [54,56–58]. Regardless of their arrangement, the phosphorylated R-SMADS will form additional complexes with SMAD4 and translocate to the nucleus to regulate various aspects of endothelial function [15]. The SMAD1/5/8 and SMAD4 complex, for example, has been shown to promote EC functions such as angiogenesis [56,59,60]. During EndMT, the SMAD2/3/4 complexes are thought to interact with SNAI1/2, ZEB1/2 and TWIST1, resulting in the regulation of endothelial, mesenchymal and other EndMT-associated genes (Figure 1.2) [14,48,61].

Additionally, TGF-β can activate other, non-canonical, SMAD-independent pathways also implicated in EndMT. For example, exposure to TGF-β2 can induce an *in vitro* EndMT profile associated with the activation of p38 MAPK signalling, which when blocked will supress SNAI1 expression and inhibit transition [48]. While the mechanisms by which p38 MAPK signalling regulates EndMT are yet to be defined, a similar role has been described in EMT both during development and disease [62]. Further, in TGF-β1-induced EndMT, both protein kinase C-delta (PKC-δ) and c-Abl can regulate the transition via the inactivation of GSK3β [63]. Crucially, GSK3β has been shown to directly bind SNAI1 ultimately increased its degradation [64,65].

Activation of TGF-β signalling, be it via canonical or non-canonical pathways, while important cannot fully explain the appearance of EndMT. Several publications have

demonstrated this, highlighting that only partial transition phenotypes are possible and that additional pathways must also be activated in parallel.
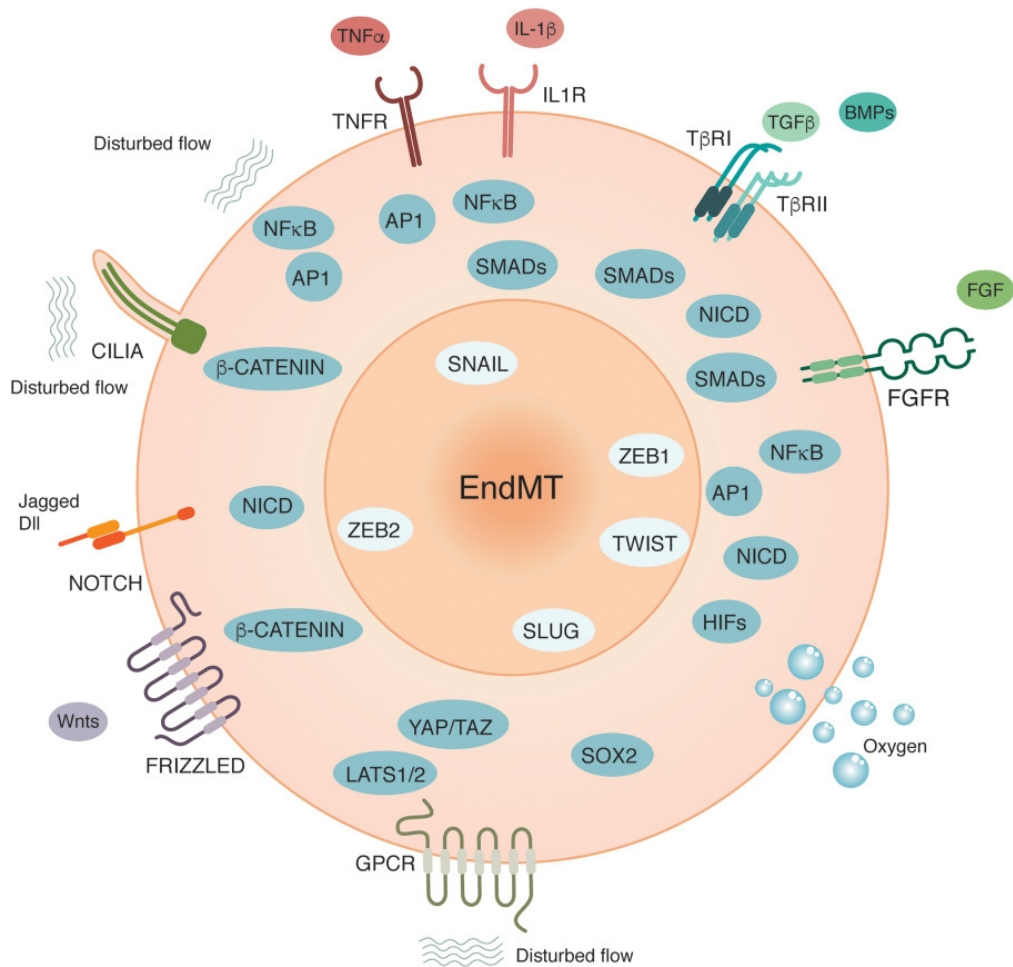
Figure 1.2: Summary of the Regulatory Signaling Cascades Involved in EndMT.
EndMT can be induced by a variety of signal transduction pathways, including those activated by TGF-β, Wnt and Notch ligands, as well as shear stress, hypoxia and inflammation. These in turn act to regulate gene expression by interacting with transcription factors such as SNAI1, SNAI2, ZEB1/2 and TWIST1. Figure adapted from Sánchez-Duffhues *et al*, 2018 [66] .

1.2.1.3 Inflammation

Under chronic inflammation, continuous endothelial activation by tumour necrosis factor alpha (TNF-α), interleukin-1 beta (IL-1β) and other pro-inflammatory cytokines, can drastically impair normal EC function contributing to the initiation and progression of vascular disease [67,68]. Interestingly, a growing number of studies have found this increase in inflammatory signalling to also be associated with EndMT (Figure 1.2) [69]. For instance, 48 hour exposure of aortic valve ECs to TNF-α induced loss of PECAM1 and VE-cadherin, along with increased α-SMA expression and the acquisition of an invasive phenotype in a dose-dependent manner [28]. While this was only seen in a specific subset of ECs, a follow-up study by the same group went on to demonstrate that this EndMT subset also expressed higher levels of Notch1 and TGF-β1 [70].

EMT studies have suggested regulation can happen via the TNF-α mediated activation and nuclear translocation of the transcription factor NF-κB (Nuclear factor kappa-light-chain-enhancer of activated B cells), which can then regulate SNAI1 expression [71–73]. A recent study as also suggested that the downregulation of the BMP Receptor Type 2 (BMPR2) is also necessary for TNF-α to induce EndMT. In fact, ectopic overexpression of BMPR2 in primary human coronary aortic endothelial cells (HAoECs) was sufficient to partially prevent EndMT [74]. Nevertheless, while often alluded to in the literature, these complex cross-regulatory pathways have yet to be thoroughly explored during EndMT.

Several publications have also demonstrated that activated inflammatory and TGF-β signalling can act in a synergistic manner to induce EndMT. For example, as reported by Rieder et al, human intestinal microvascular ECs (HIMECs) display a stronger EndMT

profile when treated with a combination of TNF-α and TGF-β1 [75]. Exposure to either factor alone only elicited partial EndMT, where TGF-β1 was associated with increased αSMA expression while TNF-α mainly induced a downregulation of PECAM1 and VE-cadherin [75].

IL-1β expression has long been known to induce significant phenotypical changes to endothelial and epithelial cells [76–79], more recently associated with EndMT-mediated fibrosis [77,80]. This critical role has been confirmed further by Maleszewska and colleagues, showing that IL-1β will act synergistically with TGF-β2 to induce a pronounced EndMT profile in human umbilical vein endothelial cells (HUVEC) [81]. This strategy was also associated with progressive increase in TGF-β2 expression and the activation of NF-κB signalling [81].

1.2.1.4 Other Signalling Pathways

Despite the well-defined pathway, TGF-β signalling does not fully encompass the totality of the mechanisms involved in the activation of EndMT. Crucially, while some of the pathways described here appear to partially converge with that of TGF-β, other independent pathways have been characterised (Figure 1.2).

Notch Signalling Pathway

Notch signalling, active during cardiac development and disease [82,83], is an important regulator of EC proliferation, migration and differentiation [84–87]. Dependent on cell–cell contact, the interaction between Notch receptors (NOTCH1-4) and their respective transmembrane ligands (i.e. Jagged1) will initiate the canonical Notch signalling pathway. Ligand–receptor interaction leads to 2 successive cleavage events by a disintegrin and metalloproteinases (ADAM) and presenilin, releasing the Notch

receptor extracellular subunit and the Notch intracellular domain (NICD), respectively. This facilitates the translocation of NICD into the cell nucleus where it will form a transcriptional complex able to recruit various transcription factors to regulate gene expression [82,83]. In ECs, activation of NOTCH1 for instance, can repress expression of p21, an endogenous cell cycle inhibitor, to increase proliferation in pulmonary arterial endothelial cells (PAEC) and regulate vessel remodelling in a pulmonary hypertension (PH) model [88]. Additionally, Notch signalling has also been shown to directly interact with the runt-related transcription factor 3 (RUNX3) to induces the expression of several EndMT genes, including SNAI2 [89]. In fact, overexpression of activated NOTCH4 NICDs is sufficient do induce both the loss of endothelial phenotype and acquisition of mesenchymal markers in ECs from different vascular beds [30].

Wnt/β-Catenin Signalling

Wnt/β-Catenin signalling is another conserved pathway reported to be activated during EndMT. Canonical Wnt signalling broadly functions by facilitating the accumulation and eventual nuclear translocation of β-catenin, where it acts on gene transcription [90,91]. In the absence of Wnt activation, a constitutively active 'destruction complex' will phosphorylate cytoplasmic β-catenin, which eventually leads to its degradation. Activation via binding of extracellular Wnt ligands to Frizzled (Fzd) and LRP5/6 receptor proteins will promote disassembly of this 'destruction complex' and inhibit the phosphorylating activity of GSK3β, ultimately allowing for signalling transduction via β-catenin [90,91].

Both *in vivo* and *in vitro* studies describe this process as essential for EndMT-mediated heart cushion formation in mice, which can be halted by the deletion of β-catenin [92].

Using an endothelial reporter mouse line, expressing a Cre-activated fluorescent marker in cells of endothelial origin, a recent publication has also demonstrated that endocardial ECs give rise to mesenchymal progenitors in response to paracrine Wnt signalling [93]. During disease, again using an endothelial-lineage tracing reporter mouse line, myocardial infarction has been shown to trigger EndMT via canonical Wnt signalling [27].

It is important to highlight, however, that the transcriptional mediators of EndMT activated by β-catenin have not yet been elucidated. Nonetheless, Wnt signalling can inhibit GSK3β activity which, as previously emphasised, may result in the stabilisation of SNAI1 and SNAI2 [64].

Lastly, despite our growing understanding of the mechanisms behind EndMT and the role that specific genes and signalling cascades can have on its initiation, there is still vast scope to expand on the underlying components that regulate EndMT and its associated pathologies. One such component has started to gather attention as research increasingly acknowledges the several types of ncRNAs involved in the regulation of gene expression and translation. Evidence now suggests that small non-coding microRNAs (miRNA) and long non-coding RNAs (lncRNA) may act as regulators of key cellular functions including differentiation, migration and proliferation. As such, the role ncRNA during the initiation and maintenance of EndMT may be a significant one. We will continue to explore the role of ncRNAs in endothelial function and EndMT, in Sections 1.4 and 1.5 of this thesis.

1.2.2   Role in Cardiac Development

EndMT was originally reported in series of seminal studies by Markwald and colleagues, describing a subset of endocardial ECs undergoing a phenotypical change

between embryonic development days E8.5 and E12. These would subsequently invade the adjacent cardiac jelly and initiate the formation of the cardiac cushions in the chick embryo [94–96]. Now described in a variety of animal models, EndMT has been known to be one of the key events driving endocardial cushion formation, leading to the development of the cardiac valves [97]. Crucially, endothelial-specific deletion of EndMT signalling pathways in mouse embryos results in decreased endocardial cushion size, followed by critical defects in atrioventricular septa and valves [92,98,99].

The development of endothelial-lineage tracing reporter mouse lines has greatly advanced our understanding of this process by allowing for the identification of EndMT-derived cells that no longer express endothelial markers [100]. For example, using Tie2-Cre transgenic mice line, in which expression of Cre recombinase is driven by an endothelial-specific promoter (Tie2), de Lange *et al* demonstrated that a large proportion of the mesenchymal cells present in mature valve and septal structures of the heart where in fact of endothelial origin [101].

Further, while controversial, some studies have suggested that EndMT may also be active throughout vascular development. This has been reported during the formation of the dorsal aorta in quail embryos [102] and pulmonary artery development of chicken embryos [103]. Of note, using the more rigorous lineage tracking approach, a recent publication has demonstrated that endocardial ECs act as progenitors for pericytes and vascular SMCs (VSMC) assembled within the wall of coronary vessels in mouse embryos [93].

1.2.3   Role in Vascular Pathophysiology

The past decade has seen an increasing number of reports showing that EndMT can occur not only as a mechanism for heart development but also as a driver for a wide

range of adult pathologies ranging from cancer to cardiovascular disease [27,104]. The development of cardiac fibrosis, for example, is now believed to be significantly affected by EndMT with murine studies estimating that up to 35% of all fibroblasts in fibrotic heart tissue are of endothelial origin [14,37]. Similarly, recent studies have highlighted the contribution of EndMT to neointima formation and intimal hyperplasia often present during vascular remodelling.

1.2.3.1 Vein graft remodelling

Vascular remodelling can occur in patients after percutaneous coronary interventions such as angioplasty or coronary artery bypass graft surgery, resulting in neointimal hyperplasia, restenosis and ultimately vessel occlusion. Coronary artery bypass surgery using saphenous vein grafts, for example, is reported to have a fail rate of up to 40% within 18 months of implantation [105,106]. Despite the commonly described accumulation of SMCs and excessive deposition of extracellular matrix seen in these occluded vessels, much of the aetiology associated with vein graft remodelling stems from the damaged endothelial lining failing to adapt to arterial pressure [107,108]. Nevertheless, the cellular contribution to the development of this population in remodelled vessels has often been considered to be lineage restricted [109].

Challenging the prevailing dogma, an influential publication by Cooley et al showed that endothelial lineage-derived cells significantly contribute to active vein graft remodelling as a result of EndMT (Figure 1.3) [14]. Using cell lineage-tracing in mice, where Cre-activated yellow fluorescent protein (YFP) expression allowed for the detection of ECs regardless of subsequent changes in cellular phenotype, the group demonstrated that 51% of the vascular neointimal cell population developing after vein graft surgery were of endothelial origin [14]. Post-grafting, this YFP+ neointimal

population progressively loss the expression of all examined endothelial markers and by day 35 uniformly expressed αSMA and SM22α, indicative of active EndMT. Of interest, the group also found that while the number of non-endothelial lineage (YFP⁻) neointimal cells seemed to plateau after14 days, the number of YFP⁺ neointimal cells continued to increase up to 35 days post-surgery [14].
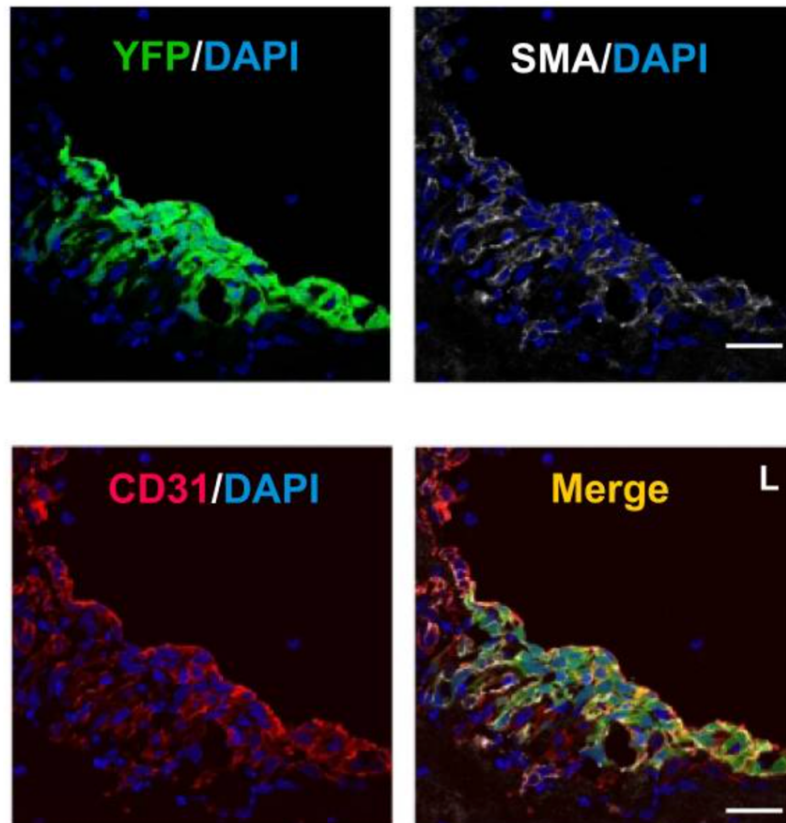
Figure 1.3: EndMT during vein graft remodelling.
Immunofluorescence confocal microscopy showing the expression of CD31 (PECAM1) and SMA in endothelial lineage derived (YFP⁺) cells 7 days after grafting. Scale bars 10 μm. L, lumen. Figure adapted from Cooley *et al*, 2014 [14].

Given the well documented activation of TGF-β signalling during vascular remodelling [110,111], Cooley *et al* go on to also show the presence of phosphorylated Smad2/3 in YFP⁺ neointimal cells by day 7, accompanied by increased SNAI2 and TWIST1 expression. Treatment of mouse veins with either a TGF-β neutralising

antibody prior to grafting, resulted in decreased Smad2/3 activation, accompanied by less EndMT and consequently a reduced neointimal area [14].

1.2.3.2 Atherosclerosis and Plaque formation

Leading from the initial disruption of normal endothelial function, atherosclerotic plaque formation is thought to involve a complex sequence of biological events within the intima of the vessel wall; this includes the recruitment of circulating leukocytes and accompanying macrophage formation, lipid deposition, along with the accumulation of mesenchymal cells and increased ECM deposition [112,113]. This mesenchymal cell population actively patriciates in plaque calcification and fibrous cap formation facilitating disease progression [112]. Historically, the origin of these cells was thought to be limited to a phenotypically distinct, albeit migratory, SMC and fibroblast population from the local vasculature. However, with a growing number of publications highlighting the presence of EndMT in atherosclerosis, these assumptions are quickly changing.

Moonen and colleagues were among the first to report the presence of EndMT in human atherosclerotic plaque and neointimal lesions. Here the group described a population of ECs, expressing both PECAM1 and αSMA, present in the inner lining as well as deeper in the neointimal tissue [114]. Given the more frequently observed incidence of atherosclerotic plaques at branching points within the vasculature [115], the paper also looked at regions exposed to either uniform laminar shear stress or disturbed oscillatory flow in the porcine abdominal aortic trifurcation. In the outer walls of the vessel trifurcation, showing pronounced neointimal thickening, the group reported the appearance of EndMT-like cells expressing αSMA, transgelin, and calponin [114].

Further, the use of endothelial-specific lineage-tracking has vastly expanded our understanding of the contribution of EndMT-derived in atherosclerosis. By combining a tamoxifen-inducible endothelial-specific lineage tracing system in a pro-atherosclerotic apolipoprotein E knockout (ApoE$^{-/-}$) mouse line, Evrard *et al* showed that EndMT contributed to approximately 45% of the fibroblast population within the intima of advanced atherosclerotic lesions (**Figure 1.4**) [53]. Interestingly, EndMT-derived cells expressing VSMC markers were significantly less within these lesions. This was further confirmed in human atherosclerotic lesions, showing that cells expressing both endothelial and fibroblast proteins are increasingly present in advanced unstable lesions [53].

Several reports, using either disease models or patient samples, have now shown a clear accumulation of EndMT-derived mesenchymal cells during atherosclerosis. Nonetheless, whether EndMT plays a functional role or is simply a consequence of atherosclerotic disease still remains somewhat unclear. This was partially addressed in a recent study by Chen et al showing that endothelial TGF-β signalling is one of the primary drivers of atherosclerosis-associated vascular inflammation [116]. Using an endothelial-specific inducible deletion of TGFBR1/2 in an atherosensitive ApoE$^{-/-}$ mouse line, the authors were able to fully block endothelial TGF-β signalling either prior or during the induction of atherosclerosis. Receptor knockout prior to induction of atherosclerosis, significantly delayed disease onset and reduced plaque size by over 60%. Suppression of endothelial TGF-β signalling during established atherosclerosis, successfully reduced disease progression.

Crucially, single-cell RNA-sequencing (scRNA-seq) of aortic ECs obtained from these models identified a multifaceted EndMT phenotype in atherosclerotic mice with

intact TGF-β signalling. Conversely, in the absence of TGF-β signalling, this EndMT signature was dramatically reduced despite the ApoE$^{-/-}$ background and high fat diet[116].
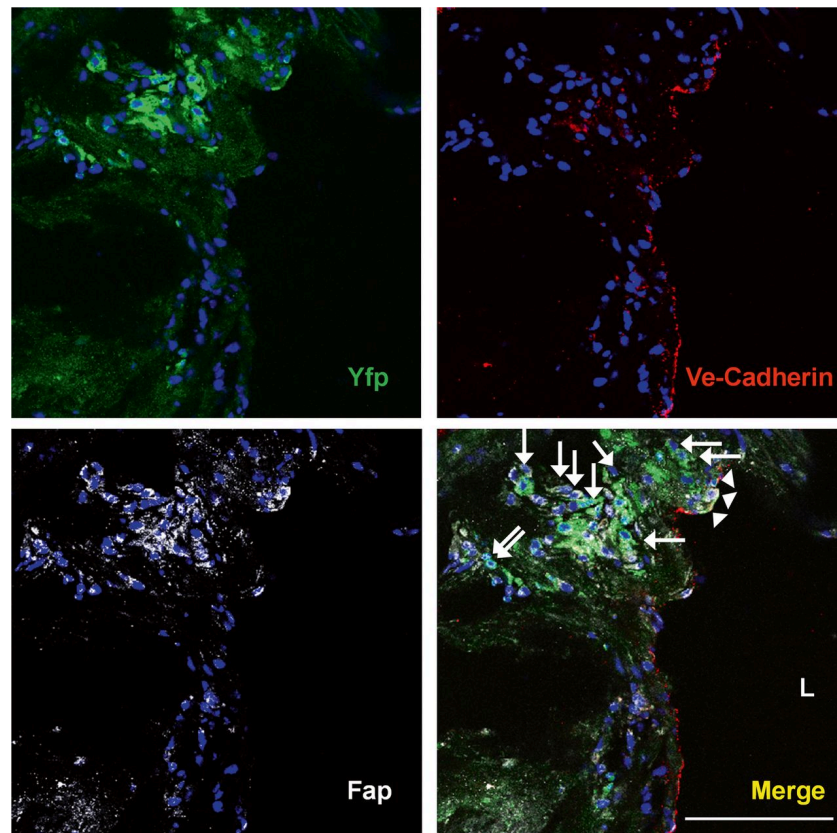
Figure 1.4: EndMT during atherosclerosis.

Immunofluorescence confocal microscopy of thoracic aortic sections from endothelial lineage tracing ApoE$^{-/-}$ mice after 18 weeks high fat diet, showing expression of endothelial lineage marker YFP, fibroblast-specific marker Fap and Ve-Cadherin. Figure adapted from Evrard *et al*, 2016 [53].

1.2.3.3 Pulmonary Arterial Hypertension

The vascular remodelling observed during the development of PAH broadly involves gross changes to the endothelial, smooth muscle and fibroblast population within the local vasculature [14]. Whilst the aetiology of PAH is fairly complex and multifactorial in nature, endothelial dysfunction and EndMT have started to gain attention as possible mediators during the development and progression of the disease [117]. Accumulating

data now suggests that the well described expansion and build-up of α-SMA and collagen-producing cells, contributing to the pathophysiological vascular remodelling observed in all forms of PAH, is not solely mediated by pulmonary artery SMCs and local fibroblasts. These cells may also result from pulmonary artery ECs which have undergone EndMT [118–120].

EndMT was first identified during PAH using an experimental disease model supported by endothelial-specific lineage-tracking in mice, reporting the presence of endothelial lineage cells expressing αSMA and SM-MHC within the neointima of remodelled pulmonary vessels [118]. Follow up studies by Ranchoux *et al* went on to show similar *in situ* evidence of EndoMT in human PAH [120]. Here the group identified cells co-expressing endothelial (PECAM1, VE-cadherin) and mesenchymal (α-SMA) markers in intimal and plexiform lesions from PAH lungs. Interestingly, this was also accompanied by increased endothelial expression of TWIST1 in PAH lungs, which was not present in control lungs [119,120]. Incidentally, overexpression TWIST1 has been found to effectively induce EndMT in human PAECs (HPAECs), while knockdown not only inhibits hypoxia-induced EndMT both *in vitro* and *in vivo* but can also attenuate hypoxia-induced PH [121].

Further, mutations in the BMPR2 gene are commonly found in patients with heritable PAH, these are associated with earlier diagnosis and more severe disease progression [122]. Notably, disrupted or impaired BMPR2 signalling has been associated with increased activation of canonical and noncanonical TGF-β pathways [123–125]. When tested *in vitro*, BMPR2 knockdown in HPAECs results in an increase in SNAI2, but not SNAI1, and higher αSMA expression. This effect is again replicated in mice lacking endothelial BMPR2 expression, leading to an EndMT-like profile in

pulmonary ECs and an enhanced PAH profile [126].

Lastly, the activation of hypoxia-inducible factors (i.e HIF1α and HIF2α) is known to play an important role in the pathogenesis of PAH [127,128]. Acting as transcription factors, increased intracellular levels of HIFs can results in the activation of transcriptional programs associated with vascular remodelling and inflammation [129]. In ECs, this pathway was recently reported to directly activate SNAI1/2 expression to induce EndMT independently of TGF-β signalling, while its knockdown was sufficient to inhibit transition during hypoxia [26,130,131]. Remarkably, endothelial-specific deletion of HIF-2α significantly attenuated or abolished the progression of hypoxia-induced PAH in mice [131].

Crucial to this thesis, we have continued to expand on the pathophysiological factors that define PH and current treatment options in Section 1.3.

1.3   Pulmonary Arterial Hypertension

1.3.1   Classification and Epidemiology

PH is a chronic condition characterised by the progressive obstructive remodelling of the pulmonary vasculature leading to a rise in pulmonary arterial pressure (PAP) and pulmonary vascular resistance (PVR). Cumulatively these structural and functional changes impose a sustained increase in pressure load on the right ventricle (RV) resulting in its hypertrophy and ultimately failure [132–134]. Despite being generally defined as an increase in mean pulmonary arterial pressure (mPAP) ≥25 mmHg at rest, it can arise from a variety of aetiologies and often presents with a wide spectrum of severities and symptoms. The most up-to-date classification strategy, based on the 5th World Symposium [132] and latest ESC/ERS Guidelines [133], categorises PH into five major groups: 1) PAH; 2) PH due to left heart disease; 3) PH due to interstitial lung diseases and/or hypoxia, including chronic obstructive pulmonary disease and sustained exposure to high altitude; 4) chronic thromboembolic PH; and 5) PH with unknown and/or multifactorial mechanisms, including hematologic, systemic and metabolic disorders (Table 1.1). These are largely grouped based on shared pathological mechanisms, clinical features, and therapeutic options, allowing for in-depth study and sub-categorisation of each group.

Given the complex heterogeneity of PH, and in order to avoid excessive disease characterisation, this thesis will largely focus on the more physiologically relevant PAH. PAH itself can be categorised into several subgroups such as idiopathic pulmonary arterial hypertension (IPAH), heritable PAH (HPAH) if there is a family history of PAH; or associated PAH (APAH) if other causes, such as HIV infection, portal hypertension or congenital heart disease, are involved (Table 1.1) [132–134].

1 Pulmonary arterial hypertension
1.1 Idiopathic PAH
1.2 Heritable PAH
1.2.1 BMPR2
1.2.2 Other mutations
1.3 Drug and toxin induced
1.4 Associated with:
1.4.1 Connective tissue disease
1.4.2 HIV infection
1.4.3 Portal hypertension
1.4.4 Congenital heart disease
1.4.5 Schistosomiasis
1' Pulmonary veno-occlusive disease and/or pulmonary capillary hemangiomatosis
1'.1 Idiopathic PAH
1'.2 Heritable PAH
1'.2.1 EIF2AK4
1'.2.2 Other mutations
1'.3 Drug, toxin and radiation induced
1'.4 Associated with:
1'.4.1 Connective tissue disease
1'.4.3 HIV infection
1''Persistent pulmonary hypertension of the newborn (PPHN)
2 Pulmonary hypertension due to left heart disease
2.1 Left ventricular systolic dysfunction
2.2 Left ventricular diastolic dysfunction
2.3 Valvular disease
2.4 Congenital/acquired left heart inflow/outflow tract obstruction and congenital Cardiomyopathy
2.5 Congenital/acquired pulmonary vein stenosis
3 Pulmonary hypertension due to lung diseases and /or hypoxia
3.1 Chronic obstructive pulmonary disease
3.2 Interstitial lung disease
3.3 Other pulmonary diseases with mixed restrictive and obstructive pattern
3.4 Sleep-disordered breathing
3.5 Alveolar hypoventilation disorders
3.6 Chronic exposure to high altitude
3.7 Developmental lung diseases
4 Chronic thromboembolic pulmonary hypertension
5 Pulmonary hypertension with unclear multifactorial mechanisms
5.1 Hematologic disorders: chronic hemolytic anaemia, myeloproliferative disorders, splenectomy
5.2 Systemic disorders: sarcoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis
5.3 Metabolic disorders: glycogen storage disease, Gaucher disease, thyroid disorders
5.4 Others: tumoral obstruction, fibrosing mediastinitis, chronic renal failure, segmental PH

Table 1.1: Comprehensive clinical classification of pulmonary hypertension.

Table adapted from the 2015 ESC/ERS Guidelines for the diagnosis and treatment of PH [133].

These subgroups, despite the different aetiologies, share several similar pathophysiological features and thus treatment strategies.

One of the larger European datasets on PAH comes from a 2006 French multicentre registry, initiated in 17 hospitals throughout the country and collecting data on 674 adult patients with the disease over a 3-year period [135]. The register estimates a population prevalence 15 cases per million adult inhabitants with an annual incidence of 2.4 cases per million. Interestingly, a large Scottish dataset collecting data on 374 PAH patients hospitalised between 1986 and 2001, suggests a much higher population prevalence of 52 cases per million population with an incidence of 7.1 cases per million [136]. Although these numbers have lowered considerably by 2005, with a prevalence of 26 cases per million population, these estimates still remain comparatively high to the French cohort [136]. A similar trend is seen in the United Kingdom and Ireland with a total of 482 patients were diagnosed in 2009, showing an annual incidence of 6.6 cases per million [137].

Further, based on current estimates, the vast majority of patients are diagnosed with IPAH. From the French 674 PAH patient register, 39.2% were diagnosed with IPAH, 3.9% with HPAH and the remaining 56.9% with a variety of other associated conditions [135]. Other large cohorts from the US show similar distributions with 46.5% cases diagnosed as IPAH, 2.9% as HPAH and 50.6% as APAH [138].

Lastly, while variable, the majority of published studies also suggest a significantly higher prevalence in females. The French cohort, for example, suggested a female predominance at a ratio of 1.6:1 for IPAH, 2.2:1 for HPAH and 14.9:1 for APAH [135].

1.3.2 Genetic and Molecular Factors

Heritable forms of PAH, although less common, are often of interest in the study of

the disease as they may provide additional clues to how it develops and progresses. Although previously mapped to a 3-cM region on chromosome 2q33 [139–141], termed primary pulmonary hypertension 1 (PPH1), a genetic linkage analysis of 35 families with a history of PAH published in 2000 was the first to report shared mutations within the BMPR2 gene locus [122]. Since then, not only over 300 mutations in the BMPR2 gene have been reported [142–145], but also within the activin receptor-like kinase-1 (ACVRL1) [146] and endoglin (ENG) genes [147].

BMPR2 mutations in particular have been found to be present in up to 75% of HPAH patients [148]. Compared with noncarriers, patients with BMPR2 mutations are younger at diagnosis and present with more severe hemodynamic compromise [149]. The BMPR2 gene itself encodes for a serine/threonine kinase receptor which is specifically recognised by BMPs belonging to the TGF-β superfamily. Activation by BMP ligands will lead to the recruitment of receptor-regulated SMAD proteins, typically SMAD 1/5/8, which in a complex with SMAD4, will enter the nucleus to promote the expression of transcription factors like ID1/2/3 to regulate cell function [150–152]. In the lung this signalling cascade is believed to play an important role in maintaining vascular homeostasis. With vascular ECs, for example, impaired BMPR2 signalling has now been linked to increased pro-inflammatory response [153], loss of barrier function [154] and apoptosis [155,156]. Recent reports also show an increase activation of canonical TGFβ-SMAD2/3 and lateral TGFβ-SMAD1/5 signalling, which may facilitate activation of processes such as EndMT [125]. Notably, BMPR2 expression has also been found to be significantly decreased not only in other forms of PAH but also other PH groups [145,157], as well as in a variety of experimental models of PH [158,159]. While less common, mutations to the ACVRL1 gene, also known as ALK1, have been

described in PAH [146,160,161]. Similar to BMPR2, ACVRL1 is a serine/threonine-protein kinase cell-surface receptor for the TGF- β superfamily. It interacts with four major ligands, TGF-β1 and TGF-β3, in a complex with TGF-βRII; and with BMP-9 and BMP-10, in a complex with the ActRIIA or BMPRII [56,59,60,110,162]. Activation of these pathways has been linked to variety of EC functions including developmental and pathological angiogenesis [59]. Unsurprisingly, ACVRL1 defects been also been suggested to have a role in the arterialisation and remodelling of arteries in other vascular disorders such as hereditary haemorrhagic telangiectasia (HHT), characterised by development of arteriovenous malformations [163]. Crucially, activation of ACVRL1 by TGF-β can also antagonise ALK5 activity and with it the activation of the SMAD2/3 cascade [56,57] linked to EndMT (described in Section 1.2).

It's worth nothing that given the significantly lower prevalence of heritable forms of PAH, it is clear that other factors are also involved in the initiation and progression of PAH. As referenced previously, PAH also has a higher prevalence in females. This, however, seems to happen irrespective of BMPR2 mutation status [123], again suggesting that additional factors may also be at play.

1.3.3   Vascular Remodelling in PAH

Despite its complex aetiology, likely to vary depending on the underlying genetic or pathogenic cause [164], PAH is largely associated with progressive structural and mechanical changes to the vascular wall, leading to sustained vasoconstriction and narrowing of both proximal and distal pulmonary arteries [4,165–167]. Described as pulmonary vascular remodelling, this process is generally characterised by alterations to the migratory and proliferative profile of the resident endothelial, smooth muscle

cell and fibroblast population, often accompanied by increased apoptotic resistance and glycolytic metabolism [168,169].
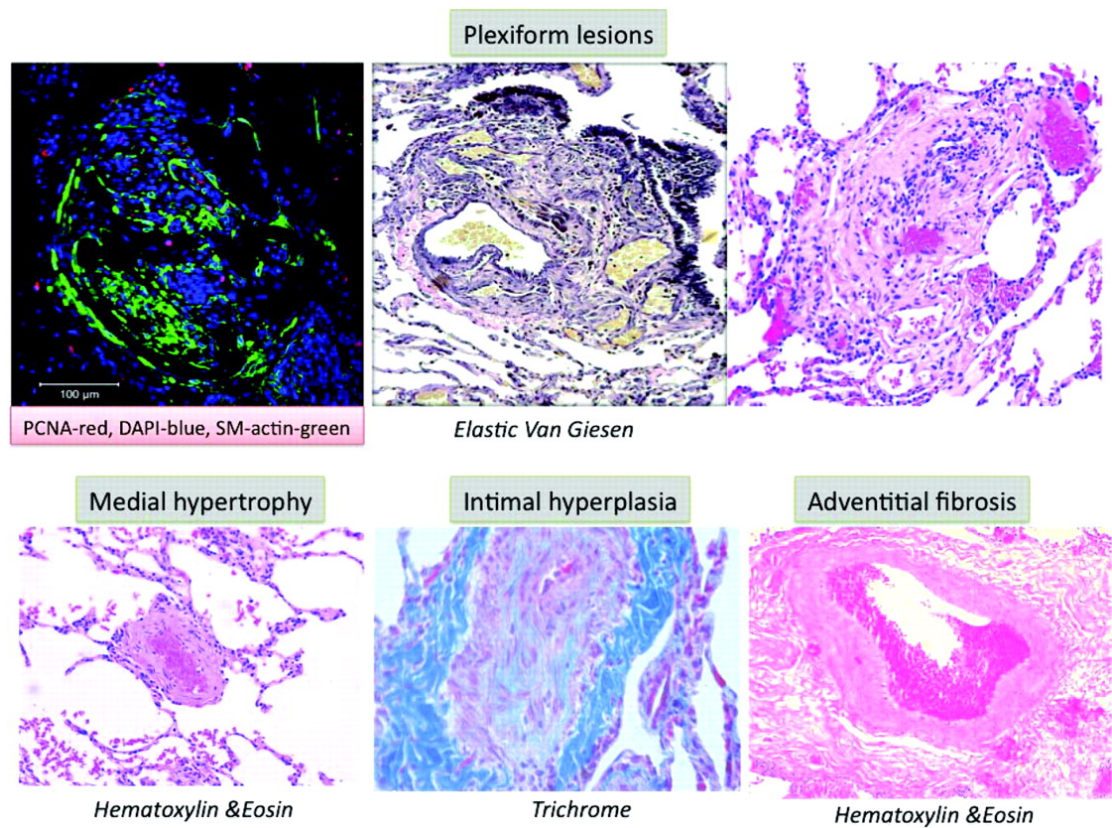
Figure 1.5: Histology of vascular remodelling in PAH.
Top panel: Plexiform lesions. Upper left, evidence of cell proliferation as shown by proliferating cell nuclear antigen (PCNA) in red and αSMA (SM) in green. Bottom panel: Medial hypertrophy, intimal hyperplasia and adventitial fibrosis. Figure adapted from Archer *et al*, 2011 [168].

Cumulatively, these changes result in intimal hyperplasia, medial thickening (hypertrophy) and adventitial fibrosis across muscular pulmonary vessels (500-70 μm in diameter) leading to eventual occlusion, along with the formation of plexiform and complex concentric lesions (Figure 1.5) [166–170]. With distal pre-capillary arterioles (70-20 μm in diameter), in addition to loss and obliteration, we also see increased vessel muscularisation with appearance of cells expressing mesenchymal markers [119,131,170–172]. Adding to the changes seen across the distal pulmonary arteries and arterioles,

pathological vascular remodelling can also affect the larger main, lobar and segmented pulmonary arteries of patients with PAH [167,173]. Increased wall thickness has been shown to occur in these vessels, leading to decreased vascular compliance and ultimately stiffening of the elastic proximal pulmonary arteries, which is strongly related to mortality in patients with PAH [173–177].

1.3.3.1 Intima Remodelling and Lesion Formation

Based on histological analysis of pulmonary artery cross sections, the intima will normally account for 10% of the vascular wall. However, this has been found to be significantly increased in PAH, irrespective of cause (Figure 1.6). As demonstrated by Statcher and colleagues, the PAH lung vasculature can present with up to 3-fold increase in intimal fractional thickness [165]. This thickened intima was associated with a 40-fold increase in pulmonary vascular resistance, making these structural changes of significance importance during disease progression [165].

During the early stages of PAH, triggered by circulating toxins [178], increased production of reactive oxygen species [179] and local autoimmune response [180], resident PAEC start to undergo apoptosis which allows for the emergence of a proliferative, apoptosis-resistant, SMC-like EC population [181–183]. In addition to contributing to intimal hyperplasia, this will also stimulate release of mediating factors, such as TGF-β1, which facilitate increased pulmonary arterial SMC (PASMC) proliferation [182,184]. Moreover, as they proliferate in increasingly disordered patterns, this EC population may also contribute to the formation of plexiform lesions, one of the hallmarks of advanced PAH (Figure 1.6 F-H) [181,185–187]. Nonetheless, we still have a limited understanding of how ECs contribute to the formation of these lesions. Interestingly, studies have shown the presence of a several angiogenetic markers throughout these

lesions, such as vascular endothelial growth factor (VEGF), VEGF receptor 2 (VEGFR2), and HIF-1α [187]. This was confirmed in a recent study using laser-assisted microdissection of plexiform lesions from PAH patient lung samples, showing that the lesions consisted of vascular channels lined by continuously proliferating ECs surrounded by tightly layered mesenchymal/myogenic cells [188]. The authors go on to show an up-regulation of HIF-1a and TGF-β1 along with variety of other tissue remodelling-associated genes, such as Angiopoietin-1 and Thrombospondin-1 [188]. This was accompanied by increased availability of sprouting-associated markers, such as NOTCH4 and several matrix metalloproteinases (MPP) [188].

Additional intima lesions, historically described as intimal fibrosis [189], are also frequent in PAH lungs [190]. Presenting in laminar or non-laminar formations, these are composed of fibroblasts, myofibroblasts and other mesenchymal-like cells, along with large deposits of connective tissue including ECM (Figure 1.6 B-E) [191]. Laminar intimal fibrosis, distinct due to their concentrically arranged fibrotic layers, occlude the lumen of small distal pulmonary arteries in patients suffering from IPAH and PAH associated with connective tissue diseases such as systemic sclerosis [191,192]. Interestingly, these are mostly paucicellular despite containing high amounts of ECM. However, it is important to note that the extent to which the pulmonary intima is compromised during PH can largely vary depending on its aetiology [166]. For example, intimal fractional thickness has been found to be greater in IPAH pulmonary samples compared to other forms of associated PAH [165]. This was found to be even higher in PAH patients with a known BMPR2 mutation, which may justify the earlier diagnosis and more pronounced disease progression seen in these patients [123,165].
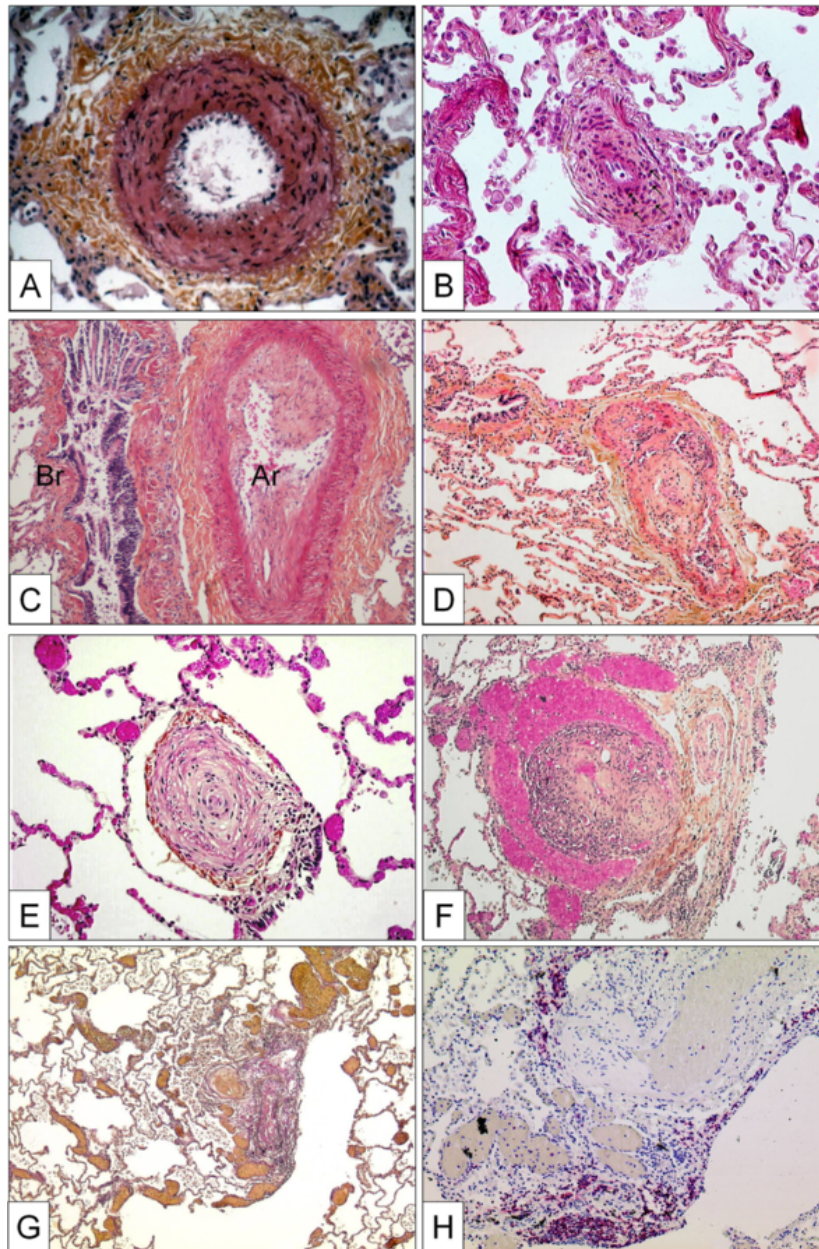
Figure 1.6: Pulmonary artery remodelling during PAH.

A) Medial hypertrophy showing SMC proliferation and adventitial fibrosis (Magnification x200, Weigert-hematoxylin-phloxine-saffron staining). **B)** Concentric non-laminar intimal fibrosis showing myofibroblast accumulation (arrows). **C)** Eccentric intimal fibrosis corresponding to organised thrombotic material (Br: bronchus, Ar: pulmonary artery) (Magnification x100, HES staining). **D)** Thrombotic lesion with partial recanalisation by microvessels (Magnification x100, HES). **E)** Concentric laminar intimal fibrosis (Magnification × 200, HES). **F)** Plexiform lesion with proliferation of small sinusoid-like vessels on a fibrotic matrix (Magnification x100, HES). **G)** Multiple dilation lesions emerging of the central plexiform lesion (Magnification × 40, Elastica-van-Gieson staining). **H)** The same plexiform lesion show in G) after immunohistochemical staining with anti-CD3, a T-lymphocytic marker (Magnification x100). Figure adapted from Montani *et al*, 2013 [191].

1.3.3.2 Media Remodelling and Smooth Muscle Proliferation

In addition to intimal changes, the media layer of the vascular wall will also undergo significant remodelling during most forms of mild to moderate PAH. Of note, while the normal media thickness will correspond to approximately 5% of the total diameter of pulmonary artery wall, during PAH this will increase up to 20% (Figure 1.6 A) [165,166]. This is widely believed to correspond to enhanced PASMC growth and proliferation within the media. During this process PASMC are reported to present with enhanced proliferative phenotypes along with increased resistance to known apoptosis inducers [193–196]. They have also been shown to acquire a synthetic rather than a contractile phenotype, increasing production of collagen and elastin [197], further supporting the role of a dysfunctional PASMC population which facilitates the remodelling pulmonary vasculature. Additionally, factors implicated in the pathogenesis of PAH, such as TGF-β1 [198], BMP-4 [199], serotonin [200] and endothelin-1 (ET-1) [201], can each induce PASMC hypertrophy, leading to increased cell size, contractile protein expression and fractional cell shortening [202]. Moreover, PASMC behaviour can be greatly influenced by dysfunctional vascular ECs. During PH these can decrease production of factors that suppress PASMC proliferation, such as apelin [203,204] and increase expression of growth factors that enhance proliferation, such as TGF- β1 [182,184] and fibroblast growth factor-2 (FGF-2) [205,206].

Nonetheless, despite the identified disease predisposing increase in PASMC proliferation, migration and decreased apoptosis, structural changes to the media during PAH will seldom happen in isolation. In fact, isolated medial hypertrophy has so far only been described as an early, and partially reversible, event seen in high-altitude PAH [207]. Interestingly, using the well-established Sugen 5146/Hypoxia

(SuHx) model of PAH (induced based on a combination of the VEGF receptor blocker SU5416 and chronic hypoxia), Taraseviciene-Stewart and colleagues demonstrated that PASMC proliferation in the pulmonary vessels largely happened during the early stages of the disease and waned at later, more advanced, stages [181,208]. Further, while previous studies have found the excessive growth of pulmonary vascular cells to be associated with predisposing genomic changes [185]. Using array-based comparative genomic hybridisation (aCGH), a recent publication has demonstrated that the genomes of PASMCs explanted from IPAH patient lungs were comparably stable with no net gain or loss of genetic material [195]. Indicating that the reported altered phenotype of PASMC during PAH may not be associated to specific chromosomal abnormalities and thus influenced by other factors.

Lastly, while it is clear that media remodelling and muscularisation of the distal pulmonary vascular occurs during PAH, the idea that this is solely due to the increased proliferation of a single resident cell population is now continuously being challenged as evidence supporting the role of active EndMT during the PAH grows.

1.3.4   Current PAH-specific Therapeutic Strategies

Given its complexity and pathological heterogeneity, the treatment of PH often requires a tailored multidisciplinary approach in order to improve patient outcomes [209]. This is particularly true regarding PAH which, despite advancements in treatment options, retains unreasonably high mortality rates. Current PAH-specific therapeutic approaches mostly target three major signalling pathways that regulate vascular tone and are often altered during PAH. Namely, the prostacyclin (PGI2), ET-1 and nitric oxide (NO) signalling pathways [191,209]. Approved PAH treatments targeting the PGI2 pathway include a variety of PGI2 synthetic analogues and derivates, along with PGI2

receptor (IPr) agonists [133]. PGI2 is a major active metabolite of arachidonic acid (AA) produced by ECs, acting as an agonist of adenylate cyclase [210]. Generally, PGI2 synthetic analogues and IPr agonists, like endogenous PGI2, will activate IPr which in turn signals adenylyl cyclase to produce cyclic AMP (cAMP) and thus direct vasodilation and inhibition of platelet aggregation in PAH [209].

Highly expressed in the lung, ET-1 is a potent vasoconstrictor released by the ET-1 converting enzyme (ECE) found on the membrane of ECs. Activation of the ET-1 receptors, ET-A and ET-, on PASMCs leads to increased formation of inositol triphosphate (IP3) [211]. Cytoplasmic IP3 will in turn stimulate the release of calcium by the sarcoplasmic reticulum leading to smooth muscle contraction, making it an attractive target for the management of PAH [211]. Commonly used ET-1 pathway targeting strategies include endothelin receptor antagonists (ERAs) such as bosentan, which selectively antagonise the binding of ET-1 to ET-A and/or ET-B receptors, blocking their activity and thus decreasing pulmonary and systemic vascular resistance [209,212,213].

The use of phosphodiesterase type 5 (PDE-5) inhibitors and sGC stimulators is also recommended to directly target and modulate the dysregulated NO pathway in PAH [133,214]. As with most vascular beds, NO is essential in the regulation of pulmonary vascular tone. NO acts as a SMC relaxant by stimulating soluble guanylate cyclase (sGC) and the subsequent conversion of guanosine triphosphate (GTP) to cyclic guanosine monophosphate (cGMP). The effects of NO are normally terminated by the breakdown cGMP by phosphodiesterase (PDE) enzymes like PDE-5, which are abundantly expressed in the PASCMs [214–216]. PDE5 inhibitors, such as sildenafil, are effective pharmacological agents for the treatment of PAH due to either ability to bind

to the PDE5 enzyme and inhibit the breakdown of cGMP, allowing for longer NO availability in PASMCs [217,218]. Alternatively, sGC stimulators, such as Riociguat, can also be used [133,214]. Instead of slowing cGMP degradation, these instead mimic the effects of endogenous NO and thus increase the production of cGMP.

Current PAH-specific treatment strategies, recommended by the 5th World Symposium [132] and latest ESC/ERS Guidelines [133], suggested targeting at least one of the three pathways mentioned above. Patients at low or intermediate risk can be initially treated with monotherapy, while high risk patients require tailored combination therapy using drugs with different action mechanisms. Despite some of the conflicting results seen in earlier short-term randomised controlled trials, recent evidence on the long-term outcomes indicates that combined PAH treatment strategies can help improve patient quality of life [219].

Further, in patients that remain unresponsive to maximum combined therapy, bilateral lung transplantation continues to be an important treatment option recommended by all up-to-date guidelines [133,220]. While successful transplantation often leads to considerable improvements in patient quality of life, survival rates remain limited. According to the International Society for Heart and Lung Transplantation 5-year survival transplantation survival rates were of approximately 45 to 50% in 2006 [221]. Notably, recent reports now place 5-year survival rates at 52-75% [222,223] and 45-66% at 10 years [224].

Ultimately, while the development of PAH-specific therapies has greatly improved patient symptoms, these remain largely palliative, mostly slowing down the progression rate of the disease with limited long-term patient survival. Data collected from a 2003 French national register comprised of 354 adults patients with idiopathic,

heritable, or anorexigen-associated PAH, estimated one, two and three year survival rates at 82.9%, 67.1% and 58.2%, respectively [225]. Recent data from the UK and Ireland accounting for 479 patients in the national pulmonary hypertension registry, reports one, two, three and five year survival rates of 92.7%, 84%, 73.3%, and 61.1%, respectively [137]. Crucially, patients with associated co-morbidities tend to also have worse outcomes. In a recent US population study, patients with systemic sclerosis (SSc) and incident group I PAH had cumulative survival rates of 93%, 88%, and 75% at one, two and three years after diagnosis, respectively [226]. Pregnancy is considered one of the major contradictions in PAH, associated with 30-50% increased mortality rates in patients, thus current guidelines continue to recommend the use of effective contraception methods for women of childbearing age with PAH [227].

In conclusion, it is clear that our understanding of the mechanism underlying PAH remains far from complete and we must continue to explore novel therapeutic avenues.

1.4 Non-Coding RNA Biology and Function

With only an estimated 1.2% of the human genome coding for proteins [228], focus is positioned to assess the possible functional roles for the other 98.8% of the genome with little to no protein-coding capacity [229]. While their function is often debated [230], ncRNA transcripts compose approximately 70 to 80% of our genome and include thousands of operationally significant RNAs implied in all manner of biological processes [231–233]. These include, not only the thoroughly described transfer RNAs (tRNA) and ribosomal RNAs (rRNA), but also small nucleolar RNAs (snoRNA) and PIWI-interacting RNAs (piRNAs), among others. From these novel ncRNA categories, the miRNA and lncRNA gene families have sparked great interest within the research community as they have been found to be critical in development and dysregulated in disease.

The development of high-throughput RNA sequencing (RNAseq) approaches now adds extraordinary range and depth to ncRNA gene profiling; making it easier to distinguish miRNAs from other similar transcriptional products and to identify any post-transcriptional effects they may have on specific genes. The same applies to the highly abundant, yet poorly characterised, lncRNAs.

1.4.1 MicroRNA

1.4.1.1 MicroRNA Biogenesis and Function

Since their discovery in 1993, miRNAs have emerged as regulators for several cellular processes, from apoptosis and proliferation, to stress response reactions [234]. Ubiquitously expressed in all human cells, miRNAs regulate mRNA translation by binding to their complementary base-pair sequences on the 3'UTR of mRNA

transcripts and ultimately suppressing protein synthesis [235,236] (Figure 1.7). MiRNA biogenesis is a complex multi-step process and, in a rapidly evolving field, our understanding of it continues to develop. The basis of this process, however, has now been thoroughly described.

MirRNA host genes are first transcribed in the nucleus by RNA polymerases to produce ~1 kb long primary miRNAS (pri-miRNA), normally arranged in a stem-loop structure with two long single-stranded flanking regions [237,238]. This transcription process is commonly performed by the RNA polymerase II (Pol II) multiprotein complex, however in specific cases where Alu elements are present within the promoter region this can also be carried out by RNA polymerase III (Pol III) [239,240]. Interestingly, multiple miRNA clustered in the same region can often be transcribed within a single pri-miRNA and processed separately at later stages [241].

Once transcribed, the pri-miRNA structure is further processed into a 70 to 120 nucleotide-long premature miRNA (pre-miRNA) hairpin stem loop by a nuclear microprocessor complex composed of the RNase type III enzyme Drosha, and the RNA binding protein DGCR8, also known as Pasha [241,242].

Next, the pre-miRNA hairpin is moved into the cytosol for further maturation. Transportation into the cytoplasm is accomplished through Exportin-5 (XPO5) in complex with the GTP binding nuclear cofactor Ran-GTP. A defined length of the stem-loop overhangs is crucial for successful binding by XPO5, ensuring that only correctly processed pre-miRNAs are translocated [241]. In the cytoplasm, the pre-miRNA is further processed by the RNase III enzyme Dicer, which cleaves the loop regions of the pre-miRNA to form a miRNA duplex [243]. This is achieved in complex with the transactivation response RNA binding protein (TRBP) and protein activator

of PKR (PACT) and Argonaute 2 (Ago2), which form the RNA-induced silencing complex (RISC) [244–247]. Crucially, after incorporation only one miRNA strand remains active (guide strand) within the RISC complex while the second strand (passenger strand) is often degraded. With regards to strand selection, while both the 5p end (from the 5' arm of the pre-miRNA duplex) and the complementary 3p end (from the 3' arm) have the potential to function as a guide strand, preference seems to be given to strands with a less stably paired 5' ends possessing either A or U as their terminal nucleotides [248–251].

Finally, the mature miRNA guides the RISC complex to complementary sites on messenger RNAs (mRNA), leading to epigenetic regulation via mRNA degradation or translation inhibition. Target recognition by the RISC complex will happen via the binding of the miRNA seed sequence to its complementary target normally found within the 3' UTR of mRNAs [241]. This canonical seed/target-site complementarity forms the basis for most miRNA target prediction algorithms available today [252–254]. Generally, the miRNA-bound RISC complex can silence gene expression via target mRNA degradation (through mRNA cleavage or deadenylation) or translation repression. While mRNA cleavage by Ago2 is possible, this is rare in mammals requiring seed sequences perfectly complementary to that of the target mRNA[255]. MiRNA-mediated degradation by deadenylation is present in 66-90% of events, accounting for the vast majority of mRNA silencing [256]. Via this pathway the target mRNA is first deadenylated by the Ccr4–Not and Pan2–Pan3 deadenylation complexes [257]. Once deadenylated, the mRNA is decapped by the aptly named decapping protein 2 (DCP2) and ultimately degraded through the action of the cytoplasmic nuclease exoribonuclease 1 (XRN1) [258,259].

Less common, translation repression is also another form of miRNA-mediated gene silencing seen in mammalian cells. While this process is now believed to happen through the inhibition of cap-dependent translation via RNA helicases such as eIF4A, our understanding of the precise molecular mechanisms that define this remain unclear [260–263].
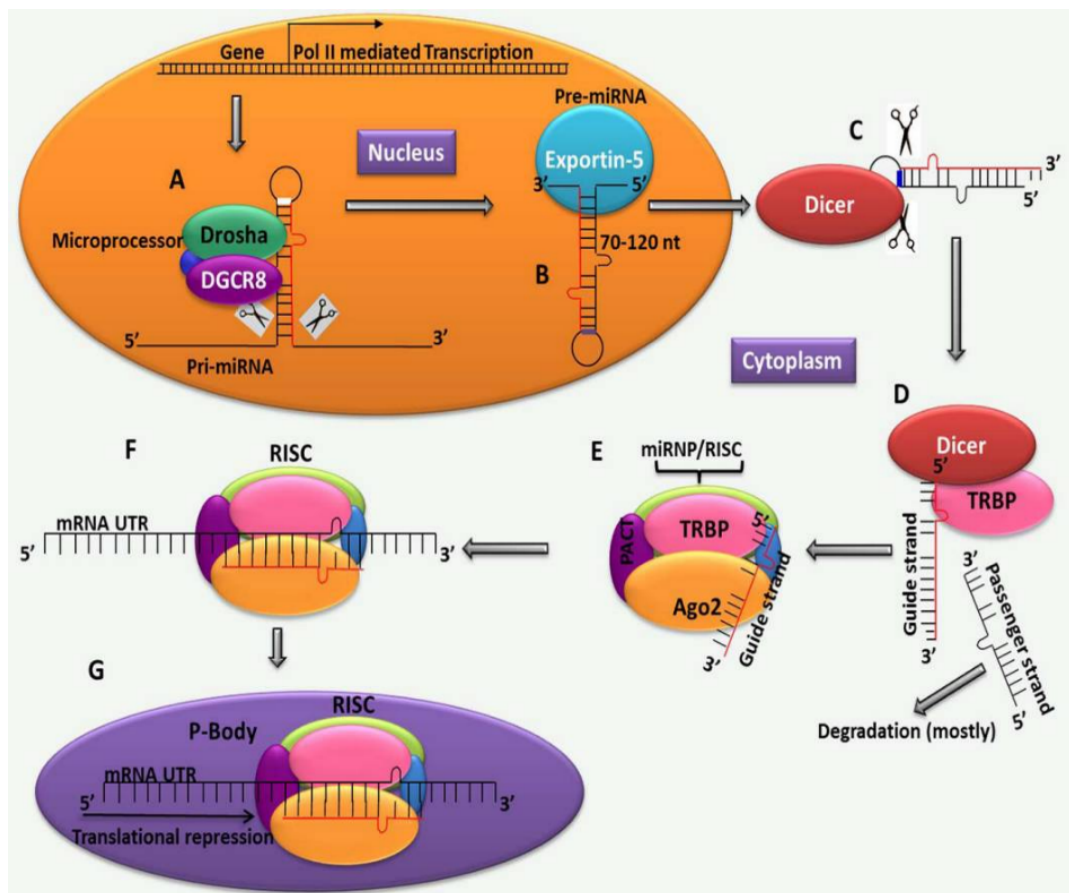
Figure 1.7: Standard microRNA biogenesis and function.
MiRNA genes are first transcribed in the nucleus by RNA polymerases as pri-miRNA, and further processed into a long oligonucleotide stem loop pre-miRNA by a protein complex composed of drosha, and DGCR8. In an alternate pathway, miRNA genes are transcribed directly as short hairpin pre-miRNA. The pre-miRNA is then pumped into the cytosol by Exportin-5, where it is further processed by Dicer, which cleaves the loop regions of the pre-miRNA to form a miRNA duplex. Facilitated by a TRBP, the two strand microRNA duplex is incorporated into the RISC along with Ago2 and other RNA binding proteins. After incorporation, only one miRNA strand remains active (guide strand) within the RISC complex while the second strand (passenger strand) is degraded. The miRNA guides the RISC complex to complementary sites on mRNAs, leading to epigenetic regulation via mRNA degradation or translation inhibition. Figure adapted from Bhaskaran & Mohan, 2014 [264].

1.4.1.2 Role of MicroRNAs in Vascular Remodelling and PH

Due to their prevalence, miRNAs have been put forward as possible biomarkers and treatment targets for many pathologies, from cancer to renal fibrosis and CVD [265]. Numerous miRNA signatures have now been found to be associated with various types of cardiovascular disease, including atherosclerosis [6], myocardial infarction [266], heart failure [267] and hypertension [268]. This is of particular significance in the endothelium, with multiple miRNAs starting to be identified in ECs, regulating their behaviour and significantly affecting the process of vascular remodelling. For example, within the vascular endothelium, the targeted disruption of the Dicer gene and consequent loss of miRNA maturation, leads to significant abnormalities in endothelial biology, directly impacting on EC survival and migration [269,270].

Harris and colleagues [271] demonstrated that miRNAs may also impact the development vascular inflammation. The endothelial-specific miR-126, abundantly found in resting adult ECs, was shown to suppress vascular inflammation by inhibiting the expression of the vascular cell adhesion molecule 1 (VCAM-1). Expression of VCAM-1 itself is increased by pro-inflammatory cytokines and allows for leukocyte adhesion to the cell surface as an inflammatory response such as that observed during early phase atherosclerotic disease [271,272].

Developmental studies in ECs have also shown miR-221 and 222 to control the activity of stem cell factor (SCF) by targeting its receptor c-Kit [273]. Often a marker for progenitor cells, stem cell factor/c-kit signalling has been repeatedly associated with increased cell survival, migration and capillary tube formation [274]. Recently, both miR-221/222 and miR-155 were found to negatively correlate to the expression of ETS1

genes, involved in the control of various endothelial-mediated inflammatory molecules during vascular inflammation and remodelling [275,276].

