Transcriptional regulation at the angiogenic sprout



Kira M. Chouliaras

St Catherine's College

University of Oxford

Supervisor: Dr Sarah De Val

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2019

Abstract

Transcriptional regulation at the angiogenic sprout

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY, HILARY TERM 2019
KIRA M. CHOULIARAS, ST CATHERINE'S COLLEGE, UNIVERSITY OF OXFORD

Angiogenesis, the formation of new blood vessels from existing ones, depends on very precise spatio-temporal gene expression patterns in different sub-types of endothelial cells. These patterns are controlled by pro-angiogenic stimuli but the transcriptional mechanism by which these stimuli induce specific gene expression in endothelial cells is not well understood. In this thesis I identify and characterize a novel sprouting-specific enhancer upstream of the angiogenesis-specific H2.0-like homeobox protein (*HLX/hlx*) gene. This conserved enhancer, termed HLX-3/hlx-3, is able to direct gene expression specifically and selectively to angiogenic endothelial cells in both transgenic zebrafish and mice during physiological and pathological angiogenesis. Analysis of the transcription factor binding sites within the HLX-3/hlx-3 enhancer reveals a requirement for myocyte enhancer factor-2 (MEF2) family members for enhancer activity. Endothelial-specific *Mef2A/C* knockout results in reduced sprouting in the mouse retina, confirming an important role for MEF2 factors in the regulation of both *Hlx* and sprouting angiogenesis more generally. Further, I use the hlx-3:GFP transgenic zebrafish line to generate a genome-wide map of accessible chromatin regions from isolated sprouting cells, and compare them to arterial and venous endothelial cells. My results demonstrate that accessible chromatin is not sufficient to reliably predict enhancer activity patterns, and that within the endothelial-lineage, regulatory elements are constitutively open irrespective of the sub-endothelial cell type. Interestingly, despite these observations, binding motifs for MEF2C and angiogenic-associated transcription factors are enriched in sprouting endothelial cells. These results provide novel insights into the transcriptional regulatory landscape of endothelial cells and confirm the importance of MEF2 factors during angiogenesis.

Declaration

This thesis is submitted to the Nuffield Department of Clinical Medicine, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. This thesis is entirely my own work and describes my own research, except where otherwise stated. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Any data that are not my own work are acknowledged and referenced appropriately. This work has not been submitted for a degree in this or any other University.

Chapter 3

- Transgenic mice were generated by Dr Ke Liu (Institute of Ageing and Chronic Disease, Liverpool).
- Embryos were sectioned by Indrika Ratnayaka (Ludwig Institute, Oxford).

Chapter 4

- Immunostaining on MEF2 factors was carried out by Dr Martin Fritzsche (Figure 41).
- qRT-PCR experiments were performed jointly with Dr Leonid Nikitenko (Figure 42).
- Chromatin immunoprecipitation was performed in collaboration with Dr Natalia Sacilotto (Figure 45 and 48).
- Work on the Dll4in3 enhancer was performed by Dr Natalia Sacilotto (Figure 47).
- Identification of other angiogenic enhancers was carried out in collaboration with Dr Marsha Wallace, Dr Natalia Sacilotto and Dr Svanhild Nornes (Figure 49).
- *Mef2A/C*^{iECKO} mice were provided by Dr Yao Wei Lu (Center for Cardiovascular Sciences, Albany) and experiments were performed jointly with Dr Natalia Sacilotto (Figure 50 and 51).

Chapter 5

- FACS was performed by the Flow Cytometry Facility (WIMM, Oxford).
- RNA and ATAC sequencing was carried out by the High-Throughput Genomics Group (WTCHG, Oxford).
- Pan-endothelial and non-endothelial sequencing data (Figure 54) were provided by Dr Rui Monteiro and Dr Tomasz Dobrzycki (WIMM, Oxford).

Kira Chouliaras

Word count: 44,713

Acknowledgements

Sarah, you have been the best supervisor I could have hoped for. Your infectious enthusiasm and knowledge are an inspiration. I am incredibly privileged to have worked for and learned from you. Your crisp and clear thinking saved me from going down the rabbit hole. Thank you for infusing me with the energy needed to accomplish my DPhil. And thank you especially for your understanding and support through many challenging times.

I would like to thank everyone from the De Val lab for creating such productive and positive atmosphere. I have been very lucky to have shared the last years with you Alice, Svanhild and Sophie. Talking with you over lunch brought happiness, sparked curiosity and solved many experimental problems. Natalia, you have been an outstanding mentor. Your focused and thorough scientific approach set a high benchmark and will always be my inner compass. Thank you so much, dear De Val lab, for sharing your knowledge with me and for being role models. Thank you, Leonid, for your critical input and interesting discussions. Thank you, Philip, Zhe, Lucija, Martin and Martin, for creating an enriching environment and for sharing the joy of science. Emma and Fede – sitting next to you might have been the best that happened to me in Oxford. Thank you for so much laughter throughout our DPhil journey. Thank you, Colin, for agreeing to be my second supervisor and for providing stimulating conversations that made me think outside my horizon.

Thank you, Rui, for hosting me in your lab and for the guidance to succeed with the sequencing project. Thank you, Tomek, for letting me be part of the thoroughly enjoyable ATAC adventure. I am grateful to Kevin and everyone at the FACS facility for the most entertaining sorting experience. Thanks to Nicki and Emmanouela at the CBRG for answering my naïve questions when I was analyzing my sequencing data. Thanks to Alison, Andy, Greg, Kelly and Adam for looking after my zebrafish and for sharing the passion for this remarkable species. Thanks to Hayley, Kelly and Tom for taking care of our much-loved mice. Thank you, Indi, for your help sectioning. Thank you, Mark and Richard, for helping me create the most stunning images and for bringing beauty to my research.

I would like to thank my family for their unconditional support. Throughout my life, my parents encouraged me to work hard and achieve my dreams (through self-discipline and responsibility). Danke, allerliebste Mama, für deine unerschöpfliche Energie und Liebe. I could not have finished my thesis without you. Weltbester Bruder, mit dir zu sein und von dir zu lernen erfüllt mein Herz. Ευχαριστώ, Αριστοτέλη και Κατερίνα, για όλη την αγάπη και βοήθεια σας. Thank you so much, my dearest love, for sharing the excitement of science and for your unwavering devotion. Rina, you are the essence of my life. Deine Strahlkraft erleuchtet mein Leben. Thank you, beloved family, for your invaluable advice and for making my life richer.

THIS THESIS IS DEDICATED TO MY PARENTS, ESPECIALLY MY FATHER, WHO WOULD HAVE BEEN SO PROUD TO WITNESS THE COMPLETION OF MY DPHIL.

Abbreviations

3D	Three-dimensional
α^{32}P-dCTP	Deoxycytidine triphosphate, labelled with 32P
αSMA	Smooth Muscle Actin
A	Artery
ACTB	β-Actin
AF	Angiogenic front
AP1	Activator protein 1
AT	Aortic trunk
ATAC-seq	Assay for Transposase-Accessible Chromatin sequencing
ATP	Adenosine triphosphate
BAM	Binary alignment map
BED	Browser extensible data
BME	β-mercaptoethanol
BMK	big MAPK
BMP	Bone morphogenetic proteins
bp	Base pairs
BRD4	Bromodomain-containing protein 4
BSA	Bovine serum albumin
BuA	Bulbus arteriosus
°C	degree Celsius
CA	Carotid artery
CBP	CREB binding protein
CD	Cluster of Differentiation
CDH5	VE-Cadherin
cDNA	Complementary DNA
CEBP	CCAAT/enhancer-binding protein
CF	Carrier Free
ChIP-seq	Chromatin immunoprecipitation sequencing
CN	Cervical nerve
CNS	Central nervous system
COUP-TFII	Chicken ovalbumin upstream-transcription factor II
CpG	Cytosine guanine dinucleotides
CREB1	cAMP response element-binding protein
CtA	Central arteries
CTCF	CCCTC-binding factor
Ctrl	Control
CV	Cardinal vein
DA	Dorsal aorta
dab2	DAB adaptor protein 2
DAPI	4’,6-diamidino-2-phenylindole

dpf	Days post fertilization
DIG	Digoxigenin
DLAV	Dorsal longitudinal anastomotic vessel
DLL4	Delta like ligand 4
Dll4i3	Enhancer located in intron 3 of Dll4
DMEM	Dulbecco's modified Eagle's medium
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase HS	DNase I hypersensitivity
dNTP	Deoxynucleotide
DSL	Delta, Serrate, Lag-2
DTT	Dithiothreitol
E	Embryonic day
e.g.	*exempli gratia*
EC	Endothelial cell
ECL	Enhanced chemiluminescence
EDTA	Ethylenediamine tetraacetic acid
ef1α	elongation factor 1α
EFNB2	Ephrin B2
EGF	Epidermal growth factor
Egfr	Egf receptor
EGTA	Ethylene glycol tetraacetic acid
ELF1	E74-like factor 1
EMCN	Endomucin
EMSA	Electrophoretic mobility shift assay
ENCODE	Encyclopaedia of DNA Elements
ephb2a	eph receptor B2a
EPHB4	Eph receptor B4
EphB4-2	Enhancer located 2 kb upstream of EphB4
ERG	ETS-Related Gene
ERK	Extracellular-signal-regulated kinase
eRNA	enhancer RNA
ERT2	Estrogen receptor ligand binding domain
et al.	*Et alia*
ETS	E26 transformation-specific
ETS1DBD	ETS1 DNA binding domain
etsrp	ets1 related protein
ETV2	ETS variant 2
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FDR	False discovery rate
FGF	Fibroblast growth factor

Fgfr	Fgf receptor
FLK1	Fetal liver kinase 1
FLT4	Fms related tyrosine kinase 4
FOX	Forkhead
g	Gravity
GAPDH	Glyceralde-hyde-3-phosphate dehydrogenase
GFP	Green fluorescent protein
GLM LRT	Generalized Linear Mixed Likelihood Ratio Test
GTF	General transcription factor
H3K122ac	Acetylation of lysine 122 of histone H3
H3K27ac	Acetylation of lysine 27 of histone H3
H3K27me3	Trimethylation of lysine 27 of histone H3
H3K4me1	Monomethylation of lysine 4 of histone H3
H3K4me3	Trimethylation of lysine 4 of histone H3
H3K9ac	Acetylation of lysine 9 of histone H3
H3K9me3	Trimethylation of lysine 9 of histone 3
HAAEC	Human abdominal aortic endothelial cells
HAT	Histone acetyltransferase
HBSS	Hanks' Balanced Salt Solution
HDAC	Histone deacetylases
HEK293T	Human embryonic kidney 293 T
HEPES	4-2-Hydroxyethyl-1-piperazine-ethanesulfonic acid
HIF	Hypoxia-inducible factor
HLX	H2.0-like homeobox
HLX-3	Enhancer located 3 kb upstream of HLX
HMGA	High mobility group protein AT-hook
HOMER	Hypergeometric Optimization of Motif EnRichment
hpf	Hours post fertilization
HPV	Hepatic portal vein
HRP	Horseradish-peroxidase
hsp68	Heat shock protein 68
HUVEC	Human umbilical vein endothelial cells
IB4	Isolectin B4
IC	Intercostal vessels
IgG	Immunoglobulin G
IGV	Integrative Genomics Viewer
IL1	Interleukin 1
IOC	Inner optic circle
IRX3	Iroquois-class homeodomain protein
ISV	Intersegmental vessel
JAG1	Jagged1
JL	Jugular lymph sac
JV	Jugular vein

kb	Kilobase pairs
KDR	Kinase insert domain receptor
kdrl	kdr-like
KLF2	Krüppel-like factor 2
KO	Knockout
L	Liver
LacZ	β-galactosidase
LMO2	Lim domain only 2
logCPM	log2 counts-per-million
LV	Lateral ventricle surrounded by the ependymal layer
µM	Micromolar
MACS2	Model-based Analysis of ChIP-Seq
Mad	Mothers against decapentaplegic
MADS	MCM1, Agamous, Deficiens, Serum response factor
MAM	Mastermind
MAPK	Mitogen-activated protein kinase
MDS	Multidimensional scaling
MEF2	Myocyte enhancer factor-2
Mef2A/C^{iECKO}	Cdh5(PAC)Cre-ER^{T2};Mef2A$^{fl/fl}$;Mef2C$^{fl/fl}$
min	Minutes
MLC	Myosin Light Chain
mM	Millimolar
MO	Morpholino
MOI	Multiplicity of infection
MS222	Tricaine mesylate
MSV	Mesencephalic vein
MUT	Mutant
N	Notochord
NC	Nuclear layer of neopallial cortex
NFAT	Nuclear factor of activated T-cells
NICD	Intracellular domain of NOTCH
NLR	Neural layer of retina
nM	Nanomolar
non-EC	Non-endothelial (kdrl:HRAS-mCherry^{-})
NP	Perineural vascular plexus
Nr2f2	Nuclear receptor subfamily 2 group F member 2
NRP	Neuropilin
NT	Neural tube
O	Oesophagus
O/N	Overnight
OCT	Optimal cutting temperature
OCT4	Octamer-binding transcription factor 4
P	Post-natal day

P-TEFb	Positive transcription elongation factor-b
PA/PV	Pectoral artery and vein
pan-EC	Pan-endothelial (kdrl:HRAS-mCherry^{+})
PAV	Parachordal vessels
PBS	Phosphate-buffered saline
PCeV	Posterior cerebral veins
PCR	Polymerase chain reaction
PDGFB	Platelet-derived growth factor B
PFA	Paraformaldehyde
PHBC	Primordial hindbrain channel
PHD	Hydroxylase domain proteins
PHS	Primary head sinus
PIGF	Placental growth factor
PLCγ	Phospholipase Cγ
PRC2	Polycomb repressive complex 2
PT	Pulmonary trunk
PTU	N-Phenylthiourea
PVDF	Polyvinylidene difluoride
PVP	Polyvinylpyrrolidone
qRT-PCR	quantitative Reverse Transcription PCR
RasGRP3	Ras guanine-releasing protein 3
RBPJ	Recombining binding protein suppressor of hairless
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA Pol II	RNA Polymerase II
RNA-seq	RNA sequencing
rpm	Revolutions per minute
RT	Room temperature
RUNX1	Runt-related transcription factor 1
SAM	Sequence alignment map
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SH3TC1	SH3 domain and tetratricopeptide repeat-containing protein 1
SHH	Sonic hedgehog
shRNAmir	microRNA-adapted short hairpin RNA
SINEs	Short interspersed nuclear elements
SIV	Subintestinal vein
SMAD	Sma/Mad homolog
SOX	SRY-related HMG box
SRF	Serum response factor
SSC	Saline-sodium citrate
STAR	Spliced Transcripts Alignment to a Reference
TA	Thoracic artery

TAD	Topologically associating domain
TAL1	T-cell acute lymphoblastic leukemia
TBE	Tris/Borate/EDTA
tbx20	T-box 20
TCF	T-cell factor
TF	Transcription factor
TMM	Trimmed mean of M-values
tnnt2	troponin t2
TSS	Transcriptional start site
TTS	Transcription termination site
ubiC	ubiquitin C
UCSC	University of California, Santa Cruz
UNC5B	Uncoordinated-5 homologue B
V	Vein
VE	Vascular endothelial
VEGF	Vascular endothelial growth factor
VEGFR2/3	VEGF receptor 2 and 3
VN	Vagal nerve
vs.	*versus*
WIMM	Weatherall Institute of Molecular Medicine
WT	Wild-type
WTCHG	Wellcome Trust Centre for Human Genetics
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
ZBTB33	Zinc finger and BTB domain containing 33
ZFIN	Zebrafish Information Network

Table of contents

Abstract I
Declaration II
Acknowledgments III
Abbreviations IV
Table of contents X

1 Introduction 1

1.1 Transcriptional regulation ... *2*
1.1.1 Chromatin ... 2
1.1.2 Transcription ... 4
1.1.3 Three-dimensional nuclear architecture ... 12

1.2 Blood vessel development ... *16*
1.2.1 Vasculogenesis ... 16
1.2.2 Angiogenesis ... 17
1.2.3 Arteriovenous differentiation ... 20
1.2.4 Vascular heterogeneity ... 21

1.3 Molecular mechanism of angiogenesis ... *23*
1.3.1 Major signalling pathways involved in angiogenic sprouting ... 23
1.3.2 Transcription factors involved in sprouting angiogenesis ... 27

1.4 Hypothesis and aim of the study ... *32*

2 Materials and methods 34

2.1 Transgenic mice ... *35*
2.1.1 Generation and breeding of transgenic mouse lines ... 35
2.1.2 *LacZ* Genotyping ... 36
2.1.3 Neo-angiogenesis assays ... 36

2.2 Histology ... *40*
2.2.1 Fixing and staining of embryos ... 40
2.2.2 Paraffin embedding and sectioning ... 41
2.2.3 Immunostaining followed by X-Gal staining ... 41
2.2.4 *In situ* hybridization ... 42
2.2.5 Preparation of whole-retina flatmounts ... 44
2.2.6 Preparation of hindbrain flatmounts ... 46
2.2.7 Preparation of tibia ... 46

2.3 Transgenic zebrafish ... *47*
2.3.1 Husbandry ... 47
2.3.2 Tol2 transposon transgenesis in zebrafish ... 48
2.3.3 Morpholinos and small molecule inhibitors ... 48
2.3.4 N-Phenylthiourea treatment and embedding ... 49
2.3.5 Image acquisition and analysis ... 49
2.3.6 Whole-mount *in situ* hybridization ... 49

2.4 Cloning of enhancer constructs ... 51
2.4.1 Gateway cloning ... 51
2.4.2 Enhancer sequences ... 52
2.4.3 Generating vectors for three-colour zebrafish for simultaneous analysis ... 58

2.5 In vitro experiments ... 60
2.5.1 Cell culture ... 60
2.5.2 Chromatin Immunoprecipitation ... 63
2.5.3 Western Blotting ... 65
2.5.4 Quantitative RT-PCR ... 66

2.6 Electrophoretic mobility shift assay ... 69
2.6.1 Plasmids for generation of recombinant proteins ... 69
2.6.2 Probe labelling and gel shifts ... 69

2.7 RNA and ATAC sequencing on sorted cells from transgenic zebrafish ... 72
2.7.1 Fluorescence-activated cell sorting ... 72
2.7.2 Library preparation ... 74
2.7.3 Bioinformatic analysis ... 77

3 Identification of an angiogenic-specific enhancer 84

3.1 Rationale: Using enhancers to delineate regulatory pathways ... 85
3.1.1 Using enhancers as a tool to understand gene regulation in sprouting angiogenesis ... 85
3.1.2 Aim: Finding an enhancer specifically active during sprouting angiogenesis ... 87

3.2 Finding an angiogenic sprouting-specific enhancer ... 88
3.2.1 *In silico* analysis of putative angiogenic-specific enhancers ... 88
3.2.2 Putative angiogenic-specific enhancers in transient transgenic zebrafish ... 90
3.2.3 Putative HLX and JAG1 enhancers in transient transgenic mice ... 93
3.2.4 Testing the zebrafish orthologue of HLX-3 in transgenic mice ... 94

3.3 Analysis of hlx-3:GFP activity during development in zebrafish ... 95
3.3.1 Optimization of the hlx-3 construct in transient transgenic zebrafish ... 95
3.3.2 hlx-3:GFP is specifically expressed in sprouting endothelial cells ... 96
3.3.3 Comparison of hlx-3 with endogenous *hlx1* expression in developing zebrafish ... 100

3.4 Analysis of a stable HLX-3:LacZ line in mouse ... 101
3.4.1 HLX-3:*LacZ* is active in sprouting endothelial cells during embryonic development ... 101
3.4.2 HLX-3:*LacZ* is active in selective endothelial cells in the post-natal mouse ... 112
3.4.3 Endogenous *Hlx* in the mouse ... 118

3.5 Discussion ... 120
3.5.1 Role of the HLX and JAG1 putative enhancers ... 120
3.5.2 The HLX-3/hlx-3 enhancers drive reporter gene expression in angiogenic sprouting endothelial cells ... 121
3.5.3 Conclusion ... 125

4 Regulation of the HLX enhancer during sprouting angiogenesis 126

4.1 Rationale ... 127
4.1.1 Aim: To identify the upstream regulators controlling HLX-3 activity in angiogenic endothelial cells ... 127

4.2 HLX-3 enhancer activity is downstream of VEGF signalling ... 128
4.2.1 HLX-3:*LacZ* transgene activity during neo-angiogenic growth in adult mice ... 128
4.2.2 Effect of inhibition of known angiogenic signalling pathways on hlx-3:GFP activity pattern in transgenic zebrafish ... 130
4.2.3 Flow does not regulate *hlx1* and its enhancer hlx-3:GFP in zebrafish ... 132

4.3 Identification of the transcription factors required for HLX-3/hlx-3 activity during angiogenesis ... 135
4.3.1 *In silico* analysis of the functionally conserved sequence ... 135
4.3.2 ETS transcription factors can directly bind the hlx-3/HLX-3 enhancer ... 137
4.3.3 ETS binding sites are required for hlx-3 driven reporter gene expression ... 143
4.3.4 Mutational analysis of the hlx-3 enhancer reveals a potential role for MEF2 transcription factors ... 144
4.3.5 MEF2 transcription factors bind the MEF2 motifs within the hlx-3/HLX-3 enhancer ... 148
4.3.6 MEF2 binding sites are required for HLX-3:*LacZ* endothelial expression in mice ... 154

4.4 Mef2 factors bind and regulate other angiogenic genes ... 156
4.4.1 MEF2A and MEF2C directly bind the DLL4in3 angiogenic enhancer ... 156
4.4.2 MEF2C binding is enriched around many genes which are upregulated during angiogenesis ... 158

4.5 MEF2 factors regulate sprouting angiogenesis during murine development ... 160
4.5.1 *Hlx* and *Dll4* expression is reduced in *Mef2A/C^{iECKO}* mice ... 160
4.5.2 Defect in angiogenic sprouting in post-natal *Mef2A/C^{iECKO}* retina ... 163

4.6 Discussion ... 165
4.6.1 HLX-3 is downstream of VEGF signalling ... 165
4.6.2 MEF2 factors are required for sprouting angiogenesis and regulate enhancer HLX-3 ... 166
4.6.3 ETS factors direct gene expression in all endothelial cells ... 167
4.6.4 Increase of H3K27 acetylation and p300 binding at angiogenic-specific enhancers ... 168
4.6.5 VEGF signalling and MEF2 factors in sprouting angiogenesis ... 169
4.6.6 Outlook ... 170
4.6.7 Conclusion ... 171

5 ATAC-seq and RNA-seq on arterial, venous and sprouting endothelial cells in zebrafish 172

5.1 Rationale: Identification of enhancers in zebrafish ... 173
5.1.1 Do different endothelial cell types have differentially accessible chromatin? ... 173

5.2 Sorting sub-endothelial cell types ... 175
5.2.1 Zebrafish lines used for RNA-seq and ATAC-seq at 27 hpf and 43 hpf ... 175
5.2.2 Sorted arterial and venous cells largely cluster as expected ... 178

5.3 Search for endothelial-specific enhancers ... 180
5.3.1 Accessible chromatin regions identified by ATAC-seq detects known endothelial enhancers ... 180
5.3.2 ATAC-seq cannot consistently predict active angiogenic-specific enhancers ... 184

5.4 Integration of ATAC-seq and RNA-seq ... 188
5.4.1 Differential gene expression in veins and sprouts at 27 hpf ... 188
5.4.2 ATAC-seq peaks correlate with differentially expressed genes in veins and sprouts ... 191
5.4.3 Motif analysis identifies known and novel transcription factors ... 195

5.5 Discussion ... 200
5.5.1 Accessible chromatin does not necessarily predict an active enhancer ... 200
5.5.2 Chromatin surrounding differentially transcribed genes is more accessible ... 201
5.5.3 Use of endogenous promoter instead of pE1b to test putative enhancers ... 202
5.5.4 Lack of known venous and sprouting genes in RNA-seq ... 203
5.5.5 Conclusion ... 205

6 Conclusion 207

6.1 Overall conclusion ... 208
6.2 Future directions ... 210

7 Bibliography 212

1 Introduction

1.1 Transcriptional regulation

Every organism contains hundreds of different cell types[1]. Each cell contains the same genome and yet morphologically and functionally different cells are formed during embryonic development. During differentiation, each cell expresses a unique set of genes which determines cell fate and identity. Cell-type-specific gene transcription is largely controlled by distal-acting regulatory deoxyribonucleic acid (DNA) regions called enhancers. Enhancers carry clusters of DNA binding motifs for transcription factors (TFs) which bind the DNA at these motifs, recruit co-factors and activate transcription. As many as 1,639 transcription factors have been identified in the human genome, with several hundreds of these being active in any cell type at any given time and collaboratively orchestrating transcription[2]. When cells differentiate, thousands of enhancers get activated and thereby change the global epigenetic landscape and determine spatio-temporally precise gene expression[3].

In this thesis, I will first discuss chromatin architecture and transcription, focusing on how enhancers achieve spatio-temporal gene expression. Then I will address blood vessel development and, more specifically, the cellular pathways that mediate sprouting angiogenesis. Last, I will describe the transcription factors which may be responsible for regulating sprouting-specific transcription.

1.1.1 Chromatin

Chromatin consists of DNA wrapped around nuclear proteins, and compacts eukaryotic genomic DNA in the nucleus. Heterochromatin refers to the compacted version of

chromatin, in which dense folding of nucleosomes supresses gene activation. Euchromatin describes the accessible, open form of chromatin which allows the interaction between DNA and transcriptional activators and other regulatory proteins[4]. Chromatin architecture determines cell identity by altering gene expression. At the smallest scale, nuclear organization can be defined by nucleosomes. At the larger scale, DNA looping determines enhancer-promoter interactions and domain boundaries.

1.1.1.1 Nucleosomes

Each nucleosome comprises approximately 147 base pairs (bp) of DNA wrapped around an octamer of histones in a left-handed superhelix[5]. An octamer constitutes of a pair of each histone H2A, H2B, H3 and H4. Linker histone H1 stabilizes nucleosomes and linker DNA[6]. Nucleosome units are folded into higher order structures until compaction is more than 10,000 fold and hence roughly two metres of DNA fit in a nucleus which is only around six micrometers in size[4]. Histone modifications are responsible for DNA packaging and determine higher order chromatin structure. While the C-terminal forms interactions between DNA and histones, the basic amino acid N-terminal tail of histones is subjected to post translational modifications[7].

1.1.1.2 Histone modifications

Post-translational histone modifications such as acetylation, methylation, phosphorylation, ubiquitylation, sumoylation and ADP ribosylation play a pivotal role in regulating the accessibility of DNA and therefore gene expression. Histone modifications, especially lysine acetylation or methylation and serine or threonine phosphorylation, integrate incoming signals, alter the chromatin landscape and increase binding affinity of

chromatin remodelling factors[8]. Histone binding proteins are recruited to their specific histone binding site and establish a repressive or active environment[8]. In addition to serving as a signalling platform, histone phosphorylation or acetylation can change electrostatic interactions. While phosphorylation causes a negative charge, acetylation neutralizes the positive charge of lysine. A neutralized histone tail is less tightly associated with the underlying DNA, making it more accessible for transcription factors[9].

One of the most frequently occurring histone modifications is acetylation of lysine residues. Histone acetylation, mediated *via* histone acetyltransferase (HATs), makes chromatin more accessible and has a positive effect on gene expression[10]. Conversely, histone deacetylases (HDACs) erase this acetylation mark, creating a more repressive condensed chromatin environment. During transcriptional activation, sequence-specific transcription factors recruit a HAT, such as CREB binding protein (CBP) and its paralogue p300, which dynamically acetylate histone H3 lysine residues at open chromatin. On the other hand, methyltransferases, such as polycomb group protein polycomb repressive complex 2 (PRC2), methylate histone H3 lysine residues and repress transcription by blocking acetylation.

1.1.2 Transcription

Cells respond to the extra-cellular signals from their environment by activating a specific set of transcription factors which subsequently bind regulatory elements, such as promoters and enhancers. Precise spatio-temporal gene expression is not just essential for any cell at any given time, but especially during development, where gene expression changes determine cell fate. Gene regulation is a very dynamic but tightly controlled

process. Promoters primarily bind ubiquitous factors required for transcriptional initiation and tend to remain static. On the other hand, enhancers integrate cell-type-specific signals and mediate on/off patterns of spatio-temporal gene expression[11]. Promoters rely on enhancers to increase the burst fraction (frequency of transcription) above basal levels, with a minor effect on the burst size (number of ribonucleic acid (RNA) molecules)[12].

1.1.2.1 Transcriptional initiation

This section will describe the traditional view of transcription characteristic for TATA-containing promoters. In response to a stimulus, transcription factors bind an enhancer leading to an increase in nucleosome displacement and more accessible DNA. A pre-requisite for transcription is the assembly of the transcription initiation complex. One of the six general transcription factors (GTFs) TFIID recognizes and binds a core-promoter initiator element located in the promoter, such as the TATA-box. Subsequently, the remaining GTFs bind the promoter upstream of the transcriptional start site (TSS) to form the pre-initiation complex. RNA Polymerase II (RNA Pol II) gets recruited at the 5' end of the gene[13]. Enhancers serve as a binding platform for transcription factors and tether chromatin remodelling complexes. Transcription factors can recruit co-regulators (including HATs, HDACs and even RNA Pol II), mediate protein-protein interactions and can have enzymatic activity. The Mediator complex bridges enhancer-bound transcription factors with the transcriptional machinery located at the promoter. The Mediator complex and HAT p300 then stimulate TFIIH-mediated phosphorylation of RNA Pol II and thereby its escape from the TSS[13]. Remarkably, targeted recruitment of p300 to enhancers has been shown the be sufficient for gene induction[14]. RNA Pol II is stalled 20 - 60 bp downstream of the TSS. RNA Pol II pausing occurs at virtually all promoters and is

considered to be the rate-limiting step[13]. At many enhancers, Bromodomain-containing protein 4 (BRD4) is recruited to acetylated histones[15]. Preceding promoter-proximal pause release, BRD4 and other transcription factors then recruit Positive transcription elongation factor-b (P-TEFb)[13]. At the onset of transcription, P-TEFb phosphorylates RNA Pol II, which triggers RNA Pol II pause release. RNA Pol II transitions from a stalled to an actively elongating state[1]. In an adenosine triphosphate (ATP)-consuming process, DNA gets unwound and RNA is synthesized[1].

1.1.2.2 Promoters

Promoters reside upstream in close proximity to the transcriptional start site (TSS) and can in principle autonomously initiate transcription. Transcriptional initiation varies depending the type of core-promoter and in mammals only a subset (around 14%) of promoters contain a TATA-box[16]. TATA-containing promoters are frequently found at key developmental genes[17]. They have been correlated to large transcriptional burst size and are thus used frequently in minimal promoters for transgenic reporter constructs[17]. In the preceding section I focussed on TATA box-guided transcription, because the heat shock protein 68 (hsp68) and pE1b minimal promoter which I will use throughout the experiments carried out in this thesis also contain the classical TATAAA sequence[18,19].

Depending on the gene, different core promoter motifs and therefore several models of transcription have been described. Promoters of constitutively, broadly expressed housekeeping genes are in an active configuration. In contrast, promoters of cell-type-specific developmental genes can be poised, waiting for the cell-type-specific stimulus-induced RNA Pol II pause release with subsequent strong and rapid gene transcription[20].

It has been suggested that promoters with high levels of paused RNA Pol II are enriched for, and dependent on, CBP/p300[21]. CBP/p300 may aid in RNA Pol II recruitment and additionally acetylate the most proximal nucleosome which reduces this transcriptional barrier, leading to pause release.

The majority of promoters are rich in cytosine and guanine dinucleotides, known as CpG islands. While methylation of cytosine in these CpG islands is linked to transcriptional silencing, hydroxymethylation and de-methylation is linked to active transcription[22]. Histone modifications and DNA methylation can occur independently, however HDACs can get recruited by methyl-CpG binding domain proteins, which additionally leads to chromatin condensation and repression[23]. Histone modifications reflect the activity of regulatory elements and can be used to distinguish promoters from other regulatory elements. Classically, inactive promoters are marked by repressive trimethylation of lysine 9 and 27 of histone 3 (H3K9me3 and H3K27me3)[24]. On the other hand, active promoters are flanked by acetylation of lysine 9 and 27 of H3 (H3K27ac and H3K9ac) and trimethylation of H3K4 (H3K4me3)[9].

1.1.2.3 Enhancers

An estimated 20 - 30% of the human genome (3.2 billion base pairs) are regulatory elements which regulate the 1.2% of protein-coding DNA[25,24]. Even though numbers vary depending on the method used, roughly 1 million regulatory elements have been predicted[26,24]. Regulatory elements comprise of promoters, enhancers, insulators and silencers[27]. Out of these regulatory elements, around 300,000 are estimated to be enhancers[28]. The estimated number of enhancers is based on DNA accessibility, chromatin

structure and transcription factor binding. Here, DNA accessibility was measured by DNase I hypersensitivity, chromatin structure was measured by chromosome conformation capture and transcription factor binding was measured by chromatin immunoprecipitation sequencing (ChIP-seq). But verification for these putative enhancer sequences is lacking. Less than 50% of all enhancers that were tested *in vivo* increased transcription of reporter genes[29,24]. Up to date, only approximately 1,600 functional developmental enhancers have been verified in transgenic mice, although this list is incomplete[30].

Enhancers are 200 – 1,000 bp sequences located anywhere in the genome, but primarily in intergenic, intronic and untranslated regions[27]. Most enhancers (around 99%) are *cis*-regulatory regions[31] and only very few inter-chromosomal *trans*-enhancers have been described[32]. Enhancers can be located up- or downstream of the target gene. They act independent of orientation and distance from the TSS. On average, enhancers are located 20 - 50 kilobase pairs (kb) upstream of the TSS[11]. It has been predicted that 10% of enhancers interact with the most proximal promoter[28], while also 10% of enhancers interact with a promoter more than 1,000 kb away[31,33]. Clusters of densely located, highly active enhancers that collectively control gene expression, such as the Hox or ß-globin loci, have been termed super-enhancers[1].

During differentiation, the DNA of the most active enhancers is de-methylated or hydroxymethylated where it correlates with an increase in chromatin accessibility[34]. However, the DNA methylation status at enhancers is variable and highly context-specific depending on the genomic location and cell type. Enhancer DNA hydroxymethylation has

also been linked to an increase in bidirectional enhancer RNA (eRNA) transcription[35]. These short and unstable non-coding RNAs are the result of fortuitous transcription. Even though eRNA transcription can be used as a measure for enhancer activity, it is not consistently linked to increased enhancer activity[36]. Enhancer RNA has been hypothesized to facilitate transcription factor binding to the enhancer[37] or to enhance DNA accessibility[38,39]. Many other functions have been suggested, but the role(s) of eRNAs remain largely elusive.

Enhancers can broadly be classified into active, poised and repressed enhancers. Especially during differentiation, enhancers dynamically transition between these states of activity. The transformation from a repressed or poised enhancer to one that activates transcription is accompanied by changes in transcription factor occupancy and histone modifications. In a simplified view, enhancers carry the following histone modifications: Active enhancers are flanked by H3K27ac and/or H3K122ac with or without H3K4 mono-methylation (H3K4me1) co-occupancy[40]. Poised enhancers may lack additional binding of signal-dependent transcription factors or activation by co-activators[1]. Poised enhancers traditionally show H3K4me1 peaks (without H3K27ac). Repressed elements show compact chromatin and repressive H3K27me3 mark[41]. Active and poised enhancers and promoters are either nucleosome-free or occupied by dynamic histone variants H3.3 and H2A.Z[42]. DNA accessibility can therefore be used to infer the binding of TFs and to predict regulatory elements. The following figure depicts the chromatin signature of promoters, poised and active enhancers.

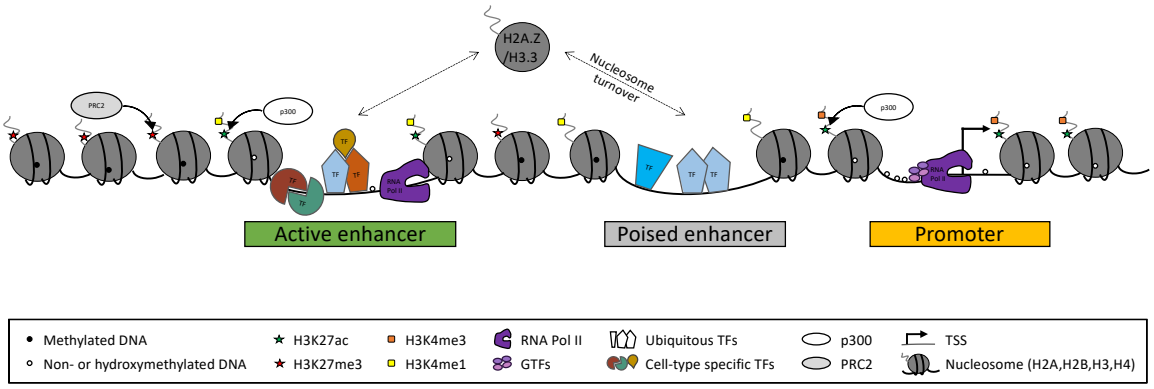


Figure 1: Chromatin signature at active and poised enhancers and the promoter. *Cell-type-specific and ubiquitously expressed TFs bind the active enhancer. HAT p300 binds both the active enhancer and the promoter and acetylates H3K27. The active enhancer can also be enriched for H3K4me1 and marked by non-methylated or hydroxymethylated DNA. RNA Pol II activity at active enhancers can lead to eRNA transcription. The poised enhancer lacks the H3K27ac mark, but can be flanked by H3K4me1. Active promoters are characterised by DNA hypomethylation and distinctly marked by H3K4me3. All regulatory elements show DNA accessibility, either because of nucleosome depletion or dynamic histone H3.3/H2A.Z turnover. PRC2 methylates H3K27 at inactive regions. Adapted from Long et al.[43].*

1.1.2.4 Transcription factor binding

In response to an extra-cellular signal, the activity and/or concentration of enhancer-bound transcription factors is increased[1]. Transcription factors regulate precise spatially- and/or temporally-specific gene expression and thereby drive differentiation. Transcription factors can broadly be clustered based on their DNA binding domain (DBD), such as homeobox, helix-loop-helix and cysteine-rich zinc-finger. While most transcription factors tend to bind accessible chromatin, some do have the ability to bind to nucleosomal DNA. These pioneer transcription factors diffuse through the nucleus or slide across the chromatin to find a specific DNA binding motif[44]. Upon binding they initiate local opening of chromatin and linker histone displacement[45]. Nucleosomes can also be displaced when multiple transcription factors cooperatively bind to an enhancer and compete with nucleosomes for binding to the DNA[46]. After nucleosome eviction, other transcription factors which preferentially bind to regions of open chromatin are recruited. The initial binding of a transcription factor promotes binding of other transcription factors to an enhancer, either caused by conformational changes in its protein structure (protein-

protein interaction) or by changing the physical property of the DNA (DNA-facilitated interactions)[47].

Enhancers harbour multiple degenerate DNA binding motifs for transcription factors (around 6 - 10 bp)[46]. Three quarters of all transcription factors recognize and bind to a specific DNA sequence. The preference for the DNA binding domain of a transcription factor to bind to its specific DNA motif can be around 1,000 fold compared to other sequences[2]. In general, transcription factors bind enhancers in an orientation-independent manner[11]. Nevertheless, transcription factors which form heterodimers and cooperatively bind DNA do have distinct preferences for the orientation and spacing of their binding site[48]. This orientation, spacing and order of binding sites is called the enhancer motif grammar. Highly conserved enhancers are more likely to be dependent on a rigid motif grammar[43]. While only a subset of enhancers is estimated to be conserved between metazoans, super-enhancers and enhancers that regulate developmentally important genes are especially highly conserved[49]. Interestingly, enhancer conservation depends on the type of tissue in which they are active; for example, enhancer conservation across vertebrae was 44% for forebrain-specific enhancers compared to only 6% of heart-specific enhancers[50]. So apart from transcription factor binding and dynamic histone modifications, a high level of sequence conservation in non-coding genomic regions can also help in the identification of enhancers[51].

Several transcription factors collectively bind to one enhancer with either repressing, additive or synergistic effects[47]. Depending on the context and cell type, the same transcription factor can have either an activating or repressing function and with that,

enhancers can turn into silencers[52]. Repressors may block binding of activating transcription factors and silencers can compete with enhancers for promoter binding. Transcriptional regulation has evolved to be highly robust and transcription factors are frequently redundant, especially within transcription factor families. Members of one transcription factor family often bind similar motifs and can compete for enhancer binding. Even though it is unlikely that a single transcription factor is responsible for the acquisition of a certain cell fate, important key players are often referred to as master regulators. Combinatorial control of gene regulation ensures specificity given that less than half of all transcription factors are expressed in a tissue-specific manner[53]. This synergistic interaction also provides increased sensitivity, ensuring that subtle changes in the concentration of one co-factor results in robust transcriptional activity[54]. In general, the number of transcription factors bound to an enhancer indicate its accessibility and positively correlate with gene transcription[43].

1.1.3 Three-dimensional nuclear architecture

For transcription to actively occur, promoter and enhancers do not directly need to touch, but need to be in close proximity (100 - 330 nm)[55]. The three-dimensional (3D) configuration of chromatin determines the physical proximity between enhancers and promoters and therewith regulates gene expression[56]. And nuclear architecture is reciprocally shaped by gene activity.

1.1.3.1 Chromatin domains

The 3D architecture of the nucleus broadly separates the genome into active and inactive compartments[1]. Chromatin is further assembled into domains called topologically

associating domain (TAD) which exhibit high levels of internal interactions. At least two thirds of promoter-enhancer interactions occur within TADs[31], and only some long-range interactions are able to cross TAD boundaries[57]. TADs are formed by loops and are separated from each other by regions of low interaction (where the loops connect) termed boundary elements. TAD boundaries serve as chromatin insulators and are marked by the presence of CCCTC-binding factor (CTCF) transcription factor and are stabilized by Cohesin. Insulators are crucial for inhibiting enhancer activity and can even prevent allele-specific promoter-enhancer interaction[27].

Higher chromatin order structure formed by TADs is largely unchanged across lineage commitment or even species[58]. TADs increase proximity and thus probability of specific promoter-enhancers communication. These increased intra-TAD promoter-enhancer interactions correlate with an increase in transcriptional activation[58]. Intra-TAD chromatin microarchitecture and thus promoter-enhancer interactions are dynamic and cell-type-specific[31]. During differentiation, around 60% of promoter-enhancer interactions are rewired[59] and specific to one cell type, which points towards cell-type-specific 3D folding of chromatin[57].

1.1.3.2 Promoter-enhancer interaction

The interaction between enhancers and promoters mediates correct spatial and temporal gene expression. Many promoters of developmental genes are regulated by several enhancers which act together in a dose-dependent manner[60]. The interaction of multiple enhancers increases specificity and restricts ectopic expression[60]. Around half of all promoters are regulated by more than one enhancer[26,57]. One promoter can even interact

with up to 20 enhancers, although it is of note that these data were derived from cell homogenates and not single cells[57]. But *vice versa*, 10% of enhancers can interact with promoters of four or more genes, with 55% of enhancers regulating only one gene within the same cell type[31]. This illustrates the complexity of interconnected chromatin.

In long-range gene regulation, promoter and distal enhancers are brought into physical proximity by DNA looping[61]. Cohesin mediates these distal element interactions by extrusion and maintains the loops. In addition to Cohesin, tissue-specific transcription factors can also assist looping[11]. While looping has been shown to be sufficient for gene activation at some loci (as exemplified at the super-enhancer of ß-globin)[62], more static and stable promoter-enhancer interactions can also be observed prior to transcription[63,64]. Poised enhancers are often in a pre-established contact with the promoter, with a promoter-proximal paused RNA Pol II waiting for the incoming release-triggering signal[63,65]. Dynamic looping of a second enhancer to an otherwise stable promoter-enhancer interaction can also initiate transcription[64].

Proximal enhancers frequently do not loop, but bridge the short distance between promoters and enhancers (less than 20 kb) *via* the tracking or linking model[66]. According to the linking model, several transcription factors sequentially polymerize towards a closely located promoter thereby opening up chromatin and stimulating transcription[39]. In the tracking model, enhancer-bound transcription factors and RNA Pol II track along the DNA unidirectionally towards the promoter. As RNA Pol II moves along the DNA, it essentially elongates and eRNA is produced as a by-product of tracking enhancer interaction.

Figure 2 is an example of intra-TAD looping, tracking or linking promoter-enhancer interaction leading to precise spatio-temporal gene expression. Enhancers achieve cell-type-specific gene expression *via* recruitment of transcription factor and co-activators.

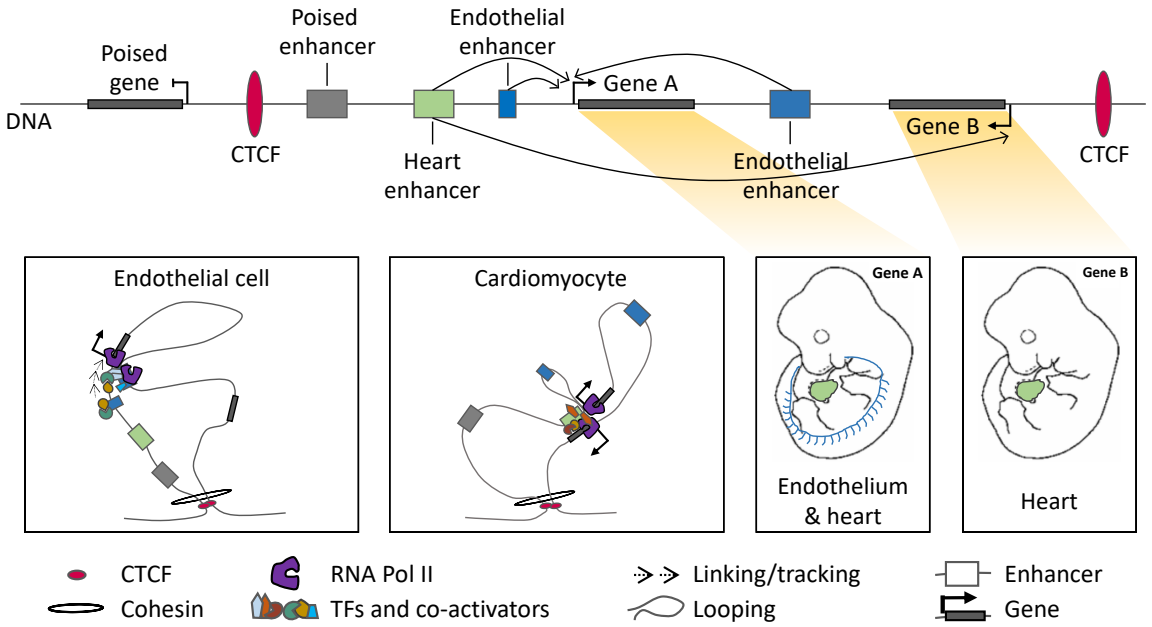


Figure 2: Illustration of typical regulatory elements regulating spatio-temporal-specific transcription of nearby genes. *This example depicts intra-TAD promoter-enhancer interactions which do not cross the TAD boundary to induce the gene outside the TAD ("poised gene"). CTCF marks TAD boundaries. Transcription factors collectively bind DNA motifs located at the endothelial or cardiomyocyte-specific enhancers which leads to RNA Pol II recruitment and subsequent expression of gene A and B. An enhancer with transcriptional inactivity during this step of development is poised. In endothelial cells, two endothelial-specific enhancers drive transcription of gene A. This overlapping activity provides robustness to gene expression. The proximal endothelial enhancer interacts with the promoter of gene A via linking or tracking, while the second, more distal endothelial enhancer loops to the promoter. In cardiomyocytes, an enhancer drives expression of two genes, gene A and B. This interaction between the promoters of the two genes and the cardiomyocyte-specific enhancer is formed by looping out intervening DNA. Alternative scenarios in which poised enhancers are already in a stable contact with the respective enhancer independent of gene expression is equally possible.*

1.2 Blood vessel development

Blood vessels in the human body form a 100,000 km maze consisting of arteries, veins and capillaries. Endothelial cells (ECs) line the entire circulatory system and provide a crucial interface between blood and the surrounding tissues. In humans, the total number of endothelial cells has been estimated to be 2.5 trillion[67]. The cardiovascular system is the first functional organ to form to ensure oxygen and nutrient supply for the developing embryo. Vascular development starts with vasculogenesis, the differentiation of mesodermal precursor into endothelial cells to form a primitive vessel network. Thereafter, the predominant method of vascular growth is angiogenesis where new vessels form, branch and prune from pre-existing vessels.

1.2.1 Vasculogenesis

Vasculogenesis describes the fundamental process during which angioblasts differentiate into endothelial cells. In mammals, *de novo* specification occurs first in the yolk sac and later in the embryo proper. In the yolk sac, precursor cells of the haematopoietic and endothelial lineages assemble into blood islands which will form a primitive vascular plexus. Within the embryo, angioblasts, endothelial precursor cells, arise from the mesoderm in the embryo during gastrulation. These cells proliferate and assemble to form the first primitive vascular plexus. By embryonic day (E) 8 angioblasts form the endocardium, paired dorsal aortae and cardinal veins[68]. After lumen formation and before the first heartbeat, a functional network is established[69]. Finally, mural cells are recruited and surround the endothelial layer.

In zebrafish, vasculogenesis only occurs intra-embryonically. Angioblasts from the lateral plate mesoderm form a precursor vascular cord[70]. Through ventral migration of endothelial cells from the primitive aorta, the cardinal vein is formed[70].

1.2.2 Angiogenesis

During angiogenesis new blood vessels form from a pre-exisiting vascular plexus. Angiogenesis is derived from the Greek word αγγείων meaning vessels and γένεση meaning birth. Angiogenesis occurs during development, but also in the adult during physiological processes such as wound healing and placental development, and during pathological processes such as retinopathy, arthritis and tumour growth[71]. Angiogenesis is a metabolically demanding process where previously quiescent endothelial cells are switched on and rapidly start proliferating and migrating in response to a pro-angiogenic stimulus[72].

1.2.2.1 Sprouting angiogenesis

Sprouting angiogenesis is the most studied and consequently best understood type of angiogenesis. During angiogenic sprouting, endothelial cells migrate towards a stimulus and proliferate, ultimately fusing with other sprouts to initiate circulation. To maintain integrity of the existing parent vessel, only a small number of specialised endothelial cells respond to the pro-angiogenic stimulus. These endothelial cells are called tip cells. Tip cells secrete metalloproteinases to degrade the extracellular matrix and the endothelial basement membrane of the existing vessels. These tip cells then migrate and invade the hypoxic avascular tissue[73]. Tip cells are rich in filopodia and sense repellent or attractant cues similar to axonal growth cones[74,75]. Following the leading edge of the sprout is a

second specialised type of endothelial cells, known as stalk cells. These are less motile, but more proliferative and elongating[76].

Angiogenic sprouts grow in response to pro-angiogenic signals secreted by oxygen- and nutrient-deficient tissue, such as vascular endothelial growth factor A (VEGFA). Several receptors are located on tip cells to guide angiogenic sprouting, including netrin receptor uncoordinated-5 homologue B (UNC5B) and VEGFA receptor 2 and 3 (VEGFR2/3). VEGFA signalling to tip cells triggers lateral inhibition *via* Notch signalling (see section 1.3.1.3). While VEGFA signalling promotes the tip cell phenotype, Notch signalling selects the stalk cell phenotype. This ensures that vessel integrity is maintained with only one or two tip cells positioned at the leading edge responding to the VEGF gradient. However, this tip/stalk cell arrangement is dynamic and VEGFR – Delta like ligand 4 (DLL4) – Notch signalling is constantly evaluated. Uninhibited cells in the stalk shuffle up to compete with tip cells for the tip position[77]. Figure 3 shows the key steps during sprouting angiogenesis. Further detail about tip and stalk cell selection is contained in section 1.3.1.3.

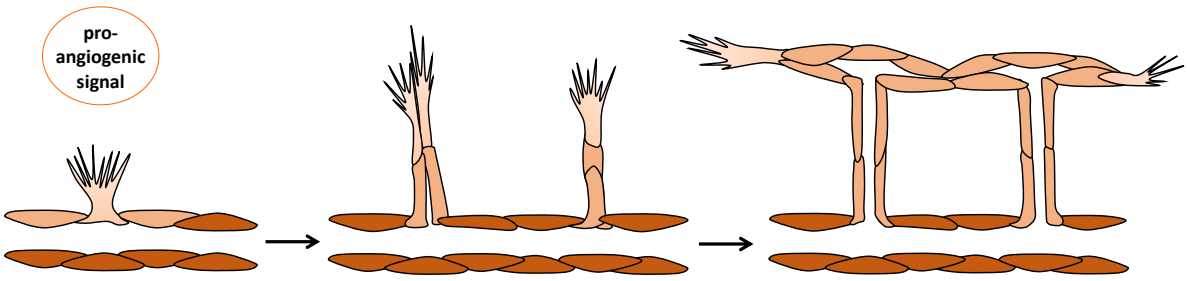


Figure 3: Steps of sprouting angiogenesis. *In response to a pro-angiogenic stimulus, an endothelial cell within the existing vessel gets selected as a tip cell. The neighbouring stalk cells form the trunk of the sprout and remain in contact with the vessel. Filopodia-rich tip cells migrate towards the stimulus and compete for the tip cell position. Trailing stalk cells proliferate and elongate. Tip cells ultimately anastomose with each other or another existing vessel. After lumen formation is complete, the newly formed vessel gets perfused and remodelling begins. Adapted from Potente et al.[78].*

Lumenization of unicellular tubes occurs either *via* intracellular vacuolization or invagination, while multicellular tubes lumenize *via* a cord-hollowing mechanism of tip cells[79]. Filopodia-guided sprouts ultimately anastomose to form a circulatory system. Anastomosis can be aided and accelerated by macrophages[80]. Thereafter, these microvascular vessels remodel to form a functional, newly perfused vascular bed. Once pro-angiogenic signals cease, vascular endothelial (VE)-Cadherin in endothelial junctions recruit phosphatases to deactivate VEGFR2 and thus reinforce quiescence[78]. The adult vasculature is ultimately quiescent with only 0.5% of all endothelial cells showing mitotic activity[81].

1.2.2.2 Intussusceptive angiogenesis

A less studied mechanism of angiogenesis is intussusception, where one blood vessel splits into two or more daughter vessels. Intussusception is characterized by insertion of trans-capillary pillars and, in contrast to sprouting, always occurs in already perfused vascular beds[82]. Intussusceptive microvascular growth allows the expansion of an existing capillary network. In the zebrafish caudal plexus, intussusceptive angiogenesis complements sprouting angiogenesis in areas with low haemodynamic forces[83].

1.2.2.3 Remodelling and mural cell recruitment

After angiogenesis is complete, an existing vascular network is remodelled and pruned. One of the mechanisms regulating remodelling is intussusception. During intussusceptive branching remodelling, blood vessels are fused and pruned in the proximity of bifurcations. Intussusceptive arborization is a process whereby a primitive vascular bed is restructured into a more mature, hierarchical vascular tree[82].

Already during angiogenesis, tip cells secrete Platelet-derived growth factor B (PDGFB) which causes pericyte recruitment to endothelial cells[84]. Pericytes loosely cover and therewith increase stability of the newly formed vessels and capillaries. As the vasculature matures, vascular smooth muscle cells are recruited to vessels with a larger diameter. These wrap around the endothelium in a continuous layer. While arteries have a thick layer of smooth muscle and a matrix of elastic fibres that protects from high blood pressure and pulsatile flow, veins are covered by a thinner layer of mural smooth muscle cells[78]. Endothelial and mural cells are additionally lined by an extracellular matrix of connective tissue.

1.2.3 Arteriovenous differentiation

Arteries carry oxygen-rich blood to the capillary network, where metabolites and waste are exchanged with the surrounding tissues. The oxygen-depleted, low pressure blood then gets transported back to the heart in veins. Veins are larger in diameter and contain valves to facilitate flow against gravity. Although genetically pre-determined, haemodynamic forces and oxygen levels are thought to reinforce arterial and venous identity[85].

Arterial or venous fate is already genetically specified in angioblasts before the onset of blood flow. While members of the Notch pathway, Neuropilin1 (NRP1) and Ephrin B2 (EFNB2) are arterial-specific, NRP2, Eph receptor B4 (EPHB4) and Chicken ovalbumin upstream-transcription factor II (COUP-TFII, encoded by the gene Nuclear receptor subfamily 2 group F member 2 (*Nr2f2*)) are venous-specific[85]. COUP-TFII is thought to downregulate Notch signalling, which allows the transcription of venous marker *EphB4*[85].

Arterial EFNB2 binding to its receptor EPHB4 promotes repulsion and mediates the segregation of arterial and venous vessels[74]. The mechanism promoting arterial identity has been extensively studied. In contrast, venous differentiation is not well understood. Also initially assumed to be acquired passively as the default endothelial state, the loss of venous identity after deletion of COUP-TFII suggests some degree of independent fate acquisition, although the upstream regulators are not well understood[86].

1.2.4 Vascular heterogeneity

As is already clear from the preceding sections, endothelial cells are far from forming a homogeneous population and actively assume a diverse array of different identities and functions. Already very early in embryonic development, endothelial cells adapt to accommodate tissue-specific functions depending on the surrounding environment. Arteries, veins and sprouting endothelial cells describe the most basic classification of specialized sub-endothelial cell types during development. The different types of endothelial cells found during sprouting angiogenic cells are also a functionally and morphologically separate population and are expected to have distinct molecular properties. Vascular heterogeneity is controlled by tissue-specific signalling and sub-endothelial cell-specific transcription. So far, genome-wide analysis of arterial and venous identity has only been studied in cell culture[87]. Detailed organ-specific transcriptional profiling in the adult mouse has also established that endothelial cells in the different vascular beds are fundamentally different from each other, adding another level of complexity[88].

Transcription factors are thought to play a key role in the establishment and maintenance of this heterogeneous vasculature, directly binding to *cis*-regulatory regions in the genome to modulate gene expression and thus drive differentiation towards different endothelial fates. However, our understanding of the gene regulatory pathways that direct endothelial differentiation and heterogeneity *in vivo* are incomplete.

1.3 Molecular mechanism of angiogenesis

1.3.1 Major signalling pathways involved in angiogenic sprouting

1.3.1.1 Hypoxia

Hypoxia and nutrient deprivation are the primary reasons for angiogenesis. The diffusion distance of oxygen is only 100 - 200 µm. The idea that blood vessels form in order to alleviate hypoxia was already suggested in 1962[89]. Under normoxic conditions Hypoxia-inducible factor (HIF) is hydroxylated by Hydroxylase domain proteins (PHD1–3) and therewith marked for degradation. During oxygen deprivation HIF is stabilised[90]. HIF1α binds HIF-responsive elements leading to the transcription of genes which will improve oxygen supply, such as pro-angiogenic factors, including VEGFA (see section 1.3.1.2)[91]. This transcriptional response will result in the formation of new blood vessels and re-establish oxygen levels to meet the demand[92].

1.3.1.2 VEGF signalling

VEGF signalling mediates a plethora of responses, including vascular permeability, invasiveness, chemotaxis, proliferation and differentiation. VEGFA-F and Placental growth factor (PIGF) form a large family of ligands than can bind to one or several of the three VEGF receptor (VEGFR) tyrosine kinases: VEGFR3, encoded as Fms related tyrosine kinase 4 (FLT4), and VEGFR2, also denoted as Kinase insert domain receptor (KDR) or Fetal liver kinase 1 (FLK1) and VEGFR1, encoded as FLT1, which possesses only weak kinase activity. In zebrafish, there is a fourth, diverged Vegfr named kdr-like (kdrl) which was lost in evolution[93]. Upon ligand binding, VEGFR form hetero- or homodimers, auto-phosphorylate and activate downstream signalling.

VEGFA is the most powerful and best-characterized pro-angiogenic ligand. Extracellular VEGFA binds to VEGFR2 located on endothelial cells. VEGFA-VEGFR2 signalling is absolutely essential for vasculogenesis, arteriogenesis and angiogenesis[94]. VEGFR2 can activate the Phospholipase Cγ (PLCγ)/Protein kinase C (PKC)/Mitogen-activated protein kinase (MAPK) pathway, but also the Phosphatidyl inositol 3-kinase (PI3K)/AKT pathway. Alternative splicing results in at least 16 VEGFA isoforms, with different solubilities and matrix-binding properties and varying abilities for co-receptor binding. The most prominent and potent isoform is $VEGFA_{165}$ ($VEGFA_{164}$ in mouse)[95]. Gradient-forming $VEGFA_{165}$ binds heparan sulphate on the extracellular matrix, but is also partially diffusible. With its intermediate properties, $VEGFA_{164}$ is the only isoform that sustains normal angiogenesis in the absence of other isoforms by promoting branching, proliferation and guided tip-cell migration[96]. There is also an anti-angiogenic, inhibitory $VEGFA_{165b}$ which is upregulated in the quiescent vasculature[95,97].

During arterial specification, VEGFA signalling upregulates the expression of arterial markers *EFNB2*, *NRP1* and components of Notch signalling[78,98]. Co-receptor NRP1 in turn enhances VEGFA/C-VEGFR2/3 signalling in endothelial cells[99]. During angiogenesis, VEGFA signalling is absolutely necessary for tip cell-typical filopodia morphology and endothelial proliferation. VEGFA establishes the tip and stalk cell pattern by inducing *DLL4* transcription (see section 1.3.1.3)[100]. VEGFR1 exists as a full length and a short soluble receptor. VEGFR1 binds VEGFA with a ten-fold higher affinity than VEGFR2 and functions as a decoy receptor. Classically, VEGFC/D-dependent VEGFR3 signalling is known to be required for lymphangiogenesis[101]. But VEGFR3 signalling also plays an important role in tip cells[101]. In the absence of VEGFR2, VEGFR3 is not sufficient to sustain sprouting

angiogenesis[102], although these findings were inconsistent[103]. A combinatorial regulation of VEGFR2/3 heterodimers has been proposed and further research is required to fully understand the contribution of VEGFC-VEGFR3 in sprouting angiogenesis[75].

1.3.1.3 Notch signalling

Notch signalling is crucial for arterial specification and angiogenesis. In arteries, Notch ligand *DLL4*, which is one of the five Delta, Serrate, Lag-2 (DSL) ligands, is the first arterial-specific gene expressed in the forming dorsal aorta. Notch signalling triggered by DLL4 induces arterial-specific genes, such as *EFNB2*[104]. In the absence of any component of the DLL4-Notch signalling pathway, endothelial differentiation fails. For example, *Dll4*$^{-/-}$ mutant mice die of severe vascular defects due to failure of arterial differentiation[105]. DLL4-Notch signalling is not confined to arteriovenous specification, as it is also crucial for tip and stalk cell selection.

In the tip cell during sprouting angiogenesis, VEGFA activates expression of a variety of genes including *DLL4 via* the PI3K/AKT pathway[106,107]. NOTCH transmembrane receptors are expressed in stalk cells further down the sprout and are activated by the DLL4 of adjacent tip cells. Upon activation in neighbouring stalk cells, NOTCH is proteolytically cleaved twice which releases the intracellular domain of NOTCH (NICD) that translocates to the nucleus[108]. Nuclear NICD binds to the transcription factor Recombining binding protein suppressor of hairless (RBPJ) and activates transcription through the formation of a trimeric complex with Mastermind (MAM)[109]. Upon Notch signalling, VEGFR2/3 levels are decreased, while VEGFR1 expression increased. In the stalk cell, the decoy receptor VEGFR1 (encoded by *FLT1*) sequesters VEGFA away and reduces VEGFA-responsive gene

induction due to its low tyrosine kinase activity[72,110]. This inhibits the filopodia rich, highly VEGFA-responsive tip cell phenotype in the stalk cell to maintain the newly formed sprout[107,111]. Additionally, Notch ligand Jagged1 (JAG1) expressed in stalk cells competes with DLL4 for NOTCH receptor binding. NOTCH receptors get glycosylated by Manic Fringe glycosaminyltransferase, which represses signalling upon JAG1 binding. JAG1 thereby antagonises Notch signalling *via* DLL4 in tip cells[112,113]. Genetic depletion or chemical inhibition of DLL4 lead to a hypersprouting defect[114,105]. In contrast, endothelial-specific *Jag1* knockout leads to impaired angiogenesis[112]. In line with these findings *Dll4* is expressed in tip cells, while *Jag1* is predominantly expressed in stalk cells[112]. Even though angiogenic sprouting is a very dynamic process, the respective cell types acquire distinct transcription profiles at least transiently[77,115]. Figure 4 illustrates the discussed tip and stalk cell key players.

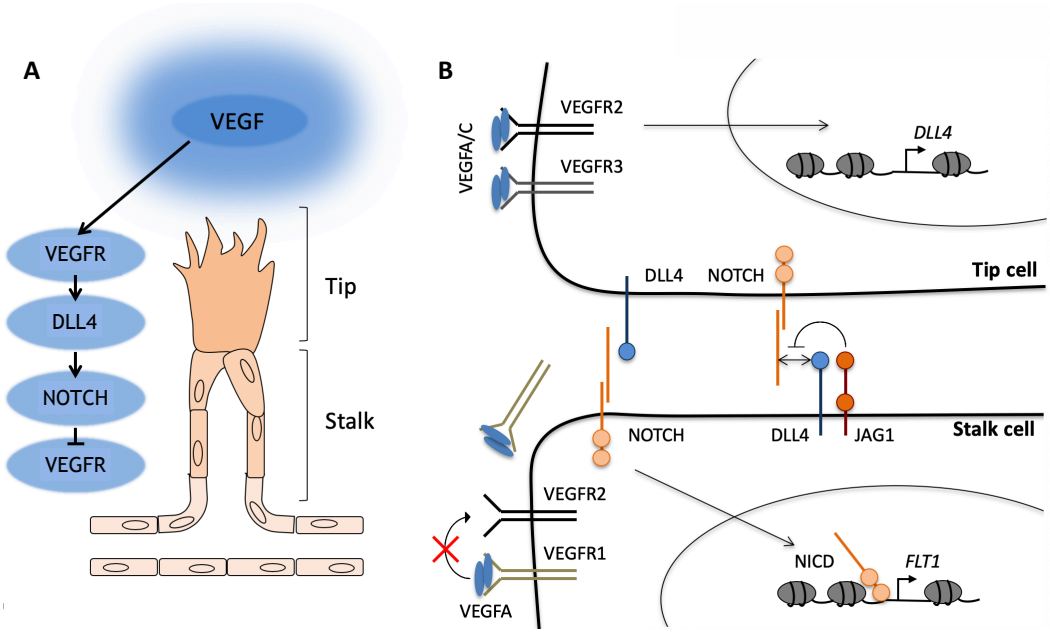


Figure 4: Tip and stalk specification in endothelial sprouting. A) *During angiogenic sprouting, a VEGF gradient activates tip cells while DLL4-NOTCH signalling represses the tip cell phenotype in stalk cells.* ***B)*** *Upon VEGFA/C signalling VEGFR2/3 form hetero- or homodimers to induce transcription of DLL4 in tip cells. DLL4 activates Notch signalling in adjacent stalk cells which in turn induces VEGFR1 expression (encoded by FLT1). Soluble and membrane-bound VEGFR1 serves as a decoy receptor and reduces VEGF binding to VEGFR2. JAG1 blocks the interaction between DLL4 and NOTCH, thereby reducing the tip cell phenotype. For consistency, I used the nomenclature adopted for human genes in the figures. Figure was adapted from Herbert and Stainier[74].*

1.3.2 Transcription factors involved in sprouting angiogenesis

The transcriptional regulation of vasculogenesis has been extensively studied[116], but our understanding of the transcriptional mechanism regulating sprouting angiogenesis is more limited[117]. In total, around 20 transcription factor families have been suggested to play a role during sprouting angiogenesis[118]. Amongst these are Activator protein 1 (AP1) heterodimers[119] and members of the SRY-related HMG-box (SOX)[120,121], the Sma/Mad (Mothers against decapentaplegic) homolog (SMAD)[122] and Forkhead (FOX) family[123]. But evidence for their requirement during sprouting angiogenesis *in vivo* and understanding of the underlying transcriptional mechanism is incomplete[118]. The following section only focusses on the transcription factors relevant for this thesis.

1.3.2.1 E26 transformation-specific factor family

E26 transformation-specific (ETS) transcription factors are crucial at every step during vessel formation by directing haematopoiesis, vasculogenesis and angiogenesis. Upon VEGF or Fibroblast growth factor (FGF) signalling[124], ETS factors bind to regulatory elements and upregulate genes known to be important for endothelial identity, such as *VEGFR2*, VE-Cadherin (encoded by *CDH5*), T-cell acute lymphoblastic leukemia (*TAL1*) and Myocyte enhancer factor-2 C (*MEF2C*)[125]. There are 27 members of the ETS family, of which at least 19 are expressed in human endothelial cells[126]. The ETS family can further be subdivided into clades based on homology of the winged helix–turn–helix ETS DNA binding domain[126]. While single *Ets1*$^{-/-}$ or *Ets2*$^{-/-}$ do not show a severe vascular phenotype in mice, homozygous *Ets1* and *Ets2* knockout causes embryonic lethality due to defective blood vessel branching during angiogenesis[127]. As a result of combinatorial gene regulation and redundancy, depletion of single ETS transcription factors often do not show

a profound endothelial defect[128]. Even though other family members can usually compensate for severe defects, the ETS factor ETS variant 2 (ETV2) has a crucial, non-redundant function. *Etv2*$^{-/-}$ mice die at E9.5 due to haematopoietic defects and a complete lack of blood vessel formation[129]. Knockdown of functional *Etv2* zebrafish orthologue ets1 related protein (*etsrp*) in zebrafish results in similarly extensive vascular defects and loss of all endothelial markers[130]. ETV2 has been considered a pioneer factor, opening up chromatin of regulatory elements at the onset of endothelial cell fate acquisition[131,132]. *Etv2* is expressed during embryonic angiogenesis and neo-angiogenesis during wound healing[133]. But since *etsrp* expression decreases during zebrafish intersegmental vessel formation and conditional endothelial-specific *Etv2*$^{-/-}$ mice lack a phenotype during later angiogenesis stages, ETV2 is generally not thought to play a unique crucial role during physiological angiogenesis[134]. Depending on post-translational modifications, ETS factors can repress or activate transcription. ETS domain-containing protein ELK3 for instance can act as a transcriptional repressor in quiescent cells, but turns into a transcriptional activator upon MAPK-dependent phosphorylation during the angiogenic switch[135,136].

