ST HUGH'S COLLEGE, UNIVERSITY OF OXFORD

Transcription Regulation of Nrp1 during Endothelial Cell Differentiation

A thesis submitted to the University of Oxford in candidature for the degree of Doctor of Philosophy

Zhe Zhao

Trinity, 2014

Supervisor: Dr Sarah De Val, Ludwig Cancer Research, Nuffield Department of Clinical Medicine, University of Oxford

ZHAO, Zhe
St Hugh's College

Doctor of Philosophy
Trinity, 2014

Transcription Regulation of Nrp1 during Endothelial Cell Differentiation

Abstract

Various diseases, including cancer, stroke and heart attack, are associated with disruption of the vascular system. However, lack of a profound understanding of the transcription regulation during vascular development hinders the formation of effective molecular intervention strategies targeting angiogenesis. Here we describe an enhancer of Neuropilin1 (Nrp1) from the second intron of the gene that directs arterial and coronary endothelial cell-specific expression. Mice transgenic for either human or mouse sequences of the Nrp1in2 enhancers drove expression of the LacZ reporter gene specifically in the endothelial cells within the arterial compartment from early in development, while no expression was detected in veins. In addition, the hNrp1in2 enhancer directed expression to the endothelial cells in the developing coronary vasculature, with the initial expansion from around the sinus venosus at E11.5, and eventually contributed to the capillary, venous and arterial compartments of the coronary vessels but not the endocardium. This expression pattern is consistent with that reported in the Apelin-nlacZ line (Red-Horse et al., 2010), making the Nrp1 enhancer the first identified mammalian regulating enhancer of the coronary endothelial cell. Phylogenetic footprinting, and a tissue culture reporter assay suggested that this enhancer contains a 184bp minimal core region hNrp1in2peakA2 that recapitulates the expression profile of the full length enhancer. hNrp1in2peakA2 has conserved and *in vitro* validated recognition sites for Gata, Ets, and Fox. The validated Fox and Ets sites form a functional FOX:ETS motif, and the FOX:ETS motif is responsible for synergistic activation ofthe enhancer by FoxC2 and Etv2 in reporter assays. Mutation introduction to the functional Ets sites or compound ablation of the Gata and Fox site in hNrp1in2peakA2 result in total loss of vascular expression, in terms of both arterial and coronary expression. The Fox, Ets and Gata recognition sites may be sufficient to achieve arterial- and coronary- specific expression of the hNrp1in2peakA2.

This thesis is submitted to the Department of Medical Science, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. This thesis is entirely my own work, and except where otherwise stated, describes my own research.

Word count: 41,909

Zhe Zhao Doctor of Philosophy

St Hugh's College Trinity, 2014

Acknowledgements

First and foremost, I would like to express my utmost gratitude towards my supervisor Dr Sarah De Val for all her guidance and patience; the intellectual, technical and motivational support throughout my DPhil, and on various occasions carried me through challenging situations.

I am also deeply indebted to Dr Natalia Sacilotto for the help on transgenic zebrafish works, Martin Fritzsche for the tips on image processing, and everyone in the tumour angiogenesis group for not only the useful discussions but also tolerating me for such a long time.

I would also like to thank Dr Ben Davies from Wellcome Trust Human Genetics and Dr Ke Liu from Professor George Bou-Gharios group for the great help on generating transgenic mice, Professor Anna Randi for providing the *Erg* Flox breeder, Dr Lammets van Bueren from Professor Brian Black group for making one of the Hsp68-hNrp1in2 *LacZ* line, and last but not the least, Dr Indrika Ratnayaka and Mark Shipman for the prompt and tireless assistance on histology preparation and visualization.

List of Abbreviations

β-Gal .....β Galactosidase

BLAT .....BLAST-Like Alignment Tool

BMP ..... Bone Morphogenic Protein

DHS..... DNaseI Hypersensitive Site

DLL4..... Delta-Like Ligand 4

DNA..... Deoxyribonucleic acid

EC .....Endothelial cells

EMCN ..... Endomucin

EMSA .....Electrophoresis mobility shift assay

EMT .....Epithelial-mesenchymal transition

H3K27Ac .....Histone H3 Lys 27 acetylation

H3K4Me1 .....Histone H3 lysine4 monomethylation

H3K4Me3 ..... Histone H3 lysine4 trimethylation

Hprt .....Hypoxanthine phosphoribosyl transferase

HSC ..... Haematopoitic Stem Cell

HSP .....Heat shock protein

HUVEC..... human umbilical vein endothelial cell

IB4 ..... Isolection B4

IF ..... Immunofluorescence

IHC ..... Immunohistochemistry

LEC ..... Lymphatic endothelial cells

mESC ..... Mouse embryonic stem cells

mTK ..... Minimal promoter of thymidine kinase gene

NECD ..... Notch extracellular domain

NICD ..... Notch intracellular domain

Nrp ..... Neuropilin

PBS ..... Phosphate Buffered Saline

PCR ..... Polymerase Chain Reaction

PFA ..... Paraformaldehyde

PEO ..... Proepicardial organ

RNA ..... Ribonucleic Acid

ROSA26 ..... ROSAβgeo26

RT ..... Room temperature

SEMA ..... Semaphorin

sv ..... *sinus venosus*

TGF-β ..... Transforming Growth Factor-β

VEGF ..... Vascular endothelial growth factor

WT ..... Wild type

X-Gal ..... 5-Bromo-4-chloro-3-indolyl- β-D-galactopyranoside

Contents

1. Introduction 1

1.1. Establishment and specification of the vasculature 2

1.1.1. Vasculogenesis and angiogenesis in embryonic development 2

1.1.2. Arterial, venous and lymphatic differentiation 8

1.1.3. Origin and specification of the coronary endothelium 11

1.1.4. Major signalling pathways regulating vascular differentiation 15

1.2. Regulation of endothelial cell specification and differentiation 23

1.2.1. Mechanisms of gene regulation 23

1.2.2. Gene enhancers 24

1.2.3. Transcription factors involved in vascular regulation 27

1.3. Neuropilin 1 in the vasculature 39

1.3.1. The Neuropilin family 39

1.3.2. Neuropilin as a co-receptor 42

1.3.3. Role of Nrp1 in the VEGF pathway during EC specification and differentiation. 46

1.4. Project objectives 47

2. Materials and Methods 49

2.1. Mouse work 50

2.1.1. Genotyping 50

2.1.2. Preparation of constructs for transgenic mouse generation 52

2.1.3. Single copy Rosa26-locus targeted transgenic mice 58

2.1.4. Analysis of reporter gene expression in transgenic mice 58

2.2. Zebrafish 62

2.2.1. Tol2 transgenesis 62

2.2.2. Analysis of transgene expression 62

2.3. Tissue culture 64

2.3.1. Cell lines and Maintenance 64

2.3.2. Cell Transfection 64

2.3.3. Analysis of transient transfections in cell lines 65

2.4. Phylogenetic foot-printing and *in silico* analysis 66

2.5. Electrophoretic mobility shift assay 67

2.6. immunohistochemistry and immunofluorescence 77

2.7. Stock solutions 73

3. Identification and validation of the Nrp1in2 enhancer 75

3.1. The hNrp1 enhancer 76

3.1.1. Identification of an enhancer within the Nrp1 locus....76

3.1.2. hNrp1in2 enhancer expression in arteries and veins throughout development .82

3.1.3. hNrp1in2 enhancer expression in the coronary vasculature throughout development....87

3.2. Conclusions....93

4. Single Copy Transgenesis Using The R26/PhiC31 System....98

4.1. Introduction....99

4.1.1. Complications with regular pronuclear microinjection....99

4.1.2. Targeted single copy transgenesis methods....99

4.1.3. The R26/PhiC31 transgenesis system....101

4.2. Results....103

4.2.1. Arterial specific expression of hNrp1in2Rosa....103

4.2.2. Coronary expression....112

4.2.3. Alternative minimal promoter....117

4.3. Conclusion....120

5. Characterising the key elements within the Nrp1 intronic enhancer....125

5.1. Conservation of the Nrp1in2 enhancer....126

5.2. The Mouse Nrp1in2 enhancer drives similar expression patterns to the human version....132

5.3. Identification of a small region within hNrp1in2 that is sufficient to drive arterial and coronary endothelial-specific expression....137

5.3.1. Investigation of the PeakA and Peak B regions of hNrp1in2....137

5.3.2. Identification of the key sequence within hNrp1in2 PeakA 146

5.4. Conclusions 152

6. Identifying the transcription factors that regulate the vascular expression of the hNrp1in2A2 enhancer 156

6.1. Introduction 157

6.2. *In silico* analysis of the core hNrp1in2peaA2 enhancer 157

6.3. Validation of putative binding sites using electrophoretic mobility shift assay (EMSA) 158

6.3.1. FoxC2 bind Nrp1 Fox2 site but not the Fox1 site 160

6.3.2. Ets sites Ets 1, Ets5, Ets7 and Ets8 binds recombinant ETS proteins in vitro .162

6.3.3. FoxC2 binds the Nrp1 enhancer 170

6.3.4. FoxC2 and Etv2 synergistically activate hNrp1in2 172

6.3.5. GATA4 can bind the Nrp1 Gata motifs 174

6.4. Mutational analysis 176

6.4.1. Depletion of all the Ets site prevent endothelial expression 178

6.4.2. Functional Ets site mutation deplete intra-embryonic vascular expression 180

6.4.3. Role of Erg on hNrp1in2peakA enhancer expression 182

6.5. Fox and Gata sites compound mutation of hNrp1in2peakA2 switch off vascular expression 184

6.6. Conclusion 186

6.6.1. Functional Fox:Ets motif in the Nrp1 enhancer. 186

6.6.2. ETS factor 188

6.6.3. Coronary endothelial cell expression ablated in Ets, and Fox-Gata mutation . 192

7. Discussion 193

7.1. Summary of results 194

7.2. Vascular Nrp1 expression under Ets factor transcription regulation 195

7.3. Arterial specific expression regulation 196

7.3.1. Sox and Notch 197

7.3.2. Ets factor in EC compartmentalization 198

7.4. Nrp1 in coronary endothelial cell differentiation 199

7.4.1. Multiple cellular origins of the coronary endothelial cells 199

7.4.2. Transcription regulation of coronary endothelial cell differentiation 201

List of Figures

Figure 1.1 Embryonic development of vascular growth. 7

Figure 1.2 Schematic diagram describing the two alternative hypotheses of coronary endothelial cell origin. 14

Figure 1.3 Transcription regulation of vascular endothelial cell differentiation. 38

Figure 1.4 Neuropilin1 protein structure 41

Figure 1.5 NRPs in the VEGF pathway for endothelial cell differentiation 44

Figure 3.1 Alignment of the human and mouse Nrp1in2 sequences at the FOX:ETS site 77

Figure 3.2 Schematic representation of the Nrp1 locus. 79

Figure 3.3 Schematic representation of the Nrp1in2 enhancer. 81

Figure 3.4 hNrp1in2*LacZ* transgenic mouse reporter gene expression in the arterial compartment of the vasculature. 84

Figure 3.5 Arterial-specific expression of hNrp1in2*LacZ* is consistent with endogenous NRP1 86

Figure 3.6 Coronary expression of hNrp1in2*LacZ* in embryonic and neonatal stages. 89

Figure 3.7 Immunofluorescence of PECAM, NRP1 and β-Gal on hNrp1in2*LacZ* heart coronal sections. 92

Figure 4.1 R26/PhiC31 Transgenesis Diagram. 102

Figure 4.2 Arterial-specific expression of the Nrp1in2 enhancer at E9.5. 105
Figure 4.3 Arterial-specific expression of Nrp1in2 enhancers at E11.5.....108
Figure 4.4 Arterial-specific expression of Nrp1in2 enhancers at E13.5.....110
Figure 4.5 Coronary expression of the Nrp1in2 enhancers. 114
Figure 4.6 Postnatal background expression of the Rosa line. 116
Figure 4.7 Altenative promoter miniTK has a strong background expression post-nataly. ..119
Figure 5.1 Conservation of the human Nrp1 enhancer in its orthologous sequences. 127
Figure 5.2 Schematic representation of the mouse Nrp1 locus 129
Figure 5.3 Schematic representation of the mNrp1in2 fragment.....130
Figure 5.4 mNrp1in2*LacZ* transgenic embryos express in arterial and coronary vessels in the vasculature. 134
Figure 5.5 Expression profile validation of mNrp1 in a single copy line.....136
Figure 5.6 hNrp1in2peakB*LacZ* lacks peripheral vascular expression 138
Figure 5.7 hNrp1in2peakA*LacZ* embryonic expression in the vasculature.....139
Figure 5.8 Expression profile validation of Nrp1in2peakA in a single copy transgenic line 141
Figure 5.9 hNrp1in2peakA *LacZ* line direct arterial specific expression in early embryonic development.....142
Figure 5.10 hNrp1in2peakA *GFP* transgenic Zebrafish vascular expression 145
Figure 5.11 PeakA2 truncation of the hNrp1in2peakA alignment 147
Figure 5.12 hNrp1in2peaA2 *LacZ* vascular and arterial expression at E11.5 148
Figure 5.13 hNrp1in2PeakA2 *LacZ* arterial and coronary expression at E13 149
Figure 5.14 peakB is not enriched with DHS in endothelial cell line HUVEC. 154
Figure 6.1 hNrp1in2peakA2 annotation of putative recognition sites. 159
Figure 6.2 Nrp1 enhancer Fox site binds recombinant FoxC2 protein. 161

Figure 6.3 ETS1-DBD binds some but not all conserved Ets sites in hNrp1in2peakA enhancer. 164
Figure 6.4 ETS1-DBD can directly bind Ets sites within the hNrp1in2peakA2 enhancer....165
Figure 6.5 ETV2 can bind conserved Ets sites in the hNrp1in2peakA2 enhancer. 166
Figure 6.6 ETV2 can directly bind some but not all *Nrp1* Ets sites 167
Figure 6.7 ERG binds conserved Ets sites in Nrp1in2peakA2. 169
Figure 6.8 Fox2:Ets7 composite site can bind both FOXC2 and ETV2.....171
Figure 6.9 Synergistic activation of the Nrp1 enhancers by FOXC2 and ETV2.....173
Figure 6.10 GATA4 can bind the conserved Gata site in the Nrp1in2peakA2 enhancer. 175
Figure 6.11 Design of mutational analysis of hNrp1in2peakA2. 177
Figure 6.12 ALL Ets mutation prevent vascular expression of the hNrp1in2peakA2 enhancer 179
Figure 6.13 pA2 FUNC Ets mut does not express in intra-embryonic vasculature.....181
Figure 6.14 Mutation of both Fox and Gata sites in the hNrp1in2peakA2 core enhancer resulted in loss of vascular expression.....185

List of Tables

Table 2.1 Primers sets for Genotyping 51

Table 2.2 Constructs Used for Generation of Transgenic Mice 53

Table 2.3 Oligos for cloning the mNrp1in2 and hNrp1in2peakA2 fragments 56

Table 2.4 Synthetic DNA strings 57

Table 2.5 Constructs for transient transfection 65

Table 2.6 EMSA oligos.....69

Table 2.7 Table of Antibodies used in IHC or IF 72

Table 2.8 Recipe of stock solutions 73

Table 5.1 Summary of reporter gene expression detected in transient transgenic mice.....151

Table 6.1 $Erg^{fl/fl}$ -Cdh5(PAC)iCreERT2-hNrp1in2peakA$^{Rosa/+}$ breading scheme 190

1. Introduction

1.1. Establishment and specification of the vasculature

The mammalian vascular system is responsible for distributing nutrients and oxygen to the whole body, as well as maintaining homeostasis. It is composed of blood and lymphatic vessel systems, both of which are specified into a collection of distinct compartments.

1.1.1. Vasculogenesis and angiogenesis in embryonic development

Vasculogenesis

Blood vessel formation during mammalian embryonic development occurs through a series of interconnected steps that will be divided here into two steps: vasculogenesis and angiogenesis.

The first step in embryonic vessel development is the *de novo* formation of the primitive blood vessels from their precursors. This process is termed vasculogenesis, and is considered to be the earliest stage in vascular development (summarised by Ferguson *et al*. (*1*)). At the end of gastrulation, a subpopulation of mesodermal progenitor cells colonise the primitive yolk sac, forming the blood islands that mark the initiation of vasculogenesis. These cells aggregate and proliferate to form a primitive plexus. Sometimes termed haemangioblasts, these cells possess the capacity to further differentiate (*2-4*). The outer cells become angioblast cells, and proliferate and differentiate into endothelial cells that line the vascular structure of the plexus. Further morphogenesis subsequently takes place, and the vessels start to tubulise to give rise to a volume inside the plexus. The inner cells take up the surrounding volume of the newly tubulised plexus and form the first wave of erythrocyte progenitors. These primitive erythrocytes are then joined by a second wave of haematopoietic stem cells

(HSCs). These definitive HSCs have been shown to be derived from phenotypical endothelial cells of the dorsal aorta wall, termed haemogenic endothelium (*5*), and are responsible for the life-long production of all lineages of blood cells. This common ancestry of angioblasts and HSCs is a phenomenon also shared with other non-mammalian models. In chickens, xenograft transplantation studies have established the primitive streak as the common origin of angioblasts and HSCs (*6*). The yolk sac precursors primarily form the angioblasts and the primitive erythrocytes (*7, 8*). The splanchnopleural mesoderm and paraxial mesoderm progenitors provide the cellular make-up of the dorsal aorta, and the definitive HSCs bud out from the ventral floor of the newly formed aorta (*9*). In zebrafish, lineage tracing techniques have helped to identify the haemangioblasts in the ventral mesoderm (*10*). Nevertheless, it is of note that haemangioblasts only contribute to a fraction of the endothelial and haematopoietic lineages in mouse, zebrafish and chicken embryos and that independent progenitors in the mesoderm give rise to the vasculature and haematopoiesis (Figure 1.1).

During adulthood, a neo-vascularisation process that is induced in response to ischemia has been suggested to share the same mechanism as vasculogenesis, and has been associated with "circulating endothelial progenitor cells" that are thought to be derived from the bone marrow (*11*). However, the cellular origin and function of such cells remains under debate. The circulating endothelial progenitor cells have been found to be macrophages presenting endothelial antigens following the endocytosis of endothelial cells, and may have been involved in focal vascular growth due to an alternative mechanism (*12*) (also summarised by Timmermans et al.(*13*)).

Angiogenesis

After the first vessels are formed in vasculogenesis, new vessels form from existing ones. This process, called angiogenesis, is the mechanism that accounts for the formation of the majority of blood vessels. Additionally, during angiogenic vessel growth, the surrounding mesenchymal cells differentiate into smooth muscle cells and other support cells. These are recruited to the vessel tubes to form the outer layer of the vessels, tightly enclosing and mechanically enhancing the tube. The extracellular matrix surrounding the vessels also undergoes remodelling. The endothelial lining is eventually surrounded by a basement membrane that is composed of collagens and microfibrils. Some of the vessels are further surrounded by layers of connective tissue and smooth muscle cells (*14*, *15*). As the vasculature matures and the vessels compartmentalise into arteries, veins and capillaries, the molecular and cellular make-up of these surrounding layers becomes distinct among these compartments (*16*).

There are two major mechanisms for angiogenesis. The first one occurs through sprouting and elongation. Certain endothelial cells of a vessel detect chemoattractants from avascular tissues, and in response they develop filopodia and migrate away from the vascular base (*17*). These cells are termed tip cells. Behind tip cells are more proliferative neighbour cells that elongate as migration takes place. These cells are termed stalk cells. The identities of the tip cells and the stalk cells are interchangeable (*18*). In fact, they are under constant competition for the tip cell position, and delicate molecular signalling pathways are involved in modulating the process to ensure that only one tip cell thrives in each sprouting stalk (*19*). As

multiple sprouts reach the destination region, only the first sprout matures into the major blood supply, the other sprouts ceasing to grow when the chemoattractant gradient collapses. Some of the vessels are retained, tubulised, allow inflow of blood and form a microvascular network that perfuse the region, whereas most others undergo pruning and withdrawal from the final network (*20*). The molecular signalling mechanism that governs these processes will be detailed later in session 1.1.4, especially the "Wnt pathway" paragraphs.

The second angiogenesis mechanism is intussusception. Intussusception is a process during which the existing vessels remodel and split into two vessels. This process was first discovered in chicken chorioallantoic membrane vascularisation, since the systemic perfusion hierarchy demonstrates a clear series of step by step bifurcation events at each branching point (*21*). During intussusception, the endothelial cells make inward protrusions and gain contact with the endothelial cell on the opposite side of the tube. This endothelial contact across the vascular lumen is then further formalised by remodelling of the intercellular junction. The junction then receives the invasion of smooth muscle cells and pericytes, and remodelling of the extracellular matrix (*22*).Further morphological changes expand to neighbouring regions and the vessel tube wall is further refined until a new vessel is formed. However, intussusception is not restricted to this model; it has also been implicated in the augmentation of capillary bed density in mammalian models, as well as tissue repair and lymphatic system development (*22, 23*) (Figure 1.1).

Regardless of the mechanism that vessels go through for angiogenesis, a series of events are experienced by the endothelial cells, basement membrane and the surrounding tissue to

facilitate the growth. The blood vessel permeability is inevitably compromised when under mobilization. The endothelial cells also secrete molecular signals to loosen pericyte and smooth muscle cell contacts. Proteinases and matrix metalloproteinases are also secreted to digest the basement membrane and extracellular matrix. As new vessels fuse and take form, it is necessary to terminate endothelial cell migration and reform the basement membrane before the focal blood flow is initiated. While the new vessel matures, some vessels are pruned and regress away.

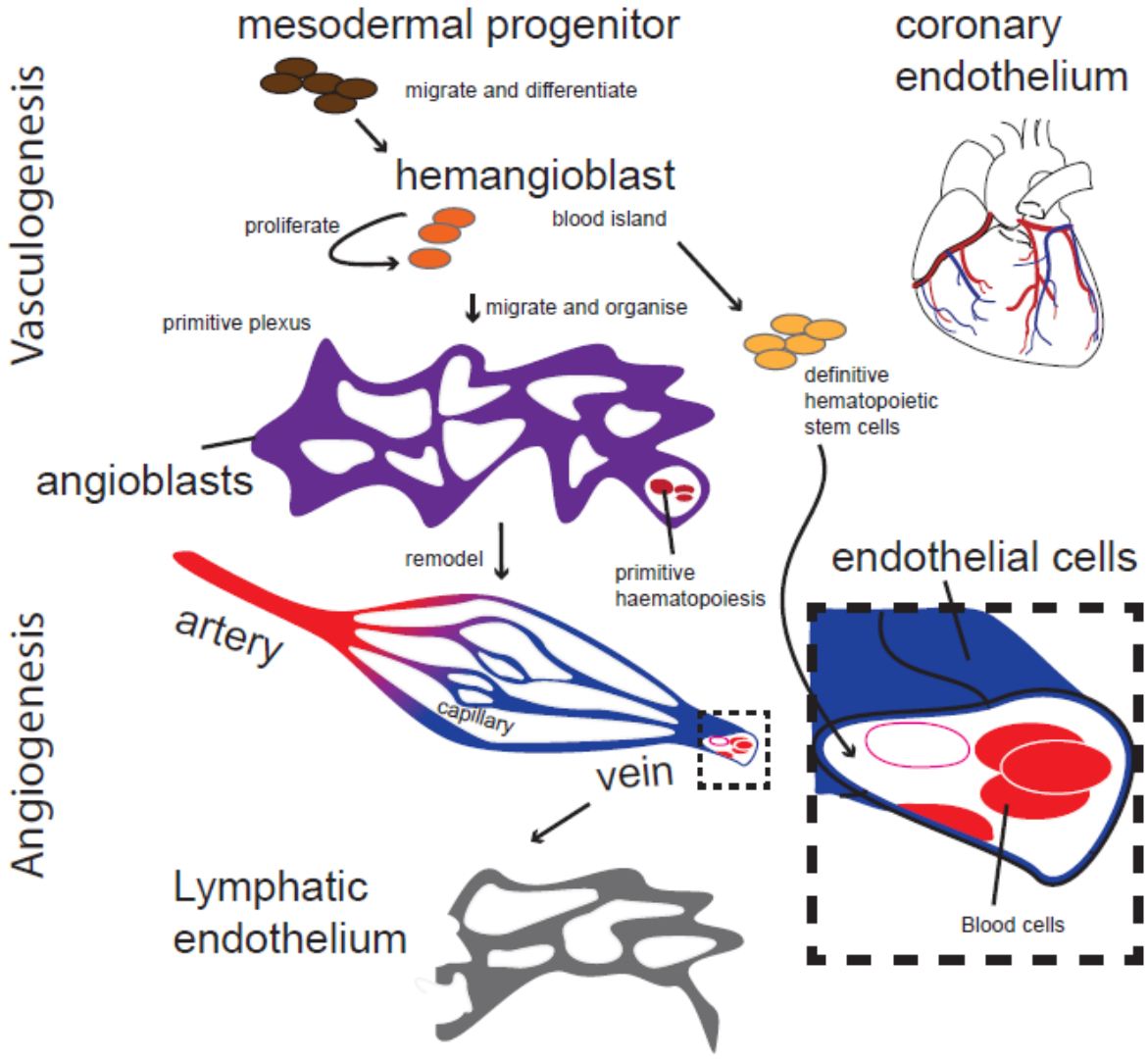


Figure 1.1 Embryonic development of vascular growth.

The differentiation hierarchy of vascular endothelial cells. As illustrated, the mesodermal progenitor cells form the whole mature vasculature in roughly two phases, the vasculogenesis and the angiogenesis. The angioblasts for a primitive vascular plexus base *de novo* during vasculogenesis, from where the angiogenesis process further the vascular growth. Two waves of haematopoiesis make up the cellular origin of highly differentiated blood cells. Arterial venous differentiation takes place during vascular maturation, and a subset of venous EC transdifferentiate to form the lymphatics. Coronary EC may be of multiple origins, and the differentiation route map is under debate. For detailed explanation, please follow main text.

1.1.2. Arterial, venous and lymphatic differentiation

Specification of the first arteries, veins and lymphatic vessels

The dorsal aorta and the cardinal vein are the first artery and vein formed during embryonic development, and they serve as a comprehensible model with which to study arterial-venous differentiation. The endothelial lining of these vessels is the key to the specification process. In mice, the first intra-embryonic vessel is the dorsal aorta, and it is formed by the amalgamation of dorsal aortae that are differentiated from the angioblasts formed in the primitive plexus during vasculogenesis (*24*). The cardinal vein is formed later than the dorsal aorta: its cellular origin and the exact stage at which its venous identity is established remains under debate. The aggregating angioblast-formed plexus has been identified as the region where the cardinal vein later forms, although they do not exhibit the molecular and morphological characteristics of venous endothelial cells at this stage. There has also been a report showing that some dorsal aorta endothelial cells demonstrate gene expression profiles with venous characteristics, and migrate laterally to the cardinal vein, suggesting that a fraction of the cardinal vein endothelium may be derived from the dorsal aorta (*25*).

In zebrafish, angioblasts originally are formed within the lateral plate mesoderm, before migrating to the midline in the trunk to form the dorsal aorta and posterior cardinal vein (*26, 27*). It has been suggested that this represents a fused progenitor vessel from which these first two proper vessels take shape via selective cell sorting (*28*). However, conflicting data suggests that alternative waves of angioblast migration may be responsible for these two vessels (*27*, *29*).

Differences in the morphological and expression profiles of arteries and veins become definite only at the end of angiogenesis, after the remodelling processes have taken place. It was therefore thought that these occurred in response to differential haemodynamic conditions, such as the shear stress of blood flow (*30*) (reviewed by Lawson and Weinstern (*31*)). However, we now know that early markers of A-V specification can be detected during vasculogenesis, at the time when the endothelial progenitors are still being formed and the blood flow has not even started (*32*). The mechanisms controlling the establishment and maintenance of early arteries and veins must therefore be, at least in part, determined by genetic programming. A number of these mechanisms are discussed in later sections.

Lymphatic specification

Occurring alongside angiogenesis, a subpopulation of cardinal vein endothelial cells (ECs) deviate from the remodelling process and start to express lymphatic markers such as the transcription factor Prox1 (*33*, *34*). These ECs bud off from the vein to form primitive lymph sacs, which later sprout and migrate into the periphery to mature into the entire lymphatic system. The survival and initial migration of these lymphatic endothelial cells (LECs) is dependent on external chemotaxis clues (summarised by Adams and Alitalo (*16*)). At the same time, the intrinsic molecular profile also contributes to their differentiation. For example, the venous master regulator CoupTFII may play a pivotal role in the transcriptional regulation of LEC-specific markers including Prox1 (*33*), as well as regulating VEGF signalling (*35*).

Blood vessel morphogenesis

Mature blood vessels recruit pericytes and smooth muscle cell, and remodel the extracellular matrix to re-form the basement membrane. The capillaries tend to have only a loose (and even scarce) lining of smooth muscle cell–like pericytes, whereas arteries and veins have more complex arrangements of the surrounding layers. The veins are surrounded by a layer of smooth muscle cells and an outer layer of adventitia. The adventitia mostly consists of connective tissues, and is also embedded with nerves. Arteries also have the smooth muscle cell and connective tissue layers, although the smooth muscle cell layer is much thicker than in corresponding veins, and is further sandwiched by two layers of elastic lamina. These structures enable the arteries to be stronger, more elastic and have greater contraction capacity, reflecting the fact that the arteries need to endure much higher blood pressure and a higher velocity of blood flow. The veins are specialised to prevent reverse blood flow, particularly where blood drainage can be challenged by muscle movement or the direction of flow is against gravity. The venous endothelial cells therefore have a different structure, in that they are equipped with one-way valves. Moreover, the expression profiles of these endothelial cells differ from each other. These differences are the result of a series of hierarchical differentiation events that are tightly regulated by various forces.

Arteriogenesis

Another phenomenon in vascular formation is the formation of mature arteries from pre-existing arterioles, particularly in response to arterial occlusion. The arteriole's endothelial cells proliferate to enlarge its lumen, and undergo extracellular remodelling to mature into artery collateral to the occluded one. This mechanism, termed arteriogenesis, was

previously regarded to be a special case of angiogenesis, although it has recently been appreciated that arteriogenesis can fully replace an occluded artery while angiogenesis cannot. Moreover, angiogenesis is closely correlated with hypoxic and ischemic scenarios, whereas arteriogenesis occurs in normoxia (*36*). Arteriogenesis is particularly important during a time in development when the volume of the capillary plexus needs to be quickly augmented, to accommodate the rapid growth of the embryo.

1.1.3. Origin and specification of the coronary endothelium

The specification of the coronary vascular system begins within a time-window similar to that of arterial, venous and lymphatic specification. However, unlike these other endothelial subtypes, the position of coronary ECs in the endothelial differentiation hierarchy is still under debate.

The coronary vasculature is the part of the cardiovascular system that is responsible for the perfusion of the heart. In spite of intensive investigation of the coronary vasculature in a pathological context, the position of coronary endothelial cells in the vascular differentiation hierarchy is unclear. Experiments in avian models have long suggested an epithelial-mesenchymal transition (EMT) hypothesis for coronary endothelial cell formation (Figure 1.2 A), in which the proepicardial organ (PEO) attaches to the looping heart and forms an epithelial sheath called the epicardium that covers the heart, from where certain populations go through an EMT process, migrate into the myocardium and colonise it. Chimera explant studies in avian models have repeatedly shown the contribution of these cells to the coronary vasculature, including ECs (*37*, *38*). This would imply that the origin of

coronary ECs is entirely different from that of the rest of the vasculature, although they share molecular and morphological similarities. However, recent lineage tracing studies in mice, while not excluding the possibility of an epicardial EMT process, have suggested that epicardium-derived cells contribute only to cardiomyocytes, cardiac fibroblasts and smooth muscle cells, and not to coronary endothelial cells (*39*).

A recent lineage-tracing study in mice has now suggested an entirely different origin, indicating the sinus venosus as the source of coronary ECs (Figure 1.2 B). Red-Horse and colleagues used an Aplin knock-in mouse model that specifically marks the endothelial layer of the coronary (*40*). By *in vitro* transplantation and clonal analysis of coronary development, they described a venous lineage of cells in the sinus venosus that expands into the developing heart and migrates throughout the organ to form a primitive vascular network. This network further matures into the coronary ECs. This study therefore suggests that the coronary ECs, both arteries and veins, differentiate from venous-committed ECs. Interestingly, some blood-island-like colonies that are initially discontinuous from the sinus venosus expansion were also seen in the lower part of the ventricles. These cells later form a plexus that connects into the coronary vasculature, although their origin is under debate. In a recent report, one possible explanation has been raised. The endocardial layer of the looped heart can respond to VEGF pathway signals, and migrate into the myocardial layer of the developing heart in a tissue culture model (*41*). It has, therefore, been suggested that the cells make a major contribution to the cellular make-up of the coronary artery, and that coronary veins mostly originate from the sinus. Early establishment of these migrant cells might have been the blood island described by Red-Horse (*40*). Nonetheless, the vasculogenic potential of the

endocardium *in vivo* during embryonic development has again been challenged, since a recent lineage tracing study showed that endocardial ECs undergo transdifferentiation and migration to contribute to the coronary endothelial plexus after birth (*42*). The endocardium-derived vasculature accommodates perfusion of the inner layer of the myocardium and interventricular groove exclusively (*42*), leaving the rest of the heart irrigated by sinus venosus-derived vessels (*43*). How exactly the sinus and endocardium descendants divide and conquer the coronary EC is to be further elucidated. On the other hand, it is still possible that the PEO origin hypothesis is correct. No study to date has conclusively ruled out the possibility of a cellular contribution of PEO to the endothelial linage of these blood islands. Nevertheless, it is the consensus that the cellular make-up of the coronary vessels (coronary endothelial cells in particular) has multiple developmental origins (Figure 1.2).

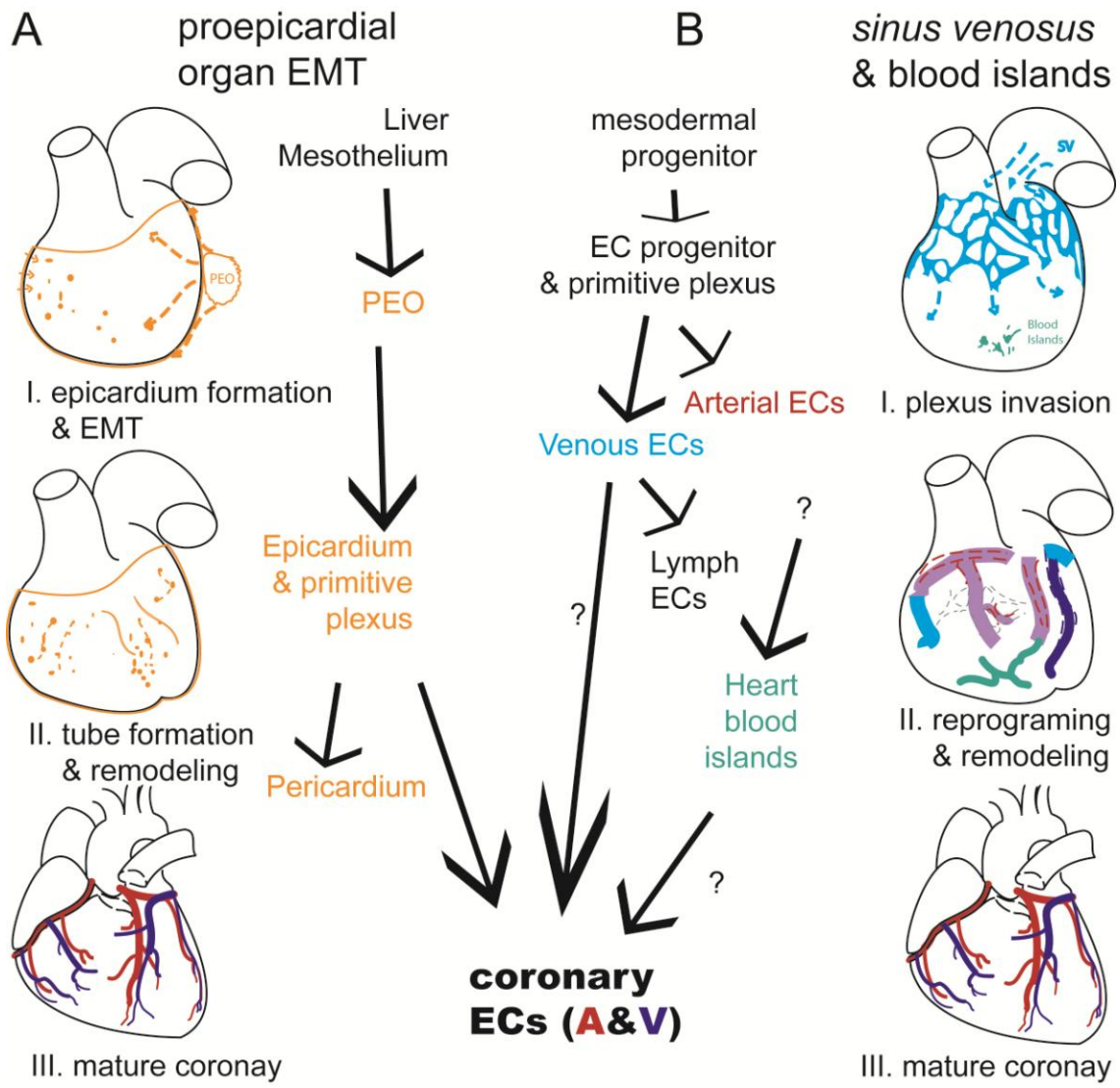


Figure 1.2 Schematic diagram describing the two alternative hypotheses of coronary endothelial cell origin.

A. The PEO EMT hypothesis suggests that coronary ECs originate from the proepicardial organ (PEO, orange).

B. The sinus venosus hypothesis argues that venous-committed cells expand from the sinus venosus (sv), to form an EC plexus in the myocardium that is subsequently reprogrammed into coronary arteries and veins (light blue). In this hypothesis, a minor population of the coronary ECs also arises from the inter-ventricular groove near the primordial septum (green, blood islands). These cells are of unknown origin, but the plexus they form connects to the plexus and matures into the coronary vasculature.

1.1.4. Major signalling pathways regulating vascular differentiation

Sonic hedgehog (Shh) pathway

Shh plays a morphogenic role in embryonic development and is the ligand for the Shh pathway, which is very important in organogenesis. As a cellular signalling pathway, Shh is involved in various processes during embryonic development including limb bud formation, central nervous system development and skin appendages. In vascular development, Shh has been implicated in various stages (*44*). During vasculogenesis, shh is expressed and secreted by notochord and floor plate to induce dorsal aorta formation in fish (*26*). Depletion of the signal by inhibitor treatment impairs dorsal aorta formation, and gain-of-function experiments show that enhanced *shh* expression converts the posterior cardinal vein to an arterial fate, indicating that shh is required for the establishment of arterial identity (*26*). The pathway is also involved in angiogenesis in various organs, especially through collaboration with other pathways. Shh-null mutant mice have defective pulmonary airway development in the lungs due to impaired angiogenesis (*45*). In the heart, Shh is expressed by the epicardium and released to the sub-epicardial space, where it induces coronary plexus growth (*46*). These functions of Shh may be due to its impact on endothelial cell migration and proliferation, which has also been shown on capillary ECs and putative EPCs and during wound healing (*47*, *48*, *49*).

VEGF pathway

The Vascular Endothelial Growth Factor (VEGF) pathway plays a central role in multiple aspects of vascular development. The impact of the pathway spans vasculogenesis and

angiogenesis, and includes roles in proliferation, differentiation and patterning of the vasculature. As a major chemoattractant for the avascular region to recruit angiogenic sprouts, VEGF is important in tip cell activation and migration. Moreover, the pathway also plays an important role in arterial-venous differentiation.

In zebrafish, VEGF has been shown to be downstream to Shh but upstream of Notch during arterial differentiation (*26*). In the absence of VEGF signalling, endothelial cells form a single fused vessel instead of the clearly differentiated dorsal aorta and cardinal vein. Expression of arterial markers is present, but expression of venous markers is not (*28*).

In mice, the arterial venous and lymphatic consequences of VEGF perturbation are confounded by its early vascular phenotype, rendering it hard to interpret VEGF's involvement. Over-expression of VEGF-A in cardiomyocytes can up-regulate the arterial marker Efnb2 in mice (*50*). However, the apparent primary receptors for VEGF-A, Flk1 and Flt1, are ubiquitously expressed in endothelial precursors of both arterial and venous fates, making it challenging to understand the mechanism through which VEGF-A affects the differentiation process. Nonetheless, there is evidence that other VEGF pathway components are differentially expressed. For example, ligand VEGF-C and D, receptor Flt4, as well as co-receptor Neuropilin2 (Nrp2) are primarily venous specific, whereas Nrp1 is arterial (*51*). It is consequently hypothesised that intracellular signalling may be triggered differently in future arteries compared to future veins, potentially leading to A-V specification upon differential downstream regulation of Dll4 and Notch receptor expression (*52-54*). Further

details of the VEGF pathway's involvement in vascular development via Nrp co-receptors will be discussed later in this introduction.

The VEGF signalling pathway is also implicated in lymphatic development. The survival and migration of LECs is dependent on VEGF signalling, which is transduced by VEGF-C and VEGFR3 (*16*, *55*).

Notch pathway

Loss of VEGF signalling in fish can be rescued by the activation of Notch signalling, indicating that Notch lies downstream of VEGF in a common genetic pathway (*26*). The canonical Notch pathway is well conserved among multi-cellular organisms, and serves as a juxtacrine cellular signalling mechanism, in which both the ligand-donating cell and the receptor-residing cell are simultaneously affected (*56*). In mammals, the Notch pathway ligands include members of the Delta-like (DLL1, DLL2, DLL4) and Jagged (JAG1, JAG2) protein families. The ligands can be recognised by the single-pass transmembrane receptors NOTCH1, 2, 3 and 4, which induce cleavage of the extracellular (NECD) and intracellular domain (NICD) of the receptor. The ligand-NECD complex remains intact and undergoes endocytosis into the ligand-donating cell. Meanwhile, in the receptor-residing cell, NICD is released from the membrane and translocates to the nucleus. Here, NICD binds to the transcription factor RBP-Jκ and its co-activator Mastermind/MAML to form an activation complex. Although not all Notch targets are clearly understood, the HES and Hey families of transcription factors are well-known examples of vascular direct targets of Notch (*57*).

As a juxtacrine cellular signal, the lateral inhibition mechanism is the major way by which the Notch pathway exhibits its function. Lateral inhibition is widely employed in organogenesis during embryonic development, cellular differentiation and body planning, ensuring that once a cell takes up a certain trait or cell fate, a similar event should be prevented in neighbouring cells to avoid redundancy. In particular, DLL4 is present at angiogenic sprouts, where lateral inhibition is thought to prevent the formation of too many tip cells, and consequently prevent excessive angiogenic sprouting (*17, 58*). Tip cells have higher levels of Dll4 compared with stalk cells, which have higher levels of Notch signalling. The high tip cell Dll4 levels, modulated by high expression of the VEGF receptors Flk1 and Nrp1, signal through the Notch1 receptor in the neighbouring stalk cells to inhibit tip cell formation, therefore maintaining the integrity of the sprout. Conversely, the Notch ligand Jag1 is preferentially expressed in stalk cells and is a positive regulator of tip cell formation (*17, 59*).

While perturbation of Dll4 in the sprout leads to excess, incorrect sprout formation, Dll4-null mice clearly demonstrate disruption of the arterial network, marked by expansion of EphB4 and loss of EphrinB2 expressions (*60*). In fact, genetic disruption of any of the Notch pathway opponents, including *Notch, Dll4, Rbpj,* or even the target genes *Hey1/2,*all result in vascular developmental defects, particularly arteriovenous malfunctions (*60-63*). These results suggest that, in addition to regulating sprouting, the Notch pathway plays a central role in suppressing venous identity and promoting arterial differentiation.

In terms of arterial venous differentiation, the Notch pathway may be wired in a different way. A lot of Notch pathway components are specifically expressed in arteries, including both receptors and ligands Notch1, Notch4, Jag1, Jag2, and Dll4; although others are also expressed in vessels (*58, 60*). In fish, loss of Notch signalling disrupts arteriovenous differentiation by up-regulating the expression of venous endothelial markers in the dorsal aorta, whereas ectopic activation of the Notch pathway suppresses vein development (*26*). It has been suggested that, in the scenario of arteriovenous differentiation, recognition of the DLL ligand or activation of the VEGF pathway achieve Notch pathway activation and, in turn, maintain such activation among its neighbouring arterial ECs in a feed-forward loop (*64, 65*).

EphrinB2 and EphB4

The mutually exclusive expression of transmembrane ligand EphrinB2 (Efnb2) and its receptor tyrosine kinase receptor EphB4, found in arterial and venous ECs respectively, is established before blood flow and provides key evidence that the arterial and venous commitment of ECs has already been established during vasculogenesis in both mouse and zebrafish models. In mice, both EfnB2 and EphB4-null models are embryonic lethal. Although the mutants undergo relatively normal vasculogenesis, they have defects in the remodelling of both arteries and veins (*32, 66, 67*). In zebrafish, alterations in the levels of Efnb2 and the Ephb4 orthologous also affect the correct formation of the first artery and vein (*28*).

Ephrins and their receptors are both membrane bound proteins, which means that intercellular contact is essential for ligand-receptor recognition and their function in a juxtacrine signalling manner. Ephrin signalling is primarily involved in axon guidance, chemo-attraction, and chemo-repulsion processes. They are expressed widely across all tissues during embryonic development, although tissue specific expression may be observed for particular isoforms (*68*). Ephrin B2 (EFNB2) has the highest affinity to its receptor EphB4 (*69*), and phosphorylation can be induced upon ligand-receptor recognition on the tyrosine residues in each other's cytoplasmic domain. Their direct downstream targets have not been fully elucidated, although reports indicate the involvement of the G-protein-coupled receptor CXCR4 and the AKT phosphorylation cascade pathway (*70*). Besides tyrosine kinase activation, the Ephrin/receptor complex may also induce *trans*-endocytosis in either a forward or a reverse direction, that in neuronal cells is known to activate the sorting and segregation of the ligand donor and receptor host cells by enacting cell-cell repulsion (reviewed by Pitulescu et al. (*71*)). In endothelial cells, EFNB2 expression is activated by VEGF and Notch pathways (*72, 73*). Interestingly, some of the VEGF receptors, namely VEGFR2 and VEGFR3, are dependent on the presence of EfrnB2 for their endocytosis. Impaired VEGFR internalisation and recycling caused by EfrnB2 ablation leads to receptor desensitisation, which prevents sustained VEGF signal transduction (*74, 75*). The exact mechanism by which EFNB2 intervenes in VEGFR endocytosis is not known yet, although this interaction may render EFNB2/EphB4 relevant to the process of cardinal vein segregation from the dorsal aorta, where the VEGF pathway dictates through its specifically paired ligand-receptor binding (*28*).

Wnt pathway

The Wnt pathway is involved in a wide range of embryonic development processes and carcinogenesis, because of its role in cell differentiation, migration and growth. The canonical Wnt pathway features the translocation of β-catenin from the cell membrane to the nucleus for the transcriptional activation of downstream target genes, upon binding of Wnt ligand to its heterodimeric transmembrane receptor complex. Depletion of β-catenin is embryonic lethal to mice during gastrulation (*76, 77*). β-catenin is expressed in endothelial cells during development, and endothelial-specific knockout of β-catenin results in vascular malformations in the central nervous system due to impaired EC migration (*78*). In fact, similar loss of function analyses of Wnt pathway components have establish Wnt's involvement in various stages of vascular growth (*79*). Of note, consistent Wnt activation leads to impairment of arteriovenous specification, in that the veins acquire strong Notch activation and gain a characteristic arterial expression profile (*80*). Cross-talk between the Wnt and Notch signalling pathways has been implicated widely, and it is not surprising that the combinatorial regulation of NICD, RBPJ-κ and β-catenin in arterial endothelial cells, through the formation of a transcription regulation complex, has been found to activate a list of arterial-specific genes (*81*). Thus, the Wnt pathway's involvement in arterial differentiation is mediated through Notch signalling.

TGF-β pathway

The TGF-β super family of ligands primarily includes TGF-β (Transforming Growth Factor-β) and the BMPs (Bone Morphogenic Proteins), as well as the GDFs, AMH, Activin, Nodal and

others. The TGF-β pathway signal is conducted by ligand binding and recruitment of the serine/threonine receptor kinase complex, inducing auto-phosphorylation of the receptors and activation of its cytoplasmic amplifier SMADs, causing their translocation to the nucleus to act as transcription factors (*82*). Deactivation of both the BMP and TGF-β pathways by EC-specific knockout of SMAD4 in mice results in enhanced numbers of haematopoietic cluster cells and haemogenic endothelial differentiation, shifting the endothelial cell fate to a haematopoietic progenitor one, and BMP is required for HSC formation in fish (*83*). Alk1 is an EC-specific TGF-β receptor, and Endoglin, which is also highly expressed in proliferating ECs, can serve as a co-receptor of the pathway for fine-tuning the signalling (*84, 85*). TGF-β is largely considered to be anti-angiogenic due to its ability to suppress endothelial cell proliferation (*86*), and it can induce the trans-differentiation of the endothelial cells resulting in permeable blood vessels and extracellular matrix remodelling (*87*). On the other hand, conditional knockout of BMP receptor Alk3 from mesodermal haemangioblast progenitors results in defective endothelial and smooth muscle cell formation (*88*). The BMPs (BMP2, BMP4 and BMP6) have been implicated as angiogenic cues for venous vessels (*89*) (reviewed by Asahara and Wiley (*11*)). The preferential expression of transmembrane receptors and some of the SMADs accounts for the differential response from arterial and venous ECs (*90, 91*).

1.2. Regulation of endothelial cell specification and differentiation

1.2.1. Mechanisms of gene regulation

The production, localization, activation and maintenance of a protein are regulated at various levels. The production of a protein, or the expression of a coding gene, begins with the synthesis of an RNA copy, its processing and export, followed by protein translation and folding. The RNA copy synthesis process, termed transcription, is where the RNA polymerase binds to the DNA template strand and creates a complementary RNA strand.

In eukaryotes, RNA polymerase binding is dependent on the assistance of a set of transcription factors, known as basal transcription factors. These recognise specific sequence motifs, bind to them and recruit the RNA polymerase to form a transcription complex to initiate transcription (*92*). Many basal transcription factors possess functions that are essential for the initiation of transcription, such as helicase activity to untangle the two strands of the DNA double helix. However, in eukaryotes, these basal transcription factors are not sufficient to stimulate the transcription of genes alone. They require additional factors, binding both nearby and at a distance. These transcription factors can modulate the rate of transcription though a variety of mechanisms including the recruitment of co-activators and co-repressors, catalysing histone modifications, remodelling chromatin structure, and stabilizing the formation of the RNA polymerase complex by anchoring on the DNA sequence and binding to the complex. The presence, activation and recruitment of transcription factors to the transcription complex is often in response to certain temporal and spatial cellular situations, and therefore this process is an important point at which protein expression is regulated (*93*).

During embryonic development, transcription factor control of gene expression is vital for differentiation and the specification of cellular fate, and the transcription regulation signalling network plays a big role in this hierarchical and genetically programmed process.

1.2.2. Gene enhancers

Enhancers

Eukaryotic gene transcription is typically regulated by two types of DNA regulatory elements: promoters and enhancers. Promoters are located near the transcription start site, and the core promoters are of binding affinity to polymerase and basal transcription factors. Enhancers are the distal regulatory elements that are capable of modulating gene transcription probability by recruiting tissue-specific transcription factors to accomplish spatial-temporal gene expression patterns during development, although regulatory elements of similar structure and function are referred to as proximal promoters if they are located nearby a core promoter. Enhancers can be located in the intron of a gene, and may also be distant from its coding region (*94*). A gene can have various enhancers, and the collaboration among them can be critical for the transcription regulation of the gene.

Many functional enhancers, particularly those directing specific patterns of gene regulation during key developmental processes, have emerged early in evolution. Consequently, the sequences of homologous enhancers are often conserved between species, although their location relative to the gen can be un-conserved. However, that is not to say that the majority of the conserved elements have clear enhancer activity, or that all enhancers share conservation between species. In fact, some recent studies have suggested that non-conserved

enhancers, although harder to identify, are more common than conserved ones (*95- 99*). This may be because the non-conserved region may exert its regulatory function in a way specifically important of the environmental circumstances of the particular species. In addition, gene enhancers often share seemingly redundant function in the so-called "shadow enhancer phenomenon" (*100*). Nevertheless, high levels of sequence conservation through evolution is still an indicator of the possible involvement of the non-coding region in the regulation of gene expression, and phylogenetic foot printing analysis has shown its power in identifying regulatory elements (*98, 101*).

Epigenetic markers

In addition to the direct binding of transcription factors to DNA motifs, modifications such as DNA methylation and chromosomal modifications also regulate gene expression. This could be through the creation of a physical blockage, or a specific acidic/alkaline or hydrophilic/hydrophobic microenvironment that interferes with RNA polymerase-DNA interaction. The conformational change incurred may result in condensed DNA wrapping, DNA loop formation or naked DNA exposure that alters the binding accessibility and affinity of the transcription factors to the regulatory elements (*102*). In particular, these modifications are epigenetic rather than genetic. This means, on the one hand, that some of the information is inheritable between generations of mitosis, leaving cells of the same direct differentiation lineage sharing the same pattern of epigenetic modification. This provides cells of common origin with a coherent transcription regulation cue, and also makes it very convenient for researchers to look at the epigenetic profile in search of common features of transcription regulation. On the other hand, epigenetic information is subject to modification in the

presence of kinases and phosphatases, histone deacetylase, DNA or histone methyl transferase or other epigenetic modifiers. Therefore, some chromosomal modifications can also serve as real-time markers of the regulatory involvement of enhancers. Among all of the histone epigenetic modifications studied to date, histone H3 lysine4 monomethylation (H3K4Me1) or trimethylation (H3K4Me3) are known to be able to recruit chromatin remodelers and cause conformational changes to the DNA double helix around the nucleosome, so as to make the DNA sequence accessible to transcription factor recognition (*103*, *104*). Both modifications have been associated with transcriptional activation (*105*). H3K4Me1 has been shown to mark all regulatory elements (both promoters and enhancers), whereas H3K4Me3 is only associated with promoter regions (*106*, *107*). Moreover, histone acetylation is also found to facilitate transcription activation by lowering the energy expense of chromosomal remodelling, hence increasing the efficiency of transcription initiation (*108*). H3K27Ac (histone H3 Lys 27 acetylation) in particular is associated with active enhancer function, and can differentiate active functioning enhancers from those of poised status (*106*, *109*).