Within the wider spectrum of disease, accumulating evidence suggests that miRNAs can significantly modulate many of the pathological vascular remodelling processes which lead to PH. The miR-143/145 axis, for instance, appears to be dysregulated in mouse models of PH being over expressed in remodelled vessels [277,278]. Furthermore, previously published work by our group as described a novel miR-143-3p-mediated cell-to-cell communication pathway between pulmonary vascular cells which contributes to altered cell migration in PH. Inhibition of this pathway, by miR-143-3p knockdown, effectively blocked experimental PH in mice exposed to chronic hypoxia [279].

Outside of PH, the key regulatory roles of miRNAs continue to emerge in other vascular remodelling scenarios such as in-stent restenosis [280] and vein graft remodelling [281].

1.4.1.3 Role in EndMT

Given the role of EndMT during vascular remodelling, the past decade as also seen increasing numbers of publications citing miRNA-based regulatory pathways. Published data from microarray analysis has mostly revealed tenuous links between specific miRNAs and EndMT during different cardiovascular disease states with no specific mode of action [282]. Nevertheless, a recent report has shown that the constitutively active miR-31 was required for the modulation of mesenchymal markers such as α-SMA, actin reorganisation and myocardin-related transcription factor A (MRTF-A) activation. A putative mechanism for this phenomenon includes the guanine nucleotide exchange factor VAV3, associated with actin remodelling and

MRTF-A activity, which was identified as a possible target for miR-31 [283]. Tumour-associated EC studies further suggest that the cell-specific miR-302c, when over expressed in HUVEC, reduces cell motility and alters the expression levels of EndMT markers – namely, VE-cadherin is up-regulated, whereas β-catenin, FSP1, and α-SMA are down-regulated [284]. Reporter assays also revealed that miR-302c may directly inhibit metadherin (MTDH) expression by binding to the 3′UTR of its mRNA and suppressing its translation. In fact, silencing of MTDH expression in HUVECs also leads to lower levels of β-catenin, FSP1 and α-SMA [284].

1.4.2   Long Non-Coding RNAs

1.4.2.1 LncRNA Biogenesis, Classification and Function

Although the regulatory role small ncRNAs is now well established, the concept of widespread control of cell function by lncRNAs has only been advocated within the last decade [285]. Originally believed to be non-functional transcriptional by-products, lncRNAs have started to spark great interest with an increasing number of studies describing them as critical in development and dysregulated in disease [286].

Generally defined as transcripts longer than 200 bp in length without any protein coding potential, lncRNAs are significantly less abundant than their mRNA counterparts and have a higher rate of evolutionary turnover [287]. Recent publications now estimate that nearly 27,000 lncRNA transcripts are produced in humans with over 1,000 of those conserved in mammals [287,288].

Biogenesis

Unlike miRNA biogenesis, which has been extensively studied, our understanding of the lncRNA transcriptional process is still somewhat limited. Nonetheless, several

studies commonly draw parallels between lncRNA biogenesis and that of protein-coding RNAs. Indeed, much like mRNAs, most known lncRNAs have distinct multi-exonic structures, which are subject to alternative splicing, polyadenylation and 5' capping [289–292]. For instance, in a ground-breaking genome-wide chromatin-state analysis, Guttman and colleagues suggest that transcribed lncRNAs can present with epigenetic marks at their promoter regions similar to that of mRNAs, including increased trimethylation of lysine 4 of histone 3 (H3K4me3) [293]. LncRNA transcription is also reported to be largely carried out by the action of the RNA polymerase (RNAP) II complex [294–296]. This was demonstrated in a recent experiment where HeLa cells were exposed to the RNAP II inhibitor, α-amanitin, and their transcriptional profile analysed using a custom oligoarray interrogating thousands of lncRNA and protein-coding transcripts [297]. In a similar fashion to mRNA transcripts, lncRNAs showed significantly supressed expression levels following α-amanitin treatment [297]. Interestingly, RNAP III-dependent mechanisms may also be possible as similar experiments report continued expression of some lncRNAs despite α-amanitin treatment [294].

With regards to polyadenylation, while several transcripts retain polyadenylation signals and are subsequently processed by canonical poly(A) polymerase activity, non-polyadenylated lncRNAs have also been described [298,299]. Notably, it is possible that the number of non-polyadenylated lncRNAs has been historically underestimated when analysing RNAseq results due to the common use of oligo(dT) primers for cDNA synthesis, which select for poly(A)-containing transcripts.

Lastly, a crucial feature that discriminates lncRNA from mRNA transcripts is the lack of sequence similarities between lncRNAs and low conservation across different

species, making their transcriptional patterns hard to predict and classify [300]. Nonetheless, lncRNAs such as Xist (X-inactive specific transcript), despite having poor linear sequence conservation, share key common functional features with homologues present in other species [301–304]. Conservation analysis studies also show that the promoter regions of lncRNAs are generally conserved at levels similar to that of protein-coding genes, suggesting a retention of the regulatory machinery rather than their transcribed sequence [300,305].

Classification

While collectively lncRNAs may evade an exact definition other than their transcriptional length, they can be classified into a variety of subclasses based on the attributes originally used to detect and study them. Among others, this includes their genomic association to neighbouring protein coding genes, the chromatin signatures of their transcriptional start sites (promoter, enhancer), their mode of regulation (transcriptional vs post-transcriptional), cellular localisation and functional mechanism (Figure 1.8) [306].

Serving as the foundation for the GENCODE categorisation of lncRNAs, their genomic location with respect to neighbouring protein-coding genes allows for the use of five distinct categories: sense, antisense, bidirectional, intronic or intergenic [289,306,307]. Sense and antisense lncRNAs, for example, are RNA molecules partially overlapping one or more exons of a protein-coding gene and are transcribed from either the same (sense) or opposite (antisense) strand. Sense-overlapping lncRNAs can be broadly thought of as non-coding transcript variants of protein-coding genes, as they overlap with same genomic strand while lacking the needed open reading frames (ORF) for protein translation [306]. This category includes un-spliced partially intronic

RNAs, known as PINs, and mRNA-like spliced transcripts [306]. Antisense-overlapping lncRNAs have on average 10-fold lower expression levels and are often less likely to be spliced [308,309]. Further, bidirectional lncRNAs are RNA molecules whose transcription can be initiated in a divergent fashion from a promoter of a protein coding gene. While there is no exact cut-off distance, bidirectionality is usually defined if the two transcription start sites are within a few hundred bases pairs and 1 kb distance from each other [310,311]. Interestingly, while bidirectional lncRNAs seem to mostly share similar expression patterns with their protein coding neighbouring locus [311], inverse expression patters are also possible [312,313].

In contrast to the overlapping-sense and antisense categories, intronic lncRNAs are located exclusively within the intronic region of protein-coding genes without intersecting any exons [294]. Found in a sense or antisense orientation, these can either be independently transcribed or display transcriptional patterns similar to that of the overlapping protein-coding transcript, indicating a shared transcriptional regulation [294]. Finally, intergenic lncRNAs, also known as long intergenic non-coding RNAs (lincRNAs), are located between the loci of two protein-coding genes but have separate transcriptional units. Most definitions of this category require the lincRNA to be 5-1 kb away from the nearest protein-coding genes [293].

Figure 1.8: LncRNA Classification and Function.

Several LncRNA categories exist based on their transcription start site, surrounding coding genes and regulatory function. Based on their transcriptional start site, these can be either promoter-driven or enhancer-derived. Relative to the surrounding coding genes and transcriptional direction, lncRNAs can be categorised as sense-overlapping, intronic, antisense, bidirectional or intergenic. LncRNAs whose function has been characterised tend to fall into distinct groups based on their molecular interactions. By interacting with a variety of RNA binding proteins, lncRNAs can act to control activity or localisation of a specific protein or play a structural role within a larger protein complex. LncRNAs can also act as transcriptional regulators by interacting with transcription factors, or by recruiting chromatin modifying complexes to a specific gene locus. In addition to acting as miRNA host genes, lncRNA can also act as miRNA sponges, titrating away specific miRNAs for degradation. At the post-transcriptional level, lncRNAs can also participate in mRNA translation and degradation. Figure adapted from Monteiro *et al*, 2019 [314].

Cellular localisation

LncRNAs also show a variety of subcellular distributions allowing for an additional subset of lncRNA categories to be described. LncRNAs are found to accumulate predominantly in either the cytoplasmic or nuclear cellular fractions, but can also be distributed between both [315]. For example, in ECs, different subpopulations can express high levels of the nuclear enriched lncRNAs such as TUG1, MEG3, and MALAT1 [316], while the lncRNA SENCR shows both cytoplasmic and nuclear accumulation [317].

Additionally, as demonstrated by the 2012 ENCODE consortium, lncRNAs appear to localise to distinct nuclear subdomains, adding a further layer of complexity to lncRNA categorisation [232]. The extensively described nuclear lncRNAs NEAT1 and MALAT1, despite the proximity of their respective genomic locus (approximately 53 kb apart within Chromosome 11), appear to localise in distinct nuclear compartments. MALAT1, for example, can be found within splicing factor-enriched subnuclear compartments known as nuclear speckles, where it is reported to regulate alternative splicing of pre-mRNAs [318–320]. On the other hand, NEAT1 has been shown to be involved in nucleating and maintaining the paraspeckle nuclear domain [321–324]. Additionally, the lncRNA Xist is found distributed along the target X-chromosome, where it silences gene expression by triggering repressive chromatin modifications [325–328].

While many of these lncRNAs may be catalytically inactive on their own, by binding and coordinating specific regulatory protein complexes they aid in shaping nuclear compartmentalisation [329]. Unsurprisingly, this subcellular localisation is often indicative of lncRNA function within the cell, which we will describe in more detail.

Function

In comparison to the tens of thousands of annotated lncRNAs genes, relatively few have been functionally characterised. Nonetheless, those who have demonstrate a remarkable mechanistic diversity with epigenetic, transcriptional and post-transcriptional effects, able to activate or suppress gene expression and translation (Figure 1.8) [292,330].

At the transcriptional level, several lncRNAs have been shown to bind chromatin-modifying complexes either activating or repressing gene expression [331,332]. Over the past decade many publications have started to elucidate the mechanism of several such lncRNAs, including Xist [333] and HOTAIR (HOX transcript antisense RNA) [334]. Xist, for example, is now established as one of the master regulators of X-chromosome inactivation (XCI) in female eutherian mammals [335]. Upon expression, the lncRNA is retained within its transcriptional region where it will start to coat the X-chromosome with the aid of RNA and DNA-binding proteins such as SAF-A (scaffold attachment factor-A, also known as HNRNPU) [336–338]. The lncRNA can then recruit Polycomb group complexes (PcG), such as PRC2, catalysing H3K27 histone trimethylation leading to chromatin modification and transcription inactivation [335,339,340]. It is important to mention, however, we still do not have a complete understanding of how Xist carries out its function and, in addition to the interaction described here, several other accompanying XCI models have been proposed [341]. Similarly, in a seminal paper by Rinn *et al*, HOTAIR was also shown to regulate PcG-mediated transcription inactivation of the HoxD gene cluster [334]. In a series of *in vivo* experiments, Li and colleagues further demonstrated that HOTAIR binds to PRC2 and Lsd1, catalysing H3K27 methylation and H3K4 demethylation, respectively [342].

In addition, lncRNAs can also regulate gene expression by interacting with specific DNA sequences and forming RNA-DNA hybrid duplex or RNA-DNA triplex structures with key regulatory regions [343–345]. A recent study by Mondal and colleagues demonstrated that the chromatin-interacting lncRNA MEG3 (Maternally expressed gene 3) will bind GA-rich target sequences which allow for the formation of RNA–DNA triplex structures. The group showed that this allowed for the recruitment of chromatin modifiers to promoter regions of TGF-β pathway genes, including TGFBR1 [346]. The same study went on to suggest that these RNA-DNA triplex structures are widespread *in vivo* and may be a common mechanism for target-site recognition and gene regulation by lncRNAs [346].

In some cases, lncRNA transcription itself can have an *in cis* regulatory function that affects the expression of neighbouring genes. This was recently demonstrated in a large-scale study by Engreitz and colleagues looking the effects of manipulating the genetic locus of several lncRNAs on the expression of nearby genes. Using variety of CRISPR/Cas9-mediated genome editing strategies, the group demonstrated that 5 out of 12 lncRNA loci studied had an impacted gene expression in a sequence-independent manner [347]. For example, changes to the length of the transcribed region of the lncRNA Blustr (Bivalent Locus Sfmbt2) affected the total amount of Sfmbt2 activation (located 5 kb upstream of the Blustr locus), but changes to the lncRNA sequence did not [347].

Another proposed mechanism involves the reannealing of the nascent RNA to the DNA template, which gives rise to RNA-DNA hybrid structures known as R-loops [348]. For example, as it is transcribed the antisense lncRNA VIM-AS1 forms an R-loop structure around the promoter region of the gene coding for vimentin (VIM). In epithelial cell lines, this lead to chromatin structural remodelling which in turn

facilitated the binding of transcriptional activators of the NF-κB pathway [349].

Additionally, they can also act as decoys for DNA-binding proteins and prevent their association to a target gene [350,351]. The lncRNA H19 was previously shown to directly inhibits p53 activation, resulting in altered gene expression profiles and with that promoting gastric cancer progression [352]. Further, the growth arrest-specific 5 (Gas5) lncRNA regulates cancer cell growth by directly interacting with the DNA-binding domain of the glucocorticoid receptor (GR) acting as a decoy glucocorticoid response element (GRE) and thus suppressing its transcriptional activity [353–355].

At the post-transcriptional level, a growing number of studies have also implicated lncRNAs at various stages of control by regulating mRNA stability [356] or enhancing mRNA translation[357]. For instance, the lncRNA TINCR (Terminal differentiation-induced ncRNA) has been shown to directly interact with the protein Staufen 1 (STAU1) to mediate mRNA stabilisation. This was seen in human cancer cell lines, where the TINCR-STAU1 complex is able to impair Krüppel-like factor 2 (KLF2) mRNA stability, thus supressing its growth-inhibitory and pro-apoptotic functions [358]. Interestingly, previous studies have suggested that a 25-nucleotide motif, aptly named *TINCR box*, is strongly enriched in mRNAs targeted by TINCR [359].

Historically, some lncRNA have also been described as host genes for miRNA found within their locus. One of the first published examples of this was H19, whose expression was strongly associated with that of the H19-derived miR-675 [360]. Conversely, lncRNAs can additionally show the ability to function as decoy molecules for mature miRNAs [361,362]. Several publications have started to describe biologically significant cross-regulatory interactions between noncoding RNAs classes, adding a further functional role for lncRNAs which we will discuss further in Section 1.4.3.

1.4.2.2 Role of LncRNAs in Vascular Remodelling and PH

While examples of miRNA-mediated vascular remodelling are vast, they are not the only ncRNAs implicated throughout the pathological process. It is now becoming clear that abnormal lncRNA levels are linked to aberrant cell migration, proliferation, and function, all defining features of active vascular remodelling. As one of the foremost cellular basis for the vascular remodelling process and the focus of early lncRNA research, the past two decades have seen several lncRNA emerge as regulators of SMC function. One of the first examples of this is the lncRNA ANRIL (Antisense noncoding RNA in the INK4 locus), identified by disease genomewide association studies (GWAS) within one of the best known genetic susceptibility locus for coronary artery disease (CAD), atherosclerosis and type 2 diabetes [363–366]. ANRIL has since been shown to regulate VSMC growth and proliferation through neighbouring genes CDKN2A/B [367,368]. More recently, the lncRNA GAS5 (Growth arrest specific 5) has also emerged as a negative regulator for VSMC survival in vascular remodelling, able to suppress cell proliferation and neointima formation in a vascular injury model [369].

Within ECs, the first regulatory lncRNA to be studied was tie-1AS, transcribed antisense to the tie-1 gene (tyrosine kinase containing immunoglobulin and epidermal growth factor homology domain-1) in zebrafish, mouse, and humans. Expressed during embryonic development, the lncRNA was shown to regulate tie-1 levels by selectively binding and degrading the mRNA, resulting in specific defects in EC contact junctions [370]. Contemporary publications also identified the ubiquitously expressed nuclear lncRNA MALAT1, whose knockdown in ECs was able alter cell proliferation and induce a pro-migratory response [316,371]. Crucially, *in vivo* both genetic

deletion and pharmacological inhibition of MALAT1 effectively inhibited proliferation of ECs leading to reduced neonatal retina vascularisation [316]. Building on these earlier studies, a growing number of lncRNAs have since been linked to all manners of endothelial function, including not only EC proliferation and migration (e.g. MALAT1, H19), but also angiogenesis (e.g. MEG3, MANTIS), inflammatory response (e.g. UMLILO, LISPR1) and, more recently, EndMT (e.g. GATA6-AS). We will continue to explore these EC-specific regulatory functions in Section 1.5.

Lastly, in the setting of pulmonary hypertension, lncRNA research is still limited. Despite the emerging cell-specific, regulatory function suggested, these have yet to be translated into the clinical setting. However, preliminary studies have already identified over 300 lncRNAs differentially expressed among rat models of pulmonary hypertension [372]. Additionally, lncRNA profiling of endothelial tissues from patients with chronic thromboembolic pulmonary hypertension also revealed the differential expression of 185 lncRNAs compared to healthy control tissues [373].

1.4.3   Non-Coding RNA Interaction and Cross-Talk

As mentioned, lncRNAs can act as endogenous miRNA sponges by binding and sequestering miRNAs away from a particular target mRNA and allowing for its translation [273,343]. Take the muscle-specific lncRNA linc-MD1, it carries a miRNA-binding site in its 3′-UTR able to 'sponge' miR-133 away and abolish its suppressive effect on transcription factors that activate muscle-specific gene expression [374].

By competing for the same miRNA-binding sites, antisense lncRNAs can further reduce mRNA destabilisation. More specifically, as shown by Faghihi *et al*, the BACE1as lncRNA can compete with miR-485-5p for the same mRNA binding site

and increase translation of beta-secretase-1 (BACE1), an enzyme involved in Alzheimer's disease pathophysiology [375].

As suggested, lncRNAs can also serve as host transcripts for small ncRNAs, including miRNAs. For example, the LOC554202 gene, which is effectively transcribed into a recently discovered lncRNA, also functions as a host gene for miR-31 – a major contributor to breast cancer progression and metastasis [376]. Similarly, Dey et al expanded on the previously unknown role of the H19 lncRNA in skeletal muscle differentiation by reporting that the first H19 exon encodes for miRNAs 675-5p and 675-3p – both miRNAs were shown to effectively induce muscle cell differentiation and regeneration [377]. Conversely, microRNA-mediated regulation of lncRNA can also occur. Reports show that miR-9 can target the MALAT1 for degradation in the nucleus by binding directly to miRNA recognition elements on the lncRNA itself [378]. Nonetheless, despite the growing body of evidence linking miRNAs and lncRNAs through a series of cross-regulatory networks, our mechanistic understanding of these processes is still in its infancy and there is much to be learned.

1.5 LncRNAs as Regulators of Endothelial Function

1.5.1 Identifying LncRNAs in Endothelial Cells

As interest in lncRNA function grows, a wealth of established *in vitro* and *in vivo* models are now being used to study lncRNA expression patterns. These, however, are yet to be thoroughly examined and a full representation of endothelial lncRNA expression throughout the cardiovascular system is still to be obtained. Nonetheless, a number of studies on EC function have yielded genes with high functional impact, which we will describe here in further detail (Table 1.2).

For instance, several studies now point to several lncRNAs that are likely to directly impact on endothelial function through pro-apoptotic or pro-migratory effects on ECs. Stimulation of ECs using lipopolysaccharides, for example, will induce endothelial dysfunction, apoptosis and sepsis, eventually leading to elevated CVD risk. This process, as demonstrated by Singh *et al*, also happens in association with the differential expression of hundreds of so far uncharacterised lncRNAs [379]. Another key feature of innate EC response is the release of chemotactic intermediaries produced from the CXCL locus involved in neutrophil recruitment. Recently, these were found to be primed for activation in TNFα-stimulated HUVECs by UMLILO, a proximal enhancer-transcribed lncRNA [380].

Further, hypoxic conditions commonly observed in myocardial infarction, peripheral ischemia and stroke, often triggering distinct endothelial responses to prevent further tissue damage and restore blood supply. Several hundred genes with hypoxia-sensitive expression have been independently reported *in vitro* using HUVECs, leading to the identification and validation of lncRNA such as MALAT1, H19, MIR503HG and LINC00323.

Gene Name	LncRNA Type	Sequence/ Synteny Conservation (GENCODE aligned to PLAR annotation)	Identification of LncRNA in Endothelial Context				Characterisation of LncRNA in Endothelial Context			
			Cell/ Tissue Type	Study Design	ID Tech.	Selection Strategy	Cell/ Tissue Type	Phenotypic Effect	Key Effectors + Interactions	Proposed Mechanism
AGAP2-AS1 (PUNISHER)	antisense	Eutheria (Synteny: Ray-finned Fish)	Differentiated ECs	Differentiation from hESC to Vascular EC	RNAseq	Enriched in final stage of differentiation	Zebrafish/Mouse Embryo, HUVECs	Early vessel branching in zebrafish embryonic development, Maintenance of endothelial differentiation	TAL1, FOXC1	Unknown
GATA6-AS	antisense	Mammalia (Synteny: Ray-finned Fish)	HUVEC	Normoxia to Hypoxia (12 + 24 hours)	RNAseq	High abundance and upregulation relative to other lncRNA	HeLa/HUVEC	EndMT/Tip cell formation/Migration	LOXL2	Directs nuclear portion of LOXL2 to remove H3K4me3
HAGLR (STEEL, HOXD-AS1)	antisense	Ray-finned fish (Synteny: Ray-finned fish)	HUVEC/ HMVEC	Profiling of Primary Vessel Cell Types	LncRNA Custom Microarray	Enriched expression vs 4 Non EC types, proximity to Hox locus	HUVEC	Migration/Proliferation/ Apoptosis/Angiogenesis	KLF2, eNOS PARP1	Recruits epigenetic regulator PARP1 to target promoters
H19 (ASM)	lincRNA	Eutheria (Synteny: Eutheria)	HUVEC	Normoxia to Hypoxia (24 + 48 hours)	RNAseq	Hypoxia-sensitive mouse ortholog, sensitive to vascular injury	HAOEC	Supports hypoxia-induced Angiogenesis	Not examined here	Not examined here
LINC00520 (LEENE)	eRNA	No sequence ortholog (Synteny: Ray-finned fish)	HUVEC	Physiological (Pulsatile) vs Pathological (Oscillatory) shear stress	RNAseq	High upregulation relative to other lncRNA, correlation with eNOS	HUVEC	eNOS expression	PolII, KLF4, Med1	Recruitment of Pol II to eNOS promoter
LINC00323 (C21orf130)	antisense	Catarrhini (Synteny: Eutheria)	HUVEC	Normoxia to Hypoxia (12 + 24 hours)	RNAseq + Non Coding Microarray	Strongest upregulation	HUVEC/HCAEC	Supports hypoxia-induced Angiogenesis	eIF4A3	Scaffold, indirect binding to GATA2
MALAT1 (HCN, LINC00047)	lincRNA	Vertebrates (Synteny: Ray-finned fish)	HUVEC	Profiling of HUVEC	RNAseq	Presence in other EC types	HUVEC	Supports hypoxia-induced Angiogenesis	S-phase cyclins, p21	Not examined in ECs
MANTIS (AK125871, ANXA4-AS)	antisense	Absent from GENCODE	HUVEC	siRNA targeting histone demethylase JARID1B	Exon-array	Upregulated after Histone Demethylase Depletion	HUVEC	Supports endothelial angiogenic function	BRG1, BAF155	Nucleosome remodelling via BRG1 interaction
MEG3 (GTL2, LINC00023)	lincRNA	Eutheria (Synteny: Ray-finned Fish)	HUVEC	Profiling of HUVEC during hypoxia + senescence	RNAseq/qRT-PCR	Upregulated in high HUVEC passage number	RF/6A (Primate retinal EC)	Supports Proliferation, Migration, Glucose-induced Apoptosis	Akt, EZH2/JARID2 (documented in ESCs)	Recruitment of EC-enriched histone demethylase JARID2
MIR503HG (H19X)	miR host	Eutheria (Synteny: Lobe-finned fish)	HUVEC	Normoxia to Hypoxia (12 + 24 hours)	RNAseq + Non Coding Microarray	Strong upregulation	HUVEC/Ea.Hy926	Supports proliferation and migration but no angiogenic effect	Not examined here	Host for miR-503
SENCR (FLI1-AS1)	antisense	Eutheria (Synteny: Mammalia)	HUVEC/ hESC-EC	Hemogenic and directed differentiation of EC from hESC	qRT-PCR	Known enrichment in vascular cells	HUVEC/hESC-EC	Mesodermal/Endothelial Commitment, supports VEGF-induced angiogenesis + membrane integrity	CCL5, CXCL3L1, CKAP4, CDH5	Binds CKAP4, freeing CDH5 to stabilise adherens junctions
UMLILO (AC112518.3)	eRNA	Catarrhini (Synteny: Ray-finned fish)	HUVEC	TNFα treatment	Hi-C, ChIA-PET, ChIP-Seq	Interaction with chromosomal locus of CXCLs	HUVEC	Chemokine Expression	CXCL Locus	Facilitates MLL1-priming of CXCL locus via H3K4me3

Table 1.2: LncRNAs Associated with Endothelial Function in Cardiovascular Disease.

List of lncRNAs reported to have endothelial regulatory functions in cardiovascular disease. LncRNAs are presented together with their type and evolutionary conservation, followed by details regarding their identification and characterisation in ECs. Table adapted from Monteiro *et al*, 2019 [314].

With additional *in vitro* and *in vivo* characterisation, several of these lncRNAs were shown to regulate a variety of hypoxia-induced EC functions such as proliferation, migration and angiogenesis [316,381,382].

Interestingly, lncRNAs initially identified and studied in different cell types, such as MEG3, MIAT and RNCR3, have been linked to retinal EC angiogenesis and are now associated with microvascular visual impairment [383–385] (Figure 1.9).

In the study of endothelial development, several models using human embryonic stem cells (hESCs) have also been widely implemented to identify lncRNA expressed at early stages during cell development and cell fate commitment. Crucially, these models have so far identified lncRNAs as such SENCR [386] and PUNISHER [387], both of which are not only upregulated during EC fate commitment but may also be important in maintaining endothelial homeostasis at later stages. For example, knockdown of the endothelial-associated lncRNA SENCR in HUVEC was found to trigger VE-cadherin internalisation, leading to perturbed adherens junctions and defective membrane integrity [388]. Interestingly, the expression of SENCR was also shown to be altered in vascular samples from patients with critical limb ischemia and coronary artery disease [386].

Additionally, given the wide variety of vascular microenvironments to which EC are exposed, it is important to consider the extracellular cues that may also guide their function [389]. For example, the release of extracellular vesicles (EVs) by surrounding neighbouring cells is known to be a major factor involved in regulating endothelial function and dysfunction [390]. Unsurprisingly, recent studies have started to show that lncRNAs can be selectively packaged, , along with other regulatory molecules, into EVs to facilitate cell-to-cell communication and induce a variety of phenotypic

changes[391]. For example, the release of EVs carrying the lncRNA RNCR3, produced by ECs, was shown to prompt an increase in VSMC proliferation and migration[392]. While conversely, exposure to exosomes released by monocytes containing the atherosclerosis-associated lncRNA GAS5 was found to enhance apoptosis in ECs *in vitro* [393].
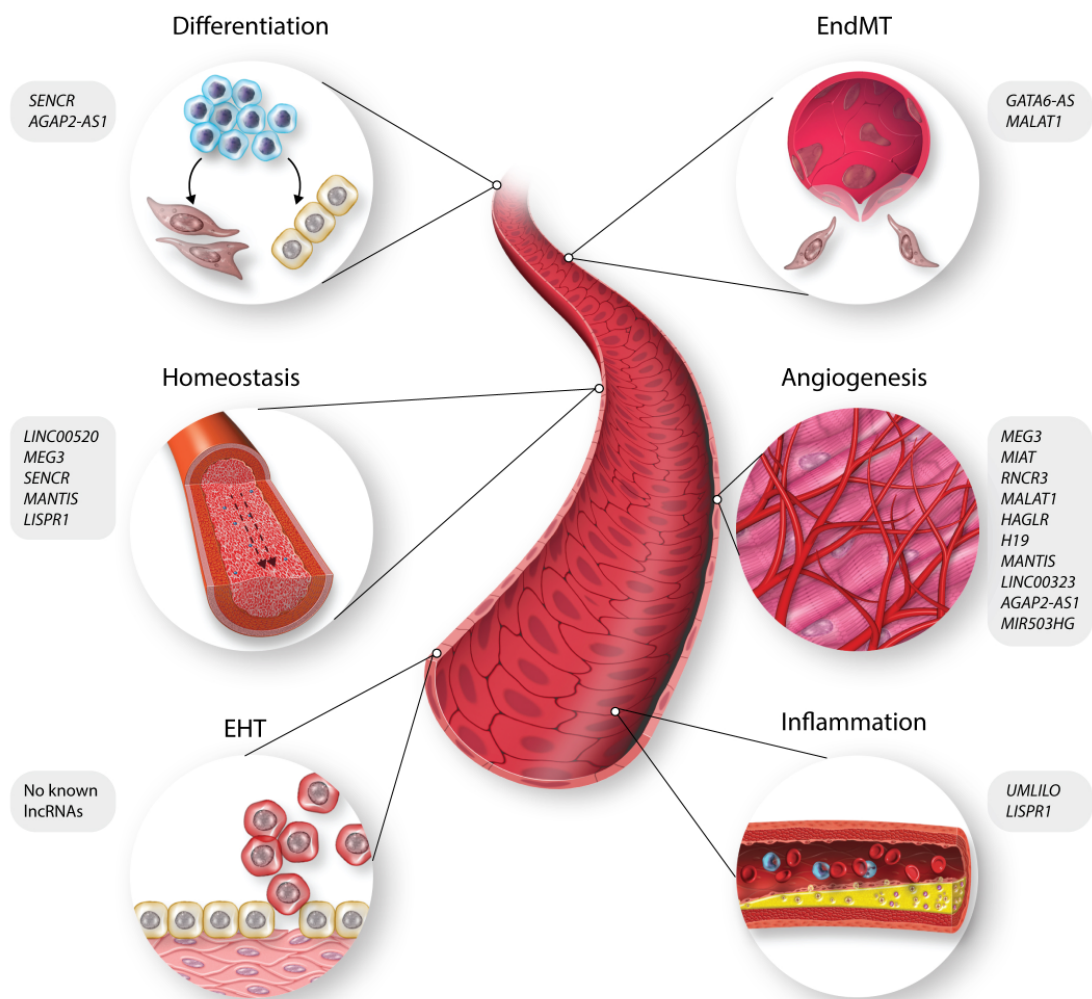
Figure 1.9: Known Function of Endothelial LncRNAs.

List of lncRNA reported to have endothelial regulatory functions with impact on cell differentiation, EndMT, angiogenesis, inflammation, EHT and vessel homeostasis. Figure adapted from Monteiro *et al*, 2019 [314].

Additionally, exosomes isolated from liver cancer cells enriched with the lncRNA H19 promoted an angiogenic phenotype and increase cell adhesion by stimulating VEGF and VEGFR1 production in HUVEC [394].

It is important to acknowledge however, that while the studies described have facilitated the discovery and characterisation of EC-associated lncRNAs, several of these publications do not account for EC heterogeneity and have a clear bias towards HUVEC-based models. Further highlighting this issue, to this day only one lncRNA-screening study actually included primary EC subpopulations other than HUVEC [395]. More specifically, by applying custom lncRNA microarrays, representing 23155 putative lncRNAs, to HUVEC and human dermal microvascular ECs (HMVECs). Crucially, the groups analysis identified a total of 116 EC-enriched lncRNAs, of which 29 where common between HUVECs and HMVECs, while 28 were unique to HUVEC and 59 to HMVEC [395]. This is of particular importance given the cell-specific and context-dependent nature of lncRNAs [396]. Manipulation of the lncRNA MIR503HG, for example, was shown to reduce cell proliferation and migration in HUVEC but had no effect on HPAECs [381]. Nevertheless, as the use of RNA sequencing approaches becomes common place, lncRNA research addressing EC heterogeneity is likely to increase over the coming years.

1.5.2 LncRNAs and EndMT

While the influence of lncRNAs on cell function and dysfunction is starting to become apparent, research surrounding EC-based mechanisms other than those supporting angiogenesis and homeostasis has been, comparatively, limited. This is particularly true regarding EndMT, with no published data of transcriptome-wide shifts in lncRNA expression, the transition process has so far only been associated with two known

lncRNAs (Table 1.2; Figure 1.9). The lncRNA MALAT1 (metastasis associated lung adenocarcinoma transcript 1), for example, was the first to be implicated in the modulation of EndMT. Upregulated in human endothelial progenitor cells (EPC) treated with TGF-β1, the lncRNA was shown to repress the expression of DICER and with it the availability of miR-145. Effects on EndMT progression, however, were found to be largely mediated by miR-145 which directly targets TGFBR2 and SMAD3 [397]. Recently, the lncRNA GATA6-AS (GATA6 antisense RNA) was shown to suppress TGF-β-induced EndMT *in vitro* via the lysyl oxidase homolog 2 (LOXL2) to regulate chromatin remodelling. Although this interaction may be present, the study relied largely on vein ECs and results may not be applicable to EndMT in other ECs [398]. While these studies do represent a step forward in the field, further comprehensive reports across multiple endothelial phenotypes are still needed.

1.6 LncRNA Potential for Clinical Translation

While the putative therapeutic applications of lncRNAs are often mentioned throughout the literature, the majority of clinical studies tend to focus on their utility as markers of disease. For example, the lncRNA HOTAIR has now been described by a growing number of studies as a potential biomarker for breast, liver, gastric, lung, and oesophageal cancer[399]. Further, the expression of MALAT1, originally associated with lung cancer metastasis, has now been linked to the development and progression of a variety other cancers [400–404]. As shown in a recently published meta-analysis using data from 14 independent studies, MALAT1 expression was found to an independent predictor of overall survival rates in patients with respiratory, digestive and other system cancers [402]. Similarly, the lncRNA SNHG15 (Small Nucleolar RNA Host Gene 15) was also reported as a potential prognostic marker in hepatocellular carcinoma, with increased tissue levels of the lncRNA shown to be associated with decreased survival rates [405].

Despite not being endothelial-specific, several lncRNAs with biomarker potential have also been described in CVD. For example, plasmas levels of the lncRNA H19 and LIPCAR (long intergenic non-coding RNA predicting cardiac remodelling) were found to be significantly increased in patients with coronary artery disease (CAD) [406]. Increased expression of LIPCAR in particular has consistently been described as an independent predictor of cardiovascular-related death in patients with heart failure [407,408]. Similarly, when comparing lncRNA expression in peripheral blood cells collected from patients with acute MI, the expression of ANRIL, along with MALAT1, were shown to be not only upregulated but also accurate predictors of left ventricular dysfunction after MI [409]. Expressed throughout the vasculature, increased levels of

ANRIL have also been reported in plaque and plasma samples of patients with atherosclerosis [365,410] and associated with the incidence of in-stent restenosis [411]. Given the importance of promoting re-endothelialisation after vascular stenting [412] and considering the lncRNAs reported regulatory functions in EC, including VEGF expression [413], ANRIL may be a potential prognostic factor for vascular remodelling. Interestingly, several single nucleotide polymorphisms (SNPs) in ANRIL locus itself have also been associated with increased susceptibility to CAD and diabetes [414,415]. Nonetheless, despite their promise, the introduction of lncRNA targeting strategies into the clinical setting comes with several challenges and real-world examples of therapeutic applications remain limited. For instance, due to the pleiotropic nature of most lncRNAs, tissue-specific delivery is essential to guarantee not only treatment efficiency but also minimising potential off-target effects. While increased ANRIL expression may link to ISR and atherosclerosis, unrestrained systemic modulation strategies may ultimately be just as harmful, given that the lncRNA is also commonly altered during tumour development and progression [416]. Further, with an average of $\simeq 4$ different isoforms per lncRNA, the transcriptional complexity of any particular lncRNA locus must be carefully considered and characterised before considering translation into the clinical setting [417]. Thus, in order to be effective, clinical strategies targeting or using lncRNAs must take into account not only possible off-target effects, the route of delivery used, drug immunogenicity, treatment dosage and duration but also sub-cellular transcript location, transcript size and sequence. Nevertheless, given the transient and tissue-specific expression patterns of certain lncRNAs, the implementation of tissue-selective approaches may not always be necessary. This is particularly true for the cardiac fibroblast-specific lncRNA WISPER (Wisp2 super-

enhancer–associated lncRNA), making it an attractive candidate for antifibrotic therapies [418].

Lastly, it is also important to consider the lack of sequence conservation seen with lncRNAs in particular, which may prevent the translation of pre-clinical animal studies. The lncRNAs MIRT1 (myocardial infarction-associated transcript 1) and MIRT2 for example, despite being shown to have a robust up-regulation during MI in association with multiple genes known to be involved in left ventricular remodelling, have no corresponding human homologs and thus limited options for translation into the clinical setting [419]. On the other hand , lncRNAs such as SMILR (smooth muscle enriched lncRNA), which are conserved only in humans, may prove difficult to research due to the lack of preclinical animal models [420]. It is possible, however, that while primary sequence conservation may not be readily apparently, homolog secondary and tertiary structures may still exist [421,422]. Thus, adding the possibility to develop strategies to target these structures instead.

1.7 Hypothesis and Aims

With emerging evidence showing that EndMT is active during pathological vascular remodelling, new therapeutic avenues may also appear. Given the regulatory functions of lncRNAs, we hypothesise that the highly conserved lncRNA MIR503HG plays a crucial role in the initiation and progression of EndMT. To address this hypothesis, the work presented within this thesis aimed to:

- To investigate the role of MIR503HG in the progression of EndMT;

- Identify shared cross-regulatory features of the MIR503HG miRNA locus;

- Determine whether MIR503HG is dysregulated during vascular remodelling in PAH.

Chapter 2: Materials and Methods

2.1 Ethical Approval

2.1.1 Human Ethical Information

All experimental procedures using human cells conform to the principles outlined in the Declaration of Helsinki (Ethics 15/ES/0094). For experiments involving blood outgrowth endothelial cell (BOEC) generation, all blood donors provided informed consent in accordance with human study 07/H0306/134 (Cambridgeshire 3 Research Ethics Committee). All human tissues shown were obtained from the Papworth NHS Foundation Trust Hospital Tissue Bank (Papworth Everard, UK). Papworth Hospital ethical review committee approved the use of human tissues (Ethics Ref. 08/H0304/56+5) and informed consent was obtained from all subjects. Both BOEC and patient lung tissues were kindly provided by Professor Nicholas W. Morrell (BHF Cambridge Centre of Excellence, University of Cambridge, UK).

Paraffin wax–embedded lung samples were available from patients with PAH who had undergone lung transplantation and from controls. Control samples comprised tissue from pneumonectomy specimens resected for malignancy, but distant from the site of tumour (Table 2.1).

Tissue Number	Sample	Age	Sex
TB16.0389.A7	Control	55 years 3 months	Female
TB16.0484.B	Control	68 years 0 months	Female
TB16.0526.B	Control	75 years 8 months	Female
TB16.0733.C	Control	74 years 7 months	Male
TB07.0029.A1	IPAH	41 years	Female
TB07.0272.6	IPAH	45 years	Male
TB07.0956.6	IPAH	40 years	Female
TB15.0859.I1	IPAH	24 years 5 months	Female

Table 2.1: **List of donor PAH and control patients.**
Tissue bank number, age and sex of donors with PAH who had undergone lung transplantation and control patients whose tissue samples were used for in situ hybridisation and immunohistochemistry staining. All tissues were obtained from the Papworth NHS Foundation Trust Hospital Tissue Bank.

2.1.2 Animal Ethical Information

All animal procedures conform to the United Kingdom Animal Procedures Act (1986) and with the "Guide for the Care and Use of Laboratory Animals" published by the US National Institute of Health (NIH publication No. 85-23, revised 1996). Animal approval was granted by the University of Edinburgh Committee Board.

2.2 General Laboratory Practice

The laboratory reagents and equipment used were all of the highest commercially available standard. All chemicals, unless otherwise stated, were purchased from Sigma Aldrich (Dorset, UK). Any hazardous chemicals were handled and disposed of in compliance with Control of Substances Hazardous to Health (COSHH) guidelines. Laboratory coats, nitrile powder-free gloves and fume hoods were used where appropriate.

2.3 Cell Culture Methods and Reagents

All procedures described were carried out under sterile conditions in a standard class II biological safety vertical laminar flow cabinet. Cabinets used were appropriately cleaned, both before and after use, with 1% virkon and 70% ethanol. During culture, all cell lines here described were kept at 37°C in a humidified atmosphere containing 5% CO_2.

2.3.1 Human Umbilical Vein-derived Endothelial Cells (HUVEC)

All HUVEC used were purchased from Lonza (#C2519A; Lonza, Slough, UK) and contained pooled male and female donor cells. HUVEC were kept in either T75 or

T150 cell culture flaks (Corning, Poole, UK) and maintained in endothelial cell growth medium (EGM-2 BulletKit™) (Lonza, Slough, UK) supplemented with 10% foetal bovine serum (FBS) (Life Technologies, Paisley, UK) and 100 U/mL penicillin/streptomycin (Gibco, Paisley, UK). Cells were typically passaged after 90% confluence was reached. In order to passage cultured cells, after a wash with warm Dulbecco's calcium and magnesium free phosphate buffered saline (PBS) (Gibco, Paisley, UK), 3 mL of Trypsin-EDTA (Gibco, Paisley, UK) were added to each flask and allowed to incubate at 37°C for 2 min. After total cell detachment from the flask, 7mL of supplemented EGM-2 were added to neutralise the trypsin. The cell culture was then pelleted by centrifugation at 1500 rpm for 5 min. Any supernatant was discarded and the cells re-suspended in normal culture medium before being distributed at a ratio of 1:4. If needed at a specific density, the cells were first counted using a haemocytometer and then seeded in a new cell culture vessel. All cell lines used were held between passages 3-8.

2.3.2 Human Pulmonary Artery-derived Endothelial Cells (HPAEC)

All HPAEC used where purchased from Lonza (#CC-2530; Lonza, Slough, UK) and contained single donor cells, batches from multiple donors were used throughout. HPAEC were kept in either T75 or T150 cell culture flaks (Corning, Poole, UK) and maintained in supplemented EGM-2 as described in Section 2.3.1. Cells were typically passaged after 90% confluence was reached. Passage and seeding of cultured cells was performed following the protocol described in Section 2.3.1. All cell lines used were held between passages 3-8.

2.3.3　Blood Outgrowth Endothelial Cells (BOEC)

Human blood outgrowth endothelial cells (BOEC) used were generated and provided by Prof. Nicholas W. Morrell (University of Cambridge, Cambridge, UK). These were derived from peripheral venous blood isolated from donors diagnosed with PAH or healthy control, as previously described by Toshner et al [423]. Patients cells were kept in either T25 or T75 cell culture flaks (Corning, Poole, UK) and maintained in supplemented EGM-2 as described in Section 2.3.1. Cells were typically passaged after 90% confluence was reached. Passage and seeding of cultured cells was performed following the protocol described in Section 2.3.1. All cells used were held between passages 6-8.

2.3.4　Human Embryonic Kidney (HEK) 293T Cell

All human embryonic kidney (HEK) 293T cells were kept in T150 cell culture flaks (Corning, Poole, UK) and maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS (Life Technologies, Paisley, UK) and 100 U/mL penicillin/streptomycin (Gibco, Paisley, UK). Passage of cultured cells was performed by washing attached cells in warm PBS (Gibco, Paisley, UK) and incubating with 3mL of Trypsin-EDTA (Gibco, Paisley, UK) at 37°C for 2 min. After total cell detachment from the flask, 7mL of supplemented DMEM were added to neutralise the trypsin. The cell culture was then pelleted by centrifugation at 1500 rpm for 5 min. Any supernatant was discarded and the cells re-suspended in normal culture medium before being distributed at a ratio of 1:5. If needed at a specific density, the cells were first counted using a haemocytometer and then seeded in a new cell culture

vessel. For large-scale lentiviral vector production HEK293Ts were instead passaged at a ratio of 1:2, this was done 24 h prior to triple transfection.

2.4 Endothelial to Mesenchymal Transition *in vitro* Models

2.4.1 TGF-β2 and IL1- β 7 Day Model

Before cytokine stimulation, cells were plated at a density of 7×10^4 cells/well using 6-well culture plates (Corning, Poole, UK) and left to incubate for 24 h. HUVEC and HPAEC lines, cultured in supplemented EGM-2 as described above, were exposed to a daily dose of recombinant human TGF-β2 (10 ng/mL) and/or IL1- β (1 ng/mL) (PeproTech, NJ, USA). Cytokine co-stimulation was repeated daily for 7 days. Both recombinant human TGF-β2 and IL1- β were re-constituted in RNase/DNase free H_2O at 20 ug/mL and 10 ug/mL, respectively. To minimise freeze thaw, all reagents were aliquoted at single use quantities and stored at -80°C.

2.4.2 TGF-β2 and H_2O_2 7 Day Model

Prior to treatment, cells were plated at a density of 7×10^4 cells/well using 6-well culture plates (Corning, Poole, UK) and left to incubate for 24 h. HUVEC, cultured in supplemented EGM-2 as described above, were exposed to a daily dose of recombinant human TGF-β2 (50 ng/mL) and/or H_2O_2 (200nM) (PeproTech, NJ, USA; SigmaAldrich, Dorset, UK). Cell stimulation was repeated daily for a total of 7 days. Recombinant human TGF-β2 was re-constituted and stored as described above. H_2O_2 stocks were kept at 4°C and diluted in RNase/DNase free H_2O prior to use.

2.5 Cell Transfection

2.5.1 Double Stranded Dicer-Substrate Short Interfering RNA

Double stranded dicer-substrate short interfering RNA (DsiRNA) targeting specific MIR503HG transcript groups were designed and synthesised following the instructions provided by Integrated DNA Technologies (IDT, Leuven, Belgium) and selected based on gene knockdown effectiveness (Table 2.2). Additionally, a commercial DsiRNA control (IDT, Leuven, Belgium) was used in parallel with all RNA interference experiments as it does not target any sequence in the human, mouse, or rat transcriptome. Transient transfection was performed using Lipofectamine RNAiMAX Tansfection Reagent (Life Technologies, Paisley, UK) according to the manufacturer's instructions, along with 20nM of either si503HG or siControl. 24 h post transfection, the transfection medium was removed and cell culture maintained for a total of 7 days using supplemented EGM-2, as described.

2.5.2 Antisense LNA GapmeR

Antisense LNA GapmeR targeting MIR503HG were generated using Qiagen's LNA GapmeR design tool (Qiagen, Hilden, Germany) and selected based on gene knockdown effectiveness (Table 2.2). A commercial Antisense LNA GapmeR Negative Control (Qiagen, Hilden, Germany) was used in parallel with all LNA GapmeR based RNA interference experiments as it did not target any sequence in the human transcriptome. Transient transfection was performed using Lipofectamine RNAiMAX Transfection Reagent (Life Technologies, Paisley, UK) according to the manufacturer's instructions, along with 20nM of Gap_503HG or gapControl. 24 h post transfection, the transfection medium was removed and cell culture maintained for a

total of 7 days using supplemented EGM-2, as described

Name	Target	Sense Strand (5' → 3')	Antisense Strand (5' → 3')
siCONTROL	Negative Control 1	CGUUAAUCGCGUAUAA UACGCGUAT	AUACGCGUAUUAUACGCGA UUAACGAC
si503HG	MIR503HG_2/4/5	CAAAUAGAAGGGUAA UAUAUAAUCA	UGAUUAUAUAUUACCCUUC UAUUUGGG
GapCONTROL	Negative Control A	AACACGTCTATACGC	
Gap503HG	MIR503HG_2/4/5	TTGGAACAAAGAAGTG	

Table 2.2: Cell Transfection Methods.

List of DsiRNA and GapmeR sequences used.

2.5.3   miRCURY LNA miRNA Inhibitors

All miRNA silencing experiments were conducted using miRCURY LNA miRNA Inhibitors targeting either miR-424-5p or miR-503-5p (Qiagen, Hilden, Germany). An appropriate miRCURY LNA miRNA Inhibitor Negative Control (Qiagen, Hilden, Germany) was used in parallel with all miRNA silencing experiments as it did not have known microRNA targets. See Table 2.3 for a complete list of miRNA inhibitors used. Transient transfection was performed using Lipofectamine RNAiMAX Transfection Reagent (Life Technologies, Paisley, UK) according to the manufacturer's instructions, along with 25nM of either miRNA inhibitor. 24 h post transfection, the transfection medium was removed and cell culture maintained for a total of 7 days using supplemented EGM-2.

2.5.4   Pre-miR miRNA Precursors

All miRNA overexpression experiments were conducted using Pre-miR miRNA Precursors for either miR-424-5p or miR-503-5p (Life Technologies, Paisley, UK). An appropriate Pre-miR miRNA Precursor Negative Control (Life Technologies, Paisley, UK) was used in parallel with all miRNA overexpression experiments as it

did not have known microRNA targets. See Table 2.3 for a complete list of miRNA precursors used.

Transient transfection was performed using Lipofectamine RNAiMAX Transfection Reagent (Life Technologies, Paisley, UK) according to the manufacturer's instructions, along with 15nM of either miRNA precursor. 24 h post transfection, the transfection medium was removed and cell culture maintained for a total of 7 days using supplemented EGM-2.

Name	Target	Assay ID
Anti-miR424	hsa-miR-424-5p	YI04100987
Anti-miR503	hsa-miR-503-5p	YI04100899
Anti-miR CT	Negative Control	YI00199006
Pre-miR424	hsa-miR-424-5p	PM10306
Pre-miR503	hsa-miR-503-5p	PM10378
Pre-miR CT	Negative Control	AM17110

Table 2.3: Cell Transfection Methods.

List of miRNA inhibitors and miRNA precursors used.

2.6   General Cloning Techniques

The following general cloning techniques were employed in order to create a LNT_SFFV plasmid construct containing MIR503HG isoform 2.

2.6.1   Restriction Digest

Plasmid restriction digest was used for both inserting target DNA sections into plasmids and also as diagnostic tools to ensure that the DNA insert was of the correct size. Depending on the reaction size, 100-1000ng of plasmid DNA was used. Small, diagnostic, reactions were set up using 1 µL NEB buffer, 0.5 µL of each restriction enzyme, 100 ng plasmid DNA and made up to a final volume of 10 µL using

RNase/DNase free H_2O. Large reactions used 5 µL NEB buffer, 1 µL of each restriction enzyme, 1 µg plasmid DNA and made up to a final volume of 50 µL using RNase/DNase free H_2O. Each restriction enzyme use was acquired from New England Biolabs (Ipswich, UK).

Restriction digestion was then performed by incubating each reaction in a water bath set at 37°C for 2 h. When required, gel purification was performed as described in Section 2.6.2.

2.6.2   Agarose Gel Electrophoresis and Extraction

Agarose gel electrophoresis was used to separate DNA molecules according to their molecular size. All samples were mixed with 6X purple gel loading dye (New England Biolabs, Hitchin, UK) for a final 1X concentration and loaded onto a pre-cast agarose gel, along with the appropriate DNA size marker (100 base pair and/or 1 kilobase pair DNA ladders) (New England Biolabs, Hitchin, UK).

Agarose powder (Invitrogen, Paisley, UK) was dissolved in 1X Tris-Borate EDTA (TBE) buffer (Gibco, Paisley, UK) to a suitable percentage (1-2% w/v) based on the sample DNA fragment size. In order to visualise the DNA, SYBR Safe DNA gel stain (1 ng/100mL) (Invitrogen, Paisley, UK) was added to molten agarose before casting.

The loaded agarose gel was electrophoresed at a constant voltage of 80-100 V in TBE buffer in BIO-RAD electrophoresis tanks and using a BIO-RAD Power Pac 300. DNA bands were later visualised by UV transillumination on a ChemiDoc XRS+ Imaging System.

After agarose gel electrophoresis, the desired DNA fragments, visualised by UV transillumination, were excised from the gel using a clean scalpel and purified using the PureLink Quick Gel Extraction Kit (Invitrogen, Paisley, UK) following the

instructions provided by the manufacturer's protocol. The purified DNA was stored at -20°C until required.

2.6.3   Dephosphorylation and Ligation

Digested DNA will possess a 5' phosphate group required for ligation, in order to prevent recipient plasmid self-ligation the 5' phosphate can be removed prior to the ligation reaction. This was performed using Calf Intestinal Alkaline Phosphatase (CIP) (New England Biolabs, Hitchin, UK), following the manufacturers' protocol.

Ligation of DNA insert into the dephosphorylated plasmid backbone was performed using T4 DNA ligase (New England Biolabs, Hitchin, UK), with a molar ratio of 1:3 plasmid to insert along with a 1:0 negative control reaction. Molar ratios were calculated using the following equation:

$$\frac{ng(Vector) \times Kb(Insert)}{Kb(Vector)} \times Ratio$$

Ligation reactions were prepared in 200 uL PCR tubes combining 150 ng of dephosphorylated plasmid vector, specific amount of insert DNA (as calculated above), 1 µL T4 DNA ligase enzyme, 2 µL 10x T4 DNA ligase buffer and made up to a final volume of 20 µL using RNase/DNase free H_2O. Samples were incubated at 16°C overnight in a thermal cycler. Ligations were subsequently subjected to further diagnostic digests and sequencing before being transformed into competent *Escherichia coli* (*E coli*).

2.6.4   DNA Sequencing

DNA sequencing was used to confirm that the plasmid construct contained the correct DNA sequence and that no mutations were present. Each sequencing reaction was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems); containing 200 ng of plasmid DNA, 2 nM of forward or reverse sequencing primer, 1 µL of v3.1 Ready Reaction Mix (Applied Biosystems), 4 µL of 5x v3.1 BigDye Sequencing Buffer (Applied Biosystems) and made up to a final volume of 20 µL using RNase/DNase free H_2O. Each reaction was assembled in a 200uL PCR tube and performed in duplicate along with a control sequencing reaction missing the template DNA.

Each sequencing reaction was then subject to 25 PCR cycles at 96°C for 50 seconds to denature the plasmid, 50-55°C for 20 seconds to facilitate primer annealing and 60°C for 4 min to extend the target DNA fragment. Samples were then sent to Edinburgh Genomics (University of Edinbrugh, Edinburgh, UK) for sequencing analysis.

2.6.5   Bacterial Transformation and Plasmid Purification

All plasmid DNA was replicated using bacterial transformation of competent *E coli*. Bacterial transformations were performed using a standardised 'heat shock' protocol combining the desired plasmid DNA and commercially available One Shot TOP10 (Thermo Fisher, Paisley, UK) or Stellar (Clontech, California, USA) Chemically Competent *E coli*. In summary, 1 ng of the desired plasmid DNA was added to 50 µL aliquots of chemically competent bacteria were and placed on ice.   The bacteria/plasmid mix was left on ice for 5 min, then submerged in a water bath at 42

$^{\circ}$C for exactly 45 seconds, removed and immediately placed on ice for an additional 2 min.

450 µL of S.O.C. medium (room temperature) was added to each tube, these were incubated at 37°C while shaken at 200 rpm for 1 h. In order to ensure antibiotic efficiency, a negative control containing only *E coli* (without plasmid DNA) was included. Luria agar (Thermo Fisher, Paisley, UK) plates containing 100 µg/mL of ampicillin were inoculated using different amounts of E coli/S.O.C medium. The inoculated plates were then incubated overnight at 37 °C to allow for colony formation. The following morning, single bacterial colonies were selected and allowed to grown in Luria broth (Thermo Fisher, Paisley, UK), containing 100 µg/mL ampicillin, for either mini or maxiprep plasmid purification. All purified plasmids were screened using diagnostic restriction digests, DNA electrophoresis and plasmid sequencing, as described.

2.6.6   Mini and MaxiPrep

Small- and large-scale plasmid DNA purification was performed using either a PureLink® Quick Plasmid Miniprep Kit or HiPure Plasmid Maxiprep Kit (Thermo Fisher, Paisley, UK), according to the manufacturer's instructions. Once extracted, DNA quantification and quality control was conducted using a NanoDrop ND-1000 Spectrophotometer (Nano-Drop Technologies, Wilmington, DE). When necessary, samples were further diluted by adding RNase/DNase free H_2O. All plasmid DNA samples were stored at -20°C until required.

2.6.7   Plasmid Glycerol Stocks

After transformation, bacteria carrying a specific plasmid were saved for long-term storage as glycerol stocks. Prior to maxiprep plasmid isolation, 500 μL of overnight culture was added to a cyrovial containing 500 μL of 50% glycerol and stored at -80℃. In order to recover stored bacteria, glycerol stocks were thawed and streaked on agar plates with the appropriate antibiotic for single colonies selection.

2.7   Lentiviral Manipulation of LncRNA Expression

2.7.1   MIR503HG Lentivirus Plasmid

A synthetic gene MIR503HG_2 was assembled from synthetic oligonucleotides, based on the human genome annotation GRCh38, and inserted into a MA-T plasmid backbone (pMA-T) by GeneArt Gene Synthesis (Invitrogen, Paisley, UK).
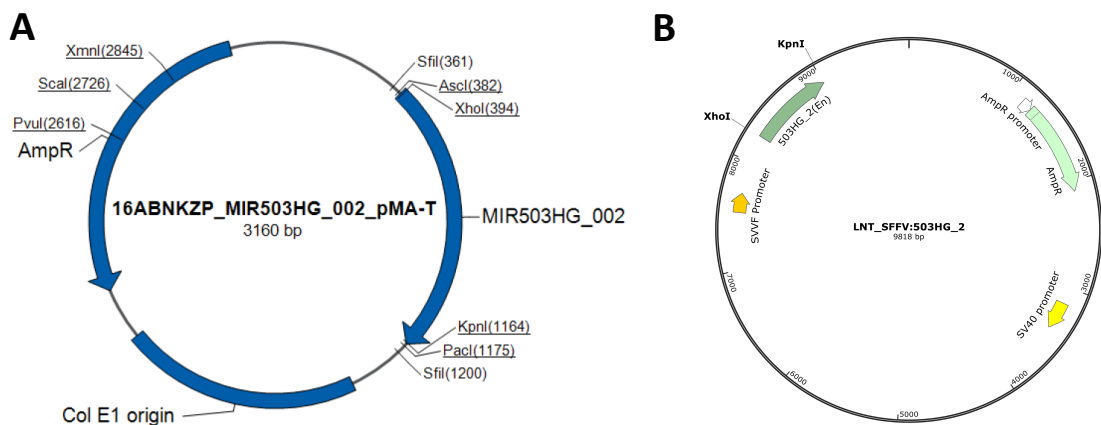
Figure 2.1: Plasmid Map.

A) MIR503HG_2_pMA-T plasmid map containing a synthetic MIR503HG_002 sequence, containing the XhoI and KpnI restriction sites, assembled from synthetic oligonucleotides and inserted into a MA-T plasmid (pMA-T). **B)** LNT_SFFV:503HG plasmid map containing MIR503HG_2 sequence between XhoI and KpnI restriction sites. The LNT_SFFV plasmid is under the control of the SFFV promoter and exhibits ampicillin resistance.

The MIR503HG_2 gene was spliced and cloned into the LNT_SFFV plasmid utilising the XhoI and Kpnl restriction sites at 5' and 3' ends of the transcript. Cloning was performed according to Section 2.6. The MIR503HG_pMA-T and LNT_SFFV:503HG plasmid maps are shown in Figure 2.1.

2.7.2   Production of Lentivirus Via Triple-Transfection

Large-scale lentiviral vector production was conducted using a transient triple-transfection method, whereby HEK293T cells were transfected using 3 plasmids, required for lentiviral production. Namely, an expression plasmid (pHR'SIN-cPPT-SFFV-MCS-WPRE; pSFFV Lenti MCS) containing the gene interest and a spleen focus-forming virus (SFFV) promoter (provided by Prof. Adrian Thrasher, Institute of Child Health, University College London, London, UK), a lentiviral packaging plasmid (pCMVΔ8.74) containing the Gag, Pol, Tat and Rev genes, and an envelope plasmid encoding the vesicular stomatitis virus protein (VSVg) (pMDG). Polyethylenimine (PEI) was used as a plasmid transfection reagent.

All lentivirus preparations were performed in batches of twelve 150 cm^2 cell culture flasks. For each flask, two separate mixes were prepared; one containing the 3 plasmids for transfection, and second mix containing PEI. The plasmid mix contained 50 µg expression plasmid, 17.5 µg PMDG and 32.5 µg pCMVΔ8.74, diluted in 5 mL of Opti-MEM reduced serum medium with GlutaMAX supplement (Gibco, Paisley, UK), which was filtered using a 0.22 µm sterile filter. The second mix contained 1 µL PEI diluted in 5 mL of Opti-MEM reduced serum medium, which was finally sterile filtered and added directly to the plasmid mixture. The plasmid/PEI mixture was then incubated at room temperature with minimal light exposure for 20 min. Incubation

allowed for the formation of DNA:PEI complexes, which have endosomolytic activity, and are protected from lysosomal degradation.

Before transfection, the culture medium was removed from HEK293Ts and the cells were washed in fresh Opti-MEM medium. This was then removed and 10 mL of medium containing DNA:PEI complexes was added to each flask. The cells were allowed to incubate for 4 h at 37°C and 5% CO_2. Later, the transfection medium was removed, replaced with 20 mL fresh complete culture medium and cells were returned to the incubator for 48 h. During this period, after a successful transfection, lentiviral particles are produced and released into the medium by the cells. After 48 h the virus containing medium was collected and filtered using a 0.22 μm sterile filter unit. A further 10 mL of fresh complete culture medium was added to each flask and the cells cultured for a further 24 h. The medium was again collected, filtered and combined with the medium removed at 48 h.

2.7.3   Concentration of Lentivirus

The lentiviral particles were collected and concentrated using ultracentrifugation. Briefly, medium collected from triple-transfected cells was aliquoted into Beckman 14 x 95 mm (14 mL) plastic tubes (Beckman Coulter, London, UK), previously rinsed with 70% ethanol, and loaded into a Sorvall WX+ Ultracentrifuge (Thermo fisher, Paisley, UK). The collected medium was centrifuged at 23,000 rpm for 1 h at 4°C. The supernatant was then discarded and the process was repeated until all virus-containing medium had been used. After the final centrifugation, the supernatant was again discarded and tubes placed up-side down to remove all traces of medium. All collection tubes were placed on ice and loaded with 100 μL of Opti-MEM reduced serum medium with GlutaMAX supplement (Gibco, Paisley, UK), they were then

incubated on ice for 30 min. The lentivirus pellets were finally suspended in the medium, aliquoted at single use quantities and stored at -80°C until required.

2.7.4   Calculation of Lentivirus Titre

Viral titre, in particle infectious units per mL (PIU/mL), was quantified by TaqMan® qRT-PCR based detection. See Table 2.4 for sequence of primers and probe used. Briefly, HEK293T cells were first cultured and then seeded into a 12-well cell culture plate (Corning, Poole, UK) at a density of 5x104 cells per well. The next day, decreasing concentrations of the lentivirus produced were added to each well and left for 72 h. The cells were washed in 1X PBS (Gibco, Paisley, UK) and left with 200 µL of 1x PBS in each well.

Name	Sequence (5' → 3')
PRIMER (FW)	5'-TGTGTGCCCGTCTGTTGTGT-3'
PRIMER (RV)	5'- GAGTCCTGCGTCGAGAGAGC-3'
PROBE	5'-(FAM)-CAGTGGCGCCCGAACAGGGA- (TAMRA)-3'

Table 2.4:   Viral titre primer and probe sequence list.