ETS factors regulate gene expression in multiple sub-endothelial cell types. For example, the ETS-related gene (*Erg*) is constitutively expressed in all endothelial cells, including during mesoderm specification and angiogenesis[137,138]. But out of all ETS factors, ERG is most highly expressed in the quiescent vasculature[139]. Another example is ETS1 and E74-like factor 1 (ELF1), both expressed throughout the vasculature and known to regulate both embryonic vascular development and tumour angiogenesis[140]. Given that ETS factors are expressed throughout vascular development, and have been shown to be essential for various steps of endothelial differentiation[138], any specificity of gene expression in

different endothelial sub-types is thought to be achieved *via* the recruitment of co-factors[134]. For example, GATA2 serves as a co-factor for ETV2 during haematopoietic and endothelial lineage commitment[141]: ETS factors can form a large transcriptional complex with GATA2, TAL1 and Lim domain only 2 (LMO2) during endothelial differentiation[116]. FOX factors also work combinatorially with ETS factors during vasculogenesis: ETS factors, in particular ETV2, bind synergistically with FOX factors, particularly FOXC1/2[125], to a composite binding motif located in regulatory elements of endothelial genes (the FOX:ETS motif) and induce transcription of many essential early vascular development genes. Similar to the GATA2:ETS and FOX:ETS transcriptional complex, other cooperative protein-protein interactions are likely to mediate transcriptional specificity in different sub-endothelial cell types, but these have not yet been elucidated.

1.3.2.2 Myocyte enhancer factor-2 family

The Myocyte enhancer factor-2 (MEF2) family of MADS (MCM1, Agamous, Deficiens, Serum response factor)-box transcription factors control differentiation and organogenesis[142]. There are four MEF2 family members: A, B, C and D. MEF2 factors form homo- and heterodimers and have hence partially overlapping, but also distinct, functions[143,144]. Depending on the post-translational modifications and co-factor interactions, MEF2 factors can either activate or repress transcription. Out of the four family members, *MEF2A* and *MEF2C* share the highest sequence homology with 98% of the MADS domain being conserved. In contrast, *MEF2B* is the most divergent and shows the lowest degree of conservation with other family members (with only 68% conserved between the MEF2 domain of *MEF2A* and *B*)[145]. MEF2 factors are widely expressed, but have higher levels of expression in myogenic and neural lineages[142].

In mice, *Mef2A*, *Mef2C* and *Mef2D* family members are expressed in the cardiovasculature[146,147], with at least *Mef2A* and *Mef2C* known to be expressed specifically in endothelial cells[148,149]. While *Mef2D*$^{-/-}$ knockout mice appear normal, *Mef2A*$^{-/-}$ knockout mice show 80% peri-natal mortality with cardiovascular defects[142] and *Mef2C*$^{-/-}$ knockout mice die at E9.5 from severe cardiovascular defects[148]. Interestingly, endothelial-specific *Mef2C*$^{-/-}$ mice do not suffer from obvious vascular defects, possibly due to redundancy between the MEF2 family members[150]. Such redundancy has been reported in skeletal muscle, where combined deletion of *Mef2A*/*Mef2C* and *Mef2D* is required to see any phenotype[151].

MEF2C is the most studied protein of the MEF2 family and its expression in murine embryonic endothelial cells is mediated by ETS and FOX factors[152]. In human cell culture, VEGFA increases MEF2C transcriptional activity and *MEF2C* mRNA induction[153,154]. *In vitro* studies have shown that MEF2C is required for endothelial cell survival and migration[155,156]. In *Dll4*$^{-/+}$ post-natal mouse retinas, endothelial cells are hypersprouting and *Vegfr2* and *Mef2C* mRNA were upregulated[114].

1.3.2.3 H2.0-like homeobox factor

The H2.0-like homeobox (HLX) transcription factor is a non-clustered homeobox protein, which has been implicated in embryogenesis and tumour growth[157,158]. Like all homeobox family members, HLX contains a helix-turn-helix DNA-binding homeodomain. *HLX* is overexpressed in in 87% of patients with acute myeloid leukaemia where it promotes proliferation[159,160]. *Hlx*$^{-/-}$ knockout is lethal in mice at E15. These studies have demonstrated a role for HLX during yolk sac remodelling[161], the enteric nervous system

and visceral organogenesis[162,163]. Up to E11, HLX is not essential for the development of embryonic vasculature, although the investigated time points might have been too early to detect a defect[161]. This emphasizes the need for endothelial-specific *Hlx* knockdown.

More recently, studies in zebrafish have shown that the orthologous *hlx1 is* expressed specifically in the angiogenic front in actively sprouting endothelial cells (see Figure 5)[164]. Moreover, hlx1 is absolutely required for intersegmental vessel (ISV) formation[164]. Herbert *et alia* (*et al.*) treated zebrafish embryos with VEGF receptor inhibitor SU5416 which inhibited sprouting and altered gene expression of 109 genes[164]. The most VEGF-dependent gene was *hlx1*. *In vitro*, human HLX was also strongly and specifically induced by pro-angiogenic VEGFA *via* VEGFR2, but not by PIGF, Epidermal growth factor (EGF) or Interleukin 1 (IL1)[154]. Moreover, *HLX* knockdown inhibited sprouting angiogenesis *in vitro*[161].

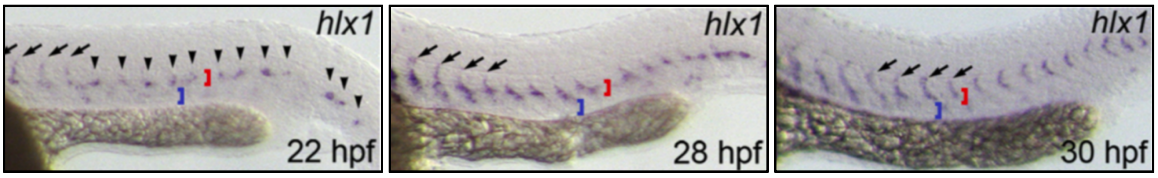


Figure 5: In zebrafish, hlx1 is expressed in sprouting endothelial cells. *Sprouting endothelial cells (black arrows) and regions of future angiogenic remodelling (arrowheads) show hlx1 expression, while hlx1 is absent from the dorsal aorta (red brackets) and cardinal vein (blue brackets). Whole-mount in situ hybridization of hlx1 at 22 hours post fertilization (hpf), 28 hpf and 30 hpf. Images were taken from Herbert et al.[164].*

1.4 Hypothesis and aim of the study

Although many transcription factors have been shown to influence sprouting angiogenesis, a comprehensive understanding of the transcriptional mechanism regulating sprouting-specific gene expression is lacking. If we want to intervene with angiogenic growth during pathological conditions (for example during cancer), it is paramount that we understand the regulation of angiogenic endothelial cells. We hypothesize that understanding the inherent logic of a sprouting-specific enhancer will allow us to pinpoint important transcription factors and extrapolate the upstream regulatory pathway during angiogenesis. However, no sprouting-specific *cis*-regulatory elements have been identified[165]. I shall therefore screen the loci of the sprouting-specific *HLX* and stalk cell-specific *JAG1* for putative enhancers *in silico*, aiming to find an enhancer that specifically drives reporter gene expression in sprouting endothelial cells. Once a sprouting-specific enhancer has been identified and verified in both zebrafish and mouse transgenic models, I plan to analyse the DNA sequence of this conserved enhancer. I hypothesize that motif analysis combined with mutational analysis will allow me to find the key transcription factors that mediate the specificity of this enhancer, and consequently of angiogenic genes more generally. The effects of genetic deletion of this identified transcription factor will then be examined in angiogenesis model systems. Lastly, I will use this angiogenic-specific enhancer to generate a transgenic zebrafish line expressing reporter genes specifically in sprouting endothelial cells, allowing me to isolate sprouting endothelial cells for genome-wide chromatin accessibility and gene expression analysis.

My thesis has the following aims:

- Identification of an angiogenic-specific enhancer (Chapter 3).
- Elucidation of the transcriptional pathway that regulates sprouting fate through analysis of the angiogenic-specific enhancer (Chapter 4).
- Generation of a genome-wide map of *cis*-regulatory elements in endothelial arterial, venous and sprouting cells *in vivo* through the identification of open chromatin regions and relate these open chromatin regions to transcription (Chapter 5).

2 Materials and methods

2.1 Transgenic mice

2.1.1 Generation and breeding of transgenic mouse lines

All animal procedures were approved by local ethical review and licensed by the UK Home Office in accordance with the Animals (Scientific Procedures) Act 1986 (PPL 30/3324). Mice were bred and housed at the Functional Genomics Facility in the Wellcome Trust Centre for Human Genetics (WTCHG). Transgenic mice were generated on a mixed CBA x C57BL/6 background by pro-nuclear injection into the oocyte by Dr Ke Liu (Musculoskeletal Biology, Institute of Ageing and Chronic Diseases, University of Liverpool)[166]. Sequences for all transgene enhancer sequences are provided below (section 2.4.2). Dll4in3:*LacZ* mice were already established and generated by Dr Natalia Sacilotto[167]. Stable transgenic lines were generated by crossing HLX-3:*LacZ* positive male founders with wild-type C57BL/6JOlaHsd mice (Envigo, Cat# 057). Induced endothelial cell-specific *Mef2A/C* knockout mice [*Cdh5(PAC)Cre-ER^{T2};Mef2A$^{fl/fl}$;Mef2C$^{fl/fl}$*, denoted here as *Mef2A/C^{iECKO}*][168,169,170] were generated and bred by Dr Yao Wei Lu and Dr John Schwarz (Center for Cardiovascular Sciences, Albany Medical College). For post-natal retina analysis, recombination was induced by 50 µl intra-gastric tamoxifen injections (1 mg/ml in sunflower-seed oil containing 2.5% ethanol) at post-natal day (P)1, P2 and P3 for harvest at P5 and P6. For embryonic hindbrain analysis, 100 µl intra-peritoneal tamoxifen injection (10 mg/ml in sunflower-seed oil containing 5 mg/ml progesterone and 5% ethanol) occurred at E9 and harvest at E12.

2.1.2 *LacZ* Genotyping

Tissue samples (either from ear biopsies or from embryonic yolk sacs) were digested overnight (O/N) at 55 degree Celsius (°C) in 180 µl GNT buffer [50 millimolar (mM) KCl, 1.5 mM MgCl2, 10 mM Tris-HCl pH 8.5, 10% gelatin, 0.45% Nonidet P-40, 0.45% Tween-20] with 10 µl 10 mg/ml proteinase K (Cat# 3115801001, Sigma). The solution was heat-inactivated at 95°C for 10 minutes (min). Subsequently, 0.8 µl (for ear notches) or 2 µl (for yolk sacs) of the digested supernatant was used for polymerase chain reaction (PCR) with the GoTaq G2 Green master mix (Cat# M7823, Promega) together with primers as indicated in Table 1 (total volume 20 µl). PCR products were run on 1.5% agarose gel at 180 V and genotype was determined by the presence of two bands.

Table 1: Primers and cycling conditions used for LacZ genotyping.

Cycling conditions	
94°C, 3min	
94°C, 30sec 55°C, 30sec 72°C, 30sec	Repeat 28 times
72°C, 5min	

Primers	
Hsp68LacZ F	CAAGCGAAGACAAGAGAAGCAGAGCGAGC
Hsp68LacZ R	AAACCAGGCAAAGCGCCATTCGCCATTCAG
Internal Control F	TGGACAGGACTGGACCTCTGCTTTCCTAGA
Internal Control R	TAGAGCTTTGCCACATCACAGGTCATTCAG
Use 1 µl of 10 µM primer per reaction	

2.1.3 Neo-angiogenesis assays

Prior to neo-angiogenesis studies, male transgenic mice were crossed with wild-type (WT) females to confirm genotype through analysis of transgene activity in embryonic F1 offspring [as assessed by 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) staining due to β-galactosidase activity at previously characterized embryonic time points].

2.1.3.1 FGF Matrigel plug

Matrigel Basement Membrane Matrix (Cat# 354234, BD) was mixed with basic Fibroblast growth factor (bFGF, Cat# 450-33, Peprotech) to a final concentration of 2 µg/ml bFGF. Transgenic HLX-3:*LacZ* mice were anesthetized with isoflurane and subcutaneously injected in the flank with 400 µl of bFGF/Matrigel. Matrigel plugs were harvested eight days (n = 3), nine days (n = 5), eleven days (n = 1), 14 days (n = 7) after injection. Internal organs were also harvested from each mouse and subjected to X-Gal staining, providing an internal control.

2.1.3.2 B16F10 allograft

B16F10 melanoma cells were kindly provided by Professor Adrian Harris (Department of Oncology, Weatherall Institute of Molecular Medicine, University of Oxford) and grown in Dulbecco's modified Eagle's medium (DMEM; Glutamax, Gibco) containing 10% fetal bovine serum (FBS) at 37°C in 5% CO_2 to near confluency. B16F10 melanoma cells were counted and resuspended in serum-free DMEM to a concentration of $2x10^6$ cells/ml. Next, 50µl of Matrigel Basement Membrane Matrix (BD) were mixed with 50 µl containing $1x10^5$ cells. HLX-3:*LacZ* transgenic mice were anesthetized with isoflurane and subcutaneously injected in the flank with 100 µl B16F10 cells/Matrigel. B16F10 allografts were measured daily by Vernier calliper and animals were monitored for distress using the Mouse Grimace Scale. Allografts were harvested when tumours reached a size of 10 - 11.5 mm. This corresponded to eight days (n = 2) and eleven days (n = 6) after injection. As with Matrigel experiments, internal organs were also harvested from each mouse and subjected to X-Gal staining, providing an internal control.

2.1.3.3 Staining of Matrigel plugs and B16F10 allografts

Both Matrigel plugs and B16F10 allografts were cut into two, one half was used for cryosections and the other half was used for X-Gal staining and paraffin-embedding. Bladder and ribs from each animal were also harvested and X-Gal stained to confirm transgene expression.

For X-Gal staining, allografts and Matrigel plugs were fixed in 2% PFA, 0.2% glutaraldehyde in Phosphate-buffered saline (PBS) for 20 min on ice and then washed twice in PBS for 20 min at room temperature (RT). Allografts were stained with X-Gal staining solution and imaged as described in section 2.2.1. For histological analysis, paraffin-embedded allografts and Matrigel plugs were cut into 5 µm sections and counterstained with nuclear fast red (Cat# 26356-02, Electron Microscopy Sciences).

For cryosections, allografts and Matrigel plugs were fixed in 4% Paraformaldehyde (PFA) in PBS for one hour on ice and then washed twice in PBS for 20 min at RT. Tissue was taken through 15% and subsequently 30% sucrose in PBS until the tissue sunk to the bottom of the container at 4°C. Tissue was then washed in a 50/50 mix of 30% sucrose/Optimal cutting temperature compound (OCT, TissueTek). After two washes in OCT, tissue was embedded over dry ice and stored at -80°C until sections were cut at 15 µm. To verify co-localization of X-Gal staining with endothelial cells in B16F10 allografts, 15 µm cryosections were first stained with X-Gal and then analysed by immunofluorescence. Slides were fixed in 4% PFA in PBS, 2 mM $MgCl_2$, 5 mM ethylene glycol tetraacetic acid (EGTA) for 10 min at RT. Slides were then stained with X-Gal staining solution [1 mg/ml X-Gal in dimethylformamide (DMF) dissolved in 0.1 M PBS pH 7.3, 2 mM $MgCl_2$, 0.01%

sodium deoxycholate, 0.02% Nonidet P-40, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide] for 30 min at RT. Stained sections were rinsed twice in PBS and post-fixed in 4% PFA in PBS for 10 min at RT. Slides were rinsed three times in PBS and then washed for 15 min in PBS containing 0.5% Triton X-100. Slides were blocked in 10% donkey serum in PBS and 0.5% Triton X-100 for one hour at RT. Antibodies were diluted in blocking buffer. Endothelial cells were visualized using anti-Cluster of Differentiation 31 (CD31) antibody (1:300, Cat# DIA-310, Dianova) and biotinylated Isolectin B4 (IB4, 1:100, Cat# B-1205, Vector Laboratories) O/N at 4°C. After three PBS washes, secondary Alexa Fluor®-conjugated 488 donkey anti-rat Immunoglobulin G (IgG, 1:300, Cat# A21208, Thermo Fisher) and Streptavidin AF546 (1:300, Cat# S11225, Thermo Fisher) were incubated at RT for one hour. Tissue was covered with a drop of Vectashield mounting medium containing 4',6-diamidino-2-phenylindole (DAPI, Cat# H-1200, Vector Laboratories) for imaging with the Zeiss 710MP confocal microscope using the Zen software. To compare the activity of enhancer:*LacZ* transgenes in sprouting blood vessels between Matrigel plugs and B16F10 allografts, X-Gal positive vessels were quantified per CD31 positive vessels. X-Gal positive vessels were counted in Matrigel plugs and B16F10 allografts from HLX-3:*LacZ* (n = 2, each).

2.2 Histology

2.2.1 Fixing and staining of embryos

For timings of embryonic development, vaginal plugs were used as an estimate for the onset of gestation. Transgenic embryos were collected at indicated time points and dissected from uteri and rinsed in ice-cold PBS. Embryos were fixed in 2% PFA, 0.2% glutaraldehyde in PBS at 4°C: E8 embryos were fixed for 10 min, E9 embryos were fixed for 30 min, E10 and E11 embryos were fixed for one hour, E13 embryos were fixed for two hours, E14 embryos were fixed for two and a half hours and E15 were deskinned and fixed for three hours. After fixation, embryos younger than E11 were incubated for 30 min in PBS and then stained while embryos older than E11 were rinsed twice for 20 min in rinse solution (0.1% sodium deoxycholate, 0.2% Nonidet P-40, 2 mM $MgCl_2$ in PBS) at 4°C. Embryos were stained O/N at RT with 1 mg/ml X-Gal (Cat# 437132J, VWR) dissolved in DMF in stain solution (5 mM $K_4Fe(CN)_6$, 5 mM $K_3Fe(CN)_6$, 0.1% sodium deoxycholate, 0.2% Nonidet P-40, 2 mM $MgCl_2$ in PBS). The following day, embryos were rinsed in PBS and fixed O/N in 4% PFA at 4°C. Skin was taken from E13 X-Gal stained embryos. For the clearing protocol, E14 embryos were X-Gal stained as normal and then cleared. First, embryos were fixed O/N in 4% PFA at 4°C. Thereafter, embryos were incubated in a series of glycerol (20%, 50%, 80%, 100%) in 1% KOH solution for three to five days each at 37°C[171]. Whole-mount images were taken with a binocular dissecting microscope (M165C, Leica) equipped with a ProGres CF Scan camera and ProgRes CapturePro software (Jenoptik). Thereafter, embryos were processed for histological analysis.

2.2.2 Paraffin embedding and sectioning

First, embryos were dehydrated through a series of ethanol washes, cleared with xylene and incubated in liquid paraffin using a tissue processor (Excelsior AS, Cat# A82300001, Thermo Scientific). Embryos were then embedded in desired configuration in paraffin using the HistoStar Embedding Workstation (Cat# A81000002, Thermo Scientific). Embryos were sectioned by Indrika Ratnayaka (Ludwig Institute for Cancer Research, University of Oxford) using HM 355S Automatic Microtome (Cat# 905200, Thermo Scientific) and 5 µm sections were mounted onto slides. Sections were deparaffinised through a series of Histoclear II (National Diagnostics, HS-202) and ethanol washes (5 min each), rehydrated in water for 2 min and counterstained with nuclear fast red for 30 seconds (Cat# 26356-02, Electron Microscopy Sciences). Sections were then dehydrated through a series of ethanol and Histoclear II washes (5 min each), then mounted with Vectamount permanent mounting medium (Cat# H-5000, Vector Laboratories). Nuclear fast red stained slides were imaged with a Leica DM5500 B Upright Microscope using a ProgRes C5 microscope camera (and ProgRes CapturePro software, Jenoptik) or the NanoZoomer S210 with the NDP.view2 viewing software (Hamamatsu).

2.2.3 Immunostaining followed by X-Gal staining

E11 and E13 embryos were fixed in 4% PFA in PBS at 4°C, then taken through a sucrose gradient as described in section 2.1.3.3. Cryosections were allowed to thaw at RT before washing in PBS to remove the OCT embedding medium. Cryosections were then permeabilised in 0.5% Triton X-100 in PBS for 10 minutes. They were then washed in PBS and incubated for one hour in blocking solution [10% donkey serum (Cat# D9663, Sigma),

4% FBS, 0.2% Triton X-100 in PBS] at RT. Sections were incubated O/N at 4°C in primary antibodies diluted in the blocking solution. The following antibodies were used: armenian hamster anti-CD31 (1:100, Cat# ab119341, abcam), mouse anti-α-Smooth muscle actin (αSMA)-Cy3™ (1:200, Cat# C6198, Sigma) and rat anti-Endomucin (EMCN, 1:50, Cat# sc-65495, Santa Cruz). After further PBS washes, sections were incubated for one hour at RT with AF-conjugated secondary antibodies (1:500, Life Technologies or abcam) in blocking solution. After final PBS washes, DAPI staining and mounting of slides using Fluoromount™ aqueous mounting medium (Cat# F4680, Sigma), confocal images were acquired with a Zeiss 710 MP confocal microscope.

To compare β-galactosidase expression with these endothelial markers, X-Gal staining was performed immediately following this immunostaining protocol. After confocal imaging, slides were soaked in MilliQ water overnight at 4°C, cover slips removed and the X-Gal staining protocol performed (see section 2.2.2).

2.2.4 *In situ* hybridization

The pCRII Hlx plasmid (*Hlx* mRNA target: 1330 – 1956 bp) was linearized by restriction enzyme digest (with BamHI for *Hlx* antisense probe and with EcoRV for sense negative control probe). T7 RNA polymerase was used for *in vitro* transcription of the antisense probe and Sp6 for the control probe. Probes were labelled with digoxigenin (DIG)-UTP using the DIG RNA Labelling Kit (Cat# 11175025910, Roche).

For *in situ* hybridization, transverse sections of wild-type E13 embryos were cleared from paraffin twice for 5 min in xylene. Slides were rehydrated through 1 min washes through

an ethanol gradient (100%, 90%, 70%, 40%, each step twice). Slides were washed in PBS for 5 min and treated with 20 µg/ml proteinase K (Cat# 3115801001, Sigma) in 50 mM Tris pH 8, 5 mM ethylenediamine tetraacetic acid (EDTA) for 7 min at RT. After one wash in PBS for 5 min, slides were fixed in 4% PFA for 20 min at RT. Subsequently, slides were washed again for 5 min in PBS and 2x Saline-sodium citrate [SSC (300 mM NaCl, 30 mM sodium citrate, pH 7)]. Last, slides were dehydrated through a series of ethanol washes (40%, 70%, 90%, 100%) and left to dry. Subsequently, 1 µg/ml of Hlx probe was diluted in 100 µl hybridization solution [50% formamide, 10% dextran sulphate, 1 mg/ml yeast tRNA, 1x salt (0.2M NaCl, 10 mM Tris HCl pH 7.5, 5 mM $NaH_2PO_4.2H_2O$, 5 mM Na_2HPO_4, 5 mM EDTA), and 1x Denhardt's solution (Cat# J62416, Alfa Aesar)]. The diluted Hlx *in situ* probe was incubated on the slide in a moist chamber filled with a 1:5 mixture of 20x SSC and formamide O/N at 65°C. The following day, sections were washed in 50% formamide in 1x SSC/0.1% Tween-20 at 65°C for 20 min. After two 15 min washes in 2x SSC at 37°C, 10 µg/ml RNase A was incubated for 15 min. Thereafter, sections were washed again in 2x SSC and 0.1x SSC for 15 min each at 37°C. After a 20 min wash in MABT (150 mM NaCl, 100 mM Maleic acid, 0.1% Tween-20, pH 7.5), slides were incubated in blocking solution [2% (w/v) blocking reagent (Cat# 11096176001, Roche), 10% sheep serum in MABT] for one and a half hours. Anti-DIG antibody conjugated to alkaline phosphatase (Cat# 11093274910, Roche) was diluted in blocking solution (1:400 in 2% blocking reagent and 1% sheep serum in MABT) and slides were incubated O/N at RT in a water based moist chamber. The next day, slides were washed twice for 15 min in MABT and then 5 min in NTMT (100 mM Tris pH 9.5, 100 mM NaCl, 10 mM MgCl2, 1% Tween-20). For the staining, slides were covered in BM Purple (Cat# 11442074001, Roche) and 3 mM levamisole at

37°C a water based moist chamber for three days. Sections were dehydrated and imaged as previously described.

2.2.5 Preparation of whole-retina flatmounts

Transgenic HLX-3:*LacZ* mice at P4, 5, 6, 7, 8, 15 and 18 were sacrificed directly by cervical dislocation. Eyes were enucleated and fixed for one hour in 4% PFA in PBS on ice. After two washes in PBS, retinas were dissected by removing the lens, cornea, pigment epithelium, sclera and hyaloid vessels (as previously described[172,173]). Eyes from *Mef2A/C*^{iECKO} mice were harvested at P5 and P6. Eyes were fixed for 90 min in 4% PFA in PBS on ice and stained as described in section 2.2.5.2.

2.2.5.1 Retina X-Gal staining

Retinas from HLX-3:*LacZ* mice were rinsed twice for 20 min in rinse solution (0.1 M PBS pH 7.3, 2 mM MgCl2, 0.01% sodium deoxycholate, 0.02% Nonidet P-40) at RT. Retinas were stained O/N in 500 µl of staining solution (1 mg/ml X-Gal in DMF dissolved in 0.1 M PBS pH 7.3, 2 mM MgCl2, 0.01% sodium deoxycholate, 0.02% Nonidet P-40, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide) at 37°C to visualize β-galactosidase activity. After two 20 min washes in rinse solution and three PBS washes, retinal cups were post-fixed for 10 min in 4% PFA. Retinas were further stained with Isolectin B4 which detects α-D-galactosyl residues on glycoproteins on the basement of endothelial cells[174].

2.2.5.2 Retina immunostaining

Endothelial cells were stained with biotinylated *Griffonia simplicifolia* Isolectin B4 (1:100, Cat# B-1205, Vector Laboratories) in PBS containing 0.1% Triton X-100 O/N at 4°C. Retinas

were washed in PBS containing 0.1% Triton X-100 and then incubated with specific secondary antibodies Streptavidin AF546 (1:300, Cat# S11225, Thermo Fisher) in PBS containing 0.1% Triton X-100 O/N at 4°C. Where applicable, P5 retinas were also stained with anti-DLL4 antibody (10 µg/ml, Cat# AF1389, R&D systems) and anti-goat AF488-conjugated secondary antibody (1:300, Cat# A11055, Thermo Fisher). Following the staining, the retina was cut 2/3 from the periphery down to the optic nerve creating four leaves and flattened onto the microscopy slide. Retinas were covered with a drop of Vectashield mounting medium containing DAPI (Cat# H-1200, Vector Laboratories) for imaging with the Zeiss 710MP confocal microscope. In order to merge the X-Gal staining recorded in colour transmission images with the fluorescent channel, images were unmixed using ImageJ[175]. An ImageJ unmixing macro was kindly developed by Mark Shipman (Ludwig Institute for Cancer Research, University of Oxford). The blue X-Gal stain was transformed into false-coloured green which was then overlaid with the pan-endothelial red Isolectin B4 staining.

Whole-mount P7 retinas were also stained with anti-HLX antibody. Here, wild-type retinas were fixed in methanol and kept at -20°C. Retinas were rehydrated through a gradient of Methanol and PBS containing 0.2% Triton X-100 and ultimately washed in PBS containing 0.2% Triton X-100. Whole-mount retinas were blocked in 10% goat serum in 0.2% PBS-Triton for two hours at RT. Retinas were stained with anti-HLX (1:4000, Cat# 09-084, Millipore) and anti-CD31 (1:50, Cat# DIA-310, Dianova) in 0.1% PBS-Triton X-100, O/N at 4°C in the dark. After three washes in 2% goat serum in 0.2% PBS-Triton, secondary AF546 goat anti-rabbit IgG (1:300, Cat# A11035, Thermo Fisher) and secondary AF488 donkey anti-rat IgG (1:300, Cat# A21208, Thermo Fisher) were incubated in 10% goat in 0.2% PBS-

Triton at RT for one hour. Retinas were rinsed five times in 2% goat serum in 0.2% PBS-Triton for 15 min each and imaged as described in the section above.

2.2.5.3 Phenotypic analysis of post-natal retinal angiogenesis

The total number of branch points, tip cells and retinal outgrowth length were measured with ImageJ software from pooled images of P6 *Mef2A/C*^{iECKO} retinas from at least two independent litters. The ImageJ plugin AnalyzeSkeleton by Ignacio Arganda Carreras[176] was used to perform branching analysis of Isolectin B4 stained retinas. Tip cells were counted and marked manually with the ImageJ cell counter and the outgrowth length was measured for all quadrants of each retina with the straight line tool.

2.2.6 Preparation of hindbrain flatmounts

Hindbrains from E11 - E12 embryos were dissected from embryos without prior fixation, the roofplate was opened and the hindbrain was flat-mounted[177]. Hindbrains from HLX-3:*LacZ* and HLX-3mutMEF2:*LacZ* were X-Gal stained according to the same protocol as embryos (see section 2.1). Hindbrains from *Mef2A/C*^{iECKO} mice and wild-type controls were further processed for Western blotting (see section 2.5.3.2).

2.2.7 Preparation of tibia

Tibia and fibula from four week old pups were dissected. Bones were fixed for four hours in 4% PFA on ice. After three washes in PBS, bones were decalcified in 0.5 M EDTA at 4°C for three days. After three washes in PBS, bones were incubated for 24 hours in 20% sucrose and 2% polyvinylpyrrolidone (PVP 360,000 g/mol, Cat# P5288, Sigma) at 4°C. Bones were incubated and then embedded in OCT (TissueTek) for cryosectioning. Then,

15 µm cryosections were cut by Indrika Ratnayaka. Slides were X-Gal stained and counterstained with nuclear fast red (Cat# 26356-02, Electron Microscopy Sciences).

2.3 Transgenic zebrafish

2.3.1 Husbandry

All animal procedures were approved by local ethical review and licensed by the UK Home Office in accordance with the Animals (Scientific Procedures) Act 1986 (PPL 30/3324). Adult zebrafish were maintained at 28.5°C in system water at the Biomedical Services housed in the Weatherall Institute of Molecular Medicine (WIMM). Zebrafish embryos were maintained in E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$ and 10^{-5}% Methylene Blue) at 28.5°C. Stable transgenic zebrafish lines contained an enhancer:fluorescent reporter construct, *exempli gratia* (e.g.) green fluorescent protein (GFP): *tg(Dll4in3:GFP), tg(Ephb4-2:GFP*) and *tg(kdrl:HRAS-mCherry*) fish lines were already established and generated as previously described[167,178,179]. The stable line *tg(hlx-3:GFP)* was generated by outcrossing adult F0 founders generated using the Tol2 system[180]. Adult positive F1s were inter-crossed to achieve brighter reporter gene expression and transgene expression was assessed in the more stable F2 generation. To visualize the entire vasculature together with arterial and non-arterial cells, *tg(Dll4in3:GFP)* was crossed with *tg(kdrl:HRAS-mCherry).* To visualize sprouting cells as well as pan-endothelial cells, *tg(hlx-3:GFP)* was also crossed with *tg(kdrl:HRAS-mCherry).*

2.3.2 Tol2 transposon transgenesis in zebrafish

F0 transient mosaic transgenic zebrafish embryos were generated using Tol2-mediated integration[180]. Wild-type embryos were injected at the one-cell stage with 0.5 nl of 50 ng/µl Tol2 transposase and 60 ng/µl pE1b-enhancer-GFP/Citrine/mCherry/Cerulean expression vector using a FemtoJet® microinjector (Eppendorf). For simultaneous analysis of two putative enhancers in one transient transgenic fish, 25 ng/µl of pE1b-enhancer-Citrine and 25 ng/µl pE1b-enhancer-mCherry expression vectors were injected (per 0.5 nl). Tol2 transposase capped mRNA was generated from the linearized pCS-TP plasmid using the mMESSAGE mMACHINE SP6 kit (Cat# AM1340, LifeTechnologies).

2.3.3 Morpholinos and small molecule inhibitors

Control (CCTCTTACCTCAGTTACAATTTATA) and Tnnt2 morpholinos[181] (ATGTTTGCTCTGATCTGACACGCA) were dissolved in ultrapure water following the published guidelines[182]. Morpholinos (MO) were injected into the yolk of one-cell *tg(hlx-3:GFP)* embryos at a concentration of 1 nl of 4 ng/µl morpholino per embryo. Morpholino experiments were conducted at least three times. For pharmacological inhibition of Vegf and Notch signalling pathways, embryos were manually dechorionated at 22 hours post fertilization (hpf) and incubated with 0.63 micromolar (µM) SU5416, 2.5 µM of SU5416 (Cat# S8442, Sigma) or 100 µM DAPM (Cat# 565777, Calbiochem). Control embryos were treated with identical volumes of dimethyl sulfoxide (DMSO) only. Inhibitor experiments were conducted two times.

2.3.4 N-Phenylthiourea treatment and embedding

For imaging zebrafish embryos older than 48 hpf, fish were treated with 0.03 mg/ml N-Phenylthiourea (PTU, Cat# P7629, Sigma) at 24 hpf to prevent melanogenesis. Embryos were embedded in 0.4% TopVision low melting point agarose (Cat# R0801, Thermo Scientific) in 0.14 mg/ml Tricaine mesylate (MS222, Cat# A5040, Sigma) in glass bottom multi-well culture plates (MatTek) as previously described[170].

2.3.5 Image acquisition and analysis

Embryos were dechorionated and anesthetized with 0.1% MS-222 tricaine (Sigma). To test putative and mutant enhancers, transient transgenic zebrafish were assessed at 30 – 34 hpf using the Olympus MVX10 Macroscope. For each mutant/putative enhancer, the percentage of fish with any detectable reporter gene expression per total number of injected fish and percentage of fish with reporter gene expression in ISVs per transgenic fish was recorded. Qualitative assessment of reporter gene expression was performed on at least 24 fish per transgene. Zebrafish were imaged as tile scans of z-stacks collected using a confocal microscope Zeiss LSM 710 MP.

2.3.6 Whole-mount *in situ* hybridization

The zebrafish hlx1 *in situ* hybridization probe was kindly provided by Dr Shane Herbert[164]. For antisense probe generation, the pCR-BluntII-TOPO-*hlx* plasmid was digested with BsaI, and T7 polymerase was used for *in vitro* transcription. *hlx1* antisense probe and sense negative control probe were labelled with DIG using the DIG RNA Labelling Kit (Cat# 11175025910, Roche). Embryos were collected at 25, 27, 28 and 30 hpf and fixed in 4%

PFA at 4°C O/N. Fixed embryos were dehydrated and stored in 100% methanol at -20°C. Embryos were rehydrated using a methanol/PBS containing 0.1% Tween-20 (PBST) gradient and washed four times in PBST for 5 min. To allow probe penetration, embryos were treated with 10 µg/ml proteinase K (Cat# 3115801001, Sigma) for 15 min (< 28 hpf) or 30 min (30 hpf). After three quick washes in PBST, embryos were post-fixed with 4% PFA for 20 min. After five washes in PBST, embryos were blocked in hybe buffer [50% formamide, 5× SSC, 0.1% Tween-20, 50 µg/ml heparin, 500 µg/ml torula yeast RNA (Sigma), 10 mM citric acid] for two hours at 65°C. Hybe buffer was exchanged with similar buffer containing 1 ng/µl of the antisense probe for an O/N incubation at 65°C. The following washes were performed in an *in situ* robot (Biolane HT1, intavis). First, samples were taken through a hybe buffer/2× SSC (containing 0.1% Tween-20) gradient at 65°C. Then, embryos were washed for 10 min in 2× SSC, followed by four 15 min washes in 0.2× SSC (containing 0.1% Tween-20) at 65°C. Embryos were taken through a dilution series of 0.2× SSC/MABT (0.1 M Maleic Acid, 0.15 M NaCl, 0.1% Tween-20, pH 7.5) ending with a 5 min wash in 100% MABT at RT. Zebrafish were incubated for one hour with MAB block reagent (MABT with 2% Boehringer block reagent), followed by an O/N incubation with anti-DIG antibody (1:2500, Roche) at 4°C. Embryos were washed three times in MABT and subsequently in AP buffer (100 mM Tris-HCl pH 9.5, 50 mM MgCl2, 100 mM NaCl, 0.1% Tween-20). Zebrafish were stained for eight hours with BM Purple (Roche) at 37°C and the reaction was stopped by fixation in 4% PFA. To remove pigments, embryos were bleached (10% H_2O_2, 5% formamide, 0.5% SSC) and washed three times in PBST for 5 min before being transferred to 80% glycerol for imaging.

2.4 Cloning of enhancer constructs

2.4.1 Gateway cloning

GeneArt® Strings™ (LifeTechnologies), which are custom-made double stranded DNA fragments, were cloned into reporter vectors with Gateway technology (Invitrogen). To add adenosine overhangs (tailing), 16.2 µl of 10 ng/µl of DNA fragments/putative gene regulatory regions were incubated with 0.5 µl Taq polymerase (Cat# 10342-053, LifeTechnologies) together with 2 µl 10x Taq buffer, 0.6 µl 10 mM deoxynucleotides (dNTPs) and 0.7 µl 50 mM $MgCl_2$ at 72°C for 20 min. Next, 4.5 µl of the PCR product were cloned into pCR8/GW/TOPO TA entry vectors (Cat# K2500-20, Invitrogen) according to the manufacturer's instructions. Putative enhancers were cloned from pCR8/GW entry vectors into the destination vector using LR recombination following manufacturer's instructions for the Gateway LR Clonase II Enzyme mix (Cat# 11791-100, LifeTechnologies). This generated expression vectors containing the fragment upstream of a silent promoter and a reporter gene. The hsp68:*LacZ*-Gateway vector was used for mouse transgenesis while the E1b-GFP-Tol2 vector was the primary vector used for zebrafish transgenesis (both provided by N. Ahituv). Additionally, the E1b:Citrine/Cerulean/mCherry Tol2 vectors used for zebrafish transgenesis were made as described in section 2.4.3. For pro-nuclear injections, hsp68-based plasmids were digested with NotI and the linear DNA fragment was gel purified with the QIAquick gel extraction kit (Cat# 28704, Qiagen) according to the manufacturer's instructions, then resuspended in 30 µl H_2O at a concentration of at least 120 ng/µl. No digestion was required for zebrafish vectors.

2.4.2 Enhancer sequences

2.4.2.1 HLX putative enhancer (human sequences)

HLX-35:GFP/LacZ (tested in zebrafish and mouse)
CACTTTAGCGTGAAATTCTCTGATACCAGCAGGGAGTTTTTAAACACAACTGGCAACAAGAAGCATATGGTCAGAGA
TTCAAGTTACTGCTTTCTTTCTTTCTTTTTTTTTTTTTTAATTGTTGTCTGCCAAGGACAATATATTTTTTTCCCTTCTTCAA
AATAAACATTGTGCAGGAATGTGCCTCAGCAAATGGCTGGGGCTCTATTATATTTCATGCAGCTTTCTAAACAATTTT
CAACCCTGGGCTGGCCTGGTGTCAAACAGAAGGCAACCAACAATAGCAAAATCCCATTAAAGCCAAGAAGGTTCCTT
CCTGAGGTGAACCTGGCCAGAAATAGTCATCTTGGCTCATTCTCCAGTTTGGCACTGGGGACTAATAATGCACGGTCT
CCTAATGGAAATTATCTAATTTCATCCAAGAGCTCCTCCAGAGAGAGCCGGGCTGACGCACTAACAGTCAATATAACA
ACATTCTTCTTGAAGTGCTGAACAATAGGCCTTTGGAGAGACTGTGATAAACCACGTCAGACTCCGGACAAACACGA
GTGTAAAAATAGGCATCAGAGGCATTGACACAGCC

HLX-5:GFP (tested in zebrafish)
CCAGCCCTGTCCCAGCCCACCAGGCCCGGGCATCTCCCTCCCTGGAGGAGCCCACCTAGTACTGAAAGTTGGTTCTCT
GCTGACACCTGCTGGACCATCTCTACATTGTTTTATGGCAGTAACCATCCTGTGCCAGCAACATTGTTATGATGAAAG
TTGACTCTTCCTTTTGCAAAAATGCAGGTTTCTAAAACTGGCTTTGACTAGCAGCTGTTTCCTTCCATATTGCAACAAT
ATGGGAAGATCTTGTATTGTGTGCTTAAAATCACTAGACCGGGCAAGGGAAGCAGTCATGGTTCAGGAATAATACTG
TGTACATTTTACAGATCTCAAGGTCAATTCACATTTATTATCTCACTTTTTTAAAATAACCAACTGTAGGATGCCGACG
AGGGATTATTATCCCATTTGACTTTAGGAAGTTGATCCTCAGAAAGGTTAAGTGGCTTGCCCAAGGTGACGTAGTTG
GGAAATGGTGTTGGAGGGAACAGAGGCTAGGTTTTCTGGTTCATCAACCGGTATTATACCATCCAGCCAATTTCTTCA
GGGGAA

2.4.2.2 HLX-3 enhancer (human sequences)

HLX-3:GFP (tested in zebrafish)
CAGAACCAAGAAAGCAGCAGGAAACCCTGGGCTCGCCGGAGTTCAGGTCATCTCCTGTTTTTTGATGGGCTCAAACT
CGGGACTAAATGAGCGGATGTGCCTGCGAGATTGGTCTCAAAATAGAAGCGGCCTTAGTATTTTAGGATACTGCAG
AAATCCTTATTTCTTCCGGAGAAAAACTCTATAATGACTCCTAGCACTTCCCCTTTCAATATTTGTGCAACTTTATCTCC
ACCGGCAAGTTGCTCGGGTGACGCAAATCCGCCCTTATGTATTACAACTACCAAAATAAATGCACGGAAATCAATTCC
CTTCTAGCAATGAA AATATAAATATACAGCTGCGAC

HLX-3:LacZ (tested in mouse) – used to generate the stable line
TCTCCGCTGCGTAGACCAAGCTCAAGTTAATTCGGTTTAGCCAGGCCCTTTTCTTTCCTGTTAAATTCTCAACCCCCAT
GGCTACCAGCCCCTGGGGAGCCCTCGCCTCTACCCCCAGGACCTCCCTGCCCGGCTCAGTAGCTCTGGGCCAGTTCAT
TTCGCCACAGCCCGGAGCCTCTCGGAAAGCAAAGATTAAAGGGGAAAGTCGCAGCTGTATATTTATATTTTCATTGCT
AGAAGGGAATTGATTTCCGTGCATTTATTTTGGTAGTTGTAATACATAAGGGCGGATTTGCGTCACCCGAGCAACTTG
CCGGTGGAGATAAAGTTGCACAAATATTGAAAGGGGAAGTGCTAGGAGTCATTATAGAGTTTTTCTCCGGAAGAAA
TAAGGATTTCTGCAGTATCCTAAAATACTAAGGCCGCTTCTATTTTGAGACCAATCTCGCAGGCACATCCGCTCATTTA
GTCCCGAGTTTGAGCCCATCAAAAAACAGGAGATGACCTGAACTCCGGCGAGCCCAGGGTTTCCTGCTGCTTTCTTG
GTTCTGAAGAGTGGGGAGTAAGGAGGGCGGGAGTCTGCGGGCTCAGAACTCGGCGAGGGGCCTGCAGGGGCCAG
GCTTGGGCCTGGGGAAGGGGTAGAGGGGGCGGCGGGGGTCGCTCCAAAGACTTGTATTTCGCGTTTGCCTCCGGG
AGCTGGGAGTAAGGCCTTGGATGGCGCCGACGCGGTTGCGAGGAAGCTGAGGCCTGGGAGAGCAAGGGGCGCGC
AGGCGAAGTTGCAACTTGCACTCCAGCCGCGGGCCTGGCGGAGAAAGGGAGGCTCGCGGCGCCGCGAGGAGTCGG
CGGGCCTCGGGGCCTCGCTTTCGCCGCATCTGCCCAGCGCTCCGGGCCTTGAATCTCGGCAGATGCGAGTTGTGGGC
ACCTAGGGAACCCTGAGGACTCGCATTCCCCCGGGTCTGTATCCCGTGCCCACCCCGGAGCGTCGCAAACCTTGGAA
AGGGTGAAAGCTGATAGGGGAGTCTTATTTCTTAAGAGAGAGGGTGAGGGCGCGCGGCGCTGCCTCGGGTAAGGA
GTAAATGCACCTCTTTCACAACCGTGAGTTTGGGCTTAGGCTGAGGTCTAAATGTTTGCAG

HLX-3mutMEF:LacZ (tested in mouse)
TCTCCGCTGCGTAGACCAAGCTCAAGTTAATTCGGTTTAGCCAGGCCCTTTTCTTTCCTGTTAAATTCTCAACCCCCAT
GGCTACCAGCCCCTGGGGAGCCCTCGCCTCTACCCCCAGGACCTCCCTGCCCGGCTCAGTAGCTCTGGGCCAGTTCAT
TTCGCCACAGCCCGGAGCCTCTCGGAAAGCAAAGATTAAAGGGGAAAGTCGCAGCTGTATATTTACGCCCTCATTGC

TAGAAGGGAATTGATTTCCGTGCATTCGCCCTGGTAGTTGTAATACATAAGGGCGGATTTGCGTCACCCGAGCAACT
TGCCGGTGGAGATAAAGTTGCACAAATATTGAAAGGGGAAGTGCTAGGAGTCATTATAGAGCCCCCCTCCGGAAGA
AATAAGGATTTCTGCAGTATCCTAGGGCGCTAAGGCCGCTTCCGCCCTGAGACCAATCTCGCAGGCACATCCGCTCAT
TTAGTCCCGAGTTTGAGCCCATCGGGGGGACAGGAGATGACCTGAACTCCGGCGAGCCCAGGGTTTCCTGCTGCTTTC
TTGGTTCTGAAGAGTGGGGAGTAAGGAGGGCGGGAGTCTGCGGGCTCAGAACTCGGCGAGGGGCCTGCAGGGGC
CAGGCTTGGGCCTGGGGAAGGGGTAGAGGGGGCGGCGGGGGTCGCTCCAAAGACTTGTATTTCGCGTTTGCCTCC
GGGAGCTGGGAGTAAGGCCTTGGATGGCGCCGACGCGGTTGCGAGGAAGCTGAGGCCTGGGAGAGCAAGGGGCG
CGCAGGCGAAGTTGCAACTTGCACTCCAGCCGCGGGCCTGGCGGAGAAAGGGAGGCTCGCGGCGCCGCGAGGAGT
CGGCGGGCCTCGGGGCCTCGCTTTCGCCGCATCTGCCCAGCGCTCCGGGCCTTGAATCTCGGCAGATGCGAGTTGTG
GGCACCTAGGGAACCCTGAGGACTCGCATTCCCCCGGGTCTGTATCCCGTGCCCACCCCGGAGCGTCGCAAACCTTG
GAAAGGGTGAAAGCTGATAGGGGAGTCTTATTTCTTAAGAGAGAGGGTGAGGGCGCGCGGCGCTGCCTCGGGTAA
GGAGTAAATGCACCTCTTTCACAACCGTGAGTTTGGGCTTAGGCTGAGGTCTAAATGTTTGCAG

2.4.2.3 hlx-3 enhancer (zebrafish sequences)

hlx-3:GFP/LacZ (tested in zebrafish and mouse) – used to generate the stable line
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATAGATTTCCTCTGTTTATTATAAGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGGAAGTGCTAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGGAAATAAGGGTGCTCTGCTGTATCCTAAAATACTAAAGCATCTTCTATTTTTGGCCTAAATCTTGCAGGAATAT
CCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3:GFP 650 bp (tested in zebrafish)
GCATGTGTGCATTCTCTTCAGTGCCGACCGTGTTTTTATTTTTTATTATTATTTTTTTATTATTATGCTAACAAACTATGA
TTAAAAAATGATCCACAATCAAATGCTAACACTCCGACTCGTGTGTGTGTGTATGTGTGTGTGTTTTCTGCTCTGTGG
AATTAGGAAACAGGAGTTTCTCGATGCAGGTCATCCTGTTTTTGATGAGCTCTGATTCAGGCTAAATGAGCGGATATT
CCTGCAAGATTTAGGCCAAAAATAGAAGATGCTTTAGTATTTTAGGATACAGCAGAGCACCCTTATTTCCTCCGACTC
ACAGAGGTGGAACTTTATAATGACTGTTAGCACTTCCTTTTTTCAATATTTGTGCAACTTTATCTCAACTCGCAAGTCG
CACGGTGACGCAAATCCGTCCTTATGCTTATAATAAACAGAGGAAATCTATTCTCCGCTGAAACGAAAACACAATTAA
CTACTGAGGCGATACTTCATCCATATGCACATGACAACTTTACATTGCTGTGCCAAAAATAGAGCGTGCTGGAAAGAC
GAAAACAGATTTGAAGAAAAAAAAATGCATGTATATAGATAAGCCACATATGCATGACAACTTTACATTACTGGCAA
AAAATAGAGCGCTTTACACAATG

hlx-3:GFP 850 bp (tested in zebrafish)
TAACCATTTATTCTAATAGATGCATTATATTTATACAAAAAAAAAGTTTCATTTAGATTAACACACAATATACTTCAAAT
AGAAAAAATACTGTTTTACACGCATGTGTGCATTCTCTTCAGTGCCGACCGTGTTTTTATTTTTTATTATTATTTTTTTAT
TATTATGCTAACAAACTATGATTAAAAAATGATCCACAATCAAATGCTAACACTCCGACTCGTGTGTGTGTGTATGTG
TGTGTGTTTTCTGCTCTGTGGAATTAGGAAACAGGAGTTTCTCGATGCAGGTCATCCTGTTTTTGATGAGCTCTGATTC
AGGCTAAATGAGCGGATATTCCTGCAAGATTTAGGCCAAAAATAGAAGATGCTTTAGTATTTTAGGATACAGCAGAG
CACCCTTATTTCCTCCGACTCACAGAGGTGGAACTTTATAATGACTGTTAGCACTTCCTTTTTTCAATATTTGTGCAACT
TTATCTCAACTCGCAAGTCGCACGGTGACGCAAATCCGTCCTTATGCTTATAATAAACAGAGGAAATCTATTCTCCGC
TGAAACGAAAACACAATTAACTACTGAGGCGATACTTCATCCATATGCACATGACAACTTTACATTGCTGTGCCAAAA
ATAGAGCGTGCTGGAAAGACGAAAACAGATTTGAAGAAAAAAAAAATGCATGTATATAGATAAGCCACATATGCATG
ACAACTTTACATTACTGGCAAAAAATAGAGCGCTTTACACAATGCATTAAAAACCACAACATATATTAAAAATTAAAA
TGCATCTATAGAAATAATCTAAATTTGTTACAATGCATTAAAAACAACAACAGATTTTGAAAGT

2.4.2.4 JAG1 putative enhancer (human sequences)

JAG1in2:GFP (tested in zebrafish)

ACTTAGGCTGCTGTGCCAAGGTTGTGTTCTCTGCCCTGTCCCACTCCATGGCGGTTACCTTGGCACTGGGCAGAGCGG
TAAGCACTTAATTGCCACATTAAGAAGGGGGGATGTAAGAAACAAAGGTAATTGTGTTGGCAGCCGGGAGCTTCCTT
TTCCTGGCAACAGTGACAGCCAAACCAAGCCCTGAGAAGCGGCTGCCTCCCCTCCTGCCAAGCTGACGGACACAAAT
AAACCGCCATCCAAACGGCACGCGTGCGGCGTCCAAAACCCCAGGCTGCTGCGACAGGCAGGCCCGCCGGCGGCAC
GCACATCCTGGGAATGCCCGACGTGCCAACGCGGCCATTGTGCACAGTCAGTCGGCTCCCAGGGTTCGTCAGGGGC
CCTGAGCTCTGCTGACTCAACTGTTGCTTCCTGTCGCTTGTCCTGTGTTAGCATTATTTTTGGAGAGGCCCTATTATTCT
CCATAATCCTCATTTCAGTCAATTCGGCCAACTCTATTGTCCTGCCTGTGTCTGGGTGGCCAGAGTATAAAATATTCCA
GTCACAAAACATATTAAGCAGAAAAGGTTGCAGGGAGTCCATGTGTGGCGCATGCACACGTACTCATGCACACAATG
CGAG

JAG1in5:GFP (tested in zebrafish)

GAGCGTGGAGGGTGGAGGGTGGAGGGTGGGGTGCGGGTGGGATGGGCAACATGTTTCCAGAACGGGACCTGGGA
AAAGCCACGATCAGCGCCAACCCCCTAAGCTGCTCAGCTGTAACTCTGGAAAACAACCCCCGCCCCCTCTAGCTCCCC
AGAAGGAAAGGCTTTCGGTTTCCCAAGAGCGGGAGGCAGGCTAGCAAATCCTTGTCATGGGAAAATTATAGGGATC
TAGCTAATAGCTAATCACCCACGGCTGAGAGTCAACTTAACCCCTGAGAATACTGACTAACCACTTCCTAACTTCTGA
AACCAAGCTATGCACACCCACATGCAAAGAAATCACTAGAATGTGTGGTTTAAAGTTAGACACTGGATTCTGGTTTAT
GTAAAATGGAATGTCAACGTGACAAGTGACTAAAGGAGTAAGATAAGAACAAAATTACAATCTAAATAATTTCTCAA
AAGCGGACACCCACGTTAGTGGGAGGTAAGGGTGTAGCCATTCTGTCAAGGC

JAG1in5:LacZ (tested in mouse)

CCAGCCGACTCCTGACAAAAGGAGGGCTGTTTGGGAAGAGGCCAGCATGGGGATGCCCTTCACAAAGCCACTCTCT
TCTCAGGCTCACACGAGTGAGCCAAAGGTGAAGGCATGGAATAAATCAACCCGTGCACCTAGTTGCGACAGACTATA
GAGGAAAGGACTCTCTCGGTGACAAGAGAGCTATGCCAAAGTCTCTGCAGACGCTGCACTTGCCACTCTGCTCTTTG
TCTGCACAGAAAGTCATGACCCCAGTGCACAGAAATGCAACAAGCCACCAAGTGCAGCAGGAAACTGCCTCACGAT
GACGCTTTATTTAAAAAAACAAGGAGAGAAAGGGGCACCCTCCCTTGGCCAAAGCCAGAGTAGGCTCTCTAAACCCA
CTTCTATAGGGAAGGAGAAGGAAATGACAACACTTGGAACCAACTCCATAATTTGTGTCCCATTACTGTGTGCTAATT
TATAGTGATTTCTTGCTCTGCCAAGAAAAGAAAAAAAAAAGAAAGCAGTCTGCTTAATTTATCTAGTATTGGCGGAGT
TGAGGTTCTTTTATATAAGCACCTTGATTAATCATACTCTTTTGGGGCCAAATAGTGCCTCTGATCCCCAAGCAAAAAC
AAAAACAAATACAACAACTTTCATCTGGGGTGGGAGAGGGGGCAGCGTCTCAAACCACTCTCCCTTGAATTTTCAAG
CCCTCTGAAACTTTAATTGCCTTCTCAAGAGCCAAGCTTTAACAAGGGTGGGGGAAGGCAGGGAGAAACATCAACTA
GCCTCCCGGTGGAGCGTGGAGGGTGGAGGGTGGAGGGTGGGGTGCGGGTGGGATGGGCAACATGTTTCCAGAAC
GGGACCTGGGAAAAGCCACGATCAGCGCCAACCCCCTAAGCTGCTCAGCTGTAACTCTGGAAAACAACCCCCGCCCC
CTCTAGCTCCCCAGAAGGAAAGGCTTTCGGTTTCCCAAGAGCGGGAGGCAGGCTAGCAAATCCTTGTCATGGGAAA
ATTATAGGGATCTAGCTAATAGCTAATCACCCACGGCTGAGAGTCAACTTAACCCCTGAGAATACTGACTAACCACTT
CCTAACTTCTGAAACCAAGCTATGCACACCCACATGCAAAGAAATCACTAGAATGTGTGGTTTAAAGTTAGACACTGG
ATTCTGGTTTATGTAAAATGGAATGTCAACGTGACAAGTGACTAAAGGAGTAAGATAAGAACAAAATTACAATCTAA
ATAATTTCTCAAAAGCGGACACCCACGTTAGTGGGAGGTAAGGGTGTAGCCATTCTGTCAAGGCTGATATTCTTAGC
CTGAAAGATAGTAAGAGGGAACTGATTACGAAATGGCAATATTCTCACCGTCTTCCCTTATTCTCTTCCAAATTCAGG
ATCACTAGCCCTCTTCAAGTTGCCTCATAAAGGCAAGGAGAGGCAGCTGAGGACCCAACCTCACCTAATTAAAGAGA
CCTCAATCCCCAAGCATAAAAAAGCAACCACAAGAGGCCATGGATGACTTCAAAAGAAAACAAATAAAAATACTGAC
CCTAAAAGAGGCACCAACCAAGTTATCCCAGTTTAATAATATATGATTAAGATTCTCATGAGAAATTGAAGTGACTTT
CCTCAGATGATGATTCATAATTAAGTTTGATATTATTGGCAGGCTTGCAATAAGCAAGCATGTATACCCAGGGAAAG
GGACTCAGTCTACTTGGGAACAACAAAAACCAGTAAATTTTCTCATGTGGGGGTGAAAAAAAAAATAACCTGTGTAA
TACATAATCCAAAAGACAACTTGCAGGTAACTTGCCATTCAACTATTTCTGCAGCAGGACACAACCTGAACTTCCTCA
GCGTGAAGGTTCTGTTTTGACTTTCTTTCAAGGTAGGGAGGGAAAACTGGGCTGGGAGACTATGCAAATGATGTGT
GGGAAAGTCGCACCTTAAGTAAATAAGGCCATCAGTGCGTGTCTGGGCAGGGAGAGGGAGGTGGGACTTCGAAAG
AGAAAGGGCTAAATTCTGGAAACAAAGCCTGTTGTTCTGGCTGGGAGGGCCCTCCCCTGTATAGCTGTATGTGTGTT
CTGATGACAGCAGCACAAATCACAGTCGGCCTC

JAG1in20:GFP (tested in zebrafish)

AAAGAAGAAACAGATCTATTGCAAAAACCTCCAATCACTGGCTCCCAAAGACCACCACCATCCTGAGTTCTAACAGTG
GAGAACCATTTTGGGAGGCAGCAATTTTTAATTTGATTCCTTTGACAGTGCAATTTACAGCAAGCCTGGCAAAGTTTC
CTGGAAGGGGTCAGACAGTAAACATCTTAGGCTTTGTGGGGTATATAGTCTCTGTCACAACTATCCAACTCTTGTGGT

ATCATGAAAGTAGCCATAGACAATACATAAATGAATGTGTGGCTGTGTGCCAAGAAAACTGTATTTATAAAAACAGA
CAGTGGGCTGGATACGGCCCCTGGGCCACAGTTTGCCAACTGTTTATTTAGAAACATGAAACAGATACAACCTGCCC
CCATCTAGTGCTGAAGACTAGTTAATCCCAGACTTGGATTTCAAGCACCAGTAAAATTTCAGACAATTTGGGTACAAG
AATCAAACCATAACTACTCTAGCACCTTTTTTGTTCTGTAAGAGAAAGGGCTTGTAAAAAAGTAGAGAAAAGGCTCA
AAATAATATGCTTTAAGCCTCTTGAAAAACTGCGTTGCCTGC

JAG1+46:GFP (tested in zebrafish)
CTAATTTTGATTTTTTACTTGCTTCAACGTGTTTCAGTTCATTAATGTTTCAATAATAAGGAATGAATTCTCTTGTCTTG
ACAAACCAAGCATGGAAATTTCCCCATAAGAGAAAAGAAGAAATTAAAATGATTTCCATATTCAAGTCACAGCACGC
TGAAAAGAGAACTTTGAGTTCTTAGGAAAGGCATCTGTTACCTCCAAAGGAGGGCCACTCATTACAGTTTTGTGTAAT
ATTACAGTAAGGTCTGCCAAGAAGGAATTCAAGGAAATTGGGAATGACGACACTGTGATCCCAGGCCACTCTCTTGA
GCAGAGGTTTCCTTCCCTTATGTGTTCTGCCCTCGCCAAGCTCTTTCTGAAATCAGAGGCCACTGACATGCATAACAA
ACCATGCTAAAATCCAAACATGACACCAACAGCAAACAATTACACTGTGATCCCATTGTGGTCCCCAAGATATATCTT
GAGATACTTCAGAGAGAAGCTATATTCTTCATAAATCAAAATACAGAGGGTTTTTATGATTCTTGCTTTTTATAAAATG
AAATACACTTCTATAGTATTAAAATAAGTGAACAAAAGAACTACAGC

2.4.2.5 Mutated hlx-3 zebrafish sequences (tested in zebrafish)

hlx-3 mut all EMSA binding ETS-b, d, e, h, i:Citrine
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATAGATTCTCTCTGTTTATTATAAGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGAGAGTGCTAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGAGAATAAGGGTGCTCTGCTGATCCTTAAAATACTAAAGCATCTTCTATTTTTGGCCTAAATCTTGCAGGAATAC
TCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTCTCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut 4 out of 5 EMSA binding ETS-b, d, e, h:Citrine
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATAGATTCTCTCTGTTTATTATAAGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGAGAGTGCTAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGAGAATAAGGGTGCTCTGCTGTATCCTAAAATACTAAAGCATCTTCTATTTTTGGCCTAAATCTTGCAGGAATAC
TCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut F3, 4, 6, 7, 8:Citrine (with wild-type ETS binding sites)
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATATGATTCCTCTGTTTATTATAAGCATAAGGACGGATAGTCCTCGTCGTGCGACTTG
CGAGTTAAGATCTATGAGAGTGCAAAAAATATAAAGGAATTGCGAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGGAAATGAGAGTGCTCTGCTTGATCCTTAATAACAAAAGCATTTATCTTCTTTGGCCTAAATCTTGCAGGAATAT
CCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut F2, 5, 9, 10:Citrine (with wild-type ETS binding sites)
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATAGATTTCCTCTGTAATTATATTGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGGAAGTGCAAGTTATCAATTACAAGTTCCACCTCTGTGAGTC
GGAGGAAATAAGGGTGCTCTGCTGTATCCTAAAATACTAAAGCATCTTCTATTTTTGGCCTATCATATGCAGGAATAT
CCGCTTACTATGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut F3, 4, 5, 9:Citrine
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATAGATTTCCTCTGTTTATTATAAGCATAAGGACGTATCGCGTGTCCGTGCAGACTTG
CGAGTTATAGATTCAGATAAGTCGAAGTAAAAAAAGGAAGTGGAACTTTAAACTCATAAGTTCCACCTCTGTGAGTC
GGAGGAAATAAGGGTGCTCTGCTGTATCCTAAAATACTAAAGCATCTTCTATTTTTGGCCTGATCGAATTAGCAATAT

CCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut F1, 7, 8:Citrine
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTTGTTATGTC
TATTAGTTGTTCAGCGGAGAATAGATTTCCTCTGTTTATTATAAGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGGAAGTGCTAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGGAAATAAGGGTGCTCTGCTAATACAATATCTTGCAAAGCATTCTTTTTGGCCTAAATCTTGCAGGAATATCCGC
TCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACAGAGC
AGAAAACACACACACATACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut F1:Citrine
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTGTTAATCTT
GTTGTTGTATTCAGCGGAGAATAGATTTCCTCTGTTTATTATAAGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGGAAGTGCTAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGGAAATAAGGGTGCTCTGCTGTATCCTAAAATACTAAAGCATCTTCTATTTTTGGCCTAAATCTTGCAGGAATAT
CCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACACGAGTCGGAGTGTTAGCATTTGAT

hlx-3 mut F7, 8:Citrine
CAGCACGCTCTATTTTTGGCACAGCAATGTAAAGTTGTCATGTGCATATGGATGAAGTATCGCCTCAGTAGTTAATTG
TGTTTTCGTTTCAGCGGAGAATAGATTTCCTCTGTTTATTATAAGCATAAGGACGGATTTGCGTCACCGTGCGACTTG
CGAGTTGAGATAAAGTTGCACAAATATTGAAAAAAGGAAGTGCTAACAGTCATTATAAAGTTCCACCTCTGTGAGTC
GGAGGAAATAAGGGTGCTCTGCTAATACAATATCTTGCAAAGCATTCTTCTTATTTGGCCTAAATCTTGCAGGAATAT
CCGCTCATTTAGCCTGAATCAGAGCTCATCAAAAACAGGATGACCTGCATCGAGAAACTCCTGTTTCCTAATTCCACA
GAGCAGAAAACACACACACATACACACACACACGAGTCGGAGTGTTAGCATTTGAT

2.4.2.6 Enhancer sequences identified by MEF2C ChIP-seq

(human sequences injected into zebrafish)

RASGRP-0.5:GFP
TATTTGAGGCCAAATTTTTAGTTTGTTTGTTGGTATTCATTTGAAAGAAGCACACATGCACTCATATACCTAAGGCTGC
TTTCTGAAATCAAATACATTTCCTGCCTTCTTCTTTCCTCAGTACTCTTACCACCAGCCTATTTATAGCCCAGGAAAGGA
AAAAGCTCCACTGGAGGCAGGAAAACTAAGTTCGGAAAGAATAGGAGGAAGTCTGCCTGCCATTTAGCCAAACAAC
CTTTAATTCATTGAGCTCAAAGCTTAGCAGCTTCAATGTTGTTGTTAGGCATAAACAGCTTGCAAATTTCAATTTTTTTT
CTCTCCTCACTTTTGAACTAGATTTGGGCAATGGAGGTGATGACAGCATTTTAGAAGACAGAACAGGAATTCTAGGT
GTTTAGGCACGTTGAAGAGAAACTTGCTAGAGGTGATGAAGTGG

ETS1+194:GFP
GAGCATTGATGAAGCAGCTCGATGAACTCTTGTAAGACCTTTCTCCAAGTGAGTTATCATTCCAAGGTTCCTGTTGTT
ACTTTCTGAAGGGCTTTGTGCTTATGAAGCAGATAGGGACAGAAGGGCGGGCCGGACCCCATGGTCATCTGGCAGC
CTTGCTTGTGTTGGGTCAGCGTGCCAGCATTCATGCCCCTGGGGGTGTGAGCACCCTTTCTGGAATCTTGAAAGGAA
ATGCAACCTCCCCAAACCCCACTACTCTCTCAAGGAGTCATGGGTGATGAAGATGTTTCCTGCCATTCTCTGCAGGAC
AGAGAGTTGCTTAACCCTCGTTTCATTATCAGTGGAGGAGCCAGCCATTGGACCCCCACCCCTCAGAGTGAGCCAGC
AGCCAGAACTAGCCCTTTGGCATCCCCCGTCTTGATTGCAAGCAAGCAAACCTGTCACCTTCCTGTTTTATCTACTCAT
GGTGTTTGTTTCCTTTGACGTGGATTCTTACACTTAGGGTGGGAGGGATTTCTCGAAAAACATTTGTTGTATGCCACT
GATCCTGGCCCCACAACGCCCCTTTCTCTTTCCTTTTCAAGGCTCCAGGAATTGCTAAATGGCATATCCTGAACACCCA
GCACTCTGGGGCTAAAAATACCAGAGGGTGTGTTTATCTAGATTGGGGCTTCCTCTCCAGGGGCTTTACAGACGTTA
TCTCATTTATCCTTGAAAACACTCTGCGAGGCAGCATGTAGTCAAATAACTTTACCCCAGAAAGGTACGGAAGGCTTG
AAGAGGAAAGGTCATCGGTGGAGGTTCTATTCTGAGGCCCAGTAGACAGGCTCAGGTCTTTAATTTCTCAGCTT

ELK3-29:GFP

TTTAAATGTGATTAATTTGCAATTCTGAAAGATGTCATTTGTGTCAATCCAGGAACGCATTCATTTTGGGGTGAGCGG
AAATGTCCCCAATTTCTGGTGGTAGCGGATACATCTGAAAAATGGAATCGAATCAATTTTGCTCCCGAGAGGGACAG
GTGGAGGACTCCTATTTCCTGCCAATTCAGTTGGAAACAATGAGGAGAGAGGAAAAAAAATGGAGCCCTCCCTTCTC
TTTCTGGCCTTAGACATGAACTCCTGATAAGATCCTGGAAAACAGGCAGCGTTTTTAACAGCCGGTTGCCCTCTTCAG
CTGAGAAAACAGCACGAGCCAAAACTAGGAACCGGGGCCACCTCTATTGTGAATGTGGCCAGGCTGACTTATTGTCG
CCAAACAGCAAATCATGAAATGACTGAGATAAGGAATATAATCACATTTTCACATGAACCTTACATTCTTACCTCTTCT
ATAGGAAAGGATATCCCTGCAGGAAGGAGGAAGGTGGGGCCTGCCCATGTTCTGCAGCAGGGCTGGGACTGCATT
GTCCAGGCTTTTGTAGGGACACAGGAAGAGGGTACGGAAGTTTCTACAGGAAAACAGAGAAAAGTTCAAACATAAA
CAGCATCCAAAAAAATTTACCAACAGCAAACTAGATTTCCTCTTTGGGCCTCCCATGCTTCCATGCACCCCTGGACTGG
CAGATTTTAATCTTGGATTCTGTTCACCTTGGTAACCTCCAGCGCGTGTGATTATGCTGTTTCCCTACTTATTAAGTTAA
AAGGATGTGTTTT

SH3TC1-10:GFP

CCTGTATAAAATGCTGTCTCTGCTGAACGCAGTGTTTCCCAGAAACACAAACAAGCAAAGATGGCCCCGCACACCTCC
CAAGGCATTATCAAGCCCTACTCAGACCCAAGAATGGGGGGAAGGACCTGCACAAATCTCTCCTGAGGTTCCTTCCT
GAGTTGATTCCTCTGCGCACCCAAGTTTGAGGACGGGTGTATCCGGGCCGGCATGATAAGCCGCTAGCCTGGAGCCC
CGTACCAGCTGGTGCTGGAATGGTAGCATGATGAGTCATCCCCCGGAAAGGGCTAGGACGGTGCCCTATTCAGCAG
CGCTCACACCATGTGTTATTTCTGGCCTTTGCTGCTGAAACCGGCCCCTGATCAGAAGAGCAGAGCTCTAGGCCCCAT
GATTTCAGGAACTGCTCTTTGCAAAACGAGGACACGGGGCTCATAGGCTTTCCAGATGACTGATGGCCACCCCAGCC
GGATGGGTGGCTGGGCCGCCGGGTCGGAGGCCTGCAGGAAGTGGTCAGCAGTTCTCGTGCCTGGGCTGGAACTGA
AGACACCCACCCCCGAACTGGCCGGAATGTCCTGATTTGTCGGGGCAGGCGTCAGCCTGGAGCCCAGGGAATCGCT
GACTCTGCGGCCCAGCAGGTCCACAGTCCTTCCTTCTCCCTGGGTCTGCCTGACCTGCTGACCAGGTGCCTGAGACTC
CTTAGCACAGTGCTCCTCATCTGGGGGTGGTGAGGGAAGCGAGCTGTTTCCCCAGAGGAGATTTCAGAAGAGGGAC
CTTCCTTCCTTGGAGGAAGCTGAC

2.4.2.7 Putative sprouting-specific enhancers identified by ATAC-seq

(zebrafish sequences tested in zebrafish)

ATAC-seq top hlx-3		**ATAC-seq top ISV vs nonISV**	
putative enhancer	danRer10 coordinates	putative enhancer	danRer10 coordinates
ggps1:Citrine	chr13:49,430,485-49,431,084	esm1:Citrine	chr21:9,893,841-9,894,440
igf2bp1:mCherry	chr3:23,417,322-23,417,921	vegfc:Citrine	chr1:38,504,522-38,505,121
ppp2r2d:Citrine	chr12:40,918,720-40,919,319	fgfr1a:mCherry	chr8:51,509,931-51,510,530
ephb2b:Citrine	chr11:28,058,682-28,059,281	mxtx2:Citrine	chr12:29,105,347-29,105,946
elf2b:mCherry	chr1:13,280,282-13,280,881	ptp4a3:mCherry	chr19:25,516,390-25,516,989
mych:mCherry	chr6:50,476,401-50,477,000	etfb:Citrine	chr16:50,235,759-50,236,358
sfrp1a:mCherry	chr5:26,911,879-26,912,478	rnpc3:mCherry	chr24:28,947,267-28,947,866
pin1:mCherry	chr1:58,281,520-58,282,119	siae:Citrine	chr18:36,545,855-36,546,454
psmb11b:Citrine	chr23:1,125,122-1,125,721	gpx7:mCherry	chr23:42,826,542-42,827,141
ifngr1:Citrine	chr1:53,126,681-53,127,280		
ATAC-seq hypothesis-driven approach			
esm1+7:Citrine	chr21:9,887,312-9,887,914	esm1+10:mCherry	chr21:9,884,205-9,884,807

2.4.3 Generating vectors for three-colour zebrafish for simultaneous analysis

In order to screen two putative enhancers simultaneously in transient transgenic fish, pE1b:mCherry, pE1b:Cerulean and pE1b:Citrine Tol2 vectors were generated. All three plasmids can be injected into the one-cell stage zebrafish for parallel screening of putative enhancers (Figure 6A). Transient transgenic zebrafish expressing all three fluorophores simultaneously can specifically be detected using the emission fingerprinting function of the Zeiss 710MP confocal microscope. I tested the new vectors by inserting the previously tested arterial and sprouting-specific Dll4in3 enhancer[167] upstream of the E1b minimal promoter (see Figure 6B). When screening putative enhancers identified with Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) or testing hlx-3 mutant sequences, pE1b-Dll4in3:Cerulean served as an internal control for successful injection while enhancer-pE1b-mCherry and enhancer-pE1b-Citrine were used to test putative enhancers.