The accessibility of enhancers is essential for motif recognition and the binding of transcription factors. When an enhancer is active, the chromatin structure exposes the DNA and dissociates, allowing transcription factor complexes to impose conformational changes on the naked DNA string. As a result, these enhancers can cooperate with core promoters and other gene enhancers, both elsewhere in the coding region and further away. Since naked DNA is subject to DNaseI recognition and digestion, DNaseI digestion patterns can identify open chromatin that may represent coding regions or transcription regulatory elements.

Consequently, active enhancers are not only likely to be marked by H3K4Me1 and H3K27Ac, but also by DNaseI hypersensitivity (*110*).

Another way to identify regulatory elements is through the recognition of specific sequence motifs. Genes that play collaborative roles in a particular cell type or process are usually regulated by a handful of the same transcription master regulators. Therefore, the enhancers of these genes tend to share similar recognition motif signatures. For instance, using reported and putative transcription factor binding specificities, Narlikar and colleagues built a classifier that can identify cardiac-specific enhancers at a validated accuracy of 62% (*111*).

1.2.3. Transcription factors involved in vascular regulation

1.2.3.1. Transcription factors regulating early vascular specification and differentiation

Ets factors

Members of the Ets transcription factor family, which share a highly conserved Ets DNA-binding domain, have long been known to be crucially important for endothelial gene expression. In the vasculature, Ets binding motifs are found within all known vascular gene enhancers, and dictate every stage of endothelial cell differentiation (summarised by De Val, 2011 (*112*)). Among the Ets factors, Ets1, Elf1, Fli1, Erg and Tel are all shown to direct the activation of endothelial gene expression. Fli1 is expressed early on in haemangioblast cells in both zebrafish and mice (*113*), and is thought to be upstream of Tal1 and Gata2 gene activation, at least in zebrafish and xenopus (*114*). Erg expression begins during mesodermal angioblast formation, and remains highly expressed in endothelial cells throughout adulthood

(*115*). Erg expression is important for angioblast proliferation, and its continued expression is important for the activation of endothelial specific genes including angiopoietin 2, endoglin, Vwf and VE-cadherin (*116*). Erg ablation in mammals does not cause significant vascular defects during vasculogenesis or early angiogenesis, as the embryo and the vasculature remain intact until its disruption of the cardiac cushion and cardiac valve formation cause lethality to the embryo (*117*).

Although Ets factors are highly expressed in the vasculature, they are not vascular-specific. Ets expression can be found widely during mesodermal and epidermal differentiation, for example, in haematopoiesis, bone formation and neuronal maturation. *Ets* motifs are found in numerous other regulatory elements, therefore are not by themselves particularly predictive of vascular-specific expression patterns. The consensus *Ets* recognition site is the simple4bp GGA(A/T) (*118*); this appears to be shared among all Ets factors. This shared recognition site provides the biochemical basis of functional redundancy among Ets factors, and may explain why deletion of a single Ets factor tends not to result in morphological defects in the vasculature. For example, gene disruption of Fli1 does not result in significant vasculature defects in either zebrafish or mouse models, likely due to the fact that its closely related factor Erg continues to be expressed and plays its role instead (*114*, *119-121*). Similarly, mice null for Ets1 or Ets2 alone have few vascular defects, while the compound Ets1:Ets2-null mouse dies at mid-gestation with a severely malformed vascular network (*122*).

However, the Ets factor Etv2 (also known as ER71 and Etsrp) is an exception. Etv2 is critical for endothelial specification, as Etv2-null mice do not have any endothelial cells (*123*). Etv2

expresses from the earliest stages of vasculogenesis, as early as the mesodermal haemangioblast stem cell (*123*). Ablation of *Etv2*completely abolishes the expression of Fli1, CD31 and other early vascular markers (*124*). Later on, *Etv2* expresses less and less as the vasculature develops. In mice, Etv2 eventually ceases to be expressed at around E10.5, indicating that its role in vascular development may be restricted to the early phase (*124, 125*). Consistent with this notion, several early endothelial genes have been found to be under direct transcriptional regulation of Etv2: Flk1, Tal1, CD31 and Tie2 all have enhancers with crucial binding motifs recognised by Etv2 (*124, 126*). The role of Etv2 in early vascular development is also preserved in zebrafish. The *Etv2* orthologous gene *etsrp* is expressed in blood islands and the vascular endothelial cells (*120, 127*). Gain-of-function experiments have found that increased Etsrp expression augments proliferation of the haemangioblast, and that the endothelial cell pool expands accordingly. Ablation of the gene, on the other hand, disrupts the vascular morphogenesis process. Consequently, it is clear that Etv2 plays a central role in the transcription regulation network of endothelial cell differentiation; however, its short-lived expression profile restricts the direct impact of the gene to haemangioblast differentiation and the early phase of vascular development.

Forkhead factors

Various members of the Forkhead (Fox) family transcription factors are expressed in endothelial cells or their progenitors, and may play a critical role in vascular development. FoxO1 knockout mice die at mid-gestation, and display vascular remodelling defects (*128, 129*). The direct downstream target of FoxO1 during endothelial cell differentiation is not fully known. However, FoxO1 is most commonly regarded as a transcriptional repressor, and

there is evidence that in endothelial cells, FoxO1 regulation of the gene can be either positive or negative, suggesting that it may play the role of a transcriptional switch in endothelial cells (*130, 131*). FoxH1 over-expression results in impaired vascular development and a role for FoxH1 in down-regulating *flk1* expression through the suppression of transcription has been shown in zebrafish (*132*). FoxF1 knockout mice also suffer from defective mesodermal differentiation and yolk sac vascular patterning, even though endothelial cells in the mature vasculature do not express the gene (*133*). However, it is known to be expressed in mesodermal haemangioblast cells, indicating a role of FoxF1 in early vascular development (*134*). The FoxC sub-family is also essential for vascular development. FoxC1 and FoxC2 compound knockout mice suffer from severely impaired vascular development (*135*). Even though endothelial cell specification is not completely absent in the double knockout, arterial and lymphatic differentiation is impaired, and specific markers are not expressed. FoxC1 and FoxC2 have therefore been implicated in arterial and lymphatic EC differentiation, primarily through the Notch pathway (*136*).

Mef2 factors

Mef2 proteins, members of the MADS family of transcription factors, are expressed in both the endothelial and smooth muscle cells of the developing vascular system. The MADS box transcription factor MEF2C is expressed in developing endothelial cells and smooth muscle cells, but is also involved in cardiogenesis, myogenesis, neurogenesis and in the development of the cortical architecture.Mef2c is the first Mef2 factor expressed in endothelial cells and, although it may be required for vascular development, it is dispensable for vasculogenesis (*112*). In mice, Mef2C is expressed in yolk sac blood islands, and its expression is maintained

in the mature vasculature (*137*). *Mef2c* ablation in mice results in E9.5 lethality and vascular defects are evident. The angioblasts are still specified in the absence of *Mef2c*, but fail to amalgamate to form a vascular plexus (*137*). Mef2C is thought to be downstream of the VEGF pathway. The endocardium of *Mef2c* knockout mice is defective, which coincides with the fact that myocardial VEGF expression is down-regulated (*138*). In the retina, VEGF stimulation can activate the transcription of Mef2c, and this activation is dependent on the MAPK and RhoA/ROCK cascades. It is thought that VEGF pathway targets are activated in a Mef2C-dependent fashion (*139, 140*). It is therefore suggested that Mef2C is involved in VEGF-induced endothelial migration.

Gata factors

Gata2 is highly expressed in endothelial cells (*141*), and Gata2 is responsible for the activation of various endothelial specific genes (*125*). However, Gata2 is also important for the early development of haematopoietic stem cells, and may even be involved in the establishment of haemangioblast cell identify from differentiation of the mesodermal progenitor (*114, 142*). Since ECs and haematopoietic stem cells share a common origin, it is unsurprising that the transcriptional regulation of the early development of the two lineages is intertwined. Both Gata2 and Tal1, another key haemangioblast transcription factor, have been jointly implicated in the activation of a number of transcription factors also important in later endothelial specification (*142*). Other Gata factors are also expressed in the vasculature, but relatively little is known about them (*112*). For example, Gata4 is also expressed in the endothelial cells of the heart and some related vasculature (*143*).

Klf factors

Our knowledge of the transcription regulation mechanism that directs the specification of ECs into arterial, venous or lymphatic ECs is more fragmented. A major aspect of EC differentiation is haemodynamic regulation. The onset of blood flow is followed by a change in blood pressure and an increase of shear force, both important in terms of shaping the blood vessel remodelling that is vital for the eventual blood distribution. The haemodynamic response of the endothelial cells is also controlled, at least in part, through transcriptional regulation (*144*). In the arteries, blood pressure and blood flow shear are much higher, and the need for volume augmentation, stronger mechanical support and higher elasticity is sensed by the endothelial cell lining, internalised and passed on to recruited support cells remodelling the extracellular matrix. Expression of the transcription factor Kruppel-like factor 2 (Klf2) is activated during embryonic development when blood flow emerges and increases, and is in turn thought to be responsible for the activation of angiogenic genes in ECs (*30*, *145*). Klf2 knockout mice are embryonically lethal with pandemic periphery haemorrhage, although no arteriovenous malfunction is detected, suggesting that while Klf2 may play a role in early endothelial proliferation, it is not required for arterial-venous differentiation (*30*). In addition to Klf2, Klf4 may also be involved in response to shear force. Klk4 is involved in remodelling of the vessels in response to shifting haemodynamic cues, yet plays a protective role for endothelial cells, preventing them from transdifferentiating into mesodermal and fibrotic cells (*146*).

COUP-TFII

As introduced earlier, commitment of ECs to an arterial or venous fate is established before the angioblasts form clear separate structures, which is well ahead of plexus tubulisation and the onset of blood flow. Therefore, as the role of Klf2 and Klf4 transcription regulation important as part of the mechanism of endothelial differentiation, early arterial and venous cell fate has mostly been predetermined before the Klf proteins exert their impact on EC maturation. Nuclear receptor Coup-TFII (also known as Nr2f2), which acts as a transcription factor, is venous and lymphatic-specific and vital for venous identity. Coup-TFII expression can be traced in the angioblast cells, and its expression is maintained until the maturation of the vasculature (*147*). EC-specific knockout of *Coup-TFII* results in venous vessels expressing a list of arterial-specific marker genes. Coup-TFII has been found to suppress Notch and Nrp1 expression, although neither correlation has been established as direct transcription regulation (*148*). In fact, no direct transcription target of the Coup-TFII has yet been discovered in veins. In the lymphatic vessels, where Coup-TFII is also expressed, there is evidence that Coup-TFII regulates Prox1 and Nrp2 via their promoters during lymphatic endothelial cell differentiation (*33, 35*), although the proposed promoter has not been tested *in vivo*, and no functional enhancer has been identified.

SoxF factors

The SoxF superfamily transcription factors Sox7, Sox17 and Sox18 have also been implicated in vascular growth. In zebrafish, sox7 and sox18 are both expressed in mesodermal progenitors and the midline axial vessels. Compound disruption of the two genes' expression results in defective arteriovenous differentiation (*149-151*). In mice, Sox17 is

involved in early haemangioblast differentiation, and also demonstrates arterial specific expression in ECs. EC-specific knockout of Sox17 results in defective arterial endothelial cell differentiation and vascular remodelling, as well as embryonic lethality (*152*). The Notch pathway has been found to transcriptionally regulate Sox factor expression in the vasculature to impact on angiogenesis (*153*, *154*). The position of Sox factors in the signalling network and their collaboration with other transcription factors is introduced in the next section.

Hey factors

Hey2 is a direct target of the Notch pathway via Rbpj (*149-151*). Gridlock (Grl), a zebrafish orthologue of the Hey proteins, has been studied extensively to illustrate the Notch pathway's involvement in arterial specification. The *grl* mutant is characterised by arterial-venous shunts and dorsal aorta defects, with augmented expression of venous markers and loss of expression of arterial markers (*155*, *156*). *Hey1;Hey2* double knockout mice are embryonic lethal, marked with severe vascular defects and loss of expression of arterial markers (*61*, *62*).

1.2.3.2. Combinatorial regulation of endothelial transcription

As described previously, the transcription factors involved in the transcriptional regulation of endothelial differentiation are not exclusively expressed in endothelial cells. It has been suggested that combinatory regulation of multiple transcription factors is needed to achieve endothelial specificity. This hypothesis is based on the fact that most endothelial-specific enhancers identified to date contain recognition motifs for specific combinations of endothelial transcription factors. A prominent example is the collaboration of Gata factors

and Ets factors. Several endothelial and haematopoietic enhancers contain multiple Ets and Gata recognition sites, and the synergistic effect of Ets and Gata activation has been noted (*142*, *157*). Indeed, protein-protein binding of Fli1 and Elf1 to Gata factors has also been implicated in megakaryocytes (*158*). However, overlapping expression of Ets and Gata factors is not restricted to ECs. A genome-wide scan of common Ets/Gata bound loci was not particularly enriched in haemangioblast-expressing genes, suggesting that extra factors are needed for endothelial specification. Gata and Tal1 have been suggested to collaboratively regulate endothelial-specific genes through the formation of a multi factorial transcription regulation complex with Lmo2 (*159*). Lmo2, Gata2, and Tal1 are co-expressed in haemangioblasts in zebrafish (*114*), and the three collectively activate *Ve-cadherin*in mammals (*160*). Interestingly, the *Ve-cadherin* promoter also harbours Ets recognition sites, although the exact mechanism that combines the transcriptional regulation of Gata, Tal1 and Ets factors is yet to be elucidated (*161*).

Endothelial enhancers tend to contain clusters of Ets sites, leading to the possibility that a combination of Ets family factors may collaborate in vascular-specific regulation. In fish, simultaneous knock down of four Ets factors results in a more severe vascular defect than that of singular knockdowns (*120*). Such a phenomenon can be explained by multiple Ets factors needing to bind the enhancer at the same time, or that the binding of each individual factor fluctuates during vascular development. However, it is also possible that these factors are simply redundant. The exact mechanism of Ets factor combinatory regulation is, therefore, yet to be fully understood.

Similar to Ets, no Fox proteins are expressed in an endothelial-specific cell manner. Work by De Val *et. al* identified a cognate FOX:ETS binding motif that is common to most, though not all, vascular enhancers (*126*). This Fox:Ets site is composed of neighbouring Fox and Ets recognition sites, and can be robustly bound by Ets factors in combination with FoxC proteins. Disruption of the composite site in the endothelial-specific *Mef2C* enhancer prevents its endothelial specificity. FOX:ETS composite sites are found in various endothelial enhancers, and harbouring of a FOX:ETS site is powerful enough to predict endothelial enhancer regions within conserved genomic sequences (*126*). In particular, FoxC2 and Etv2 can robustly bind the composite site, and synergistic activation of FOX:ETS-containing enhancers has been seen in tissue culture and in *Xenopus*. It would, therefore, be interesting to determine whether such a combinatory regulation mechanism is a common feature of endothelial specification, whether the FoxC proteins are replaceable, and how binders of the composite site collaborate with other factors in specific compartmentalisation or maturation events.

Sox factors are genetically closely related to TCF factors, and can be regulated by the canonical Wnt pathway via β-catenin recruitment in various developmental and pathological contexts (*162*). In endothelial cells, Sox17 is known to be essential for Wnt ability to induce angiogenesis *in vitro*, although protein-protein binding of the transcription factors is not fully established in this scenario (*163*). The SoxF factors may have redundant or collaborative roles among each other, as compound heterozygosity of Sox17 and Sox18 leads to defective heart and early vascular development (*164*). In our laboratory, Sacilotto *et al*. have found that Dll4 arterial expression is under combinatorial regulation by SoxF factors and Notch pathway

transcription factor Rbpj, and that compound knockdown of *sox7*, *sox18* and *rbpj* results in dll4 expression ablation and an absence of dorsal aorta formation (*65*). During sprouting angiogenesis, Notch activation down-regulates Sox17 expression, and sox17 promotes tip cell formation and migration by activating tip cell-specific gene transcription (*153*).

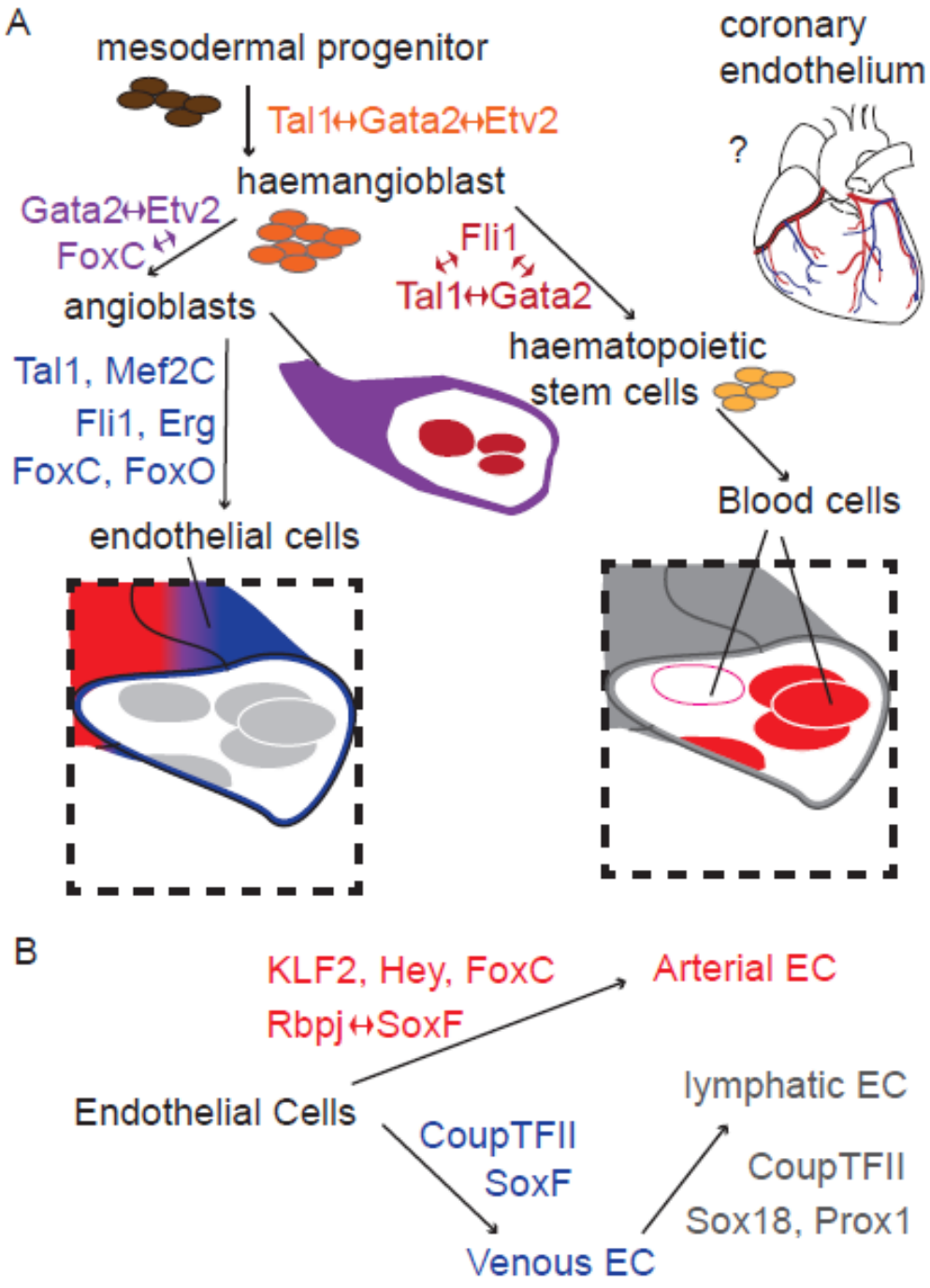


Figure1.3 Transcription regulation of vascular endothelial cell differentiation.

A. Early endothelial cell differentiation. Combinatory regulation partnership among FoxC, Ets and Gata factors are important for angioblast and haematopoietic progenitor formation.

B. Endothelial cell compartmentalization. The SoxF sub-family factors and the Rbpj mediated Notch signal determines by large the arterial commitment and the CoupTFII serve as a venous master regulator. Runx1 and Prox1 make central role in the lymphatic differentiation.

1.3. Neuropilin1 in the vasculature

1.3.1. The Neuropilin family

The Neuropilin (Nrp) family consists of Nrp1 and Nrp2, both of which are transmembrane glycoproteins. Neuropilins have elaborate extracelluar regions, including two CUB domains (complement C1r/C1s, Uegf, Bmp1), two Factor V/VIII homology (b1/b2) domains and a MAM domain (Meprin/A5-protein/PTPmu) (Figure 1.4, adapted from Pellet-Many, 2008 (*165*). The CUB domains mediate protein-protein binding in the complement system (*166*), and for NRPs these domains are responsible for binding axonal guidance polypeptide semaphorins. The b1/b2 domains are comprised of two tandem regions, homologous to blood coagulation factor V/VIII C-terminal domains, which are important for binding to membrane phospholipids on the surface of ECs. The b1/b2 domains are essential for semaphorin binding. Additionally, in NRP1, these domains are required for optimal binding of the VEGF ligands. The exact function of MAM domains remains undetermined, but they have been associated with NRP1 oligomerisation (*167*).

The NRPs have a single transmembrane domain and a rather smaller cytoplasmic one. The cytoplasmic domains were thought to have no functionality other than NIP1 (Nrp-interacting protein 1) recruitment, although recent studies suggest that the Nrp1 cytoplasmic PDZ domain (Psd95/Dlg1/Zo-1) binding motif SEA is required for ERK and VEGF-A dependent arteriogenesis signalling (*168*, *169*).

Neuropilin genes have been found in all vertebrates examined so far, but not in non-vertebrate organisms. Mouse (mammal, *Musmusculus*), chicken (avian, *Gallus gallus*)

and the zebrafish (bony fish, *Daniorerio*) all have Nrp genes, and their amino acid sequences are highly conserved.

Human *NRP1* is located on chromosome 10p12, spans 120kb and is composed of 17 exons. Besides the full length protein, there is one isoform that is membrane-associated and at least another two soluble isoforms. The soluble isoforms vary in size, but both lack the MAM domain, the transmembrane domain and the cytoplasmic region. Human Nrp2 has two major isoforms, NRP2a and NRP2b. The two isoforms do not differ in their extracellular domains. However, NRP2b lacks the SEA motif, and its transmembrane and cytoplasmic regions are only 11% homologous to NRP2a, a phenomenon that coincides with the differential tissue specific expression of the two isoforms (*170*). Functional differences between the two have, therefore, been suggested but not yet fully explored. Another soluble NRP2 isoform is s_9NRP2, known to result from alternative splicing (*170*).

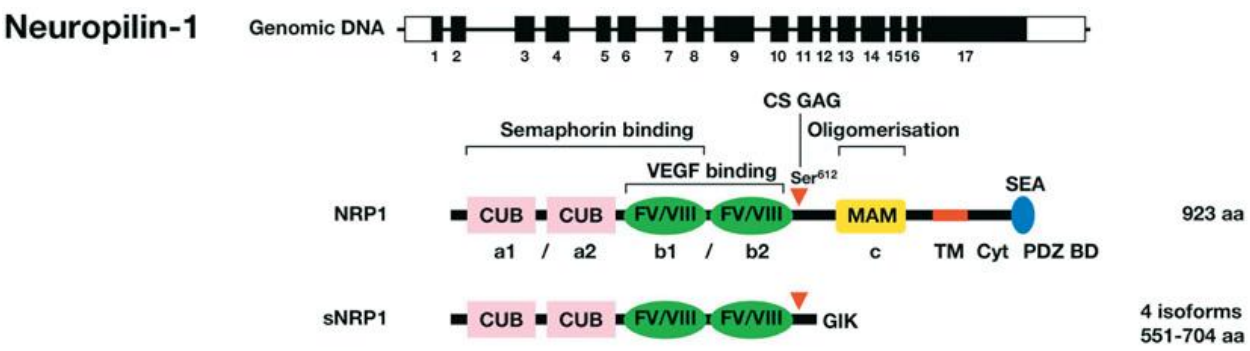


Figure 1.4 Neuropilin1 protein structure

The human NRP1 gene encodes a full-length protein of 923 amino acids (aa) consisting of an extracellular region with two CUB (a1 and a2) domains, two tandem Factor V/VIII homology (b1 and b2) domains, a linker region, a MAM (c) domain, a single transmembrane domain (TM), and a 44-amino-acid cytoplasmic (Cyt) domain containing a C-terminal PDZ-binding domain motif, with the sequence, SEA. The diagram is adapted from Pellet-Many et al., 2008 (*165*).

1.3.2. Neuropilin as a co-receptor

As briefly mentioned in the previous section, the neuropilin receptors are capable of binding two structurally unrelated types of ligands; the class 3 semaphorin (SEMA3) and the VEGF family ligands.

VEGF is well-known for its key role in vasculogenesis and angiogenesis during both development and adulthood (see earlier in Introduction). The VEGF family consists of VEGF-A, -B, -C, -D, and PlGF (Placental growth factor), and each member has various isoforms that differ in their affinity for NRP receptors. The Factor V/VIII (b1/b2) extracellular domains of NRPs are essential for VEGF binding, and VEGF ligand variants binding rely on b1/b2 recognition specificity. Of note, the most abundant VEGF-A isoform $VEGF_{165}$ binds both NRP1 and NRP2, and NRP1 binding can be further enhanced by the presence of heparin (*171*, *172*). VEGF-C and VEGF-D bind to NRP2 with greater affinity, although they interact with both NRPs (*173*).

Primarily, the neuropilins function in the VEGF pathway as co-receptors of the VEGF receptors. Co-expression of NRP1 with VEGFR2 enhances VEGF-induced chemotaxis and VEGFR2 phosphorylation, through increased binding of VEGF to VEGFR2 (*171*). Inhibition of NRP1-VEGFR2 complex formation is associated with down-regulation of VEGF signalling and EC proliferation, migration and angiogenesis (*171*). NRP1 antibodies that block VEGF binding effectively inhibit VEGF-induced sprouting or neovascularisation (*174*). Another biochemical proof of NRP1-mediated VEGF signalling is the observation that VEGF-A_{165}b is both incapable of inducing the VEGF pathway and of binding NRP1 (*175*).

However, VEGFR2 phosphorylation and downstream signalling can be induced in the absence of NRP1 (*176*). Specific inhibition of VEGF-NRP1 binding cannot prevent VEGFR2 tyrosine phosphorylation, even though it attenuates VEGF signalling (*175*). It has also been demonstrated that NRP1 can promote the recycling of VEGFR2 from endocytosis vesicles, therefore enhancing the VEGF signal's endurance (*177*). These findings suggest that NRP1 is important, but not indispensible, in VEGF signalling. Interestingly, disruption of VEGF-NRP1 binding reduced VEGFR2 phosphorylation and downstream ERK activation, but did not affect Akt activation or VEGF-induced EC proliferation (*175*), suggesting a fine-tuning role of NRP1 in VEGF signalling (Figure 1.5). NRP1 is also able to bind VEGFR1, and the formation of the NRP1-VEGFR1 complex may compete with the binding of VEGF to form the NRP1-VEGFR2 complex (*172*).

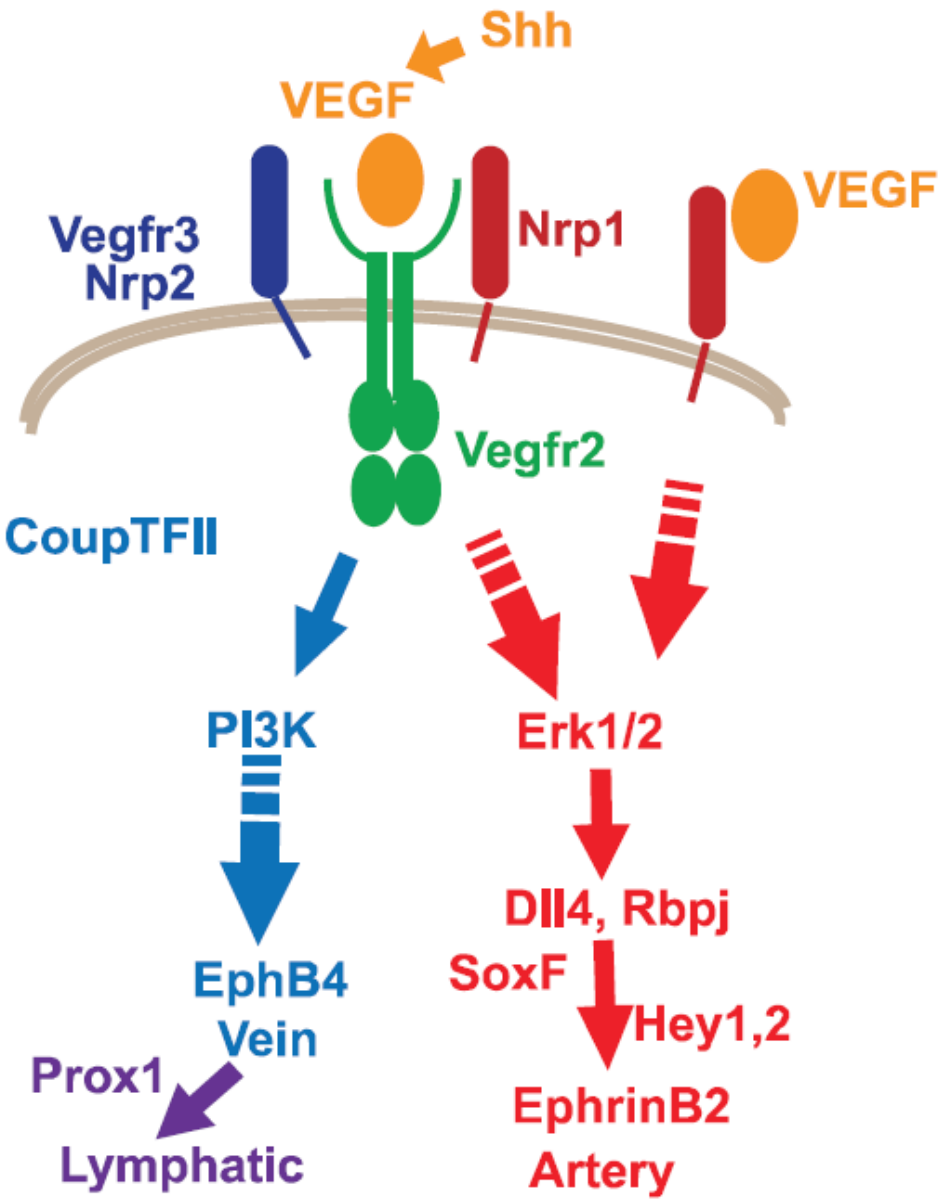


Figure1.5 NRPs in the VEGF pathway for endothelial cell differentiation

In the VEGF pathway, NRPs primarily function as co-receptors on the cell membrane. Recruitment of Nrp2, alongside Vegfr3, is associated with PI3K cascade activation. The (direct or indirect) target genes include CoupTFII and EphB4, which are venous endothelial cell identity regulators. The Nrp1 on the other hand can be recruited to the VEGF-VEGFR2 complex and promote the MAPK cascade activation. This signal results in Dll4, Rbpj activation, and further associated to arterial EC specification. This MAPK cascade can also be activated by Nrp1 binding to the VEGF ligand independent from VEGFR2, majorly in the context of promoting EC migration.

Nrp1 may also transduce cellular signals in a VEGFR-independent manner *(178)*. Recombinant NRP1 cytoplasmic region activation can promote endothelial cell migration without the recruitment and activation of VEGFR2 phosphorylation *(75)*. It has also been suggested that Nrp1 may act on VEGFR2 heterotypically, in that angiogenic endothelial cells may present NRP1 *in trans* to neighbouring cells for their VEGFR2 signal (*179*). However, such a capacity of NRP1 may be irrelevant in physiological circumstances during developmental angiogenesis (*180*). NRP2 binds VEGF-C and -D, and its binding is followed by VEGFR3 recruitment (*181*). This NRP2-VEGFR3 signal is important in lymphogenesis (*182*). NRP2 can also bind to VEGFR1 or VEGFR2 (*183, 184*) (Figure 1.5).

NRPs also serve as obligate co-receptors for SEMA3 binding to its receptor PLEXIN A (PLXNA). The SEMA3A-NRP1-PLXNA4 complex is essential for axonal guidance during neurogenesis, whereas NRP2 is essential for recognition of SEMA3B, C, F, and G and signalling through PLXNA3 (*185, 186*). The sophisticated ligand affinity variation, diverse partnership with transmembrane receptors, and tissue specific expression observed permits the NRPs to have various levels of regulation potential and to dynamically exert their biological functions. Though the NRPs are well-studied with regard to their role in chemo-repulsion during axonal guidance with SEMA3 and PLXNA (*187, 188*), this Introduction will now focus on the roles of the NRPs in vascular development, especially those executed through the VEGF pathway. SEMA3 knockout mice have vascular defects, although this is thought to be due to competitive binding of NRP1 and interference of the VEGF pathway, rather than a result of SEMA3's direct contribution (*189*).

1.3.3. Role of Nrp1 in the VEGF pathway during EC specification and differentiation

A number of studies have revealed the significance of the neuropilins in vascular differentiation and angiogenesis. Deletion of *Nrp1* results in embryonic lethality at E12.5 due to both neural system dysfunction and vascular disruption (*190*, *191*). Both the central and peripheral nervous system are affected; the efferent fibres suffer severe abnormalities in their connections and trajectories. However, in addition to neural defects, the extra-embryonic yolk sac vasculature is disorganised, and the capillary plexus is not fully developed. The development of the embryonic vasculature is also widely disrupted, the brachial arches and dorsal aorta failing to form in the *Nrp1* knockout, alongside malformations in various areas of the arterial vasculature. Morpholino knockdown of orthologous of *Nrp1* in zebrafish resulted in excessive sprouting of the intersomitic vessels, but normal trunk vessel formation, suggesting that NRP1 was not involved in early vascular development (*192*). Specific conditional knockout in ECs resulted in survival to a later embryonic stage, although cardiovascular defects were still observed, including angiogenic sprouting (*51*). The *Nrp1*$^{sema/-}$ knockout mice vasculature was thinly branched, and cardiac malformation was evident.

Nrp1 is expressed in arterial but not venous ECs in zebrafish, chick and mouse, and is therefore thought to play an important role in arterial-venous patterning. Endothelial-specific knockout of *Nrp1*in mice demonstrated a decreased expression of arterial markers (*193*). As a transmembrane protein, NRP1 possesses the highest binding affinity to VEGF-A165 amongst other VEGF ligands in the vascular context (*194*, *165*, *195*). It is thought that Nrp1 may

promote spatial segregation of arteries from veins via the VEGF pathway, possibly by VEGF modulation to favour the Dll4/Notch pathway. However, recent data suggests that NRP1 may also transduce VEGF signalling in a VEGFR-independent manner (*178*, *196*).

NRP1 also plays an important role in tip cell guidance in newly formed sprouts. Gerhardt *et al.* 2004 performed hindbrain assay analysis in wild type and *Nrp1* knockout mice. The angiogenic sprouts in the hindbrain had tip cells with long extended filopodia, and the E9.5 –E11.5 embryos saw the hindbrain sprouts make a stereotypical series of growth and movements, where the radial glia in the subventricular zone moved laterally to form the tufts. In contrast, *Nrp1* knockout tip cell filopodia did form, but were incapable of sensing the VEGF gradient and making the normal lateral movement across the subventricular zone *(213)*. This report demonstrates that Nrp1 is not needed for tip cell identity, filopodia formation or migration, but is instead essential for path finding of the sprouts.

NRP1 is also important for support cell recruitment and stabilisation of the developing vasculature during maturation. Blocking of VEGF binding to NRP1 in neonatal retinas prevents the recruitment of pericytes to the new vessels *(174)*. Since NRP1 is also expressed in smooth muscle cells, it is unclear whether NRP1 affects smooth muscle cell adhesion and migration, or that endothelial NRP1 plays a chemoattractant role in recruiting support cells and stimulating their adherence.

1.4. Project objectives

Nrp1 is a key co-receptor of VEGF that is specifically expressed during, and plays a key role in, arterial differentiation and angiogenic growth. In this study, I have identified and studied

an enhancer within the Nrp1 gene locus that directs expression to arterial and angiogenic ECs during vascular establishment and differentiation. In addition, this enhancer specifically directs expression to coronary endothelial cells as they differentiate from the sinus venosus. Since the transcriptional regulation of Nrp1 in ECs is not currently understood, it is anticipated that analysis of the enhancer will reveal the transcription regulation pathways directly upstream of Nrp1 during these stages of vascular differentiation and, consequently, the manner in which this crucial component of the VEGF pathway is regulated. Additionally, by identifying the transcription factors that contribute to arterial and coronary specific gene expression, this project will shed light on the transcriptional mechanisms that regulate arterialisation, as well as the reprogramming and re-differentiation of ECs during coronary vessel formation.

2.Materials and Methods

2.1. Mouse work

All animal procedures were supervised under local ethical review and approved by Home Office of the UK (PPL 30/2783). All transgenic lines were generated on a C57BL/6 background.

2.1.1. Genotyping

Genotyping was typically achieved by PCR followed by electrophoresis analysis. DNA was first extracted from ear notches collected from newly weaned pups, or from fragments of yolk sac from embryos. Where the recipient mother was known not to harbour the transgene (e.g when collecting transient transgenic embryos), the placenta was used, enabling me to preserve the yolk sac for other purposes.

The tissue sample was incubated over-night at 55℃with 400 µL of GNT buffer (50mmol/L KCL, 1.5mmol/L $MgCl_2$, 10mmol/L Tris-pH8, 0.01% gelatin, 0.45% nonidet P40, 0.45% Tween) and proteinase K (10mg/ml). The solution was then heated to 100°C for 5 minutes, then cooled to room temperature and centrifuged to remove debris of the sample floating in the extract. 2 µL of tissue extract was then used for a PCR reaction with the GoTag Green master mix (Promega, M7122).Genotyping PCR primers are listed below in Table 2.1.

PCR programmes were as follows:

LacZ genotyping

94℃ for 3 min, followed by 35 cycles of: 94℃ for 30 sec, 58℃ for 30 sec, 72℃ for 30 sec, completed with 72℃for 10 mins.

Cre genotyping

94℃ for 3 min, followed by 29 cycles of: 94℃ for 40 sec, 60℃ for 40 sec, 72℃ for 40sec, completed with 72℃ for 2 mins.

Erg genotyping

94℃ for 3 min, followed by 34 cycles of: 94℃ for 30 sec, 55℃ for 30 sec, 72℃ for 60 sec, completed with 72℃ for 10 mins.

Rosa26LacZ genotyping

94℃ for 3 min, followed by 34 cycles of: 94℃ for 30 sec, 63℃ for 30 sec, 72℃ for 60 sec, completed with 72℃ for 10 mins.

Table 2.1 Primers sets for Genotyping

Names	Sequences	Amplicon
LacZ primer set	LacZ3: GTTGCAGTGCACGGCAGATACACTTGCTGA LacZ4:GCCACTGGTGTGGGCCATAATTCAATTCGC	400bp
	Internal Ctrl F: TGGACAGGACTGGACCTCTGCTTTCCTAGA Internal Ctrl R: TAGAGCTTTGCCACATCACAGGTCATTCAG	200bp (internal control)
Cre primer set	Cre_F CATTTGGGCCAGCTAAACAT Cre_R ATTCTCCCACCGTCAGTACG	294bp
	Actin_F GGTGTCATGGTAGGTATGGGT Actin_R CGCACAATCTCACGTTCAG	600bp (internal control)
Erg floxP primer set	Flp (ERG) Fwd: AGATTTTGTTCTGGTTAACAAGCCGTGC Flp (ERG) Rev: AATGAGACAGAGCCATGAGGTAGATGGG	171bp (WT) and/or 279bp(*FloxP*)

Rosa26	oIMR3054	ATCCTCTGCATGGTCAGGTC	315bp
primer	oIMR0040	CGTGGCCTGGATTCATTCC	
set	oIMR8745	GTCAGTCGAGTGCACAGTT	210bp (internal control)
	oIMR8744	CAAATGTTGCTTGTCTGGTG	

2.1.2. Preparation of constructs for transgenic mouse generation

Table 2.2 describes all the enhancer constructs used for this project. The 1.8kb hNrp1in2 enhancer was generated by PCR on human DNA (extracted from HeLa cells) and then sub-cloned into the transgenic reporter plasmid *HSP68-LacZ*using NotI and XhoI restriction sites as previously reported (*126*). PCR primers used for cloning truncated versions of hNrp1in2, and mNrp1in2, are listed in Table 2.3. Half way through project, the GeneArt gene synthesis service became available and financially viable. Consequently, hNrp1in2A2 enhancer mutation constructs were generated by direct synthesis. A table of DNA strings that has been achieved in this manner is listed below in Table 2.4. The resultant DNA strings were treated for A-overhang addition upon arrival, and then treated in the same manner as PCR fragments. Both then underwent TOPO TA cloning and gateway recombination cloning using the pCR8/GW/TOPO TA Cloning Kit (Invitrogen, K2500-20) to allow entry of the DNA into the destination vector HSP68-LacZ-Gateway (provided by NadavAhituv).

To prepare for transgenic generation, theHSP68-enhancer-LacZ constructs were digested with XhoI (NEB, R0146S) and NotI (NEB, R3189S as NotI-HF) to create a linearized fragment. It was then gel-purified by electrophoresis, cleaning using Qiaquick gel extraction kits, and provided to our collaborators.

The pronuclear injection procedure was performed by Dr. Ke Liu at our collaborating group of Prof. George Bou-Gharios's in the Institute of Ageing and Chronic Diseases at University of Liverpool. A thorough description of the procedure can be found in (*197*). Briefly, zygotes are harvested from female donors' oviduct, and then positioned at the opening of a holding capillary. The diluted linearized DNA is then injected into the zygote nuclears under astereoscopic microscope. The injected zygotes are then re-implanted into the oviduct infundibulum of an anesthetized female that has been plugged by a vasectomised male. This pseudo pregnant recipient mother carries the injected zygotes. The resulting pups or embryos will be then genotyped to screen for successful transgenesis.

Table 2.2 Constructs Used for Generation of Transgenic Mice

Full Descriptive Name	Description of construct	Transgenic line name
HSP68- LacZ	The Hsp68 minimal promoter fused to a LacZ reporter gene and a SV40 polyA tail.	Hsp68-LacZ
HSP68-hNrp1in2-LacZ	The human Nrp1 intron 2 enhancer linked to the Hsp68 minimal promoter fused to the LacZ reporter gene, and a SV40 polyA tail.	hNrp1in2
HSP68-mNrp1in2-LacZ	The mouse Nrp1 intron 2 enhancer linked to the Hsp68 minimal promoter fused to the LacZ reporter gene, and a SV40 polyA tail.	mNrp1in2
HSP68-hNrp1in2peakA-LacZ	The human Nrp1 intron 2 enhancer peakA region linked to the Hsp68 minimal promoter fused to the	hNrp1in2-peakA/ hNrpin2A

	LacZ reporter gene, and a SV40 polyA tail.	
HSP68-hNrp1in2 peakA2-LacZ	The human Nrp1 intron 2 enhancer peakA2 core enhancer region linking to the Hsp68 minimal promoter fused to the LacZ reporter gene, and a SV40 polyA tail.	hNrp1in2-peakA2/ hNrp1in2A2
HSP68-hNrp1in2 _peakA2_ALL _ETS_mut-LacZ	The human Nrp1 intron 2 enhancer peakA2 core enhancer with all 8 Ets sites mutated linked to the Hsp68 minimal promoter fused to the LacZ reporter gene, and a SV40 polyA tail.	hNrp1in2A2-ALL-ETS_mut
HSP68-hNrp1in2 _peakA2_Func _ETS_mut-LacZ	The human Nrp1 intron 2 enhancer peakA2 core enhancer with4 functional Ets site mutated, linked to the Hsp68 minimal promoter fused to the LacZ reporter gene, and a SV40 polyA tail.	A2-Func-ETS_mut-LacZ
HSP68-hNrp1in2-Rosa26-CB92	The human Nrp1 intron 2 enhancer linked to the Hsp68 minimal promoter fused to the LacZ reporter gene, flanked by attB attachment sites in the backbone of a CB92 plasmid.	hNrp1in2-Rosa
HSP68- Rosa26-CB92	The Hsp68 minimal promoter fused to the LacZ reporter gene, flanked by attB attachment sites in the backbone of a CB92 plasmid.	HSP-Rosa

HSP68-mNrp1int2-Rosa26-CB92	The mouse Nrp1 intron 2 enhancer linked to the Hsp68 minimal promoter fused to the LacZ reporter gene, flanked by attB attachment sites in the backbone of a CB92 plasmid.	mNrp1in2-Rosa
HSP68-hNrp1int2-peakA-Rosa26-CB92	The human Nrp1 intron 2 enhancer peakA region linked with a Hsp68 minimal promoter, a LacZ reporter gene, flanked by attB attachment sites in the backbone of a CB92 plasmid.	hNrp1in2peakA-Rosa
miniTK-Rosa26-CB92	The Tyrosine Kinase minimal promoter fused to LacZ reporter gene, flanked by attB attachment sites in the backbone of a CB92 plasmid.	mTK-Rosa
miniTK-hNrp1in2-Rosa26-CB92	The human Nrp1 intron2 enhancer linked with the Tyrosine Kinase minimal promoter fused to LacZ reporter gene, flanked by attB attachment sites in the backbone of a CB92 plasmid.	mTK-hNrp1-Rosa

Table 2.3 Oligos for cloning the mNrp1in2 and hNrp1in2peakA2 fragments

Name	sequence
hNrp1in2A2F1	GTTACTTCCTGCCCTTCTGG
hNrp1in2A2 F2	TTGTTACTTCCTGCCCTTCTG
hNrp1in2A2 F3	ATGTTGTTACTTCCTGCCCTTC
hNrp1in2A2 R1	ATCAAGGAGGAGCCCACAAC
hNrp1in2A2 R2	CCCACAACCTCAGAGGAAAT
mNrp1-FL F1	ACTGCACTTGTGCTGGTGAC
mNrp1-FLR1	CCAACACTTCACCCATTCCT
mNrp1-FLR2	CCCTGTGAGCATTCCCTCTA
mNrp1-FLF2	CCATTCTCTCCTCCCTGTGA

Table 2.4 Synthetic DNA strings

Name of string	Sequence
hNrp1in2A2 All ETS mut	CTTATGTTGTTACT*gg*CTGCCCTTCTGGGT*gg*CTATGATATGCAGT T*gg*CA*gg*CTGGTCTCA*gg*CTGTTTGGCCCCTTTGCGTAAAAGCTAA GccTGGCTGGCCCACAGCCTACCAGGGCTTTCAAAGCCCCAGAA GTCTGTAAACAtctcAGATT*gg*CTCTGAGGTTGTGGGCTCCTCCTTG A
hNrp1in2A2 Func ETS mut	CTTATGTTGTTACT*gg*CTGCCCTTCTGGGTTCCTATGATATGCAGT TTCCATCCTGGTCTCA*gg*CTGTTTGGCCCCTTTGCGTAAAAGCTA AGGATGGCTGGCCCACAGCCTACCAGGGCTTTCAAAGCCCCAG AAGTCTGTAAACAtctcAGATT*gg*CTCTGAGGTTGTGGGCTCCTCCT TGA
Nrp1in2A2 FOXGATAmut	CTTATGTTGTTACTTCCTGCCCTTCTGGGTTCCTAT*tcaa*TGCAGTT TCCATCCTGGTCTCATCCTGTTTGGCCCCTTTGCGTAAAAGCTAA GGATGGCTGGCCCACAGCCTACCAGGGCTTTCAAAGCCCCAGA AGTCTGT*ggg*CAGGAAAGATTTCCTCTGAGGTTGTGGGCTCCTCC TTGA

2.1.3. Single copy Rosa26-locus targeted transgenic mice

We collaborated with Ben Davies (Wellcome Trust Centre for Human Genetics, Oxford) to use a recombinase-mediated cassette exchange protocol to generate single copy, Rosa26-locus targeted transgenic mice. In brief, the enhancer to be test was cloned into a plasmid such that it was surrounded by an insulator, and upstream the HSP68 minimal promoter and the LacZ reporter gene, with the entire construct flanked by attB sties. Neomycin and hygromycin resistance genes were used for positive selection of successful integration events. The construct was electroporated together with PhiC31 mRNA into Rosa26 target-carrying mouse ES cells. Homologous recombination took place between the plasmid attB and the host genomic attP sites, directing the attB flanked region of the construct into the ES genome, and the positive colonies of ES that survived selection were injected into the blastocyst to create chimera carrier. Germ-line integration events permitted the establishment of the mouse line. The eventual transgenic mouse contained a single copy of hNrp1 enhancer linked to the Hsp68 minimal promoter and *LacZ* coding region at the Rosa26 locus flanked by an insulator.

2.1.4. Analysis of reporter gene expression in transgenic mice

Extraction and X-gal staining

Transient transgenic mouse embryos were taken at the time point specified by culling recipient females and removing embryos. For stable transgenic lines, male transgenic mice were crossed with wild type females for time mating, and the female mouse sacrificed at the required embryonic stage. To take embryos, the uterus was removed to a Petri dish of chilled

1XPBS-A, and embryos harvested. For embryos that wereE13.5 days or older, the skin was also removed to facilitate X-gal infiltration, and the hearts dissected away from the body. Once dissected, samples were rinsed in 1 X PBS- before fixing in FIX solution (see Table 2.8) at 4℃. Fixation time was subject to the size and stage of the embryo, with E8.5-9.0 embryos requiring 20 mins, E13.5 requiring 1hr, E15.5 requiring 3hr. The embryos were then washed in RINSE solution for 30 mins on ice twice, and then incubated with MAGIC STAIN working solution overnight in the dark in room temperature to allow the X-gal reaction to occur. After X-gal staining, embryos were again rinsed in 1 X PBS-A then fixed overnight in 4%PFA at 4℃.

For neonatal retina harvest, the eyes were firstly harvested using bent-tip forceps from new pups at postnatal day 5 or day 7. They were then fixed in 1% PFA over night at 4℃ before the retina was removed from the cornea for further process (*198*). The transgene expression was detected by X-Gal staining, same as in embryos.

Imaging

Whole mount embryos were visualized using ProGres CF Scan under a stereo microscope (Leica M165C). The E13.5 hearts pictures were imaged as a "Z-stack" at several of sequential focusing planes. The pictures were then loaded in Adobe Photoshop as a stack, followed by the "auto-align layers" command to align the images. The images are then merged using "auto-blend layers", to allow sharp visualization of the heart at depth.

Paraffin embedding and sectioning

To paraffin embed whole embryos or organs, the tissues underwent an automated program of a series of steps in a paraffin embedding machine. In brief, the samples are firstly dehydrated

through being placed in an ethanol series: 70% ethanol, 1h;80% ethanol, 1h;95% ethanol, 1h;100% ethanol, 2 changes, 1h each; xylene, 2 changes, 1h each; and paraffin wax (56-58℃), 2 changes, 1.5h each. The paraffin infused tissues were then embedded into paraffin blocks using the embedding working space of a Thermo Scientific HistoStar. The blocks were placed in the desired orientation and sectioned using a Thermo Scientific Microm HM355S.This was done in collaboration with Dr Indrika Ratnayaka.

Frozen embedding and histology

Embryos were fixed in 4% paraformaldehyde overnight, then stored in 1xPBS. They were then placed into a 10% sucrose solution and incubate for 3-15 hours with motion at 4℃. This was repeated with 20% and 30% sucrose solutions. For smaller embryos this time can be reduced down to around 15 minutes for each sucrose solution. Samples were then brought to room temperature, removed from sucrose solution and placed in small plastic moulds. Samples were immersed in O.C.T. compound (Tissue-Tek, Cat. #4583), placed on dry ice and set to freeze at -20℃ for 2 hours-overnight. Sections were cut at 10microns using cryostat (Bright OTF5000) by Dr Indrika Ratnayaka.

Histological analysis

Samples were first rehydrated, 5mins X 2 of Histo-clear II (National Diagnostics, HS-202), 3; 3 mins X 2100% ethanol; 1 min in 90% ethanol; 1min in 70% ethanol; and 1 min in distilled water. For counterstaining X-Gal stained embryo sections, the slides were then placed in nuclear fast red (Electron Microscopy science, Cat#26356-02) for 30 seconds, and then rinsed in water extensively, before going through dehydration process of ethanol series: 1mins in 70% ethanol;1min in 90% ethanol;3 min X 2 in 100% ethanol; and 5 mins X 2 in Histoclear.

The slides were then mounted by Vectashield HardSet mounting medium (Vetorlabs, H-1004) before pictures are taken by ProgRes C5under the histology microscope Zeiss Axioplan.

This paraffin embedding and sectioning procedure was also used to embed and prepare non X-Gal stained embryos for *in situ* hybridization, immunohistochemistry or immunofluorescence analysis (see section 2.6).

2.2. Zebrafish

2.2.1. Tol2 transgenesis

All Home Office regulated zebrafish work was performed by Natalia Sacilotto on PPL 30/2783. All embryo analysis was performed by me.

Zebrafish were raised and maintained at 28.5℃in systemically circulating water. Embryos were maintained in E3 medium (5 mMNaCl, 0.17 mMKCl,0.33 mM $CaCl_2$,0.33 mM $MgSO_4$.) at 28.5℃. Typically, the embryos are obtained by timed spawning of wild type adult zebrafish.

