Molecular Regulation and Function of Gata2 in the Programming of Haemogenic Endothelium



Tomasz Dobrzycki

Green Templeton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2017

For my niece Julia

Acknowledgements

Working in academia may feel lonely at times. We do collaborate, but in the end, each project belongs to one person. This thesis comes out of a project that was, in theory, entirely my own. But no man is an island and I must thank all those who contributed to the production of this thesis and made those four years a genuinely pleasant and rewarding experience.

First, I want to thank Roger. For offering me lab space and a zebrafish-based project, for sharing advice and impressive amount of knowledge whenever I needed them, and for critical comments that kept me engaged and motivated throughout. For helping me improve my English and making sure I come out of this degree as an all-rounded scientist. And… for assigning me to Rui.

Rui, I cannot thank you enough for all you have done for me. From the beginning, you have made me really enthusiastic about my project and somehow managed to maintain this enthusiasm all the way until the end! I gladly recall those times when we would go for tea and coffee with big smiles, satisfied with the work we had put in. I loved how supportive you were of my activities outside of the lab and how you always made sure I maintained a healthy and productive balance. You have been more than just a project supervisor, you have been an excellent mentor and a friend. I nominated you for the McMichael medal without any doubt. I am sure you will be formally recognised for your excellent mentoring one day.

Working in the Patient lab has been a very pleasant experience and I thank Aldo, Arif, Jana, Lei, Mukesh, Phil and Tessa for providing stimulating discussions during lab meetings, but also silly chats in the canteen and loads of fun during lab outings. Linda, thank you for keeping us all organised and especially for arranging my viva. I was also really lucky to share my time with other DPhil students. Caroline and Michael, although I never really understood your projects, I am glad we could take that journey together. Kira, those months you spent in our lab were filled with your positive energy. I am very happy that we managed to nail those ATAC-seq protocols together. Florian… when you meet someone already at the interviews and end up in the same lab, working on very similar projects, you are bound to get on. Thanks for the chats (scientific, personal or simply ridiculous), Feel-Good Fridays with Radio Heart in the microscope room and appreciating my sense of humour. Remember our alternative career plan!

I must especially thank those who directly contributed to the work presented in this thesis. Rossella, I do not know how we would survive without you. Thank you for your patience when I flooded you with sequencing data. Monika, your ability to find simple solutions to little, annoying problems is impressive. Thank you for the beautiful images, hundreds of *in situ* hybridized embryos and pushing me to write our paper. And as crazy as it was for me to 'supervise' someone who had already done a PhD, I learnt a lot from you and our work together.

I must also acknowledge all the people whose work made my project possible. All the people from the fish facility, who made sure my thousands of fish always found their way to my racks, happy and healthy. Kevin from the FACS facility, who not only did all the sorting, but also provided entertainment throughout. Neil from the Single Cell Facility and Esther from the Imaging Suite, for teaching me how to use really advanced equipment, and Emmanouela from the CBRG, who analysed my data when time was very precious.

I thank all the faculty members who contributed to my work and time at Oxford. Liz and Rob, thank you for forming a fantastic Thesis Advisory Committee, my project would have

been very different without you. Sarah, thank you for giving me the last push for the final months during the Confirmation. Leanne, thank you for your ongoing support that I always enjoyed during our termly dinners. I would especially like to thank Catherine, for agreeing to examine my Transfer, my Confirmation and, finally, my DPhil viva! I also thank Shane for coming all the way from Manchester to examine me. I hope this thesis is a good read!

Thanks to Bryony and Emma for helping me explore all the science communication and public engagement options. You made me rediscover what was really exciting about my work!

I thank all the past members of the lab and temporary visitors, who I shared the time with. A special mention goes to Nicola, whose sense of humour and loveliness turned sitting in the distant lab corner or through individual meetings with Rui into very happy memories.

Now, out of the Patient lab… being in Oxford gave me a chance to be surrounded by a fantastic community: in the WIMM, at GTC (including the Overtones, the GCRC and GTBC), within the CDB programme, in the Innovation Forum, in the OUTriC, or at various events. There were so many people who made my experience so much more meaningful that listing them all would exceed my word limit of 50,000 words. But I want to mention those who have been much more than mates. Andra, it was great to go through the undergraduate and postgraduate journey together, side by side. I truly appreciate your mentorship, which helped me get the perfect job. But even more, I value our friendship. I am sure we will continue to support each other through our future endeavours outside of academia. Ben, thanks for sharing the same exact Oxford journey, I knew we would be bros already when we first met at the interviews, almost 5 years ago! Thanks for being a great coursemate, housemate, conference mate, and a friend. Andrzej, thanks for welcoming me to Oxford, for the evenings of recalling and dissing high school times, for putting us on the front cover of a newspaper and for joining forces in the last weeks of our DPhil. I think we made each other's theses better! Clarissa, thank you for giving me the chance to share the last months of my degree with you. For the deep conversations, trouble-free distractions from writing and for making me rediscover my artistic soul. And finally… Juan. My singing partner and co-founder of Frameshift, my brother from another mother, my 'second wife' and a true friend who I was lucky to have around for two years. I know the both of us can make this world a better place!

Last but not least, I would like to thank my family. My parents, for whom the distance between Poland and the UK was never too long. Thank you for supporting me in all my DPhil-related adventures. My grandmothers, who probably never really understood what I did, but they definitely made me feel that I was doing something very worthwhile! And everyone else: aunts, uncles, cousins – it was always a pleasure to share the DPhil experience with you.

I want to give special thanks to my wife Karolina, for being there for me through weakness and strength, happiness and sorrow, despite the distance between us. You made me realise that there was so much more to life than work and I will always appreciate that. Thank you for making me get out of Oxford to explore the world and cherish the time away. Thank you for making me chase my dreams. Thank you for making me a better person.

And finally… I thank my dear little sister, Ania. I would not be the same person without you by my side. Thank you for teaching me what is important in life. You and Julia are the living proof that even when life is not perfect, it is possible to beat all the odds with a smile. In Poland, we say that the family is one thing one cannot choose. Well, I am glad I did not have to choose, because I would not have made a better choice even if I wanted to.

Research declaration

The work presented in this thesis is my own unless indicated. All text, figures, tables, data or results which are not wholly my own work are indicated and the sources acknowledged.

Chapter 3

ATAC-seq reads were quality checked, mapped to the genome and further analysed by Rossella Rispoli at the MRC MHU within the MRC WIMM, Oxford. The transcription factor motifs in the putative *gata2a* enhancer were annotated by Sarah de Val from Ludwig Cancer Research, Oxford. Confocal images and their analysis were provided by Monika Krecsmarik, MRC MHU, MRC WIMM, Oxford.

Chapter 4

In situ hybridization were performed and their quantitation was designed in collaboration with Monika Krecsmarik, MRC MHU, MRC WIMM, Oxford.

Chapter 5

RNA-seq reads were quality checked, mapped to the genome and analysed using differential expression gene analysis by Rossella Rispoli at the MRC MHU within the MRC WIMM, Oxford. The statistical analysis of Fluidigm Biomark™ data was performed by Emmanouela Repapi, Computational Biology Research Group, MRC WIMM.

For the work in all the chapters, FACS was performed by Kevin Clark at the MRC WIMM, Oxford.

This project was funded by grants from Wellcome Trust, British Heart Foundation and Medical Research Council.

Work presented in chapters 3 and 4 is currently under preparation for publication. Additional work presented in Appendix IV has been submitted for publication.

Word count: 46,303

Abbreviations

Ac-LDL	Acetylated Low Density Lipoprotein
AGM	Aorta-Gonad-Mesonephros
ALM	Anterior Lateral Plate Mesoderm
AML	Acute Myeloid Leukaemia
ANOVA	Analysis of Variance
AP	Alkaline Phosphatase
ATAC	Assay for Transposase-Accessible Chromatin
BIO	6-bromoindirubin-3'-oxime
BMP	Bone Morphogenetic Protein
bp	base pair
CD	Cluster of Differentiation
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
CHT	Caudal Haematopoietic Tissue
CNS	Central Nervous System
Co-IP	Co-Immunoprecipitation
CPM	Counts Per Million
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Ct	Cycle threshold
d.f.	degrees of freedom
DA	Dorsal Aorta
DAVID	Database for Annotation, Visualization and Integrated Discovery
DE	Differential Expression
DH	Definitive Haemangioblast
DIG	Digoxygenin
Dkk	Dickkopf
Dll	Delta-like
DLP	Dorsal Lateral Plate
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DNAse	Deoxyribonuclease
dpf	days post fertilisation
dUTP	Deoxyuridine Triphosphate
Dvl	Dishevelled
E	Embryonic day (mouse development)
EB	Elution Buffer
EDTA	Ethylenediaminetetraacetic acid
EHT	Endothelial-to-Haematopoietic Transition
EMP	Erythromyeloid Progenitor
ESC	Embryonic Stem Cell
FACS	Fluorescence-Activated Cell Sorting
FAIRE	Formaldehyde-Assisted Isolation of Regulatory Elements
FC	Fold Change
Fgf	Fibroblast growth factor
FSC	Forward Scatter
Fzd	Frizzled
Gcsf	Granulocyte colony stimulating factor
GFP	Green Fluorescent Protein
GO	Gene Ontology
GPR	G Protein-coupled Receptor
GRN	Gene Regulatory Network
HBSS	Hank's Balanced Salt Solution
HE	Haemogenic Endothelium

HEPES	Hydroxyethylpiperazineethanesulfonic acid
Hes	Hairy/Enhancer of split
Hh	Hedgehog
HH	Hamburger-Hamilton stage
HotSHOT	Hot Sodium Hydroxide and Tris
HPC	Haematopoietic Progenitor Cell
hpf	hours post fertilisation
HSC	Haematopoietic Stem Cell
HSPC	Haematopoietic Stem and Progenitor Cell
i4	intron 4
ICM	Intermediate Cell Mass
IFN	Interferon
iPSC	induced Pluripotent Stem Cell
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISH	*In Situ* Hybridization
ISV	Intersomitic Vessel
IWR-1	Inhibitor of Wnt Response 1
Jag	Jagged
Jam	Junction adhesion molecule
kb	kilobase
LB	Luria-Bertani broth
Ldb	Lim domain binding
Lmo	Lim domain only
LPM	Lateral Plate Mesoderm
MAB	Maleic Acid Buffer
MEL	Mouse Erythroleukaemia
MNase	Micrococcal Nuclease
MO	Morpholino Oligonucleotide
mRNA	messenger RNA
MS222	Tricaine methanesulphonate
n.s.	not significant
nc	notochord
NEB	New England Biolabs
NF	Nuclear Factor
NT	Neural Tube
NTP	Nucleoside Triphosphate
OE	Overexpressed
PBI	Posterior Blood Island
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline with Tween
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PCV	Posterior Cardinal Vein
PFA	Paraformaldehyde
PLM	Posterior Lateral Plate Mesoderm
PTU	Phenylthiourea
Q-Q	Quantile-Quantile
qRT-PCR	quantitative Reverse Transcriptase Real-Time PCR
RBI	Rostral Blood Island
RBP	Recombining Binding Protein
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
ROI	Region of Interest
RPKM	Reads Per Kilobase per Million
RT	Room Temperature
RVD	Repeat Variable Diresidue

SD	Standard Deviation
-seq	sequencing
sgRNA	single guide RNA
SOC	Super Optimal broth with Catabolite repression
ss	-somite stage
SSC	Saline-Sodium Citrate
TAE	Tris-acetate-EDTA
TALENs	Transcription Activator-Like Effector Nucleases
Tcf	T-cell factor
Tcf	T-cell Factor
TE	Tris-EDTA
TF	Transcription Factor
Tgf	Transforming growth factor
TNF	Tumour Necrosis Factor
UCSC	University of California, Santa Cruz
UTR	Untranslated Region
VBI	Ventral Blood Island
VE	Vascular Endothelium
Vegf	Vascular endothelial growth factor
wt	wild type
ZFIN	Zebrafish Model Organism Database
ZFNs	Zinc-Finger Nucleases
Zv9	Zebrafish genome version 9

Table of Contents

Acknowledgements iii
Research declaration v
Abbreviations vi
Table of Contents ix
List of Figures and Tables xiv
Abstract xvi
1. Introduction 1
1.1. An overview of embryonic haematopoiesis 2
1.1.1. The haemangioblast 2
1.1.2. Primitive and pro-definitive waves of haematopoiesis 5
1.1.3. The haemogenic endothelium 8
1.1.4. Haematopoietic stem cells 11
1.2. Molecular specification of the HE and HSCs 16
1.2.1. Transcriptional regulatory network of HE and HSCs 16
1.2.1.1. *Runx1* 17
1.2.1.2. *Tal1/Scl* 19
1.2.1.3. *Etv2* 20
1.2.1.4. *Lmo2* 21
1.2.1.5. *Erg* 21
1.2.1.6. *Sox7* and *Sox17* 22
1.2.1.7. *Cmyb/myb* 23
1.2.1.8. *Gata2* 23
1.2.2. Signalling pathways specifying the HE and HSCs 26
1.2.2.1. Notch 26
1.2.2.2. Vegf 32
1.2.2.3. Tgfβ 34
1.2.2.4. Wnt 35
1.3. Zebrafish as a model system 38
1.3.1. Development of the zebrafish embryo 40
1.3.2. Definitive haematopoiesis in the zebrafish 41
1.3.3. *Gata2* paralogues in the zebrafish 46
1.3.3.1. *Gata2a* 46
1.3.3.2. *Gata2b* 48
1.4. Objectives 48
2. Materials and Methods 51
2.1. Solutions, oligonucleotides and molecular reagents 52