Prior to DNA extraction, transduced cells underwent one freeze thaw cycle placing the plates at -20°C for 10 min and removing them. Total DNA was then extracted using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). Once extracted, DNA quantification was conducted using a NanoDrop ND-1000 Spectrophotometer (Nano-Drop Technologies, Wilmington, DE), all samples were then diluted to 250 ng/µL using RNase/DNase free H_2O. In order to calculate the viral titre, serial dilutions of the purified expression plasmid (SFFV-LV) were used to generate a standard curve of $1x10^{13}$ to $1x10^4$ plasmid copies.

The calculations required to determine the μL of virus needed to generate the top standard are shown below:

1. **Molecular weight of 1 copy of expression plasmid**

$$\frac{Size\ of\ plasmid\ (bp) \times size\ of\ 1\ bp\ (660Da)}{Avogadro's\ Constant} = g\ per\ molecule$$

Daltons = g/mole

Avogadro's Constant = 6.023 x 10^{23} molecules/mole

2. **Determine copy number of plasmids in 1 mL stock.**

$$\frac{Concentration\ of\ stock\ plasmid\ (g/mL)}{g\ per\ molecule} = no.\ of\ molecules\ per\ mL$$

3. **Preparation of top standard (1x10^{13} copies)**

$$\frac{No.\ of\ molecules\ per\ 1\ mL}{1 \times 10^3} = initial\ dilution\ factor$$

$$\frac{1000}{Initial\ dilution\ factor\ for\ top\ standard}$$

$$= μL\ of\ plasmid\ needed\ for\ top\ standard$$

A TaqMan® qRT-PCR reaction master mix was prepared containing 2x TaqMan® Universal Master Mix, 2 μL Primer/Probe mix, 3.125 μL Nuclease-free H2O, 1 μL of DNA standard or DNA samples collected from lentiviral titre. Finally, 11.5 μL of master mix was added to each well of a 384-well PCR plate. All reactions were performed in triplicate using a QuantStudio 7 Flex Real-Time PCR System real time PCR system (Thermo Fisher, Paisley, UK) as described under Section 2.16. The titre

of each sample was calculated by plotting the CT values from the standard curve and solving the equation of the line for each sample.

The number of copies plasmid DNA per cell was then identified by using the equation below:

$$\text{Copies of plasmid in sample} \times \text{no. of cells used in 250 ng} = \frac{\text{Copies of plasmid}}{\text{per cell}}$$

The total copies of plasmid per cell was then used to generate both the PIU/mL and the MOI (multiplicity of infection), respectively, shown in 1. and 2. below:

1. **PIU/mL Calculation**

$$\frac{\text{Copies of plasmid}}{\text{per cell}} \times \frac{\text{Dilution factor of virus stock used} \times 1000}{\text{µL virus added to cells}}$$

$$= \text{PIU/mL}$$

2. **µL of virus per well**

$$\frac{Cell\ no. \times\ MOI}{Viral\ Titre} = \text{µL of virus / well}$$

2.7.5 Lentivirus Transduction of HUVEC

Lentiviral transduction was performed over a period of 24 h on a confluent monolayer of HUVEC. Briefly, healthy HUVEC, maintained in supplemented EGM-2, were plated at a density of 7×10^4 cells/well on a 6-well cell culture plate (Corning, Poole, UK). The cells were allowed to attach to the plate surface overnight at 37°C and 5% CO_2. A negative control (LNT_CT) or lncRNA containing lentivirus (LNT_503HG)

was added directly to the culture medium of each well at a multiplicity of infection (MOI) of 5 (MOI is the ratio of the number of virus particles to the number of target cells). 24 h post transduction, the transduction medium was removed and cell culture maintained in supplemented EGM-2 alone or containing recombinant human TGF-β2 (10 ng/mL) and IL1- β (1 ng/mL) (PeproTech, NJ, USA). The medium was replaced daily, along with cytokine co-stimulation, for 7 days.

2.8   Animal Models

2.8.1   Inducible Endothelial Tracking Model

Inducible endothelial tracking (Ind.EndoTrack) transgenic mice were developed based previously published endothelial lineage tracing models by Madisen et al and Monvoisin et al [424–426]. In brief, Cdh5-Cre-ER^{T2}-TdTomato (Ind.EndoTrack) transgenic mice were generated by crossing Cdh5-Cre-ER^{T2} (strain Tg(Cdh5-cre/ER^{T2})^{1Rha}) with a ROSA-TdTomato reporter mouse line (strain B6.Cg-Gt(ROSA)26Sor$^{tm9(CAG-tdTomato)Hze}$/J). The Cdh5-Cre-ER^{T2} line was produced using a P1 artificial chromosome (PAC) vector containing the tamoxifen-inducible Cre-ER^{T2} sequence and a VE-cadherin (Cdh5) promotor to direct endothelial-specific gene expression (Ralf H Adams, MGI:3848982) [424]. ROSA-TdTomato reporter line was produced by inserting a Rosa-CAG-LSL-tdTomato-WPRE targeting vector into the Gt(ROSA)26Sor locus (#007909, The Jackson laboratory) [426].

Induction of Cre-recombinase activity was achieved by administering 400 mg/kg body weight per dose of tamoxifen dissolved in sterile corn oil by gavage on 3 alternate days for 5 days, this was followed by two weeks of rest before initiating the Hypoxia/SU5416 PH model.

2.8.2   Hypoxia/SU5416 Model of Pulmonary Hypertension

Eight to eleven-week-old female mice were injected once weekly with Sugen 5416 at 25 mg/kg body weight per dose for a total of 21 days. Stock Sugen 5416 powder was suspended in CMC solution (0.5% [w/v] carboxymethylcellulose sodium, 0.9% [w/v] sodium chloride, 0.4% [v/v] polysorbate 80 and 0.9% [v/v] benzyl alcohol in deionized water), then vigorously vortexed and sonicated to produce a homogenous solution. Any age/sex matched controls received the same volume of vehicle alone.

During this period the mice were exposed to either normoxic conditions or chronic hypobaric hypoxia inside a ventilated plexiglass chamber with a controlled atmosphere of 10% oxygen. Excess ammonia was removed by ventilation and activated charcoal filtration through an air purifier. At the end of the treatment period the animals underwent terminal anaesthesia and right ventricular systolic pressure (RVSP) measurement. Following RVSP measurement, lung tissues were harvested for cell isolation or histology.

2.9   Intranasal Delivery of Lentivirus

The lentivirus construct was delivered to eight to eleven-week-old female mice on 3 alternate days for a total of 5 days, this was followed by a 2 day rest period before initiating the Hypoxia/SU5416 PH model as described below. Prior to the intranasal procedure, each mouse was anesthetised in an induction chamber containing 3% (v/v) isoflurane (Abbot Laboratories, Berkshire, United Kingdom) supplemented with O_2 at a flow rate of 0.5 L/min. Post-induction, the mouse was gently restrained by scruffing and held in a vertical position while the liquid was administered directly through the

nostrils. Each delivery consisted of 25uL of Opti-MEM medium (Gibco, Paisley, UK) containing either LNT_503HG_2 or LNT_Control at 4×10^8 piu/mL.

2.10 Assessment of Pulmonary Hypertension

2.10.1 Anaesthetic Induction

Anaesthesia was initiated in an induction chamber containing 3% (v/v) isoflurane (Abbot Laboratories, Berkshire, United Kingdom) supplemented with O2 at a flow rate of 0.5 L/min. Post-induction, mice were fitted with facemask supplying 1.5% (v/v) isoflurane supplemented with O2 at flow rate of 0.5 L/min. Appropriate levels of anaesthesia were confirmed by the absence of hind limb reflex before and during surgery.

2.10.2 Right Ventricular Systolic Pressure Measurement

RVSP of anaesthetised mice was measured by catheterisation of the right ventricle (RV) of the heart to allow for the measurement of right ventricular pressure. RV catheterisation was achieved via the right jugular vein using a 1.4 F Millar catheter (SPR-671, Millar, Houston TX, USA). Once in place, continuous RVSP measurement was carried out using a calibrated 25 mm gauge heparinised saline filled needle attached to an Elcomatic E751A pressure transducer connected a MP100 data acquisition system (BIOPAC Systems Inc, Santa Barbara, USA). Mean RVSP, systolic and diastolic RVSP were measured at three independent areas of the steady trace and PH was determined when reaching values exceeding 30 mm Hg.

2.11 Mouse Lung Cell Isolation

TdTomato+ and/or CD31+ endothelial mouse lung cells were isolated as previously described Fehrenbach et al [427]. In brief, exposed lungs were first filled with and later fully submerged in digestion buffer. The lung digestion buffer consisted of a PBS mix containing 1mg/ml of Liberase™ (SigmaAldrich, Dorset, UK), 0.1mM EDTA (Thermo Fisher, Paisley, UK) and 1U/uL DNase (Thermo Fisher, Paisley, UK). Submerged lung samples were then incubated in a shaking water bath for a total of 15 min at 37°C. The lung tissue was then transferred to a gentle MACS C Tube, topped with wash buffer, and further processed using the gentleMACS Dissociator (Miltenyi Biotec, Auburn, CA). The wash buffer added consisted of a PBS mix containing 1% BSA (Thermo Fisher, Paisley, UK) and 1mM EDTA (Thermo Fisher, Paisley, UK). The dissociated cell mix filtered through a 70 μm nylon mesh (Miltenyi Biotec, Auburn, CA) and centrifuged at 2000 rpm for 5 min at 4°C. Once pelleted, the supernatant was removed, the cell mix resuspended in cold ACK lysis buffer (Thermo Fisher, Paisley, UK) and left for 10 min on ice; this was again centrifuged at 2000 rpm for 5 min. The pelleted cell mix was resuspended in wash buffer and finally sorted using the BD FACSAria II cell sorter (BD Biosciences) based on their CD31$^+$ or TdTomato$^+$ status.

2.12 Histology

2.12.1 Immunohistochemistry

Prior to staining, the formalin-fixed and paraffin embedded (FFPE) tissue sections were de-paraffinised and rehydrated by washing in Xylene for 5 min, followed by sequential 5 min washes in 100%, 96%, 70% ethanol and water, respectively. The

rehydrated tissue sections were then submerged in warm sodium citrate buffer (10mM, pH6.0) and boiled for 10 min in a microwave for antigen retrieval. After boiling, these were left to cool at room temperature in sodium citrate buffer for 15 min. All samples were then washed under running water for 10 min, followed by two washes using tris-buffered saline plus 1% tween-20 (TBS-T) for 5 min each.

Next, the sections were blocked in 3% goat serum/TBS for 1 h at room temperature to reduce non-specific background staining. Primary antibodies, diluted in in 3% goat serum/TBS, were then added and left for 1 h further at room temperature. The antibody dilutions used were as follows: 1/100 vWF (#A0082, Dako) and 1/1000 αSMA (MO851, Dako), matching Mouse IgG (#10400C, Invitrogen) and Rabbit IgG (#ab172730, Abcam) negative controls were used at the same concentrations. After the incubation period, the sections were washed in TBS-T three times for 5 min and incubated with their corresponding fluorescent conjugated secondary antibodies for 1 h at room temperature. The secondary antibody dilutions were as follows: 1/500 Goat Anti-Mouse (#A11001, Alexa Fluor 488, Invitrogen) and 1/500 Goat Anti-Rabbit (#A11010, Alexa Fluor 546, Invitrogen). The stained tissue sections were finally washed in TBS-T three times for 5 min and mounted using ProLong™ Gold antifade reagent (#P36935, Invitrogen). All slides were imaged using an Axio Scan slide scanner (Zeiss) and ZEN imaging software (Zeiss).

2.12.2 Immunocytochemistry

Cells were first in PBS containing 4% (w/v) paraformaldehyde for 5 min and the permeabilised in PBS with 0.2% (w/v) Triton X-100 for a further 5 min. This was followed by incubation in PBS/3% BSA for 30 min, then overnight with primary

antibodies anti-Slug (#IC7408P, R&D systems), anti-αSMA (#IC1420A, R&D systems), anti-CD31 (#555445, BD Biosciences) at 4°C. The next day cells were incubated with secondary antibodies conjugated with Alexa Fluor (Life Technologies, Paisley, UK) at room temperature for 1 h, followed by three washes with PBS. Finally, the before being mounted using glass coverslips and ProLong® Gold Antifade Mountant. Immunofluorescence images of HUVEC and HPAEC were acquired with Andor Revolution XDi spinning disk confocal microscope and analysed with Image J Software. All settings for the microscope and software were optimised and then maintained for each set of experiments, so that set of cells can be compared accurately.

2.12.3 *In Situ* Hybridisation

Prior to staining, the formalin-fixed and paraffin embedded (FFPE) tissue sections were deparaffinised and rehydrated by washing in Xylene for 5 min, followed by sequential 5 min washes in 100%, 96%, 70% ethanol, DEPC treated water and PBS respectively. The miRCURY LNA miRNA ISH Buffer Set (#339450; Qiagen, Hilden, Germany) was then utilised for tissue digestion and probe hybridisation. Specifically, the rehydrated tissue sections were first incubated with proteinase K (1:1000) in PBS 4 min at 37°C then rinsed under flowing DEPC-treated water and washed twice in PBS for 5 min. The 1x hybridisation buffer was used to dilute the ISH probes (Table 2.5) to 25nM. The tissue sections were then incubated with the probe/buffer dilution overnight at 55°C in a humidified chamber, the probe dilution was sealed against the tissue using coverslips and rubber glue to prevent drying out. The following morning, the sections were washed three times in 5x saline-sodium citrate buffer (SSC) for 5 min at 55°C and a fourth time at room temperature, this was followed by a 2.5x SSC

wash and a final wash in PBS for 5 min. Next, the sections were incubated using the Roche Digoxigenin wash and block buffer set (#11585762001, Roche Applied Science, Mannheim, Germany) for 1 h. This was followed by a 1 h incubation with an anti-Digoxigenin-antibody (#11093274910; Roche Applied Science, Mannheim, Germany) diluted in the Digoxigenin wash and block buffer set at 1:500. After, the tissue sections were washed three times for 5 min with TBS-T. In order to visualise the probes, an NBT/BCIP AP solution (#11697471001; Roche Applied Science, Mannheim, Germany) with the addition of levamisole (Vector Laboratories, SP-500) was added to each section and left at room temperature until detection occurred (approx. 2-3 h). After detection, the sections were rinsed in DEPC-treated water and counterstained with nuclear fast red. This was followed by sequential 20 second washes of 70%, 96% and 100%, followed by 2 min in xylene. Finally, the sections were mounted using glass coverslips and Pertex mounting media.

Name	Sequence
503HG__ISH	/5DigN/CGGATGGCGCGGGCTTGGT/3Dig_N/

Table 2.5: *In Situ* **RNA Detection.**

List of MIR503HG detection *in situ* probe sequence.

2.13 Flow Cytometry

Single-cell suspensions were first in PBS containing 4% (w/v) paraformaldehyde for 5 min and the permeabilised in PBS with 0.2% (w/v) Triton X-100 for a further 5 min. This was followed by incubation in PBS/1% BSA containing anti-CD31 (FITC) (#555445, BD Biosciences), anti-αSMA (APC) (#IC1420A, R&D systems) and anti-SLUG/SNAI2 (PE) antibodies (#IC7408P, R&D systems), alongside unstained and single stain controls for 1 h at 4°C. Cells were then centrifuged at 2000 rpm for 5 min

and resuspended in PBS/1% BSA three times. Resuspended samples were then analysed using the BD LSR5 Fortessa Analytic Flow Cytometer (BD Biosciences). Cells were first separated from debris by plotting forward versus side scatter, single cells were then selected using forward scatter area (FSC-A) versus forward scatter height (FSC-H) and side scatter area (SSC-A) versus side scatter height (SSC-H). From the single cell population, viable cells were selected by gating for DAPI negative events. Marker expression was then determined from this gated population compared to unstained and single stain controls. Representative histograms and mean fluorescence intensity (MFI) calculations were generated using FlowJo v10 (BD Biosciences).

2.14 Total RNA Isolation and Quantification

Total RNA isolation was performed using QIAzol Lysis Reagent and miRNEasy Mini Kit (Qiagen, Hilden, Germany), following the manufacturer's protocol. An additional in-column DNase digestion step was added to remove any further traces of genomic DNA. RNA quantification and quality control were conducted using a NanoDrop ND-1000 Spectrophotometer (Nano-Drop Technologies, Wilmington, DE). All RNA samples were stored at -80°C until required.

2.14.1 Evaluation of RNA Quality

RNA quality was evaluated by RNA integrity number (RIN) analysis using an Agilent® 2100 Bioanalyser in conjunction with RNA 6000 Nano LabChip kits (Agilent Technologies, Berkshire, UK). Loaded RNA molecules were separated by automated electrophoresis to produce electropherograms and gel-like images, allowing for the assessment of RNA quality and RIN calculation. RIN values were calculated

based on the size and distribution of RNA particles within each sample. This technique was used to evaluate the quality of RNA samples prior to RNA-seq, with RIN values of 8 and above being deemed as high quality.

2.15 RNA Sequencing Library Construction and Analysis

All RNA samples used for RNAseq experiments were obtained in biological triplicates. For the "EndMT" RNAseq, RNAs were obtained from HUVEC and HPAEC (treated with TGF-β2 and IL-1β separately or in combination as well as control untreated cells). Ribosomal-depleted stranded libraries were prepared by Beckman Coulter Genomics using the TruSeq Stranded Total RNA kit with Ribo-Zero Gold. Libraries were sequenced with HiSeq Illumina at an average of 50 million reads per samples (paired end 2x125 bp). For "MIR503HG overexpression" RNAseq, polyA stranded libraries were prepared and sequences on HiSeq Illumina at an average of 20 million reads per samples (paired end 2x150bp).For both RNAseq analysis, mapping was performed on the human genome reference sequence GRCh38 (GENCODE Release 26 primary assembly), using STAR (version 2.5.1b) [428]. Gene quantification (read count and normalised expression value as FPKM) was obtained using RSEM [429] (options: -bowtie2 -forward-prob 0 -paired-end), based on GENCODE annotation version 26 (primary assembly). For the "EndMT" RNAseq, the differential expression was assessed using DESeq2 [430] by comparing treated conditions with the untreated control cells. For the RNAseq of MIR503HG overexpression samples, the differential expression analysis was done using edgeR, While DESeq2 is more specific, edgeR [431] has a better sensitivity [432] and is therefore more appropriate for samples that can present coverage variability. We considered a threshold of absolute fold Change >= 2

and adjusted p-value<0.01 to identify significant changes between two conditions. We also applied an expression value threshold of 1 FPKM (average of the three replicates) in the considered groups. Sample clustering was evaluated using the Principal component analysis (PCA) tool available in DESeq2 on the regularised log transformed data. The 3D plot was obtained using the rgl R package. The gene ontology analysis was done using topGO on enriched genes over a background of expressed genes (FPKM>2 in at least one condition).

2.16 Quantitative real time PCR for gene expression

2.16.1 Reverse Transcription Polymerase Chain Reaction

For gene expression analysis, cDNA was synthesised from total RNA (400ng per reaction) using the MultiScribe™ Reverse Transcriptase kit (Life Technologies, Paisley, UK). Synthesis was performed by subjecting each sample to 10 min at 25°C for primer annealing, 30 min at 48°C to allow for reverse transcription and 5 min at 95°C for reverse transcriptase inactivation. After synthesis, all samples were stored at -20°C until required. Target dependant, quantitative real-time polymerase chain reaction (qRT-PCR) was later performed using either Power SYBR green (Life Technologies, Paisley, UK) or TaqMan® (Thermo Fisher, Paisley, UK) gene expression assays.

2.16.2 SYBR Green qRT-PCR Analysis of Gene Expression

SYBR Green based qRT-PCR analysis was conducted utilising Power SYBR Green PCR Master Mix (Thermo Fisher, Paisley, UK) and custom PCR primers (Eurofins MWG, Ebersberg, Germany). Each custom primer pair was designed to target unique,

transcript specific, sequences of MIR503HG. Additionally, a primer pair targeting Ubiquitin protein C (UBC) was used as an endogenous control in order to allow for normalisation of changes in gene expression. See Table 2.6 for a list of all primers used. Further, each reaction was conducted in triplicate using a QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher, Paisley, UK), subject to 10 min at 95°C followed by 40 cycles of 15 seconds at 95°C and 60 seconds at 60°C for primer annealing and extension. Primer specificity was confirmed both by melting curve analysis and gel electrophoresis.

Primer Name	Gene Name	Sequence
UBC_Fw	UBC	5' TTGCCTTGACATTCTCGATG 3'
UBC_Rv	UBC	5' ATCGCTGTGATCGTCACTTG 3'
503HG_1_Fw	MIR503HG_001	5' CCACAGGACAACGAAGAAAACC 3'
503HG_1_Rv	MIR503HG_001	5' TACTCCTTTTCCAGTCCTCCCC 3'
503HG_2_Fw	MIR503HG_002	5' GTGGAACCCCACACAGGAAA 3'
503HG_2_Rv	MIR503HG_002	5' GGACAGTTGCCCATATTAACGG 3'
503HG_3_Fw	MIR503HG_003	5' GGACCTGAGCTGTCGATTTCA 3'
503HG_3_Rv	MIR503HG_003	5' GAGGGATGGAGGTGGCTTTA 3'
Gm28730_Fw	Gm28730	5' CCCCAAATCTAGGCTCCTTTGT 3'
Gm28730_Rv	Gm28730	5' ATCAGGACTGACTCATTTGGTGG 3'
18S_Fw	18S	5' GGCCCTGTAATTGGAATGAGTC 3'
18S_Rv	18S	5' CCAAGATCCAACTACGAGCTT 3'
NEAT1_Fw	NEAT1	5' CTTCCTCCCTTTAACTTATCCATTCAC 3'
NEAT1_Rv	NEAT1	5' CTCTTCCTCCACCATTACCAACAATAC 3'
CDC25A_Fw	CDC25A	5' TAAGACCTGTATCTCGTGGCTG 3'
CDC25A_Rv	CDC25A	5' CCCTGGTTCACTGCTATCTCT 3'

Table 2.6: SYBR Green qRT-PCR Analysis of Gene Expression.
List of MIR503HG transcript-specific and housekeeper primers used for SYBR Green-based qRT-PCR analysis.

2.16.3 Taqman® qRT-PCR Analysis of Gene Expression

TaqMan® qRT-PCR was performed using the TaqMan® Universal Master Mix II and TaqMan® Gene Expression probes following the manufacturer's protocol (Thermo

Fisher, Paisley, UK) (Table 2.7 and Table 2.8). Further, each reaction was conducted in triplicate using a QuantStudio 7 Flex Real-Time PCR System real time PCR system (Thermo Fisher, Paisley, UK) , subject to 10 min at 95°C followed by 40 cycles of 15 seconds at 95°C and 60 seconds at 60°C for primer annealing and extension.

2.16.4 MicroRNA Reverse Transcription

For microRNA analysis, cDNA was generated from total RNA (5ng per reaction) using the TaqMan® miRNA Reverse Transcription kit and specific TaqMan® miRNA reverse transcription primers (Thermo Fisher, Paisley, UK) (Table 2.7 and Table 2.8). Synthesis was performed by subjecting each sample to 30 min at 16°C for primer annealing, 30 min at 42°C to allow for reverse transcription and 5 min at 85°C for reverse transcriptase inactivation. After synthesis, all samples were stored at -20°C until required. Note that, RNU48 was selected as an endogenous control miRNA to allow for normalisation of changes in miRNA expression.

Gene ID	Assay ID		microRNA ID	Assay ID
UBC	Hs01871556_s1		RNU48	001006
PECAM1	Hs01065282_m1		miR-424-5p	000604
CDH5	Hs00901465_m1		miR-424-3p	002309
SNAI1	Hs00195591_m1		miR-450a-5p	002303
SNAI2	Hs00161904_m1		miR-450a-1-3p	469835
S100A4	Hs00243202_m1		miR-450b-5p	002207
ACTA2	Hs00426835_g1		miR-450b-3p	002208
COL1A1	Hs00164004_m1		miR-503-5p	001048
COL3A1	Hs00943809_m1		miR-503-3p	476380
VIM	Hs00958111_m1		miR-542-5p	002240
FGF2	Hs00266645_m1		miR-542-3p	001284
FGFR1	Hs00241111_m1			

Table 2.7: Taqman qRT-PCR Analysis of Gene Expression in Humans.
List of gene and microRNA TaqMan probes used for qRT-PCR analysis

Gene ID	Assay ID
18S	Mm03928990_g1
PECAM1	Mm01242576_m1
CDH5	Mm00486938_m1
NOS3	Mm00435217_m1
SNAI2	Mm00441531_m1
ACTA2	Mm00725412_s1
COL1A1	Mm00801666_g1
VIM	Mm01333430_m1

microRNA ID	Assay ID
U6	
miR-322-5p	001076
miR-503-5p	002456

Table 2.8: Taqman qRT-PCR Analysis of Gene Expression in Mice.

List of gene and microRNA TaqMan probes used for qRT-PCR analysis

2.16.5 MicroRNA qRT-PCR

MicroRNA qRT-PCR was performed using specific TaqMan® miRNA RT-PCR probes (Table 2.7 and Table 2.8), and TaqMan® Universal Master Mix II following the manufacturer's protocol (Thermo Fisher, Paisley, UK).

As with previous assays, each reaction was conducted in triplicate using a QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher, Paisley, UK), subject to 10 min at 95°C followed by 40 cycles of 15 seconds at 95°C and 60 seconds at 60°C for primer annealing and extension.

2.16.6 qRT-PCR Data Analysis

All qRT-PCR data obtained (SYBR, mRNA and microRNA) was analysed using the 2-ΔΔCt method, as described by Livak & Schmittgen. 2-ΔΔCt values were used to calculate the fold change (described as FC or RQ) in gene expression between the experimental and control groups.

2.17 Endothelial barrier integrity assay

The Electric Cell-substrate Impedance Sensing (ECIS) assay was implemented using 8-well 8W1E arrays via a ECIS® Z-Theta station (Applied Biophysics), as per the manufacturer's instructions. All arrays were treated using 10 mM L-cysteine (C7352-25G, Sigma-Aldrich) and coated with Collagen Type I (C3867-1VL, Sigma-Aldrich). A total of 4×10^4 HUVEC were seeded into each well of the array and allowed to adhere overnight. The assay was then conducted using a monitoring frequency of 4 kHz for a total of 5 h. All impedance measurements were analysed using ECIS mathematical modelling software (Applied Biophysics) to determine cell barrier resistance (Rb), expressed as Ohm x cm^2 [433,434]. All data was analysed using GraphPad Prism 8 and presented as mean ± SEM.

2.18 Transwell Migration Assay

The migration assay was carried out in 24-well migration chambers with polycarbonate 8-μm pore membrane filters (#CLS3422; Corning, Poole, UK). Prior to use, each chamber was coated with a 10ug/ml fibronectin in 0.1% Gelatine mixture and allowed to dry for 2 h. A total of 3×10^4 HUVEC were seeded with 200 μl 0.1% BSA basal medium in the upper layer of each transwell chamber, a total of 400 μl of either 10% FBS basal medium or 0.1% BSA basal medium was also added to the lower portion of each well. The cells were allowed to incubate at 37°C with 5% CO_2 for a total of 6 h, the migrated cells then fixed using methanol and stained with DAPI. Images of migrated cells were obtained by fluorescence microscopy (Zeiss AxioSkop Upright Fluorescence Microscope) and counted in 5 random fields (x10) of each well.

2.19 EdU Proliferation Assay and Cell Cycle Analysis

To measure HUVECs proliferation Click-iT® EdU Flow Cytometry Cell Proliferation Assay (Invitrogen) was used according to manufacturer's instructions. 24 h prior harvest cells were counted and re-plated in equal numbers. Cells were incubated with 10 uM EdU for 4 h alongside an unstained negative control, detached using trypsin and labelled with live/dead dye according to manufacturer's recommendations (Zombie NIR™ Fixable Viability Kit (Biolegend, 1:500) before being fixed and permeabilised. EdU incorporation was quantified using Click-it chemistry conjugated with an Alexa Fluor 488 antibody according to the manufacturer's protocol. To separate the cell cycle phases cells were incubated with 2.5 ug/ml DAPI briefly before acquiring the samples on BD LSRFortessa™ cell analyser.

2.20 Statistical Analysis

All data given was shown as mean ± standard error of the mean (SEM) on biological or technical replicates as detailed in the figure legends. All biological replicates using human primary cells correspond to independent experiments from distinct expansions and passage numbers. All experimental data was analysed by paired two-tailed t-test, one-way analysis of variance (ANOVA) or two-way ANOVA with a Bonferroni multiple comparisons test, as indicated. A P value of less than 0.05 was considered statistically significant. Statistical analysis was performed using ΔCt values, which reflect the difference in cycle threshold (Ct) between the target gene and appropriate housekeeping gene.

Chapter 3: Identifying the Function of MIR503HG in EndMT

3.1 Introduction

Whilst EndMT can be activated by several signalling pathways, its precise transcriptional profile and molecular mechanisms are yet to be rigorously defined. This is particularly true with the discovery of lncRNA as multifaceted regulators of cell function, both at the transcriptional and post-transcriptional level [286]. Nonetheless, the role lncRNAs in EndMT is yet to be thoroughly explored and relatively few have been reported in the regulation of EndMT so far. The lncRNA MALAT1, for example, was the first to be implicated in the modulation of TGF-β-induced EndMT, still this was show to happen through the regulation of miR-145 on the TGFBR2/SMAD3 axis [397]. Recently, the lncRNA GATA6-AS was shown to suppress TGF-β-induced EndMT *in vitro* via LOXL2 to regulate chromatin remodelling. Although this interaction may be present, the study relied largely on vein ECs and results may not be applicable to EndMT in other ECs [398].

Ultimately, while these studies represent a step forward in the field, more comprehensive reports across multiple endothelial phenotypes are needed. Our group used high-throughput RNA sequencing (RNAseq) of HUVEC and HPAEC undergoing EndMT *in vitro*. The project aimed to investigate the common transcriptional architecture of EndMT across ECs originated from different vascular beds and with it identify previously unreported lncRNAs (data presented as part of Section 3.3). Initial bioinformatics analysis revealed 103 differentially expressed lncRNAs during EndMT in both HUVEC and HPAEC. Among them, the significant loss of the lncRNA MIR503HG was found to be a consistent event during the initial stages of EndMT. The present chapter will expand on these initial results, describe the MIR503HG locus and explore is function and mechanism.

3.1.1 The MIR503HG locus
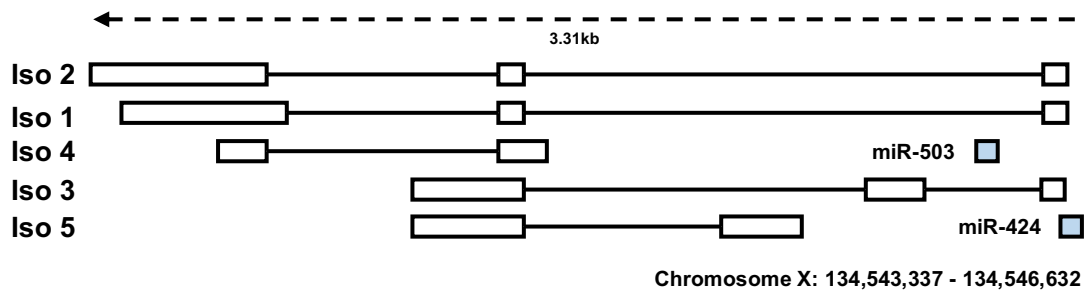
Figure 3.1: MIR503HG locus.

Schematic representation of the MIR503HG, miRNA-424 and miRNA-503 locus (ChrX q26.3) (Ensembl Genome Browser (GRCh38.p10).

MIR503HG is an intergenic lncRNA located on chromosome X (Xq26) with five isoforms reported in GENCODE (Figure 3.1). First described in 2014 solely as a miRNA precursor, the lncRNA locus was found to be highly conserved across 11 tetrapod species (human, chimpanzee, bonobo, gorilla, orangutan, macaque, mouse, opossum, platypus, chicken and frog) and believed to originate from a common tetrapod ancestor [435]. Since this first report, a wealth of studies have started to emerge showing that the lncRNA not only had its own independent transcriptional pattern but also a functional role in both epithelial and endothelial cells [381,436], making it an attractive candidate.

3.1.2 Role of MIR503HG in cell proliferation and migration

Early reports have described MIR503HG as an hypoxia-sensitive non-coding RNA with a potential role in EC function. Gain- and loss-of-function *in vitro* studies by Fiedler and colleagues were the first to show that MIR503HG-deficient ECs exhibited reduced proliferative and migratory capacity, showing potential pro-angiogenic

effects. Similarly, silencing MIR503HG in a engineered heart tissue *ex vivo* model resulted in impaired endothelial biology and tissue vascularisation [381].

The potential of these pro-angiogenic regulatory features led subsequent studies to look at impact of MIR53HG in tumorigenesis and tumour progression. However, unlike earlier EC reports, siRNA-mediated knockdown of MIR503HG was found to significantly enhance migration and invasion in human hepatocarcinoma cell (HCC) lines, while its overexpression had the opposite effect both *in vitro* and *in vivo* [437]. Similar effects were also observed in choriocarcinoma [438], triple-negative breast cancer [439] and human colorectal cancer cell lines [440].

Conflicting reports were also seen regarding the lncRNAs effect on cell proliferation. While MIR503HG knockdown significantly enhanced trophoblast cell proliferation, this was supressed in anaplastic lymphoma kinase (ALK)-negative tumour cells [441,442]. Ultimately, these results highlight cell-specific and context-dependent mechanisms of MIR503HG, a common feature of several lncRNA [396].

3.1.3   MIR503HG as a regulator of EMT and EndMT

Despite the reported differences in MIR503HG function across cell types, the lncRNA has often been associated with cell migration and proliferation [381,437,441,442], all key features of EndMT [61]. Recently, MIR503HG has also been associated to epithelial-mesenchymal transition with overexpression of the lncRNA increasing the availability of E-cadherin and decreasing the expression and translation of several transition-associated genes, such as ZEB1, SNAI1, N-cadherin, and vimentin [436,441]. This, along with our preliminary results showing a remarkable EndMT *in vitro* phenotype after MIR503HG depletion, warrant further functional analysis.

3.2 Aims

Previous RNAseq analysis by our group showed a marked downregulation of MIR503HG in both HUVEC and HPAEC during EndMT, which warranted further research. As such, the aims of this chapter were as follows:

- To replicate a baseline EndMT *in vitro* model and delineate key markers;

- To validate the expression profile of MIR503HG during EndMT *in vitro*;

- Confirm differential MIR503HG expression in multiple EC phenotypes;

- To identify the function of MIR503HG in ECs;

- To identify the molecular mechanism governing MIR503HG function.

3.3 Results

3.3.1 EndMT *in vitro* model validation

In order to replicate EndMT *in vitro* and delineate its key molecular markers, we used a 7 day TGF-β2 and IL-1β co-stimulation model developed by our lab based on previous publications [81]. This served to guide and interpret the results presented throughout this thesis.

First, we performed gene expression analysis of HUVEC and HPAEC co-stimulated with a daily dose of TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) for a total of 7 days. In line with previous validation work performed by our lab (unpublished), this protocol induced a significant increase in the expression of SNAI2, ACTA2 and COL1A1, accompanied by a marked repression of PECAM1 (Figure 3.2). These changes were confirmed by flow cytometry analysis of co-stimulated cells again showing a similar decrease in PECAM1 expression, accompanied by an increase in ACTA2 and SNAI2, compared to an untreated control (Figure 3.3). Crucially, only TGF-β2 and IL-1β co-stimulation was able to induce significant changes, suggesting that both cytokines are essential to trigger the appearance of an EndMT-like phenotype.

Consistent with our results, immunofluorescence microscopy showed that the loss of intercellular junctions was accompanied disruption of the endothelial monolayer, along with an increase in mesenchymal marker expression and the acquisition of a fibroblast-like phenotype (Figure 3.4). Together, these findings demonstrate that the *in vitro* model used here can induce a robust EndMT profile in both arterial and venous ECs.
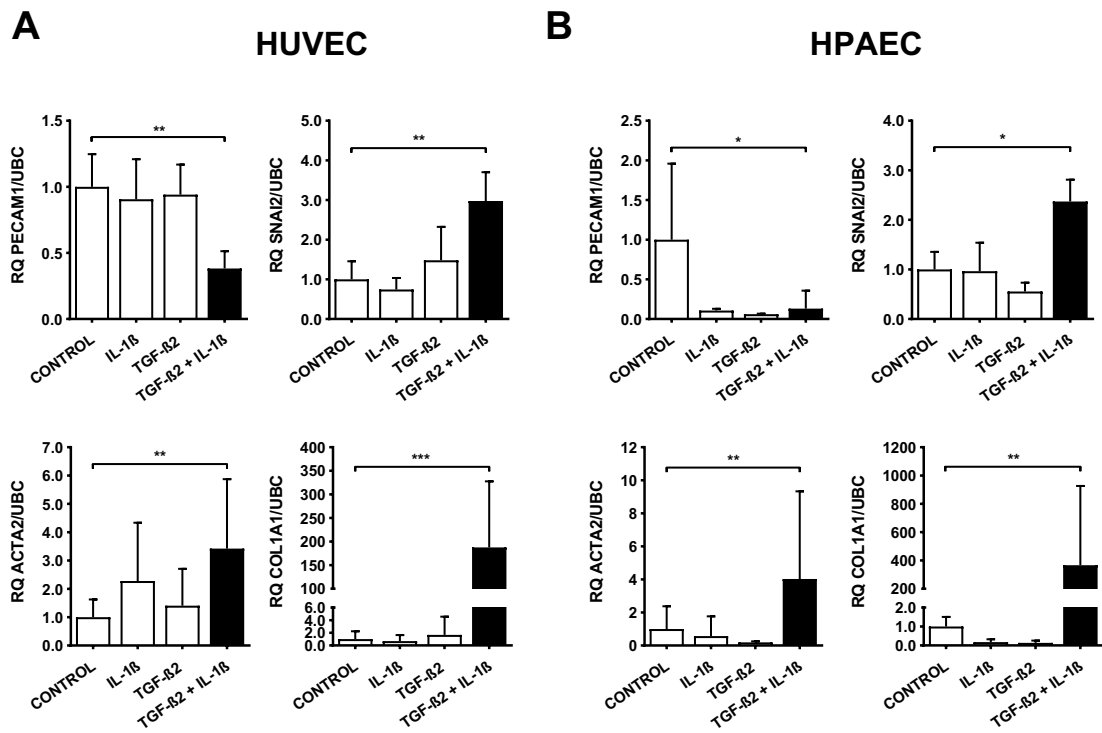
Figure 3.2: EndMT *in vitro* model in arterial and venous ECs.

ECs were treated with TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) in combination or alone for 7 days. Change in expression of EndMT markers in **(A)** HUVEC and **(B)** HPAEC expressed as RQ. RQ values for gene expression were quantified by qRT-PCR assay relative to UBC (n=3 biological replicates). Analysis by one-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001. Data represented as mean ± SEM.
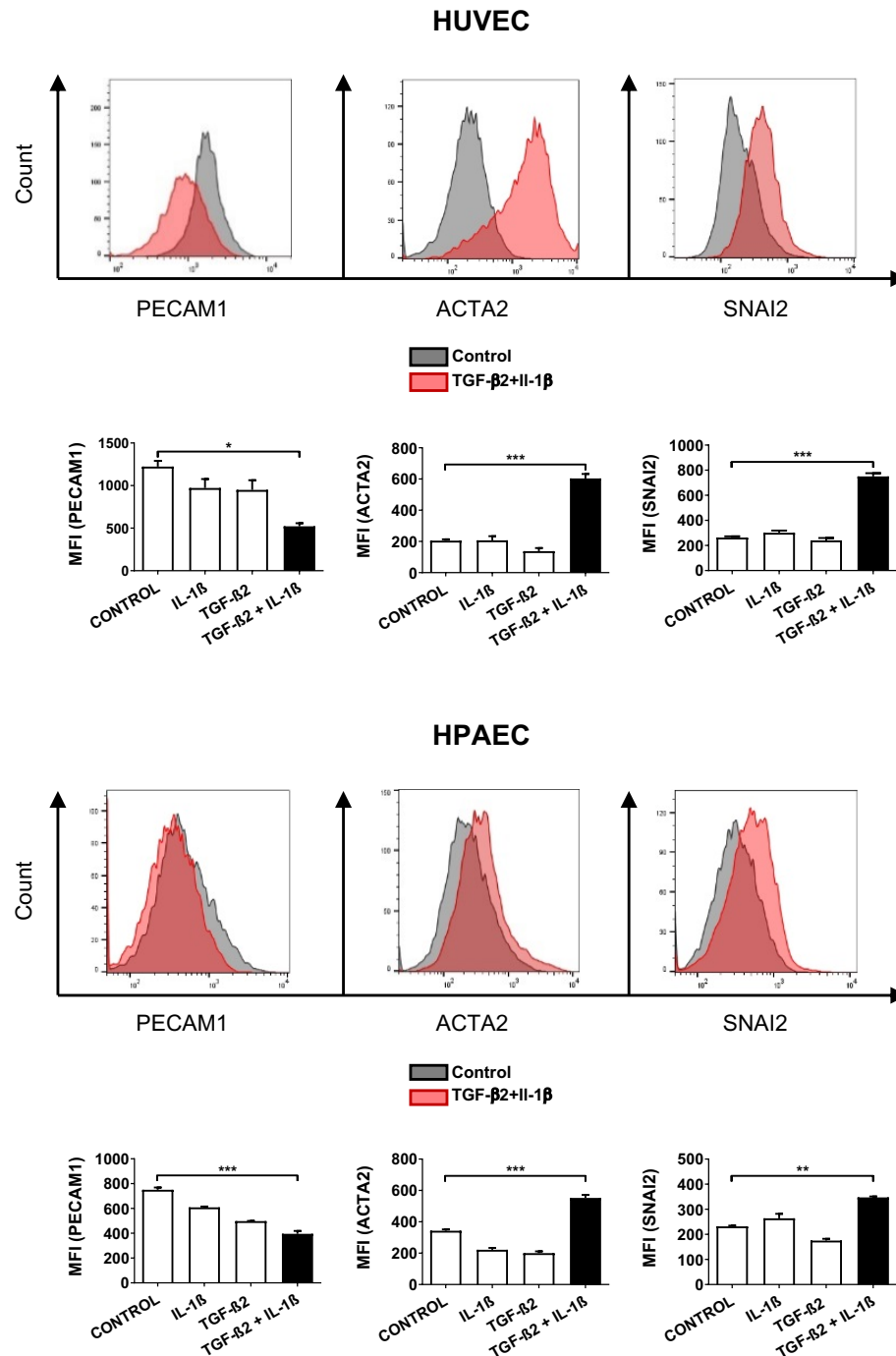
Figure 3.3: EndMT *in vitro* model in arterial and venous EC.

EndMT marker protein expression changes analysed by flow cytometry in HUVEC and HPAEC following TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) treatment in combination or alone for 7 days quantified by mean fluorescence intensity (MFI). Data presented with representative histograms of untreated control cells (grey) compared to co-treated cells (red) (n=3 biological replicates). Analysis by one-way ANOVA; *$p \leq 0.05$, ** $p \leq 0.01$, and *** $p \leq 0.001$. Data represented as mean ±SEM. This work was performed in collaboration with Dr Axelle Caudrillier.
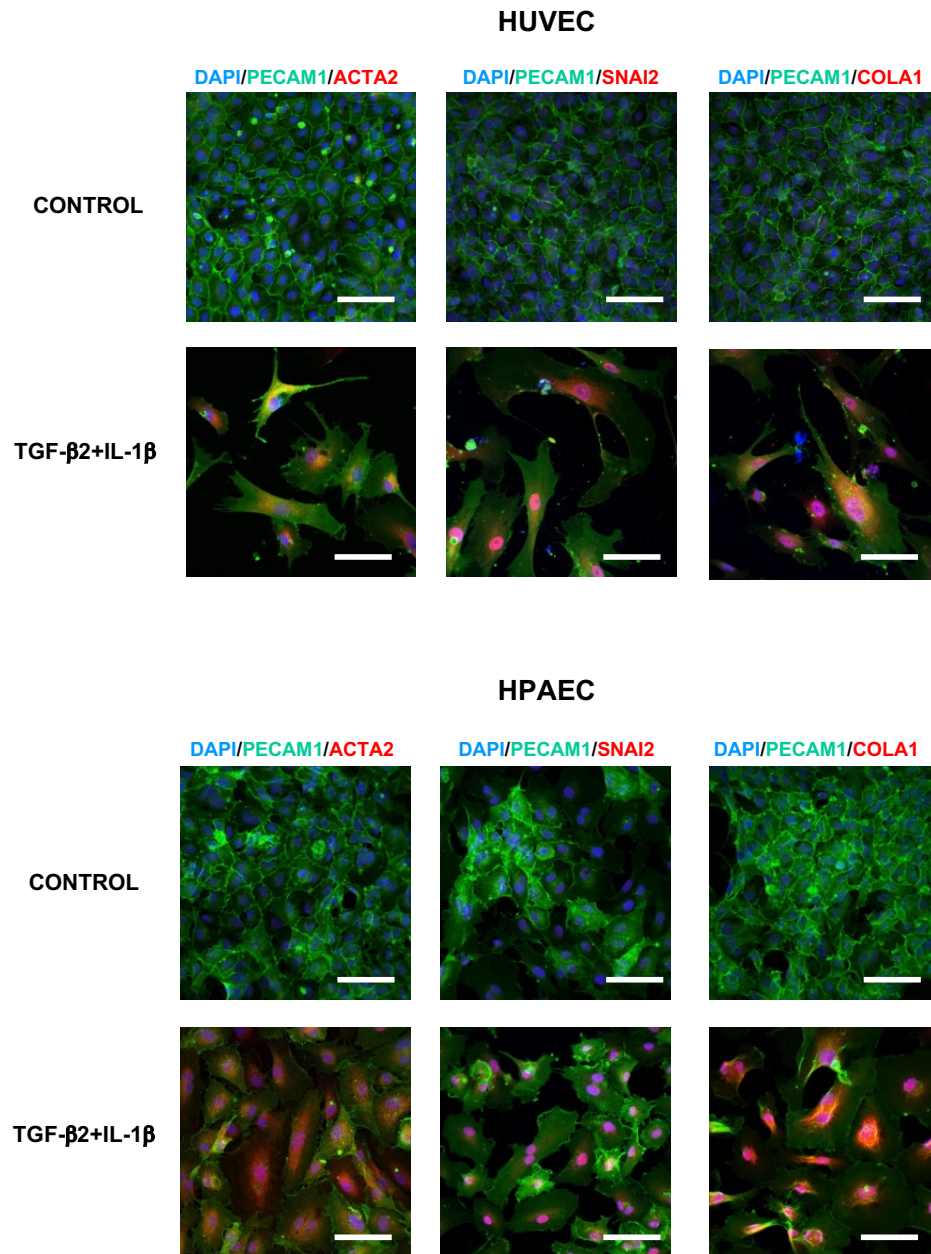
HUVEC

Figure 3.4: EndMT *in vitro* model in arterial and venous EC.
Representative immunofluorescence images of EndMT marker expression in HUVEC and HPAEC following TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) treatment in combination or alone for 7 days (20X, scale bar 50 μm). PECAM1 in green, ACTA2/COL1A1/SNAI2 in red and DAPI (nucleus) in blue. Images acquired with Andor Revolution XDi spinning disk confocal microscope and analysed with Image J Software. This work was performed in collaboration with Dr Axelle Caudrillier.

3.3.2 Identifying the transcriptional profile of EndMT

In order to further characterise our *in vitro* EndMT model and identify key transcriptional changes, we performed high-throughput RNAseq of both HUVEC and HPAEC undergoing transition. Principal component analysis (PCA) of the differentially expressed transcriptome revealed that both treatment groups clustered separately from their respective untreated control (Figure 3.5 A). Interestingly, both HUVEC and HPAEC treatment groups also clustered separately from each other, clearly suggesting transcriptional heterogeneity in the induction of EndMT across different vascular beds (Figure 3.5 A). Despite the heterogenous EndMT profiles, our analysis revealed a pool of expressed genes regulated in both transitioning HUVEC and HPAEC. More precisely, 606 genes were up-regulated in both groups, while 583 were down-regulated (Figure 3.5 B-C).

Within this shared transcriptional profile, we have further identified a group of known lncRNAs whose expression was altered during EndMT. A total of 69 lncRNAs were up-regulated and 34 down-regulated across both different vascular bed ECs. Of these, based on fold change (FC), FPKM levels , genomic location, as well as available publications, we selected 9 top candidates: MIR503HG, HOTAIRM1, AC123023.1, CTC-378h22.1, LINC00702, MIR3142HG, RP-37B2.1, AC147651.4 and RP11-79h23.3 (Figure 3.5 D-F). Despite their individual merits, for the purpose of this thesis, only MIR503HG was selected for further validation and downstream analysis.
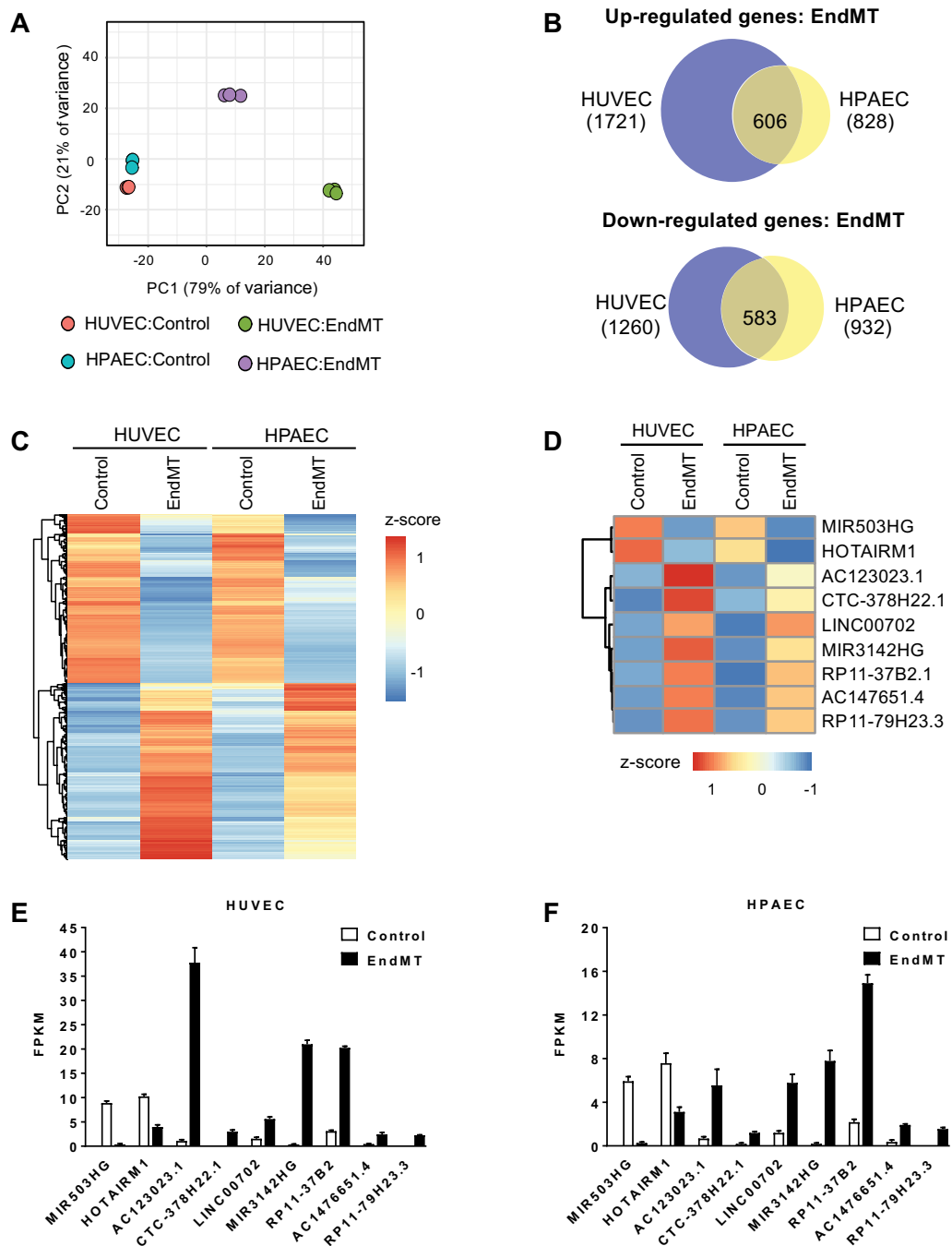
Figure 3.5: RNAseq analysis of EndMT identifying a common lncRNA signature.
(A) PCA plot of Control and EndMT HUVEC and HPAEC RNAseq samples. **(B)** Venn diagram of the overlap between up-regulated genes and down-regulated genes during EndMT in HUVEC and HPAEC. **(C)** Heatmap showing the expression data (as z-score) of differentially expressed genes, including lncRNAs. **(D)** List of selected top candidate lncRNAs for downstream analysis. Representative expression of candidate lncRNAs in **(E)** HUVEC and **(F)** HPAEC ± EndMT treatment. Expression values shown as FPKM. Data represented as mean ±SEM. RNAseq sample preparation was carried out by Dr Axelle Caudrillier and bioinformatic analysis by Dr Julie Rodor.

131

3.3.3 Validation of MIR503HG expression profile during EndMT

RNA sequencing data from our EndMT *in vitro* model showed a clear downregulation of all MIR503HG isoforms. However, only those reaching a minimum FPKM value above 1 (isoforms 2, 3 and 5) were selected for further qRT-PCR validation (Figure 3.6). As with the RNAseq dataset, all selected isoforms exhibited a significant downregulation in both HUVEC and HPAEC at day 7 after EndMT co-treatment (Figure 3.7).

To further assess the broad relevance of MIR503HG down-regulation in EndMT, we extended our analysis to ECs from different vascular beds, including Human Saphenous Vein Endothelial Cells (HSVEC) and Human Coronary Artery Endothelial Cells (HCAEC), in addition to HUVEC and HPAEC. Down-regulation of MIR503HG after TGF- β2 and IL-1β co-treatment was found to be consistent across all studied vascular beds (Figure 3.8).

Additionally, we also found decreased MIR503HG expression in a previously published alternative *in vitro* model of EndMT, consisting of TGF- β2 and H_2O_2 co-stimulation [53] (Figure 3.9). Ultimately, our results show that independent of endothelial origin or *in vitro* model used, the loss of MIR503HG expression is strongly associated with the process EndMT.
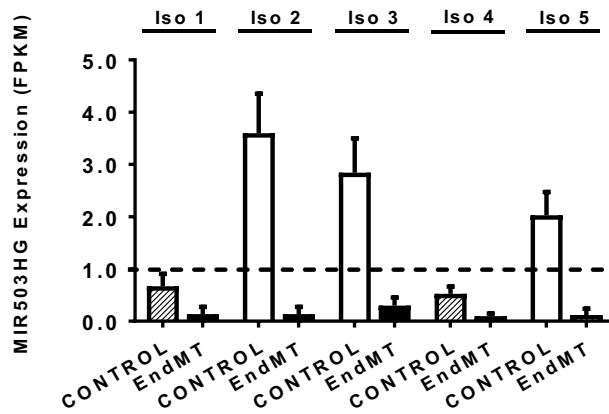
Figure 3.6: MIR503HG expression during EndMT.

Representative expression of MIR503HG by isoform in HUVEC after ± EndMT treatment. Expression values shown as FPKM. Data represented as mean ±SEM.
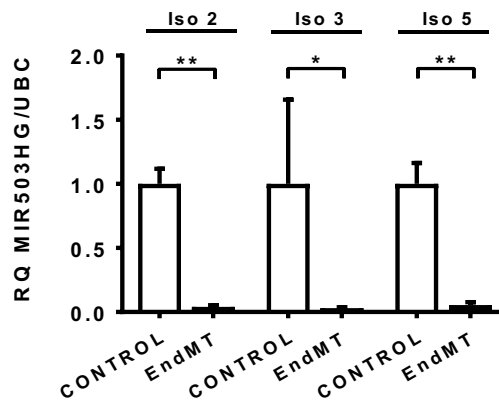
Figure 3.7: Validation of MIR503HG expression during EndMT.

Expression of MIR503HG_2/3/5 in HUVEC ± EndMT treatment (n=4 biological replicates). RQ value for gene expression was quantified by qRT-PCR assay relative to untreated control cells. Analysis by two-tailed t-test; *p≤0.05 and ** p≤0.01 vs paired control. Data represented as mean ±SEM.
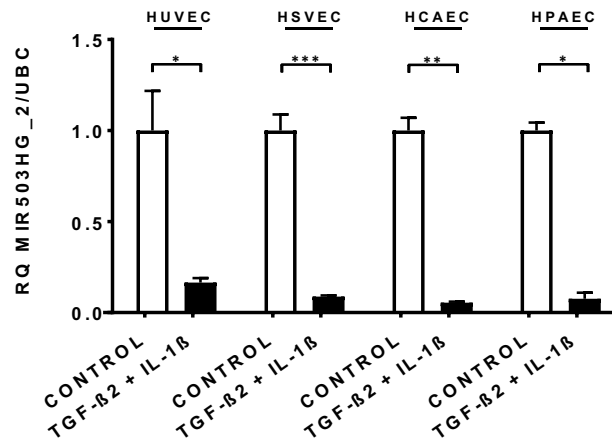
Figure 3.8: MIR503HG expression during EndMT across different EC beds.

Expression of MIR503HG_2 in HUVEC, HSVEC, HCAEC and HPAEC ± EndMT treatment (n=3 biological replicates). RQ value for gene expression was quantified by qRT-PCR assay relative to untreated control cells. Analysis by two-tailed t-test; *p≤0.05 and ** p≤0.01 vs paired control. Data represented as mean ±SEM. This work was performed in collaboration with Dr Axelle Caudrillier and Ms Shiau Haln Chen.
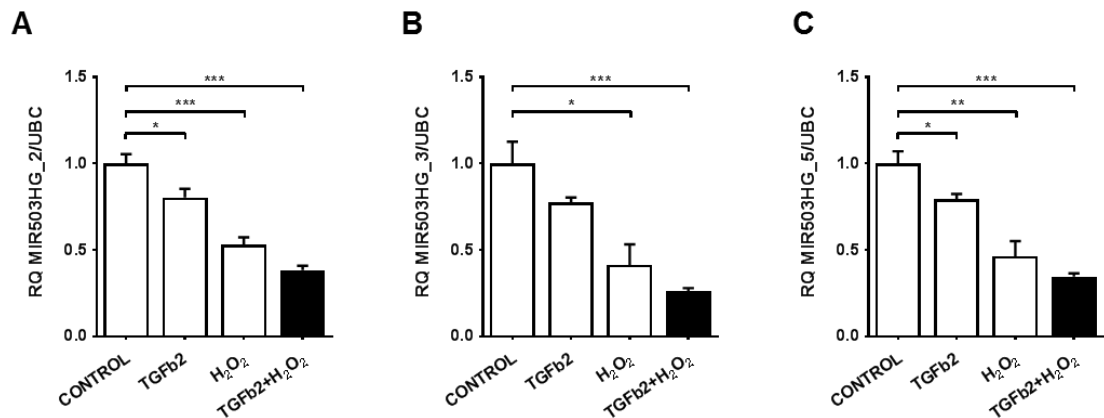
Figure 3.9: TGF-β2 and H_2O_2 *in vitro* model of EndMT.

Expression of **(A)** MIR503HG_2, **(B)** MIR503HG_3 and **(C)** MIR503HG_5 in HUVEC after treatment with TGF-β2 (50 ng/mL) ± H_2O_2 (200 nM) (n=3 biological replicates). RQ values for gene expression were quantified by qRT-PCR assay relative to untreated cells. Analysis by one-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired control. Data represented as mean ±SEM.

3.3.4 Subcellular localisation of MIR503HG

Subcellular localisation is of key importance to understand the putative function and mechanism of action of a lncRNA [314]. Fractionation of EC nuclear and cytoplasmic compartments demonstrated that MIR503HG was enriched in the nucleus under both basal and TGF-β2 and IL-1β co-stimulation conditions (Figure 3.10).
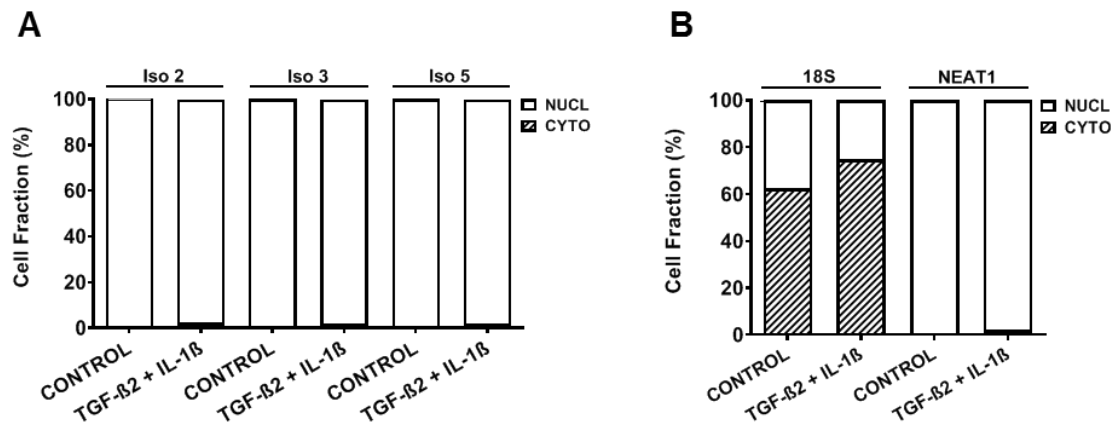
Figure 3.10: MIR503HG Subcellular Localisation

Subcellular localisation of **(A)** MIR503HG_2, MIR503HG_3 and MIR503HG_5, and **(B)** fractionation controls 18S and NEAT in HUVEC ± EndMT treatment. Cell fraction value for gene expression was quantified by qRT-PCR assay relative to UBC (n=3 technical replicates).