Generation of Tol2-Mediated Mosaic Transgenic Fish

Transient mosaic transgenic zebrafish embryos were generated using the Tol2 transposon system (*199*),. In brief, 0.5 nl of Tol2 transposase (50 ng/µl), capped mRNA (mMESSAGE mMACHINE SP6 Kit, AM1340) and pE1b-enhancer-GFP expression vector (60ng/µl) were combined for injection. Microinjecting was performed under dissection microscope Leica Wild M8 using the micromanipulator Leitz M-type by Natalia Sacilotto. Standard glass wall capillaries (Clark Electromedical Instruments, GC100F-15) were pulled using pipe puller (David Kopf, Narishige PB-7) to allow backfilling of the sample and fine insertion into the fish egg. Microinjection is performed into the embryos at the one-cell stage.

2.2.2. Analysis of transgene expression

The embryos are dechorionated prior to analysis. Typically 48 hours post fertilization (hpf) and 72 hpf were the time points the reporter gene expression is determined. The embryos

were anesthetized with tricainemesylate treatment, to minimize fish movement and to allow microscopy visualization.

The fish embryos were transferred each into a flat bottom 96-well plate well, and GFP reporter gene expression screened under UV light using an up-right microscope. The total survival number of the injected fish, the total GFP expressers number, the vascular expresser number were all noted, and the "typical" transgenic embryos were recorded by microscopy. The whole fish were imaged using the "tile scan" command, combined with Z-stack collection under a confocal microscope Zeiss LSM 710 MP at 488nm excitation and 509nm emission. The fish were then removed to a petri-dish and recovered in E3.

2.3. Tissue culture

2.3.1. Cell lines and Maintenance

Cos-1 (ATCC, CRL-1650) and bEnd.3 (ATCC, CRL-2299) cells were purchased from ATCC. These monolayer cell cultures were grown at 37℃ in humidified air containing 10% CO_2, in Gibco high glucose DMEM medium (Gibco, 11965-084) with 10% FBS, supplemented with L-glutamine and penicillin/streptomycin (100U/mL Penicillin; 100 µg/mL Streptomycin).

Splitting Cells

10cm plate of 80% confluent Cos-1 cells was split by first adding 10ml 1XPBS to wash, and then incubate at 37℃with 0.5-1.0ml with Trypsin/EDTA (0.25%; Cellgro, MT25053CI) for 3-5 minutes. The plate is then bang gently to dissociate cells. 10ml medium is then added to terminate the digestion, pipette up and down to remove cells and place in tube. The cells are then spun down and 10ml fresh medium is added to the pellet to re-suspend and form the cell

solution. 1ml of cell solution will then be added to 9ml fresh media (with FBS) in a new 10cm plate.

Cryopreservation and thawing of the cells

Each sub-confluent 10cm plate can be made into 1 vial. First adding 10ml 1XPBS to wash, and then incubate at 37℃with 0.5-1.0ml with Trypsin/EDTA (0.25%; Cellgro, MT25053CI) for 3-5 minutes. The plate is then bang gently to dissociate cells. 10ml medium is then added to terminate the digestion, pipette up and down to remove cells and place in tube. The cells are then spun down, re-suspend pellet in 500µl 100% FBS (ice cold). DMSO with 15% FBS was then added to cells and aliquot into pre-chilled freezer tubes. Cells freeze in -80℃ overnight before being transferred to liquid nitrogen storage. For thawing, frozen vial of cells were placed in 37℃ water bath and watch until thaw. 1ml medium are mixed into the cells before transferred into fresh medium in plate.

2.3.2. Cell Transfection

Cells were plated into 12 well plates at 70% confluence. Allow to grow overnight in incubator. Medium is refreshed and carefully drip transfection reaction complex (MirusTransIT®-LT1, MIR2300) onto each well as appropriate, then swirl in figure of "8" to mix. Incubate cells for 40-48 hrs before analysis. The constructs used in this study is listed below in Table 2.5

.

Table 2.5 Constructs for transient transfection

Name	Description
pRK5	The empty vector of pRK5. Harbouring a CMV promoter that can promote mammalian cell translation.
pRK5-FoxC2	Recombinant FoxC2 coding region on the pRK5 backbone.
pRK5-Etv2	Recombinant ETV2 coding region on the pRK5 backbone.
hNrp1in2-TKβGal	hNrp1in2enhancer upstream of TKβGal reporter gene coding region.
mNrp1in2-TKβGal	mNrp1in2enhancer upstream of TKβGal reporter gene coding region.
hNrp1in2peakA-TKβGal	hNrp1in2peakAenhancer upstream of TKβGal reporter gene coding region.
hNrp1in2peakB-TKβGal	hNrp1in2peakB upstream of TKβGal reporter gene coding region.

2.3.3. Analysis of transient transfection in cell lines

Protein Extraction:

Plates were removed from incubator to bench. Media was aspirated, and the plate washed once with 1XPBS..1ml of 1XPBS-A was added and left for 5 minutes. Cells were removed by pipetting and/or cell scraper, and spun for 1 minute at 14,000 rpm, after which the media was aspirated. If necessary, the cell pellets were frozen. 100µl 0.1M $NaPO_4$ (pH7.4) was added to each pellet which was re-suspended by pipetting and placed on ice. Cells were broken open by freeze/thawing as follows: Tubes placed in liquid nitrogen (1 min) then in 37℃

water bath (1min)three times, then spun for3-10mins at 14,000rpm. The Bradford assay was used to establish the amount of protein in each sample, using 2-10µl extract and 200-1000ml Bradford reagent (Bio-Rad, 500-0205).

Luminescent β-Gal assay

To analyse expression of the LacZ gene, I detected β-Galactosidase levels using the Luminescent β-Galactosidase detection kit II from Clontech (631712). The reaction used12µl cell extract with 1µl substrate and50µl buffer, incubated for 1hr in the dark before reading using Microplate Multimode ReaderPromegaGloMax.

2.4. Phylogenetic foot-printing and *in silico* analysis

The human *Nrp1* intron 2 1,860bp fragment (chr10:33589869-33591728, GRCh37/HG19) was used for this analysis. The fragment was used to BLAT for orthologous fragments in mouse (chr8:130911050-130912307), opossum (chr8:239603846-239605139) and chick (chr2:13,822,886-13,824,248). Justification of species selection is explained later in Chapter V. The sequence of human Nrp1 intron2 enhancer (hNrp1in2) was both aligned against all other three orthologous fragments at the same time, and individually against the mouse sequence. Alignment analysis was done in EBI ClustalW2.

For *in silico* analysis, the hNrp1in2 was loaded into BioBase Transfac database for transcription factor motif recognition. The conservation (alignment) and functional (motif) information were then compared to create a list of candidate *cis* elements in this hNrp1in2 fragment.

2.5. Electrophoresis mobility shift assay

Making protein for EMSA

Vector contained the coding sequence for the protein required (or DNA binding domain) with a functional ATG sequence. Vectors used were pCITE2, pcDNA, pCS2 or pCR2.1vector. Constructs were sequenced to ensure that coding sequence was in frame with ATG. Recombinant protein was generated using the TNT T7/SP6 Coupled Reticulocyte Lysate System (Promega, L5020) according to manufacturer instructions. Briefly, 25µl TNT lysate, 2µl reaction buffer, 1µl RNA polymerase (usually T7 or Sp6), 0.5µl leucine amino acid mix, 0.5µl methionine amino acid mix was incubated with the template DNA 1µg (circular)at 30℃ for 1.5 hrs.

Preparing EMSA probe

5µl top strand oligo (1mg/ml), 5µl bottom strand oligo (1mg/ml)were mixed with 10µl 2X annealing buffer and incubated for 10 mins at 95℃. Heat block was then switched off and samples left to cool. 1 µl of this annealed oligo mix was incubate with 2.5 µl buffer 2 (New England Biolab)1 µl 0.5mM dGTP, 1 µl 0.5mM dATP, 1 µl 0.5mM dTTP, 17 µl distilled water, 0.5 µl Klenow (New England Biolab) and 1 µl 32P-dCTP (PerkinElmer EasyTides, NEG502A250UC)at RT for 30-60 minutes. The radiolabeled EMSA probe was then ran on anon-denaturing acrylamide gelin 0.5 X TBE buffer for 1-2 hours at 150V, covering in Saran wrap an imaged using radiographic image on an overlaid film, aligned with pen marks. The labelled probe was extracted from the gel, placed in a tube with 100ul 40mM KCl, and incubated at 37℃overnight to elute. The sample was then spun down and supernatant collected as the probe. 2µl of sample was added into 1ml scintillation liquid and analysed in a

micro plate counter (PerkinElmer 2450 MicroBeta2 LumiJet). 10-40,000 cpm of each probe was used for each reaction.

EMSA

Most samples were run on a 6% acrylamide gel. For very small proteins such as ETS1_DBD, I used a 10% acrylamide gel. Gels were poured and allowed 10-20 minutes to set, then pre-run in 0.5X TBE buffer at 150V at RT, or 200V at 4ºC.For each EMSA reaction, the following was added:2µl 10X binding buffer, 0.2-1.0µl poly dIdC (Roche,Cat10108812001), 1-5µl protein from TNT reaction, (1µl competitor) and distilled water up to 18µl.Samples were incubated for 10 mins at room temperature, after which 2µl hot probe at 10-40,000 cpm/µl was added, and further incubated for 20-30 mins at RT. To run on a gel, 2ul 10X DNA loading buffer was added. After gel was run, it was dried at 80℃ for 30-45 minutes, and exposed overnight onto radiography film.

Table 2.6 EMSA oligos

Name	sequence
hNrp1in2A2 ETS1 F	ctagTTATGTTGTTACTTCCTGCCCTTCTGGG
hNrp1in2A2ETS1 R	ctagCCCAGAAGGGCAGGAAGTAACAACATAA
hNrp1in2A2 ETS1 mu F	ctagTTATGTTGTTACTggCTGCCCTTCTGGG
hNrp1in2A2 ETS1 mu R	ctagCCCAGAAGGGCAGccAGTAACAACATAA
hNrp1in2A2 ETS2 F	ctagTGCCCTTCTGGGTTCCTATGATATGCAG
hNrp1in2A2 ETS2 R	ctagCTGCATATCATAGGAACCCAGAAGGGCA
hNrp1in2A2 ETS2 mu F	ctagTGCCCTTCTGGGTggCTATGATATGCAG
hNrp1in2A2 ETS2 mu R	ctagCTGCATATCATAGccACCCAGAAGGGCA
hNrp1in2A2 ETS3-4 F	ctagTATGATATGCAGTTTCCATCCTGGTCTC
hNrp1in2A2 ETS3-4 R	ctagGAGACCAGGATGGAAACTGCATATCATA
hNrp1in2A2 ETS3 mu F	ctagTATGATATGCAGTTggCATCCTGGTCTC
hNrp1in2A2 ETS3 mu R	ctagGAGACCAGGATGccAACTGCATATCATA
hNrp1in2A2 ETS4nc mu F	ctagTATGATATGCAGTTTCCAggCTGGTCTC
hNrp1in2A2 ETS4nc mu R	ctagGAGACCAGccTGGAAACTGCATATCATA
hNrp1in2A2ETS3-4 mu F	ctagTATGATATGCAGTTggCAggCTGGTCTC
hNrp1in2A2 ETS3-4 mu R	ctagGAGACCAGccTGccAACTGCATATCATA
hNrp1in2A2 ETS5nc F	ctagCGTAAAAGCTAAGGATGGCTGGCCCACA
hNrp1in2A2 ETS5nc R	ctagTGTGGGCCAGCCATCCTTAGCTTTTACG
hNrp1in2A2 ETS5nc mu F	ctagCGTAAAAGCTAAGccTGGCTGGCCCACA
hNrp1in2A2 ETS6 F	ctagAACAGccAAGATTTCCTCTGAGGTTGTG

hNrp1in2A2ETS6 R	ctagCACAACCTCAGAGGAAATCTTggCTGTT
hNrp1in2A2 ETS6 mu F	ctagAACAGccAAGATTggCTCTGAGGTTGTG
hNrp1in2A2ETS6 mu R	ctagCACAACCTCAGAGccAATCTTggCTGTT
mef2c-F10 F	CTAGcatctccttttcccggggtttcc
mef2c-F10 R	CTAGccttctcccgggcccatctctgt
hNrp1in2A2 FE2mut-F	CTAGTGGGTCTGTACTCCTGCCCCTT
hNrp1in2A2FE2mut-R	CTAGAAGGGGCAGGAGTACAGACGACCA
hNrp1in2A2 GATA fwd	ctagGGGTTCCTATGATATGCAGTTTCC
hNrp1in2A2 GATA Rev	ctagGGAAACTGCATATCATAGGAACCC
hNrp1in2A2 GATA mutfwd	ctagGGGTTCCTATtcaaTGCAGTTTCC
hNrp1in2A2 GATA mut Rev	ctagGGAAACTGCAttgaATAGGAACCC
Erg control site F	CTAGgggagagtgcaccggaagtcagttaa
Erg control site R	CTAGttaactgacttccggtgcactctccc
hNrp1in2A2ETVmut Fwd2	ctagAAGTCTGTAAACAtctcAGATCGACT
hNrp1in2A2ETVmut Rev2	ctagAGTCGATCTgagaTGTTTACAGACTT
hNrp1in2A2 FE2-F	CTAGTGGTCTCATCCTGTTTGGCCCCTT
hNrp1in2A2 FE2-R	CTAGAAGGGGCCAAACAGGATGAGACCA

2.6. immunohistochemistry and immunofluorescence

Immunohistochemistry on paraffin sections

Tissues were fixed in 4% buffered paraformaldehyde and then dehydrated in an ethanol series, cleared in xylene and embedded in paraffin wax (see 2.1.4). Sections were cut at 5µm thickness and either stained with Nuclear Fast Red (Sigma-Aldrich, N3020) or processed for immunostaining. Rehydrated paraffin-embedded sections were processed in 10mM sodium citrate buffer, pH 6 for antigen retrieval, and then incubated in 3% hydrogen peroxide in methanol, washed in PBS and then blocked with 5% normal goat serum in 1 X PBS-A for 1hr at RT. The sections were then incubated at 4℃overnight with the primary antibody diluted in blocking solution. Next, sections were washed with 1 X PBS-A, and incubated with either biotinylated or Alexa Fluor (Molecular Probes)labelled secondary antibodies, always in blocking solution, for another 30 minutes. Biotinylated conjugation was detected by or further amplified by Elite ABCkit (Vectorlabs, PK-6100) was visualized using the peroxide substrate solution DAB (diaminobenzidine, Vectorlabs SK-4100).

Immunofluorescence on cryosections

Slides were incubated in 1 X PBS-A for 5 mins, then fixed for 10 mins in 4% PFA, and rinsed for 5 mins in 1 X PBS-A. Samples were blocked in 3% normal goat serum for 30 minutes, after which 100-200µl primary antibody (diluted 1 in 1000 in 3% normal goat serum) was added to each slide, carefully covered with cover slip. Samples were incubated at room temperature for 1-3 hours OR at 4℃ overnight. To prevent the slides drying out, this was done in humid conditions.

After incubation, slides were rinsed in 1 X PBS-A for 1 hour. For each slide, 100-200 µl diluted secondary antibody was added and covered with cover slip, incubated at RT for 1 hour in humid conditions. Samples were then rinsed in 1 X PBS-A for 30-60 minutes. If fluorescent antibodies were used, these we then mounted using Slow-Fade media (Vector Labs). The slides were then visualized using Q1CAM Fast1394 under Zeiss Axiophot. If HRP conjugated secondary antibodies were used, the image was developed using a peroxidise substrate kit (Vector Labs).

Primary antibodies used and relative concentrations are listed in Table 2.7.

Table 2.7 Table of Antibodies used in IHC or IF

Antibody	Source	Application
Beta-Gal	MP Biomedicals, 559761	IF, 1:300
Endomucin	Santa Cruz, sc-65495	IF, 1:500
Isolectin B4-Biotinylated	Vector Laboratories, B-1205	IF, 1:1000
Neuropilin1	Abcam, ab81321	IF, 1:150
PECAM (CD-31)	Dianova, DIA-310	IF, 1:150

2.7. Stock solutions

Table 2.8 Recipe of stock solutions

Name	Recipe
Rinse solution:	For 500ml, 50ml 10X PBS-A, 1 ml 1M $MgCl_2$, 500mg sodium deoxycholate (deoxycholic acid), 1ml NP40, To 500ml with distilled water. Store at 4℃.
Stain solution:	500ml Rinse solution, 825mg potassium ferricyanide, 920mg potassium ferrocyanide, Store this at 4℃ in the dark.
Working stain solution:	10ml Magic Stain, 100ul 100mg/ml X-Gal in N,N,dimethylformamide
10% paraformaldehyde:	For 1L, Heat 600ml water in microwave to around 60-70℃, add 100g paraformadehyde (sigma), stir on a hot plate (you must do this in a fume hood); add 1M NaOH until powder goes into solution (around 1-2mls); add 100ml 10XPBS-A; make up to 1L with distilled water, and pour into 10ml and 50ml aliquots and freeze until needed.
4% paraformaldehyde (PFA):	For 10ml, 4ml 10% paraformaldehyde, 6ml 1xPBS-A and adjust to PH 7.6.
FIX (2% PFA and 0.2% glutaraldehyde):	For 10ml, 2ml 10% paraformaldehyde, 80µl glutaraldehyde, 8ml 1 X PBS-A, adjust to PH 7.6.
Tricaine stock Solution:	400 mg tricaine in 97.9 ml E3, 0.26 g Tris in 2.1 ml E3, Adjust

	pH to 7 and store this solution in a 50 ml tube in the freezer. To use as an anaesthetic: 4.2 ml tricaine solution in 100 ml E3.
Black buffer, 10X	400mM KCl, 150mM HEPES (pH 7.9), 10mM EDTA, 5mM DTT 50% glycerol, make up with distilled water; Store at -20℃ (it will not freeze).
Annealing buffer 2X(1ml):	100mM Tris pH7.6, 10mM DTT, 20mM $MgCl_2$, 2mM Spermidine, distilled water (850µl)
SDS running buffer:	Tris base 30.3 g, Glycine 144 g, SDS 10 g
Non-denaturing acrylamide gel:	10 ml 40% acrylamide/bis 19:1 (Biorad #1610144), 4 ml 5 X TBE, 26 ml distilled water, to polymerize, add: 500µl 10% Ammonium persulfate (APS; Fisher, BP179-25), 25µl TEMED (Amresco, 0761-25ML).
3X SDS loading buffer 10ml (store 4℃):	1M Tris-Cl pH 6.8 2.4 ml, 20% SDS 3 ml, Glycerol (100%) 3 ml, B-mercaptoethanol 1.6 ml, Bromophenol blue 0.006g 10 ml
Acrylamide gel 6%:	8ml 30% Acrylamide/Bis, 37.5:1 (Biorad #1610158), 4ml 5XTBE buffer, 28ml distilled water, 450µl 10% APS, 26µl TEMED.
Acrylamide gel 10%:	13.3ml 30% Acrylamide/Bis, 37.5:1, 4ml 5XTBE buffer, 22.7ml distilled water, 450µl 10% APS, 26µl TEMED.

3. Identification and validation of the Nrp1in2 enhancer

3.1.The hNrp1 enhancer

3.1.1. Identification of an enhancer within the Nrp1 locus

De Val et al. previously demonstrated that Etv2 can bind and function synergistically with FoxC2 on a combinatorial binding site, known as FOX:ETS, a motif found in various established endothelial regulatory elements of vascular differentiation, growth and remodelling genes (*126*). A human vascular enhancer within the Nrp1 locus was first identified via a screen for potential endothelial-specific regulators using this knowledge. This screen searched non-coding sequences with human to mouse sequence homology for the conserved FOX:ETS motif. A FOX:ETS motif was located at Chr10:33591342-33591353 (GRCh37/HG19) in the second intron of the Nrp1 gene (Figure 3.1). Consequently, a 1,860bp fragment of the human sequence (chr10:33589869-33591728) surrounding the motif was cloned in front of theHSP68 minimal promoter and the *LacZ* reporter gene, and transient transgenic mice were analysed at embryonic day 9.5 (*126*), At this time point, the anterior cardinal veins are substantially developed, whereas fusion and formation of the midline dorsal aorta has already occurred. Whole-mount analysis of hNrp1in2 enhancer-directed *LacZ* expression was apparent in vessels all around the embryo, with no specific restriction to particular organ. Some microvasculature in the developing lung, central nervous system and elsewhere in the periphery was also found to express the reporter gene. Sections through transgenic mice suggested that the hNrp1in2 enhancer directed EC-specific expression in all vessels, including both arteries and veins. However, the endogenous Nrp1 gene is known to be expressed in an arterial-specific manner (*24*). Therefore, this pan-vascular and microvascular expression is inconsistent with endogenous mouse Nrp1 expression.

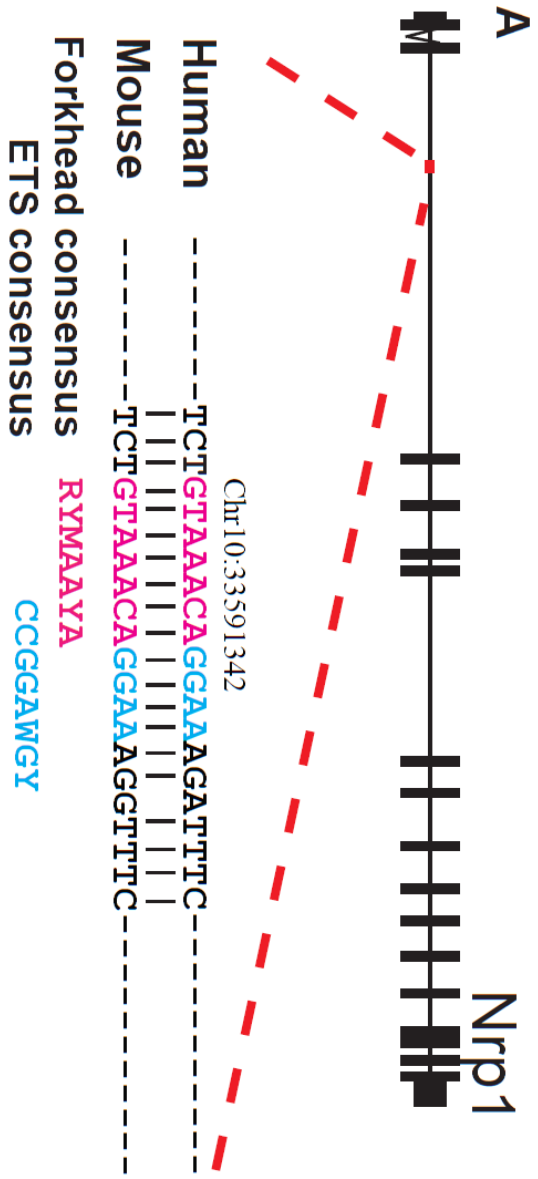


Figure 3.1 Alignment of the human and mouse Nrp1in2 sequences at the FOX:ETS site.

Magenta letters denote a non-consensus Forkhead-binding element, and the cyan letters denote a core ETS-binding site. The FOX:ETS sites located at Chr10:33591342-33591353 (GRCh37/HG19) and is conserved between human and mouse genome data sets. Consensus Forkhead- and Ets-binding sites (*200,133*) are denoted.

This putative hNrp1in2 enhancer lies in a region that is highly conserved from humans to most sequenced mammals and avian species, which indicates an early origin during evolution, and that this sequence may fundamentally participate in key developmental functions. Furthermore, the enhancer region is located within a cluster of DNaseI hypersensitive sites found in cell lines including HUVEC, but not in closely related haematopoietic cell lineages (Figure 3.2). Since DNaseI hypersensitivity indicates a naked structure, and accessibility of the DNA having lost its condensed chromatin structure, it is associated with an active involvement in the transcription regulation process in the respective cellular context (*110*). Moreover, various histone markers associated with transcription regulatory elements, namely H3K4Me1 and H3K27ac marks (*106*), are also present in this fragment in HUVEC cells (Figure 3.3). However, these markers are not present in this region in the leukaemia cell line K562, indicating that the fragment is likely to exert its functions in a tissue-specific manner. It is therefore not surprising that this region can act as an enhancer, although it is worth noting that other regions of the locus have these chromatin marks that suggest enhancer activity, despite lacking conserved FOX:ETS motifs.

The hNrp1in2 putative enhancer, therefore, presented itself as a good candidate to be a functional transcriptional regulation element directing Nrp1-specific expression in an endothelial context. However, the data from E9.5 suggested that its expression did not match that of Nrp1 in the vascular context, where specificity to arteries is observed. In addition, no information was obtained regarding the expression of the reporter gene at different stages of embryonic development. Subsequently, analyses were performed to investigate and characterise this DNA fragment.

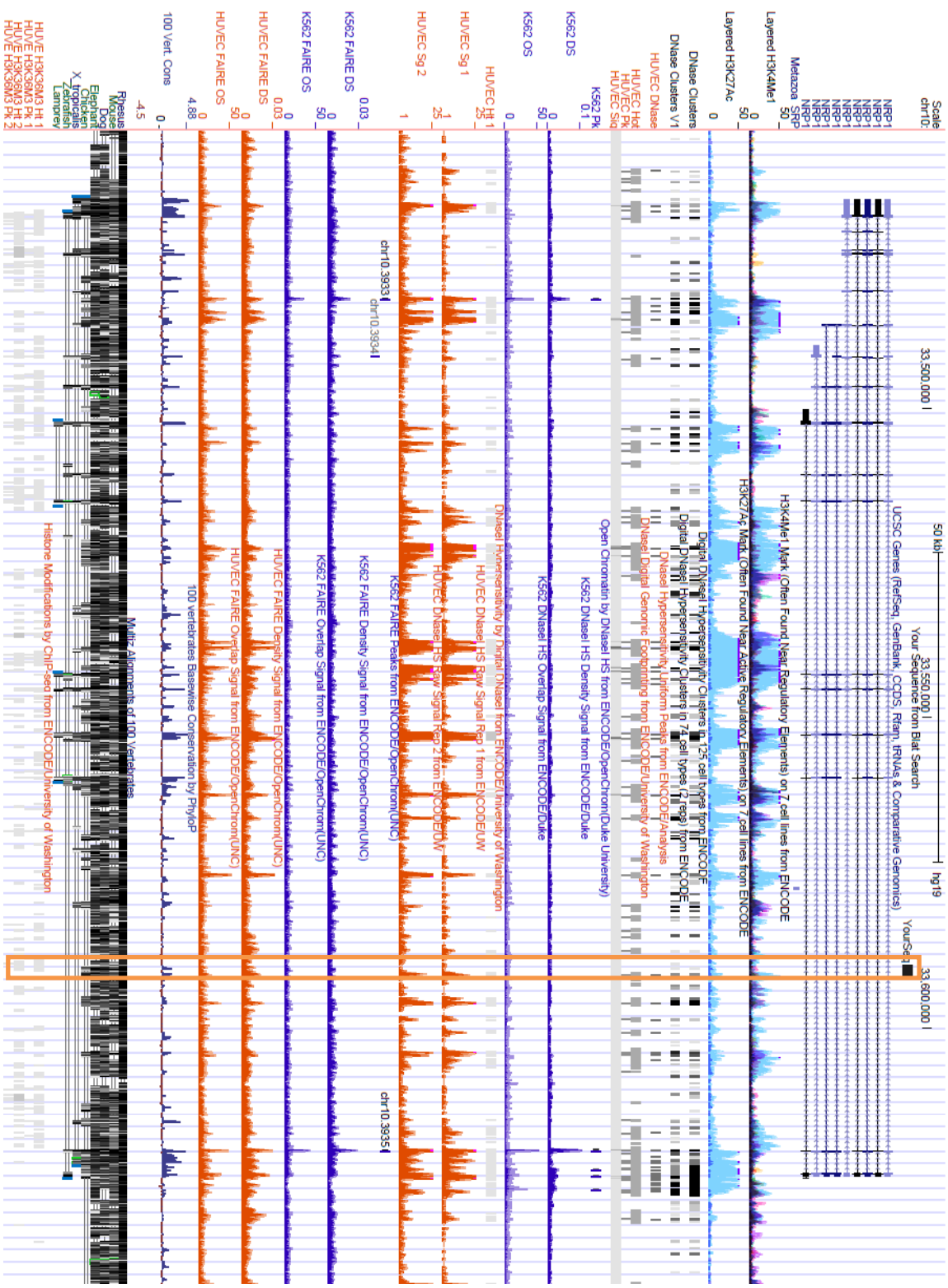


Figure 3.2 Schematic representation of the Nrp1 locus.

Figure3.2 Schematic representation of the Nrp1 locus.

The human Nrp1 locus as presented in the UCSC ENCODE Browser (*201*). HUVEC-specific H3K4me1 and H3K27ac marks are labelled in light blue, while the K562 leukemia cell line is in pink. DNaseI Hypersensitive Sites (DHS) from 125 cell lines are shown in black blocks in the DNase Cluster row. Representative DHS of HUVEC and K562 cells are shown in the orange and dark blue histogram. The vertebrate conservation score (dark blue histogram) and conservation to other animals (black clusters) are presented at the bottom. The Nrp1in2 enhancer is boxed in light orange.

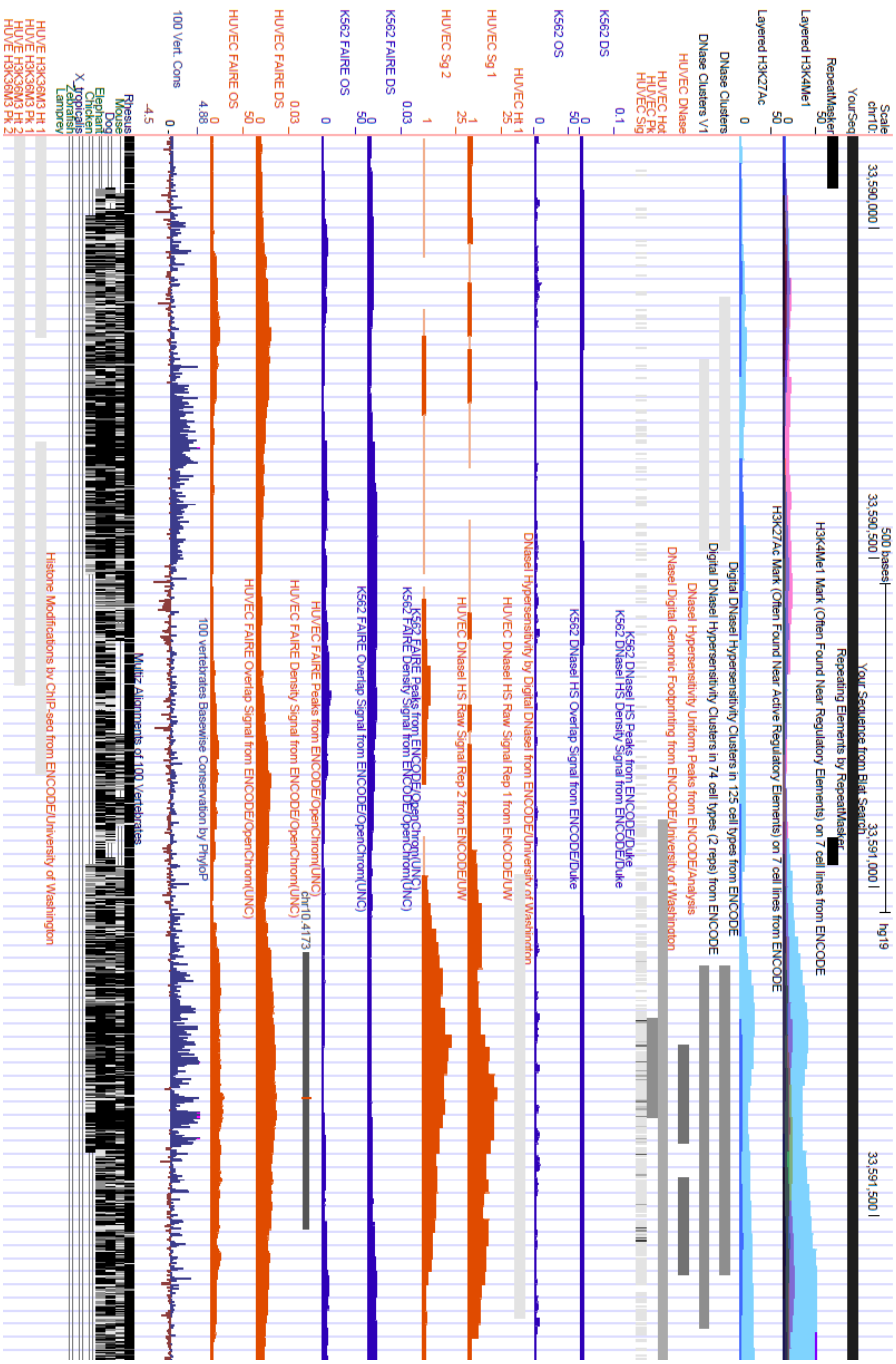


Figure 3.3 Schematic representation of the Nrp1in2 enhancer.

A more focused presentation of the human sequence surrounding the hNrp1in2 enhancer in the UCSC ENCODE Browser (*201*).The layout and colour code used is the same as in Figure 3.2. Notice that while conservation spans the whole fragment, there is deeper and more continuous conservation in two roughly 500bp regions. The HUVEC-specific DHS, the H3K27ac and H3K4me1 markers, preferentially overlap one of these conserved regions.

3.1.2. hNrp1in2 enhancer expression in arteries and veins throughout development

The pronuclear microinjections for hNrp1in2 enhancer reporter transgenesis were repeated, with the fragment constructed with a silent Hsp68 minimal promoter and the *LacZ* reporter gene (Figure 3.4 A). Transgenic embryos were analysed at various stages of embryonic development. Data were obtained from embryos of an established line and various transient transgenic embryos. It was initially intended that the data set would be obtained from one single line, however the line turned out to be difficult to breed. This may have been due to extensive expression of the reporter gene in the central nervous system. Consequently, extra data for time points had to be obtained from transient transgenic embryos.

First, the embryonic stage E9.5 (9.5 days post-coitus) was re-examined (Figure 3.4B). Dissection of the X-Gal-stained embryos (from line) consistently revealed expression in both arteries and veins. Although there was slightly stronger expression in the arterial compartment of the brachial arch artery, weak expression in the cardinal vein could also be detected, and the micro-vasculature was also positive (Figure 3.4B'). At later stages, previously unseen arterial specificity of the enhancer began to take shape. Beginning at E10.5, and continuing in the E11.5 and E13.5 time points, the dorsal aorta had clear expression of the report gene whereas the cardinal vein was silent (Figure 3.4 C, D, E). This was the case in the vessels in other organs at later phases. For example, at E18.5, the intercostal vessels (the blood vessels between the ribs) establish stereotypical arrangements and are therefore easy to identify. Expression could be detected in the intercostal arteries, but not in the veins (Figure

3.4 F). A cross-section of the rib cage further illustrated the arterial-specific expression (Figure 3.4, F1, F1').

To further confirm that reporter gene expression was found in arterial endothelial cells but not venous endothelial cells, immunofluorescence analysis was performed. By E13.5, the extracellular matrix sialoglycoprotein gene Endomucin (EMCN) becomes venous-specific and is consequently known as a marker for venous vessels [ref: morgan 1999 blood]. Immunofluorescence of EMCN, NRP1 and β-Gal was performed on unstained hNrp1in2*LacZ*sections. Since endogenous NRP1 expression is arterial while EMCN is venous, and the two do not overlap (Figure 3.5A). The reporter gene β-Gal's expression was also distinct from that of EMCN (Figure 3.5B), showing that the hNrpin2*LacZ* transgene drives an arterial-specific expression pattern, at least at these later stages of embryonic development.

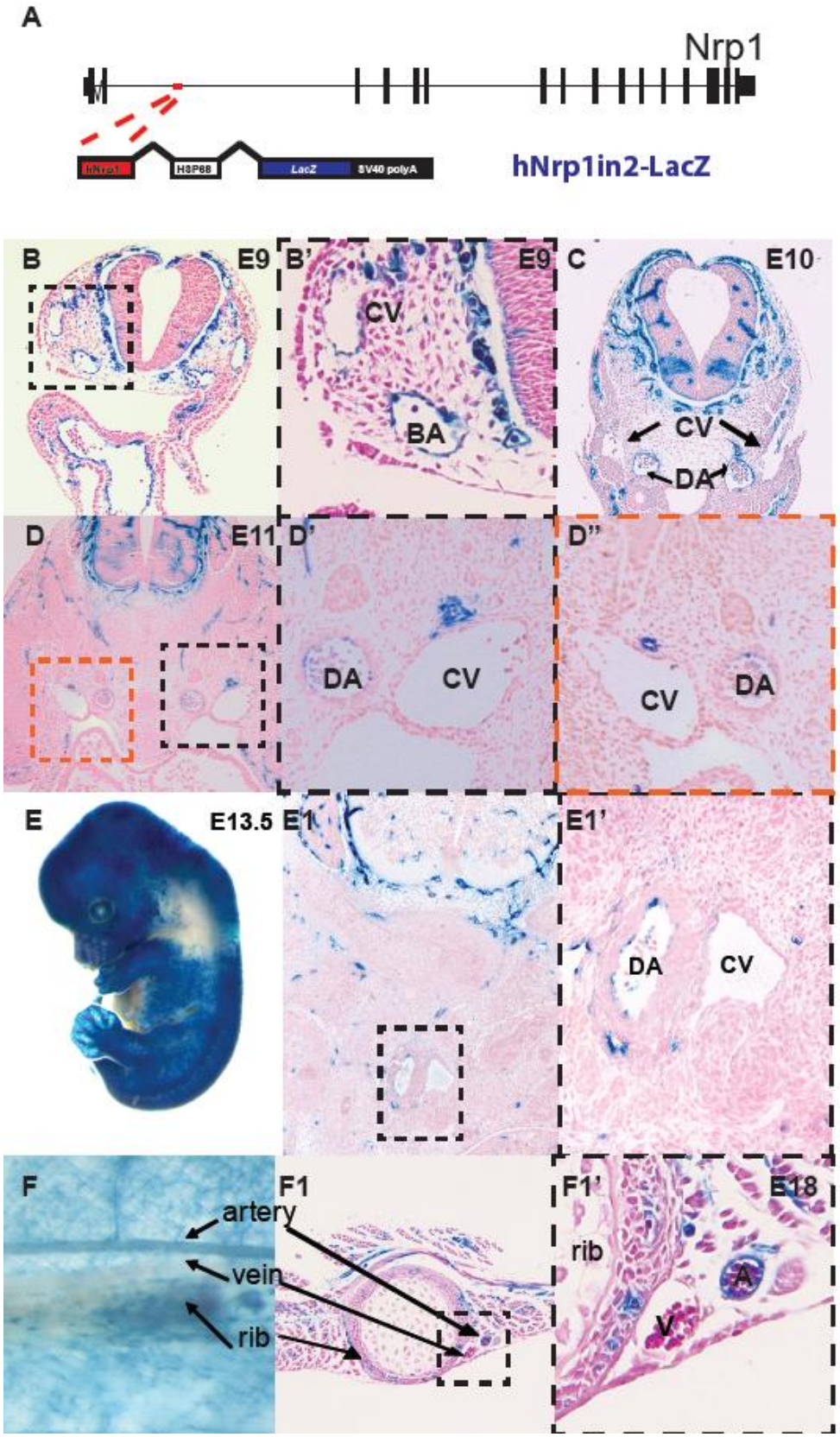


Figure 3.4 hNrp1in2*LacZ* transgenic mouse reporter gene expression in the arterial compartment of the vasculature.

Figure 3.4 hNrp1in2*LacZ* transgenic mouse reporter gene expression in the arterial compartment of the vasculature.

A. A schematic representation of the transgene construct. An enhancer from the second intron of the human Nrp1 gene was cloned in front of the silent Hsp68 minimal promoter and the reporter gene *LacZ*, forming the hNrp1in2*LacZ* transgene.

B. Mice transgenic for hNrp1in2*LacZ* at E9.5 express *LacZ* in a pan-vascular pattern. The transverse section of the X-Gal-stained transgenic embryo demonstrates that reporter gene expression can be detected in both the arterial (BA, *branchial arch artery*) and venous (CV, *cardinal vein*) compartments. Various forming vascular plexus along the neural tube were also positive. B', a magnified view of the boxed region of B.

C. The reporter expression is largely restricted to arteries in E10.5transgenic embryos. The *dorsal aorta* (DA) is positive whereas the *cardinal vein* (CV) is not.

D. The reporter expression is arterial in E11.5 transgenic embryos. Similar to at E10.5, the *dorsal aorta* (DA) is positive whereas the *cardinal vein* (CV) is not. D' and D", magnified view of the boxed region in D in the corresponding colour.

E. Arterial expression in the E13.5 embryo. E, whole mount E13.5 expression of the reporter gene β-Gal is detected by X-Gal staining. E1, transverse section of the embryo in E. E1', magnified view of the boxed region in E1, which illustrates that the dorsal aorta is positive, while the cardinal vein is negative for the reporter gene.

F. The reporter expression is arterial in E18.5 transgenic embryos. Whole mount X-Gal staining of E18.5 intercostal vessels appears arterial-specific, although staining can also be seen in the microvasculature. F1, transverse section of the rib cage.F1', the magnified view of the boxed region of F1.

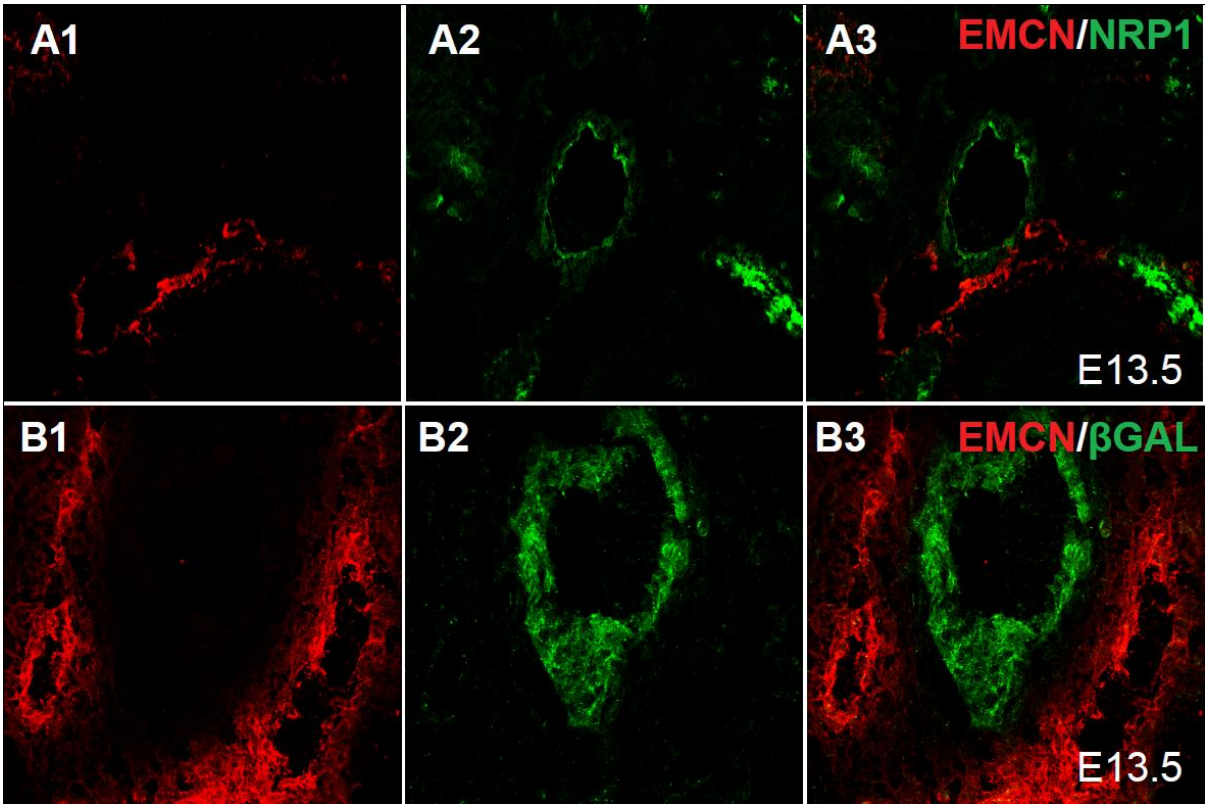


Figure 3.5 Arterial-specific expression of hNrp1in2*LacZ* is consistent with endogenous NRP1

A. Immunofluorescent staining of Endomucin (EMCN) and Nrp1 in E13.5 paraffin sections. Endogenous Nrp1 (green) is positive in arterial endothelial cells whereas Endomucin staining (EMCN, red) is venous-specific.

B. Immunofluorescence staining of Endomucin and β-Gal in E13.5 paraffin sections of hNrp1in2*LacZ*. The reporter gene β-Gal (green) is detected in arterial endothelial cells, whereas EMCN (red) is venous-specific.

3.1.3. hNrp1in2 enhancer expression in the coronary vasculature throughout development

Hearts were removed from embryos at various stages of embryonic and neonatal development, stained with X-Gal, and analysed for hNrp1in2 expression in the different vascular compartments of the heart. Sections from E10.5 embryos demonstrated that expression of the hNrp1in2 enhancer could be detected in some, but not all, of the endocardium mass of the heart (Figure 3.6 A). At this time point the coronary vessels have not been formed. At E11.5, expression was no longer detected clearly in the endocardium, but a few positive cells could be detected in the septum (Figure 3.6 B). Whole-mount analysis of E12.5hearts revealed that the hNrp1in2 enhancer specifically directs expression in vascular regions (Figure 3.6 C1, C2), highly similar to that described by Red Horse et al. as the origin of all coronary vessels (Figure 3.6 D). By E13.5, both whole-mount and sagittal sections clearly demonstrated that, while endocardial expression of the transgene was gone, the transgene was now expressed in the developing coronary vasculature (Figure 3.6 E). This was further emphasised in sections of E15 and E18 hearts. For example, a coronal section of the E18.5 heart demonstrated expression in the vascular plexus under the epicardium (Figure 3.6 G1, sub-epicardial layer) where the coronary veins reside, as well as expression in the arterial coronaries that are embedded in the myocardium (Figure 3.6 G1'). Such an expression pattern was maintained into the neonatal stage P2 (Figure 3.6 H).

Interestingly, the coronary vasculature was not the only site of expression seen in the heart. From as early as E11.5, endothelial cells in the midline of the primordial septum also started

to express the reporter gene, which could also clearly be seen at E13 and E15 (Figure 3.6 E2, F). This expression is similar to the pattern described as “blood islands” by Red Horse et al., and these can also form coronary vessels. However, unlike those that originate from the sinus venosus, the origin of these cells is unclear. Possible sources of the cells range from the proepicardial organ to the endocardium, although their discontinuity from the sinus plexus and the fact that these cells are distal from the sinus venosus suggests that they are from a secondary source. Additionally, by E15, there is intense staining of the X-Gal reporter in the endothelial cells of the valves (Figure 3.6 F”, G2 and H”).

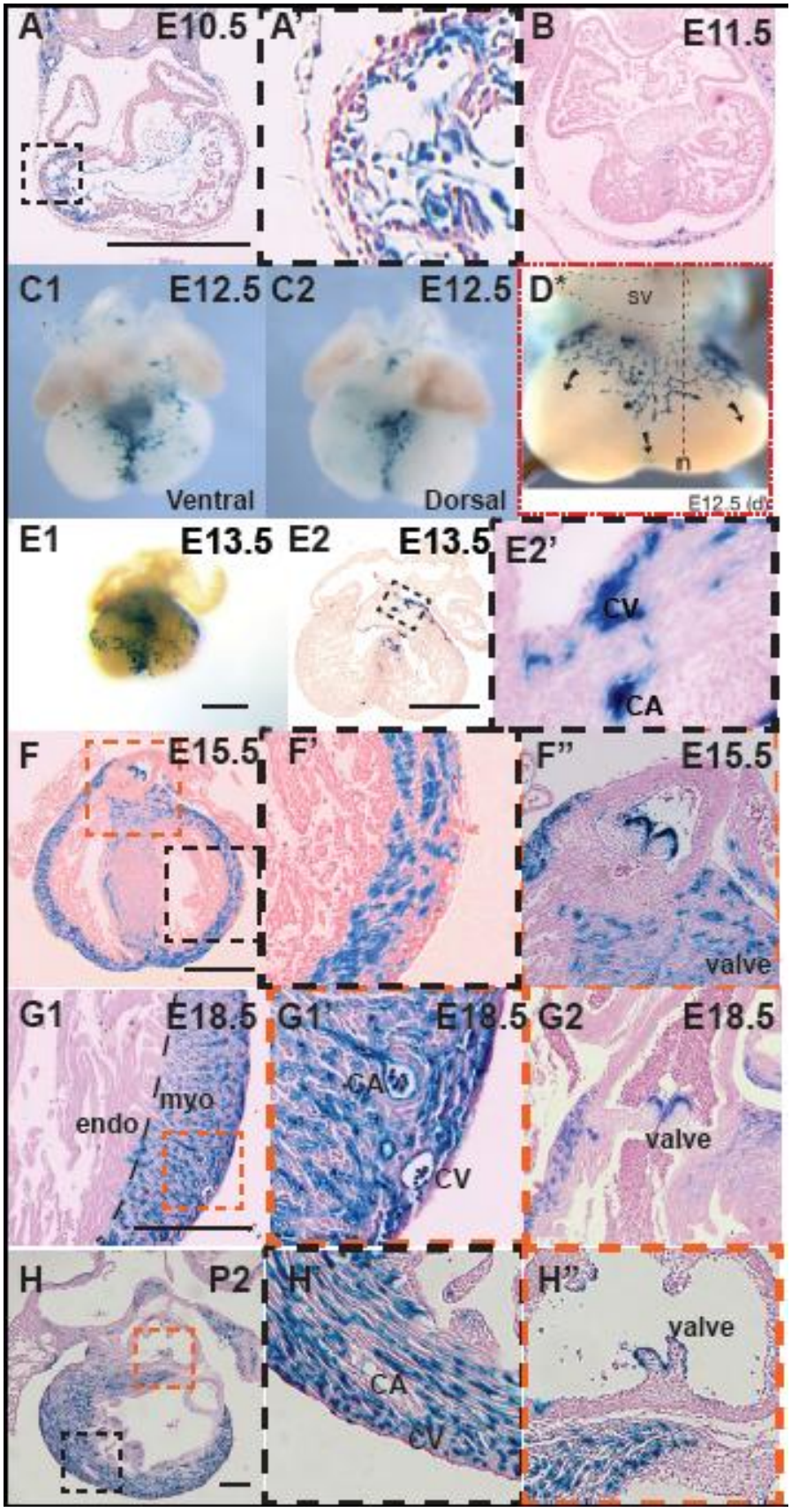


Figure3.6 Coronary expression of hNrp1in2*LacZ* in embryonic and neonatal stages.

Figure 3.6 Coronary expression of hNrp1in2*LacZ* in embryonic and neonatal stages.

A. The cardiac loop region of an E10.5 embryo transverse section. The avascular heart is void of X-Gal staining besides the positive endocardial lining. A', a close-up illustrating the endocardial expression.

B. The cardiac loop region of an E11.5 embryo transverse section. The primordial septum expresses the reporter gene.

C. hNrp1in2*LacZ* heart at E12.5 stained with X-Gal. C1, the ventral view of the heart. C2, the dorsal view of the heart.

D*. Previously reported Aplin knock-in *LacZ* expression in the E12.5 heart embryo (*40*). The Nrp1in2*LacZ* enhancer shares similarities in the expression pattern of the downward invasion of the vascular plexus.

E. Coronary expression in an E13.5 embryo. E1, whole mount E13.5 heart was dissected and stained on its own. The plexus of the developing coronary is positive for X-Gal staining. E2, coronal section of the E1 heart. E2', the enlarged view of the boxed region in E2; notice the coronary artery and vein are both positive.

F. The E15.5 heart expresses hNrp1in2*LacZ*in various endothelial cells. Coronal section of the X-Gal-stained heart. The midline of the primordial septum, the vessels in the myocardial layer, and the valves are all positive for X-Gal staining. F' and F", magnified view of the boxed region in D with corresponding colour, illustrating expression in the coronary vasculature (F') and that the epithelial layer of the valves is positive (F").

G. The E18.5 heart expresses hNrp1in2*LacZ*in various endothelial cells. G1, the coronary vessels in the myocardial layer are positive. G1', magnified view of the boxed region in G1, illustrating the expression in the coronary arteries and veins.G2, the valves are positive for X-Gal staining.

H. Coronary vascular expression of hNrp1in2 is maintained in post-natal stage P2. H' and H", magnified view of the boxed region in H of the corresponding colour, illustrating the expression in coronary vessels and heart valves, respectively. Scale bar = 2.5mm.

To test whether the enhancer drives endothelial-specific expression in the same way as endogenous Nrp1, immunofluorescence was performed against NRP1, β-Gal and CD31/PECAM on cryosections of the heart (Figure 3.7). In the myocardium, the endothelial cells of the coronary vessels expressed Nrp1, which co-localised with the endothelial marker PECAM. β-Gal was also detected in endothelial cells in the coronary and co-localised with PECAM (Figure 3.7 A1, A2). However, in the endocardium, endogenous NRP1 was detected, although the β-Gal immunofluorescence from thehNrp1in2 enhancer was absent (Figure 3.7 B1, B2). Interestingly, this indicates that the hNrp1in2 enhancer has vascular endothelial-selective expression within the developing heart, and may indicate that different enhancers, and maybe different transcription factors, are in charge of directing gene expression in the different types of cardiac endothelial cells.