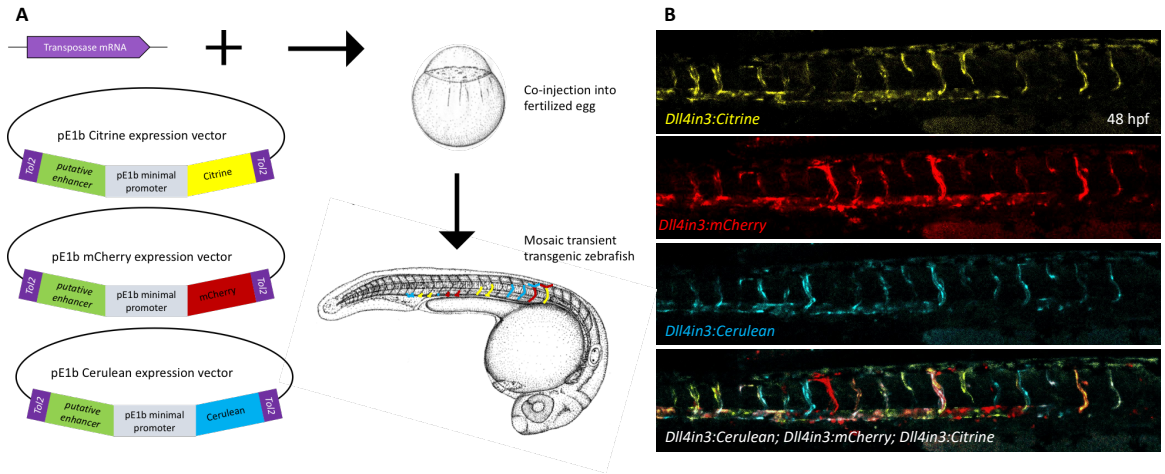


Figure 6: pE1b:mCherry/Cerulean/Citrine Tol2 vectors for the generation of transgenic zebrafish. A) *One-cell stage embryos were injected with Tol2 transposase and equal amounts of three pE1b Tol2 expression vectors: Dll4in3:Cerulean, Dll4in3:mCherry and Dll4in3:Citrine. Schematic illustration was adapted from[183] and zebrafish sketches were modified from[184].* ***B)*** *Tol2-mediated mosaic zebrafish transgenic for the Dll4in3 enhancer driving Cerulean, mCherry and Citrine expression in the dorsal aorta and ISVs at 48 hours post fertilization (hpf). Tol2 transposase-mediated integration events occur at random locations in the genome are independent for all three constructs. Cells show hence varying levels of mCherry, Cerulean and Citrine reporter gene expression.*

pE1b:mCherry/Cerulean/Citrine Tol2 vectors were cloned as followed: Sequences with mCherry, Citrine and Cerulean sequences containing ClaI and AgeI restriction enzyme sites were ordered as GeneArt strings. Sequences were cloned into pCR8 GW using TOPO TA cloning. mCherry, Citrine and Cerulean inserts and pE1b Tol2 GFP vector were cut with ClaI/AgeI. Vector was treated with Antarctic Phosphatase (Cat# M0289, NEB) and gel-purified with the QIAquick gel extraction kit (Cat# 28704, Qiagen). Vector and insert were ligated using T4 DNA ligase (Cat# M0202, NEB) at molar vector/insert ratio of 1:3. One shot ccdB resistant chemically competent E.coli (Cat# A10460, Invitrogen) were transformed. Colonies were picked from LB agar plates and cultured O/N in LB medium and Ampicillin. Midiprep (Cat#12145, Qiagen) was performed according to the manufacturer's instructions. Correct and in frame insertion of mCherry, Cerulean and Citrine into the pE1b Tol2 (and thereby replacing GFP) was validated by automated Sanger sequencing.

2.5 *In vitro* experiments

2.5.1 Cell culture

2.5.1.1 VEGF stimulation

Human umbilical vein endothelial cells (HUVECs, pooled donors, Lonza) were cultured for up to four passages in EBM-2 (supplemented with EGM-2 bulletkit containing hEGF, hydrocortisone, GA-1000, FBS, VEGF, hFGF-B, IGF-1, ascorbic acid, heparin) at 37°C in a humidified incubator with 5% CO_2. For experiments, cells were seeded at a density of $1x10^5$ cells per 3.5 cm-well in EGM-2 medium. HUVECs grew for 12 hours and were then quiesced with 0.1 % FBS for 24 hours before treatment with 25 ng/µl recombinant human $VEGFA_{165}$, Carrier Free (CF, R&D). Cells were harvested for protein and RNA extraction as described below.

2.5.1.2 Shear Stress

HUVECs were grown to confluency and quiesced in 0.1% FBS for 24 hours. Cells were put in an orbital shaker at 11.72 dynes/cm^2 with 60 revolutions per minute (rpm) which is within the range of arterial shear stress[185]. Non-stressed static controls were taken at each time point. RNA was extracted at the indicated time points and processed for quantitative Real Time PCR (qRT-PCR).

2.5.1.3 Protein extraction

Cells were lysed by the addition of RIPA buffer (50 mM Tris-HCl pH 7.4, 1% Nonidet P-40, 0.5% Na-deoxycholate, 0.1% sodium dodecyl sulfate (SDS), 150 mM NaCl, 2 mM EDTA, phosphatase inhibitor cocktail (PhosSTOP, Roche) and a protease inhibitor cocktail

(cOmplete, Roche)). Extracts were put on ice for 20 min and centrifuged at 12,300 times gravity (xg) at 4°C for 10 min. The supernatant was added to 4x Laemmli buffer [63 mM Tris-HCl pH 6.8, 10% glycerol, 1% SDS, 0.005% bromophenol blue, 355 mM β-mercaptoethanol (BME)] for analysis by SDS polyacrylamide gel electrophoresis (SDS-PAGE).

2.5.1.4 Lentiviral infection and *HLX* short hairpin RNA

2.5.1.4.1 Lentiviral production and HUVEC infection

Human embryonic kidney 293 T (HEK293T) cells were transfected with pCMV-d8.91 envelope plasmid, pCMV-VSV-G packaging plasmid (Addgene) and pGIPZ-GFP-Puromycin transfer plasmid. Two different HLX and one non-silencing control pGIPZ lentiviral microRNA-adapted short hairpin RNA (shRNAmir) clone libraries were kindly provided by Dr Alessandro Ceroni (Target Discovery Institute, University of Oxford). pGIPZ HLX shRNA antisense sequences were TCGTGAGCACAGCATAGGG and ACAGTAAGCACAAAACCCA. HEK293T cells were transfected using Fugene 6 (Cat# E2691, Promega) in Opti-MEM (Cat# 11058021, LifeTechnologies) according to the manufacturer's instruction. Lentiviral production and supernatant harvest were done with the help of Dr Leonid Nikitenko. Multiplicity of infection (MOI) was assessed by infecting HEK293T cells. HUVECs were infected with the filtered (0.45 µm), viral supernatant and cells were selected with 1 µg/ml puromycin (Sigma). For the *HLX* knockdown, efficiency of retroviral-mediated pGIPZ-GFP-HLX shRNA infection was confirmed by confocal microscopy and FACS (BD FACS Canto) (Figure 7A and B).

2.5.1.4.2 HLX knockdown to verify specificity of anti-HLX antibody

HUVECs were seeded at a density of $1x10^4$ cells per well in a 6-well plate. HUVECs were grown for 24 hours after which RNA and protein were extracted as described elsewhere. Knockdowns were confirmed with qRT-PCR and used to asses anti-HLX antibody specificity. Several antibodies were tested for HLX detection by Western Blot, but only Millipore anti-HLX (1:2000, Cat# 09-084) was able to detect a band at the correct size (50 kDa) and also showed a decrease in *HLX* shRNAmir samples (Figure 7C). Western Blot band intensity was measured using ImageJ and relative HLX protein expression levels were visualized (Figure 7D). RNA samples were processed for qRT-PCR using Taqman probes and *HLX* expression levels are shown in Figure 7D. RNA levels are decreased by around 50% which corresponds to the decrease observed by Western Blot. Millipore anti-HLX antibody will be used for further Western Blot experiments.

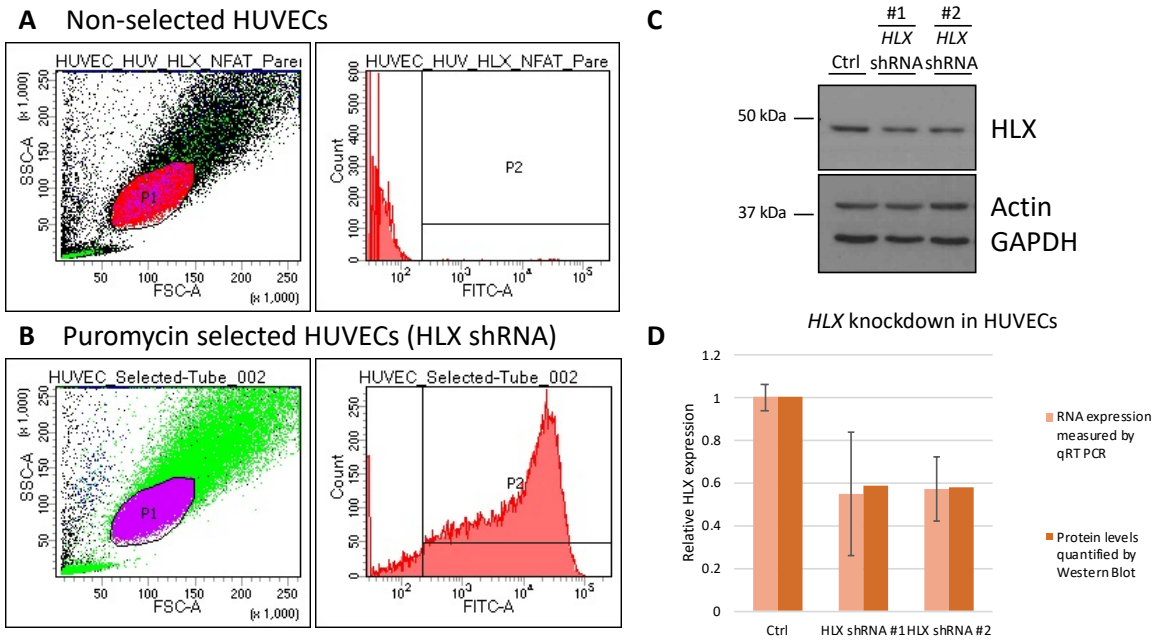


Figure 7: HLX knockdown in HUVECs confirms anti-HLX antibody specificity. A) *Fluorescence profile of non-selected HUVECs. P1 selects viable, single cells. Gating of P2 shows GFP positive cells (only 3%).* ***B)*** *Fluorescence profile of infected and puromycin selected HUVECs. 92.8% of cells are GFP and hence pGIPZ HLX shRNAmir positive. FACS plots are representative for technical triplicates of both pGIPZ HLX shRNAmir libraries.* ***C)*** *Western Blot showing HLX knockdown in HUVECs detected with anti-HLX Millipore antibody. Glyceralde-hyde-3-phosphate dehydrogenase (GAPDH) and ß-*

Actin show equal loading between samples. ***D)*** *Relative levels of HLX protein and RNA expression are shown for control (Ctrl) and knockdown samples (HLX shRNA #1 and #2). Relative protein levels were measured by quantifying band intensity of the HLX Western Blot. HLX RNA was normalized to β-Actin. The error bar shows the standard error for two biological replicates.*

The same protocol as described in section 2.1.3.3 was also used for HLX staining of E13 embryonic cryosections, where the following antibodies were used: anti-αSMA (1:100, Cat# A5228, Sigma) and anti-HLX antibody (1:4000, Cat# 09-084, Millipore).

2.5.2 Chromatin Immunoprecipitation

Chromatin Immunoprecipitation (ChIP) assays were performed under the supervision of and jointly with Dr Natalia Sacilotto. Sub-confluent HUVECs were starved for 18 hours in EBM-2 medium (Lonza) and then stimulated for one hour with EBM-2 supplemented with 25 ng/ml $VEGFA_{165}$ (Peprotech, Lonza). Cells were cross-linked with 1% formaldehyde at RT for 7 min, and the crosslinking reaction was quenched with 0.125 M glycine. Cells were collected in ice-cold PBS, lysed in cell lysis buffer (10 mM NaCl, 3 mM MgCl2, 30 mM sucrose, 10 mM EDTA, 0.5% Nonidet P-40, 10 mM Tris-HCl pH 7, supplemented with Roche protease inhibitor cocktail), and the resulting nuclei were lysed in nuclei lysis buffer (10 mM EDTA, 1% SDS, 50 mM Tris-HCl pH 8.1). Chromatin was sonicated using the Covaris S220 sonicator (8.5 min, 160 W peak incident power, 5% duty cycle and 200 cycles per burst, at 4°C) to obtain populations of fragments with an average size of 250 bp. The size of the resulting chromatin fragments was analysed after each independent sonication as follows: chromatin formaldehyde crosslinks were reversed at 65°C in the presence of proteinase K (Fisher Scientific) O/N, and the DNA purified using the PCR purification kit (Qiagen) following the manufacturer's instructions. The resulting purified DNA fragments were analysed by agarose gel electrophoresis.

Ten-fold diluted chromatin was pre-cleared with Bovine serum albumin (BSA, Sigma) and tRNA (Ambion) saturated Dynabeads protein-G/protein A (Life Technologies) for two hours at 4°C. Pre-cleared chromatin was then incubated O/N with 3 µg of anti-MEF2A (Cat# ab109420, abcam) or anti-MEF2C (Cat# 5030S, Cell Signalling) or anti-MEF2D (Cat# 610774, BD) antibodies and Isotype IgG controls, in low-binding tubes (Thermo Fisher Scientific), and 1/10th of the diluted chromatin used in the IP was kept as an input sample, for each condition. Saturated Dynabeads protein-G/protein A (Life Technologies) were added to the samples and incubated for two hours at 4°C. Immuno-complexes were recovered with a DynaMag-2 magnet (Thermo Fisher Scientific), extensively washed twice with low salt buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), twice with high salt buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 2mM EDTA, 1% Triton X-100, 0.1% SDS), twice in LiCl buffer (10 mM Tris-HCl pH 8, 250 mM LiCl, 1 mM EDTA, 1% sodium deoxycholate, 1% Nonidet P-40), once with TE (pH 8) and eluted twice with 100 µl of elution buffer (10 mM EDTA, 1% SDS, 50 mM Tris-HCl) at 65°C for 10 min. The resulting 200 µl were incubated at 65°C for one hour with RNase A (Sigma) followed by incubation with proteinase K (Fisher Scientific) at 65°C O/N. DNA was purified by PCR purification kit (Qiagen) following the manufacturer's instructions, and used as a template for qPCR with the following primers[170].

DLL4in3 ChIP F3 GTCGGACGCTCGGATTC	R3 GAGCACGGCGGTGAGAAAG
HLX-3 ChIP F2 GGGAAGTGCTAGGAGTCATTATAGAG	R2 CGGGACTAAATGAGCGGATGTG

2.5.3 Western Blotting

2.5.3.1 Western Blot on HUVECs

The high salt lysis fraction was separated by a 10% SDS-PAGE, 1/4th of a 6-well was loaded and run at 150 V. Protein concentration was measured using the Bradford assay (Cat# 500-0205, BioRad) and adjusted if required. Proteins were transferred to polyvinylidene difluoride (PVDF, Hybond-P, Amersham) membranes by Western blotting in Towbin buffer[186] (25 mM Tris, 192 mM glycine, 20% methanol) at 200 mA O/N. PVDF membranes were activated in 100% methanol. Following transfer, PVDF membranes were stained in Ponceau-S solution (0.5% Ponceau S, 1% acetic acid). Wet membranes were incubated in blocking buffer for one hour [TBS containing 0.1% Tween-20 (TBST), 5% Marvel skimmed milk powder for most antibodies or 5% BSA for all anti-phospho-antibodies] and then incubated O/N at 4°C with the primary antibody, diluted in blocking buffer. Following three 10 min washes in TBST, membranes were incubated with the secondary antibody. Blots were washed again and incubated with enhanced chemiluminescence (ECL) substrate Clarity (Cat# 1705060, BioRad) for 1 min. Membranes were wrapped in Saran wrap and exposed to Fuji X-ray films. The following antibodies were used: HLX (1:2000, Cat# 09-084, Millipore), VEGF Receptor 2 (D5B1, Cat# 9698), phospho-VEGFR2-Tyr1059 (D5A6, Cat# 3817), phospho-VEGFR2-Tyr996 (Cat# 2474) (all 1:1000, Cell Signalling), Extracellular-signal-regulated kinase (ERK, 1:3000, E6, Cat# sc-271270, Santa Cruz), phospho-ERK-Thr202/Tyr204 (1:2000, 197G2, Cat# 4377, Cell Signalling), phospho-Akt-Ser473 (1:1000, 193H12, Cat# 4058, Cell Signalling), Akt (1:1000, A9, Cat# 9272, Cell Signalling). As all primary antibodies were raised in rabbit, the same secondary polyclonal goat anti-rabbit antibody coupled to horseradish-peroxidase (HRP) was used for detection (1:3000, Cat# P0448, Dako).

2.5.3.2 Western Blot on hindbrains

Mef2A/C^{iECKO} E12.5 embryos and control littermates were kindly provided by Dr Yao Wei Lu and Dr John Schwarz (Center for Cardiovascular Sciences, Albany Medical College). Hindbrains were dissected as described in section 2.2.6. *Mef2A/C*^{iECKO} and control hindbrains were lysed in Tris buffer (20 mM Tris-HCl pH 9, 2% SDS, supplemented with protease and phosphatase inhibitor cocktail, Roche), boiled at 100°C for 20 min followed by a two hour incubation at 80°C at 750 rpm[187]. Insoluble material was pelleted at 14,000 rpm for 20 min. Protein concentration was determined with the Pierce BCA Protein Assay kit (Cat# 23252, Thermo Scientific) and 10 µg of protein per lane was separated by 10% SDS-PAGE run for two hours at 100 V. Proteins were transferred in Towbin buffer for four hours at 60 V unto PVDF membranes (0.45 PVDF Hybond-P, Amersham). After one hour blocking in 5% Marvel in TBST, DLL4 was detected using anti-DLL4 (1:1000, Cat# 7280, abcam) for an O/N incubation at 4°C in 5% Marvel/TBST. Secondary anti-rabbit HRP (1:3000,Cat# P0448, Dako) was incubated for two hours at RT. Blots were developed with ECL and visualized on a ChemiDoc imaging system (XRS+, BioRad). Membranes were stripped and re-probed with anti-β-Actin antibody (1:100000, Clone AC-15, abcam) and detected with anti-mouse HRP (1:2000, Cat# P0447, Dako). Band intensities of DLL4 and β-Actin Western blots were quantified with the Image Lab™ Software (BioRad). DLL4 intensity was adjusted for levels of β-Actin and blotted relative to band intensity of the control littermate.

2.5.4 Quantitative RT-PCR

HUVECs were washed in PBS and RNA was extracted using the illustra RNAspin kit (Cat# 25050071, GE Healthcare) following the manufacturer's instructions.

Whole lungs were dissected from *Mef2A/C*^{iECKO} mice at post-natal day P5 and P6, rinsed in PBS, and then snap-frozen in liquid nitrogen for subsequent RNA extraction. Lungs were homogenized in liquid nitrogen using a pestle and mortar and lysed in R1 buffer of the RNAspin kit. RNA was extracted using the illustra RNAspin kit (Cat# 25050071, GE Healthcare). Relative gene expression levels of *Hlx*, *Dll4*, *Mef2A*, *Mef2C*, and *Mef2D* were analysed by qRT-PCR in P5 and P6 lungs taken from control (Cre-negative) and *Mef2A/C*^{iECKO} pups.

RNA from either HUVECs, retinas or FAC sorted zebrafish (see section 2.7) was used for qRT-PCR. Briefly, RNA concentrations were measured using a Nanodrop spectrophotometer and adjusted, followed by reverse transcription with SuperScript III (Cat# 18080093, Invitrogen). Complementary DNA (cDNA) was run on an Applied Biosystems 7500 real-time thermal cycler. Either pre-designed Taqman assays (Applied Biosystems) and TaqMan Universal PCR Master Mix (Cat# 4304437, LifeTechnologies) or primers and SYBR Green Mastermix (Cat# 4385612, LifeTechnologies) were used. Expression levels were normalized to an internal control: β-Actin (*ACTB*) or VE-Cadherin (*CDH5*) for human and mouse samples, the geomean of elongation factor 1α (*ef1α*) and ubiquitin C (*ubiC*) for zebrafish samples. Table 2 shows the used Taqman probes and PCR primers.

Table 2: Taqman probes and primers used for qRT-PCR. *For PCRs on RNA derived from homo sapiens and mus musculus, β-Actin (ActB) and VE-Cadherin (Cdh5) were used as internal controls. For PCRs on RNA derived from sorted danio rerio, ubiC and ef1a[188] were used as internal controls. Danio rerio primers used for ephb2a[189], dll4[190] and kdrl[191] detection were previously described. Dab2, ephb4a, flt4, tbx20 and ubiC were kindly provided by Dr Rui Monteiro and Dr Florian Bonkhofer.*

Homo sapiens	**Taqman probe**
DLL4 Taqman (FAM)	Hs00184092_m1
MEF2A Taqman (FAM)	Hs01050409_m1
MEF2C Taqman (FAM)	Hs00231149_m1
MEF2D Taqman (FAM)	Hs00954735_m1
HLX Taqman (FAM)	Hs00172035_m1
KLF2 Taqman (FAM)	Hs00360439_g1
ACTB Taqman (FAM)	Hs01060665_m1
CDH5 Taqman (FAM)	Hs00901465_m1
Mus musculus	**Taqman probe**
Hlx Taqman (FAM)	Mm00468656_m1
Dll4 Taqman (FAM)	Mm00444619_m1
Mef2D Taqman (FAM)	Mm00504931_m1
ActB Taqman FAM)	Mm02619580_g1
Cdh5 Taqman (FAM)	Mm00486938_m1

qRT PCR using primers and SYBR Green		
Mus musculus	**Forward primer**	**Reverse primer**
Mef2A Primer	AGACAAGGTGACTGAAAATGGG	CAGAGCACACTGAGTTCATAGG
Mef2C Primer	CAGAGCACACTGAGTTCATAGG	ATTACGAGGATAATGGATGAGCG
Danio rerio	**Forward primer**	**Reverse primer**
kdrl Primer	CTCCTGTACAGCAAGGAATG	ATCTTTGGGCACCTTATAGC
dab2 Primer	GGCTGAGCCTCTAGTCATTGAT	GCCATTTCCTTCACCCTTTT
ephb4a Primer	CCTGATGAACACGAAAACGGA	TGATAGGTCCGCACACTGTT
flt4 Primer	ACAGAGGAGCCATGTTGACA	GTCTGGCCTGAGAGTTGAGT
ephb2a Primer	CCCATTTCCCCCAAAGACTA	CTTCCCCATGAGGAGATGC
tbx20 Primer	CCAAGATTTCCTGCAGCCTC	ACTCCGGAAAATGACACACG
dll4 Primer	GAATACTGCGAAGAACCGATCTG	GTGTTTACAGGCTGGATATGTTTTACA
ubiC Primer	AAGAGACTCCCATACACCGC	ATTCTCAATGGTGTCGCTGG
ef1a Primer	GAGAAGTTCGAGAAGGAAGC	CGTAGTATTTGCTGGTCTCG

2.6 Electrophoretic mobility shift assay

2.6.1 Plasmids for generation of recombinant proteins

To make recombinant proteins for the Electrophoretic mobility shift assay (EMSA), the following plasmids were used to *in vitro* transcribe and translate protein with the TNT Quick Coupled Transcription/Translation System (Cat# L1170, Promega): pCITE2:ETS1-DNA-binding-domain (ETS1DBD), pCITE2:ELF1, pCITE2:ERG, pCS2:ETV2 and pcDNA3:MEF2A, pcDNA3:MEF2C, pcDNA3:MEF2D. ETS1DBD and full-length ERG, ELF1 and MEF2A/C/D were transcribed using T7 polymerase. ETV2 was transcribed using Sp6 polymerase according to the manufacturer's instructions.

2.6.2 Probe labelling and gel shifts

Oligos (see Table 3) were annealed to generate double stranded DNA fragments. Probes were labelled with α^{32}P-dCTP [deoxycytidine triphosphate, labelled with ^{32}P (Easy Tides, Perkin Elmer] using Klenow (#Cat 9PIM220, Promega) to fill in overhanging 5' ends and purified on a non-denaturing polyacrylamide-Tris/Borate/EDTA (TBE) gel as previously described[170,152]. Probe concentration was measured (2450 MicroBeta2 LumiJet, PerkinElmer) and diluted to 20 - 40,000 counts per minute/µl (in 40 mM KCl).

For the gel shift, recombinant protein and probe were bound in a 20 µl mix consisting of 2 µl 10x binding buffer [40 mM KCl, 15 mM 4-2-hydroxyethyl-1-piperazine-ethanesulfonic acid (HEPES) pH 7.9, 1 mM EDTA, 0.5 mM dithiothreitol (DTT), 5% glycerol), 0.2 - 2µl polydeoxyinosinicdeoxycytidylic acid (dI-dC, Cat# 10108812001, Roche)] and 100 fold excess of competitor DNA, if indicated. MEF2 EMSAs were run on a 6%, while ETS family

EMSAs were run on a 10% non-denaturing polyacrylamide gel at 4°C for two hours in the HoeferTM SE400 Air-cool vertical electrophoresis system. Gels were dried and visualized by autoradiography.

Table 3: Oligo sequences used for EMSA. *As a positive control for ETS binding, the previously identified Mef2C F10 - 44 ETS-A enhancer sequence was used[18], while the Myosin Light Chain (MLC) promoter sequence containing the MEF2 site was used as a Mef2C control[19]. WT = wild-type, MT = mutant.*

Mef2 control (MLC)	ctagGATCCTCATCTTTTAAAAATAACTTTTCAAAAG	ctagCTTTTGAAAAGTTATTTTTAAAAGATGAGGATC
HLX-3 MEF2-a WT	ctagGTATATTTA**TATTTTCA**TTGCTAG	ctagCTAGCAA**TGAAAATA**TAAATATAC
HLX-3 MEF2-a MT	ctagGTATATTTA**CGCCCTCA**TTGCTAG	ctagCTAGCAA**TGAGGGCG**TAAATATAC
HLX-3 MEF2-b WT	ctagTGCAGTATCC**TAAAATA**CTAAGGCCGC	ctagGCGGCCTTAG**TATTTTA**GGATACTGCA
HLX-3 MEF2-b MT	ctagTGCAGTATCC**TAGGGCG**CTAAGGCCGC	ctagGCGGCCTTAG**CGCCCTA**GGATACTGCA
HLX-3 MEF2-c WT	ctagGGCCGCTTC**TATTTTGA**GACCAATCT	ctagAGATTGGTC**TCAAAATA**GAAGCGGCC
HLX-3 MEF2-c MT	ctagGGCCGCTTC**CGCCCTGA**GACCAATCT	ctagAGATTGGTC**TCAGGGCG**GAAGCGGCC
hlx-3 MEF2-a WT	ctagAGTTAATTG**TGTTTTCG**TTTCAGCGG	ctagCCGCTGAAA**CGAAAACA**CAATTAACT
hlx-3 MEF2-a MT	ctagTGTTATGTCTATTAGTTGTTCAGCGG	ctagCCGCTGAACAACTAATAGACATAACA
hlx-3 MEF2-b WT	ctagCTGTATCC**TAAAATA**CTAAAGCAT	ctagATGCTTTAG**TATTTTA**GGATACAG
hlx-3 MEF2-b MT	ctagCTAATACAATATCTTGCAAAGCAT	ctagATGCTTTGCAAGATATTGTATTAG
hlx-3 MEF2-c WT	ctagAAAGCATCTTC**TATTTTTG**GCCTAAA	ctagTTTAGGC**CAAAAATA**GAAGATGCTTT
hlx-3 MEF2-c MT	ctagAAAGCATTCTT**CTTATTTG**GCCTAAA	ctagTTTAGGC**CAAATAAG**AAGAATGCTTT
ETS control (Mef2)	ctagGGAGTTACTCTCTTCCTGTTATG	ctagTGTCATAACAGGAAGAGAGTAA
HLX-3 ETS-a WT	ctagATTGCTAGAAG**GGAA**TTGATTCGCGT	ctagACGCGAATCAA**TTCC**CTTCTAGCAAT
HLX-3 ETS-a MT	ctagATTGCTAGAAG**GCGA**TTGATTCGCGT	ctagACGCGAATCAA**TCGC**CTTCTAGCAAT
HLX-3 ETS-b WT	ctagAGGCGATTGAT**TTCC**GTGCATTTAT	ctagATAAATGCAC**GGAA**ATCAATCGCCT
HLX-3 ETS-b MT	ctagAGGCGATTGAT**TCGC**GTGCATTTAT	ctagATAAATGCAC**GCGA**ATCAATCGCCT
HLX-3 ETS-c WT	ctagACATAAGGGC**GGAT**TTGCGTCACC	ctagGGTGACGCAA**ATCC**GCCCTTATGT
HLX-3 ETS-c MT	ctagACATAAGGGC**GCGT**TTGCGTCACC	ctagGGTGACGCAA**ACGC**GCCCTTATGT
HLX-3 ETS-d WT	ctagTATTGAAAGG**GGAA**GTGCTAGGAG	ctagCTCCTAGCAC**TTCC**CCTTTCAATA
HLX-3 ETS-d MT	ctagTATTGAAAGG**GCGA**GTGCTAGGAG	ctagCTCCTAGCAC**TCGC**CCTTTCAATA
HLX-3 ETS-e WT	ctagGTTTTTCTCC**GGAA**GAAATAAGG	ctagCCTTATTTC**TTCC**GGAGAAAAAC
HLX-3 ETS-e MT	ctagGTTTTTCTCC**GCGA**GAAATAAGG	ctagCCTTATTTC**TCGC**GGAGAAAAAC
HLX-3 ETS-f WT	ctagGAAATAA**GGAT**TTCTGCAGT	ctagACTGCAGAA**ATCC**TTATTTC
HLX-3 ETS-f MT	ctagGAAATAA**GCGT**TTCTGCAGT	ctagACTGCAGAA**ACGC**TTATTTC
HLX-3 ETS-g WT	ctagTTCTGCAGT**ATCC**TAAAATACTAAG	ctagCTTAGTATTTTA**GGAT**ACTGCAGAA
HLX-3 ETS-g MT	ctagTTCTGCAGT**ACGC**TAAAATACTAAG	ctagCTTAGTATTTTA**GCGT**ACTGCAGAA
HLX-3 ETS-h WT	ctagTCGCAGGCAC**ATCC**GCTCATTTAG	ctagCTAAATGAGC**GGAT**GTGCCTGCGA
HLX-3 ETS-h MT	ctagTCGCAGGCAC**ACGC**GCTCATTTAG	ctagCTAAATGAGC**GCGT**GTGCCTGCGA
HLX-3 ETS-i WT	ctagAGCCCAGGGT**TTCC**TGCTGCTTTC	ctagGAAAGCAGCA**GGAA**ACCCTGGGCT
HLX-3 ETS-i MT	ctagAGCCCAGGGT**TCGC**TGCTGCTTTC	ctagGAAAGCAGCA**GCGA**ACCCTGGGCT
hlx-3 ETS-b WT	ctagGAGAATAGAT**TTCC**TCTGTTTATT	ctagAATAAACAGA**GGAA**ATCTATTCTC
hlx-3 ETS-b MT	ctagGAGAATAGAT**TCTC**TCTGTTTATT	ctagAATAAACA**GAGA**GAATCTATTCT
hlx-3 ETS-c WT	ctagGCATAAGGAC**GGAT**TTGCGTCACC	ctagGGTGACGCAA**ATCC**GTCCTTATGC
hlx-3 ETS-c MT	ctagGCATAAGGAC**GATG**TTGCGTCACC	ctagGGTGACGCA**ACAT**CGTCCTTATGC
hlx-3 ETS-d WT	ctagATTGAAAAAA**GGAA**GTGCTAACAG	ctagCTGTTAGCAC**TTCC**TTTTTTCAAT
hlx-3 ETS-d MT	ctagATTGAAAAAA**GAGA**GTGCTAACAG	ctagCTGTTAGCAC**TCTC**TTTTTTCAAT
hlx-3 ETS-e WT	ctagGTGAGTCGGA**GGAA**ATAAGGGTGC	ctagGCACCCTTAT**TTCC**TCCGACTCAC
hlx-3 ETS-e MT	ctagGTGAGTCGGA**GAGA**ATAAGGGTGC	ctagGCACCCTTAT**TCTC**TCCGACTCAC
hlx-3 ETS-g WT	ctagGCTCTGCTGT**ATCC**TAAAATACTA	ctagTAGTATTTTA**GGAT**ACAGCAGAGC
hlx-3 ETS-g MT	ctagGCTCTGCTGT**ACTC**TAAAATACTA	ctagTAGTATTTTA**GAGT**ACAGCAGAGC
hlx-3 ETS-g.2 WT	ctagAAATCTTGCA**GGAA**TACTCGCTCA	ctagTGAGCGAGTA**TTCC**TGCAAGATTT
hlx-3 ETS-g.2 MT	ctagAAATCTTGCA**GAGA**TACTCGCTCA	ctagTGAGCGAGTA**TCTC**TGCAAGATTT
hlx-3 ETS-h WT	ctagTTGCAGAGAT**ATCC**GCTCATTTAG	ctagCTAAATGAGC**GGAT**ATCTCTGCAA
hlx-3 ETS-h MT	ctagTTGCAGAGAT**ACTC**GCTCATTTAG	ctagCTAAATGAGC**GAGT**ATCTCTGCAA
hlx-3 ETS-i WT	ctagAAACTCCTGT**TTCC**TAATCTCACA	ctagTGTGAGATTA**GGAA**ACAGGAGTTT
hlx-3 ETS-i MT	ctagAAACTCCTGT**TCTC**TAATCTCACA	ctagTGTGAGATTA**GAGA**ACAGGAGTTT
hlx-3 ETS-i.2 WT	CTAGTGTTCTCTAA**TTCC**ACAGAGCAGA	ctagTCTGCTCTGT**GGAA**TTAGAGAACA
hlx-3 ETS-i.2 MT	CTAGTGTTCTCTAA**TCTC**ACAGAGCAGA	ctagTCTGCTCTGT**GAGA**TTAGAGAACA

2.7 RNA and ATAC sequencing on sorted cells from transgenic zebrafish

2.7.1 Fluorescence-activated cell sorting

For both ATAC-seq and RNA sequencing (RNA-seq) experiments, embryos from transgenic zebrafish lines *tg(Ephb4-2:GFP)*, *tg(hlx-3:GFP)*, *tg(Dll4in3:GFP;kdrl:HRAS-mCherry)* were allowed to develop to the desired stage (27 hpf) at 28.5°C and selected by expression pattern. A later time point (43 hpf) was also concurrently collected but is not the focus of this thesis. Figure 8 illustrates the whole protocol from Fluorescence-activated cell sorting (FACS) to RNA-seq and ATAC-seq library preparation.

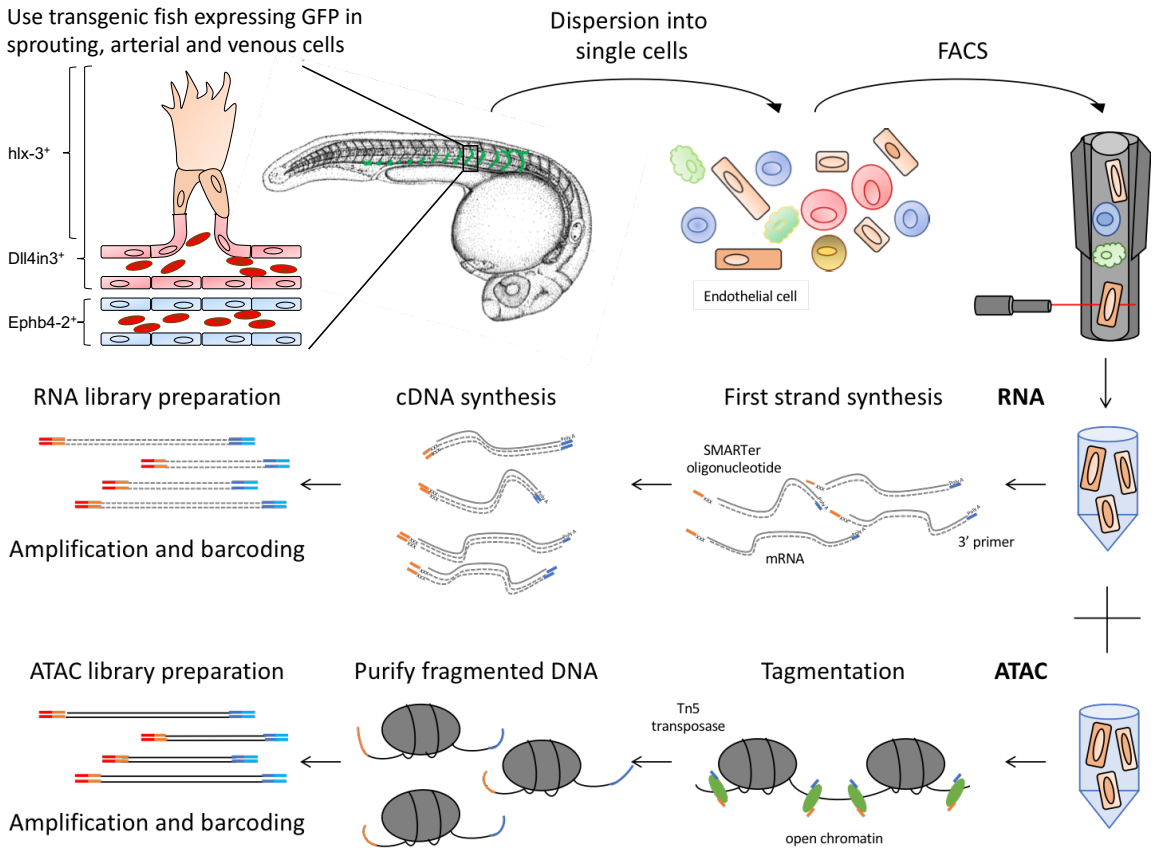
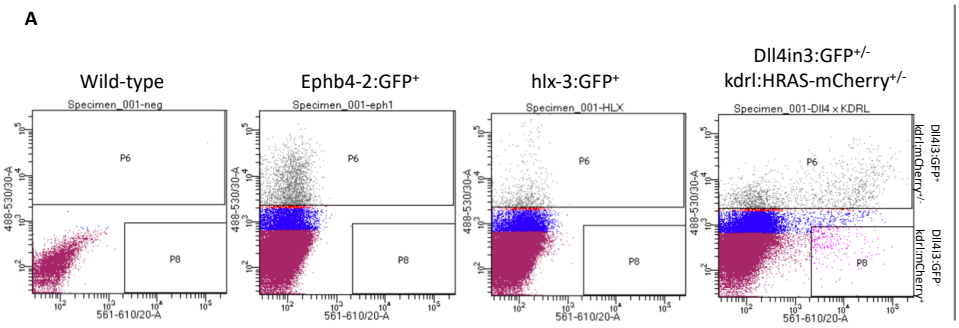


Figure 8: Schematic illustration of FACS, ATAC-seq and RNA-seq library preparation.

To prepare zebrafish embryos for FACS, each pool of 80 - 120 zebrafish was manually dechorionated and deyolked in 500 µl of Ringer solution (116 mM NaCl, 2.9 mM KCl, 5 mM HEPES and 1 mM EDTA). The fish embryos were pelleted at 500 ×g in a tabletop centrifuge and supernatant was removed. Fish were dissociated in 300 µl of Hank's Balanced Salt Solution (HBSS, Cat# Gibco) supplemented 0.05% Trypsin (with EDTA, Gibco) and 2% collagenase (Sigma) and left shaking at 300 rpm for 15 min at 30°C with gentle pipetting every 2 min to disrupt embryos. The solution was then pipetted to make a homogeneous solution and the reaction was stopped by adding 50 µl FBS and 650 µl Hank's solution (1x HBSS, 10 mM HEPES (Gibco), 0.25% BSA (Sigma)). The cells were pelleted at 500 ×g for 5 min at 4°C, washed with cold Hank's solution and passed through a 40 µm filter to obtain a single cell solution. Cells were spun at 750 ×g for 5 min at 4°C. Pellet was resuspended in 400 µl Hank's solution with 0.01% Hoechst and transferred to 5 ml polystyrene round bottom tube for FACS. Cells were sorted on a FACS Aria Fusion (BD Bioscience) at the WIMM Flow Cytometry Facility with the help of Kevin Clark and Dr Sally-Ann Clark. Wild-type, non-fluorescent fish were used as a control to set the gates. By gating high fluorescence, we obtained enriched populations of Dll4in3:GFP^{+}, Dll4in3:GFP^{-} x kdrl:HRAS-mCherry^{+}, Ephb4-2:GFP^{+} and hlx-3:GFP^{+} (see Figure 9).



Sorted sub-endothelial cell type	Percentage of collected cells
hlx-3:GFP+	0.80%
Ephb4-2:GFP+	2.12%
Dll4in3:GFP+	1.00%
Dll4in3:GFP- *x* kdrl:HRAS-mCherry+	0.32%

Figure 9: Fluorescence-activated cell sorting on zebrafish wild-type, tg(Ephb4-2:GFP), tg(hlx-3:GFP) and tg(Dll4in3:GFP;kdrl:HRAS-mCherry) lines. A) *Diagnostic fluorescence profile prior to sorting of cells from tg(Ephb4-2:GFP), tg(hlx-3:GFP), tg(Dll4in3:GFP;kdrl:HRAS-mCherry) or non-transgenic wild-type embryos prior to sorting.*

Representative for all three replicates at 27 hpf. Y-axis is GFP, X-axis is mCherry fluorescence. ***B)*** *Average of collected sprouting (hlx-3:GFP^{+}), venous (Ephb4-2:GFP^{+}), arterial (Dll4in3:GFP^{+}) and non-arterial (Dll4in3:GFP^{-} x kdrl:HRAS-mCherry^{+}) endothelial cells constitute 0.32 - 2.12% of all embryonic cells in zebrafish at 27 hpf.*

Dependent on the sample, 10,000 to 50,000 sorted cells were collected into safeseal low-binding eppendorf tubes (Sorenson) and maintained on ice until further processing.

2.7.2 Library preparation

2.7.2.1 Assay for transposase accessible chromatin

For the Assay for transposase accessible chromatin sequencing (ATAC-seq), cells were sorted directly into cold PBS. Cells were spun down at 500 xg for 5 min at 4°C. Dissociated cells were lysed [10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, and 0.1% IGEPAL (Sigma CA-630)] and spun down immediately at 500 xg at 4°C. Supernatant was removed and the pellet was resuspended and incubated for 30 min at 37°C with the Tn5 enzyme using the Illumina Nextera kit (Cat# FC-121-1030)[192]. Amount of Tn5 transposase was calculated based on the number of cells (0.03 µl per 1,000 cells) with a total volume of 50 µl per reaction. DNA was purified using the PCR purification MinElute kit (Cat# 28006, Qiagen). To achieve optimal DNA binding to the column, 25 µl 3 M sodium acetate (pH 5) was added per 1 ml of binding buffer and to achieve a greater yield, DNA was eluted twice in 10 µl elution buffer. Tagmented DNA was amplified using 25 µM universal Nextera primer Ad1_noMX, 25 µM sample-specific indexing primer (Ad2.1-13) and the NEBNext High-Fidelity 2x PCR Master Mix (Cat# M0541) for five cycles (see Figure 10A, PCR cycle). The amount of primer was adjusted to the number of isolated cells (0.05 µl of 25 µM universal/barcoding primer per 1,000 cells) with a total volume of 50 µl per reaction. After fives cycles of initial PCR, I calculated the required number of remaining cycles for each sample individually by qPCR to stop amplification prior to saturation (5 µl PCR amplified

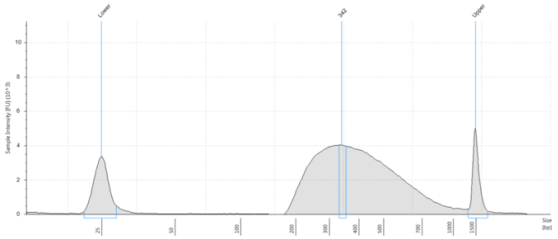
DNA, 5 µl NEBNext High-Fidelity 2x Master Mix, 0.25 µl 25 µM universal/barcoding primer, 0.06 µl 100x SYBR Green (Cat# S-7563, Invitrogen) and 4.4 µl H_2O). The required number of remaining cycles was determined by reading the cycle number on the multicomponent plot [linear fluorescence *versus* (vs.) cycle] which corresponds to 1/4th of the maximum SYBR fluorescent intensity. PCR cycling conditions for the initial, remaining PCR and qPCR can be found in Figure 10A. The amplified library was purified using the PCR purification MinElute kit (Qiagen) as described previously. Excess primers were removed using AMPpure beads (Cat# A63880, Beckman Coulter). Therefore, 20 µl of resuspended beads were added to the tagmented library and incubated at RT for 5 min. Slurry was spun down and left in magnetic rack to clear for 5 min. Supernatant was removed and beads with DNA were washed in 200 µl of 80% ethanol. Beads were dried and library DNA was eluted in 20 µl 0.1x TE Buffer (1 mM Tris-HCl and 0.1 mM EDTA, both pH 8, passed through a 0.22 µM filter). As a control, genomic DNA from endothelial cells (kdrl:HRAS-mCherry^{+}) was extracted using the Blood and Tissue kit (Cat# 69504, Qiagen) and eluted in elution buffer. A total of 92 ng of naked genomic DNA was then tagmented and processed as described. Tagmentation efficiency of all samples was assessed using D1000 or D5000 TapeStation (buffer Cat# 5067-5582 and ladder Cat# 5583, Agilent). Nucleosomal profiles representative for all triplicates can be found in Figure 10C and TapeStation profile for the genomic input control can be found in Figure 10B. Tagmented DNA and subsequently pooled libraries were quantified with the KAPA Library Quantification Kit (Cat# KK4835). The concentration of ATAC libraries was adjusted for the average fragment length of each sample. Multiplexed library pool [13 nanomolar (nM) total concentration] was sequenced

using paired-end 75 bp runs on the Illumina® HiSeq4000 platform by the WTCHG Oxford Genomics Centre.

A

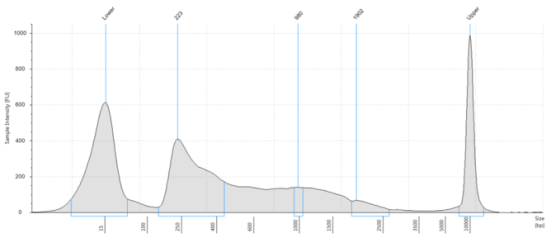
	PCR cycle	PCR continued	qPCR cycle
1	72°C, 5min	98°C, 30sec	98°C, 30sec
2	98°C, 30sec	98°C, 10sec	98°C, 10sec
3	98°C, 10sec	63°C, 30sec	63°C, 30sec
4	63°C, 30sec	72°C, 1min	72°C, 1min
5	72°C, 1min		
	Repeat steps 3-5, 5 times	Repeat steps 2-4, x times	Repeat steps 2-4, 40 times

B Genomic input

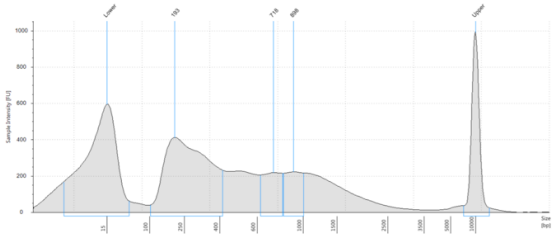


C

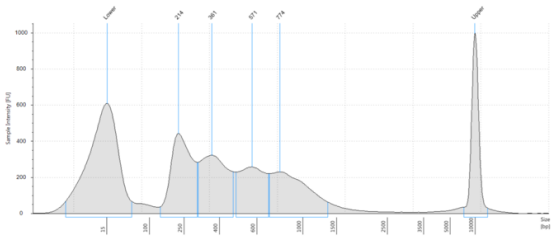
Dll4in3:GFP⁻ x kdrl:HRAS-mCherry⁺



hlx-3:GFP⁺



Dll4in3:GFP⁺



Ephb4-2:GFP⁺

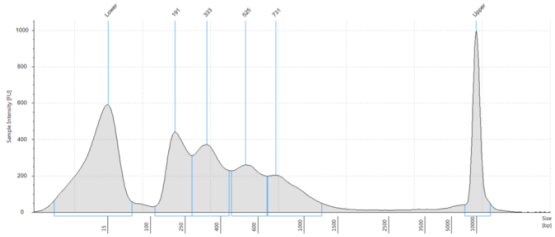


Figure 10: PCR amplification and tagmented DNA profiles. A) *PCR cycling conditions for the initial PCR amplification, the continued PCR on the remaining tagmented DNA and the qPCR used to determine the required number of remaining cycles.* ***B)*** *D1000 TapeStation profile of genomic input DNA generated from sorted kdrl:HRAS-mCherry⁺ cells at 27 hpf.* ***C)*** *Nucleosomal profiles generated by D5000 TapeStation for ATAC libraries of FAC sorted Ephb4-2:GFP⁺, hlx-3:GFP⁺, Dll4in3:GFP⁺ and Dll4in3:GFP⁻ x kdrl:HRAS-mCherry⁺ cells show a periodicity of roughly 150 bp, which corresponds to the fragments protected by nucleosomes. These profiles are representative for all three replicates at 27 hpf.*

2.7.2.2 RNA

For RNA-seq, cells were sorted directly into 350 µl RLT plus buffer containing 1% BME. Amount of needed lysis buffer was adjusted by measuring final volume of sorted cells, 175 µl of RLT/BME were added per 120 µl of sorted cells (which roughly corresponds to

35,000 cells). RNA was isolated according the manufacturer's instruction using the RNEasy Micro Plus kit (Cat# 74034, Qiagen). Quality and quantity of extracted RNA was determined by 2100 Bioanalyzer (Agilent). RNA integrity number (RIN) values ≥ 8 were attained for all samples sequenced (see Table 4).

Table 4: Amount of total extracted RNA (in ng) and RIN value for each sample at 27 hpf.

	Dll4in3:GFP+			**Dll4in3:GFP- *x* kdrl:HRAS-mCherry+**			**Ephb4-2:GFP+**			**hlx-3:GFP+**		
Replicate #	1	2	3	1	2	3	1	2	3	1	2	3
Total [ng]	51.8	74.0	9.5	17.1	9.9	4.5	53.1	126.1	8.2	7.3	10.0	12.4
RIN	8	9.6	9.3	8.4	9.7	9.9	9.8	8.6	9.6	10	8.8	10

For the library construction, 4 ng of total RNA per sample was used for the SMARTer® Universal Low Input RNA for Illumina® Sequencing kit (Cat# 634828, Takara). Library preparation and sequencing was performed at the WTCHG Oxford Genomics Centre. Three biological replicates were generated for each sub-endothelial group and the resulting 12 libraries were size selected and multiplexed before 75 bp paired-end sequencing over two lanes on an Illumina® HiSeq4000 flow cell.

2.7.3 Bioinformatic analysis

2.7.3.1 ATAC-sequencing

Raw sequencing reads were de-multiplexed using the indexing primers. Read quality was evaluated using FastQC[193] and reads were trimmed to remove low quality bases using Sickle[194]. Trimmed reads were mapped to the zebrafish September 2014 assembly (danRer10) reference genome[195]. Technical replicates were merged and mapped using Bowtie2 (bowtie2 -q -N 1 -x $danRer10/genome -X 2000 -1 $F_lane1,F_lane2 -2 $R_lane1,

R_lane2 -S $Output.sam)[196]. On average, 50 Million 75 bp reads per sample were sequenced with an alignment rate of 84.38%. Sequence alignment map (sam) files were converted to binary alignment map (bam) files and then sorted by name using SAMtools[197]. Duplicates were removed with Picard tools using the feature MarkDuplicates[198]. Bam files were converted to bigwig files using bamToGBrowse.pl. Statistically significant peaks were called using Model-based Analysis of ChIP-Seq [MACS2 (macs2 callpeak -t $sample.bam -c $genomicinput.bam -f BAMPE -g 1371719383 -n $Output -B -q 0.0])[199]. To visualize ATAC-seq data, bigwig files and called peaks for each sample were uploaded as a session to the University of California, Santa Cruz (UCSC) genome browser (https://genome-euro.ucsc.edu/cgi-/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=kira.fischer%40ludwig.ox.ac.uk&hgS_otherUserSessionName=danRer10)[200]. Peaks summits were annotated with Hypergeometric Optimization of Motif EnRichment (HOMER) annotatepeaks.pl[201]. Browser extensible data (Bed) files were merged with BEDtools to create one peak set per sub-endothelial cell type. Figure 11 shows clustering of all samples according to their sub-endothelial cell type and expression pattern at 27 hpf.

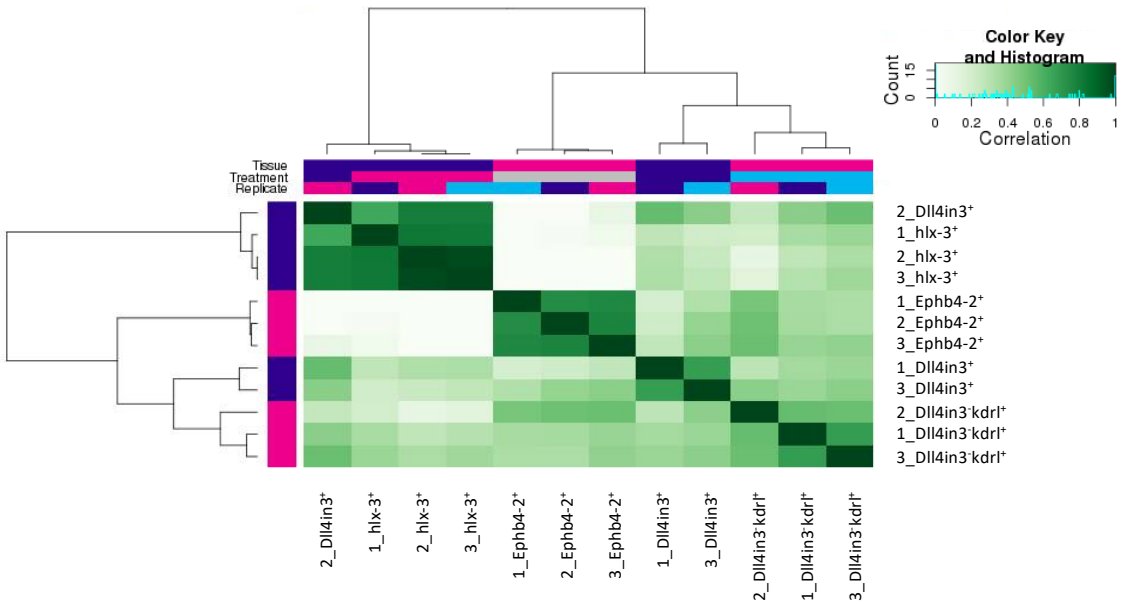


Figure 11: Correlation heatmap of all 12 ATAC-seq samples using peak caller score data of significantly differentially open regions *(Tissue = ISV and non-ISV, Treatment = sub-endothelial cell types). Biological triplicates cluster together with the exception of Dll4in3⁺ replicate #2. This could be explained by the overlapping expression pattern as both hlx-3⁺ and Dll4in3⁺ are expressed in sprouting ISVs.*

Annotated peaks were separated into promoters, introns, exons, TTS or intergenic regions according to their HOMER annotation (see Figure 12A). Motif analysis was performed only on peaks that were annotated as either intergenic, intronic or promoters with HOMER findMotif.pl[201]. Homer analysis for ISV-specific DNA motifs was performed on all hlx-3⁺ peaks that also had a peak in Dll4in3⁺ using bedtools intersect. ATAC-seq peaks from all sub-endothelial cell types were enriched within 500 bp of the TSS (see Figure 12B)[202].

Two approaches were used to test whether ATAC-seq can predict angiogenic-specific enhancers. A schematic illustration of the two methods can be found in Figure 12C and D. For this analysis I integrated ATAC-seq data generated from kdrl:HRAS-mCherry⁺ (pan-EC) and kdrl:HRAS-mCherry⁻ (non-EC) cells from *tg(kdrl:HRAS-mCherry)* fish at 26 hpf[203]. Both pan-EC and non-EC datasets were kindly provided by Dr Rui Monteiro and Dr Tomasz

Dobrzycki (Molecular Haematology Unit, WIMM, University of Oxford). ATAC-seq on these two populations was performed according the same protocol as described above.

To find putative angiogenic-specific enhancers, MACS2 peaks from hlx-3:GFP^{+} (hlx-3^{+}) triplicates were merged using BEDtools[204]. Peaks that were located in +/- 1,000 bp of the TSS, exons, histone clusters and repetitive elements [e.g. short interspersed nuclear elements (SINEs), microsatellites or tandem repeats] were excluded. Peaks were then intersected using BEDtools[204] (using parameter -wa) with ATAC-seq peaks found in at least two out of the four pan-endothelial kdrl:HRAS-mCherry^{+} replicates. I then ranked hlx-3^{+} accessible regions by -log10pvalue independent of the presence or absence of peaks in other sub-endothelial cell types [such as Dll4in3:GFP^{+} (Dll4in3^{+}), Ephb4-2:GFP^{+} (Ephb4-2^{+}) or Dll4in3:GFP^{-} x kdrl:HRAS-mCherry^{+} (Dll4in3^{-}kdrl^{+})]. See Figure 12C.

The second approach is illustrated in Figure 12D. Correlation analysis on ATAC-seq peaks was performed in R[205] using DiffBind[206]. Significant peaks comparing ISV (hlx-3^{+} and Dll4in3^{+}) vs. non-ISV (Ephb4-2^{+} and Dll4in3^{-}kdrl^{+}) were found using DESeq2[207]. Heatmaps were generated in R using ggPlot2 (see Figure 12D)[208]. The resulting 3,266 peaks were then intersected using BEDtools with ATAC-seq peaks found in pan-endothelial kdrl:HRAS-mCherry^{+} cells (same as for the first approach). Moreover, ATAC-seq peaks that overlapped with peaks found in non-endothelial kdrl:HRAS-mCherry^{-} cells were excluded, ending up with a total of 28 peaks.

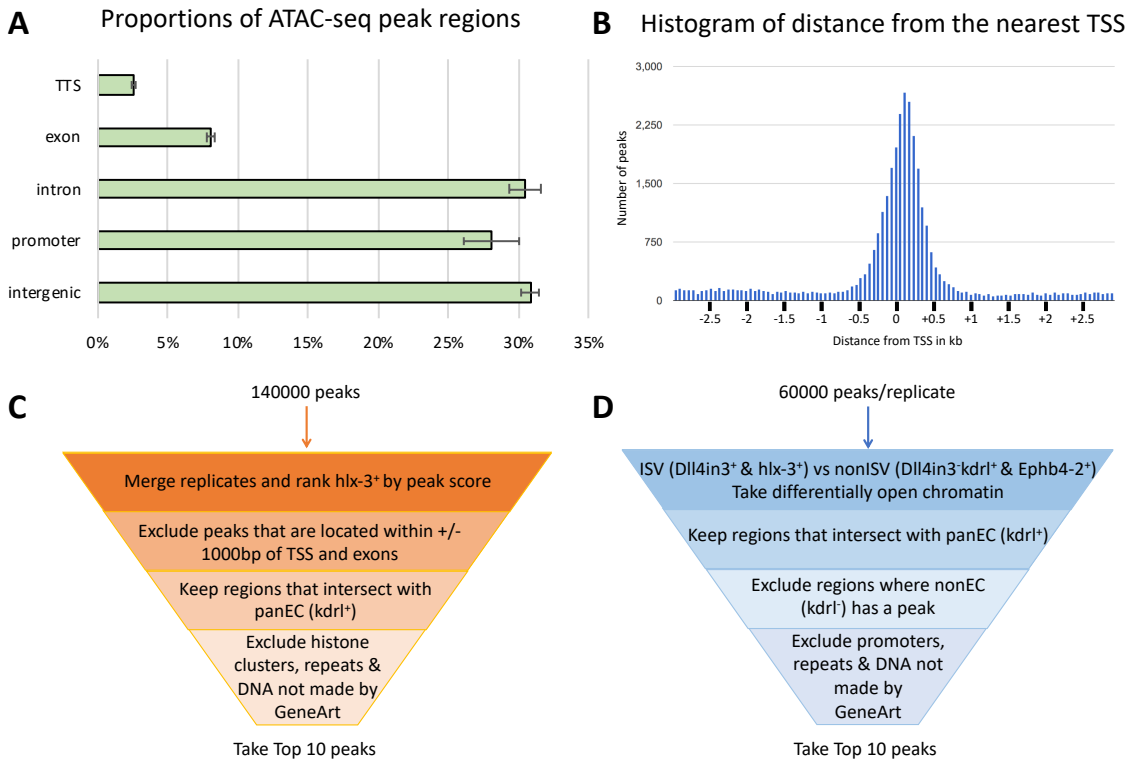


Figure 12: ATAC-seq data from sorted sub-endothelial Dll4in3-kdrl+, Dll4in3+, Ephb4-2+, hlx-3+ cell types. A) *Average proportions of the ATAC-seq peak regions identified in Dll4in3-kdrl+, Dll4in3+, Ephb4-2+, hlx-3+ cells at 27 hpf that are annotated as intergenic regions, introns, exons, TTS and promoter-TSS. Error bar shows standard error.* ***B)*** *Histogram of ATAC-seq peaks within 3 kb of the nearest TSS that were identified in hlx-3+. Representative of all ATAC-seq samples.* ***C)*** *Filtering of hlx-3+ ATAC-seq peaks and exclusion criteria to get top ten most open chromatin regions.* ***D)*** *Filtering of ISV vs. non-ISV ATAC-seq peaks to get top ten differential open chromatin regions.*

To test enhancer activity of the top ten peaks identified in the first and second approach, 600 bp surrounding the peak summit were ordered as GeneART strings.

2.7.3.2 RNA-sequencing

RNA-seq reads were adaptor and quality trimmed using TrimGalore[209] (SMARTer® IIa oligonucleotide was trimmed from every reverse read). Reads were mapped to the zebrafish genome assembly danRer10[195] using STAR [Spliced Transcripts Alignment to a Reference (STAR --genomeDir $danRer10/genome --outSAMtype BAM SortedByCoordinate --readFilesIn $F_lane1.fq,F_lane2.fq -2 $R_lane1.fq, R_lane2.fq --outSAMattributes All --outFileNamePrefix $Output --runThreadN 4)][210]. On average,

sequencing was performed to a depth of 30 million 75 bp paired-end reads per library with an alignment rate of 88.4%. Bam files were converted to bigwig files using bamtoscaledbigwig.pl, this script takes into account the total number of reads and makes it possible to compare samples between each other. Bigwig files for each sample were were uploaded onto UCSC browser. SAMtools[197] was used to filter reads for unique mapping and properly paired reads (-q 255 -f 0x2). PCR duplicates were removed using SAMtools rmdup. Aligned reads were counted with Subread featureCounts[211]. Differential expression analysis between sub-endothelial cell types was performed using EdgeR[212]. Briefly, reads were normalized for library size by calculating trimmed mean of M-values (TMM)[213]. The Benjamini-Hochberg method was applied on the p-values[214]. Differentially regulated transcripts were identified using the Generalized Linear Mixed Likelihood Ratio Test (GLM LRT)[215]. Multidimensional scaling (MDS) plots were generated in R using gplots[216].

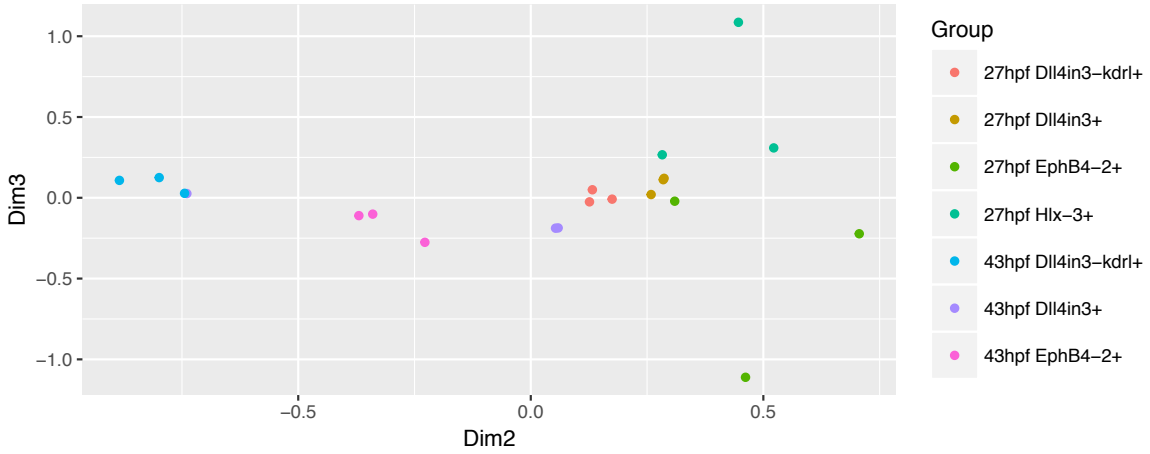


Figure 13: MDS plot showing the leading log-fold-changes between samples. *Biological replicates from the same sub-endothelial cell type cluster largely together. Dimension 2 separates samples according to time point (27 and 43 hpf).*

The MDS plot (see Figure 13) showed clustering of biological replicates, but also of two samples which had a low input. Lower amount of input RNA was hence used as a blocking

variable in the design matrix. The second factor denoted the sub-endothelial cell type. Comparisons were made by defining the following contrasts: sprouting ISVs (hlx-3^{+} and Dll4in3^{+}) vs. non-ISVs (Ephb4-2^{+} and Dll4in3^{-}kdrl^{+}) and veins (Ephb4-2^{+}) versus arteries (Dll4in3^{+}). ISV-specific hits were filtered for genes that are also overexpressed in pan-endothelial vs. non-endothelial cells. A gene was called overexpressed in a sub-endothelial cell type (ISV or vein) if the log2-fold-change (logFC) > 2, with a false discovery rate (FDR) < 0.05 and a $\log_2$ counts-per-million (logCPM) > 1. Non-differentially expressed genes were arbitrarily defined as having an average rpkm > 3 with a logFC < 1.

3 Identification of an angiogenic-specific enhancer

3.1 Rationale: Using enhancers to delineate regulatory pathways

3.1.1 Using enhancers as a tool to understand gene regulation in sprouting angiogenesis

Delineating the precise regulatory cascade controlling angiogenesis has been challenging. While the ability to generate endothelial-specific knockout mice has greatly increased our understanding of vascular biology, loss of many key endothelial genes often results in early lethality as the vasculature is first formed, making it hard to determine the precise angiogenic defects that occur. It is also hard to decouple potential angiogenic problems from other related vascular defects such as failure of vasculogenesis, smooth muscle recruitment or arteriovenous patterning[217]. While inducible endothelial-specific gene deletion can help avoid vasculogenic defects, these have the added challenge of choosing the correct time point to target angiogenesis without affecting other vascular events, and can give contradictory results depending on Cre driver and tamoxifen dosage[103,102]. Further complicating analysis by gene deletion, knockout mouse models often suffer from redundancy due to compensatory effects, especially for transcription factors that are part of a larger gene family with multiple endothelial-expressed factors. Therefore, it is often hard to determine the cause of vascular failure in mouse models with a phenotype, yet often hard to rule out a potential role for a potentially redundant gene when no phenotype is seen.

To gain further insight into the regulatory processes governing angiogenesis, I hypothesised that analysis of enhancers could be used to identify transcriptional regulators of genes switched on during angiogenesis, thus identifying transcription factors

likely to be important for switching on genes during this process without the initial requirement of analysing a knockout mouse model. Angiogenesis is a highly dynamic process where correct spatial and temporal gene expression is crucial. Analysis of mRNA expression of many key angiogenic genes such as *Jag1* and *Hlx* have clearly demonstrated transcription of these genes is specifically activated during angiogenic sprouting, pointing to a crucial role for transcriptional regulation in angiogenesis[164,112]. Since enhancers are the primary regulators of gene activity, analysis of transcription factor binding sites within angiogenic sprouting-specific enhancers could identify the transcription factors regulating its gene expression. Additionally, it may be possible to use this information to further determine the upstream signals involved in this process.

The success of enhancer-centred approaches to study endothelial heterogeneity has already been demonstrated by work from my group focused on arteriovenous differentiation[167,218,219,179]. For example, two enhancers for *Dll4* (one located in intron 3 (Dll4in3) and the other 12 kb upstream) were identified which were both active in arterial endothelial cells but silent in venous cells[167]. Analysis of these enhancers found essential binding motifs for RBPJ (the transcriptional effector of Notch signalling) and SOXF transcription factors. Further analysis demonstrated that these two factors combinatorially activate arterial transcription of *Dll4,* and many other arterial-specific genes[167]. This work therefore demonstrated that enhancer analysis can be used to identify transcriptional regulators controlling gene expression within the heterogeneous endothelium.