3.3.5 MIR503HG *in vitro* knockdown

3.3.5.1 Dicer substrate siRNA (DsiRNA)-based knockdown

To further asses the contribution of MIR503HG to EndMT we used a Dicer substrate siRNA (si503HG) to knockdown the lncRNA in HUVEC. The resulting cellular phenotype was evaluated based on our defined EndMT markers (i.e. PECAM1, ACTA2, SNAI2, COL1A1). The effectiveness of the knockdown was first confirmed in untreated HUVEC 3 days after si503HG transfection, with the expression of all MIR503HG isoforms significantly downregulated (Figure 3.11). Further, MIR503HG

knockdown was found to be sustained for 7 days after transfection and lead to significant changes in EndMT markers in the absence of TGF-β2 and IL-1β co-stimulation (Figure 3.12 A). RT-qPCR analysis of HUVEC transfected with si503HG revealed a decrease of *PECAM1*, along with an increase in *SNAI2, ACTA2* and *COL1A1* at day 7 (Figure 3.12 B). These effects were mirrored at the protein level as demonstrated by flow cytometry analysis (Figure 3.13 A-B). Further, immunofluorescence imaging showed a loss of monolayer formation and clear cell morphology changes, as shown by reduced PECAM1 expression, and the gain of mesenchymal protein expression with SNAI2 and COL1A1 (Figure 3.13 C).
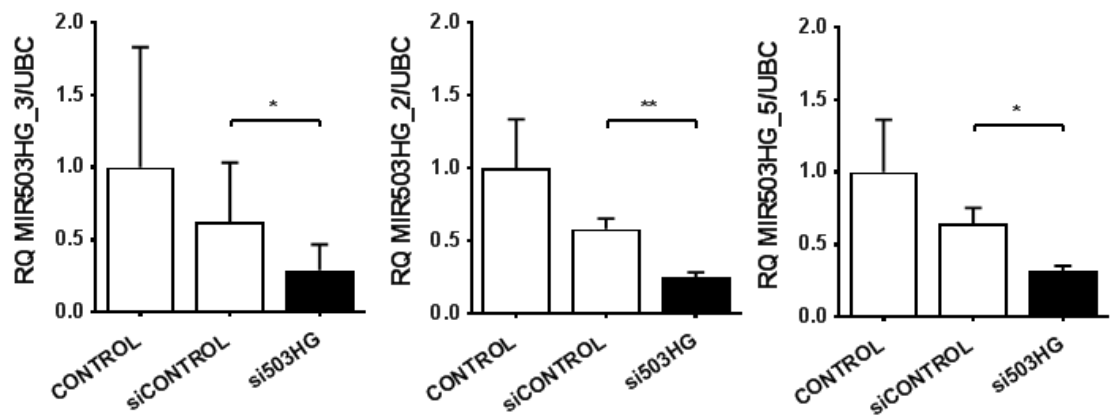
Figure 3.11: MIR503HG siRNA-mediated knockdown at day 3.
Expression of MIR503HG_2, MIR503HG_3 and MIR503HG_5 in HUVEC at day 3 after transfection with si503HG (20 nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=3 technical replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired siControl. Data represented as mean ±SEM.
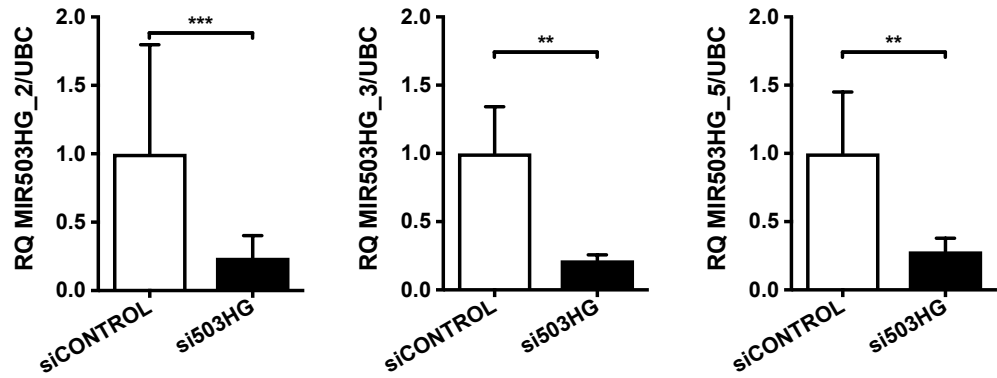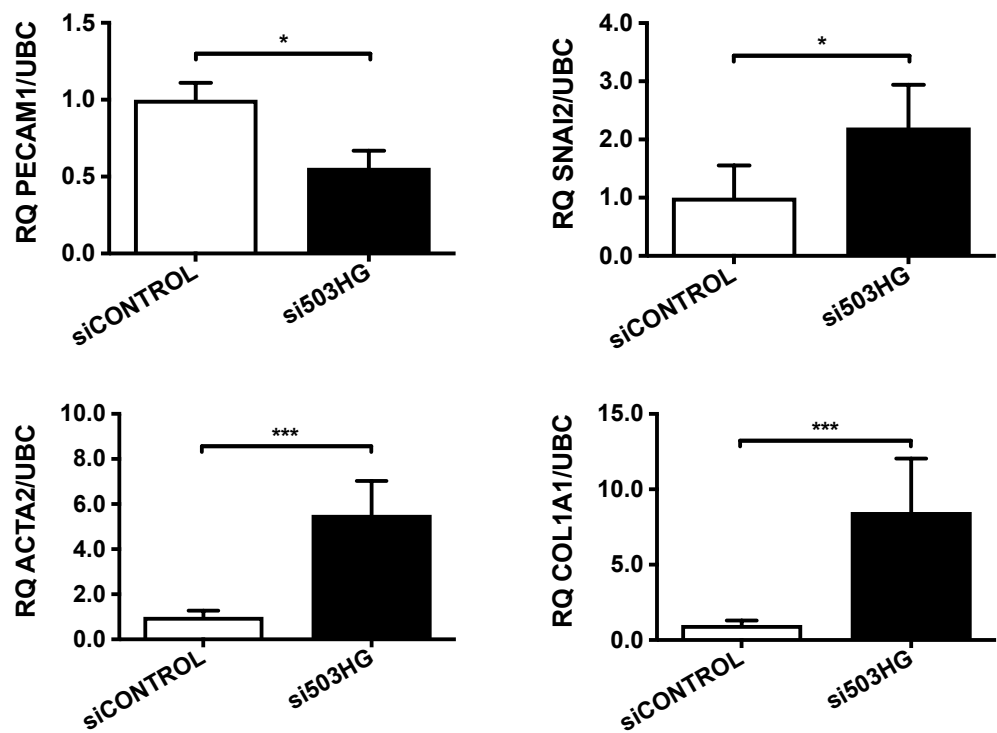
Figure 3.12: MIR503HG siRNA-mediated knockdown induces EndMT.

Expression of **(A)** MIR503HG_2/3/5 and **(B)** EndMT marker genes in HUVEC 7 days after knockdown using si503HG (20 nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=5 biological replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired siControl. Data represented as mean ±SEM.
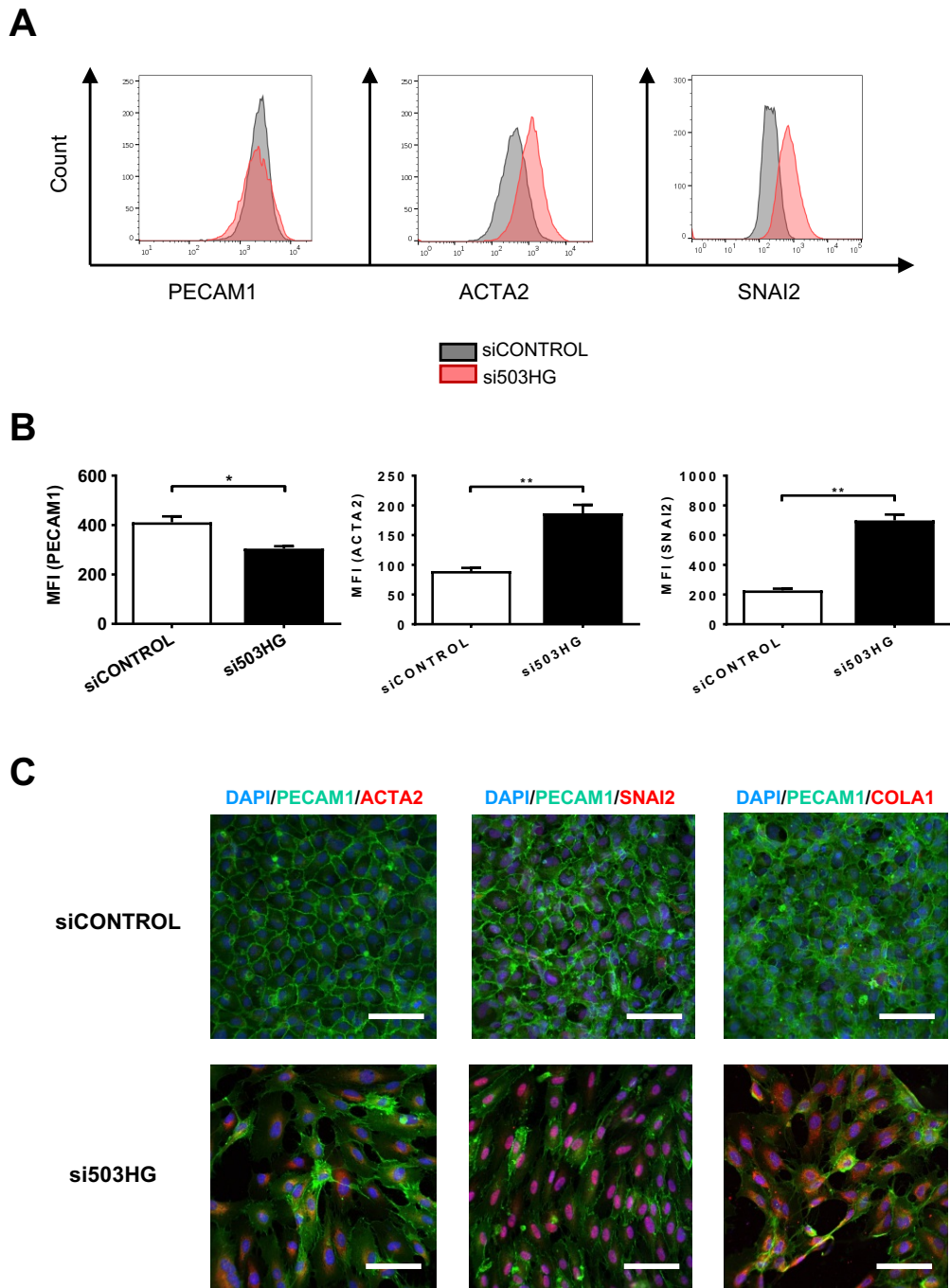
Figure 3.13: MIR503HG siRNA-mediated knockdown induces EndMT.

(A) EndMT marker expression analysed by flow cytometry 7 days after knockdown using si503HG. Representative histograms for controls cells (grey) compared to si503HG-cells (red) and **(B)** MFI quantification (n=3 biological replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired siControl. Data represented as mean ±SEM. **(C)** Representative immunofluorescence images of EndMT markers expression in HUVEC (20X, scale bar 50 μm) 7 days after knockdown using si503HG (20 nM) compared to paired control. PECAM1 in green, ACTA2/COL1A1/SNAI2 in red and DAPI in blue. Images acquired with Andor Revolution XDi spinning disk confocal microscope and analysed using Image J Software. This work was performed in collaboration with Dr Axelle Caudrillier.

3.3.5.2 Antisense GapmerR-based knockdown

Given its nuclear localisation, an antisense GapmeR (gap503HG) was also used to manipulate MIR503HG expression. Knockdown was validated in untreated HUVEC 3 days after transfection, confirming that the expression of all MIR503HG isoforms was significantly downregulated (Figure 3.14). Knockdown using gap503HG was found to induce a robust EndMT profile in the absence of TGF-β2 and IL-1β co-stimulation at 7 days, show by RT-qPCR analysis (Figure 3.15).

Duplicating the effect seen with si503HG, this was mirrored at the protein level as confirmed by both flow cytometry analysis and immunofluorescence imaging (Figure 3.16).
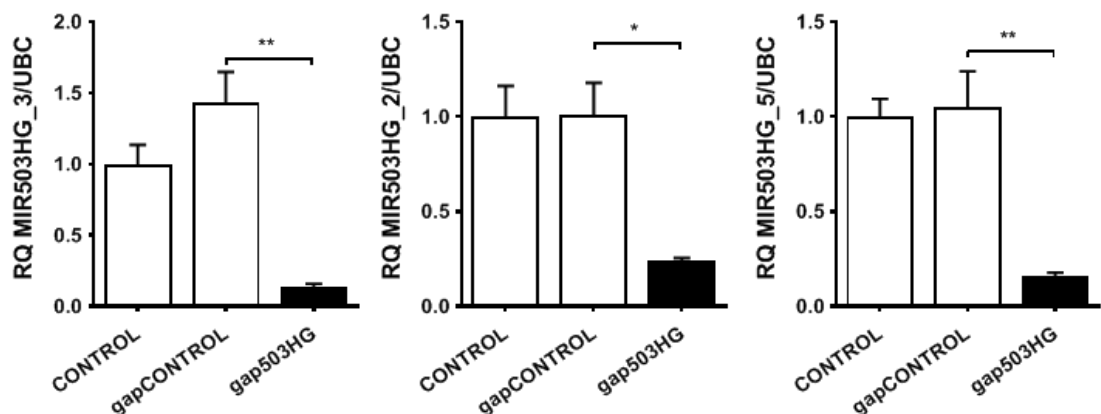
Figure 3.14: MIR503HG gapmer-mediated knockdown at day 3.
Expression of MIR503HG_2, MIR503HG_3 and MIR503HG_5 in HUVEC at day 3 after transfection with gap503HG (20 nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=3 technical replicates). Analysis by two-tailed t-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired gapControl. Data represented as mean ±SEM.
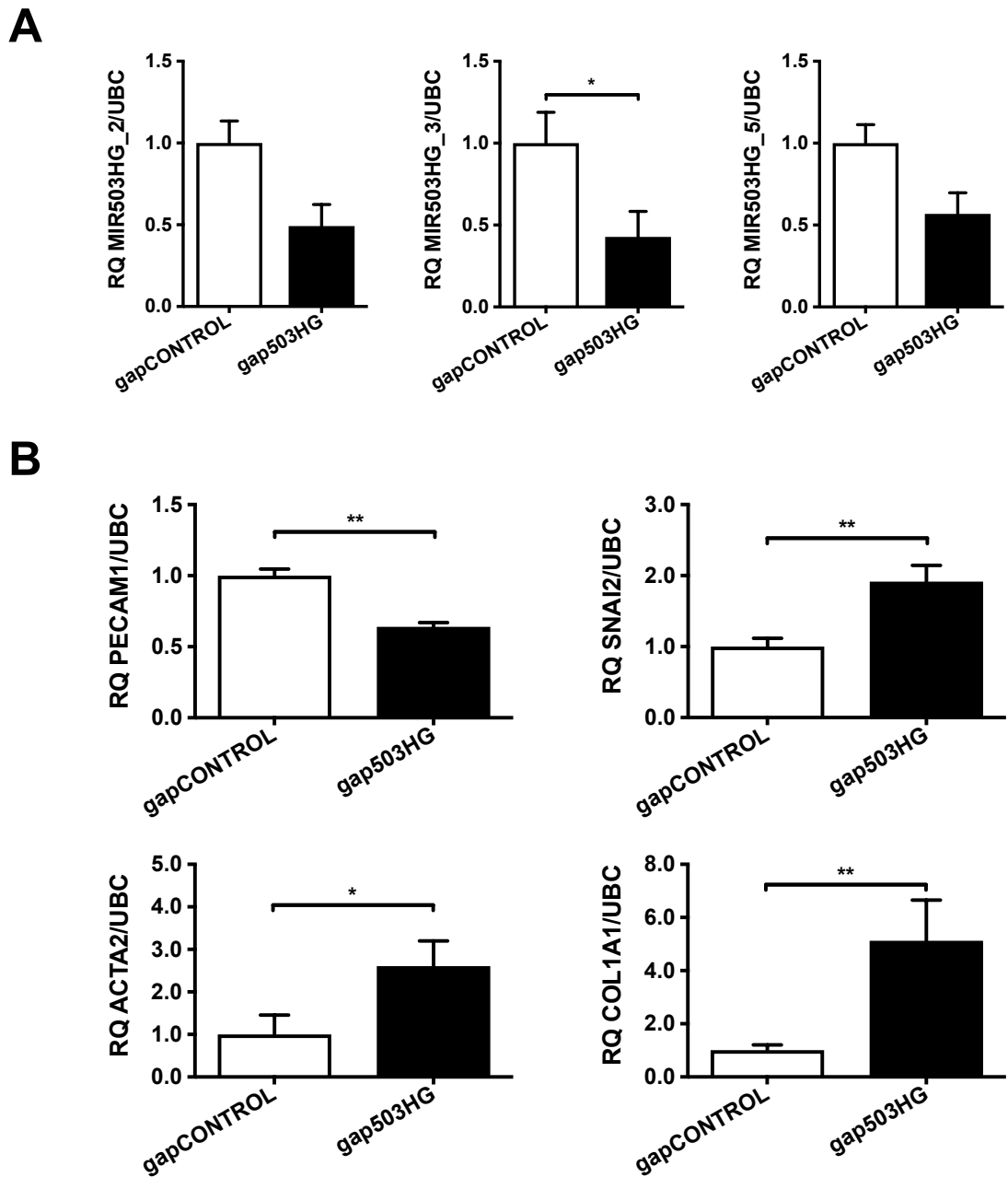
Figure 3.15: MIR503HG gapmeR-mediated knockdown induces EndMT.
Expression of **(A)** MIR503HG_2/3/5 and **(B)** EndMT marker genes in HUVEC 7 days after knockdown using gap503HG (20 nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=4 biological replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired gapControl. Data represented as mean ±SEM.
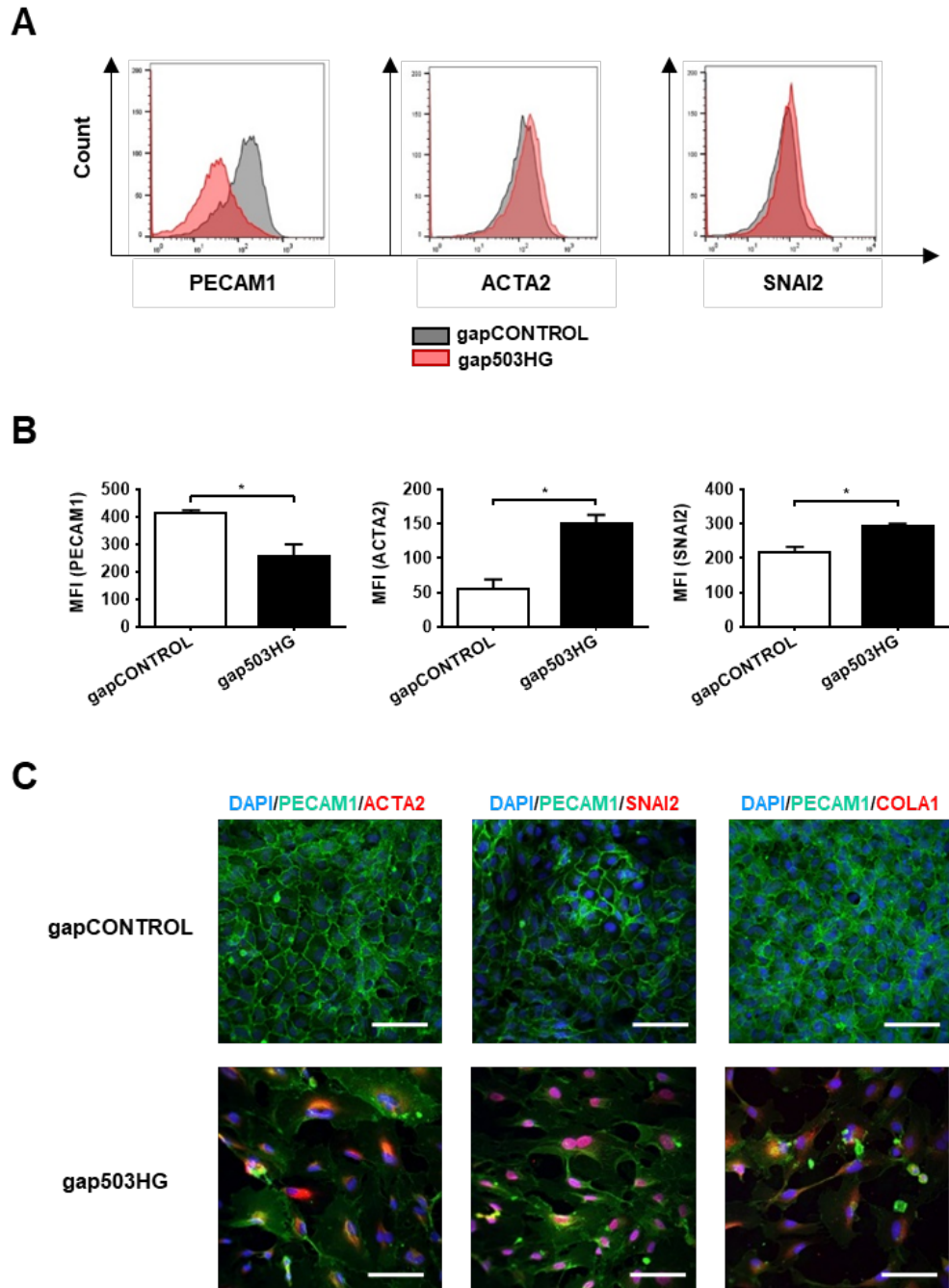
Figure 3.16: MIR503HG gapmeR-mediated knockdown induces EndMT.

(A) EndMT marker expression analysed by flow cytometry 7 days after knockdown using gap503HG. Representative histograms for controls cells (grey) compared to gap503HG-cells (red) and **(B)** MFI quantification (n=3 biological replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired gapControl. Data represented as mean ±SEM. **(C)** Representative immunofluorescence images of EndMT markers expression in HUVEC (20X, scale bar 50 μm) 7 days after knockdown using gap503HG. PECAM1 in green, ACTA2/COL1A1/SNAI2 in red and DAPI in blue. Images acquired with Andor Revolution XDi spinning disk confocal microscope and analysed with Image J Software. This work was performed in collaboration with Dr Axelle Caudrillier.

3.3.6 MIR503HG *in vitro* overexpression

3.3.6.1 Transcript selection and lentiviral overexpression

By comparing the MIR503HG locus sequence against that of 100 vertebrates, we found that the final 595 base-pair region of isoform 2 (MIR503HG_2) was highly conserved (Figure 3.17), as was its secondary substructure [438]. As such, we chose MIR503HG_2 as our transcript of interest. To further understand the role of MIR503HG during EndMT, we designed a lentiviral vector carrying the entire 760bp sequence of the MIR503HG_2 mature transcript (LNT_503HG), which does not include the sequences encoding for miRNA-424 and miRNA-503 found within the lncRNA locus (Figure 3.1). The lentiviral construct was initially validated in untreated HUVEC 3 days after transduction, this resulted in an increase in MIR503HG_2 expression but not MIR503HG_3 or 5 (Figure 3.18). In order to confirm that the overexpression was restricted to the cell nucleus, given the endogenous lncRNAs nuclear localisation, we carried out cellular fractionation and subsequent qPCR analysis on HUVECs overexpressing MIR503HG_2. In control conditions, MIR503HG_2 is predominately nuclear, with 2.7% cytoplasmic expression. Whilst this does increase to 7.3% after overexpression, MIR503HG_2 remains predominately nuclear (Figure 3.19).
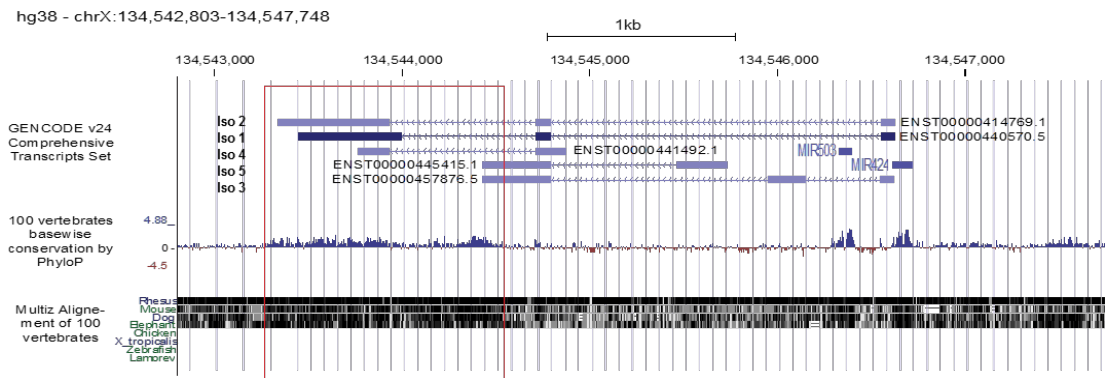
Figure 3.17: MIR503HG locus conservation and annotation in vertebrates.

MIR503HG transcript annotation and conservation based on UCSC genome Browser data (GENCODE v24) from 100 vertebrates (PhyloP score and multiple alignment with Multiz). The red box highlights the higher conservation regions present in the downstream of the MIR503HG locus.
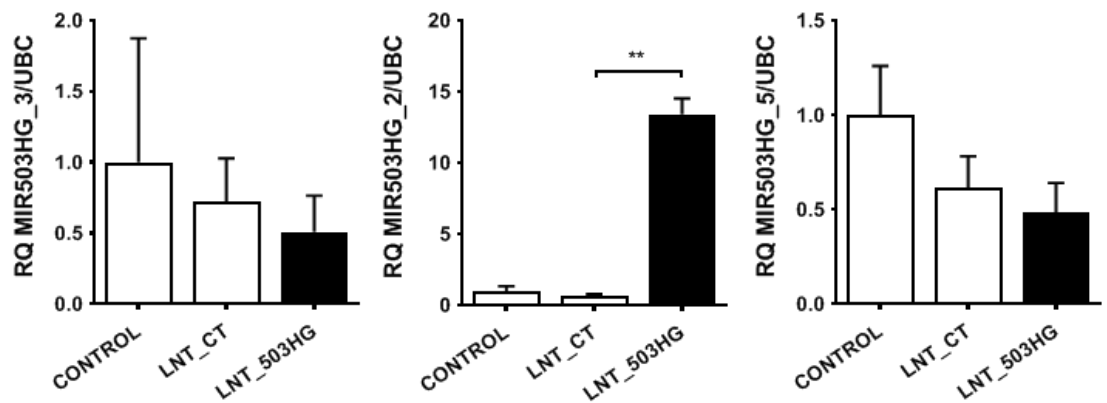
Figure 3.18: Lentiviral overexpression of MIR503HG_2 at day 3.

Expression of MIR503HG_2, MIR503HG_3 and MIR503HG_5 in HUVEC at day 3 after lentiviral transduction with LNT_503HG (isoform 2) with a MOI 5 compared to a paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=3 technical replicates). Analysis by one-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001. Data represented as mean ±SEM.

143
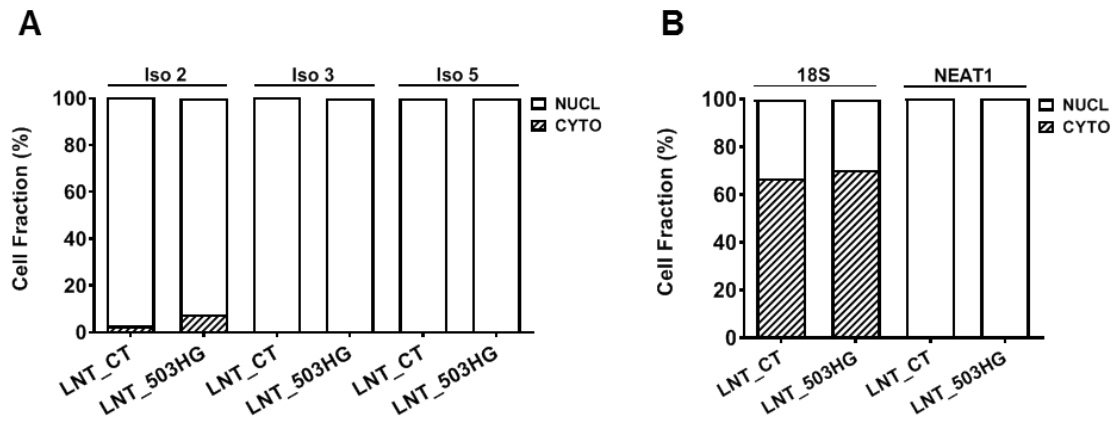
Figure 3.19: Subcellular Localisation of MIR503HG after lentiviral overexpression.
Subcellular localisation of **(A)** MIR503HG_2, MIR503HG_3 and MIR503HG_5, and **(B)** fractionation controls 18S and NEAT in HUVEC at day 3 after lentiviral transduction with LNT_503HG (isoform 2) with a MOI 5 compared to a paired control. Cell fraction value for gene expression was quantified by qRT-PCR assay relative to UBC (n=3 technical replicates).

3.3.6.2 Overexpression of MIR503HG during EndMT *in vitro*

By day 7, untreated HUVEC infected by LNT_503HG showed a 4.8 ($\pm$1.3) fold increase in MIR503HG_2 expression compared to an empty lentiviral control (LNT_CT) (Figure 3.20 A), without any changes in endothelial or mesenchymal markers (Figure 3.20 B). Under EndMT conditions, MIR503HG overexpressing cells presented a significantly higher PECAM1 expression, and a marked suppression of SNAI2 and COL1A1 expression by day 7 when compared to co-treated LNT_CT cells (Figure 3.20). Immunofluorescence (Figure 3.21) and flow cytometry (Figure 3.22) further validated these results at the protein level.
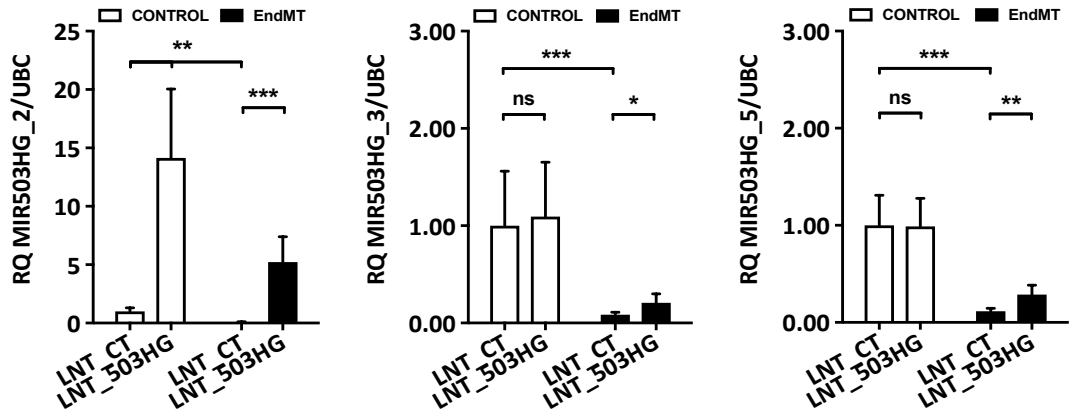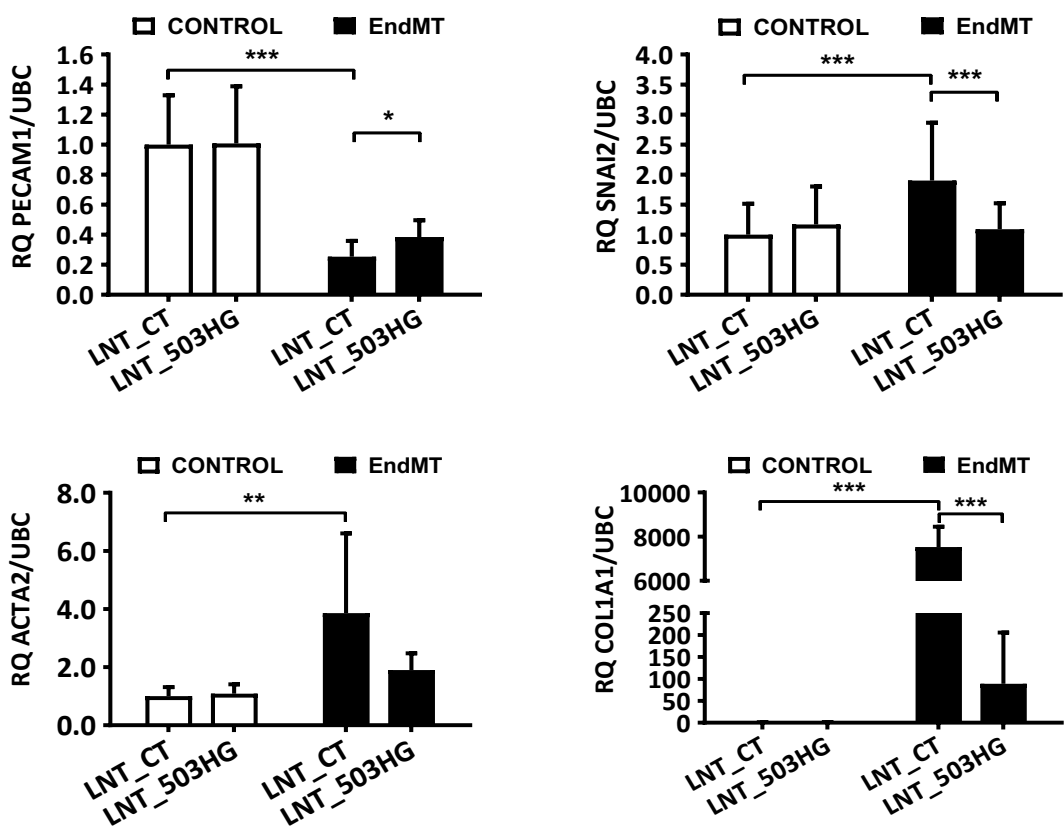
Figure 3.20: MIR503HG overexpression represses EndMT *in vitro*.

Expression of **(A)** MIR503HG_2/3/5 and **(B)** EndMT marker genes in HUVEC 7 days following MIR503HG overexpression with LNT_503_2 (MOI 5) ± EndMT (n=5 biological replicates). Analysis by two-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001. Data represented as mean ±SEM.
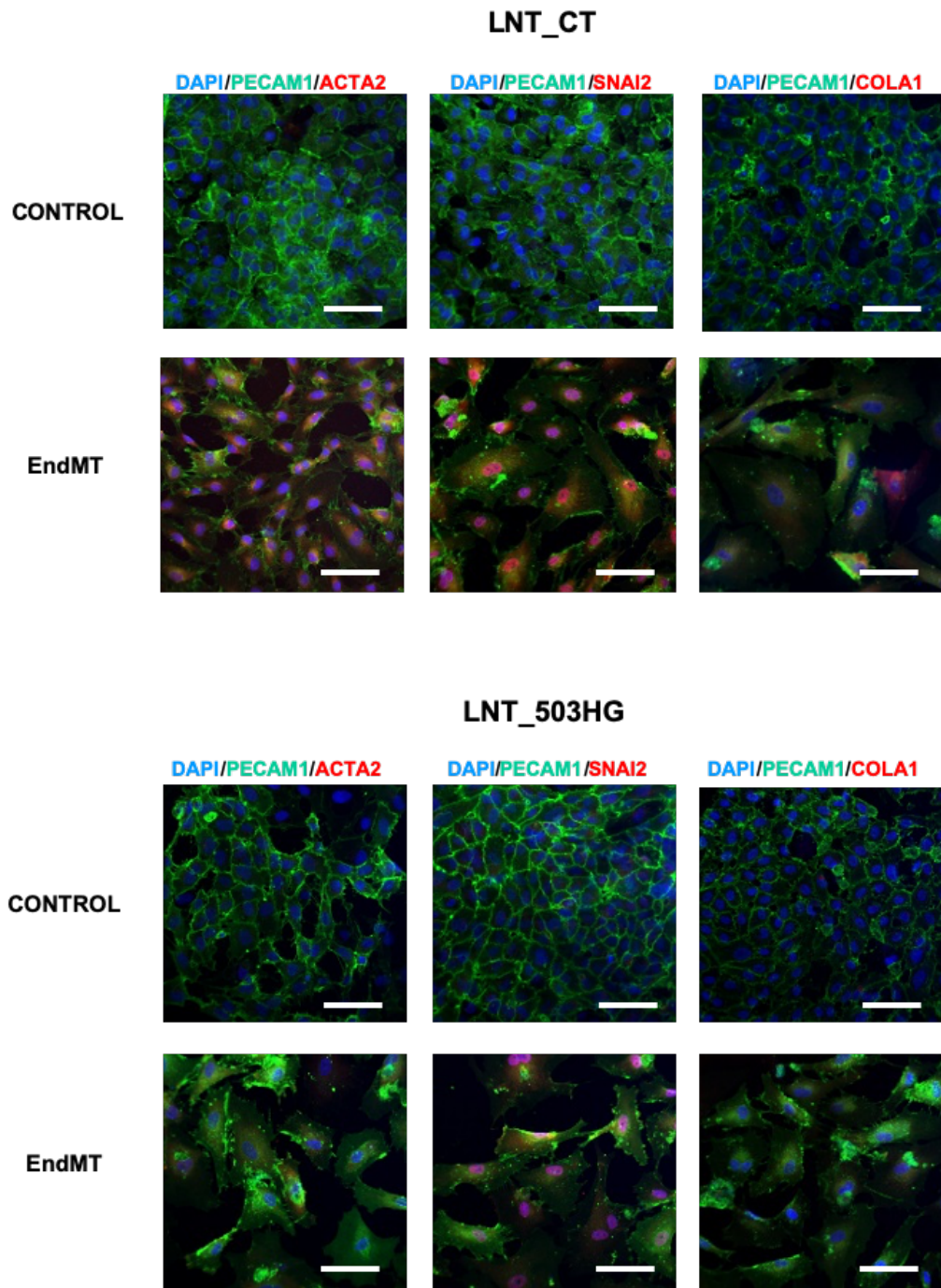
Figure 3.21: MIR503HG overexpression represses EndMT *in vitro*.

Representative immunofluorescence images of EndMT marker expression in HUVEC 7 days following MIR503HG overexpression with LNT_503_2 (MOI 5) ± EndMT (20X, scale bar 50 μm). PECAM1 in green, ACTA2/COL1A1/SNAI2 in red and DAPI (nucleus) in blue. Images acquired with Andor Revolution XDi spinning disk confocal microscope and analysed with Image J Software. This work was performed in collaboration with Dr Axelle Caudrillier.
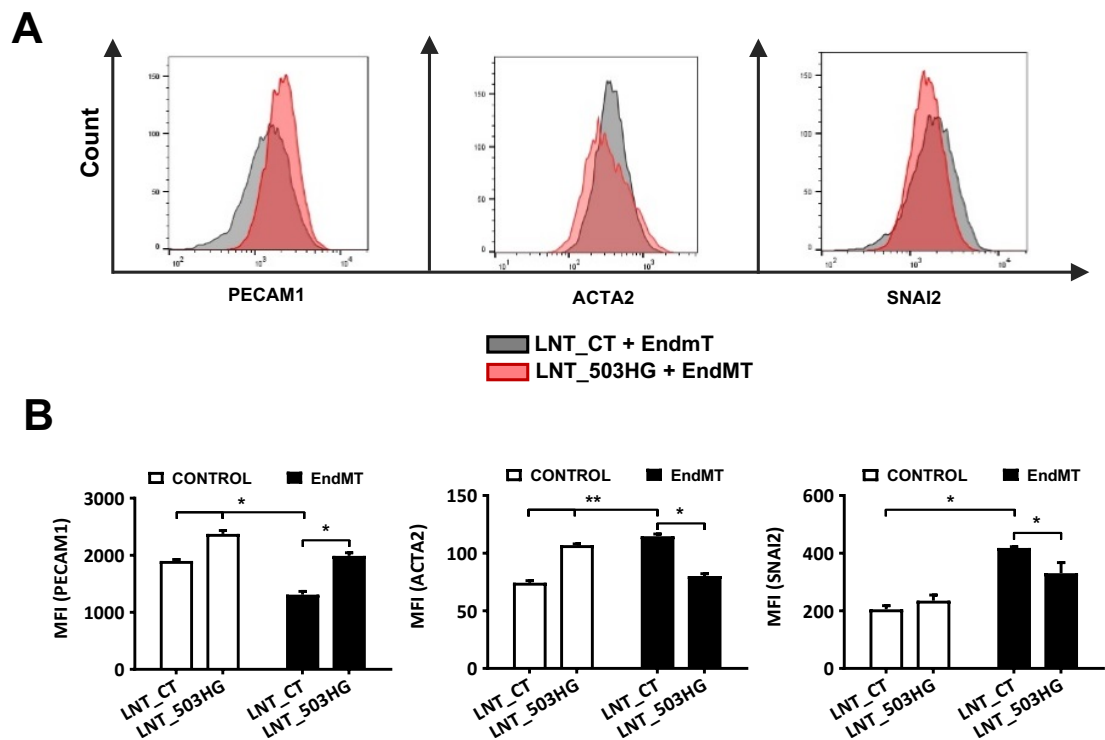
Figure 3.22: MIR503HG overexpression represses EndMT *in vitro*.

Flow cytometry analysis of EndMT marker protein expression in HUVEC 7 days following MIR503HG overexpression with LNT_503_2 (MOI 5) ± EndMT. **(A)** Representative histograms for EndMT-LNT_CT (grey) compared to EndMT- LNT_503_2 (red) and **(B)** MFI quantification (n=3 biological replicates). Analysis by two-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001. Data represented as mean ±SEM. This work was performed in collaboration with Dr Axelle Caudrillier.

3.3.6.3 Transcriptome analysis of MIR503HG overexpression during EndMT

To identify the contribution of MIR503HG to endothelial function and which regulated pathways may be involved in EndMT progression, we performed deep RNAseq on cells overexpressing MIR503HG_2 (LNT_503HG_2) during TGF-β2 and IL-1β co-treatment. PCA revealed that MIR503HG_2 overexpression alone had no major effect on the transcriptome of untreated cells (CT_LNT_503HG), However, when over expressed during EndMT (EndMT_LNT_503HG), treated cells clustered in proximity to their respective untreated controls (Figure 3.23 A). Additionally, in EndMT_LNT_503HG cells, we found that a total of 877 genes were significantly up-regulated and 1047 down-regulated by day 7 (Figure 3.23 B-C). Of these, there was a substantial overlap in genes that were regulated in opposing directions during EndMT prevention and induction, with approximately 28% of EndMT-associated genes affected by MIR503HG_2 overexpression (Figure 3.23 B-C). Gene ontology (GO) term analysis of genes affected by MIR503HG, relative to all EndMT-associated changes, showed an overlap in functions linked to both cell adhesion and migration (Figure 3.23 D). Migration-associated genes such as MMP1, ITGA6 and HNRNPA2B1 were downregulated, whereas those associated with proliferation such as MKI67, AURKA/B or CENPF increased.
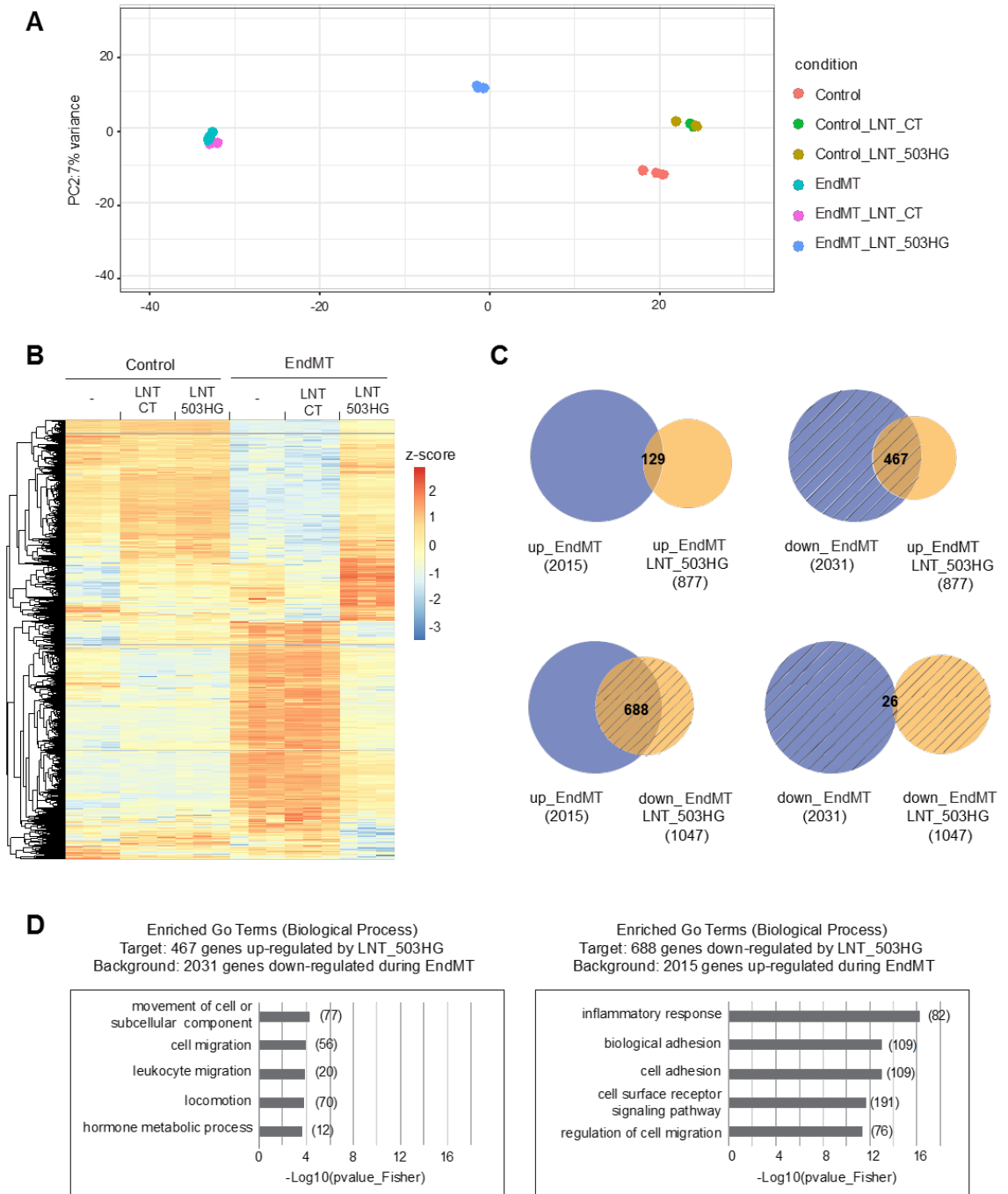
Figure 3.23: Transcriptome analysis of MIR503HG overexpression during EndMT

RNAseq samples represented as **(A)** PCA plot and **(B)** Heatmap (z-score). **(C)** Venn diagram of the overlap between significant change after LNT_503HG vs LNT_CT in EndMT-cells and EndMT changes (EndMT vs untreated). **(D)** Gene Ontology analysis of the genes regulated by LNT_503HG compared to the genes affected by EndMT (background). Bioinformatic analysis shown was carried out by Dr Julie Rodor.

3.3.7 Effects on endothelial cell proliferation and cell cycle

Along with key transcriptional changes, EndMT is often associated with abnormal cell proliferation [61,172]. As such, we used an EdU cell proliferation assay to stain live cells and flow cytometry analysis to quantify EdU uptake after a 4 h incubation period. Our results showed that knockdown of the MIR503HG locus led to a significant decreased in cell proliferation, with only 15.4% of si503HG transfected cells incorporating EdU, compared to 35.9% with siControl (Figure 3.24). This mirrored our 7-day *in vitro* model of EndMT, showing similar differences in EdU incorporating between co-treated and untreated cells (Figure 3.25).

These changes were found to be associated with abnormal cell cycle progression. Cell cycle analysis demonstrated that MIR503HG locus knockdown induced G1-phase retention, acutely blocking S-phase entry (Figure 3.26). Again, fully mirroring our 7-day *in vitro* model of EndMT (Figure 3.27).

Notably, these changes were also accompanied by a substantial increase in forward scatter (FSC) and side scatter (SSC) intensity as measured by flow cytometry (Figure 3.28 and Figure 3.29). Changes to FSC intensity are proportional to cell diameter and SSC to cell granularity, thus indicating not only changes in cell size, but also the appearance of a distinct cell population [443].
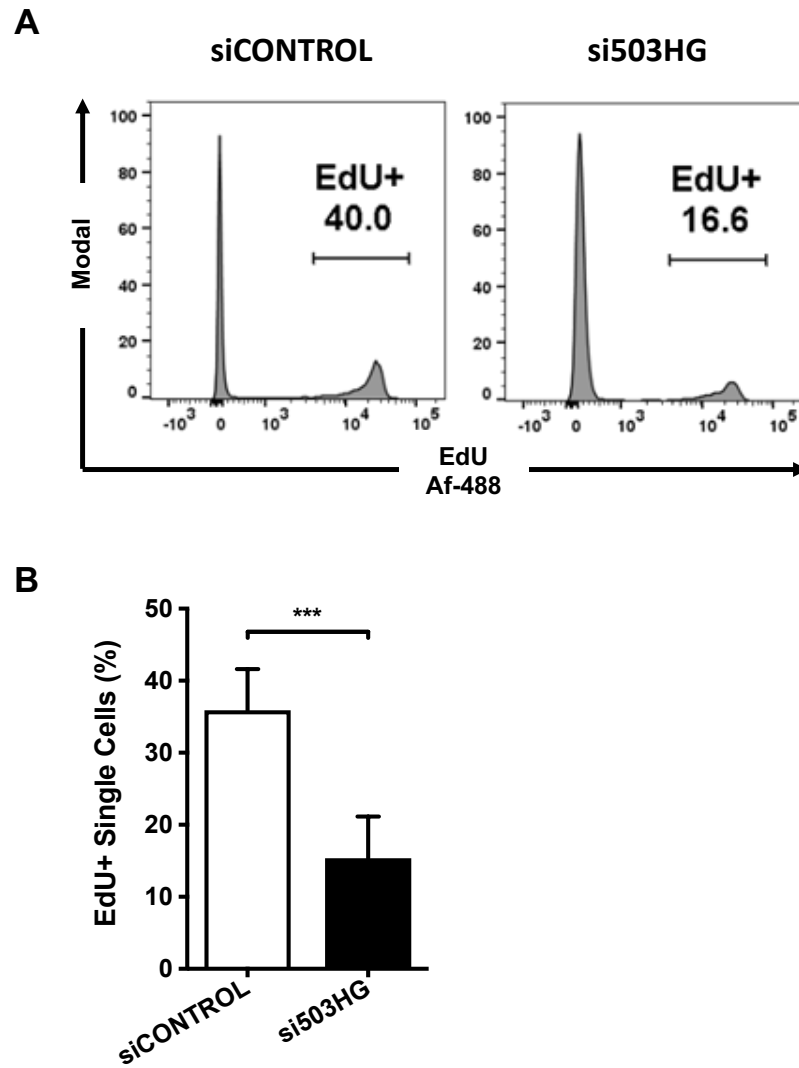
Figure 3.24: MIR503HG knockdown mediated changes to cell proliferation.

(A) Representative FACS histogram plots for EdU incorporation by HUVEC at Day 7 following DsiRNA-mediated knockdown. EdU+ box highlights the intensity of EdU signal used as an indicator of the proliferative state of the cells. **(B)** Bar chart showing mean changes for EdU incorporation (n=3 biological replicates). Data represented as mean ±SEM. Analysis by two-tailed t-test; ***$p \leq 0.001$ vs paired siControl.
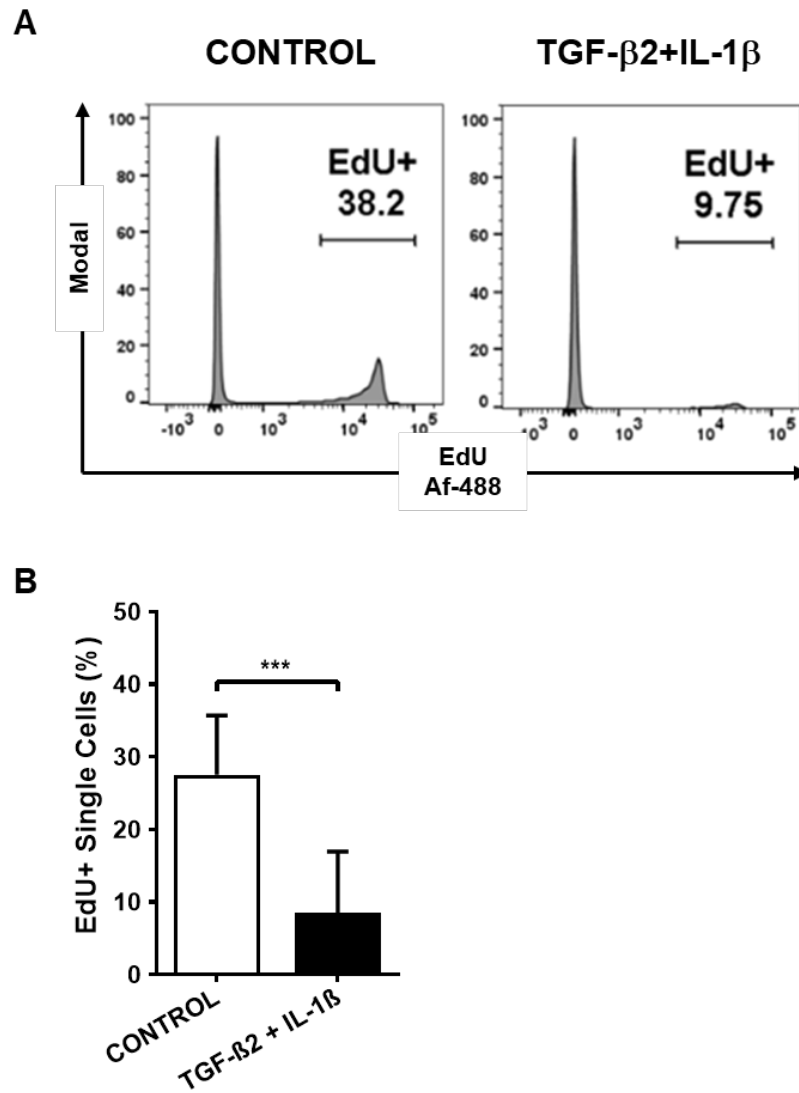
Figure 3.25: EndMT mediated changes to cell proliferation.

(A) Representative FACS histogram plots for EdU incorporation by HUVEC at Day 7 following TGF-β2+IL-1β treatment. *EdU+* box highlights the intensity of EdU signal used as an indicator of the proliferative state of the cells. **(B)** Bar chart showing mean changes for EdU incorporation (n=3 biological replicates). Data represented as mean ±SEM. Analysis by two-tailed *t*-test; ***p≤0.001 vs paired untreated control.
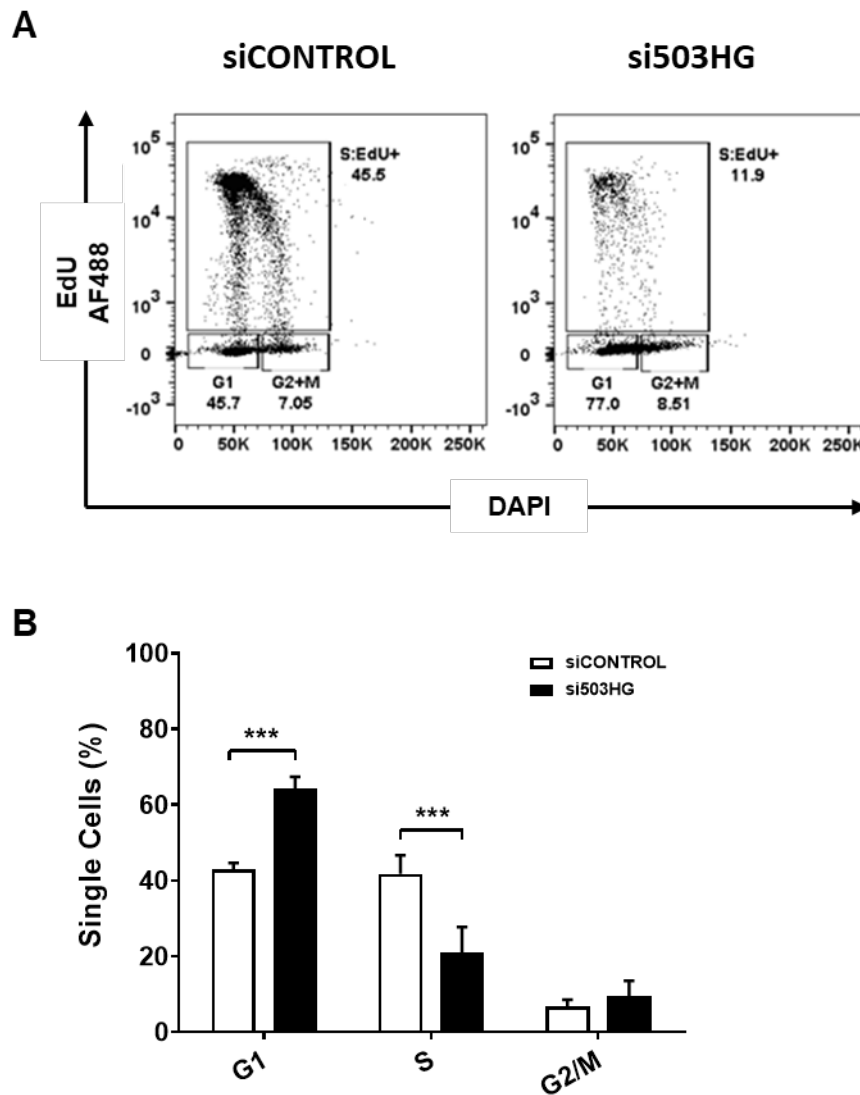
Figure 3.26: MIR503HG knockdown mediated changes to cell cycle.
(A) Representative FACS dot plots of cell cycle profile of HUVECs treated with EdU and DAPI at Day 7 following DsiRNA-mediated knockdown. *EdU+* box highlights the intensity of EdU signal used as an indicator of the proliferative state of the cells. (B) Bar chart showing mean changes in distribution of cells between cell cycle phases (n=3 biological replicates). Data represented as mean ±SEM. Analysis by two-tailed *t*-test; ***p≤0.001 vs paired siControl. This work was performed in collaboration with Dr Alena Shmakova.
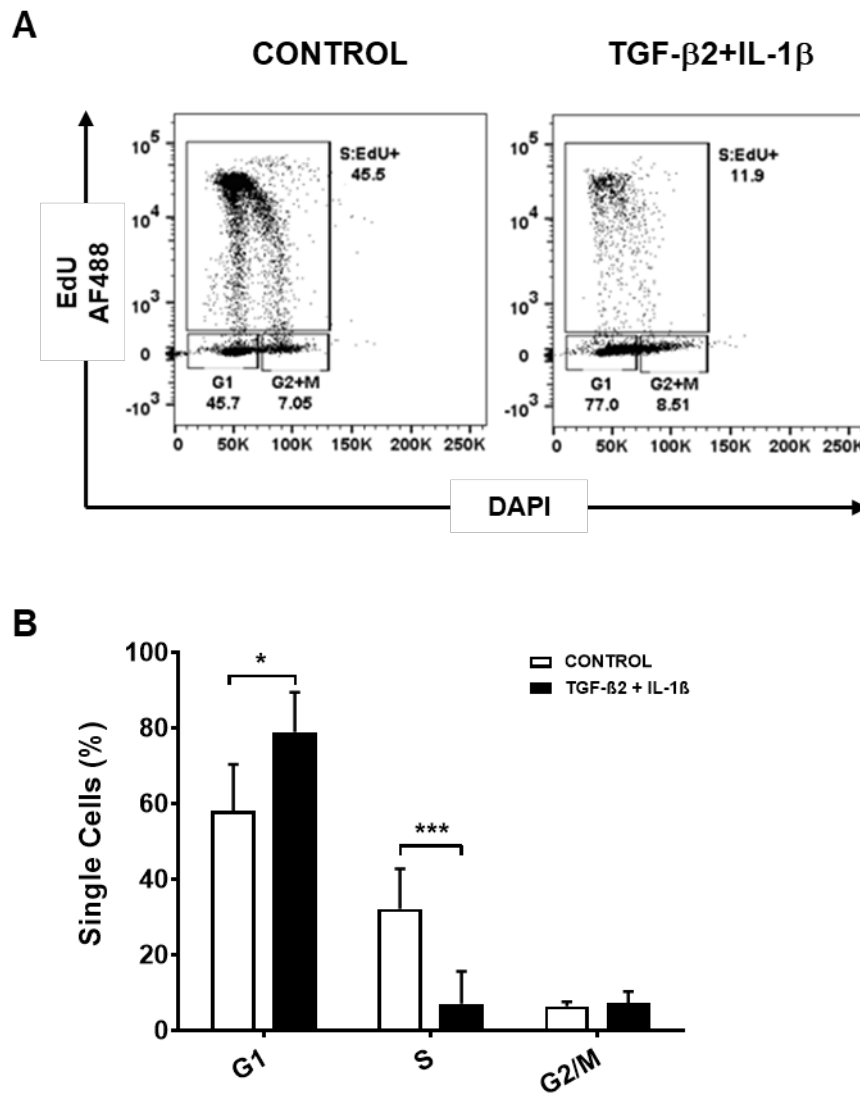
Figure 3.27: EndMT mediated changes to cell cycle.

(A) Representative FACS dot plots of cell cycle profile of HUVECs treated with EdU and DAPI at Day 7 following TGF-β2+IL-1β treatment. *EdU+* box highlights the intensity of EdU signal used as an indicator of the proliferative state of the cells. **(B)** Bar chart showing mean changes in distribution of cells between cell cycle phases (n=3 biological replicates). Data represented as mean ±SEM. Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and ***p≤0.001 vs paired untreated control. This work was performed in collaboration with Dr Alena Shmakova.
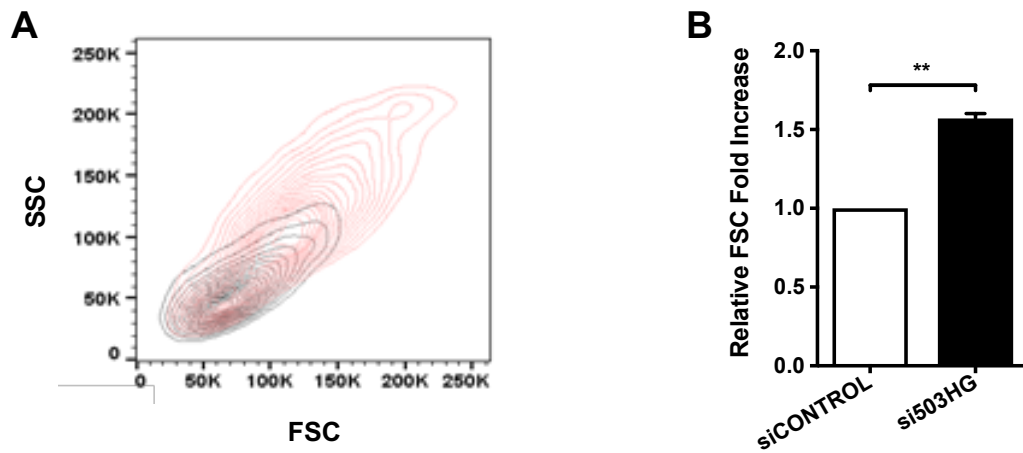
Figure 3.28: MIR503HG knockdown mediated changes to cell size.

(A) Representative FACS contour plots depicting cell distribution of HUVEC 7 days after knockdown using si503HG (Red) compared to paired siControl (Grey). **(B)** Bar chart showing mean fold change in relative cell size calculated as a ratio of FSC-A of si503HG treated cells to siControl cells. Data represented as mean ±SEM (n=3 biological replicates). Data analysed as one sample t-test compared to hypothetical value of 1; **p≤0.01 vs paired siControl. This work was performed in collaboration with Dr Alena Shmakova.
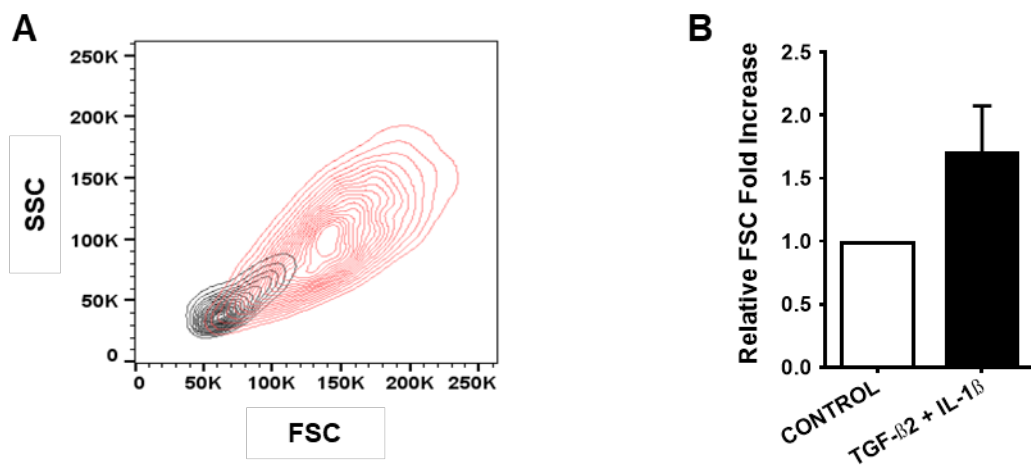
Figure 3.29: EndMT mediated changes to cell size.