2.1.1. Solutions 52
2.1.2. Morpholino oligonucleotides (MOs) 53
2.1.3. Primers 53
2.1.4. CRISPR/Cas9 sgRNAs 54
2.1.5. Cas9(D10A) mRNA 54
2.1.6. *Tol2* mRNA 55
2.1.7. TaqMan® assays 55
2.1.8. In situ hybridization probes 57
2.2. Maintenance and manipulation of zebrafish 58
2.2.1. Maintenance of zebrafish 58
2.2.2. Collection and staging of zebrafish embryos 58
2.2.3. Zebrafish lines 59
2.2.4. Generation of new zebrafish lines 59
2.2.4.1. *Gata2a* i4 enhancer transgenic lines 59
2.2.4.2. *Gata2a* i4 enhancer mutant lines 59
2.2.4.3. *Lmo4a* mutant lines 62
2.2.4.4. Maintenance of new mutant lines 62
2.2.5. Microinjection of embryos 62
2.2.6. Chemical treatment of embryos 63
2.2.7. Dechorionation and fixation of embryos 63
2.2.8. Fluorescence-activated cell sorting (FACS) 63
2.3. Preparation, manipulation and analysis of DNA 64
2.3.1. Transformation of competent *E. coli* 64
2.3.2. *E. coli* culture 64
2.3.3. Isolation of zebrafish genomic DNA 64
2.3.4. Isolation and sequencing of plasmid DNA 65
2.3.5. Polymerase chain reaction (PCR) and colony PCR 65
2.3.6. Agarose gel electrophoresis and gel purification 65
2.3.7. Restriction digestion, ligation and In-Fusion® 66
2.3.8. Site-directed mutagenesis 66
2.3.9. Construction of new plasmids 66
2.3.9.1. *Gata2a*-i4-GFP-Tol2 vectors 66
2.3.9.2. *Gata2b* tissue-specific overexpression vectors 67
2.3.9.3. *Lmo4a* TALENs 67
2.3.9.4. DNA templates for new in situ hybridization probes 67
2.3.10. ATAC-seq library preparation 68
2.4. Preparation, manipulation and analysis of RNA 68
2.4.1. RNA extraction 68

2.4.2. RNA-seq sample preparation and sequencing 69
2.4.3. Reverse transcription of RNA 69
2.4.4. Real-time quantitative PCR (RT-qPCR) 70
2.4.4.1. SYBR® Green qRT-PCR 70
2.4.4.2. Fluidigm Biomark™ qRT-PCR 70
2.4.5. Whole mount in situ hybridization (ISH) and immunohistochemistry 71
2.5. Imaging zebrafish embryos 71
2.5.1. Bright-field microscopy 71
2.5.2. Confocal microscopy 72
2.5.3. Live imaging 72
2.6. Bioinformatic and statistical tools 72
2.6.1. Genomic data 72
2.6.2. Quality check and mapping of high-throughput sequencing data 73
2.6.3. ATAC-seq differential peak analysis 73
2.6.4. RNA-seq differential gene expression analysis 73
2.6.5. In silico DNA analysis 74
2.6.6. In situ hybridization image quantification 74
2.6.7. Statistical analysis 74
2.6.7.1. In situ hybridization analysis 74
2.6.7.2. qRT-PCR analysis 76
2.6.8. Preparation of figures 77
3. Identification and characterisation of a conserved endothelial enhancer of *gata2a* 78
3.1. Introduction 79
3.2. A highly conserved region within the zebrafish gata2a locus is accessible to transposase specifically in the endothelium 84
3.3. The *gata2a*-i4 enhancer contains conserved putative transcription factor binding sites 86
3.4. The *gata2a*-i4 enhancer drives GFP expression in haematopoietic precursors 89
3.4.1. The *gata2a*-i4:GFP transgenic zebrafish embryos are characterised by GFP expression in the PLM and in the dorsal aorta 89
3.4.2. The *gata2a*-i4-GFP$^+$ cells include *gata2a*$^+$ and *runx1*$^+$ haematopoietic precursors 90
3.4.3. The *gata2a*-i4-GFP$^+$ cells express endothelial and haematopoietic markers 94
3.5. Activity of the *gata2a*-i4 enhancer increases upon *gpr65* knockdown 95
3.6. Discussion 98
3.6.1. Overview 98
3.6.2. Genomic locus of *gata2a* but not *gata2b* contains a highly conserved intronic enhancer previously characterized in mice 99
3.6.3. The *gata2a*-i4 enhancer drives GFP expression in haematopoietic cell populations 101
3.6.4. The *gata2a*-i4:GFP transgenic line provides a new tool to assay the activity of the i4 enhancer in vivo 103

4. The role of *gata2a* in the programming of the haemogenic endothelium revealed by genomic deletion of its endothelial enhancer 105

4.1. Introduction 106

4.2. Generation and detection of CRISPR/Cas9-induced deletions of *gata2a*-i4 enhancer 108

4.3. Deletion of the *gata2a*-i4 enhancer results in decreased expression of *gata2a* in the embryonic endothelium 111

4.4. Loss of the *gata2a*-i4 enhancer does not affect endothelial development 112

4.5. *Gata2a*$^{\Delta i4/\Delta i4}$ mutants exhibit defects in the expression of haematopoietic genes 113

4.5.1. Deletion of the *gata2a*-i4 enhancer results in reduced expression of *runx1* at early stages of HE specification 113

4.5.2. Deletion of the *gata2a*-i4 enhancer results in reduced expression of *gata2b* at early stages of HE specification 118

4.5.3. Deletion of the *gata2a*-i4 enhancer does not result in complete loss of HE specification 121

4.6. Lack of endothelial *gata2a* does not affect the activity of the *gata2a*-i4 enhancer, while overexpression of *gata2b* increases the activity of the enhancer 122

4.7. *Gata2a*$^{\Delta i4/\Delta i4}$ mutants recover from the early haematopoietic phenotype 124

4.7.1. Expression of late markers of embryonic HSC activity is unaffected in *gata2a*$^{\Delta i4/\Delta i4}$ mutants 124

4.7.2. The haematopoietic recovery of *gata2a*$^{\Delta i4/\Delta i4}$ mutants is likely mediated by *gata2b* and may involve Wnt signalling pathway 126

4.7.3. Adult *gata2a*$^{\Delta i4/\Delta i4}$ mutants suffer from infections and heart oedemas 128

4.8. Discussion 131

4.8.1. Overview 131

4.8.2. A conserved intronic enhancer is required for the endothelial expression of *gata2a* but its deletion does not affect endothelial development 133

4.8.3. Functional analysis of the *gata2a*-i4 reveals non-redundant roles for Gata2a and Gata2b in HE programming 134

4.8.4. Recovery of definitive haematopoiesis in *gata2a*$^{\Delta i4/\Delta i4}$ mutants likely results from the activity of *gata2b* 137

5. The role of *lmo4a* in the specification of HSCs 141

5.1. Introduction 142

5.2. Characterisation of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cell population for the transcriptomic profiling of endothelium at the stage of HE specification 145

5.3. RNA-seq in *lmo4a* morphants reveals changes in expression of known haemogenic endothelial genes 150

5.4. Expression of *gata2a* is regulated by Lmo4a 151

5.4.1. *Gata2a* mRNA levels are elevated in *lmo4a* morphants 151

5.4.2. *Lmo4a* mutants exhibit increased levels of *gata2a* expression 152

5.4.3. The activity of the *gata2a*-i4 enhancer increases upon *lmo4a* knockdown 154

5.5. Knockdown of *lmo4a* affects the expression of several genes involved in HE and HSC programming 155

5.5.1. Expression of *gata2b* mRNA is not majorly affected in *lmo4a* morphants 155

5.5.2. Selected early HE markers are upregulated upon *lmo4a* knockdown 156
5.5.3. *Lmo4a* morphants express reduced levels of *gfi1aa* and *cmyb* 160
5.6. Specification of the main embryonic vessels is not affected in the *lmo4a* morphants 162
5.7. Wnt signalling likely acts downstream of *lmo4a* and upstream of *gata2a* 165
5.7.1. Expression of Wnt signalling genes is increased in *lmo4a* morphants 165
5.7.2. Inhibition of Wnt signalling reduces *gata2a* expression in *lmo4a* morphants 166
5.7.3. Chemical modulation of Wnt signalling affects the activity of the *gata2a*-i4 enhancer 167
5.8. Loss of *lmo4a* function – differences between morphants and mutants 168
5.8.1. Expression of *runx1* is reduced in *lmo4a* morphants 168
5.8.2. Expression of *runx1* increases in *lmo4a* mutants 172
5.9. Discussion 173
5.9.1. Overview 173
5.9.2. Lmo4a is a negative regulator of *gata2a* expression in the endothelium 173
5.9.3. Wnt signalling is a potential mediator of the *lmo4a-gata2a* regulation 175
5.9.4. A genetic network responsible for precise regulation of *gata2a* expression is required for HSC specification 177
5.9.5. Compensatory mechanisms potentially rescue *runx1* expression in *lmo4a* mutants 179
Conclusions 181
References 186
Appendix I: Supplementary Figures 208
Appendix II: Differentially expressed genes identified in the *lmo4a* RNA-seq experiment (related to Chapter 5.3). 212
Appendix III: Supplementary Table 5.1 247
Appendix IV: Manuscript submitted for peer-reviewed publication describing the method to quantitatively analyse in situ hybridization in zebrafish embryos. 267
Appendix V: Haematopoiesis rap 293

List of Figures and Tables

Figure 1.1. The haemangioblast, localised in close association to pronephric ducts across vertebrates, represents a bipotent state of competence to generate blood and endothelium. 3
Figure 1.2. Blood islands in the yolk sac of a mouse embryo consist of closely associated erythroid and endothelial progenitors. 6
Figure 1.3. The linear model of HSC specification from vascular endothelium involves a haemogenic endothelial intermediate. 8
Figure 1.4. Generation of patient-specific HSCs *in vitro* holds great clinical potential. 12
Figure 1.5. The first definitive HSCs emerge in intra-aortic clusters from the haemogenic endothelium in the floor of the dorsal aorta. 13
Figure 1.6. Selected stages of zebrafish embryonic development. 39
Figure 1.7. Definitive haematopoiesis in the zebrafish originates from the PLM and progresses through the CHT to kidney. 45
Figure 1.8. Both paralogues of zebrafish *gata2* are expressed in the DA endothelium prior to the emergence of HSCs. 47

Figure 2.1. The pipeline to generate zebrafish transgenic lines. 60
Figure 2.2. The pipeline to generate zebrafish mutant lines. 61
Figure 2.3. An overview of the method to quantify *in situ* hybridization (ISH) signal. 75

Figure 3.1. Analysis of *gata2b* locus with ATAC-seq does not reveal endothelial enhancers. 85
Figure 3.2. Identification of a conserved endothelial enhancer of zebrafish *gata2a*. 88
Figure 3.3. *Gata2a*-i4 enhancer drives GFP expression in PLM, DA and endothelial cells. 91
Figure 3.4. *Gata2a*-expressing endothelial cells in DA are GFP-positive in *gata2a*-i4:GFP transgenic embryos. 92
Figure 3.5. GFP-positive cells in *gata2a*-i4:GFP transgenic fish include *runx1*$^{+}$ cells. 93
Figure 3.6. GFP^{+} cells in Tg(*gata2a*-i4-1.1kb:GFP) and Tg(*gata2a*-i4-450bp:GFP) transgenic fish express similar levels of haematopoietic and endothelial genes compared to GFP^{-} cells. 96
Figure 3.7. Expression of *gfp* mRNA in Tg(*gata2a*-i4-450bp:GFP) transgenic fish depends on the expression of *gpr65*. 97

Figure 4.1. Generation of *gata2a*$^{\Delta i4}$ zebrafish mutants with CRISPR/Cas9. 109
Figure 4.2. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type, express reduced levels of *gata2a* mRNA specifically in the endothelium. 110
Figure 4.3. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants exhibit no defects in endothelial development. 115
Figure 4.4. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type embryos, express lower levels of *runx1* in the dorsal aorta during the specification of HE. 117
Figure 4.5. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type embryos, express lower levels of *gata2b* in the dorsal aorta during the specification of HE. 119
Figure 4.6. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants do not express significantly reduced levels of *gfi1aa*, *etv2* or *tal1*, compared to wild type embryos. 120
Figure 4.7. Expression of *gata2a*-i4:*gfp* is not affected by the loss of endothelial *gata2a*, but increases upon ectopic overexpression of *gata2b*. 124
Figure 4.8. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants recover the expression of markers of definitive haematopoiesis. 126
Figure 4.9. Endothelial overexpression of *gata2b* rescues the early *runx1* phenotype of *gata2a*$^{\Delta i4/\Delta i4}$ mutants. 128

Figure 4.10. Endothelial expression of Wnt signalling genes recovers in *gata2a*$^{\Delta i4/\Delta i4}$ mutants. 129
Figure 4.11. Endothelial expression of selected Notch and Tgfβ signalling genes is not affected in *gata2a*$^{\Delta i4/\Delta i4}$ mutants. 130
Figure 4.12. *Gata2a*$^{\Delta i4/\Delta i4}$ adult fish suffer from infections and heart oedemas. 131
Figure 4.13. Zebrafish *gata2a* regulates the early HE programme through *gata2b*. 138

Figure 5.1. Genetic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells reveals their endothelial identity. ... 147
Figure 5.2. Transcriptomic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells with RNA-seq reveals multiple differentially expressed genes upon *lmo4a* knockdown. 148
Figure 5.3. Transcriptomic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells with RNA-seq confirms their endothelial identity and reveals differentially expressed genes upon *lmo4a* knockdown. 149
Figure 5.4. Expression of *gata2a* increases in the endothelium of *lmo4a* morphants. 153
Figure 5.5. Expression of *gata2a* mRNA increases in the endothelium of *lmo4a* morphants and mutants after 26hpf, which is also associated with increased *gata2a*-i4 enhancer activity. 154
Figure 5.6. Expression of *gata2b* mRNA is not affected by *lmo4a* knockdown. 158
Figure 5.7. Expression of selected genes required for HE specification increases upon *lmo4a* knockdown. 159
Figure 5.8. Expression of *gfi1aa* and *cmyb* is reduced in *lmo4a* morphants. 161
Figure 5.9. Specification of the dorsal aorta endothelium is not affected in *lmo4a* morphants. 163
Figure 5.10. Gene Ontology term analysis suggests a role of *lmo4a* in regulating Wnt signalling. 164
Figure 5.11. Expression of few Wnt signalling genes is increased in *lmo4a* morphants. 167
Figure 5.12. Chemical inhibition of Wnt signalling rescues *gata2a* overexpression in *lmo4a* morphants, potentially involving the i4 enhancer. 169
Figure 5.13. Expression of *runx1* and *rag1* is reduced upon *lmo4a* knockdown. 170
Figure 5.14. Expression of *runx1* is increased in the dorsal aorta of *lmo4a* maternal homozygous mutants. 171

Supplementary Fig. 4.1. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type embryos, express lower levels of *dnmt3bb.1* in the dorsal aorta during the specification of HE. 209
Supplementary Fig. 5.1. Expression of *gata2a* increases in the endothelium of *lmo4a* morphants. .. 210
Supplementary Fig. 5.2. Generation of *lmo4a*$^{\Delta 25\text{-}29}$ zebrafish mutants with TALENs. 211

Table 2.1. List of solutions used in this thesis. 52
Table 2.2. Details of morpholino oligonucleotides used in this thesis. 53
Table 2.3. Sequences of PCR primers used in this thesis. 53
Table 2.4. Sequences of oligonucleotides used to synthesize sgRNA templates. 54
Table 2.5. Catalogue numbers of TaqMan® assays used in this thesis. 55
Table 2.6. *In situ* hybridization probes used in this thesis. 58

Supplementary Table 5.1. Expression levels in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells of selected genes identified as differentially expressed upon *lmo4a* morphants in the RNA-seq experiment, measured with Fluidigm Biomark™ qRT-PCR at 22hpf and 28hpf. 248

Abstract

Haematopoietic stem cells (HSCs) maintain the vertebrate blood system throughout life. Exploiting their clinical potential requires a thorough understanding of the natural origins of the HSCs. They first arise from the haemogenic endothelium (HE), located in the main embryonic artery, the dorsal aorta. Our understanding of the genetic mechanisms underlying HE specification remains incomplete, but one of the crucial transcription factors is Gata2.