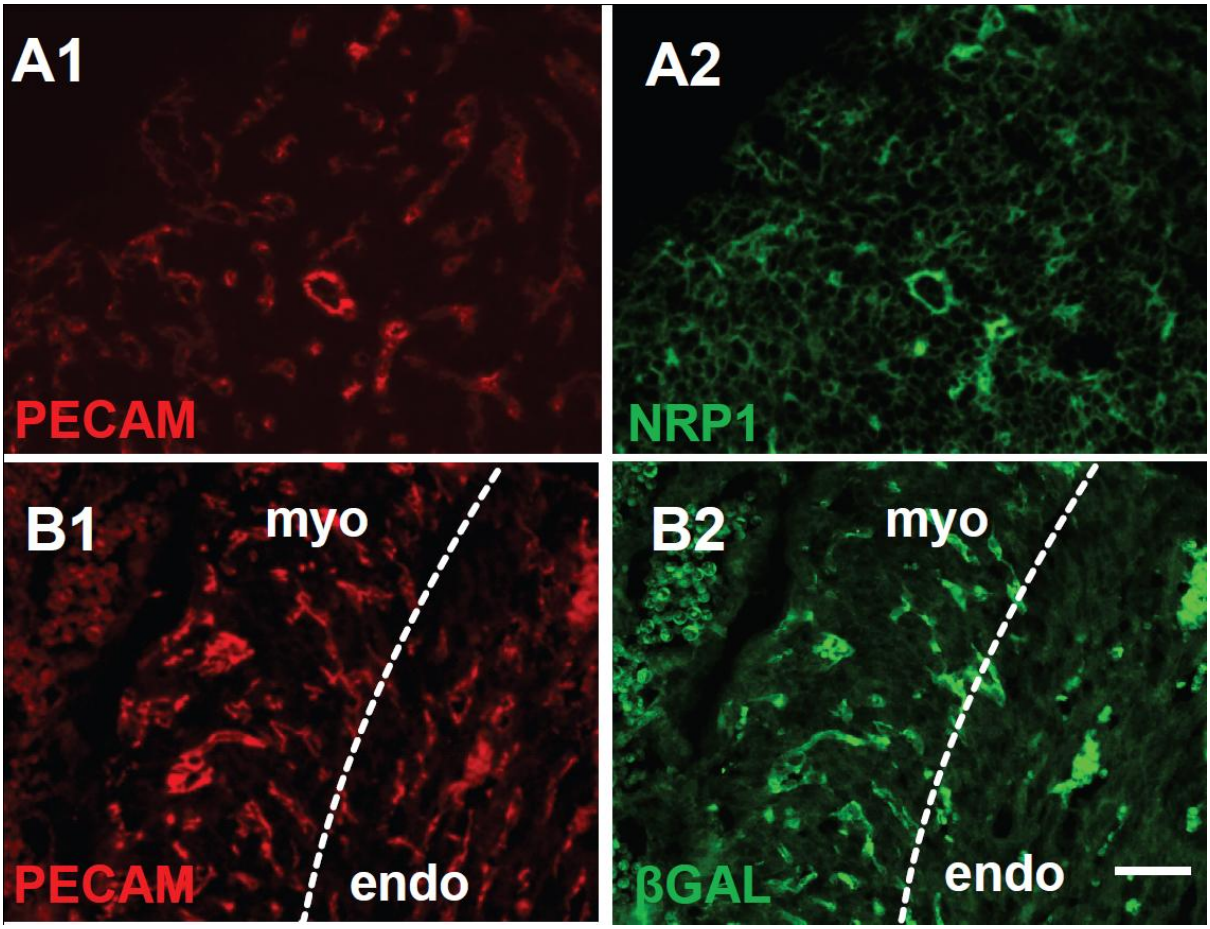


Figure 3.7 Immunofluorescence of PECAM, NRP1 and β-Gal on hNrp1in2*LacZ* heart coronal sections.

A. Immunofluorescence of PECAM (red) and NRP1 (green) on ahNrp1in2*LacZ* heart (E13.5) coronal section.

B. Immunofluorescence of PECAM and β-Gal on an hNrp1in2*LacZ* heart (E13.5) coronal section. PECAM labelled the endothelial cells in both the coronary vessels and endocardium. The coronary endothelial cells that were embedded in the myocardial layer were positive for the reporter gene, whereas those in the endocardium were not.

3.2. Conclusions

The data presented in this chapter demonstrate the identification of an enhancer of the NRP1 gene that is involved in its specific expression in the vasculature. In the early stages of embryonic development, during the formation of the dorsal aorta and cardinal vein, the enhancer drives expression in both arterial and venous endothelial cells. In contrast, in the later stages when angiogenesis dominates vascular growth, expression of the enhancer is restricted to arterial endothelial cells and the microvasculature. Such vanishing venous expression may be due to one of two mechanisms. Firstly, it is possible that the cardinal vein NRP1-positive cells migrate to the dorsal aorta during maturation of the vessels. It has previously been reported that some endothelial cells co-residing with the dorsal aorta are venous marker CoupTFII-positive, and later relocate to the cardinal vein (*25*). The Nrp1-positive cardinal vein cells may undergo a similar pathway in the opposite direction. Such cross-over migration represents delayed completion of the ongoing process of selective cell sorting of the previously mixed endothelial progenitor population. On the other hand, it is possible that the Nrp1 enhancer ceases to be expressed in the cardinal vein at a later stage. One possibility is that the enhancer is dependent on recognition of the Fox:Ets motif by the Ets factor Etv2 at an early stage of development, as Etv2 is expressed at this point. Etv2 is known for integrating cytokine pathway signals for transcriptional regulation, in particular playing an essential role in the differentiation of the vascular mesoderm from the primitive mesoderm (*123, 202*). However, as Etv2 expression decreases at a later stage, the enhancer needs an alternative ETS factor to maintain its transcriptional activity. Such a factor may be only present in arteries and not in veins, determining the specificity of the enhancer. The

Fox:Ets motif may or may not be essential for this specificity. If it is, alternative Ets factors would be needed to cover the absence of Etv2. A second wave of Ets factors starts to be expressed as the dorsal aortae amalgamate, including VEGF signal-induced elf2a, elf2b, elk4, erg, ets1, and fli1a. DLL4 has been shown to acquire arterial precursor cell-specific expression through activation of the ETS factor Erg, particularly in later stages of development (*203*). This Ets factor is a candidate for the factor responsible for reinforcing VEGF co-receptor Nrp1 expression in arterial endothelial cells. It is also possible that Fox:Ets is irrelevant to the compartmentalisation of vascular expression. In this case, transcriptional activation of the enhancer in the artery would be dependent on other motifs in the enhancer. Nonetheless, the discovery of such a factor would shed light on the involvement of the VEGF signal in arterial development.

In heart, the expression of the enhancer is detected in coronary endothelial cells throughout embryonic development in different compartments. Contrary to the arterial-specific expression seen in peripheral blood vessels, expression in the coronary vessels occurs in both arterial and venous vessels, as well as the primary plexus during early cardiac invasion and the microvasculature in the mature coronary vessels. In fact, the looping heart was also positive for this enhancer in the primitive endocardial inner layer. This may be a continuation of the endocardial expression in the linear heart tube when the cardiac region is newly defined and fused amid the continuous amalgamated artery near the first pharyngeal arch artery. Such expression vanishes as the primitive endocardium matures, since no endocardial expression was observed at time points later than E9.5. This could again be due to the fact that Etv2 expression decreases in the mature endocardium. It is also possible that some

Nrp1in2-positive endocardial cells maintain their expression but migrate out of the tissue. This may provide an explanation for the “blood island”-like structure in the lower ventricle. Indeed, it has also been reported that a ventricular endocardial lineage is able to become angiogenic in response to a VEGF signal and contribute to the coronary vascular endothelial cell population (*41*). Elucidation of the role of Nrp1 in the migration of these cells, and identification of the transcription profiles that enable them to express Nrp1 and other genes to maintain their coronary endothelial cell progenitor identity, would be of great interest.

Embryonic heart valve formation is of endocardial origin (*204*). The endocardial cushions undergo an epithelial-mesodermal transition (EMT) to become valve precursor cells during heart looping, and the cushion colonized cells proliferate, remodel and eventually mature into membranous septa heart valves. It is possible that the endocardial origin gives the valve the trace of Nrp1 enhancer expression. At this point it is difficult to understand why the expression is maintained while the valve precursor cells are still positive for reporter gene expression even though it has been through EMT already and the transcription profile must have been re-formed.

A key problem with the analysis outlined in this chapter is that it has been conducted on embryos generated by pronuclear injection. This technique of transgenesis relies on random genomic integration, and therefore results invariable and unknown copy number, insertion location and insertion orientation of the transgene. Each time a different transgenic line or transient transgenic embryo is examined, the subject represents the outcome of a different transgenesis event, with the outcome of the transgene expression likely influenced by the

nature of its insertion and its neighbouring regions. The first analysis of Nrp1 illustrated by Chong et al. (*24*), that Nrp1 expressed in a pan-vascular manner at E9.5, but no information was provided for later time points. While the analysis presented here agrees with this expression pattern in the arteries, veins and microvasculature at E9.5, the pan-vascular expression in the first transgenic line is not precisely preserved in the second. While the second line does express in all compartments in E9.5, the differential strength is noticeable (compare Figure 3.4 of this thesis to De Val et al. 2008 Figure 7C (*126*)). Arterial specificity at later time points was consistent across organs and stages, the only exception being expression in both the coronary arteries and veins. However, most of these results come from one line and one set of transient transgenic embryos. Because of the problems of generating transgenic mice using pronuclear injection, to be sure that this expression pattern truly represented our enhancer, I would need to either generate multiple additional enhancer lines or perform transient transgenic analysis at each time point. A problem could still come from trying to interpret differences in enhancer strength with differences in enhancer specificity. To further the analysis, transcription recognition motif mutations will be made on the enhancer to generate transgenic embryos, which would enable us to pinpoint the arterial specificity to a certain *trans* factor(s) recognition event. The difference in expression pattern and expression intensity incurred would then need to be interpreted against the wild type enhancer expression pattern. It would be unreliable to have a reference pattern that has severe variation within it. On top of that, it is impossible to elucidate what is the wild type "starting point" of this particular transgenesis, which the mutation-caused alteration has deviated from. It is essential to produce transient transgenesis with a good number of biological replicates, to

establish the most likely representation of the enhancer. It also avoids making conclusions from hasty generalisations that result from a technical artefact, although it is a very expensive way to do so. It is therefore desirable to use a transgenesis technique that has its integration event in a locus- and copy number-controlled manner. In such a situation, the integration target location can be designed to avoid influence from neighbouring genomic information or focal chromatin configuration. The gene dosage effect can also alter the reporter expression pattern and intensity to a great extent, and such a technical limitation could also be circumvented. A targeted single copy transgenic technique would, therefore, be of benefit to this project.

4. Single Copy Transgenesis Using The R26/PhiC31 System

4.1. Introduction

4.1.1. Complications with regular pronuclear microinjection

As discussed in the previous chapter, the method of generating transgenic mice by injecting DNA into the pronucleus of a one-cell embryo results in a random integration pattern. That is, neither the genomic loci nor the number of integrated copies of the transgene are predetermined. The variation in copy number of the transgene could lead to a change in its expression pattern or intensity. For example, changes in transgene copy number in *Drosophila* can alter chromatin structure by creating differential affinities for heterochromatin, which may then subject the region to gene silencing and eventually affect the phenotype (*205*). On the other hand, an unfortunate placement of the transgene into the realm of influence of certain chromatin configurations or a nearby transcription regulation element can also severely interfere with it, resulting in a misleading expression pattern that is difficult to interpret. Furthermore, transgene insertion can also cause mutagenesis to the host loci, resulting in unexpected phenotypes. It has been estimated that about 10% of transgene integration events may cause phenotypes (*206*), which would compromise the use of these lines in the original research plan.

4.1.2. Targeted single copy transgenesis methods

It is apparent that control over the position and copy number of the transgene would be preferable to standard pronuclear injection-generated transgenesis, particularly when comparing mutations. A well-defined location in the genome that is independent of heterochromatin or endogenous genes is ideal as an insertion target. A copy number of 1 or 2

that resembles most cases of an endogenous gene would be advisable. The homologous recombination technique can be useful for targeting chosen genomic loci. By introducing a transgenic sequence to embryonic stem cells through microinjection, the homologous recombination process occurs in the host cell at a designated genomic locus. Strategies applying such techniques have been developed in several models. In fish, aTol2-based single insertion transgenesis system has been used (*207*). However, the integration of the introduced fragment into the host genome occurs amid early divisions of the fertilised egg. Therefore, not all cells received with the transgene, resulting in mosaic transgenesis and transgenic expression. It is not until the next generation of a germ line-transmitted transgenic fish that homogenous transgenesis can be achieved, losing the advantage of the speedy nature of the model animal. In mice, the housekeeping gene hypoxanthine phosphoribosyl transferase (Hprt) locus has been used as a homologous recombination target for transgene introduction (*208*), and the fact that the locus can be selected for using hypoxanthine-aminopterin-thymidine medium makes it a convenient choice for targeting. However, recent evidence demonstrates the possibility of partnership between transgene transcription regulatory elements and the HPRT environment (*209*), which may cause confusion in the interpretation of the reporter expression pattern. Additionally, targeting to the Hprt locus is not particularly robust, requiring extensive screening before transgenic mice are generated, which has consequently prevented this technique from entering common use for the generation of mouse lines. *Ex vivo*, such a technique can be adapted to ES tissue culture, making it possible to perform high throughput scanning of enhancer analysis at relatively low cost and without ethical issues (*210*). However, although this technique has

promise, there is not yet a reliable way to differentiate ES cells into arterial, venous and coronary endothelial cell niches, making this technique unsuitable for use in this instance.

4.1.3. The R26/PhiC31 transgenesis system

PhiC31 integrase, a member of the resolvase family of bacteriophage φC31 recombinases, is a DNA site-specific recombinase that integrates DNA into chromosomes between two specific attachment sites called attB and attP. When the attB sites containing donor plasmid are electroporated together with PhiC31 mRNA into mouse embryonic stem cells (mESC) containing attP sites, homologous recombination occurs between the attB and attP sites, leading to the integration of the region flanked by attB into the mESC genome at a defined site. In collaboration with Dr Ben Davies, who runs the transgenic core at the Wellcome Trust Centre for Human Genetics, a plasmids was created that contained an enhancer construct of interest fused to the Hsp68 minimal promoter and the *LacZ* reporter gene, along with an insulator to prevent endogenous *Rosa26* (ROSAβgeo26)(*211*) enhancer/promoter activity, and a neomycin resistance gene to help with selection (Figure 4.1). After electroporation of the transgene plasmid in combination with PhiC31 mRNA, the resulting mESC were subjected to neomycin treatment for positive selection of successful transgenesis. This method of targeting is so successful that only 8-12 colonies need to be picked for each construct to ensure that a correct clone is found. Following sequence confirmation, the targeted ES cells were injected into a mouse blastocyst, which was then inserted into a surrogate mouse to generate chimeric founders. Chimeric founders that have germ line transgenesis can then be bred with wild type (WT) partner to enable the study of the reporter gene's expression pattern.

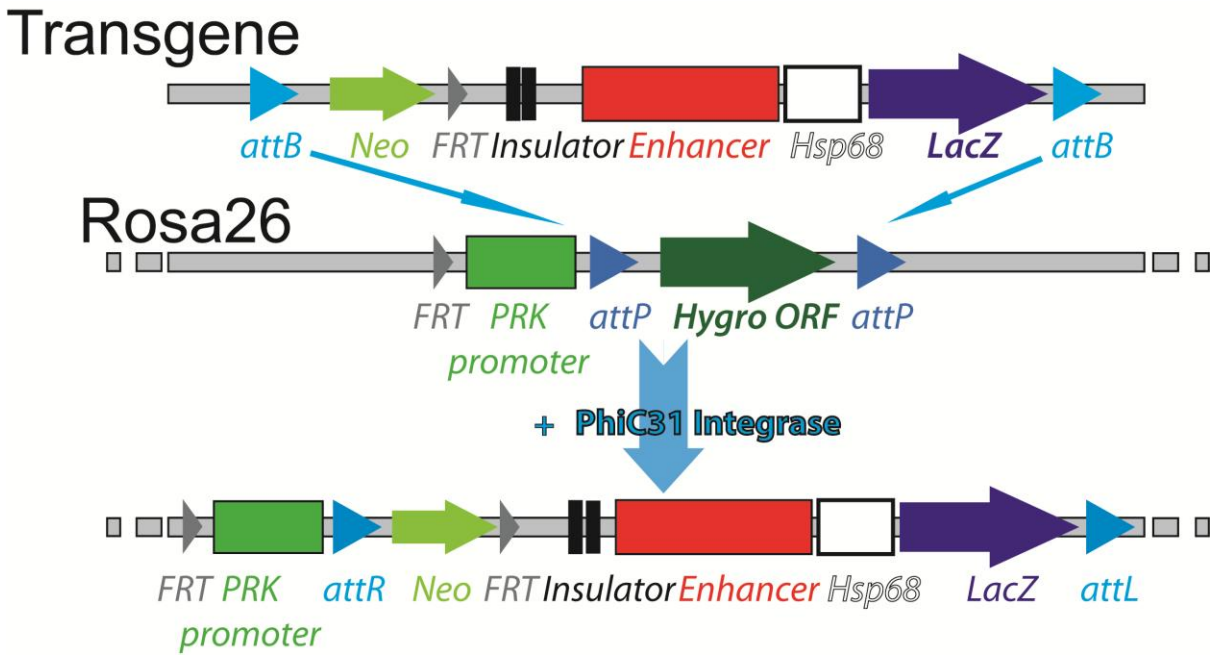


Figure 4.1 R26/PhiC31 Transgenesis Diagram.

The enhancer constructed with an insulator, the hsp68 minimal promoter and the *LacZ* reporter gene flanked by attB was electroporated together with PhiC31 mRNA into *Rosa26* target-carrying mouse ES cells. Homologous recombination took place between the attB and the attP sites, directed the attB flanked region of the construct into the ES genome, and positive colonies were injected into the blastocyst. Neomycin and hygromycin were used to select for positive colonies with successful integration events. The eventual transgenic line contains a single copy of the transgene at the *Rosa26* locus.

4.2. Results

Because this targeting strategy had not previously been used to examine enhancer activity using a minimal promoter, theR26/PhiC31 transgenesis strategy was first used to generate two stable mouse lines, one containing the Hsp68 minimal promoter alone upstream of the *LacZ* reporter gene, and the other with the human Nrp1in2 enhancer cloned upstream of the minimal promoter. This was for two reasons: firstly, to identify whether the minimal Hsp68-LacZ construct directed any expression alone at this locus; and secondly, to see whether a single copy of the hNrp1in2 enhancer could direct arterial- and coronary-specific expression patterns. For convenience male founders were bred with WT female mice to enable the examination of the expression patterns of these two transgenes at different embryonic time points.

4.2.1. Arterial-specific expression of hNrp1in2Rosa

The expression pattern of the reporter gene driven by hNrp1in2-Hsp68-LacZ in the Rosa26 locus (abbreviated to hNrp1Rosa) and Hsp68-LacZ-Rosa26 (abbreviated to HspRosa), as detected by X-Gal staining, was studied during embryonic development at time points E9.5, E11.5, E13.5 and E15.5 (Fig 4.2-4.5).

At E9.5, the multi-copy pronuclear-injected hNrp1in2 embryos described in Chapter 3 directed expression to both venous and arterial endothelial cells. Expression of hNrp1Rosa initially appeared to be similar to that of these earlier transgenic embryos, with X-Gal expression spread out in the vasculature throughout the whole body, and not restricted to particular organ (Figure4.2 A, Right). However, transverse sections of the embryos suggested

that the expression observed, even at this early time-point, was arterial-specific. For example, X-Gal staining was only detected in the arterial branchial arch, but not in the cardinal vein (Figure 4.2 B). The midline dorsal aorta was also positive (data not shown). In addition, there was strong expression in the floor plate of the neural tube, however, expression in this region was also detected in the HspRosa transgenic embryos (Figure 4.2 B, Left). The Hsp68 minimal promoter has previously been reported to occasionally drive some expression in the central nervous system (*212*), but this limitation has been largely ignored by the research community. Nevertheless, no vascular expression of the reporter gene was detected in Hsp68Rosa transgenic embryos at E9.5. It can, therefore, be concluded that arterial-specific expression of the hNrp1in2 enhancer is established very early in vascular endothelial cell specification, and that the wider expression pattern seen in multi-copy transgenics might be attributed to copy number or position effects.

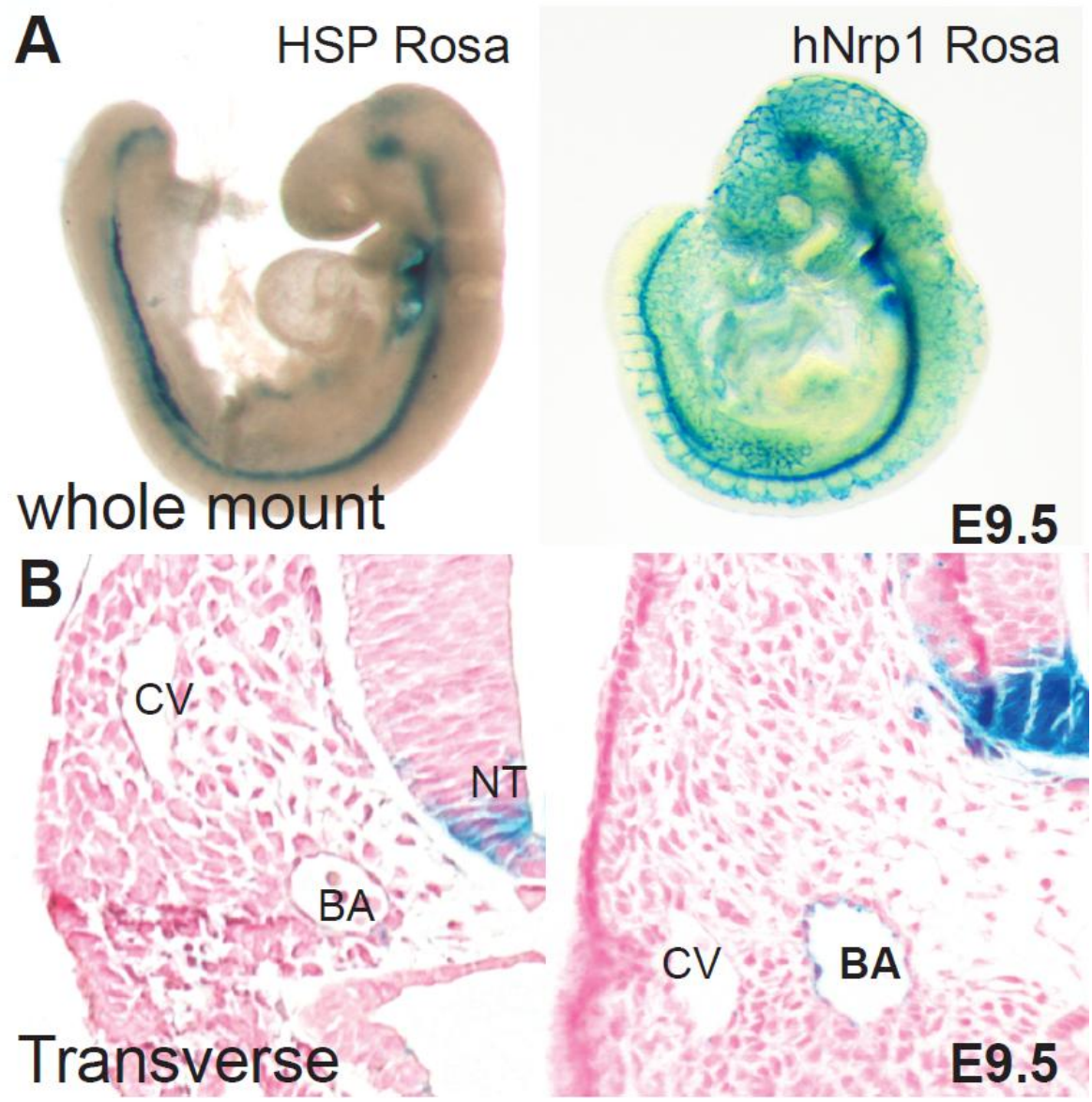


Figure 4.2 Arterial-specific expression of the Nrp1in2 enhancer at E9.5.

A. The lateral view of the corresponding whole-mount embryos of the *Rosa* lines.

B. Transverse section of X-Gal-stained E9.5 embryos of the lines indicated. The arterial branchial arches of the hNrp1in2 line sections were positive, whereas the cardinal veins were negative. Hsp68Rosaexpresses in the floor plate of the neural tube (NT), but otherwise expression was negative in the vasculature.

Abbreviations: CV, *cardinal vein*; DA, *dorsal aorta;* BA, *branchial arch*. NT, *neural tube.*

At E11.5, the arterial specificity of the hNrp1Rosa transgene was maintained. In contrast, in the Hsp68Rosa control, expression was only detected in the neural tube, and not anywhere in the circulatory system (Figure 4.3 A, A'). In hNrp1Rosa embryos, vascular expression is again observed throughout the embryo (Figure 4.3 B1). For example, the newly vascularised front and rear palms of the embryos are positive in the plexus (Figure 4.3 B2), and X-Gal staining is also seen in the vascularising pericardium and hindbrain (Figure 4.3 B3, B4). At this developmental stage, the sub-ventricular zone of the hindbrain undergoes classical angiogenic sprouting, receiving radical sprouts from the perineural plexus. These sprouts branch laterally, fuse and form a plexus, which Nrp1 is known to play a crucial role in guiding (*213*). Thus, this enhancer is directing expression in the sprouts and the resulting plexus (Figure 4.3 B5). In more mature, differentiated vessels, the dorsal aorta expresses the reporter gene in the hNrp1Rosa transverse sections, whereas the cardinal veins do not (Figure 4.3 B').

At E13.5, the detectable X-Gal staining was of greater intensity in the peripheral microvasculature, although arterial-specific expression was still detectable, as shown in the tail ventral artery rather than in the caudal vein (Figure 4.4 B). Because it was thought that the lack of X-Gal staining inside the embryo might be due to problems with the infiltration of the reagent, and to further confirm arterial-specific expression, the vascular expression of the reporter gene was also investigated by immunolabelling against β-Gal on cryosections from the E13.5 hNrp1Rosa embryos. Endothelial expression of β-Gal was detected that was arterial-specific, since the β-Gal-labelled vessels were mutually exclusive with EMCN labelled veins (Figure 4.4 C1). On a serial section, a similar pattern of NRP1

immunofluorescence could be seen in the vasculature, suggesting that the hNrp1in2 enhancer directs overlapping expression as the endogenous Nrp1 with blood vessel endothelial cells (Figure 4.4 C2).

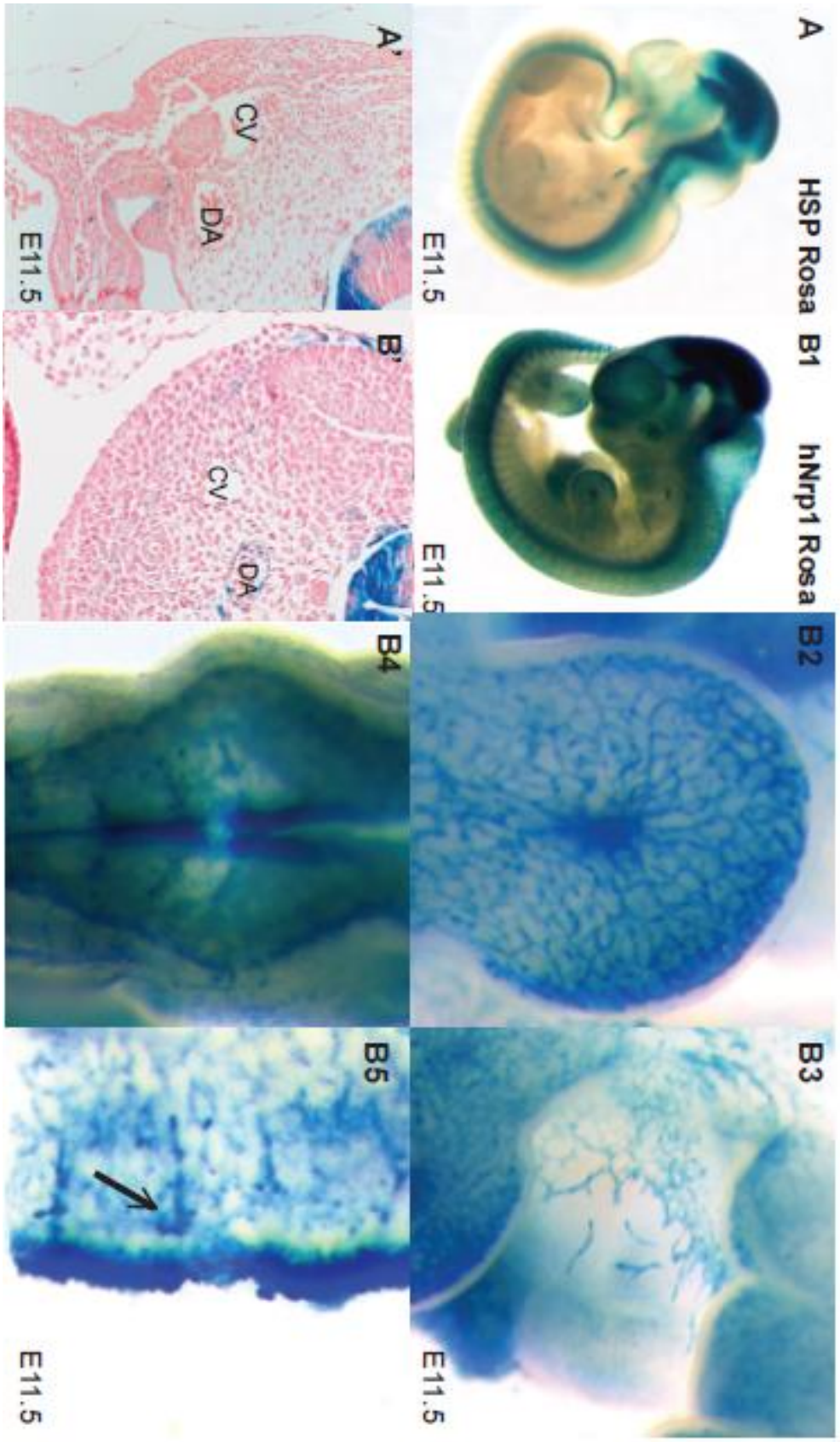


Figure 4.3 Arterial-specific expression of Nrp1in2 enhancers at E11.5

Figure 4.3 Arterial-specific expression of Nrp1in2 enhancers at E11.5

A. Side views of whole-mount X-Gal-stained embryos of Hsp68Rosaat E11.5. A', transverse section indicating that expression can be detected in the floor plate of the neural tube, but not in the dorsal aorta (DA), cardinal vein (CV) or anywhere else in the vasculature.

B. Side views of whole-mount X-Gal-stained embryos of hNrp1Rosaat E11.5. B', transverse section indicating that expression can be detected in the floor plate, but also in the dorsal aorta (DA). The cardinal vein (CV) is negative. B2, the newly vascularised front palm.B3, pericardium.B4, dorsal view of the hindbrain, subventricular zone.B5, dissected hindbrain. Notice the lateral sprouts toward the midline (arrow).

Abbreviations: CV, *cardinal vein*; DA, *dorsal aorta*..

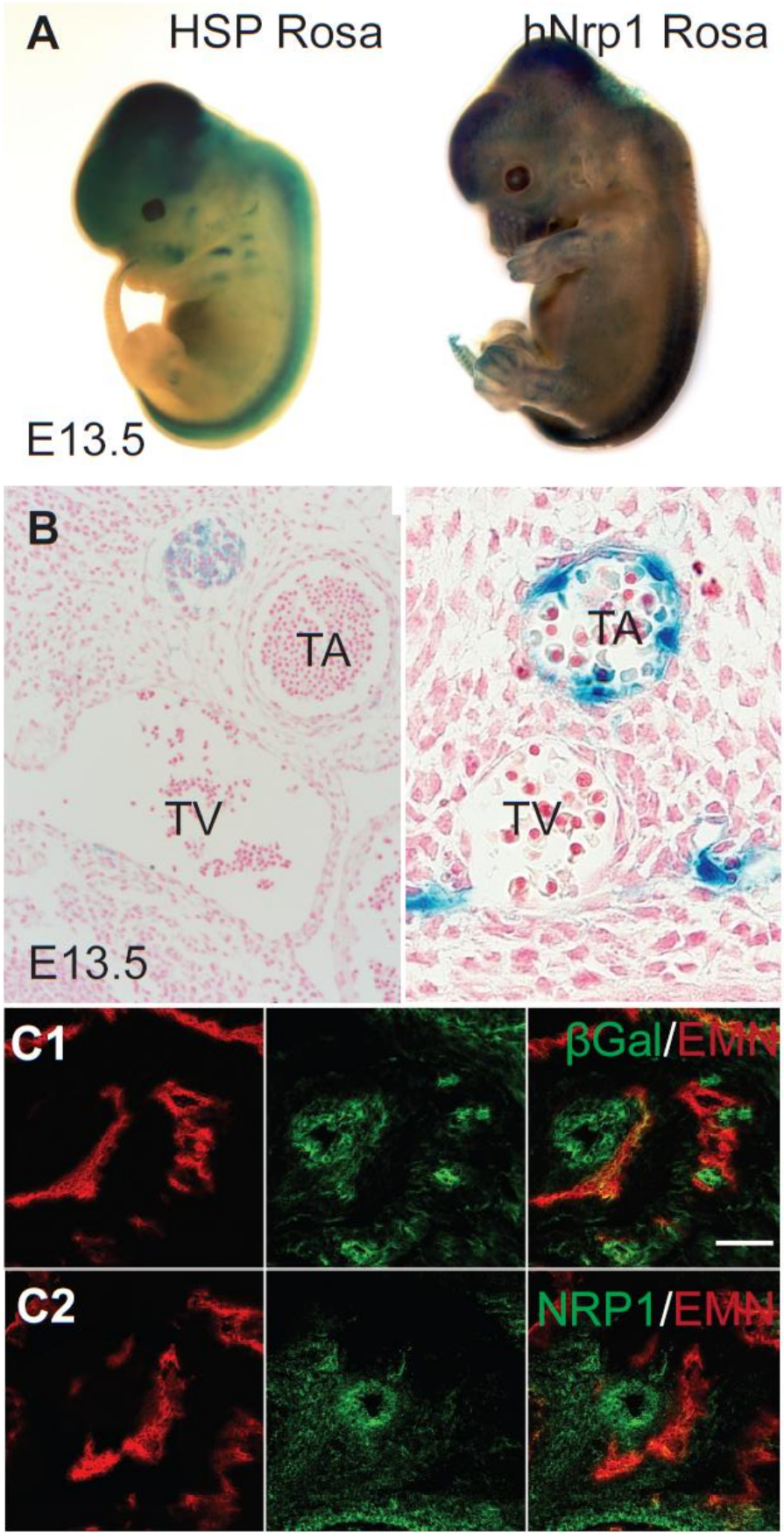


Figure 4.4 Arterial-specific expression of Nrp1in2 enhancers at E13.5.

Figure 4.4 Arterial-specific expression of Nrp1in2 enhancers at E13.5.

A. Lateral view of the corresponding whole-mount embryos at E13.5.

B. Cross section of X-Gal stained tails, counter-stained with nuclear fast red. The ventral tail artery in the hNrp1Rosa embryo is positive for reporter expression, whereas the caudal tail vein is negative. Hsp68Rosa is negative in the vasculature.

C. Arterial-specific expression directed by the hNrp1 enhancer overlaps with endogenous expression of Nrp1. C1, Immunofluorescence staining on cryosections of 13.5E hNrp1Rosa embryos showed Nrp1 expression (green, middle) in the artery but not in the vein (red, marked by Endomucin, EMN, left). C2, serial section. An identical β-Gal expression pattern was observed (green, middle). The expression pattern was similar between endogenous Nrp1 and β-Gal. Scale bar, 100 μm.

4.2.2. Coronary expression

Analysis of the coronary vasculature in hNrp1Rosa embryos demonstrated that, similar to that seen in the earlier transgenics, both arterial and venous vessels were β-Gal positive. However, unlike in the earlier multi-copy mice, expression was not detected in the endocardium at any time point.

Expression in the primitive coronary vasculature in the hNrp1in2 line could be detected as early as E11.5, and was maintained throughout embryonic development (Figure 4.5B). The two stereotypical layers of primordial coronary vessels, the outer layer (in the sub-epicardium region) being venous and the inner layer (embedded in the myocardium) being arterial, both expressed the reporter gene (Figure 4.5 C1, E13.5; enlarged in C1'right panel). Some cells in the midline of the primordial septum also expressed β-Gal and, at E13.5, the cells seemed to form a continuous structure with vessels undergoing tubularisation in the inter-ventricular groove (Figure 4.5 C1", right panel).The growing coronary endothelial network formed a spider's web-shaped plexus at E13.5, covering the upper half of the ventricles (Figure 4.5 C). By E15.5 the coronary vessels, all apparently expressing the reporter gene, covered the entire heart. This expression pattern was similar, if not the same, to that of endogenous Nrp1 expression in the coronary vasculature.

Unexpectedly, some expression of the reporter gene was also found in the hearts of HspRosa mice, although only at later stages. At E11.5, very little expression was seen, but at E13.5 the minimal promoter alone was able to drive X-Gal expression in the endothelial cells of tricuspid and mitral valves, such that the expression pattern in the valves between the two

lines was very similar (E13.5, Figure 4.5 C; E15.5, Figure 4.5 D, D1”). This clearly suggests that valve expression of the reporter gene in both cases is likely to have been directed by the Hsp68 minimal reporter, and should not be counted as coming from the hNrp1in2 enhancer fragment. However, HspRosa hearts did not exhibit substantial expression in the coronary vessels, even at this stage.

However, due to the effect of Rosa transactivity particularly dominant in the postnatal stages, all lines demonstrated expansive expression in various tissues (Figure 4.5), rather than in a vascular specific manner. It disparaged the line for the retinal angiogenesis model.

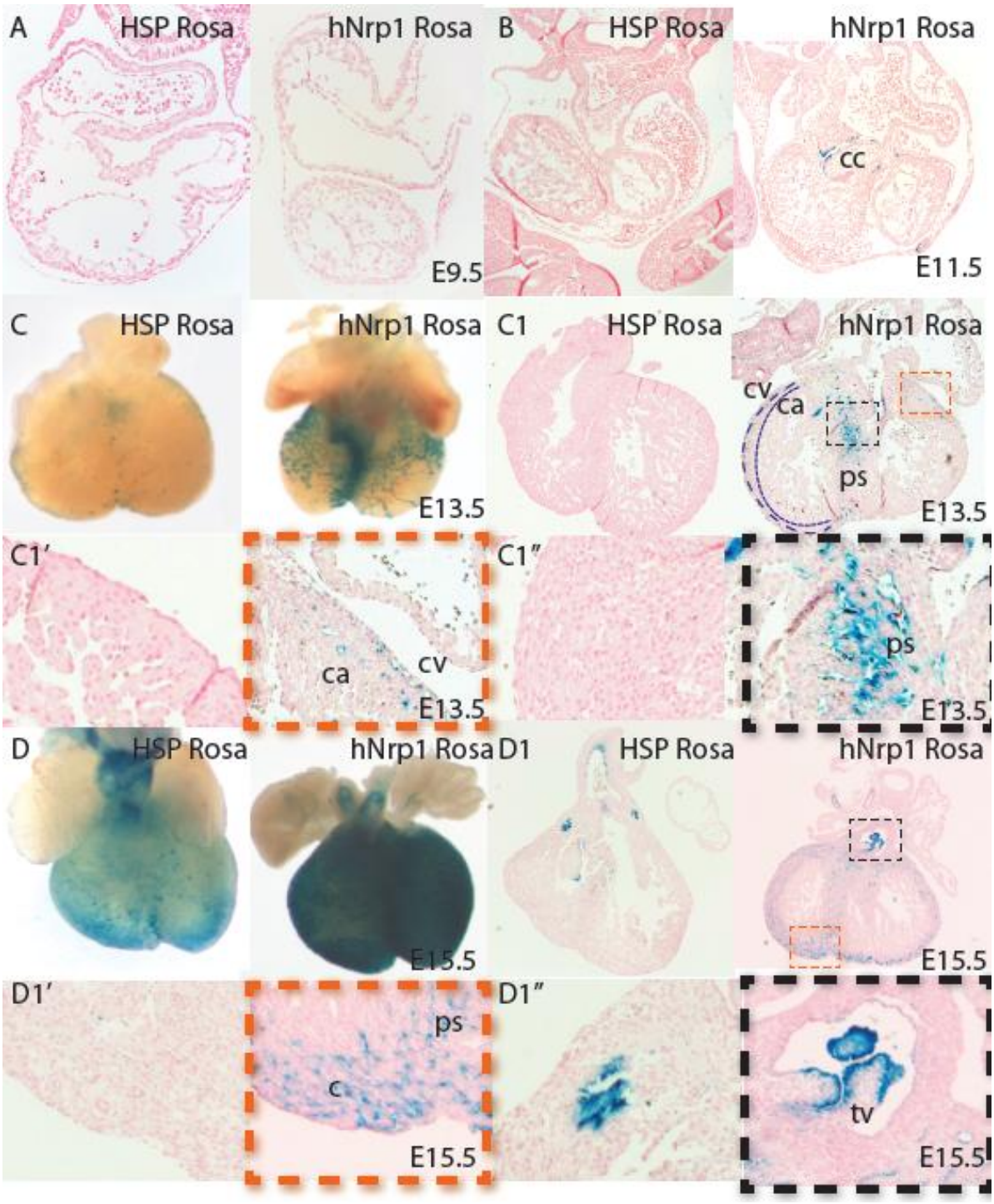


Figure 4.5 Coronary expression of the Nrp1in2 enhancers.

Figure 4.5 Coronary expression of the Nrp1in2 enhancers.

A. Transverse sections of Hsp68Rosa or hNrp1in2Rosa embryos at E9.5. The embryos were stained with X-Gal and counter-stained with nuclear fast red. No noticeable expression was found at this time point in either line.

B. Transverse sections of Hsp68Rosa or hNrp1in2Rosa embryos at E11.5. Hsp68Rosa was silent, while hNrp1in2 expression could be seen in the lining of the cardiac cushion.

C. Whole mount and coronal sections of Hsp68Rosa or hNrp1in2Rosa hearts at E13.5. Hearts were dissected from the embryos to be fixed and X-Gal-stained separately.C1, the coronary veins (indicated as the dashed line in the hNrp1Rosa picture) that reside in the outer line of the myocardium or the so-called sub-epicardial region, the coronary arteries (indicated by a dotted line) that are embedded in the deeper layer of the myocardium, and some endothelial cells in the primordial septum of the hNrp1in2 sections, were positive, whereas Hsp68 was negative in the coronary vasculature. C1' and C1", magnified view of the boxed region in C1of the hNrp1in2 Rosa E13.5 heart, and the comparative region in the Hsp68Rosa heart. The coronary arteries, veins and endothelial cells in the primordial septum are all stained in the hNrp1in2 Rosa, but not the Hsp68, heart.

D. Whole-mount hearts of both lines at E15.5. Hsp68Rosa hearts do not have plexus network staining at either stage, whereas hNrp1in2Rosa expresses in the enveloping plexus of coronary endothelial cells. D1' and D1", magnified view of the boxed region in D1 of the hNrp1in2 Rosa E15.5 heart, and its comparative region in the Hsp68Rosa heart. D1', the coronary vessels of the Nrp1in2Rosa heart is positive for X-Gal staining.D1", valves are stained in both Hsp68Rosa and hNrp1in2Rosa hearts.

Abbreviations: cc, *cardiac cushion*. ps, *primordial septum*; cv, *coronary vein*; ca, *coronary artery*.c, coronary. tv, *tricuspid valves*.

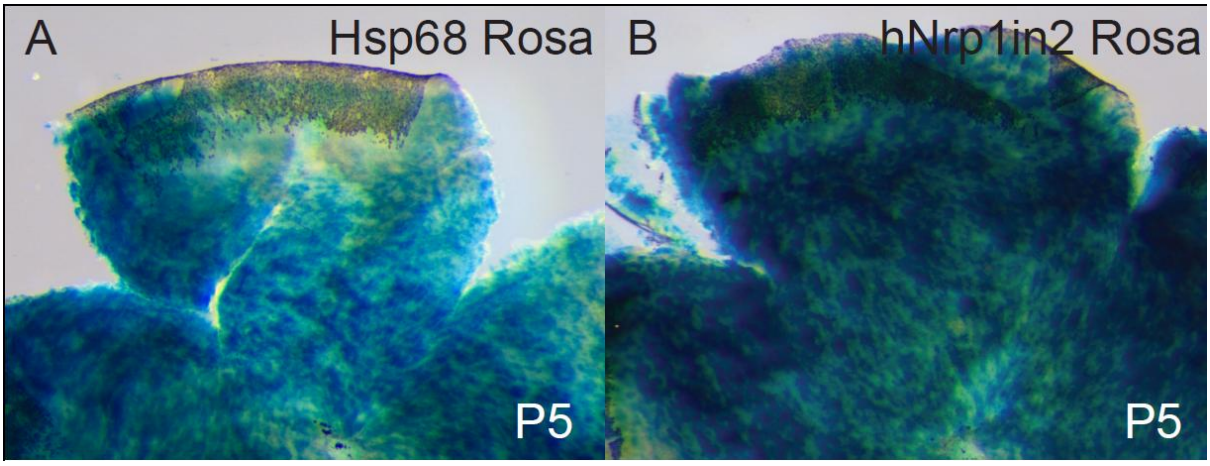


Figure 4.6 Postnatal background expression of the Rosa line.

A. Hsp68 Rosa expression at post natal day 5 (P5).

B. hNrp1in2 Rosa line retina at P5. Background expression in sample of both lines express is high and not in particular of characteristic of any vascular structure.

4.2.3. Alternative minimal promoter

As described previously, reporter gene expression driven by the Hsp68 minimal promoter when recombined at the Rosa26 locus included expression in the neural tube and heart valves during embryonic development and in a much wider spectrum of tissues in the postnatal stages (Figure 4.6). For example, extensive reporter gene expression in the neural tissue of the post-natal retina, a widely used model for angiogenesis and endothelial differentiation, made it impossible to detect any vascular-specific expression of the tested enhancer.

I therefore decided to look for an alternative minimal promoter that would, by itself, direct no or a non-interfering expression pattern. In addition to Hsp68, the other promoter often used in enhancer assays (particularly in the luciferase assay), and generally deemed silent in normal mammalian settings, is the 62bp minimal region of the HSV thymidine kinase promoter *(215)*. To determine whether this promoter would be more appropriate and more silent for our assays, I established two lines using this thymidine kinase minimal promoter (termed mTK) fused to the *LacZ* reporter gene integrated into the Rosa26 locus, with or without the hNrp1in2 enhancer, respectively. I then examined and compared the expression patterns of the reporter gene in these two lines. The mTK promoter alone appeared to drive much less background expression in the embryonic developmental stages, and serve as good minimal promoter for the tested enhancer (Figure 4.7 A1, B1, B1'). However, in the post-natal stage in the retina, mTK again directed expression widely; and the hNrp1in2mTK line also expressed beyond the vasculature too, although the background expression was not of characteristic vascular structure (Figure 4.7 B2, B2').This makes it impossible to study the vascular

specificity of the enhancer, and the lines are also therefore not useful for retinal model analysis.

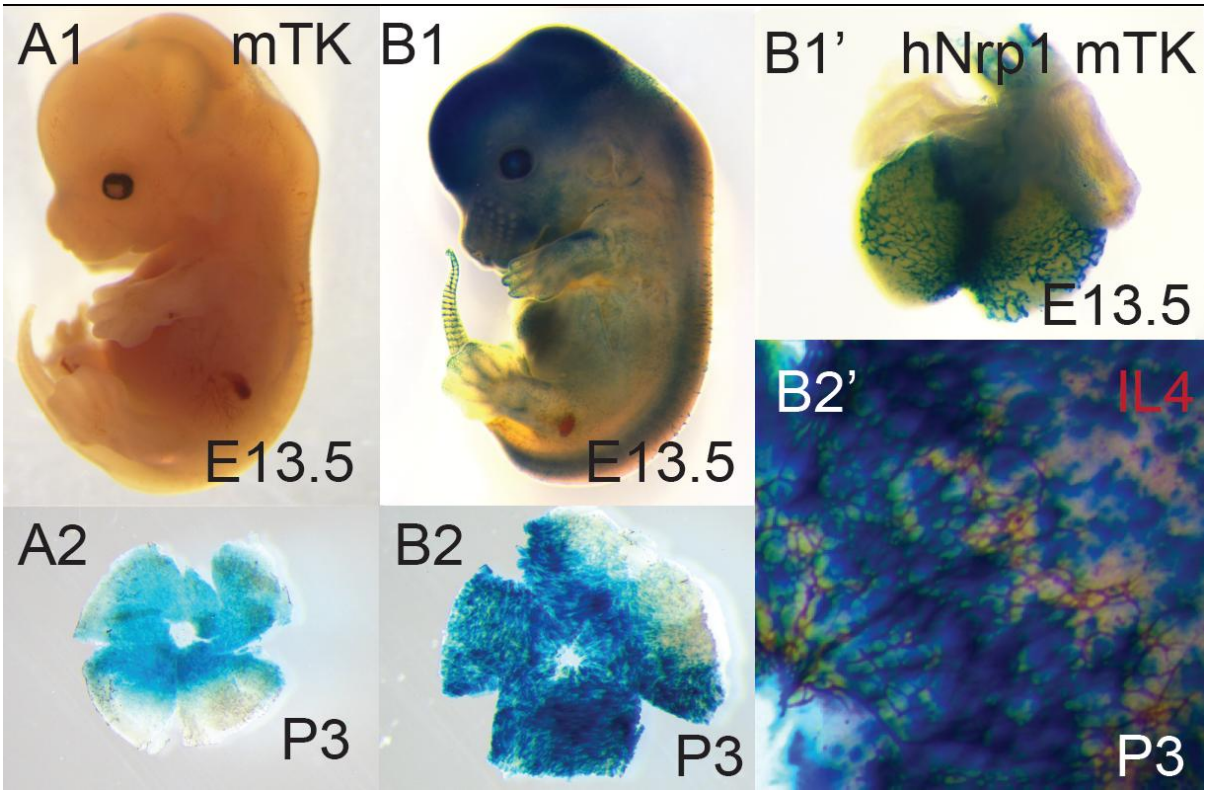


Figure 4.7 Alternative promoter miniTK has a strong background expression post-natally.

A. mTK*LacZ* E13.5 embryo and P3 retina X-Gal stained. A1, the E13.5 embryo has almost no background expression. A2, the P3 retinal has a wide and non-characteristic background expression.

B. hNrp1in2 mTK*LacZ* embryo and retina. B1, the E13.5 embryo demonstrate similar expression pattern as in hNrp1in2 Rosa. B1', coronary expression at the same time point. B2, the Postnatal expression in the retina is a more intense, yet still non-characteristic expression. B2', Isolectin4 IHC of P3 retina labels the vascular plexus.

4.3. Conclusion

I have established a novel method with which to study the expression pattern of enhancers at a single location. Moreover, this technique allows us to generate a single-copy targeted integration of the transgene that is free from the difficulties of interpretation usually associated with random integration, resulting from dosage effects due to transgene copy number and the impact of random integration. As the data shown here demonstrates, multi-copy pronuclear injection transgenic lines are inconsistent in expression patterns, with different lines of non-targeted transgenesis hNrp1in2 LacZ pan-vascular in E9.5, rather than showing arterial specificity as in the other lines. This variation between lines caused confusion and made them difficult to interpret, particularly as changes in signal intensity could not easily be separated from changes in arterial-specification. The lines from the R26/PhiC31 system on the other hand are, by definition, consistent so the expression pattern does not vary between litters, generations or lines. This is of particular value when studying mutant versions of an enhancer, as any change in expression is due to the mutation made, and not to any change in transgene insertion events.

Arterial-specific expression in the endothelial cells directed by the hNrp1Rosa enhancer was established by E9.5 in mice, and maintained throughout embryonic development. Arterial specificity was demonstrated in the dorsal aorta, limb vessels, tail arteries, and intercostal arteries. In the developing heart, expression was detected in the primordial septum, coronary artery and coronary vein, as well as in the valves. The fact that the hNrp1in2Rosaline demonstrate an arterial- and coronary-specific expression pattern, recapitulating most of the

hNrp1in2-LacZ lines as well as the vascular expression of endogenous Nrp1, suggests that the enhancer is responsible at least in part for the transcriptional regulation of the gene in this context.

However, although this targeted approach had many advantages, a major disadvantage was reporter gene expression directed by theHsp68 minimal promoter. This promoter has been long accepted by the transcription research community as a silent promoter during embryonic development, and has been widely adopted for the characterisation of transcription regulation elements *(221).* Even though there have been reports indicating that the promoter drives detectable expression, this problem has been largely ignored, especially when non-neural tissues have been studied. Unfortunately, in this experimental setting, Hsp68-directed expression significantly affected our interpretation of enhancer function. Expression from this minimal promoter occurs in the neural tube soon after gastrulation (E8.5 ~ 9.5) and continues at least until the neonatal stages. In the heart, Hsp68 also directs expression in various places, especially in both the tricuspid valve and the mitral valves. This is not surprising, since some heat shock proteins have been suggested to play roles in mesodermal lineage migration, contributing to the cardiac cushions and, in turn, the heart valve's cellular makeup *(222).* Such migrating cells can be morphologically confused with the migrating sinus venosus endothelial cells in the neighbouring regions. It is therefore difficult to study the expression of an enhancer upstream of Hsp68 in the early phase of cardiac organogenesis without the risk of contaminating Hsp68-driven expression. Furthermore, HspRosa transgene expression in the retina was not restricted to endothelial cells, and it was therefore not possible to study the expression of the Nrp1 enhancer in angiogenesis. Therefore, the relatively restricted

expression of this Hsp68 minimal promoter, which might have been considered negligible in other scenarios, compromises the use of this targeting strategy in coronary vascular expression analysis at later developmental time points, or vascular analysis in postnatal models.