3.1.2 Aim: Finding an enhancer specifically active during sprouting angiogenesis

Until now, the transcriptional regulation of genes expressed in the angiogenic sprout has not been well understood. This project aims to find and analyse angiogenic-specific enhancers to understand how pro-angiogenic signals are integrated into gene activation, and how common signalling pathways, in particular the Notch and VEGF pathways, are able to enact this specific behaviour.

3.2 Finding an angiogenic sprouting-specific enhancer

Enhancers regulate spatio-temporal expression of gene transcription and contain many transcription factor binding sites[220]. My aim was to find an enhancer driving reporter gene expression at the angiogenic front. This enhancer would then be able to be used as a tool to dissect the transcriptional mechanism during sprouting angiogenesis. I chose two candidate genes which are highly expressed in sprouting endothelial cells, *HLX* and *JAG1*.

3.2.1 *In silico* analysis of putative angiogenic-specific enhancers

The first step in my approach was to screen the genomic landscape of human *HLX* and *JAG* genes for putative regulatory elements *in silico* on the UCSC browser (http://genome.ucsc.edu). Enhancers are marked by cell-type-specific DNase I hypersensitivity, p300 binding and histone modifications, such as H3K4me1 and H3K27ac[221,222,223]. Therefore, to identify putative enhancer elements in these two loci, I first looked for endothelial cell-specific histone modifications, provided by the Encyclopaedia of DNA Elements (ENCODE) project[224]. Specifically, this database provided information on regions enriched with H3K4me1 and H3K27ac marks in HUVECs[224]. While H3K27ac primarily marks active enhancers, H3K4me1 marks enhancers independent of whether they are active or poised, and also marks promoter regions. I also investigated H3K4me3, which marks promoter regions but not enhancers[222], to rule out promoter regions. Additionally, I also looked at DNase I hypersensitivity in 14 different cell types, including human umbilical vein and microvascular endothelial cell types.

For the *HLX* locus, three putative enhancer regions were identified. All three putative enhancers lie upstream of the TSS. These were named HLX-3, HLX-5 and HLX-35, with the numbers denoting the distance in kilobases from the transcriptional start site (Figure 14A). H3K27 acetylation (light blue) and DNase I hypersensitivity (DNase HS) sites in endothelial cells were concentrated around the three putative enhancers. Conservation between the mouse sequence and opossum, platypus, chicken, xenopus and zebrafish is indicated by black vertical lines in Figure 14. Both HLX-3 and HLX-35 were conserved between mammals and zebrafish, whereas the HLX-5 enhancer was only conserved down to egg-laying mammals such as platypus.

For the *JAG1* locus, I identified four putative enhancers using enhancer-associated histone marks and endothelial-specific DNase I hypersensitivity. All putative *JAG1* enhancers had endothelial H3K27ac, but varying degrees of DNase I hypersensitivity and sequence conservation. Three of these putative regions were located in introns and were consequently labelled JAG1in2, JAG1in5 and JAG1in20, while the last candidate region, JAG1+46, was located +46 kb downstream of the TSS (Figure 14B).

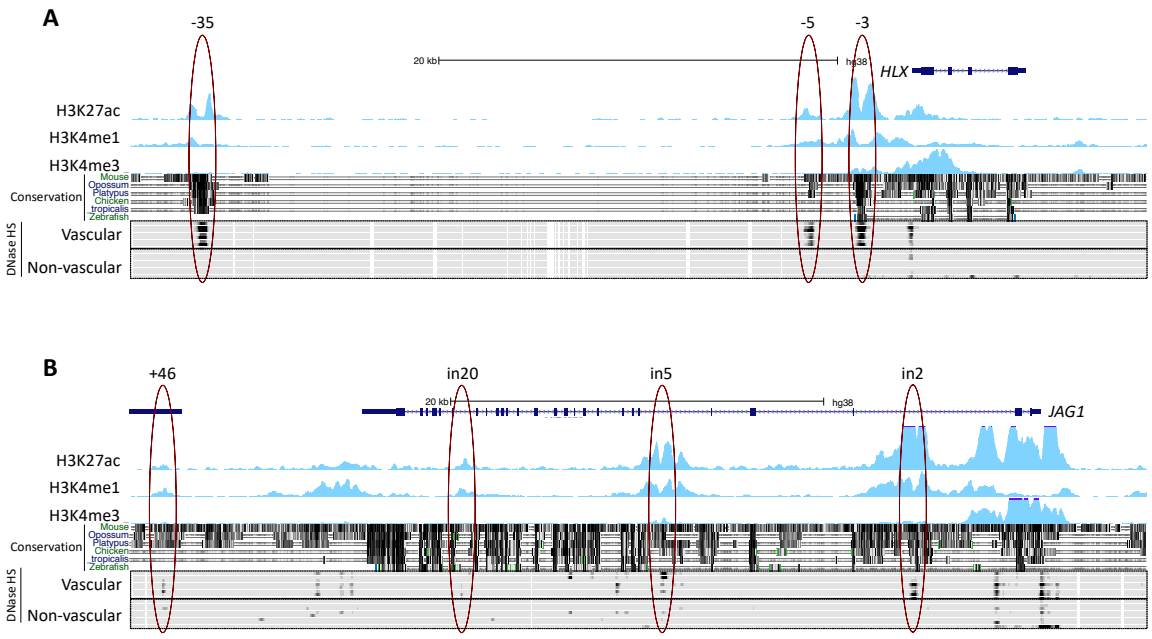


Figure 14*: *In silico analysis of putative angiogenic-specific enhancers. A) *The human HLX gene loci annotated with histone marks, sequence conservation and DNase I hypersensitivity information provided by UCSC ENCODE Genome Browser (http://genome.ucsc.edu)[224]. Putative angiogenic enhancers HLX-35 kb, HLX-5 kb and HLX-3 kb (indicating position with respect to the TSS) are circled in red.* ***B)*** *The JAG1 gene loci annotated with histone marks, sequence conservation and DNase I hypersensitivity information provided by UCSC ENCODE Genome Browser. Putative angiogenic endothelial enhancers JAG1in2, JAG1in5, JAG1in20 (indicating intronic position), and JAG1+46 kb are circled in red. Conservation of underlying DNA sequence is shown for mouse, opossum, platypus, chicken, xenopus and zebrafish. HUVEC-specific H3K27ac, H3K4me1 and H3K4me3 are depicted in light blue. DNase I hypersensitivity sites for vascular cells (HUVECs and HMVECs) and non-vascular cells (K562, HepG2, HeLa, Th2, SK-N-MC, NHEK and hES) are highlighted by dotted boxes.*

3.2.2 Putative angiogenic-specific enhancers in transient transgenic zebrafish

3.2.2.1 Rationale

To establish if the *in silico* identified putative enhancer regions were *bona-fide* developmental enhancers active during angiogenesis, I first tested their ability to drive GFP reporter gene expression in transgenic fish. Although I am ultimately interested in mammalian angiogenesis, the processes occurring during sprouting angiogenesis (see Introduction, section 1.2.2.1) are thought to be largely conserved between zebrafish and mammals. However, tip cell mitosis occurs more frequently in the sprouting ISVs of zebrafish than in the post-natal retina of mice[225,226,76]. In zebrafish, the dorsal aorta and cardinal vein (known collectively as the axial vessels) form by vasculogenesis, and

circulation starts at approximately 25 hpf. The first angiogenic sprouts are seen in the ISVs, which are formed within 36 hours after fertilization[227]. First, at around 23 hpf, Vegfa-guided sprouts grow bilaterally and dorsally from the dorsal aorta along vertical somite boundaries. When sprouting tip cells reach the dorsal side of the embryo, sprouts grow longitudinally. From around 28 hpf, sprouts connect, anastomose, to form the dorsal longitudinal anastomotic vessel (DLAV)[70,228]. Thereafter, vessels lumenize and circulation can be observed in intersegmental arteries and the DLAV. In a second wave from 32 hpf, sprouts from the posterior cardinal vein grow to form intersegmental veins and lymphatics[229]. A proportion of these venous sprouts connect with the base of ISVs to form intersegmental veins, while the rest grows further to the parachordal region to contribute to lymphatic vessels[230].

Because of the ease of manipulation in zebrafish, I ideally wanted to work on an enhancer that was downstream of an evolutionary conserved pathway and was active in both zebrafish and mouse transgenic models. Therefore, as I established whether the putative HLX and JAG enhancers were able to drive expression during sprouting angiogenesis, I investigated their ability to drive GFP expression in ISVs as they sprouted between 24 – 32 hpf in transgenic zebrafish. Since I used the human annotation to predict these enhancers, I tested the human sequence of each enhancer in zebrafish. For this assay, I used Tol2-mediated transgenesis, and analysed GFP activity in F0 mosaic zebrafish. Each putative enhancer was cloned upstream of the E1b core promoter and the GFP reporter gene in a Tol2 vector, and co-injected into one-cell zebrafish embryos along with RNA encoding the Tol2 transposase enzyme[183]. Tol2-based transgenesis results in mosaic but

highly efficient genome integration of the transgene. Reporter gene expression occurs only when the enhancer is active and transcription occurs.

3.2.2.2 Putative HLX enhancer

HLX-5:GFP transgenic fish did not have any GFP expression in the vascular compartment. Conversely, both HLX-3:GFP and HLX-35:GFP were able to drive limited GFP expression in ISVs in a small number of transgenic F0 zebrafish (Figure 15). Interestingly, neither drove any GFP expression in the axial vessels. [Please note that the HLX-35:GFP F0 transgenic fish were generated by Dr Natalia Sacilotto].

3.2.2.3 Putative JAG1 enhancer

Of the four putative *JAG1* enhancers, only JAG1in5:GFP showed any expression in endothelial cells (Figure 15). Similar to HLX-3 and HLX-35, JAG1in5 showed no GFP expression in the dorsal aorta or cardinal vein. [Please note that the intronic JAG1 constructs were cloned by Branwen Brockley, a summer intern who worked with me].

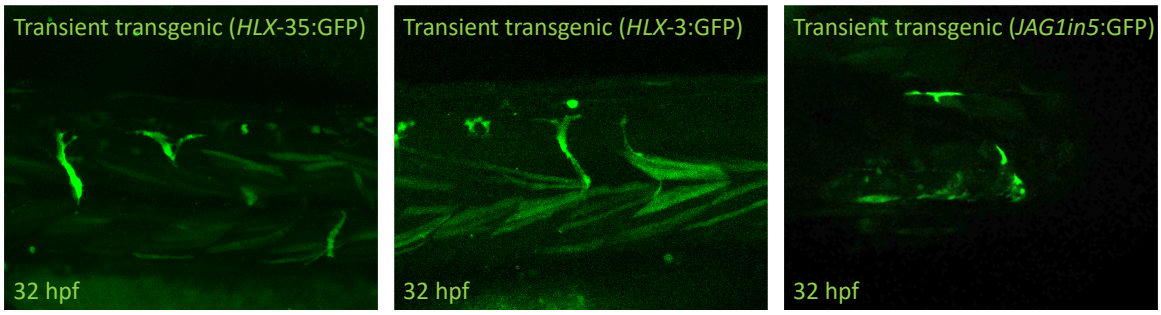


Figure 15: Putative angiogenic-specific human enhancers for HLX and JAG1 in transient transgenic zebrafish. *All three images show a zoom of the developing zebrafish trunk at 32 hpf. GFP is expressed in sprouting ISVs in HLX-35:GFP and HLX-3:GFP, while an already anastomosed endothelial cell shows GFP expression in JAG1in5:GFP. The fish also show ectopic transgene expression in the muscle. A minimum of n = 47 zebrafish were injected with each construct. All zebrafish images in this thesis are orientated in the same way: a lateral view with the rostrum of the zebrafish to the right is shown.*

3.2.3 Putative HLX and JAG1 enhancers in transient transgenic mice

Next, the zebrafish positive enhancers were analysed in mouse transgenic models, to establish if they could drive reporter gene expression during mouse vascular embryonic development[231]. For this assay, the same human enhancer sequences were cloned upstream of the silent hsp68 minimal promoter and β-galactosidase (*LacZ*) reporter gene. *LacZ* reporter gene expression and thus enhancer activity can be visualized by blue X-Gal staining[51]. I analysed transgene expression at E11.5, a time point which allowed the evaluation of enhancer activity in the intersegmental vessels and growing microvasculature.

The HLX-3, HLX-35 and JAG1in5 putative enhancers were all active in some ISVs in transgenic zebrafish. However, only HLX-3 was able to drive robust *LacZ* expression in endothelial cells of transgenic mice embryos at E11.5 (Figure 16A). HLX-35:*LacZ* transgenic embryos did not show any detectable reporter gene expression, while JAG1in5 drove *LacZ* expression but not in endothelial cells; instead activity appeared to be specific for cartilage (see Table 5). Since the HLX-3 enhancer was the only one to drive reporter gene expression in sprouting endothelial cells in both zebrafish and mice, I decided to focus my further investigation of angiogenic transcription on the HLX-3 enhancer.

Table 5: Summary of reporter gene expression detected in E11.5 mice transgenic for the HLX-3:LacZ, HLX-35:LacZ and JAG1in5:LacZ transgenes.

Transgene	**HLX-35:*LacZ***	**HLX-3:*LacZ***	**JAG1in5:*LacZ***
Number of mice injected	9	18	9
Number of transgenic mice	3	11	4
Transgenic mice with X-Gal expression in endothelial cells	0	7	0

3.2.4 Testing the zebrafish orthologue of HLX-3 in transgenic mice

The HLX-3 enhancer was highly conserved, and I was able to identify an orthologous enhancer region in the zebrafish *hlx1* genome. In zebrafish, the enhancer was located 1 kb upstream of the TSS. But to maintain consistency with the human locus, I labelled the zebrafish sequence hlx-3. I have already established that the human HLX-3 enhancer is active in transgenic zebrafish. To further confirm that the key regulatory motifs driving HLX-3 expression are conserved in the zebrafish hlx-3 enhancer, I next tested whether the zebrafish hlx-3 enhancer could drive endothelial angiogenic reporter gene expression in transgenic mice. Therefore, I cloned it upstream of hsp68 and *LacZ* and tested its activity in transgenic mice. Further confirming the conservation of key transcriptional information between HLX-3 and hlx-3, the vascular-specific expression pattern of hlx-3:*LacZ* was very similar to the human HLX-3:*LacZ* transgene (Figure 16). [Please note that all pro-nuclear injections for the generation of transgenic mice were performed by Dr Ke Liu].

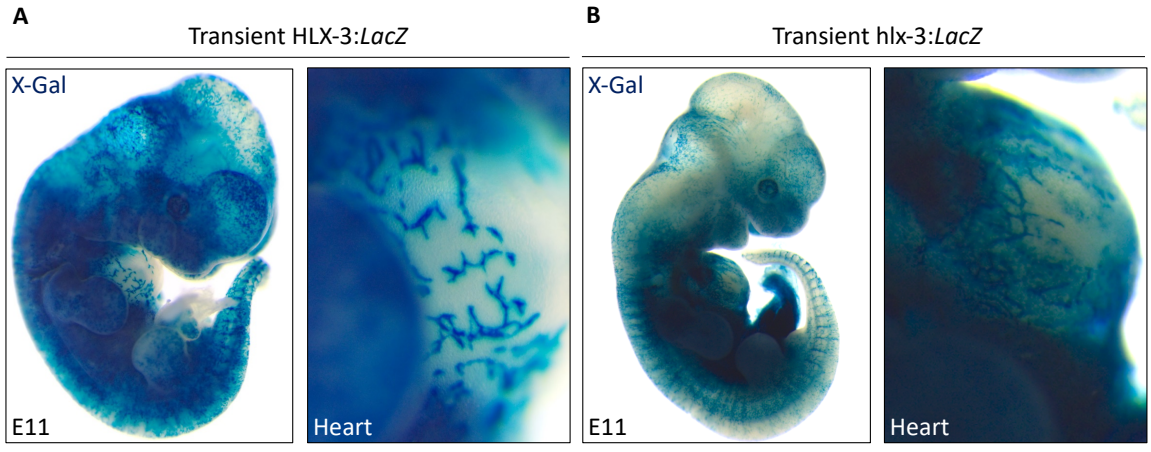


Figure 16: HLX-3 enhancer is functionally conserved in transient transgenic mice at E11. A) *The human HLX-3 enhancer sequence drives LacZ expression in sprouting endothelial cells.* ***B)*** *Zebrafish hlx-3:LacZ is also expressed in sprouting endothelial cells.*

3.3 Analysis of hlx-3:GFP activity during development in zebrafish

3.3.1 Optimization of the hlx-3 construct in transient transgenic zebrafish

I next aimed to examine the activity of the hlx-3 enhancer in transgenic zebrafish, compared to the known angiogenic sprout-specific expression of the endogenous hlx1 gene[164]. In order to find the optimal hlx-3 construct for zebrafish transgenesis, I compared three different lengths of the zebrafish enhancer sequence in transient transgenic zebrafish. The shortest hlx-3 sequence (445 bp) resulted in the most consistent GFP expression in sprouting ISVs. For the 445 bp construct 72% of transient transgenic fish showed GFP expression in ISVs, while for the 650 bp construct only 13% showed GFP expression in ISVs and for the 850 bp construct 20% showed ISV expression (see Table 6). It is unclear exactly why the shorter hlx-3 enhancer was better at driving GFP expression, although the longer sequences might have included binding sites for inhibitory proteins and repressed angiogenic-specific activity. However, since the 445 bp hlx-3:GFP was robustly active in the angiogenic sprouts, I used this short hlx-3 sequence to generate the stable line and for all further experimental work.

Transgene	hlx-3 445 bp	hlx-3 650 bp	hlx-3 850 bp
Number of injected zebrafish	79	28	61
% of transgenic fish with GFP expression in ISVs	72%	13%	20%

Table 6: Effect of transgene hlx-3 enhancer length on reporter gene expression in sprouting ISVs in transient hlx-3:GFP zebrafish.

3.3.2 hlx-3:GFP is specifically expressed in sprouting endothelial cells

To establish if hlx-3:GFP was specifically active during sprouting angiogenesis, I analysed its expression pattern over a time-course during the first five days of embryonic development (Figure 17). Transgene activity was first detected at 26 hpf, with GFP expression detected in the first sprouting ISVs. Stalk cells as well as filopodia bearing tip cells were GFP positive. At 29 hpf, hlx-3:GFP transgene expression was stronger and remained specific to sprouting endothelial cells. At 31 hpf, hlx-3:GFP was active as sprouts migrated and proliferated to anastomose with the neighbouring cell. At 48 hpf, the hlx-3:GFP transgene was active in the regions where caudally located endothelial cells in the DLAV interconnect with the opposite cell. Occasionally (in 13% of all ISVs) hlx-3:GFP was active in the intersegmental veins that formed from the PCV during the second wave of venous sprouting. As GFP expression was most pronounced at the root of the newly formed intersegmental veins, it is likely that the hlx-3 enhancer was active as venous sprouts connected to intersegmental arteries.

Importantly, from three days post fertilization (dpf), hlx-3:GFP expression faded from endothelial cells in ISVs and the DLAV during maturation. During later stages of zebrafish development, hlx-3:GFP expression was active in the root of some intersegmental arteries. However, from 2 dpf, hlx-3:GFP activity continued in tissues that undergo angiogenesis later, such as the embryonic intestine, brain and future gills[228]. At 48 hpf, hlx-3:GFP became increasingly expressed in the developing cranial vasculature (such as the central arteries (CtA), mesencephalic vein (MSV), primordial hindbrain channel (PHBC) and the primary head sinus (PHS)). hlx-3:GFP activity increased as the cranial trunk and caudal head vasculature connected *via* the basilary artery and PHBC. These endothelial

cells in the head vasculature are known to express *Vegfr2*/*Kdr* and *Vegfr3*/*Flt4*[232]. As the eye vasculature formed, hlx-3 drove GFP expression in active inner optic circle (IOC) and the plexus within the optic cup. Also, hlx-3:GFP was already visible in the circulatory vessel of the developing pectoral artery and vein (PA/PV).

At 2.5 dpf, hlx-3:GFP was active in a few cells of the ventral floor of the cardinal vein and this expression increased until 4.5 dpf (Figure 17). These cells could correspond to the previously described angioblasts that give rise to the sub-intestinal plexus and hence organ-specific vasculature[233]. hlx-3:GFP was also active in the sub-intestinal plexus, including the sub-intestinal vein (SIV) and interconnecting vessels, throughout gastrointestinal organ formation. The formation of interconnecting vessels from the SIV *via* dorsal migration was shown to be Vegf-dependent, while ventral migration was Bone morphogenetic protein (Bmp)-dependent[233].

At 3 dpf, hlx-3:GFP was strongly expressed in all developing six aortic arches of the bulbus arteriosus (BuA) and in sprouts that will form the branchial laminar arteries of the future gills. Conversely, expression of hlx-3:GFP decreased in the largely unaltered cranial vasculature, except in the posterior cerebral veins (PCeV)[228]. Between 3 to 5 dpf, the PCeV and associated vessels grow by angiogenic sprouting to connect to the forming myelencephalic choroid plexus[234].

At 4 dpf, hlx-3:GFP enhancer expression was also weakly visible in parachordal vessels (PAV) and even fainter in the vertebral arteries, both of which form horizontally from ISVs. Enhancer activity was still observed in yolk-absorbing SIVs and gut-associated vessels

which drain into the hepatic portal vein (HPV). There, the HPV empties into the reticular hepatic sinusoid which will form the future liver (L).

Interestingly, hlx-3:GFP was always active in the most caudal tip of the aorta where it drains into the caudal vein. Further, GFP expression was frequently detected in the heart at all stages. Taken together, these results indicate that the hlx-3 enhancer provides all the information necessary to drive GFP expression specifically in active angiogenic endothelial cells in the zebrafish embryo.

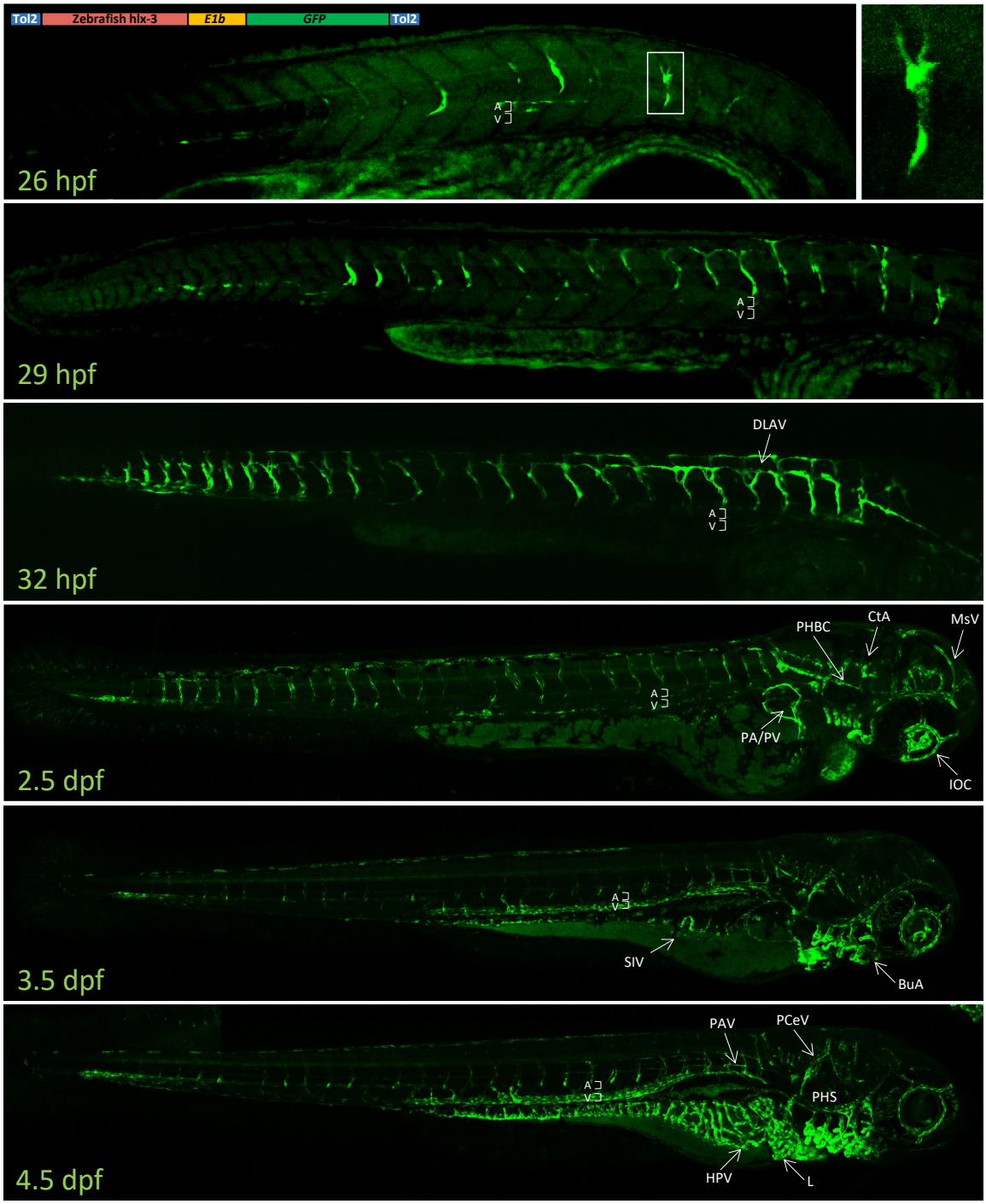


***Figure 17: Time-course of stable transgenic tg(hlx-3:GFP) zebrafish line**. At 26 hours post fertilization (hpf), GFP was detected in a few sprouting ISVs. Zoom shows GFP expression in both stalk and tip cells. At 29 hpf, hlx-3 drove GFP expression in sprouting ISVs, anastomosing tip cells which will form the DLAV and in endothelial cells in the caudal artery which were about to sprout. At 32 hpf, hlx-3 was active in ISVs and the DLAV. At 2.5 days post fertilization (dpf), GFP expression was fading from ISVs, but GFP expression became more pronounced in the cranial vasculature and the pectoral artery/vein. At 3.5 dpf, hlx-3:GFP expression was observed in the bulbus arteriosus, sub-intestinal plexus (inlcuding the sub-intestinal vein) and the floor of the cardinal vein. At 4.5 dpf, hlx-3:GFP is strongly expressed in the hepatic sinusoid and occasionally in the parachordal vessels. A = dorsal aorta, V = cardinal vein, DLAV = dorsal longitudinal anastomotic vessel, PHBC = primordial hindbrain channel, CtA = central arteries, MsV = mesencephalic vein, IOC = inner optic circle, PA = pectoral artery, PV = pectoral vein, BuA = bulbus arteriosus, SIV = sub-intestinal vein, PHS = primary head sinus, PCeV = posterior cerebral vein, HPV = hepatic portal vein, L = liver, PAV = parachordal vessels.*

3.3.3 Comparison of hlx-3 with endogenous *hlx1* expression in developing zebrafish

Zebrafish *hlx1* has been reported to be expressed specifically in the angiogenic sprout in zebrafish[164]. My results so far suggest that hlx-3:GFP is also specifically active during sprouting angiogenesis, suggesting that it may closely mimic the endogenous gene. To confirm this, I compared hlx-3:GFP activity with *hlx1* mRNA expression at 25, 27 and 30 hpf (Figure 18). As expected, *hlx1* was expressed in sprouting endothelial cells at all time points. Both hlx-3:GFP and *hlx1* were detected in ISVs and the angiogenic mid-cerebral vein. *hlx1* and hlx-3:GFP were also seen in endothelial cells in the roof of the caudal artery which are about to undergo ISV sprouting. hlx-3 driven GFP expression was more pronounced than *hlx1* in tip cells and anastomosing vessels which will form the DLAV. In conclusion, endogenous *hlx1* expression detected by *in situ* hybridization at 30 hpf correlated fairly well with hlx-3 enhancer driven GFP expression.

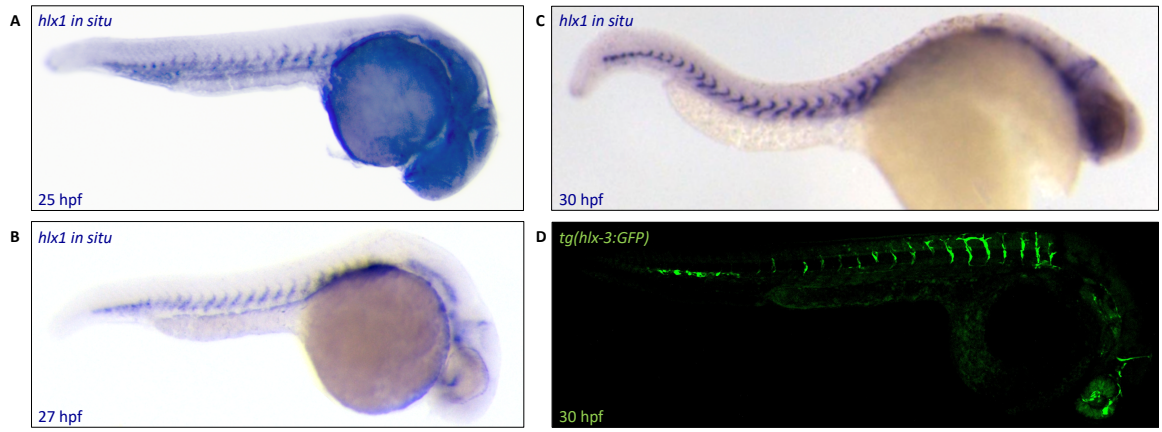


***Figure 18: hlx-3:GFP expression correlates with hlx1 expression**. **A)** hlx1 expression was detected by in situ hybridization at 25 hpf. **B)** at 27 hpf. **C)** at 30 hpf. At least 30 zebrafish embryos were analysed at each time point. **D)** Zebrafish hlx-3 drives reporter GFP expression in transgenic zebrafish line tg(hlx-3:GFP) in intersegmental vessels in the trunk and in the mid-cerebral vein in the head at 30 hpf. Very early during development, only at one day post fertilization, weak GFP expression was observed in the retina.*

3.4 Analysis of a stable HLX-3:*LacZ* line in mouse

3.4.1 HLX-3:*LacZ* is active in sprouting endothelial cells during embryonic development

3.4.1.1 Time-course of HLX-3 driven reporter gene expression

Experiments in the previous section confirmed angiogenic-specific expression of the zebrafish hlx-3 enhancer in transgenic zebrafish. To analyse the spatio-temporal expression pattern of the HLX-3 enhancer activity in mice, I generated a number of stable transgenic mouse lines using the same human HLX-3:hsp68:*LacZ* transgene tested earlier in transient transgenic mouse embryos. In total, six different independent stable HLX-3:*LacZ* transgenic lines were analysed during embryonic development and post-natal retinal angiogenesis. In general, all HLX-3:*LacZ* lines showed a similar expression pattern but had different staining intensities. Whole-mount staining of E11 embryos showed X-Gal staining in intersegmental arteries and the developing microvasculature (Figure 19). This suggests that the pattern of HLX-3:*LacZ* expression in endothelial cells was largely independent of copy number and location within the genome where the transgene was inserted.

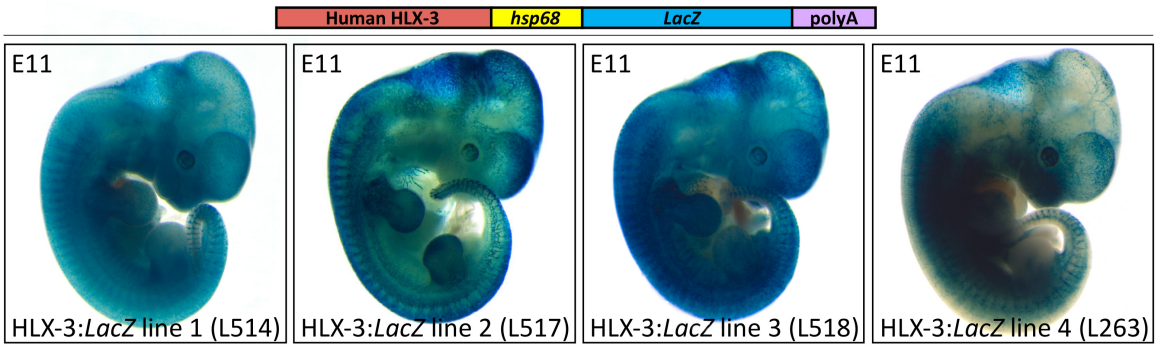


Figure 19: Representative whole-mount X-Gal stained embryos at E11 from four independent stable HLX-3:LacZ transgenic mouse lines. *All four HLX-3:LacZ mouse lines (L514, L517, L518 and L263) demonstrated strong X-Gal staining in the growing microvasculature. For each transgenic line, the image is representative for at least four samples. For all following experiments, line L263 is shown. Although all experiments were also performed and validated in L514.*

Throughout embryonic development, the HLX-3:*LacZ* transgene was active in the dorsal aortae, intersegmental arteries and in the developing microvasculature, giving rise to a punctuate staining. The HLX-3 enhancer did not drive reporter gene expression in the yolk sac vasculature and was largely absent from veins. Vessels were determined by location and, additionally, arteries and veins were distinguished by morphology[235].

The HLX-3:*LacZ* transgene was found to be active in endothelial cells from very early in vascular development. At E8.25, HLX-3:*LacZ* was active in the bilateral, paired dorsal aortae, but absent in the vitelline veins. At E8.5, HLX-3 drove reporter gene expression into sprouting intersegmental arteries, primitive maxillary arteries and in disconnected cells, particularly in the developing cephalic plexus[236]. Transverse sections of the E8.5 embryo also showed X-Gal staining of the endocardial region of the primitive heart tube (Figure 20A). Surrounding the pharyngeal region of foregut diverticulum, HLX-3:*LacZ* was active in four vessels. The upper two vessels are the dorsal aortae, while the lower two will coalesce to form the first branchial arch artery. At E9 and E9.5, HLX-3:*LacZ* activity was clearly detected in the dorsal aortae, maxillary arteries, perineural vascular plexus, branchial arches and developing limbs (Figure 20B). In addition to transgene activity in capillaries and the fusing dorsal aortae, I occasionally noted HLX-3:*LacZ* activity in single endothelial cells in the cardinal veins. However, robust reporter gene expression was not seen in the venous compartment at these time points. At E10.5, HLX-3 reporter activity remained largely arterial-specific (Figure 20C). Interestingly, not all endothelial cells lining the dorsal aortae were X-Gal positive. Instead dorsal aortae staining was polarized towards the midline, where endothelial cells migrate to form a single vessel in the midline. Once the dorsal aorta was fused, *LacZ* expression was reduced and only appeared active

in endothelial cells at the site where intersegmental vessels sprouted. Similar to the process described in zebrafish, intersegmental arteries sprout dorsally between somites and branch longitudinally to form the DLAV. HLX-3 drove reporter gene activity in these branching vessels and in the capillary network forming the perineural vascular plexus in the trunk. However, by E13 no major vessels showed HLX-3:*LacZ* activity. These included the thoracic/abdominal aorta, vena cava, subclavian artery, umbilical arteries and azygous veins. The only larger vessels with positive X-Gal staining were the pulmonary and aortic trunks. However, intense X-Gal staining was observed in the microvasculature of the skin (Figure 20D). Small blood vessels in developing organs also had HLX-3:*LacZ* activity, including the pulmonary, mesonephric and intestinal vasculature. Strong X-Gal staining was also seen in the microvasculature of the eye, brain, nasopharyngeal cavity, tongue and developing bone (in particular the humerus, radius and scapula, data not shown). Moreover, HLX-3 was expressed in endothelial cells lining tail and intercostal arteries.

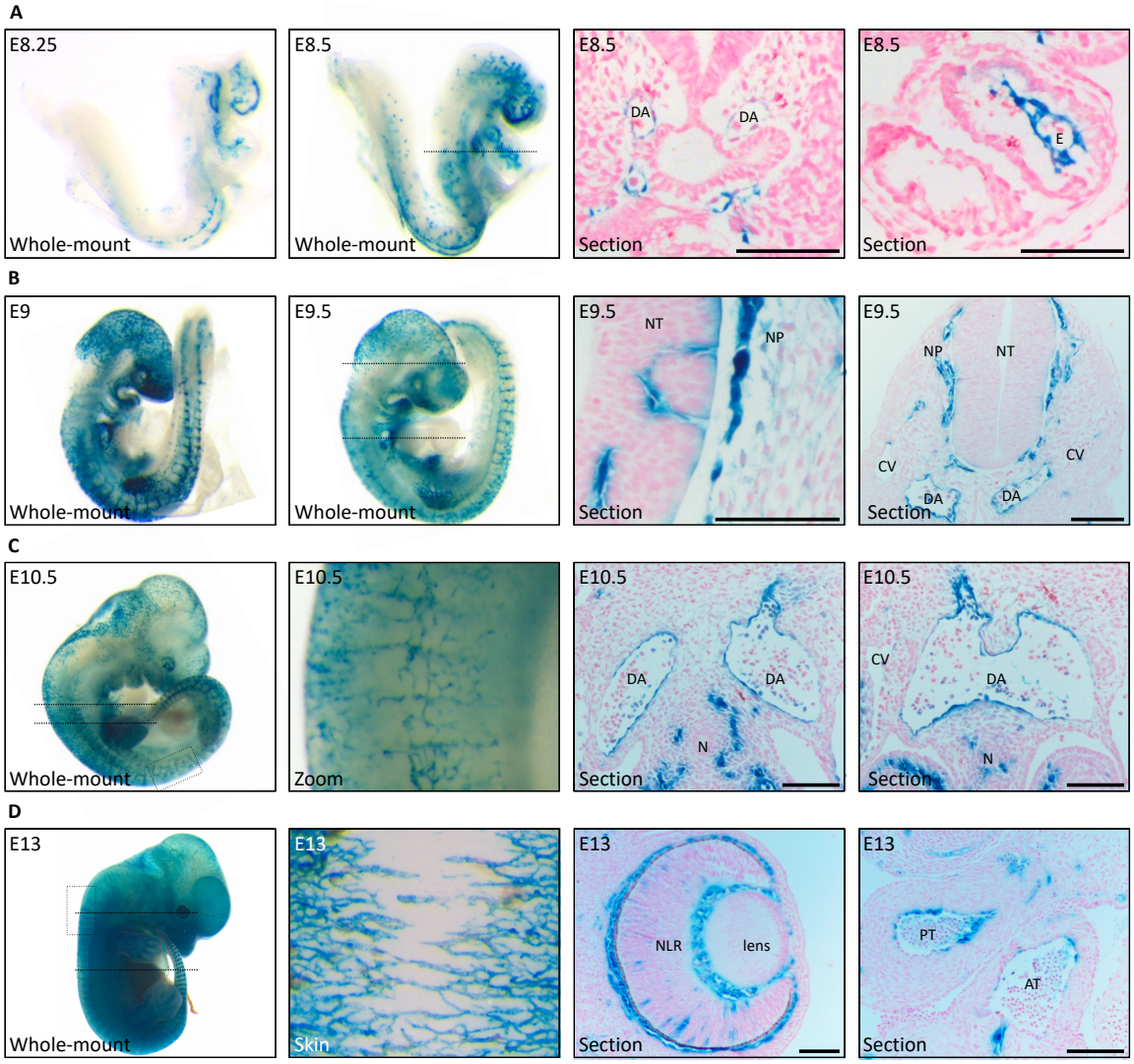


Figure 20: Analysis of HLX-3:LacZ transgene expression during early embryonic development. *Representative images of X-Gal stained (blue) whole-mount embryos are shown at E8.25, E8.5, E9, E9.5, E10.5 and E13. Transverse sections are shown for E8.5, E9.5, E10.5 and E13.* ***A)*** *HLX-3:LacZ is active very early during development at E8.25 and E8.5. Sections show X-Gal staining in the dorsal aortae, branchial arch arteries and the endocardial lining of the primitive heart tube.* ***B)*** *Strong X-Gal staining was observed in the cephalic vasculature, branchial arches and intersegmental arteries of the E9 and E9.5 embryo. The first section shows HLX-3:LacZ transgene expression in the microvasculature of the neural tube and in the perineural plexus. The second section shows X-Gal positive endothelial cells within the dorsal aortae and a single positive endothelial cell within the cardinal veins.* ***C)*** *HLX-3:LacZ transgene expression continued in the developing vasculature at E10.5. Magnified view on X-Gal positive intersegmental vessels and forming DLAV in the tail. Sections demonstrate polarized X-Gal staining in the dorsal aortae towards the midline, whereas the cardinal veins remained largely X-Gal negative.* ***D)*** *HLX-3:LacZ expression was strongly detected in the microvasculature of the skin giving rise to a blue hue. Flatmount of the skin taken from the back detects intensive X-Gal staining in the developing microvascular network. The first section shows HLX-3 driven LacZ expression in the vessels surrounding the eye. The second sections shows X-Gal positive endothelial cells within the aortic and pulmonary trunk. All transverse paraffin sections were co-stained with nuclear fast red with dorsal at the top. DA = dorsal aorta, E = endocardial region of heart, NT = neural tube, NP = perineural vascular plexus, CV = cardinal vein, N = notochord, NLR = neural layer of retina, PT = pulmonary trunk, AT = aortic trunk. Lines indicate location of sections, while boxes indicate location of the zoom. Scale bar is equal to 100 µm.*

At E11.5 and E13.5, β-galactosidase expression was compared with various endothelial markers. Therefore, I performed immunostaining followed by X-Gal staining on the same section. All HLX-3 enhancer activity overlapped with the pan-endothelial marker CD31 (see Figure 21). At E11.5, HLX-3 enhancer was active in growing microvessels (e.g. perineural vascular plexus, see Figure 21A). As described in the previous section, X-Gal staining in the αSMA positive dorsal aortae was polarized towards the midline (Figure 21A). At E13.5, HLX-3 enhancer was only active in the microvasculature (e.g. microvessels lining the thoracic vertebral body, oesophagus and bronchi, see Figure 21B). At this time point, HLX-3:*LacZ* expression was absent in more differentiated vessels and did not co-localize with EMCN positive veins and larger αSMA positive arteries.

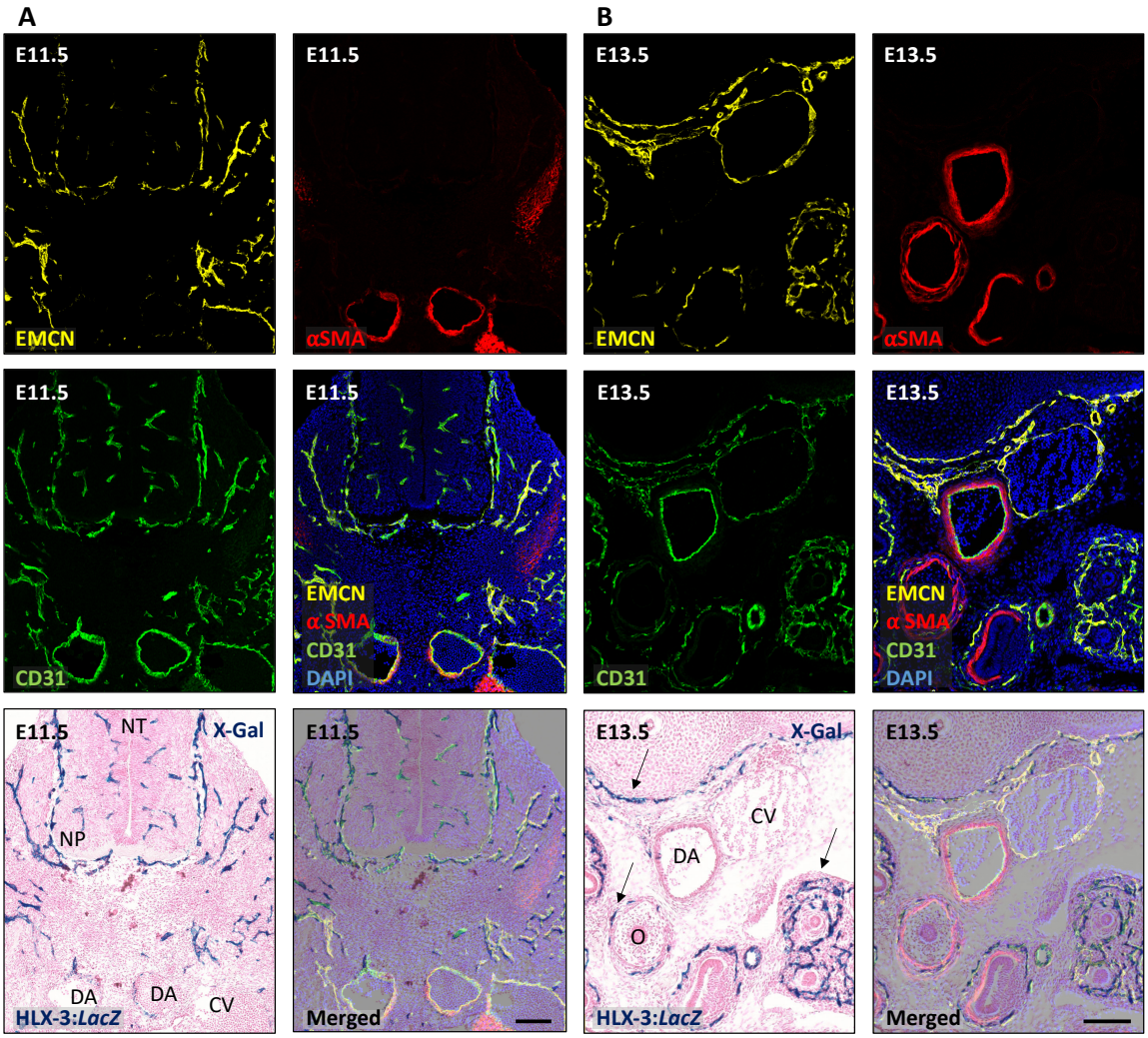


Figure 21: HLX-3:LacZ transgene expression overlaps with pan-endothelial marker CD31. *Anti-CD31 (green), anti-α smooth muscle actin (αSMA, red) and anti-Endomucin (EMCN, yellow) immunostaining and X-Gal staining (blue on nuclear fast red) was conducted on transverse section through a representative* ***A)*** *E11.5 and* ***B)*** *E13.5 HLX-3:LacZ transgenic embryo. HLX-3 activity was restricted to undifferentiated vessels. Arrows indicate location of X-Gal positive microvessels lining the thoracic vertebral body, oesophagus and bronchi. DA = dorsal (thoracic) aorta, CV = (posterior) cardinal vein, NT = neural tube, NP = perineural vascular plexus, O = oesophagus. Due to processing of the tissue between image acquisition, the tissue did not maintain the same size (swelling and/or shrinking) making it difficult to precisely match the immunofluorescent and light microscopy image. Antibody stained cryosections were imaged using the emission fingerprinting function of the Zeiss 710MP confocal microscope to eliminate autofluorescence. Scale bar is equal to 100 µm.*

Later during development, strong HLX-3 reporter gene expression in the microvasculature of the skin and the opacity of the embryo precluded a detailed analysis of whole-mount embryos. To overcome some of these issues, I cleared an E14 embryo[171] to allow visualization of the capillary network in the skin and anastomosing intersegmental vessels in the tail (Figure 22A). Again, HLX-3:*LacZ* was expressed in the vasculature of developing

organs, such as the lung. To study activity at E15, the embryo was partially dissected prior to X-Gal staining to allow better penetration of the X-Gal stain. HLX-3:*LacZ* activity was detected in the vasculature of the rib cage and in the digits of the developing limbs (Figure 22B). HLX-3 also drove reporter gene expression in the microvasculature of the brain, while it was again absent from major vessels. HLX-3:*LacZ* positive vessels were found in the choroid plexus, on the pia mater on the surface of, and in the broad nuclear layer of the neopallial cortex. At E16 and E18, HLX-3:*LacZ* activity was limited to just a few capillaries of newly formed organs (data not shown).

To summarize, HLX-3 driven *LacZ* expression was robustly detected during early stages of embryonic vascularization, where it was specific to sprouting intersegmental arteries and the perineural vascular plexus (E9.5). HLX-3:*LacZ* activity decreased in the major vessels from E10.5 and was mainly found in the microvasculature of developing organs by E13.

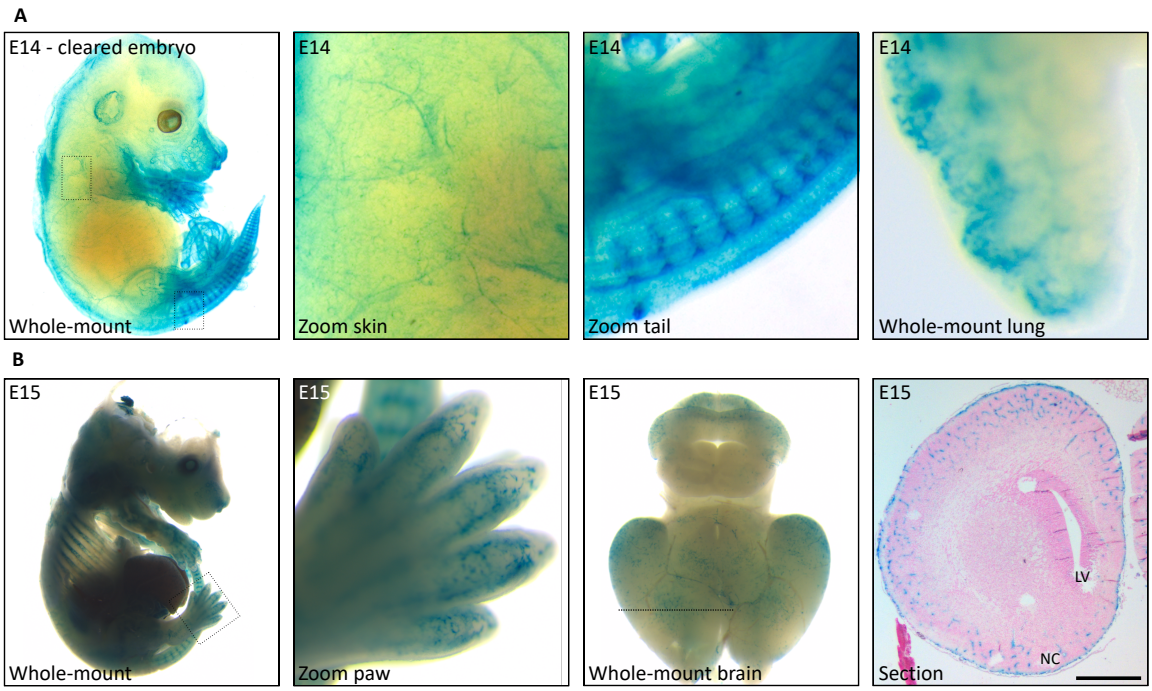


Figure 22: Analysis of HLX-3:LacZ reporter gene expression at later stages of embryonic development. *Representative images of X-Gal stained (blue) whole-mount embryos at E14 and E15.* ***A)*** *HLX-3:LacZ expression in the E14 embryo. Embryo was cleared to visualize reporter gene expression more clearly[171]. Magnified view of the skin in the shoulder region and the tail. HLX-3 was active in the intersegmental vessels. Whole-mount image of the right lung (middle and lower lobe) at E14 showing X-Gal staining in vessels surrounding the developing bronchial buds.* ***B)*** *HLX-3:LacZ expression in the E15 embryo. Embryo was partially dissected (internal organs were taken out for improved X-Gal infiltration). Magnified view on the hind-limb (paw) with X-Gal positive vessels in the developing digits. Ventral view of a whole-mount image of the brain. A coronal section (co-stained with nuclear fast red) is shown for half of the neopallial cortex of the brain at E15. LV = lateral ventricle surrounded by the ependymal layer, NC = nuclear layer of neopallial cortex. Lines indicate location of sections, while boxes indicate location of the zoom. Scale bar is equal to 500 µm.*

3.4.1.2 HLX-3:*LacZ* expression pattern in the developing heart

Analysis of the HLX-3:*LacZ* mouse lines found X-Gal staining in the developing heart, suggesting that HLX-3 may drive activity in the developing endocardium, or in the coronary vessels, both containing endothelial cells. During embryonic coronary vessel development, endothelial cells sprout mainly from two separate origins, the endocardium and the sinus venosus[237], and it is not clear which type of coronary vessels utilize classical angiogenic sprouting. It is also not clear how angiogenic sprouting relates to the formation of the endocardium.

Analysis of E8.5 HLX-3:*LacZ* mouse lines revealed early X-Gal staining in the endocardial lining of the primitive heart tube. To get greater clarity on the exact staining pattern of HLX-3:*LacZ* in the coronary vessels and endocardium at later stages, I next analysed HLX-3 activity during coronary vessel sprouting in hearts removed prior to X-Gal staining. At E12 and E13, HLX-3*:LacZ* transgene activity was primarily limited to the endothelial cells within the cardiac septum, and in vascular sprouts coming down the dorsal face of the heart (Figure 23). At E15, HLX-3:*LacZ* activity was seen mainly on the ventral side of the heart and again primarily around the septum. This population of coronary vessels has previously been reported to be made up of endocardial-derived endothelial cells. Conversely, little HLX-3:*LacZ* activity was seen in the sinus venosus-derived coronary plexus, located on the dorsal and lateral sides of the heart[238]. HLX-3:*LacZ* activity was robustly detected in coronary vessels of post-natal hearts (Figure 23).

To summarize, the HLX-3:*LacZ* transgene was active during coronary vessel growth of the embryonic and post-natal heart, and activity was primarily detected in the VEGFA-dependent endocardial-derived coronary vessels, agreeing with the known role of VEGFA in angiogenesis[239,240]. Some limited HLX-3:*LacZ* activity was also seen in the sinus venosus-derived vessels during their initial sprouting from the sinus venosus but HLX-3 activity was not maintained in this population, correlating with the observation that sinus venosus-derived coronary vessels rely more on VEGFC signalling[237].

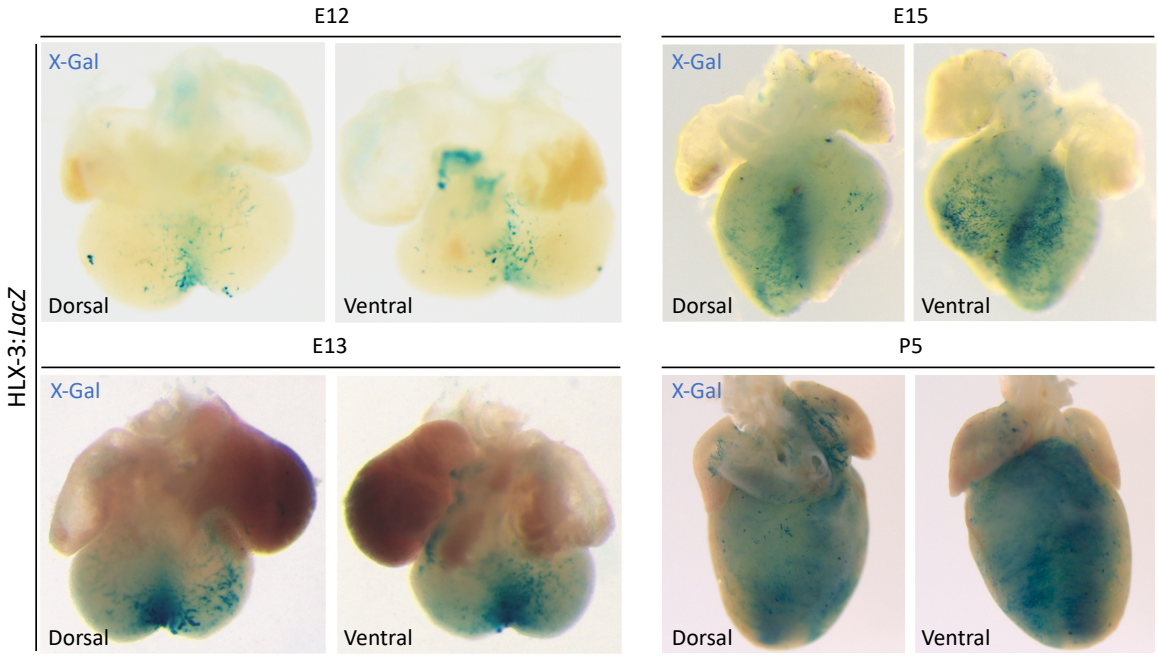


Figure 23: Representative images of HLX-3:LacZ transgene expression in the heart. *At E12 and E13, X-Gal expression was primarily limited to the base of the septum and to the dorsal site of the heart. At E15, X-Gal staining was focused on the ventral face and the septum. At P5, X-Gal expression was maintained in coronary vessels.*

3.4.1.3 HLX-3 is active during angiogenic sprouting in the hindbrain

In the previous section, I established that the HLX-3 enhancer can drive endothelial cell-specific reporter gene expression in the systemic vasculature of the developing mouse embryo, and that this expression correlates with regions undergoing angiogenic sprouting and VEGFA driven vessel growth in the heart. To further confirm that the HLX-3 enhancer was able to drive reporter gene expression, I next examined HLX-3:*LacZ* expression in the embryonic hindbrain, a known model for angiogenic sprouting[177].

The vascular network of the embryonic hindbrain is formed by angiogenesis between E10 and E13[177]. Vessels sprout from the perineural vascular plexus towards the ventricular side of the hindbrain. They then sprout laterally and anastomose to form the subventricular plexus. This sprouting is highly VEGFA-dependent, responding to a VEGF gradient secreted by neuroprogenitors and radiating from the midline[177]. To examine HLX-3:*LacZ* activity in

this model, E11 hindbrains were removed from fresh embryos, fixed and X-Gal stained and then examined as flatmounts. HLX-3 driven *LacZ* expression was seen in every actively sprouting vessel in the hindbrain (Figure 24). HLX-3 enhancer was active in endothelial cells that sprout from the perineural vascular plexus *via* the pial side of the hindbrain, as well as in the laterally sprouting cells throughout the subventricular plexus. Consistent with the end of angiogenesis and the onset of arteriovenous differentiation, HLX-3 enhancer activity in the hindbrain gradually diminished after E12.5 (data not shown). These results confirmed my previous observation that HLX-3 driven expression was restricted to actively angiogenic endothelial cells.

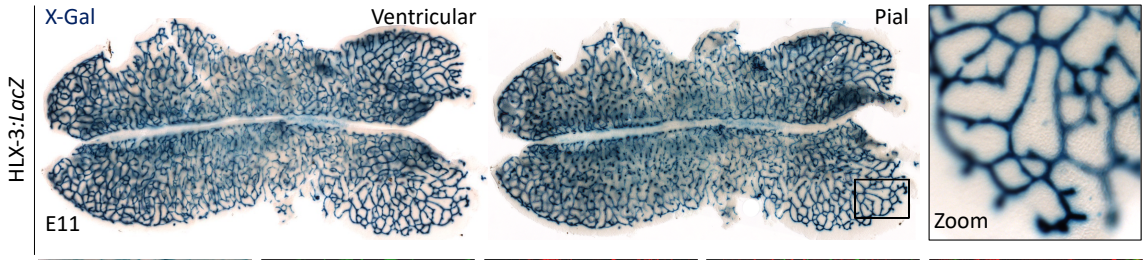


Figure 24: HLX-3:LacZ transgenic mice demonstrate enhancer activity in endothelial cells undergoing sprouting angiogenesis in the E11 embryonic hindbrain. *Representative image shows enhancer activity detected by X-Gal staining (blue). Flatmount hindbrains are shown from the ventricular and pial side.*

3.4.2 HLX-3:*LacZ* is active in selective endothelial cells in the post-natal mouse

3.4.2.1 HLX-3:*LacZ* positive endothelial cells in the growth plate during bone development

I next analysed the expression pattern of HLX-3:*LacZ* in post-natal mice. This investigation found that HLX-3:*LacZ* expression was generally downregulated in post-natal mice. The bladder, where vasculature is largely quiescent, showed very little HLX-3:*LacZ* vascular expression in the growing ureters (Figure 25A). HLX-3 enhancer activity was completely absent from the vasculature in the parietal peritoneum (Figure 25B). In the transgenic P5 rib vasculature, sprouting intercostal arteries showed strong X-Gal staining, while the thoracic artery had no HLX-3:*LacZ* activity. At P8, HLX-3 was active in lateral branches sprouting from the intercostal artery, but negative in the intercostal vein (Figure 25C).

Interestingly, while HLX-3 enhancer activity was limited in juvenile mice, it demonstrated strong activity at 4 weeks of age in the developing bone, specifically the metaphysis (Figure 25D). Recently, it was found that the bone vasculature contains a highly proliferative, angiogenic population located in the metaphysis[241]. In the metaphysis, capillaries form interconnected columnar tubes with blind-ended, filopodia-bearing protrusions towards the growth plate[242]. HLX-3:*LacZ* positive cells marked a sub-population at the most distal top of these columnar vessels (Figure 25D). In contrast, the quiescent sinusoidal capillaries in the diaphysis did not show any X-Gal staining.

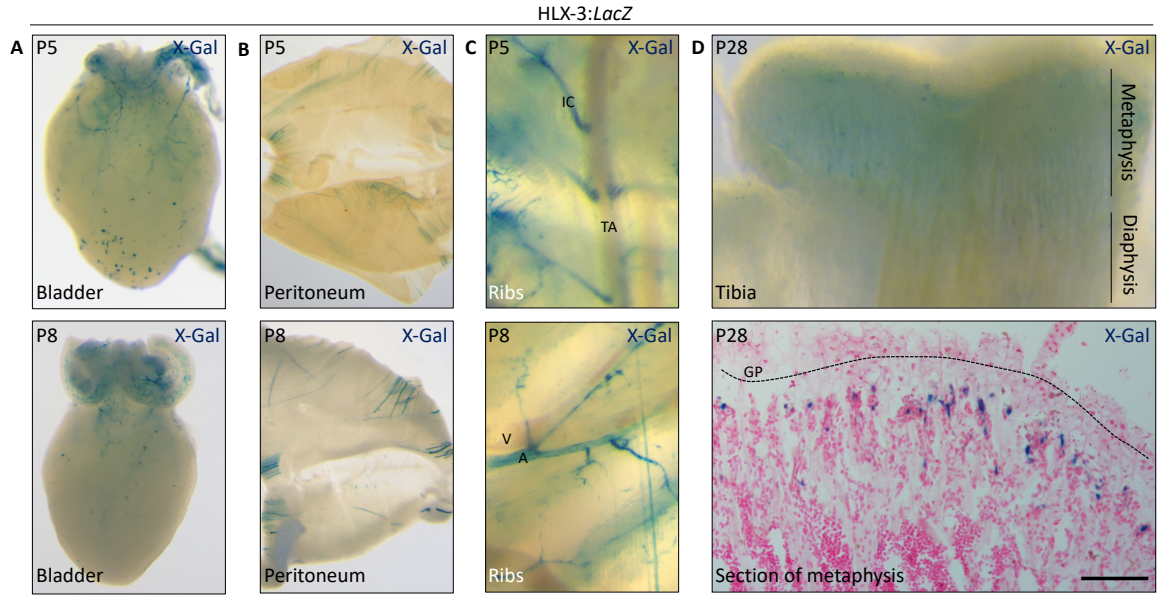


Figure 25: Representative images of HLX-3:LacZ transgene expression in the post-natal mouse. A) *HLX-3:LacZ was largely inactive in the bladder at P5 and P8. Only selective X-Gal positive cells were found in the growing ureters.* ***B)*** *HLX-3:LacZ reporter activity was completely downregulated in the parietal peritoneum at P5 and P8. Ectopic smooth muscle staining in the peritoneum might result from leaking hsp68 minimal promoter activity.* ***C)*** *At P5, sprouting intercostal vessels in the ribs were X-Gal positive, while the thoracic artery and vein were X-Gal negative. IC = intercostal vessels, TA = thoracic artery. In the ribs at P5, lateral branches sprouting from the intercostal arteries showed HLX-3 driven LacZ expression, while intercostal veins did not. A = artery, V = vein.* ***D)*** *Tibia (and fibula) from a juvenile mouse at P28 show X-Gal staining in the metaphysis, but not the diaphysis. Coronal section of the metaphysis at P28. Dashed line marks growth plate bearing chondrocytes. X-Gal positive cells are located at the most distal bulb of the arterioles. Scale bar corresponds to 100 µm.*

3.4.2.2 HLX-3:*LacZ* is active in veins and at the angiogenic front in the post-natal retina

The retina represents a classical and well-characterized model for sprouting angiogenesis[100]. The retina is the only place in the pup that vascularizes completely after birth through angiogenic sprouting[243]. From the optic disc, blood vessels grow radially on the surface of a pre-existing plexus of astrocytes into the periphery. Retinal angiogenesis is very much regulated by the demand of neurons. Increased neuronal activity causes hypoxia-dependent secretion of VEGFA by astrocytes[100]. During the first stage of vascularization the superficial primary plexus is formed (P0 - P8). Once the periphery is reached, endothelial cells sprout perpendicular into the deeper layers. With the formation of the secondary and tertiary plexuses (P8 - P25), the retinal vasculature switches from a two- into a three-dimensional vascular network[244]. First, HLX-3:*LacZ* expression was analysed in the two-dimensional stage of sprouting during formation of the primary plexus.

Upon dissection, the retina was cut into three or four leaves and flattened[245]. The first image of Figure 26A shows one leaf of the flatmount retina at P4 labelled with blood vessel marker Isolectin B4. The second image shows the HLX-3:*LacZ* retina stained with X-Gal in blue and the retina pigment epithelium in brown in the background. To allow for better visualization of the X-Gal positive vessels, transmission images of retinas were spectrally unmixed and false-coloured. All subsequent images show the pan-endothelial Isolectin B4 in red and X-Gal staining in green, although it should be noted that merged images can underestimate the X-Gal staining due to the image processing. At P4, HLX-3 driven X-Gal expression was detected in most, but not all, endothelial cells. This correlates with the highly angiogenic nature of this vascular bed at this stage.

In the P5 retina, the developing arteries and veins can be identified by morphology (the developing arteries have fewer branches than veins). Figure 26B shows the entire P5 retina, followed by a zoom of one leaf with its developing capillary free artery (A) and capillary dense veins (V) clearly distinguishable. The third image is a zoomed image showing tip and stalk cells of the angiogenic front. Expression of the HLX-3:*LacZ* transgene becomes increasingly specified to venous endothelial cells and the angiogenic front (AF) by this time point. Tip and stalk cells within the angiogenic front were identified both by location and by the DLL4 antibody (Figure 26C), which is expressed specifically in the developing arteries and tip cells[77].

At P6, X-Gal staining in veins and their branched capillaries became more pronounced, while X-Gal staining faded progressively from developing arteries (Figure 26D). Even though HLX-3:*LacZ* activity was prominent in the angiogenic front, many individual filopodia-rich tip cells did not display robust reporter gene activity. This expression pattern was maintained, and even more distinct, in the P7 retina (Figure 26E). Here, the HLX-3:*LacZ* transgene was mainly expressed in proliferating stalk cells beneath migrating filopodia-bearing tip cells. HLX-3 was still strongly active in veins and the dense peri-venous capillaries.

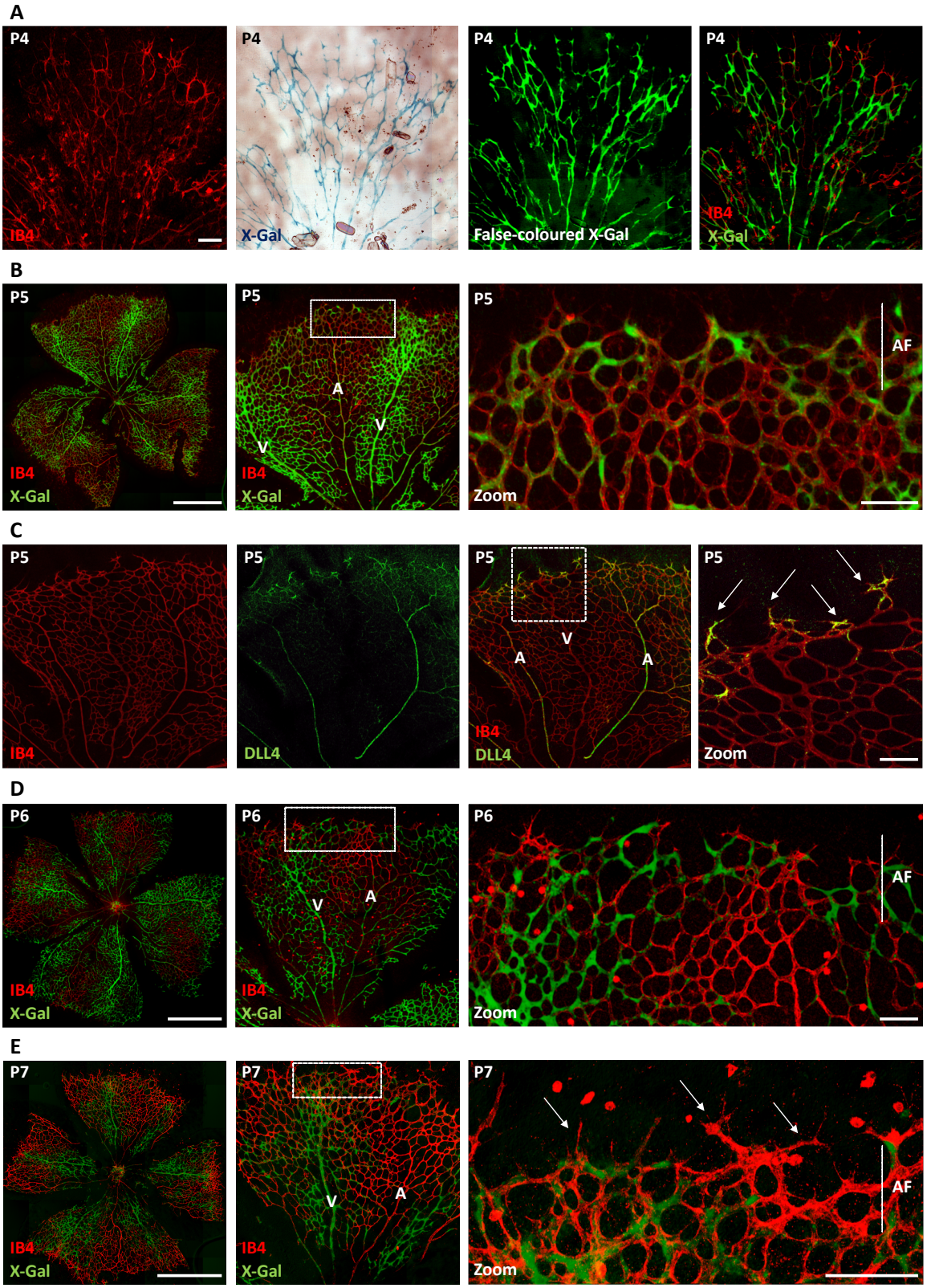


Figure 26: HLX-3:LacZ mice demonstrate X-Gal activity in veins (V) and at the angiogenic front (AF), but not arteries (A) in the post-natal retina. *Representative images from of the retina at post-natal days 4, 5, 6 and 7 are shown. Isolectin B4 (IB4) staining marks all endothelial cells. Occasionally, IB4 staining also occurred in speckles outside the vasculature, which could correspond to galactose epitopes on macrophages also detected by IB4*[246]*.* ***A)*** *The first image shows IB4 staining in red at P4. Note that preceding X-Gal staining partially masked IB4 staining. X-Gal staining is shown in blue (second image) and false-coloured green (third image). HLX-3:LacZ expression appeared patchy (overlay of X-Gal with*

IB4 in fourth image). Scale bar is equal to 100 µm. ***B)*** *At P5, strong HLX-3 driven X-Gal staining was observed in veins and at the angiogenic front. Staining in differentiating arteries appeared weaker.* ***C)*** *In the P5 retina, anti-DLL4 antibody detects DLL4 in tip cells and arteries. IB4 staining in red detects all endothelial cells. A zoom of the angiogenic front of the merged image shows DLL4 staining in tip cells (white arrow, scale bar = 100 µm).* ***D)*** *At P6, HLX-3 driven LacZ expression corresponded with the radial pattern of alternating veins and arteries. At the angiogenic front, HLX-3 became progressively specified to stalk cells.* ***E)*** *At P7, X-Gal staining was less pronounced in filopodia-rich tip cells (white arrows). Veins remained X-Gal positive, while arteries were completely X-Gal negative. For panels B), D), E): First image of each panel shows whole flatmount retina. Scale bars indicate 1 mm. The middle image shows the magnified view on one leaf. Dotted boxes indicate regions of zoom. The third image shows a magnified view of the angiogenic front. Here, scale bars correspond to 100 µm.*

At P8, the primary vascular plexus was formed and had reached the perimeter of the retina. With the decrease in proliferation and migration at the angiogenic front, HLX-3 driven *LacZ* expression faded (Figure 27A). HLX-3:*LacZ* was only expressed in the dense capillary network surrounding veins and was completely absent from the periphery of the retina and the former angiogenic front. At P15, when retinas became more developed, HLX-3:*LacZ* expression was downregulated and venous-specific expression was lost. Strong X-Gal staining was only observed in the capillary network in the proximity of veins (Figure 27B). Endothelial cells surrounding veins sprout into the deeper layer of the retina to form the secondary plexus[247]. At P18, during the final steps of retinal vascularization, HLX-3 driven *LacZ* expression was very sparse, although individual endothelial cells in peri-venous capillaries showed strong X-Gal staining (Figure 27C). Detailed investigation of the nature of these X-Gal positive cells which are penetrating the deeper layer was hindered by the limitations of X-Gal staining. In contrast to immunofluorescence staining and confocal microscopy, X-Gal staining does not allow 3D rendering.