We found that a conserved enhancer of zebrafish *gata2a* gene (i4 enhancer) is active *in vivo* specifically in endothelial cells, including the HE. To unravel the function of *gata2a* in specifying the HSCs, we have targeted the i4 enhancer with CRISPR/Cas9, generating the first reported genomic deletion of an endogenous *cis*-regulatory region in zebrafish. Deletion of the i4 enhancer leads to a decrease in endothelial *gata2a* expression and a concomitant transient decrease in the number of HSCs. This is marked by an early decrease in the expression of *gata2b*, a *gata2a* paralogue previously shown to be required for the initiation of the haematopoietic programme. Our results suggest non-redundant roles of both zebrafish *gata2* paralogues in programming of HSCs, providing insights into different roles of GATA2 throughout the programming of HSCs.

We also confirmed the previously reported loss of HSCs upon MO-mediated knockdown of *lmo4a*, associated with increased *gata2a* expression in HE. We validated the increase in *gata2a* levels in TALEN-generated *lmo4a* mutants. To identify the links between *lmo4a*, *gata2a* and the HE programming, we have profiled the transcriptome of *lmo4a*-deficient endothelial cells, including the HE. Our results suggest that Lmo4a may be a global regulator of the transcriptional programming of the HE. Moreover, Wnt signalling pathway may regulate *gata2a* downstream of *lmo4a*. This provides novel insights into the gene regulatory network orchestrating the generation of HSCs in the embryo.

Chapter 1

Introduction

1.1. An overview of embryonic haematopoiesis

Blood is one of the most heterogeneous tissues, considering both its functions and cellular content. Multiple cell types within the blood fulfil crucial body functions and disruptions of these processes frequently result in human diseases. Appearance of the circulatory system during evolution was likely critical in the generation of complex structures and multi-organ arrangement of vertebrate bodies, because it allowed independence from diffusion from the external environment. Blood is one of the first tissues specified during embryonic development and developing vertebrate embryos very early become dependent on blood circulation. The development of blood cells, termed haematopoiesis, has been extensively studied for many decades, using a variety of *in vivo* and *in vitro* models. Each model provides different advantages and recent opinions stress the importance of bringing this knowledge together to fully understand this complex developmental process (Ciau-Uitz and Patient 2016).

1.1.1. The haemangioblast

The concept linking blood and endothelial fates dates back to anatomical and histological studies from the late 19th and early 20th century (Gritz and Hirschi 2016). Many decades later, the idea of a 'haemangioblast' regained much attention after single mesodermal progenitors with both haemogenic and endothelial potential were isolated from embryonic stem cell (ESC) cultures (Choi et al. 1998). The term 'haemangioblast' was coined to describe a hypothetical progenitor cell with a dual potential, encompassing blood and vasculature.

The existence of the haemangioblast is supported by shared expression of genes linked to endothelial and blood cells, including mesodermal markers Brachyury, BMP4 and Flk1, as well as VE-cadherin (VE-cad), CD31 and CD34 (Gritz and Hirschi 2016). Haemangioblasts are believed to be specified within the mesoderm, as common mesodermal origins of blood and endothelium are found across vertebrate species (Nishikawa et al. 1998; Huber et al. 2004; Ciau-Uitz et al. 2010a).

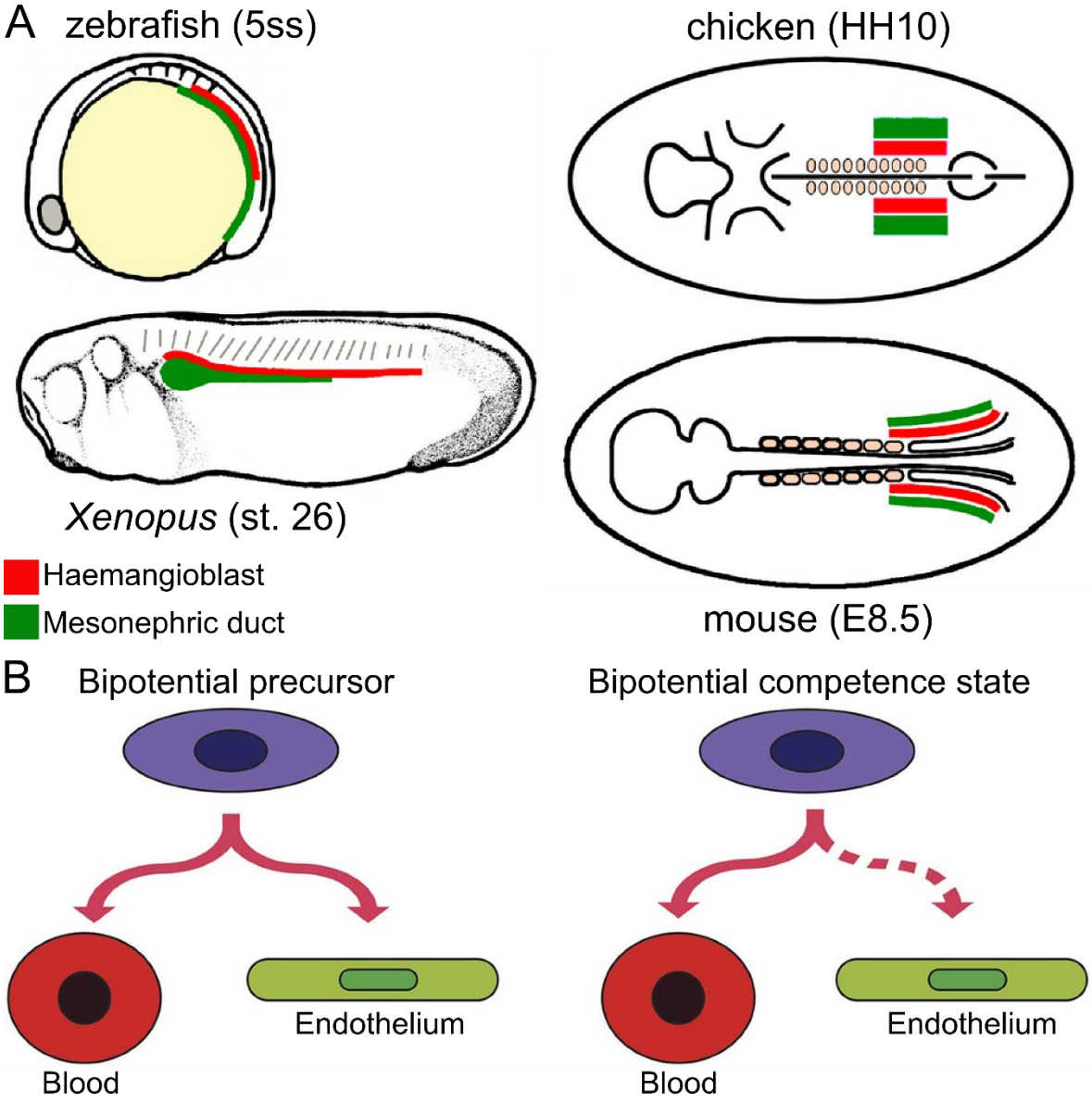


Figure 1.1. The haemangioblast, localised in close association to pronephric ducts across vertebrates, represents a bipotent state of competence to generate blood and endothelium.

(A) Populations of haemangioblasts (red) have been proposed to exist across vertebrate species, including the embryos of zebrafish at 5-somite stage (5ss), *Xenopus* at embryonic stage 26, chicken at Hamburger-Hamilton stage 10 (HH10) and mouse at embryonic day 8.5 (E8.5). They are believed to develop in close proximity to mesonephric/pronephric ducts (green). Images not to scale. (B) The haemangioblast, originally proposed as a bipotential precursor with the ability to differentiate into blood or endothelial cells, is more likely a state of competence. While these cells may have the theoretical capacity to generate both lineages, cells arising from the common haemangioblastic precursors in the *in vivo* situations are almost exclusively blood, while endothelial cells frequently originate from other mesodermal precursors. (Reproduced from Amaya 2013; Ciau-Uitz and Patient 2016)

In *in vivo* studies, a haemangioblast population, giving rise to endothelium and primitive erythrocytes, was described in the primitive streak of the gastrulating mouse embryo (Huber et al. 2004). Similarly, a Gata1- and VE-cad-expressing population isolated at mouse embryonic day 7.5 (E7.5) gave rise to erythroid and endothelial cells in culture (Yokomizo et al. 2007). A haemangioblast population has been described in the zebrafish, where common progenitors for endothelial cells and erythrocytes were identified in the ventral mesoderm of a 6hpf embryo through lineage tracing experiments (Vogeli et al. 2006). Further support for their existence was based on the analysis of *cloche* mutants, in which mutation of *npas4l* resulted in the loss of both blood and endothelial lineages (Stainier et al. 1995; Liao et al. 1998; Reischauer et al. 2016). The first evidence for a human haemangioblast came from human ESC *in vitro* culture, where progenitors of dual erythro-endothelial potential were described (Kennedy et al. 2007). Haemangioblasts are also believed to exist in *Drosophila* (Mandal et al. 2004).

In *Xenopus laevis*, the intraembryonic haemangioblast population is localised between the somites and the mesonephric duct (**Fig. 1.1A**) (Walmsley et al. 2002; Ciau-Uitz et al. 2013a). This is also true in the zebrafish (Gering et al. 1998) and has been suggested to be a conserved morphological arrangement across vertebrates (Ciau-Uitz and Patient 2016).

The controversies regarding the existence of the haemangioblast arose from studying the output of common blood-endothelial precursor populations. Collective evidence from mouse embryos suggests that while the haemangioblast population gives rise to most erythroid cells, only few endothelial cells originate from the haemangioblast (Palis 2008). Endothelial cells may arise from anterior but also from posterior mesodermal precursors and most of them do not emerge from haematovascular precursors (Ciau-Uitz and Patient 2016). In contrast, evidence strongly suggests that all haematopoietic cells develop through a bipotent progenitor stage (Ciau-Uitz and Patient 2016). Furthermore, it has been suggested that all blood cells in the mouse embryo originate from $Flk1^+$ mesoderm (Lugus et al. 2009). These observations, backed

by further *in vivo* experiments, led to the conclusion that the haemangioblast represented a state of competence rather than a single-cell bipotent progenitor (**Fig. 1.1B**) (Amaya 2013). While such cells could be isolated or generated *in vitro*, shared haematoendothelial precursor populations displayed strong preference for the haematopoietic fate *in vivo*. Furthermore, clonal fate mapping studies in mouse embryos suggested independent specification of early endothelial and blood precursors (Ueno and Weissman 2006; Padron-Barthe et al. 2014). However, these studies have been criticised for suboptimal lineage tracing, potentially resulting in mis-labelling of the actual haemangioblasts (Lacaud and Kouskoff 2017). Therefore, the existence of the embryonic haemangioblast *in vivo* remains to be demonstrated beyond doubt.

1.1.2. Primitive and pro-definitive waves of haematopoiesis

The development of blood occurs in distinct waves, each of them characterised by the generation of blood progenitors in a specific location and restricted in time (Ciau-Uitz et al. 2014). The terminology used to describe multiple stages of haematopoietic specification continues to be a topic of debate. Different definitions of primitive versus definitive haematopoiesis have been discussed recently (Lacaud and Kouskoff 2017). Lacaud and Kouskoff argue that the most widely approved definition assigns the term 'primitive haematopoiesis' only to the initial wave taking place before E8 in the mouse embryo, as only this wave generates blood cells that are transient and do not persist into adulthood (Lacaud and Kouskoff 2017). The authors classify all the other sources, including erythromyeloid progenitors (EMPs) and haematopoietic stem cells (HSCs), as definitive. Although the aim of this binary distinction is to simplify the understanding of developmental haematopoiesis, it does not reflect the functional output and the clinical significance that distinguishes the multilineage, long-term engrafting HSCs from other blood progenitors. Therefore, this thesis adopts the idea of three waves, with only the HSCs representing the 'definitive' wave, preceded by primitive and pro-definitive waves (Ciau-Uitz et al. 2014).