Notably, the background expression of Hsp68 observed in this study was much stronger than that previously reported or experienced in our own laboratory in multi-copy pronuclear analysis, suggesting that part of the problem may be the extremely permissive ROSA26 site (*214*), especially in postnatal stages. HspRosa drives expression throughout the whole body, regardless of the transgenic regulatory elements involved. Knowing this, insulator elements were added to our construct, but they may not have been strong enough later in development.

Possible ways to circumvent the problems caused by permissive Hsp68 expression were explored by using an alternative silent promoter. The minimal promoter of the thymidine kinase gene has been widely used to combine with other regulatory elements for reporter assays in tissue culture (*215*). Its intrinsic activity is very low and has been considered silent on its own. Embryonic expression of the promoter is indeed much lower, and restricted to far fewer tissues than the HSP68. However, as for the retinal model, the Nrp1-mTK expression is still generic rather than vascular or endothelial specific. In fact the mTK expression in the retina is also as generic; it rendered the model not superior to the previously established Hsp68 system.

An obvious way to solve such a dilemma would be to utilise the endogenous Nrp1 promoter in place of an artificial minimal promoter. A successful use of such a strategy has been

performed by our collaborator Ben Davies with the Tie2 enhancer. The Tie2 promoter-enhancer combination was known to direct vascular-specific expression in traditional multi-copy transgenesis. Ben Davies replaced the Nrp1enhancer-Hsp68 promoter part of our construct with the Tie2enhancer-promoter construct, and recapitulated the exact vascular specific expression pattern, although admittedly at a lower expression level (data not shown). However, using the endogenous promoter for enhancer analysis would, in this case, prove challenging. Although the classical definition of a core promoter states that it directs spatial- and temporal-specific expression (*216*), it would be difficult to prove that a specific expression pattern was only directed by a particular regulatory element motif in the studied enhancer, since it could be cooperating or overlapping with those in the promoter. Indeed, the area around the Nrp1 promoter does contain conserved ETS recognition sites as well as sites of SP1, AP1 and STAT, and these sites can influence transactivation *in vitro* (reporter assay (*216*)). Since the expression specificity of the promoter and the binding motifs contained within it has not been determined in an embryonic developmental context, it remains a major caveat for any interpretation in enhancer characterisation analysis.

A further, though minor, issue is that a single copy of the transgene appears to drive rather modest expression. While this allows the *Rosa* lines to avoid the confusion of expression pattern variation that comes with intensity changes, at the same time the weak expression of the reporter gene means that it cannot be easily detected by X-Gal at all developmental stages, and it is a major challenge to detect β-Gal with an antibody.

Lastly, a major problem with the targeted *Rosa* lines became the amount of time that it took to make them. Obviously, the F0 generation of embryos are mosaic, since they are a mixture of host blastocyst and transgenic mES cells. More useful research subjects would be at least the F1 generation embryos, if the transgenic mES cells entered the germ line. To establish such a line for multiple time points and allow continuous breeding, F1 embryos tend to be screened for breeders to make the F2 embryos the major supply of research subjects. The time span of such an experiment stretches to 4-6 months. Where the unfortunate scenario of F0 chimera formation with no germ line integration occurs, it can take 3-4 months to discover the situation and transgenic mES cells would have to be re-established from scratch. Compared to the 2-3 week time line of the transient transgenesis assay, any progress and conclusions made through the Rosa26 lines would take significantly longer.

In conclusion, although this method has many advantages, the disadvantages sometimes outweighed them. In particular, the amount of time taken to make each target line, and problems with 'leaky' expression patterns, meant that both the single copy approach and multiple transient transgenic embryos had to be used for the further analysis of Nrp1in2. When the single copy targeted approach was used, only early time points between E11-E13.5 were studied to avoid the problems caused by leaky Hsp68 promoter expression. For this reason, transient transgenic mice made using multi-copy pronuclear injected transgenesis were also examined at E13.5.

5. Characterising the key elements within the Nrp1 intronic enhancer

5.1. Conservation of the Nrp1in2 enhancer

Although conservation between species is not always a feature of functional enhancers, it can often be used as a mechanism to identify key regulatory regions within larger enhancer elements. Consequently, the Nrp1in2 enhancer sequences found in a variety of different species were compared. These included the previously validated human sequence (*Homo sapiens*, Genome Reference Consortium Human Build 37, 2009) and the widely used, well characterised and fellow *placentalia* mammal, mouse (*Musmusculus*, NCBI37/mm9, 2007). Additionally, the opossum sequence (*Monodelphisdomestica*, Broad prelim/monDom1, 2004) was chosen to represent *marsupials* on a different infraclass of the mammalian evolutionary tree. Conservation between the infraclasses would indicate a common origin of the enhancer among mammals, and a shared regulatory mechanism. Avian models have also been widely used in vascular development research, and the molecular networks of vascular development in the chick (*Gallus gallus*), especially those regulating transcription, are often comparable to mammals. An enhancer with shared sequence conservation in such species would suggest that the elements within played a central role in a conserved transcription regulation process in the vasculature. Chicken WUGSC 1.0/galGal2, 2004 was used for this study.

First, an alignment analysis was conducted using the mVista alignment tool and visualised in the Vista browser. Conservation of each species sequence against that of human Nrp1in2 is depicted below (Figure 5.1).

Surprisingly, among the animals chosen for the conservation analysis, the fellow *placentalia* (mouse) shared the least alignment (Figure 5.1).

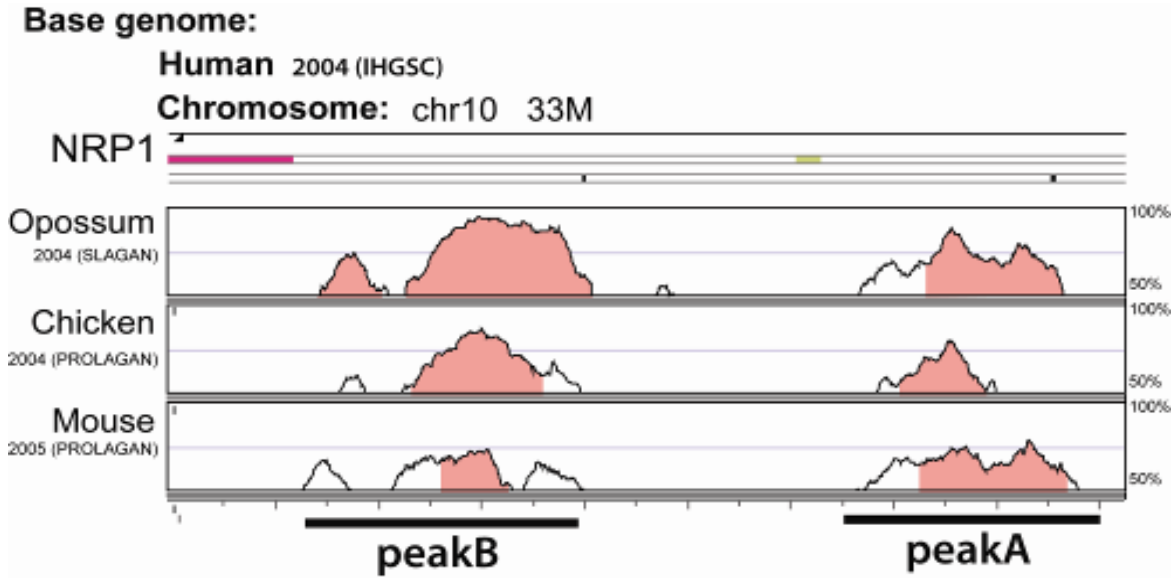


Figure 5.1 Conservation of the human Nrp1 enhancer in its orthologous sequences.

Homologous sequences of the enhancer from opossum, chicken and mouse were drawn from UCSC by BLAT alignment to the genomic sequence of the most recent assembly of each species. Conservation is depicted in the diagram along the fragment; high scores of continuous conservation are labelled in pink. Conserved regions of 510 bp (peakA) and 538 bp (peakB) were identified with deep conservation. Alignment visualization was adopted from analysis with the 2004 release of data on the VISTA server.

As the least conserved Nrp1in2 enhancer, I investigated whether the homologous mouse Nrp1 sequences could function in an EC-specific manner, similar to that seen with the human Nrp1in2 (hNrp1in2) enhancer. First, we looked at the orthologous mouse sequence for epigenetic markers (chr8:130911050-130912307) which is also in the second intron of the Nrp1 gene. Investigation of the ENCODE database for DNaseI hypersensitive sites (DHS) around this sequence revealed high levels of DHS in samples from heart, the myeloid cell line 416B, and the kidney, which is a highly vascularised organ (Figure 5.2, the maroon, the blue and the pink tracks). Fibroblasts did not have any DHS peak in this region, suggesting that DNaseI hypersensitivity in this sequence is tissue-specific rather than ubiquitous in the whole embryonic body (Figure 5.2 black tracks). Using the Ludwig Histone track, whether the mouse sequence was enriched for enhancer-associated histone marks was also investigated. H3K4Me1 and H3K27ac signals were found in samples from both the heart and liver at embryonic day 14.5 (E14.5) and postnatal (8 weeks) stages, suggesting that this sequence is actively involved in transcription activation in vascular development (Figure 5.2, maroon and yellow tracks). When the enhancer was investigated in more detail, the epigenetic marks of the enhancer demonstrated unequal distribution across the entire fragment (Figure 5.3). Interestingly, the H3K4 methylation marks and the H3K27 acetylation mark enrichments appeared to coincide with the mammalian conservation score peaks, and also with the relevant conservation peaks identified among human, mouse, opossum and chick. In particular, the sequence marked “PeakA” was also marked with an enriched DNS cluster. This is strongly consistent with what we had previously discovered in the human sequence (Chapter III, Figure 3.3).

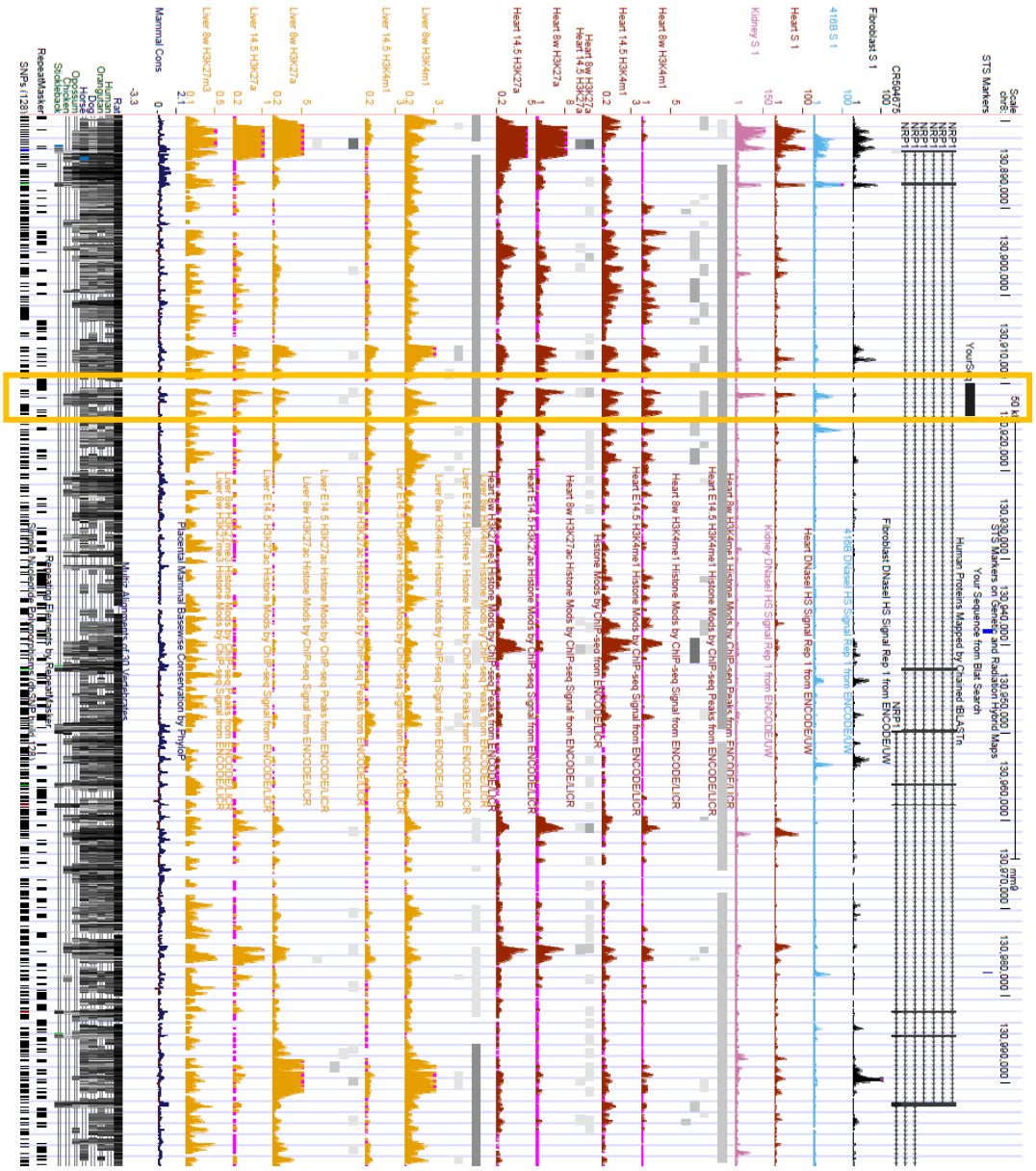


Figure 5.2 Schematic representation of the mouse Nrp1 locus

The mouse Nrp1 locus is presented in the UCSC ENCODE Browser (*201*). DHS recognized in fibroblasts, myeloid cell line 416B, embryonic heart and kidney are shown in the DHS histogram tracks. Embryonic heart and liver specific H3K4me1 and H3K27ac marks are labeled in maroon and orange, respectively. The vertebrate conservation score (dark blue histogram) and conservation to other animals (black clusters) are presented at the bottom. The mNrp1in2 enhancer is boxed in light orange.

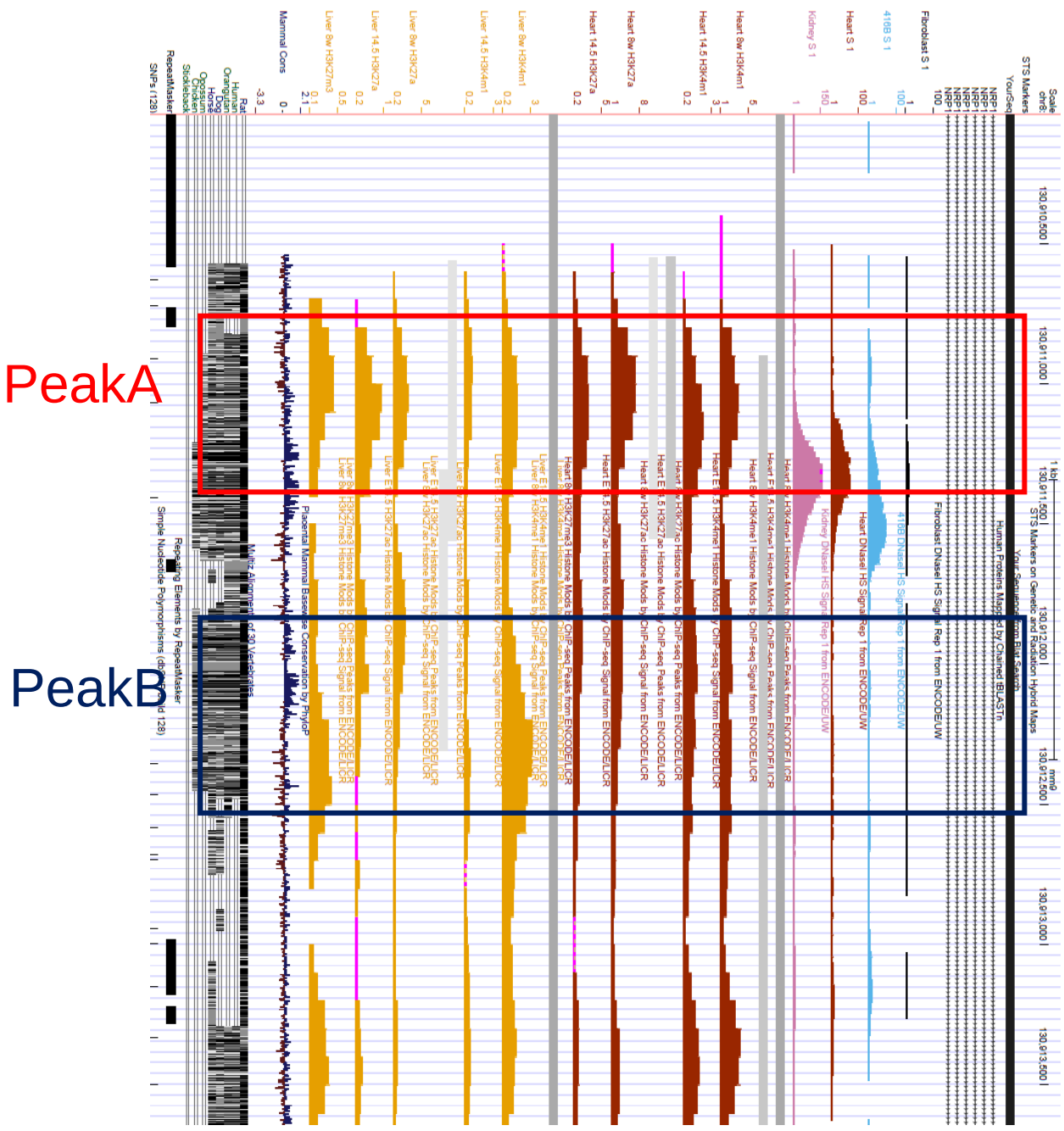


Figure 5.3 Schematic representation of the mNrp1in2 fragment

Figure 5.3 Schematic representation of the mNrp1in2 fragment

A more focused representation of the human sequence surrounding the mNrp1in2 enhancer in the UCSC ENCODE Browser (*201*). The layout and colour coding used is the same as in Figure 5.2. Notice that while conservation spans the whole fragment, there is deeper and more continuous conservation in two roughly 500bp regions: the H3K27ac and H3K4me1 markers in embryonic heart and kidney overlap the two conservation peaks. The 416B/heart/kidney specific DHS preferentially overlaps one of these conserved regions, which is orthologous to the peakA region (as marked in red box; note that the mouse sequence orientation is the opposite to that of the human).

5.2. The mouse Nrp1in2 enhancer drives similar expression patterns to the human version

If the conserved regions between the human and mouse sequences of Nrp1in2represented functional motifs, it would be expected that the expression pattern directed by the mouse sequence and human Nrp1in2 enhancers would be similar. This would then justify phylogenetic foot printing analysis between the two as an efficient strategy to identify *cis* regulatory elements in the enhancer.

To confirm that the mouse Nrp1in2 enhancer was similar in function to hNrp1in2, we generated two mouse models. First, we used the standard pro-nuclear injection method of transgenic mouse analysis, using a construct consisting of the mNrp1in2 fragment with the Hsp68 minimal promoter and the *LacZ* reporter gene, similar to the way the hNrp1in2 enhancer was initially examined (Figure 5.4 A). These transgenic mice, labelled mNrp1in2*LacZ*, expressed the *LacZ* reporter gene in arterial and coronary endothelial cells, in a manner similar to the hNrp1in2*LacZ* expression pattern (Figure 5.4 B). Out of the eight transgenic mNrp1in2*LacZ* E13.5 embryos, six were confirmed with vascular expression and 4 are of arterial-specific patterns of β-Gal (Figure 5.4, B1, B2; Table 5.1 on page 151). Only one expressed β-Gal in a subset of periphery veins, but not in major veins such as the cardinal vein. In the heart, all six mNrp1in2*LacZ* hearts expressed the reporter gene in the developing coronary plexus (Figure 5.4 C), as well as in the inter-ventricular groove and valves, a similar expression pattern to that of hNrp1in2*LacZ* (Figure 5.4 C1). This data suggests that

mNrp1in2 drives a very similar expression profile to the human orthologous enhancer in the vasculature, in an arterial- and coronary-specific fashion.

To confirm this observation, a transgenic mouse line was generated that carried the Hsp68-mNrp1in2-*LacZ* transgene at the Rosa26 locus (mNrp1Rosa). This line was generated using the same techniques used to make the hNrp1Rosa model, permitting the analysis of the expression profile of one single copy of the mNrp1in2 enhancer in a targeted transgenic manner, and allowing the efficient examination of the expression profile in multiple embryonic stages.

As shown in Figure 5.5, the expression of mNrp1in2Rosa was comparable to that of hNrp1in2Rosa. The transgenic embryos all had vascular expression at E9.5, E11.5 and E13.5 (Figure 5.5 A). The dorsal aorta of mNrp1in2 was positive for X-Gal staining, while the cardinal vein was negative at both E9.5 and E11.5. At E13.5, the tail artery was positive, whereas the tail vein was negative (Figure 5.5 B). The arterial- specific expression of this Nrp1 enhancer was very similar to that previously observed in hNrp1Rosa, although the expression was slightly lower at later embryonic stages. To conclude, the expression of mNrp1in2 recapitulated that of hNrp1in2 in single copy transgenic mice. Therefore, although fairly divergent from hNrp1in2, it was concluded that the mNrp1in2 sequences till carries the essential information for transcription regulation directing arterial and coronary EC-specific expression. The two regions most highly conserved between mouse and human Nrp1in2 are the most likely to carry this information and, thus, were prioritised for the search for functional *cis* elements.

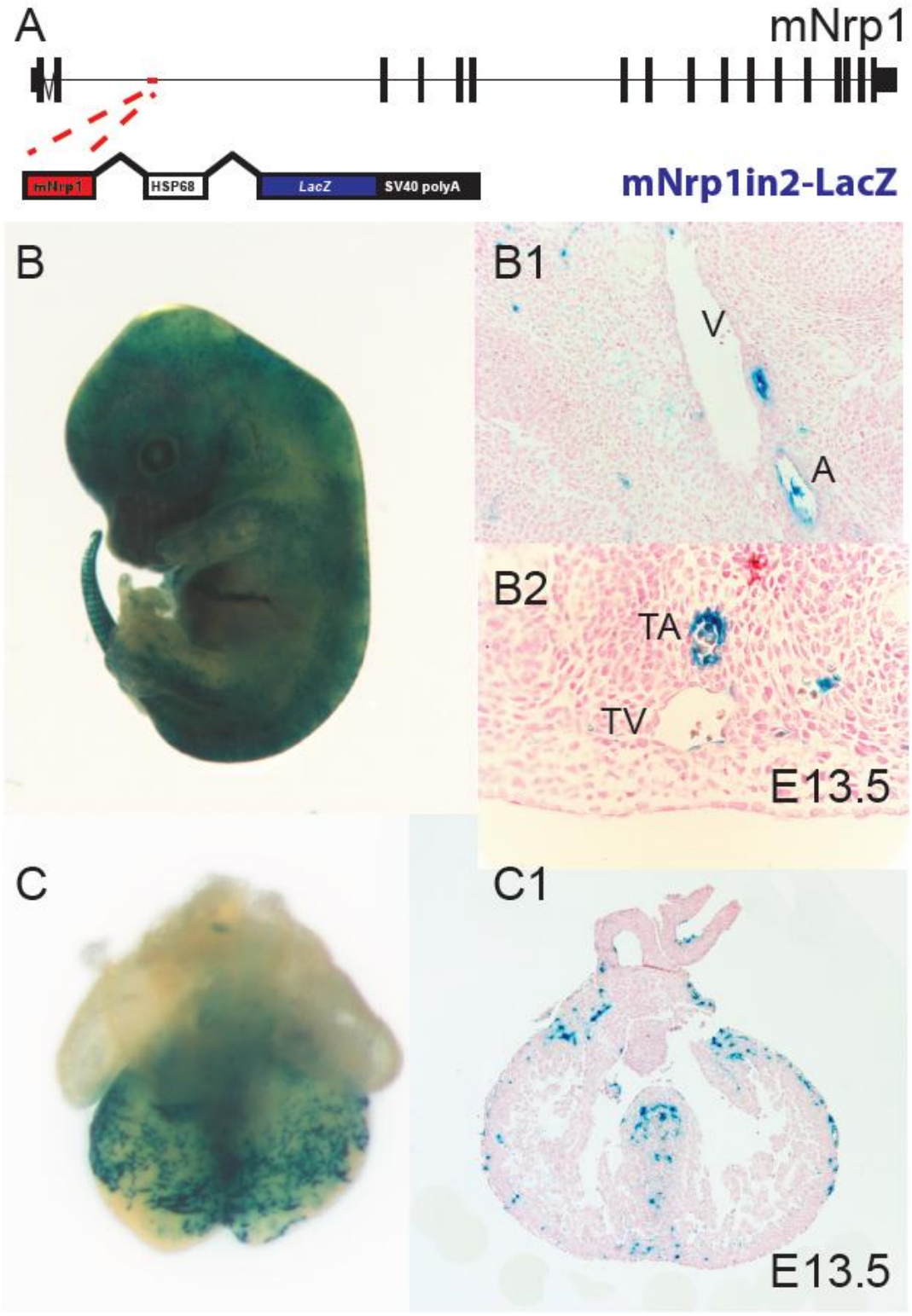


Figure 5.4 mNrp1in2*LacZ* transgenic embryos express in arterial and coronary vessels in the vasculature.

Figure 5.4 mNrp1in2*LacZ* transgenic embryos express in arterial and coronary vessels in the vasculature.

A. Schematic representation of the transgene construct. An enhancer from the second intron of the mouse Nrp1 gene was cloned in front of the silent Hsp68 minimal promoter and the reporter gene *LacZ*, forming the mNrp1in2*LacZ* transgene.

B. Arterial expression in the E13.5 embryo. B, whole mount E13.5 expression of the reporter gene β-Gal is detected by X-Gal staining. B1, transverse section of the embryo in B, illustrating that the tail artery is positive, while the tail vein is negative for the reporter gene. B2, transverse section of embryo tail.

C. Coronary expression in an E13.5 embryo. C, whole mount E13.5 heart was dissected and stained on its own. The plexus of the developing coronary is positive for X-Gal staining. C1, coronal section of the heart in C. Notice the coronary artery and vein are both positive for expression, as well as the interventricular groove.

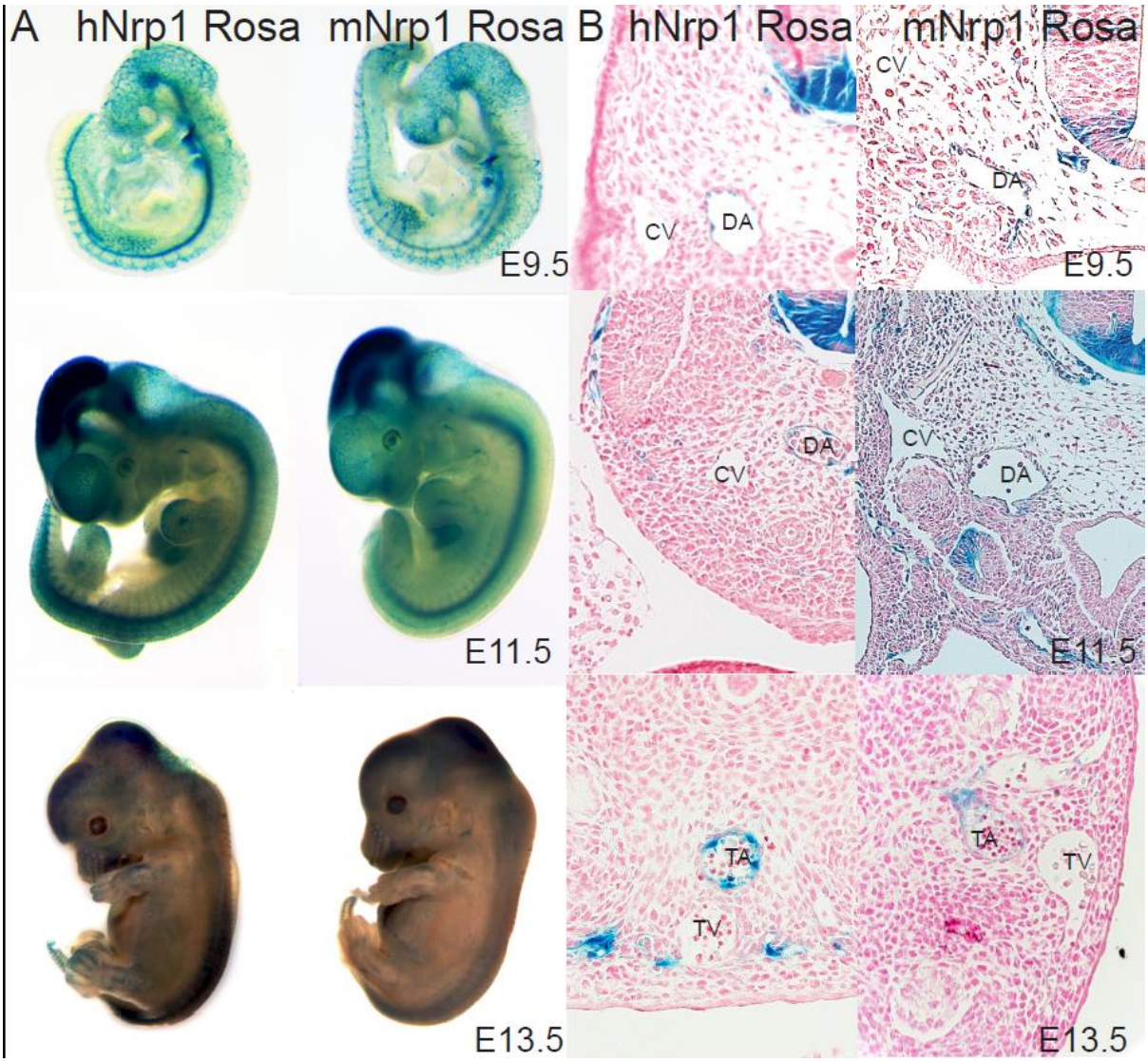


Figure 5.5 Expression profile validation of mNrp1 in a single copy line

A. Mice transgenic for mNrp1in2*LacZ* at E9.5, E11.5 and E13.5 express *LacZ* in the vasculature. The expression of mNrp1in2 (right panel) is compared to that of hNrp1in2Rosa (left panel) at the same embryonic developmental stages.

B. Mice transgenic for hNrp1in2*LacZ* at E9.5, E11.5 and E13.5 express *LacZ* arterial specifically. The transverse section of the X-Gal-stained transgenic embryo demonstrates that reporter gene expression can be detected in the arterial (BA, *branchial arch artery*), but not the venous (CV, *cardinal vein*), compartments. The reporter expression is arterial in E11.5 transgenic embryos. Similar to at E11.5, the *dorsal aorta* (DA) is positive whereas the *cardinal vein* (CV) is not. Lower row, tail arterial expression in the E13.5 embryo whereas there is no expression in the tail vein, as in hNrp1in2Rosa.

5.3. Identification of a small region within hNrp1in2 that is sufficient to drive arterial and coronary endothelial-specific expression

5.3.1. Investigation of the Peak A and Peak B regions of hNrp1in2

To validate the association between conservation of sequence and functional importance, and to further focus on a smaller region, the functionality of the two conserved peaks of hNrp1in2was tested separately in transgenic mice. First, hNrp1in2peakB was analysed, with the sequence covering the PeakB conserved region fused to the Hsp68 minimal reporter and the *LacZ* reporter gene using transient transgenesis. E13.5 transgenic mouse embryos were analysed for reporter expression using X-Gal staining. However, none of the three transgenic embryos showed expression in the vasculature (Figure 5.6 A). Other versions of the Nrp1in2 enhancer showed 70-100% expression (Table 5.1, Page 151). It is of note that the apex of transgenic hearts demonstrated blood island-like structures that were positive for reporter gene expression (Figure 5.6 B, C); but this region is discontinuous from the sinus venosus which is anterior to the ventricle, and similar expression was not seen in the full hNrp1in2 enhancer. Something similar was also seen in the Hsp68Rosa single copy mice, suggesting that the Hsp68 minimal promoter may be responsible for this expression. In conclusion, it appears that the conserved PeakB region is not capable of arterial and coronary vascular expression in a manner comparable to the hNrp1in2 enhancer.

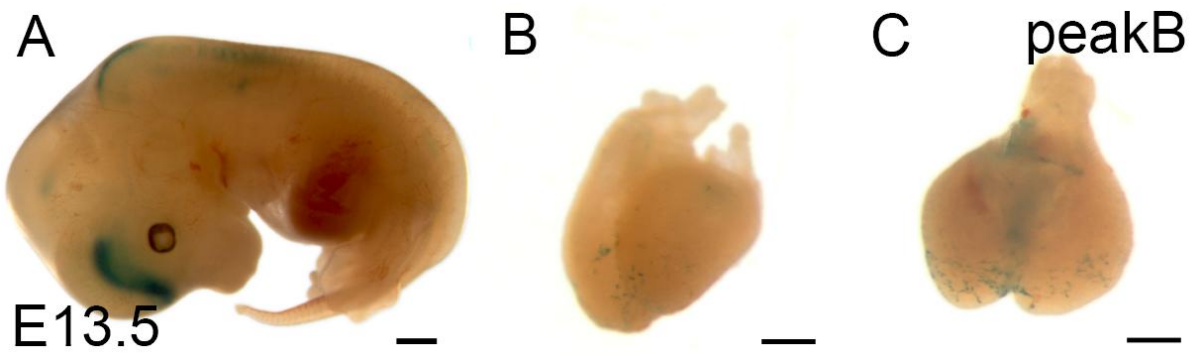


Figure 5.6 hNrp1in2peakB*LacZ* lacks peripheral vascular expression

A. Whole mount PeakB*LacZ* transgenic embryo. Expression is not seen in the vasculature.

B. and C. Ventral view of the E13.5 heart. The stereotypical spider net downwards-developing plexus is not positive for X-Gal staining. However, there is blood island-like staining in the apex of the heart. n=3. Scale bar, A, 2mm; B and C, 1mm.

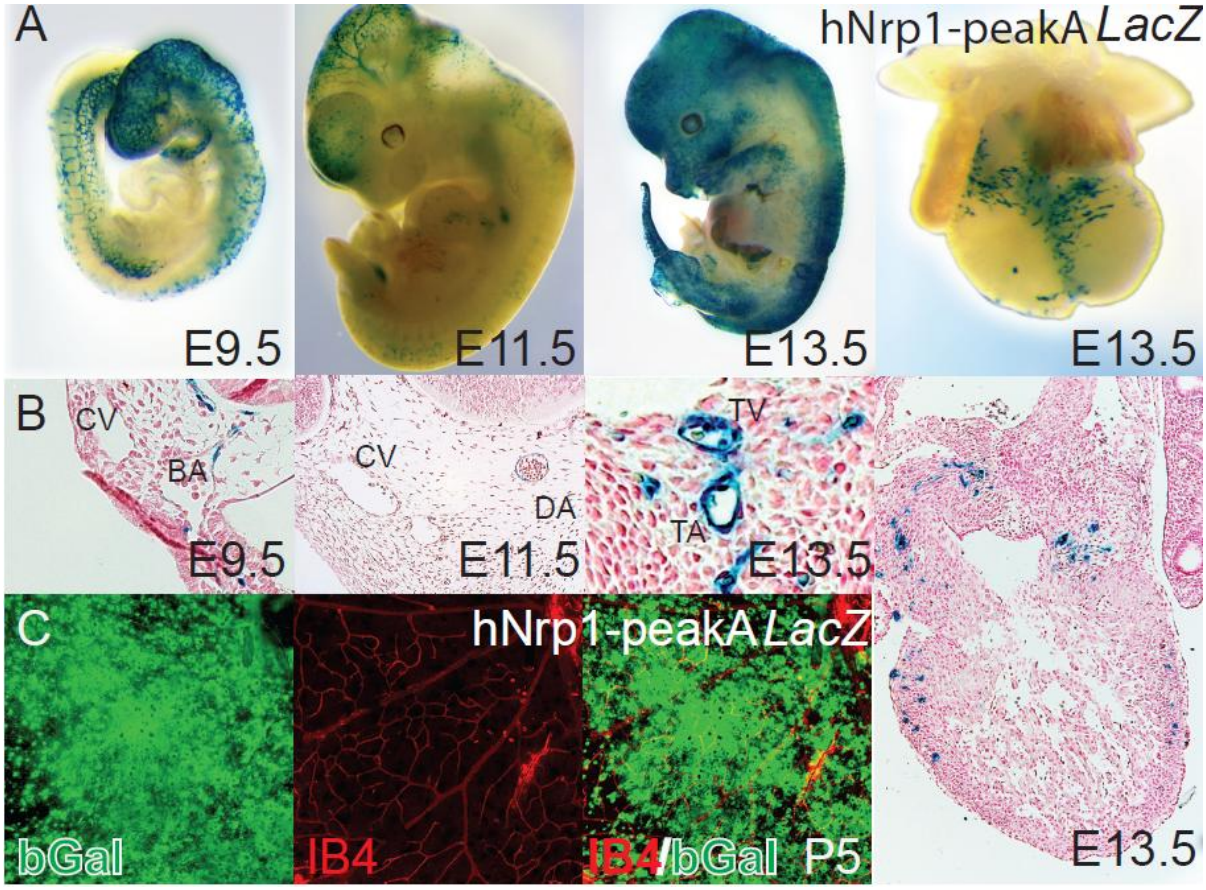


Figure 5.7 hNrp1in2peakA*LacZ* embryonic expression in the vasculature

A. Whole mount peakA embryos at various stages of embryonic development. E9.5, E11.5 and E13.5 embryos are harvested and analysed for reporter gene expression detected by X-Gal staining. The E13.5 heart is dissected and stained separately.

B. Transverse section of the X-Gal-stained embryo, counter-stained with nuclear fast red. E9.5, the dorsal aorta and the arterial branchial arch of the E9.5 embryo are positive for the stain, whereas the cardinal veins are not. E11.5, the dorsal aorta expresses and the cardinal veins do not. E13.5, both the tail arteries and the tail veins are positive. Coronary expression is in both the arterial and the venous layers.

C. the P5 retina of the Nrp1in2peakA LacZ does not have an endothelial-specific expression pattern. Green, X-Gal staining pseudo-colouring; Red, Isolectin4 immunofluorescence labelling of all vessels. Right panel, a merge of the two stainings.

The second region of conservation, peak A, was next investigated. For this, transient transgenic E13.5 embryos were first generated using an hNrp1in2peakA Hsp68 with *LacZ* construct with pronuclear injection (Figure 5.7). Although the results were variable, in total, 4 out of 7 transgenic embryos demonstrated arterial specificity at E13.5, similar to that seen with the hNrp1in2 full length enhancer (as summarised in Table 5.1). The variability in expression was a concern, so a line carrying the Hsp68-hNrp1in2peakA-*LacZ* transgene at the Rosa26 locus (PeakA Rosa) was also established, allowing direct comparison with the other single copy transgenic lines. This analysis is seen in Figure 5.7. In general, the expression of hNrp1in2peakARosa was similar to that of hNrp1in2Rosa. Transgenic embryos at E9.5, E11.5 and E13.5 all had vascular expression, although a decrease in visible vascular expression was noted at E13.5, similar to the mNrp1in2Rosa model (Figure 5.8). The dorsal aorta of hNrp1in2peakA was positive for X-Gal staining, while the cardinal vein was negative. This arterial-specific expression at E9.5 and E11.5 was maintained in later phases. In E13.5, the dorsal aorta was positive, whereas the cardinal vein was negative (Figure 5.8 B). The arterial-specific expression of this truncated enhancer was very similar to what was observed for full-length enhancer-driven expression in hNrp1Rosa, suggesting that the expression of hNrp1in2peakA largely recapitulates that of hNrp1in2 in single copy transgenic mice, and that the transcription factors that direct these specific patterns of expression are likely binding motifs found within the Nrp1 Peak A sequence.

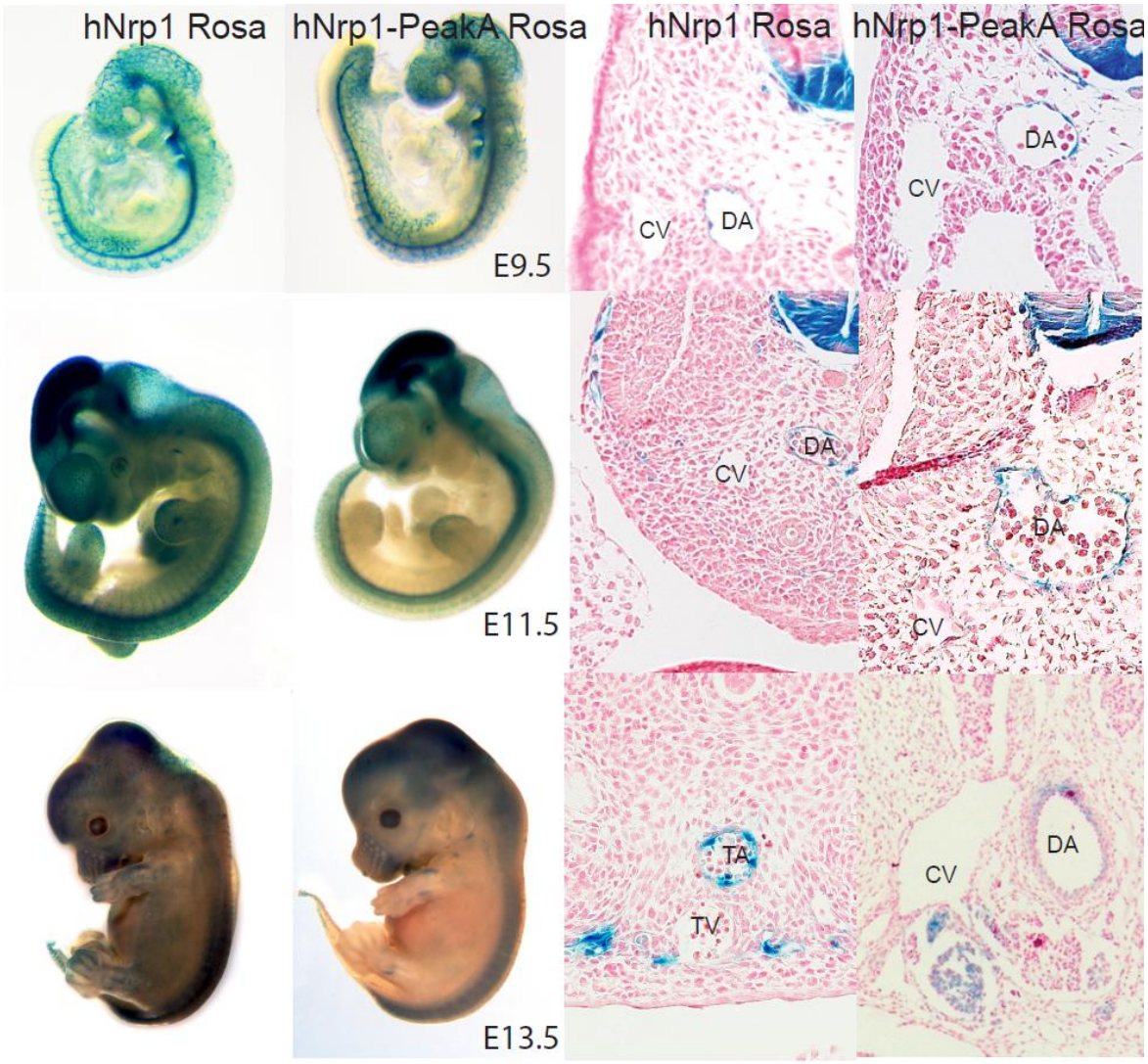


Figure 5.8 Expression profile validation of Nrp1in2peakA in a single copy transgenic line

A. Mice transgenic for hNrp1peakA*LacZ* at E9.5, E11.5 and E13.5 express *LacZ* in the vasculature. PeakA (right panel) expression is compared to the expression pattern of full-length hNrp1in2 Rosa (left panel) at the same embryonic developmental stages.

B. Mice transgenic for hNrp1peakA*LacZ* at E9.5, E11.5 and E13.5 express *LacZ* arterial specifically. The transverse section of the X-Gal-stained transgenic embryo demonstrates that reporter gene expression can be detected in the arterial (BA, *branchial arch artery*), but not the venous (CV, *cardinal vein*), compartments. Reporter expression is arterial (DA, *dorsal aorta*) in E11.5 transgenic embryos when the venous *cardinal vein* (CV) is not positive. Lower row, dorsal aorta expression in the E13.5 embryo whereas cardinal veins do not show expression.

In addition to arterial and coronary vessel models, angiogenic sprouting was also examined. The ideal model for this is the post-natal retina. However, as discussed in Chapter 4, the single copy Hsp68Rosa model is not suitable for this purpose as the Hsp68-directed expression switches on the reporter gene in all tissues in the retina, making analysis impossible. Consequently, a multi-copy transgenic line was generated using pronuclear injection. Unfortunately, as demonstrated in Figure 5.9, the reporter gene expression intensity in this line was quite low. However, it indeed directs arterial-specific expression similar to that seen in the other mice at earliest time points (E11.5, Figure 5.9 A), so the expression in the P5 retina was investigated. The postnatal retina is an established model of angiogenic sprouting (*198*), and at P5 the sprouting plexus can be stained with the EC marker Isolectin B4 so that the developing vasculature can be clearly visualized. However, even though the vasculature was clearly detected, the X-Gal staining was not specific to the vascular compartment (Figure 5.9 C). It is possible that the hNrpin2peakA enhancer does not drive expression in this region. Conversely, it may be that the Hsp68 minimal promoter was also the cause of this problem, directing X-Gal expression elsewhere in the retina, similar to that seen in the Hsp68Rosa single copy mouse line.

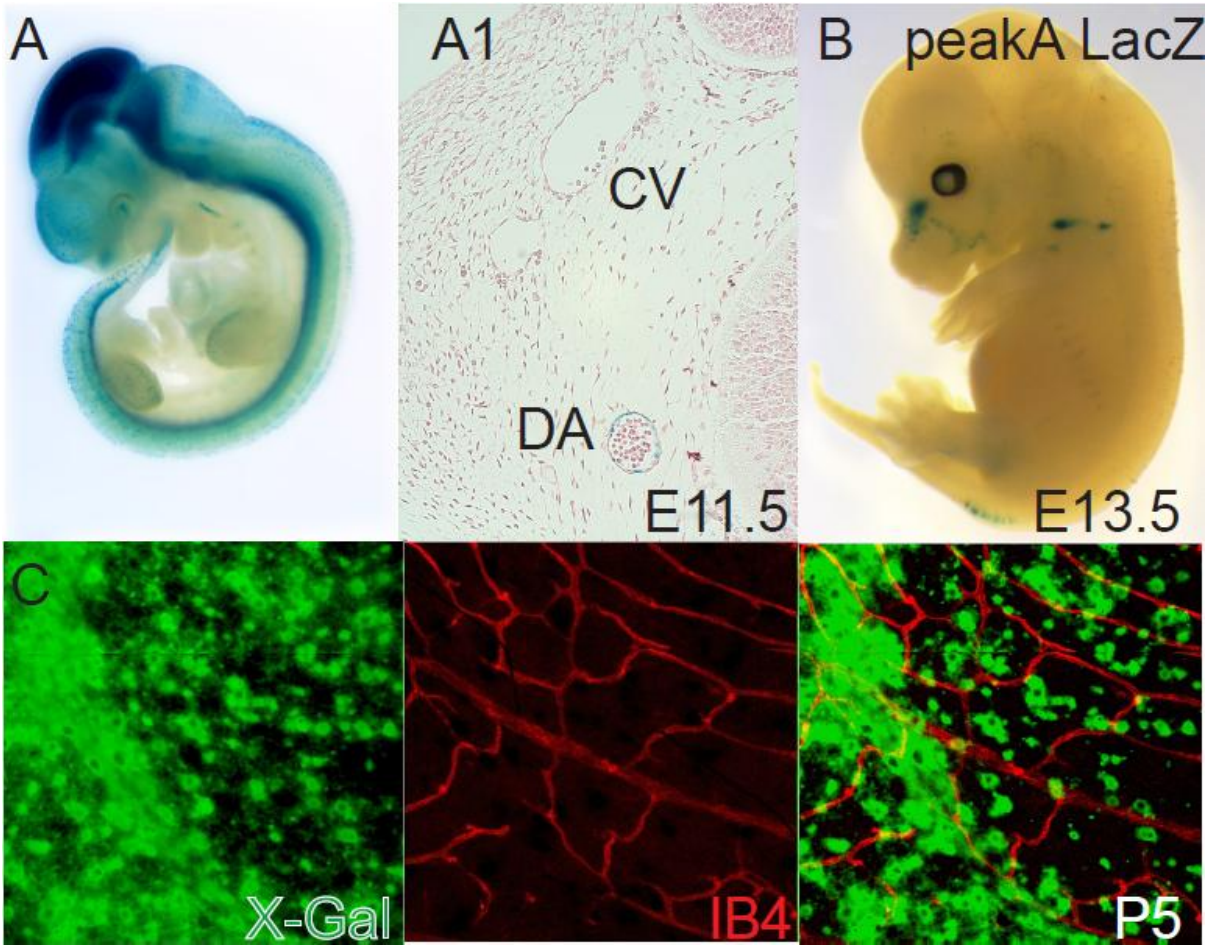


Figure 5.9 hNrp1in2peakA *LacZ* line directs arterial-specific expression in early embryonic development

A. E11.5 whole mount transgenic embryo of hNrp1in2peakA *LacZ*. A1, transverse section of the E11.5 embryo, showing dorsal aorta (DA) positive for X-Gal staining whereas the cardinal vein negative.

B. E13.5 whole mount embryo show the staining intensity is low.

C. IsolectionB4 (IB4, red) IF and X-Gal pseudo-colouring of P5 retina. Merged image of the two staining show little if not no overlaps.

Another way to look at arterial-venous differentiation and angiogenic sprouting is using the zebrafish. Having a transparent body at embryonic stages during vascular development, zebrafish serve as a great model for real-life tracing of vascular-specific expression using fluorescent reporter transient transgenesis. Our lab often analyses mouse and human enhancers in zebrafish by fusing them to the E1b minimal promoter and GFP reporter gene, and surrounding this construct with Tol2 elements. When injected into fertilised fish embryos it generates a mosaic transient transgenic fish. Dr Natalia Sacilotto, a postdoctoral scientist in my lab, injected a construct containing hNrp1peakA-E1b-GFP into fish, following which I analysed the expression pattern at 48hpf (Figure 5.10). Although the expression was detected in some intersomitic vessels, only 32 of 98 fish injected had any reporter expression (33%, compared to the 50-60% that would be expected with efficient enhancers in our laboratory). This suggests that the hNrp1peakA enhancer does not work efficiently in zebrafish; therefore, this model could not be used to look at Nrp1 enhancer activity. This may be due to the fact that the transcription factors needed for hNrp1peakA expression are not available in zebrafish. Significantly, this enhancer is not conserved in zebrafish, BLAT the human sequence of Nrp1peakA in zebrafish genome returns with queries no longer than 29 base pairs. It is also possible that the other conserved region (peakB) is needed for vascular specific expression of the enhancer in zebrafish, although the 1.8kb length of the enhancer is longer than the TOL2 system technical range (less than 1kb). The zebrafish system is therefore abandoned from further analysis options.

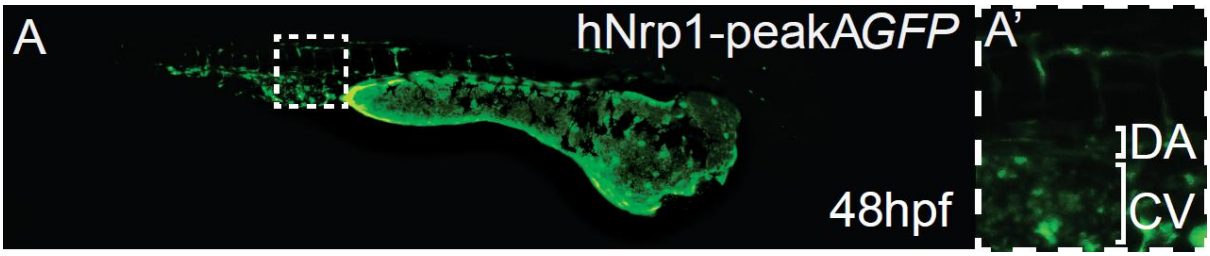


Figure 5.10 hNrp1in2peakA *GFP* transgenic zebrafish vascular expression

A. Representative E1b hNrp1peakA *GFP* zebrafish express in the vasculature. A', magnified view of the boxed region in A. Notice both the dorsal aorta and the posterior cardinal vein are positive for reporter gene expression. 22 fish had vascular expression (including DA, CV and intersomitic vessels); 8 expressed in DA, 7 in CV. n=32. DA, dorsal aorta; CV, caudal vein.

5.3.2. Identification of the key sequence within hNrp1in2 PeakA

Using the ClustalW program, the entire sequence of the hNrpin2PeakA enhancer was aligned with the sequences from mouse, chick and opossum (Figure 5.11). Since the mouse and human sequences direct similar arterial- and coronary- specific expression patterns, it was hypothesised that sequence motifs conserved between the two species would be the most likely regions to bind the transcription factors that regulate arterial- and coronary- specific expression. It would, therefore, be preferable to identify the smallest number of conserved motifs necessary for these expression patterns. Consequently, it was decided to further truncate the Nrp1in2peakA enhancer into a shorter DNA sequence termed hNrp1in2peakA2, consisting of 184bp (Fig 5.11 red text), of which 117bp was conserved with the mouse orthologue (Materials and Methods, Table 2.2). The fragment was fused to the Hsp68 minimal promoter and LacZ reporter gene and used in pronuclear injections. Embryos were investigated at both E11.5 and E13.5. At both time points, embryos demonstrated arterial-specific reporter gene expression (Figure 5.11 and Figure 5.12). The coronaries of the E13.5 heart expressed X-Gal similar to the human and mouse enhancers (Figure 5.12). In conclusion, although we did not have time to confirm this result in the single copy model, the hNrp1in2PeakA2 sequence was sufficient to drive arterial and coronary EC- specific reporter gene expression. It is therefore likely that motifs within this short sequence are able to bind transcription factors that together are able to regulate gene expression during the development of arteries and coronary vessels.