(A) Representative FACS contour plots depicting cell distribution of HUVEC after TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) co-treatment for 7 days (Red) compared to an untreated control (Grey). **(B)** Bar chart showing mean fold change in relative cell size calculated as a ratio of FSC-A of TGF-β2+IL-1β treated cells to control cells. Data represented as mean ±SEM (n=3 biological replicates). Data analysed as one sample t-test compared to hypothetical value of 1; **p≤0.01 vs paired control. This work was performed in collaboration with Dr Alena Shmakova.

3.3.8 Effects on endothelial cell migration

A

B
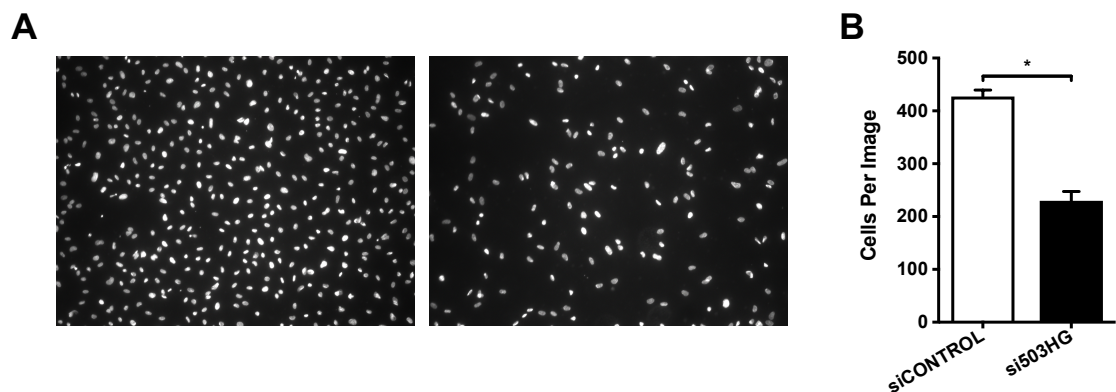
Figure 3.30: MIR503HG knockdown mediated changes to cell migration.

(A) Representative transwell migration assay images of HUVEC 7 days after knockdown using si503HG (20nM) compared to siControl (20 nM). Fixed migrated cells were stained using DAPI and imaged by confocal microscopy. **(B)** Migrated cell average after siRNA transfection. Average calculated by counting cell number in five random x10 fields (n=3 biological replicates). Data represented as mean ±SEM. Analysis by two-tailed t-test; *$p \leq 0.05$ vs paired siControl.

A

B
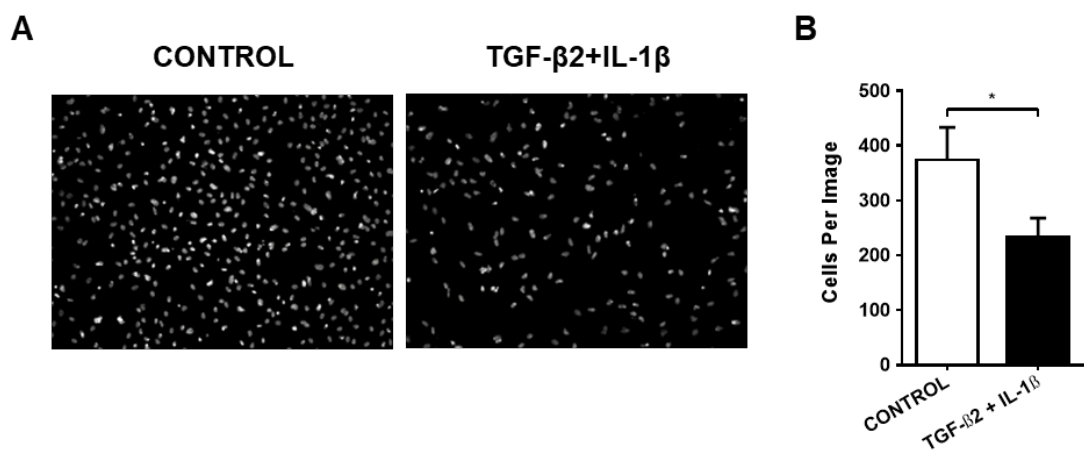
Figure 3.31: EndMT mediated changes to cell migration.

(A) Representative transwell migration assay images of HUVEC 7 days after TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) co-treatment compared to an untreated control. Fixed migrated cells were stained using DAPI and imaged by confocal microscopy. **(B)** Migrated cell average after co-treatment. Average calculated by counting cell number in five random x10 fields (n=3 biological replicates). Data represented as mean ±SEM. Analysis by two-tailed t-test; *$p \leq 0.05$ vs paired control.

Another key feature of cells undergoing EndMT is the acquisition of a migratory phenotype. As such, we analysed the effect of MIR503HG depletion on HUVEC using a Transwell migration assay. Cell migration was significantly decreased after 7 days after MIR503HG knockdown, with the average number of cells per viewing field dropping to 216.6±23.5 as opposed to 392.3±50 when using siControl (Figure 3.30). Nearly identical results were seen after TGF-β2 and IL-1β co-treatment, with the average number of cells per viewing field dropping to 235.6±32.4 as opposed to 376.6±56.9 when untreated (Figure 3.31).

3.4 Discussion

In line with the EndMT-associated lncRNA transcriptional profile described by our exploratory RNAseq analysis, in this chapter we identify the loss of the lncRNA MIR503HG as a key event in the induction of EndMT *in vitro* across multiple EC phenotypes. Crucially, based on available publications, in the context of this thesis we have defined EndMT as a significant decrease in the expression of the endothelial marker PECAM1, accompanied by an increase in the transcription factor SNAI2, as well as the mesenchymal markers ACTA2 and COL1A1 [10,14,15,25,53]. However, it is important to highlight that there are a wealth of other valid markers used to define this transition. For example, expression of the transcription factors SNAI1, TWIST1 a ZEB1/2 has been shown to be associated with the loss of EC markers and an EndMT profile [37,38,40,444]. Unsurprisingly, as shown in Chapter 3.3.2, our RNAseq analysis identifies a pool of over 1000 differentially expressed genes in both transitioning HUVEC and HPAEC. However, such an exhaustive validation process both at the RNA and protein level could not be implemented, thus our choice to limit this to a

well-established group of markers whose change in expression strongly supports the appearance of an EndMT phenotype.

Further, while the loss of MIR503HG after TGF-β2 and IL-1β co-treatment was consistent across ECs from different vascular beds, conditions such as PAH often involve the remodelling of pre-capillary micro-vessels [445]. Additional validation studies should therefore include a panel of relevant pulmonary microvascular ECs such as pulmonary microvascular ECs (PMVEC), previously shown to undergo EndMT [130,159,446].

3.4.1 MIR503HG Conservation

Although MIR503HG was first considered to function solely as a miRNA precursor gene, it is now understood that the lncRNA has its own independent transcriptional pattern and functional role. Importantly, despite the lack of evolutionary conservation commonly seen with lncRNA genes, the MIR503HG locus displays remarkable conservation across vertebrates. As described by Necsulea and colleagues, the genetic locus of MIR503HG likely originated from a common tetrapod ancestor 370 million years ago [435]. Our own analysis of the locus showed that the final exonic 595 base-pair region of MIR503HG_2, largely separated from the conserved miRNA cluster, was the transcriptional area with the highest conservation.

Further, lncRNAs also present secondary structure conservation which may rely on short sequences preserved across species rather than large regions [421,422,447]. Interestingly, when analysed, the secondary structure formed by the final exonic region of MIR503HG_2 was similar to that of several non-human primates [438]. Ultimately, this remarkable sequence and structural conservation suggests the existence of functional transcriptional domains beyond its role as a host gene.

3.4.2   MIR503HG Regulates EndMT *in vitro*

DsiRNA-mediated knockdown of MIR503HG induced a robust EndMT profile in the absence of any other treatment, including TGF-β2 and IL-1β co-stimulation. While this strategy was clearly effective, knockdown of lncRNAs may be heavily impacted by their cellular localisation. Given the nuclear accumulation of MIR503HG, the reduced presence of RNA interference (RNAi) machinery in the nucleus may limit the efficacy of our knockdown strategy. Thus, while nuclear RNAs can still be targeted by siRNAs [448,449], the use of an alternative strategy was needed to confirm our results. As an alternative, we used an antisense GapmeR target MIR503HG in HUVEC. While there was a strong downregulation of all transcripts by day 3, this was not sustained for a full 7 days. Nonetheless, this was sufficient to replicate our initial results. As with our DsiRNA, changes were seen both at the mRNA and protein level, with a clear downregulation of PECAM1, and increased expression of SNAI2, ACTA2 and COL1A1.

Despite the clear association between MIR503HG knockdown and the induction of EndMT, we should acknowledge that MIR503HG is encoded and likely transcribed alongside miR-424(322) and miR-503. While DsiRNA and GapmeR-mediated knockdown are commonly used strategies to target functional RNAs, these reagents may also promote the cleavage of unspliced nascent RNA transcripts [450]. Thus, our knockdown strategy could additionally target the miRNA cluster and confound the results show in this chapter. We have attempted to address these issues fully in Chapter 4.

Given the complexity of the MIR503HG locus and risk of DsiRNA/Gapmer-induced off-target effects, the use of gain-of-function experiments was imperative to further

confirm our results. As such, we have also attempted to examine the effects of MIR503HG overexpression on EC function and phenotypic commitment. In contrast to our knockdown experiments, by increasing the availability of MIR503HG_2 alone during our EndMT *in vitro* model we effectively supressed the appearance of mesenchymal markers, while preventing further depletion of endothelial cell-cell adhesion molecules. This, along with our knockdown model, suggests that MIR503HG may have a key regulatory role in maintaining endothelial homeostasis and that its absence allows for the adoption of a mesenchymal phenotype.

Despite the success of MIR503HG_2 overexpression in repressing EndMT after continuous exposure to TGFβ2 and IL1β, this was still a partial inhibition of transition. Seeing that this may be due to the over expression of a single isoform and not the totality of the MIR503HG locus, in Chapter 6 of this thesis we have outlined a series of future studies that may expand on the results presented here.

3.4.3 Endothelial Migration and Proliferation

Along with crucial transcriptional changes to EndMT marker expression, depletion of MIR503HG also induced a significant decrease in cell migration and proliferation. This appears, in part, to be due to a delay in cell cycle progression leading to an accumulation of G1-phase cells, resulting in a significant increase in cell size. This, not only fully mimics our TGF-β2 and IL-1β EndMT model, but largely corroborates previous published data by Fiedler and colleagues, showing similar changes to EC proliferation and migration profiles after MIR503HG knockdown [381].

Distinct stages of EndMT with different characteristics have also been reported [52]. In particular, cells with partial versus complete EndMT have been isolated based on the level of endothelial markers, with those cells with complete EndMT showing high

proliferative and migratory capacity [172]. While advanced EndMT may lead to increased mesenchymal proliferation, early stage EndMT models, like ours, have been associated with a decrease in endothelial proliferation [53,119,451]. Despite canonical TGF-β signalling being one of the major drivers of EndMT, its effect will largely depend on which receptor pathway is activated. In ECs, TGF-β can act via two distinct type I receptors, ALK1 and ALK5 [14,56,59]. Activation of ALK1, for example, will lead to Smad1/5 phosphorylation promoting endothelial migration and proliferation at the expense of EndMT [56,59,452,453]. On the other hand, despite inducing EMT and EndMT, ALK5 signalling and consequent SMAD2/3 phosphorylation will inhibit EC proliferation and migration [14,56,454,455]. Crucially, as demonstrated by Cooley and colleagues, EndMT in a murine vein graft model is dependent on the early activation of ALK5 signalling. The appearance of phosphorylated SMAD2/3 was detected in transitioning EC at day 3 and 7 but declined after day 14, eventually becoming undetectable by day 35 [14].

Of note, while both our EndMT model and MIR503HG knockdown strategy induce the expression of mesenchymal markers, by day 7 these cells are still in the initial stages of the transition and continue to express endothelial markers. One can speculate that, once in a more advanced mesenchymal state, these cells would ultimately reach an enhanced proliferative and migratory state. Further in-depth studies are necessary to confirm this however.

3.4.4 MIR503HG Mechanistic insight

Given their functional versatility, uncovering the cell-specific molecular interactions that dictate the function of lncRNAs is crucial. Unlike mRNAs, which are generally exported to the cytoplasm for further processing, lncRNAs can be functionally active

in both the cytoplasmic or nuclear compartments. Thus, the nuclear enrichment of MIR503HG may narrow down its window of putative mechanistic interactions. For example, several nuclear lncRNAs have been shown to regulate gene transcription via recruitment of chromatin modifying complexes or chromosome spatial conformation or acting as scaffolds joining distal enhancer elements to gene promoters [456].

With reports suggesting that 60% of annotated LncRNAs are chromatin enriched, future studies may implement additional fractionation techniques to isolate chromatin-associated RNA, in addition to the cytoplasmic and nuclear portions [457]. While beyond the scope of this thesis, this may guide future RNA-centred studies using techniques such as RNA antisense purification (RAP) and chromatin isolation by RNA purification (ChIRP) to isolate MIR503HG-associated RNAs, proteins or genomic regions [458–460].

The use of RNAseq-based transcriptome analysis of MIR503HG overexpression during EndMT also confirmed its key role in maintaining EC commitment, with over 28% of transition genes affected by its expression. However, while these transcriptional changes may help dictate further experiments, they were too broad to find any specific candidates.

Additionally, our analysis also highlighted several migration regulatory functions. As ECs undergoing EndMT have bene reported to migrate into the surrounding tissue this may be a biologically conserved function of MIR503HG [14,53]. In fact, overexpression of MIR503HG has also been previously linked to hepatocellular carcinoma cell migration through its interaction with the heterogeneous nuclear ribonucleoprotein A2/B1 (HNRNPA2B1) and consequent inhibition of NF-κB, central for migration [437]. Similarly, during EMT, increased HNRNPA2B1 expression has been shown to

promote cell invasion and migration [461] again connecting MIR503HG with regulation of migration. It's important, however, to note that given the cell-specific functions of MIR503HG suggested earlier in this chapter, this interaction may not be present in ECs. Thus, confirming an interaction between MIR503HG and HNRNPA2B1 in EC via RNA immunoprecipitation would be necessary.

Lastly, while we have provided initial findings on the effect of MIR503HG on gene expression and cell function, it is important to point out that further comprehensive mechanistic studies are required to identify the molecular targets of MIR503HG.

Chapter 4: Identifying the Function of the MIR503HG Locus in EndMT

4.1 Introduction

Spanning a total of ~7 kb on the minus strand of Xq26.3, the MIR503HG locus encodes 6 district miRNAs (miR-424, -503, -542, -450-1, - 450-2 and -450b). Unlike its other miRNA counterparts, the miR-424 and miR-503 cluster is found within the intronic and exonic regions of MIR503HG and share common regulatory regions (

Figure 4.1). Both the ENCODE and GeneHancer projects have identified a variety of regulatory elements approximately 2 kb upstream from the MIR503HG locus, including the gene regulatory histone modification marker H3K27Ac and DNAse I-hypersensitive clusters [462]. Over the past decade miR-424 and miR-503 have also been shown, either individually or in unison, to drive a variety of processes key to tissue differentiation and remodelling both during development and disease.

During embryogenesis, for example, miR-424 and miR-503 expression was found to be largely restricted to the developing heart and skeletal muscles. This expression pattern was found to match that of the transcription factor Mesp1, a master regulator of cardiovascular progenitor cell specification and skeletal myogenic differentiation [463]. In an embryonic stem cell (ESC) differentiation model, ectopic expression of miR-424/-503 was able to specifically drive cardiomyocyte lineage programming while its inhibition repressed cardiomyocyte formation [463]. Crucially, during heart development EMT and EndMT are believed to control formation the mitral and tricuspid valves in the atrioventricular canal and the aortic and pulmonary valves in the outflow tract [464–468]. Unsurprisingly, several publications now show that regulating EMT is one of the key functions of miR-424/-503 able to deregulate epithelial function and enhance mesenchymal programming [469–472]. A mechanism by which miR-424/503 amplify

TGF-β signalling, and with it EMT, has also been reported. For example, during TGF-β-induced EMT in human lung epithelial cells miR-424 targeted the E3 ubiquitin ligase SMURF2, preventing the degradation of phosphorylated SMAD2 and thus enhance TGF-β signalling [473]. A similar mechanism was seen with miR-503, which again directly interacted with SMURF2 mRNA potentiating TGF-β/SMAD2 signalling and in turn disrupting intestinal epithelial homeostasis [474]. Nonetheless, and central the work described herein, a link between the miRNA cluster and EndMT is yet to be established.

Further, much like their lncRNA host, miR-424/503 have also been described as hypoxia-sensitive non-coding RNAs with a role in cell proliferation, migration and angiogenic response. A seminal paper by Ghosh *et al* was the first to demonstrate this in a variety of human primary ECs, where miR-424 levels were significantly increased under hypoxic conditions [475]. The paper went on to show that the rodent homolog of human miR-424, miR-322, was significantly upregulated in experimental rat myocardial infarction and mouse hindlimb ischemia models. Similarly, the expression of miR-503 was also strongly upregulated *in vitro* in ECs after hypoxia and *in vivo* in ischemic limb muscles of diabetic mice [381,476].
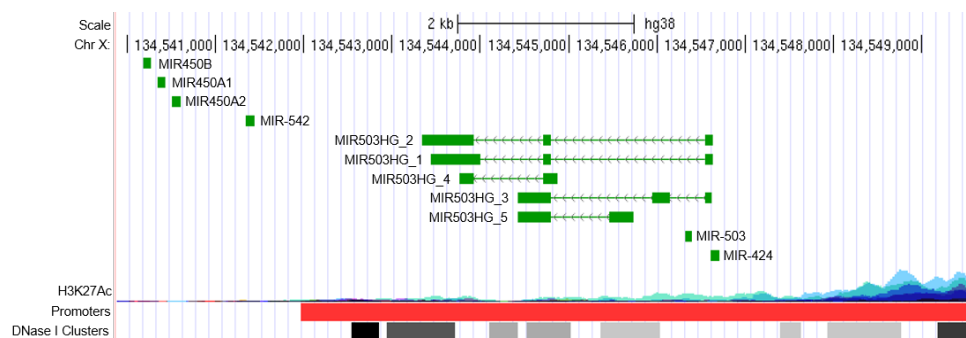
Figure 4.1: Regulatory regions of the MIR503HG locus.

MIR503HG locus annotation and regulatory regions based on UCSC genome Browser data (GENCODE v24).

However, while several papers have continued to show an acute increase in the expression of the miRNAs in response to ischemic injury in a variety of tissues [477,478], this seems to be largely a context dependent effect. For example, in a cerebral ischemia/reperfusion injury mouse model, the expression levels of miR-424 where significantly upregulated in the surrounding infarct area at 1 and 4 h after injury but downregulated at 24 h [478]. The expression of miR-503 was also found to be decreased in endothelial progenitor cells (EPCs) under hypoxic conditions, and when overexpressed the miRNA suppressed EPC proliferation, migration and tube formation [479]. Similarly, mice with an endothelial-specific deletion of miR-424(322) and miR-503 were reported to have augmented angiogenic responses as demonstrated by increased neovascularisation [480].

Interestingly, in hypoxia/Sugen-induced PH the expression of both miR-424 and miR-503 was also found to be significantly decreased in association with disrupted apelin signalling [203]. Apelin is itself is known to maintain pulmonary vascular homeostasis and reduced expression and has been linked to hyperproliferative EC phenotypes and the development of PAH [203,481,482]. As demonstrated by Kim and colleagues, miR-424/503 are able to exert antiproliferative effects in both pulmonary ECs and SMCs by directly targeting FGF2 and FGFR1 [203]. In fact, restoration of miR-424 and miR-503 was able to both decrease right ventricular systolic pressures (RVSP), reduce the number of muscularised lung microvasculature and drop FGF2/FGFR1 expression in a hypoxia/Sugen rat model of PH [203].

Finally, given the body of evidence presented here, both miR-424 and miR-503 present within similar regulatory networks of MIR503HG. Thus, further dissection of the locus is necessary to validate the function miRNA cluster independently of MIR503HG.

4.2 Aims

A growing body of evidence now suggests that miRNAs and lncRNAs may be co-transcribed and/or co-regulated. However, despite the range of studies published, the transcriptional and functional overlap between the miRNA cluster and its lncRNA host has been largely ignored. As such, the aims of this chapter were as follows:

- To quantify the expression of the miR-424 and miR-503 during EndMT;

- To verify if the miRNA cluster is modulated by MIR503HG manipulation;

- Evaluate the function of miR-424 and miR-503 in EndMT to see if their manipulation phenocopies the effect of MIR503HG modulation.

4.3 Results

4.3.1 MIR503HG miRNA Locus Expression Profile in EndMT

As with previous experiments, miRNA expression was measured in HUVEC co-stimulated with a continuous dose of TGF-β2 (10 ng/mL) and IL-1β (1 ng/mL) for a total of 7 days. qRT-PCR analysis showed a significant down regulation of both 5' and 3' arms (guide and passenger strand) of miR-424(322) and miR-503, much like their lncRNA counterpart (Figure 4.2). Further, while both DsiRNA and Gapmer strategies described in Chapter 3 were successful in depleting MIR503HG, these may also led to the cleavage of nascent RNA transcripts prior to pre-miRNA excision [450]. Consistent with our experimental EndMT model, knockdown of MIR503H*G* via si503HG was associated with a significant drop in expression of the 5' arms of miR-424(322) and miR-503 (Figure 4.3). However, while mean RQ values were lower, this was not significant with gap503HG (Figure 4.3).
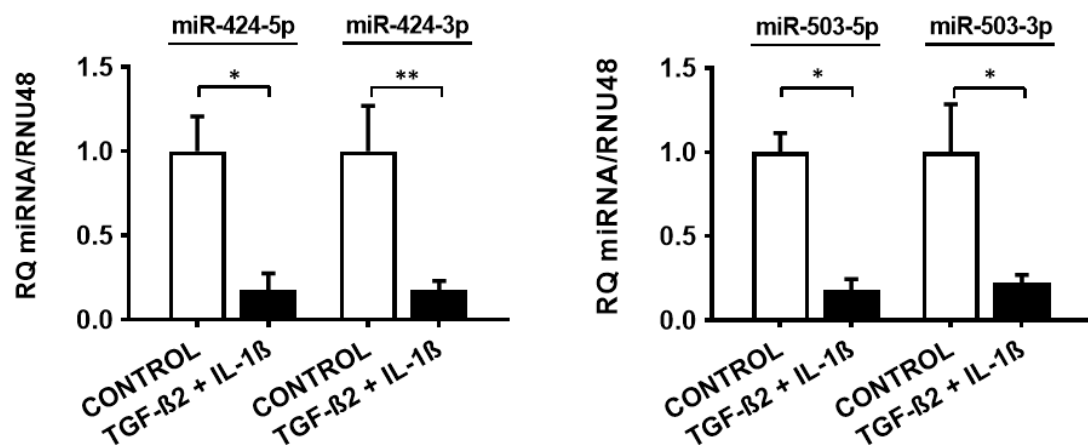
Figure 4.2: MIR503HG expression during EndMT *in vitro*.
Expression of miRNA-424-5p/3p and miR-503-5p/3p in HUVEC ± EndMT treatment. RQ value for gene expression was quantified by qRT-PCR assay relative to untreated control cells (n=4 biological replicates). Analysis by two-tailed t-test; *p≤0.05, ** p≤0.01 and *** p≤0.001 vs paired control. Data represented as mean ±SEM.
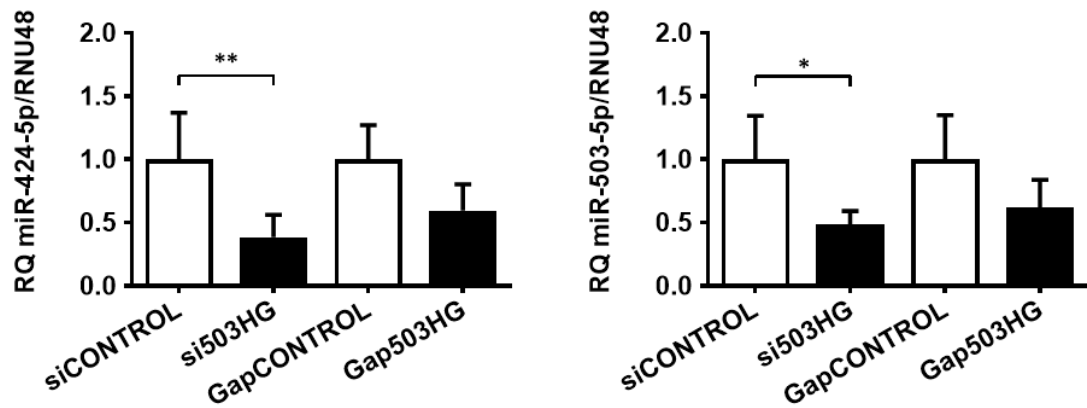
Figure 4.3: miR-424 and miR-503 expression after MIR503HG knockdown.

Expression of miRNA-424-5p/3p and miR-503-5p/3p in HUVEC 7 days after knockdown using si503HG (20 nM) or Gap503HG (20 nM) compared to a paired control. RQ values for gene expression were quantified by qRT-PCR assay relative to RNU48 (n=3 biological replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired control. Data represented as mean ±SEM.
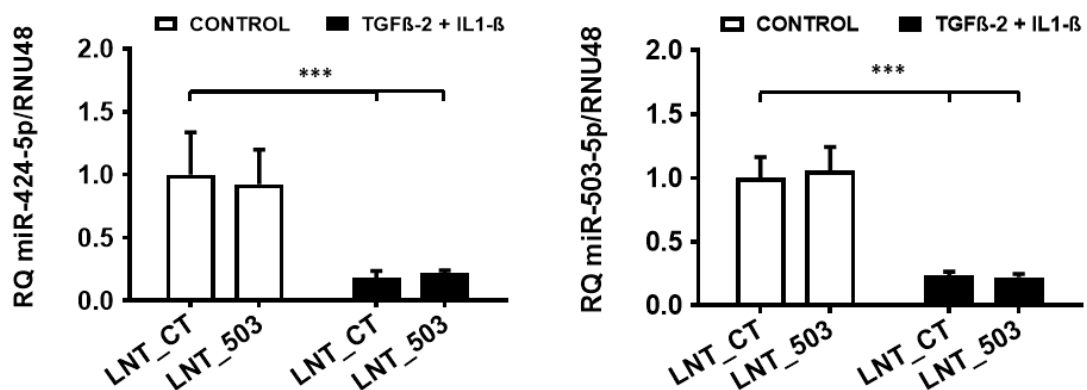
Figure 4.4: miR-424 and miR-503 expression after MIR503HG overexpression.

Expression of miRNA-424-5p/3p and miR-503-5p/3p in HUVEC 7 days following MIR503HG overexpression with LNT_503_2 (MOI 5) ± EndMT. RQ values for gene expression were quantified by qRT-PCR assay relative to RNU48 (n=3 biological replicates). Analysis by two-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired control. Data represented as mean ±SEM.

In order to understand the role of MIR503HG during EndMT, we designed a lentiviral vector carrying the entire 760 bp sequence of the MIR503HG_2 mature transcript (LNT_503HG), which did not include the sequences encoding the two miRNAs.

Crucially, overexpression of MIR503HG in untreated HUVEC did not affect miR-424-5p or miR-503-5p levels compared to LNT_CT alone (Figure 4.4). In EndMT-treated cells, MIR503HG overexpression again did not affect miRNA levels, compared to EndMT-LNT_CT cells; suggesting that the effects of MIR503HG on EndMT may be independent of miRNA regulation.

4.3.2   MIR503HG miRNA Locus Target Analysis
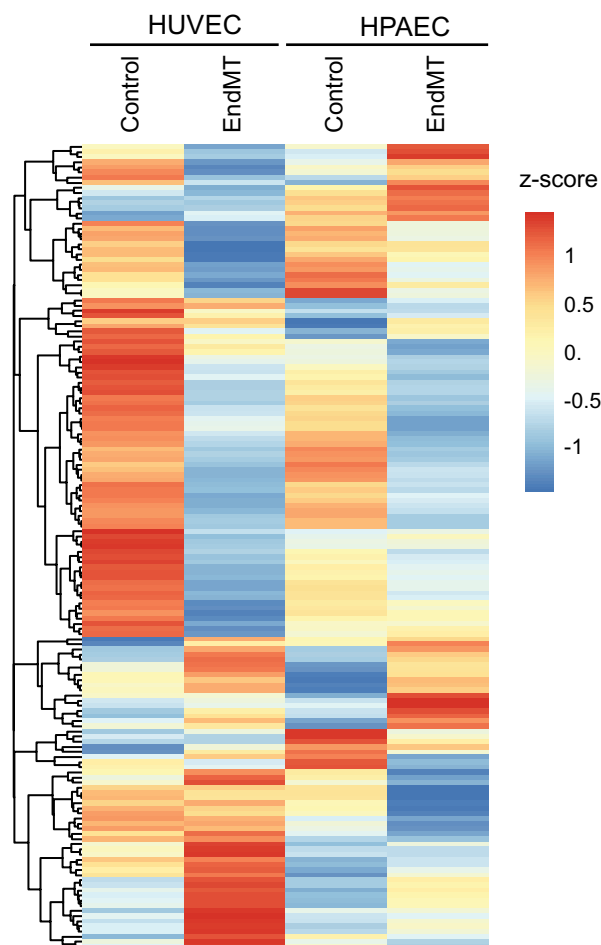
Figure 4.5: Gene expression profile of miR-503 and miR-424 mRNA targets during EndMT.
Heatmap of predicted miR-503 and miR-424 targets relative to their expression levels in the RNAseq. Target list obtained using miRTarbase. Z-score represents the deviation from the mean by standard deviation units of the Log2 (FPKM) for each gene. Bioinformatics analysis shown was carried out by Dr Julie Rodor.

Loss of MIR503HG, be it via targeted knockdown or EndMT co-treatment, was associated with a significant reduction in the expression of both miR-424 and miR-503. If the transcriptional changes observed during EndMT were to be driven by the associated miRNAs, it would stand to reason that their respective miRNA-targets should be enriched in the up-regulated gene pool of our RNAseq analysis. Thus, using miRTarBase v7.0 [483], we selected 183 previously validated targets common to miR-424 and miR-503 (Figure 4.5). Of the 183 selected miRNA target genes, 157 were detected in our RNA-seq dataset for both HUVEC and HPAEC. However, of these, only a total of 8 were up-regulated in co-treated HUVEC and HPAEC. Given the small percentage of upregulated target genes and lack of target overlap between EC populations, this led us to believe that miR424/503 were unlikely to be major drivers of EndMT. Nonetheless, further research using gain- and loss-of-function strategies were necessary to confirm our findings.

4.3.3 MIR503HG miRNA locus *in vitro* knockdown

Given the pronounced downregulation observed during EndMT, an anti-miR-based approach was used to knockdown either miR-424-5p or miR-503-5p. Knockdown was first confirmed at day 3 after transfection, showing a significant depletion of both miRNA in untreated HUVEC (Figure 4.6). Given that both miRNA have been previously shown to target phosphatases of the cell division cycle 25 family (CDC25) and fibroblast growth factor (FGF) signalling via FGF2 and FGFR1, these targets were used to validate our knockdown strategy [476,484,485]. While only FGF2 was increased at day 3 after miR-424 knockdown, the availability of CDC25A1, FGF2 and FGFR1 was significantly increased after miR-503 knockdown (Figure 4.7).

When allowed to culture for a total of 7 days, our miRNA knockdown strategy did not

induce an EndMT profile in HUVEC, as no changes in endothelial or mesenchymal markers were observed (Figure 4.8). Notably, unlike our si503HG and gap503HG-based strategies, knockdown of either miRNA did not affect MIR503HG expression (Figure 4.9). Based on these results, we conclude that depletion of either miR-424 or miR-503 alone is not sufficient to induce EndMT. Thus, the effects observed after loss of MIR503HG appear to be independent of changes in canonical miR-424 and miR-503 function.
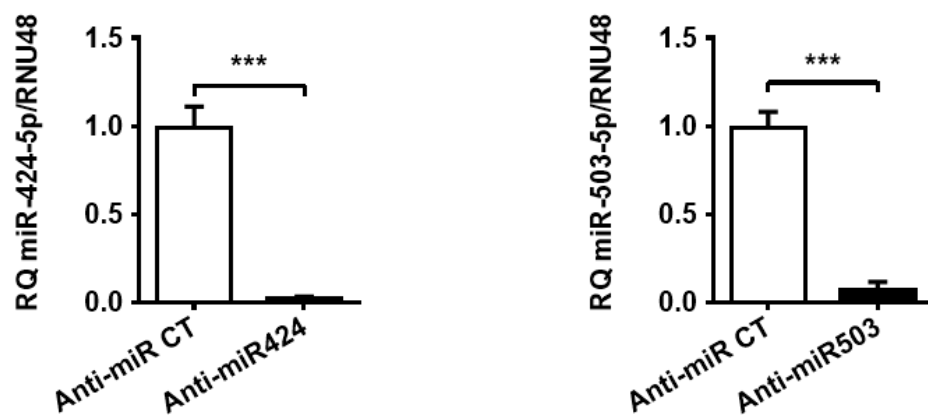
Figure 4.6: Anti-miR mediated knockdown of miR-424-5p and mirR-503-5p at day 3.
Expression of miR-424-5p and miR-503-5p in HUVEC at day 3 after transfection with either anti-miR-424 (25nM) or anti-miR-503 (25nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 technical replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Anti-miR Control (Anti-miR CT). Data represented as mean ±SEM.
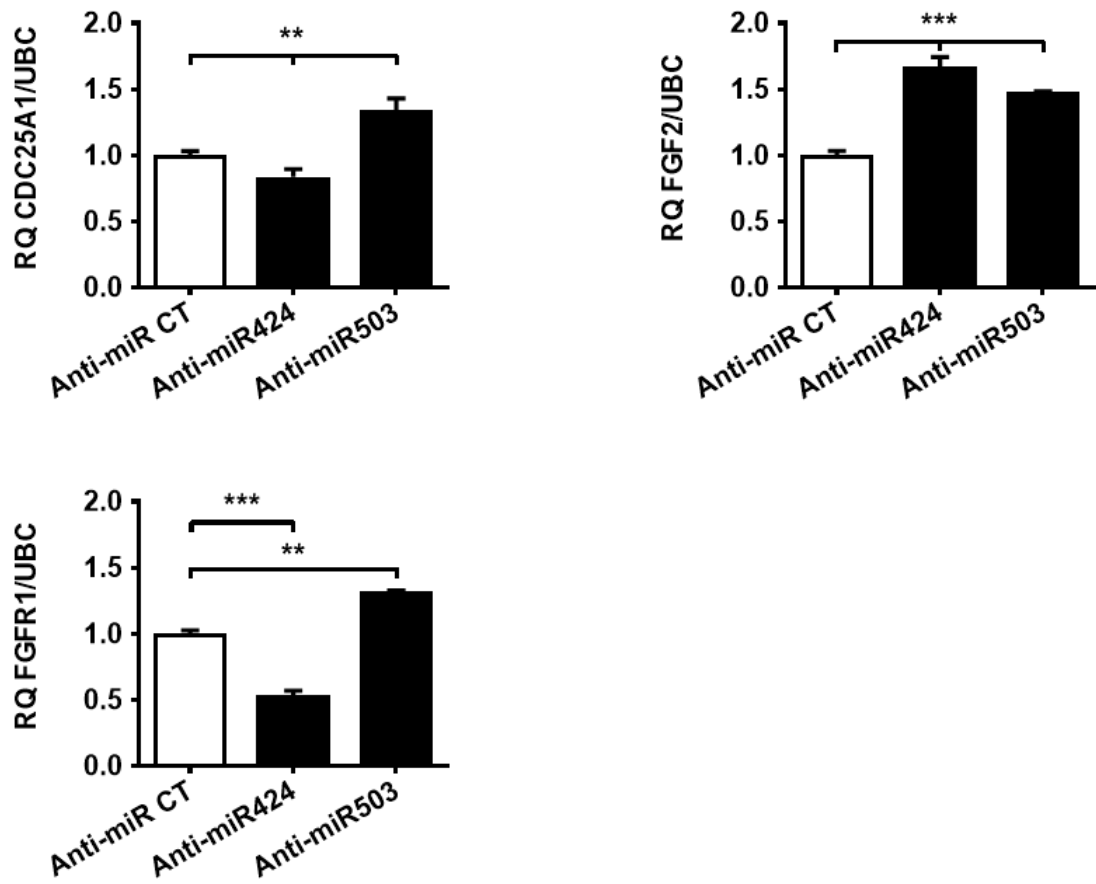
Figure 4.7: Anti-miR mediated knockdown of miR-424-5p and mirR-503-5p at day 3.

Expression of CDC25A, FGF2 and FGFR1 in HUVEC at day 3 after transfection with either anti-miR-424 (25nM) or anti-miR-503 (25nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 technical replicates). Analysis by one-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Anti-miR Control (Anti-miR CT). Data represented as mean ±SEM
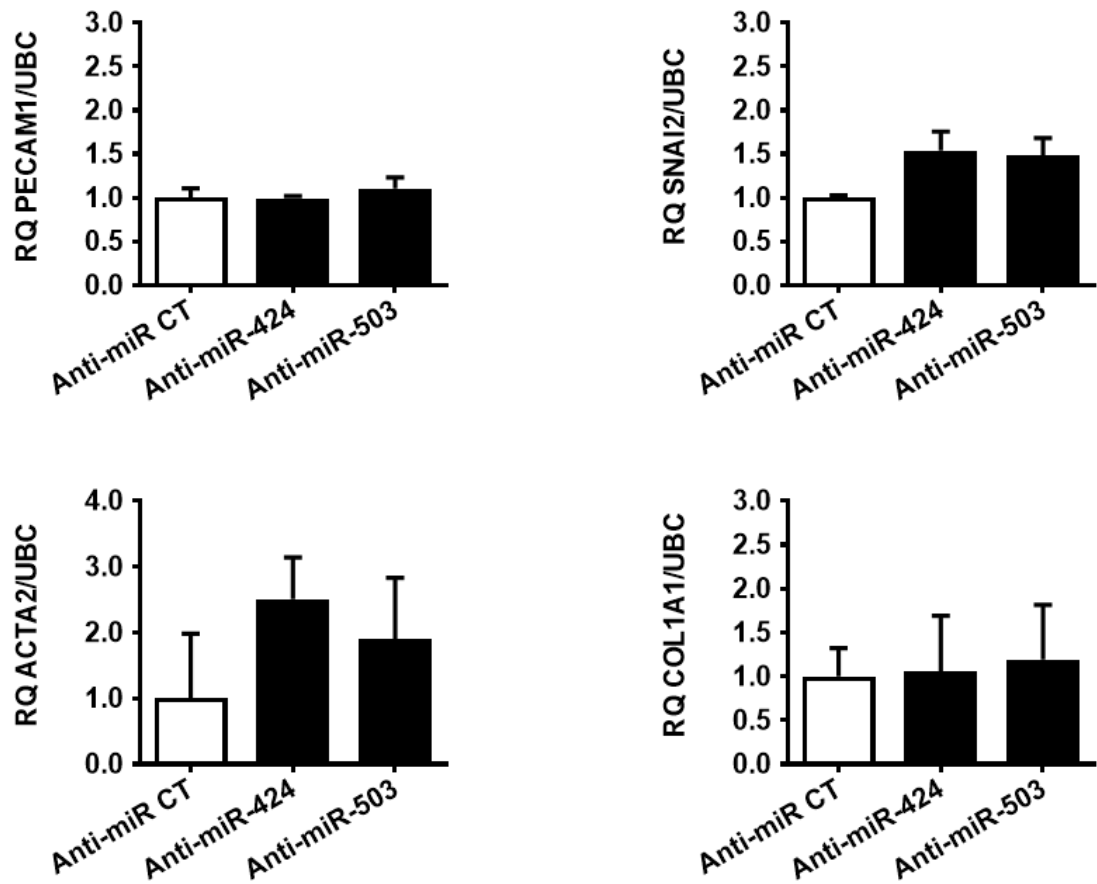
Figure 4.8: Anti-miR mediated knockdown of miR-424-5p and mirR-503-5p at day 7.
Expression of PECAM1, SNAI2, ACTA2 and COL1A1 in HUVEC at day 7 after transfection with either anti-miR-424 (25nM) or anti-miR-503 (25nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 biological replicates). Analysis by one-way ANOVA; $*p \leq 0.05$, $** \, p \leq 0.01$, and $*** \, p \leq 0.001$ vs paired Anti-miR Control (Anti-miR CT). Data represented as mean ±SEM.
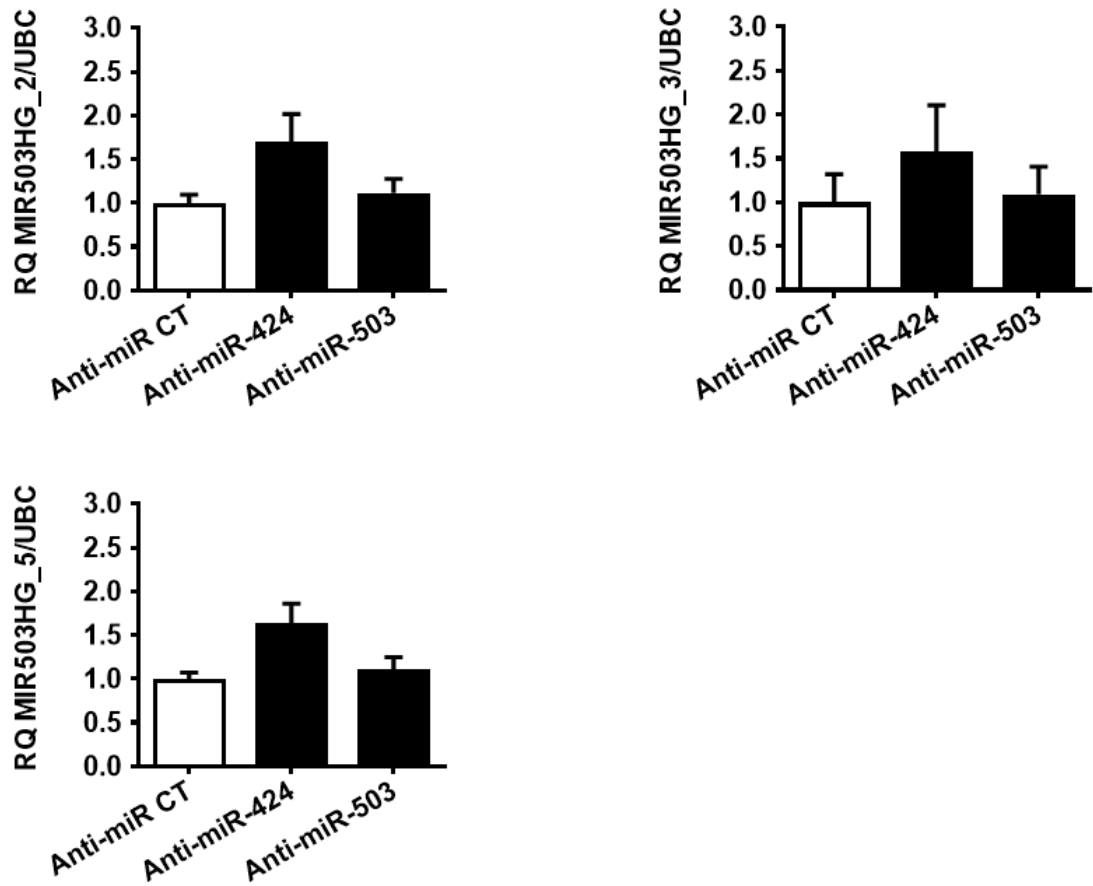
Figure 4.9: Anti-miR mediated knockdown of miR-424-5p and mirR-503-5p at day 7.

Expression of MIR503HG_2/3/5 in HUVEC at day 7 after transfection with either anti-miR-424 (25nM) or anti-miR-503 (25nM) compared to paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 biological replicates). Analysis by one-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Anti-miR Control (Anti-miR CT). Data represented as mean ±SEM

4.3.4 MIR503HG miRNA locus *in vitro* overexpression

In an attempt to mimic our MIR503HG overexpression experiments, we next overexpressed both miR-424-5p and miR-503-5p in the context of EndMT. In order to do this, we used synthetic pre-miR RNA precursors carrying the mature sequence for either miR-424-5p or miR-503-5p and introduced them during EndMT treatment. Overexpression was first confirmed at day 3 after transfection in untreated HUVEC, with both miRNAs showing significant upregulation compared to a control pre-miR (Figure 4.10). In contrast to our anti-miR strategy, overexpression of either miRNA induced a significant downregulation of CDC25A1 and FGFR1 (Figure 4.11). While there were no significant changes to FGF2 expression, this confirmed our overexpression strategy was functionally successful.

At 7 days after transfection, in contrast to MIR503HG overexpression, increased miRNA availability in untreated HUVEC led to a pronounced EndMT-like profile (Figure 4.12). Overexpression of either miR-424 or miR-503 was accompanied by a significant loss in PECAM1 expression along with increased SNAI2 and COL1A1. Again, in direct contrast to MIR503HG overexpression, this EndMT profile was further potentiated by TGF-β2 and IL-1β co-treatment (Figure 4.12). Notably, expression MIR503HG appeared to be decreased in all pre-miR samples (Figure 4.13). While only MIR503HG_3 was significantly downregulated, this suggests that the EndMT phenotype observed after miRNA overexpression may be due to regulation of MIR503HG.

Based on these results, despite being downregulated during EndMT, overexpression of either miR-424 or miR-503 does not supress endothelial transition. However, a cross-regulatory mechanism between the miRNA cluster and MIR503HG may exist.
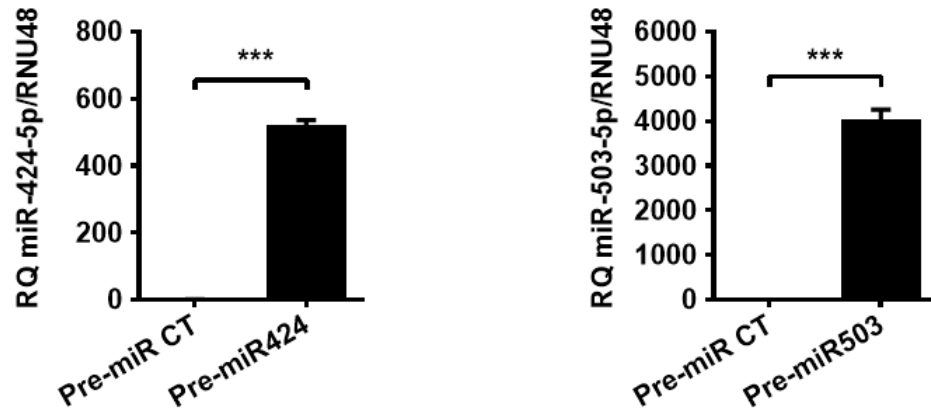
Figure 4.10: Pre-miR mediated overexpression of miR-424-5p and mirR-503-5p at day 3.
Expression of miR-424-5p and miR-503-5p in HUVEC at day 3 after transfection with either Pre-miR-424 (15nM) or Pre-miR-503 (15nM) compared to a paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 technical replicates). Analysis by two-tailed *t*-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Pre-miR Control (Pre-miR CT). Data represented as mean ±SEM.
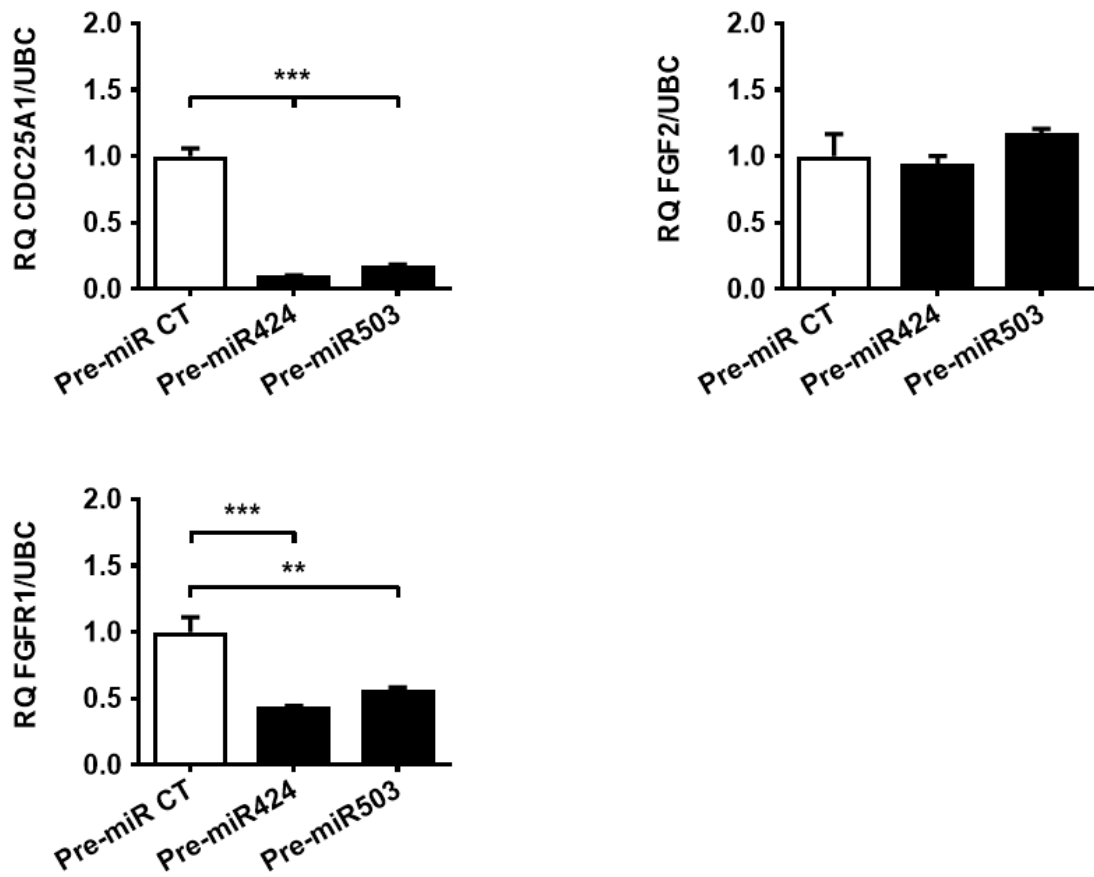
Figure 4.11: Pre-miR mediated overexpression of miR-424-5p and mirR-503-5p at day 3.

Expression of CDC25A, FGF2 and FGFR1 in HUVEC at day 3 after transfection with either Pre-miR-424 (15nM) or Pre-miR-503 (15nM) compared to a paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 technical replicates). Analysis by one-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Pre-miR Control (Pre-miR CT). Data represented as mean ±SEM.
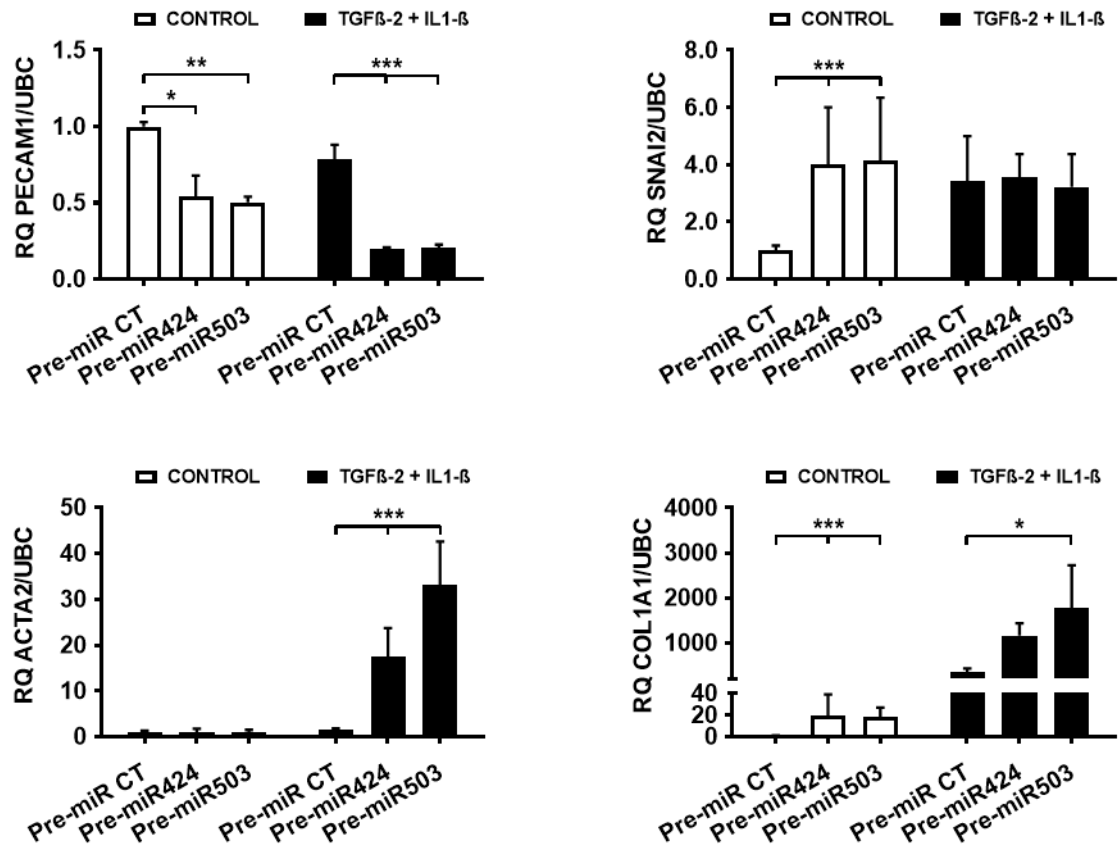
Figure 4.12: Pre-miR mediated overexpression of miR-424-5p and mirR-503-5p at day 7.
Expression of PECAM1, SNAI2, ACTA2 and COL1A1 in HUVEC at day 7 after transfection with either Pre-miR-424 (15nM) or Pre-miR-503 (15nM) ± EndMT compared to a paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 biological replicates). Analysis by two-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Pre-miR Control (Pre-miR CT). Data represented as mean ±SEM.
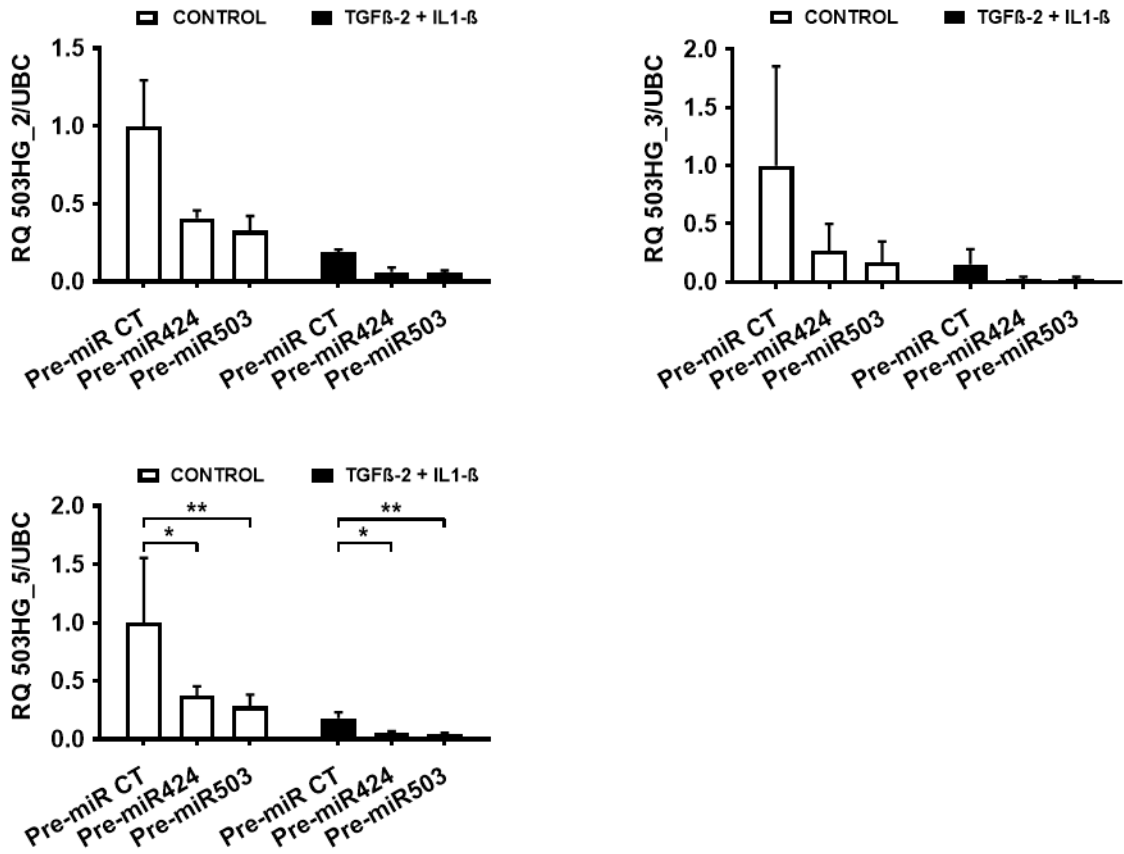
Figure 4.13: Pre-miR mediated overexpression of miR-424-5p and mirR-503-5p at day 7.
Expression of MIR503HG_2/3/5 in HUVEC at day 7 after transfection with either Pre-miR-424 (15nM) or Pre-miR-503 (15nM) ± EndMT compared to a paired control. RQ value for gene expression was quantified by qRT-PCR assay relative to RNU48 (n=3 biological replicates). Analysis by two-way ANOVA; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs paired Pre-miR Control (Pre-miR CT). Data represented as mean ±SEM.

4.4 Discussion

This chapter crucially demonstrates that both EndMT treatment and targeted knockdown of MIR503HG are accompanied by a reduction in miR-424/-503 expression. Conversely, overexpression of MIR503HG_2 had no effect on miRNA availability while still inhibiting EndMT. Moreover, targeted depletion of either miRNA using anti-miRs is not sufficient induce an EndMT phenotype. Interestingly, overexpression of either miRNA, while triggering the appearance of a moderate EndMT phenotype, was accompanied by a decrease in MIR503HG expression.

Studies on miR-424 report a similar positive correlation between increased miRNA expression and the appearance of EMT [469,470]. For example, stable overexpression of miR-424 was accompanied by increased production of mesenchymal markers, which could be reversed when miRNA expression was lowered [470].

Moreover, miR-424/503 have also been shown to supress FGF signalling by directly targeting FGF2 and FGFR1 [203]. FGF signalling itself is known to be critical in the maintenance of endothelial homeostasis and the prevention of EndMT [486,487]. For example, knockdown of FGFR1 alone in HUVEC activated TGFβ signalling and resulted in decreased CD31 levels and the increased mesenchymal marker expression [488]. Similarly, FGF2 has also been recently shown to prevent TGF-β-induced EndMT by repressing the α-SMA promoter activity via Elk1 [489]. However, a direct link between miR-424/503 overexpression and EndMT is yet to be shown and, as demonstrated by our own model, the miRNA cluster is clearly downregulated during transition. Nevertheless, while beyond the scope of this thesis, further studies should continue to explore the impact of miR-424/50 mediated suppression of FGF signalling and its role during EndMT.

Further, in chapter 3 we described a clear decrease in endothelial proliferation and G1/S-phase transition arrest after MIR503HG knockdown. In contrast, while no proliferation assays were conducted in this chapter, previous studies have shown that *in vitro* overexpression of miR-424/503 in non-tumorigenic epithelial cells was associated with decreased cell proliferation mediated by CDC25 targeting [490]. Notably, these changes were also accompanied by cell cycle arrest at the G1-phase [490]. miR-424 has also shown to be decreased in a variety of cervical cancer tissues and cells, which are often associated highly proliferative cell phenotypes [491,492]. Crucially, when overexpressed in these cell types, miR-424 suppressed proliferation and again blocked G1/S-phase transition [491,492]. Similarly, miR-503 has also been reported to suppress cell proliferation and metastasis in multiple cancers , including glioma [493], osteosarcoma [494], colorectal cancer [495], breast cancer [496], prostate cancer [497] and oesophageal squamous cell carcinoma [498].

While a miRNA/lncRNA cross-regulatory axis may be possible, the link between miR-424/503 overexpression and loss of MIR503HG was not confirmed in any of these studies. However, based on the results presented in this chapter, induction of EndMT during miRNA overexpression may be due to the accompanying downregulation of MIR503HG.

4.4.1 MIR503HG locus regulation

Despite the reported functional differences between the miR-424/503 cluster and MIR503HG, both in this chapter and throughout the prevailing literature, several studies suggest that miRNAs are generally transcribed in parallel with their host transcripts [499,500]. Unsurprisingly, the locus was shown to be consistently downregulated during EndMT. Similarly, the reported upregulation of MIR503HG

during hypoxia by Fiedler and colleagues was also accompanied by increased miR-424 and miR-503 availability [381], suggesting that the locus may be activated and repressed by the same transcriptional mechanisms.

In ECs, miR-424 and miR-503 expression was reported to be directly regulated by the peroxisome proliferator-activated receptor gamma (PPARγ), which is decreased in response to inflammatory factors known to promote EndMT, such as TNFα. Work by Lee *et al* demonstrated the expression of miR-424/503 was not only increased upon treatment using a PPARγ agonist, but that this could be reversed after PPARγ knockdown in HUVEC [480]. Additionally, the group described the presence of two binding sites in the miR-424/503 promotor region to which PPARγ binds acting as a transcriptional regulator [480]. Interestingly, PPARγ activation has been shown to attenuate EndMT both *in vitro* and *in vivo* [501–504].

Of note, studies by Llobet-Navas et al have also highlighted the existence SMAD and E2F-binding sites, which the group linked to TGF-β1 and TGF-β3 mediated upregulation of the miR-424/503 cluster in non-tumorigenic epithelial cell lines [485]. This was further emphasised in a recent study showing that the transcription factor E2F1 could bind to the miR-424 promoter and directly activate its transcription during G1/S transition, but not G2/M progression, in oesophageal squamous cell carcinoma lines [505]. Strikingly, this is at odds with the results presented throughout this chapter, showing that both miR-425 and miR-503 were significantly downregulated in ECs after TGF-β2 and IL-1β co-treatment. This highlights not only the complexity of MIR503HG locus but also a finely regulated, context and cell-specific transcription pattern.

Ultimately, it is clear that further in-depth dissection of the locus is necessary to

unravel any cross-regulatory functional mechanisms present. With its simplicity of implementation and high efficiency, a CRISPR/Cas9 gene editing system could be introduced to manipulate the locus in such precise manner. For example, this could be used to specifically target and delete the final 595 base-pair exonic region of MIR503HG_2 without disturbing the miRNA cluster upstream. Such a strategy would prospectively generate truncated MIR503HG transcripts containing only the pri-miRNA portion of the gene and lacking other key functional regions. Nevertheless, caveats may emerge from deleting such a large genomic region, leading to spurious results. Deletion may lead to unpredicted wider changes to chromatin organisation [506] or locally by influencing the expression of neighbouring genes by removing cis-regulatory elements overlapping with the lncRNA gene [507,508]. This was elegantly demonstrated by Engreitz and colleagues, which after manipulating 12 lncRNA-producing genomic loci, found that 5 of these had an impact on the expression of neighbouring genes in a manner not mediated by the lncRNA transcripts themselves [347].