Overall, these observations suggest that the HLX-3 enhancer was active in the angiogenic front, in venous and peri-venous endothelial cells during sprouting angiogenesis in the post-natal retina. When vascularization of the retina was nearly complete, HLX-3 activity decreased. Combined with the observations in transgenic mouse embryos, these data

suggest that the angiogenic-specific expression pattern of the hlx-3 enhancer observed in zebrafish was similar to the HLX-3 enhancer in the mouse.

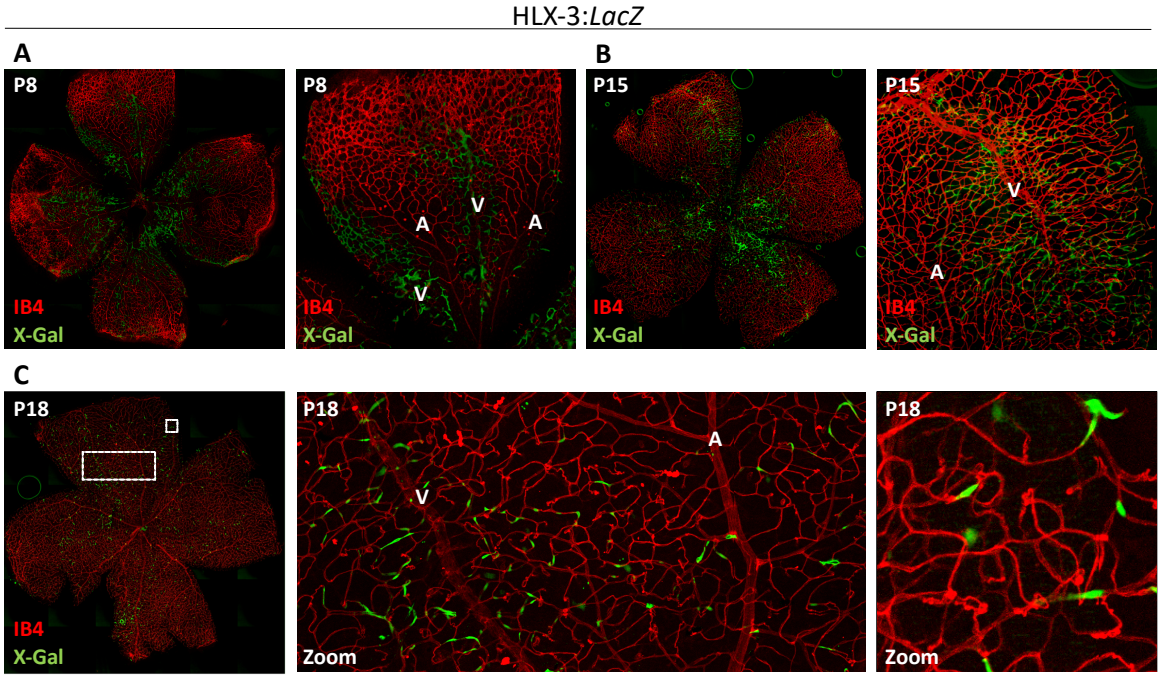


Figure 27: Representative images from HLX-3:LacZ transgenic retinas at post-natal days 8, 15 and 18 show overall reduction of HLX-3 activity*. During secondary and tertiary sprouting, only endothelial cells surrounding veins remained X-Gal positive. Isolectin B4 (IB4) staining marked all endothelial cells. HLX-3 driven LacZ activity was detected by X-Gal staining (false-coloured in green).* ***A)*** *At P8, the superficial primary plexus was formed and there was no radially extending angiogenic front. While HLX-3 was strongly expressed in the dense capillary network surrounding veins, it was absent from arteries and the periphery of the retina.* ***B)*** *At P15, overall X-Gal staining was reduced with strong X-Gal staining in the capillary network in the proximity of veins.* ***C)*** *At P18, IB4 staining shows the depth of the nearly completed vascular network. HLX-3:LacZ was only active in a subset of endothelial cells surrounding veins. V = veins. A = arteries. Dotted box indicates region of magnified view.*

3.4.3 Endogenous *Hlx* in the mouse

So far, no conclusive expression pattern has been shown for *Hlx* expression in mouse, leaving unanswered whether activity of the HLX-3 enhancer correlates with *Hlx* RNA or HLX protein expression. In mouse embryos, expression of *Hlx* has only been studied by *in situ* hybridization. In this analysis, *Hlx* was detected in liver and intestines at early embryonic time points, and in the angiogenic sprout at P5[161]. However, the quality of these *in situs* was relatively poor. I made numerous attempts to repeat this analysis with

several *in situ* probes, but did not succeed (e.g. Figure 28A and B). Likewise, my attempts to perform immunohistochemistry analysis to detect HLX expression in mouse embryos was unsuccessful as I was only able to detect HLX in astrocytes and neurons (e.g. Figure 28C and D). I was therefore unable to confirm that human HLX-3 enhancer drives expression in the mouse in a manner that is recapitulating endogenous *Hlx* expression. Transcription factors are frequently expressed at very low levels[248], HLX expression levels might thus have been too low to detect endogenous HLX in endothelial cells. Nonetheless, my analysis of the HLX-3 enhancer confirms that it is active during angiogenesis and quiescent in the differentiated vasculature, similar to the pattern of endogenous *Hlx* reported by Prahst *et al.*[161]. This angiogenic-specific expression pattern in mouse, combined with the expression pattern relative to endogenous *hlx1* in zebrafish, makes it a suitable enhancer to study angiogenic regulation.

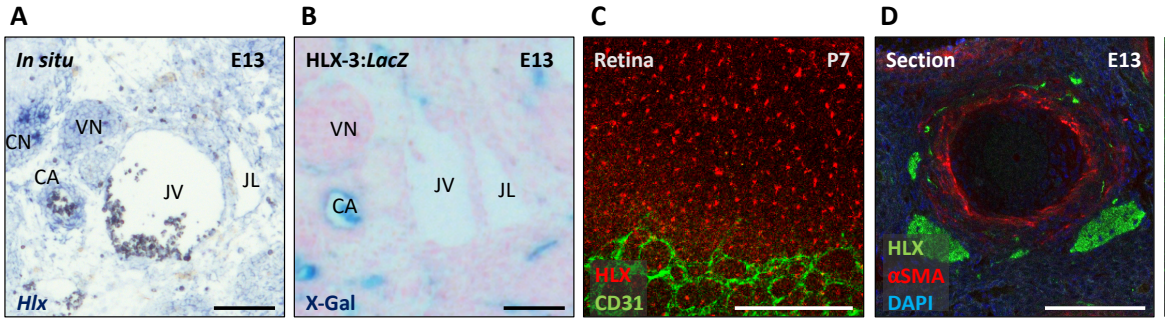


Figure 28: Attempts to detect endogenous Hlx expression in the mouse. A) *Hlx in situ on a transverse section of an E13 embryo. Hlx was detected in the vagal nerve (VN), cervical nerve (CN) and the cells surrounding the carotid artery (CA). Hlx was absent from the jugular vein (JV) and the jugular lymph sac (JL).* ***B)*** *Transverse section of a HLX-3:LacZ E13 embryo stained with X-Gal for comparison. HLX-3 drove LacZ expression in the carotid artery and the microvasculature, but was absent in the nervous, venous and lymphatic compartment.* ***C)*** *Post-natal retina at P7 stained with pan-endothelial marker anti-CD31 (green) and anti-HLX antibody (red). HLX was only detected in astrocytes in the the retina. No endothelial-specific staining could be detected.* ***D)*** *Transverse cryosection of an E13 embryo was stained for vascular smooth muscle marker αSMA (in red) and HLX (in green) and DAPI (blue). HLX staining was detected in phrenic nerves surrounding the αSMA positive artery. Antibody stained section and retina whole-mount were imaged using the emission fingerprinting function of the Zeiss 710MP confocal microscope to eliminate red blood cell autofluorescence. All scale bars correspond to 100 µm.*

3.5 Discussion

3.5.1 Role of the HLX and JAG1 putative enhancers

The first aim of my project was to find an enhancer that is exclusively expressed in endothelial cells during angiogenesis. Although all putative enhancers tested here shared HUVEC-specific H3K27ac or H3K4me1, only the HLX-3 enhancer was able to consistently drive endothelial reporter gene activity in both mice and zebrafish. Interestingly, this HLX-3 enhancer region was previously independently identified by a different research group as a putative enhancer based on its high degree of conservation across species, but it was not tested *in vitro* or *vivo*[249].

After my analysis of this enhancer was completed, a new feature appeared on the UCSC browser called GeneHancer. This feature annotates regulatory elements and predicts gene-enhancer interactions[28]. The negative putative enhancers JAGin5, JAGin20 and JAG+46 were not annotated as enhancers, while JAGin2 was annotated as part of an enhancer/promoter region. Although the JAGin2 construct was unable to drive reporter gene activity in transgenic zebrafish in my analysis, it is therefore possible that a larger region including both JAGin2 and the *JAG1* endogenous promoter would be able to drive sprouting-specific expression. GeneHancer also predicted that the HLX-35 enhancer (which in my analysis drove GFP expression in zebrafish but not in mouse) interacted with the TSS of *HLX*, while the negative HLX-5 was only shown to interact with HLX-35 and the TSS of a long non-coding RNA. It could therefore be speculated that HLX-5 and HLX-35 may have an important role for 3D architecture of the *HLX* locus, but do not direct transcription independently in endothelial cells during development. Interestingly, GeneHancer gave

HLX-3 the highest enhancer score, although it was not predicted to loop with the TSS. While this *in silico* prediction has not been validated in endothelial cells, it is therefore possible that the HLX-3 enhancer may act *via* the linking or tracking model[66]. Supporting this theory, GeneHancer annotated a non-coding RNA at the HLX-3 locus as an eRNA[28]. In the tracking model, RNA Pol II moves from the proximal enhancer to the promoter which produces eRNA as a by-product.

3.5.2 The HLX-3/hlx-3 enhancers drive reporter gene expression in angiogenic sprouting endothelial cells

In zebrafish, the hlx-3:GFP transgene was expressed in actively proliferating intersegmental vessels, the head vasculature and in angiogenic vessels during the later vascularization of organs. Until 2 dpf, hlx-3 driven GFP expression was entirely specific to endothelial cells within the sprouting ISVs, strongly indicating that all hlx-3 driven expression was in angiogenic endothelial cells, although this was more challenging to establish in other vascular beds. To my knowledge, the *tg(hlx-3:GFP)* zebrafish line is the first-ever described zebrafish line in which angiogenic endothelial cells are specifically labelled[165].

Within the early ISVs, GFP expression was observed in both stalk and tip cells with varying intensities depending on the ISV examined, likely correlating with the mitotic activity of the endothelial cell[226], although the half-life of GFP (approximately 26 hours) makes it challenging to prove this experimentally. Angiogenic sprouting from the dorsal aorta is already known to be mediated by Vegf/Notch signalling, suggesting a potential role for these pathways in activating the hlx-3 enhancer[250,251,252]. This is supported by the later

expression patterns of this enhancer: hlx-3:GFP was also active when endothelial cells in the DLAV interconnected with the opposite cell, a process known to be dependent on Vegfa signalling[253]. Conversely, the hlx-3:GFP transgene was not active during the Vegfa-independent maturation of ISVs, and during venous caudal angiogenic sprouting, a process known to be dependent on Vegfc, not Vegfa[230].

Endogenous *hlx1* expression in zebrafish predominantly correlated with hlx-3 enhancer driven GFP expression, although hlx-3:GFP expression was more pronounced than *hlx1* in anastomosing vessels which will form the DLAV[164]. GFP expression in these cells might result from actual enhancer activity or might be remnants of prior hlx-3:GFP expression due to the stability of GFP. It is also possible that synthesis and degradation of reporter GFP was lagging behind that of endogenous *hlx1*. Likewise, GFP protein from hlx-3 positive stalk cells could have been carried over to daughter cells. Nonetheless, the expression pattern of the hlx-3:GFP transgene and its proximity to *hlx1* gene strongly indicate that hlx-3 likely regulates *hlx1* expression during angiogenic sprouting.

During murine embryonic development, the HLX-3 enhancer was able to drive spatially and temporally restricted *LacZ* expression in the developing microvascular cells in most vascular beds examined. Further HLX-3:*LacZ* expression was also observed in the dorsal aortae of early embryos (E8 - E11), in the sprouting intersegmental arteries, and in the angiogenic vascular beds of the hindbrain. During mid-gestation, X-Gal staining in the dorsal aortae was polarized towards the midline. This coincided with migration of dorsal aortae towards the VEGFA gradient secreted by the notochord[254]. Dorsal aortae fusion into a single vessel is thought to be mediated by downregulation of BMP antagonist

Chordin allowing endothelial cells to respond to Sonic hedgehog (SHH) and VEGFA midline-positioning signal[255,256]. This agrees with an angiogenic sprout-specific role for the HLX-3 enhancer, since angiogenesis and endogenous *Hlx* expression require VEGFA signalling[154,164]. Interestingly, while the expression of HLX-3:*LacZ* continued in the microvasculature throughout development, the activity of this transgene stopped in the dorsal aorta after E12. Arterial endothelial development has recently been closely linked with angiogenesis in both the post-natal retina and regenerating fin tissue in fish, where studies have demonstrated that angiogenic endothelial cells migrate to form the early arteries[257,258]. Therefore, early arterial activity of HLX-3 may either directly relate to its angiogenic activity, or indicate a second domain of activity, although the latter explanation seems less likely in light of the absence of arterial activity in zebrafish.

It is unclear why some endothelial cells in the cardinal veins also expressed HLX-3:*LacZ* at the early embryonic time points. This could reflect HLX-3 activity in differentiating lymphangiogenic cells which are about to sprout[259,260]. Alternatively, these could be endothelial cells leaving the venous lineage to participate in future angiogenic behaviour: previous studies in zebrafish fin and post-natal mouse retina both suggest that angiogenic endothelial cells are venous in origin[257,261] .

HLX-3:*LacZ* transgene activity was also found at the vascular front in the post-natal retina. These results further suggest that the specificity of hlx-3/HLX-3 activity to angiogenic endothelial cells may be responsive to VEGFA signalling. However, while HLX-3:*LacZ* expression could be observed in tip cells at the angiogenic front of P5 and P6 retinas, by P7 expression was increasingly restricted to only stalk cells at the angiogenic front. It is

not clear whether this initial reporter activity in tip cells was caused by active transgene transcription or by carry-over of β-galactosidase from its earlier transgene expression in stalk cells, or whether the lack of staining in tip cells is a result of a time-lag in β-galactosidase expression and build-up. However, consistent with my results, Prahst *et al.* also detected *Hlx* by *in situ* in some, but not all, sprouting tip cells in the retina[161], suggesting that the HLX-3:*LacZ* expression pattern was largely consistent with endogenous *Hlx*.

Interestingly, HLX-3:*LacZ* expression in the retina was also found in venous endothelial cells, and therefore could be seen as exactly matching that of actively proliferating cells in this vascular bed. HLX-3:*LacZ* expression in veins in the post-natal retina and HLX-3:*LacZ* expression in the arteries in the systemic vasculature might seem contradictive. But a recent study has shown that highly proliferative venous endothelial cells contribute to cells in angiogenic front[261]. Moreover, it has been demonstrated that pro-angiogenic VEGFA was specifically expressed in the plexus surrounding veins[100]. A study performed in 1948 has already postulated that the capillary network in the cat retina develops from veins[262].

To summarize, my data suggest that HLX-3 enhancer was specifically transcribed in a highly proliferative, but not migrating, subset of angiogenic cells. In the retina, the most proliferative cells are stalk cells, followed by venous endothelial cells[100]. In both regions I observed especially strong HLX-3 activity. In concordance with this observation, it has previously been hypothesized that HLX enhances proliferation[160] by repressing $p16^{INK4a}$, a

cell cycle inhibitor[263]. Additionally, HLX reduces migration by activating repulsive cell-guidance molecule UNC5B[264].

3.5.3 Conclusion

In conclusion, this descriptive analysis of enhancer activity in transgenic models shows that the evolutionary conserved human HLX-3 and zebrafish hlx-3 enhancers are specifically transcribed in newly formed vessels and angiogenic endothelial cells, but not in the mature quiescent vasculature. Further, the hlx-3/HLX-3 enhancers predominantly drive reporter gene activity in highly proliferative endothelial cells known to depend on, or respond to, VEGFA signalling in both mouse and zebrafish (including zebrafish ISVs, the mouse polarized dorsal aortae, endocardial-derived coronary vessels in the heart, and angiogenic endothelial cells in the mouse hindbrain, tibia and retina). This therefore suggests that sequences found within both the HLX-3 and hlx-3 enhancers are capable of receiving information from VEGFA signalling and translating this into gene activity specifically during angiogenesis. Therefore, in the next chapter, the HLX-3 and hlx-3 enhancers, and cognate HLX-3:*LacZ* and hlx-3:GFP transgenic lines will be used to pinpoint the key transcriptional player(s) regulating angiogenic sprouting.

4 Regulation of the HLX enhancer during sprouting angiogenesis

4.1 Rationale

4.1.1 Aim: To identify the upstream regulators controlling HLX-3 activity in angiogenic endothelial cells

In the previous chapter, I described novel gene enhancers within the mammalian *HLX* and zebrafish *hlx1* loci, and demonstrated that these enhancers are specifically active in endothelial cells during angiogenesis. Here, I aim to identify the upstream regulators of the hlx-3/HLX-3 enhancer. This will include investigating the role of pro-angiogenic VEGFA signalling in the regulation of the hlx-3/HLX-3 enhancer, identifying the crucial transcription factor binding motifs within the enhancer sequence required for angiogenic activity, and confirming a role for the cognate binding factors in angiogenic sprouting.

4.2 HLX-3 enhancer activity is downstream of VEGF signalling

4.2.1 HLX-3:*LacZ* transgene activity during neo-angiogenic growth in adult mice

After establishing the HLX-3:*LacZ* transgene activity pattern during physiological angiogenesis in development, I next examined neo-vascular growth during pathological conditions, such as cancer[71]. Similarly to developmental vascular growth, tumour blood vessels grow *via* angiogenesis. To determine if my HLX-3 enhancer is also active during this process, the expression of the HLX-3:*LacZ* transgene was analysed in adult mice during neo-angiogenesis in Matrigel plugs (mediated by bFGF[265]) and during neo-angiogenesis in subcutaneous B16F10 melanoma allografts[266] (mediated primarily by VEGFA[267]).

The HLX-3 enhancer was unable to direct robust reporter gene expression during FGF-stimulated neo-vascularization in Matrigel plugs. Even though the angiogenic front protruding into Matrigel plugs was clearly visible, the majority of sprouts did not co-localize with HLX-3:*LacZ* expression (Figure 29A). Conversely, HLX-3:*LacZ* was robustly expressed in actively sprouting endothelial cells in B16F10 melanoma allografts, although HLX-3 driven *LacZ* expression was not detected in every endothelial cell (Figure 29B). Co-staining with CD31 revealed that 60.4% of the visible vascular network was X-Gal positive in B16F10 tumour allografts (Figure 29C), while only 12.5% of CD31 positive vessels showed an X-Gal positive cell in bFGF-containing Matrigel plugs (Figure 29D). The otherwise quiescent vasculature of the adult mouse was X-Gal negative (data not shown).

These results demonstrate that the HLX-3 enhancer is more consistently active in neo-angiogenesis in response to VEGFA than by bFGF.

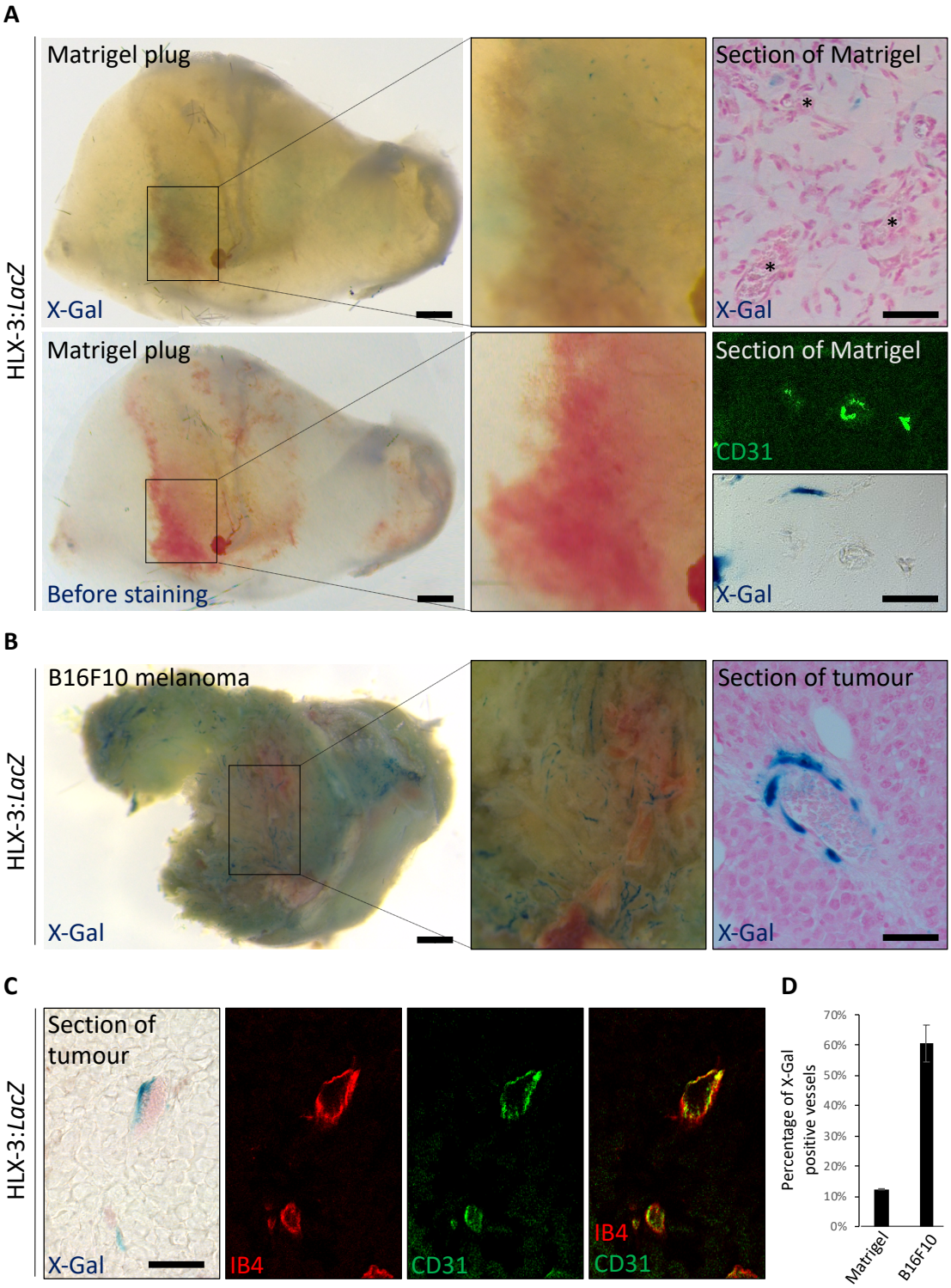


Figure 29: The HLX-3:LacZ transgene is active during melanoma-induced, but less active during FGF-mediated adult neo-vascularization. A) *Representative image of HLX-3:LacZ transgene activity during neo-vascularization into Matrigel plugs after nine days before and after X-Gal staining (n = 16). Nuclear fast red stained section of the Matrigel plug shows*

*X-Gal negative blood vessels (denoted by *). Immunostaining on sectioned X-Gal-stained bFGF-containing Matrigel for CD31 (green).* ***B)*** *Representative image of HLX-3:LacZ transgene activity in B16F10 melanoma tumours after eleven days (n = 8). B16F10 melanoma tumours show robust vascular X-Gal staining.* ***C)*** *Immunostaining on cryosectioned X-Gal-stained tumour with pan-endothelial markers IB4 (red) and CD31 (green).* ***D)*** *Quantification of X-Gal positive blood vessels identified with pan-endothelial marker CD31 in Matrigel plugs and B16F10 allografts (n = 2, each). Error bars indicate standard error of the mean. Scale bars for whole-mounts correspond to 1 mm. Scale bars for sections and cryosections correspond to 50 µm.*

4.2.2 Effect of inhibition of known angiogenic signalling pathways on hlx-3:GFP activity pattern in transgenic zebrafish

Developmental angiogenesis in the zebrafish ISVs, mouse embryonic hindbrain and mouse post-natal retina are all known to be VEGFA-dependent. Given that the hlx-3/HLX-3 enhancers are active during each of these processes, as well as during melanoma cell-induced pathological angiogenesis, my data so far strongly indicate that the hlx-3/HLX-3 enhancer is activated by VEGFA signalling.

To further test this hypothesis, I next investigated the activity of the hlx-3:GFP transgene in zebrafish in which VegfA signalling has been inhibited. To do this, I used SU5416 (Semaxanib), a selective Vegfr tyrosine kinase inhibitor, which does not have an effect on Egf receptor (Egfr) or Fgf receptor (Fgfr) tyrosine kinases[268]. Therefore, SU5416 selectively inhibits all Vegf-mediated signalling and transcription. Hlx-3:GFP was expressed as expected in sprouting ISVs in control DMSO treated *tg(hlx-3:GFP; kdrl:HRAS-mCherry)* zebrafish (Figure 30A), while partial inhibition of Vegf signalling using 0.63 µM SU5416 treatment resulted in substantially reduced hlx-3:GFP expression, although some sprouting ISVs could still be seen (Figure 30B, ISVs are marked by pan-endothelial kdrl:HRAS-mCherry expression). At greater concentrations of Vegf signalling inhibitor, there was a near total loss of both hlx-3:GFP transgene activity and ISV sprouting (Figure 30C). These results suggest that the hlx-3 enhancer is not transcriptionally activated when

Vegf signalling is inhibited. Similar results have been reported for endogenous *hlx1*, with Herbert *et al.* reporting that addition of the Vegf signalling inhibitor SU5416 decreased *hlx1* expression in wild-type zebrafish at 30 hpf[164]. Of course, my experiments cannot evaluate whether the effect of Vegf signalling on the hlx-3 enhancer is direct or indirect.

During sprouting angiogenesis, VegfA signalling interacts with Notch signalling at several levels. In tip cells, Vegf signalling induces the expression of Notch ligand *Dll4* leading to the activation of Notch signalling in the neighbouring stalk cell. There, activated Notch acts to inhibit expression of pro-angiogenic *Vegfr2/Kdr* and *Vegfr3/Flt4,* which encode the VEGF receptors VEGFR2 and VEGFR3. Consequently, inhibition of Notch in both zebrafish embryos and mouse post-natal retina results in increased angiogenic sprouting and endothelial cell hyperproliferation with excessive filopodia formation[164,269,270]. To establish whether my hlx-3 enhancer also responds to alterations in Notch signalling, I next investigated whether reduced Notch signalling affected hlx-3 activity in transgenic zebrafish. To do this, I blocked Notch signalling by inhibiting NICD with the γ-secretase inhibitor DAPM, added to the fish water in combination with DMSO, similar to the SU5416 inhibitor. In line with the increased angiogenic sprouting seen, hlx-3 driven GFP expression was robustly detected in the excessively and aberrantly sprouting ISVs. Similar results have been reported for endogenous *Hlx*/*hlx1* in both mice and zebrafish: in *Dll4*$^{-/+}$ post-natal mouse retinal cells, *Hlx* mRNA is upregulated[114], while zebrafish treated with Notch signalling inhibitor increased endogenous *hlx1*[164].

In conclusion, my results suggest that *hlx-3:GFP* expression, similar to endogenous *hlx1*, is strongly correlated with sprouting behaviour in endothelial cells, is repressed by Notch signalling and is activated by Vegf signalling[164,114].

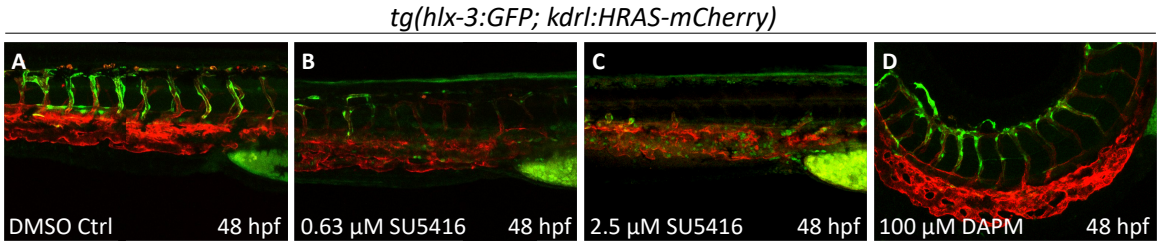


Figure 30: Effect of chemical inhibition on tg(hlx-3:GFP; kdrl:HRAS-mCherry) at 48 hpf. A) *In the DMSO treated control (Ctrl), hlx-3:GFP labels sprouting ISVs (green) and kdrl:HRAS-mCherry all endothelial cells (red).* ***B)*** *0.63 µM of the Vegf signalling inhibitor SU5416 reduced enhancer activity and angiogenic sprouting.* ***C)*** *2.5 µM SU5416 ablated enhancer activity and ISV formation.* ***D)*** *100 µM DAPM inhibited Notch signalling and lead to an increase in hlx-3 reporter activity in ectopic ISVs. DAPM caused dorsalization, irregular branching, hyperproliferation and excessive sprouting. All zebrafish were treated at 22 hpf with the indicated concentrations. Images are representative for n = 40 in all conditions.*

4.2.3 Flow does not regulate *hlx1* and its enhancer hlx-3:GFP in zebrafish

In the previous section, I have shown that the hlx-3 activity in transgenic zebrafish requires intact Vegf signalling. I next investigated whether blood flow itself also regulates the activity of this enhancer. Previous work has suggested that flow is required for sprouting of the aortic arches *via* activation of Vegf signalling in zebrafish[271] (a location where hlx-3:GFP was strongly expressed), and shear stress is known to activate Vegfr2 in some circumstances[272]. However, the hlx-3:GFP transgene is robustly active in ISVs at early time points where flow is not yet established[273]. To establish if blood flow plays any role in hlx-3 activity, I examined whether the hlx-3 enhancer was able to drive GFP expression in the absence of blood circulation.

Blow flow was ablated in zebrafish by depletion of cardiac troponin t2 (*tnnt2*) using morpholino (MO) oligonucleotides, a process known to induce cardiac arrest but not

embryonic death at early developmental time points[181]. At 28 hpf, expression of hlx-3:GFP in ISVs was reduced in *tnnt2* MO treated embryos (*tnnt2* morphants) (Figure 31A). However, by 3 dpf all *tnnt2* morphant embryos expressed hlx-3:GFP within the ISVs at levels comparable to control MO treated embryos (Figure 31B), suggesting that earlier decreased hlx-3 enhancer activity was due to a developmental delay caused by *tnnt2* MO. Likewise, hlx-3:GFP expression was absent within the sub-intestinal plexus at 3 dpf, but this was also likely to be secondary to reduced growth of the sub-intestinal plexus in embryos lacking circulation, which has been described previously[274] but remains controversial[275,276].

Expression of endogenous *hlx1* mRNA in zebrafish was also unchanged in *tnnt2* morphants (Figure 31C). Further, in mammalian cell culture, endogenous *HLX* was also unaffected by shear stress while shear stress-responsive Krüppel-like factor 2 (*KLF2*)[277] was upregulated as expected (Figure 31D).

In conclusion, my results clearly indicate that neither shear stress nor the ablation of blood flow regulate the hlx-3 enhancer nor endogenous *hlx1* or *HLX* expression. This also rules out the possibility that loss of hlx-3:GFP expression in SU5416 treated embryos was secondary to alterations in blood flow due to malformed vasculature.

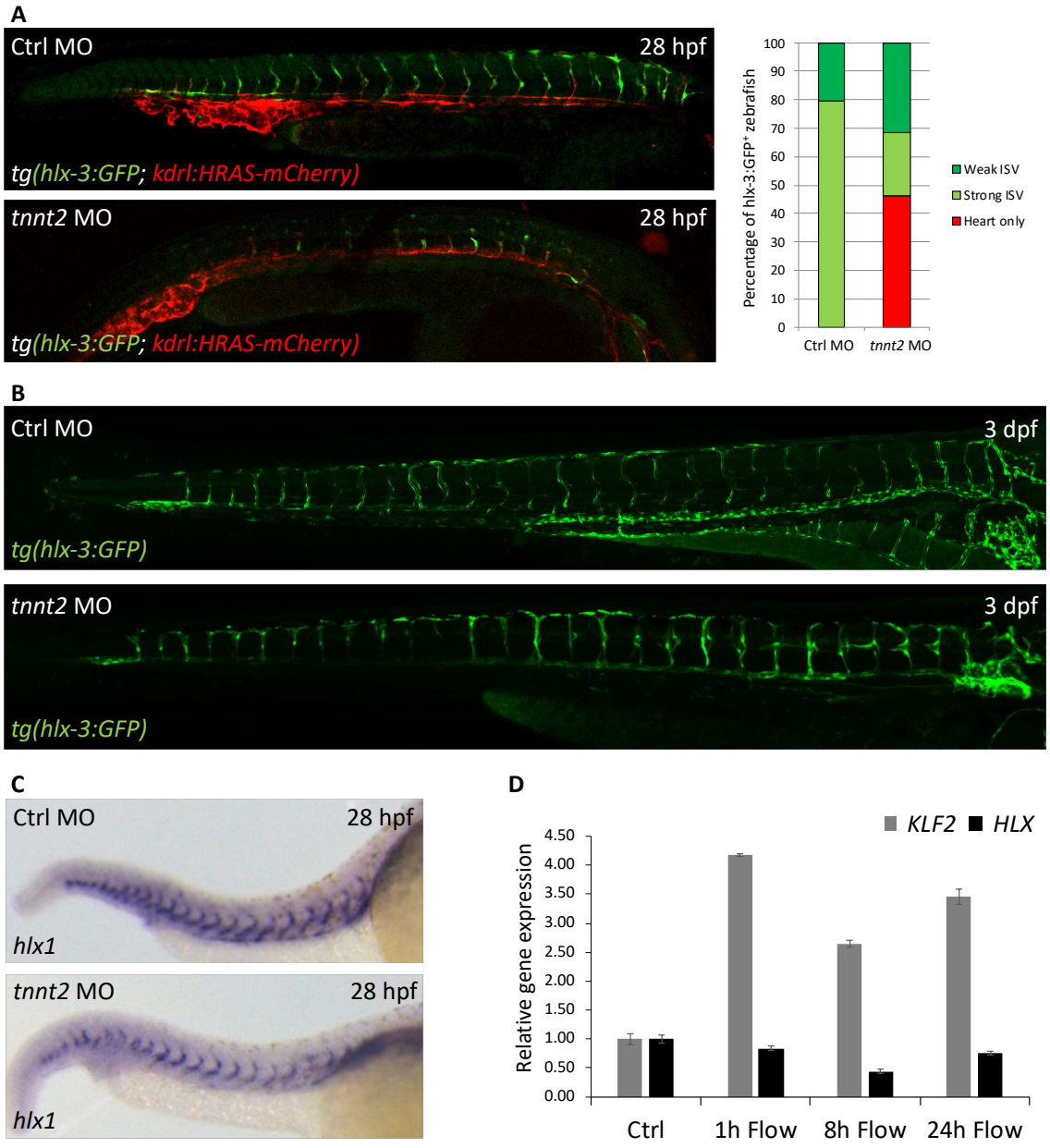


Figure 31: Expression of hlx-3:GFP in zebrafish after MO-mediated depletion of tnnt2 and HLX expression after shear stress in HUVECs. A) *tg(hlx-3:GFP; kdrl:HRAS-mCherry) treated with 4 ng tnnt2 MO (n = 107) or control (Ctrl) MO (n = 108) at 28 hpf. 80% of injected fish showed reporter gene expression. Graph depicts observed expression pattern in zebrafish showing any detectable GFP expression for control or tnnt2 MO injected embryos: light green represents strong GFP expression in ISVs, dark green represents weak GFP expression in ISVs, red represents expression in the heart only. tg(hlx-3:GFP; kdrl:HRAS-mCherry) treated with tnnt2 MO is representative for category "weak ISV".* ***B)*** *Representative tg(hlx-3:GFP) treated with either tnnt2 MO or control MO at 3 dpf.* ***C)*** *Expression of hlx1 as detected by whole-mount in situ hybridization after MO-mediated tnnt2 depletion (n = 106) or control MO (n = 50).* ***D)*** *qRT-PCR analysis shows increase in KLF2 but not in HLX after shear stress (11.72 dynes/cm²) in HUVECs. Expression levels were normalized to GAPDH. Non-stimulated control was set as reference for each time point separately. Error bars indicate standard deviation for three technical replicates.*

4.3 Identification of the transcription factors required for HLX-3/hlx-3 activity during angiogenesis

4.3.1 *In silico* analysis of the functionally conserved sequence

Both experiments in this thesis and published have demonstrated that endogenous *HLX/hlx1*, the HLX-3/hlx-3 enhancers and, more generally, angiogenesis are all downstream of VEGFA/VegfA stimulation. I next wanted to investigate the link between this inductive signal and gene expression, as it is currently unclear how VEGFA signalling is able to transcriptionally regulate gene expression during angiogenesis. To investigate this process, I first analysed the enhancer sequence in order to establish which transcription factor binding sites are responsible for the expression of the enhancer during angiogenesis.

Both the orthologous zebrafish and the human sequences of HLX-3/hlx-3 were able to drive angiogenic-specific expression, and were both able to direct expression equally in either transgenic mouse and zebrafish models. Therefore, I considered it highly likely that the crucial transcriptional binding motifs required for activation of these enhancers during angiogenesis would be conserved across species. To identify conserved transcription factor binding sites, the human, mouse, opossum, chicken and zebrafish sequences orthologous to the HLX-3 enhancer were aligned with ClustalW[278] (Figure 32A). Sequence comparison across species revealed five highly conserved regions (labelled Con1 to Con5 and marked in red). Separate alignment of the hlx-3 zebrafish sequence with other teleost (tetrodon, fugu, medaka and stickleback) hlx-3 sequences with ClustalW[278] revealed ten

highly conserved regions (marked in blue) of which five regions where also conserved up to human (Figure 32B).

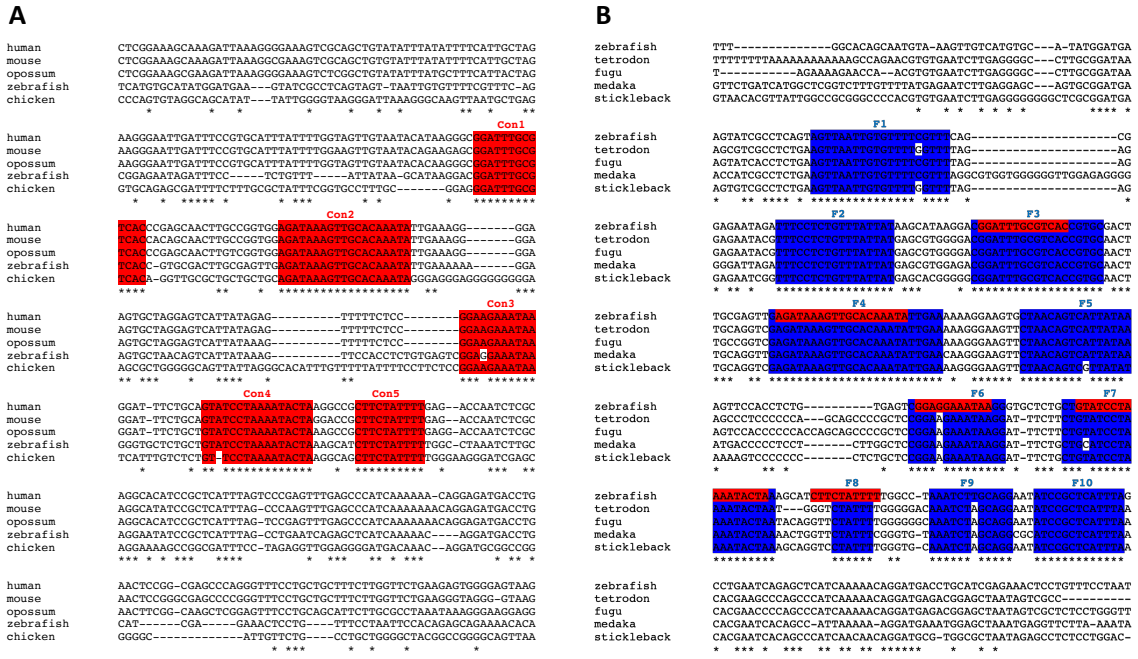


Figure 32: HLX-3 sequence alignment with ClustalW. A) *Aligned human, mouse, opossum, zebrafish and chicken sequence of the HLX-3 enhancer. Conserved regions are marked in red and labelled Con1 - Con5 in a 5' to 3' direction.* ***B)*** *Aligned zebrafish, tetrodon, fugu, medaka, stickleback sequence of the hlx-3 enhancer. Regions that are conserved in all fish species are marked in blue and labelled F1 - F10. Conserved base pairs indicated with *.*

Given the known importance of ETS transcription factors for vascular growth, and for the endothelial-specific expression of all endothelial enhancers and promoters[279,280,125,218], I first established whether these enhancers contained conserved consensus ETS binding sites in the enhancer sequences by looking for the core ETS GGA$^{A}/_{T}$ sequence motif. The HLX-3 enhancer contained nine core ETS motifs, of which seven were conserved to zebrafish. Two additional ETS motifs were found in zebrafish which were not present in mammals (Figure 33).

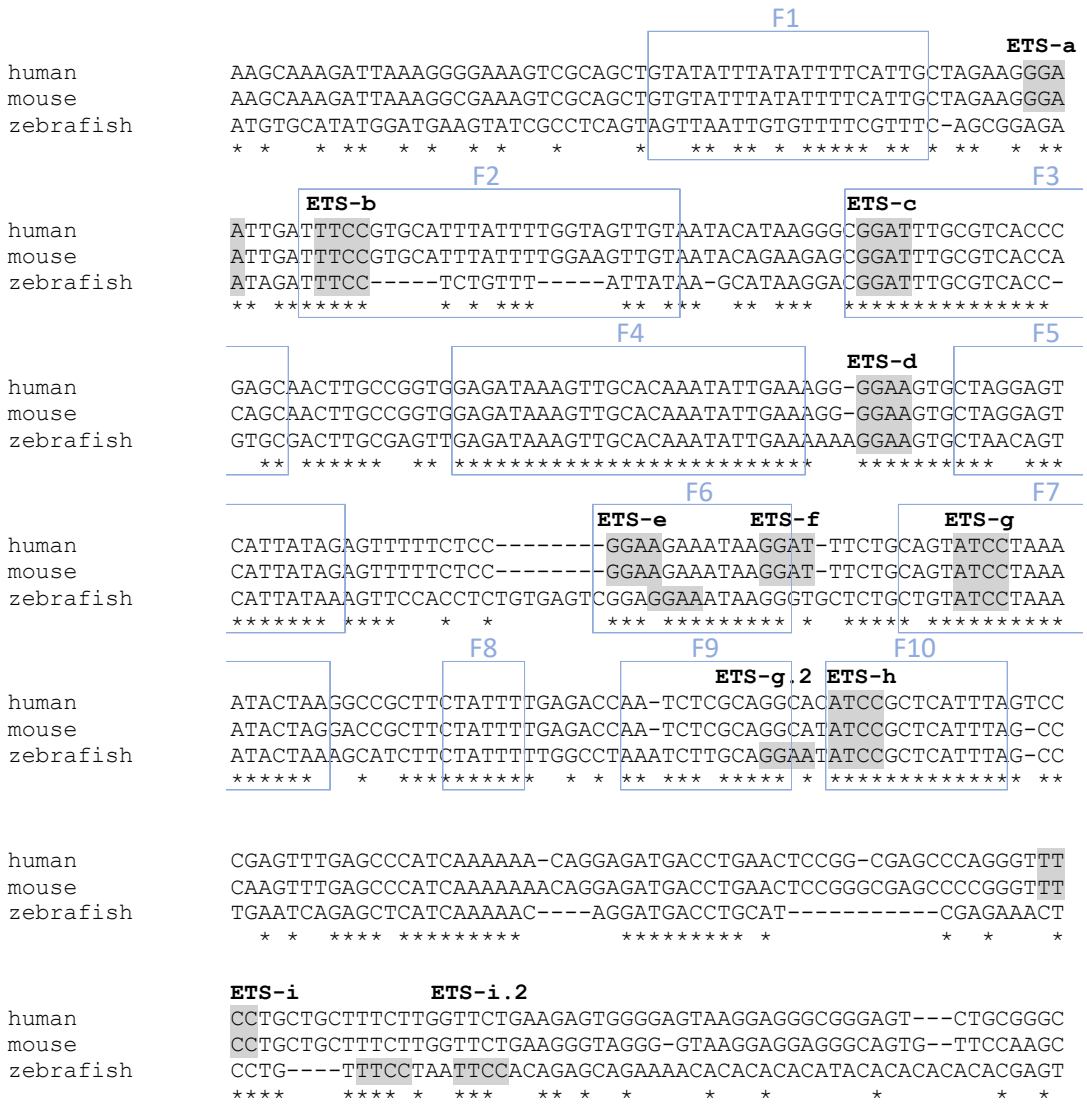


Figure 33: Putative ETS transcription factor binding sites conserved between human, mouse and zebrafish. *All putative ETS binding motifs were either identified by manual inspection or MatInspector (highlighted in grey). Sequences of human HLX-3, mouse Hlx-3 and zebrafish hlx-3 enhancers were aligned using ClustalW[278]. Conserved base pairs are indicated with *. Regions conserved in teleost are marked by blue boxes and labelled F1 - F10.*

4.3.2 ETS transcription factors can directly bind the hlx-3/HLX-3 enhancer

In silico enhancer sequence analysis provides a good prediction for transcription factor binding but actual binding has to be confirmed *in vitro* or *in vivo*. This is particularly important for ETS transcription factors, as they require only a fairly common GGA$^{A}/_{T}$ motif while their specific requirements for flanking sequences are not understood[281]. To investigate which consensus and near-consensus ETS binding sites can actually be bound

by endothelial-expressed ETS transcription factors, electrophoretic mobility shift assays (EMSA) were performed.

All nine HLX-3 ETS sites that were identified *in silico* were first investigated to see whether they were able to robustly compete with a known ETS binding site for binding of an ETS DNA-binding domain (ETS1DBD, taken from ETS1). ETS1DBD protein was transcribed and translated *in vitro*, then incubated with a radiolabelled probe containing a known ETS binding motif and an unlabelled excess of the putative ETS motif to be tested (or mutant version, in which the putative ETS site was altered). After incubation, this mix was then run on a gel to establish whether binding has occurred by a change in running distance. If the HLX-3 ETS motif was able to bind the ETS protein, the labelled control ETS would be unable to bind, and no shift is seen on the gel. This analysis indicates that the HLX-3 ETS sites -b, d, e, h and i were able to compete for binding of ETS1DBD (Figure 34A).

Having established which putative ETS motifs were able to compete for ETS1 binding with a known ETS binding site, I next tested whether HLX-3 ETS sites could directly bind different endothelial ETS proteins. Directly radiolabelled oligos containing each ETS motif were incubated separately with ETS1DBD or the ETS factor ETV2, which is absolutely required for endothelial differentiation[117], as well as abundant endothelial ETS proteins ELF1 and ERG (Figure 35). HLX-3 ETS-c bound ETV2, but only weakly, while ETS binding sites ETS-a, f and g were unable to bind any endothelial ETS site (ETS1DBD, ERG, ETV2 or ELF1). HLX-3 ETS sites -b, d, e, h and i were bound strongly by both ETS1DBD and ETV2. Last, validated ETS binding sites -b, d, e, h and i were tested for direct binding of ETS1DBD.

ETS binding sites -b, d, e, h and i were competed by an excess of self-probe, but not by mutant self-probe (Figure 36A).

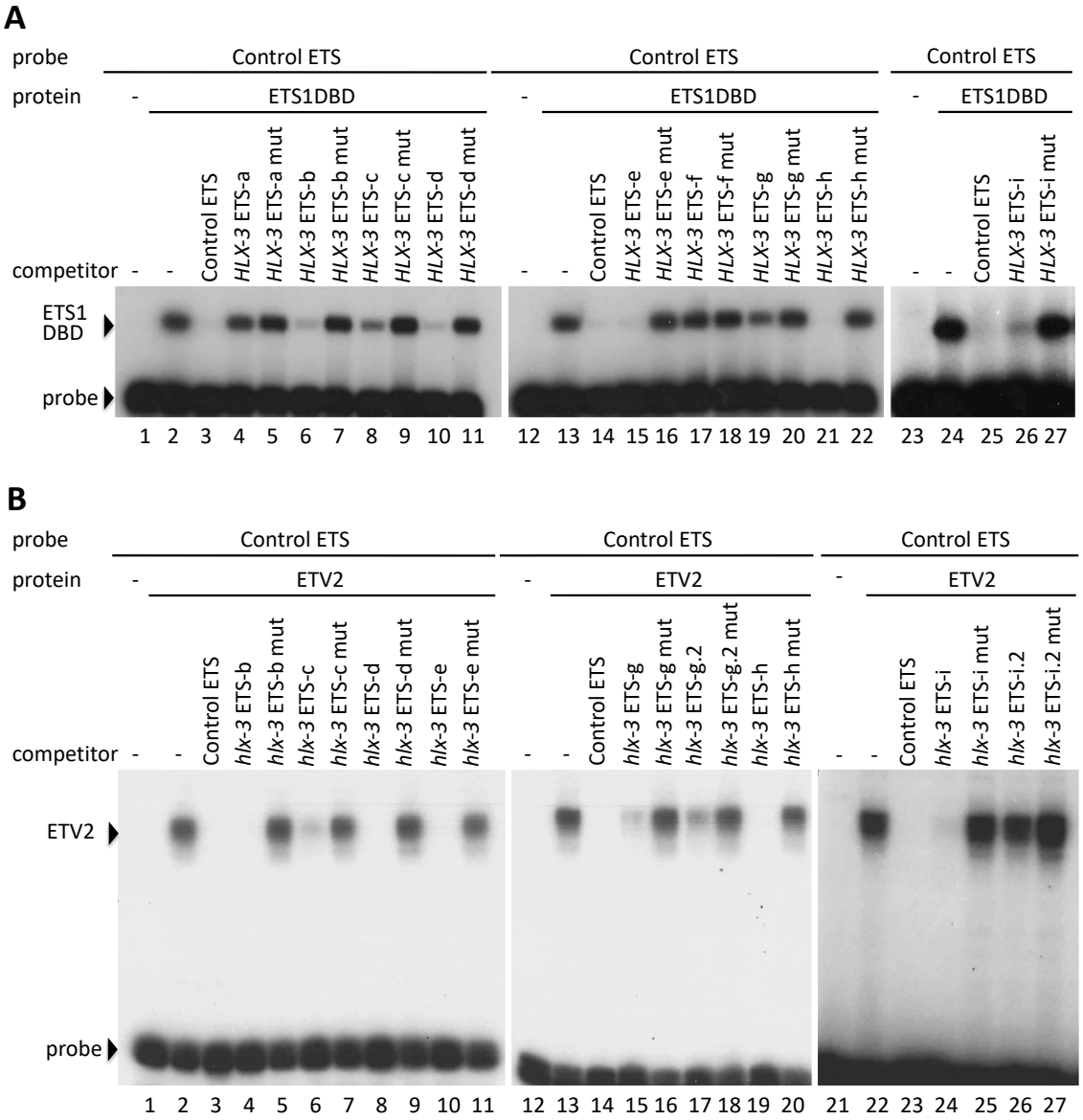


Figure 34: HLX-3 and hlx-3 enhancer are bound by ETS transcription factors. A) *Human HLX-3 ETS binding sites b, d, e, h and i specifically compete with binding of ETS1DBD to labelled control ETS (lanes 6, 10, 15, 21 and 26) in an indirect EMSA. HLX-3 ETS-a, c, f and g were not able to compete with ETS control for ETS1DBD binding (lanes 4, 8, 17 and 19). ETS1DBD binds labelled control (lanes 2, 13 and 24). Mutating HLX-3 ETS sites re-establishes ETS1DBD binding to the labelled control, implicating that the interaction is specific to the motif (lanes 5, 7, 9, 11, 16, 18, 20, 22 and 27).* ***B)*** *Zebrafish hlx-3 ETS binding sites b, d, e, h and i specifically compete with binding of ETV2 to labelled control ETS (lanes 4, 8, 10, 19 and 24) in an indirect EMSA. ETS1DBD binds labelled control (lanes 2, 13 and 22). Mutating HLX-3 ETS sites re-establishes ETV2 binding to the labelled control, implicating that the interaction is specific to the motif (lanes 5, 7, 9, 11, 16, 18, 20, 25 and 27). Zebrafish-specific ETS sites hlx-3 ETS-g.2 and hlx-3 ETS-i.2 were not conserved in human and mouse (see putative ETS binding sites marked in grey in Figure 33). hlx-3 ETS-c, g, g.2 and i.2 were not able to compete with ETS control for ETS1DBD binding (lanes 6, 15, 17 and 26). White vertical lines indicate where regions of EMSA gel have been deleted from image for clarity. All EMSAs were performed three times and probes were used at 20 - 40,000 counts/minute.*

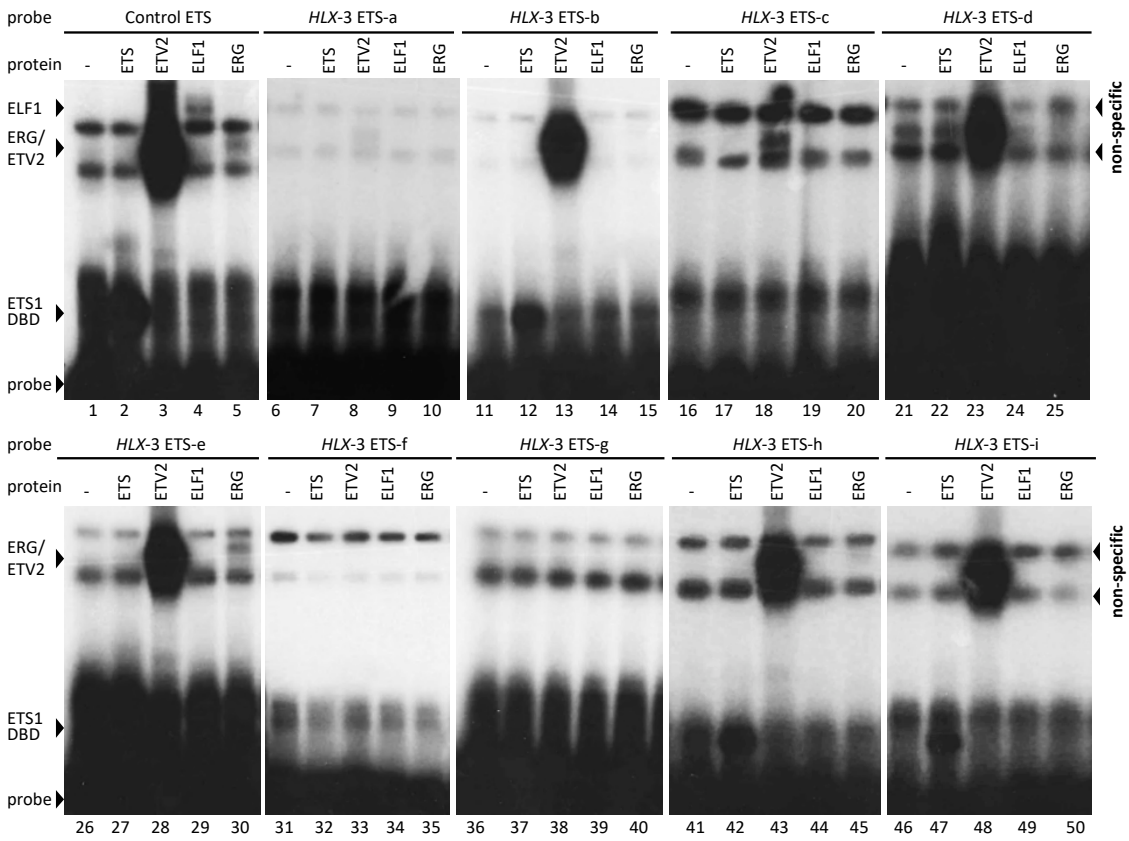


Figure 35: EMSA showing binding of ETS family proteins to human HLX-3 enhancer. *Radiolabelled oligonucleotide probes encompassing the control ETS (lanes 1 - 5), HLX-3 ETS-a (lanes 6 - 10), HLX-3 ETS-b (lanes 11 - 15), HLX-3 ETS-c (lanes 16 - 20), HLX-3 ETS-d (lanes 21 - 25), HLX-3 ETS-e (lanes 26 - 30), HLX-3 ETS-f (lanes 31 - 35), HLX-3 ETS-g (lanes 36 - 40), HLX-3 ETS-h (lanes 41 - 45), HLX-3 ETS-i (lanes 46 - 50) were incubated with recombinant ETS1 DNA binding domain (ETS), ETV2, ELF1 and ERG. HLX-3 ETS-b, d, e, h, i probes bound to recombinant ETV2 in a manner similar to control ETS, whereas no binding was detected for HLX-3 ETS-a, f, g. HLX-3 ETS-c only weakly bound ETV2. HLX-3 ETS-e, h probes also bound ERG. In all cases, white vertical lines indicate where regions of EMSA gel have been deleted from image for clarity. All EMSAs were performed two times and probes were used at 20 - 40,000 counts/minute.*

I next validated the putative zebrafish hlx-3 ETS binding motif sequences in EMSA, to ensure both human and zebrafish sequences could bind. The HLX-3 ETS-a and -f motifs were not conserved in zebrafish and hence only zebrafish ETS-b, c, d, e, g, g.2, h, i and i.2 were tested for binding ETV2 (only ETV2 was used here, as the ETV2 results in human were directly indicative of the other ETS proteins). As with the human sequences, the zebrafish versions of hlx-3 ETS-b, d, e, h and i were able to bind ETV2 in both indirect and direct EMSAs (Figure 34B and Figure 36B).

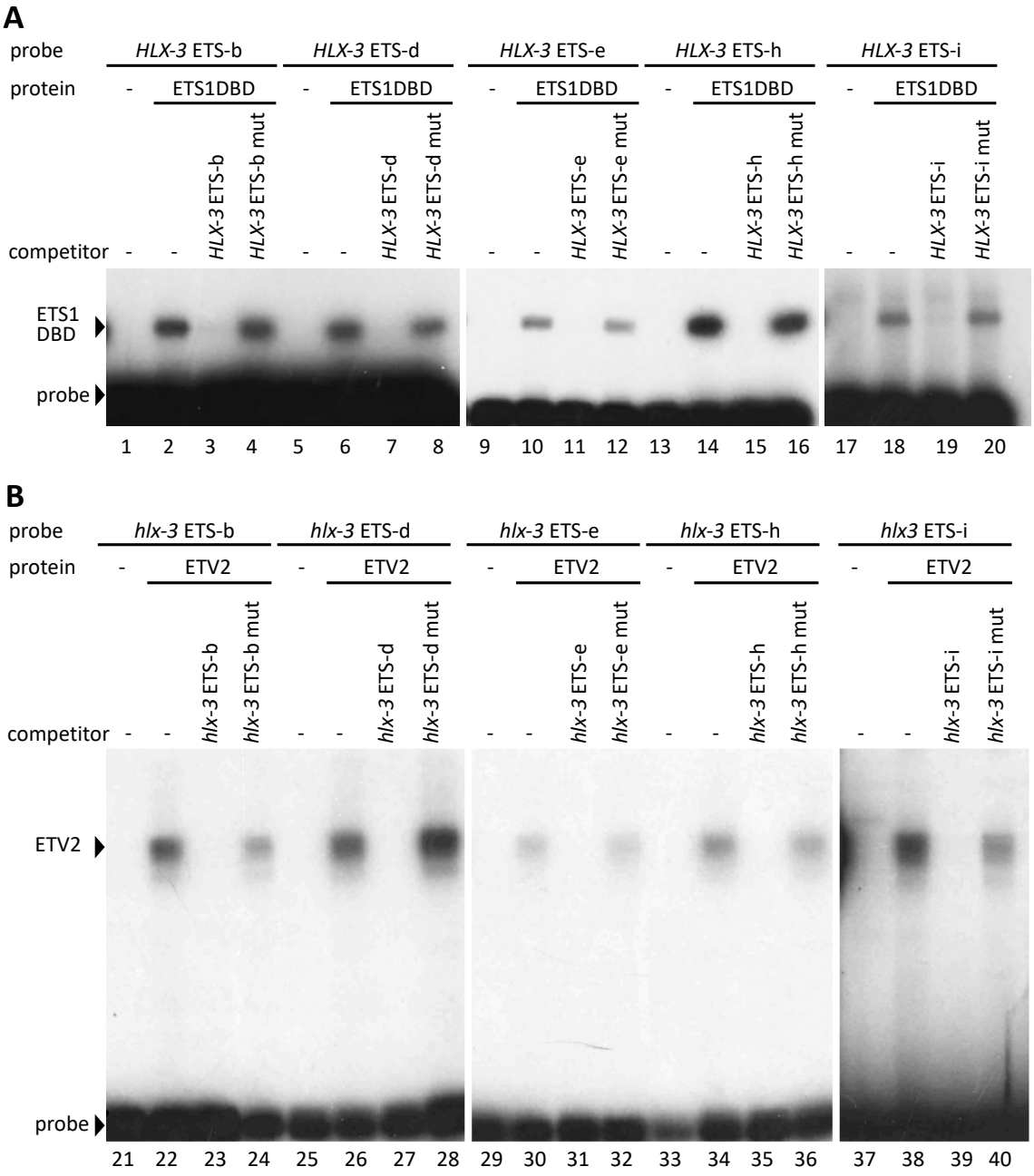


Figure 36: The human HLX-3 and zebrafish hlx-3 enhancers robustly bind the ETS factors ETS1DBD and ETV2 in EMSA analysis. A) *Radiolabelled oligonucleotide probe encompassing the human sequence of five ETS binding motifs (ETS-b, d, e, h and i) were bound by recombinant ETS1DBD protein (lanes 2, 6, 10, 14 and 18), were competed by an excess of self-probe (lanes 3, 7, 11, 15 and 19), but not by mutant self-probe (lanes 4, 8, 12, 16 and 20).* ***B)*** *Radiolabelled oligonucleotide probe encompassing the zebrafish sequence of five ETS binding motifs (ETS-b, d, e, h and i) were also bound by recombinant ETV2 protein (lanes 22, 26, 30, 34 and 38), were competed by an excess of self-probe (lanes 23, 27, 31, 35 and 39), but not by mutant self-probe (lanes 24, 28, 32, 36 and 40). White vertical lines indicate where regions of EMSA gels have been deleted from image for clarity. All EMSAs were performed three times and probes were used at 40,000 counts/minute. Adapted from Sacilotto, Chouliaras et al.*[170].

In conclusion, my results confirm that, similar to all known endothelial enhancers, the HLX-3 and hlx-3 enhancer contain conserved functional ETS binding sites and can be bound by ETS family members[139]. Figure 37 summarizes these results and shows the five functional out of the nine putative ETS binding sites of the HLX-3/hlx-3 enhancer sequence.

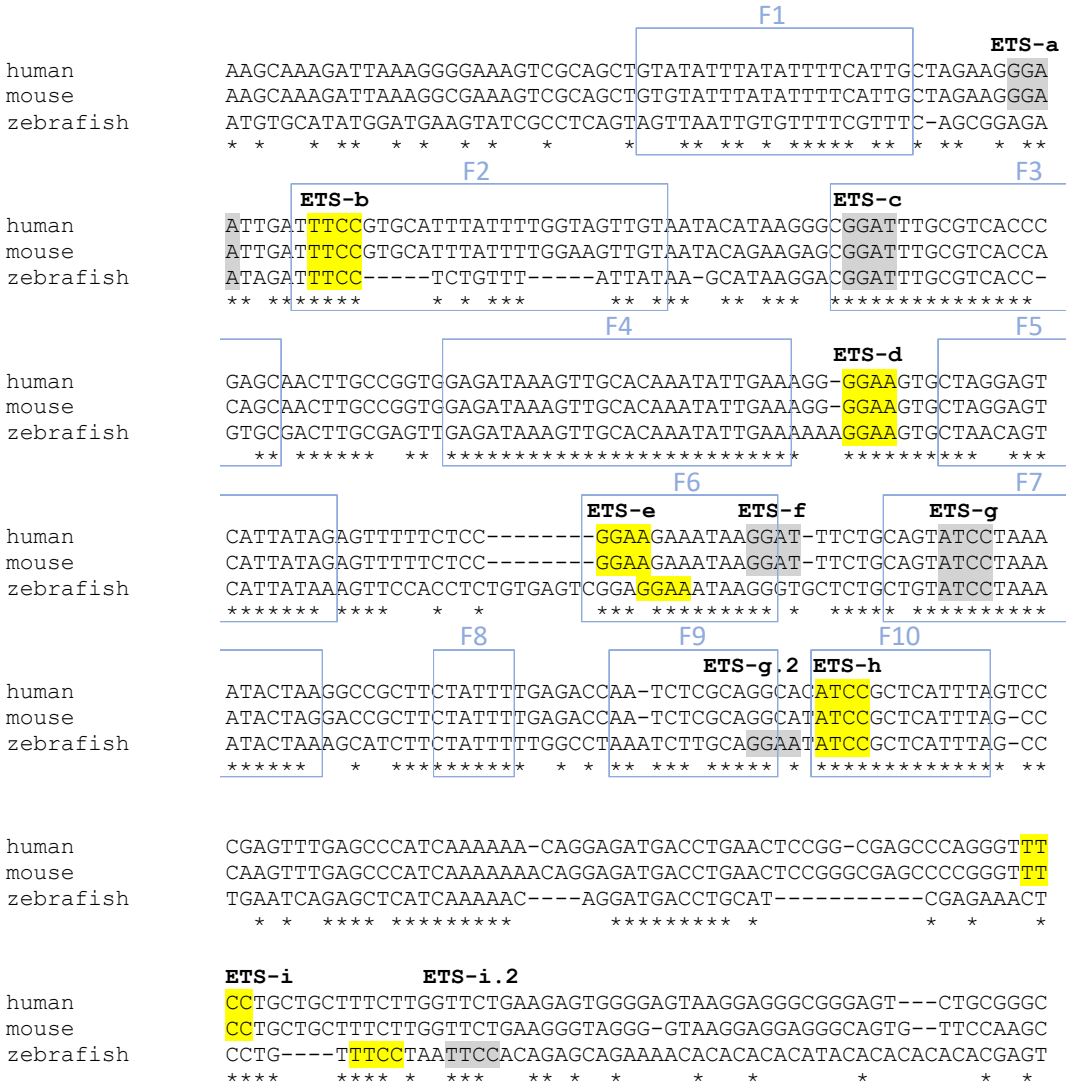


Figure 37: Functional ETS transcription factor binding sites conserved between human, mouse and zebrafish. *Sequences of human HLX-3, mouse Hlx-3 and zebrafish hlx-3 enhancers were aligned using ClustalW[278]. Conserved base pairs are indicated with *. Regions conserved in teleost are marked by blue boxes and labelled F1 - F10. Verified functional ETS binding motifs are marked in yellow, while ETS motifs that were not bound by ETS factors in EMSAs are marked in grey.*

4.3.3 ETS binding sites are required for hlx-3 driven reporter gene expression

So far my analysis has demonstrated that the HLX-3 and hlx-3 enhancers contain functional ETS motifs. Next, I examined whether these ETS motifs were essential for hlx-3 enhancer activity. I generated a mutated version of hlx-3 in which all functional ETS binding sites (b, d, e, h, i) were mutated to preclude ETS binding. These mutations were the same ones verified by EMSA and resulted in total ablation of ETS binding (Figure 36). The mutated hlx-3 transgene was tested by generating transient transgenic zebrafish and evaluating reporter gene expression in sprouting ISVs at 30 - 32 hpf (see Figure 38). As expected, mutation of all five functional ETS motifs resulted in a total ablation of hlx-3:Citrine activity (I replaced the GFP reporter gene with Citrine in these assays). Interestingly, mutation of 4 out of 5 ETS binding sites (leaving ETS-i intact) did not ablate transgene activity as the hlx-3 mut ETS b, d, e, h enhancer was still able to drive expression in ISVs, although at reduced levels (Figure 38).

In conclusion, this mutational analysis confirms that ETS binding sites are essential for the angiogenic activity of hlx-3, as they are for every endothelial enhancer.