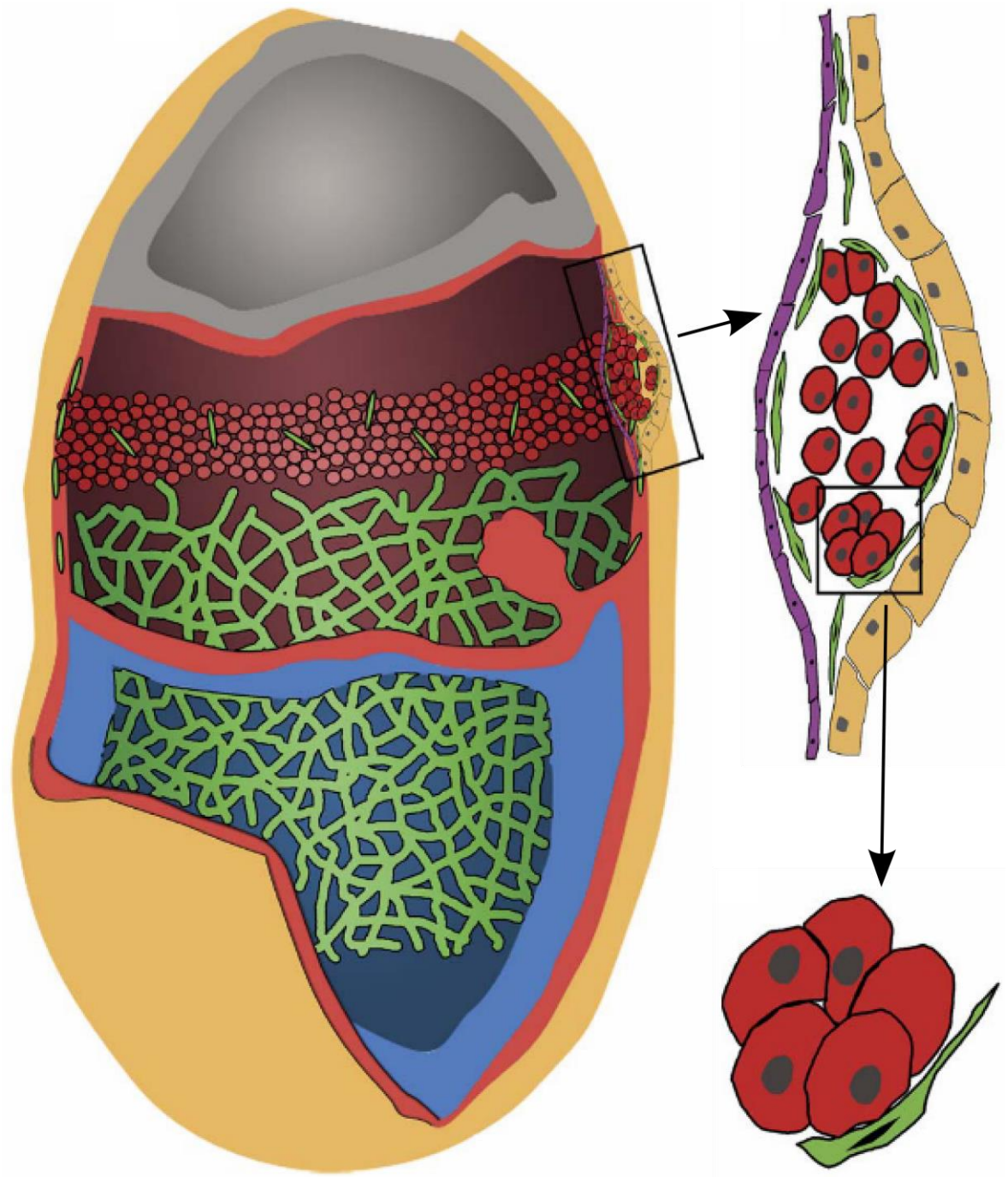


Figure 1.2. Blood islands in the yolk sac of a mouse embryo consist of closely associated erythroid and endothelial progenitors.

An anatomically distinct band of primitive erythroid progenitors (red) and endothelial cells (green) defines the blood island region of an E7.25 mouse embryo. The blood islands are located between the mesodermally-derived epithelial layer (purple) and visceral endoderm (yellow). They are composed of primitive erythroid progenitors in close association with endothelial cells. (Reproduced from Ferkowicz and Yoder 2005)

The first blood cells in the developing embryo define the primitive wave. It initiates soon after gastrulation with the formation of blood islands from the extraembryonic mesoderm, located in the yolk sac (**Fig. 1.2**). The blood islands, first defined in the chick embryo and later identified in the mouse, are assembled from committed erythroblasts closely associated with endothelial cells (**Fig. 1.2**) (Ferkowicz and Yoder 2005). CD41 accurately marks the earliest blood progenitors in the developing mouse embryo around E7.25 (Ferkowicz et al. 2003; Mikkola et al. 2003). Its expression, together with *Tal1*, *Gata1* and *β-globin*, marks an anatomically distinct band of haematopoietic cells in the blood island region of the murine yolk sac (**Fig. 1.2**) (Ferkowicz and Yoder 2005). They give rise to primitive erythrocytes, unique to the embryonic stage, which enter circulation at E8.25, undergo maturation and, after the onset of circulation, fulfil a critical role in delivering oxygen to the developing embryo (Palis 2008). The primitive erythrocytes differ from their definitive equivalents on the morphological and functional levels (Palis et al. 1999). In fact, one of very few ways to distinguish primitive and definitive waves in *in vitro* systems, where there is no spatial and little temporal separation between them, is by erythrocyte morphology (Lacaud and Kouskoff 2017). These primitive cells are subsequently replaced by erythrocytes arising from pro-definitive and definitive waves (Palis 2008). In addition to the erythrocytes, the primitive wave generates myeloid cells (Palis et al. 1999; Dzierzak and Speck 2008).

The *Xenopus* ventral blood island (VBI), the equivalent of the murine and chick yolk sac, also gives rise to primitive erythrocytes and macrophages (Ciau-Uitz et al. 2010a). The precursors of definitive and primitive lineages are plastic in their fate only until gastrulation, as demonstrated by grafting experiments in the *Xenopus* embryo (Turpen et al. 1997). Afterwards, they are separate in their fates.

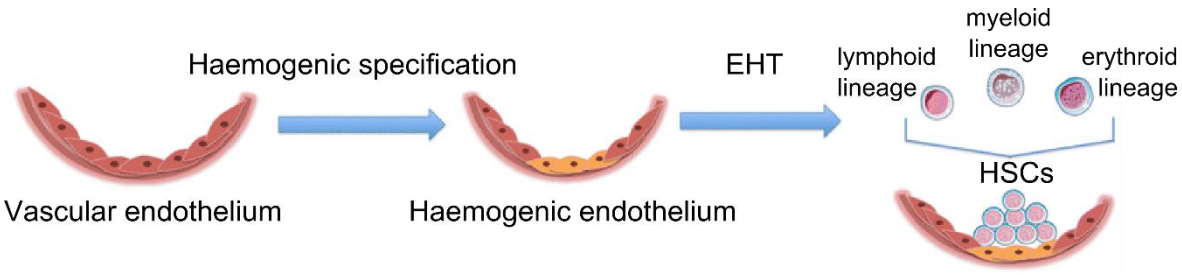


Figure 1.3. The linear model of HSC specification from vascular endothelium involves a haemogenic endothelial intermediate.

A subset of vascular endothelium, located in the ventral wall of the dorsal aorta, undergoes haemogenic specification. The specified haemogenic endothelial cells undergo endothelial-to-haematopoietic transition (EHT) and generate haematopoietic stem cells (HSCs) with the potential to give rise to all the adult blood lineages. (Adapted from Gritz and Hirschi 2016)

The second, pro-definitive wave of haematopoiesis, is characterised by the presence of erythromyeloid progenitors (EMPs) in the yolk sac. We term it 'pro-definitive' to reflect the persistence of blood cells generated in this wave beyond embryogenesis, as opposed to the earliest primitive progenitors. In the mouse, the EMPs include yolk sac-derived progenitors that mature in the foetal liver (Frame et al. 2013; McGrath et al. 2015). The EMPs are *cmyb*$^{+}$ and generate tissue-resident macrophages that remain there throughout adult life (Gomez Perdiguero et al. 2015; Hoeffel et al. 2015; Frame et al. 2016). This wave generates mostly erythrocytes and myeloid cells, but populations of B and T lymphocytes also arise from the yolk sac around E9-9.5 (Yoshimoto et al. 2011; Yoshimoto et al. 2012).

1.1.3. The haemogenic endothelium

A complementary model to the existence of bipotent haemangioblast precursors suggests a more linear progression of haematopoietic specification, where blood cells arise through an endothelial intermediate (**Fig. 1.3**) (Lancrin et al. 2009). Dye-labeling experiments in chicken and mouse embryos showed that circulating blood cells were derived from vasculature and laid the ground for the idea of the 'haemogenic endothelium' (HE) (Jaffredo et al. 1998; Sugiyama et al. 2003). These observations were further validated by lineage tracing experiments, where tissue-specific inducible Cre-mediated recombination of *VE-cad*$^{+}$ cells demonstrated that HSCs

derive from cells that were initially part of the primary embryonic vasculature (Zovein et al. 2008).

The definition of haemogenic endothelium has been a subject of intense debate. In the definitive programme, there is a consensus that the HE encompasses cells within the wall of the main embryonic artery, the dorsal aorta (DA). These cells express endothelial markers, have endothelial morphology and the potential to become haematopoietic stem and progenitor cells (HSPCs) in the process of endothelial-to-haematopoietic transition (EHT) (**Fig. 1.3**) (Lacaud and Kouskoff 2017). Yolk sac-derived EMPs and lymphoid progenitors also arise through an endothelial stage (Yoshimoto et al. 2011; Yoshimoto et al. 2012; Frame et al. 2016), suggesting a wider employment of the HE in developmental haematopoiesis. This pro-definitive HE, however, differs from its definitive analogue in that it gives rise to haematopoietic cells that lack the multilineage, long-term engrafting potential of definitive HSCs (Chen et al. 2011; Gritz and Hirschi 2016). Only the earliest wave of primitive haematopoiesis has not been traced back to endothelial origins, since the earliest blood potential in the embryo arises before the morphological differentiation of vasculature. However, these blood progenitors do express endothelial markers, further confirming the close molecular relationship of endothelial and blood lineages throughout the embryonic development. In this thesis, we refer to the haemogenic endothelium (HE) as the definitive HE, capable of generating the HSCs.

Haemogenic endothelium is a very rare population within all the endothelial cells, not only because of low cell count, but also because of its transient nature (Gritz and Hirschi 2016). Several early studies attempted to distinguish haemogenic from non-haemogenic vascular endothelium (VE) using flow cytometry for cell surface markers, including VE-cad, CD31, CD34, CD45 or Flk1 (Gritz and Hirschi 2016). Although they confirmed the endothelial origins of cells with haematopoietic potential, distinction between VE and HE has proven difficult. Additional insights came from the transgenic studies on *cis*-regulatory elements, suggesting that

an enhancer of *Flk1* was active specifically in the non-haemogenic fraction of endothelial cells, both in *in vitro* differentiation and in mouse embryos (Hirai et al. 2003). However, this idea was never employed to its full potential. Instead, a more recent single-cell sequencing study relied on the distinction of HE and VE based on an enhancer of *Runx1*, active specifically in the HE (Bee et al. 2010; Ng et al. 2010; Swiers et al. 2013a). This analysis revealed different unique profiles of endothelial subpopulations, dissecting separate stages during the maturation of HE and HSC specification (Swiers et al. 2013a). At the molecular level, it showed gradual silencing of the endothelial programme, loss of endothelial characteristics and acquisition of haematopoietic features (Swiers et al. 2013a). It also demonstrated that the definitive HE cells are specified to become either endothelial or haematopoietic, and single HE cells progressively lose the capacity to generate both lineages (Swiers et al. 2013a).

Advances in imaging allowed direct *in vivo* visualisation of the emergence of HSPCs from the HE by a process termed endothelial-to-haematopoietic transition (EHT) (Eilken et al. 2009; Bertrand et al. 2010a; Boisset et al. 2010; Kissa and Herbomel 2010; Lam et al. 2010). During EHT, HE cells lose their endothelial sheet morphology and molecular characteristics (VE-cad, Claudin5^{+} tight junctions, Ac-LDL uptake) and acquire in parallel the expression of the haematopoietic marker CD41, followed by CD45, as well as myeloid CD11b (Eilken et al. 2009). CD41 is regarded as the first marker of committed HE cells *in vivo* (Ferkowicz et al. 2003) and in ESC *in vitro* differentiation model (Mikkola et al. 2003).

Collectively, there is strong evidence that the specification of definitive HSCs initiates in cells with arterial characteristics and proceeds through an HE intermediate (**Fig. 1.3**) (Gritz and Hirschi 2016). The HE may exist at further stages of embryonic development or even postnatally, suggesting that it may not be as transient as currently believed (Gritz and Hirschi 2016). However, this idea requires further validation before it could be employed to derive clinically relevant HE/HSCs from postnatal stages.

1.1.4. Haematopoietic stem cells

Haematopoietic stem cells (HSCs) are the source of all blood cells that are produced throughout the lifetime of the organism. In the human, they reside in the bone marrow and are capable of self-renewal and differentiation into progenitor cells that generate erythrocytes, leukocytes and thrombocytes in the process of lineage-specific differentiation. These cells have been used clinically for decades in bone marrow transplants, primarily because of their capacity to replenish the whole blood system of a patient with a haematological disorder (Medvinsky et al. 2011). Currently, around 20,000 transplants per year are performed around the world (Daniel et al. 2016). However, identification of suitable donors still limits the success rate of these therapies (Garza et al. 2017). The only time during ontogenesis when the HSCs are generated *de novo* (rather than self-renewed) is during embryonic development, in the definitive wave of haematopoiesis. Therefore, recapitulation of this process holds great clinical promise, as it would allow *in vitro* production of patient-specific HSCs (**Fig. 1.4**). Moreover, such HSCs could be used for drug screening to optimise them for each patient (**Fig. 1.4**) (Daniel et al. 2016). However, generation of HSCs from human iPSCs has been challenging mostly because of the presence of very similar haematopoietic programmes in the developing embryo. For instance, both definitive and pro-definitive haematopoiesis programmes proceed through an HE stage, but at the molecular level these HE cells have different requirements for the protein Cbfβ (Chen et al. 2011). Further distinction of these programmes is critical for the future success in generating multi-lineage, long-term engrafting HSCs, characterised by homing to bone marrow and generation of all blood lineages (Daniel et al. 2016; Lacaud and Kouskoff 2017). Very recent proof-of-concept reports suggest that such cells can be generated *in vitro*, but the described protocols still lack efficiency and depend on the mis-expression of oncogenes (Lis et al. 2017; Sugimura et al. 2017). Hence, we need to continue making efforts to understand the specific cellular processes governing the appearance of the first HSCs during ontogenesis.