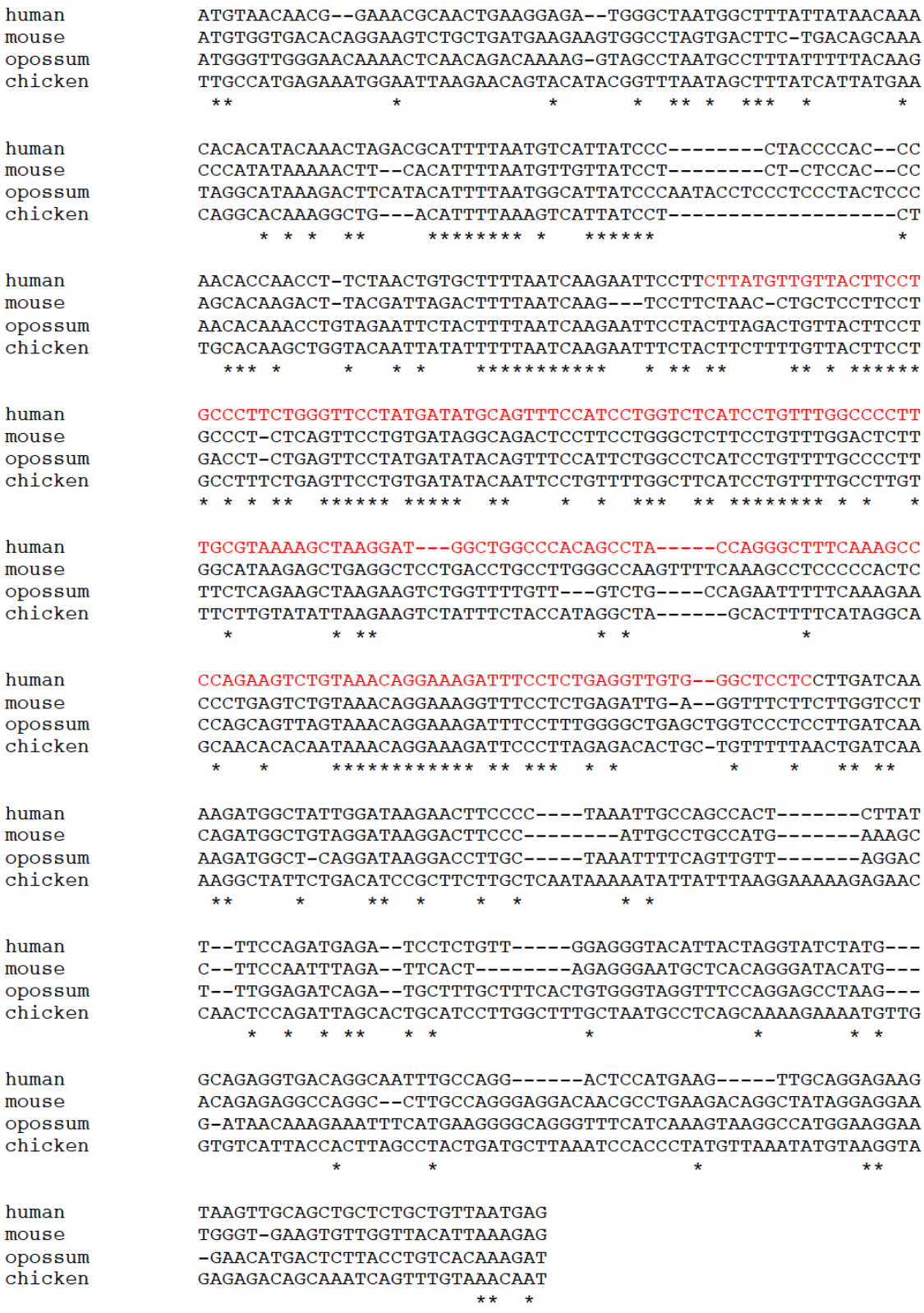


Figure 5.11 PeakA2 truncation of the hNrp1in2peakA alignment

The red lettered human sequence is the super conserved peakA2 region of the hNrp1in2 peakA enhancer.

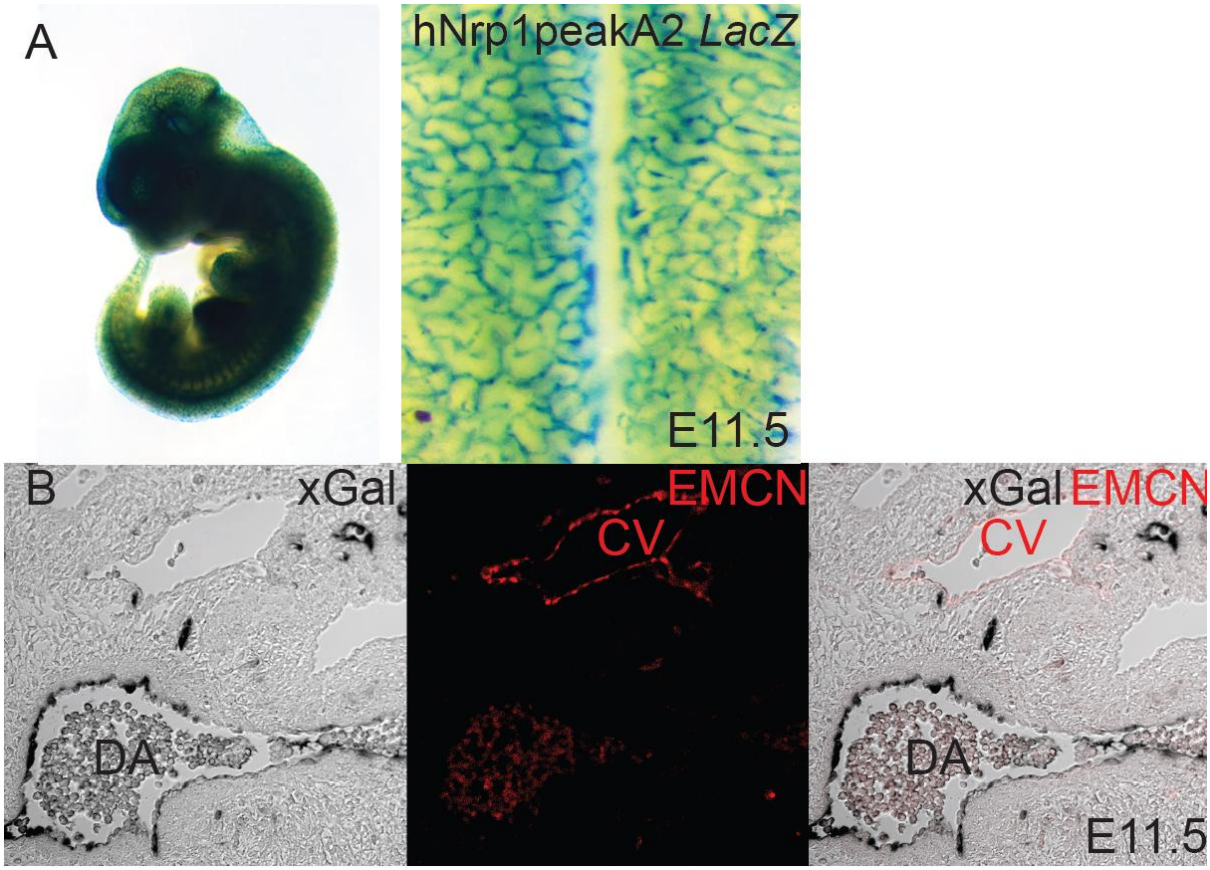


Figure 5.12 hNrp1in2peakA2*LacZ* vascular and arterial expression at E11.5

A. PeakA2*LacZ* whole-mount embryo demonstrates vascular expression. Right, hindbrain dissection of peakA2 *LacZ* demonstrates vascular expression.

B. Arterial-specific expression of reporter gene in hNrp1in2peakA2. The dorsal aorta is positive for X-Gal staining, whereas the cardinal vein (EMCN, red) is negative. n=4.

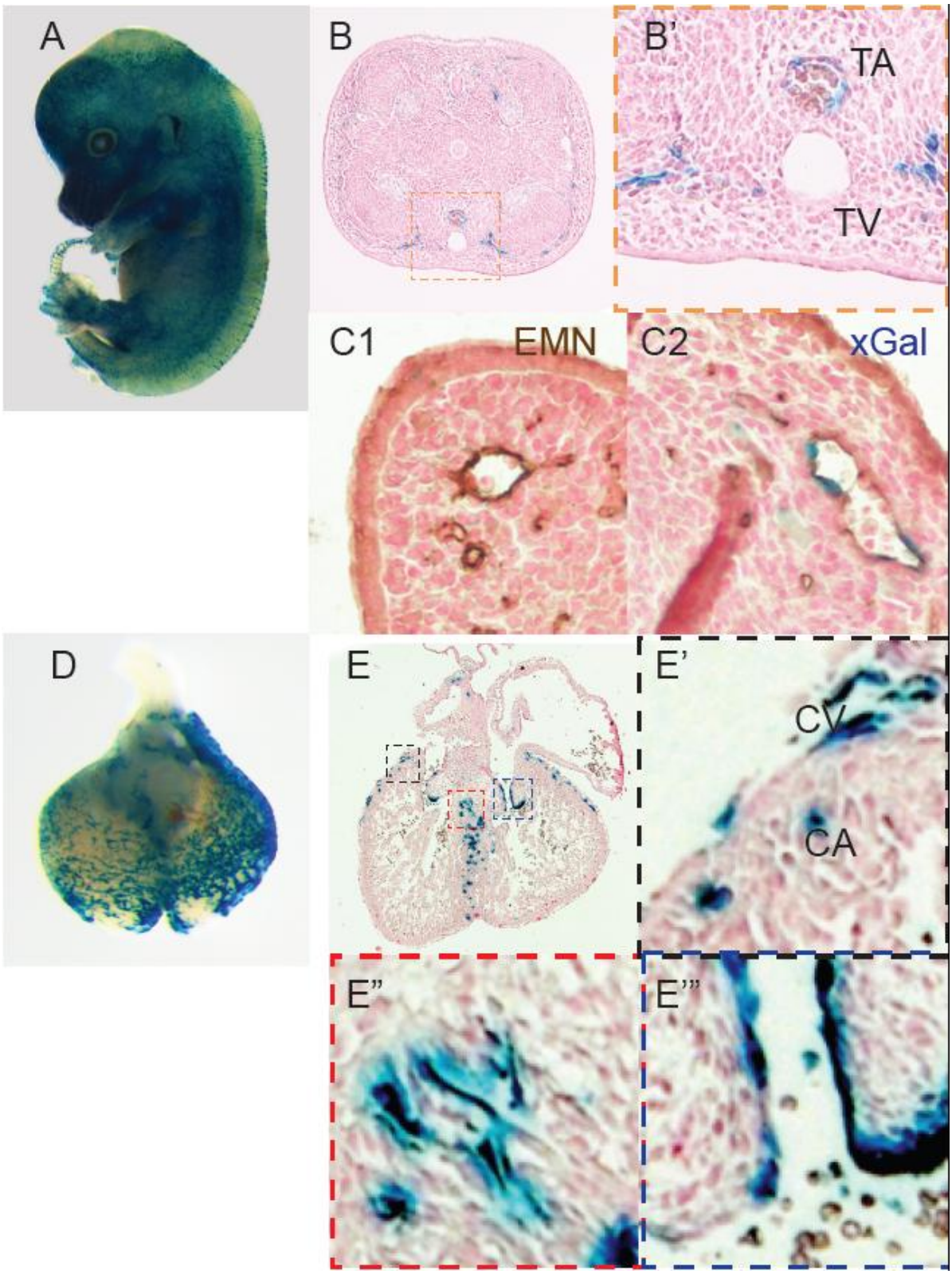


Figure 5.13 hNrp1in2PeakA2*LacZ* arterial and coronary expression at E13

Figure 5.13 hNrp1in2PeakA2 *LacZ* arterial and coronary expression at E13

A. Whole mount embryos of hNrp1in2PeakA2 *LacZ* (peakA2 *LacZ*) transgenic mouse at E13.5. Notice the pandemic vascular expression.

B. Transverse section of peakA2 *LacZ* embryo tail. B', the magnified view of the boxed region in B. Tail artery (TA) is positive for reporter gene expression, whereas tail vein (TV) is negative.

C. Transverse section of peakA2 *LacZ* embryo hind limb.C1, Endomucin (EMCN) immunohistochemistry labelling of the veins.C2, X-Gal staining labelling the arteries.

D. Whole mount heart of peakA2 *LacZ* transgenic embryo at E13.5. Notice the pandemic expression in the developing plexus.

E. Coronal section of the heart. E', E" and E"', magnified views of the boxed region in E of corresponding colour. Notice the expression in the coronary veins and coronary arteries (E'), the interventricular groove (E") and the cardiac valves (E"').

n=12.

Table 5.1 Summary of reporter gene expression detected in transient transgenic mice

Construct	Stage	Expresser/ transgenic (%)	Vascular expressers	% vascular expresser embryos with arterial specific staining	Maximal expression intensity (* least; ***** most)	% vascular expresser embryos with coronary specific staining	Widest cardiac expression (* specks; ***coronary; ***** pan cardiac)
Nrp1in2							
Human WT	E13	1/1 (100%)	1	100%	****	1/1 (100%)	***
mouse WT	E13	8/8 (100%)	6	4/6 66.7%	****	6/6 (100%)	******
hPeakA	E13	6/7 (85.7%)	5	4/5 (80%)	****	5/5 (100%)	***
hPeakB	E13	0/3 (0%)	0	N.A.	N.A.	0	N.A.
PeakA 2	E13	12/12 (100%)	9	5/9 (55.5%)	*****	9/9 (100%)	***
PeakA 2	E11	4/4 (100%)	4	4/4 (100%)	****	N.A.	N.A.

5.4. Conclusions

The mouse Nrp1in2orthologue recapitulates human Nrp1in2 enhancer expression

I have demonstrated that the orthologous mouse sequence of the Nrp1 intron2 enhancer can direct the same vascular endothelial expression profile as the hNrp1in2 enhancer. This is true for both arterial specification between E9.5 to E13.5, and also in the coronary arteries and veins during cardiac development. This demonstrates the importance of conserved elements within the Nrp1 enhancer, and their likely functionality in transcriptional regulation. In particular, the fact that mNrp1in2Rosa directs arterial and coronary expression throughout embryonic developmental stages strongly suggests that the conserved sequences within the hNrp1in2 enhancer harbour the recognition *cis* motif essential for the transcriptional regulation of Nrp1 during arterial and coronary specification. It is surprising that the fellow *placentalia* mouse apparently shares less conservation with the human enhancer than opossum or chick, since of these species; mice should have diverged at the latest point during the evolution of humans. It is, however, possible that this apparent low conservation is due to the fact that a long, non-conserved sequence has recently become inserted into the mouse sequence, breaking up the two conserved regions (peakA and peakB), and messing up conservation algorithms. Supporting this hypothesis, the functional peakA fragment itself is actually well conserved between the human and mouse sequences.

In the analysis described here, the peakB region of hNrp1in2 does not have a prominent functional role in vascular expression during embryonic development. This agrees with the DHS data, where the peakB region was rarely associated with DHS clusters in endothelial

cells or vascular rich organs. The peakB sequence was marked by DNaseI hypersensitive clusters in only 4 cell types in the ENCODE Analysis Working Group database of 125 cell lines (Figure 5.14), and none of these 4 lines are endothelial (HFF, HFF-Myc, and WI38 are fibroblast cells, while Huh-7 is a hepatocellular carcinoma cell line). Since DNaseI hypersensitivity is indicative of open chromatin structure and active involvement in transcriptional regulation of the respective cell, this also suggests that the peakB region of the hNrp1in2 enhancer is not involved in Nrp1 transcriptional regulation in the endothelial cell context. Further, the peakB fragment was analysed for motif recognition, and non-conserved functional ETS binding site was discovered *in silico*. Since Ets motifs are ubiquitous among endothelial enhancers, this agrees with the transgenic analysis. However, the hNrp1in2 full enhancer did appear to direct stronger vascular expression when compared to hNrp1in2PeakA, although when investigated in sections, both were detected in arterial ECs. It is therefore possible that sequences within peakB exert their function through collaboration with other elements. Additionally, peakB may be involved in other scenarios rather than in embryonic vascular development, since it is well conserved between species, and is associated with the chromosomal conformation of active transcription regulation involvement in some tissues and cell types. However, it is also possible that this conservation is merely an evolutionary legacy with no functional indication whatsoever. It has been reported that some ultra-conserved elements, deeply conserved between zebrafish and humans, appear to not always have clear transcriptional or other regulatory functions (*217*).

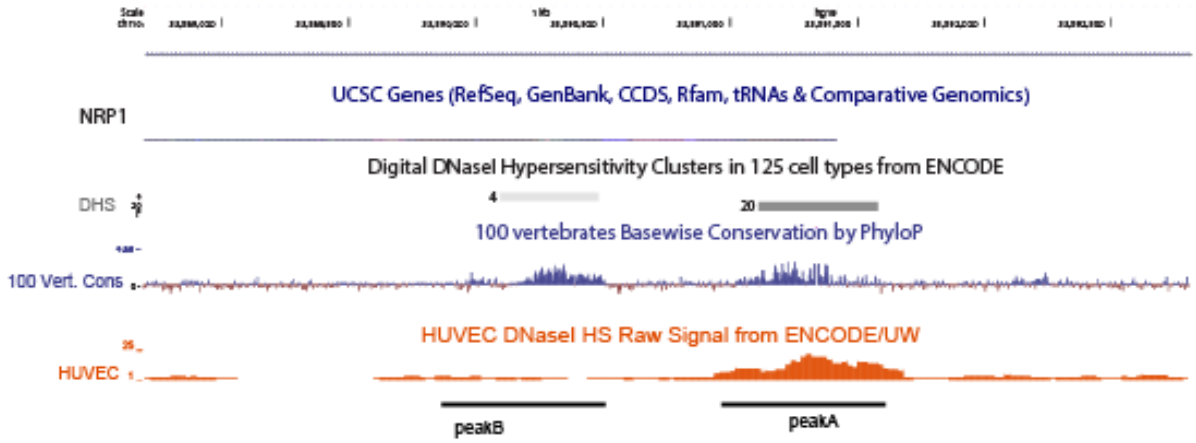


Figure 5.14 peakB is not enriched with DHS in endothelial cell line HUVEC.

Grey track, DNaseI Hypersensitivity Clusters from ENCODE. Only 4 of the 125 cell lines examined are enriched with DHS in the peakB region of the enhancer, whereas there are 20 enriched in peakA. For example, in the endothelial cell line HUVEC, DHS enrichment is pronounced in the conserved peakA region but not in peakB. (This track shows data from the University of Washington and Duke ENCODE groups, with uniform processing by the ENCODE Analysis Working Group. The clustering was performed at UCSC. For additional credits and references, see the Uniform DNaseI HS page.)

Core enhancer essential for transcription regulation function during vascular development

This study reports the identification of a core conserved region of the human Nrp1in2 enhancer, peakA2, consisting of 184bp. The region is highly conserved between human and mouse, and is marked with epigenetic modifications indicating *trans* factor accessibility and transcription regulation involvement in endothelial cells. It is demonstrated that this core region is fully capable of recapitulating transcriptional activation in arterial and coronary endothelial cells similar to the full-length hNrp1in2 enhancer. The examination of the sequences of this core enhancer identified a small number of conserved motifs. Because both mouse and human hNrp1in2 enhancers directed similar patterns of arterial and coronary EC expression, it is hypothesised that these conserved motifs are likely to be important for hNrpin2 expression. These motifs include binding sites for Ets, Fox, and Gata factors. Both of the Fox motifs are adjacent to Ets sites, forming composite Fox:Ets motifs (please see detailed description in Figure 6.1). As discussed in the Introduction chapter, these *trans* factors are commonly seen alone and in combination in endothelial cell development, although mostly in early, pan-vascular expression. Etv2 is expressed during the early phase of endothelial development, can bind with the Gata2 protein, and share a composite site with FoxC during the transcriptional regulation of endothelial or endothelial progenitor development. The role of such regulation, both alone and in combination, has not been fully explored in terms of endothelial cell differentiation into arteries and veins, or in coronary vessel formation.

6. Identifying the transcription factors that regulate the vascular expression of the hNrp1in2peakA2 enhancer

6.1.Introduction

As discussed in the Introduction chapter, a number of transcription factors are commonly involved in endothelial cell differentiation. All known vascular enhancers contain binding motifs for Ets factors, and combinatorial regulation by Fox and Ets is also a common theme. The core conserved region of the human Nrp1in2 enhancer consists of 184 bp of DNA (hNrp1in2peakA2). This region is highly conserved between human and mouse, and is marked with epigenetic modifications indicating *trans* factor accessibility and transcription regulation involvement in endothelial cells. Although the expression pattern directed by this minimal enhancer region varied slightly from the full length hNrp1in2 enhancer, it is as capable of recapitulating transcription activation in arterial and coronary ECs as the full length 1.8 kb (Chapter 5).

6.2.*In silico* analysis of the core hNrp1in2peakA2 enhancer

I used the Transfac database (Biobase) to identify the putative transcription factors that could bind to the hNrp1in2peakA2 sequence. Because both the mouse and the human Nrp1in2 enhancers directed arterial- and coronary-specific expression, I concentrated on the elements conserved between the two species. I also compared the conserved regions to a list of putative binding motifs for endothelial-associated transcription factors. This was assisted by the fact that this core element contained very few conserved motifs. The results of this search are highlighted in Figure 6.1. This included eight Ets sites (Ets1-8). Of these, five were fully conserved (Ets1, Ets2, Ets5 Ets7 and Ets8), and another three Ets sites (Ets3, Ets4 and Ets6) found in the human sequence were not fully aligned with all the other three species, yet

conserved between human and at least one of the other three species. I included these in my analysis to ensure an exhaustive approach. Additionally, I identified two putative FoxC sites and one putative Gata site (Figure 6.1). Each of the two putative FoxC sites were adjacent to a conserved Ets site, forming two composite Fox:Ets sites (Fox1:Ets5 and Fox2:Ets7). Since both mouse and human sequences directed arterial- and coronary- specific expression, and these were the only motifs within the conserved regions that bound known transcription factors, I hypothesised that a subset of these sites would be the essential *cis* elements responsible for the expression of the enhancer in vascular endothelial cell transcription.

6.3.Validation of putative binding sites using electrophoresis mobility shift assay (EMSA)

To test the putative transcription factor recognition motifs in the enhancer, I screened them *in vitro*. Using an *in vitro* transcription and translation system, I synthesized recombinant proteins for Ets, FoxC and Gata, and incubated them with radiolabeled oligos in an EMSA assay. Mobility shift will result in band formation in the electrophoresis, which indicates affinity between the protein and the oligo.

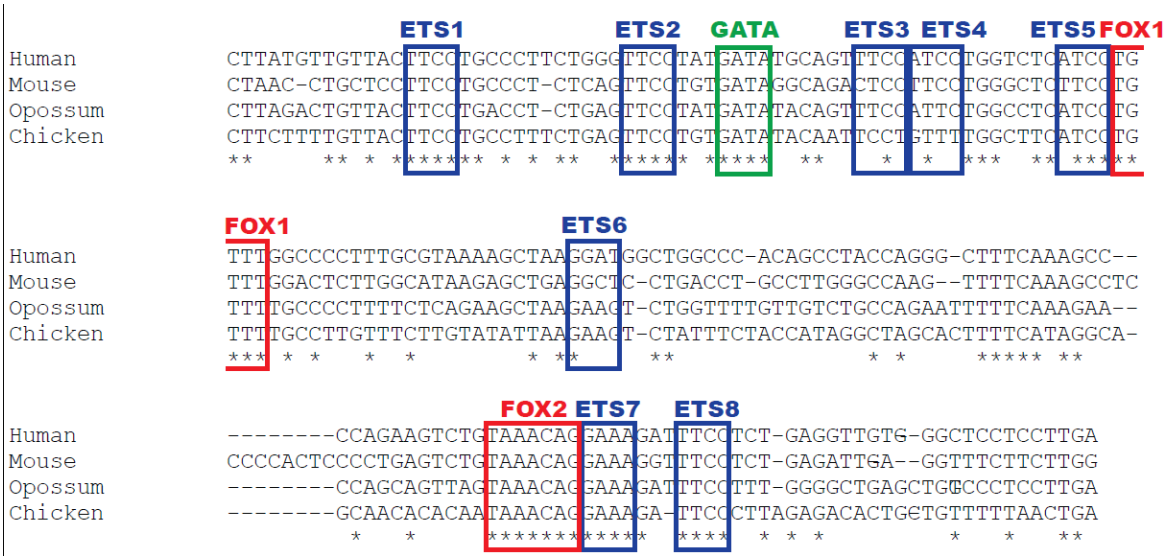


Figure 6.1 hNrp1in2peakA2 annotation of putative recognition sites.

The hNrp1in2peakA2 fragment and its orthologous sequence in mouse, opossum and chicken were aligned using the ClustalW multiple sequence alignment tool. Conserved Ets1, 2, 5, 7, 8, Fox1, 2 and Gata putative sites are observed in this sequence. Ets3, 4, and 6 putative sites were found in the human sequence, but are not conserved to all other species.

6.3.1. FoxC2 binds the Nrp1 Fox2 site but not the Fox1 site

The Fox sites were tested using recombinant FOXC2 protein. I used the Mef2c-F10 FoxC binding site as positive control of FOXC2 binding (*126*). In the indirect EMSA experiment, excessive unlabeled oligos of both hNrp1in2A2 FoxC sites compete against the radiolabeled Mef2C-F10 Fox site binding to FOXC2, ablating the shifted bands (Figure 6.2, left). This result suggests that both Nrp1 Fox sites can effectively compete with the known Fox site for binding of FoxC2. I also radiolabeled oligos containing each of the twoNrp1 Fox sites, and directly incubated them with the FOXC2 protein. The Nrp1 Fox1 site did not generate a band shift (Figure 6.2, right), suggesting that it is a poor binding site for FoxC2. However, the oligos for the Nrp1 Fox2 site robustly bound to FOXC2 protein; further, while excessive unlabeled WT Nrp1 FoxC2 oligo can compete away the radiolabeled probe binding to ablate the band shift, such competition did not occur when using excessive unlabeled oligo with a mutation in the putative Nrp1 Fox2 site (Figure 6.2 right). These results suggest that FOXC2 can bind the hNrp1in2A2 enhancer by recognizing the Fox2 binding motif, whereas the Nrp1 Fox1 site does not strongly bind FOXC2. It is however possible that the Fox1 site can be recognized by FOXC2 only with the assistance of another factor (e.g. ETS factors), or it may be recognized by other FOX factors. It is also possible that although the Fox1 site looks like a Forkhead binding motif, it is not actually recognized and bound by any Forkhead. Nevertheless, the Nrp1 Fox2 site appears to be a legitimate candidate as a functional *cis* element in the enhancer.

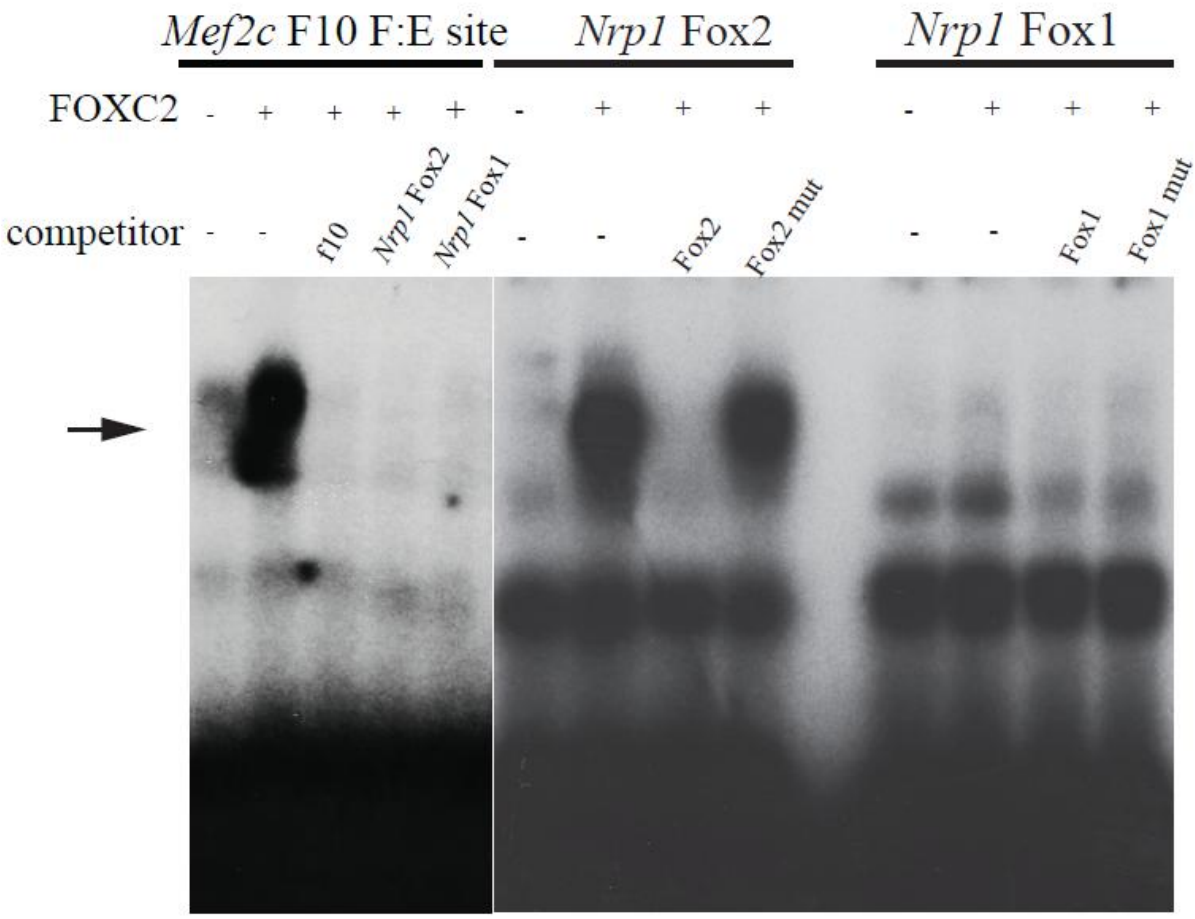


Figure 6.2 Nrp1 enhancer Fox site binds recombinant FoxC2 protein.

EMSA of FOXC2 binding to oligos containinghNrp1in2peakA2 Fox sites. Radiolabeled oligos including Fox binding sites of the *Mef2C*-F10 enhancer competed with, and directly bound to FOXC2 recombinant protein (arrow). FOXC2 robustly binds the control *Mef2C*-F10 Fox site in lane 2. However, in lane3, 4 and 5, excess unlabelled self-probe (F10) and *Nrp1* Fox1 and Fox2 oligos are able to compete away the binding, suggesting that all three oligos can bind FOXC2. In lanes 5-12, radiolabeled oligos containing the *Nrp1* Fox1 and Fox2 sites were incubatedwithFOXC2 recombinant protein. The*Nrp1* Fox2 probe is able to robustly bind FOXC2 (lane 6), which is competed away by excess unlabelled self-probe (lane 7) but not by mutated self-probe (lane 8). However, the*Nrp1*Fox1 oligo probe was unable to bind FOXC2 directly (lane 10).

6.3.2. The Ets sites Ets 1, Ets5, Ets7 and Ets8 bind recombinant ETS proteins in vitro

I identified five conserved Ets sites in the hNrp1in2peakA2 enhancer. I also investigated three Ets sites of the human sequence that were not conserved in all of the other three species, in order to cover all possible Ets binding sites in the enhancer (Figure 6.1, blue boxes).

ETS1

The full length Ets1 protein contains an auto-inhibitory domain, and is not suitable for EMSA. Consequently, I used the ETS1 DNA Binding Domain (ETS1-DBD) instead, as in previous publication (*126*). ETS1-DBD showed robust binding to a number of Ets sites in the enhancer.Ets1, Ets5, Ets7 and Ets8all clearly compete for binding of ETS1-DBD with the radiolabeled *Mef2C*-F10 control Ets site (Figure 6.3). Among the conserved Ets sites, only the Ets2 site oligo did not compete. None of the less conserved sites competed well. This again demonstrates that the conservation information is a good indicator of functionality of the non-coding sequences. When I tested directly the ability of the radiolabeled Nrp1 Ets sites to bind ETS1-DBD, a band shift was clearly observed with Ets1 and Ets5, as well as less intense shifted bands with Ets8 and Ets7 site oligos (Figure 6.3).

ETV2

As a consensus site, the ETS factors recognize the same 4-base-pair site. Functional Ets sites tend to share affinity to various ETS factors, but some variations can be seen (*65*). Because of its importance in early vascular identity, I also performed EMSA testing the ability of recombinant protein for ETV2 to bind the Nrp1 Ets motifs. The conserved Ets1, Ets5 and

Ets7 all bound ETV2 strongly, and Ets8bound mildly (Figure 6.4 and Figure 6.5). It was interesting that both the Ets5 and Ets7 sites, part of putative composite Fox:Ets site, bound ETV2.

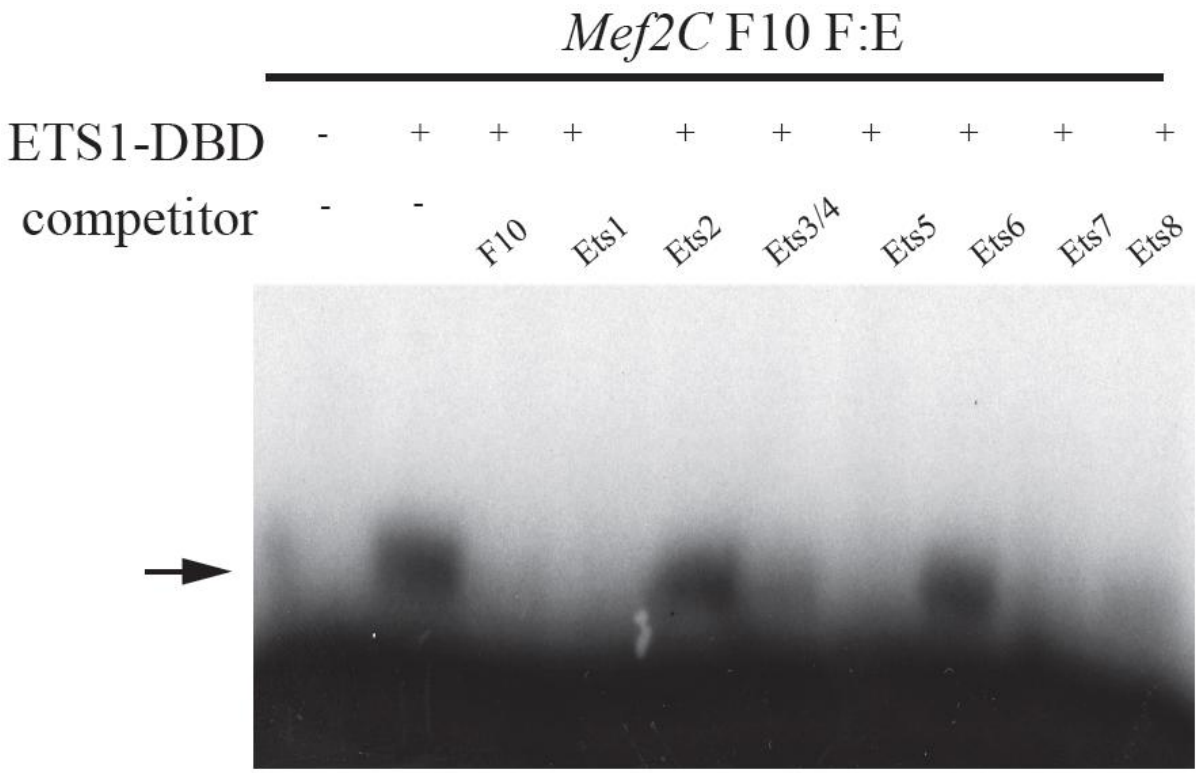


Figure 6.3 ETS1-DBD binds some but not all conserved Ets sites in hNrp1in2peakA enhancer.

Radiolabeled oligos containing the control Ets binding site of the *Mef2C*-F10 enhancer robustly bound ETS1-DBD (lane 2). This binding was successfully competed away by oligos containing putative Ets motifs *Nrp1* Ets1, Ets5, Ets7 and Ets8 (lanes 4, 6, 7, 9 and 10). The *Nrp1* Ets2, Ets3/4 or Ets6 sites do not compete with the ETS1-Mef2C F10 complex band shift for binding with ETS1-DBD.

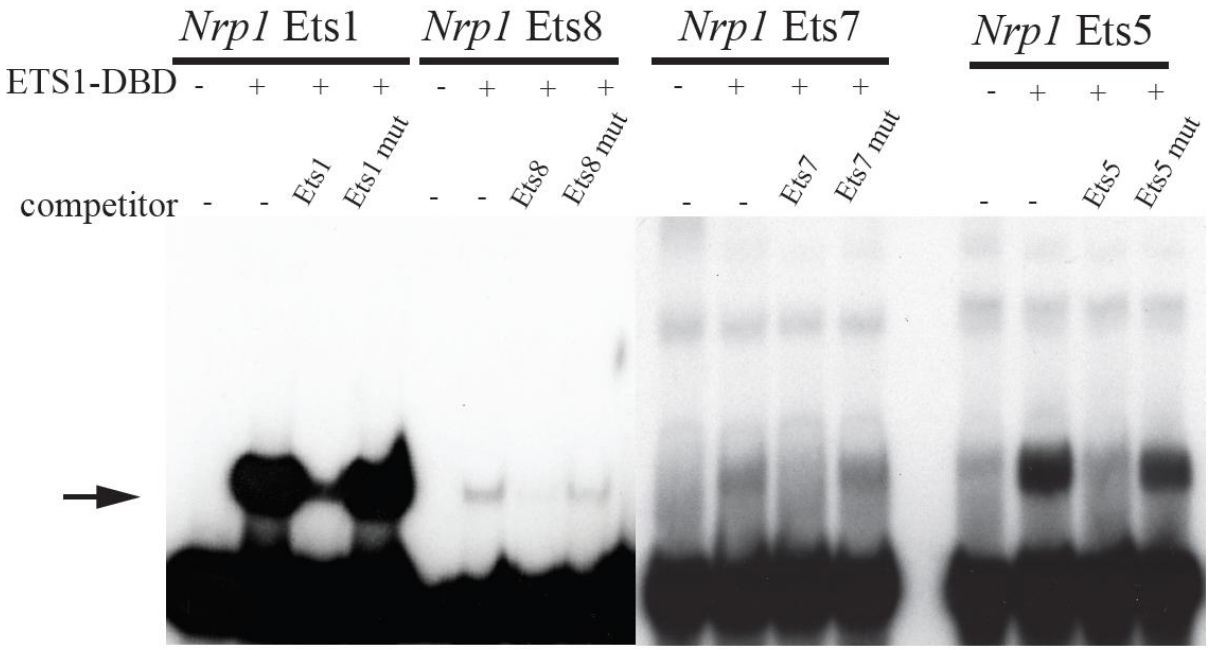


Figure 6.4 ETS1-DBD can directly bind Ets sites within the hNrp1in2peakA2 enhancer

EMSA of ETS1-DBD binding of hNrp1in2peakA2 conserved Ets sites. Radiolabeled oligos containing the Ets binding sites 1, 8, 7 and 5 were tested as to whether they bound ETS1-DBD. *Nrp1*sites Ets1, Ets7 and Ets5, were able to directly bind ETS1-DBD (arrow), and this shift was competed by excess unlabelled self-probe but not by mutated self-probe. *Nrp1*Ets8 was also able to shift ETS1-DBD, but only very weakly.

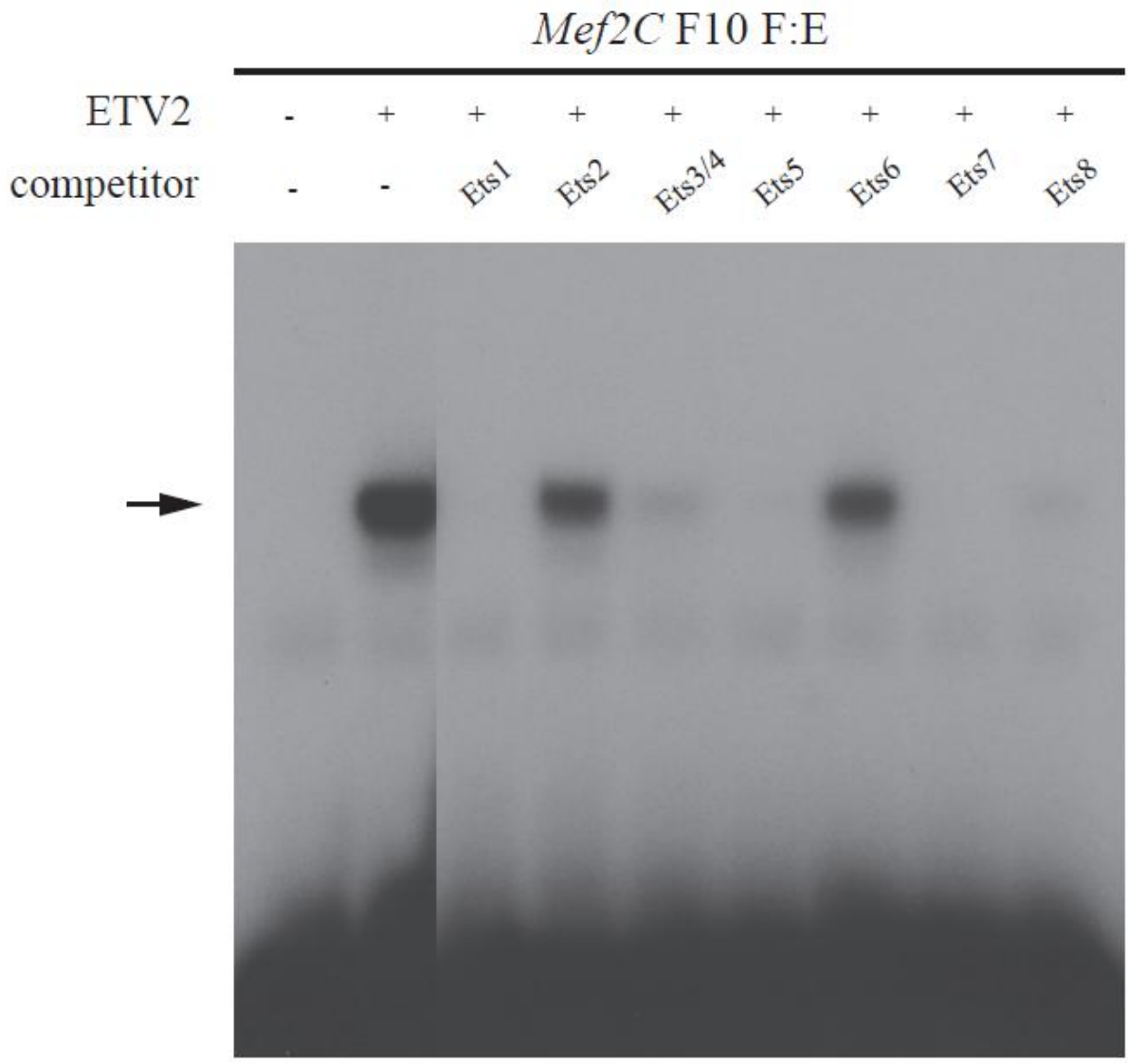


Figure 6. 5 ETV2 can bind conserved Ets sites in the hNrp1in2peakA2 enhancer.

EMSA of ETV2 recombinant protein with Nrp1in2peakA2 ETS sites. Radiolabeled oligos containing the sequence of the Ets binding sites of the *Mef2C*-F10 enhancer robustly bound ETV2 (lane 2). This binding was competed by excess unlabelled *Nrp1* Ets1, (arguably partially by) Ets3/4, Ets5, Ets7 and Ets8 oligos. The *Nrp1*Ets2 or Ets6 probes do not compete against the ETS1-Mef2C F10 complex band shift.

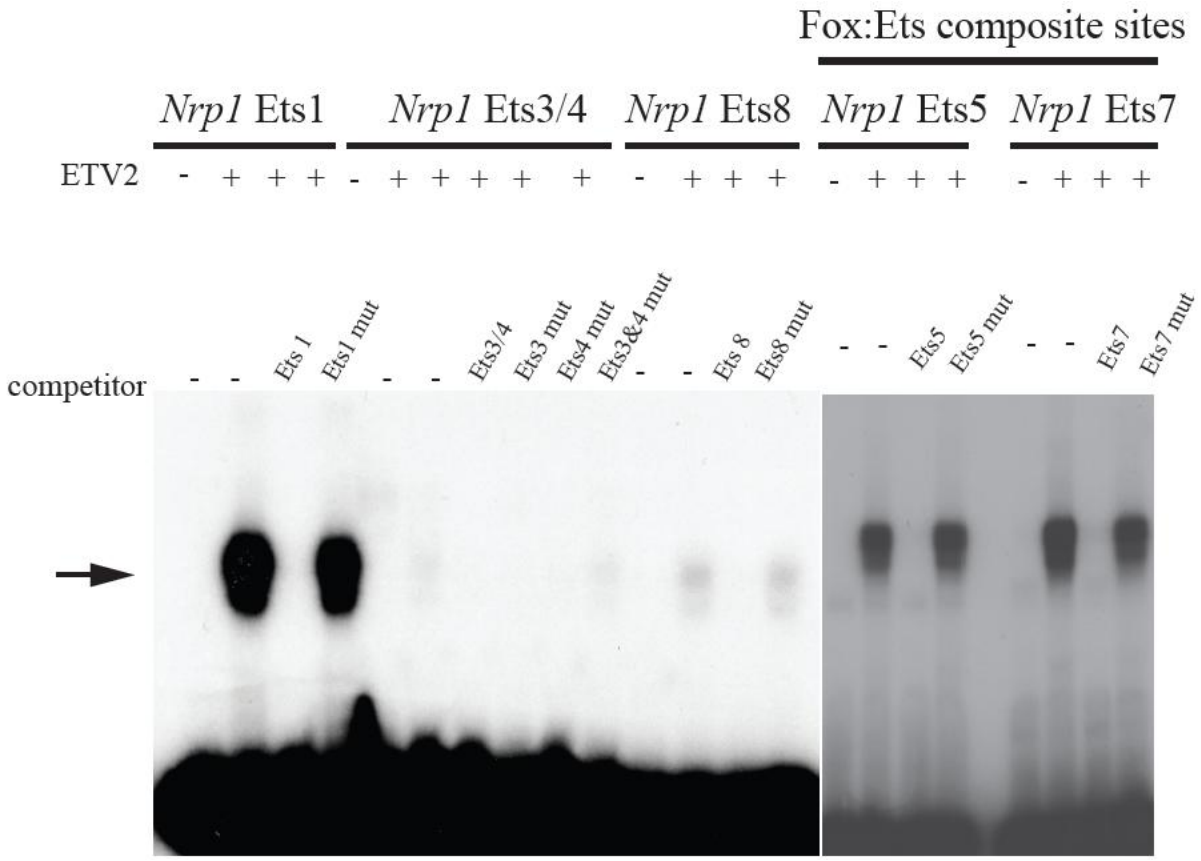


Figure 6.6 ETV2 can directly bind some but not all *Nrp1*Ets sites

EMSA of ETV2 against Nrp1in2peakA2 Ets sites. Radiolabeled oligos containing each Ets binding sites within the Nrp1 enhancer was incubated with ETV2 recombinant protein. *Nrp1*Ets1 was able to bind ETV2 in lane 2, which is in lane3 competed by excess unlabelled self-probe but not by mutated self-probe in lane4. Similar affinity for ETV2 is seen in with Nrp1 Ets5, and 7, as well as less strongly with *Nrp1*Ets8. Sites Ets3/4 displays no binding to the ETV2 protein.

ERG

Because of the role of Erg in arterial specification, I wanted to determine whether the Ets factor Erg also bound the Nrp1 Ets sites. Consequently, I tested the Nrp1 Ets sites that were able to bind ETS1-DBD for their binding affinity to the ERG recombinant protein in EMSA. Nrp1Ets1, Ets5, Ets7 and Ets8 sites were also able to bind ERG (Figure 6.7). The fact that these Ets sites promiscuously bind multiple ETS factors is consistent with the fact that ETS factors share a consensus recognition motif. The Ets5 and Ets7 sites, both part of composite Fox:Ets motifs, also bind various ETS factors. These may therefore represent FOX:ETS composite sites.

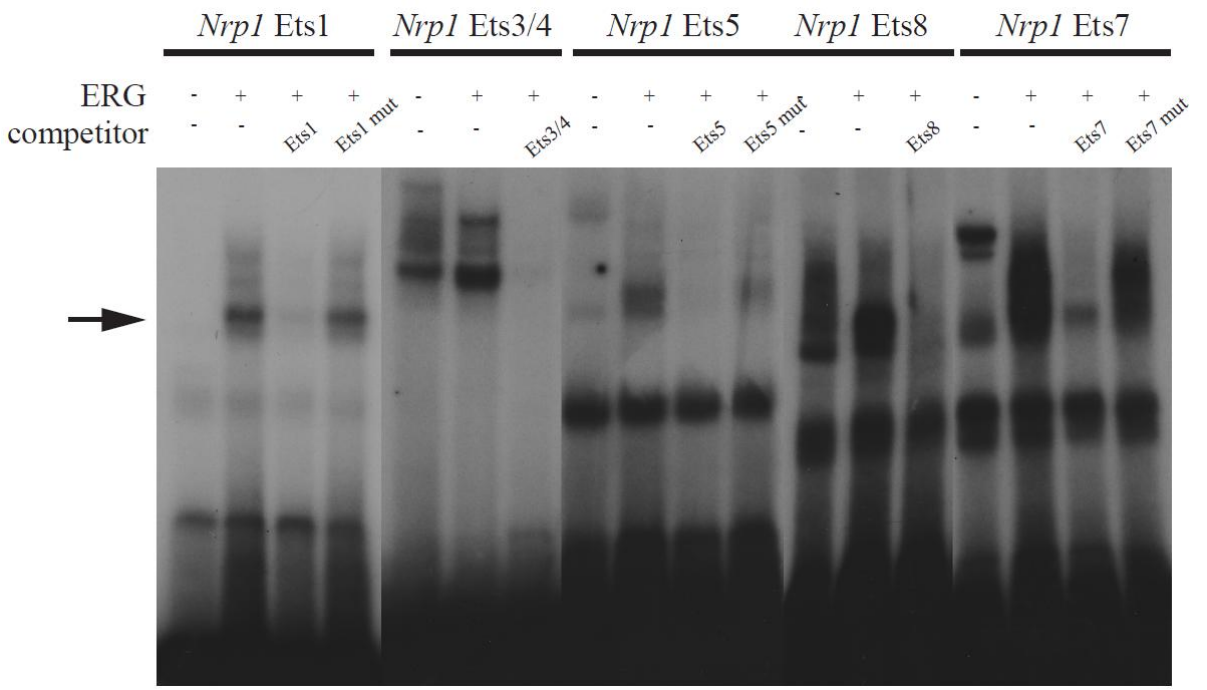


Figure 6.7 ERG binds conserved Ets sites in Nrp1in2peakA2.

EMSA of ERG against Nrp1in2peakA2 Ets sites. Radiolabeled oligos harbouring Ets binding sites of the Nrp1 enhancer bind to ERG recombinant protein. Such affinity is seen in Ets1 mobility shift forming a band in lane 2, which is in lane3 competed by excess unlabelled self-probe but not by mutated self-probe in lane4. ETV2 affinity is seen in Ets1, 5, 7 and 8. Site Ets3/4 displays no binding to the ERG protein. Ets5 and 7 are in composite Fox:Ets site in the enhancer (Fox1:Ets5 and Fox2:Ets7).

6.3.3. FoxC2 binds the Nrp1 enhancer

The Ets5 site within the hNrp1in2A2 core enhancer is adjacent to the Fox1 site, and so as Ets7 to the Fox2 site. Since both Ets5 and Ets7 are functional sites in terms of binding to ETV2 and ERG in EMSA, I next wanted to establish whether the Fox sites could bind FOXC proteins in vivo (Figure 6.8). The EMSA demonstrated at Nrp1 Fox2 site was able to bind FOXC2, but the function of the Fox1 site was not confirmed. Fox2:Ets7 is therefore the only putative composite Fox:Ets site within the Nrp1in2A2 enhancer (Figure 6.8). However, this could not be clearly validated as a functional FOX:ETS motif through EMSA alone.