A

	hlx-3	mut ETS-b,d,e,h,i	mut ETS-b,d,e,h
% of fish with Citrine expression in ISVs	81%	0%	22%
# of fish with detectable Citrine expression (%)	61 (60%)	14 (56%)	27 (75%)
Number of fish analysed	102	25	36

B

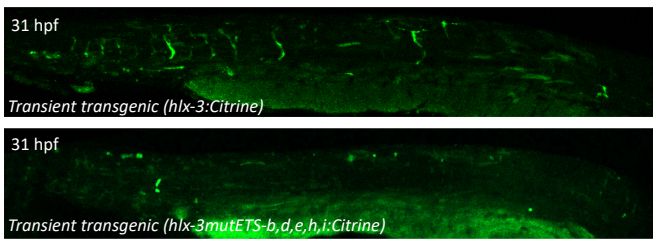


Figure 38: ETS factor binding sites are crucial for angiogenic-specific hlx-3 enhancer activity. A) *Summary of reporter gene expression detected in Tol2-mediated mosaic transient transgenic zebrafish at 30 - 32 hpf. A transient transgenic zebrafish was scored as positive when at least one ISV showed Citrine expression. In all tested mutants, no novel binding motifs were created.* ***B)*** *Transient transgenic 31 hpf zebrafish embryos representative for the hlx-3:Citrine and hlx-3 mut ETS b, d, e, h, i:Citrine transgenes.*

4.3.4 Mutational analysis of the hlx-3 enhancer reveals a potential role for MEF2 transcription factors

All genes expressed in endothelial cells require ETS factors for transcriptional activation. Independent of the sub-endothelial cell type, ETS transcription factors have been shown to be essential for all developing blood vessels[282,152]. In line with these studies, ETS factor binding is an essential feature for every endothelial-specific enhancer so far investigated[167,179,218]. These enhancers drive a range of different expression patterns within the vasculature, including expression in all endothelial cells, or specific expression to either arterial and venous cells. Therefore the ETS factors are unlikely to be responsible alone for driving expression in angiogenic endothelial cells. Angiogenic-specific enhancers are thus likely to receive their angiogenic-specific information from a different transcription factor.

Having confirmed the ETS-dependency of hlx-3, I next wanted to examine which region within the HLX-3/hlx-3 enhancer conveys its unique angiogenic-specific transcriptional information. Figure 39 shows the putative transcription factor binding sites within the conserved regions of the hlx-3 enhancer.

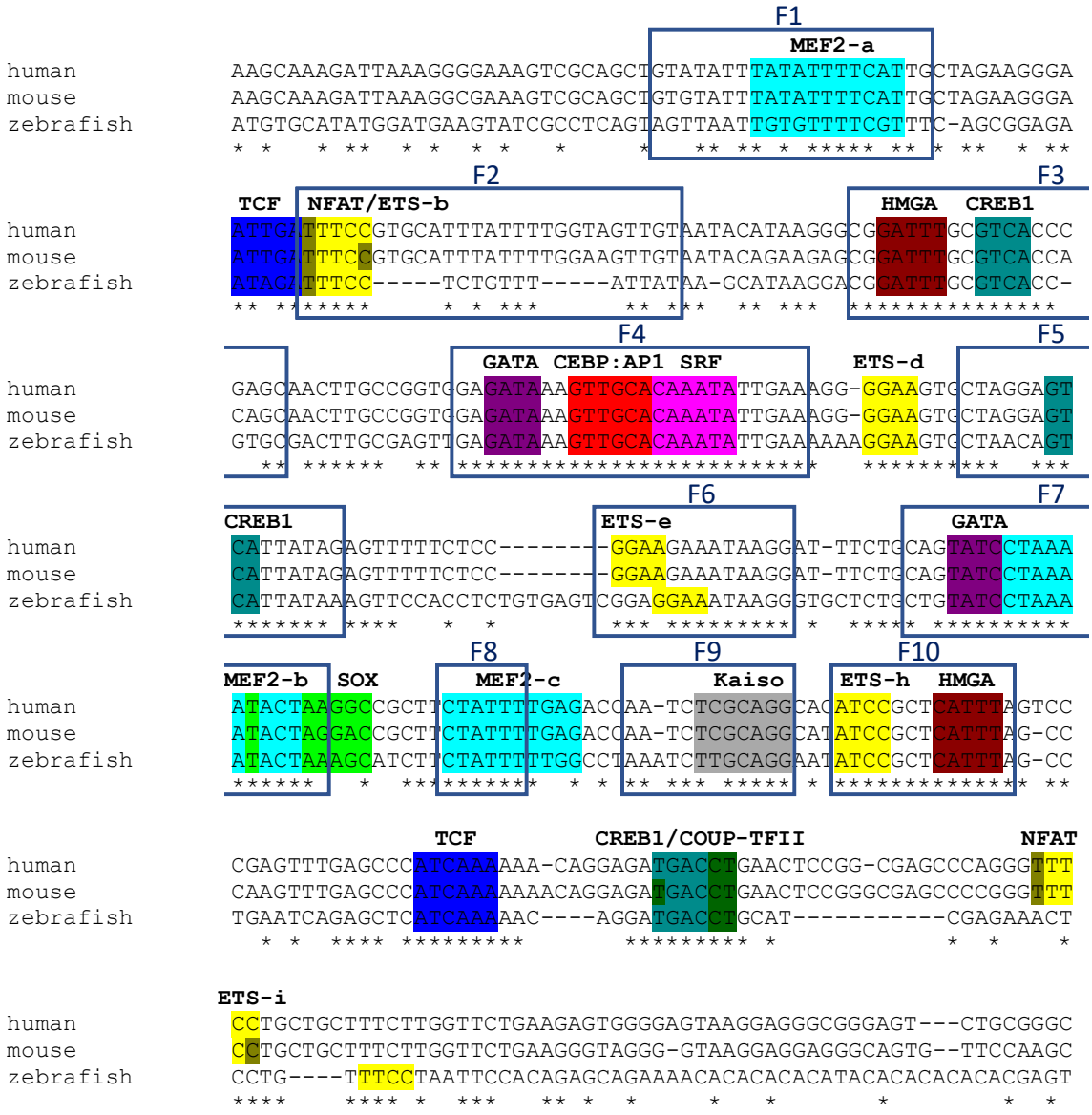


Figure 39: Consensus and near-consensus transcription factor binding sites conserved between human, mouse and zebrafish. *All sites were either identified by inspection or MatInspector. Transcription factor binding sites are highlighted with different colours. Transcription factors which have an overlapping binding sites are marked with " / ". Sequences of human HLX-3, mouse Hlx-3 and zebrafish hlx-3 enhancers were aligned using ClustalW[278]. Conserved base pairs are indicated with *. Regions conserved in teleost are marked by blue boxes and labelled F1 - F10. Verified ETS binding motifs are marked in yellow. The binding sites for the following transcription factors were found: MEF2 = Myocyte enhancer factor-2, NFAT = Nuclear factor of activated T-cells, ETS = E26 transformation-specific, TCF = T-cell factor, HMGA = High mobility group protein AT-hook, CREB1 = cAMP response element-binding protein, CEBP:AP1 = CCAAT/enhancer-binding protein:Activator protein 1, SRF = Serum response factor, SOX = sry-related HMG box, COUP-TFII = chicken ovalbumin upstream-transcription factor II, GATA and Kaiso (encoded by Zinc finger and BTB domain containing 33 (ZBTB33)).*

As previously discussed, the HLX-3/hlx-3 enhancer contains a number of deeply conserved regions. Since angiogenic enhancer activity was conserved between the HLX-3 and hlx-3 enhancers, I hypothesised that these deeply conserved regions may bind one or more

transcription factors important for the angiogenic-specific activity of the enhancer. Therefore, I first applied a combinatorial mutational approach, where multiple conserved regions were mutated simultaneously and the subsequent activity of each mutated hlx-3:Citrine transgene was assessed in transient transgene zebrafish. This allowed me to narrow down which transcription factors are responsible for hlx-3 angiogenic-specific expression in a few steps. Because I was primarily interested in finding one or more transcription factors that were absolutely required for angiogenic-specific gene transcription, I was looking for sequences within the hlx-3 enhancer that, when mutated, would result in complete loss of enhancer activity.

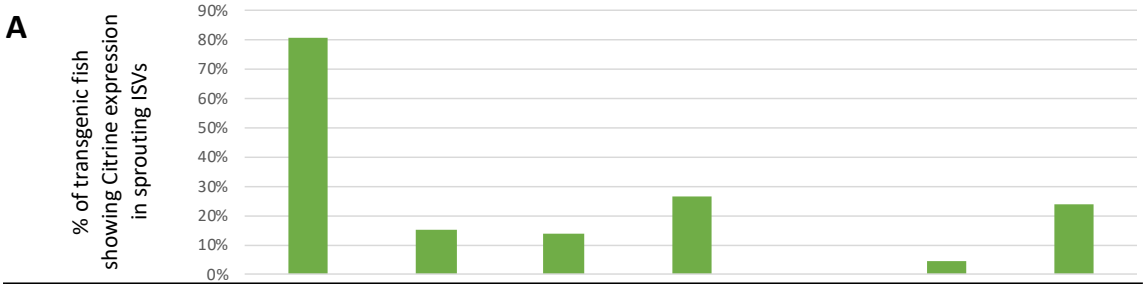
First, the five regions of the hlx-3 enhancer that were conserved across mammalian and teleost species were mutated and tested for enhancer activity (hlx-3 mut F3, 4, 6, 7, 8 see Figure 39 for sequences mutated). When mutating these deeply conserved regions while leaving all ETS binding sites intact, hlx-3 reporter gene expression dropped from 80% to 18% of transgenic zebrafish in sprouting ISVs at 30 - 31 hpf (see Figure 40A). Although the likelihood of transcription factors binding to the enhancer and activating reporter gene expression was reduced, it was not ablated. Thereafter, I mutated the conserved regions which lie adjacent to functional ETS binding sites (hlx-3 mut F2, 5, 9, 10). The resultant hlx-3 mut F2, 5, 9, 10 enhancer was still able to drive Citrine expression in ISVs although in a reduced number of transgenic fish, suggesting that these regions did not contain binding sites that were crucial for hlx-3 enhancer angiogenic specificity (see Figure 40A).

Next, I investigated any potential role for the conserved regions found in both hlx-3 and HLX-3. These were investigated into two groups based on the putative transcription factor

binding sites they contain. In the hlx-3 mut F3, 4, 5, 9:Citrine mutant transgene, I mutated conserved regions containing putative CREB1 binding sites, while in the hlx-3 mut F1, 7, 8:Citrine mutant transgene, I mutated conserved regions containing putative MEF2 binding sites (see Figure 39). The hlx-3 mut F3, 4, 5, 9:Citrine mutant transgene was still able to direct activity in ISVs, suggesting that none of the mutated regions were absolutely required for enhancer activity in sprouting endothelial cells (see Figure 40A). In contrast, the hlx-3 mut F1, 7, 8:Citrine mutant transgene was unable to direct any enhancer activity in the vasculature (Figure 40B). Since F1, F7 and F8 all contained binding motifs for members of the MEF2 transcription factor family, this result suggested a potential involvement of MEF2 factors in activity of the hlx-3 enhancer. Supporting this, mutating either F1 or F7/8 regions separately was not sufficient to abolish hlx-3 enhancer activity in ISVs (Figure 40A). This result suggests that leaving at single functional MEF2 site in hlx-3 can maintain its ability to drive angiogenic gene expression, but loss of all three MEF2 motifs removes the ability of hlx-3 to be active during angiogenesis.

To summarize, these results show that MEF2 binding sites are absolutely required to drive angiogenic-specific expression of hlx-3, indicating a potential crucial role for MEF2 factors in sprouting endothelial cells. It should also be noted that mutating the other conserved regions in hlx-3 also significantly reduced angiogenic-specific enhancer expression, suggesting that other TFs also provide some information required for robust enhancer activity and are required for the maximal functioning of the hlx-3 enhancer.

A



	hlx-3	mut F3,4,6,7,8	mut F2,5,9,10	mut F3,4,5,9	mut F1,7,8	mut F1	mut F7,8
% of fish with Citrine expression in ISVs	81%	15%	14%	26%	0%	5%	24%
# of fish with detectable Citrine expression (%)	61 (60%)	40 (71%)	49 (58%)	34 (79%)	62 (67%)	22 (33%)	21 (31%)
Number of fish analysed	102	56	84	43	93	66	68

B

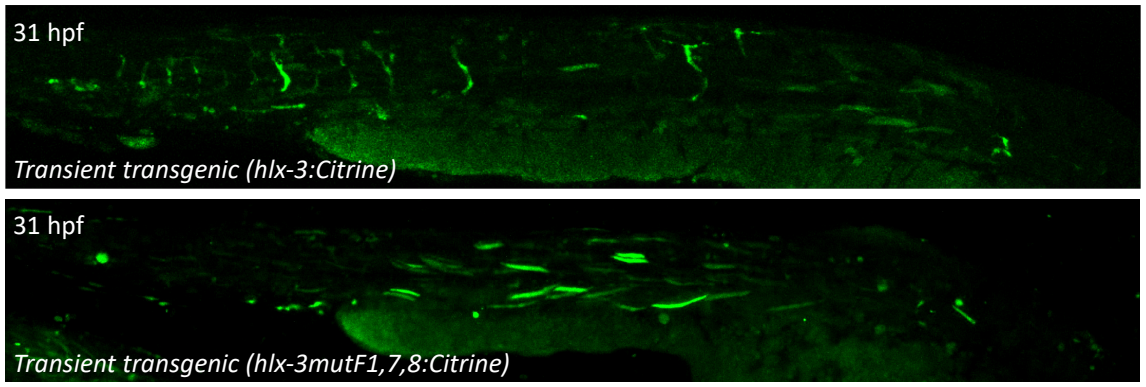


Figure 40: Mutational analysis of hlx-3 conserved regions to find transcription factor binding motifs crucial for angiogenic-specific enhancer activity. A) *Summary of reporter gene expression detected in Tol2-mediated mosaic transient transgenic zebrafish at 30 - 32 hpf. A transient transgenic zebrafish was scored as positive when at least one ISV showed Citrine expression. Analysis of transgene expression was qualitative not quantitative, therefore no statistical analysis was applied to the observations of reporter gene intensity. In all tested mutants, no novel binding motifs were created and ETS binding sites were left intact.* ***B)*** *Transient transgenic 31 hpf zebrafish embryos representative for the hlx-3:Citrine and the hlx-3 mut F1, 7, 8:Citrine transgenes. hlx-3 mut F1, 7, 8:Citrine shows ectopic expression in skeletal muscle fibres.*

4.3.5 MEF2 transcription factors bind the MEF2 motifs within the hlx-3/HLX-3 enhancer

4.3.5.1 HLX and MEF2 factor expression increases after VEGFA stimulation of HUVECs

There are four members of the MEF2 family (MEF2A, B, C and D) which are nearly ubiquitously expressed yet play a crucial role during differentiation and development of many different cell lineages, including cardiovascular development[148]. Out of the four

MEF2 factors, MEF2B shows the lowest degree of sequence homology with the other family members and is the least well studied[142]. MEF2A and C are expressed in the vasculature, although endothelial-specific ablation of MEF2C did not show vascular defects during embryogenesis[150]. Analysis in our laboratory by Dr Martin Fritzsche has found that MEF2A, C and D protein can be detected in both E11 hindbrain and HUVECs (Figure 41).

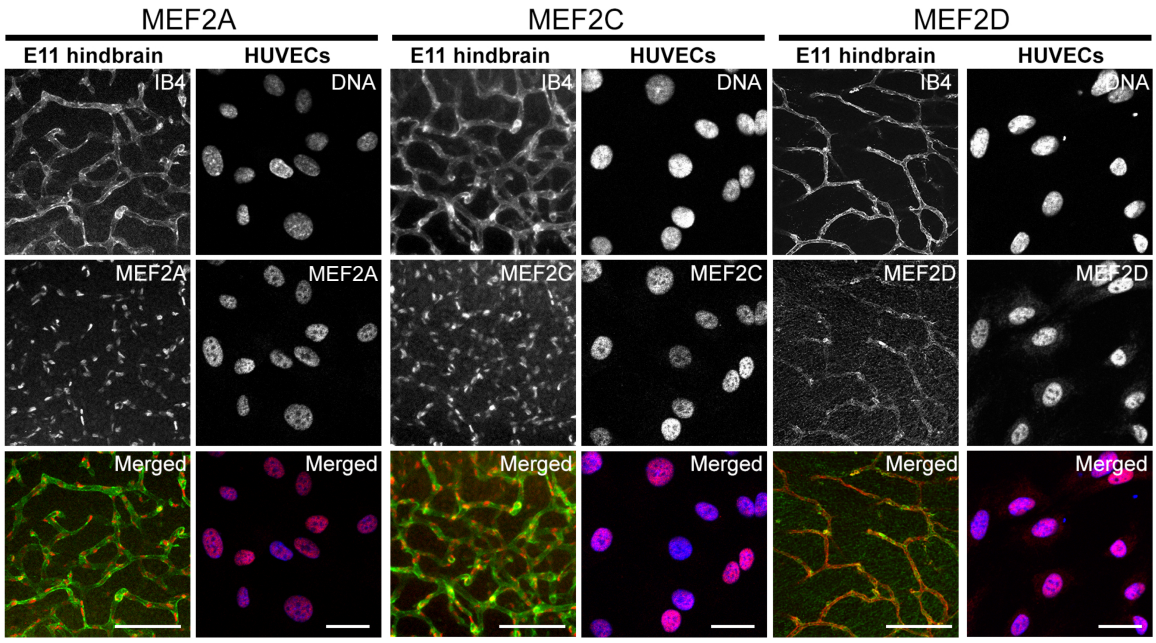


Figure 41: MEF2A, C and D are expressed in endothelial cells. *Whole-mount immunostaining demonstrates expression of MEF2A, MEF2C, and MEF2D transcription factors in endothelial cells from E11 hindbrains and in HUVECs. Bars correspond to 100 µm. Data were provided by Dr Martin Fritzsche and the figure was taken from Sacilotto, Chouliaras et al.[170].*

To determine if MEF2 transcription factors were expressed in endothelial cells after VEGFA induction, I investigated the expression of *MEF2A*, *C* and *D* in HUVECs before and after VEGFA stimulation, relative to *HLX* expression and *DLL4* expression (as a marker for angiogenesis). HUVECs were serum-starved in 0.1% FBS for 24 hours and then treated with $VEGFA_{165}$ over a time-course of 12 hours. Upon VEGFA stimulation, I first established that

the cells elicited the expected biological response, thus phosphorylation of VEGFR2 after 5 minutes, followed by ERK phosphorylation after 5 - 10 minutes and Akt phosphorylation after 15 minutes (Figure 42A). This analysis also found that $VEGFA_{165}$ induces both *HLX* and *DLL4* expression peaking at 30 minutes after stimulation and decreasing over time. Likewise, *MEF2A*, *MEF2C* and *MEF2D* were induced upon VEGFA stimulation, although the *MEF2A/C/D* induction was slower (Figure 42B). DLL4, MEF2A, MEF2C and MEF2D protein levels were also increased upon VEGFA (data not shown). [qRT-PCR experiments were performed jointly with Dr Leonid Nikitenko.]

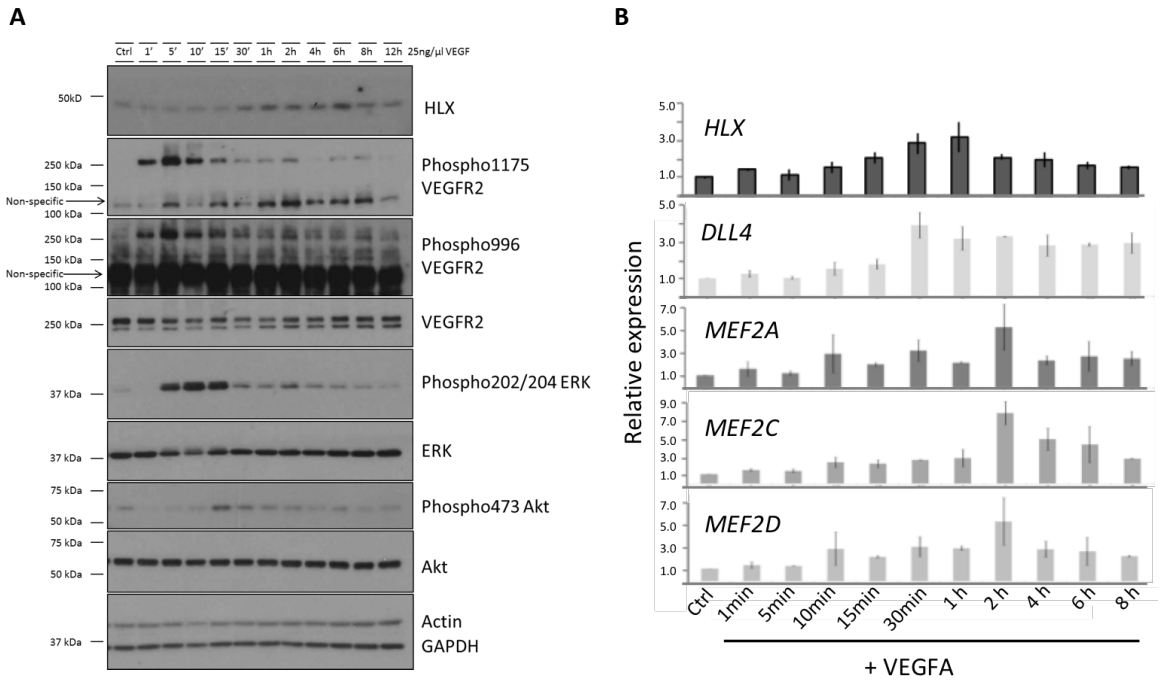


Figure 42: HLX, DLL4 and MEF2A/C/D were activated by $VEGFA_{165}$ in HUVECs*. Quiescent HUVECs were treated for one minute to 12 hours with 25 ng/µl $VEGFA_{165}$. Protein and RNA was harvested at each time point as indicated.* ***A)*** *High salt lysis extracts were analysed by immunoblotting with anti-HLX, anti-phospho-tyrosine 1175-VEGFR, anti-phospho-tyrosine 996-VEGFR, anti-VEGFR, anti-phospho-threonine 202/tyrosine 204-ERK, anti-ERK, anti-phospho-serine 437-Akt and anti-Akt antibodies. Total levels of VEGFR2, ERK1/2 and Akt remain constant after stimulation whereas changes in HLX, pERK and pAkt indicate successful VEGFA stimulation. Actin and GAPDH were used as a loading control.* ***B)*** *qRT-PCR analysis demonstrates increased expression of HLX, DLL4, MEF2A, MEF2C and MEF2D in HUVECs 0 - 8 hours after VEGFA stimulation. Error bars indicate standard error of the mean of three biological replicates for HLX and two biological replicates for DLL4, MEF2A, MEF2C and MEF2D. Ctrl indicates the unstimulated control. Expression levels were normalized to ACTB. Figure adapted from Sacilotto, Chouliaras et al.[170]. Results are representative of three independent experiments.*

4.3.5.2 MEF2 factors can bind HLX-3 and hlx-3 sites by EMSA analysis

MEF2 family members bind to the consensus DNA motif $^{C}/_{T}TA\ (^{A}/_{T})^{4}\ TA^{A}/_{G}$[283]. As previously discussed, the hlx-3 enhancer contains three highly conserved putative MEF2 motifs, the mutation of which ablated all activity of the hlx-3 enhancer in endothelial cells. To determine if these putative motifs were able to bind MEF2 protein *in vitro*, we investigated the ability of both human HLX-3 and zebrafish hlx-3 MEF2 motif sequences to compete for, and directly bind, the MEF2C protein using EMSA analysis. This clearly demonstrated that each of the three MEF2 binding sites were able to binding MEF2C (Figure 43). This binding was conserved from human to zebrafish, although binding of MEF2C to zebrafish hlx-3 MEF2-a was weaker than that to human HLX-3 MEF2-a. Out of the three MEF2 binding sites tested, MEF2-c showed strongest binding to MEF2C and was comparable between human and zebrafish.

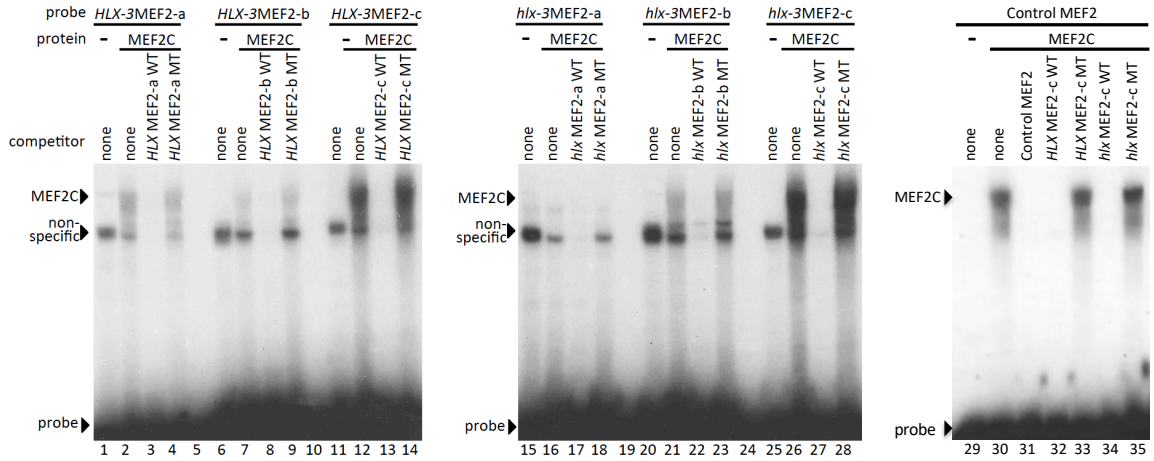


Figure 43: MEF2C factor binds human HLX-3 and zebrafish hlx-3. *Radiolabelled oligonucleotide probe encompassing the human sequence of the three HLX MEF2 site were directly bound by recombinant MEF2C protein (lanes 2, 7 and 12), were competed by an excess of unlabelled self-probe (lanes 3, 8 and 13) but not mutant self-probe (lanes 4, 9 and 14). Radiolabelled oligonucleotide probe encompassing the orthologous zebrafish sequence of the three hlx MEF2 site were directly bound by recombinant MEF2C protein (lanes 16, 21, 26), were competed by an excess of unlabelled self-probe (lanes 17, 22 and 27) but not mutant self-probe (lanes 18, 23 and 28). Radiolabelled MEF2C control probe bound to MEF2C (lane 30), was competed by an excess of unlabelled HLX-3 and hlx-3 MEF2-c probe (lanes 32 and 34), but not mutant HLX-3 and hlx-3 MEF2-c probe (lanes 33 and 35). All EMSAs were performed two times. Probes were used at 40,000 counts/minute. Adapted from Sacilotto, Chouliaras et al.*[170].

Although this analysis was principally done with MEF2C, the literature and our data indicate that other MEF2 factors are also expressed in endothelial cells. MEF2 factors are highly redundant and are known to bind similar DNA motifs[142], although MEF2B has been reported to have a reduced binding affinity for the core MEF2 motif[145]. My results confirm these observations, as the HLX-3 and hlx-3 MEF2-c sites were also able to robustly bind MEF2A and MEF2D proteins in EMSA analysis (Figure 44).

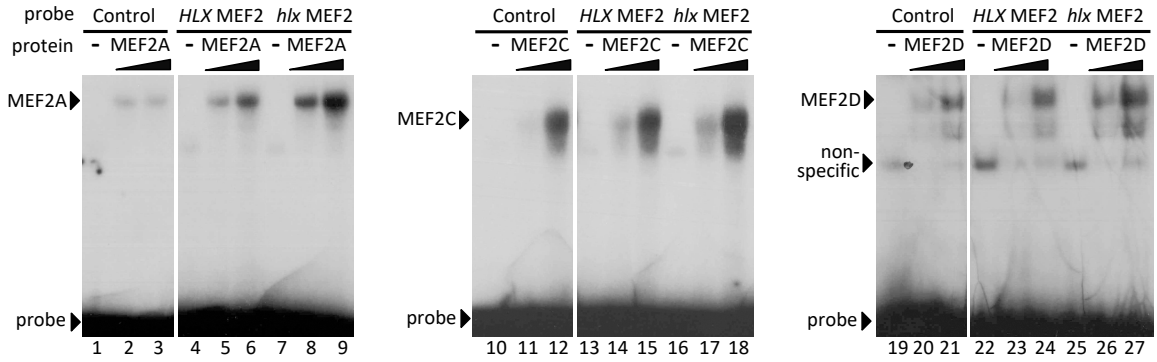


Figure 44: MEF2A and MEF2D also bind HLX-3 and hlx-3. *Radiolabelled oligonucleotide probe encompassing the HLX-3 MEF2-c and hlx-3 MEF2-c sites were able to bind 2 µl and 4 µl MEF2A (lanes 1 - 9), MEF2C (lanes 10 - 18) and MEF2D (lanes 19 - 27) proteins at higher affinity than the control MLC MEF2 site. All probes were used at 40,000 counts/minute. EMSAs were performed two times. White vertical lines indicate where regions of EMSA gels have been deleted from image for clarity. Adapted from Sacilotto, Chouliaras et al.[170].*

4.3.5.3 ChIP analysis demonstrates enhanced MEF2 binding at the HLX-3 enhancer

The EMSA analysis provides useful information on potential binding, and identifies the precise nucleotides involved in MEF2 binding. However, this is a very artificial system. Therefore, we also confirmed the ability of MEF2A, C and D to bind the HLX-3 enhancer by ChIP analysis. First, we looked in HUVECs after stimulation with VEGFA (as described in 4.3.5.1). ChIP-qPCR demonstrated that HLX-3 was preferentially bound by all three MEF2 proteins compared to the IgG control (Figure 45A). [This analysis was done together with Dr Natalia Sacilotto.]

To confirm these observations, I next interrogated a previously published ChIP-seq database of global MEF2C bound regions in HUVECs. This publicly available database also provided the corresponding H3K27ac ChIP-seq data, to identify putative enhancers[284]. [This database was downloaded and transferred onto the Integrative Genomics Viewer 2 (IGV2) platform by Dr Marsha Wallace.] Analysis of this database also confirmed enriched binding for MEF2C specifically over the HLX-3 enhancer region, corresponding to a region of H3K27ac (Figure 45B). In the 200 kb surrounding the *HLX* gene, the only significant MEF2C peak was found at the HLX-3 enhancer, whereas no peaks were found around the other two putative HLX enhancer sequences which did not drive robust expression in sprouting cells (HLX-5 and HLX-35, see Chapter 3 section 3.2.2). A previous paper also investigated the pattern of histone acetyltransferase p300 binding after VEGFA treatment around the *HLX* locus[285]. Enhancers are known to be bound by p300 and CBP, and p300 ChIP-seq has been used to predict enhancers[286]. Figure 45B shows the p300 peak reported by Zhang *et al*., which directly corresponds to the HLX-3 sequence and the region of enriched MEF2C binding[285].

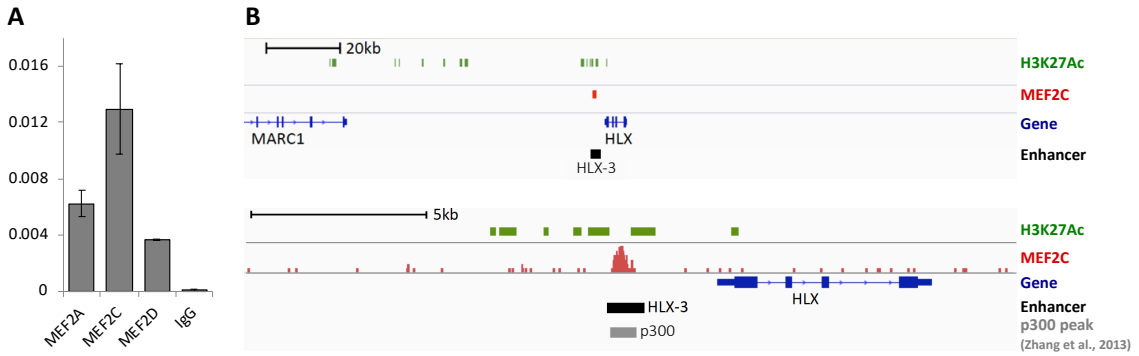


Figure 45: MEF2 factors are bound to HLX-3 enhancer. A) *MEF2 factor binding at the HLX-3 enhancer analysed by ChIP-qPCR after one hour of VEGFA stimulation in HUVECs. The percentage of input DNA for MEF2A, C, D and control IgG is shown. Technical triplicates were used for each sample and error bars indicate the standard error. Graph is representative of three biological replicates.* ***B)*** *MEF2C binding profile around the HLX locus. Red peaks indicate MACS2 bedgraph MEF2C peaks visualized in IGV, green lines indicate statistically significant H3K27ac regions, black lines indicate location of HLX-3, grey line indicates region dynamically bound by p300 and acetylated after VEGFA stimulation shown by Zhang et al.[285]. Figure adapted from Sacilotto, Chouliaras et al.[170].*

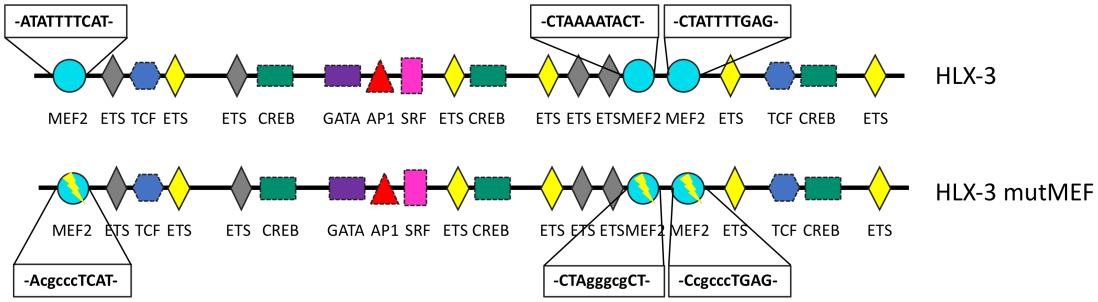
In conclusion, this analysis strongly indicates that MEF2 proteins directly bind the HLX-3 enhancer over the three MEF2 binding motifs previously identified, and that this binding correlates with enhancer-associated marks.

4.3.6 MEF2 binding sites are required for HLX-3:*LacZ* endothelial expression in mice

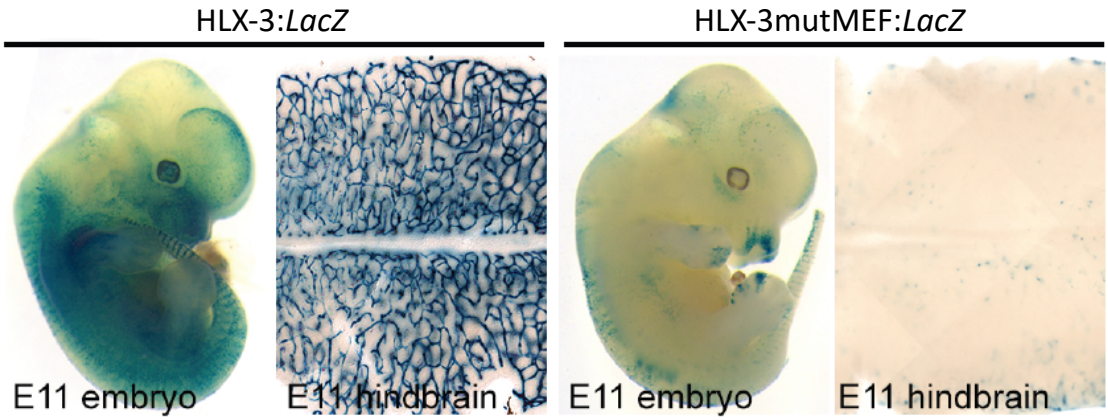
My previous results strongly suggest that *bona-fide* MEF2 binding sites are critical for the expression of hlx-3 enhancer in zebrafish. Next, I wanted to verify the importance of MEF2 binding motifs for the activity of the HLX-3 enhancer during angiogenesis. To investigate, the three verified MEF2 binding sites were mutated in the human HLX-3 enhancer sequence (Figure 46A, these mutations were identical to those shown to ablate binding in EMSA, see section 4.3.5.2). The resultant HLX-3mutMEF:*LacZ* transgene was then tested for its ability to drive *LacZ* expression in transient transgenic mice at E11.5 (Figure 46B). To study angiogenic endothelial cells specifically, I analysed the hindbrain of HLX-3*:LacZ* and HLX-3mutMEF:*LacZ* mice for reporter gene expression. This analysis found that while the wild-type HLX-3:*LacZ* transgene directed robust angiogenic expression in the vast majority of transient transgenic E11 embryos, mutation of the MEF2 motifs resulted in the near-total loss of transgene expression, and the absence of reporter gene activity in the angiogenic hindbrain (Figure 46B and C).

Similar to the zebrafish results presented in section 4.3.4, MEF2 binding sites are thus absolutely required for the angiogenic-specific expression of HLX-3 in the mouse angiogenic hindbrain.

A



B



C

Transgene	Number of transgenic mice	Mice with any detectable X-Gal expression	Mice with X-Gal expression in angiogenic cells
HLX-3:*LacZ*	11	9	7
HLX-3mutMEF:*LacZ*	10	10	0*

Figure 46: MEF2 binding motifs are absolutely required for HLX-3 driven expression. A) *Schematic illustration of the wild-type and mutated HLX-3 enhancer driving LacZ reporter gene expression. ETS binding sites in yellow were bound by ETS in EMSA, while ETS binding sites in grey did not bind ETS family proteins in EMSAs. Transcription factor binding sites marked with a dotted line were not verified in vitro.* ***B)*** *Whole-mount embryos and hindbrains of representative E11.5 X-Gal-stained transient transgenic embryos expressing HLX-3:LacZ and HLX- 3mutMEF:LacZ transgenes.* ***C)*** *Summary of reporter gene expression detected in E11.5 mice transgenic for the HLX-3:LacZ and HLX-3mutMEF:LacZ transgenes. * denotes transgenic mouse that expressed LacZ throughout embryo in all tissues, including all endothelial cells. Figure adapted from Sacilotto, Chouliaras et al.*[170].

4.4 Mef2 factors bind and regulate other angiogenic genes

4.4.1 MEF2A and MEF2C directly bind the DLL4in3 angiogenic enhancer

I next determined whether MEF2 factors bound to enhancers of other tip cell enriched genes. Concurrent to my analysis of the HLX-3 enhancer, a second angiogenic enhancer had been characterized in our laboratory by Dr Natalia Sacilotto. This enhancer, known as Dll4in3, was previously identified as an arterial-specific enhancer[167], but was later found to also be active in angiogenic endothelial cells (Figure 47). This is consistent with the known expression of *Dll4* in both arterial and angiogenic endothelial cells[257].

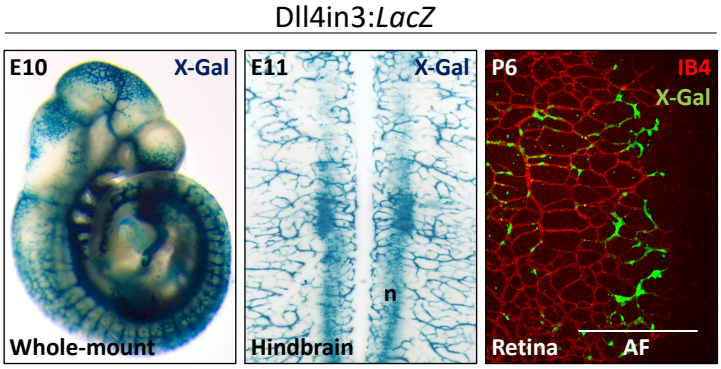


Figure 47: Dll4in3 enhancer drives LacZ expression in endothelial cells in the E10 embryo, E11 hindbrain and P6 retina during angiogenesis. *Representative image of the Dll4in3:LacZ E10 embryo, with reporter gene expression (blue) in arterial and sprouting endothelial cells. In the E11 hindbrain, Dll4in3 demonstrates LacZ expression in sprouting angiogenic endothelial cells and neuronal tissue (n). The Dll4in3:LacZ P6 post-natal retina shows X-Gal activity (false-coloured green) at the angiogenic front (AF). The vasculature was detected by Isolectin B4 (IB4 in red). Data were provided by Dr Natalia Sacilotto. The figure was adapted from Sacilotto, Chouliaras et al.[170].*

I next investigated whether MEF2 factors also bound the Dll4 enhancer. In concordance with my hypothesis, the DLL4in3 enhancer region also showed a robust enrichment of MEF2C binding (Figure 48A). We wanted to verify these findings by ChIP-qPCR (Figure 48B, done in collaboration with Dr Natalia Sacilotto). [ChIP-qPCR experiments were mainly performed by Dr Natalia Sacilotto.]

Therefore, MEF2 binding to HLX-3 and DLL4in3 enhancers was measured *in vitro* before and after stimulation with VEGFA. Interestingly, MEF2A/C were already bound to the HLX-3 and DLL4in3 enhancer prior to VEGFA stimulation (Figure 48B). As with HLX-3, MEF2C was the most bound MEF2 factor, while MEF2D was detected at only very low levels. This is consistent with the Western Blot analysis in section 4.3.5.1, where *MEF2A, C* and *D* were already detectable before VEGFA stimulation. These results suggest that MEF2 proteins are bound to the angiogenic DLL4in3 and HLX-3 enhancers independent of VEGFA stimulation.

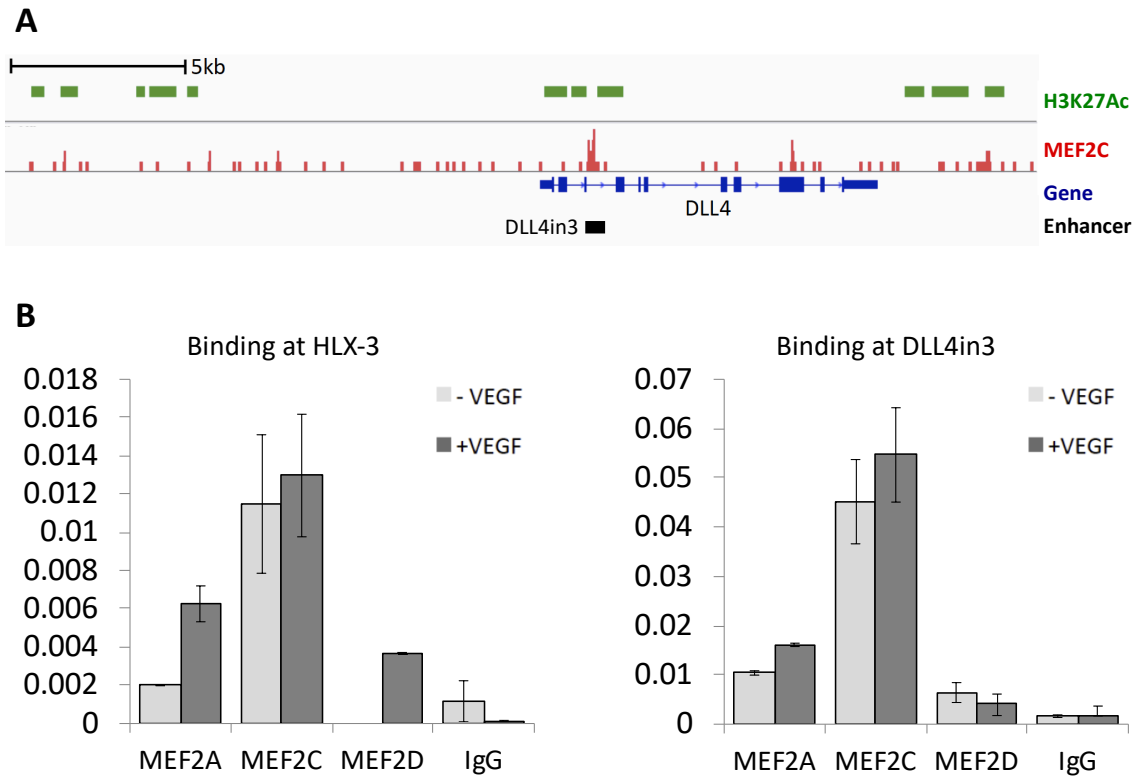


***Figure 48: ChIP reveals binding of MEF2 factors to HLX-3 and DLL4in3 enhancer**. **A)** MEF2C binding profile around the DLL4 locus. Red peaks indicate MACS2 bedgraph MEF2C peaks visualized in IGV, green lines indicate statistically significant H3K27ac regions, black lines indicate location of the DLL4in3 enhancer. **B)** MEF2 factor binding at the HLX-3 and DLL4in3 enhancer analysed by ChIP-qPCR before and after one hour of $VEGFA_{165}$ stimulation in HUVECs. The percentage of input DNA for MEF2A, C, D and control IgG is shown. MEF2D binding at HLX-3 was not consistent and MEF2D levels were at the border of being detectable. Error bar shows standard error of the mean. Graph is representative of three biological replicates. Figure was adapted from Sacilotto, Chouliaras et al.[170].*

4.4.2 MEF2C binding is enriched around many genes which are upregulated during angiogenesis

We hypothesized that the shared MEF2C binding at both HLX-3 and DLL4in3 angiogenic enhancers was an indication that MEF2 factors regulate angiogenesis more generally. A recent study by Toro *et al.* published a list of genes which are upregulated in the hypersprouting *Dll4*$^{-/+}$ post-natal mouse retinas[114]. Virtually all endothelial cells in *Dll4*$^{-/+}$ post-natal mouse retinas show filopodia extensions and exhibit a tip cell phenotype. Dr Marsha Wallace screened the gene loci of these tip cell enriched genes for the presence of MEF2C binding peaks which coincided with H3K27ac[114,284]. Dr Marsha Wallace discovered that MEF2C peaks were enriched within 200 kb of the TSS of genes that are known to play a role during angiogenesis, including SH3 domain and tetratricopeptide repeat-containing protein 1 (*SH3TC1*), RAS guanine-releasing protein 3 (*RASGRP3*)[287], *ETS1*[288,289] and *ELK3*[290,291].

In contrast, no MEF2C peaks were found around pan-endothelial or venous genes, e.g. *CD31*, *SMAD4*, *NRP2* and *COUP-TFII*. To test whether MEF2C binding can predict gene enhancers that are active during angiogenesis, the human sequence surrounding the MEF2C binding peak was injected to generate mosaic transient transgenic zebrafish. In a joint lab effort, together with Dr Natalia Sacilotto and Dr Svanhild Nornes, we found that the four tested regions surrounding the *SH3TC1*, *RASGRP3*, *ETS1* and *ELK3* loci showed reporter gene expression in sprouting ISVs at 27 - 30 hpf (Figure 49). All sequences contained a functional MEF2C binding site and conserved ETS binding sites, which were hence sufficient to direct endothelial angiogenic-specific expression.

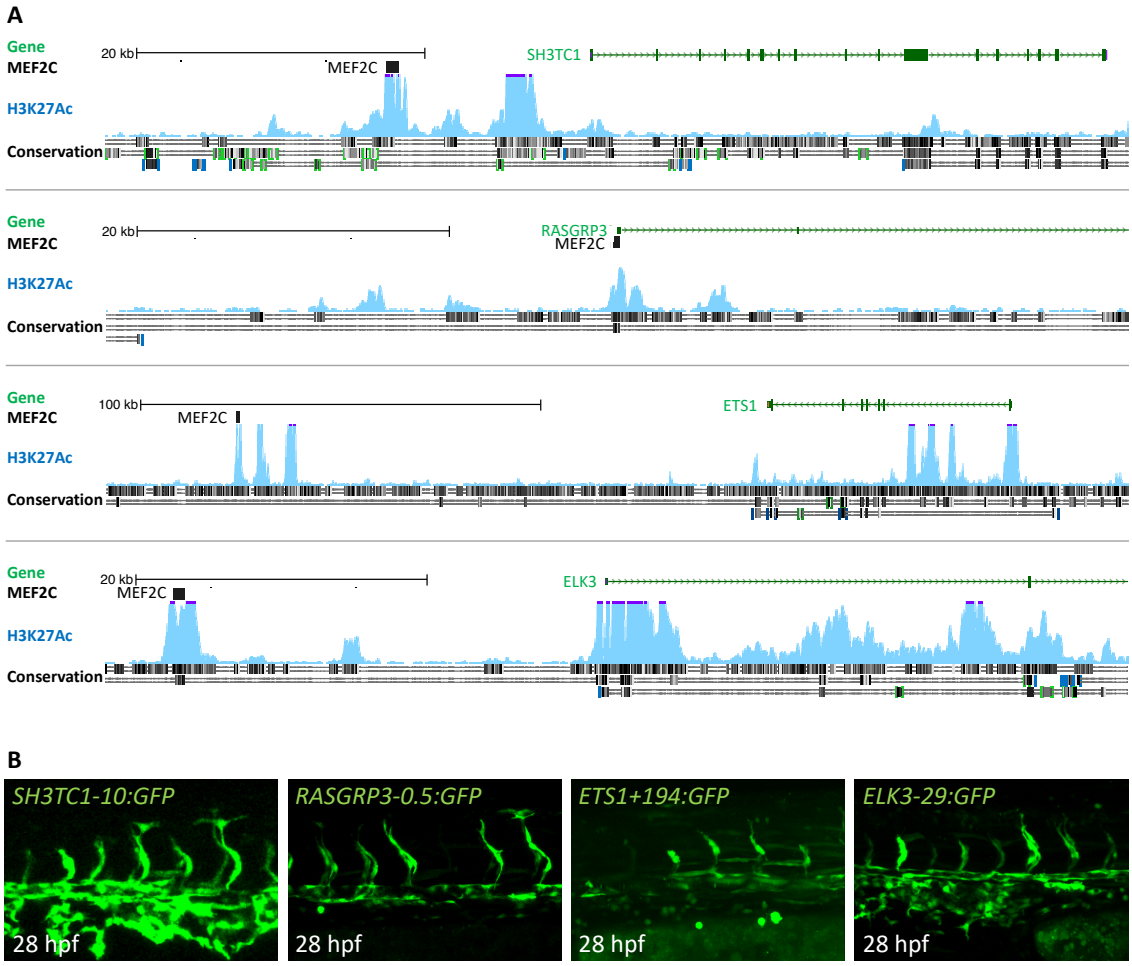


Figure 49: Novel human enhancers identified by MEF2C binding and validation in transient transgenic zebrafish. ***A)*** *Genomic snapshot showing the statistically significant MEF2C binding site (black) within the loci of SH3TC1, RAGRP3, ETS1 and ELK3.* ***B)*** *SH3TC1-10, RASGRP-0.5, ETS1+194 and ELK3-29 were validated in mosaic transgenic zebrafish at 27 - 30 hpf. Blue bedgraph show H3K27ac in HUVECs, black boxes indicate locations of MEF2C peaks and tested enhancers. Conservation is shown for the mouse, chicken and zebrafish sequence. RASGRP3 is likely a promoter, but start ATG codon was not included in the RASGRP3-0.5:GFP sequence.*

4.5 MEF2 factors regulate sprouting angiogenesis during murine development

4.5.1 *Hlx* and *Dll4* expression is reduced in *Mef2A/C^{iECKO}* mice

My results so far clearly demonstrate that mutation of MEF2 binding motifs within the HLX-3/hlx-3 enhancers ablates their ability to drive reporter gene activity during angiogenesis in transgenic zebrafish and mice. I now wish to investigate whether MEF2 factors play a role in angiogenesis more generally. Loss of MEF2C in endothelial cells has no reported effects on vascular development in embryos[150,292]. However, loss of MEF2C resulted in increased and aberrant vascular sprouting in post-natal retina after oxygen-induced retinopathy, closely resembling the results seen in *Dll4*-deficent mice after similar insult[150]. MEF2 factors are known to be functionally redundant in other tissue types[151] and I have shown that multiple MEF2 factors are expressed in angiogenic endothelial cells. Therefore, I next investigated whether depletion of multiple MEF2 factors had an effect on angiogenesis.

This analysis was done using the inducible endothelial-specific Cdh5(PAC)Cre-ER^{T2} line, in which the *Cdh5* locus activates Cre-ER^{T2} expression in all endothelial cells[293]. Upon tamoxifen injection, tamoxifen and its metabolites bind to the cytosolically located mutant estrogen receptor ligand binding domain (ER^{T2}). This allows Cre DNA recombinase to translocate to the nucleus and excise the floxed region. The Cdh5(PAC)$^{Cre\text{-}ERT2}$ line was crossed into mice carrying both *Mef2C$^{fl/fl}$* and *Mef2A$^{fl/fl}$* alleles, creating Cdh5(PAC)Cre-ER^{T2};Mef2A$^{fl/fl}$;Mef2C$^{fl/fl}$ mice, also denoted here as *Mef2A/C^{iECKO}* mice. In *Mef2A/C^{iECKO}* mice, Cre-mediated recombination excises coding exon 2 of both *Mef2A* and *Mef2C*[294,168].

Exon 2 encodes for the MADS-box and MEF2 domain. So even though *Mef2A* and *Mef2C* are still transcribed and translated, the floxed MEF2 proteins are not functional as they are unable to bind DNA and co-factors. To measure the efficiency of *Mef2A/C* deletion, we measured exon 2 RNA levels by qRT-PCR. [All *Mef2A/C*^{iECKO} mice were provided by Dr Yao Wei Lu.]

We first investigated *Hlx* and *Dll4* expression in *Mef2A/C*^{iECKO} mice in which tamoxifen was injected at P1, P2 and P3. Whole lung extracts were collected from *Mef2A/C*^{iECKO} mice at P5 and P6. As expected when using the Cdh5(PAC)Cre-ER^{T2} Cre-driver line, the extent of *Mef2A/Mef2C* exon 2 mRNA deletion varied between samples[102,295]. Therefore, only lung lysate samples where exon 2 of *Mef2A* and *Mef2C* were reduced more than 50% compared to control wild-type litter mates were used for qRT-PCR. This analysis reveals that *Hlx* mRNA levels were reduced significantly in lung lysates derived from *Mef2A/C*^{iECKO} mice (Figure 50A). *Dll4* levels were reduced significantly in *Mef2A/C*^{iECKO} lung lysates at P5, but not P6. Given that we analysed mRNA levels in lung homogenates which included non-endothelial cells and that both *Hlx* and *Dll4* are known to be expressed in other cell types, only a partial decrease was expected.

Next, I also investigated whether a loss of *Mef2A/Mef2C* would affect target gene transcription during embryonic angiogenesis. To do this, I analysed the angiogenic hindbrain *of Mef2A/C*^{iECKO} embryos. Although we were unable to investigate HLX levels because of difficulties detecting endogenous murine *Hlx* RNA and HLX protein (see Chapter 3, section 3.4.3), this analysis found that DLL4 protein levels were reduced in all E12 hindbrains three days after induction of gene deletion (Figure 50B and C). Again, the

level of reduction varied significantly between independent samples, potentially due to different levels of reduction of *Mef2A*/*Mef2C*.

In conclusion, despite the inclusion of non-vascular cells in lung and hindbrain homogenates, *Dll4* and *Hlx* levels were significantly reduced in *Mef2A/C*^{iECKO} mice. These results support a key role for MEF2 factors in the regulation of endogenous *Hlx* and *Dll4* during sprouting angiogenesis.

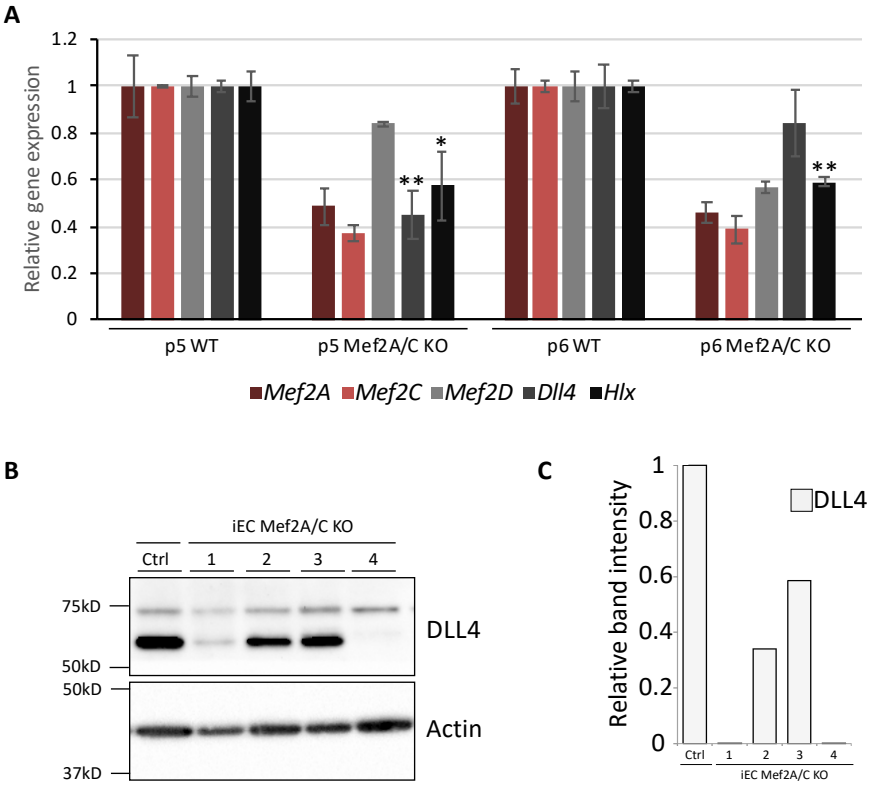


Figure 50: Effect of Mef2A/C knockout on Dll4 and Hlx expression. A) *Relative gene expression levels of Hlx, Dll4, Mef2A, Mef2C, and Mef2D analysed by qRT-PCR in five different lungs taken from control (Cre-negative; Ctrl) and knockout (iEC Mef2A/C KO) pups at P5 and P6. Cre recombination induced by tamoxifen injection was performed once every 24 hours at P1, P2 and P3 as previously described*[245]*. Hlx was reduced significantly (at P5, P = 0.01 and at P6, P = 0.0018). Dll4 was reduced significantly at P5 (P = 0.008), but not at P6 (P = 0.23). Asterisks denote statistically significant difference compared with control (* = P < 0.05, ** = P < 0.01). Samples were normalized to ActB. Three technical replicates were used for each pup. Error bars indicate standard error. RNA from lung lysate samples were extracted jointly with Dr Natalia Sacilotto. Dr Natalia Sacilotto also performed qRT-PCR on Mef2A, C and Dll4.* ***B)*** *Induced endothelial cell-specific (iEC) Mef2A/C double knockouts (Mef2A/C KO) show a significant reduction in DLL4 expression, as detected by Western blot of E12 hindbrain extracts. Cre recombination was induced by tamoxifen injection in pregnant mice at E9. Ctrl = Cre-negative; (1 - 4) four different Mef2A/C*^{iECKO} *littermates. Note the variability in Dll4 reduction among littermates, as expected from Cre recombination in response to tamoxifen.* ***C)*** *Comparative band intensity was normalized to Actin levels and made relative to wild-type. Relative protein levels were measured by quantifying band intensity of the DLL4 Western Blot. Figure was adapted from Sacilotto, Chouliaras et al.*[170]*.*

4.5.2 Defect in angiogenic sprouting in post-natal *Mef2A/C*^{iECKO} retina

Given that both *Dll4* and *Hlx* expression was reduced in *Mef2A/C*^{iECKO} mice, we wanted to establish whether *Mef2A/C* knockout has an effect on sprouting angiogenesis more generally. We therefore analysed the highly angiogenic post-natal retina.

Endothelial-specific *Mef2A/C* knockout was induced at P1, P2 and P3 by tamoxifen injection as before. Retinas were collected from control wild-type littermates and *Mef2A/C*^{iECKO} mice at P5 and P6 and stained with endothelial marker Isolectin B4. Induced endothelial deletion of *Mef2A* and *Mef2C* resulted in reduced retinal angiogenesis at both time points (Figure 51A). *Mef2A/C*^{iECKO} retinas displayed a significantly reduced number of tip cells at the angiogenic front (tip cells were counted manually and are marked with asterisks, see Figure 51A and B). We also observed significantly fewer branching points and reduced vascular outgrowth (Figure 51B). These results indicate that the loss of *Mef2A/C* results in an overall reduction of vascular coverage and angiogenic sprouting. Considering that not all retinal vascular growth was abolished, more detrimental effects might have been compensated by MEF2B and MEF2D. [Quantitative phenotypic analysis of *Mef2A/C*^{iECKO} retinas was performed by Mark Shipman and Dr Martin Fritzsche].

In summary, we have shown that MEF2A and MEF2C do not only bind angiogenic-specific enhancers and thereby induce *Hlx* and *Dll4* expression *in vivo*, but also affect sprouting angiogenesis in general.

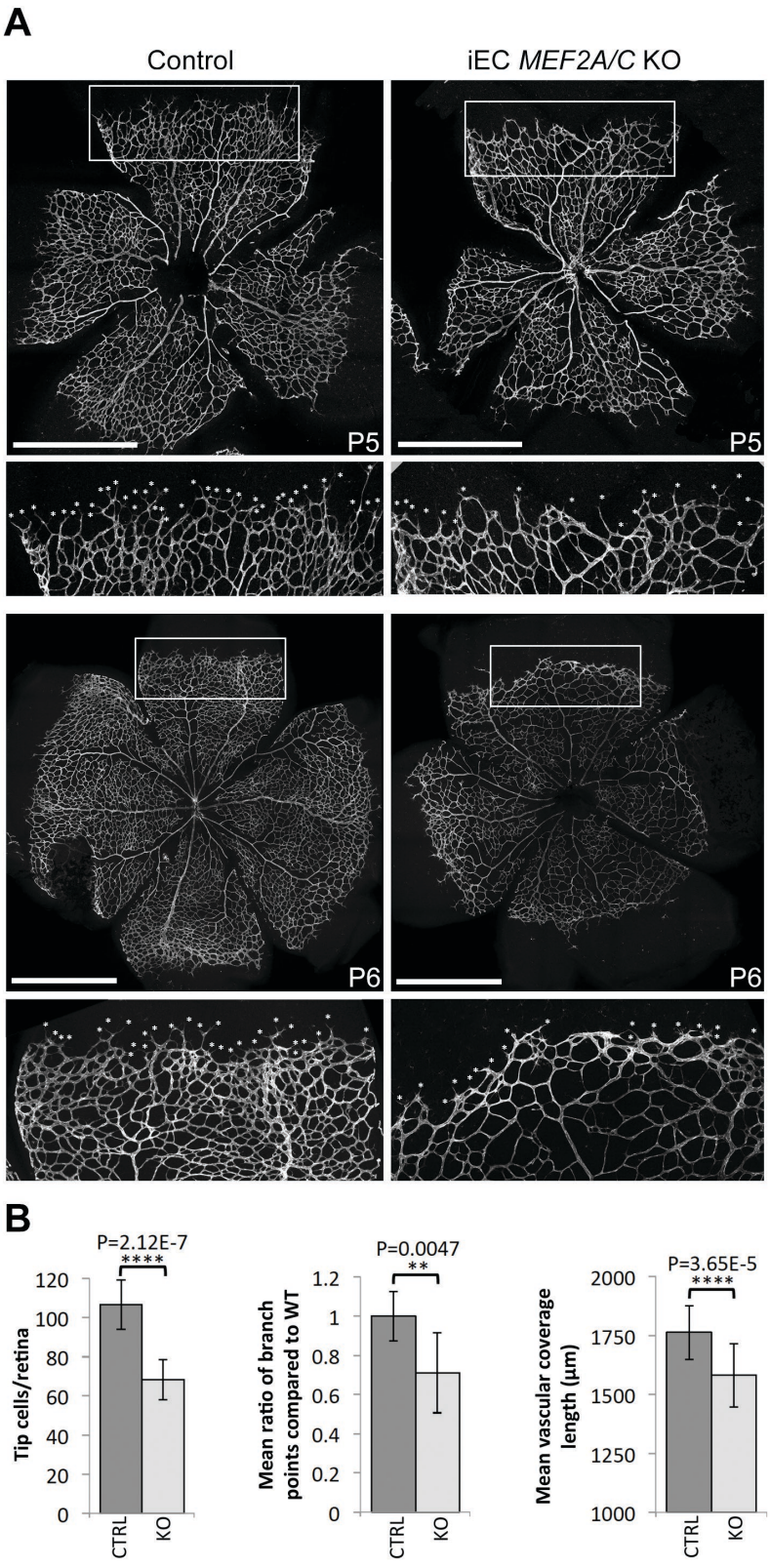


Figure 51: Post-natal retinas of Mef2A/C^{iECKO} mice show reduced sprouting. A) *Representative P5 and P6 retinas taken from control (Cre-negative) and iEC Mef2A/C knockout (KO) pups and stained with Isolectin B4. Tip cells were detected by filopodia and are indicated by asterisks. The white box indicates region taken for zoom image. Bars correspond to 1 mm. Cre recombination induced by tamoxifen injection was performed once every 24 hours at P1, P2 and P3 as previously described[245].* ***B)*** *Graphs demonstrating the mean number of tip cells, the ratio of the mean number of branch points, and the mean outgrowth length as measured by actual distance covered by the growing vessel from the centre of the retina. Statistical analysis was performed on seven wild-type and nine knockout pups pooled from two different litters. Figure was taken from Sacilotto, Chouliaras et al.[170].*

4.6 Discussion

4.6.1 HLX-3 is downstream of VEGF signalling

In the first part of this chapter, my data verified that *hlx-3:GFP* expression was supressed by Notch but enhanced by Vegf signalling in zebrafish. hlx-3 enhancer was dependent on Vegf-signalling, which is concordance with published data on endogenous *HLX*/*hlx1* regulation[164,264]. My results also demonstrated that HLX-3 enhancer directs *LacZ* gene expression in mice in pathological Vegf-stimulated neo-angiogenesis during melanoma growth.

Even though expression of angiogenic *Vegfr2* can get activated by mechanotransduction, neither shear stress nor the ablation of blood flow had an effect on *hlx1* or *HLX* expression. These data contradict Prahst *et al.* who have reported HLX protein upregulation in response to shear stress in human endothelial cell culture and Hlx protein downregulation in the absence of blood flow in quail embryos[161]. There are several possible explanations for the different results. First, *in vivo* experiments were performed in quail embryos and not in zebrafish. Second, human abdominal aortic endothelial cells (HAAECs), and not HUVECs, were used for *in vitro* experiments. Last, Prahst *et al.* used lower levels of laminar flow. The contradicting results may therefore reflect the different methods used. In conclusion, while shear stress may reinforce *hlx1* and hlx-3:GFP expression, my experiments showed that flow is not required for *HLX*/*hlx1* and hlx-3:GFP expression.

4.6.2 MEF2 factors are required for sprouting angiogenesis and regulate enhancer HLX-3

The main finding of this chapter is that *HLX-3/hlx-3* lie downstream of VEGFA signalling, and that its angiogenic-specific expression pattern is mediated by MEF2 transcription factors. MEF2 binding sites were essential for *HLX-3/hlx-3* transgene expression in sprouting endothelial cells. In addition, MEF2 factors directly bound HLX-3 and regulatory elements of other genes known to be implicated in sprouting angiogenesis (DLL4, SH3TC1, RASGRP3[287], ETS1[288,289] and ELK3[290,291]). This MEF2C binding was able to accurately identify novel angiogenic regulatory elements driving expression in sprouting endothelial cells. Knockout of MEF2 factors *in vivo* resulted in reduced levels of *Hlx* and Notch-ligand *Dll4*. Lastly, I have also clearly shown that *Mef2A* and *Mef2C* are required for correct angiogenic sprouting in the post-natal retina.

The decreased levels of sprouting angiogenesis in *Mef2A/C*^{iECKO} retinas contrasts with the increase in angiogenesis and vessel density seen after *Dll4* reduction[296]. This disparity may be because MEF2 factors not only regulate *Dll4* but also the activity of other angiogenic genes, such that the overall loss of these genes cumulatively has an anti-angiogenic effect[164,127,291]. It is of note that endothelial-specific knockout mice of *Mef2C* alone did not show disturbed vascular development in the post-natal retina[150]. Similar studies in skeletal muscle found that only the combined deletion of *Mef2A/C/D* ablated muscle regeneration[151], and it is therefore likely that redundancy is also the cause here. In cell culture, overexpression of *Mef2C* resulted in decreased sprouting, implicating an anti-angiogenic effect[297]. These knockout experiments reiterate the importance of context-specificity (*in vivo* vs. *in vitro* experiments) and redundancy.

4.6.3 ETS factors direct gene expression in all endothelial cells

After I published the results of Chapter 3 and Chapter 4 (jointly with Dr Natalia Sacilotto including the work on Dll4 regulation) in Genes and Development[170], a different group also published a paper looking at the HLX-3 enhancer. Although this study did not generate stable mouse and zebrafish transgenic lines showing HLX-3 activity was specific to sprouting endothelial cells, they were able to make limited conclusions about its regulation from work in tissue culture. Further, they also used a CRISPR-mediated deletion of one copy of the HLX-3 enhancer to demonstrate reduced responsiveness of *HLX* expression to VEGFA in cell culture[298]. This paper, Fish *et al.*[298], did not uncover a role for MEF2, instead focusing on the importance of ETS factor ERG for HLX-3 activity. As previously discussed, all genes expressed in endothelial cells require ETS factors for transcriptional activation. Independent of the sub-endothelial cell type, ETS transcription factors have been shown to be essential for all developing blood vessels[218,282,152]. In the same way, ETS factors were required for the hlx-3 enhancer. However, my work demonstrates that the enhancer was unable to drive angiogenic expression when the MEF2 motif was ablated even if the ETS motifs are intact, further demonstrating that ETS factors alone do not regulate HLX. No functional MEF2 binding sites were found in non-angiogenic endothelial enhancers (such as Dll4-12[170], EphB4-2 or CoupTFII-965[179]), whereas ETS binding sites were abundant in those arterial and venous enhancers. Additionally, the ETS factor ERG was abundantly expressed in all endothelial cells in the post-natal retina and not restricted to actively proliferating and sprouting cells. These results point towards a cooperative crucial role for MEF2 and ETS factors with ETS factors providing pan-endothelial information, whereas MEF2 factors direct angiogenic-specificity.

4.6.4 Increase of H3K27 acetylation and p300 binding at angiogenic-specific enhancers

The addition of VEGFA to media rapidly increased the transcription of *HLX* and *DLL4* in HUVECs, concomitant with a delayed induction of *MEF2A/C/D*. Despite this delayed induction of MEF2 factors, HLX-3 and DLL4in3 were already bound by MEF2 factors independent of VEGFA stimulation. MEF2 factors are hence already present in quiescent endothelial cells, bound to angiogenic enhancers and immediately induce transcription in response to VEGFA. We speculated that a VEGFA-responsive transcriptional co-activator may be recruited allowing MEF2 to activate *HLX* and *DLL4* transcription. Indeed, it has been previously published that VEGFA results in p300 recruitment at the HLX-3 enhancer concomitant with an increase of H3K27 acetylation[285]. Subsequently, *HLX* was upregulated in response to VEGFA and this induction was ablated in the presence of p300 inhibitor C646[285].