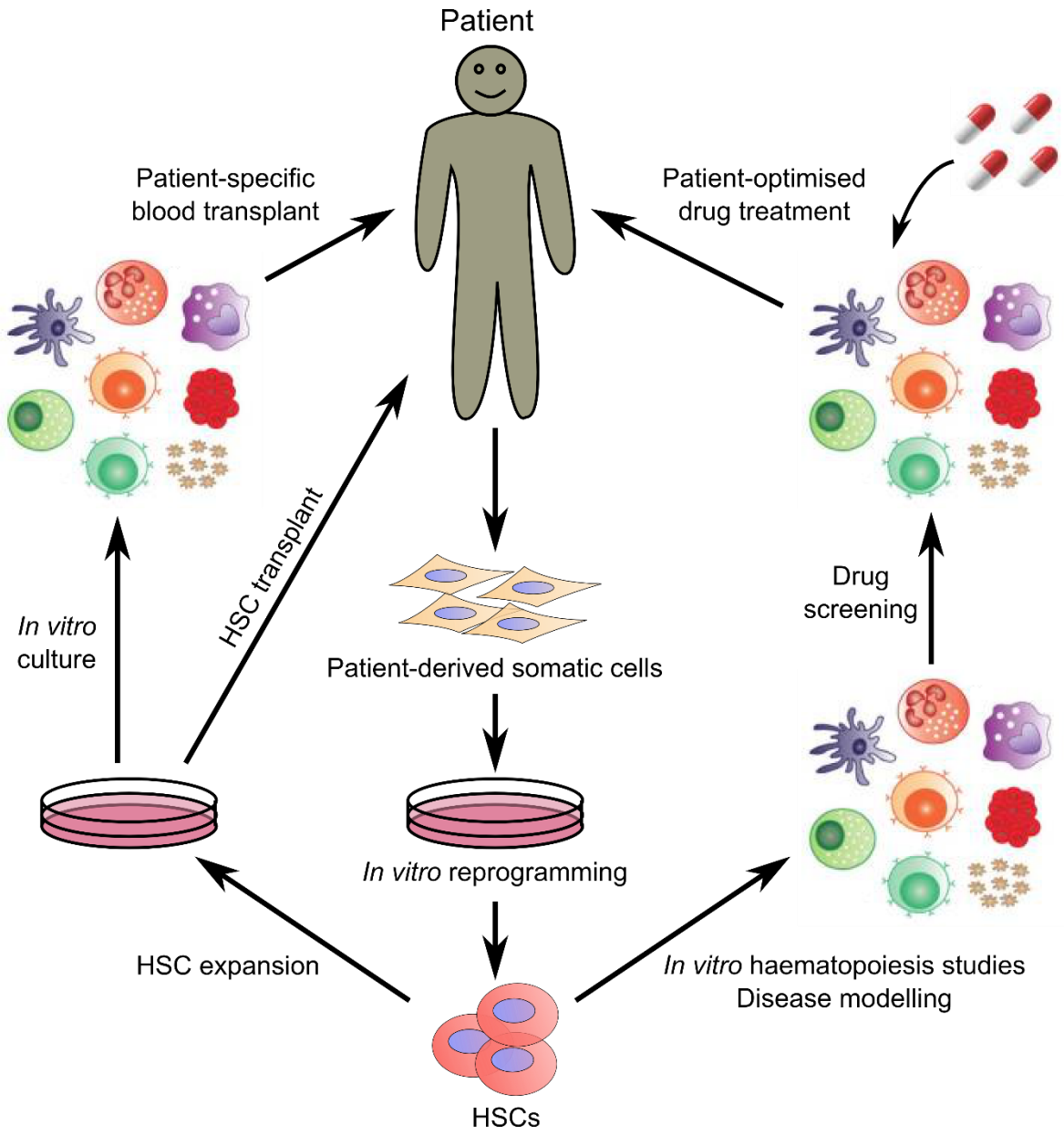


Figure 1.4. Generation of patient-specific HSCs *in vitro* holds great clinical potential.

The current efforts in the haematopoietic research field are concentrated on generating multilineage, long-term engrafting haematopoietic stem cells (HSCs) by *in vitro* reprogramming of patient-derived somatic cells. Such cells would, after their expansion, allow production of patient-specific HSCs or blood cells for transplant-based therapies. In addition, *in vitro* cultures of HSCs would allow further studies on the biology, self-renewal and differentiation of the HSCs, with the potential to model and study disease states *in vitro*. This would also allow drug screening *in vitro* to optimise drug-based therapies for each patient. (Adapted from Daniel et al. 2016)

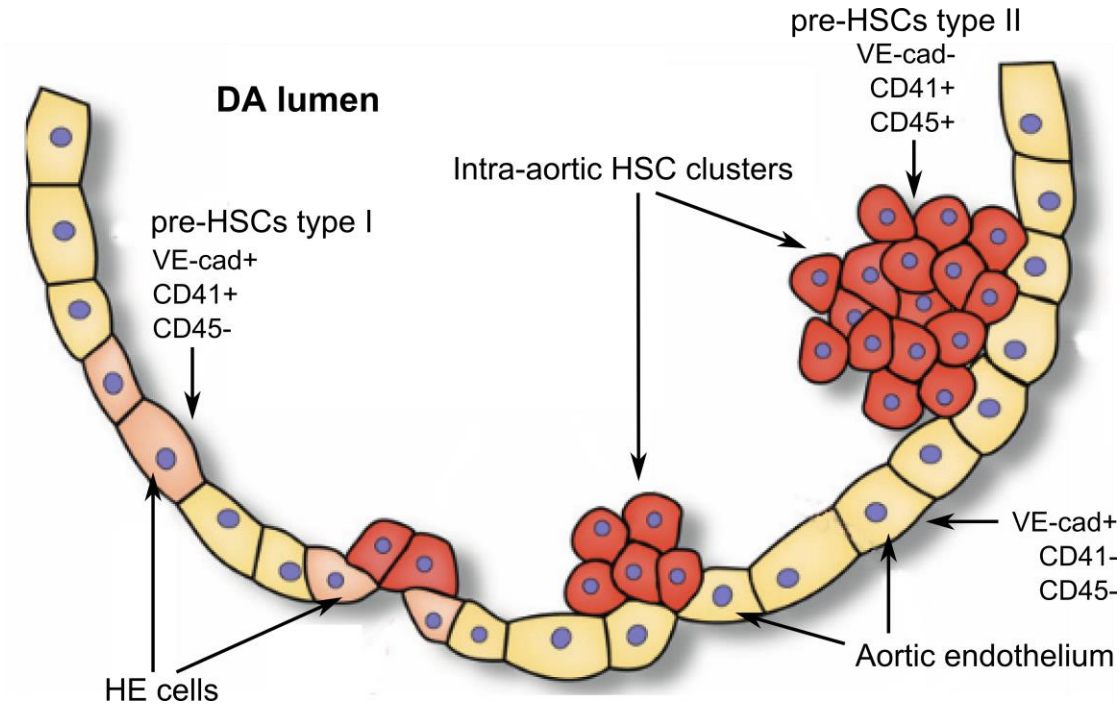


Figure 1.5. The first definitive HSCs emerge in intra-aortic clusters from the haemogenic endothelium in the floor of the dorsal aorta.

The vascular endothelium (yellow) lining the walls of the dorsal aorta (DA) is characterised by expression of VE-cad and the absence of CD41 and CD45. Haemogenic endothelium (HE; pink) can be identified by the expression of VE-cad and CD41 (pre-HSCs type I). These cells express transcription factor Runx1, required by the HE cells to give rise to intra-aortic clusters of pre-HSCs type II, characterised by the loss of VE-cad expression and acquisition of the CD45 marker. (Adapted from Costa et al. 2012a)

As opposed to extraembryonic yolk sac haematopoiesis, the HSCs have intraembryonic origins and this phenomenon is conserved across the model vertebrates studied (Ciau-Uitz and Patient 2016). The first HSCs are organised in haematopoietic 'clusters' (**Fig. 1.5**), which have been observed inside the main embryonic artery, the dorsal aorta (DA), in chicken (Dieterlen-Lièvre and Martin 1981), mouse (Garcia-Porrero et al. 1998), *Xenopus* (Ciau-Uitz et al. 2000) and human (Tavian et al. 1996) embryos. In the mouse embryo, these intra-aortic haematopoietic clusters first appear around E9.5 in the ventral wall of the dorsal aorta and reach their maximum prevalence at E10.5 (Medvinsky and Dzierzak 1996; de Bruijn et al. 2000; Yokomizo and Dzierzak 2010). Considering their antero-posterior localisation, the clusters are found in the aorta-gonad-mesonephros (AGM) region in the trunk, between the forelimbs and hindlimbs (Garcia-Porrero et al. 1998; Yokomizo et al. 2011).

As evidenced by *in vivo* imaging, the HSC clusters derive from the HE through EHT (Boisset et al. 2010). The clusters exhibit a loss of vascular identity, as evidenced by decreased expression of *Dll4* and *VE-cad* (**Fig. 1.5**) (Richard et al. 2013). Molecular profile of the HSCs emerging from the dorsal aorta includes the presence of CD31, CD41, c-Kit and Ly6a (Boisset et al. 2010). Based on the expression of a *Ly6a* transgene, these cells have been unambiguously localised in the endothelial layer of the DA and not in the surrounding mesenchyme (de Bruijn et al. 2002). However, the clusters were subsequently observed to be heterogeneous. They contain so-called 'pre-HSC' precursors that already possess long-term reconstitution capacity before they subsequently differentiate into mature HSCs (Boisset et al. 2015). Lineage potential analysis revealed that the pre-HSCs were generated from the HE independently of expression of *Ly6a*, suggesting that *Ly6a* was not fully indicative of haemogenic potential in the endothelium (Boisset et al. 2015). However, all HSCs become *Ly6a*$^{+}$ by E11.5 and the expression of this marker seems to be acquired after pre-HSC specification (Boisset et al. 2015). Importantly, this study provided insights into the stepwise generation of HSPCs from the HE. However, the full ensemble of different cells forming the intra-aortic clusters has not been characterised to date (Gritz and Hirschi 2016).

HSCs achieve their final numbers through rapid expansion in the AGM region between days 9.5 and 11.5, as demonstrated in a recent quantitative approach (Rybtsov et al. 2016), confirming the role of the DA as the primary source of definitive HSCs. They subsequently migrate with the bloodstream to colonize the foetal liver by E12 (Kumaravelu et al. 2002; Boisset et al. 2010) and the bone marrow by E16.5 (Coşkun et al. 2014), establishing the adult pool of HSCs (Zovein et al. 2008). This journey of HSCs proceeds through a number of niches, beginning in the AGM, through the foetal liver to the bone marrow. The adult HSC niche in the mammalian bone marrow has been well characterised and has served as a paradigm in understanding stem cell niches (Mendelson and Frenette 2014). However, transcriptomic meta-

analysis of various HSC niche cells revealed significant differences between the three different stages of developmental HSC niches, reflecting different requirements of HSCs at these stages (Charbord et al. 2014).

A full understanding of the generation of HSCs requires the dissection of multiple earlier programming steps that establish the DA and HE before the EHT (Ciau-Uitz and Patient 2016). The DA, the source of the first HSCs, originates from the lateral plate mesoderm (LPM) by migration of cells towards the midline to form a vascular cord, which later develops into the DA (Poole and Coffin 1988; Poole and Coffin 1989; Chen and Turpen 1995; Zovein et al. 2010; Garriock and Mikawa 2011; Medvinsky et al. 2011). The mesodermal precursors undergo similar rearrangements across vertebrates, first acquiring endothelial or blood fate and then migrating from their bilateral origins to the midline (Medvinsky et al. 2011; Ciau-Uitz et al. 2014).

Within the LPM, the definitive precursors in *Xenopus* originate from the dorsal lateral plate (DLP), in close proximity to the somites (Ciau-Uitz et al. 2010a). The DLP and its derivatives, the DA and HSCs, all originate from a single pair of blastomeres in the 32-cell embryo, which does not contribute to the VBI (see Chapter 1.1.2) (Ciau-Uitz et al. 2000; Walmsley et al. 2002). Therefore, the primitive and definitive waves of haematopoiesis are spatially separated from very early stages in *Xenopus* embryos.

Because of these common origins of the DA and HSCs, the *Xenopus* LPM has been defined as the source of so-called 'definitive haemangioblasts' (DHs) (Walmsley et al. 2002; Ciau-Uitz et al. 2013a). The DH, defined as a cell with the potential to generate the endothelium of the DA, the definitive HE and the HSCs, has only been identified in the *Xenopus* embryos (Ciau-Uitz and Patient 2016). Specification of the DH in the LPM initiates the HSC programming in the *Xenopus* embryo (Ciau-Uitz and Patient 2016). The transcriptional profile of the DH is not very different from that of other blood progenitors, but it is unique in its spatial ontogeny and environment (Ciau-Uitz and Patient 2016). Studies in *Xenopus* also suggest

that the location of mesodermally-derived progenitors along the anterio-posterior axis influences whether they acquire haemogenic potential or not (Ciau-Uitz and Patient 2016).

To date, there is no clear evidence for the existence of DH in the zebrafish (Ciau-Uitz and Patient 2016). The zebrafish haemangioblast population in the posterior LPM gives rise to primitive blood cells as well as the precursors of DA and HSPCs (Ciau-Uitz et al. 2010a). The blood and endothelial precursors migrate separately, with DA precursors migrating before the haematopoietic precursors (Gering et al. 1998; Davidson and Zon 2004). Limited evidence for the existence of DH in the mouse LPM comes from the studies showing *Tal1* expression in the DA precursors and from lineage tracing, assigning the origins of the DA and subsequently derived HSPCs to the LPM (Elefanty et al. 1999; Zovein et al. 2010).

1.2. Molecular specification of the HE and HSCs

1.2.1. Transcriptional regulatory network of HE and HSCs

Molecular specification of the haematopoietic tissues in the embryo depends on complex genetic programmes. Transcriptional control is key in defining expression of multiple lineage-specific genes. In general, transcription is modulated by the activity of transcription factors (TFs) that bind to DNA at promoter and enhancer regions. A few recent publications have pointed out the importance of interactions among haematopoietic TFs, encouraging efforts to generate gene regulatory networks (GRNs) responsible for the programming of haematopoietic structures (Pimanda et al. 2007b; Liu et al. 2008; Wilson et al. 2011; Ciau-Uitz et al. 2013a; Ciau-Uitz and Patient 2016; Kirmizitas et al. 2017). The GRNs will help simplify the complex molecular profile of embryonic haematopoiesis that includes several important transcription factors.

1.2.1.1. *Runx1*

The expression and function of transcription factor Runx1 is conserved across vertebrates (Swiers et al. 2010). It is present in endothelial cells, including yolk sac endothelium, and in the ventral, but not dorsal, wall of the DA (North et al. 1999). Expression of *Runx1* is indicative of HSC progenitors in the AGM region (North et al. 2002). In the zebrafish, *runx1* is one of the earliest markers of the HE (Burns et al. 2002; Kalev-Zylinska et al. 2002). Expression of *Runx1* in the *Xenopus* dorsal aorta is restricted to the trunk (Ciau-Uitz et al. 2010a) and is believed to mark the endothelium with EHT capacity. Expression of *Runx1* is regulated by a multiprotein complex through its +23.5 enhancer (Nottingham et al. 2007). The activity of this enhancer is regarded as a faithful marker of the HE (Bee et al. 2010; Ng et al. 2010; Swiers et al. 2013a). Furthermore, Runx1 likely maintains its own expression in the endothelium (North et al. 1999).