Alternatively, in order to avoid any structural alteration of the locus, a short synthetic poly(A) (SPA) site could be introduced downstream of the miRNA locus thus prematurely stopping transcription of MIR503HG. This strategy takes advantage of a homology directed repair (HDR) a mechanism in cells used to repair double-strand DNA lesions such as those produced after Cas9-induced cleavage. In this case, a single sgRNA would be used to guide the Cas9 endonuclease, where it would then produce a double-stranded break (DSB). Paired with a single-stranded DNA (ssDNA) insert containing our gene of interest and homology arms with sequences matching the region on either side the break, this would enable homologous recombination by HDR after

cleavage [509–512]. However, while conceptually possible, this has never been attempted in such a manner to dissect a miRNA host gene locus and thus would require lengthy and thorough validation. Additionally, transfection of primary cell lines with CRISPR/Cas9 components can often be challenging due to their sensitivity and finite number of division cycles compared to immortalised cell lines. As such, this strategy may be largely limited to clonal cell lines or mutant animal models. We have further explored the application of these strategies in Chapter 5 to generate a viable MIR503HG mouse homolog knockout model.

In conclusion, although miR-424 and miR-503 have been shown to been involved in EMT and vascular remodelling [203,470,513], the data presented in this chapter has demonstrated that their contribution to EndMT was mostly associated with the loss of MIR503HG. Nevertheless, the possibility of cross-regulatory activity within the locus still exists.

Chapter 5: Identifying the role MIR503HG in vascular remodelling

5.1 Introduction

The study of EndMT during disease has remained largely obscure and mostly limited to observational or *in vitro* work. However, with the emergence of novel cell lineage tracing strategies and development of better disease models both *in vitro* and *in vivo*, this has allowed for new insights into the presence and relevance of EndMT during disease [14,52,53].

The reduced availability of ECs isolated from patients with vascular disease has often been one of the main limiting factors when studying the underlying molecular mechanisms of endothelial function and dysfunction. However, the identification of blood outgrowth endothelial cells (BOECs), generated from circulating endothelial progenitors in adult peripheral blood, has started to emerge as a way to circumvent these limitations offering a patient-derived EC surrogate. In the study of PAH, for example, a seminal paper by Toshner and colleagues demonstrated that BOEC derived from HPAH patients with mutations in the gene encoding for BMPRII presented with an hyperproliferative phenotype and impaired ability to form vascular networks [423]. Later studies by the same group went on to show that normal proliferation could be restored in these cells by overexpressing miR-124, or by silencing its molecular target PTPB1 [514]. Interestingly, despite the link between EndMT and BMPR2 deficiency, this is yet to be studied in PAH patient-derived BOEC [120,126,515].

Further, early reports of EndMT during disease have largely focused on histological analysis of tissue samples from both patients and animal models showing the presence of cells with a mixed mesenchymal and endothelial phenotype seen throughout the remodelled vasculature. This is true for several publications, not only on PAH [10,119,120], but also idiopathic portal hypertension [516], atherosclerosis [29], vein graft restenosis [14]

and in artery transplant rejection [487]. However, while these reports may be important to demonstrate the presence of active EndMT during disease and indicative of its contribution to vascular remodelling, they cannot show the full extent of EndMT in these vessels. This is largely due to the fact that despite co-expressing endothelial and mesenchymal markers during the early stages of EndMT, cells that have complete transition will be undiscernible from the surrounding mesenchymal population. This has led to the eventual widespread use of lineage tracing transgenic animal models in the study of vascular remodelling, which allow for the continued expression of endothelial-specific reporter genes, regardless of subsequent changes in cellular phenotype. When implemented along with a 35-day vein graft mouse model, prone to the development of an occlusive neointima, Cooley and colleagues demonstrated a progressive increase in the number of reporter cells co-expressing endothelial and early SMC markers. By day 35, however, reporter cells within the neointima mostly lost their expression of endothelial markers and thus would be undetectable using classical histological staining strategies [14]. This was also recently established in a mouse model of PH, where over 14% of lung reporter cells of endothelial origin had undergone complete EndMT and no longer colocalised with $CDH5^+$ ECs. Interesting, cells that had undergone partial EndMT and still retained endothelial markers only accounted for 5% of the population [172]. Additionally, when isolated and cultured, complete EndMT reporter cells presented with a significantly higher proliferative and migratory phenotype than non-endothelial lung mesenchymal cells [172].

Lastly, despite its high evolutionary conservation in mammals, the growing number of publications on MIR503HG have been largely limited to *in vitro* work while its function *in vivo* remains unexplored. In contrast, with publications ranging from

overexpression during PH to knockdown animal models, the impact of miR-424 and miR-503 manipulation *in vivo* has been thoroughly explored over the past decade. However, as demonstrated throughout this thesis, there may be a cross-regulatory axis between the two miRNAs and the host lncRNA, making it particularly difficult to draw definite conclusions about their function *in vivo*. Nonetheless, the emergence of CRISPR/Cas9-based genome manipulation techniques have now made it possible to thoroughly dissect and study such a complex genomic locus.

5.2 Aims

Given the wide range of publications reporting on the presence of active EndMT during PAH, along with the advancements in EC lineage tracing and disease modelling strategies, the aims of this chapter were as follows:

- To explore the link between the loss of MIR503HG and vascular remodelling in PAH patients.

- To elucidate the MIR503HG expression profile *in vivo* during vascular remodelling in association with EndMT.

- Target MIR503HG expression *in vivo*.

- To further dissect the MIR503HG locus *in vivo* using CRISPR/Cas9.

5.3 Results

5.3.1 MIR503HG expression profile during disease

As described throughout this thesis, the decreased expression of MIR503HG has been consistently linked to the appearance of an EndMT profile *in vitro*. This, however, does not provide a functional association between MIR503HG and EndMT during vascular remodelling or the expression in patient samples vs control to quantify association in human disease. As such, it is crucial that we also assess changes to MIR503HG expression in patient samples and relevant disease models.

5.3.1.1 Blood Outgrowth Endothelial Cells

The use of BOECs derived from the adult circulation has proven to be a clinically relevant, patient-derived, EC surrogate for the investigation of vascular disease, particularly PAH [514]. As such, in order to assess EC transcriptional differences in PAH, we compared BOECs isolated from PAH patients harbouring BMPR2 mutations and from healthy controls. As with our previous *in vitro* work, these were cultured in endothelial growth media for a total of 7 days. Notably, BOECs derived from PAH patients presented with a partial, but meaningful, EndMT phenotype when compared to healthy controls. As demonstrated by qRT-PCR, these cells showed an enhanced expression of *ACTA2* and *COL1A1*, and reduced expression of *PECAM1* (Figure 5.1). However, due to limited cell availability, this was not confirmed at the protein level. Nonetheless, further qRT-PCR analysis revealed that MIR503HG_ 2 expression alone was significantly lowered in cells from PAH patients compared to controls, again showing an association between the lncRNA and EndMT (Figure 5.2).
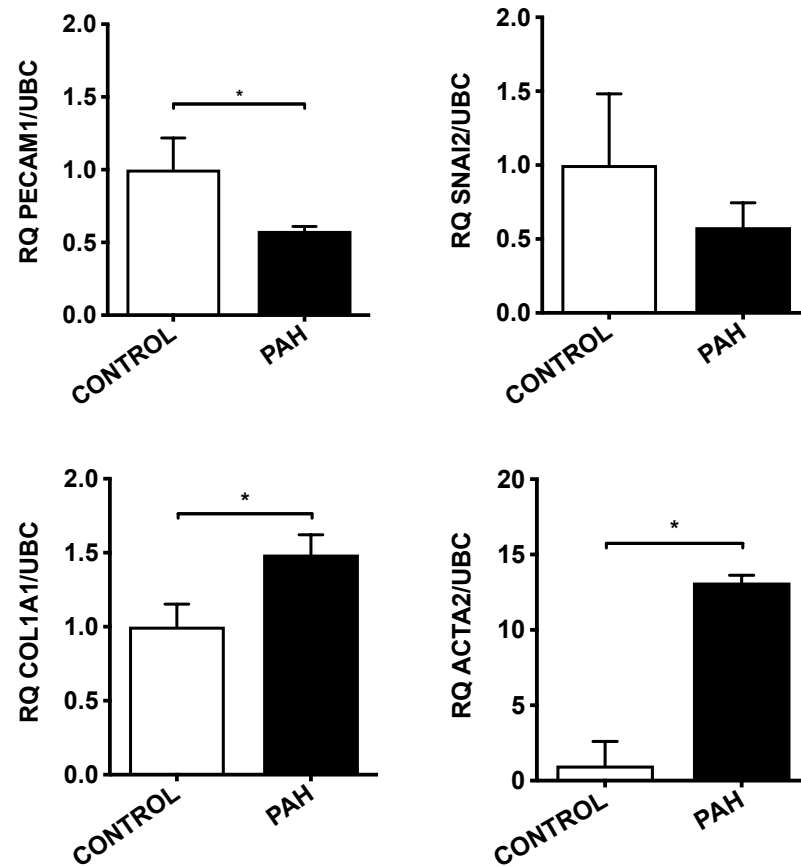
Figure 5.1: EndMT profile analysis in PAH patient-derived BOEC.

Expression of PECAM1, SNAI2, ACTA2 and COL1A1 in PAH patient-derived BOEC at day 7 compared to healthy controls. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=4 biological replicates). Analysis by two-tailed t-test; *$p \leq 0.05$ vs control. Data represented as mean ± SEM.
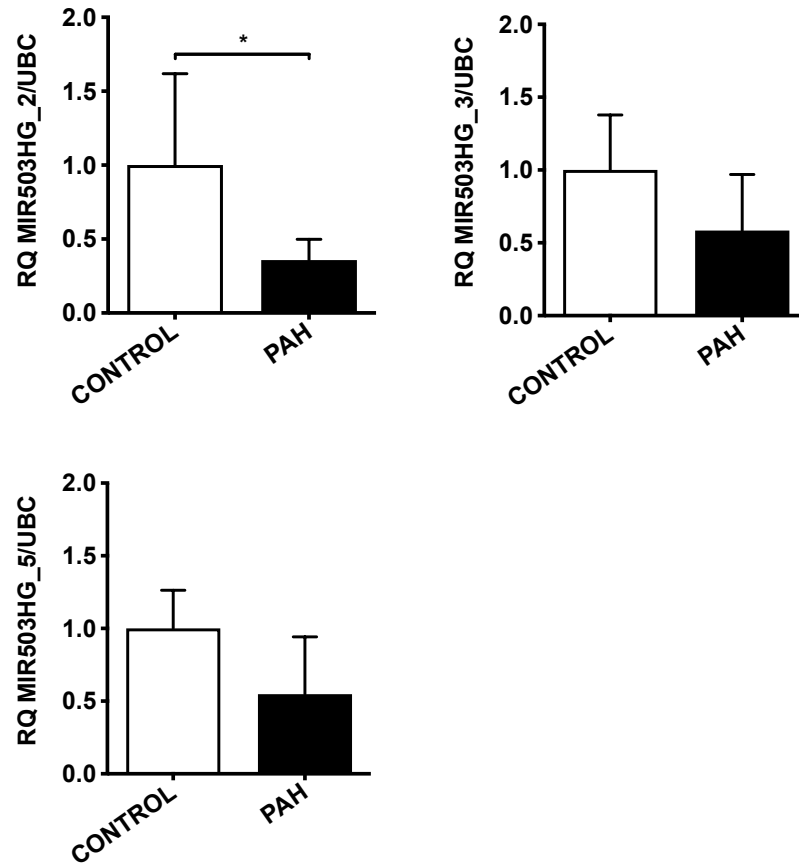
Figure 5.2 : MIR503HG expression in PAH patient-derived BOEC.

Expression of MIR503HG_2/3/5 in PAH patient-derived BOEC at day 7 compared to healthy controls. RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=4 biological replicates). Analysis by two-tailed t-test; *p≤0.05 vs control. Data represented as mean ± SEM.

5.3.1.2 Pulmonary arterial hypertension (PAH)

Having shown that MIR503HG expression is significantly reduced in ECs from PAH patients in association with an EndMT profile, we sought to examine if these changes were also present during vascular remodelling in human PAH. In order to determine the association between the downregulation of MIR503HG and vascular remodelling during PAH, we analysed lung tissue sections collected from patients diagnosed with idiopathic PAH (IPAH) and healthy donor control lungs. MIR503HG expression was detected by *in situ* hybridisation (ISH) using a small, digoxigenin (DIG)-labelled, RNA probe designed to target an 84bp exonic region common to all MIR503HG transcripts. To further assess its vascular localisation, ISH for MIR503HG expression was performed along with immunofluorescence staining for endothelial and mesenchymal markers on matching serial sections.

In healthy control lungs, we found that MIR503HG expression was present throughout the lung vasculature, co-localising with von Willebrand Factor (vWF) in the endothelium (Figure 5.3 A-B; Figure 5.4 A-B). Conversely, this expression was largely absent from the PAH lung vasculature, particularly in small remodelled vessels with constrictive and complex lesions (Figure 5.3 A-B; Figure 5.4 A-B). Crucially, these lesions were associated not only with generalised medial hypertrophy but also altered intimal endothelial structure (Figure 5.3 B; Figure 5.4 B). Ultimately, this demonstrates not only that MIR503HG is expressed in the human lung vasculature but also that its loss is associated with vascular remodelling found in PAH patients.
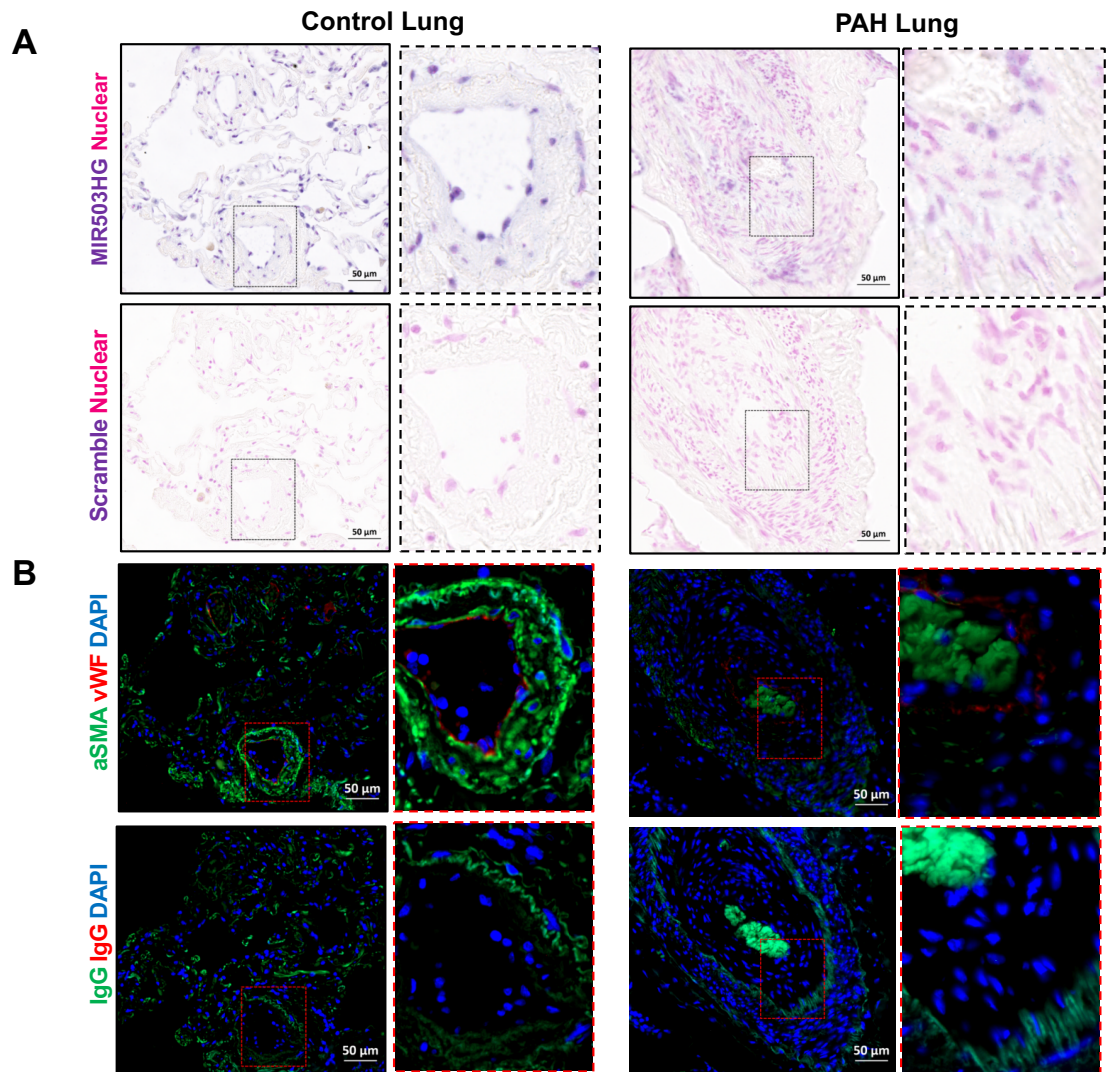
Figure 5.3: MIR503HG expression during lung vascular remodelling in PAH patients.

Representative images from human PAH patient and healthy control lung tissue samples on serial sections. **(A)** Tissue *in situ* hybridisation for MIR503HG expression with brightfield staining showing MIR503HG (purple) and cell nucleus (pink). **(B)** Accompanying tissue immunofluorescence for smooth muscle cells (αSMA, green), ECs (vWF, red), cell nucleus (DAPI, blue) and IgG controls. Dotted squares denote high-power view of vessels. Images acquired using an Axioscan slide scanner and processed with Zen software. Scale bar 50μm. This work was performed in collaboration with Dr Jessica Scanlon.

Figure 5.4: MIR503HG expression during lung vascular remodelling in PAH patients.

Representative images from human PAH patient and healthy control lung tissue samples on serial sections. **(A)** Tissue *in situ* hybridisation for MIR503HG expression with brightfield staining showing MIR503HG (purple) and cell nucleus (pink). **(B)** Accompanying tissue immunofluorescence for smooth muscle cells (αSMA, green), ECs (vWF, red), cell nucleus (DAPI, blue) and IgG controls. Dotted squares denote high-power view of vessels. Images acquired using an Axioscan slide scanner and processed with Zen software. Scale bar 50μm. This work was performed in collaboration with Dr Jessica Scanlon.

5.3.1.3 Sugen 5146/ Hypoxia model of PAH

Next, we sought to confirm that the changes in MIR503HG expression were associated not only with vascular remodelling but also active EndMT. In order to do this, an inducible endothelial tracking (Ind.EndoTrack) transgenic mouse line (Cdh5-Cre-ER^{T2}-TdTomato) was generated. In brief, this was achieved by crossing Cdh5-Cre-ER^{T2} (strain Tg(Cdh5-cre/ER^{T2})^{1Rha}) with a ROSA-TdTomato reporter mouse line (strain B6.Cg-Gt(ROSA)26Sor$^{tm9(CAG-tdTomato)Hze}$/J). Here, induction of Cre-recombinase activity by tamoxifen allows for the VE-cadherin (Cdh5) promotor to direct endothelial-specific gene expression of the TdTomato reporter gene in pan-EC regardless of subsequent changes in cellular phenotype. Next, to model PH *in vivo*, we used a well-established Sugen 5146/Hypoxia (SuHx) mouse model of PH in which active vessel remodelling has previously been shown to involve EndMT [119,120]. The presence of active EndMT was further confirmed by our group using confocal microscopy, which established that cells in the pulmonary vascular intima of SuHx mouse lungs presented with a mixed mesenchymal/endothelial phenotype (Figure 5.5). After 1 week of tamoxifen treatment and a 2 week washout period, the Ind.EndoTrack mice were subjected to either SuHx or normoxic conditions for a total of 3 weeks, total and TdT^{+} lung cells were then isolated by flow cytometry (Figure 5.6). TdT^{+} cells isolated from SuHx lungs presented with a distinct EndMT profile, showing increased expression of the transcription factor SNAI2, as well as the mesenchymal markers ACTA2 and COL1A1. This was accompanied by the loss of endothelial specificity as reflected by decreased PECAM1 expression (Figure 5.7). Notably, expression of the MIR503HG mouse homolog, Gm28730, was found to be significantly reduced in the SuHx TdT^{+} cell population (Figure 5.8).

Comparable to our human PAH *in situ* analysis, Gm28730 expression was decreased in SuHX whole lung RNA samples with a mean relative quantification (RQ) of 0.5 (Figure 5.8). However, these changes were more pronounced in TdT[+] cell of endothelial origin, where Gm28730 expression was largely absent with mean RQ of 0.1 (Figure 5.8).

NORMOXIA/VEHICLE
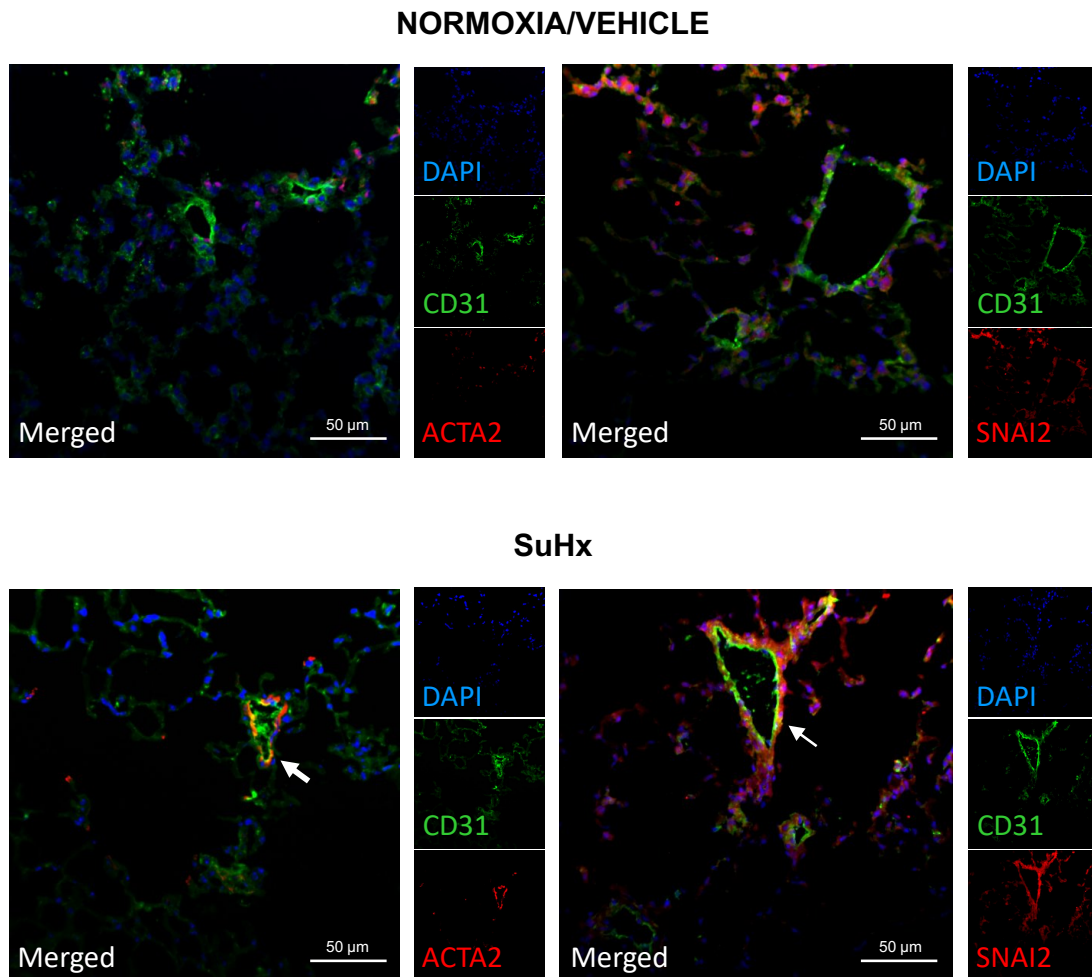
SuHx

Figure 5.5: EndMT profile during lung vascular remodelling in PH mouse model.
Representative immunofluorescence staining of normoxia/vehicle or SuHx mouse lung tissue samples for CD31 (green), Dapi (blue) and ACTA2 or SNAI2 (red). Single-channel (right) and merged (left) images acquired with Andor Revolution XDi spinning disk confocal microscope and analysed with Image J Software. Scale bar 50 μm. This work was performed by Dr Axelle Caudrillier using tissue samples provided by Dr Lin Deng.
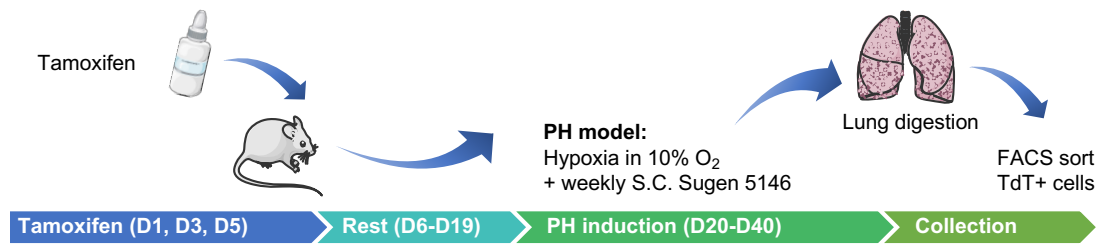
Figure 5.6: Diagram Sugen 5146/Hypoxia PH model in inducible endothelial-lineage tracing mice.
Induction of Cre-recombinase activity was achieved by administering 400 mg/kg body weight per dose of tamoxifen dissolved in sterile corn oil by gavage on 3 alternate days for 5 days (D1, D3 and D5). This was followed by a two week (D6-D19) washout period before initiating the Hypoxia/SU5416 PH model for a total of 3 weeks (D20-D40). Lung tissues were then harvested for cell isolation by flow cytometry.
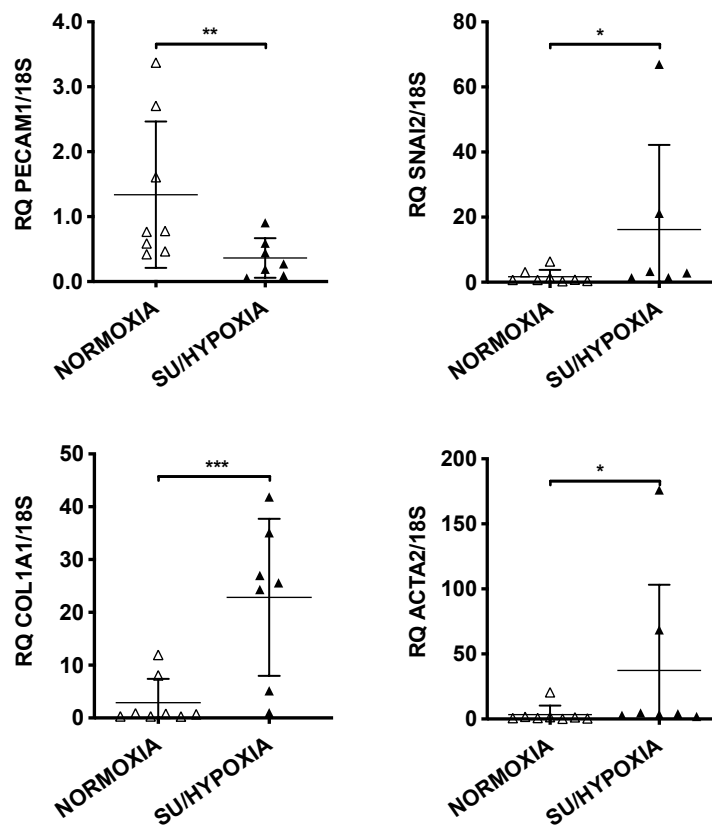
Figure 5.7: EndMT profile analysis during lung vascular remodelling in SuHx mice.
Expression of Pecam1, Snai2, Acta2 and Col1a1 in TdT+ cells isolated from normoxia/vehicle (n=8mice/group) or SuHx mouse lung tissue (n=7mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=4 biological replicates). Analysis by two-tailed t-test; *$p \leq 0.05$, ** $p \leq 0.01$, and *** $p \leq 0.001$ vs normoxic control (Normoxia). Data represented as individual data points and mean ± SEM.
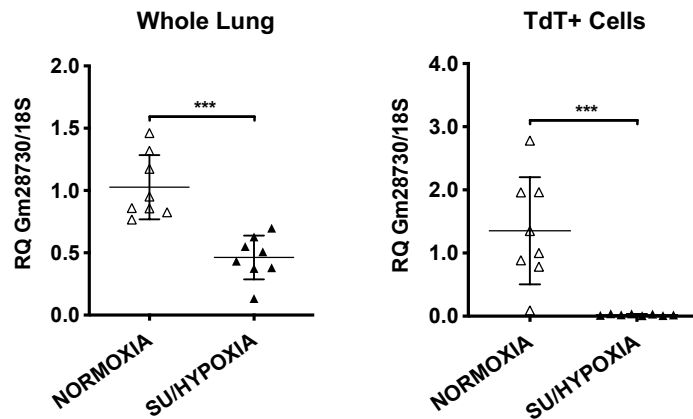
Figure 5.8: MIR503HG expression during lung vascular remodelling in SuHx mice.
Expression of MIR503HG_2/3/4 in TdT+ cells isolated from normoxia/vehicle (n=8mice/group) or SuHx mouse lung tissue (n=7mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to UBC (n=4 biological replicates). Analysis by two-tailed t-test; *p≤0.05, ** p≤0.01, and *** p≤0.001 vs normoxic control (Normoxia). Data represented as individual data points and mean ± SEM.

5.3.2 Overexpression of MIR503HG *in vivo*

As both the human and *in vivo* disease data sets presented so far were largely associative, showing no causation between the loss of MIR503HG and EndMT, we attempted to mimic our initial *in vitro* results by overexpressing the lncRNA in SuHX mice only and assess its impact on lung ECs. Crucially, as a proof of concept, we first tested lentiviral uptake efficiency in the lung vasculature of normoxic mice. A GFP lentiviral construct was delivered intranasally to 8 to 10-week-old C57BL/6 mice, which were then given a 2 week rest period prior to tissue collection in order to confirm that viral overexpression could be maintained throughout the period required for the induction of PH during subsequent studies (Figure 5.9). Flow cytometry analysis confirmed that 5% of isolated lung cells were GFP+, of which 13.7% were of endothelial origin (CD31+) (Figure 5.10).

Despite our use of a global Gm28730 probe for RT-qPCR analysis, so far there are no publications on the mouse Gm28730 locus, and its predicted transcript sequence arrangement is yet to be validated. As such, given the depth of sequence conservation across vertebrates, we chose to overexpress the human MIR503HG_2 transcript used throughout Chapter 3 of this thesis. Either LNT_503HG or LNT_CT were delivered intranasally to 8 to 10-week-old C57BL/6 mice, which were then given a 2 day rest period prior to PH induction by SuHx (Figure 5.11). Finally, both endothelial (CD31$^+$) and non-endothelial (CD31$^-$) lung cells were isolated by flow cytometry and the expression of MIR503HG/Gm28730 and EndMT markers was assessed (Figure 5.12). In line with our initial GFP proof of concept, RT-qPCR analysis showed that the expression of the human MIR503HG_2 was significantly higher in CD31$^+$ lung EC of LNT_503HG mice compared to the LNT_CT group (Figure 5.13). Interestingly, overexpression of MIR503HG also led an increased availability of the endogenous mouse homologue, Gm28730 (Figure 5.13). Concordantly, while there were no changes to *Pecam1* expression, we observed a significant downregulation of the mesenchymal markers *Acta2* and *Col1a1* upon MIR503HG overexpression in CD31$^+$ ECs (Figure 5.14). Interestingly, while not significant, an increase in *Snai2* was observed in CD31$^+$ cells. These results were then compared to CD31$^-$ cell, which despite the increased availability of MIR503HG in non-ECs from LNT_503HG mice (Figure 5.15), showed no changes in Gm28730 or mesenchymal maker expression (Figure 5.15; Figure 5.16).

Finally, this initial set of results provides robust preliminary evidence for a MIR503HG-mediated mechanism in the suppression of EndMT during vascular remodelling.
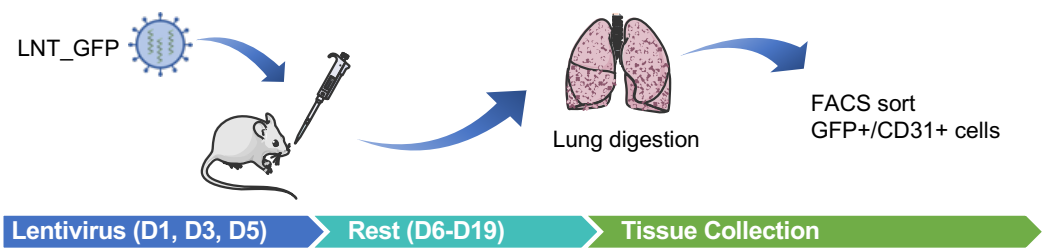
Figure 5.9: Diagram of GFP lentivirus intranasal delivery to mice lung cells.

GFP lentiviral construct was delivered by administering intranasally a 25uL dose at 4.00E+08 piu/mL on 3 alternate days for 5 days (D1, D3 and D5). This was followed by a 2 week (D6-D19) rest period before lung tissues were then harvested for cell isolation by flow cytometry.
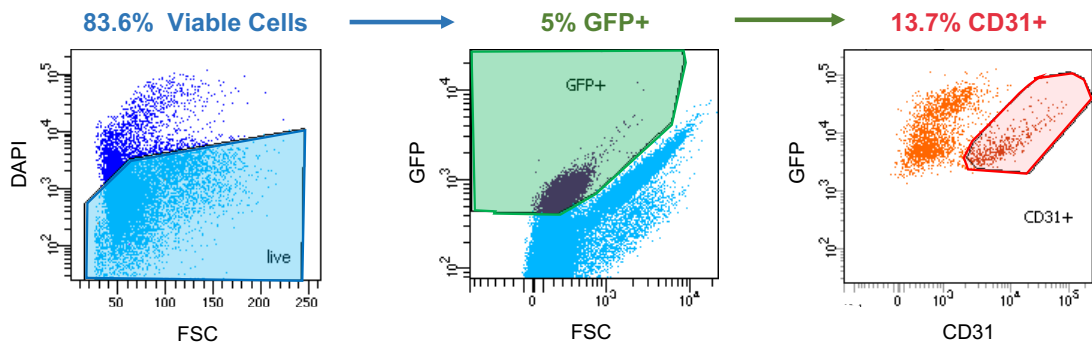
Figure 5.10: Flow cytometry gating strategy used to identify mouse lung GFP+ cells.

Viable (live) total lung cells were first selected by gating for DAPI negative events. Transduced cells were then selected by gating for GFP+ events, from which ECs were selected by gating for CD31+ events. Flow cytometry analysis performed by Dr Axelle Caudrillier using samples provided by Dr Lin Deng.
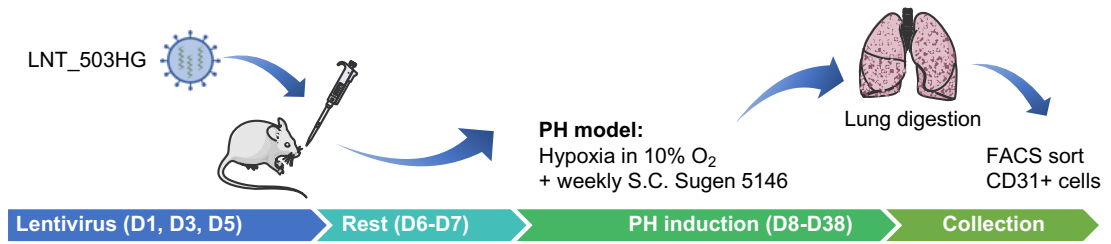
Figure 5.11: Diagram of Sugen 5146/Hypoxia PH model after MIR503HG_2 overexpression.
Lentiviral constructs were delivered by administering intranasally a 25uL dose at 4.00E+08 piu/mL on
3 alternate days for 5 days (D1, D3 and D5). This was followed by a 2 day (D6-D7) rest period before
initiating the Hypoxia/SU5416 PH model for a total of 3 weeks (D20-D40). Lung tissues were then
harvested for cell isolation by flow cytometry.
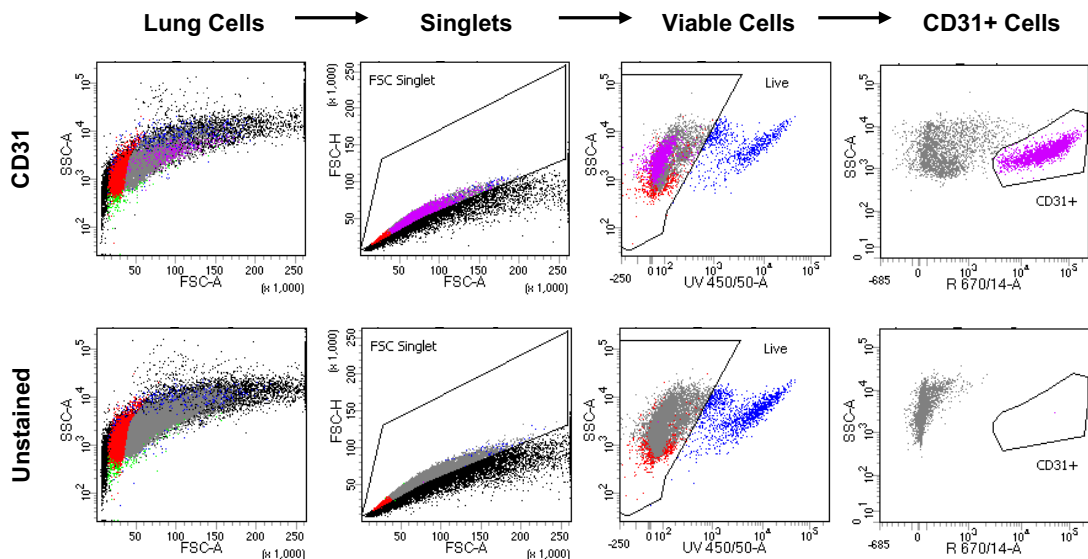
Figure 5.12: Flow cytometry gating strategy used to identify SuHx mouse lung CD31$^+$ cells.
Total lung cells were first separated from debris by plotting forward versus side scatter. Single cells
(singlets) were then selected using forward scatter area (FSC-A) versus forward scatter height (FSC-H)
and side scatter area (SSC-A) versus side scatter height (SSC-H). From the singlet population, viable
(live) cells were selected by gating for DAPI negative events. ECs from the viable cell population were
finally selected by gating for CD31$^+$ events compared to control unstained samples.
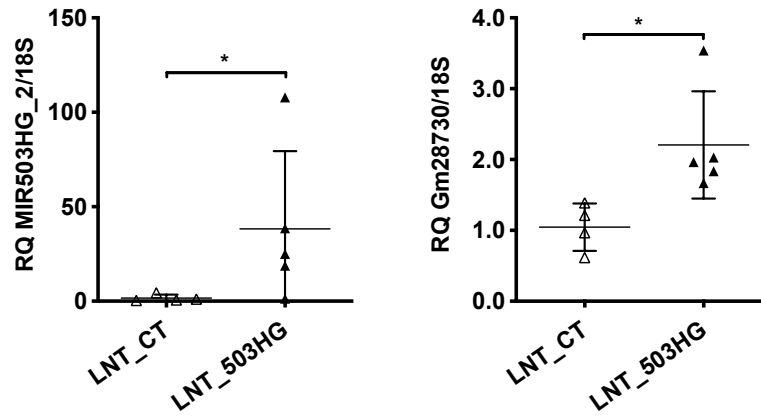
Figure 5.13: MIR503HG_2 overexpression in SuHx mouse lung CD31$^+$ cells.

Expression of MIR503HG_2 and Gm28730 in CD31$^+$ cells isolated only from SuHx mouse lung tissue after intranasal delivery of LNT_MIR503HG compared to LNT_Control (n=5 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to 18S. Analysis by two-tailed t-test; *p≤0.05 vs LNT_Control. Data represented as individual data points and mean ± SEM.
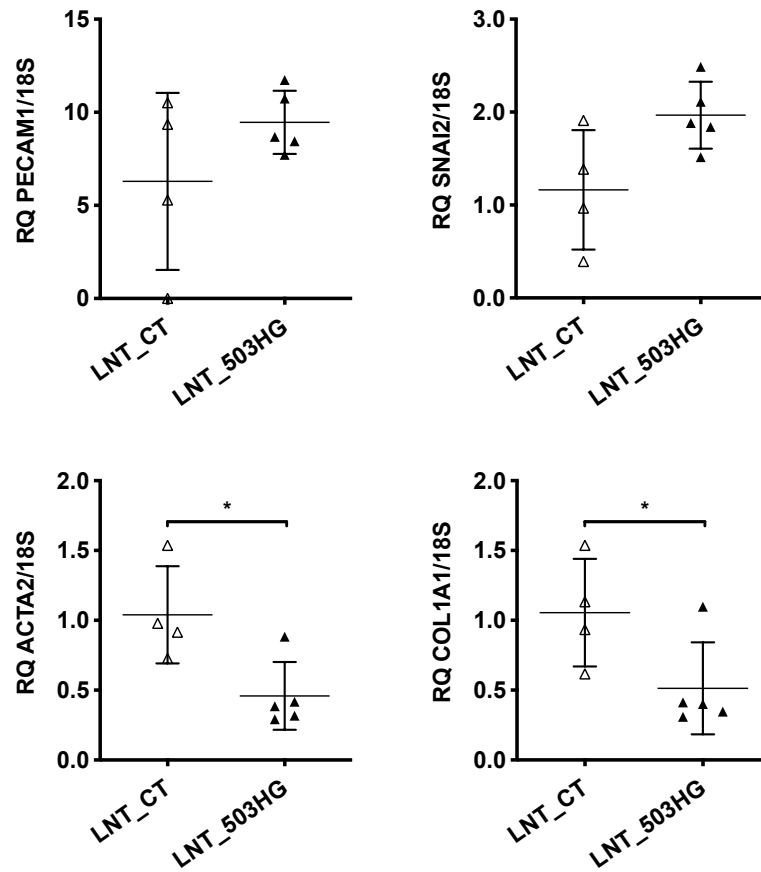
Figure 5.14: MIR503HG_2 overexpression in SuHx mouse lung CD31[+] cells.

Expression of Pecam1, Snai2, Acta2 and Col1a1 in CD31[+] cells isolated only from SuHx mouse lung tissue after intranasal delivery of LNT_MIR503HG compared to LNT_Control (n=5 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to 18S. Analysis by two-tailed t-test; *p≤0.05 vs LNT_Control. Data represented as individual data points and mean ± SEM.
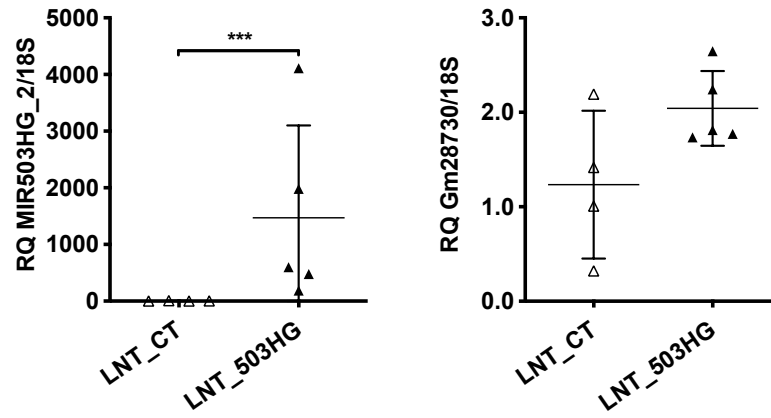
Figure 5.15: MIR503HG_2 overexpression in SuHx mouse lung CD31⁻ cells.

Expression of MIR503HG_2 and Gm28730 in CD31⁻ cells isolated only from SuHx mouse lung tissue after intranasal delivery of LNT_MIR503HG compared to LNT_Control (n=5 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to 18S. Analysis by two-tailed t-test; ***p≤0.001 vs LNT_Control. Data represented as individual data points and mean ± SEM.
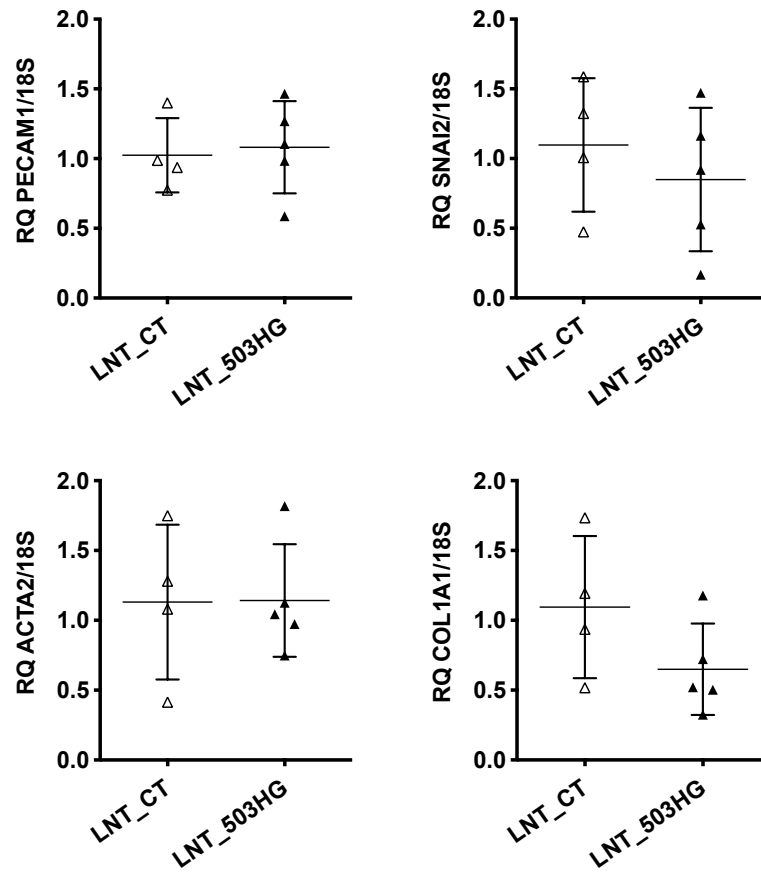
Figure 5.16: MIR503HG_2 overexpression in SuHx mouse lung CD31⁻ cells.

Expression of Pecam1, Snai2, Acta2 and Col1a1 in CD31⁻ cells isolated only from SuHx mouse lung tissue after intranasal delivery of LNT_MIR503HG compared to LNT_Control (n=5 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to 18S. Analysis by two-tailed t-test; *p≤0.05 vs LNT_Control. Data represented as individual data points and mean ± SEM.

5.3.3 Gm28730 knockout mouse model

In order to further dissect the lncRNA/miRNA axis and understand the role of MIR503HG *in vivo*, a CRISPR/Cas9 system was used to generate a Gm28730 null allele in mice. To avoid unpredicted alterations to the locus structure and function, often seen with large genomic deletions [347,506–508], we have developed a Gm28730[-/-] mutant containing a synthetic polyadenylation signal aimed at prematurely stopping transcription of Gm28730 without disrupting miRNA expression. Based on the work of Ballarino *et al,* which generated a functional knockout for the lncRNA Charme in mice [517], our strategy used a SPA-MAZ cassette containing a compact and highly efficient synthetic poly(A) (SPA) [518] followed by two MAZ protein binding sites known to further promote transcriptional termination [519,520]. The SPA-MAZ site was introduced within Exon 1 of the Gm28730 locus, down-stream from the miR-322(424)/503 cluster allowing for the expression of the miRNA alone (Figure 5.17).

Despite current advancements to CRISPR/Cas9-sgRNA systems, the on-target efficiency of sgRNA to distinct DNA loci can vary widely. As such, prior to mutant generation, the effectiveness of three sgRNAs designed to target Exon 1 of Gm28730 was first tested *in vitro*. This was done using human embryonic kidney (HEK) 293 cells co-transfected with a pX330 vector carrying the sgRNA and a mCherry/eGFP dual reporter vector containing the corresponding target sequence. Of the tested guides, sgRNA #2 showed the highest efficiency with 21.4% of cells transfected cells positive for both fluorescent markers, which was then selected for downstream application (Figure 5.18).

Genomic insertion of the SPA-MAZ cassette was first confirmed using DNA extracted from each F0-generation mouse. This was analysed by PCR amplification using two

different primer-pairs targeting either the cassette sequence or the genomic region on either side of it. Gel electrophoresis analysis and subsequence sequencing of the PCR products showed that the insertion of the entire SPA-MAZ sequence was present on a single allele of pup #6, which was selected as a founder for successive breeding (Figure 5.19). The Gm28730 null allele was then backcrossed and carefully monitored for a total of four generations to obtain homozygous animals.

Preliminary analysis of RNA isolated from heart and lung tissue of 8-week-old Gm28730$^{-/-}$ mutants confirmed a significant a depletion in Gm28730 availability compared to WT littermates (Figure 5.20). Crucially, the expression of miR-322(424) and miR-503 was not altered in Gm28730$^{-/-}$ mutants (Figure 5.21 and Figure 5.22).

While beyond the scope of this thesis, extensive work is still required to fully characterise and phenotype the Gm28730$^{-/-}$ mutant both during development and disease.
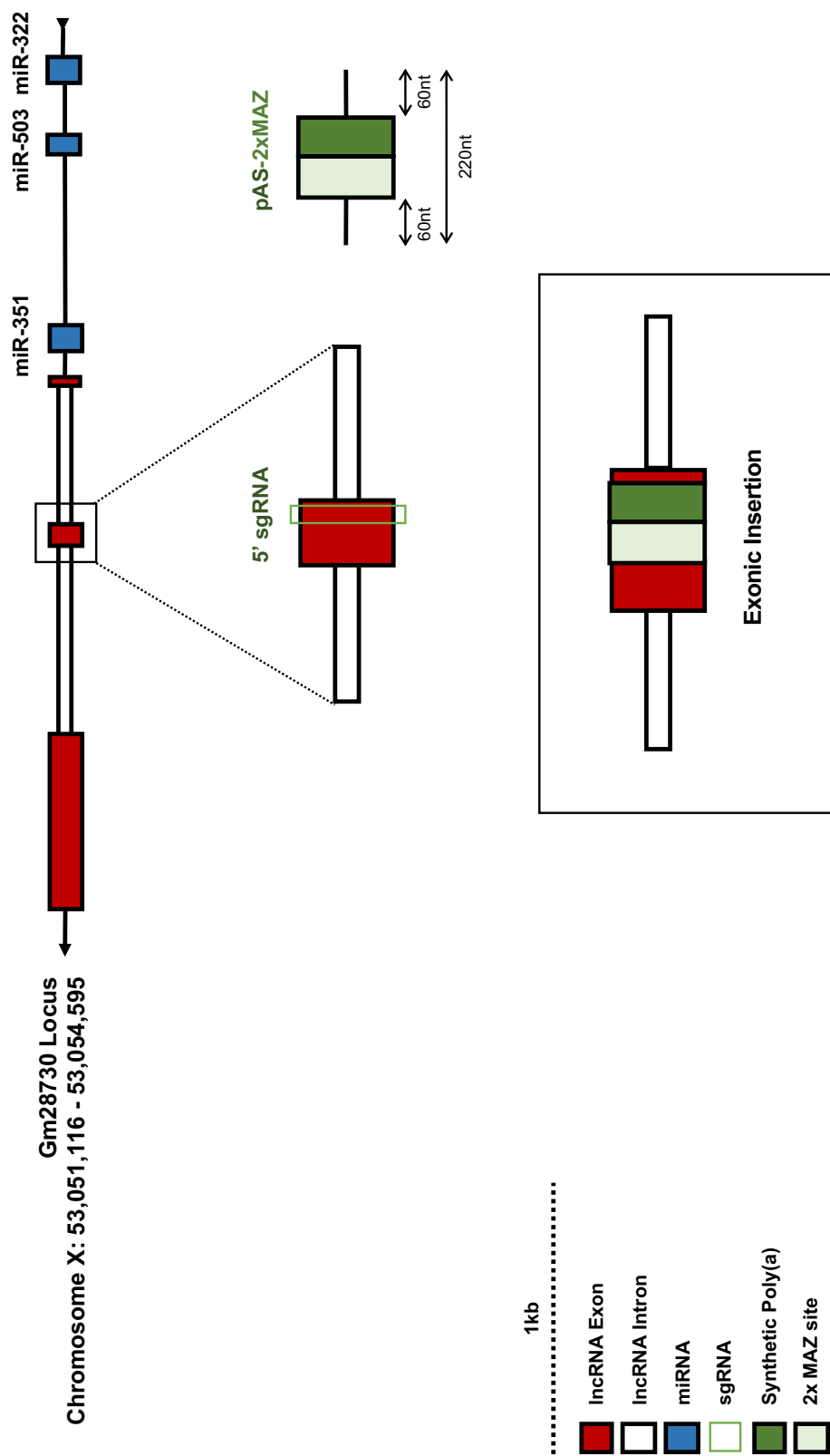
Figure 5.17: Gm28730 knockout mice generation strategy.

Diagram of Gm28730 locus shown together with sgRNA target position on exon 1. Design of the single-stranded oligodeoxynucleotide (ssDNA) containing two 60 bp homology arms (black lines) and a 100 bp SPA-MAZ insertion cassette along with predicted exonic insertion structure.
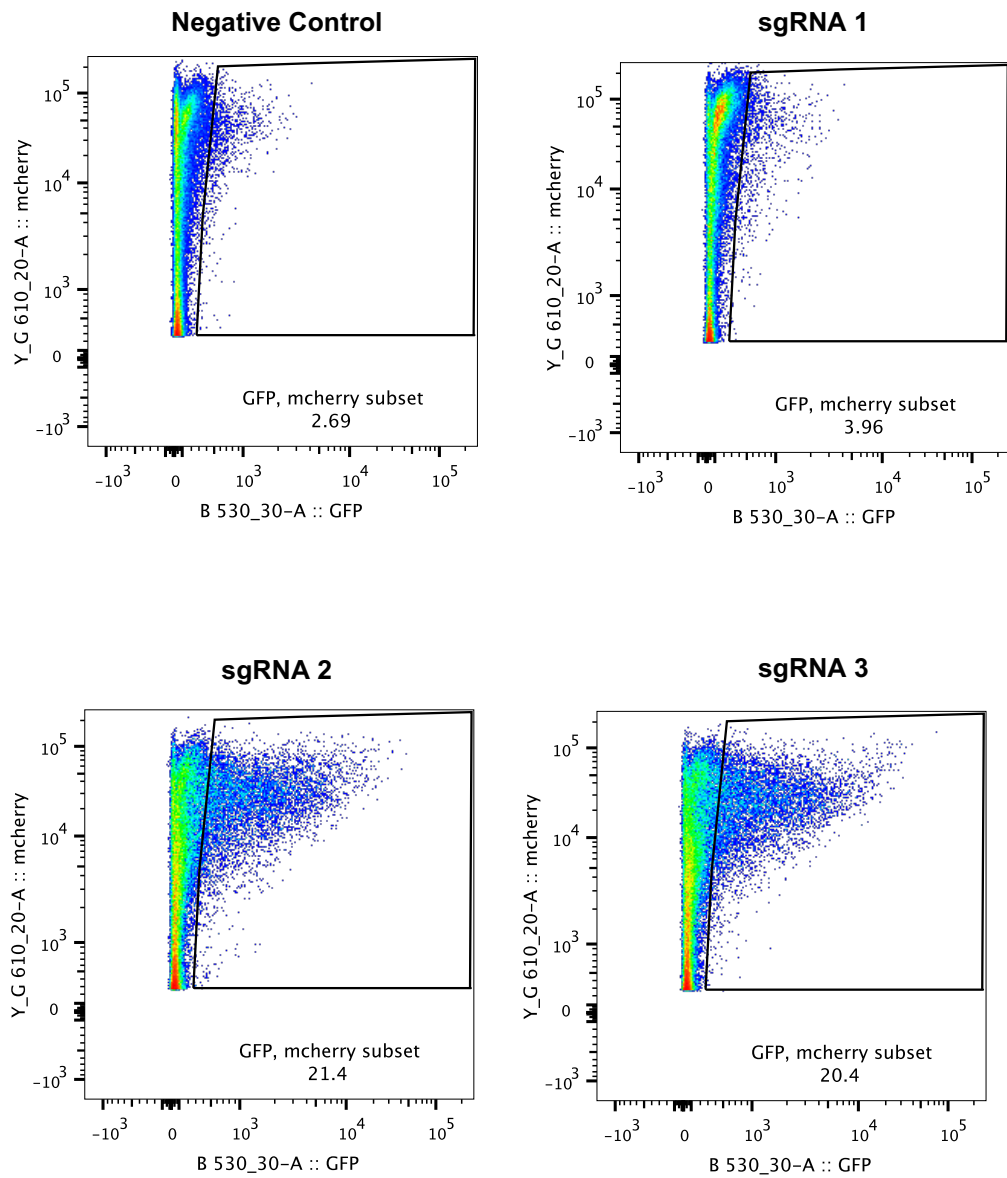
Figure 5.18: sgRNA on-target efficiency.

Flow cytometry analysis of sgRNA on-target efficiency in HEK293 at day 3 after co-transfection with px330 and pBS-SK-mCherryY66HeGFP-forw vectors. sgRNA/Cas9 on-target cleavage was assessed based on mCherry$^+$/GFP$^+$ cell subset compared to a no-sgRNA negative control. This work was performed in collaboration with Dr Francesco Severi.

A

B
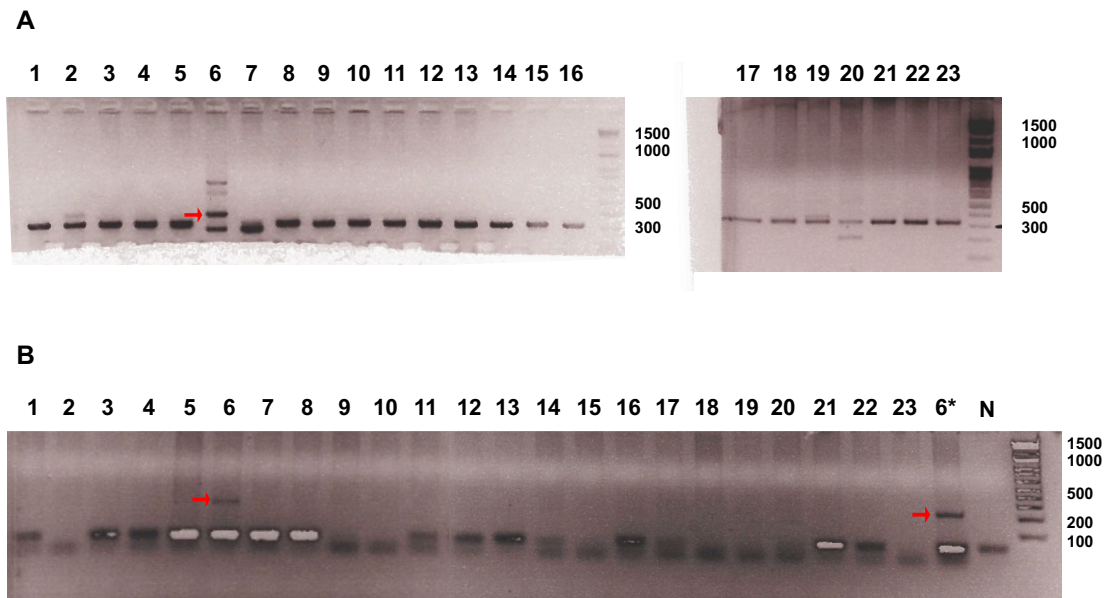
Figure 5.19: Gm28730 knockout F0-generation genotyping.

Gel electrophoresis analysis of PCR amplification products from F0-generation mice DNA. PCR primers-pair were designed to amplify (A) the genomic region on either side of the insertion site or (B) the SPA-MAZ cassette sequence. SPA-MAZ cassette insertion was confirmed on a single allele in pup #6 (red arrow), which was selected for successive breeding. This work was performed in collaboration with Dr Francesco Severi.
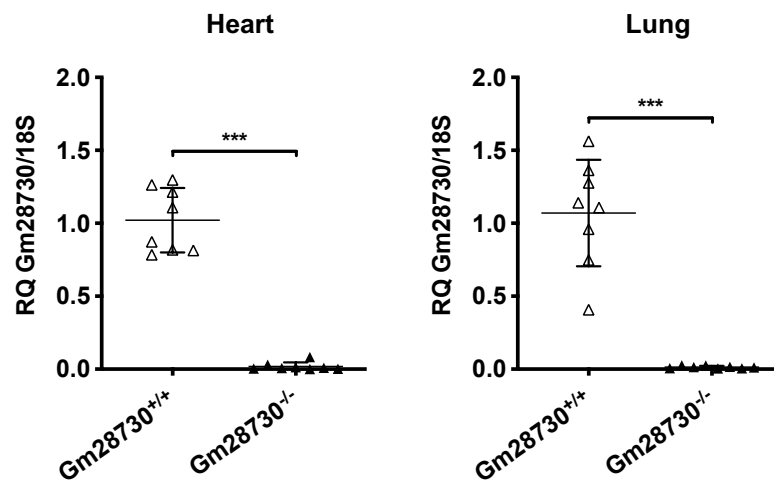
Figure 5.20: Gm28730 expression in Gm29730$^{-/-}$ mice.

Expression of Gm28730 in heart and lung tissue from homozygous Gm28730$^{-/-}$ mice compared to wildtype Gm28730$^{+/+}$ littermates (n=8 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to 18S. Analysis by two-tailed t-test; ***p$\leq$0.001 vs Gm28730$^{+/+}$. Data represented as individual data points and mean ± SEM.
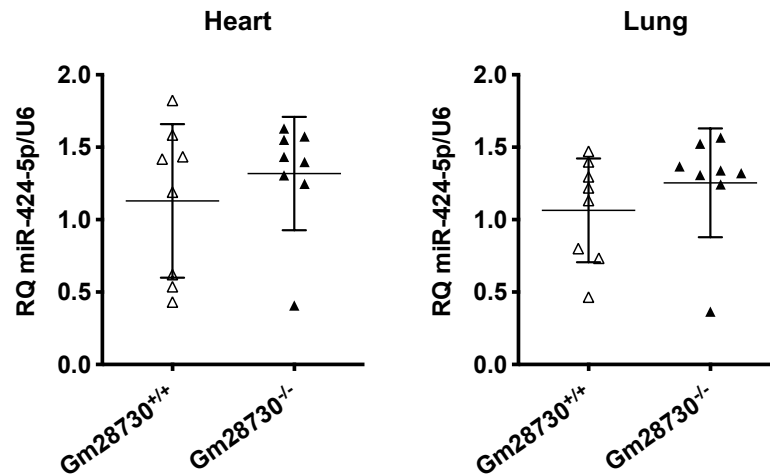
Figure 5.21: miR-424(322) expression in Gm29730^{-/-} mice.

Expression of miR-424(322)-5p in heart and lung tissue from homozygous Gm28730^{-/-} mice compared to wildtype Gm28730^{+/+} littermates (n=8 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to U6. Analysis by two-tailed t-test; ***p≤0.001 vs Gm28730^{+/+}. Data represented as individual data points and mean ± SEM.
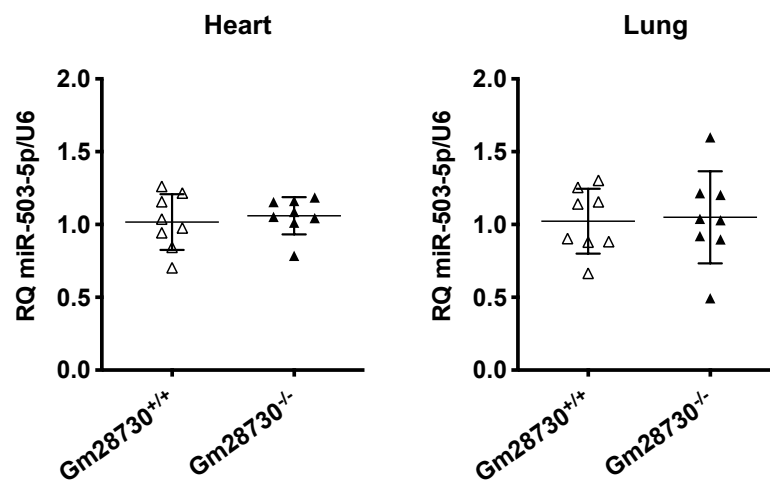
Figure 5.22: miR-503 expression in Gm29730^{-/-} mice.