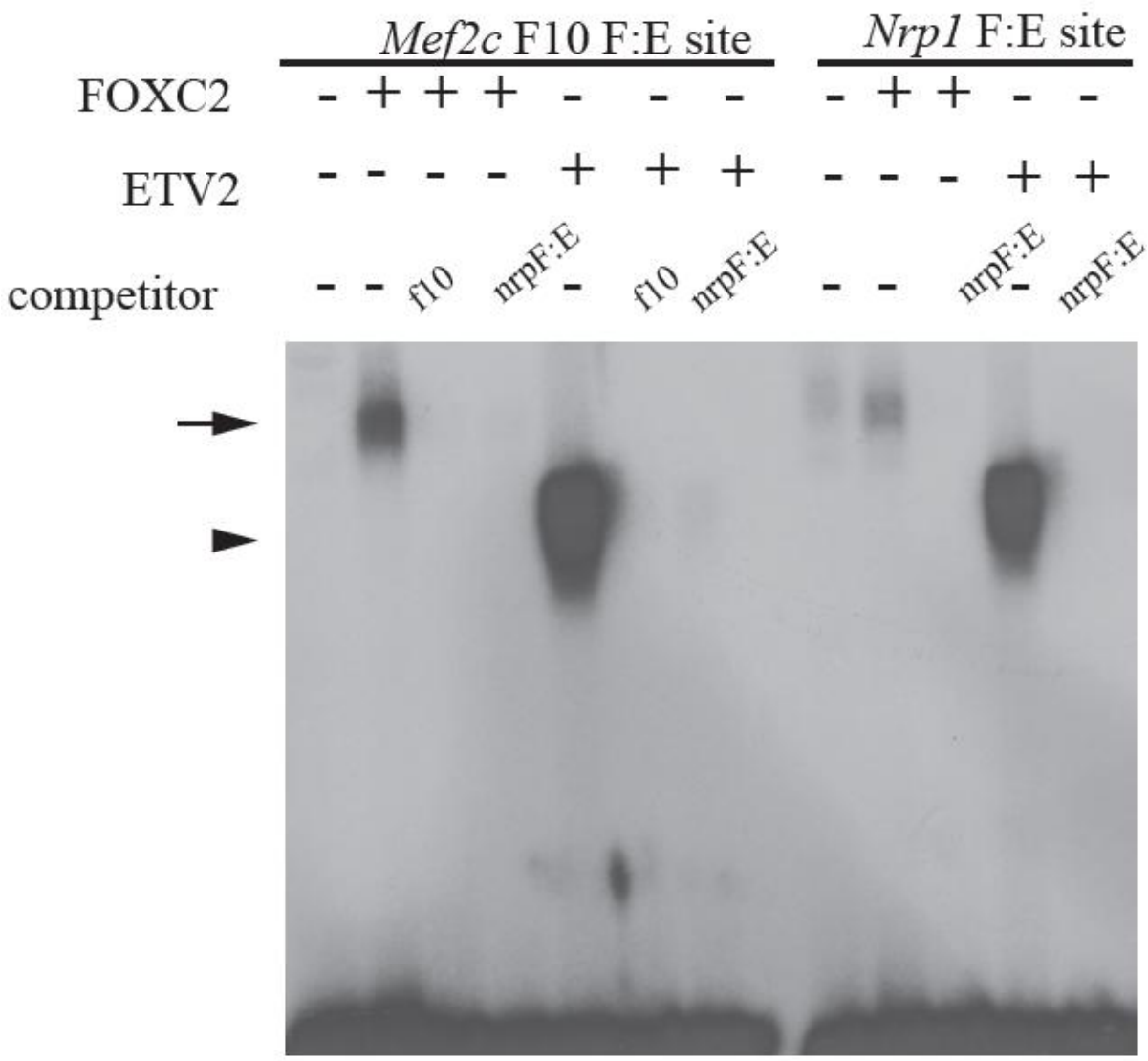


Figure 6.8 Fox2:Ets7 composite site can bind both FOXC2 and ETV2.

EMSA of FOXC2 and ETV2 against Nrp1in2peakA2 putative Fox:Ets motif. Radiolabeled oligos harbouring *Mef2C* F10 Fox:Ets motif and the Fox2:Ets7 site of the *Nrp1* enhancer bind to FOXC2 (arrow) and ETV2 (arrowhead) recombinant protein. Such affinity is seen in *Mef2C* F10 Fox:Ets mobility shift forming a band in lane 2 and 5 to FOXC2 and ETV2 respectively, which is in lane3 and 6 competed by excess unlabelled self-probe and in lane4 and 7 *Nrp1*Fox:Ets. Direct *Nrp1* Fox:Ets site (Fox2:Ets7) band shift against FOXC2 and ETV2 affinity is seen in Lane 9 and 11.

6.3.4. FoxC2 and Etv2 synergistically activate hNrp1in2

I performed reporter assays in a non-endothelial cell line COS-1 to test the putative enhancer's activation in response to FOXC2 and ETV2 expression. The confirmed Fox:Ets motifs such as Mef2c-F10 are synergistically activated by FOXC2 and ETV2. The enhancers were cloned into TK-βGal-basic backbone and transfected into COS-1 cells. Meanwhile, FOXC2 and ETV2 coding regions were in the pRK5 backbone, a vector with a CMV promoter that drives high level transient expression in mammalian cell lines. The FOXC2 and/or ETV2 in pRK5 were co-transfected with the enhancer-reporter constructs, and reporter gene expression was measured by β-galactosidase luminescent detection, Activation of the hNrp1in2, mNrp1in2 and hNrp1in2peakA enhancers by a combination of both FOXC2 and ETV2, analysed by reporter gene β-Gal expression detection was found to be significantly higher than either of the two factors alone (Figure 6.9). In agreement with our other data, hNrp1peakB did not display this synergistic activation, reflecting the fact that there is no Fox:Ets motif identified in the peakB region (data not shown). The synergistic activation of the enhancer by FOXC2 and ETV2 is a phenomenon resembling the Mef2C endothelial enhancer reported by De Val et al. (*126*), and confirms that the Nrp1Fox2;Ets7 motif is likely to be a functional FOX:ETS motif.

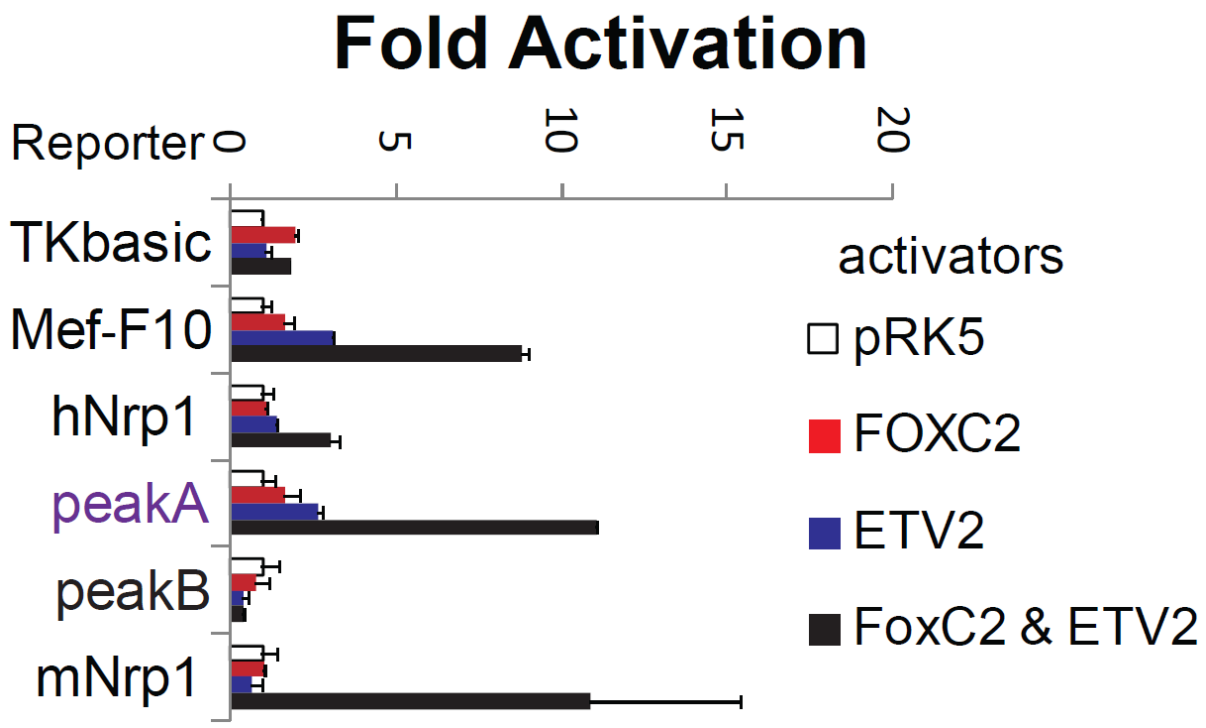


Figure 6.9 Synergistic activation of the Nrp1 enhancers by FOXC2 and ETV2

Reporter assay showing pronounced synergistic activation of FOXC2 and ETV2 on peakA, mNrp1in2 and hNrp1in2 48 hrs after transfection into COS-1 cells. Peak B shows no activation. Mef2c-F10 served as positive control of synergistic activation (*126*). The activation folds were calculated by normalizing to that of pRK5 empty vector of each corresponding reporter. N=3. Error bar, standard deviation.

6.3.5. GATA4 can bind the Nrp1 Gata motifs

There is one conserved Gata site in the hNrp1in2peakA2 core enhancer (Figure 6.10, green box). In an indirect competition EMSA against the Dll4Gata site (*65*), the Nrp1Gata site was able to compete with theDll4Gata site, forcing the radiolabelled Dll4 Gata probe and recombinant GATA4 protein to form a lesser intense band (Figure 6.10). Directly radiolabeled Nrp1 Gata site was also able to bind GATA4, and this shift was competed away by excessive un-labeled Nrp1 Gata oligo. Additionally, mutations introduced into the Gata recognition motif of the Nrp1 Gata1 oligo ablated binding. GATA2 was also tested and it did not bind (data not shown). To summarise, this Nrp1enhancer can be recognized by GATA4 and the recognition would be due to the conserved Gata site.

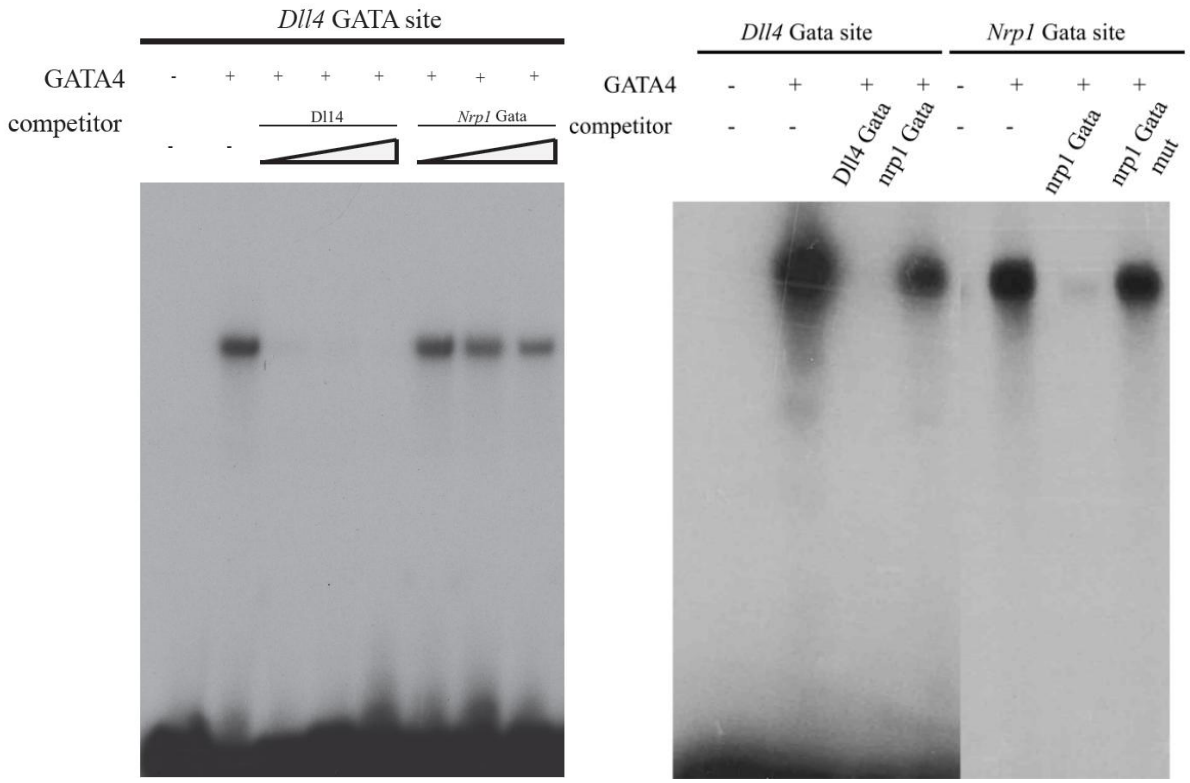


Figure 6.10 GATA4 can bind the conserved Gata site in the Nrp1in2peakA2 enhancer.

Direct and indirect EMSA of GATA4 binding Nrp1in2peakA2 Gata site. Radiolabeled oligos containing the Gata binding site from known motifs in the *Dll4*-intron3 enhancer were able to robustly bind to GATA4 recombinant protein (lane 2), which was competed away by excess unlabelled self-probe, and partially by the *Nrp1*Garaoligos, particularly with higher concentrations of unlabelled Nrp1 Gata oligo. Directly radiolabelled oligos containing the *Nrp1*Gata site were also able to bind to GATA4 recombinant protein. This was competed away by excess unlabelled self-probe but not by mutated self-probe.

6.4. Mutational analysis

To define the role of the conserved, validated by EMSA, transcription binding motifs *in vivo*, I introduced sequence mutations into the enhancer and generated a series of transient transgenic mice to enable the analysis of the expression profile driven by the disrupted enhancer. I used the same mutations that were tested in the EMSA assay, already proven to prevent the binding of the *trans* factor protein *in vitro*. I anticipated that the functional motifs that are essential for vascular endothelial cell transcription regulation would be indispensable in the enhancer sequence. Such disruption would impair the vascular-specific expression of the enhancer, and the arterial or coronary compartment-specific expression would be expanded or further restricted. Non-functional motifs on the other hand would survive mutation introduction and result in an unaffected expression profile. I therefore designed three different mutations to the hNrp1in2peakA2 enhancer (Figure 6.11). The 'pA2 ALL Ets mut' construct has all 8 putative Ets sites mutated, regardless of conservation or *in vitro* validation. The 'pA2 FUNC Ets mut' construct has only the EMSA validated sites Ets1, 5, 7 and 8 mutated. The 'pA2 Fox Gata mut' construct is a compound mutation construct with both Fox and Gata sites ablated. Our collaborator George Bou-Gharios and Ke Liu then performed transient transgenesis, and I analyzed the produced E13.5 transgenic embryos with X-Gal staining.

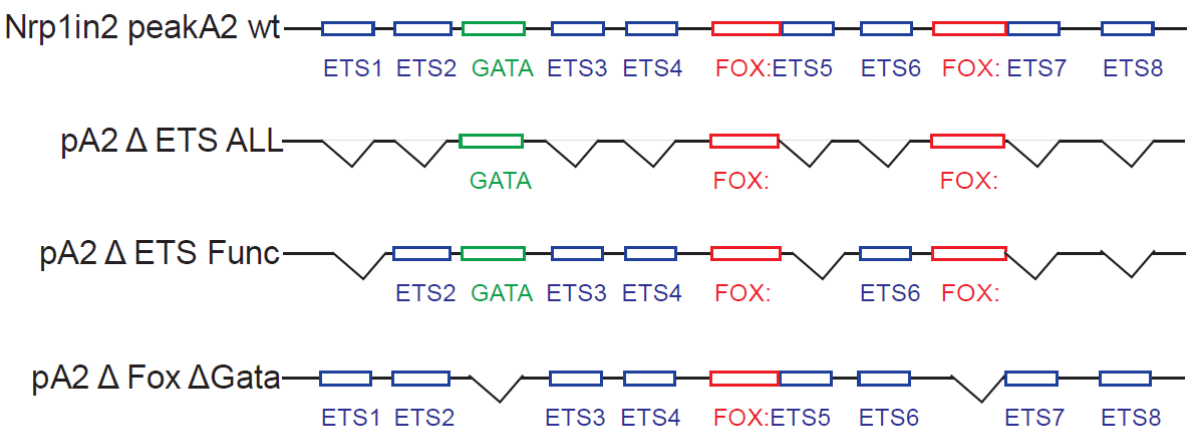


Figure 6.11 Design of mutational analysis of hNrp1in2peakA2.

Schematic representation of the mutational design of the hNrp1in2peakA2 enhancer. Mutations introduced to all putative Ets sites found in the human sequence to the hNrp1in2peakA2is termed 'pA2 All Ets mut'. Fragment with mutations introduced to validated conserved sites Ets1, 5, 7 and 8 is termed 'pA2 Func Ets mut'. Compound replacement of both Fox and Gata sites is termed 'pA2 mut Fox mut Gata'. Further experiments and discussion will refer to these sequences as shown here.

6.4.1. Depletion of all the Ets sites prevents endothelial expression

pA2 ALL Ets mut contains mutations in all the eight putative Ets binding sites (see Figure 6.11). Three transgenic embryos were found to express the reporter gene *LacZ*. However, unlike all other previous hNrp1 enhancer constructs, none of the three had any expression in the vasculature. From the whole mount picture, the stereotypical vascular plexus on the surface of the skull that is typically found positive in hNrp1in2 *LacZ* and hNrp1in2peakA2 *LacZ* transgenic embryos is not visible in these pA2 ALL Ets mutants (Figure 6.12, A1, A2, B3). In fact, none of the vasculature is positive for X-gal staining in these embryos. The dorsal aorta, the coronary vessels and the tail artery are all negative, as clearly seen in the transverse sections (Figure 6.12 B). In the heart, there was some staining in the apex of the heart, although it is not in the location of the developing plexus or the interventricular groove, where the expression was detected in the wild type hNrp1in2peakA2 transgenic mice (Figure 6.12 B2).This is the same for the tail vessels. Staining was detected in the ganglion cluster surrounding the segmental nerves, but not in the stereotypical structure of the caudal tail vein or the ventral tail artery (Figure 6.12 B4). Expression was also found in the midline neural tube (data not shown), suggesting that the neuronal expression of the Nrp1in2peakA2 enhancer was maintained in this mutant. Nevertheless, the loss of vascular expression is consistent in all three transgenic mice, indicating the essential role of Ets sites in this enhancer for its vascular transcription regulation.

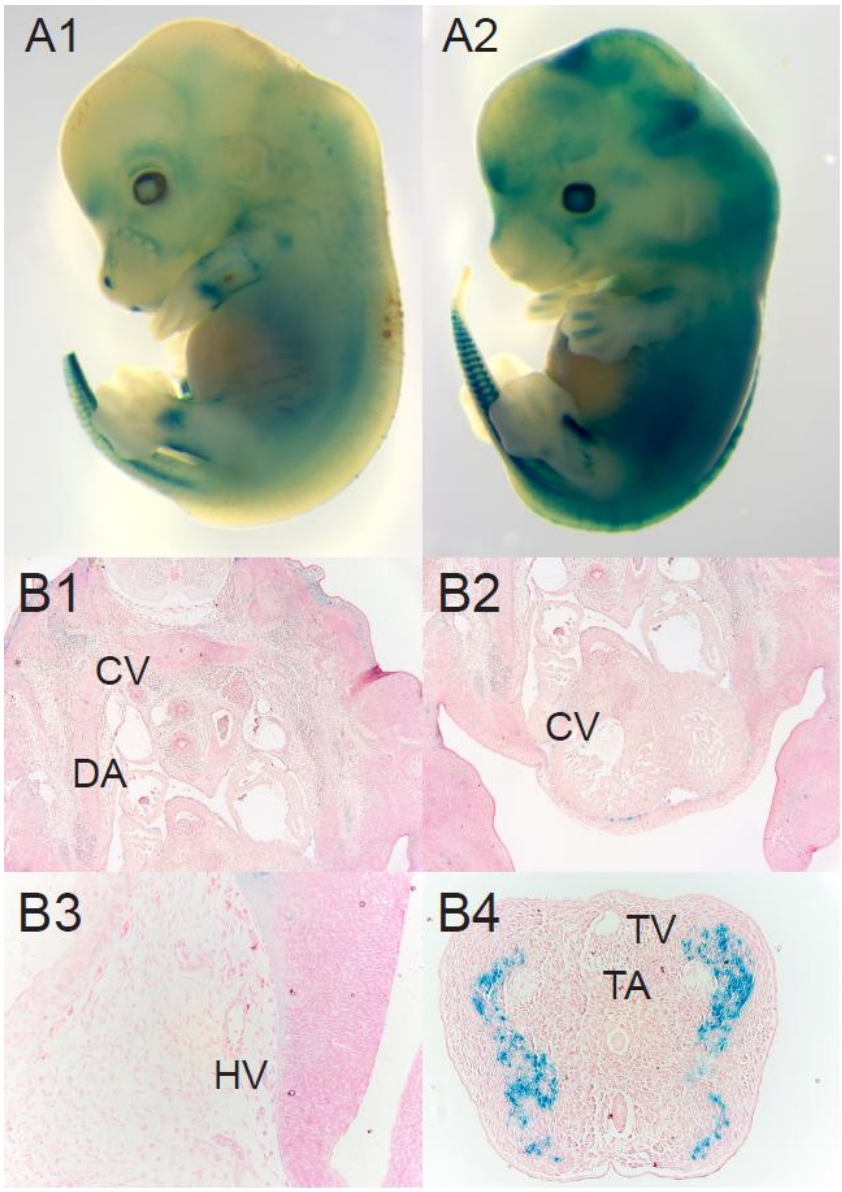


Figure 6.12 ALL-Ets mutation prevents vascular expression of the hNrp1in2peakA2 enhancer

A. Whole mountE13.5 embryo transgenic for the 'pA2 ALL Ets mut' construct. A1 and A2 are two different embryos of the same transgene.

B. Transverse sections of the embryo in A1. B1, the cardinal vein and the dorsal aorta are both negative for x-Gal staining. B2, the developing heart shows not coronary vascular expression. B3, the head vessels are negative for X-Gal staining.B4, tail arteries and veins are negative. X-Gal expression is in the non-vascular interstitial tissues. DA, dorsal aorta; CV cardinal vein; CV coronary vessels; HV, head veins; TV, (caudal) tail vein; TA, (ventral) tail artery. n=3.

6.4.2. Functional Ets site mutation deplete intra-embryonic vascular expression

Since not all the eight putative ETS sites were validated as ETS1-factor-binding in the *in vitro* assay, I decided to focus on the functional sites. Therefore, I made a construct in which mutations eliminate the Ets1, Ets5, Ets7 and Ets8 sites, termed the 'pA2 FUNC Ets mut' construct (Figure 6.11). Again, we collaborated to generate transgenicE13.5 embryos. I only received two embryos that were transgenic (according to PCR based genotyping against the LacZ sequence), and both of them were almost completely silent. There was weak expression near the region of the primordial ear, but the intra-embryonic vasculature is totally negative (Figure 6.13 A). Interestingly, however, there is some very weak staining in the yolk sac, suggesting that in this extra-embryonic tissue, either Ets motifs are not required, or that the non-conserved and not particularly functional Ets sites still retained in the Nrp1in2peakA2 enhancer are capable of directing expression. Nevertheless, these results demonstrate that the loss of the functional Ets motifs is sufficient to abolish all intra-embryonic vascular expression and most extra-embryonic expression of the enhancer.

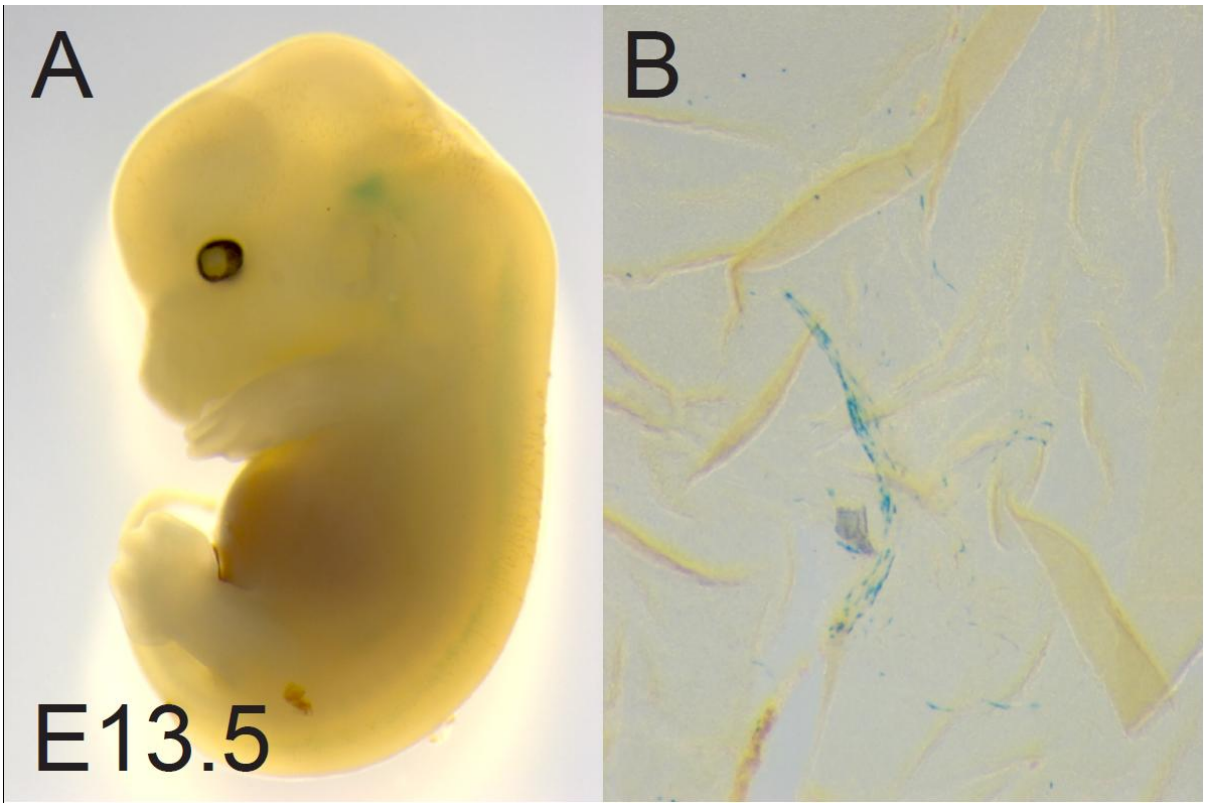


Figure 6.13 pA2FUNC Ets mut *LacZ* does not express in intra-embryonic vasculature.

A representative whole-mount embryo (A) transgenic for pA2 FUNC Ets mut construct has little visible X-gal staining, with the exception of a small amount in the large vessels of the yolk sac vasculature (B). n=2.

6.4.3. Role of Erg in hNrp1in2peakA enhancer expression

The Ets factors are well known for their essential roles in the vascular growth but have until recently not been implicated in specific expression patterns within the vasculature. However, research has now indicated that the Ets family member Erg may directly regulate arterial-specific gene expression (*203*).

Erg expression begins during mesodermal progenitor formation, but its expression is maintained until adulthood (*115*). Erg is involved in angiogenesis *in vitro* and *in vivo*, and has been implicated in a number of roles in endothelial cells, including the maintenance of endothelial homoeostasis and stability (*116*). A recent report has now demonstrated that ablation of ERG in endothelial cells results in loss of characteristic arterial marker gene expression in both mouse and zebrafish dorsal aortas (*203*). The effect was downstream of the VEGF signal induced by its ligand. Although the study focused on the role of Erg in regulating the Notch ligand Dll4, it was tantalizing to hypothesise that ERG may directly transcriptionally regulate VEGF transmembrane receptors. Therefore, I explored the possibility that ERG may directly regulate Nrp1 arterial expression through the Nrp1in2A2 enhancer. Through collaboration with Professor Anna Randi, I was permitted to cross the hNrp1in2peakA-Rosa mouse line with Erg flox mice, in which the essential coding region of the Erg gene was floxed. These mice were provided as $Erg^{fl/fl}$ homozygotes crossed into the Cdh5(PAC)-iCreERT2 mouse line. The Cdh5(PAC)-iCreERT2 mice originated from Ralf Adams, and consist of an inducible Cre gene directed by regulatory elements of the endothelial specific Cdh5 (Ve-cadherin) gene.

The aim of this work was to cross the Nrp1in2peakARosa with the $Erg^{fl/fl}$;Cdh5(PAC)-iCreERT2 mice to generate the compound genotype h$Nrp1in2peakA^{Rosa/+}$;$Erg^{fl/+}$;Cdh5(PAC)-iCreERT2. These mice would then be interbred, injected with tamoxifen to induce Cre at either E7 or E10, and sacrificed three days later. Embryos containing the genotype $Nrp1in2peakA^{Rosa/+}$;$Erg^{fl/fl}$;Cdh5(PAC)-iCreERT2 would be established by PCR genotyping, and the expression of the Nrp1 transgene in the Erg null background would be examined by X-gal staining. However, unfortunately, no mice of the correct genotype were successfully generated in the time allowed.

6.5.Fox and Gata site compound mutation of hNrp1in2peakA2 switch off vascular expression

To test the role of the Fox and Gata binding sites within the Nrp1in2A2 core enhancer, I have also created the construct ‘pA2 mut Fox mut Gata’, in which mutations were inserted into the Gata site and the functional Fox2 site (Figure 6.11). As before, this construct was used to generate transgenic mice. Two transgenic embryos were generated and neither demonstrated the stereotypical vascular expression in the subcutaneous blood vessels (Figure 6.14), as seen in hNrp1in2 *LacZ*, mNrp1in2 *LacZ* or hNrp1in2peakA2 *LacZ* embryos. Some X-gal expression was detected, but this was found to be non-vascular.



Figure 6.14 Mutation of both Fox and Gata sites in the hNrp1in2peakA2 core enhancer resulted in loss of vascular expression

Whole-mount embryos E13.5 transgenic for ‘pA2 mut Fox mut Gata’ construct had no subcutaneous vascular X-gal staining. Both of the two transgenic embryos are presented. n=2.

6.6. Conclusion

In this chapter the transcriptional regulators of Nrp1in2 have been addressed through the study of the 184bp conserved core regulatory element hNrp1in2peakA2. *In silico* analysis and *in vitro* validation of putative binding motifs within this regulatory element uncovered one *in vitro* validated Gata site, one *in vitro* validated Fox motif adjacent to an Ets motif and a cluster of three conserved and *in vitro* validated Ets sites. Reporter assays additionally demonstrated that the enhancer was synergistically activated by FoxC2 and Etv2, suggesting that the adjacent Fox and Ets sites formed a composite Fox:Ets motif. Mutational analysis revealed that mutations of either all the hNrpin2peakA2 Ets sites, or the four *in vitro* validated Ets, sites can both disrupt vascular expression of the enhancer, resulting in the loss of both the arterial and coronary expression. A compound mutation of Gata and Fox sites also prevented the enhancer from driving expression in the embryonic vasculature.

6.6.1. Functional Fox:Ets motif in the Nrp1 enhancer.

The composite Fox:Ets site plays an essential role in vascular endothelial cell transcription regulation. The Fox:Ets motif was found to be a powerful predictor of endothelial enhancer activity. In fact, this Nrp1 enhancer was first proposed to be an endothelial enhancer through a genomic Fox:Ets composite site search (*126*). The hNrp1in2peakA2 enhancer contains two putative Fox:Ets sites, and here I have demonstrated that the adjacent Fox2 and Ets7 sites form a functional Fox:Ets motif: the reporter assay demonstrated that over-expression of FoxC2 and ETV2 can synergistically activate this enhancer in a non-EC cellular environment. The Fox1 site adjacent to Ets5, which also has the sequence similarity of a Fox:Ets motif, was

found not to bind FoxC2 recombinant protein in EMSA, and was therefore deemed unlikely to be functional, although it might exert its function through FoxC1 or other Forkhead factors. Whether and how this conserved site may be involved in transcription regulation of hNrp1in2peakA2 is to be further studied.

While the FoxC and Ets partnership has been shown to be essential for vascular endothelial expression, Forkhead proteins may have other, more specific roles in the vasculature also. Although we have been using FoxC2 protein in the *in vitro* assays, other Fox factors may also be involved in transcription regulation of the hNrp1in2peakA2 enhancer, as both FoxC1 and FoxC2 have been shown to be important for arterial-specific expression (*218*). It has been suggested that FoxC factors are important for arterial regulation via the Dll4 promoter (*135*). However, these binding sites have now been shown to be dispensable for arterial-specific expression of Dll4, meaning that the direct downstream targets of FoxC in arteries are still debatable (*203, 65*). Ablation of the Nrp1in2PeakA functional FoxC site, in combination with a Gata site, was shown to disrupt vascular expression of the enhancer completely. There are a number of reasons this may have happened. Firstly, the mutation of the Fox site may have interrupted an essential composite Fox:Ets site, and therefore the loss of vascular expression may be downstream of this. It is worth noting that a similar mutation of the Fox site of the composite Fox:Ets motif in the Dll4in3 enhancer did not have a similar effect: expression was unaffected (*65*). Alternatively, it is possible that the FoxC site is responsible for directing arterial- specific expression. Even though FoxC1 and FoxC2 are expressed in both arteries and veins, the activity of these *trans* factors can be subject to phosphorylation or nuclear translocation modulation (*218*). However, the composite Fox:Ets site was shown to be of

importance in general endothelial differentiation and growth, and has little role implicated in arterial- or coronary- specific expression. Another tempting speculation would be that additional *trans* factors are needed for endothelial compartmentalisation.

However, with the current dataset, it is impossible at this point to rule out the possibility that the mutation of the functional Gata site is the cause of the loss of expression, since Gata factors have previously been shown to be essential for endothelial cell progenitor formation and differentiation too (*114*). It is therefore essential that we next elucidate whether the mutation of the Fox or the Gata site alone recapitulate the effects seen in the double knock-out, or whether only a combined mutation is responsible for the vascular expression.

6.6.2. ETS factor

Erg has been implicated in endothelial cell specification. Wythe *et al* demonstrated that Erg transcription regulates endothelial differentiation downstream or as part of the VEGF MAPK pathway (*203*). This transcription factor has been suggested to directly transcriptionally regulate Dll4 expression in the arterial compartment of the endothelium. Erg knockout mice are shown to lose arterial marker gene expression in the dorsal aorta, and consistent effects of Erg knockdown have also been reported in the zebrafish model. Since Ets factors obviously play an important role in Nrp1in2 regulation, it was therefore tempting to investigate the Erg involvement in VEGF pathway component Nrp1 transcription regulation. However, despite our best effort, we were unable to obtain mice transgenic for hNrp1peakA in an Ergfl/fl-Cdh5(PAC)-iCreERT2 mouse background..

Compound transgenic mouse breeding obstacle

As detailed in Figure 6.1, no mice of the correct genotype were successfully generated. Because the idea of creating this conditional knockout reporter cross arose only after the pA2 Ets mutational analyses results were generated, and that the mice were only made available fairly recently (these were generously provided by Prof Anna Randi), we were under time pressure to generate the compound genotype required (Nrp1in2peakA$^{Rosa/+}$;Erg$^{fl/fl}$;Cdh5(PAC)-iCreERT2/+). We therefore took the risk and made use of the F1 for inbreeding timed mating, instead of waiting until acquiring the F2 generation triple carrier and Erg$^{fl/fl}$ homozygotes, but were unlucky with genotypes. I eventually generated one male mouse (M155) of the preferred Nrp1in2peakA$^{Rosa/+}$; Ergfl$^{/fl}$; Cdh5(PAC)-iCreERT2/+genotype with which to perform another two inductions (Figure 6.1, orange). In total I received 53 embryos with a 1/16 chance each of the correct compound genotype and 12 in which I should have had a 3/16 chance of the correct genotype. This left us with the expectation of roughly:

$$53 \times \frac{1}{16} + 12 \times \frac{3}{16}$$

$= 5.56 \approx 6$ embryos with the desired genotype, and the chance not to have any one being:

$$\left(\frac{15}{16}\right)^{53} \times \left(\frac{13}{16}\right)^{12}$$

$= 0.271\%$, which is rather unlikely. I was confident of receiving sufficient numbers of samples for analysis. Unfortunately, I received none. It is however still likely that further litters will produce the genotype to provide the results I was hoping for.

Table6.1 $Erg^{fl/fl}$-Cdh5(PAC)iCreERT2-hNrp1in2peakA$^{Rosa/+}$ breading scheme

	$Erg^{fl/fl}$	iCreERT2	$Nrp1in2peakA^{R26/+}$			$Erg^{fl/fl}$	iCreERT2	$Nrp1in2peakA^{R26/+}$	
1	wt			M128xF139	**35**	wt			M125xF166
2	hyt	+			**36**	hyt	+	+	
3	wt			E10.5	**37**	hyt	+		E10.5
4	wt	+			**38**	wt			
5	homo				**39**	wt		+	
6	wt				**40**	wt		+	
7	wt				**41**	wt			
8	hyt	+			**42**	hyt		+	M128xF137
9	homo				**43**	homo	+		
10	hyt		+	M155xF151	**44**	homo	+		E10.5
11	homo	+			**45**	hyt			
12	hyt		+	E10.5	**46**	homo		+	
13	hyt	+	+		**47**	hyt			
14	hyt		+		**48**	wt			
15	hyt	+	+		**49**	hyt	+		
16	hyt		+		**50**	homo		+	
17	wt	+		M128xF186	**51**	hyt	+		M126xF141
18	hyt				**52**	wt		+	
19	hyt			E13.5	**53**	wt	+		E13.5
20	homo				**54**	wt		+	
21	wt	+	+		**55**	wt			
22	wt		+		**56**	hyt			
23	homo			M155xF158	**57**	homo	+		
24	hyt	+	+		**58**	wt			M126xF142
25	hyt	+		E13.5	**59**	wt			
26	homo	+			**60**	homo			E13.5
27	hyt	+	+		**61**	hyt			
28	hyt	+		M128xF156	**62**	wt	+		
29	homo				**63**	hyt		+	
30	homo			E13.5	**64**	homo	+		
31	hyt		+		**65**	wt		+	
32	wt								
33	homo		+	M125xF164		m128	EeCcnn	M126	EeccNn
34	wt			E13.5		M125	EeccNn	M155	EECcNn

Erg arterial specification targets could include Nrp1

It is likely that removal of only one Ets factor, Erg, may not actually reveal the precise function of this factor, due to redundancy. As has been mentioned previously, the ETS factors share a consensus recognition site, and are therefore often viewed as functionally redundant. Especially when the Ets factors are co-expressed, the role of one ETS can sometimes be substituted by another. During endothelial differentiation, there are various ETS factors expressed throughout. The way they constitute partnership with other *trans* factors has yet to be fully explored. With the exception of Etv2, which is essential for any endothelial cell formation, possibly because it is one of the first Ets to be expressed, there has been no particular evidence to single out one ETS factor to be a master regulator of vasculature differentiation. It is hence very likely that Erg may be indeed functional in arterial expression, but that loss of Erg expression alone may not result in loss of downstream gene expression. Our data suggests this may be true, as heterozygotes exhibiting a morphological defect did not switch off hNrp1PeakA-lacZ expression. Nevertheless, to position Erg under VEGF MAPK signalling and above direct transcription regulation over VEGF transmembrane receptor/co-receptor Nrp1 is tantalizing. In such a mechanism, Erg can be serving as the hub of a classical feed forward loop, activating transcription of Nrp1 under activation of MAPK mediated VEGF signal to auto-amplify this cytokine cue for arterial cell fate commitment and maintenance. Even if Erg does indeed directly regulate Nrp1 in this context, the cooperation of a Notch signal, SoxF factors and the impact of other Ets factors may still outweigh this hypothetical VEGF-ERG-Nrp1 loop in arterial identity. Nevertheless, the study of this potential transcription regulation relationship could be of great interest.

6.6.3. Coronary endothelial cell expression ablated in Ets and Fox-Gata mutation

The coronary expression of the hNrp1in2peakA2 enhancer is in the coronary artery and the vein, as well as the blood island-like area in the lower ventricle, and the endothelial cell clusters in the interventricular groove. Introduction of Ets mutations or compound Fox and Gata site mutations completely prevents any cardiac endothelial expression. In spite of the abundance of studies on the involvement of Fox, Ets or Gata factors in cardiac morphogenesis, their role in coronary endothelial cell differentiation is not well studied. It is possible that the Fox, Gata and Ets factors are rather essential even for the endothelial progenitors, making it difficult to study their timely role in coronary compartmentalisation. Mutational analysis of these sites from the hNrp1in2peakA2 enhancer will not answer the temporal function of these *cis* elements.

7. Discussion

7.1. Summary of results

In this thesis, I have described an enhancer within the second intron of Nrp1 that is capable of directing arterial and coronary EC-specific expression during embryonic development. Both orthologous human and mouse sequences of this enhancer were able to direct this specific pattern of expression. Further, a highly conserved core element of 184bp was sufficient to recapitulate the arterial- and coronary-specific expression pattern. Within this core element we identified and validated four functional Ets recognition sites, one of which forms a Fox:Ets composite site. These Ets binding sites were shown to be essential for the vascular expression of the enhancer, as mutation of the Ets binding motifs ablated the vascular expression driven by the enhancer. We further demonstrated that the Nrp1in2 core enhancer also contains functional binding sites for Gata proteins, and for FoxC factors, also previously thought to regulate arterial genes. Mutation of the Fox and Gata sites within this core enhancer also disrupted vascular specific expression of the transgene, suggesting that either one or both of these elements is essential for Nrp1in2 enhancer expression. The sinus venosus of the developing heart, the cellular origin of most of the coronary endothelial cells, also expressed the NRp1in2 transgene, as did the coronary plexus that derives from this area. This enhancer also marks a cluster of endothelial cells in the interventricular groove, but was never expressed in the endothelial cells in the endocardium.

7.2. Vascular Nrp1 expression under Ets factor transcription regulation

Nrp1 is the co-receptor of the VEGF signaling pathway, and the gene has been implicated in arterial and tip cell differentiation during vascular growth. Transcription regulation of the Nrp1 gene is of great interest in terms of vascular development regulation, cancer angiogenesis, wound healing and vascular medicine in general. Our enhancer recapitulates the endogenous arterial and coronary expression of Nrp1. By characterizing this enhancer we can discover the underlying transcription regulation mechanism. The fact that the 184bp core enhancer region is sufficient to drive the same vascular expression profile as the 1.8kb allows us to focus on the conserved motifs within. By characterizing the Ets, Fox and Gata sites through *in vitro* validation and *in vivo* mutational analyses, we manage to reach the conclusion that Ets, Fox and Gata are essential for vascular expression of Nrp1. Synergistic activation by FoxC2 and Etv2 is very likely to play a major role in transcription activation of Nrp1. We are unable to conclude whether the synergistic activation would be important at early stages or it will last into later phases of vascular growth. The technical limitations of enhancer characterization prevent us from being able to understand the role of the *trans* or *cis* factors in a temporal manner. We have failed to distinguish the arterial and coronary expression drivers in the enhancer, which may be due to the possibility that the transcription mechanisms controlling Nrp1 expression overlap between the two processes; or perhaps that there are other transcription regulatory elements in the gene responsible for this distinction. It is fair to suggest that this enhancer is not the only regulatory element of Nrp1, since there are various non-coding regions in the gene that are enriched with epigenetic markers for

favorable chromosomal configurations. There was also a report showing that there is a regulatory element near the core promoter region of the gene that can be activated *in vitro*, although the functionality of the regulatory element in any animal model is yet to be confirmed (*216*). Nevertheless, it is fair to say that the limitation of studying one enhancer of a gene prevent us from making an exhaustive conclusion on the regulation of Nrp1 vascular expression.

In addition, we are unable to draw conclusions on neuronal transcription regulation of Nrp1, since our enhancer does not express extensively in this regions as the endogenous Nrp1 does. It is likely that there is another (set of) enhancer(s) in the gene responsible for the neuronal expression, which would incorporate neuronal-specific *cis* elements; although it would be interesting to find out whether or how these enhancers collaborate, especially when faced with similar cellular processes (axon/tip cell guidance) or signaling pathways (SEMA signaling).

7.3. Arterial specific expression regulation

The tip cell expression of Nrp1 requires dynamic adaptation to the situation of the cell in real-time, since tip and stalk cell competition results in rapid exchange of role and relative position in the sprouts. The activity of a cell membrane receptor may be better regulated by phosphorylation-mediated recruitment to the receptor, or translocation between the cell membrane and endocytosis vesicle, to face the quickly changing environment. Whereas in the artery, the established identity needs a more stable gene expression profile to be maintained as well as further maturation processes. In this sense, studying arterial transcription regulation

through enhancer characterization is more realistic. It has been suggested that the master venous transcription regulator CoupTFII suppresses Nrp1 expression in the vein, and such regulation may be direct (*26*), however, there has not been any *in vivo* evidence to confirm this notion. Our Nrp1 enhancer demonstrates arterial specific expression that is similar to the endogenous Nrp1 expression profile. The 184bp hNrp1in2peakA2 does not have a discernible CoupTFII motif yet arterial specificity is achieved nonetheless. This suggests that CoupTFII suppression is not the mechanism for Nrp1 arterial-specific expression, at least not the only one.

7.3.1. Sox and Notch

Dr Sacilotto, in our lab, has demonstrated that SoxF and Rbpj also combinatorially regulate arterial-specific expression in an enhancer that has functional Fox:Ets sites. Compound knockdown of SoxF factors and Rbpj in fish ablated dorsal aorta formation. The Rbpj-mediated Notch pathway and the SoxF sub-family factors have been shown to have a large impact on arterial specificity (*65*). The Rbpj-Sox partnership, however, does not have a direct target in this Nrp1 arterial enhancer. This hNrp1in2peakA2 enhancer does not have Rbpj or SoxF sites, yet arterial-specific expression is achieved anyway. It is possible that Rbpj-Sox do have an effect on other components of the VEGF pathway, and regulate Nrp1 arterial specific expression indirectly. Future exploration of enhancer characterisation of hNrp1in2peakA2 will likely shed light on our understanding of Ets factor collaboration with other *trans* factors during endothelial specialization.It also suggests that Rbpj-Sox is not necessarily the only way to define arterial expression.

7.3.2. Ets factors in EC compartmentalization

Since Etv2 expression diminishes as the vasculature matures, other Ets factors need to be implicated in later stages if ETS *cis* elements are found to be essential for arterial-venous differentiation. Historically Ets factors are considered on top of the hierarchy of transcription regulation during vascular development, since they start expression early in the mesodermal progenitor cells and are responsible for activating and collaborate with the early transcription regulators (*112*). Very recently, a report has challenged this notion and showed that Ets factors collaboratively regulate VEGF-directed Dll4 transcription during arterial specification (*203*). The Ets factor ERG, in particular, has been found to be associated with arterial endothelial cell identity. Over-expression of ERG in mouse up-regulates the expression of arterial marker Efnb2, and disruption of the gene is accompanied by aortic vessels losing their characteristic expression both in zebrafish and in mouse. It is worth noticing that a cohort of Ets factors are also expressed at this stage in endothelial cells, and the fact that the Ets recognition site can be easily shared among the Ets factors raises the question whether and how Erg would preferentially function is at the presence of Fli1, Ets1, and Elf1 to name just a few. To regulate hNrp1in2peakA2 expression, a responsible Ets factor would most likely be partnered with FoxC and Gata, and this combinatorial regulation dynamic will be the key for arterial- specific expression. Similar partnerships have been implicated in haemangioblast differentiation and haematopoiesis processes (*159, 114*); it will be interesting to find out how the similar mechanism would be incorporated in arterial EC specification, and how the participating factors with overlapping expression exert their expected specific functions in each particular cellular scenarios.

7.4. Nrp1 in coronary endothelial cell differentiation

7.4.1. Multiple cellular origins of the coronary endothelial cells

There have been various hypotheses about the cellular origin of the coronary endothelial cells. PEO, epicardium, endocardium and sinus venosus are all suggested to form the coronary endothelial cells, and the current consensus is that coronary endothelial cells are of multiple origins. We have never detected PEO or epicardium expression of this enhancer. Additionally, previous report has shown that the majority of the PEO derived cells contribute to the smooth muscle cell layer of the coronary, making it unlikely that the PEO will be making significant contribution to the coronary endothelial cell population. Our results are consistent with this scenario.

The sinus venosus-originated plexus makes up a large proportion of the coronary endothelial cells. Our hNrp1in2, mNrp1in2 and hNrp1peakA all express in this downwards migrating and growing plexus, and both the outer layer of coronary veins and inner layer of coronary arteries are positive for X-Gal staining. As the sinus venosus endothelial cells express the venous marker CoupTFII (*40*), and as the sinus is part of the collecting inflow tract, it was considered likely to be committed venous endothelial cells. However, it is worth noticing that these cells are also positive for Nrp1 (*24*), and the venous commitment is soon reversed once the coronary forming cells migrate out of the region. It is somewhat counterintuitive that Nrp1 and Coup-TFII co-express, since various reports have suggested that Coup-TFII directly suppresses Nrp1 expression. Our enhancer is expressed in the sinus venosus, and the hNrp1in2peakA enhancer does not have a clear Coup-TFII recognition site. It is possible that

the Coup-TFII suppression is an indirect mechanism; or that our enhancer circumvents the Coup-TFII suppression in this scenario. This mechanism may lay at the foundation of the transcription regulation of reversing the venous commitment into an arterial coronary endothelial- forming differentiation path.

We have mentioned that the interventricular groove has X-Gal staining in some of our lines, i.e. hNrp1in2 Rosa, mNrp1in2 Rosa, and hNrp1in2peakA Rosa. The surface of the lower ventricles also sees separated sparse blood island-like structures positive in these lines. Both of these cell types are in a morphological structure that rises from the apex of the heart ventricle, and therefore are definitely discontinuous with the sinus venosus-originated plexus. It is fair to conclude that these cells are of a different origin from the sinus venosus cells. A recent report has suggested that these cells are of an endocardial origin, and that the migration and angiogenic commitment of these cells are responsive to VEGF signalling a tissue culture model (*41*). As a VEGF receptor/co-receptor, expression of Nrp1 in these endocardial cells leads to sensing of VEGF, and therefore migration and differentiation would have been a perfect fit for the model. The endogenous Nrp1 gene is indeed expressed in endocardial endothelial cells (*24*). However, it is worth noticing that none of our lines have demonstrated endocardial expression of the enhancer. It is possible that the *in vitro* model in the previous report did not reflect the reality of embryonic development; it is also likely that the initial VEGF-sensing endocardial cells did not need Nrp1 (using VEGFR only), or at least not the activity of this enhancer; yet later the migrated endothelial cells maintain Nrp1 expression through the activity of this enhancer, to feed the further differentiation of these endothelial cells into coronary plexus.

Nevertheless, our results are consistent with the notion that the coronary endothelium is of multiple cellular origin. As the sinus venosus appears to be a major contributor, the endocardium is likely to make up a fraction of the low ventricles and contribute to the interventricular groove coronary endothelium via migration and angiogenesis.

7.4.2. Transcription regulation of coronary endothelial cell differentiation

Given the complexity and controversy of the cellular origin of the coronary endothelial cells, and the morphological intricacy of cardiac development, it is not surprising that coronary vascular growth is impacted and confounded by multiple factors, and that the transcription regulation mechanism of the coronary endothelial cells is understudied. FoxO factor ablation results in coronary plexus malformation, but the FoxO effect has been reported to be an indirect one, through impairing vascular smooth muscle cells and under-supported coronary ECs (*219*). Epicardium-specific knockout of Gata4 and 6 reduces the endothelial cell migration into the sub-epicardial region, and the coronary plexus is severely compromised (*220*). It is interesting that epicardium-specific knockout of Gata factors does not affect coronary endothelial cell proliferation, leading to the hypothesis that the transcription regulation of coronary endothelial cell differentiation and cellular activity to be controlled in a stepwise manner. As previously mentioned, the CoupTFII positive sinus cells express Nrp1 and hNrp1in2peakA2, and some of these cells eventually transdifferentiate into coronary arterial EC. The Fox and Gata sites contained in the hNrp1in2peakA2 enhancer are recognized in both coronary arteries and veins, making it very interesting to discover whether and how they may play a role in the transdifferentiation process.

Bibliography

1. J. E. Ferguson, R. W. Kelley, C. Patterson, Mechanisms of Endothelial Differentiation in Embryonic Vasculogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 25, 2246-2254 (2005).
2. K. Choi, M. Kennedy, A. Kazarov, J. C. Papadimitriou, G. Keller, A common precursor for hematopoietic and endothelial cells. *Development* 125, 725-732 (1998).
3. S. I. Nishikawa, In vitro generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. *Immunity* 8, 761-769 (1998).
4. R. Wagner, Endothelial cell embryology and growth. *Advances in Microcirculation* 9, (1980).
5. C. Lancrin, P. Sroczynska, C. Stephenson, T. Allen, V. Kouskoff, G. Lacaud, The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* 457, 892-895 (2009)
6. V. Garcia-Martinez, G. C. Schoenwolf, Primitive-Streak Origin of the Cardiovascular System in Avian Embryos. *Developmental Biology* 159, 706-719 (1993).
7. J. Palis, S. Robertson, M. Kennedy, C. Wall, G. Keller, Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* 126, 5073-5084 (1999).
8. K. McGrath, J. Palis, in *Current Topics in Developmental Biology*, J. B. James, Ed. (Academic Press, 2008), vol. Volume 82, pp. 1-22.
9. B. Vandenbunder, L. Pardanaud, T. Jaffredo, M. A. Mirabel, D. Stehelin, Complementary patterns of expression of c-ets 1, c-myb and c-myc in the blood-forming system of the chick embryo. *Development* 107, 265-274 (1989).
10. K. M. Vogeli, S. W. Jin, G. R. Martin, D. Y. Stainier, A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. *Nature* 443, 337-339 (2006).
11. T. Asahara, A. Kawamoto, H. Masuda, Concise Review: Circulating Endothelial Progenitor Cells for Vascular Medicine. *STEM CELLS* 29, 1650-1655 (2011)10.1002/stem.745).
12. J. Rehman, J. Li, C. M. Orschell, K. L. March, Peripheral Blood "Endothelial Progenitor Cells" Are Derived From Monocyte/Macrophages and Secrete Angiogenic Growth Factors. *Circulation* 107, 1164-1169 (2003).
13. F. Timmermans, J. Plum, M. C. Yöder, D. A. Ingram, B. Vandekerckhove, J. Case, Endothelial progenitor cells: identity defined? *Journal of Cellular and Molecular Medicine* 13, 87-102 (2009).