Runx1 is critically required for blood development (Okuda et al. 1996; Wang et al. 1996; Cai et al. 2000). Loss or haploinsufficiency of Runx1 affects specification of the HE and the generation of intra-aortic clusters in the mouse (North et al. 1999; Cai et al. 2000; Yokomizo et al. 2001). In ESC-derived HE cells *in vitro*, blood cells are not formed in the absence of *Runx1* (Lancrin et al. 2009). In addition, tissue-specific knockouts have positioned Runx1 as the critical regulator of EHT. Endothelial specific, *VE-cad*-driven deletion of *Runx1* results in the loss of the HSC clusters (Chen et al. 2009b). However, using a *Vav1*-Cre deleter strain to delete *Runx1* showed no effects, suggesting that Runx1 is required before, but not after EHT (Chen et al. 2009b). In agreement to this, loss of *runx1* function in zebrafish leads to aborted EHT events and eventual apoptosis of cells that were fated to become haematopoietic (Kissa and Herbomel 2010).

Using conditional knockout mice, it was shown that while Runx1 was not essential for the acquisition of CD41 expression (pre-HSC type I), its loss abrogated the acquisition of the CD45 cell surface marker that characterises the pre-HSC type II (**Fig. 1.5**) (Liakhovitskaia et al.

2014). This study distinguished multiple molecular steps during HSPC emergence and maturation. The developing HSPCs require Runx1 until E11.5, as demonstrated in temporally regulated tissue-specific mouse knockout embryos (Tober et al. 2013). The HSPCs become Runx1-independent before they colonize the foetal liver (Tober et al. 2013). In certain instances, reduced activity of Runx1 is associated with increased potential of HSPCs (North et al. 2002; Gritz and Hirschi 2016), pointing towards a requirement for precise and context-dependent regulation of this critical gene.

Global analysis of the genetic programme regulated by Runx1 in early blood development revealed its requirements in primitive blood from E7.5 (Tanaka et al. 2012). Identification of Runx1 direct transcriptional targets within the endothelium suggested unexpectedly early roles in remodelling of endothelial morphology prior to EHT (Lie-A-Ling et al. 2014). In the definitive HSC programme, Runx1 binds to regulatory regions of multiple haematopoietic regulators (Wilson et al. 2010; Tanaka et al. 2012). It represses the endothelial programme while activating the expression of haematopoietic genes (Swiers et al. 2010; Swiers et al. 2013b). Among the direct targets of Runx1, Gfi1 and Gfi1b are transcription factors that repress endothelial genes and allow the specification of haematopoietic morphology (Lancrin et al. 2012). Importantly, *Gfi1*;*Gfi1b*-deficient haematopoietic progenitors fail to depart from the vasculature (Lancrin et al. 2012).

In the human *in vitro* models of HSC specification from both ESCs and iPSCs, overexpression of *RUNX1* induces the emergence of haematopoietic cells and results in upregulation of *TAL1* and *GATA2* (Ran et al. 2013a). Most recently, *RUNX1* was used as one of the inducing TFs in both studies that reported the generation of multilineage, long term engrafting mammalian HSCs *in vitro* (Lis et al. 2017; Sugimura et al. 2017).

1.2.1.2. *Tal1/Scl*

Tal1, also known as Scl, is a transcription factor that plays crucial roles in developmental and malignant haematopoiesis (Porcher et al. 2017). During the primitive wave of embryonic haematopoiesis, expression of CD41, *Tal1*, *Gata1* and *β-globin* marks an anatomically distinct band of the earliest blood progenitors in the blood island region of the murine yolk sac around day E7.25 (Ferkowicz et al. 2003; Mikkola et al. 2003; Ferkowicz and Yoder 2005). After E7.5, expression of *Tal1* is restricted to blood and endothelial cells and precursors (Kallianpur et al. 1994; Elefanty et al. 1999). The DA precursors in the LPM express *Tal1* (Drake et al. 1997). In the zebrafish LPM, *tal1* is expressed in the haemangioblast population (Gering et al. 1998).

Loss of *tal1* results in the loss of all blood lineages in the mouse (Shivdasani et al. 1995; Porcher et al. 1996; Robb et al. 1996). In particular, early specification of the haematopoietic fate in the mesoderm critically requires the activity of this TF (Porcher et al. 1996). This has also been observed in the zebrafish, where MO-mediated knockdown of *tal1* resulted in the failure to specify the DA and all the blood lineages (Patterson et al. 2005). Forced ectopic expression of *tal1* is sufficient to convert other mesodermal tissues into haemangioblastic precursors (Gering et al. 1998; Gering et al. 2003). *Tal1* acts downstream of Npas4l, the gene mutated in the *cloche* mutants, characterised by defective development of endothelium and blood (Stainier et al. 1995; Liao et al. 1998; Reischauer et al. 2016). In *Xenopus*, *Tal1* is a key effector of the GRN programming the definitive haemangioblast (Ciau-Uitz et al. 2013a). Silencing *TAL1* reduces the haematopoietic potential in *in vitro* cultures of human cells, while its overexpression increases the numbers of $CD45^+$ blood progenitors (Real et al. 2012). However, this overexpression alone is not sufficient to establish long-term engraftment.

Two isoforms of *tal1* in zebrafish, *scla* and *sclβ*, have different roles in zebrafish haematopoiesis (Qian et al. 2007). *Scla* is believed to play a role in the maintenance of HSCs

after EHT, while *scl*β is involved in the programming of HE beforehand, upstream of *runx1* (Zhen et al. 2013).

Tal1 also plays an important role in primitive haematopoiesis, where it prevents the specification of cardiac fate in the yolk sac endothelium, allowing the generation of haemogenic endothelium (Van Handel et al. 2012). This is achieved by differential binding to a number of haematopoietic *cis*-regulatory enhancers (Org et al. 2015). In the foetal liver, *Runx1* is directly regulated by Tal1 (Landry et al. 2008).

1.2.1.3. *Etv2*

Etv2 belongs to the Ets family of TFs, comprising multiple regulators of haematopoiesis (Ciau-Uitz et al. 2013b). In the zebrafish, Ets factors are some of the earliest genes required for haematopoietic programming (Liu and Patient 2008). *Etv2* is among the earliest markers of developing blood in mouse ESC *in vitro* cultures (Wareing et al. 2012).

Etv2 is required for specification of mesodermal lineages (blood and endothelium) both *in vitro* (ESC differentiation) and in mouse embryos (Lee et al. 2008). This is likely achieved by direct induction of the Vegf receptor *Flk1*, as Etv2 binds to the promoter of *Flk1* (Lee et al. 2008). A similar mechanism was also observed in the frog (Salanga et al. 2010). Zebrafish *etv2* morphants also exhibit reduced levels of *kdrl* (orthologue of *Flk1*) expression (Sumanas and Lin 2006). However, other lines of evidence suggest that the relationship between *etv2* and *flk1*/*kdrl* is context-dependent and both factors can regulate each other (Ciau-Uitz et al. 2013b). Furthermore, haematopoietic differentiation of mouse ESCs *in vitro* suggested that *Etv2* was downstream of VEGF signalling and required for the specification of haemangioblast progenitors from $Flk1^+$ mesoderm (Kataoka et al. 2011).

Zebrafish and *Xenopus* embryos do not critically depend on *etv2* during definitive haematopoiesis (Salanga et al. 2010; Ciau-Uitz et al. 2013b; Robertson et al. 2016), suggesting

that its role may be in the specification of the haematopoietic programme at the level of the HE, which can be rescued at later stages of definitive haematopoiesis in the absence of functional Etv2. Similar to *tal1*, *etv2* is downstream of the *cloche* gene Npas4l (Stainier et al. 1995; Liao et al. 1998; Reischauer et al. 2016). In zebrafish, Etv2 is required for the activity of Tal1 isoforms (Ren et al. 2010), which places *etv2* upstream of *tal1* in the GRN. A similar relationship was observed in *Xenopus* (Ciau-Uitz et al. 2013a).

1.2.1.4. *Lmo2*

Lmo2 belongs to the Lim domain only (Lmo) family of proteins. These transcription factors do not bind DNA and their main role is to bridge other DNA-binding factors in multiprotein complexes (Sang et al. 2014). Lmo2 is known to form a complex with Ldb1, E47, Gata1 and Tal1 (Wadman et al. 1997; El Omari et al. 2013). This complex has been shown to induce the erythroid differentiation programme in *Xenopus* (Mead et al. 2001). Another complex involving Lmo2 likely regulates the expression of *Runx1* through its HE-specific enhancer (Nottingham et al. 2007).

Lmo2 knockout mice are embryonic lethal due to defective development of blood cells (Warren et al. 1994) and the loss of *Lmo2* results in the loss of haematopoietic repopulation capacity in adult mice (Yamada et al. 1998). In the zebrafish and *Xenopus*, expression of *lmo2*, alongside *tal1*, defines the haemangioblast in the LPM (Dooley et al. 2005; Ciau-Uitz et al. 2013a; Robertson et al. 2016). MO knockdown of *lmo2* in the zebrafish reduces expression of *runx1* and abolishes the generation of HSCs in the DA (Patterson et al. 2007). Similarly, *lmo2* acts together with *tal1* as a key downstream effector of the haemogenic programme in the *Xenopus* DLP (Ciau-Uitz et al. 2013a).

1.2.1.5. *Erg*

Erg, like Etv2, belongs to the Ets protein family (Ciau-Uitz et al. 2013b). It is expressed in the developing embryonic vasculature (Baltzinger et al. 1999; Loughran et al. 2008), including

the dorsal aorta of midgestation mouse embryos (Wilson et al. 2010). In the programming of DHs in *Xenopus*, *Erg* acts downstream of *Tal1* and *Lmo2* (Ciau-Uitz et al. 2013a).

In the analysis of the DNA binding profiles of ten key haematopoietic TFs using ChIP in cultured blood progenitor cells, Erg had the highest number of binding peaks (Wilson et al. 2010), suggesting a broad role in specifying the haematopoietic programme. Erg interacts at the protein level with Gata2 and Runx1, placing it among the most crucial haematopoietic TFs (Wilson et al. 2010).

ERG was used in *in vitro* studies to reprogramme committed human myeloid precursors into short-term engrafting erythro-myeloid progenitors (Doulatov et al. 2013). In this study, the combination of *Erg* and *HoxA9* was required for expansion of the lineage potential of these precursors (Doulatov et al. 2013).

1.2.1.6. *Sox7* and *Sox17*

Sox7 and Sox17 belong to the SoxF family of TFs. *Sox7* is expressed in the DLP of the *Xenopus* embryo and is regulated by *Tal1* and *Lmo2* (Ciau-Uitz et al. 2013a). It was also shown to be expressed in the HE prior to the emergence of blood progenitors *in vitro* and its downregulation was necessary to allow haematopoietic differentiation (Gandillet et al. 2009; Costa et al. 2012b).

Sox17 was proposed to be the main member of the SoxF family to be required for HSPC specification. Its endogenous locus drives GFP expression in the HE of a mouse embryo, as observed in *Sox17*^{GFP} knock-in mice (Clarke et al. 2013). Moreover, its expression in ESC-derived HE distinguished primitive from definitive haematopoietic fates (Irion et al. 2010). Sox17 is required for the specification of haemogenic endothelium that undergoes EHT, both in mouse embryos *in vivo* and in human cells *in vitro* (Clarke et al. 2013; Corada et al. 2013; Ditadi et al. 2015). Closely followed morphological changes during EHT correlated well with

temporally orchestrated loss of Sox17 and acquisition of Runx1 (Bos et al. 2015). As discussed in Chapter 1.1.3, transition to the haematopoietic fate requires downregulation of endothelial genes in the HE, and these include *Sox17* (Lizama et al. 2015). Sox17 in the endothelium represses *Gata2* and *Runx1* and its downregulation allows progression of haematopoietic specification (Lizama et al. 2015).

1.2.1.7. *Cmyb/myb*

The role of *cmyb* in definitive haematopoiesis was first identified in mouse *Cmyb* knockouts. *Cmyb* mutants had specific loss of foetal liver haematopoiesis without affecting primitive blood (Mucenski et al. 1991). Further studies using a knockdown allele of *cmyb* in the mouse embryo showed that it was also required for the downstream lineage specification and differentiation of HSPCs (Emambokus et al. 2003). *CMYB* is expressed in the definitive haematopoietic progenitors in the aorta of the human embryo, alongside *TAL1*, *GATA2*, *GATA3*, *C-KIT* and *FLK1* (Labastie et al. 1998). *CMYB* is also required for homing to the bone marrow and short-term engraftment of erythro-myeloid progenitors generated via reprogramming of committed human myeloid precursors (Doulatov et al. 2013). Zebrafish *cmyb* mutant embryos exhibit defects in the migration of HSCs from the DA to the caudal haematopoietic tissue (CHT), the zebrafish equivalent of the foetal liver (Zhang et al. 2011b). Cmyb is also required for further differentiation of HSCs. *Dnmt3bb.1* zebrafish mutants lack methylation of the *cmyb* locus, leading to loss of *cmyb* expression and subsequent loss of terminally differentiated haematopoietic cells (Gore et al. 2016).

1.2.1.8. *Gata2*

Gata2, a member of the Gata family of TFs, plays multiple roles in haematopoietic processes (Vicente et al. 2012). Mutations in the *GATA2* gene have been associated with several human blood disorders, including the MonoMAC syndrome, myelodysplastic syndromes and immunodeficiencies (Hsu et al. 2013; Spinner et al. 2014).

Gata2 is expressed in the mouse yolk sac and is involved in the proliferation and maturation of primitive blood progenitors (Silver and Palis 1997; Palis et al. 1999). During later stages, it is expressed in the AGM region (Minegishi et al. 1999). *Gata2*, alongside *Gata1*, was used as a marker for the induction of the haematopoietic fate in *Xenopus* mesoderm (Kelley et al. 1994; Bertwistle et al. 1996). In the zebrafish, concurrent expression of *tal1*, *lmo2* and *gata2a* defines the haemangioblast population in the LPM (Dooley et al. 2005; Robertson et al. 2016).