Expression of miR-503-5p in heart and lung tissue from homozygous Gm28730^{-/-} mice compared to wildtype Gm28730^{+/+} littermates (n=8 mice/group). RQ value for gene expression was quantified by qRT-PCR assay relative to U6. Analysis by two-tailed t-test; ***p≤0.001 vs Gm28730^{+/+}. Data represented as individual data points and mean ± SEM.

5.4 Discussion

Building on the *in vitro* work presented in Chapters 3 and 4, here we have demonstrated that the expression of MIR503HG is significantly lost in the remodelled lung vasculature during PH in association with EndMT. Additionally, we provide preliminary evidence supporting a MIR503HG-mediated mechanism in the suppression of EndMT during vascular remodelling.

5.4.1 MIR503HG expression profile during disease

As demonstrated in this chapter, BOEC derived from PAH patients with BMPR2 mutations present with an EndMT-like profile, showing an increase in mesenchymal markers and decreased endothelial marker expression. Interestingly, mutations to the gene encoding for BMPR2, while being the most common cause of heritable PAH, have also been implicated in the initiation of EndMT *in vitro* [120,126]. Crucially, we have further shown that the appearance of an EndMT-like profile in in PAH BOEC also happens in association with a decrease in MIR503HG_2 availability. While largely associative, this data set reinforces the proposed role of the MIR503HG in regulating EC phenotype commitment described in chapter 3. Nonetheless, it's important to acknowledge that despite PAH patient-derived BOEC presenting with an EndMT-like phenotype, only MIR503HG_ 2 was found to be downregulated while all other transcripts remained unaltered. Additionally, the analysis shown here focused solely on baseline marker expression and did not include more in-depth analysis of these cells. Future studies would therefore require additional, in-depth, manipulation of the MIR503HG locus. For example, if the loss of MIR503HG_2 is involved in the emergence of an EndMT profile in these cells, overexpression of the lncRNA using

our LNT_503HG construct may work to reverse mesenchymal marker expression and maintain endothelial commitment. Furthermore, the experiments described were limited to BOEC between passages 6 and 8. While these have been shown to retain endothelial function [423,514], as with other primary endothelial cell lines , after long-term serial passaging these may become quiescent and start to exhibit changes gene expression patterns, as well as physiological function [521–524].

Moreover, despite using a disease relevant EC subset, the studies presented were performed *in vitro* and did not account for crucial external changes to the vascular environment that also regulate EC function and therefore EndMT [525]. We have attempted to mitigate this by observing MIR503HG *in situ* across the lung vasculature of PAH patients. With this, we have shown that MIR503HG expression was largely abolished in the remodelled lung vasculature of PAH patients in association with a loss of endothelial structure in these vessels. However, while the *in situ* protocol used provides convincing qualitative evidence, the nuclear lncRNA staining shown is diffuse and difficult to accurately quantify. Future studies may benefit from using novel ISH assays such as RNAscope, which allow for single RNA molecules to be presented as quantifiable punctate dots within the cell. Further, it is also important to acknowledge that, despite the changes to EC structure, no evidence of active EndMT was found in these vessels.

Lastly, by introducing an inducible EC lineage tracing mouse model (Ind.EndoTrack) we were able to specifically isolate lung cells of endothelial origin. Crucially, we have shown that mouse lung cells of endothelial origin presented with a clear EndMT profile in association with the decreased expression of the MIR503HG mouse homologue Gm28730. While, EndMT has previously been identified in PH, this is the first

instance where the Gm28730 lncRNA locus was linked to both EndMT and vascular remodelling during disease. Exposure to our 3 week SuHx model induces a severe PH phenotype accompanied by RVSP overload, RV hypertrophy and the development of angioobliterative lesions comparable to those seen in our advanced PAH patient samples [526,527]. Thus, at this time point, changes in Gm28730 expression may not be directly linked to EndMT and instead due to global changes to the vessel cellular architecture. This is particularly important, given that the loss of Gm28730 expression in our SuHx samples was not limited to ECs. Future studies would therefore need to also validate lncRNA expression changes at earlier time points to confirm if drop in expression matches with that of early EndMT compared to other cell types.

Unbiased exploratory RNAseq of isolated lung ECs could also prospectively be implemented, not only to validate changes to Gm28730 transcript expression, but to identify novel targets not detected in our initial *in vitro* analysis. For example, using high-throughput RNAseq of fluorescently labelled cells isolated from a Confetti EC lineage-tracing mouse line after acute myocardial infarction, Manavski *et al* demonstrated that ECs undergo clonal expansion and begin to express a variety of EndMT-associated mesenchymal markers [528]. Moreover, single-cell RNA sequencing (sgRNAseq) may also start to emerge as a powerful tool for study of EndMT *in vivo* due to its added ability to validate EC signatures while simultaneously showing mesenchymal gene up-regulation in individual cells. Nonetheless, a known caveat to the use of RNAseq is the need for sufficient read depth to reliably detect and annotate lowly expressed transcripts, such as that of lncRNAs.

5.4.2 Overexpression of MIR503HG *in vivo*

Recapitulating the *in vitro* results presented in chapter 3 of this thesis, overexpression

of the human MIR503HG isoform 2 in mice prior to PH induction was linked to a decrease of mesenchymal marker expression in mouse CD31$^+$ lung ECs, suggesting a partial prevention of the EndMT process *in vivo*. As noted, however, a non-significant increase in *Snai2* was observed in CD31$^+$ cells overexpressing MIR503HG. While this effect was not present during our *in vitro* studies, it is worth considering that MIR503HG may target pathways downstream or independent of *Snai2*. This reinforces the need for additional mechanist studies, as highlighted in Chapter 6.6. Further, given the preliminary nature of our study, the reduced sample size used (n=5 mice/group) was not sufficient to confirm changes to PH outcomes such as increased RVSP and RV hypertrophy. Additionally, overexpression of MIR503HG_2 was only carried out prior to PH induction, without a paired normoxic treatment group was also not included. With our limited observation data and use of the human MIR503HG, as opposed to the endogenous Gm28730 lncRNA, we could not ethically justify a large-scale study prior to proof-of-consent. Given the results presented here, however, a full-scale study can now be in implemented alongside a normoxic control group not exposed to SuHx conditions, to further confirm the changes at the molecular level and thoroughly assess the impact on the severity of PH. Moreover, due to the limited availability of Ind.EndoTrack mice, overexpression of MIR503HG_2 *in vivo* was carried out on a C57BL/6 mouse line. Thus, the cell sorting strategy used was restricted to endothelial surface markers shown to be lost during EndMT, which may have ultimately obscured the impact MIR503HG overexpression had on EndMT cell that no longer express these markers. As such, a future full-scale study should also include our Ind.EndoTrack mouse line, allowing us to capture all cells of endothelial origin. Despite the positive results presented, our overexpression strategy was not restricted

to the endothelium, and MIR503HG_2 expression was found present in both CD31$^+$ and CD31$^-$ lung cells. Moreover, while shown to be an effective delivery route to the lung endothelium [203,529,530], high-titre lentiviral administered intranasally will first be exposed to airway epithelium prior to reaching the underlying vasculature. Thus, given the cell-specific and context-dependent nature of the lncRNA, it is worth considering the negative impact that unrestrained global overexpression may have on airway function as well as the lung vasculature and disease progression. Nonetheless, MIR503HG_2 overexpression was accompanied by an increase in endogenous Gm28730 availability in CD31$^+$ ECs but not in CD31$^-$ cells, which highlights a pro-endothelial function.

Notably, it has also been previously established that reconstitution of miR-424(322) and miR-503 in the rat lung can reduce the number of number of muscularised microvessels and with it prevent increased RVSP and RV hypertrophy associated with the SuHx model [203]. While both miRNAs were shown to be overexpressed in CD31$^+$ SuHx lung ECs, the overexpression strategy used in the study, like ours, relied on the use of a lentivirus constructs that did not specifically target the endothelium. Thus, the cell-specific effects of miRNA overexpression most also be considered. For example, overexpression of miR-503 was found suppress proliferation and migration of EPCs *in vitro* [479], the opposite was reported in ECs and SMCs [203]. Nonetheless, these results are in conflict with our *in vitro* assessment and warrant additional dissection of the locus *in vivo*. Given the technical hurdles and limitations associated with selectively targeting the Gm28730 locus in the pulmonary endothelium, genetic manipulation approaches may be required as further proof-of-concept.

5.4.3 Gm28730 knockout mouse model

Has highlighted throughout this thesis and the prevailing literature it is likely that a cross-regulatory transcriptional and functional axis may exist between the miR-424(322)/503 cluster and its lncRNA host. While our *in vitro* lncRNA overexpression strategy suggested a function independent of the miRNA cluster, loss of MIR503HG (Gm28730) has been consistently accompanied by a significant downregulation of both miR-424(322) and miR-503. As such, additional loss-of-function experiments would require precise manipulation of the MIR503HG (Gm28730) locus. We have attempted to do this by using a novel CRIPR/Cas9 system to introduce a premature transcriptional termination site (SPA-MAZ) within exon 1 of the lncRNA locus. The presence of the SPA-MAZ cassette effectively stopped transcription of the lncRNA without disrupting miRNA processing upstream. Here, we have confirmed the effectiveness of this strategy by showing a significant reduction in Gm28730 expression in the mouse heart and lung, while retaining miRNA expression throughout. Thus, to our knowledge, generating the first true murine Gm28730 knockout. However, while our genotyping efforts have confirmed the absence of Gm28730, due to the pleiotropic effects that the KO may have on the function of different organ systems and during development, further in-depth phenotypical assessment is still necessary [531–533].

Once fully phenotyped, Gm28730$^{-/-}$ mutants can be used to assess the role of the lncRNA during vascular remodelling in disease. For example, implementing the SuHx-induced model of PH described in this chapter would be crucial in order to study the effect of Gm28730 KO on PH hallmarks such as increased RVSP and RV hypertrophy [526,527]. Additionally, given the reported role of EndMT in atherosclerosis,

a proprotein convertase subtilisin/kexin type 9 (PCSK9)-induced model of accelerated atherosclerosis in mice could also be implemented for future studies [534,535].

Lastly, we acknowledge that despite the success in generating Gm28730[-/-] mutants, there are still a number of limitations to this KO model that may confound future studies. Namely, the lack of endothelial specificity of the Gm28730[-/-] mutants. As described in the previous chapters, this is particularly important given the cell and context specific nature of lncRNA. Further, given the potential for developmental pathologies to arise, the non-inducible nature of our KO strategy may exacerbate the pathological response of Gm28730[-/-] mutants to disease models. However, given the additional SPA-MAZ cassette insertion step need, an inducible Gm28730[-/-] would require the use of multiple viral vectors in adult mice which may result only in a partial KD. As demonstrated here, no discernible changes in Gm28730 expression were present in heterozygous Gm28730[+/-], highlighting the need for biallelic insertion. Nonetheless, drug-inducible CRISPR/Cas9 systems have recently been described and may be possible to use in the near future [536].

5.4.4   Conclusions

Our analysis of both human PAH and *in vivo* animal models, has provided important insights into the role of MIR503HG in regulating EndMT during disease. Taken together with the fact that EndMT can be inhibited by the restoration of MIR503HG expression *in vitro*, the results presented in this chapter suggest a critical role played by MIR503HG in maintaining endothelial function during active vascular remodelling in PH. Our results open new avenues for targeting EndMT during disease, showing a previously unknown regulatory function of MIR503HG in driving the process.

Chapter 6: General Discussion

6.1 Summary

As evidence continues to emerge highlighting the role of EndMT in pathological vascular remodelling seen during the development of conditions such as PAH, understanding the underlying mechanisms that drive this phenotypical transition may ultimately offer novel therapeutic opportunities. In this thesis we demonstrate for the first time that the loss of the lncRNA MIR503HG is a pivotal event during the induction of EndMT both *in vitro* and *in vivo* (Figure 6.1). We show that targeted downregulation of this lncRNA can induce a spontaneous EndMT phenotype in HUVEC in the absence of any other treatment, including TGF-β2 and IL-1β co-stimulation, while upregulation of a single conserved transcript is sufficient to inhibit this process. In a mouse model of PAH, MIR503HG expression in lung ECs was also found to be lost in association with an EndMT profile. When overexpressed in the mouse lung, the data presented here shows that the presence of MIR503HG can suppress the appearance of mesenchymal markers despite induction of PAH. Similarly, the loss of MIR503HG was associated with human PAH, demonstrating a putative clinical role. Ultimately, the data presented here highlights the potential for lncRNA-based therapeutic interventions, positioning MIR503HG as a novel putative target during the treatment of vascular remodelling in the context of PAH.

Nonetheless, despite the novelty of the evidence presented, several limiting factors have been identified which must be considered in future studies. Similarly, lingering concerns still exist regarding the efficacy of lncRNA-based therapeutic strategies. We will discuss these topics in further detail through this chapter.
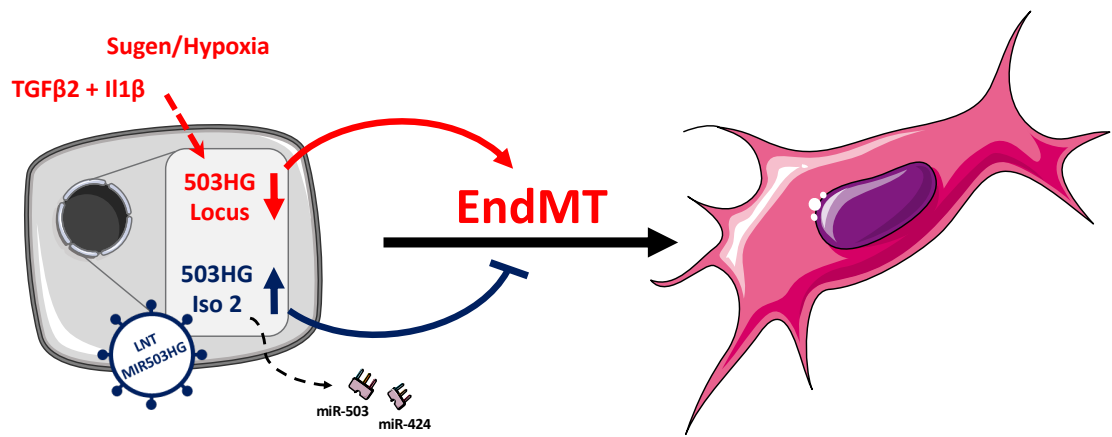
Figure 6.1: Diagram Summarising the Role of MIR503HG in Regulating EndMT.
Graphical summary of the role played by MIR503HG in regulating EC function, where loss of the lncRNA is sufficient to induce a spontaneous EndMT phenotype, in the absence of TGF-β2 and IL-1β co-stimulation. Conversely overexpression of MIR503HG during transition can supress transition, in a manner independent of miR-424 and miR-503.

6.2   MIR503HG as a Master Regulator of EndMT

The work presented in Chapter 3 of this thesis showed a large overlap between the transcriptional changes of HUVEC and HPAEC undergoing EndMT after TGF-β2 and IL-1β co-treatment, supporting the existence of shared regulatory mechanisms. Among these, we identify for the first time the loss of MIR503HG as a key event during the induction of EndMT in EC from various vascular beds.

Remarkably, DsiRNA or GapmeR-mediated knockdown of MIR503HG was sufficient to induce a robust EndMT profile *in vitro*, without the need for TGF-β2 and IL-1β co-treatment. While other lncRNAs have been previously associated with EndMT, to our knowledge, this is the first report of spontaneous EndMT after manipulation of a single lncRNA. Upregulation of the lncRNA MALAT1, for example, was previously linked to EndMT in endothelial progenitor cells (EPCs); however, this was shown to happen via a MALAT1-mediated suppression of miR-145, which in turn increased the availability of TGF-βR2 and SMAD3 [397]. Recently, GATA6-AS, a lncRNA antisense

to GATA6, was also implicated in the regulation of endothelial gene expression no direct link was made between its expression and the induction of EndMT [398].

We additionally show that the upregulation of a single MIR503HG isoform, containing a highly conserved exonic region, is sufficient to inhibit transition in HUVEC after continuous exposure to TGF-β2 and IL-1β. Nonetheless, as described in Section 3.3.6, it is important to highlight that this was still a partial inhibition of transition. Conceptually, this could be mitigated by overexpressing the totality of the MIR503HG locus using a CRISPR activation (CRISPRa) system targeting the lncRNA promotor region. This system uses a catalytically-inactive Cas9 nuclease (dCas9) with additional transcription activation domains such as VP64, which can then use a single guide RNA (sgRNA) to bind complementary DNA. Along with the desired target sequence, the sgRNA structure can be modified to recruit and bind additional transcriptional activators for an enhanced effect [537–539]. Further experiments may also include additional transcript-specific overexpression of the remaining MIR503HG isoforms, again using the same lentiviral-mediated transduction system described here. This, however, may be limited by the lack of reliable MIR503HG transcript annotations.

Lastly, several other lncRNAs candidates with a potential role in EndMT have emerged from the RNAseq analysis described in Chapter 3. While beyond the scope of this thesis, their presence highlights the existence of an unexplored regulatory network of lncRNAs which warrants further research.

6.3   Identifying the Mechanistic Interactions of MIR503HG

In order to identify the contribution of MIR503HG to endothelial function and which associated pathways may be regulated during EndMT, in Chapter 3 we performed bulk RNAseq analysis of ECs overexpressing MIR503HG_2 during EndMT treatment. This approach showed that the expression of over 28% of EndMT-associated genes was affected by the presence of MIR503HG, however these changes were too far-reaching to select specific target genes. Thus, further mechanistic studies must be employed to identify the molecular targets and pathways associated with MIR503HG in ECs.

With a wide range of described mechanisms, lncRNAs carry out their function by interacting with a variety of other nucleic acids and proteins. While large-scale databases such as StarBase v2.0 now offer a repository of thousands of experimentally validated RNA:RNA and protein:RNA interactions [540], many of these regulatory networks remain largely unknown. The application of fitting, RNA-centric, tools to elucidate such complex interactions is therefore essential. The role of lncRNA:Protein interactions in regulating gene expression has been of particular interest and several pulldown assays have been developed over the past decade to identify these interactions. RNA-affinity chromatography assays, for example, are now widely used in mechanist studies. These require the *in vitro* transcription of a labelled lncRNA transcript, which is allowed to form complexes with protein obtained from cell extracts and later purified [541]. Alternatively, methods such as RNA antisense purification (RAP)[458,459] will instead capture endogenous lncRNA transcripts using labeled antisense oligos. While their methodology can vary significantly, both techniques

ultimately allow for the purification specific lncRNA:protein complexes which are identified by quantitative mass spectrometry or western blot.

Interestingly, using *in vitro*-transcribed biotinylated MIR503HG transcripts, Wang *et al* have identified a direct interaction with the RNA-binding protein HNRNPA2B1 [437]. This was found to promote the degradation of HNRNPA2B1, resulting in the suppression of NF-kB signalling in hepatocellular carcinoma. Given that HNRNPA2B1 has been shown to promote cell invasion and migration in EMT [461], this interaction may also prove relevant during EndMT. Nonetheless, it's worth considering that a MIR503HG-mediated degradation of HNRNPA2B1 may be cell-specific and thus not present in ECs. We propose that future mechanistic studies first confirm MIR503HG binding in HUVEC or HPAEC via RNA immunoprecipitation (RIP) using HNRNPA2B1 antibodies. Additionally, this should be further validated in an independent and unbiased manner by exposing biotinylated MIR503HG_2 transcripts to EC nuclear protein lysate.

Although proven successful, it has been recently suggested that this approach may be limited to identifying highly abundant RNA binding proteins and does not provide a complete picture of other, perhaps more relevant, lncRNA:protein interactions occurring in parallel [542]. As such, the capture of endogenous MIR503HG transcripts using biotin-labelled RNA antisense probes may add an additional layer of stringency. This system was recently used to identify GATA6-AS:protein complexes in HUVEC undergoing EndMT [398]. With this, the study successfully demonstrated that the lncRNA was capable of binding the nuclear enzyme LOXL2 to partially regulate endothelial gene expression via changes in histone methylation.

Lastly, MIR503HG expression was previously found to be positively altered in ECs

by hypoxia, with its depletion resulting in proliferative defects. Crucially, depletion of MIR503HG also lead to the repression of the endothelial transcription factor GATA-binding factor 2 (GATA2) [381]. Abrogation of GATA2 alone has previously been shown to regulate Endomucin-1 (EMCN) gene expression, leading to a reduction in the expression of endothelial markers and promoting EndMT gene expression [543]. While beyond the scope of this thesis, an EndMT regulatory axis involving MIR503HG and GATA2 is plausible and warrants further research.

6.4   Dissecting the MIR503HG Locus

In Chapter 4 we have attempted to dissect the MIR503HG locus critically demonstrating that the contribution of miR-424 and miR-503 to EndMT was largely associated with the loss of MIR503HG. Although a significant decrease in miR-503 and miR-424 availability was observed during EndMT, this is may be due to the transcriptional repression of shared regulatory regions [462]. Indeed, targeted depletion of either miRNA did not change the expression of EndMT markers, unlike that seen after MIR503HG knockdown. Crucially, overexpression of a mature MIR503HG transcript, lacking the miRNA seed sequences, effectively repressed EndMT without changing miR-424/-503 expression levels.

Whilst several putative miRNA host lncRNAs have been described, publications determining the miRNA-independent roles of these transcripts are rare. Notable exceptions include the recently characterised lncRNA MIR205HG, which despite acting as the *de facto* pri-miRNA transcript for miR-205, has also been shown to independently regulate differentiation of human prostate basal cells [544]. Similarly, the lncRNA MIR100HG, host to miR-100, let7a2 and miR-125b, regulates cell cycle

through its interaction with the RNA binding protein HuR [545]. Adding to this poorly described lncRNA subset, our data provides a novel example of a miRNA host gene with an independent functional role.

Nonetheless, the potential for cross-regulatory functional mechanisms between MIR503HG and miR-424/-503 still exists. As demonstrated in Chapter 4, overexpression of either miRNA triggered the appearance of an EndMT profile, accompanied by a decrease in MIR503HG expression. Moreover, siRNA-mediated knockdown of MIR503HG also lead to a significant decrease in miRNA expression.

It is clear that further dissection of the locus is necessary to unravel any potential overlapping mechanisms present. In order to do this, future studies should take advantage of established CRISPR/Cas9 gene editing systems to manipulate the locus to recapitulate the MIR503HG knockdown phenotype without affecting miR424/503 expression. Deletion of the conserved final 595 base-pair exonic region of MIR503HG_2, for example, would prospectively generate truncated lncRNA transcripts containing only the unsliced pri-miRNAs. The implementation of this system in primary EC lines, however, comes with an additional set of challenges due to their inherent sensitivity and finite number of division cycles compared to immortalised counterparts. Nonetheless, while rare, emerging publications have demonstrated that this is possible in ECs [546,547]. Recently showcased by Leisegang *et al*, the use of lentiviral vectors to introduce Cas9 nucleases and constitutively express sgRNA strands in HUVEC was sufficient to produce distinct bi-allelic deletions of the lncRNA MANTIS, phenocopying results obtained using siRNA and GapmeR-based approaches [547].

6.5 MIR503HG Expression During Disease

The involvement of several lncRNAs has been implicated during the initiation and progression of multiple cardiovascular pathologies, however little is known about their role in PAH. In Chapter 5 we have for the first time established a link between the loss of MIR503HG expression and lung vascular remodelling in association with EndMT, using a mouse model of PH, human BOECs and lung samples obtained from PAH patients. More specifically, using *in situ* hybridisation for MIR503HG, supported by sequential staining for endothelial and mesenchymal markers, MIR503HG expression was found to be largely abolished in the remodelled lung vasculature of PAH patients. Moreover, our use of a SuHx-induced PAH model established in Ind.EndoTrack lineage tracing transgenic mouse further allowed us specifically isolate lung cells of endothelial origin regardless of any changes to cell phenotype. Consistent with our patient data, we have confirmed that the expression of the MIR503HG mouse homolog, Gm28730, is lost during PAH in association with an EndMT profile.

Nevertheless, while the data provided focused solely on vascular remodelling in PAH, EndMT has been implicated in a variety of cardiovascular pathologies. For example, as demonstrated by Evrard and colleagues, cells co-expressing endothelial and mesenchymal markers were commonly found in complex and unstable human atherosclerotic plaque [53]. Similar observations have also been reported in vein graft remodelling, valvular calcification and cardiac fibrosis [12,14,548–550]. It would therefore be important to assess if MIR503HG down-regulation is also relevant in these scenarios and if any changes to its expression are also associated with EndMT. A similar *in situ* hybridisation strategy for MIR503HG visualisation could prospectively be applied to human samples with and without atherosclerotic lesions. Given its

versatility, our Ind.EndoTrack mouse line could prospectively be exposed to other vascular remodelling-associated disease models, such as atherosclerosis, and used to validate not only the presence of EndMT but also changes in Gm28730 expression in these cells. This could be achieved by establishing the Ind.EndoTrack transgenic mouse lineage in an atherosclerosis prone ApoE$^{-/-}$ background, as recently demonstrated by Chen *et al* [116].

6.6 Manipulating MIR503HG *In Vivo* and Clinical Translation

Although persuasive, the data presented throughout Chapter 5 was largely associative and as such must be complemented with targeted manipulation of the lncRNA. We have attempted to mitigate these limitations by overexpressing the human MIR503HG_2 transcript in the mouse lung prior to induction of PAH. In line with the *in vitro* data presented, our overexpression strategy led to a significant decrease in mesenchymal marker expression in PAH lung mouse ECs, suggesting that MIR503HG can prevent the EndMT process during disease. However, several factors must be thoroughly tested before definite conclusions can be drawn.

Larger comprehensive studies, including normoxic control mice along with optimised delivery strategies, are required in order to fully elucidate the role MIR503HG in PAH. Crucially, previous publications have made conflicting reports regarding the MIR503HG effect on cell proliferation and invasion [441,442]. Given the unrestrained overexpression of MIR503HG_2 induced by our delivery strategy, the use of endothelial-specific gene transfer viral vectors should be considered during future applications [551–553]. Moreover, while the impact on mesenchymal gene expression is

meaningful, in the context of PAH future studies must also consider a full haemodynamic and histological evaluation of disease hallmarks such as increased right ventricular systolic pressure (RVSP) measurements and the formation of plexiform lesions [165,190].

We go on to describe the development of, to our knowledge, the first true murine MIR503HG (Gm28730) knockout line (Gm28730[-/-]). Here, using a novel CRIPR/Cas9 system to introduce a premature transcriptional termination site (SPA-MAZ) within exon 1 of the lncRNA locus, we have successfully stopped Gm28730 expression without disrupting that of miR-424(322) and miR-503. This unique transgenic mouse line will be crucial in discerning the role of MIR503HG (Gm28730) in EC function and EndMT. However, as suggested in Section 5.4.3, while the loss of Gm28730 expression has been confirmed in these mice, in-depth phenotypical assessment is still necessary. It is known that cardiac ECs in the atrio-ventricular (AV) canal give rise to the mesenchymal cells through active EndMT during cardiac septa and valve formation [92,95,554]. As such, if the loss of MIR503HG (Gm28730) truly drives endothelial transition, Gm28730[-/-] mice may present with cardiac abnormalities. Moreover, there are still a number of limitations to this model possibly confounding future studies, these have been discussed in-depth through Section 5.4.3.

Lastly, we must also consider the broader implications of EndMT during disease. As proposed throughout this thesis, the transition process has established itself as an important factor in the progression of vascular disease [52]. However, as is common with complex vascular diseases, including PAH, a number of different pathological and cellular processes will actively mediate the disease progression. It is clear that EndMT cannot account for the totality of structural alterations present during vascular

remodelling. In PAH, a recent study has proposed more conservative figures than those previously estimated, showing that while EndMT is present in remodelled vessels it only accounts for approximately 14% of the cell population [172]. Nevertheless, this is still a significant population that, if managed, could improve patient outcomes. As highlighted in Section 1.3, combined therapy is now regarded as the standard of care in PAH, significantly improving patient outcomes [219,555–557]. Thus, a novel therapeutic target, combined with traditional therapeutic strategies, may ultimately extend PAH patient survival and improve quality of life. This thesis offers a promising initial outline for the role of MIR503HG during EndMT in PAH, however continuous research will be necessary for it to be considered for clinical therapy.

6.7   Conclusions

Taken together, the analysis described in this thesis provides important insights into EndMT highlighting the critical role of the lncRNA MIR503HG in maintaining EC function. While our results open new avenues for targeting EndMT, showing a previously unknown regulatory function of lncRNAs in driving this process, further research is still necessary to fully validate the mechanistic interactions of MIR503HG during EndMT and its role during disease. Nevertheless, as evidence underlining the finer regulatory mechanisms controlling EndMT continues to emerge, new lncRNA-directed therapeutic opportunities may also appear for vascular remodelling.

References

1.	de Vries, M. R., Simons, K. H., Jukema, J. W., Braun, J. & Quax, P. H. A. Vein graft failure: from pathophysiology to clinical outcomes. *Nat. Rev. Cardiol.* **13**, 451–470 (2016).

2.	Hoglund, V. J., Dong, X. R. & Majesky, M. W. Neointima Formation: A Local Affair. *Arteriosclerosis, thrombosis, and vascular biology* vol. 30 1877–1879 (2010).

3.	Rabinovitch, M. Molecular pathogenesis of pulmonary arterial hypertension. *J. Clin. Invest.* **122**, 4306–4313 (2012).

4.	Shimoda, L. A. & Laurie, S. S. Vascular remodeling in pulmonary hypertension. *J. Mol. Med. (Berl).* **91**, 297–309 (2013).

5.	Stenmark, K. R., Fagan, K. A. & Frid, M. G. Hypoxia-induced pulmonary vascular remodeling: cellular and molecular mechanisms. *Circ. Res.* **99**, 675–691 (2006).

6.	Lu, X. & Kakkar, V. The roles of microRNAs in atherosclerosis. *Curr. Med. Chem.* **21**, 1531–1543 (2014).

7.	Muto, A., Model, L., Ziegler, K., Eghbalieh, S. D. D. & Dardik, A. Mechanisms of vein graft adaptation to the arterial circulation: insights into the neointimal algorithm and management strategies. *Circ. J.* **74**, 1501–1512 (2010).

8.	Wan, S., George, S. J., Berry, C. & Baker, A. H. Vein graft failure: current clinical practice and potential for gene therapeutics. *Gene Ther.* **19**, 630–636 (2012).

9.	Marx, S. O., Totary-Jain, H. & Marks, A. R. Vascular smooth muscle cell proliferation in restenosis. *Circ. Cardiovasc. Interv.* **4**, 104–111 (2011).

10.	Arciniegas, E., Frid, M. G., Douglas, I. S. & Stenmark, K. R. Perspectives on endothelial-to-mesenchymal transition: potential contribution to vascular remodeling in chronic pulmonary hypertension. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **293**, L1-8 (2007).

11.	Nguyen, A. T. *et al.* Smooth muscle cell plasticity: fact or fiction? *Circ. Res.* **112**, 17–22 (2013).

12.	Zeisberg, E. M. *et al.* Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat. Med.* **13**, 952–961 (2007).

13.	Hashimoto, N. *et al.* Endothelial-mesenchymal transition in bleomycin-induced pulmonary fibrosis. *Am. J. Respir. Cell Mol. Biol.* **43**, 161–172 (2010).

14.	Cooley, B. C. *et al.* TGF-beta signaling mediates endothelial-to-mesenchymal transition (EndMT) during vein graft remodeling. *Sci. Transl. Med.* **6**, 227ra34 (2014).

15.	van Meeteren, L. A. & ten Dijke, P. Regulation of endothelial cell plasticity by TGF-beta. *Cell Tissue Res.* **347**, 177–186 (2012).

16.	Potente, M. & Makinen, T. Vascular heterogeneity and specialization in development and disease. *Nat. Rev. Mol. Cell Biol.* **18**, 477–494 (2017).

17. Carmeliet, P. & Jain, R. K. Angiogenesis in cancer and other diseases. *Nature* **407**, 249–257 (2000).

18. Pate, M., Damarla, V., Chi, D. S., Negi, S. & Krishnaswamy, G. ENDOTHELIAL CELL BIOLOGY: ROLE IN THE INFLAMMATORY RESPONSE. *Adv. Clin. Chem. Vol 52* **52**, 109–130 (2010).

19. Eelen, G., de Zeeuw, P., Simons, M. & Carmeliet, P. Endothelial Cell Metabolism in Normal and Diseased Vasculature. *Circ. Res.* **116**, 1231–1244 (2015).

20. Dudley, A. C. Tumor Endothelial Cells. *Cold Spring Harb. Perspect. Med.* **2**, (2012).

21. Brandes, R. P. Endothelial dysfunction and hypertension. *Hypertens. (Dallas, Tex. 1979)* **64**, 924–928 (2014).

22. Gimbrone, M. A. & Garcia-Cardena, G. Endothelial Cell Dysfunction and the Pathobiology of Atherosclerosis. *Circ. Res.* **118**, 620–636 (2016).

23. Dejana, E., Hirschi, K. K. & Simons, M. The molecular basis of endothelial cell plasticity. *Nat. Commun.* **8**, 14361 (2017).

24. Potenta, S., Zeisberg, E. & Kalluri, R. The role of endothelial-to-mesenchymal transition in cancer progression. *Br. J. Cancer* **99**, 1375–1379 (2008).

25. Wang, Z. *et al.* High-throughput investigation of endothelial-to-mesenchymal transformation (EndMT) with combinatorial cellular microarrays. *Biotechnol. Bioeng.* (2015) doi:10.1002/bit.25905.

26. Xu, X. *et al.* Snail Is a Direct Target of Hypoxia-inducible Factor 1alpha (HIF1alpha) in Hypoxia-induced Endothelial to Mesenchymal Transition of Human Coronary Endothelial Cells. *J. Biol. Chem.* **290**, 16653–16664 (2015).

27. Aisagbonhi, O. *et al.* Experimental myocardial infarction triggers canonical Wnt signaling and endothelial-to-mesenchymal transition. *Dis. Model. Mech.* **4**, 469–483 (2011).

28. Mahler, G. J., Farrar, E. J. & Butcher, J. T. Inflammatory cytokines promote mesenchymal transformation in embryonic and adult valve endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* **33**, 121–130 (2013).

29. Chen, P.-Y. *et al.* Endothelial-to-mesenchymal transition drives atherosclerosis progression. *J. Clin. Invest.* **125**, 4514–4528 (2015).

30. Noseda, M. *et al.* Notch activation results in phenotypic and functional changes consistent with endothelial-to-mesenchymal transformation. *Circ. Res.* **94**, 910–917 (2004).

31. Cano, A. *et al.* The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76–83 (2000).

32. Bolos, V. *et al.* The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J. Cell Sci.* **116**, 499–511 (2003).

33. Lopez, D., Niu, G., Huber, P. & Carter, W. B. Tumor-induced upregulation of Twist, Snail, and Slug represses the activity of the human VE-cadherin promoter. *Arch. Biochem. Biophys.* **482**, 77–82 (2009).

34. Comijn, J. *et al.* The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol. Cell* **7**, 1267–1278 (2001).

35. Eger, A. *et al.* DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene* **24**, 2375–2385 (2005).

36. Korpal, M., Lee, E. S., Hu, G. & Kang, Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J. Biol. Chem.* **283**, 14910–14914 (2008).

37. Medici, D. Conversion of vascular endothelial cells into multipotent stem-like cells. vol. 16 1400–1406 (2010).

38. Medici, D. & Kalluri, R. Endothelial-mesenchymal transition and its contribution to the emergence of stem cell phenotype. *Semin. Cancer Biol.* **22**, 379–384 (2012).

39. Teng, Y. & Li, X. The roles of HLH transcription factors in epithelial mesenchymal transition and multiple molecular mechanisms. *Clin. Exp. Metastasis* **31**, 367–377 (2014).

40. Yang, J. *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927–939 (2004).

41. Barnes, R. M. & Firulli, A. B. A twist of insight - the role of Twist-family bHLH factors in development. *Int. J. Dev. Biol.* **53**, 909–924 (2009).

42. Mahmoud, M. M. *et al.* TWIST1 Integrates Endothelial Responses to Flow in Vascular Dysfunction and Atherosclerosis. *Circ. Res.* **119**, 450–462 (2016).

43. Chakraborty, S. *et al.* Twist1 promotes heart valve cell proliferation and extracellular matrix gene expression during development in vivo and is expressed in human diseased aortic valves. *Dev. Biol.* **347**, 167–179 (2010).

44. Mahmoud, M. M. *et al.* Shear stress induces endothelial-to-mesenchymal transition via the transcription factor Snail. *Sci. Rep.* **7**, 3375 (2017).

45. Qi, Q. *et al.* Geniposide inhibited endothelial-mesenchymal transition via the mTOR signaling pathway in a bleomycin-induced scleroderma mouse model. *Am. J. Transl. Res.* **9**, 1025–1036 (2017).

46. Nishimura, G. *et al.* DeltaEF1 mediates TGF-beta signaling in vascular smooth muscle cell differentiation. *Dev. Cell* **11**, 93–104 (2006).

47. Niessen, K. *et al.* Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *J. Cell Biol.* **182**, 315–325 (2008).

48. Medici, D., Potenta, S. & Kalluri, R. Transforming growth factor-beta2 promotes Snail-mediated endothelial-mesenchymal transition through

convergence of Smad-dependent and Smad-independent signalling. *Biochem. J.* **437**, 515–520 (2011).

49. August, P. & Suthanthiran, M. Transforming growth factor beta signaling, vascular remodeling, and hypertension. *N. Engl. J. Med.* **354**, 2721–2723 (2006).

50. Meng, X.-M., Nikolic-Paterson, D. J. & Lan, H. Y. TGF-beta: the master regulator of fibrosis. *Nat. Rev. Nephrol.* **12**, 325–338 (2016).

51. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation* vol. 119 1420–1428 (2009).

52. Kovacic, J. C. *et al.* Endothelial to Mesenchymal Transition in Cardiovascular Disease: JACC State-of-the-Art Review. *J. Am. Coll. Cardiol.* **73**, 190–209 (2019).

53. Evrard, S. M. *et al.* Endothelial to mesenchymal transition is common in atherosclerotic lesions and is associated with plaque instability. *Nat. Commun.* **7**, 11853 (2016).

54. Hata, A. & Chen, Y.-G. TGF-beta Signaling from Receptors to Smads. *Cold Spring Harb. Perspect. Biol.* **8**, (2016).

55. Pardali, E., Goumans, M.-J. & ten Dijke, P. Signaling by members of the TGF-beta family in vascular morphogenesis and disease. *Trends Cell Biol.* **20**, 556–567 (2010).

56. Goumans, M.-J. *et al.* Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors. *EMBO J.* **21**, 1743–1753 (2002).

57. Goumans, M. J. *et al.* Activin receptor-like kinase (ALK)1 is an antagonistic mediator of lateral TGFbeta/ALK5 signaling. *Mol. Cell* **12**, 817–828 (2003).

58. ten Dijke, P., Goumans, M.-J. & Pardali, E. Endoglin in angiogenesis and vascular diseases. *Angiogenesis* **11**, 79–89 (2008).

59. Oh, S. P. *et al.* Activin receptor-like kinase 1 modulates transforming growth factor-beta 1 signaling in the regulation of angiogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 2626–2631 (2000).

60. Gonzalez-Nunez, M., Munoz-Felix, J. M. & Lopez-Novoa, J. M. The ALK-1/Smad1 pathway in cardiovascular physiopathology. A new target for therapy? *Biochim. Biophys. Acta* **1832**, 1492–1510 (2013).

61. Souilhol, C., Harmsen, M. C., Evans, P. C. & Krenning, G. Endothelial-mesenchymal transition in atherosclerosis. *Cardiovasc. Res.* **114**, 565–577 (2018).

62. Gui, T., Sun, Y., Shimokado, A. & Muragaki, Y. The Roles of Mitogen-Activated Protein Kinase Pathways in TGF-beta-Induced Epithelial-Mesenchymal Transition. *J. Signal Transduct.* **2012**, 289243 (2012).

63. Li, Z. & Jimenez, S. A. Protein kinase Cdelta and c-Abl kinase are required for transforming growth factor beta induction of endothelial-mesenchymal transition in vitro. *Arthritis Rheum.* **63**, 2473–2483 (2011).

64. Zhou, B. P. *et al.* Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nat. Cell Biol.* **6**, 931–940 (2004).

65. Kanlaya, R. *et al.* Epigallocatechin-3-gallate prevents TGF-beta1-induced epithelial-mesenchymal transition and fibrotic changes of renal cells via GSK-3beta/beta-catenin/Snail1 and Nrf2 pathways. *J. Nutr. Biochem.* **76**, 108266 (2019).

66. Sanchez-Duffhues, G., Garcia de Vinuesa, A. & Ten Dijke, P. Endothelial-to-mesenchymal transition in cardiovascular diseases: Developmental signaling pathways gone awry. *Dev. Dyn.* **247**, 492–508 (2018).

67. Pober, J. S. & Sessa, W. C. Evolving functions of endothelial cells in inflammation. *Nat. Rev. Immunol.* **7**, 803–815 (2007).

68. Castellon, X. & Bogdanova, V. Chronic Inflammatory Diseases and Endothelial Dysfunction. *Aging Dis.* **7**, 81–89 (2016).

69. Perez, L. *et al.* Endothelial-to-mesenchymal transition: Cytokine-mediated pathways that determine endothelial fibrosis under inflammatory conditions. *Cytokine Growth Factor Rev.* **33**, 41–54 (2017).

70. Farrar, E. J. & Butcher, J. T. Heterogeneous susceptibility of valve endothelial cells to mesenchymal transformation in response to TNFalpha. *Ann. Biomed. Eng.* **42**, 149–161 (2014).

71. Wu, Y. & Zhou, B. P. TNF-alpha/NF-kappaB/Snail pathway in cancer cell migration and invasion. *Br. J. Cancer* **102**, 639–644 (2010).

72. Aggarwal, B. B. Signalling pathways of the TNF superfamily: a double-edged sword. *Nat. Rev. Immunol.* **3**, 745–756 (2003).

73. Hu, Z., Liu, X., Tang, Z., Zhou, Y. & Qiao, L. Possible regulatory role of Snail in NF-kappaB-mediated changes in E-cadherin in gastric cancer. *Oncol. Rep.* **29**, 993–1000 (2013).

74. Sanchez-Duffhues, G. *et al.* Inflammation induces endothelial-to-mesenchymal transition and promotes vascular calcification through downregulation of BMPR2. *J. Pathol.* **247**, 333–346 (2019).

75. Rieder, F. *et al.* Inflammation-induced endothelial-to-mesenchymal transition: a novel mechanism of intestinal fibrosis. *Am. J. Pathol.* **179**, 2660–2673 (2011).

76. Romero, L. I., Zhang, D. N., Herron, G. S. & Karasek, M. A. Interleukin-1 induces major phenotypic changes in human skin microvascular endothelial cells. *J. Cell. Physiol.* **173**, 84–92 (1997).

77. Chaudhuri, V., Zhou, L. & Karasek, M. Inflammatory cytokines induce the transformation of human dermal microvascular endothelial cells into myofibroblasts: a potential role in skin fibrogenesis. *J. Cutan. Pathol.* **34**, 146–153 (2007).

78. Masola, V. *et al.* In vitro effects of interleukin (IL)-1 beta inhibition on the epithelial-to-mesenchymal transition (EMT) of renal tubular and hepatic stellate

cells. *J. Transl. Med.* **17**, 12 (2019).

79.   Li, Y., Wang, L., Pappan, L., Galliher-Beckley, A. & Shi, J. IL-1beta promotes stemness and invasiveness of colon cancer cells through Zeb1 activation. *Mol. Cancer* **11**, 87 (2012).

80.   Lee, J. G., Ko, M. K. & Kay, E. P. Endothelial mesenchymal transformation mediated by IL-1beta-induced FGF-2 in corneal endothelial cells. *Exp. Eye Res.* **95**, 35–39 (2012).

81.   Maleszewska, M. *et al.* IL-1beta and TGFbeta2 synergistically induce endothelial to mesenchymal transition in an NFkappaB-dependent manner. *Immunobiology* **218**, 443–454 (2013).

82.   MacGrogan, D., Nus, M. & de la Pompa, J. L. Notch signaling in cardiac development and disease. *Curr. Top. Dev. Biol.* **92**, 333–365 (2010).

83.   Luxan, G., D'Amato, G., MacGrogan, D. & de la Pompa, J. L. Endocardial Notch Signaling in Cardiac Development and Disease. *Circ. Res.* **118**, e1–e18 (2016).

84.   Leslie, J. D. *et al.* Endothelial signalling by the Notch ligand Delta-like 4 restricts angiogenesis. *Development* **134**, 839–844 (2007).

85.   Sainson, R. C. A. *et al.* Cell-autonomous notch signaling regulates endothelial cell branching and proliferation during vascular tubulogenesis. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **19**, 1027–1029 (2005).

86.   Sahara, M. *et al.* Manipulation of a VEGF-Notch signaling circuit drives formation of functional vascular endothelial progenitors from human pluripotent stem cells. *Cell Res.* **24**, 820–841 (2014).

87.   Reichman, D. *et al.* Notch hyper-activation drives trans-differentiation of hESC-derived endothelium. *Stem Cell Res.* **17**, 391–400 (2016).

88.   Dabral, S. *et al.* Notch1 signalling regulates endothelial proliferation and apoptosis in pulmonary arterial hypertension. *Eur. Respir. J.* **48**, 1137–1149 (2016).

89.   Fu, Y. *et al.* RUNX3 maintains the mesenchymal phenotype after termination of the Notch signal. *J. Biol. Chem.* **286**, 11803–11813 (2011).

90.   Edeling, M., Ragi, G., Huang, S., Pavenstadt, H. & Susztak, K. Developmental signalling pathways in renal fibrosis: the roles of Notch, Wnt and Hedgehog. *Nat. Rev. Nephrol.* **12**, 426–439 (2016).

91.   MacDonald, B. T., Tamai, K. & He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev. Cell* **17**, 9–26 (2009).

92.   Liebner, S. *et al.* Beta-catenin is required for endothelial-mesenchymal transformation during heart cushion development in the mouse. *J. Cell Biol.* **166**, 359–367 (2004).

93.   Chen, Q. *et al.* Endothelial cells are progenitors of cardiac pericytes and vascular smooth muscle cells. *Nat. Commun.* **7**, 12422 (2016).

94. Markwald, R. R., Fitzharris, T. P. & Smith, W. N. Sturctural analysis of endocardial cytodifferentiation. *Dev. Biol.* **42**, 160–180 (1975).

95. Markwald, R. R., Fitzharris, T. P. & Manasek, F. J. Structural development of endocardial cushions. *Am. J. Anat.* **148**, 85–119 (1977).

96. Eisenberg, L. M. & Markwald, R. R. Molecular regulation of atrioventricular valvuloseptal morphogenesis. *Circ. Res.* **77**, 1–6 (1995).

97. Zhang, H., Lui, K. O. & Zhou, B. Endocardial Cell Plasticity in Cardiac Development, Diseases and Regeneration. *Circ. Res.* **122**, 774–789 (2018).

98. Wang, J. *et al.* Atrioventricular cushion transformation is mediated by ALK2 in the developing mouse heart. *Dev. Biol.* **286**, 299–310 (2005).

99. Saxon, J. G. *et al.* BMP2 expression in the endocardial lineage is required for AV endocardial cushion maturation and remodeling. *Dev. Biol.* **430**, 113–128 (2017).

100. Kisanuki, Y. Y. *et al.* Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis in vivo. *Dev. Biol.* **230**, 230–242 (2001).

101. de Lange, F. J. *et al.* Lineage and morphogenetic analysis of the cardiac valves. *Circ. Res.* **95**, 645–654 (2004).

102. DeRuiter, M. C. *et al.* Embryonic endothelial cells transdifferentiate into mesenchymal cells expressing smooth muscle actins in vivo and in vitro. *Circ. Res.* **80**, 444–451 (1997).

103. Arciniegas, E., Neves, C. Y., Carrillo, L. M., Zambrano, E. A. & Ramirez, R. Endothelial-mesenchymal transition occurs during embryonic pulmonary artery development. *Endothelium* **12**, 193–200 (2005).

104. Miyake, T. & Kalluri, R. Cardiac biology: Cell plasticity helps hearts to repair. *Nature* **514**, 575–576 (2014).

105. Goldman, S. *et al.* Long-term patency of saphenous vein and left internal mammary artery grafts after coronary artery bypass surgery: results from a Department of Veterans Affairs Cooperative Study. *J. Am. Coll. Cardiol.* **44**, 2149–2156 (2004).

106. Mehta, R. H. *et al.* Saphenous vein grafts with multiple versus single distal targets in patients undergoing coronary artery bypass surgery: one-year graft failure and five-year outcomes from the Project of Ex-Vivo Vein Graft Engineering via Transfection (PREVENT) IV trial. *Circulation* **124**, 280–288 (2011).

107. Allaire, E. & Clowes, A. W. Endothelial cell injury in cardiovascular surgery: the intimal hyperplastic response. *Ann. Thorac. Surg.* **63**, 582–591 (1997).

108. Wadey, K., Lopes, J., Bendeck, M. & George, S. Role of smooth muscle cells in coronary artery bypass grafting failure. *Cardiovasc. Res.* **114**, 601–610 (2018).

109. Bentzon, J. F. *et al.* Smooth muscle cells in atherosclerosis originate from the local vessel wall and not circulating progenitor cells in ApoE knockout mice.

Arterioscler. Thromb. Vasc. Biol. **26**, 2696–2702 (2006).

110. Goumans, M.-J., Liu, Z. & ten Dijke, P. TGF-beta signaling in vascular biology and dysfunction. *Cell Res.* **19**, 116–127 (2009).

111. Jiang, Z. *et al.* Established neointimal hyperplasia in vein grafts expands via TGF-beta-mediated progressive fibrosis. *Am. J. Physiol. Heart Circ. Physiol.* **297**, H1200-7 (2009).

112. Ross, R. Atherosclerosis--an inflammatory disease. *N. Engl. J. Med.* **340**, 115–126 (1999).

113. Mudau, M., Genis, A., Lochner, A. & Strijdom, H. Endothelial dysfunction: the early predictor of atherosclerosis. *Cardiovasc. J. Afr.* **23**, 222–231 (2012).

114. Moonen, J.-R. A. J. *et al.* Endothelial-to-mesenchymal transition contributes to fibro-proliferative vascular disease and is modulated by fluid shear stress. *Cardiovasc. Res.* **108**, 377–386 (2015).

115. Nigro, P., Abe, J.-I. & Berk, B. C. Flow shear stress and atherosclerosis: a matter of site specificity. *Antioxid. Redox Signal.* **15**, 1405–1414 (2011).

116. Chen, P.-Y. *et al.* Endothelial TGF-beta signalling drives vascular inflammation and atherosclerosis. *Nat. Metab.* **1**, 912–926 (2019).

117. Schermuly, R. T., Ghofrani, H. A., Wilkins, M. R. & Grimminger, F. Mechanisms of disease: pulmonary arterial hypertension. *Nat. Rev. Cardiol.* **8**, 443–455 (2011).

118. Qiao, L. *et al.* Endothelial fate mapping in mice with pulmonary hypertension. *Circulation* **129**, 692–703 (2014).

119. Good, R. B. *et al.* Endothelial to Mesenchymal Transition Contributes to Endothelial Dysfunction in Pulmonary Arterial Hypertension. *Am. J. Pathol.* **185**, 1850–1858 (2015).

120. Ranchoux, B. *et al.* Endothelial-to-mesenchymal transition in pulmonary hypertension. *Circulation* **131**, 1006–1018 (2015).

121. Mammoto, T., Muyleart, M., Konduri, G. G. & Mammoto, A. Twist1 in Hypoxia-induced Pulmonary Hypertension through Transforming Growth Factor-beta-Smad Signaling. *Am. J. Respir. Cell Mol. Biol.* **58**, 194–207 (2018).

122. Deng, Z. *et al.* Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. *Am. J. Hum. Genet.* **67**, 737–744 (2000).

123. Girerd, B. *et al.* Absence of influence of gender and BMPR2 mutation type on clinical phenotypes of pulmonary arterial hypertension. *Respir. Res.* **11**, 73 (2010).

124. Long, L. *et al.* Altered bone morphogenetic protein and transforming growth factor-beta signaling in rat models of pulmonary hypertension: potential for activin receptor-like kinase-5 inhibition in prevention and progression of disease. *Circulation* **119**, 566–576 (2009).

125. Hiepen, C. *et al.* BMPR2 acts as a gatekeeper to protect endothelial cells from increased TGFbeta responses and altered cell mechanics. *PLoS Biol.* **17**, e3000557 (2019).

126. Hopper, R. K. *et al.* In Pulmonary Arterial Hypertension, Reduced BMPR2 Promotes Endothelial-to-Mesenchymal Transition via HMGA1 and Its Target Slug. *Circulation* **133**, 1783–1794 (2016).

127. Yu, A. Y. *et al.* Impaired physiological responses to chronic hypoxia in mice partially deficient for hypoxia-inducible factor 1alpha. *J. Clin. Invest.* **103**, 691–696 (1999).

128. Brusselmans, K. *et al.* Heterozygous deficiency of hypoxia-inducible factor-2alpha protects mice against pulmonary hypertension and right ventricular dysfunction during prolonged hypoxia. *J. Clin. Invest.* **111**, 1519–1527 (2003).

129. Majmundar, A. J., Wong, W. J. & Simon, M. C. Hypoxia-inducible factors and the response to hypoxic stress. *Mol. Cell* **40**, 294–309 (2010).

130. Zhang, B. *et al.* Hypoxia induces endothelialmesenchymal transition in pulmonary vascular remodeling. *Int. J. Mol. Med.* **42**, 270–278 (2018).

131. Tang, H. *et al.* Endothelial HIF-2alpha contributes to severe pulmonary hypertension due to endothelial-to-mesenchymal transition. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **314**, L256–L275 (2018).

132. Simonneau, G. *et al.* Updated clinical classification of pulmonary hypertension. *J. Am. Coll. Cardiol.* **62**, D34-41 (2013).

133. Galie, N. *et al.* 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endor. *Eur. Heart J.* **37**, 67–119 (2016).

134. Kumar, A. & Neema, P. K. Severe pulmonary hypertension and right ventricular failure. *Indian J. Anaesth.* **61**, 753–759 (2017).

135. Humbert, M. *et al.* Pulmonary arterial hypertension in France: results from a national registry. *Am. J. Respir. Crit. Care Med.* **173**, 1023–1030 (2006).

136. Peacock, A. J., Murphy, N. F., McMurray, J. J. V, Caballero, L. & Stewart, S. An epidemiological study of pulmonary arterial hypertension. *Eur. Respir. J.* **30**, 104–109 (2007).

137. Ling, Y. *et al.* Changing demographics, epidemiology, and survival of incident pulmonary arterial hypertension: results from the pulmonary hypertension registry of the United Kingdom and Ireland. *Am. J. Respir. Crit. Care Med.* **186**, 790–796 (2012).

138. Benza, R. L. *et al.* Predicting survival in pulmonary arterial hypertension: insights from the Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL). *Circulation* **122**, 164–172 (2010).

139. Morse, J. H. *et al.* Mapping of familial primary pulmonary hypertension locus (PPH1) to chromosome 2q31-q32. *Circulation* **95**, 2603–2606 (1997).

140. Nichols, W. C. *et al.* Localization of the gene for familial primary pulmonary hypertension to chromosome 2q31-32. *Nat. Genet.* **15**, 277–280 (1997).

141. Deng, Z. *et al.* Fine mapping of PPH1, a gene for familial primary pulmonary hypertension, to a 3-cM region on chromosome 2q33. *Am. J. Respir. Crit. Care Med.* **161**, 1055–1059 (2000).

142. Cogan, J. D. *et al.* High frequency of BMPR2 exonic deletions/duplications in familial pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* **174**, 590–598 (2006).

143. Machado, R. D. *et al.* Mutations of the TGF-beta type II receptor BMPR2 in pulmonary arterial hypertension. *Hum. Mutat.* **27**, 121–132 (2006).

144. Fu, L. *et al.* A novel mutation in the BMPR2 gene in familial pulmonary arterial hypertension. *Chin. Med. J. (Engl).* **121**, 399–404 (2008).

145. Machado, R. D. *et al.* Genetics and genomics of pulmonary arterial hypertension. *J. Am. Coll. Cardiol.* **54**, S32-42 (2009).

146. Girerd, B. *et al.* Clinical outcomes of pulmonary arterial hypertension in patients carrying an ACVRL1 (ALK1) mutation. *Am. J. Respir. Crit. Care Med.* **181**, 851–861 (2010).

147. Pousada, G., Baloira, A., Fontan, D., Nunez, M. & Valverde, D. Mutational and clinical analysis of the ENG gene in patients with pulmonary arterial hypertension. *BMC Genet.* **17**, 72 (2016).

148. Soubrier, F. *et al.* Genetics and genomics of pulmonary arterial hypertension. *J. Am. Coll. Cardiol.* **62**, D13-21 (2013).

149. Sztrymf, B. *et al.* Clinical outcomes of pulmonary arterial hypertension in carriers of BMPR2 mutation. *Am. J. Respir. Crit. Care Med.* **177**, 1377–1383 (2008).

150. Yang, J., Li, X. & Morrell, N. W. Id proteins in the vasculature: from molecular biology to cardiopulmonary medicine. *Cardiovasc. Res.* **104**, 388–398 (2014).

151. Orriols, M., Gomez-Puerto, M. C. & Ten Dijke, P. BMP type II receptor as a therapeutic target in pulmonary arterial hypertension. *Cell. Mol. Life Sci.* **74**, 2979–2995 (2017).

152. Andruska, A. & Spiekerkoetter, E. Consequences of BMPR2 Deficiency in the Pulmonary Vasculature and Beyond: Contributions to Pulmonary Arterial Hypertension. *Int. J. Mol. Sci.* **19**, (2018).

153. Sawada, H. *et al.* Reduced BMPR2 expression induces GM-CSF translation and macrophage recruitment in humans and mice to exacerbate pulmonary hypertension. *J. Exp. Med.* **211**, 263–280 (2014).

154. Burton, V. J. *et al.* Bone morphogenetic protein receptor II regulates pulmonary artery endothelial cell barrier function. *Blood* **117**, 333–341 (2011).

155. Teichert-Kuliszewska, K. *et al.* Bone morphogenetic protein receptor-2 signaling promotes pulmonary arterial endothelial cell survival: implications for loss-of-function mutations in the pathogenesis of pulmonary hypertension. *Circ. Res.* **98**, 209–217 (2006).

156. Alastalo, T.-P. *et al.* Disruption of PPARgamma/beta-catenin-mediated regulation of apelin impairs BMP-induced mouse and human pulmonary arterial EC survival. *J. Clin. Invest.* **121**, 3735–3746 (2011).

157. Atkinson, C. *et al.* Primary pulmonary hypertension is associated with reduced pulmonary vascular expression of type II bone morphogenetic protein receptor. *Circulation* **105**, 1672–1678 (2002).

158. Takahashi, H. *et al.* Downregulation of type II bone morphogenetic protein receptor in hypoxic pulmonary hypertension. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **290**, L450-8 (2006).

159. Reynolds, A. M., Holmes, M. D., Danilov, S. M. & Reynolds, P. N. Targeted gene delivery of BMPR2 attenuates pulmonary hypertension. *Eur. Respir. J.* **39**, 329–343 (2012).

160. Jones, G., Robertson, L., Harrison, R., Ridout, C. & Vasudevan, P. Somatic mosaicism in ACVRL1 with transmission to several offspring affected with severe pulmonary arterial hypertension. *Am. J. Med. Genet. A* **164A**, 2121–2123 (2014).

161. Piao, C. *et al.* Identification of multiple ACVRL1 mutations in patients with pulmonary arterial hypertension by targeted exome capture. *Clin. Sci. (Lond).* **130**, 1559–1569 (2016).

162. David, L., Mallet, C., Mazerbourg, S., Feige, J.-J. & Bailly, S. Identification of BMP9 and BMP10 as functional activators of the orphan activin receptor-like kinase 1 (ALK1) in endothelial cells. *Blood* **109**, 1953–1961 (2007).

163. Roman, B. L. & Hinck, A. P. ALK1 signaling in development and disease: new paradigms. *Cell. Mol. Life Sci.* **74**, 4539–4560 (2017).

164. Thenappan, T., Ormiston, M. L., Ryan, J. J. & Archer, S. L. Pulmonary arterial hypertension: pathogenesis and clinical management. *BMJ* **360**, j5492 (2018).

165. Stacher, E. *et al.* Modern age pathology of pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* **186**, 261–272 (2012).

166. Tuder, R. M. Pulmonary vascular remodeling in pulmonary hypertension. *Cell Tissue Res.* **367**, 643–649 (2017).

167. Humbert, M. *et al.* Pathology and pathobiology of pulmonary hypertension: state of the art and research perspectives. *Eur. Respir. J.* **53**, (2019).

168. Archer, S. L., Weir, E. K. & Wilkins, M. R. Basic science of pulmonary arterial hypertension for clinicians: new concepts and experimental therapies. *Circulation* **121**, 2045–2066 (2010).

169. Tuder, R. M., Stacher, E., Robinson, J., Kumar, R. & Graham, B. B. Pathology of pulmonary hypertension. *Clin. Chest Med.* **34**, 639–650 (2013).

170. Stenmark, K. R., Meyrick, B., Galie, N., Mooi, W. J. & McMurtry, I. F. Animal models of pulmonary arterial hypertension: the hope for etiological discovery and pharmacological cure. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **297**, L1013-32 (2009).

171. Isobe, S. *et al.* Endothelial-Mesenchymal Transition Drives Expression of CD44 Variant and xCT in Pulmonary Hypertension. *Am. J. Respir. Cell Mol. Biol.* **61**, 367–379 (2019).

172. Suzuki, T. *et al.* Isolation and characterization of endothelial-to-mesenchymal transition cells in pulmonary arterial hypertension. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **314**, L118–L126 (2018).

173. Tian, L. & Chesler, N. C. In vivo and in vitro measurements of pulmonary arterial stiffness: A brief review. *Pulm. Circ.* **2**, 505–517 (2012).

174. Rodes-Cabau, J. *et al.* Intravascular ultrasound of the elastic pulmonary arteries: a new approach for the evaluation of primary pulmonary hypertension. *Heart* **89**, 311–315 (2003).

175. Mahapatra, S., Nishimura, R. A., Oh, J. K. & McGoon, M. D. The prognostic value of pulmonary vascular capacitance determined by Doppler echocardiography in patients with pulmonary arterial hypertension. *J. Am. Soc. Echocardiogr.* **19**, 1045–1050 (2006).

176. Gan, C. T.-J. *et al.* Noninvasively assessed pulmonary artery stiffness predicts mortality in pulmonary arterial hypertension. *Chest* **132**, 1906–1912 (2007).

177. Stevens, G. R. *et al.* RV dysfunction in pulmonary hypertension is independently related to pulmonary artery stiffness. *JACC. Cardiovasc. Imaging* **5**, 378–387 (2012).

178. Fishman, A. P. *et al.* Mechanisms of proliferative and obliterative vascular diseases. Insights from the pulmonary and systemic circulations. NHLBI Workshop summary. *American journal of respiratory and critical care medicine* vol. 158 670–674 (1998).