14. S. Kwei, G. Stavrakis, M. Takahas, G. Taylor, M. J. Folkman, M. A. Gimbrone, Jr., G. Garcia-Cardena, Early adaptive responses of the vascular wall during venous arterialization in mice. *The American journal of pathology* 164, 81-89 (2004).
15. F. le Noble, D. Moyon, L. Pardanaud, L. Yuan, V. Djonov, R. Matthijsen, C. Breant, V. Fleury, A. Eichmann, Flow regulates arterial-venous differentiation in the chick embryo yolk sac. *Development* 131, 361-375 (2004).
16. R. H. Adams, K. Alitalo, Molecular regulation of angiogenesis and lymphangiogenesis. *Nat Rev Mol Cell Biol* 8, 464-478 (2007).
17. M. Hellstrom, L.-K. Phng, J. J. Hofmann, E. Wallgard, L. Coultas, P. Lindblom, J. Alva, A.-K. Nilsson, L. Karlsson, N. Gaiano, K. Yoon, J. Rossant, M. L. Iruela-Arispe, M. Kalen, H. Gerhardt, C. Betsholtz, Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature* 445, 776-780 (2007).
18. H. M. Eilken, R. H. Adams, Dynamics of endothelial cell behavior in sprouting angiogenesis. *Current opinion in cell biology* 22, 617-625 (2010).
19. L. K. Phng, H. Gerhardt, Angiogenesis: a team effort coordinated by notch. *Dev Cell* 16, 196-208 (2009).
20. M. Potente, H. Gerhardt, P. Carmeliet, Basic and Therapeutic Aspects of Angiogenesis. *Cell* 146, 873-887 (2011)
21. V. Djonov, M. Schmid, S. A. Tschanz, P. H. Burri, Intussusceptive Angiogenesis: Its Role in Embryonic Vascular Network Formation. *Circulation Research* 86, 286-292 (2000)
22. P. H. Burri, R. Hlushchuk, V. Djonov, Intussusceptive angiogenesis: Its emergence, its characteristics, and its significance. *Developmental Dynamics* 231, 474-488 (2004)
23. V. Djonov, A. C. Andres, A. Ziemiecki, Vascular remodelling during the normal and malignant life cycle of the mammary gland. *Microsc Res Tech* 52, 182-189 (2001).
24. D. C. Chong, Y. Koo, K. Xu, S. Fu, O. Cleaver, Stepwise arteriovenous fate acquisition during mammalian vasculogenesis. *Developmental dynamics : an official publication of the American Association of Anatomists* 240, 2153-2165 (2011).
25. H. Lindskog, Y. H. Kim, E. B. Jelin, Y. Kong, S. Guevara-Gallardo, T. N. Kim, R. A. Wang, Molecular identification of venous progenitors in the dorsal aorta reveals an aortic origin for the cardinal vein in mammals. *Development* 141, 1120-1128 (2014).
26. N. D. Lawson, A. M. Vogel, B. M. Weinstein, sonic hedgehog and vascular endothelial growth factor Act Upstream of the Notch Pathway during Arterial Endothelial Differentiation. *Developmental Cell* 3, 127-136 (2002).
27. C. Williams, S.-H. Kim, T. T. Ni, L. Mitchell, H. Ro, J. S. Penn, S. H. Baldwin, L. Solnica-Krezel, T. P. Zhong, Hedgehog signaling induces arterial endothelial cell formation by repressing venous cell fate. *Developmental Biology* 341, 196-204 (2010).
28. S. P. Herbert, J. Huisken, T. N. Kim, M. E. Feldman, B. T. Houseman, R. A. Wang, K. M. Shokat, D. Y. R. Stainier, Arterial-Venous Segregation by Selective Cell Sprouting: An Alternative Mode of Blood Vessel Formation. *Science* 326, 294-298 (2009).

29. V. Kohli, Jennifer A. Schumacher, Sharina P. Desai, K. Rehn, S. Sumanas, Arterial and Venous Progenitors of the Major Axial Vessels Originate at Distinct Locations. *Developmental Cell* 25, 196-206 (2013).
30. J. S. Lee, Q. Yu, J. T. Shin, E. Sebzda, C. Bertozzi, M. Chen, P. Mericko, M. Stadtfeld, D. Zhou, L. Cheng, T. Graf, C. A. MacRae, J. J. Lepore, C. W. Lo, M. L. Kahn, Klf2 Is an Essential Regulator of Vascular Hemodynamic Forces In Vivo. *Developmental Cell* 11, 845-857 (2006).
31. N. D. Lawson, B. M. Weinstein, Arteries and veins: making a difference with zebrafish. *Nat Rev Genet* 3, 674-682 (2002).
32. H. U. Wang, Z.-F. Chen, D. J. Anderson, Molecular Distinction and Angiogenic Interaction between Embryonic Arteries and Veins Revealed by ephrin-B2 and Its Receptor Eph-B4. *Cell* 93, 741-753 (1998).
33. R. S. Srinivasan, M. E. Dillard, O. V. Lagutin, F.-J. Lin, S. Tsai, M.-J. Tsai, I. M. Samokhvalov, G. Oliver, Lineage tracing demonstrates the venous origin of the mammalian lymphatic vasculature. *Genes & Development* 21, 2422-2432 (2007).
34. K. Yaniv, S. Isogai, D. Castranova, L. Dye, J. Hitomi, B. M. Weinstein, Live imaging of lymphatic development in the zebrafish. *Nat Med* 12, 711-716 (2006).
35. F.-J. Lin, X. Chen, J. Qin, Y.-K. Hong, M.-J. Tsai, S. Y. Tsai, Direct transcriptional regulation of neuropilin-2 by COUP-TFII modulates multiple steps in murine lymphatic vessel development. *The Journal of Clinical Investigation* 120, 1694-1707 (2010).
36. W. Cai, W. Schaper, Mechanisms of arteriogenesis. *Acta Biochimica et Biophysica Sinica* 40, 681-692 (2008).
37. J. M. Perez-Pomares, D. Macias, L. Garcia-Garrido, R. Munoz-Chapuli, The origin of the subepicardial mesenchyme in the avian embryo: an immunohistochemical and quail-chick chimera study. *Dev Biol* 200, 57-68 (1998).
38. M. P. Vrancken Peeters, A. C. Gittenberger-de Groot, M. M. Mentink, R. E. Poelmann, Smooth muscle cells and fibroblasts of the coronary arteries derive from epithelial-mesenchymal transformation of the epicardium. *Anat Embryol* 199, 367-378 (1999).
39. C.-L. Cai, J. C. Martin, Y. Sun, L. Cui, L. Wang, K. Ouyang, L. Yang, L. Bu, X. Liang, X. Zhang, W. B. Stallcup, C. P. Denton, A. McCulloch, J. Chen, S. M. Evans, A myocardial lineage derives from Tbx18 epicardial cells. *Nature* 454, 104-108 (2008).
40. K. Red-Horse, H. Ueno, I. L. Weissman, M. A. Krasnow, Coronary arteries form by developmental reprogramming of venous cells. *Nature* 464, 549-553 (2010)http://www.nature.com/nature/journal/v464/n7288/suppinfo/nature08873_S1.html).
41. B. Wu, Z. Zhang, W. Lui, X. Chen, Y. Wang, A. A. Chamberlain, R. A. Moreno-Rodriguez, R. R. Markwald, B. P. O'Rourke, D. J. Sharp, D. Zheng, J. Lenz, H. S. Baldwin, C. P. Chang, B. Zhou, Endocardial cells form the coronary arteries by angiogenesis through myocardial-endocardial VEGF signaling. *Cell* 151, 1083-1096 (2012).

42. X. Tian, T. Hu, H. Zhang, L. He, X. Huang, Q. Liu, W. Yu, Z. Yang, Y. Yan, X. Yang, T. P. Zhong, W. T. Pu, B. Zhou, Vessel formation. De novo formation of a distinct coronary vascular population in neonatal heart. *Science* 345, 90-94 (2014).
43. X. Tian, T. Hu, H. Zhang, L. He, X. Huang, Q. Liu, W. Yu, Z. Yang, Z. Zhang, T. P. Zhong, X. Yang, Y. Yan, A. Baldini, Y. Sun, J. Lu, R. J. Schwartz, S. M. Evans, A. C. Gittenberger-de Groot, K. Red-Horse, B. Zhou, Subepicardial endothelial cells invade the embryonic ventricle wall to form coronary arteries. *Cell Res* 23, 1075-1090 (2013).
44. T. Nagase, M. Nagase, M. Machida, T. Fujita, Hedgehog signalling in vascular development. *Angiogenesis* 11, 71-77 (2008).
45. T. Minke van, G. Freek, W. Jinxia, K. Maciek, L. Jason, T. Dick, P. Martin, Angiogenic factors stimulate tubular branching morphogenesis of sonic hedgehog-deficient lungs. *Developmental Biology* 303, (2007).
46. K. J. Lavine, A. C. White, C. Park, C. S. Smith, K. Choi, F. Long, C.-c. Hui, D. M. Ornitz, Fibroblast growth factor signals regulate a wave of Hedgehog activation that is essential for coronary vascular development. *Genes & Development* 20, 1651-1666 (2006).
47. K. Shigeru, M. Yasushi, S. Takashi, M. Yasuyoshi, N. Koichiro, K. Hiroshi, Sonic Hedgehog Induces Capillary Morphogenesis by Endothelial Cells through Phosphoinositide 3-Kinase. *Journal of Biological Chemistry* 278, 8244-8249 (2003).
48. J. S. Hochman, G. A. Lamas, C. E. Buller, V. Dzavik, H. R. Reynolds, S. J. Abramsky, S. Forman, W. Ruzyllo, A. P. Maggioni, H. White, Z. Sadowski, A. C. Carvalho, J. M. Rankin, J. P. Renkin, P. G. Steg, A. M. Mascette, G. Sopko, M. E. Pfisterer, J. Leor, V. Fridrich, D. B. Mark, G. L. Knatterud, Coronary Intervention for Persistent Occlusion after Myocardial Infarction. *New England Journal of Medicine* 355, 2395-2407 (2006).
49. J. Asai, H. Takenaka, K. F. Kusano, M. Ii, C. Luedemann, C. Curry, E. Eaton, A. Iwakura, Y. Tsutsumi, H. Hamada, S. Kishimoto, T. Thorne, R. Kishore, D. W. Losordo, Topical sonic hedgehog gene therapy accelerates wound healing in diabetes by enhancing endothelial progenitor cell-mediated microvascular remodeling. *Circulation* 113, 2413-2424 (2006).
50. R. P. Visconti, C. D. Richardson, T. N. Sato, Orchestration of angiogenesis and arteriovenous contribution by angiopoietins and vascular endothelial growth factor (VEGF). *Proc Natl Acad Sci U S A* 99, 8219-8224 (2002).
51. C. Gu, E. R. Rodriguez, D. V. Reimert, T. Shu, B. Fritzsch, L. J. Richards, A. L. Kolodkin, D. D. Ginty, Neuropilin-1 Conveys Semaphorin and VEGF Signaling during Neural and Cardiovascular Development. *Developmental Cell* 5, 45-57 (2003).
52. Z. J. Liu, T. Shirakawa, Y. Li, A. Soma, M. Oka, G. P. Dotto, R. M. Fairman, O. C. Velazquez, M. Herlyn, Regulation of Notch1 and Dll4 by vascular endothelial growth factor in arterial endothelial cells: implications for modulating arteriogenesis and angiogenesis. *Mol Cell Biol* 23, 14-25 (2003).

53. C. C. Hong, Q. P. Peterson, J.-Y. Hong, R. T. Peterson, Artery/Vein Specification Is Governed by Opposing Phosphatidylinositol-3 Kinase and MAP Kinase/ERK Signaling. *Current Biology* 16, 1366-1372 (2006).
54. N. D. Lawson, J. W. Mugford, B. A. Diamond, B. M. Weinstein, phospholipase C gamma-1 is required downstream of vascular endothelial growth factor during arterial development. *Genes & Development* 17, 1346-1351 (2003).
55. M. J. Karkkainen, P. Haiko, K. Sainio, J. Partanen, J. Taipale, T. V. Petrova, M. Jeltsch, D. G. Jackson, M. Talikka, H. Rauvala, C. Betsholtz, K. Alitalo, Vascular endothelial growth factor C is required for sprouting of the first lymphatic vessels from embryonic veins. *Nat Immunol* 5, 74-80 (2004).
56. E. R. Andersson, R. Sandberg, U. Lendahl, Notch signaling: simplicity in design, versatility in function. *Development* 138, 3593-3612 (2011).
57. T. Iso, L. Kedes, Y. Hamamori, HES and HERP families: multiple effectors of the Notch signaling pathway. *J Cell Physiol* 194, 237-255 (2003).
58. J. R. Shutter, S. Scully, W. Fan, W. G. Richards, J. Kitajewski, G. A. Deblandre, C. R. Kintner, K. L. Stark, Dll4, a novel Notch ligand expressed in arterial endothelium. *Genes & Development* 14, 1313-1318 (2000).
59. R. Benedito, C. Roca, I. Sorensen, S. Adams, A. Gossler, M. Fruttiger, R. H. Adams, The notch ligands Dll4 and Jagged1 have opposing effects on angiogenesis. *Cell* 137, 1124-1135 (2009).
60. A. Duarte, M. Hirashima, R. Benedito, A. Trindade, P. Diniz, E. Bekman, L. Costa, D. Henrique, J. Rossant, Dosage-sensitive requirement for mouse Dll4 in artery development. *Genes Dev* 18, 2474-2478 (2004).
61. A. Fischer, N. Schumacher, M. Maier, M. Sendtner, M. Gessler, The Notch target genes Hey1 and Hey2 are required for embryonic vascular development. *Genes & Development* 18, 901-911 (2004).
62. H. Kokubo, S. Miyagawa-Tomita, M. Nakazawa, Y. Saga, R. L. Johnson, Mouse hesr1 and hesr2 genes are redundantly required to mediate Notch signaling in the developing cardiovascular system. *Developmental Biology* 278, 301-309 (2005).
63. L. T. Krebs, J. R. Shutter, K. Tanigaki, T. Honjo, K. L. Stark, T. Gridley, Haploinsufficient lethality and formation of arteriovenous malformations in Notch pathway mutants. *Genes and Development* 18, 2469-2473 (2004).
64. V. Caolo, N. M. van den Akker, S. Verbruggen, M. M. Donners, G. Swennen, H. Schulten, J. Waltenberger, M. J. Post, D. G. Molin, Feed-forward signaling by membrane-bound ligand receptor circuit: the case of NOTCH DELTA-like 4 ligand in endothelial cells. *J Biol Chem* 285, 40681-40689 (2010).
65. N. Sacilotto, R. Monteiro, M. Fritzsche, P. W. Becker, L. Sanchez-Del-Campo, K. Liu, P. Pinheiro, I. Ratnayaka, B. Davies, C. R. Goding, R. Patient, G. Bou-Gharios, S. De Val, Analysis of Dll4 regulation reveals a combinatorial role for Sox and Notch in arterial development. *Proc Natl Acad Sci U S A* 110, 11893-11898 (2013).
66. R. H. Adams, G. A. Wilkinson, C. Weiss, F. Diella, N. W. Gale, U. Deutsch, W. Risau, R. Klein, Roles of ephrinB ligands and EphB receptors in cardiovascular

development: demarcation of arterial/venous domains, vascular morphogenesis, and sprouting angiogenesis. *Genes & Development* 13, 295-306 (1999).
67. S. S. Gerety, H. U. Wang, Z. F. Chen, D. J. Anderson, Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development. *Mol Cell* 4, 403-414 (1999).
68. D. Arvanitis, A. Davy, Eph/ephrin signaling: networks. *Genes & Development* 22, 416-429 (2008).
69. J. E. Chrencik, A. Brooun, M. I. Recht, M. L. Kraus, M. Koolpe, A. R. Kolatkar, R. H. Bruce, G. Martiny-Baron, H. Widmer, E. B. Pasquale, P. Kuhn, Structure and thermodynamic characterization of the EphB4/Ephrin-B2 antagonist peptide complex reveals the determinants for receptor specificity. *Structure* 14, 321-330 (2006).
70. O. Salvucci, M. de la Luz Sierra, J. A. Martina, P. J. McCormick, G. Tosato, EphB2 and EphB4 receptors forward signaling promotes SDF-1–induced endothelial cell chemotaxis and branching remodeling. *Blood*, (2006).
71. M. E. Pitulescu, R. H. Adams, Eph/ephrin molecules--a hub for signaling and endocytosis. *Genes Dev* 24, 2480-2492 (2010).
72. J. Torres-Vazquez, M. Kamei, B. M. Weinstein, Molecular distinction between arteries and veins. *Cell Tissue Res* 314, 43-59 (2003).
73. R. E. Lamont, S. Childs, MAPping out arteries and veins. *Sci STKE* 3, (2006).
74. S. Sawamiphak, S. Seidel, C. L. Essmann, G. A. Wilkinson, M. E. Pitulescu, T. Acker, A. Acker-Palmer, Ephrin-B2 regulates VEGFR2 function in developmental and tumour angiogenesis. *Nature* 465, 487-491 (2010).
75. R. Wang, K. Chadalavada, J. Wilshire, U. Kowalik, K. E. Hovinga, A. Geber, B. Fligelman, M. Leversha, C. Brennan, V. Tabar, Glioblastoma stem-like cells give rise to tumour endothelium. *Nature* 468, 829-833 (2010).
76. H. Haegel, L. Larue, M. Ohsugi, L. Fedorov, K. Herrenknecht, R. Kemler, Lack of beta-catenin affects mouse development at gastrulation. *Development* 121, 3529-3537 (1995).
77. J. Huelsken, R. Vogel, V. Brinkmann, B. Erdmann, C. Birchmeier, W. Birchmeier, Requirement for β-Catenin in Anterior-Posterior Axis Formation in Mice. *The Journal of Cell Biology* 148, 567-578 (2000).
78. R. Daneman, D. Agalliu, L. Zhou, F. Kuhnert, C. J. Kuo, B. A. Barres, Wnt/beta-catenin signaling is required for CNS, but not non-CNS, angiogenesis. *Proc Natl Acad Sci U S A* 106, 641-646 (2009).
79. M. F. Morini, E. Dejana, Transcriptional regulation of arterial differentiation via Wnt, Sox and Notch. *Curr Opin Hematol* 21, 229-234 (2014).
80. M. Corada, D. Nyqvist, F. Orsenigo, A. Caprini, C. Giampietro, M. M. Taketo, M. L. Iruela-Arispe, R. H. Adams, E. Dejana, The Wnt/beta-catenin pathway modulates vascular remodeling and specification by upregulating Dll4/Notch signaling. *Dev Cell* 18, 938-949 (2010).
81. K. Yamamizu, T. Matsunaga, H. Uosaki, H. Fukushima, S. Katayama, M. Hiraoka-Kanie, K. Mitani, J. K. Yamashita, Convergence of Notch and β-catenin

signaling induces arterial fate in vascular progenitors. *The Journal of Cell Biology* 189, 325-338 (2010).
82. C. H. Heldin, K. Miyazono, P. ten Dijke, TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature* 390, 465-471 (1997).
83. R. N. Wilkinson, C. Pouget, M. Gering, A. J. Russell, S. G. Davies, D. Kimelman, R. Patient, Hedgehog and Bmp polarize hematopoietic stem cell emergence in the zebrafish dorsal aorta. *Dev Cell* 16, 909-916 (2009).
84. M. J. Goumans, G. Valdimarsdottir, S. Itoh, A. Rosendahl, P. Sideras, P. ten Dijke, Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors. *Embo J* 21, 1743-1753 (2002).
85. F. Lebrin, M. J. Goumans, L. Jonker, R. L. Carvalho, G. Valdimarsdottir, M. Thorikay, C. Mummery, H. M. Arthur, P. ten Dijke, Endoglin promotes endothelial cell proliferation and TGF-beta/ALK1 signal transduction. *Embo J* 23, 4018-4028 (2004).
86. C. Castanares, M. Redondo-Horcajo, N. Magan-Marchal, P. ten Dijke, S. Lamas, F. Rodriguez-Pascual, Signaling by ALK5 mediates TGF-beta-induced ET-1 expression in endothelial cells: a role for migration and proliferation. *J Cell Sci* 120, 1256-1266 (2007).
87. M. J. Goumans, A. J. van Zonneveld, P. ten Dijke, Transforming growth factor beta-induced endothelial-to-mesenchymal transition: a switch to cardiac fibrosis? *Trends Cardiovasc Med* 18, 293-298 (2008).
88. H.-G. Kang, M. H. Chae, J. M. Park, E. J. Kim, J. H. Park, S. Kam, S. I. Cha, C. H. Kim, R.-W. Park, S. H. Park, Y. L. Kim, I.-S. Kim, T. H. Jung, J. Y. Park, Polymorphisms in TGF-β1 gene and the risk of lung cancer. *Lung Cancer* 52, 1-7 (2006).
89. D. M. Wiley, J.-D. Kim, J. Hao, C. C. Hong, V. L. Bautch, S.-W. Jin, Distinct signalling pathways regulate sprouting angiogenesis from the dorsal aorta and the axial vein. *Nat Cell Biol* 13, 686-692 (2011).
90. T. Seki, J. Yun, S. P. Oh, Arterial endothelium-specific activin receptor-like kinase 1 expression suggests its role in arterialization and vascular remodeling. *Circ Res* 93, 682-689 (2003).
91. M. Kiyono, M. Shibuya, Bone Morphogenetic Protein 4 Mediates Apoptosis of Capillary Endothelial Cells during Rat Pupillary Membrane Regression. *Molecular and Cellular Biology* 23, 4627-4636 (2003).
92. A. Shilatifard, R. C. Conaway, J. W. Conaway, THE RNA POLYMERASE II ELONGATION COMPLEX. *Annual Review of Biochemistry* 72, 693-715 (2003).
93. S. T. Whiteside, S. Goodbourn, Signal transduction and nuclear targeting: regulation of transcription factor activity by subcellular localisation. *J Cell Sci* 104, 949-955 (1993).
94. D. Shlyueva, G. Stampfel, A. Stark, Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet* 15, 272-286 (2014).
95. M. Z. Ludwig, A. Palsson, E. Alekseeva, C. M. Bergman, J. Nathan, M. Kreitman, Functional Evolution of a cis Regulatory Module. *PLoS Biol* 3, e93 (2005).

96. R. Nhili, P. Peixoto, S. Depauw, S. Flajollet, X. Dezitter, M. M. Munde, M. A. Ismail, A. Kumar, A. A. Farahat, C. E. Stephens, M. Duterque-Coquillaud, W. David Wilson, D. W. Boykin, M.-H. David-Cordonnier, Targeting the DNA-binding activity of the human ERG transcription factor using new heterocyclic dithiophene diamidines. *Nucleic Acids Research* 41, 125-138 (2013).
97. D. T. Odom, R. D. Dowell, E. S. Jacobsen, W. Gordon, T. W. Danford, K. D. MacIsaac, P. A. Rolfe, C. M. Conboy, D. K. Gifford, E. Fraenkel, Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39, 730-732 (2007).
98. L. A. Pennacchio, N. Ahituv, A. M. Moses, S. Prabhakar, M. A. Nobrega, M. Shoukry, S. Minovitsky, I. Dubchak, A. Holt, K. D. Lewis, I. Plajzer-Frick, J. Akiyama, S. De Val, V. Afzal, B. L. Black, O. Couronne, M. B. Eisen, A. Visel, E. M. Rubin, In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444, 499-502 (2006).
99. M. J. Blow, D. J. McCulley, Z. Li, T. Zhang, J. A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, F. Chen, V. Afzal, J. Bristow, B. Ren, B. L. Black, E. M. Rubin, A. Visel, L. A. Pennacchio, ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* 42, 806-810 (2010).
100. S. Barolo, Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *BioEssays : news and reviews in molecular, cellular and developmental biology* 34, 135-141 (2012).
101. O. Hallikas, K. Palin, N. Sinjushina, R. Rautiainen, J. Partanen, E. Ukkonen, J. Taipale, Genome-wide Prediction of Mammalian Enhancers Based on Analysis of Transcription-Factor Binding Affinity. *Cell* 124, 47-59 (2006).
102. C. R. Clapier, B. R. Cairns, The Biology of Chromatin Remodeling Complexes. *Annual Review of Biochemistry* 78, 273-304 (2009).
103. R. J. Sims, C.-F. Chen, H. Santos-Rosa, T. Kouzarides, S. S. Patel, D. Reinberg, Human but Not Yeast CHD1 Binds Directly and Selectively to Histone H3 Methylated at Lysine 4 via Its Tandem Chromodomains. *Journal of Biological Chemistry* 280, 41789-41792 (2005).
104. J. F. Flanagan, L.-Z. Mi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad, S. Khorasanizadeh, Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 438, 1181-1185 (2005).
105. R. Schneider, A. J. Bannister, F. A. Myers, A. W. Thorne, C. Crane-Robinson, T. Kouzarides, Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* 6, 73-77 (2004).
106. N. D. Heintzman, R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, B. Ren, Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-318 (2007).
107. M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch, R. A. Young, A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell* 130, 77-88 (2007).

108. H. Ferreira, A. Flaus, T. Owen-Hughes, Histone Modifications Influence the Action of Snf2 Family Remodelling Enzymes by Different Mechanisms. *Journal of Molecular Biology* 374, 563-579 (2007).
109. M. P. Creyghton, A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, R. Jaenisch, Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences* 107, 21931-21936 (2010).
110. A. ENCODE consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74 (2012).
111. L. Narlikar, N. J. Sakabe, A. A. Blanski, F. E. Arimura, J. M. Westlund, M. A. Nobrega, I. Ovcharenko, Genome-wide discovery of human heart enhancers. *Genome Res* 20, 381-392 (2010).
112. S. De Val, Key Transcriptional Regulators of Early Vascular Development. *Arteriosclerosis, Thrombosis, and Vascular Biology* 31, 1469-1475 (2011).
113. F. Melet, B. Motro, D. J. Rossi, L. Zhang, A. Bernstein, Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. *Mol Cell Biol* 16, 2708-2718 (1996).
114. F. Liu, M. Walmsley, A. Rodaway, R. Patient, Fli1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol* 18, 1234-1240 (2008).
115. V. Vlaeminck-Guillem, S. Carrere, F. Dewitte, D. Stehelin, X. Desbiens, M. Duterque-Coquillaud, The Ets family member Erg gene is expressed in mesodermal tissues and neural crests at fundamental steps during mouse embryogenesis. *Mechanisms of Development* 91, (2000).
116. G. M. Birdsey, N. H. Dryden, V. Amsellem, F. Gebhardt, K. Sahnan, D. O. Haskard, E. Dejana, J. C. Mason, A. M. Randi, Transcription factor Erg regulates angiogenesis and endothelial apoptosis through VE-cadherin. *Blood* 111, 3498-3506 (2008).
117. P. Vijayaraj, A. Le Bras, N. Mitchell, M. Kondo, S. Juliao, M. Wasserman, D. Beeler, K. Spokes, W. C. Aird, H. S. Baldwin, P. Oettgen, Erg is a crucial regulator of endocardial-mesenchymal transformation during cardiac valve morphogenesis. *Development* 139, 3973-3985 (2012).
118. B. J. Graves, J. M. Petersen, Specificity within the ets family of transcription factors. *Adv Cancer Res* 75, 1-55 (1998).
119. A. Hart, F. Melet, P. Grossfeld, K. Chien, C. Jones, A. Tunnacliffe, R. Favier, A. Bernstein, Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. *Immunity* 13, 167-177 (2000).
120. V. N. Pham, N. D. Lawson, J. W. Mugford, L. Dye, D. Castranova, B. Lo, B. M. Weinstein, Combinatorial function of ETS transcription factors in the developing vasculature. *Dev Biol* 303, 772-783 (2007).
121. P. P. Spyropoulos. D, Lavenburg. Kim, Jackers. P, Papas. Takis, Ogawa. M, and Watson D, Hemorrhage, Impaired Hematopoiesis, and Lethality in Mouse Embryos

Carrying a Targeted Disruption of the Fli1 Transcription Factor. *Mol and Cellular Biology*, 5643–5652 (2000).

122. G. Wei, R. Srinivasan, C. Z. Cantemir-Stone, S. M. Sharma, R. Santhanam, M. Weinstein, N. Muthusamy, A. K. Man, R. G. Oshima, G. Leone, M. C. Ostrowski, Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood* 114, 1123-1130 (2009).
123. D. Lee, C. Park, H. Lee, J. J. Lugus, S. H. Kim, E. Arentson, Y. S. Chung, G. Gomez, M. Kyba, S. Lin, R. Janknecht, D. S. Lim, K. Choi, ER71 Acts Downstream of BMP, Notch, and Wnt Signaling in Blood and Vessel Progenitor Specification. *Cell Stem Cell* 2, 497-507 (2008).
124. A. Ferdous, A. Caprioli, M. Iacovino, C. M. Martin, J. Morris, J. A. Richardson, S. Latif, R. E. Hammer, R. P. Harvey, E. N. Olson, M. Kyba, D. J. Garry, Nkx2-5 transactivates the Ets-related protein 71 gene and specifies an endothelial/endocardial fate in the developing embryo. *Proc Natl Acad Sci U S A* 106, 814-819 (2009).
125. J. J. Lugus, Y. S. Chung, J. C. Mills, S. I. Kim, J. Grass, M. Kyba, J. M. Doherty, E. H. Bresnick, K. Choi, GATA2 functions at multiple steps in hemangioblast development and differentiation. *Development* 134, 393-405 (2007).
126. S. De Val, N. C. Chi, S. M. Meadows, S. Minovitsky, J. P. Anderson, I. S. Harris, M. L. Ehlers, P. Agarwal, A. Visel, S.-M. Xu, L. A. Pennacchio, I. Dubchak, P. A. Krieg, D. Y. R. Stainier, B. L. Black, Combinatorial Regulation of Endothelial Gene Expression by Ets and Forkhead Transcription Factors. *Cell* 135, 1053-1064 (2008).
127. S. Sumanas, G. Gomez, Y. Zhao, C. Park, K. Choi, S. Lin, Interplay among Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. *Blood* 111, 4500-4510 (2008).
128. T. Furuyama, K. Kitayama, Y. Shimoda, M. Ogawa, K. Sone, K. Yoshida-Araki, H. Hisatsune, S. Nishikawa, K. Nakayama, K. Ikeda, N. Motoyama, N. Mori, Abnormal angiogenesis in Foxo1 (Fkhr)-deficient mice. *J Biol Chem* 279, 34741-34749 (2004).
129. T. Hosaka, W. H. Biggs Iii, D. Tieu, A. D. Boyer, N. M. Varki, W. K. Cavenee, K. C. Arden, Disruption of forkhead transcription factor (FOXO) family members in mice reveals their functional diversification. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2975-2980 (2004).
130. C. Daly, V. Wong, E. Burova, Y. Wei, S. Zabski, J. Griffiths, K. M. Lai, H. C. Lin, E. Ioffe, G. D. Yancopoulos, J. S. Rudge, Angiopoietin-1 modulates endothelial cell function and gene expression via the transcription factor FKHR (FOXO1). *Genes Dev* 18, 1060-1071 (2004).
131. J. H. Paik, R. Kollipara, G. Chu, H. Ji, Y. Xiao, Z. Ding, L. Miao, Z. Tothova, J. W. Horner, D. R. Carrasco, S. Jiang, D. G. Gilliland, L. Chin, W. H. Wong, D. H. Castrillon, R. A. DePinho, FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell* 128, 309-323 (2007).
132. J. Choi, L. Dong, J. Ahn, D. Dao, M. Hammerschmidt, J. N. Chen, FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Dev Biol* 304, 735-744 (2007).

133. M. Mahlapuu, M. Ormestad, S. Enerback, P. Carlsson, The forkhead transcription factor Foxf1 is required for differentiation of extra-embryonic and lateral plate mesoderm. *Development* 128, 155-166 (2001).
134. J. Astorga, P. Carlsson, Hedgehog induction of murine vasculogenesis is mediated by Foxf1 and Bmp4. *Development* 134, 3753-3761 (2007).
135. T. Kume, H. Jiang, J. M. Topczewska, B. L. Hogan, The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. *Genes Dev* 15, 2470-2482 (2001).
136. H. Hayashi, T. Kume, Foxc transcription factors directly regulate Dll4 and Hey2 expression by interacting with the VEGF-Notch signaling pathways in endothelial cells. *PLoS One* 3, 0002401 (2008).
137. Q. Lin, J. Lu, H. Yanagisawa, R. Webb, G. E. Lyons, J. A. Richardson, E. N. Olson, Requirement of the MADS-box transcription factor MEF2C for vascular development. *Development* 125, 4565-4574 (1998).
138. W. Bi, C. J. Drake, J. J. Schwarz, The Transcription Factor MEF2C-Null Mouse Exhibits Complex Vascular Malformations and Reduced Cardiac Expression of Angiopoietin 1 and VEGF. *Developmental Biology* 211, 255-267 (1999).
139. D. Maiti, Z. Xu, E. J. Duh, Vascular Endothelial Growth Factor Induces MEF2C and MEF2-Dependent Activity in Endothelial Cells. *Investigative Ophthalmology & Visual Science* 49, 3640-3648 (2008).
140. C. Pagiatakis, J. W. Gordon, S. Ehyai, J. C. McDermott, A novel RhoA/ROCK-CPI-17-MEF2C signaling pathway regulates vascular smooth muscle cell gene expression. *J Biol Chem* 287, 8361-8370 (2012).
141. M. E. Lee, D. H. Temizer, J. A. Clifford, T. Quertermous, Cloning of the GATA-binding protein that regulates endothelin-1 gene expression in endothelial cells. *J Biol Chem* 266, 16188-16192 (1991).
142. B. Gottgens, A. Nastos, S. Kinston, S. Piltz, E. C. Delabesse, M. Stanley, M. J. Sanchez, A. Ciau-Uitz, R. Patient, A. R. Green, Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *Embo J* 21, 3039-3050 (2002).
143. W. Schachterle, A. Rojas, S. M. Xu, B. L. Black, ETS-dependent regulation of a distal Gata4 cardiac enhancer. *Dev Biol* 361, 439-449 (2012).
144. A. W. Orr, B. P. Helmke, B. R. Blackman, M. A. Schwartz, Mechanisms of Mechanotransduction. *Developmental Cell* 10, 11-20 (2006).
145. R. J. Dekker, S. van Soest, R. D. Fontijn, S. Salamanca, P. G. de Groot, E. VanBavel, H. Pannekoek, A. J. G. Horrevoets, Prolonged fluid shear stress induces a distinct set of endothelial cell genes, most specifically lung Krüppel-like factor (KLF2). *Blood* 100, 1689-1698 (2002).
146. A. D. Egorova, K. van der Heiden, R. E. Poelmann, B. P. Hierck, Primary cilia as biomechanical sensors in regulating endothelial function. *Differentiation* 83, S56-S61 (2012).

147. F. A. Pereira, Y. Qiu, G. Zhou, M. J. Tsai, S. Y. Tsai, The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev* 13, 1037-1049 (1999).
148. L.-R. You, F.-J. Lin, C. T. Lee, F. J. DeMayo, M.-J. Tsai, S. Y. Tsai, Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. *Nature* 435, 98-104 (2005).
149. S. Cermenati, S. Moleri, S. Cimbro, P. Corti, L. Del Giacco, R. Amodeo, E. Dejana, P. Koopman, F. Cotelli, M. Beltrame, Sox18 and Sox7 play redundant roles in vascular development. *Blood* 111, 2657-2666 (2008).
150. R. Herpers, E. van de Kamp, H. J. Duckers, S. Schulte-Merker, Redundant Roles for Sox7 and Sox18 in Arteriovenous Specification in Zebrafish. *Circulation Research* 102, 12-15 (2008).
151. H. Pendeville, M. Winandy, I. Manfroid, O. Nivelles, P. Motte, V. Pasque, B. Peers, I. Struman, J. A. Martial, M. L. Voz, Zebrafish Sox7 and Sox18 function together to control arterial-venous identity. *Developmental Biology* 317, 405-416 (2008).
152. M. Corada, F. Orsenigo, M. F. Morini, M. E. Pitulescu, G. Bhat, D. Nyqvist, F. Breviario, V. Conti, A. Briot, M. L. Iruela-Arispe, R. H. Adams, E. Dejana, Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nat Commun* 4, (2013).
153. S.-H. Lee, S. Lee, H. Yang, S. Song, K. Kim, T. L. Saunders, J. K. Yoon, G. Y. Koh, I. Kim, The Notch Pathway Targets Proangiogenic Regulator Sox17 to Restrict Angiogenesis. *Circulation Research*, (2014).
154. S. K. Ramasamy, A. P. Kusumbe, L. Wang, R. H. Adams, Endothelial Notch activity promotes angiogenesis and osteogenesis in bone. *Nature* 507, 376-380 (2014).
155. B. M. Weinstein, D. L. Stemple, W. Driever, M. C. Fishman, Gridlock, a localized heritable vascular patterning defect in the zebrafish. *Nat Med* 1, 1143-1147 (1995).
156. T. P. Zhong, S. Childs, J. P. Leu, M. C. Fishman, Gridlock signalling pathway fashions the first embryonic artery. *Nature* 414, 216-220 (2001).
157. M.-P. Simon, R. Tournaire, J. Pouyssegur, The angiopoietin-2 gene of endothelial cells is up-regulated in hypoxia by a HIF binding site located in its first intron and by the central factors GATA-2 and Ets-1. *Journal of Cellular Physiology* 217, 809-818 (2008).
158. M. Eisbacher, L. M. Khachigian, T. H. Khin, M. L. Holmes, B. H. Chong, Inducible Expression of the Megakarocyte-specific Gene Glycoprotein IX Is Mediated through an Ets Binding Site and Involves Upstream Activation of Extracellular Signal-regulated Kinase. *Cell Growth Differ* 12, 435-445 (2001)
159. R. Lahlil, E. Lecuyer, S. Herblot, T. Hoang, SCL assembles a multifactorial complex that determines glycophorin A expression. *Mol Cell Biol* 24, 1439-1452 (2004).
160. M. H. Prandini, I. Dreher, S. Bouillot, S. Benkerri, T. Moll, P. Huber, The human VE-cadherin promoter is subjected to organ-specific regulation and is activated in tumour angiogenesis. *Oncogene* 24, 2992-3001 (2005).

161. S. Gory, J. Dalmon, M. H. Prandini, T. Kortulewski, Y. de Launoit, P. Huber, Requirement of a GT box (Sp1 site) and two Ets binding sites for vascular endothelial cadherin gene transcription. *J Biol Chem* 273, 6750-6755 (1998)
162. J. D. Kormish, D. Sinner, A. M. Zorn, Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease. *Dev Dyn* 239, 56-68 (2010).
163. X. Ye, Y. Wang, H. Cahill, M. Yu, T. C. Badea, P. M. Smallwood, N. S. Peachey, J. Nathans, Norrin, frizzled-4, and Lrp5 signaling in endothelial cells controls a genetic program for retinal vascularization. *Cell* 139, 285-298 (2009).
164. Y. Sakamoto, K. Hara, M. Kanai-Azuma, T. Matsui, Y. Miura, N. Tsunekawa, M. Kurohmaru, Y. Saijoh, P. Koopman, Y. Kanai, Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos. *Biochemical and Biophysical Research Communications* 360, 539-544 (2007).
165. C. Pellet-Many, P. Frankel, H. Jia, I. Zachary, Neuropilins: structure, function and role in disease. *Biochem J* 411, 211-226 (2008).
166. C. Gaboriaud, N. M. Thielens, L. A. Gregory, V. Rossi, J. C. Fontecilla-Camps, G. J. Arlaud, Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends in Immunology* 25, 368-373 (2004).
167. M. Nakamura, S. F. Masri, A. G. Chassiakos, T. K. Caughey, A method for non-parametric damage detection through the use of neural networks. *Earthquake Engineering & Structural Dynamics* 27, 997-1010 (1998).
168. A. Fantin, Q. Schwarz, K. Davidson, E. M. Normando, L. Denti, C. Ruhrberg, The cytoplasmic domain of neuropilin 1 is dispensable for angiogenesis, but promotes the spatial separation of retinal arteries and veins. *Development* 138, 4185-4191 (2011).
169. A. Lanahan, X. Zhang, A. Fantin, Z. Zhuang, F. Rivera-Molina, K. Speichinger, C. Prahst, J. Zhang, Y. Wang, G. Davis, D. Toomre, C. Ruhrberg, M. Simons, The Neuropilin 1 Cytoplasmic Domain Is Required for VEGF-A-Dependent Arteriogenesis. *Developmental Cell* 25, 156-168 (2013).
170. M. Rossignol, M. L. Gagnon, M. Klagsbrun, Genomic Organization of Human Neuropilin-1 and Neuropilin-2 Genes: Identification and Distribution of Splice Variants and Soluble Isoforms. *Genomics* 70, 211-222 (2000).
171. S. Soker, S. Gollamudi-Payne, H. Fidder, H. Charmahelli, M. Klagsbrun, Inhibition of Vascular Endothelial Growth Factor (VEGF)-induced Endothelial Cell Proliferation by a Peptide Corresponding to the Exon 7-Encoded Domain of VEGF165. *Journal of Biological Chemistry* 272, 31582-31588 (1997).
172. G. Fuh, K. C. Garcia, A. M. de Vos, The interaction of neuropilin-1 with vascular endothelial growth factor and its receptor flt-1. *J Biol Chem* 275, 26690-26695 (2000).
173. T. Kärpänen, C. A. Heckman, S. Keskitalo, M. Jeltsch, H. Ollila, G. Neufeld, L. Tamagnone, K. Alitalo, Functional interaction of VEGF-C and VEGF-D with neuropilin receptors. *The FASEB Journal* 20, 1462-1472 (2006).
174. Q. Pan, Y. Chanthery, W.-C. Liang, S. Stawicki, J. Mak, N. Rathore, R. K. Tong, J. Kowalski, S. F. Yee, G. Pacheco, S. Ross, Z. Cheng, J. Le Couter, G. Plowman, F.

Peale, A. W. Koch, Y. Wu, A. Bagri, M. Tessier-Lavigne, R. J. Watts, Blocking Neuropilin-1 Function Has an Additive Effect with Anti-VEGF to Inhibit Tumor Growth. *Cancer Cell* 11, 53-67 (2007).

175. Y. Jia, P. W. Buehler, R. A. Boykins, R. M. Venable, A. I. Alayash, Structural Basis of Peroxide-mediated Changes in Human Hemoglobin: A NOVEL OXIDATIVE PATHWAY. *Journal of Biological Chemistry* 282, 4894-4907 (2007).
176. J. Waltenberger, L. Claesson-Welsh, A. Siegbahn, M. Shibuya, C. H. Heldin, Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor. *J Biol Chem* 269, 26988-26995 (1994).
177. K. Ballmer-Hofer, A. E. Andersson, L. E. Ratcliffe, P. Berger, Neuropilin-1 promotes VEGFR-2 trafficking through Rab11 vesicles thereby specifying signal output. *Blood* 118, 816-826 (2011).
178. A. Cariboni, K. Davidson, E. Dozio, F. Memi, Q. Schwarz, F. Stossi, J. G. Parnavelas, C. Ruhrberg, VEGF signalling controls GnRH neuron survival via NRP1 independently of KDR and blood vessels. *Development* 138, 3723-3733 (2011).
179. S. Koch, L. A. van Meeteren, E. Morin, C. Testini, S. Westrom, H. Bjorkelund, S. Le Jan, J. Adler, P. Berger, L. Claesson-Welsh, NRP1 presented in trans to the endothelium arrests VEGFR2 endocytosis, preventing angiogenic signaling and tumor initiation. *Dev Cell* 28, 633-646 (2014).
180. A. Fantin, J. M. Vieira, A. Plein, L. Denti, M. Fruttiger, J. W. Pollard, C. Ruhrberg, NRP1 acts cell autonomously in endothelium to promote tip cell function during sprouting angiogenesis. *Blood* 121, 2352-2362 (2013).
181. L. D. Covassin, J. A. Villefranc, M. C. Kacergis, B. M. Weinstein, N. D. Lawson, Distinct genetic interactions between multiple Vegf receptors are required for development of different blood vessel types in zebrafish. *Proc Natl Acad Sci U S A* 103, 6554-6559 (2006).
182. L. Yuan, D. Moyon, L. Pardanaud, C. Bréant, M. J. Karkkainen, K. Alitalo, A. Eichmann, Abnormal lymphatic vessel development in neuropilin 2 mutant mice. *Development* 129, 4797-4806 (2002).
183. Z. Gluzman-Poltorak, T. Cohen, M. Shibuya, G. Neufeld, Vascular endothelial growth factor receptor-1 and neuropilin-2 form complexes. *J Biol Chem* 276, 18688-18694 (2001).
184. B. Favier, A. Alam, P. Barron, J. Bonnin, P. Laboudie, P. Fons, M. Mandron, J. P. Herault, G. Neufeld, P. Savi, J. M. Herbert, F. Bono, Neuropilin-2 interacts with VEGFR-2 and VEGFR-3 and promotes human endothelial cell survival and migration. *Blood* 108, 1243-1250 (2006).
185. Y. Luo, D. Raible, J. A. Raper, Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones. *Cell* 75, 217-227 (1993).
186. B. J. C. Janssen, T. Malinauskas, G. A. Weir, M. Z. Cader, C. Siebold, E. Y. Jones, Neuropilins lock secreted semaphorins onto plexins in a ternary signaling complex. *Nat Struct Mol Biol* 19, 1293-1299 (2012).
187. Z. He, M. Tessier-Lavigne, Neuropilin Is a Receptor for the Axonal Chemorepellent Semaphorin III. *Cell* 90, 739-751 (1997).

188. T. Takahashi, A. Fournier, F. Nakamura, L. H. Wang, Y. Murakami, R. G. Kalb, H. Fujisawa, S. M. Strittmatter, Plexin-neuropilin-1 complexes form functional semaphorin-3A receptors. *Cell* 99, 59-69 (1999).
189. G. Serini, D. Valdembri, S. Zanivan, G. Morterra, C. Burkhardt, F. Caccavari, L. Zammataro, L. Primo, L. Tamagnone, M. Logan, M. Tessier-Lavigne, M. Taniguchi, A. W. Puschel, F. Bussolino, Class 3 semaphorins control vascular morphogenesis by inhibiting integrin function. *Nature* 424, 391-397 (2003).
190. T. Kitsukawa, M. Shimizu, M. Sanbo, T. Hirata, M. Taniguchi, Y. Bekku, T. Yagi, H. Fujisawa, Neuropilin–Semaphorin III/D-Mediated Chemorepulsive Signals Play a Crucial Role in Peripheral Nerve Projection in Mice. *Neuron* 19, 995-1005 (1997).
191. T. Kitsukawa, A. Shimono, A. Kawakami, H. Kondoh, H. Fujisawa, Overexpression of a membrane protein, neuropilin, in chimeric mice causes anomalies in the cardiovascular system, nervous system and limbs. *Development* 121, 4309-4318 (1995).
192. P. Lee, K. Goishi, A. J. Davidson, R. Mannix, L. Zon, M. Klagsbrun, Neuropilin-1 is required for vascular development and is a mediator of VEGF-dependent angiogenesis in zebrafish. *Proc Natl Acad Sci U S A* 99, 10470-10475 (2002).
193. E. A. Jones, L. Yuan, C. Breant, R. J. Watts, A. Eichmann, Separating genetic and hemodynamic defects in neuropilin 1 knockout embryos. *Development* 135, 2479-2488 (2008).
194. E. Geretti, A. Shimizu, M. Klagsbrun, Neuropilin structure governs VEGF and semaphorin binding and regulates angiogenesis. *Angiogenesis* 11, 31-39 (2008).
195. B. Larrivee, C. Freitas, S. Suchting, I. Brunet, A. Eichmann, Guidance of vascular development: lessons from the nervous system. *Circ Res* 104, 428-441 (2009).
196. A. Fantin, B. Herzog, M. Mahmoud, M. Yamaji, A. Plein, L. Denti, C. Ruhrberg, I. Zachary, Neuropilin 1 (NRP1) hypomorphism combined with defective VEGF-A binding reveals novel roles for NRP1 in developmental and pathological angiogenesis. *Development* 141, 556-562 (2014).
197. G. Bou-Gharios, L. A. Garrett, J. Rossert, K. Niederreither, H. Eberspaecher, C. Smith, C. Black, B. Crombrugghe, A potent far-upstream enhancer in the mouse pro alpha 2(I) collagen gene regulates expression of reporter genes in transgenic mice. *J Cell Biol* 134, 1333-1344 (1996)
198. M. E. Pitulescu, I. Schmidt, R. Benedito, R. H. Adams, Inducible gene targeting in the neonatal vasculature and analysis of retinal angiogenesis in mice. *Nat. Protocols* 5, 1518-1534 (2010).
199. K. Kawakami, T. Noda, Transposition of the Tol2 element, an Ac-like element from the Japanese medaka fish Oryzias latipes, in mouse embryonic stem cells. *Genetics* 166, 895-899 (2004).
200. P. C. Hollenhorst, A. A. Shah, C. Hopkins, B. J. Graves, Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev* 21, 1882-1894 (2007).
201. K. R. Rosenbloom, C. A. Sloan, V. S. Malladi, T. R. Dreszer, K. Learned, V. M. Kirkup, M. C. Wong, M. Maddren, R. Fang, S. G. Heitner, B. T. Lee, G. P. Barber, R.

A. Harte, M. Diekhans, J. C. Long, S. P. Wilder, A. S. Zweig, D. Karolchik, R. M. Kuhn, D. Haussler, W. J. Kent, ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Research* 41, D56-D63 (2013).

202. H. Kataoka, M. Hayashi, R. Nakagawa, Y. Tanaka, N. Izumi, S. Nishikawa, M. L. Jakt, H. Tarui, S.-I. Nishikawa, *Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRα+ primitive mesoderm*. *Blood*, 118: 6975-6986 (2011).
203. Joshua D. Wythe, Lan T. H. Dang, W. P. Devine, E. Boudreau, Stanley T. Artap, D. He, W. Schachterle, Didier Y. R. Stainier, P. Oettgen, Brian L. Black, Benoit G. Bruneau, Jason E. Fish, ETS Factors Regulate Vegf-Dependent Arterial Specification. *Developmental Cell* 26, 45-58 (2013).
204. A. D. Person, S. E. Klewer, R. B. Runyan, in *International Review of Cytology*., vol. 243, pp. 287-335.(2005).
205. J. F. Sabl, S. Henikoff, Copy Number and Orientation Determine the Susceptibility of a Gene to Silencing by Nearby Heterochromatin in Drosophila. *Genetics* 142, 447-458 (1996).
206. R. D. Palmiter, R. L. Brinster, Germ-line transformation of mice. *Annu Rev Genet* 20, 465-499 (1986).
207. C. Mosimann, A.-C. Puller, K. L. Lawson, P. Tschopp, A. Amsterdam, L. I. Zon, Site-directed zebrafish transgenesis into single landing sites with the phiC31 integrase system. *Developmental Dynamics* 242, 949-963 (2013).
208. S. K. Bronson, E. G. Plaehn, K. D. Kluckman, J. R. Hagaman, N. Maeda, O. Smithies, Single-copy transgenic mice with chosen-site integration. *Proceedings of the National Academy of Sciences* 93, 9067-9072 (1996).
209. M. C. Tuason, A. Rastikerdar, T. Kuhlmann, C. Goujet-Zalc, B. Zalc, S. Dib, H. Friedman, A. Peterson, Separate Proteolipid Protein/DM20 Enhancers Serve Different Lineages and Stages of Development. *The Journal of Neuroscience* 28, 6895-6903 (2008).
210. A. C. Wilkinson, D. K. Goode, Y.-H. Cheng, D. E. Dickel, S. Foster, T. Sendall, M. R. Tijssen, M.-J. Sanchez, L. A. Pennacchio, A. M. Kirkpatrick, B. Göttgens, Single site-specific integration targeting coupled with embryonic stem cell differentiation provides a high-throughput alternative to in vivo enhancer analyses. *Biology Open* 2, 1229-1238 (2013).
211. G. Friedrich, P. Soriano, Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes & Development* 5, 1513-1523 (1991).
212. J. Hornig, F. Frob, M. R. Vogl, I. Hermans-Borgmeyer, E. R. Tamm, M. Wegner, The transcription factors Sox10 and Myrf define an essential regulatory network module in differentiating oligodendrocytes. *PLoS Genet* 9, e1003907 (2013).
213. H. Gerhardt, C. Ruhrberg, A. Abramsson, H. Fujisawa, D. Shima, C. Betsholtz, Neuropilin-1 is required for endothelial tip cell guidance in the developing central nervous system. *Developmental Dynamics* 231, 503-509 (2004).
214. P. Soriano, Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* 21, 70-71 (1999).

215. J. Park, P. Collier, E. Chen, C. W. Gibson, A beta-galactosidase expression vector for promoter analysis. *DNA and cell biology* 13, 1147-1149 (1994)
216. M. Rossignol, J. Pouysségur, M. Klagsbrun, Characterization of the neuropilin-1 promoter; gene expression is mediated by the transcription factor Sp1. *Journal of Cellular Biochemistry* 88, 744-757 (2003).
217. F. Poulin, M. A. Nobrega, I. Plajzer-Frick, A. Holt, V. Afzal, E. M. Rubin, L. A. Pennacchio, In vivo characterization of a vertebrate ultraconserved enhancer. *Genomics* 85, 774-781 (2005).
218. S. Seo, H. Fujita, A. Nakano, M. Kang, A. Duarte, T. Kume, The forkhead transcription factors, Foxc1 and Foxc2, are required for arterial specification and lymphatic sprouting during vascular development. *Dev Biol* 294, 458-470 (2006).
219. K. N. Papanicolaou, Y. Izumiya, K. Walsh, Forkhead Transcription Factors and Cardiovascular Biology. *Circulation Research* 102, 16-31 (2008).
220. K. D. Kolander, M. L. Holtz, S. M. Cossette, S. A. Duncan, R. P. Misra, Epicardial GATA factors regulate early coronary vascular plexus formation. *Developmental Biology* 386, 204-215 (2014).
221. R. Kothary, R. Clapoff, S. Darling, S. Perry, M. D. Moran, L. A. Rossant, Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Developmental Biology* 105, 4, 707-14 (1989).
222. E. Abdelwahid, M. Eriksson, L. Pelliniemi, E.J. Jokinen, Heat shock proteins,HSP25 and HSP70, and apoptosis in developing endocardial cushion of the mouse heart. *Histochemistry and Cell Biology* 115, 2, 95-104 (2001).