Gata2 knockout mice are embryonic lethal and die by E10.5 due to erythroid and HSPC defects (Tsai et al. 1994). *In vitro* studies also demonstrated that Gata2 was required for the early specification of HSCs, but not differentiation of downstream lineages (Tsai and Orkin 1997). Further studies in the mouse demonstrated a decrease in HSC numbers in *Gata2* heterozygous mutants, but also a dose-dependency of adult HSCs on Gata2 (Ling et al. 2004). Conditional *Gata2* knockout under *VE-cad* control abolished the generation of intra-aortic clusters (de Pater et al. 2013), suggesting that Gata2 is required for HSPC formation. In addition, conditional deletion of Gata2 mediated by a *Vav*-Cre transgene demonstrated that Gata2 is also required for the maintenance and survival of HSCs in the foetal liver, after the specification of HE and HSC emergence (de Pater et al. 2013).

Gata2 expression in the endothelium is regulated by an enhancer element located inside its intron 4 (Khandekar et al. 2007; Gao et al. 2013). Deletion of this enhancer results in reduced expression of *Runx1* and *Tal1* and the loss of HSPC emergence from HE, leading to lethality by E14 (Gao et al. 2013). Similarly, conditional knockout of *Gata2* in the activity domain of this enhancer results in the loss of HSPCs (Lim et al. 2012). These studies positioned Gata2 as an upstream regulator of *Runx1*. In support of this, Gata2 binds to the HE-specific *Runx1* +23.5 enhancer (Nottingham et al. 2007).

As for the relationship between *Gata2* and *Tal1*, *tal1* has been positioned downstream of *gata2* in the genetic cascade responsible for primitive and definitive haematopoiesis in zebrafish

posterior lateral plate mesoderm (PLM) (Patterson et al. 2005). A similar observation was made in *Xenopus* DLP (Ciau-Uitz et al. 2013a). However, Tal1 binds to a *Gata2* enhancer in the mouse, suggesting it acts upstream of *Gata2* (Sanalkumar et al. 2014). There are several reports of cooperative action of the two proteins (Nottingham et al. 2007; Pimanda et al. 2007b; Tripic et al. 2009; Wu et al. 2014; Org et al. 2015), but direct reciprocal transcriptional regulation has not been clearly demonstrated and it may be very context-dependent.

Gata2 is regarded as one of the key haematopoietic TFs during development. Simultaneous binding of key TFs, including Gata2, Runx1 and Tal1 has been used to identify new molecular players in the haematopoietic process in the context of existing gene regulatory networks (Tijssen et al. 2011). In a recent study, the *Ly6a*:GFP transgene was used to isolate cells undergoing EHT in the mouse AGM, including HE cells, haematopoietic cells in the intra-aortic clusters and HSCs (Kartalaei et al. 2015). Transcriptional profiling of these cells, compared to non-haemogenic endothelium, revealed that all genes within the previously identified heptad of key haematopoietic TFs (*Tal1*, *Lyl1*, *Lmo2*, *Gata2*, *Runx1*, *Erg*, and *Fli1*) (Wilson et al. 2010) had increased expression during EHT (Kartalaei et al. 2015). Together, they directly regulate nearly 60 genes differentially upregulated during EHT (Kartalaei et al. 2015).

Out of all the known haematopoietic TFs, *GATA2* has been most widely used in *in vitro* attempts to generate HSCs (Daniel et al. 2016). Culturing of human iPSCs through an HE stage resulted in different haematopoietic potential depending on the overexpression of different transcription factors, suggesting varying transcriptional programmes responsible for definitive haematopoiesis (Elcheva et al. 2014). In both cases, *GATA2* was the commonly overexpressed factor. When it was accompanied by *ETV2*, it promoted myeloid fate, while co-overexpression of *TAL1* contributed to erythro-megakaryocytic specification (Elcheva et al. 2014). *Gata2* and *Gfi1b* were among four factors that induced the differentiation of mouse fibroblasts into haemogenic endothelium with blood-forming potential that resembled embryonic HSCs

(Pereira et al. 2013). Similarly, a cocktail of TFs: *Erg*, *Gata2*, *Tal1*, *Runx1* and *Lmo2* introduced into mouse fibroblasts transformed them into multilineage blood progenitors, including lymphoid potential, through an intermediate HE stage (Batta et al. 2014). However, these cells did not achieve long term engraftment.

1.2.2. Signalling pathways specifying the HE and HSCs

The above described transcription factors act in a cell-autonomous manner, meaning that they regulate the fate of the cells they are expressed in. However, embryonic development is largely orchestrated by intercellular communication, driven by signalling cues. The integration of many signalling inputs may be required for efficient differentiation of definitive HSCs *in vitro* (Espín-Palazón et al. 2014). Several conserved signalling pathways have been described in the context of definitive haematopoiesis in the developing embryos (Clements and Traver 2013). Some of them play very complex, context-dependent roles.

1.2.2.1. Notch

Notch is a type of juxtacrine signalling, which requires direct contact between the signalling and the receiving cell. Notch signalling is activated in response to binding of a ligand from the Delta/Serrate/Lag-2 family on the membrane of the signalling cell to a Notch receptor on the receiving cell. There are five Notch ligands (Dll1, Dll3, Dll4 and Jag1-2) and four Notch receptors (Notch1-4) in mammals (Andersson et al. 2011). In response to an interaction between a ligand and a receptor, the Notch receptor is proteolytically cleaved to release the Notch intracellular domain (NICD) that translocates to the nucleus and displaces transcriptional repressors (such as RBPJκ), simultaneously recruiting co-activators (Andersson et al. 2011; Butko et al. 2016). Notch signalling regulates multiple developmental processes and mutations in Notch pathway genes have been identified in various human disorders (Andersson et al. 2011). One of the most frequent patterning mechanisms regulated by Notch is lateral inhibition,

where Notch signalling from one cell represses transcription of Notch ligand in the receiving cell, establishing different fates in neighbouring cells (Butko et al. 2016).

The role of Notch in developmental haematopoiesis is incredibly context-dependent. There are multiple sources of Notch signalling that establish the HE fate (Butko et al. 2016). Moreover, Notch is silenced during the EHT to be reactivated for the maintenance of HSC fate. Out of the different waves of embryonic haematopoiesis, including pro-definitive EMPs, Notch is only essential for the generation of definitive HSCs (Kumano et al. 2003; Hadland et al. 2004; Bertrand et al. 2010b; Li et al. 2014a; Butko et al. 2016). Moreover, *in vitro*-generated HE requires Notch1 for the acquisition of lymphoid potential, in addition to the Notch-independent erythroid and myeloid fates (Clarke et al. 2013). Therefore, Notch appears to be a distinguishing factor in fate specification.

During embryogenesis, *Notch1* and *Notch4* are expressed in the endothelial cells and required for the patterning and function of the embryonic vasculature (Uyttendaele et al. 1996; Krebs et al. 2000; Limbourg et al. 2005). However, the requirement for Notch factors in the embryonic haematopoiesis seems to be separate from the arterial specification, because it depends on different members of the Notch pathway (Kanz et al. 2016). In the definitive haematopoietic wave in the AGM, the dorsal aorta endothelium expresses receptors *Notch1* and *Notch4* and ligands *Dll4*, *Jag1*, and *Jag2* (Robert-Moreno et al. 2008). Notch1 is required for HSPC generation in the HE in a cell-autonomous manner (Kumano et al. 2003; Hadland et al. 2004; Krebs et al. 2004) and induction of Notch1 promotes the generation of blood cells from the HE *in vitro* (Jang et al. 2015). Potential action of Notch involves p27-dependent shifts in the cell cycle, lengthening the G1 phase and contributing to the acquisition of stem cell identity (Marcelo et al. 2013). Another mechanism involves Foxc2, identified as one of Notch1 targets. In the zebrafish, dose-limited, suboptimal MO-mediated knockdown of *foxc1a* and *foxc1b*, the zebrafish orthologues of *Foxc2*, resulted in reduced expression of *runx1* and *cmyb* in the DA of

24hpf and 36hpf embryos, respectively, without affecting the endothelial development (Jang et al. 2015). In the fish, both paralogues of Notch1, Notch1a and Notch1b, play a role in the specification of HE from around 20hpf (Kim et al. 2014). Notch3 expression in the somites from 13hpf is required for HSC specification, but the precise mechanism remains to be resolved (Kim et al. 2014). Potentially it is important in specifying somite-derived cells that are necessary for signalling to the DA at later developmental stages (Nguyen et al. 2014).

The vasculature provides a 'niche' for the emergence of HSPCs and Notch is one of the critical players in establishing the vascular niche environment (Hadland et al. 2015). Emerging HSPCs respond to Notch ligands expressed by the endothelium and stroma (Gritz and Hirschi 2016). *In vitro*, the presence of co-cultured vascular niche, expressing Notch ligands DLL4 and JAG1, improved the haemogenic output of differentiating human ESCs (Gori et al. 2015). It activated Notch signalling and expression of its downstream targets, *GATA2* and *RUNX1* in the haematopoietic cells that were generated (Gori et al. 2015). These cells displayed engraftment, but failed to generate T lymphocytes.

Dll4 loss of function studies position it as a key regulator of arterial fate before the emergence of HSPCs, as loss of *Dll4* produces arterial defects before the onset of definitive haematopoiesis (Duarte et al. 2004; Krebs et al. 2004). On the other hand, *Jag1* mutant mice exhibit loss of definitive haematopoiesis without any effects on the endothelium (Robert-Moreno et al. 2008; Espín-Palazón et al. 2014). A similar phenotype was observed in zebrafish *jag1a* morphants (Monteiro et al. 2016). Expression of *Notch1* and *Jag1* is mutually exclusive (Robert-Moreno et al. 2005), suggesting Jag1-Notch1 signalling between adjacent cells in the aortic endothelium. This relationship is likely controlled by Sox17, a direct regulator of *Notch1* expression (Clarke et al. 2013).

Notch signalling produces different cellular outputs depending on the ligand. High Notch signalling, activated by Dll4 ligand expressed by neighbouring cells, induces arterial fate

specification in the signal-receiving cells. On the other hand, Jag1 activates low levels of Notch signalling in the neighbouring cells through the same receptor Notch1, which specifies the HSPC fate in the signal-receiving cells (Gama-Norton et al. 2015). Jag1 is important for inhibiting the arterial programme in the AGM, as demonstrated in culture experiments and in mouse embryos (Gama-Norton et al. 2015). Normally Jag1 outcompetes Dll4, but in its absence, Dll4 takes over and induces high Notch, leading to arterial differentiation over HSPC generation (Gama-Norton et al. 2015). Complementary experiments in human cells suggested that DLL4 signalling was important for the decision between the haemogenic and endothelial fate (Ayllón et al. 2015). High DLL4 signalling distinguishes endothelial-committed progenitors from low-DLL4 haematopoietic progenitors (Ayllón et al. 2015). Notch signalling between them, potentially involving JAG1, stimulates EHT and HSPC generation.

The expression of *Runx1* in the DA and specification of the definitive HE depends on Notch signalling (Burns et al. 2005; Nakagawa et al. 2006; Bertrand et al. 2010b). In the mouse, *Runx1* overexpression rescues the Notch1$^{-/-}$ haematopoietic defect (Nakagawa et al. 2006). Similarly, the loss of HSCs in *mindbomb* zebrafish Notch mutants can be rescued by Notch-induced *runx1* expression (Burns et al. 2005). How Notch induces *runx1* at the molecular level is not clear, but this regulation is unlikely to be direct (Robert-Moreno et al. 2005; Nottingham et al. 2007; Butko et al. 2016; Robertson et al. 2016). The likely cell-autonomous mediator of Notch-Runx1 regulation is Gata2, the direct target of Notch (Robert-Moreno et al. 2005), which has been suggested to act upstream of *Runx1* and binds to its +23 haemogenic enhancer (Nottingham et al. 2007). However, Gata2 is not sufficient to rescue the Notch1$^{-/-}$ haematopoietic defect (Nakagawa et al. 2006), suggesting additional inputs from Notch into *Runx1*.

Gata2 expression downstream of Notch1 is regulated by Jag1 (Robert-Moreno et al. 2005; Robert-Moreno et al. 2008). This requires the Notch co-activator RBP-Jκ. Loss of RBP-Jκ

decreases *Gata2* levels and results in defective generation of HSPCs (Robert-Moreno et al. 2005). In agreement with that, the *Gata2* promoter contains RBP-J consensus binding sites, further supporting the direct regulation by Notch (Guiu et al. 2013). In addition, *Gata2* is directly regulated by Hes repressors Hes1 and Hes5. Notch1 activates Hes1 repressor, responsible for inhibition of *Gata2* expression by direct binding to *Gata2* promoter (Guiu et al. 2013). As a result, expression of *Gata2* downstream of Notch1 appears in a pulsed manner during HE specification and its expression is maintained at intermediate levels. Loss of *Gata2* repression and its increased expression levels, concomitant with an increase in *Runx1* and *Cmyb* expression in the absence of *Hes1/5*, resulted in abnormal proliferation and defective maturation of AGM HSPCs (Guiu et al. 2013). This model, in which *Gata2* levels are modulated by critical negative inputs, was further suggested in the zebrafish, where negative regulators of *gata2* expression *gpr65* and *lmo4a* have been identified (Gao et al. 2016; K. Tuladhar, DPhil thesis).

Notch signalling itself stimulates the expression of its own genes, including arterial *Notch1*, *Jag1*, *Jag2* in the mouse (Robert-Moreno et al. 2005) and zebrafish *dlc* and *notch3* (Lawson et al. 2001). The pathway plays another crucial role in the early specification of the definitive haematopoietic fate during the migration of zebrafish progenitors from PLM to the midline. The migrating cells move in close proximity to the ventral side of the somites, exposing them to juxtacrine signalling from Notch ligands Dlc and Dld, expressed in the somitic region (Clements et al. 2011). This phenomenon is mediated by cell adhesion molecules, Jam1a expressed on the membranes of migrating mesodermal cells and Jam2a expressed on the ventral wall of the somites (Kobayashi et al. 2014). These molecules establish close contacts necessary for the Notch signal transduction.