179. Ward, J. P. T. Point: Hypoxic pulmonary vasoconstriction is mediated by increased production of reactive oxygen species. *J. Appl. Physiol.* **101**, 993–5; discussion 999 (2006).

180. Nicolls, M. R., Taraseviciene-Stewart, L., Rai, P. R., Badesch, D. B. & Voelkel, N. F. Autoimmunity and pulmonary hypertension: a perspective. *Eur. Respir. J.* **26**, 1110–1118 (2005).

181. Taraseviciene-Stewart, L. *et al.* Inhibition of the VEGF receptor 2 combined with chronic hypoxia causes cell death-dependent pulmonary endothelial cell proliferation and severe pulmonary hypertension. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **15**, 427–438 (2001).

182. Sakao, S. *et al.* Initial apoptosis is followed by increased proliferation of apoptosis-resistant endothelial cells. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **19**, 1178–1180 (2005).

183. Sakao, S., Tatsumi, K. & Voelkel, N. F. Endothelial cells and pulmonary arterial

hypertension: apoptosis, proliferation, interaction and transdifferentiation. *Respir. Res.* **10**, 95 (2009).

184. Sakao, S., Taraseviciene-Stewart, L., Wood, K., Cool, C. D. & Voelkel, N. F. Apoptosis of pulmonary microvascular endothelial cells stimulates vascular smooth muscle cell growth. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **291**, L362-8 (2006).

185. Yeager, M. E., Halley, G. R., Golpon, H. A., Voelkel, N. F. & Tuder, R. M. Microsatellite instability of endothelial cell growth and apoptosis genes within plexiform lesions in primary pulmonary hypertension. *Circ. Res.* **88**, E2–E11 (2001).

186. Masri, F. A. *et al.* Hyperproliferative apoptosis-resistant endothelial cells in idiopathic pulmonary arterial hypertension. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **293**, L548-54 (2007).

187. Tuder, R. M. *et al.* Expression of angiogenesis-related molecules in plexiform lesions in severe pulmonary hypertension: evidence for a process of disordered angiogenesis. *J. Pathol.* **195**, 367–374 (2001).

188. Jonigk, D. *et al.* Plexiform lesions in pulmonary arterial hypertension composition, architecture, and microenvironment. *Am. J. Pathol.* **179**, 167–179 (2011).

189. Balk, A. G., Dingemans, K. P. & Wagenvoort, C. A. The ultrastructure of the various forms of pulmonary arterial intimal fibrosis. *Virchows Arch. A. Pathol. Anat. Histol.* **382**, 139–150 (1979).

190. Tuder, R. M., Marecki, J. C., Richter, A., Fijalkowska, I. & Flores, S. Pathology of pulmonary hypertension. *Clin. Chest Med.* **28**, 23–42, vii (2007).

191. Montani, D. *et al.* Pulmonary arterial hypertension. *Orphanet J. Rare Dis.* **8**, 97 (2013).

192. Cool, C. D., Kennedy, D., Voelkel, N. F. & Tuder, R. M. Pathogenesis and evolution of plexiform lesions in pulmonary hypertension associated with scleroderma and human immunodeficiency virus infection. *Hum. Pathol.* **28**, 434–442 (1997).

193. Voelkel, N. F. & Tuder, R. M. Cellular and molecular biology of vascular smooth muscle cells in pulmonary hypertension. *Pulm. Pharmacol. Ther.* **10**, 231–241 (1997).

194. Savai, R. *et al.* Pro-proliferative and inflammatory signaling converge on FoxO1 transcription factor in pulmonary hypertension. *Nat. Med.* **20**, 1289–1300 (2014).

195. Perros, F. *et al.* Smooth Muscle Phenotype in Idiopathic Pulmonary Hypertension: Hyper-Proliferative but not Cancerous. *Int. J. Mol. Sci.* **20**, (2019).

196. Goncharov, D. A. *et al.* Mammalian target of rapamycin complex 2 (mTORC2) coordinates pulmonary artery smooth muscle cell metabolism, proliferation, and survival in pulmonary arterial hypertension. *Circulation* **129**, 864–874

(2014).

197. Aggarwal, S., Gross, C. M., Sharma, S., Fineman, J. R. & Black, S. M. Reactive oxygen species in pulmonary vascular remodeling. *Compr. Physiol.* **3**, 1011–1034 (2013).

198. Zaiman, A. L. *et al.* Role of the TGF-beta/Alk5 signaling pathway in monocrotaline-induced pulmonary hypertension. *Am. J. Respir. Crit. Care Med.* **177**, 896–905 (2008).

199. Yang, K., Wang, J. & Lu, W. Bone morphogenetic protein signalling in pulmonary hypertension: advances and therapeutic implications. *Exp. Physiol.* **102**, 1083–1089 (2017).

200. Marcos, E. *et al.* Serotonin-induced smooth muscle hyperplasia in various forms of human pulmonary hypertension. *Circ. Res.* **94**, 1263–1270 (2004).

201. Cacoub, P. *et al.* Endothelin-1 in the lungs of patients with pulmonary hypertension. *Cardiovasc. Res.* **33**, 196–200 (1997).

202. Deng, H. *et al.* Pulmonary artery smooth muscle hypertrophy: roles of glycogen synthase kinase-3beta and p70 ribosomal S6 kinase. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **298**, L793-803 (2010).

203. Kim, J. *et al.* An endothelial apelin-FGF link mediated by miR-424 and miR-503 is disrupted in pulmonary arterial hypertension. *Nat. Med.* **19**, 74–82 (2013).

204. Zhang, H. *et al.* Apelin inhibits the proliferation and migration of rat PASMCs via the activation of PI3K/Akt/mTOR signal and the inhibition of autophagy under hypoxia. *J. Cell. Mol. Med.* **18**, 542–553 (2014).

205. Black, S. M., DeVol, J. M. & Wedgwood, S. Regulation of fibroblast growth factor-2 expression in pulmonary arterial smooth muscle cells involves increased reactive oxygen species generation. *Am. J. Physiol. Cell Physiol.* **294**, C345-54 (2008).

206. Wedgwood, S. *et al.* Fibroblast growth factor-2 expression is altered in lambs with increased pulmonary blood flow and pulmonary hypertension. *Pediatr. Res.* **61**, 32–36 (2007).

207. Remillard, C. V & Yuan, J. X.-J. High altitude pulmonary hypertension: role of K+ and Ca2+ channels. *High Alt. Med. Biol.* **6**, 133–146 (2005).

208. Stenmark, K. R., Frid, M. G., Graham, B. B. & Tuder, R. M. Dynamic and diverse changes in the functional properties of vascular smooth muscle cells in pulmonary hypertension. *Cardiovasc. Res.* **114**, 551–564 (2018).

209. Montani, D. *et al.* Targeted therapies in pulmonary arterial hypertension. *Pharmacol. Ther.* **141**, 172–191 (2014).

210. Vane, J. & Corin, R. E. Prostacyclin: a vascular mediator. *Eur. J. Vasc. Endovasc. Surg.* **26**, 571–578 (2003).

211. Chester, A. H. & Yacoub, M. H. The role of endothelin-1 in pulmonary arterial hypertension. *Glob. Cardiol. Sci. Pract.* **2014**, 62–78 (2014).

212. Clozel, M. *et al.* Pharmacological characterization of bosentan, a new potent orally active nonpeptide endothelin receptor antagonist. *J. Pharmacol. Exp. Ther.* **270**, 228–235 (1994).

213. Wilkins, M. R. Selective or nonselective endothelin receptor blockade in pulmonary arterial hypertension. *American journal of respiratory and critical care medicine* vol. 169 433–434 (2004).

214. Chester, A. H., Yacoub, M. H. & Moncada, S. Nitric oxide and pulmonary arterial hypertension. *Glob. Cardiol. Sci. Pract.* **2017**, 14 (2017).

215. Bender, A. T. & Beavo, J. A. Cyclic nucleotide phosphodiesterases: molecular regulation to clinical use. *Pharmacol. Rev.* **58**, 488–520 (2006).

216. O'Callaghan, D. S. *et al.* Treatment of pulmonary arterial hypertension with targeted therapies. *Nat. Rev. Cardiol.* **8**, 526–538 (2011).

217. Galie, N. *et al.* Sildenafil citrate therapy for pulmonary arterial hypertension. *N. Engl. J. Med.* **353**, 2148–2157 (2005).

218. Rubin, L. J. *et al.* Long-term treatment with sildenafil citrate in pulmonary arterial hypertension: the SUPER-2 study. *Chest* **140**, 1274–1283 (2011).

219. Ghofrani, H.-A. & Humbert, M. The role of combination therapy in managing pulmonary arterial hypertension. *Eur. Respir. Rev.* **23**, 469–475 (2014).

220. Bartolome, S., Hoeper, M. M. & Klepetko, W. Advanced pulmonary arterial hypertension: mechanical support and lung transplantation. *Eur. Respir. Rev.* **26**, (2017).

221. Taylor, D. O. *et al.* Registry of the International Society for Heart and Lung Transplantation: twenty-third official adult heart transplantation report--2006. *J. Heart Lung Transplant.* **25**, 869–879 (2006).

222. Toyoda, Y. *et al.* Long-term outcome of lung and heart-lung transplantation for idiopathic pulmonary arterial hypertension. *Ann. Thorac. Surg.* **86**, 1116–1122 (2008).

223. Fadel, E. *et al.* Long-term outcome of double-lung and heart-lung transplantation for pulmonary hypertension: a comparative retrospective study of 219 patients. *Eur. J. Cardiothorac. Surg.* **38**, 277–284 (2010).

224. de Perrot, M. *et al.* Outcome of patients with pulmonary arterial hypertension referred for lung transplantation: a 14-year single-center experience. *J. Thorac. Cardiovasc. Surg.* **143**, 910–918 (2012).

225. Humbert, M. *et al.* Survival in patients with idiopathic, familial, and anorexigen-associated pulmonary arterial hypertension in the modern management era. *Circulation* **122**, 156–163 (2010).

226. Chung, L. *et al.* Survival and predictors of mortality in systemic sclerosis-associated pulmonary arterial hypertension: outcomes from the pulmonary hypertension assessment and recognition of outcomes in scleroderma registry. *Arthritis Care Res. (Hoboken).* **66**, 489–495 (2014).

227. Bonnin, M. *et al.* Severe pulmonary hypertension during pregnancy: mode of

delivery and anesthetic management of 15 consecutive cases. *Anesthesiology* **102**, 1133–7; discussion 5A-6A (2005).

228. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

229. Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

230. Palazzo, A. F. & Lee, E. S. Non-coding RNA: what is functional and what is junk? *Front. Genet.* **6**, (2015).

231. Kapranov, P., Willingham, A. T. & Gingeras, T. R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **8**, 413–423 (2007).

232. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).

233. Song, X., Shan, D., Chen, J. & Jing, Q. miRNAs and lncRNAs in vascular injury and remodeling. *Sci. China. Life Sci.* **57**, 826–835 (2014).

234. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).

235. Fiedler, J., Gupta, S. K. & Thum, T. MicroRNA-based therapeutic approaches in the cardiovascular system. *Cardiovasc. Ther.* **30**, e9–e15 (2012).

236. Brower, J. V, Clark, P. A., Lyon, W. & Kuo, J. S. MicroRNAs in cancer: glioblastoma and glioblastoma cancer stem cells. *Neurochem. Int.* **77**, 68–77 (2014).

237. Lee, Y., Jeon, K., Lee, J.-T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663–4670 (2002).

238. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).

239. Pfeffer, S. *et al.* Identification of microRNAs of the herpesvirus family. *Nat. Methods* **2**, 269–276 (2005).

240. Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13**, 1097–1101 (2006).

241. Gebert, L. F. R. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **20**, 21–37 (2019).

242. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).

243. Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57–68 (2004).

244. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC

couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631–640 (2005).

245. Haase, A. D. *et al.* TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep.* **6**, 961–967 (2005).

246. Lee, Y. *et al.* The role of PACT in the RNA silencing pathway. *EMBO J.* **25**, 522–532 (2006).

247. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V & Doudna, J. A. In vitro reconstitution of the human RISC-loading complex. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 512–517 (2008).

248. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209–216 (2003).

249. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199–208 (2003).

250. Frank, F., Sonenberg, N. & Nagar, B. Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818–822 (2010).

251. Suzuki, H. I. *et al.* Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nat. Struct. Mol. Biol.* **22**, 512–521 (2015).

252. Betel, D., Koppal, A., Agius, P., Sander, C. & Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* **11**, R90 (2010).

253. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, (2015).

254. Wong, N. & Wang, X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* **43**, D146-52 (2015).

255. Xu, K., Lin, J., Zandi, R., Roth, J. A. & Ji, L. MicroRNA-mediated target mRNA cleavage and 3'-uridylation in human cells. *Sci. Rep.* **6**, 30242 (2016).

256. Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* **16**, 421–433 (2015).

257. Wahle, E. & Winkler, G. S. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim. Biophys. Acta* **1829**, 561–570 (2013).

258. Jonas, S. & Izaurralde, E. The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev.* **27**, 2628–2641 (2013).

259. Braun, J. E. *et al.* A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation. *Nat. Struct. Mol. Biol.* **19**, 1324–1331 (2012).

260. Petersen, C. P., Bordeleau, M.-E., Pelletier, J. & Sharp, P. A. Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* **21**, 533–542 (2006).

261. Ryu, I., Park, J. H., An, S., Kwon, O. S. & Jang, S. K. eIF4GI facilitates the MicroRNA-mediated gene silencing. *PLoS One* **8**, e55725 (2013).

262. Chu, C. & Rana, T. M. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol.* **4**, e210 (2006).

263. Schutz, P. *et al.* Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9564–9569 (2008).

264. Bhaskaran, M. & Mohan, M. MicroRNAs: history, biogenesis, and their evolving role in animal development and disease. *Vet. Pathol.* **51**, 759–774 (2014).

265. Creemers, E. E., Tijsen, A. J. & Pinto, Y. M. Circulating microRNAs: novel biomarkers and extracellular communicators in cardiovascular disease? *Circ. Res.* **110**, 483–495 (2012).

266. Boon, R. A. & Dimmeler, S. MicroRNAs in myocardial infarction. *Nat. Rev. Cardiol.* **12**, 135–142 (2015).

267. Endo, K. *et al.* MicroRNA 210 as a biomarker for congestive heart failure. *Biol. Pharm. Bull.* **36**, 48–54 (2013).

268. Li, S. *et al.* Signature microRNA expression profile of essential hypertension and its novel link to human cytomegalovirus infection. *Circulation* **124**, 175–184 (2011).

269. Yang, W. J. *et al.* Dicer is required for embryonic angiogenesis during mouse development. *J. Biol. Chem.* **280**, 9330–9335 (2005).

270. Suarez, Y., Fernandez-Hernando, C., Pober, J. S. & Sessa, W. C. Dicer dependent microRNAs regulate gene expression and functions in human endothelial cells. *Circ. Res.* **100**, 1164–1173 (2007).

271. Harris, T. A., Yamakuchi, M., Ferlito, M., Mendell, J. T. & Lowenstein, C. J. MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 1516–1521 (2008).

272. Urbich, C., Kuehbacher, A. & Dimmeler, S. Role of microRNAs in vascular diseases, inflammation, and angiogenesis. *Cardiovasc. Res.* **79**, 581–588 (2008).

273. Poliseno, L. *et al.* MicroRNAs modulate the angiogenic properties of HUVECs. *Blood* **108**, 3068–3071 (2006).

274. Matsui, J., Wakabayashi, T., Asada, M., Yoshimatsu, K. & Okada, M. Stem cell factor/c-kit signaling promotes the survival, migration, and capillary tube formation of human umbilical vein endothelial cells. *J. Biol. Chem.* **279**, 18600–18607 (2004).

275. Zhan, Y. *et al.* Ets-1 is a critical regulator of Ang II-mediated vascular inflammation and remodeling. *J. Clin. Invest.* **115**, 2508–2516 (2005).

276. Zhu, N. *et al.* Endothelial enriched microRNAs regulate angiotensin II-induced endothelial inflammation and migration. *Atherosclerosis* **215**, 286–293 (2011).

277. Caruso, P. *et al.* Dynamic changes in lung microRNA profiles during the development of pulmonary hypertension due to chronic hypoxia and monocrotaline. *Arterioscler. Thromb. Vasc. Biol.* **30**, 716–723 (2010).

278. Caruso, P. *et al.* A role for miR-145 in pulmonary arterial hypertension: evidence from mouse models and patient samples. *Circ. Res.* **111**, 290–300 (2012).

279. Deng, L. *et al.* MicroRNA-143 Activation Regulates Smooth Muscle and Endothelial Cell Crosstalk in Pulmonary Arterial Hypertension. *Circ. Res.* **117**, 870–883 (2015).

280. McDonald, R. A. *et al.* Reducing In-Stent Restenosis: Therapeutic Manipulation of miRNA in Vascular Remodeling and Inflammation. *J. Am. Coll. Cardiol.* **65**, 2314–2327 (2015).

281. McDonald, R. A. *et al.* miRNA-21 is dysregulated in response to vein grafting in multiple models and genetic ablation in mice attenuates neointima formation. *Eur. Heart J.* **34**, 1636–1643 (2013).

282. Ghosh, A. K., Nagpal, V., Covington, J. W., Michaels, M. A. & Vaughan, D. E. Molecular basis of cardiac endothelial-to-mesenchymal transition (EndMT): differential expression of microRNAs during EndMT. *Cell. Signal.* **24**, 1031–1036 (2012).

283. Katsura, A. *et al.* MicroRNA-31 is a positive modulator of endothelial-mesenchymal transition and associated secretory phenotype induced by TGF-beta. *Genes Cells* **21**, 99–116 (2016).

284. Zhu, K. *et al.* MiR-302c inhibits tumor growth of hepatocellular carcinoma by suppressing the endothelial-mesenchymal transition of endothelial cells. *Sci. Rep.* **4**, 5524 (2014).

285. Deveson, I. W., Hardwick, S. A., Mercer, T. R. & Mattick, J. S. The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome. *Trends Genet.* **33**, 464–478 (2017).

286. Schmitz, S. U., Grote, P. & Herrmann, B. G. Mechanisms of long noncoding RNA function in development and disease. *Cell. Mol. Life Sci.* **73**, 2491–2509 (2016).

287. Hezroni, H. *et al.* Principles of Long Noncoding RNA Evolution Derived from Direct Comparison of Transcriptomes in 17 Species. *Cell Rep.* **11**, 1110–1122 (2015).

288. Hon, C. C. *et al.* An atlas of human long non-coding RNAs with accurate 5 ' ends. *Nature* **543**, 199-+ (2017).

289. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789 (2012).

290. Geisler, S., Lojek, L., Khalil, A. M., Baker, K. E. & Coller, J. Decapping of long noncoding RNAs regulates inducible genes. *Mol. Cell* **45**, 279–291 (2012).

291. Zhang, X.-O., Yin, Q.-F., Chen, L.-L. & Yang, L. Gene expression profiling of non-polyadenylated RNA-seq across species. *Genomics data* **2**, 237–241 (2014).

292. Quinn, J. J. & Chang, H. Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **17**, 47–62 (2016).

293. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).

294. Nakaya, H. I. *et al.* Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biol.* **8**, R43 (2007).

295. Jensen, T. H., Jacquier, A. & Libri, D. Dealing with pervasive transcription. *Mol. Cell* **52**, 473–484 (2013).

296. Bunch, H. Gene regulation of mammalian long non-coding RNA. *Mol. Genet. Genomics* **293**, 1–15 (2018).

297. Ayupe, A. C. *et al.* Global analysis of biogenesis, stability and sub-cellular localization of lncRNAs mapping to intragenic regions of the human genome. *RNA Biol.* **12**, 877–892 (2015).

298. Furuno, M. *et al.* Clusters of internally primed transcripts reveal novel long noncoding RNAs. *PLoS Genet.* **2**, e37 (2006).

299. Cheng, J. *et al.* Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **308**, 1149–1154 (2005).

300. Noviello, T. M. R. *et al.* Detection of long non-coding RNA homology, a comparative study on alignment and alignment-free metrics. *BMC Bioinformatics* **19**, 407 (2018).

301. Nesterova, T. B. *et al.* Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome Res.* **11**, 833–849 (2001).

302. Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A. & Panning, B. Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* **36**, 233–278 (2002).

303. Cantrell, M. A., Carstens, B. C. & Wichman, H. A. X chromosome inactivation and Xist evolution in a rodent lacking LINE-1 activity. *PLoS One* **4**, e6252 (2009).

304. Sprague, D. *et al.* Nonlinear sequence similarity between the Xist and Rsx long noncoding RNAs suggests shared functions of tandem repeat domains. *RNA* **25**, 1004–1019 (2019).

305. Chen, J. *et al.* Evolutionary analysis across mammals reveals distinct classes of long non-coding RNAs. *Genome Biol.* **17**, (2016).

306. St Laurent, G., Wahlestedt, C. & Kapranov, P. The Landscape of long noncoding RNA classification. *Trends Genet.* **31**, 239–251 (2015).

307. Harrow, J. *et al.* GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).

308. Katayama, S. *et al.* Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564–1566 (2005).

309. Lin, S., Zhang, L., Luo, W. & Zhang, X. Characteristics of Antisense Transcript Promoters and the Regulation of Their Activity. *Int. J. Mol. Sci.* **17**, (2015).

310. Latge, G., Poulet, C., Bours, V., Josse, C. & Jerusalem, G. Natural Antisense Transcripts: Molecular Mechanisms and Implications in Breast Cancers. *Int. J. Mol. Sci.* **19**, (2018).

311. Sigova, A. A. *et al.* Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2876–2881 (2013).

312. Han, J., Kim, D. & Morris, K. V. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 12422–12427 (2007).

313. Hu, H. Y., He, L. & Khaitovich, P. Deep sequencing reveals a novel class of bidirectional promoters associated with neuronal genes. *BMC Genomics* **15**, 457 (2014).

314. Monteiro, J. P. *et al.* Endothelial function and dysfunction in the cardiovascular system: the long non-coding road. *Cardiovasc. Res.* **115**, 1692–1704 (2019).

315. Ulitsky, I. & Bartel, D. P. lincRNAs: Genomics, Evolution, and Mechanisms. *Cell* **154**, 26–46 (2013).

316. Michalik, K. M. *et al.* Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ. Res.* **114**, 1389–1397 (2014).

317. Boulberdaa, M. *et al.* A role for the long non-coding RNA SENCR in commitment and function of endothelial cells. *Mol. Ther.* (2016) doi:10.1038/mt.2016.41.

318. Miyagawa, R. *et al.* Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA* **18**, 738–751 (2012).

319. Hutchinson, J. N. *et al.* A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**, 39 (2007).

320. West, J. A. *et al.* The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* **55**, 791–802 (2014).

321. Mao, Y. S., Sunwoo, H., Zhang, B. & Spector, D. L. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* **13**, 95–101 (2011).

322. Nakagawa, S., Naganuma, T., Shioi, G. & Hirose, T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J. Cell Biol.* **193**, 31–39 (2011).

323. Naganuma, T. *et al.* Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* **31**, 4020–4034 (2012).

324. Wang, Y. *et al.* Genome-wide screening of NEAT1 regulators reveals cross-regulation between paraspeckles and mitochondria. *Nat. Cell Biol.* **20**, 1145–1158 (2018).

325. Heard, E. & Disteche, C. M. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev.* **20**, 1848–1867 (2006).

326. Erwin, J. A. & Lee, J. T. New twists in X-chromosome inactivation. *Curr. Opin. Cell Biol.* **20**, 349–355 (2008).

327. Payer, B. & Lee, J. T. X chromosome dosage compensation: how mammals keep the balance. *Annu. Rev. Genet.* **42**, 733–772 (2008).

328. Cerase, A., Pintacuda, G., Tattermusch, A. & Avner, P. Xist localization and function: new insights from multiple levels. *Genome Biol.* **16**, 166 (2015).

329. Engreitz, J. M., Ollikainen, N. & Guttman, M. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat. Rev. Mol. Cell Biol.* **17**, 756–770 (2016).

330. Ransohoff, J. D., Wei, Y. N. & Khavari, P. A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **19**, 143–157 (2018).

331. Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11667–11672 (2009).

332. Geisler, S. & Coller, J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat. Rev. Mol. Cell Biol.* **14**, 699–712 (2013).

333. Engreitz, J. M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).

334. Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).

335. da Rocha, S. T. & Heard, E. Novel players in X inactivation: insights into Xist-mediated gene silencing and chromosome conformation. *Nat. Struct. Mol. Biol.* **24**, 197–204 (2017).

336. Hasegawa, Y. *et al.* The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* **19**, 469–476 (2010).

337. Sakaguchi, T. *et al.* Control of Chromosomal Localization of Xist by hnRNP U Family Molecules. *Developmental cell* vol. 39 11–12 (2016).

338. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).

339. Escamilla-Del-Arenal, M., da Rocha, S. T. & Heard, E. Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum. Genet.* **130**,

307–327 (2011).

340. Brockdorff, N. Polycomb complexes in X chromosome inactivation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372**, (2017).

341. Lu, Z., Carter, A. C. & Chang, H. Y. Mechanistic insights in X-chromosome inactivation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372**, (2017).

342. Li, L. *et al.* Targeted disruption of Hotair leads to homeotic transformation and gene derepression. *Cell Rep.* **5**, 3–12 (2013).

343. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43**, 904–914 (2011).

344. Li, T., Mo, X., Fu, L., Xiao, B. & Guo, J. Molecular mechanisms of long noncoding RNAs on gastric cancer. *Oncotarget* **7**, 8601–8612 (2016).

345. Simon, M. D. *et al.* The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20497–20502 (2011).

346. Mondal, T. *et al.* MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* **6**, (2015).

347. Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).

348. Santos-Pereira, J. M. & Aguilera, A. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583–597 (2015).

349. Boque-Sastre, R. *et al.* Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5785–5790 (2015).

350. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).

351. Hung, T. *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.* **43**, 621–629 (2011).

352. Yang, F. *et al.* Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS J.* **279**, 3159–3165 (2012).

353. Kino, T., Hurt, D. E., Ichijo, T., Nader, N. & Chrousos, G. P. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.* **3**, ra8 (2010).

354. Lucafo, M. *et al.* Long noncoding RNA GAS5: a novel marker involved in glucocorticoid response. *Curr. Mol. Med.* **15**, 94–99 (2015).

355. Sharma, N. S. *et al.* Long non-coding RNA GAS5 acts as proliferation 'brakes' in CD133+ cells responsible for tumor recurrence. *Oncogenesis* **8**, 68 (2019).

356. Gong, C. & Maquat, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284–288 (2011).

357. Nishizawa, M., Okumura, T., Ikeya, Y. & Kimura, T. Regulation of inducible gene expression by natural antisense transcripts. *Front. Biosci. (Landmark Ed.* **17**, 938–958 (2012).

358. Xu, T.-P. *et al.* SP1-induced upregulation of the long noncoding RNA TINCR regulates cell proliferation and apoptosis by affecting KLF2 mRNA stability in gastric cancer. *Oncogene* **34**, 5648–5661 (2015).

359. Kretz, M. *et al.* Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**, 231–235 (2013).

360. Cai, X. & Cullen, B. R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **13**, 313–316 (2007).

361. Franco-Zorrilla, J. M. *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **39**, 1033–1037 (2007).

362. Hansen, T. B. *et al.* Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).

363. Samani, N. J. *et al.* Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* **357**, 443–453 (2007).

364. Broadbent, H. M. *et al.* Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum. Mol. Genet.* **17**, 806–814 (2008).

365. Holdt, L. M. *et al.* ANRIL Expression Is Associated With Atherosclerosis Risk at Chromosome 9p21. *Arterioscler. Thromb. Vasc. Biol.* **30**, 620-U518 (2010).

366. Pasmant, E., Sabbagh, A., Vidaud, M. & Bieche, I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **25**, 444–448 (2011).

367. Congrains, A. *et al.* Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. *Atherosclerosis* **220**, 449–455 (2012).

368. Motterle, A. *et al.* Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum. Mol. Genet.* **21**, 4021–4029 (2012).

369. Tang, R. *et al.* LncRNA GAS5 regulates vascular smooth muscle cell cycle arrest and apoptosis via p53 pathway. *Biochim. Biophys. acta. Mol. basis Dis.* **1865**, 2516–2525 (2019).

370. Li, K. *et al.* A noncoding antisense RNA in tie-1 locus regulates tie-1 function in vivo. *Blood* **115**, 133–139 (2010).

371. Liu, J.-Y. *et al.* Pathogenic role of lncRNA-MALAT1 in endothelial cell dysfunction in diabetes mellitus. *Cell Death Dis.* **5**, e1506 (2014).

372. Wang, X. *et al.* Long noncoding RNA expression profiles of hypoxic pulmonary hypertension rat model. *Gene* **579**, 23–28 (2016).

373. Gu, S. *et al.* Aberrant expression of long noncoding RNAs in chronic

thromboembolic pulmonary hypertension. *Mol. Med. Rep.* **11**, 2631–2643 (2015).

374. Cesana, M. *et al.* A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **147**, 358–369 (2011).

375. Faghihi, M. A. *et al.* Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* **11**, R56 (2010).

376. Augoff, K., McCue, B., Plow, E. F. & Sossey-Alaoui, K. miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. *Mol. Cancer* **11**, 5 (2012).

377. Dey, B. K., Pfeifer, K. & Dutta, A. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev.* **28**, 491–501 (2014).

378. Leucci, E. *et al.* microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Scientific Reports* vol. 3 (2013).

379. Singh, K. K. *et al.* Investigation of novel LPS-induced differentially expressed long non-coding RNAs in endothelial cells. *Mol. Cell. Biochem.* **421**, 157–168 (2016).

380. Fanucchi, S. *et al.* Immune genes are primed for robust transcription by proximal long noncoding RNAs located in nuclear compartments. *Nat. Genet.* **51**, 138-+ (2019).

381. Fiedler, J. *et al.* Development of Long Noncoding RNA-Based Strategies to Modulate Tissue Vascularization. *J. Am. Coll. Cardiol.* **66**, 2005–2015 (2015).

382. Voellenkle, C. *et al.* Implication of Long noncoding RNAs in the endothelial cell response to hypoxia revealed by RNA-sequencing. *Sci. Rep.* **6**, (2016).

383. Qiu, G. Z., Tian, W., Fu, H. T., Li, C. P. & Liu, B. Long noncoding RNA-MEG3 is involved in diabetes mellitus-related microvascular dysfunction. *Biochem. Biophys. Res. Commun.* **471**, 135–141 (2016).

384. Yan, B. A. *et al.* lncRNA-MIAT Regulates Microvascular Dysfunction by Functioning as a Competing Endogenous RNA. *Circ. Res.* **116**, 1143-+ (2015).

385. Shan, K., Li, C. P., Liu, C., Liu, X. & Yan, B. A. RNCR3: A regulator of diabetes mellitus-related retinal microvascular dysfunction. *Biochem. Biophys. Res. Commun.* **482**, 777–783 (2017).

386. Boulberdaa, M. *et al.* A Role for the Long Noncoding RNA SENCR in Commitment and Function of Endothelial Cells. *Mol. Ther.* **24**, 978–990 (2016).

387. Kurian, L. *et al.* Identification of Novel Long Noncoding RNAs Underlying Vertebrate Cardiovascular Development. *Circulation* **131**, 1278–1290 (2015).

388. Lyu, Q. *et al.* SENCR stabilizes vascular endothelial cell adherens junctions through interaction with CKAP4. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 546–555 (2019).

389. Nolan, D. J. *et al.* Molecular signatures of tissue-specific microvascular

endothelial cell heterogeneity in organ maintenance and regeneration. *Dev. Cell* **26**, 204–219 (2013).

390. Jansen, F., Li, Q., Pfeifer, A. & Werner, N. Endothelial- and Immune Cell-Derived Extracellular Vesicles in the Regulation ofCardiovascular Health and Disease. *JACC. Basic to Transl. Sci.* **2**, 790–807 (2017).

391. Dragomir, M., Chen, B. & Calin, G. A. Exosomal lncRNAs as new players in cell-to-cell communication. *Transl. Cancer Res.* **7**, S243–S252 (2018).

392. Shan, K. *et al.* Role of long non-coding RNA-RNCR3 in atherosclerosis-related vascular dysfunction. *Cell Death Dis.* **7**, (2016).

393. Chen, L. *et al.* Exosomal lncRNA GAS5 regulates the apoptosis of macrophages and vascular endothelial cells in atherosclerosis. *PLoS One* **12**, (2017).

394. Conigliaro, A. *et al.* CD90+liver cancer cells modulate endothelial cell phenotype through the release of exosomes containing H19 lncRNA. *Mol. Cancer* **14**, (2015).

395. Man, H. S. J. *et al.* Angiogenic patterning by STEEL, an endothelial-enriched long noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 2401–2406 (2018).

396. Balas, M. M. & Johnson, A. M. Exploring the mechanisms behind long noncoding RNAs and cancer. *Non-coding RNA Res.* **3**, 108–117 (2018).

397. Xiang, Y., Zhang, Y., Tang, Y. & Li, Q. MALAT1 Modulates TGF-beta1-Induced Endothelial-to-Mesenchymal Transition through Downregulation of miR-145. *Cell. Physiol. Biochem.* **42**, 357–372 (2017).

398. Neumann, P. *et al.* The lncRNA GATA6-AS epigenetically regulates endothelial gene expression via interaction with LOXL2. *Nat. Commun.* **9**, (2018).

399. Tang, Q. & Hann, S. S. HOTAIR: An Oncogenic Long Non-Coding RNA in Human Cancer. *Cell. Physiol. Biochem.* **47**, 893–913 (2018).

400. Gutschner, T., Haemmerle, M. & Diederichs, S. MALAT1- a paradigm for long noncoding RNA function in cancer. *J. Mol. Med.* **91**, 791–801 (2013).

401. Zhang, X., Hamblin, M. H. & Yin, K.-J. The long noncoding RNA Malat1: Its physiological and pathophysiological functions. *RNA Biol.* **14**, 1705–1714 (2017).

402. Tian, X. & Xu, G. Clinical value of lncRNA MALAT1 as a prognostic marker in human cancer: systematic review and meta-analysis. *BMJ Open* **5**, (2015).

403. Zhou, Q. *et al.* lncRNAs as potential molecular biomarkers for the clinicopathology and prognosis of glioma: A systematic review and meta-analysis. *Gene* **668**, 77–86 (2018).

404. Chen, J. *et al.* LncRNAs act as prognostic and diagnostic biomarkers in renal cell carcinoma: a systematic review and meta-analysis. *Oncotarget* **7**, 74325–74336 (2016).

405. Zhang, J.-H., Wei, H.-W. & Yang, H.-G. Long noncoding RNA SNHG15, a potential prognostic biomarker for hepatocellular carcinoma. *Eur. Rev. Med. Pharmacol. Sci.* **20**, 1720–1724 (2016).

406. Zhang, Z. *et al.* Increased plasma levels of lncRNA H19 and LIPCAR are associated with increased risk of coronary artery disease in a Chinese population. *Sci. Rep.* **7**, 7491 (2017).

407. Kumarswamy, R. *et al.* Circulating Long Noncoding RNA, LIPCAR, Predicts Survival in Patients With Heart Failure. *Circ. Res.* **114**, 1569–1575 (2014).

408. Li, M. *et al.* Circulating Long Noncoding RNA LIPCAR Acts as a Novel Biomarker in Patients with ST-Segment Elevation Myocardial Infarction. *Med. Sci. Monit.* **24**, 5064–5070 (2018).

409. Vausort, M., Wagner, D. R. & Devaux, Y. Long Noncoding RNAs in Patients With Acute Myocardial Infarction. *Circ. Res.* **115**, 668-U180 (2014).

410. Folkersen, L. *et al.* Relationship between CAD Risk Genotype in the Chromosome 9p21 Locus and Gene Expression. Identification of Eight New ANRIL Splice Variants. *PLoS One* **4**, (2009).

411. Wang, F., Su, X., Liu, C., Wu, M. & Li, B. Prognostic Value of Plasma Long Noncoding RNA ANRIL for In-Stent Restenosis. *Med. Sci. Monit.* **23**, 4733–4739 (2017).

412. Inoue, T. *et al.* Vascular Inflammation and Repair Implications for Re-Endothelialization, Restenosis, and Stent Thrombosis. *Jacc-Cardiovascular Interv.* **4**, 1057–1066 (2011).

413. Zhang, B., Wang, D., Ji, T.-F., Shi, L. & Yu, J.-L. Overexpression of lncRNA ANRIL up-regulates VEGF expression and promotes angiogenesis of diabetes mellitus combined with cerebral infarction by activating NF-kappa B signaling pathway in a rat model. *Oncotarget* **8**, 17347–17359 (2017).

414. Cheng, J. *et al.* Variants in ANRIL gene correlated with its expression contribute to myocardial infarction risk. *Oncotarget* **8**, 12607–12619 (2017).

415. Xu, B., Fang, Z., He, S., Wang, J. & Yang, X. ANRIL polymorphism rs4977574 is associated with increased risk of coronary artery disease in Asian A meta-analysis of 12,005 subjects. *Medicine (Baltimore).* **97**, (2018).

416. Aguilo, F., Di Cecilia, S. & Walsh, M. J. Long Non-coding RNA ANRIL and Polycomb in Human Cancers and Cardiovascular Disease. *Long Non-Coding Rnas Hum. Dis.* **394**, 29–39 (2016).

417. Kornienko, A. E. *et al.* Long non-coding RNAs display higher natural expression variation than protein-coding genes in healthy humans. *Genome Biol.* **17**, (2016).

418. Micheletti, R. *et al.* The long noncoding RNA Wisper controls cardiac fibrosis and remodeling. *Sci. Transl. Med.* **9**, (2017).

419. Zangrando, J. *et al.* Identification of candidate long non-coding RNAs in response to myocardial infarction. *BMC Genomics* **15**, (2014).

420. Ballantyne, M. D. *et al.* Smooth Muscle Enriched Long Noncoding RNA (SMILR) Regulates Cell Proliferation. *Circulation* **133**, 2050–2065 (2016).

421. Zampetaki, A., Albrecht, A. & Steinhofel, K. Long Non-coding RNA Structure and Function: Is There a Link? *Front. Physiol.* **9**, (2018).

422. Novikova, I. V, Hennelly, S. P. & Sanbonmatsu, K. Y. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res.* **40**, 5034–5051 (2012).

423. Toshner, M. *et al.* Evidence of dysfunction of endothelial progenitors in pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* **180**, 780–787 (2009).

424. Sorensen, I., Adams, R. H. & Gossler, A. DLL1-mediated Notch activation regulates endothelial identity in mouse fetal arteries. *Blood* **113**, 5680–5688 (2009).

425. Monvoisin, A. *et al.* VE-cadherin-CreERT2 transgenic mouse: a model for inducible recombination in the endothelium. *Dev. Dyn.* **235**, 3413–3422 (2006).

426. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).

427. Fehrenbach, M. L., Cao, G., Williams, J. T., Finklestein, J. M. & Delisser, H. M. Isolation of murine lung endothelial cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **296**, L1096-103 (2009).

428. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

429. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, (2011).

430. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).

431. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

432. Rajkumar, A. P. *et al.* Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics* **16**, 548 (2015).

433. Giaever, I. & Keese, C. R. Micromotion of mammalian cells measured electrically. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7896–7900 (1991).

434. Lo, C. M., Keese, C. R. & Giaever, I. Cell-substrate contact: another factor may influence transepithelial electrical resistance of cell layers cultured on permeable filters. *Exp. Cell Res.* **250**, 576–580 (1999).

435. Necsulea, A. *et al.* The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**, 635–640 (2014).

436. Qiu, F., Zhang, M.-R., Zhou, Z., Pu, J.-X. & Zhao, X.-J. lncRNA MIR503HG functioned as a tumor suppressor and inhibited cell proliferation, metastasis and epithelial-mesenchymal transition in bladder cancer. *J. Cell. Biochem.* **120**, 10821–10829 (2019).

437. Wang, H. *et al.* Long noncoding RNA miR503HG, a prognostic indicator, inhibits tumor metastasis by regulating the HNRNPA2B1/NF-kappaB pathway in hepatocellular carcinoma. *Theranostics* **8**, 2814–2829 (2018).

438. Muys, B. R. *et al.* Placenta-Enriched LincRNAs MIR503HG and LINC00629 Decrease Migration and Invasion Potential of JEG-3 Cell Line. *PLoS One* **11**, e0151560 (2016).

439. Fu, J. *et al.* LncRNA MIR503HG inhibits cell migration and invasion via miR-103/OLFM4 axis in triple negative breast cancer. *J. Cell. Mol. Med.* **23**, 4738–4745 (2019).

440. Chuo, D., Liu, F., Chen, Y. & Yin, M. LncRNA MIR503HG is downregulated in Han Chinese with colorectal cancer and inhibits cell migration and invasion mediated by TGF-beta2. *Gene* **713**, 143960 (2019).

441. Cheng, D. *et al.* The Increased lncRNA MIR503HG in Preeclampsia Modulated Trophoblast Cell Proliferation, Invasion, and Migration via Regulating Matrix Metalloproteinases and NF-kappaB Signaling. *Dis. Markers* **2019**, 4976845 (2019).

442. Huang, P.-S. *et al.* The Long Non-Coding RNA MIR503HG Enhances Proliferation of Human ALK-Negative Anaplastic Large-Cell Lymphoma. *Int. J. Mol. Sci.* **19**, (2018).

443. Stern, A. D., Rahman, A. H. & Birtwistle, M. R. Cell size assays for mass cytometry. *Cytometry. A* **91**, 14–24 (2017).

444. Li, Z. *et al.* MKL1 promotes endothelial-to-mesenchymal transition and liver fibrosis by activating TWIST1 transcription. *Cell Death Dis.* **10**, 899 (2019).

445. Ranchoux, B. *et al.* Endothelial dysfunction in pulmonary arterial hypertension: an evolving landscape (2017 Grover Conference Series). *Pulm. Circ.* **8**, 2045893217752912 (2018).

446. Huang, J. *et al.* Transplantation of Mesenchymal Stem Cells Attenuates Pulmonary Hypertension by Normalizing the Endothelial-to-Mesenchymal Transition. *Am. J. Respir. Cell Mol. Biol.* **62**, 49–60 (2020).

447. Torarinsson, E., Sawera, M., Havgaard, J. H., Fredholm, M. & Gorodkin, J. Thousands of corresponding human and mouse genomic regions unalignable in primary sequence contain common RNA structure. *Genome Res.* **16**, 885–889 (2006).

448. Meister, G. RNA Interference in the Nucleus. *Science (80-.).* **321**, 496 LP – 497 (2008).

449. Castel, S. E. & Martienssen, R. A. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* **14**, 100–112 (2013).

450. Kopp, F. & Mendell, J. T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell* **172**, 393–407 (2018).

451. Charytan, D. M. *et al.* Increased concentration of circulating angiogenesis and nitric oxide inhibitors induces endothelial to mesenchymal transition and myocardial fibrosis in patients with chronic kidney disease. *Int. J. Cardiol.* **176**, 99–109 (2014).

452. Lebrin, F. *et al.* Endoglin promotes endothelial cell proliferation and TGF-beta/ALK1 signal transduction. *EMBO J.* **23**, 4018–4028 (2004).

453. Lebrin, F., Deckers, M., Bertolino, P. & Ten Dijke, P. TGF-beta receptor function in the endothelium. *Cardiovasc. Res.* **65**, 599–608 (2005).

454. Mercado-Pimentel, M. E., Hubbard, A. D. & Runyan, R. B. Endoglin and Alk5 regulate epithelial-mesenchymal transformation during cardiac valve formation. *Dev. Biol.* **304**, 420–432 (2007).

455. Correia, A. C. P., Moonen, J.-R. A. J., Brinker, M. G. L. & Krenning, G. FGF2 inhibits endothelial-mesenchymal transition through microRNA-20a-mediated repression of canonical TGF-beta signaling. *J. Cell Sci.* **129**, 569–579 (2016).

456. Vance, K. W. & Ponting, C. P. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends in Genetics* vol. 30 348–355 (2014).

457. Werner, M. S. & Ruthenburg, A. J. Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. *Cell Rep.* **12**, 1089–1098 (2015).

458. Simon, M. D. Insight into lncRNA biology using hybridization capture analyses. *Biochim. Biophys. Acta-Gene Regul. Mech.* **1859**, 121–127 (2016).

459. McHugh, C. A., Russell, P. & Guttman, M. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biol.* **15**, (2014).

460. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions. *Mol. Cell* **44**, 667–678 (2011).

461. Dai, S. *et al.* HNRNPA2B1 regulates the epithelial-mesenchymal transition in pancreatic cancer cells through the ERK/snail signalling pathway. *Cancer Cell Int.* **17**, 12 (2017).

462. Fishilevich, S. *et al.* GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford).* **2017**, (2017).

463. Shen, X. *et al.* miR-322/-503 cluster is expressed in the earliest cardiac progenitor cells and drives cardiomyocyte specification. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 9551–9556 (2016).

464. Nakajima, Y., Yamagishi, T., Hokari, S. & Nakamura, H. Mechanisms involved in valvuloseptal endocardial cushion formation in early cardiogenesis: roles of transforming growth factor (TGF)-beta and bone morphogenetic protein (BMP). *Anat. Rec.* **258**, 119–127 (2000).

465. Timmerman, L. A. *et al.* Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev.* **18**, 99–115 (2004).

466. Mercado-Pimentel, M. E. & Runyan, R. B. Multiple transforming growth factor-beta isoforms and receptors function during epithelial-mesenchymal cell transformation in the embryonic heart. *Cells. Tissues. Organs* **185**, 146–156 (2007).

467. MacGrogan, D., Luna-Zurita, L. & de la Pompa, J. L. Notch signaling in cardiac valve development and disease. *Birth Defects Res. A. Clin. Mol. Teratol.* **91**, 449–459 (2011).

468. de la Pompa, J. L. & Epstein, J. A. Coordinating tissue interactions: Notch signaling in cardiac development and disease. *Dev. Cell* **22**, 244–254 (2012).

469. Banyard, J. *et al.* Regulation of epithelial plasticity by miR-424 and miR-200 in a new prostate cancer metastasis model. *Sci. Rep.* **3**, 3151 (2013).

470. Drasin, D. J. *et al.* TWIST1-Induced miR-424 Reversibly Drives Mesenchymal Programming while Inhibiting Tumor Initiation. *Cancer Res.* **75**, 1908–1921 (2015).

471. Zhao, Z. *et al.* miR-503-3p promotes epithelial-mesenchymal transition in breast cancer by directly targeting SMAD2 and E-cadherin. *J. Genet. Genomics* **44**, 75–84 (2017).

472. Jiang, S.-P. & Li, Z.-R. MiR-503-5p regulates cell epithelial-to-mesenchymal transition, metastasis and prognosis of hepatocellular carcinoma through inhibiting WEE1. *Eur. Rev. Med. Pharmacol. Sci.* **23**, 2028–2037 (2019).

473. Xiao, X. *et al.* Regulation of myofibroblast differentiation by miR-424 during epithelial-to-mesenchymal transition. *Arch. Biochem. Biophys.* **566**, 49–57 (2015).

474. Cao, S. *et al.* Inhibition of Smurf2 translation by miR-322/503 modulates TGF-beta/Smad2 signaling and intestinal epithelial homeostasis. *Mol. Biol. Cell* **25**, 1234–1243 (2014).

475. Ghosh, G. *et al.* Hypoxia-induced microRNA-424 expression in human endothelial cells regulates HIF-alpha isoforms and promotes angiogenesis. *J. Clin. Invest.* **120**, 4141–4154 (2010).

476. Caporali, A. *et al.* Deregulation of microRNA-503 contributes to diabetes mellitus-induced impairment of endothelial function and reparative angiogenesis after limb ischemia. *Circulation* **123**, 282–291 (2011).

477. Chen, S. *et al.* MiR-424 is over-expressed and attenuates ischemia-reperfusion kidney injury via p53 and death receptor 6 pathway. *Am. J. Transl. Res.* **11**, 1965–1979 (2019).

478. Liu, P. *et al.* MicroRNA-424 protects against focal cerebral ischemia and reperfusion injury in mice by suppressing oxidative stress. *Stroke* **46**, 513–519 (2015).

479. Wen, Y. *et al.* MiR-503 suppresses hypoxia-induced proliferation, migration and angiogenesis of endothelial progenitor cells by targeting Apelin. *Peptides* **105**, 58–65 (2018).

480. Lee, A. *et al.* A PPARgamma-dependent miR-424/503-CD40 axis regulates inflammation mediated angiogenesis. *Sci. Rep.* **7**, 2528 (2017).

481. Chandra, S. M. *et al.* Disruption of the apelin-APJ system worsens hypoxia-induced pulmonary hypertension. *Arterioscler. Thromb. Vasc. Biol.* **31**, 814–820 (2011).

482. Yang, P. *et al.* A novel cyclic biased agonist of the apelin receptor, MM07, is disease modifying in the rat monocrotaline model of pulmonary arterial hypertension. *Br. J. Pharmacol.* **176**, 1206–1221 (2019).

483. Chou, C.-H. *et al.* miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **46**, D296–D302 (2018).

484. Sarkar, S., Dey, B. K. & Dutta, A. MiR-322/424 and -503 are induced during muscle differentiation and promote cell cycle quiescence and differentiation by down-regulation of Cdc25A. *Mol. Biol. Cell* **21**, 2138–2149 (2010).

485. Llobet-Navas, D. *et al.* The miR-424(322)/503 cluster orchestrates remodeling of the epithelium in the involuting mammary gland. *Genes Dev.* **28**, 765–782 (2014).

486. Chen, P.-Y., Qin, L., Tellides, G. & Simons, M. Fibroblast growth factor receptor 1 is a key inhibitor of TGFbeta signaling in the endothelium. *Sci. Signal.* **7**, ra90 (2014).

487. Chen, P.-Y. *et al.* FGF regulates TGF-beta signaling and endothelial-to-mesenchymal transition via control of let-7 miRNA expression. *Cell Rep.* **2**, 1684–1696 (2012).

488. Li, J. *et al.* FGFR1 is critical for the anti-endothelial mesenchymal transition effect of N-acetyl-seryl-aspartyl-lysyl-proline via induction of the MAP4K4 pathway. *Cell Death Dis.* **8**, e2965 (2017).

489. Akatsu, Y. *et al.* Fibroblast growth factor signals regulate transforming growth factor-β-induced endothelial-to-myofibroblast transition of tumor endothelial cells via Elk1. *Mol. Oncol.* **13**, 1706–1724 (2019).

490. Llobet-Navas, D. *et al.* The microRNA 424/503 cluster reduces CDC25A expression during cell cycle arrest imposed by transforming growth factor beta in mammary epithelial cells. *Mol. Cell. Biol.* **34**, 4216–4231 (2014).

491. Zhou, Y. *et al.* miR424-5p functions as an anti-oncogene in cervical cancer cell growth by targeting KDM5B via the Notch signaling pathway. *Life Sci.* **171**, 9–15 (2017).

492. Xu, J. *et al.* Suppressed miR-424 expression via upregulation of target gene Chk1 contributes to the progression of cervical cancer. *Oncogene* **32**, 976–987 (2013).

493. Liu, H. *et al.* miR-503 inhibits cell proliferation and invasion in glioma by targeting L1CAM. *Int. J. Clin. Exp. Med.* **8**, 18441–18447 (2015).

494. Guo, X. *et al.* MicroRNA-503 represses epithelial-mesenchymal transition and inhibits metastasis of osteosarcoma by targeting c-myb. *Tumour Biol.* **37**, 9181–9187 (2016).

495. Chang, S.-W., Yue, J., Wang, B.-C. & Zhang, X.-L. miR-503 inhibits cell proliferation and induces apoptosis in colorectal cancer cells by targeting E2F3. *Int. J. Clin. Exp. Pathol.* **8**, 12853–12860 (2015).

496. Long, J., Ou, C., Xia, H., Zhu, Y. & Liu, D. MiR-503 inhibited cell proliferation of human breast cancer cells by suppressing CCND1 expression. *Tumour Biol.* **36**, 8697–8702 (2015).

497. Guo, J., Liu, X. & Wang, M. miR-503 suppresses tumor cell proliferation and metastasis by directly targeting RNF31 in prostate cancer. *Biochem. Biophys. Res. Commun.* **464**, 1302–1308 (2015).

498. Wu, J. *et al.* miR-503 suppresses the proliferation and metastasis of esophageal squamous cell carcinoma by triggering autophagy via PKA/mTOR signaling. *Int. J. Oncol.* (2018) doi:10.3892/ijo.2018.4320.

499. Baskerville, S. & Bartel, D. P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241–247 (2005).

500. Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **14**, 1902–1910 (2004).

501. Meng, Z., Yu, X., Chen, J., Li, L. & Li, S. Curcumin attenuates cardiac fibrosis in spontaneously hypertensive rats through PPAR-gamma activation. *Acta Pharmacol. Sin.* **35**, 1247–1256 (2014).

502. Jin, Y.-G. *et al.* Puerarin Protects against Cardiac Fibrosis Associated with the Inhibition of TGF-beta1/Smad2-Mediated Endothelial-to-Mesenchymal Transition. *PPAR Res.* **2017**, 2647129 (2017).

503. Lee, S.-W. *et al.* Snail as a potential target molecule in cardiac fibrosis: paracrine action of endothelial cells on fibroblasts through snail and CTGF axis. *Mol. Ther.* **21**, 1767–1777 (2013).

504. Wei, W.-Y. *et al.* Pioglitazone Alleviates Cardiac Fibrosis and Inhibits Endothelial to Mesenchymal Transition Induced by Pressure Overload. *Cell. Physiol. Biochem.* **45**, 26–36 (2018).

505. Wen, J. *et al.* miR-424 coordinates multilayered regulation of cell cycle progression to promote esophageal squamous cell carcinoma cell proliferation. *EBioMedicine* **37**, 110–124 (2018).

506. Kim, K., Eom, J. & Jung, I. Characterization of Structural Variations in the Context of 3D Chromatin Structure. *Mol. Cells* **42**, 512–522 (2019).

507. Groff, A. F. *et al.* In Vivo Characterization of Linc-p21 Reveals Functional cis-

Regulatory DNA Elements. *Cell Rep.* **16**, 2178–2186 (2016).

508. Font-Cunill, B., Arnes, L., Ferrer, J., Sussel, L. & Beucher, A. Long Non-coding RNAs as Local Regulators of Pancreatic Islet Transcription Factor Genes. *Front. Genet.* **9**, 524 (2018).

509. Miura, H., Gurumurthy, C. B., Sato, T., Sato, M. & Ohtsuka, M. CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA. *Sci. Rep.* **5**, 12799 (2015).

510. Yoshimi, K. *et al.* ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat. Commun.* **7**, 10431 (2016).

511. Miura, H., Quadros, R. M., Gurumurthy, C. B. & Ohtsuka, M. Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. *Nat. Protoc.* **13**, 195–215 (2018).

512. Mangeot, P. E. *et al.* Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat. Commun.* **10**, 45 (2019).

513. Yan, W. *et al.* MiR-503 modulates epithelial-mesenchymal transition in silica-induced pulmonary fibrosis by targeting PI3K p85 and is sponged by lncRNA MALAT1. *Sci. Rep.* **7**, 11313 (2017).

514. Caruso, P. *et al.* Identification of MicroRNA-124 as a Major Regulator of Enhanced Endothelial Cell Glycolysis in Pulmonary Arterial Hypertension via PTBP1 (Polypyrimidine Tract Binding Protein) and Pyruvate Kinase M2. *Circulation* **136**, 2451–2467 (2017).

515. Stenmark, K. R., Frid, M. & Perros, F. Endothelial-to-Mesenchymal Transition: An Evolving Paradigm and a Promising Therapeutic Target in PAH. *Circulation* vol. 133 1734–1737 (2016).

516. Kitao, A. *et al.* Endothelial to mesenchymal transition via transforming growth factor-beta1/Smad activation is associated with portal venous stenosis in idiopathic portal hypertension. *Am. J. Pathol.* **175**, 616–626 (2009).

517. Ballarino, M. *et al.* Deficiency in the nuclear long noncoding RNA Charme causes myogenic defects and heart remodeling in mice. *EMBO J.* **37**, (2018).

518. Levitt, N., Briggs, D., Gil, A. & Proudfoot, N. J. Definition of an efficient synthetic poly(A) site. *Genes Dev.* **3**, 1019–1025 (1989).

519. Yonaha, M. & Proudfoot, N. J. Transcriptional termination and coupled polyadenylation in vitro. *EMBO J.* **19**, 3770–3777 (2000).

520. Gromak, N., West, S. & Proudfoot, N. J. Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol. Cell. Biol.* **26**, 3986–3996 (2006).

521. Grillari, J., Hohenwarter, O., Grabherr, R. M. & Katinger, H. Subtractive hybridization of mRNA from early passage and senescent endothelial cells. *Exp. Gerontol.* **35**, 187–197 (2000).

522. Liao, H. *et al.* Effects of long-term serial cell passaging on cell spreading,

migration, and cell-surface ultrastructures of cultured vascular endothelial cells. *Cytotechnology* **66**, 229–238 (2014).

523. Groeneveld, D. J. *et al.* Angiogenic characteristics of blood outgrowth endothelial cells from patients with von Willebrand disease. *J. Thromb. Haemost.* **13**, 1854–1866 (2015).

524. Ormiston, M. L. *et al.* Generation and Culture of Blood Outgrowth Endothelial Cells from Human Peripheral Blood. *J. Vis. Exp.* e53384 (2015) doi:10.3791/53384.

525. Abaci, H. E., Shen, Y.-I., Tan, S. & Gerecht, S. Recapitulating physiological and pathological shear stress and oxygen to model vasculature in health and disease. *Sci. Rep.* **4**, 4951 (2014).

526. Ciuclan, L. *et al.* A novel murine model of severe pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* **184**, 1171–1182 (2011).

527. Vitali, S. H. *et al.* The Sugen 5416/hypoxia mouse model of pulmonary hypertension revisited: long-term follow-up. *Pulm. Circ.* **4**, 619–629 (2014).

528. Manavski, Y. *et al.* Clonal Expansion of Endothelial Cells Contributes to Ischemia-Induced Neovascularization. *Circ. Res.* **122**, 670–677 (2018).

529. Zhang, Y., Jiang, G., Sauler, M. & Lee, P. J. Lung endothelial HO-1 targeting in vivo using lentiviral miRNA regulates apoptosis and autophagy during oxidant injury. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **27**, 4041–4058 (2013).

530. Zhang, Y. *et al.* Endothelial PINK1 mediates the protective effects of NLRP3 deficiency during lethal oxidant injury. *J. Immunol.* **192**, 5296–5304 (2014).

531. Mossbrugger, I., Hoelzlwimmer, G., Calzada-Wack, J. & Quintanilla-Martinez, L. Standardized morphological phenotyping of mouse models of human diseases within the German Mouse Clinic. *Verh. Dtsch. Ges. Pathol.* **91**, 98–103 (2007).

532. Fuchs, H. *et al.* Mouse phenotyping. *Methods* **53**, 120–135 (2011).

533. Meehan, T. F. *et al.* Disease model discovery from 3,328 gene knockouts by The International Mouse Phenotyping Consortium. *Nat. Genet.* **49**, 1231–1238 (2017).

534. Kumar, S., Kang, D.-W., Rezvan, A. & Jo, H. Accelerated atherosclerosis development in C57Bl6 mice by overexpressing AAV-mediated PCSK9 and partial carotid ligation. *Lab. Invest.* **97**, 935–945 (2017).

535. Bjorklund, M. M. *et al.* Induction of atherosclerosis in mice and hamsters without germline genetic engineering. *Circ. Res.* **114**, 1684–1689 (2014).

536. Sun, N. *et al.* Development of drug-inducible CRISPR-Cas9 systems for large-scale functional screening. *BMC Genomics* **20**, 225 (2019).

537. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).

538. Zhang, Y. *et al.* CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. *Sci. Rep.* **5**, 16277 (2015).

539. Chavez, A. *et al.* Comparison of Cas9 activators in multiple species. *Nat Meth* **13**, 563–567 (2016).

540. Li, J.-H., Liu, S., Zhou, H., Qu, L.-H. & Yang, J.-H. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **42**, D92–D97 (2014).

541. Marin-Bejar, O. & Huarte, M. RNA Pulldown Protocol for In Vitro Detection and Identification of RNA-Associated Proteins. *Regul. Non-Coding Rnas Methods Protoc.* **1206**, 87–95 (2015).

542. McDonel, P. & Guttman, M. Approaches for Understanding the Mechanisms of Long Noncoding RNA Regulation of Gene Expression. *Cold Spring Harb. Perspect. Biol.* **11**, (2019).

543. Kanki, Y. *et al.* Epigenetically coordinated GATA2 binding is necessary for endothelium-specific endomucin expression. *EMBO J.* **30**, 2582–2595 (2011).

544. Profumo, V. *et al.* LEADeR role of miR-205 host gene as long noncoding RNA in prostate basal cell differentiation. *Nat. Commun.* **10**, 307 (2019).

545. Sun, Q. *et al.* MIR100 host gene-encoded lncRNAs regulate cell cycle by modulating the interaction between HuR and its target mRNAs. *Nucleic Acids Res.* **46**, 10405–10416 (2018).

546. Abrahimi, P. *et al.* Efficient Gene Disruption in Cultured Primary Human Endothelial Cells by CRISPR/Cas9. *Circulation research* vol. 117 121–128 (2015).

547. Leisegang, M. S. *et al.* Long Noncoding RNA MANTIS Facilitates Endothelial Angiogenic Function. *Circulation* **136**, 65-+ (2017).

548. Murdoch, C. E. *et al.* Endothelial NADPH oxidase-2 promotes interstitial cardiac fibrosis and diastolic dysfunction through proinflammatory effects and endothelial-mesenchymal transition. *J. Am. Coll. Cardiol.* **63**, 2734–2741 (2014).

549. Balachandran, K. *et al.* Cyclic strain induces dual-mode endothelial-mesenchymal transformation of the cardiac valve. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19943–19948 (2011).

550. Hjortnaes, J. *et al.* Valvular interstitial cells suppress calcification of valvular endothelial cells. *Atherosclerosis* **242**, 251–260 (2015).

551. Nicklin, S. A. *et al.* Analysis of cell-specific promoters for viral gene therapy targeted at the vascular endothelium. *Hypertens. (Dallas, Tex. 1979)* **38**, 65–70 (2001).

552. Miller, W. H. *et al.* Targeting endothelial cells with adenovirus expressing nitric oxide synthase prevents elevation of blood pressure in stroke-prone spontaneously hypertensive rats. *Mol. Ther.* **12**, 321–327 (2005).

553. Morecroft, I. *et al.* Gene therapy by targeted adenovirus-mediated knockdown of pulmonary endothelial Tph1 attenuates hypoxia-induced pulmonary hypertension. *Mol. Ther.* **20**, 1516–1528 (2012).

554. Potts, J. D., Dagle, J. M., Walder, J. A., Weeks, D. L. & Runyan, R. B. Epithelial-mesenchymal transformation of embryonic cardiac endothelial cells is inhibited by a modified antisense oligodeoxynucleotide to transforming growth factor beta 3. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1516–1520 (1991).

555. Burks, M., Stickel, S. & Galie, N. Pulmonary Arterial Hypertension: Combination Therapy in Practice. *Am. J. Cardiovasc. Drugs* **18**, 249–257 (2018).

556. Simonneau, G. *et al.* Addition of sildenafil to long-term intravenous epoprostenol therapy in patients with pulmonary arterial hypertension: a randomized trial. *Ann. Intern. Med.* **149**, 521–530 (2008).

557. McLaughlin, V. V *et al.* Addition of inhaled treprostinil to oral therapy for pulmonary arterial hypertension: a randomized controlled clinical trial. *J. Am. Coll. Cardiol.* **55**, 1915–1922 (2010).