Experiments performed in our laboratory by Dr Natalia Sacilotto confirmed these findings at the DLL4in3 locus[170]. In the absence of VEGFA, class IIa HDACs (HDAC4, HDAC7, and HDAC9) and MEF2 factors form a complex[299,300]. Class IIa HDACs thereby repress MEF2C and prevent binding of transcriptional co-activators to the angiogenic-specific enhancer[142]. VEGFA stimulation leads to phosphorylation and subsequent cytoplasmic accumulation of HDACs[301,302]. Chaperone protein 14-3-3 binds the phosphorylated HDAC and mediates the nuclear export[303]. This is concurrent with the recruitment of p300, and consequently the transcriptional activation of MEF2 during sprouting angiogenesis[302,301]. In addition to the transcriptional activation of MEF2 factors, phosphorylation and p300-

mediated acetylation of histone tails causes chromatin relaxation allowing MEF2-dependent transcription to occur.

Fish *et al.* proposed a similar mechanism, in which VEGFA-ERK-dependent ERG phosphorylation and activation resulted in p300 recruitment to the DLL4in3 and HLX-3 enhancers and subsequent induction of *DLL4* and *HLX*[298]. CBP/p300 in turn then acetylate histones and ETS family members which additionally stimulates RNA Pol II release through recruitment of BRD4 and thus transcription[288]. Taken together, our and the conclusion of Fish *et al.* are not mutually exclusive, suggesting that MEF2 and ETS factors interact to coordinate transcription of VEGFA-dependent angiogenic-specific genes.

4.6.5 VEGF signalling and MEF2 factors in sprouting angiogenesis

A variety of signalling pathways are downstream of VEGFA receptor VEGFR2. Our experiments have not directly addressed the exact signalling pathway by which VEGFA enhances the transcriptional activation of MEF2 and regulates angiogenic-specific genes. First, VEGFA enhances the transcriptional activity of MEF2 factors. This effect is likely to be mediated *via* MAP kinases p38, ERK or ERK5/big MAPK (BMK) and calcium signalling. Phosphorylation of the transactivation domain of MEF2 enhances its transcriptional activity[304], while phosphorylation of HDAC leads to dissociation and hence a release of its repressive effect (see previous section)[303]. The second effect of VEGFA signalling is the delayed induction of *MEF2* RNA. *MEF2C* has been described as a late response gene in the context of VEGFA stimulation[288]. Increase in *MEF2* RNA and therefore protein levels in the angiogenic cell then reinforces MEF2-dependent transcription. We have shown that enhancers in the proximity of ETS factors *ETS1* and *ELK3* are bound by MEF2C.

Interestingly, it has been demonstrated that *MEF2C* expression in endothelial cells itself was mediated by ETS1[152,288]. This feed forward loop of *MEF2C* and *ETS1* activation would result in increased levels of MEF2C and ETS1 during sprouting angiogenesis. Both would then be able to induce transcription of other sprouting-specific genes, such as *DLL4*, *HLX*, and *RASGRP3*.

4.6.6 Outlook

4.6.6.1 *HLX* as an immediate early gene

Based on the dynamics of *HLX* expression (rapid induction and transient nature) and its p300-dependency, I hypothesize that *HLX* is an immediate early gene. Immediate early genes are transiently and rapidly induced in the absence of newly formed proteins. This means that all required transcriptional machinery is already present in the cell or poised at the enhancer/promoter. Immediate early gene induction is concomitant with hyperacetylation and phosphorylation of surrounding histones. I therefore propose to study H3K9acS10ph both at the HLX-3 enhancer and its promoter in response to VEGFA.

In line with our MEF2C ChIP and ATAC-seq experiments (see Chapter 5), it is plausible that HLX-3 enhancer is already bound by transcription factors in all endothelial cells and that RNA Pol II is already stalled at the promoter/enhancer ready to start transcription just waiting for the incoming activating signal. Indeed, VEGFA-dependent *HLX* transcription has been shown to be highly dependent on RNA Pol II pause release[20].

4.6.7 Conclusion

We have demonstrated that MEF2 factors directly activate the expression of *DLL4* and *HLX* and are the transcriptional effectors of VEGFA signalling during sprouting angiogenesis. Our findings might therefore provide potential therapeutic benefits for the future. Transcription factors are difficult to target due to their ubiquitous expression pattern. However, the interaction between MEF2 factors and inhibitory HDAC or activating p300 can selectively be targeted with small molecule modulators providing a therapeutic target to inhibit sprouting angiogenesis specifically[170].

5 ATAC-seq and RNA-seq on arterial, venous and sprouting endothelial cells in zebrafish

5.1 Rationale: Identification of enhancers in zebrafish

5.1.1 Do different endothelial cell types have differentially accessible chromatin?

So far, efforts to understand gene regulation during endothelial differentiation in zebrafish have focused on single genes or enhancers. Additionally, endothelial cells have frequently been treated as a homogenous population, ignoring the fact that they comprise functionally and morphologically different arteries, veins and angiogenic cells *in vivo*[305,306]. The goal of this chapter was to take advantage of the zebrafish models generated in earlier chapters, and by other research in the De Val lab to instead generate a genome-wide map of all active promoter and enhancer regions in different sub-types of endothelial cells *in vivo.* This would be done through the identification of open chromatin regions (ATAC-seq) in endothelial cells expressing different types of enhancer:reporter transgenes, and to relate these open chromatin regions to transcription (RNA-seq).

The zebrafish genome is not as well annotated as the genome of other model organisms[307]. So even though histone modifications derived from whole zebrafish are publicly available, endothelial-specific information is lacking[308,309]. Until now, it was therefore only possible to identify zebrafish vascular enhancers based on phylogenetic footprinting, but most enhancers are not under selection pressure and hence not conserved[43,310]. My first aim was to provide a resource which allows the easy identification of putative endothelial enhancers in the zebrafish genome. It is already known that all experimentally verified regulatory elements are marked by increased DNA accessibility[311]. Given the low number of endothelial cells in the zebrafish embryo, I wanted to assess

chromatin accessibility using ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing). In contrast to DNase I hypersensitivity assays, ATAC-seq allows the identification of regulatory elements using only a few thousand cells[312]. During ATAC-seq, adapters are inserted into accessible chromatin, which allows genome-wide sequencing and thus mapping of open chromatin regions, including promoters and enhancers.

My second aim was to investigate whether specific sub-types of endothelial cells have differentially open chromatin early during development, or whether the chromatin landscape is similar for all types of endothelial cells. It is already known from comparing ATAC-seq data across distinctly different cell types that chromatin is differentially accessible[313]. But it has not been investigated whether differential opening of regulatory regions contributes to heterogenic gene expression within a single cell type. It may be hypothesized that regulatory regions are accessible in all endothelial cells due to the binding of a master regulator such as ETS, but only active in the single sub-endothelial cell type due to a requirement for additional factors. Alternatively, these regions may have a closed chromatin confirmation until all regulatory factors are available for binding.

5.2 Sorting sub-endothelial cell types

5.2.1 Zebrafish lines used for RNA-seq and ATAC-seq at 27 hpf and 43 hpf

Cells from different stable transgenic zebrafish lines were sorted by FACS to isolate four sub-endothelial cell populations representing sprouting, arterial, non-arterial and venous cells (Figure 52A). Isolated endothelial cells from the four sub-populations were then used for qRT-PCR, RNA-seq and ATAC-seq.

I used three transgenic zebrafish lines, which have been generated and characterized in our laboratory, where an enhancer directs transgene expression in the specific sub-endothelial cell type. First, I used the angiogenic-specific *tg(hlx-3:GFP)* zebrafish line (described in Chapter 3), which is activated by the VEGFA-MEF2 pathway (described in Chapter 4)[170,176]. I also used a second VEGFA-MEF2-dependent *tg(Dll4in3:GFP)* zebrafish line (as described in Chapter 4, section 4.4)[170]. In addition to its activity in sprouting endothelial cells, Dll4in3 enhancer is also active in arterial cells *via* RBPJ and SOXF[167]. I crossed this line with *tg(kdrl:HRAS-mCherry),* a zebrafish line generated by Jin *et al.* where the *kdrl* promoter directs transgene expression in all endothelial cell[184]. Last, I used a BMP-SMAD-dependent *tg(Ephb4-2:GFP)* zebrafish line, which is specific to venous endothelial cells[179].

At 27 hpf, GFP^+ cells from *tg(hlx-3:GFP)* were sorted to get actively sprouting endothelial cells (Figure 52B). GFP^+ cells from *tg(Dll4in3:GFP;kdrl:HRAS-mCherry)* were sorted to get sub-endothelial cell populations corresponding to sprouting and arterial cells (termed arterial for simplicity). GFP^-/$mCherry^+$ cells were also sorted from

tg(Dll4in3:GFP;kdrl:HRAS-mCherry) embryos to get non-arterial cells. GFP^+ cells from *tg(Ephb4-2:GFP)* were also sorted at 27 hpf to get venous cells. While non-arterial Dll4in3:GFP^- x kdrl:HRAS-$mCherry^+$ are mainly found in the caudal plexus and cranial vasculature but not the cardinal vein, venous Ephb4-2:GFP^+ cells are located in the caudal plexus and cardinal vein. As shown in Chapter 2 Figure 9, the average of collected sprouting (hlx-3:GFP^+), venous (Ephb4-2:GFP^+), arterial (Dll4in3:GFP^+) and non-arterial (Dll4in3:GFP^- x kdrl:HRAS-$mCherry^+$) endothelial cells constitute 0.32 - 2.12% of all embryonic cells in zebrafish at 27 hpf. This is in concordance with the observation that all endothelial cells constitute less than 5% of all cells in the zebrafish at this time point[306].

We also collected arterial, non-arterial and venous endothelial cells at 43 hpf (Figure 52C). This second time point is outside the remit of this thesis, so I shall only discuss the results of the first time point which is when angiogenesis and endothelial fate decisions are actively occurring.

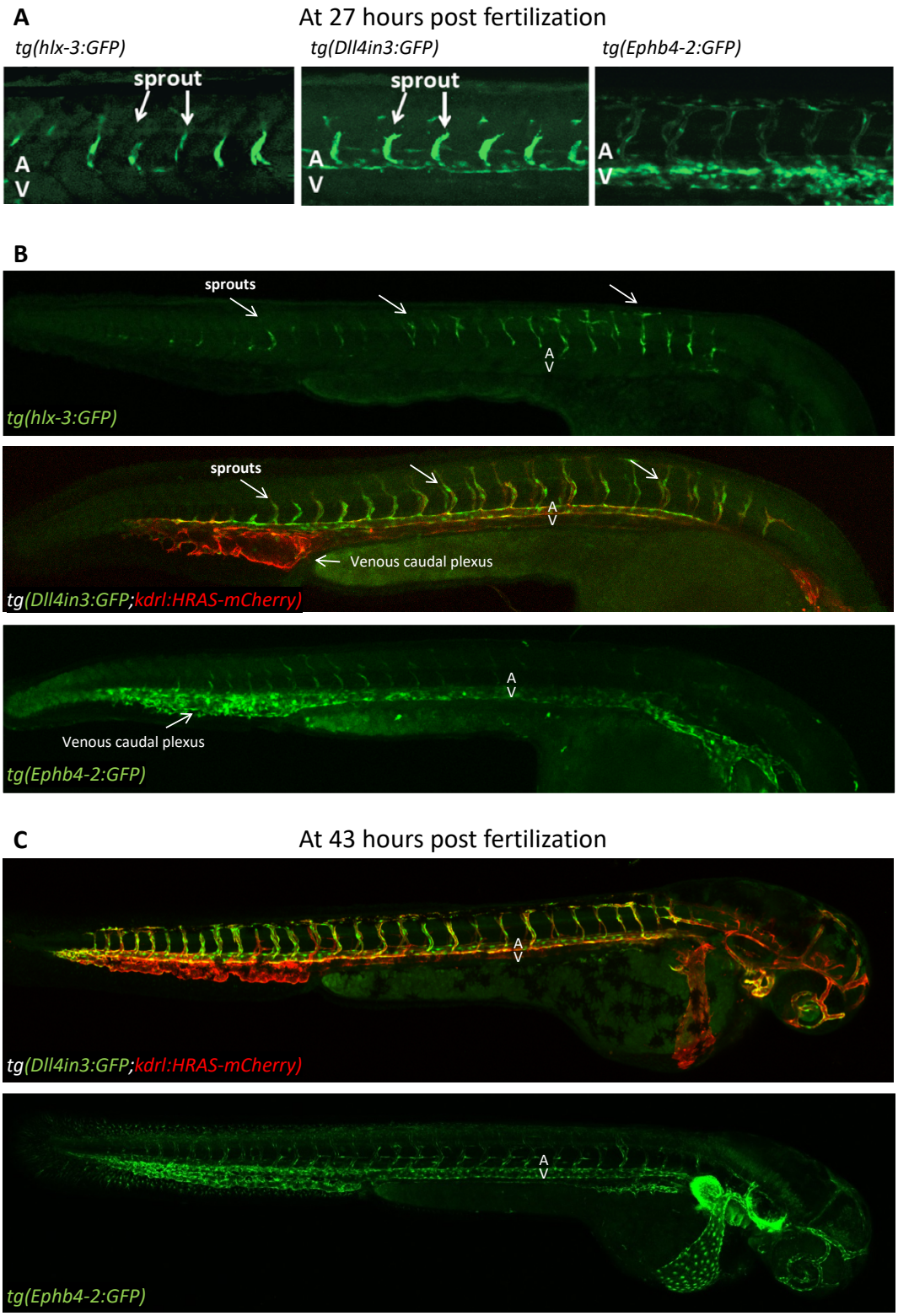


Figure 52: Stable transgenic zebrafish lines used for ATAC-seq and RNA-seq. A) *Zoom of trunk of stable transgenic zebrafish lines at 27 hpf. Tg(hlx- 3:GFP) directs expression specifically to sprouting endothelial cells in the intersegmental vessels (sprouts). Tg(Dll4in3:GFP) directs expression in the dorsal aorta (A) and sprouts. Tg(Ephb4-2:GFP) is expressed in the posterior cardinal vein (V).* ***B)*** *Tg(hlx-3:GFP) as described before. Tg(Dll4in3:GFP) crossed with tg(kdrl:HRAS-mCherry) which marks all endothelial cells weakly but is strongly expressed in the venous caudal plexus. Tg(Ephb4-2:GFP) is specifically expressed in the posterior cardinal vein (V) and caudal venous plexus.* ***C)*** *Stable transgenic zebrafish lines at 43 hpf. Tg(Dll4in3:GFP) is active in arterial cells, namely intersegmental vessels and dorsal aorta (A). Tg(kdrl:HRAS-mCherry) marks all endothelial cells and the endocardium. Tg(Ephb4-2:GFP) directs the GFP reporter gene expression in the posterior cardinal vein (V), intersegmental veins and the endocardium.*

5.2.2 Sorted arterial and venous cells largely cluster as expected

To ascertain that arterial and venous endothelial cells were successfully sorted, qRT-PCR was performed comparing the expression levels of arterial markers [*dll4*, eph receptor B2a (*ephb2a*) and T-box 20 (*tbx20*)] and of venous markers [*flt4*, *ephb4a* and DAB adaptor protein 2 (*dab2*)] in RNA isolated from Dll4in3:GFP$^+$ and Ephb4-2:GFP$^+$ cells (Figure 53). All markers were enriched in Dll4in3:GFP$^+$ and Ephb4-2:GFP$^+$ cells compared to non-endothelial kdrl:HRAS-mCherry$^-$ cells (data not shown).

As expected, Dll4in3:GFP$^+$ cells show strong expression of all arterial markers. Venous Ephb4-2:GFP$^+$ cells only showed enrichment for *tbx20* and *dab2*, but not *ephb4a*. The lack of clustering of venous marker *ephb4a* could be a result of the chosen early time point when venous and arterial cells are still differentiating. Pan-endothelial marker *kdrl* was expressed at higher levels in arterial Dll4in3:GFP$^+$ than in venous Ephb4-2:GFP$^+$ cells which corresponds to its expression pattern. As discussed in the previous section*, kdrl* is only weakly expressed in the posterior cardinal vein, where Ephb4-2:GFP$^+$ is strongly expressed.

To summarize, even though arterial markers do not differ significantly in isolated Dll4in3:GFP$^+$ and Epb4-2:GFP$^+$ cells (and *vice versa* for venous markers), they did follow the expected trend. Moreover, multidimensional scaling and correlation analysis revealed differences between the four sub-endothelial cell populations and clustering of the three biological replicates (see Chapter 2, Figure 11 and Figure 13).

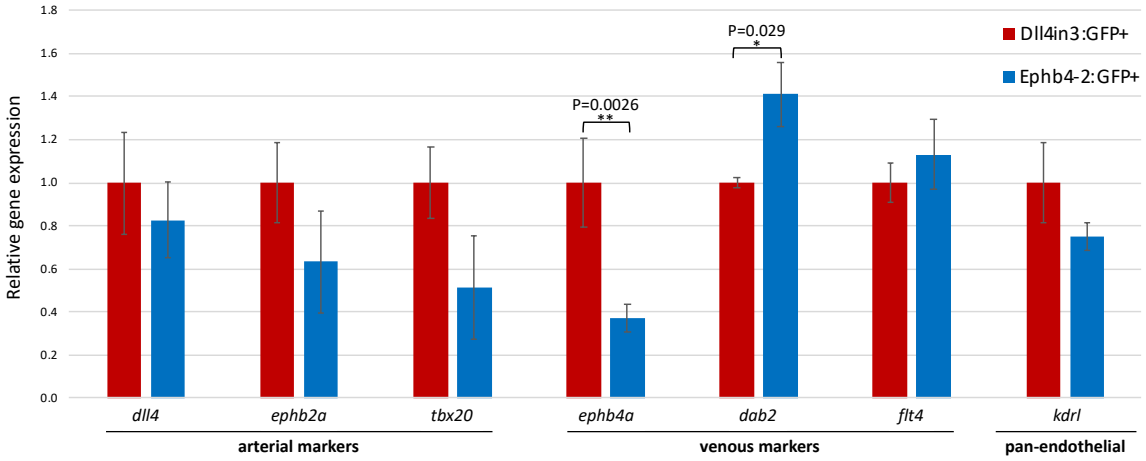


Figure 53: qRT-PCR on FAC sorted cells at 27 hpf. *Results are representative of four biological replicates. Three technical replicates were used for each sample and error bars indicate the standard error. Transcripts were normalized to eIF1a and ubiC. Expression of Ephb4-2:GFP⁺ was normalized using Dll4in3:GFP⁺ as a reference. Arterial markers ephb2a, dll4 and tbx20 are more expressed in Dll4in3:GFP⁺ than in Epb4-2:GFP⁺ cells. Venous markers dab2 and flt4 are overexpressed in Epb4-2:GFP⁺ cells, but ephb4a was only weakly expressed in venous Epb4-2:GFP⁺ cells. Please note that in zebrafish, several genes have two paralogues due to whole genome duplication during evolution 350 million years ago. Here, only one of the two ephb2 and ephb4 paralogues were analysed. Endothelial marker kdrl was expressed at higher levels in Dll4in3:GFP⁺ than in Epb4-2:GFP⁺ cells.*

5.3 Search for endothelial-specific enhancers

5.3.1 Accessible chromatin regions identified by ATAC-seq detects known endothelial enhancers

ATAC-seq was performed on all sub-endothelial cell types. Tn5 transposase inserts sequencing adaptors into open accessible chromatin making it possible to identify potential *cis*-regulatory elements. In my analysis, I also included ATAC-seq peaks from kdrl^{+} pan-endothelial (pan-EC) and kdrl^{-} non-endothelial cells (non-EC) which were kindly provided by our collaborator Dr Rui Monteiro and Dr Tomasz Dobrzycki[203]. Chromatin accessibility was more similar between sub-endothelial cells than to non-ECs but this difference was not especially striking, with 74% of sub-endothelial peaks overlap with pan-EC while only 67% overlap with non-EC.

First, I assessed whether peaks were called at loci of known endothelial enhancers. In *tg(Ephb4-2:GFP) and tg(Dll4in3:GFP)* it is the mouse enhancer sequence that drives GFP expression. The murine Ephb4-2 enhancer is not conserved in zebrafish. The Dll4in3 enhancer sequence is conserved in zebrafish, but the Dll4in3 region was not accessible in any of the tested cell types - neither in endothelial nor non-endothelial cells. I next examined the conserved *hlx1* locus. Over the region representing the hlx-3 enhancer, chromatin is accessible in pan-endothelial cells, but not in non-endothelial cells. Interestingly, despite the angiogenic-specific expression of hlx-3, an ATAC-seq peak was called in at least one replicate of all sub-endothelial cell types (Figure 54A). However, angiogenic-specific sub-endothelial cell populations hlx-3^{+} and Dll4in3^{+} had more replicates with significant peaks and the highest peak scores at this sprouting-specific

locus (Table 7). I next analysed peak scores at a previously published venous/lymphatic enhancer located 7 kb downstream of *mafbb* (mafbb+7)[306]. Here, only pan-EC and Ephb4-2$^+$ cells had a significant ATAC-seq peak. At another recently identified venous enhancer, nr2f2-965[179], a significant ATAC-seq peak was called in at least one replicate in each sub-endothelial cell type, but with venous Ephb4-2$^+$ again having the highest peak score.

After having investigated sprouting and venous-specific enhancers, I next compared chromatin accessibility between sub-endothelial cell types at known pan-endothelial enhancers. I first analysed the previously published *etv2* enhancer located 2 kb upstream of the TSS (etv2-2)[314]. Figure 54B shows that all sub-endothelial and pan-endothelial cells had a peak at the etv2-2 enhancer. Unlike the sub-type-specific enhancers, there was no difference in peak score in the sub-endothelial cell types (Table 7). I also analysed ATAC-seq peaks at the previously published pan-endothelial *fli1a* intron 1 enhancer (fli1ain1)[315]. All sub-endothelial cell types, pan-EC, but not non-EC, identified the endothelial-specific open chromatin at the fli1ain1 enhancer. Similarly, as for etv2-2, peak scores were comparable for all sub-endothelial cell types at the pan-endothelial fli1ain1 enhancer.

To summarize, I observed equal ATAC read density at pan-endothelial enhancers between sub-endothelial cell types. But at loci which drive venous or angiogenic-specific expression, peak scores and the number of replicates having a significant called peak were higher in the respective sub-endothelial cell type, although some degree of chromatin accessibility was always seen in all sub-types.

Table 7: ATAC-seq peak scores for each sub-endothelial cell type at known enhancers: *hlx-3, mafbb+7, nr2f2-965, etv2-2 and fli1ain1*[306,314,315,179]*. Peak scores that showed significant ATAC-seq peaks in at least two replicates per sub-endothelial cell type were marked in bold. The presence or absence of a significant ATAC-seq peak in kdrl*$^{+}$ *pan-endothelial cells (pan-EC) and kdrl*$^{-}$ *non-endothelial cells (non-EC) is indicated with a "+/-". Peak scores were higher in sprouts, or veins respectively, compared to the other sub-endothelial cell types in which the associated gene is not expressed.*

Sub-endothelial cell type in which enhancer is active	Enhancer	Peak score in hlx-3+	Peak score in Dll4in3+	Peak score in Ephb4-2+	panEC	nonEC
Sprout	hlx -3	**20.0**	**46.8**	9.9	+	-
Vein	mafbb +7	14.5	38.0	**52.7**	+	-
	nr2f2 -965	14.0	**21.0**	**35.0**	+	+
panEC	etv2 -2	**20.3**	**19.2**	**27.9**	+	-
	fli1a in1	**34.7**	**47.7**	**69.3**	+	-

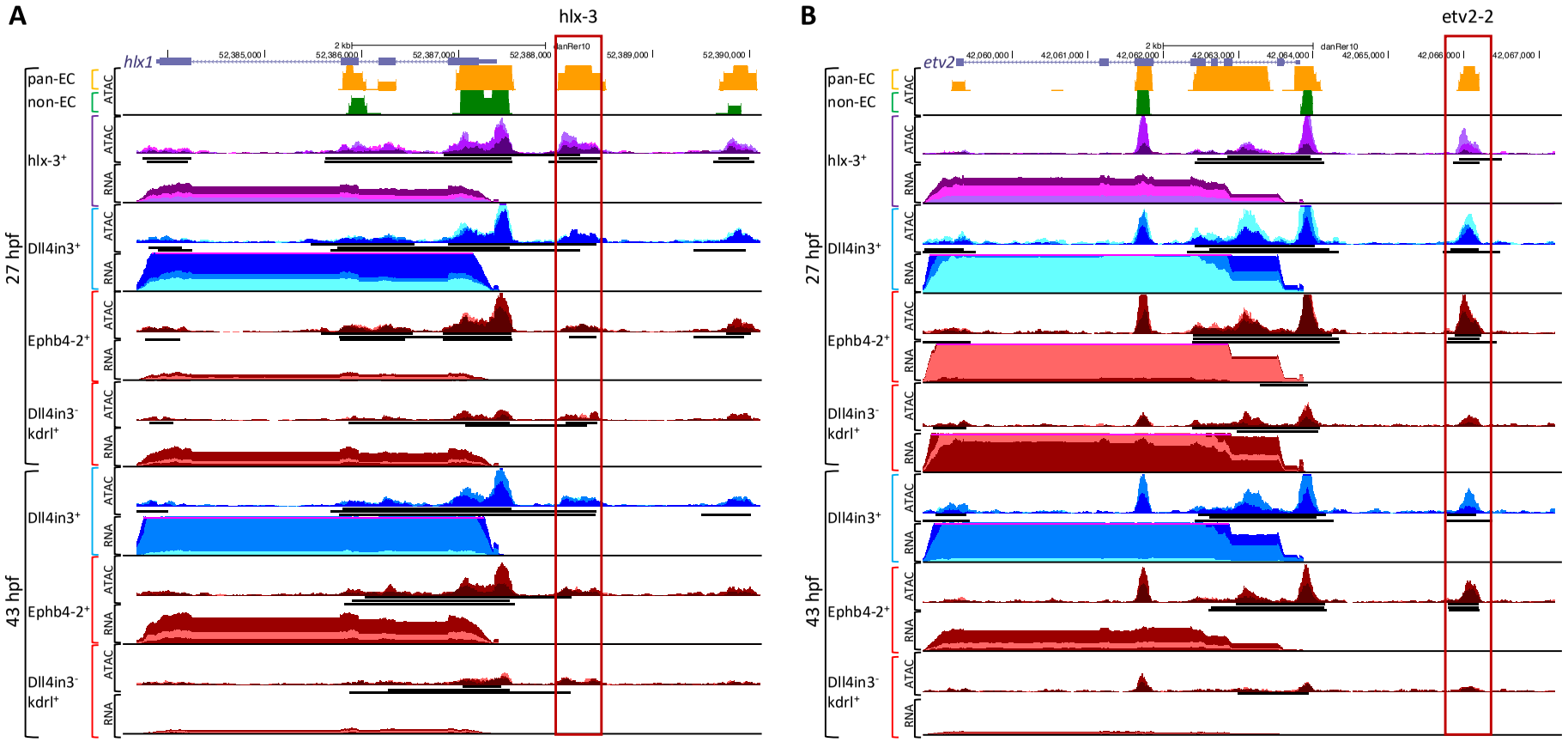


Figure 54: Mapping of ATAC-seq reads at known zebrafish enhancers. *Visualization of ATAC-seq and RNA-seq at 27 and 43 hpf. The first two lanes show ATAC-seq peaks from kdrl⁺ pan-endothelial (pan-EC) and kdrl⁻ non-endothelial cells (non-EC). For these two samples peak height corresponds to the number of replicates in which that peak was significant. The following lanes show ATAC-seq peaks and RNA-seq read map density for hlx-3⁺, Dll4in3⁺, Dll4in3⁻kdrl⁺ and Ephb4-2⁺ cells. ATAC-seq peaks are visualized by bigwig files, while significant ATAC-seq peaks in each replicate are represented by black bars. Three biological replicates are overlaid for each group.* ***A)*** *Red box shows the angiogenic-specific hlx-3 enhancer .* ***B)*** *Red box shows the pan-endothelial etv2-2 enhancer*[314]*.*

5.3.2 ATAC-seq cannot consistently predict active angiogenic-specific enhancers

Next, I tested whether I can use the ATAC-seq data in an agnostic approach to find novel angiogenic-specific enhancers. As opposed to a hypothesis and target driven analysis, I here wanted to analyse the genome-wide data in an unbiased way. Two approaches were used to validate *cis*-regulatory elements *in vivo* in transient transgenic zebrafish.

For the first agnostic approach, hlx-3^{+} ATAC-seq peaks were ranked by p-values generated by MACS2[199]. Open chromatin regions do not necessarily have to be regulatory elements. I tried to enrich my top ATAC-seq hlx-3^{+} peaks for putative enhancers by filtering ATAC-seq peaks for intergenic and intronic regions and removing repetitive elements, promoters, TTS and exons. To test the activity of the top hlx-3^{+} accessible regions, 600 bp sequences were synthesized and cloned into the pE1b Citrine or mCherry Tol2 plasmids. Putative enhancers were screened *in vivo* in transient transgenic zebrafish using Tol2 transgenesis. None of the ten tested accessible regions were functional enhancers and specific to the angiogenic sprout (Figure 55A). Even though *tg(hlx-3:GFP)* predominantly marks sprouting ISVs, GFP expression was weakly detected in the endocardium and occasionally in the retina very early during development. This inclusion of non-endothelial cells might be one of the reasons why the top hlx-3^{+} ATAC-seq peaks were not specific to endothelial cells. It is also plausible that these regions denoted constitutively open chromatin independent of cell type and were not regulatory regions capable of driving transcription. This is further supported by the fact that all tested sequences also had a significant peak in non-endothelial cells. So for the second approach, I only selected peaks that did not overlap with a peak in non-endothelial cells.

For the second approach, I analysed differentially open chromatin in sprouting cells. Therefore, the ATAC-seq data were clustered into two groups: ISV (hlx-3^{+} and Dll4in3^{+}) and non-ISV (Ephb4-2^{+} and Dll4in3^{-}kdrl^{+}). In total 67,697 peaks were shared between all replicates of hlx-3^{+} and Dll4in3^{+} cells. Comparing group sprouting "ISV" with "non-ISV" 3,266 peaks were significantly differentially accessible. To reduce the risk of including non-endothelial peaks, I filtered the data based on the criteria described in the Materials (see section 2.7.3.1) and reduced the number of peaks to 28. The top nine differentially accessible regions represented putative angiogenic-specific enhancers and were screened for their ability to activate reporter gene expression in transient transgenic zebrafish (Figure 55B). However, only one sequence, located 210 kb upstream of *etfb* (etfb-210), was able to drive reporter gene expression in sprouting endothelial cells *in vivo* (Figure 55C). Etfb-210 was the only locus with a significant peak in all three sub-endothelial cells and a higher peak score in both hlx-3^{+} and Dll4in3^{+} (Figure 55E). Interestingly, my list of top angiogenic-specific differentially accessible regions contained several genes that have been shown to be expressed in sprouting endothelial cells, such as *esm1* and *vegfc*[316,251]. Moreover, *ptpt4a3* and *fgfr3* have been implicated in angiogenesis[317,318,319]. ATAC-seq was hence largely able to find open chromatin surrounding genes that are associated with angiogenesis, but the tested sequences were not able to activate reporter gene transcription.

Lastly, I wanted to know whether novel angiogenic-specific enhancers could be identified using a hypothesis driven approach. For this approach, I looked for open chromatin regions near one of known sprouting-enriched genes. I selected *esm1*, which is known to

be expressed in retinal tip cells[316,320,321]. My RNA-seq data also showed that *esm1* was differentially overexpressed in sprouting compared to venous endothelial cells. I manually screened the *esm1* locus for pan-endothelial ATAC-seq peaks and tested two accessible regions in transient transgenic zebrafish. This approach had some success. Out of the two putative enhancers, esm1+7:Citrine was able to drive reporter gene expression in sprouting ISVs and ectopically in the muscle (Figure 55D).

To summarize, even though the ATAC-seq data were not able to reliably predict angiogenic-specific enhancers in an agnostic approach, it was possible to discover a regulatory element activating transcription in sprouting ISVs in a hypothesis driven approach. While both the hypothesis-free and hypothesis driven approaches found a single enhancer that drove some ISV expression, the success rate appears to be greater when we restricted analysis to a loci where we expected the gene to be expressed. Therefore, it does not appear that ATAC-seq peaks alone can successfully predict enhancer activity in a specific endothelial sub-type.

A

Top hlx-3

Top 10 accessible regions in hlx-3+

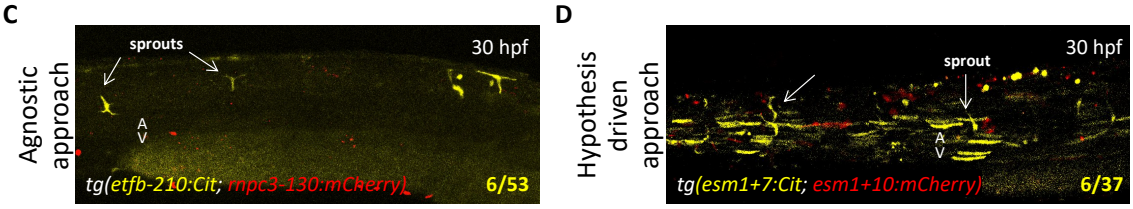
Coordinates of injected sequence	Closest gene	Relative to TSS	Number of fish analysed	% detectable reporter gene expression	% expression in ISV	Reporter activity
chr3:23,417,322-23,417,921	**igf2bp1**	**-6 kb**	29	66%	0%	ectopic
chr13:49,430,485-49,431,084	**ggps1**	**+15 kb (intron 2)**	40	85%	0%	ectopic (neuronal)
chr12:40,918,720-40,919,319	**ppp2r2d**	**+300 kb (intron 9)**	36	72%	0%	ectopic (neuronal)
chr1:13,280,282-13,280,881	**elf2b**	**+1.2 kb (intron 1)**	67	93%	0%	ectopic (neuronal)
chr23:1,125,122-1,125,721	**psmb11**	**-6 kb**	27	48%	0%	ectopic
chr6:50,476,401-50,477,000	**mych**	**-20 kb**	29	66%	0%	ectopic (heart)
chr11:28,058,682-28,059,281	**ephb2b**	**+90 kb (intron 2)**	29	93%	0%	ectopic (retina)
chr5:26,911,879-26,912,478	**sfrp1a**	**-13 kb**	29	72%	0%	ectopic (heart)
chr1:53,126,681-53,127,280	**ifngr1**	**-55 kb**	31	87%	0%	ectopic (neuronal)
chr1:58,281,520-58,282,119	**pin1**	**-4 kb**	36	67%	4%	ISV and ectopic

B

ISV vs non-ISV

Top 9 differentially open sprouting specific ATAC-seq peaks

Coordinates of injected sequence	Closest gene	Relative to TSS	Number of fish analysed	% detectable reporter gene expression	% expression in ISV	Reporter activity
chr21:9,893,841-9,894,440	**esm1**	**-1.2 kb**	44	59%	0%	ectopic
chr1:38,504,522-38,505,121	**vegfc**	**-310 kb**	37	86%	0%	ectopic (neuronal)
chr19:25,516,390-25,516,989	**ptp4a3**	**+12 kb (intron 3)**	44	61%	0%	ectopic
chr8:51,509,931-51,510,530	**fgfr3**	**+27 kb (intron 2)**	37	81%	0%	ectopic (muscle)
chr12:29,105,347-29,105,946	**mxtx2**	**-16 kb**	67	67%	0%	ectopic
chr16:50,235,759-50,236,358	**etfb**	**-210 kb**	**55**	**96%**	**11%**	**ISVs and ectopic**
chr24:28,947,267-28,947,866	**rnpc3**	**-130 kb**	55	84%	0%	ectopic
chr18:36,545,855-36,546,454	**siae**	**-56 kb**	24	63%	0%	ectopic
chr23:42,826,542-42,827,141	**gpx7**	**-52 kb**	24	58%	0%	ectopic



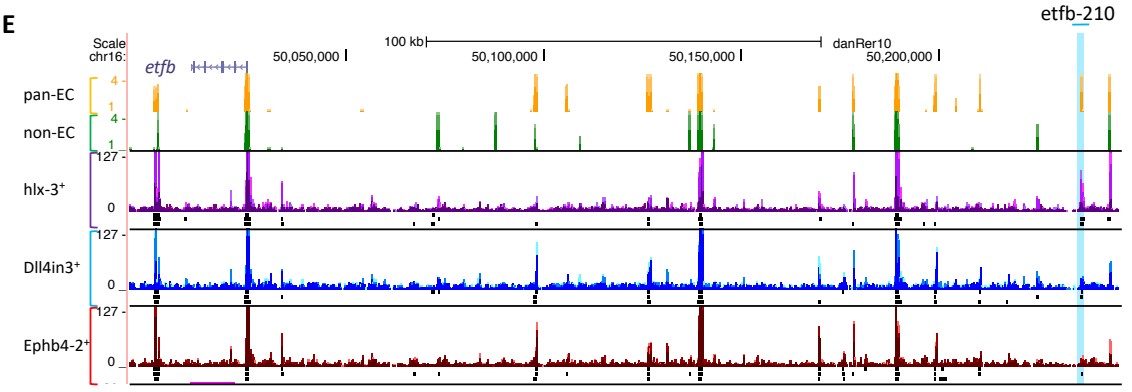


Figure 55: ATAC-seq regions tested for reporter gene expression in transient transgenic zebrafish. A) *Table listing top accessible regions ranked by p-value in hlx-3⁺ and summary of reporter gene expression detected in mosaic transient transgenic zebrafish at 30 hpf.* ***B)*** *Table listing top differentially accessible regions comparing ISV (hlx-3⁺ and Dll4in3⁺) with non-ISV (Ephb4-2⁺ and Dll4in3-kdrl⁺) and summary of reporter gene expression detected in mosaic transient transgenic zebrafish at 30 hpf. Reporter gene expression was scored as "ISV" if at least one intersegmental vessel was positive.* ***C)*** *Example of two tested differentially open ATAC-seq peaks. Overlay of yellow and red fluorescent image from embryo injected with etfb-210:Citrine and rnpc3-130:mCherry. Only etfb-210 directs expression in sprouting ISVs.* ***D)*** *Two*

constructs tested in hypothesis driven approach: esm1+7:Citrine and esm1+10:mCherry. Only esm1+7 directs expression in sprouting ISVs and ectopically in the muscle resulting from pE1b activity in the neural crest lineage. Number at the bottom right of the image indicates the number of fish with reporter gene expression in ISVs per number of transgenic fish. (A) denotes the dorsal aorta and (V) the cardinal vein. ***E)*** *ATAC-seq data of the etfb locus for pan-EC, non-EC, hlx-3^{+}, Dll4in3^{+}, Dll4in3^{-}kdrl^{+} and Ephb4-2^{+} cells at 27 hpf. ATAC-seq peaks are visualized by bigwig files, while significant ATAC-seq peaks in each replicate are represented by black bars. The injected sequence is marked with a blue box.*

5.4 Integration of ATAC-seq and RNA-seq

5.4.1 Differential gene expression in veins and sprouts at 27 hpf

As mentioned in the introduction, promoter/enhancer activity determines gene expression levels and DNA accessibility is a predictor of enhancer activity[322]. I hypothesized that chromatin accessibility is increased in regions surrounding genes that are overexpressed in a specific sub-endothelial cell type. To determine the correlation between accessible chromatin and transcription, RNA-seq was performed on hlx-3^{+}, Dll4in3^{+}, Dll4in3^{-}kdrl^{+} and Ephb4^{+} cells with three biological replicates each. Global gene expression data of the four isolated sub-endothelial cell types were used for a comprehensive comparison.

Overall, 122 genes were expressed at significantly higher levels in sprouting ISV (hlx-3^{+} and Dll4in3^{+}) versus non-sprouting ISV (Ephb4-2^{+} and Dll4in3^{-}kdrl^{+}) cells. These hits include genes known to play a role during angiogenesis, such as *dll4*[296], *il6*[323], *vegfc*[324], *tbx20*[325] and *flt1*[326]. To analyse the RNA-seq results in an unbiased way, ISV-specific transcripts were ranked by p-value and the most 15 overexpressed genes are shown in Table 8 below. While some of the transcripts listed in the top 15 genes overexpressed in ISVs were sprouting-specific (*slc8a1a*[327], *fhl2a* and *rgs5a*[114]), some of the transcripts have only been implicated to play a role in endothelial cells more generally (*dbx1b*[328], *lft2* and *six3a*[87], *aldh1a3*[329], *nkx2.5*[330,331]). In contrast, differentially expressed genes (such as *tnnt2a*[181],

tnni1b[332], *myh6*[333] and *cx36.7*[334]), which are well documented to be crucial for heart development, were not considered to be sprouting-specific as both hlx-3 and Dll4in3 enhancers drive expression in the endocardium. Even though we were enriching for endothelial-specific genes, we also found transcription factors associated with neural differentiation (*barhl2*[335], *otx2*[336], *dmrta2*[337]).

I intersected my ISV-enriched genes with a list from a recently published study where RNA-seq was performed on Hlx protein overexpressing, hyperproliferating endothelial cells in zebrafish at 48 hpf[338]. Out of my top 15 ISV-enriched hits, 12 genes (or closely related family members and paralogues) were also upregulated in Hlx overexpressing endothelial cells; including *tnnt2c*, *tnni1b*, *slc8a1b*, *fhl2a*, *rgs5b*, *dmrt2a*, *lft1*, *aldh1a2*, *nkx2.2*, as well as *connexins*, *myh* and *six* factors[338].

Table 8: Top 15 significantly overexpressed genes in ISVs (hlx-3$^+$/Dll4in3$^+$) compared to non-ISVs (Ephb4-2$^+$/Dll4in3$^-$ kdrl$^+$) at 27 hpf ranked by p-value. *For each transcript, log fold change (logFC) and log counts per million (logCPM) are given. Information shown in the column "expression pattern" refers to in situ data provided by the Zebrafish Information Network (ZFIN) between pharyngula stages prim-5 and prim-25. The information in the column "potential role in endothelial cells based on literature" was based on a species-independent PubMed search with the respective gene and/or orthologue name and the terms angiogenesis, sprouting, endothelial, vasculature, artery and vein. CNS = central nervous system.*

	logFC	logCPM	p-value	Expression pattern	Potential role in endothelial cells based on literature	Reference
tnnt2a	10.16	5.70	7.22E-19	heart	-	Sehnert 2002
tnni1b	10.93	5.64	6.48E-15	heart	-	Cai 2008
dbx1b	7.13	5.02	2.31E-12	CNS	hlx1 family member, previous name: *hlx2*	Dean 2014
slc8a1a	12.24	2.29	4.56E-12	heart	required for VEGF-dependent angiogenesis, previous name: *ncx1*	Andrikopolous 2017
myh6	17.93	4.06	2.70E-11	heart	-	Singleman 2012
fhl2a	7.48	5.73	5.21E-11	heart	tip cell enriched gene	del Toro 2010
rgs5a	9.30	2.91	7.72E-11	heart	tip cell enriched gene	del Toro 2010
barhl2	8.64	4.66	1.57E-10	CNS	-	Parish 2016
otx2	5.23	4.37	1.05E-09	CNS	-	Mori 1994
dmrta2	7.94	4.56	2.11E-09	CNS	-	Yoshizawa 2012
lft2	12.56	4.74	6.45E-09	heart	expressed in cultured venous endothelial cells	Chi 2003
aldh1a3	11.08	3.04	9.60E-09	eye	required for cardiac atrial appendage stem cells (incl. endothelial progenitor)	Puttini 2018
six3a	5.32	4.50	3.38E-08	CNS, eye	expressed in cultured nasal endothelial cells	Chi 2003
cx36.7	12.92	1.03	3.65E-08	heart	-	Miyagi 2015
nkx2.5	6.23	3.30	3.72E-08	heart	required for yolk sac vasculature and angioblast differentiation	Tanaka 1999, Paffett-Lugassy 2013

Given that not much is known about venous cell fate acquisition, I next examined upregulated genes in developing veins compared to arteries at 27 hpf. Overall, 86 genes

were expressed at significantly higher levels in venous (Ephb4-2$^+$) compared to arterial (Dll4in3$^+$) cells. Unexpectedly, the top venous hits were not the classical genes known to be important for venous differentiation (see Table 9). Nonetheless, 10 out of the top 15 venous transcripts have been implicated to play a role during endothelial specification (*hoxa13b* and *hoxa13a*[339], *nr3c1*[340], *hoxb13a*[341], *hoxc10a*[342], *tmed1a*[343], *chrna1*[344], *sp5l*[345], *msgn1*[346] and *crabp2b*[347]). The remaining genes have so far not been studied or are not known to be expressed in endothelial cells (*spata6l*, *musk*[348], *her7*[349], *eve1*[345], *ved*[350]).

Table 9: Top 15 genes significantly overexpressed in veins (Ephb4-2$^+$ vs. Dll4in3$^+$) at 27 hpf ranked by p-value. *For each transcript, log fold change (logFC) and log counts per million (logCPM) are given. Information shown in the column "expression pattern" refers to in situ data provided by ZFIN between pharyngula stages prim-5 and prim-25. The information in the column "potential role in endothelial cells based on literature" was based on a species-independent PubMed search with the respective gene and/or orthologue name and the terms angiogenesis, sprouting, endothelial, vasculature, artery and vein.*

	logFC	logCPM	p-value	Expression pattern	Potential role in endothelial cells based on literature	Reference
hoxa13b	3.60	2.62	5.38E-09	tail and fin bud	required for vascular patterning of placenta	Shaut 2013
nr3c1	2.69	3.79	2.75E-08	whole organism	expressed in cultured endothelial cells	Mata-Greenwood 2015
sp5l	4.35	3.14	3.52E-06	tail bud, CNS	controls tail development (and possibly arteriovenous specification)	Thorpe 2005
hoxa13a	3.93	3.10	3.76E-06	fin bud	required for vascular patterning of placenta	Shaut 2013
msgn1	6.51	3.44	4.15E-06	tail bud	required for paraxial mesoderm specification, and thus vasculature	Row 2016
crabp2b	2.80	4.01	7.32E-06	tail and fin bud, eye	expressed in paraxial mesoderm	Sharma 2005
hoxb13a	5.71	1.85	9.53E-06	-	expressed in cultured endothelial cells	Toshner 2014
hoxc10a	2.74	2.30	1.22E-05	tail bud, CNS	expressed in kidney-specific endothelial cells	Marcu 2018
spata6l	2.28	2.28	2.56E-05	-	paralogue *spata6* expressed in KLF2/4 KO cardiac endothelial cells	Sangwung 2017
musk	3.85	2.82	2.71E-05	myotome, CNS	-	Zhang 2004
her7	4.74	1.94	3.04E-05	tail bud, mesoderm	-	Sieger 2003
tmed1a	2.54	4.67	3.38E-05	myotome	*tmed1* expressed in KLF2/4 KO cardiac endothelial cells and in HUVECs	Sangwung 2017
chrna1	2.76	4.79	4.03E-05	myotome	expressed in cultured endothelial cells	Wick 2007
eve1	5.87	2.62	4.29E-05	tail bud	-	Thorpe 2005
ved	2.65	4.13	4.86E-05	tail bud, CNS	-	Thisse 2001

Of the 86 genes that are differentially expressed in the venous Ephb4-2$^+$ cells, 88% had at least one ATAC-seq peak associated with the upregulated transcript. Of these, 9.7% genes had a unique venous Ephb4-2$^+$ associated peak, while 79% of genes had a peak in all sub-endothelial cell types. Of the 122 genes that are differentially expressed in sprouting ISVs (hlx-3$^+$ and Dll4in3$^+$), 87% had at least one ATAC-seq peak associated with the upregulated transcript. Of these, only 4.9% genes had a unique sprouting associated peak with no peak

identified in venous Ephb4-2^{+}. This suggests that independent of gene expression, chromatin accessibility is largely similar across sub-endothelial cell types.

To conclude, the list of the top 15 differentially expressed genes for both the venous and sprouting ISV cell type did contain some genes that are known to be expressed in endothelial cells. This confirms that we did indeed collect the correct cell type. While classical venous and angiogenic markers were not enriched, several novel venous and sprouting-specific genes might have been uncovered. Unfortunately, due to time constraints, I was not able to validate the correct spatial expression pattern of these most significantly enriched transcripts by *in situ* hybridization. A more detailed analysis of the global gene expression data generated from of all four sub-endothelial cells types in the future is required (such as GO term analysis and intersecting our RNA-seq with already published endothelial-specific datasets).

5.4.2 ATAC-seq peaks correlate with differentially expressed genes in veins and sprouts

As described in section 5.3.2, only 28 differentially accessible regions were identified in sprouting ISVs computationally. Apart from *vegfc*, none of the genes associated with angiogenic-specific ATAC-seq peaks were found to be differentially expressed by RNA-seq, I thus wanted to assess the correlation between ATAC-seq and RNA-seq qualitatively. Therefore, I manually compared the ATAC-seq peak scores flanking an overexpressed gene (until the TSS/TTS of the neighbouring gene) between the sub-endothelial cell types. The most highly and significantly differentially expressed genes for sprouts (ISV vs. non-ISV)

and veins (Ephb4-2$^+$ vs. Dll4in3$^+$) were taken from the RNA-seq lists described in the previous chapter. A set of genes that were non-differentially expressed and common to all sub-endothelial cell types were used as controls. “Similarly” open chromatin was defined as having a peak score that differs not more than 50% between hlx-3$^+$, Ephb4-2$^+$ and Dll4in3$^+$. For the peak score analysis I excluded the non-arterial Dll4in3-kdrl$^+$ sub-endothelial cell type, as DNA was slightly overtagmented and peak scores were hence not comparable between sub-endothelial cell types.

Table 10 lists ATAC-seq peak scores surrounding non-differentially expressed endothelial genes, venous and sprouting overexpressed genes. 61% of peaks surrounding non-differentially expressed genes had similarly accessible chromatin (Table 10A). In contrast, peaks surrounding angiogenic-specific overexpressed genes only had 21% similarly accessible chromatin when comparing the highest ATAC peak score to the lowest peak score between sub-endothelial cell types (Table 10B). Peaks surrounding venous-specific overexpressed genes had 33% similarly accessible chromatin (Table 10C). As predicted, 73% of ATAC-seq peaks had the highest peak score in the respective sub-endothelial cell type in which the gene was overexpressed. It was not just the peak score that differed between sub-endothelial cell types, but also the number of significant peaks. For example, genes that were significantly enriched in sprouting ISVs compared to non-ISVs, the number of peaks surrounding their TSS was higher in sprouting hlx-3$^+$ than in non-sprouting Ephb4-2$^+$ cells. In contrast, there was no correlation between the number of peaks in the sub-endothelial cell types flanking non-differentially expressed endothelial genes.

Collectively, these data indicate that not just the absence/presence of a significant ATAC-seq peak, but even the ATAC-seq peak score of a sub-endothelial cell-type-specific regulatory element correlates with cell-type-specific transcription.

Table 10: Qualitative peak score analysis of hlx-3$^+$, Ephb4-2$^+$ and Dll4in3$^+$ ATAC-seq peaks surrounding differentially expressed genes. *Chromatin is denoted as similarly open if the ratio between the largest and smallest peak score is ≤ 2 (green) and as not similarly open if the ratio is > 2 (red). The absence of a peak is indicated with "–". In pan-EC and non-EC where no peak scores were available, the presence of an ATAC-seq peak is indicated with "+". **A)** Peak scores surrounding non-differentially expressed genes that are common to all sub-endothelial cell types. **B)** Peak scores surrounding top eight differentially expressed sprouting genes. **C)** Peak scores surrounding top 15 differentially expressed venous genes.*

A Control
ATAC-seq peaks surrounding non-differentially expressed pan-endothelial genes

Peak Scores	hlx-3+	Dll4in3+	Ephb4-2+	panEC	nonEC	Similarly open chromatin
kdrl peak1	19.7	-	6.8	-	-	-
kdrl peak2	19.5	14.5	13.0	+	-	1.5
kdrl peak3	-	14.0	10.0	+	+	-
kdrl peak4	41.0	28.5	53.0	+	-	1.9
dusp5 peak1	19.6	15.9	19.5	+	-	1.2
dusp5 peak2	49.2	56.0	86.0	+	+	1.7
dusp5 peak3	21.3	43.9	17.8	+	+	2.5
dusp5 peak4	37.4	71.4	93.5	+	-	2.5
dusp5 peak5	-	17.4	23.4	+	-	1.3
dusp5 peak6	15.0	14.7	16.2	+	-	1.1
etv2 peak	20.3	19.2	27.9	+	-	1.5
fli1a peak1	44.8	49.8	70.6	+	+	1.6
fli1a peak2	3.6	7.4	6.8	+	-	2.0
fli1a peak3	71.1	28.2	14.2	+	+	5.0
fli1a peak4	41.1	54.2	96.4	+	-	2.3
fli1a peak5	4.8	15.5	14.7	+	+	3.2
fli1a peak6	6.0	21.5	22.6	+	+	3.7
fli1a peak7	13.5	15.5	14.9	+	+	1.1
fli1a peak8	22.2	26.9	28.3	+	+	1.3
fli1a peak9	40.3	30.1	62.3	+	+	2.1
fli1a peak10	14.0	11.4	12.3	+	+	1.2
fli1a peak11	15.9	14.1	7.8	-	-	2.0
fli1a peak12	8.2	7.2	14.1	+	-	2.0
fli1a peak13	82.4	90.6	126.9	+	+	1.5
fli1a peak14	6.9	10.6	8.2	-	-	1.5
fli1a peak15	34.7	47.7	69.3	+	-	2.0
fli1a peak16	2.0	-	4.1	+	+	-
fli1a peak17	15.3	8.6	10.0	-	+	1.8
erg peak1	15.3	-	-	-	-	-
erg peak2	17.0	22.7	7.1	+	+	3.2
erg peak3	3.0	5.4	4.0	-	+	1.8
erg peak4	75.9	63.6	85.9	+	+	1.3
erg peak5	9.9	25.7	34.5	+	-	3.5
erg peak6	41.8	77.3	110.0	+	-	2.6
erg peak7	34.4	30.1	41.4	+	+	1.4
erg peak8	14.3	22.1	21.1	+	+	1.5

n = 36
61% of peaks have similarly open chromatin in all sub-endothelial cell types

B ISV vs non-ISV
ATAC-seq peaks surrounding top 8 differentially expressed sprouting genes

Peak Scores	hlx-3+	Dll4in3+	Ephb4-2+	panEC	nonEC	Similarly open chromatin
tnnt2a peak1	12.5	8.2	14.3	+	-	1.7
tnnt2a peak2	-	16.7	19.1	+	+	-
tnni1b peak	5.8	9.2	-	-	-	-
dbx1b peak1	6.5	5.8	3.5	-	-	1.9
dbx1b peak2	8.9	15.5	21.8	+	-	2.5
dbx1b peak3	65.4	60.5	81.0	+	+	1.3
dbx1b peak4	33.3	14.0	4.4	+	+	7.5
dbx1b peak5	12.5	21.7	-	-	+	-
dbx1b peak6	3.9	8.7	8.5	-	+	2.2
dbx1b peak7	15.7	6.1	10.9	-	+	2.6
dbx1b peak8	7.5	2.6	6.5	-	-	2.5
dbx1b peak9	46.1	23.2	13.1	+	+	3.5
dbx1b peak10	51.9	51.8	59.9	+	+	1.2
dbx1b peak11	18.1	15.5	-	-	+	-
dbx1b peak12	11.0	-	-	+	+	-
dbx1b peak13	13.6	22.8	18.5	+	-	1.7
dbx1b peak14	26.7	15.9	12.1	+	+	2.2
dbx1b peak15	52.2	-	-	-	+	-
dbx1b peak16	31.7	13.8	10.8	+	+	2.9
fhl2a peak	3.8	5.3	7.1	-	-	1.9
rgs5a peak1	3.2	-	-	-	-	-
rgs5a peak2	12.3	4.7	6.7	-	-	2.6
barhl2 peak1	111.8	26.5	-	-	+	-
barhl2 peak2	10.8	4.8	-	-	-	-
barhl2 peak3	17.8	-	-	-	-	-
barhl2 peak4	8.5	-	-	-	-	-
barhl2 peak5	90.1	37.7	31.4	+	+	2.9
barhl2 peak6	9.4	8.1	6.2	-	-	1.5
barhl2 peak7	7.7	35.6	-	-	+	-
barhl2 peak8	12.5	10.3	-	-	-	-
barhl2 peak9	13.6	-	-	-	-	-
barhl2 peak10	6.6	-	-	-	-	-
barhl2 peak11	7.7	-	-	+	+	-

n = 33
21% of peaks have similarly open chromatin in all sub-endothelial cell types

C Ephb4-2+ vs Dll4in3+
ATAC-seq peaks surrounding top 15 differentially expressed venous genes

Peak Scores	hlx-3+	Dll4in3+	Ephb4-2+	panEC	nonEC	Similarly open chromatin
hoxa13b peak1	6.3	-	-	-	-	-
hoxa13b peak2	40.3	22.3	32.4	+	+	1.8
hoxa13b peak3	9.7	6.8	8.5	-	-	1.4
nr3c1 peak1	24.2	15.3	34.9	+	+	2.3
nr3c1 peak2	-	-	13.3	-	-	-
nr3c1 peak3	-	15.0	18.1	-	+	-
nr3c1 peak4	15.4	-	30.0	-	-	-
nr3c1 peak5	47.6	34.3	87.9	-	+	2.6
nr3c1 peak6	6.9	4.4	8.1	-	-	1.8
sp5l peak1	-	10.3	8.3	+	-	-
sp5l peak2	5.3	14.8	12.7	-	-	2.8
sp5l peak3	49.2	38.6	43.5	-	-	1.3
sp5l peak4	22.7	-	28.5	+	-	-
sp5l peak5	-	10.4	10.1	+	-	-
hoxa13a peak	37.1	14.9	23.7	+	+	2.5
msgn1 peak	20.9	61.8	78.5	-	-	3.8
crabp2b peak	41.3	11.3	-	-	+	-
hoxb13a peak1	11.0	10.8	17.6	-	-	1.6
hoxb13a peak2	55.9	16.5	20.3	+	+	3.4
hoxb13a peak3	5.3	3.1	5.5	-	-	1.8
hoxb13a peak4	4.4	-	10.1	-	-	-
hoxc10a peak	24.3	30.8	31.0	+	+	1.3
spata6l peak	11.5	-	19.2	-	+	-
musk peak1	-	3.7	5.6	-	-	
musk peak2	4.2	4.0	3.7	-	-	1.1
musk peak3	4.1	7.6	3.8	-	-	2.0
musk peak4	18.7	5.0	12.3	-	-	3.7
tmed1a peak	-	-	4.8	-	-	-
chma1 peak1	-	11.7	13.4	+	-	-
chma1 peak2	7.1	4.2	7.9	-	-	1.9
eve1 peak1	5.5	11.3	-	-	+	-
eve1 peak2	11.0	10.8	17.6	-	-	1.6
ved peak	12.6	17.4	2.5	-	+	6.8

n = 33
33% of peaks have similarly open chromatin in all sub-endothelial cell types

73% of peaks have highest peak score in sub-endothelial cell type in which respective gene is overexpressed

5.4.3 Motif analysis identifies known and novel transcription factors

To gain more insights into which transcription factors are responsible for sprouting, arterial or venous specificity, I wanted to determine whether accessible chromatin from sub-endothelial cell types were enriched for particular transcription factor binding sites. Therefore, overrepresented motifs at the TSS and intergenic and intronic regions were investigated to identify novel transcriptional regulators. I determined the enrichment of transcription factor motifs within open regions of sprouting hlx-3^{+} (sprouting), Dll4in3^{+} (sprouting and arterial) and Ephb4-2^{+} (venous) cells. The most significant *de novo* transcription factor binding site motifs are listed in Table 11.

Strikingly, the transcription factor binding site for MEF2C was enriched in both the hlx-3^{+} and Dll4in3^{+} sub-endothelial cell type (Table 11A and B). This reiterates my findings in Chapter 3 and 4 linking MEF2 with the induction of angiogenic sprouting, and further highlights the importance of MEF2 family members for sprouting angiogenesis[170]. The transcription factor binding sites for SOXF family members (SOX8, SOX17, SOX18)[351] were also enriched in both the hlx-3^{+} and Dll4in3^{+} sub-endothelial cell type. SOXF family members are crucial for sprouting angiogenesis[120]. SOX17 is absolutely required for correct retinal sprouting, *Sox17* knockout results in a severe hypersprouting defect[352]. Given that it has been shown that SOX18 can interact with MEF2C[353], it is tempting to speculate that SOXF and MEF2 family members synergistically activate angiogenic-specific transcription. Another sprouting-enriched DNA binding motif (shared between Dll4in3^{+} and hlx-3^{+}) was MAFB. Two recently published studies identified MAF factors as the master regulators for angiogenesis[354,355]. Wang *et al.* found increased MAF transcription factor occupancy at distal regulatory elements in response to VEGFA in mammalian

endothelial cell culture[355], while Jeong *et al.* found *MafB* to be indispensable for sprouting angiogenesis in the mouse retina[354].

Other notable sprouting-enriched DNA binding motifs (overrepresented in both Dll4in3$^+$ and hlx-3$^+$ sub-types) have also been shown to play a role during sprouting angiogenesis, but are less well described, such as Iroquois-class homeodomain protein (IRX3)[356], Runt-related transcription factor 1 (RUNX1)[357,358] and heterodimer CCAAT/enhancer-binding protein (C/EBP)/Activator protein 1 (AP1)[359,119]. CEBP/AP1 could reflect the increased proliferation that angiogenic sprouting cells undergo. Last, hlx-3$^+$ was also enriched for GATA binding sites, a transcription factor known to be crucial for endothelial differentiation[116,131].

As discussed previously, ETS factors are essential for endothelial gene expression[138] and indeed 17% of the total peaks present in the Dll4in3$^+$ and EphB4-2$^+$ population revealed an enrichment for the motif of the ETS family member ERG1. With a binding site in 17% of all regions, ERG1 constitutes the most highly enriched hit in both Dll4in3$^+$ and EphB4-2$^+$. This confirms that we successfully collected the correct cell type and reiterates the requirement of ETS factors for the endothelial lineage. Across all sub-endothelial cell types an average of 10% of all open chromatin was found to contain binding motif for cAMP response element binding protein (CREB1). CREB1 binds to promoters and enhancers and stimulates transcription through subsequent CBP/p300 recruitment[360]. CREB1 binding motif may hence serve as a proxy for CBP/p300, given that it does not directly bind DNA.

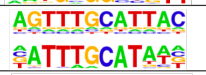
The average of 10% CREB1 binding motifs in our data is consistent with a study that found that 12% of all distal accessible chromatin was bound by CBP/p300[361].

In the venous Ephb4-2^{+} sub-endothelial cell type, I found overrepresented motifs for two transcription factors RBPJ[218] and SMAD[179], both which have recently been shown to be important for venous cell fate acquisition (Table 11C). Even though Notch-mediated *Rbpj* activation is crucial for arterial gene activation, RBPJ acts as a repressor in the absence of Notch signalling and has been shown to repress the activity of arterial enhancers in veins[218]. Likewise, although SMAD transcriptional factors are often associated with later arterial differentiation, a recent paper from my lab has now shown that venous endothelial cells acquire venous identity *via* a BMP4-ALK3 driven signalling pathway through the direct binding of SMAD1/5 to venous-specific loci[179]. In addition to RBPJ and SMAD, the binding motif for the gene insulator CTCF was also enriched in 4.2% of all ATAC-seq peaks. CTCF occupies most of its DNA binding motifs[2]. My data showed that, independent of the sub-endothelial cell type and cell-specific gene expression, ATAC-seq peaks can be found in the proximity of arterial genes. The increase in CTCF-associated boundary domains in venous Ephb4-2^{+} cells might thus prevent transcription of arterial genes[362].

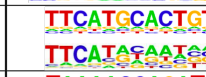
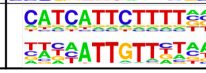
Taken together, these motif lists contained the expected enrichment of transcription factor binding motifs already known to play important roles during the differentiation of each specific endothelial sub-type. Further, these data also uncovered several previously unrecognized transcription factor signatures. For example, both hlx-3^{+} and Ephb4-2^{+} cells shared a binding motif for Octamer-binding transcription factor 4 (OCT4)[363]. OCT4 is

associated with the pluripotent state of stem cells, which might hint towards the progenitor-like character of differentiating endothelial cells.

A Transcription factor binding motifs enriched in sprouting-specific open chromatin regions (hlx-3+)

Factor	Binding Motif	% Regions	P-value	Fold-enriched
CREB1		11.89	1e-6194	3.66
GATA		0.88	1e-1746	29.33
CEBP:AP1		0.98	1e-1731	19.60
RUNX1		0.87	1e-1730	29.00
OCT4		0.97	1e-1678	19.40
MAFA/B		0.96	1e-1574	16.0
MEF2C		0.91	1e-1471	18.2
IRX3/4/5		0.95	1e-1346	13.6
SOX8		0.96	1e-1343	12.0

B Transcription factor binding motifs enriched in arterial and sprouting-specific open chromatin regions (Dll4in3+)

Factor	Binding Motif	% Regions	P-value	Fold-enriched
CREB1		9.42	1e-2145	3.4
ERG1		17.82	1e-575	1.5
MAFA/B		0.59	1e-561	19.7
MEF2C		0.57	1e-557	28.5
SOX17		0.61	1e-547	20.3
RUNX1		0.51	1e-522	25.5
CEBP:AP1		0.63	1e-510	15.8
IRX3/4/5		0.62	1e-506	15.5
SOX18		0.59	1e-405	11.8

C Transcription factor binding motifs enriched in venous-specific open chromatin regions (Ephb4-2+)

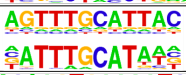
Factor	Binding Motif	% Regions	P-value	Fold-enriched
CREB1		8.8	1e-2355	3.4
ERG1		17.7	1e-1038	1.6
CTCF		4.2	1e-691	2.5
RBPJ		0.6	1e-528	12.6
SMAD2/3/4		0.5	1e-528	16.7
OCT4		0.6	1e-481	11.8
IRF1		0.6	1e-475	11.8
NFATc1		0.5	1e-439	13.3
MAFK		0.6	1e-412	11.2

Table 11: Transcription factor binding motifs enriched in venous, sprouting and arterial ATAC-seq regions ranked by p-value. *The overrepresented sequence is shown in the upper line of the binding motif. Factor denotes the transcription factor associated with the overrepresented sequence. % Regions denotes the percentage of accessible regions that contain the enriched DNA motif. Fold enrichment denotes the enrichment of the overrepresented sequence in accessible regions compared to background sequences derived from Zv10 genome.* ***A)*** *Enriched DNA motifs in sprouting (hlx-3⁺) accessible regions that also had a peak in Dll4in3⁺.* ***B)*** *Enriched DNA motifs in arterial and sprouting (Dll4in3⁺) accessible regions.* ***C)*** *Enriched DNA motifs in venous (Ephb4-2⁺) accessible regions.*

5.5 Discussion

We hypothesized that different sub-types of endothelial cells would have a different chromatin architecture due to their different molecular properties and gene expression patterns. We therefore attempted to correlate accessible *cis*-regulatory regions (ATAC-seq peaks) with global gene expression profiles (RNA-seq) in the developing zebrafish. My aim was to determine if these data can be used as a resource to specifically identify sprouting, venous and arterial-specific enhancers.

5.5.1 Accessible chromatin does not necessarily predict an active enhancer

It has previously been published that virtually all enhancers have accessible chromatin[311]. However, not only active enhancers and promoters have accessible chromatin, but also poised enhancers, silencers and gene insulators have been shown to assume an open chromatin structure[311]. I tested 19 regions identified as open chromatin through ATAC-seq for their ability to activate transcription and drive reporter gene expression in zebrafish, but only a single putative enhancer, etfb-210, was able to drive reporter gene expression in sprouting ISVs. Therefore, I concluded that ATAC-seq in sub-endothelial cell types was not a good predictor for the identification of enhancers active in a specific sub-type of endothelial cells. This may be because ATAC-seq cannot differentiate between the different types of regulatory regions. Active enhancers show p300 binding and are flanked by histone modifications such as H3K4me1 and H3K27ac[286]. Chromatin of poised enhancers is marked by H3K27me3 and H3K4me3. It is therefore possible that I would have improved my success rate if I combined my ATAC-seq data with endothelial-specific

H3K27ac and/or p300 ChIP-seq, which has been proven to detect endothelial-specific enhancers with a greater sensitivity[364].

5.5.2 Chromatin surrounding differentially transcribed genes is more accessible

Overall not many significantly different accessible regions were found across the different sub-endothelial cell types. ATAC-seq peaks surrounding differentially expressed genes were found in all sub-endothelial cell types, and chromatin regions were even accessible in sub-endothelial cell types where the associated gene is not transcribed. However, in 73% of the analysed regions, the peak score was highest in the sub-endothelial cell type in which the respective, most proximal gene was overexpressed, suggesting some correlation between the extent of open chromatin and transcription. It could be speculated that these results occur because, irrespective of the sub-endothelial cell type, all endothelial enhancers are open and bound by ETS transcription factors. These pre-established promoter-enhancer interactions could be defined and maintained by endothelial-lineage-specific ETS transcription factors[64]. In this scenario, sub-type-specific endothelial enhancers may be poised, and would only become active when additional transcription factors bind. The slightly larger ATAC-seq peak in these sub-endothelial cell types might thus reflect further binding of either arterial, venous or angiogenic-specific transcription factors or aggregation of the RNA Pol II/Mediator complex. This additional co-factor binding may occur concomitantly with the recruitment of a second, more dynamic enhancer. The observed subtle correlation between ATAC-seq peak score and

transcription will potentially help to identify sub-endothelial cell-type-specific enhancers in the future.

Despite these nuances in chromatin accessibility, my data indicate that endothelial-lineage regulatory elements are constitutively open irrespective of the sub-endothelial cell type. This proposed mechanism of sub-endothelial cell-type-specific increased chromatin accessibility is similar to the finding for the long-range ZRS enhancer and the *Sonic hedgehog* promoter. Here, promoter and enhancer are already in proximity in all limb bud cells, but in the cells where transcription is actively occurring, the posterior limb bud, promoter-enhancer interaction is even tighter[365].

Given that all endothelial cells had similar chromatin accessibility regardless of being artery, sprouting or vein-derived, alternative methods such as DNase footprinting would be needed to find the differential transcription factor binding sites within these regions of open chromatin[366].

5.5.3 Use of endogenous promoter instead of pE1b to test putative enhancers

After our study was completed, a paper was published describing ATAC-seq in endothelial cells at 24 hpf in zebrafish[306]. Quillien *et al.* did not investigate sub-endothelial cell types and only compared endothelial with non-endothelial ATAC-seq peaks. This study tested putative pan-endothelial enhancers identified by ATAC-seq. In a hypothesis driven

approach, 12 putative enhancers were identified of which nine were able to drive endothelial-specific reporter gene expression in zebrafish. For the enhancer-reporter construct Quillien *et al.* used the endogenous promoter instead of a basal minimal promoter, which increased the proportion of positive embryos and the robustness of the expression levels[306]. It could therefore be speculated that by using the endogenous promoter instead of the pE1b minimal promoter, my putative enhancers identified by ATAC-seq would have been able to activate transcription in sprouting endothelial cells.

5.5.4 Lack of known venous and sprouting genes in RNA-seq

Although my motif analysis identified the expected angiogenic and venous binding motifs in the correct endothelial sub-populations, I did not detect many known sprouting or venous-specific genes in the RNA-seq data from my sorted hlx-3^{+} and Ephb4-2^{+} populations. There are two potential explanations. First, it might be a result of PCR biases which are associated with SMARTer sequencing. Highly expressed short lengths transcripts are preferentially amplified. By correcting for gene length, I tried to minimize the occurrence of false negatives. Even though there might be a trend for detecting highly amplified shorter genes, the mean length of all zebrafish transcripts did not differ significantly from the mean length of differentially expressed genes. Finally, transcription factors are known to be expressed at low numbers which additionally hinders detection by RNA-seq[46].

Second, the lack of overlap between the observed overexpressed genes and genes known to be important for endothelial differentiation could be the chosen time point. Sub-

endothelial cell types at 27 hpf are still in the process of differentiation and do not strongly express venous and arterial markers. Specifically, despite murine enhancer Ephb4-2 driving strong venous-specific GFP expression in zebrafish, *ephb4a* was not detected to be overexpressed by qPCR or RNA-seq at 27 hpf. It is possible that the Ephb4-2 enhancer sequence is already bound and activated by the upstream transcription factors, but that there is a transcriptional delay caused by the length of *ephb4a* (34,139 bp).

Despite endothelial cells already having distinct molecular properties at the very onset of arteriovenous specification, venous and arterial cells become increasingly specified between 27 and 43 hpf. To further analyse these distinct gene expression profiles, RNA-seq and ATAC-seq data could be analysed in a time-dependent manner comparing the sub-endothelial cell types at the two time points[367]. It would be interesting to see whether chromatin accessibility is more distinct at the second time point (43 hpf). I shall correlate accessible chromatin with RNA-seq data from sub-endothelial cell types to identify regulatory elements that may be responsible for the regulation of arterial and venous-specific genes. ATAC-seq and RNA-seq data could be integrated computationally using Morpheus[368].

5.5.5 Conclusion

5.5.5.1 Possible role of other transcription factors for combinatorial regulation of sprouting

In the previous chapter, I identified MEF2 factors as the key regulators of angiogenesis. In zebrafish, mutation of any conserved region within the hlx-3:Citrine enhancer led to decreased reporter gene expression in ISVs (Chapter 4). This indicated that other transcription factors, apart from Mef2 and Ets factors, are necessary for robust transcriptional activation of *hlx1*. This points towards an additional layer of regulation for sprouting-specific *hlx1* expression. Regions in the HLX-3 enhancer that were ultra-conserved between fugu and human harboured binding sites for MEF2, ETS, CEBP:AP1, CREB1, GATA and Serum response factor (SRF) (Chapter 4). Given that these binding motifs were under selective pressure and sustained 450 million years of evolution, any of these transcription factors could play a significant role. I therefore hypothesize that, in addition to MEF2, another transcription factor may also contribute to the regulation of the hlx-3 enhancer and angiogenesis in general. HOMER *de novo* motif analysis found 1,790 sequences with a ME2C binding site in hlx-3^{+} sprouting endothelial cells (see section 5.4.3). To identify additional transcription factors, the DNA sequence adjacent to these overrepresented MEF2C DNA binding sites could be screened for recurring motifs. Additionally, *de novo* motif analysis could be performed to find a co-factor in sprouting ISVs within validated MEF2C-bound regulatory elements (such as DLL4in3, RASGRP-0.5, SH3TC1-10, ETS1+194 and ELK3-29).

5.5.5.2 RNA-seq and ATAC-seq as a resource on the UCSC browser

This is the first study that describes the transcriptional regulatory landscape of sub-endothelial cell types and provides a unique resource for arterial, venous and sprouting endothelial-specific regulatory elements in zebrafish. My ATAC-seq and RNA-seq data are accessible on UCSC browser and allows identification of novel enhancers and transcriptional regulators of endothelial development and differentiation. Although ATAC-seq is not sufficient to unambiguously identify enhancers, this study provides the basis for further experiments and downstream applications and will enhance our understanding of the transcriptional networks controlling endothelial heterogeneity.

6 Conclusion

6.1 Overall conclusion

Angiogenesis is absolutely critical for any physiological and pathological growth. Classically, Notch-VEGF signalling determines the tip and stalk cell fate of sprouting endothelial cells. However, the intricate interplay of transcription factors orchestrating gene expression downstream of VEGFA during angiogenesis has not been studied in detail. The aim of this thesis was to investigate the transcriptional regulation of angiogenesis through the analysis of an angiogenic-specific enhancer and more generally the regulatory landscape of endothelial cells through a genome-wide approach.

In Chapter 3 I identified an angiogenic-specific enhancer upstream of the *HLX*/*hlx1* locus. During early differentiation in transgenic zebrafish, hlx-3 enhancer activity correlated with that of actively proliferating sprouts. This novel enhancer was conserved between zebrafish and humans, and the HLX-3 enhancer was also active in angiogenic endothelial cells in transgenic mouse embryos. Both during development and during pathological neo-angiogenesis, HLX-3 reliably drove reporter gene expression in sprouting endothelial cells.

In Chapter 4 I have shown that MEF2 factors are the transcriptional effectors of VEGFA during angiogenesis. Analysis of the transcription factor binding sites within the HLX-3/ hlx-3 enhancer revealed a requirement for MEF2 family members. While ETS factors are also required for hlx-3 enhancer activity, they are ubiquitous to all endothelial gene enhancers and promoters. In contrast, MEF2 binding sites are enriched around genes associated with sprouting angiogenesis. Endothelial-specific deletion of *Mef2A* and *Mef2C* resulted in reduced sprouting in the mouse retina. MEF2 factors regulate the expression of *Hlx* and other key genes upregulated during sprouting angiogenesis.

In Chapter 5, I exploited the hlx-3 transgenic reporter zebrafish line (described in Chapter 3) to isolate sprouting endothelial cells by FACS. These were compared to isolated arterial and venous endothelial cells by using arterial and venous-specific transgenic zebrafish lines that were already established in the lab[167,179]. Correlation of global gene expression profile with accessible regions revealed that regulatory elements within the endothelial lineage are constitutively open irrespective of the sub-endothelial cell type. I found that accessible chromatin is not sufficient to predict enhancer activity in an agnostic approach. Strikingly, motif analysis across all accessible regions within the angiogenic-specific cell type revealed an enrichment for MEF2C factors and thus confirmed the importance of MEF2 factors during angiogenesis and our findings of Chapter 4. DNA binding motifs of other transcription factors already known to play important roles during angiogenesis or arteriovenous differentiation were also found to be enriched in the respective endothelial cell sub-type.

Overall this work increases our understanding of chromatin accessibility of regulatory regions during endothelial differentiation and exemplifies that analysis of an enhancer can be used to identify the upstream transcription factors driving angiogenesis.

6.2 Future directions

A few transcription factors have been proposed to be the key regulator of angiogenesis[354,298]. Nevertheless, the expression patterns of all transcription factors known to play a role during sprouting angiogenesis are not restricted to angiogenic cells and are even expressed outside the vasculature[151,369,370]. It is therefore unlikely that a single transcription factor determines angiogenic cell fate during differentiation. I thus propose a model whereby several transcription factors cooperatively regulate spatio-temporal-specific transcription during angiogenesis.

In my hlx-3 transgenic enhancer studies in zebrafish I tested several mutated versions of the hlx-3 enhancer and observed varying levels of reporter gene responsiveness (Chapter 3). This indicates that not just MEF2 and ETS factors, but also other co-factors cooperatively initiate transcription. In the future I propose to determine the specific combination of transcription factors that leads to full transcriptional activity of the hlx-3 enhancer in sprouting endothelial cells. Of course, there is a plethora of co-activators and transcription factors that do not bind a specific consensus sequence which may not have been picked up by our mutational analysis. To circumvent this issue, ATAC and RNA-seq data described in Chapter 5 could be used to find transcription factors which collaborate with MEF2 factors and which are needed to achieve full transcriptional activation. Last, I propose to develop a reliable approach for the identification of functional sub-endothelial cell-type-specific enhancers on a genome-wide scale in zebrafish.

The following figure depicts a working model of the transcriptional response at angiogenic-specific enhancers in response to VEGFA. Figure 56 incorporates my findings with already published information.

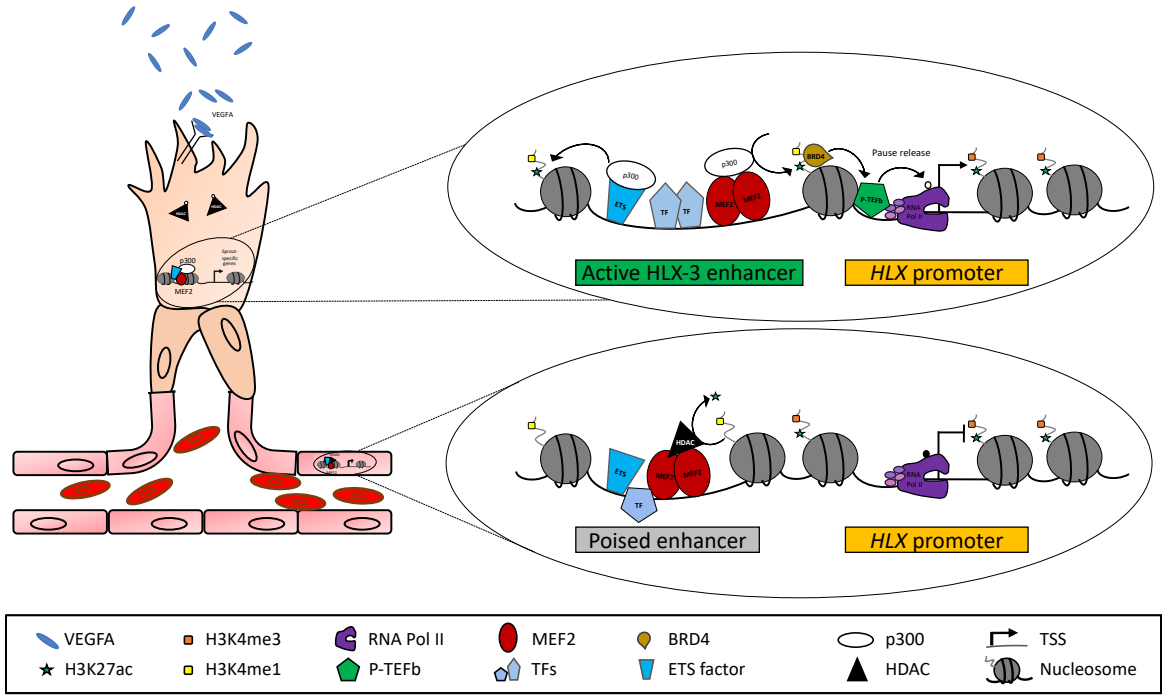


Figure 56: Working model of VEGFA-mediated transcriptional response at angiogenic-specific enhancers*. Independent of the sub-endothelial cell type, MEF2 and ETS factors are already bound to a angiogenic-specific enhancer (such as HLX-3). Chromatin is already accessible and RNA Pol II is already stalled at the promoter in quiescent, non-sprouting cells. Pro-angiogenic VEGFA signalling, probably via ERK signalling, results in HDAC phosphorylation and its cytoplasmic accumulation. This allows p300 recruitment to MEF2 and an ETS factor (potentially ERG). HAT p300 then acetylates H3K27 and ERG which subsequently stimulates BRD4 binding and activation of P-TEFb. P-TEFb phosphorylates RNA Pol II causing its pause release and transcriptional initiation. In response to VEGFA, protein occupancy increases at the HLX-3 enhancer, resulting in a further increase in chromatin accessibility in sprouting cells. In quiescent cells, transcription does not occur although the enhancer is already bound by some transcription factors. Based on my ATAC-seq data and motif analysis of the HLX-3 enhancer, the potential co-factors bound to the enhancer could be CREB1, CEBP:AP1 or SRF. This illustrates that further research is needed to find all co-factors that are responsible for full transcriptional activation of HLX-3. This figure is broadly based on Chen et al.*[288].

7 Bibliography

1. Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nature Reviews Molecular Cell Biology* **16**, 144–154 (2015).

2. Lambert, S. A. *et al.* The human transcription factors. *Cell* **172**, 650–665 (2018).

3. Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**, 642–654 (2013).

4. *Molecular biology of the cell*. (Garland Science, 2002).

5. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).

6. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421**, 448–453 (2003).

7. Luger, K. & Richmond, T. J. The histone tails of the nucleosome. *Current Opinion in Genetics & Development* **8**, 140–146 (1998).

8. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).

9. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **21**, 381–395 (2011).

10. Sterner, D. E. & Berger, S. L. Acetylation of histones and transcription-related factors. *Microbiology and Molecular Biology Reviews* **64**, 435–459 (2000).

11. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).

12. Bartman, C. R., Hsu, S. C., Hsiung, C. C.-S., Raj, A. & Blobel, G. A. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Molecular Cell* **62**, 237–247 (2016).

13. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nature Reviews Molecular Cell Biology* **19**, 621–637 (2018).

14. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology* **33**, 510–517 (2015).

15. Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).

16. Yella, V. R. & Bansal, M. DNA structural features of eukaryotic TATA-containing and TATA-less promoters. *FEBS Open Bio* **7**, 324–334 (2017).

17. Arnold, C. D. *et al.* Genome-wide assessment of sequence-intrinsic enhancer responsiveness at single-base-pair resolution. *Nature Biotechnology* **35**, 136–144 (2017).

18. Kothary, R. *et al.* Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* **105**, 707–714 (1989).

19. Huang, M., Jia, F.-J., Yan, Y.-C., Guo, L.-H. & Li, Y.-P. Transactivated minimal E1b promoter is capable of driving the expression of short hairpin RNA. *Journal of Virological Methods* **134**, 48–54 (2006).

20. Day, D. S. *et al.* Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. *Genome Biology* **17**, 120 (2016).

21. Boija, A. *et al.* Cbp regulates recruitment and release of promoter-proximal rna polymerase II. *Molecular Cell* **68**, 491-503 (2017).

22. Wernig-Zorc, S. *et al.* Global distribution of DNA hydroxymethylation and DNA methylation in chronic lymphocytic leukemia. *Epigenetics and Chromatin* **12**, 4 (2019).

23. Chouliaras, L. *et al.* Epigenetic regulation in the pathophysiology of Alzheimer's disease. *Progress in Neurobiology* **90**, 498–510 (2010).

24. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

25. Kellis, M. *et al.* Defining functional DNA elements in the human genome. *Proceedings of the National Academy of Sciences* **111**, 6131–6138 (2014).

26. Schoenfelder, S. *et al.* The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Research* **25**, 582–597 (2015).

27. Kolovos, P., Knoch, T. A., Grosveld, F. G., Cook, P. R. & Papantonis, A. Enhancers and silencers: An integrated and simple model for their function. *Epigenetics and Chromatin* **5**, 1 (2012).

28. Fishilevich, S. *et al.* GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database* **2017**, bax028 (2017).

29. Peng, Y. & Zhang, Y. Enhancer and super-enhancer: Positive regulators in gene transcription. *Animal Models and Experimental Medicine* **1**, 169–179 (2018).

30. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser-a database of tissue-specific human enhancers. *Nucleic Acids Research* **35**, D88–D92 (2007).

31. Javierre, B. M. *et al.* Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* **167**, 1369-1384 (2016).

32. Monahan, K., Horta, A. & Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* **565**, 448–453 (2019).

33. Sagai, T. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797–803 (2005).

34. Li, J. *et al.* Decoding the dynamic DNA methylation and hydroxymethylation landscapes in endodermal lineage intermediates during pancreatic differentiation of hESC. *Nucleic Acids Research* **46**, 2883–2900 (2018).

35. Rinaldi, L. *et al.* Dnmt3a and Dnmt3b associate with enhancers to regulate human epidermal stem cell homeostasis. *Cell Stem Cell* **19**, 491–501 (2016).

36. Mikhaylichenko, O. *et al.* The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes and Development* **32**, 42–57 (2018).

37. Sigova, A. A. *et al.* Transcription factor trapping by RNA in gene regulatory elements. *Science* **350**, 978–981 (2015).

38. Mousavi, K. *et al.* eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Molecular Cell* **51**, 606–617 (2013).

39. Meng, H. & Bartholomew, B. Emerging roles of transcriptional enhancers in chromatin looping and promoter-proximal pausing of RNA polymerase II. *Journal of Biological Chemistry* **293**, 13786–13794 (2018).

40. Pradeepa, M. M. *et al.* Histone H3 globular domain acetylation identifies a new class of enhancers. *Nature Genetics* **48**, 681–686 (2016).

41. Cruz-Molina, S. *et al.* PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. *Cell Stem Cell* **20**, 689-705 (2017).

42. Deaton, A. M. *et al.* Enhancer regions show high histone H3.3 turnover that changes during differentiation. *eLife* **5**, e15316 (2016).

43. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170–1187 (2016).

44. Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274–1285 (2014).

45. Iwafuchi-Doi, M. *et al.* The pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. *Molecular Cell* **62**, 79–91 (2016).

46. Biggin, M. D. Animal transcription networks as highly connected, quantitative continua. *Developmental Cell* **21**, 611–626 (2011).

47. Morgunova, E. & Taipale, J. Structural perspective of cooperative transcription factor binding. *Current Opinion in Structural Biology* **47**, 1–8 (2017).

48. Jolma, A. *et al.* DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527**, 384–388 (2015).

49. Pérez-Rico, Y. A. *et al.* Comparative analyses of super-enhancers reveal conserved elements in vertebrate genomes. *Genome Research* **27**, 259–268 (2017).

50. Blow, M. J. *et al.* ChIP-Seq identification of weakly conserved heart enhancers. *Nature Genetics* **42**, 806–810 (2010).

51. Pennacchio, L. A. *et al.* In vivo enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502 (2006).

52. Doni Jayavelu, N., Jajodia, A., Mishra, A. & Hawkins, R. D. An atlas of silencer elements for the human and mouse genomes. *bioRxiv* **2018**, 252304 (2018).

53. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A. & Luscombe, N. M. A census of human transcription factors: Function, expression and evolution. *Nature Reviews Genetics* **10**, 252–263 (2009).

54. Kim, T. K. & Maniatis, T. The mechanism of transcriptional synergy of an in vitro assembled interferon-β enhanceosome. *Molecular Cell* **1**, 119–129 (1997).

55. Lim, B., Heist, T., Levine, M. & Fukaya, T. Visualization of transvection in living drosophila embryos. *Molecular Cell* **70**, 287-296 (2018).

56. Andrey, G. & Mundlos, S. The three-dimensional genome: regulating gene expression during pluripotency and development. *Development* **144**, 3646–3658 (2017).

57. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–113 (2012).

58. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).

59. Freire-Pritchett, P. *et al.* Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells. *eLife* **6**, e21926 (2017).

60. Will, A. J. *et al.* Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nature Genetics* **49**, 1539–1545 (2017).

61. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. & Fraser, P. Long-range chromatin regulatory interactions in vivo. *Nature Genetics* **32**, 623–626 (2002).

62. Deng, W. *et al.* Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).

63. Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290–294 (2013).

64. Rubin, A. J. *et al.* Lineage-specific dynamic and pre-established enhancer–promoter contacts cooperate in terminal differentiation. *Nature Genetics* **49**, 1522–1528 (2017).

65. Ghavi-Helm, Y. *et al.* Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96–100 (2014).

66. Benabdallah, N. S. & Bickmore, W. A. Regulatory domains and their mechanisms. *Cold Spring Harbor Symposia on Quantitative Biology* **80**, 45–51 (2015).

67. Bianconi, E. *et al.* An estimation of the number of cells in the human body. *Annals of Human Biology* **40**, 463–471 (2013).

68. Drake, C. J. & Fleming, P. a. Vasculogenesis in the day 6.5 to 9.5 mouse embryo. *Blood* **95**, 1671–1679 (2000).

69. Downs, K. M. Florence Sabin and the mechanism of blood vessel lumenization during vasculogenesis. *Microcirculation* **10**, 5–25 (2003).

70. Ellertsdóttir, E. *et al.* Vascular morphogenesis in the zebrafish embryo. *Developmental Biology* **341**, 56–65 (2010).

71. Kerbel, R. S. Tumor angiogenesis. *NEJM* **358**, 2039–2049 (2008).

72. Eelen, G., Cruys, B., Welti, J., De Bock, K. & Carmeliet, P. Control of vessel sprouting by genetic and metabolic determinants. *Trends in Endocrinology and Metabolism* **24**, 589–596 (2013).

73. Van Hinsbergh, V. W. M. & Koolwijk, P. Endothelial sprouting and angiogenesis: Matrix metalloproteinases in the lead. *Cardiovascular Research* **78**, 203–212 (2008).

74. Herbert, S. P. & Stainier, D. Y. R. Molecular control of endothelial cell behaviour during blood vessel morphogenesis. *Nature Reviews Molecular Cell Biology* **12**, 551–564 (2011).

75. Nilsson, I. *et al.* VEGF receptor 2/-3 heterodimers detected in situ by proximity ligation on angiogenic sprouts. *EMBO Journal* **29**, 1377–1388 (2010).

76. Siekmann, A. F., Covassin, L. & Lawson, N. D. Modulation of VEGF signalling output by the Notch pathway. *BioEssays* **30**, 303–313 (2008).

77. Jakobsson, L. *et al.* Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nature Cell Biology* **12**, 943–953 (2010).

78. Potente, M., Gerhardt, H. & Carmeliet, P. Basic and therapeutic aspects of angiogenesis. *Cell* **146**, 873–887 (2011).

79. Pelton, J. C., Wright, C. E., Leitges, M. & Bautch, V. L. Multiple endothelial cells constitute the tip of developing blood vessels and polarize to promote lumen formation. *Development* **141**, 4121–4126 (2014).

80. Fantin, A. *et al.* Tissue macrophages act as cellular chaperones for vascular anastomosis downstream of VEGF-mediated endothelial tip cell induction. *Blood* **116**, 829–840 (2010).

81. Duran, C. L. *et al.* Molecular regulation of sprouting angiogenesis. *Comprehensive Physiology* **8**, 153–235 (2018).

82. Burri, P. H., Hlushchuk, R. & Djonov, V. C.-10. 1002/dvdy. 20184. Intussusceptive angiogenesis: Its emergence, its characteristics, and its significance. *Developmental Dynamics* **231**, 474–488 (2004).

83. Karthik, S. *et al.* Synergistic interaction of sprouting and intussusceptive angiogenesis during zebrafish caudal vein plexus development. *Scientific Reports* **8**, 9840 (2018).

84. Armulik, A. *et al.* Pericytes regulate the blood-brain barrier. *Nature* **468**, 557–561 (2010).

85. Swift, M. R. & Weinstein, B. M. Arterial-venous specification during development. *Circulation Research* **104**, 576–588 (2009).

86. Lin, F.-J., Tsai, M.-J. & Tsai, S. Y. Artery and vein formation: a tug of war between different forces. *EMBO reports* **8**, 920–924 (2007).

87. Chi, J.-T. *et al.* Endothelial cell diversity revealed by global expression profiling. *Proceedings of the National Academy of Sciences* **100**, 10623–10628 (2003).

88. Nolan, D. J. *et al.* Molecular signatures of tissue-specific microvascular endothelial cell heterogeneity in organ maintenance and regeneration. *Developmental Cell* **26**, 204–219 (2013).

89. Folkman, M. J., Long, D. M. & Becker, F. F. Tumor growth in organ culture. *Surg Forum* **13**, 81–83 (1962).

90. Wang, G. L., Jiang, B. H., Rue, E. A. & Semenza, G. L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5510–5514 (1995).

91. Forsythe, J. A. *et al.* Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Molecular and Cellular Biology* **16**, 4604–4613 (1996).

92. Hayakawa, H. & Shibasaki, F. Regulation of angiogenesis by hypoxia-inducible factors. *Biochemical Basis and Therapeutic Implications of Angiogenesis* **6**, 93–106 (2013).

93. Bussmann, J., Lawson, N., Zon, L., Schulte-Merker, S. & Zebrafish Nomenclature Committee. Zebrafish VEGF receptors: A guideline to nomenclature. *PLoS Genetics* **4**, e1000064 (2008).

94. Shalaby, F. *et al.* Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature* **376**, 62–66 (1995).

95. Peach, C. J. *et al.* Molecular pharmacology of VEGF-A isoforms: Binding and signalling at VEGFR2. *International Journal of Molecular Sciences* **19**, 1264 (2018).

96. Stalmans, I. *et al.* Arteriolar and venular patterning in retinas of mice selectively expressing VEGF isoforms. *Journal of Clinical Investigation* **109**, 327–336 (2002).

97. Nowak, D. G. *et al.* Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors. *Journal of Cell Science* **121**, 3487–3495 (2008).

98. Coultas, L. *et al.* Hedgehog regulates distinct vascular patterning events through VEGF-dependent and -independent mechanisms. *Blood* **116**, 653–660 (2010).

99. Raimondi, C., Brash, J. T., Fantin, A. & Ruhrberg, C. NRP1 function and targeting in neurovascular development and eye disease. *Progress in Retinal and Eye Research* **52**, 64–83 (2016).

100. Gerhardt, H. *et al.* VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia. *The Journal of Cell Biology* **161**, 1163–1177 (2003).

101. Tammela, T. *et al.* VEGFR-3 controls tip to stalk conversion at vessel fusion sites by reinforcing Notch signalling. *Nature Cell Biology* **13**, 1202–1213 (2011).

102. Zarkada, G., Heinolainen, K., Makinen, T., Kubota, Y. & Alitalo, K. VEGFR3 does not sustain retinal angiogenesis without VEGFR2. *Proceedings of the National Academy of Sciences* **112**, 761–766 (2015).

103. Benedito, R. *et al.* Notch-dependent VEGFR3 upregulation allows angiogenesis without VEGF-VEGFR2 signalling. *Nature* **484**, 110–114 (2012).

104. Grego-Bessa, J. *et al.* Notch signaling is essential for ventricular chamber development. *Developmental Cell* **12**, 415–429 (2007).

105. Duarte, A. *et al.* Dosage-sensitive requirement for mouse Dll4 in artery development. *Genes and Development* **18**, 2474–2478 (2004).

106. Liu, Z.-J. *et al.* Regulation of Notch1 and Dll4 by vascular endothelial growth factor in arterial endothelial cells: implications for modulating arteriogenesis and angiogenesis. *Molecular and Cellular Biology* **23**, 14–25 (2003).

107. Jakobsson, L., Bentley, K. & Gerhardt, H. VEGFRs and Notch: a dynamic collaboration in vascular patterning. *Biochemical Society Transactions* **37**, 1233–1236 (2009).

108. Bray, S. J. Notch signalling: A simple pathway becomes complex. *Nature Reviews Molecular Cell Biology* **7**, 678–689 (2006).

109. Wilson, J. J. & Kovall, R. A. Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA. *Cell* **124**, 985–996 (2006).

110. Kappas, N. C. *et al.* The VEGF receptor Flt-1 spatially modulates Flk-1 signaling and blood vessel branching. *Journal of Cell Biology* **181**, 847–858 (2008).

111. Geudens, I. & Gerhardt, H. Coordinating cell behaviour during blood vessel formation. *Development* **138**, 4569–4583 (2011).

112. Benedito, R. *et al.* The Notch Ligands Dll4 and Jagged1 Have Opposing Effects on Angiogenesis. *Cell* **137**, 1124–1135 (2009).

113. Zimrin, A. B. *et al.* An antisense oligonucleotide to the Notch ligand Jagged enhances fibroblast growth factor-induced angiogenesis in vitro. *Journal of Biological Chemistry* **271**, 32499–32502 (1996).

114. Del Toro, R. *et al.* Identification and functional analysis of endothelial tip cell-enriched genes. *Blood* **116**, 4025–4033 (2010).

115. Eilken, H. M. & Adams, R. H. Dynamics of endothelial cell behavior in sprouting angiogenesis. *Current Opinion in Cell Biology* **22**, 617–625 (2010).

116. De Val, S. & Black, B. L. Transcriptional control of endothelial cell development. *Developmental Cell* **16**, 180–195 (2009).

117. Park, C., Kim, T. M. & Malik, A. B. Transcriptional regulation of endothelial cell and vascular development. *Circulation Research* **112**, 1380–1400 (2013).

118. Hamik, A., Wang, B. & Jain, M. K. Transcriptional regulators of angiogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **26**, 1936–1947 (2006).

119. Jia, J. *et al.* AP-1 transcription factor mediates VEGF-induced endothelial cell migration and proliferation. *Microvascular Research* **105**, 103–108 (2016).

120. Kim, K. *et al.* SoxF Transcription factors are positive feedback regulators of VEGF signaling. *Circulation Research* **119**, 839–852 (2016).

121. Lee, S.-H. *et al.* Notch pathway targets proangiogenic regulator Sox17 to restrict angiogenesis. *Circulation Research* **115**, 215–226 (2014).

122. Larrivée, B. *et al.* Alk1 signaling inhibits angiogenesis by cooperating with the Notch pathway. *Developmental Cell* **22**, 489–500 (2012).

123. Seo, S. *et al.* The forkhead transcription factors, Foxc1 and Foxc2, are required for arterial specification and lymphatic sprouting during vascular development. *Developmental Biology* **294**, 458–470 (2006).

124. Sato, Y. Role of ETS family transcription factors in vascular development and angiogenesis. *Cell Struct. Funct.* **26**, 19–24 (2001).

125. De Val, S. *et al.* Combinatorial regulation of endothelial gene expression by Ets and Forkhead transcription factors. *Cell* **135**, 1053–1064 (2008).

126. Hollenhorst, P. C., Shah, A. A., Hopkins, C. & Graves, B. J. Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes and Development* **21**, 1882–1894 (2007).

127. Wei, G. *et al.* Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood* **114**, 1123–1130 (2009).

128. Pham, V. N. *et al.* Combinatorial function of ETS transcription factors in the developing vasculature. *Developmental Biology* **303**, 772–783 (2007).

129. Lee, D. *et al.* ER71 acts downstream of BMP, Notch, and WNT signaling in blood and vessel progenitor specification. *Cell Stem Cell* **2**, 497–507 (2008).

130. Sumanas, S. & Lin, S. Ets1-related protein is a key regulator of vasculogenesis in zebrafish. *PLoS Biology* **4**, e10 (2005).

131. Kanki, Y. *et al.* Dynamically and epigenetically coordinated GATA/ETS/SOX transcription factor expression is indispensable for endothelial cell differentiation. *Nucleic Acids Research* **45**, 4344–4358 (2017).

132. Liu, F. *et al.* Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. *EMBO reports* **16**, 654–669 (2015).

133. Park, C. *et al.* Injury-mediated vascular regeneration requires endothelial ER71/ETV2. *Arteriosclerosis, Thrombosis, and Vascular Biology* **36**, 86–96 (2016).

134. Craig, M. P. & Sumanas, S. ETS transcription factors in embryonic vascular development. *Angiogenesis* **19**, 275–285 (2016).

135. Ayadi, A., Suelves, M., Dollé, P. & Wasylyk, B. Net, an Ets ternary complex transcription factor, is expressed in sites of vasculogenesis, angiogenesis, and chondrogenesis during mouse development. *Mechanisms of Development* **102**, 205–208 (2001).

136. Zheng, H. The transcription factor Net regulates the angiogenic switch. *Genes and Development* **17**, 2283–2297 (2003).

137. Birdsey, G. M. *et al.* Transcription factor Erg regulates angiogenesis and endothelial apoptosis through VE-cadherin. *Blood* **111**, 3498–3506 (2008).

138. Meadows, S. M., Myers, C. T. & Krieg, P. A. Regulation of endothelial cell development by ETS transcription factors. *Seminars in Cell and Developmental Biology* **22**, 976–984 (2011).

139. Randi, A. M., Sperone, A., Dryden, N. H. & Birdsey, G. M. Regulation of angiogenesis by ETS transcription factors. *Biochemical Society Transactions* **37**, 1248–1253 (2009).

140. Huang, X. Critical role for the Ets transcription factor ELF-1 in the development of tumor angiogenesis. *Blood* **107**, 3153–3160 (2006).

141. Shi, X. *et al.* Cooperative interaction of Etv2 and Gata2 regulates the development of endothelial and hematopoietic lineages. *Developmental Biology* **389**, 208–218 (2014).

142. Potthoff, M. J. & Olson, E. N. MEF2: a central regulator of diverse developmental programs. *Development* **134**, 4131–4140 (2007).

143. Wu, Y. *et al.* Structure of the MADS-box/MEF2 Domain of MEF2A bound to DNA and its implication for Myocardin recruitment. *Journal of Molecular Biology* **397**, 520–533 (2010).

144. Desjardins, C. & Naya, F. The function of the MEF2 family of transcription factors in cardiac development, cardiogenomics, and direct reprogramming. *Journal of Cardiovascular Development and Disease* **3**, 26 (2016).

145. Pon, J. R. & Marra, M. A. MEF2 transcription factors: developmental regulators and emerging cancer genes. *Oncotarget* **7**, 2297-2312 (2016).

146. Subramanian, S. V. & Nadal-Ginard, B. Early expression of the different isoforms of the myocyte enhancer factor-2 (MEF2) protein in myogenic as well as non-myogenic cell lineages during mouse embryogenesis. *Mechanisms of Development* **57**, 103–112 (1996).

147. Edmonson, D. G., Lyons, G. E., Martin, J. F. & Olson, E. N. *Mef2* gene-expression marks the cardiac and skeletal-muscle lineages during mouse embryogenesis. *Development* **120**, 1251–1263 (1994).

148. Lin, Q. *et al.* Requirement of the MADS-box transcription factor MEF2C for vascular development. *Development* **125**, 4565–4574 (1998).

149. Wang, L., Fan, C., Topol, S. E., Topol, E. J. & Wang, Q. Mutation of MEF2A in an inherited disorder with features of coronary artery disease. *Science* **302**, 1578–1581 (2003).

150. Xu, Z. *et al.* MEF2C ablation in endothelial cells reduces retinal vessel loss and suppresses pathologic retinal neovascularization in oxygen-induced retinopathy. *American Journal of Pathology* **180**, 2548–2560 (2012).

151. Liu, N. *et al.* Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proceedings of the National Academy of Sciences* **111**, 4109–4114 (2014).

152. De Val, S. *et al.* Mef2c is activated directly by Ets transcription factors through an evolutionarily conserved endothelial cell-specific enhancer. *Developmental Biology* **275**, 424–434 (2004).

153. Rivera, C. G., Mellberg, S., Claesson-Welsh, L., Bader, J. S. & Popel, A. S. Analysis of VEGF-A regulated gene expression in endothelial cells to identify genes linked to angiogenesis. *PLoS ONE* **6**, e24887 (2011).

154. Schweighofer, B. *et al.* The VEGF-induced transcriptional response comprises gene clusters at the crossroad of angiogenesis and inflammation. *Thrombosis and Haemostasis* **102**, 544–554 (2009).

155. Maiti, D., Xu, Z. & Duh, E. J. Vascular endothelial growth factor induces MEF2C and MEF2-dependent activity in endothelial cells. *Investigative Ophthalmology and Visual Science* **49**, 3640 (2008).

156. Hayashi, M. *et al.* Targeted deletion of BMK1/ERK5 in adult mice perturbs vascular integrity and leads to endothelial failure. *Journal of Clinical Investigation* **113**, 1138–1148 (2004).

157. Lints, T. J., Hartley, L., Parsons, L. M. & Harvey, R. P. Mesoderm-specific expression of the divergent homeobox gene Hlx during murine embryogenesis. *Developmental Dynamics* **205**, 457–470 (1996).

158. Nagel, S. *et al.* Aberrant expression of NKL homeobox gene HLX in Hodgkin lymphoma. *Oncotarget* **9**, 14338-14353 (2018).

159. Bates, M. D., Schatzman, L. C., Harvey, R. P. & Potter, S. S. Two CCAAT boxes in a novel inverted repeat motif are required for Hlx homeobox gene expression. *Biochimica et Biophysica Acta - Gene Structure and Expression* **1519**, 96–105 (2001).

160. Kawahara, M. *et al.* H2.0-like Homeobox regulates early hematopoiesis and promotes acute myeloid leukemia. *Cancer Cell* **22**, 194–208 (2012).

161. Prahst, C. *et al.* The H2.0-like homeobox transcription factor modulates yolk sac vascular remodeling in mouse embryos. *Arteriosclerosis, Thrombosis, and Vascular Biology* **34**, 1468–1476 (2014).

162. Bates, M. D., Dunagan, D. T., Welch, L. C., Kaul, A. & Harvey, R. P. The Hlx homeobox transcription factor is required early in enteric nervous system development. *BMC Developmental Biology* **6,** 33 (2006).

163. Hentsch, B. *et al.* Hlx homeo box gene is essential for an inductive tissue interaction that drives expansion of embryonic liver and gut. *Genes and Development* **10**, 70–79 (1996)

164. Herbert, S. P., Cheung, J. Y. M. & Stainier, D. Y. R. Determination of endothelial stalk versus tip cell potential during angiogenesis by H2.0-like homeobox-1. *Current Biology* **22**, 1789–1794 (2012).

165. Chávez, M. N., Aedo, G., Fierro, F. A., Allende, M. L. & Egaña, J. T. Zebrafish as an emerging model organism to study angiogenesis in development and regeneration. *Frontiers in Physiology* **7**, 56 (2016).

166. Ittner, L. M. & Götz, J. Pronuclear injection for the production of transgenic mice. *Nature Protocols* **2**, 1206–1215 (2007).

167. Sacilotto, N. *et al.* Analysis of Dll4 regulation reveals a combinatorial role for Sox and Notch in arterial development. *Proceedings of the National Academy of Sciences* **110**, 11893–11898 (2013).

168. Vong, L. H., Ragusa, M. J. & Schwarz, J. J. Generation of conditional Mef2c loxp/loxp mice for temporal- and tissue-specific analyses. *Genesis* **43**, 43–48 (2005).

169. Lu, Y. W. *et al.* Endothelial Myocyte enhancer factor 2c inhibits migration of smooth muscle cells through fenestrations in the internal elastic lamina. *Arteriosclerosis, Thrombosis, and Vascular Biology* **37**, 1380–1390 (2017).

170. Sacilotto, N. *et al.* MEF2 transcription factors are key regulators of sprouting angiogenesis. *Genes and Development* **30**, 2297–2309 (2016).

171. Schatz, O., Golenser, E. & Ben-Arie, N. Clearing and photography of whole mount X-gal stained mouse embryos. *BioTechniques* **39**, 650–656 (2005).

172. Tual-Chalot, S., Allinson, K. R., Fruttiger, M. & Arthur, H. M. Whole Mount immunofluorescent staining of the neonatal mouse retina to investigate angiogenesis in vivo. *Journal of Visualized Experiments* **77**, e50546 (2013)

173. Fantin, A. & Ruhrberg, C. The embryonic mouse hindbrain and postnatal retina as in vivo models to study angiogenesis. *Methods in Molecular Biology* **1332**, 177–188 (2015).

174. Laitinen, L. Griffonia simplicifolia lectins bind specifically to endothelial cells and some epithelial cells in mouse tissues. *The Histochemical Journal* **19**, 225–234 (1987).

175. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).

176. Arganda-Carreras, I., Fernández-González, R., Muñoz-Barrutia, A. & Ortiz-De-Solorzano, C. 3D reconstruction of histological sections: Application to mammary gland tissue. *Microscopy Research and Technique* **73**, 1019–1029 (2010).

177. Fantin, A., Vieira, J. M., Plein, A., Maden, C. H. & Ruhrberg, C. The embryonic mouse hindbrain as a qualitative and quantitative model for studying the molecular and cellular mechanisms of angiogenesis. *Nature Protocols* **8**, 418–429 (2013).

178. Jin, S.-W., Beis, D., Mitchell, T., Chen, J.-N. & Stainier, D. Y. R. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**, 5199–5209 (2005).

179. Neal, A. *et al.* Venous identity requires BMP signalling through ALK3. *Nature Communications* **10**, 453 (2019).

180. Kawakami, K. Transposon tools and methods in zebrafish. *Developmental Dynamics* **234**, 244–254 (2005).

181. Sehnert, A. J. *et al.* Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. *Nature Genetics* **31**, 106–110 (2002).

182. Eisen, J. S. & Smith, J. C. Controlling morpholino experiments: don't stop making antisense. *Development* **135**, 1735–1743 (2008).

183. Kawakami, K. Tol2: A versatile gene transfer vector in vertebrates. *Genome Biology* **8**, S7 (2007).

184. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Developmental* **203**, 253–310 (1995)

185. dela Paz, N. G., Walshe, T. E., Leach, L. L., Saint-Geniez, M. & D'Amore, P. A. Role of shear-stress-induced VEGF expression in endothelial cell survival. *Journal of Cell Science* **125**, 831–843 (2012).

186. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* **76**, 4350–4354 (1979).

187. Guo, H. *et al.* An efficient procedure for protein extraction from formalin-fixed, Paraffin-embedded tissues for reverse phase protein arrays. *Proteome Science* **10**, 56 (2012).

188. Bertrand, J. Y., Kim, A. D., Teng, S. & Traver, D. CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development* **135**, 1853–1862 (2008).

189. Lee, Y. *et al.* FGF signalling specifies haematopoietic stem cells through its regulation of somitic Notch signalling. *Nature Communications* **5**, 5583 (2014).

190. Krueger, J. *et al.* Flt1 acts as a negative regulator of tip cell formation and branching morphogenesis in the zebrafish embryo. *Development* **138**, 2111–2120 (2011).

191. Bertrand, J. Y. *et al.* Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* **464**, 108–111 (2010).

192. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols in Molecular Biology* **109**, 21.29.1-21.29.9 (2015).

193. Andrews, S. FastQC: A quality control tool for high throughput sequence data. (2010).

194. Joshi, N. & Fass, J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. *https://github.com/najoshi/sickle* (2011).

195. Howe, K. *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013).

196. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).

197. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

198. Broad Institute. Picard tools. *https://broadinstitute.github.io/picard/* (2016).

199. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137 (2008).

200. Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Research* **12**, 996–1006 (2002).

201. Benner, C., Heinz, S. & Glass, C. K. HOMER - Software for motif discovery and next generation sequencing analysis. *http://Homer.Ucsd.Edu/* (2017).

202. Chen, T. W. *et al.* ChIPseek, a web-based analysis tool for ChIP data. *BMC Genomics* **15**, 539 (2014).

203. Dobrzycki, T. *et al.* A zebrafish model for MonoMAC syndrome identifies an earlier role for gata2 in haemogenic endothelium programming and generation of haematopoietic stem cells. *bioRxiv* **2019**, 516203 (2019).

204. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

205. R Development Core Team, R. *R: A Language and Environment for Statistical Computing*. *R Foundation for Statistical Computing* (2011).

206. Stark, R. & Brown, G. DiffBind : differential binding analysis of ChIP-Seq peak data. *Bioconductor* (2011).

207. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

208. Wilkinson, L. ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. *Biometrics* **67**, 678–679 (2011).

209. Krueger, F. Trim Galore. *Babraham Bioinformatics*. http://www.bioinformatics.babraham.ac.uk/projects/trim galore (2016)

210. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

211. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

212. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

213. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**, R25 (2010).

214. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* **57**, 289–300 (1995).

215. Chen, Y., Lun, A. T. L. & Smyth, G. K. Differential expression analysis of complex RNA-seq experiments using edgeR. *Statistical analysis of Next Generation sequence data* (2014).

216. Warnes, G. R. *et al.* Package 'gplots': Various R programming tools for plotting data. *R package version 2.17.0.* (2016).

217. Eklund, L., Bry, M. & Alitalo, K. Mouse models for studying angiogenesis and lymphangiogenesis in cancer. *Molecular Oncology* **7**, 259–282 (2013).

218. Becker, P. W. *et al.* An intronic Flk1 enhancer directs arterial-specific expression via RBPJ-mediated venous repression. *Arteriosclerosis, Thrombosis, and Vascular Biology* **36**, 1209–1219 (2016).

219. Li, I. M. H. *et al.* Differential tissue specific, temporal and spatial expression patterns of the Aggrecan gene is modulated by independent enhancer elements. *Scientific Reports* **8**, 950 (2018).

220. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Molecular Cell* **49**, 825–837 (2013).

221. Visel, A., Rubin, E. M. & Pennacchio, L. A. Genomic views of distant-acting enhancers. *Nature* **461**, 199–205 (2009).

222. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).

223. Ong, C. T. & Corces, V. G. Enhancer function: New insights into the regulation of tissue-specific gene expression. *Nature Reviews Genetics* **12**, 283–293 (2011).

224. Rosenbloom, K. R. *et al.* ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Research* **41**, D56–D63 (2012).

225. Aydogan, V. *et al.* Endothelial cell division in angiogenic sprouts of differing cellular architecture. *Biology Open* **4**, 1259–1269 (2015).

226. Costa, G. *et al.* Asymmetric division coordinates collective cell migration in angiogenesis. *Nature Cell Biology* **18**, 1292–1301 (2016).

227. Young, R. C. & Weinstein, B. M. Visualization and experimental analysis of blood vessel formation using transgenic zebrafish. *Birth Defects Research* **81**, 286–296 (2007).

228. Isogai, S., Horiguchi, M. & Weinstein, B. M. The vascular anatomy of the developing zebrafish: An atlas of embryonic and early larval development. *Developmental Biology* **230**, 278–301 (2001).

229. Blum, Y. *et al.* Complex cell rearrangements during intersegmental vessel sprouting and vessel fusion in the zebrafish embryo. *Developmental Biology* **316**, 312–322 (2008).

230. Hogan, B. M. *et al.* Ccbe1 is required for embryonic lymphangiogenesis and venous sprouting. *Nature Genetics* **41**, 396–398 (2009).

231. Chong, D. C., Koo, Y., Xu, K., Fu, S. & Cleaver, O. Stepwise arteriovenous fate acquisition during mammalian vasculogenesis. *Developmental Dynamics* **240**, 2153–2165 (2011).

232. Covassin, L. D., Villefranc, J. A., Kacergis, M. C., Weinstein, B. M. & Lawson, N. D. Distinct genetic interactions between multiple Vegf receptors are required for development of different blood vessel types in zebrafish. *Proceedings of the National Academy of Sciences* **103**, 6554–6559 (2006).

233. Hen, G. *et al.* Venous-derived angioblasts generate organ-specific vessels during zebrafish embryonic development. *Development* **142**, 4266–4278 (2015).

234. Bill, B. R. *et al.* Development and Notch signaling requirements of the zebrafish choroid plexus. *PLoS ONE* **3**, e3114 (2008).

235. Graham, E. *et al.* The atlas of mouse development eHistology resource. *Development* **142**, 2545–2545 (2015).

236. Walls, J. R., Coultas, L., Rossant, J. & Henkelman, R. M. Three-dimensional analysis of vascular development in the mouse embryo. *PLoS ONE* **3**, e2853 (2008).

237. Chen, H. I. *et al.* The sinus venosus contributes to coronary vasculature through VEGFC-stimulated angiogenesis. *Development* **141**, 4500–4512 (2014).

238. Red-Horse, K., Ueno, H., Weissman, I. L. & Krasnow, M. A. Coronary arteries form by developmental reprogramming of venous cells. *Nature* **464**, 549–553 (2010).

239. Wu, B. *et al.* Endocardial cells form the coronary arteries by angiogenesis through myocardial-endocardial VEGF signaling. *Cell* **151**, 1083–1096 (2012).

240. Tian, X. *et al.* De novo formation of a distinct coronary vascular population in neonatal heart. *Science* **345**, 90–94 (2014).

241. Kusumbe, A. P., Ramasamy, S. K. & Adams, R. H. Coupling of angiogenesis and osteogenesis by a specific vessel subtype in bone. *Nature* **507**, 323–328 (2014).

242. Ramasamy, S. K., Kusumbe, A. P., Wang, L. & Adams, R. H. Endothelial Notch activity promotes angiogenesis and osteogenesis in bone. *Nature* **507**, 376–380 (2014).

243. Stahl, A. *et al.* The mouse retina as an angiogenesis model. *Investigative Ophthalmology and Visual Science* **51**, 2813 (2010).

244. Ye, X., Wang, Y. & Nathans, J. The Norrin/Frizzled4 signaling pathway in retinal vascular development and disease. *Trends in Molecular Medicine* **16**, 417–425 (2010).

245. Pitulescu, M. E., Schmidt, I., Benedito, R. & Adams, R. H. Inducible gene targeting in the neonatal vasculature and analysis of retinal angiogenesis in mice. *Nature Protocols* **5**, 1518–1534 (2010).

246. Sorokin, S. P., Hoyt, R. F., Blunt, D. G. & McNelly, N. A. Macrophage development: II. Early ontogeny of macrophage populations in brain, liver, and lungs of rat embryos as revealed by a lectin marker. *The Anatomical Record* **232**, 527–550 (1992).

247. Fruttiger, M. Development of the mouse retinal vasculature: Angiogenesis versus vasculogenesis. *Investigative Ophthalmology and Visual Science* **43**, 522–527 (2002).

248. Beck, M. *et al.* The quantitative proteome of a human cell line. *Molecular Systems Biology* **7**, 549–549 (2014).

249. Bates, M. D., Wells, J. M. & Venkatesh, B. Comparative genomics of the Hlx homeobox gene and protein: Conservation of structure and expression from fish to mammals. *Gene* **352**, 45–56 (2005).

250. Isogai, S. Angiogenic network formation in the developing vertebrate trunk. *Development* **130**, 5281–5290 (2003).

251. Gore, A. V. *et al.* Rspo1/Wnt signaling promotes angiogenesis via Vegfc/Vegfr3. *Development* **138**, 4875–4886 (2011)

252. Li, W., Chen, J., Deng, M. & Jing, Q. The zebrafish Tie2 signaling controls tip cell behaviors and acts synergistically with Vegf pathway in developmental angiogenesis. *Acta Biochimica et Biophysica Sinica* **46**, 641–646 (2014).

253. Zygmunt, T. *et al.* 'In parallel' interconnectivity of the dorsal longitudinal anastomotic vessels requires both VEGF signaling and circulatory flow. *Journal of Cell Science* **125**, 5159–5167 (2012).

254. Hogan, K. A. & Bautch, V. L. Blood vessel patterning at the embryonic midline. *Current Topics in Developmental Biology* **62**, 55–85 (2004).

255. Cleaver, O. & Krieg, P. A. VEGF mediates angioblast migration during development of the dorsal aorta in Xenopus. *Development* **19**, 3905–3914 (1998).

256. Garriock, R. J., Czeisler, C., Ishii, Y., Navetta, A. M. & Mikawa, T. An anteroposterior wave of vascular inhibitor downregulation signals aortae fusion along the embryonic midline axis. *Development* **137**, 3697–3706 (2010).

257. Pitulescu, M. E. *et al.* Dll4 and Notch signalling couples sprouting angiogenesis and artery formation. *Nature Cell Biology* **19**, 915–927 (2017).

258. Xu, C. *et al.* Arteries are formed by vein-derived endothelial tip cells. *Nature Communications* **5**, 5758 (2014).

259. Karkkainen, M. J. *et al.* Vascular endothelial growth factor C is required for sprouting of the first lymphatic vessels from embryonic veins. *Nature Immunology* **5**, 74–80 (2004).

260. Yang, Y. *et al.* Lymphatic endothelial progenitors bud from the cardinal vein and intersomitic vessels in mammalian embryos. *Blood* **120**, 2340–2348 (2012).

261. Ehling, M., Adams, S., Benedito, R. & Adams, R. H. Notch controls retinal blood vessel maturation and quiescence. *Development* **140**, 3051–3061 (2013).

262. Feige, J. J., Pagès, G. & Soncin, F. Molecular mechanisms of angiogenesis: From ontogenesis to oncogenesis. (Springer, 2014).

263. Martin, N. *et al.* Interplay between Homeobox proteins and Polycomb repressive complexes in p16 INK4a regulation. *EMBO Journal* **32**, 982–995 (2013).

264. Testori, J. *et al.* The VEGF-regulated transcription factor HLX controls the expression of guidance cues and negatively regulates sprouting of endothelial cells. *Blood* **117**, 2735–2744 (2011).

265. Kleinman, H. K. & Martin, G. R. Matrigel: Basement membrane matrix with biological activity. *Seminars in Cancer Biology* **15**, 378–386 (2005).

266. Chen, L. *et al.* Ligustrazine inhibits B16F10 melanoma metastasis and suppresses angiogenesis induced by Vascular endothelial growth factor. *Biochemical and Biophysical Research Communications* **386**, 374–379 (2009).

267. Einspahr, J. G. *et al.* Expression of Vascular endothelial growth factor in early cutaneous melanocytic lesion progression. *Cancer* **110**, 2519–2527 (2007).

268. Itokawa, T. *et al.* Antiangiogenic effect by SU5416 is partly attributable to inhibition of FIt-1 receptor signaling. *Mol.Cancer Ther.* **1**, 295–302 (2002).

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

270. Geling, A., Steiner, H., Willem, M., Bally-Cuif, L. & Haass, C. A γ-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO Reports* **3**, 688–694 (2002).

271. Nicoli, S. *et al.* MicroRNA-mediated integration of haemodynamics and Vegf signalling during angiogenesis. *Nature* **464**, 1196–1200 (2010).

272. Jin, Z.-G. Ligand-independent activation of Vascular endothelial growth factor receptor 2 by fluid shear stress regulates activation of endothelial Nitric oxide synthase. *Circulation Research* **93**, 354–363 (2003).

273. Packham, I. M. *et al.* Microarray profiling reveals CXCR4a is downregulated by blood flow in vivo and mediates collateral formation in zebrafish embryos. *Physiol Genomics* **38**, 319–327 (2009).

274. Cermenati, S. *et al.* Sox 18 and Sox7 play redundant roles in vascular development. *Blood* **111**, 2657–2666 (2008).

275. Lenard, A. *et al.* Endothelial cell self-fusion during vascular pruning. *PLoS Biology* **13**, e1002126 (2015).

276. Montero-Balaguer, M. *et al.* Stable vascular connections and remodeling require full expression of VE-cadherin in zebrafish embryos. *PLoS ONE* **4**, e5772 (2009).

277. Huddleson, J. P., Ahmad, N., Srinivasan, S. & Lingrel, J. B. Induction of KLF2 by fluid shear stress requires a novel promoter element activated by a phosphatidylinositol 3-kinase-dependent chromatin-remodeling pathway. *Journal of Biological Chemistry* **280**, 23371–23379 (2005).

278. Larkin, M. *et al.* ClustalW and ClustalX version 2. *Bioinformatics* **23**, 2947–2948 (2007).

279. Kappel, A. *et al.* Role of SCL/Tal-1, GATA, and ETS transcription factor binding sites for the regulation of Flk-1 expression during murine vascular development. *Blood* **9**, 3078–3085 (2000).

280. Wythe, J. D. *et al.* ETS factors regulate Vegf-dependent arterial specification. *Developmental Cell* **26**, 45–58 (2013).

281. Karim, F. D. *et al.* The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes and Development* **4**, 1451–1453 (1990).

282. Brown, L. A. *et al.* Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. *Mechanisms of Development* **90**, 237–252 (2000).

283. Black, B. L. & Olson, E. N. Transcriptional control of muscle development by Myocyte enhancer factor-2 (MEF) proteins. *Annual Review of Cell and Developmental Biology* **14**, 167–196 (1998).

284. Maejima, T. *et al.* Direct evidence for pitavastatin induced chromatin structure change in the KLF4 gene in endothelial cells. *PLoS ONE* **9**, e96005 (2014).

285. Zhang, B. *et al.* A dynamic H3K27ac signature identifies VEGFA-stimulated endothelial enhancers and requires EP300 activity. *Genome Research* **23**, 917–927 (2013).

286. Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).

287. Roberts, D. M. *et al.* A vascular gene trap screen defines RasGRP3 as an angiogenesis-regulated gene required for the endothelial response to phorbol esters. *Mol Cell Biol* **24**, 10515–10528 (2004).

288. Chen, J. *et al.* VEGF amplifies transcription through ETS1 acetylation to enable angiogenesis. *Nature Communications* **8**, 383 (2017).

289. Watanabe, D. *et al.* Transcription factor Ets-1 mediates Ischemia- and Vascular endothelial growth factor-dependent retinal neovascularization. *American Journal of Pathology* **164**, 1827–1835 (2004).

290. Heo, S. H. & Cho, J. Y. ELK3 suppresses angiogenesis by inhibiting the transcriptional activity of ETS-1 on MT1-MMP. *International Journal of Biological Sciences* **10**, 438–447 (2014).

291. Weinl, C. *et al.* Elk3 deficiency causes transient impairment in post-natal retinal vascular development and formation of tortuous arteries in adult murine retinae. *PLoS ONE* **9**, e107048 (2014).

292. Materna, S. C., Sinha, T., Barnes, R. M., Lammerts van Bueren, K. & Black, B. L. Cardiovascular development and survival require Mef2c function in the myocardial but not the endothelial lineage. *Developmental Biology* **445**, 170–177 (2019).

293. Wang, Y. *et al.* Ephrin-B2 controls VEGF-induced angiogenesis and lymphangiogenesis. *Nature* **465**, 483–486 (2010).

294. Akhtar, M. W. *et al.* In vivo analysis of MEF2 transcription factors in synapse regulation and neuronal survival. *PLoS ONE* **7**, e34863 (2012).

295. Payne, S., De Val, S. & Neal, A. Endothelial-specific CRE mouse models: is your CRE credibile? *Arteriosclerosis, Thrombosis, and Vascular Biology* **38**, 2550–2561 (2018).

296. Hellström, M. *et al.* Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature* **445**, 776–780 (2007).

297. Sturtzel, C., Testori, J., Schweighofer, B., Bilban, M. & Hofer, E. The transcription factor MEF2C negatively controls angiogenic sprouting of endothelial cells depending on oxygen. *PLoS ONE* **9**, e101521 (2014).

298. Fish, J. E. *et al.* Dynamic regulation of VEGF-inducible genes by an ERK/ERG/p300 transcriptional network. *Development* **144**, 2428–2444 (2017).

299. Lu, J., McKinsey, T. A., Nicol, R. L. & Olson, E. N. Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proceedings of the National Academy of Sciences* **97**, 4070–4075 (2000).

300. Lu, J., McKinsey, T. A., Zhang, C. L. & Olson, E. N. Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. *Molecular Cell* **6**, 233–244 (2000).

301. Chang, H. H. *et al.* Protein kinase D-dependent phosphorylation and nuclear export of histone deacetylase 5 mediates vascular endothelial growth factor-induced gene expression and angiogenesis. *Journal of Biological Chemistry* **283**, 14590–14599 (2008).

302. Chang, H. H., Bong, S. J., Kao, H. Y. & Jin, Z. G. VEGF stimulates HDAC7 phosphorylation and cytoplasmic accumulation modulating matrix metalloproteinase expression and angiogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **28**, 1782–1788 (2008).

303. McKinsey, T. A., Zhang, C. L. & Olson, E. N. MEF2: A calcium-dependent regulator of cell division, differentiation and death. *Trends in Biochemical Sciences* **27**, 40–47 (2002).

304. Zhao, M. *et al.* Regulation of the MEF2 family of transcription factors by p38. *Molecular and Cellular Biology* **19**, 21–30 (1999).

305. Covassin, L. *et al.* Global analysis of hematopoietic and vascular endothelial gene expression by tissue specific microarray profiling in zebrafish. *Developmental Biology* **299**, 551–562 (2006).

306. Quillien, A. *et al.* Robust identification of developmentally active endothelial enhancers in zebrafish using FANS-assisted ATAC-Seq. *Cell Reports* **20**, 709–720 (2017).

307. Brown, J. B. & Celniker, S. E. Lessons from modENCODE. *Annual Review of Genomics and Human Genetics* **16**, 31–53 (2015).

308. Aday, A. W., Zhu, L. J., Lakshmanan, A., Wang, J. & Lawson, N. D. Identification of cis regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. *Developmental Biology* **357**, 450–462 (2011).

309. Bogdanović, O. *et al.* Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. *Genome Research* **22**, 2043–2053 (2012).

310. Hiller, M. *et al.* Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: Application to zebrafish. *Nucleic Acids Research* **41**, e151 (2013).

311. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).

312. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods* **10**, 1213–1218 (2013).

313. Ackermann, A. M., Wang, Z., Schug, J., Naji, A. & Kaestner, K. H. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Molecular Metabolism* **5**, 233–244 (2016).

314. Veldman, M. B. & Lin, S. Etsrp/Etv2 is directly regulated by Foxc1a/b in the zebrafish angioblast. *Circulation Research* **110**, 220–229 (2012).

315. Villefranc, J. A., Amigo, J. & Lawson, N. D. Gateway compatible vectors for analysis of gene function in the zebrafish. *Developmental Dynamics* **236**, 3077–3087 (2007).

316. Rocha, S. F. *et al.* Esm1 modulates endothelial tip cell behavior and vascular permeability by enhancing VEGF bioavailability. *Circulation Research* **115**, 581–590 (2014).

317. Zimmerman, M. W. *et al.* PTP4A3 phosphatase promotes VEGF signaling and enables endothelial cell motility. *The Journal of biological chemistry* **289**, 5904–5913 (2014).

318. Fus, Ł. P. & Górnicka, B. Role of angiogenesis in urothelial bladder carcinoma. *Central European Journal of Urology* **69**, 258–263 (2016).

319. De Smet, F. *et al.* Fibroblast growth factor signaling affects vascular outgrowth and is required for the maintenance of blood vessel integrity. *Chemistry and Biology* **21**, 1310–1317 (2014).

320. Toro, R. *et al.* Identification and functional analysis of endothelial tip cell enriched genes. *Blood* **116**, 4025–4033 (2010).

321. Strasser, G. A., Kaminker, J. S. & Tessier-Lavigne, M. Microarray analysis of retinal endothelial tip cells identifies CXCR4 as a mediator of tip cell morphology and branching. *Blood* **115**, 5102–5110 (2010).

322. Catarino, R. R. & Stark, A. Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. *Genes and Development* **32**, 202–223 (2018).

323. Gopinathan, G. *et al.* Interleukin-6 stimulates defective angiogenesis. *Cancer Research* **75**, 3098–3107 (2015).

324. Chien, M. H. *et al.* Vascular endothelial growth factor-C (VEGF-C) promotes angiogenesis by induction of COX-2 in leukemic cells via the VEGF-R3/JNK/AP-1 pathway. *Carcinogenesis* **30**, 2005–2013 (2009).

325. Meng, S. *et al.* TBX20 regulates angiogenesis through the prokineticin 2–prokineticin receptor 1 pathway. *Circulation* **138**, 913–928 (2018).

326. Shibuya, M. Vascular endothelial growth factor (VEGF) and its receptor (VEGFR) signaling in angiogenesis: A crucial target for anti- and pro-angiogenic therapies. *Genes and Cancer* **2**, 1097–1105 (2011).

327. Andrikopoulos, P., Eccles, S. A. & Yaqoob, M. M. Coupling between the TRPC3 ion channel and the NCX1 transporter contributed to VEGF-induced ERK1/2 activation and angiogenesis in human primary endothelial cells. *Cellular Signalling* **37**, 12–30 (2017).

328. Dean, B. J., Erdogan, B., Gamse, J. T. & Wu, S. Y. Dbx1b defines the dorsal habenular progenitor domain in the zebrafish epithalamus. *Neural Development* **9**, 1–9 (2014).

329. Puttini, S. *et al.* Aldh1a3 is the key isoform that contributes to aldehyde dehydrogenase activity and affects in vitro proliferation in cardiac atrial appendage progenitor cells. *Frontiers in Cardiovascular Medicine* **5**, (2018).

330. Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N. & Izumo, S. The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development* **126**, 1269–1280 (1999).

331. Paffett-Lugassy, N. *et al.* Heart field origin of great vessel precursors relies on nkx2.5-mediated vasculogenesis. *Nature Cell Biology* **15**, 1362–1369 (2013).

332. Cai, C. *et al.* Knockout of tnni1b in zebrafish causes defects in atrioventricular valve development via the inhibition of the myocardial wnt signaling pathway. *The FASEB Journal* **33**, 696–710 (2019).

333. Singleman, C. & Holtzman, N. G. Analysis of postembryonic heart development and maturation in the zebrafish, Danio rerio. *Developmental Dynamics* **241**, 1993–2004 (2012).

334. Miyagi, H. *et al.* Characterization of the zebrafish cx36.7 gene promoter: Its regulation of cardiac-specific expression and skeletal muscle-specific repression. *Gene* **577**, 265–274 (2016).

335. Parish, E. V., Mason, J. O. & Price, D. J. Expression of Barhl2 and its relationship with Pax6 expression in the forebrain of the mouse embryo. *BMC Neuroscience* **17**, 1–16 (2016).

336. Mori, H., Miyazaki, Y., Morita, T., Nitta, H. & Mishina, M. Different spatio-temporal expressions of three otx homeoprotein transcripts during zebrafish embryogenesis. *Molecular Brain Research* **27**, 221–231 (1994).

337. Yoshizawa, A. *et al.* Zebrafish Dmrta2 regulates neurogenesis in the telencephalon. *Genes to Cells* **16**, 1097–1109 (2011).

338. Piragyte, I. *et al.* A metabolic interplay coordinated by HLX regulates myeloid differentiation and AML through partly overlapping pathways. *Nature Communications* **9**, 3090 (2018).

339. Shaut, C. A. E., Keene, D. R., Sorensen, L. K., Li, D. Y. & Stadler, H. S. HOXA13 is essential for placental vascular patterning and labyrinth endothelial specification. *PLoS Genetics* **4**, e1000073 (2008).

340. Mata-Greenwood, E., Jackson, P. N., Pearce, W. J. & Zhang, L. Endothelial glucocorticoid receptor promoter methylation according to dexamethasone sensitivity. *Journal of Molecular Endocrinology* **55**, 133–146 (2015).

341. Toshner, M. *et al.* Transcript Analysis Reveals a Specific HOX Signature Associated with Positional Identity of Human Endothelial Cells. *PLoS ONE* **9**, e91334 (2014).

342. Marcu, R. *et al.* Human Organ-Specific Endothelial Cell Heterogeneity. *iScience* **4**, 20–35 (2018).

343. Sangwung, P. *et al.* KLF2 and KLF4 control endothelial identity and vascular integrity. *JCI Insight* **2**, e91700 (2017).

344. Wick, N. *et al.* Transcriptomal comparison of human dermal lymphatic endothelial cells ex vivo and in vitro. *Physiological genomics* **28**, 179–192 (2007).

345. Thorpe, C. J. Wnt/catenin regulation of the Sp1-related transcription factor sp5l promotes tail development in zebrafish. *Development* **132**, 1763–1772 (2005).

346. Row, R. H., Tsotras, S. R., Goto, H. & Martin, B. L. The zebrafish tailbud contains two independent populations of midline progenitor cells that maintain long-term germ layer plasticity and differentiate in response to local signaling cues. *Development* **143**, 244–254 (2016).

347. Sharma, M. K. *et al.* Differential expression of the duplicated cellular retinoic acid-binding protein 2 genes (crabp2a and crabp2b) during zebrafish embryonic development. *Gene Expression Patterns* **5**, 371–379 (2005).

348. Zhang, J., Lefebvre, J. L., Zhao, S. & Granato, M. Zebrafish unplugged reveals a role for muscle-specific kinase homologs in axonal pathway choice. *Nature Neuroscience* **7**, 1303–1309 (2004).

349. Sieger, D., Tautz, D. & Gajewski, M. The role of Suppressor of Hairless in Notch mediated signalling during zebrafish somitogenesis. *Mechanisms of Development* **120**, 1083–1094 (2003).

350. Thisse, Bernard *et al.* Expression of the zebrafish genome during embryogenesis. *ZFIN Direct Data Submission* (2001).

351. Matsui, T. *et al.* Redundant roles of *Sox17* and *Sox18* in postnatal angiogenesis in mice. *Journal of Cell Science* **119**, 3513–3526 (2006).

352. Corada, M. *et al.* Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nature Communications* **4**, 2609 (2013).

353. Hosking, B. M. *et al.* SOX18 directly interacts with MEF2C in endothelial cells. *Biochemical and Biophysical Research Communications* **287**, 493–500 (2001).

354. Jeong, H.-W. *et al.* Transcriptional regulation of endothelial cell behavior during sprouting angiogenesis. *Nature Communications* **8**, 726 (2017).

355. Wang, S. *et al.* A dynamic and integrated epigenetic program at distal regions orchestrates transcriptional responses to VEGFA. *Genome Research* **29**, 193–207 (2019).

356. Scarlett, K., Pattabiraman, V., Barnett, P., Liu, D. & Anderson, L. M. The proangiogenic effect of iroquois homeobox transcription factor Irx3 in human microvascular endothelial cells. *Journal of Biological Chemistry* **290**, 6303–6315 (2015).

357. Iwatsuki, K. *et al.* Runx1 promotes angiogenesis by downregulation of insulin-like growth factor-binding protein-3. *Oncogene* **24**, 1129–1137 (2005).

358. Sangpairoj, K. *et al.* RUNX1 regulates migration, invasion, and angiogenesis via p38 MAPK pathway in human glioblastoma. *Cellular and Molecular Neurobiology* **37**, 1243–1255 (2017).

359. Li, T. *et al.* Transcription factors regulate GPR91-mediated expression of VEGF in hypoxia-induced retinopathy. *Scientific Reports* **7**, 45807 (2017).

360. Mayr, B. & Montminy, M. Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nature Reviews Molecular Cell Biology* **2**, 599–609 (2001).

361. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics* **39**, 311–318 (2007).

362. Ong, C. T. & Corces, V. G. CTCF: An architectural protein bridging genome topology and function. *Nature Reviews Genetics* **15**, 234–246 (2014).

363. Mou, Y. *et al.* OCT4 remodels the phenotype and promotes angiogenesis of HUVECs by changing the gene expression profile. *International Journal of Medical Sciences* **13**, 386–394 (2016).

364. Zhou, P. *et al.* Mapping cell type-specific transcriptional enhancers using high affinity, lineage-specific Ep300 bioChIP-seq. *eLife* **6**, e22039 (2017).

365. Williamson, I., Lettice, L. A., Hill, R. E. & Bickmore, W. A. Shh and Zrs enhancer colocalisation is specific to the zone of polarising activity. *Development* **143**, 2994–3001 (2016).

366. Piper, J. *et al.* Wellington: A novel method for the accurate identification of digital genomic footprints from DNase-seq data. *Nucleic Acids Research* **41**, e201–e201 (2013).

367. Nueda, M. J., Tarazona, S. & Conesa, A. Next maSigPro: Updating maSigPro bioconductor package for RNA-seq time series. *Bioinformatics* **30**, 2598–2602 (2014).

368. Starruß, J., De Back, W., Brusch, L. & Deutsch, A. Morpheus: A user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics* **30**, 1331–1332 (2014).

369. Mohamed, A. A. *et al.* Ets family protein, erg expression in developing and adult mouse tissues by a highly specific monoclonal antibody. *J Cancer* **1**, 197–208 (2010).

370. Eichmann, A. *et al.* The expression pattern of the mafB/kr gene in birds and mice reveals that the kreisler phenotype does not represent a null mutant. *Mech. Dev.* **65**, 111–122 (1997).