Evidence strongly suggests that Notch signalling is regulated by other inputs during definitive haematopoiesis. Studies in zebrafish and mouse demonstrated that so-called 'sterile inflammation', characterised by various modes of inflammatory signalling, is required for the

generation of definitive HSCs. Developing HSCs in the mouse AGM have a strong inflammatory genetic profile, including IFN signature, in response to IFN signalling from primitive myeloid cells (Li et al. 2014a). Indeed, reduction of IFNα or IFNγ signalling in the mouse or zebrafish reduced the haemogenic potential of the DA (Li et al. 2014a; Sawamiphak et al. 2014). These effects have been replicated during the maturation of mouse AGM HSCs, showing increased long term repopulation potential in response to IFNα treatment (Kim et al. 2016). Another pro-inflammatory cytokine involved in definitive haematopoiesis is TNFα, likely acting cooperatively with IFNγ (Li et al. 2014a). Zebrafish mutants and morphants lacking *tnfa* or its receptor show reduced numbers of HSCs (Espín-Palazón et al. 2014; He et al. 2015). TNFα signalling from primitive neutrophils induces expression of *jag1a* in the endothelial cells, which signals via Notch1 to activate NFκB within the HE and promote EHT (Robert-Moreno et al. 2008; Espín-Palazón et al. 2014). In parallel to TNFα, granulocyte colony stimulating factor (Gcsf) activates NFκB within the HE to promote the emergence of HSCs (Stachura et al. 2013; He et al. 2015). The role of IFNγ in modulating the EHT was positioned downstream of Notch signalling (Sawamiphak et al. 2014). The sterile inflammation links definitive and primitive haematopoiesis, since the source of these signals are primitive neutrophils and macrophages (Espín-Palazón et al. 2014; Li et al. 2014a; He et al. 2015).

Another Notch modulatory mechanism was identified in the zebrafish. As demonstrated by the analysis of mutants deficient for the enzymes required for DNA demethylation, Tet2 and Tet3, Notch signalling and the downstream *gata2b/tal1/runx1* cascade in the HE depends on the activity of Tet2/3, suggesting a potential link to epigenetic regulatory mechanisms (Li et al. 2015). Furthermore, loss of at least one allele of Tet2 predisposes mutant mice to leukaemic transformation (Moran-Crusio et al. 2011), linking developmental haematopoiesis with adult malignancies. The Notch-dependent specification of HE also depends on retinoic acid signalling (Marcelo et al. 2013).

Following HE specification, downregulation of Notch is required for the specification of emerging HSCs, as demonstrated by the loss of intra-aortic clusters upon forced Notch signalling in the AGM endothelium (Tang et al. 2013). Other studies also showed that transient inhibition of Notch increased or accelerated the production of haematopoietic cells from the AGM (Richard et al. 2013; Lizama et al. 2015). In the mouse and chicken embryos, intra-aortic clusters have reduced Notch signalling, as opposed to aortic endothelium, characterised by maintained Notch signals (Richard et al. 2013). This is likely mediated by Sox17, required in in arterial endothelium to upregulate *Notch1* but also to repress *Gata2* and *Runx1* (Lizama et al. 2015). Expression of *Sox17* decreases upon EHT, resulting in decreased Notch and allowing expression of the HSC genes (Lizama et al. 2015). Furthermore, *Notch1* is inhibited by Gpr183 to allow the emergence of HSCs (Zhang et al. 2015). After the emergence of HSCs, Notch is required again for the maintenance of their identity in the foetal liver and bone marrow (Butko et al. 2016).

1.2.2.2. Vegf

Vascular endothelial growth factor (Vegf) signalling is an important conserved regulator of vasculogenesis and angiogenesis (generation of primary and secondary vessels, respectively). Five Vegf ligands exist in mammals, but they commonly function in various splice isoforms (Olsson et al. 2006). Three different receptor tyrosine kinases of Vegf signals (Vegfr1-3) have been described (Olsson et al. 2006). In the context of the development of the cardiovascular system, Vegfa isoforms mainly signals through Vegfr2, commonly known as Flk1/Kdr.

The crucial role of Notch in definitive haematopoiesis, described above, has been placed downstream of the Vegf signalling in various contexts, including the arterial specification preceding the generation of HSPCs. Here, the notochord and floor plate are the source of Hedgehog (Hh) signals that activate Vegfa production in the somites (Lawson et al. 2002). This phenomenon was also observed in the mouse embryo, where expression of Notch ligand *Dll4*

critically depended on Vegfa signals (Coultas et al. 2010). Their role is to induce Notch-dependent arterio-venous specification of the developing endothelium (Lawson et al. 2002), mediated by reciprocal EphrinB2/EphB4 signalling (Wang et al. 1998; Gerety et al. 1999; Grego-Bessa et al. 2007). In fact, VEGF, Hh, and Notch determine the arterio-venous fate decision already in the migrating LPM cells (Rowlinson and Gering 2010).

Furthermore, activation of Notch in response to Hh-Vegfa signalling inputs is required for the specification of the HE and HSPCs within the arterial endothelium (Gering and Patient 2005; Kim et al. 2013). *Tal1* is a key mediator of this crucial signalling axis in the DA region, directly regulating EHT in response to these signals (Gering and Patient 2005; Kim et al. 2013). Hey2 is another one of the Notch-induced TFs, required downstream of Vegfa for the expression of *notch1b* and *efnb2a* in the arteries, as well as *runx1* and *cmyb* in the HE (Lawson et al. 2001; Rowlinson and Gering 2010). Importantly, Hedgehog mutations block definitive, but not primitive haematopoiesis (Gering and Patient 2005), placing it as another cue distinguishing the two developmental programmes. In addition, Hh is an important signal polarising the DA in the mouse embryo, which is crucial for the emergence of the haematopoietic clusters (Peeters et al. 2009).

As discussed above, Vegfa is required both for arterial and haemogenic specification, but these roles are distinct and they have been dissected through experiments in frogs and fish. Initiation of the DH programme in the *Xenopus* LPM depends on paracrine Vegfa signalling from the somites (Ciau-Uitz et al. 2010b; Ciau-Uitz et al. 2013a). There, different isoforms of Vegfa act sequentially to induce the arterial and definitive haematopoietic fates in the trunk vasculature (Leung et al. 2013). The small isoform programmes the DH. Similarly, zebrafish short Vegfa isoform, $Vegfa_{121}$, is responsible for arterial patterning (Genthe and Clements 2017). Somitic activity of Eto2 drives production of medium and large Vegfa isoforms, which are specifically required for the HE specification, marked by expression of *Tal1*, *Gata2*, *Runx1* and

Gfi1 (Leung et al. 2013). This requirement was confirmed in zebrafish *eto2* morphants (Leung et al. 2013). The haemangioblast population expresses *Flk1/Kdr*, the receptor of Vegfa signalling. In response to the signals from the somites, the DLP activates autocrine Vegfa signalling through the transcription factor Etv6, which establishes the expression of *Tal1* (Ciau-Uitz et al. 2010b; Ciau-Uitz et al. 2013a).

1.2.2.3. Tgfβ

The Tgfβ superfamily of signalling peptides, composed of over 30 members, includes, among others, Tgfβ ligands and bone morphogenic proteins (BMPs). In response to ligand binding to different types of receptors, an intracellular cascade is activated, culminating in the activation of Smad transcription factors (Weiss and Attisano 2013). Different Smad proteins relay signals induced by different members of the superfamily (Weiss and Attisano 2013).

The involvement of BMP and Tgfβ signals in definitive haematopoiesis in not fully understood. The first insights came from the observation that BMP4 expression formed a gradient in the AGM of the mouse embryo, which was important for specification of the ventral wall of the DA (Marshall et al. 2000). Moreover, BMP4, localised in the stromal cells in the mouse AGM, has a pro-haemogenic effect on the cultured AGM cells (Durand et al. 2007). Further, BMP4 was found involved in establishing *Runx1* expression in the murine DA at E10.5 (Pimanda et al. 2007a).

Studies from zebrafish and *Xenopus* shed more light on the temporal and molecular mechanisms of BMP signals in definitive haematopoiesis. The BMP4 signal is required for polarisation of the DA and generation of HSCs in the zebrafish (Wilkinson et al. 2009). This effect is specific to HSC formation, as arterial specification is not affected upon loss of BMP signalling (Wilkinson et al. 2009). After 20hpf, BMP is modulated by Fgf signalling acting as a repressor of HE specification (Pouget et al. 2014). In the *Xenopus*, BMP signalling acts well before the specification of the HE. First, it is involved in the primitive haematopoiesis

programme (Kirmizitas et al. 2017). In the DLP, it is required to activate *Gata2* expression during DH specification (Kirmizitas et al. 2017). This agrees with *in vitro* data, where *Gata2* is an important target of BMP signalling in the specification of haemangioblast-derived lineages (Lugus et al. 2007). Afterwards, BMP induces expression of Kdr in the DH, which in turn relays the Vegfa signals from the somites (Kirmizitas et al. 2017).

Tgfβ ligands are also required for the specification and emergence of HSCs from the zebrafish DA (Monteiro et al. 2016). In this process, different sequential inputs are involved. Autocrine Tgfβ1a and Tgfβ1b signals within the endothelium specify the *runx1*+ HE, while Tgfβ3, produced by the notochord, is required for the subsequent EHT (Monteiro et al. 2016). The autocrine signalling is regulated by Vegfa signals from the somites, while the main downstream target of the Tgfβ signals in the DA is the Notch ligand *jag1a* (Monteiro et al. 2016). This function, like that of BMP4, was independent of arterial programming.

1.2.2.4. Wnt

Wnt signalling is highly conserved across metazoans and regulates multiple developmental processes (Logan and Nusse 2004). Disruption of Wnt processes confers disease states in the human (Logan and Nusse 2004). The main roles of Wnt signals encompass induction of cell fate decisions and establishment of polarised multicellular patterns (Loh et al. 2016). At the molecular level, Wnt ligands bind to Fzd receptors to induce intracellular signal transduction via Dvl (Logan and Nusse 2004). 19 different Wnt ligands exist that can bind to Frizzled/LPR receptors. They either activate β-catenin (canonical signalling) or JNK/PKC (non-canonical signalling) (Gritz and Hirschi 2016). The most important DNA-binding effectors of Wnt signals are TCF and LEF factors (Cavallo et al. 1998). In canonical Wnt signalling, they form a transcription activating complex upon β-catenin translocation to the nucleus (Daniels and Weis 2005). Non-canonical Wnt involves many other genes, such as Vangl, but its full mechanism and repertoire is not well understood (Richter et al. 2017).

The Wnt pathway is tightly regulated. The signalling range is limited by the chemistry of the ligand (Takada et al. 2006). Moreover, multiple negative regulators of the pathway are directly upregulated by Wnt signalling, establishing negative feedback mechanisms restricting the response to Wnt signals (Richter et al. 2017).

Wnt is required at multiple steps of definitive haematopoiesis, from HSPC specification, through emergence from HE to maturation and expansion (Richter et al. 2017). Both types of Wnt signalling (canonical and non-canonical) are involved in HSC development.

In the mouse, Wnt is already involved in EMP development (Frame et al. 2016). In the definitive wave in the AGM, Wnt components are expressed around E10, while β-catenin is restricted to the endothelium at the ventral end of intra-aortic clusters (Ruiz-Herguido et al. 2012; Chanda et al. 2013). Chemical modulation of Wnt in mouse AGM explants revealed that increased Wnt signalling increased the HSC potential, while Wnt inhibition reduced the HSPC emergence (Ruiz-Herguido et al. 2012). This effect was tissue- and stage-specific, with *VE-cad*-driven, but not *Vav*-driven, inactivation of β-catenin leading to reduced HSPC numbers (Ruiz-Herguido et al. 2012). That suggests that canonical Wnt signalling in the endothelium is required for the specification of HE and emerging HSCs in the AGM and *in vitro*, but it is not necessary at later stages of HSC maturation (Ruiz-Herguido et al. 2012). In fact, Wnt signalling needs to be switched off after the emergence of HSPCs from the HE (Chanda et al. 2013). After the HSCs depart from the DA, another Wnt input, mediated by Wnt3a, is required for HSC maintenance in the foetal liver (Luis et al. 2010).

In the zebrafish, Wnt plays a role in the early specification of the haematopoietic fate during the migration of LPM cells to the midline. Expression of Notch ligands *dlc* and *dld* in the somites depends on non-canonical Wnt signalling via Wnt16b ligand (Clements et al. 2011). This is partly mediated by Fgf signalling through Wnt16-dependent Fgf4 receptor at 14-17hpf (Lee et al. 2014). As a result, *wnt16* knockdown with MO results in reduced expression of HE

markers from 24hpf (Clements et al. 2011). The requirement for Wnt16 is modulated by Rspo1, an activator of Wnt signalling in the somites (Genthe and Clements 2017). Furthermore, constitutive Hh signalling in zebrafish *Patched* mutants promotes arterial fate at the cost of haemogenic fate, which is partly due to reduced somitic *wnt16*, *dlc* and *dld* (Wilkinson et al. 2012). Crucially, expression of *gata2b* in the HE depends on signalling by Wnt16 and Notch (Butko et al. 2015).

Different Wnt ligands have an input in the emergence of HSCs after their specification in the zebrafish. Somitic Wnt9a is required before 20hpf to establish an environment that regulates HSC emergence around 32hpf (Grainger et al. 2016). Inhibition of Wnt signalling by forced expression of its antagonist *dkk1* results in a decrease in *cmyb*$^{+}$ haematopoietic progenitors at 36hpf (Goessling et al. 2009). On the other hand, Wnt8 was sufficient to enhance the emergence of HSCs from the DA (Goessling et al. 2009). Stimulation of Wnt signalling by a chemical agonist resulted in increased prevalence of *runx1*$^{+}$ and *cmyb*$^{+}$ cells in the zebrafish DA (Wang et al. 2013; Grainger et al. 2016). Similar results were obtained by targeting the Wnt pathway at the transcriptional level with dominant negative Tcf (*dntcf*), which resulted in decreased haematopoietic specification (Goessling et al. 2009; Grainger et al. 2016). Tissue-specific overexpression of *dntcf* revealed that *gata2b*+ cells in HE critically depended on receiving the Wnt signal (Grainger et al. 2016). The reduction in HSC specification in *wnt9a* morphants can be rescued by *myca*, a known target of Wnt signalling (Grainger et al. 2016). The Wnt9a cue is likely proliferative, because *myca* is involved in cell cycle regulation.

In vitro model of differentiating human PSCs proposes that specification of different mesodermal populations from primitive streak cells depends on early inputs from Nodal and Wnt pathways, with pre-HSC mesoderm requiring high levels of Wnt (Ditadi et al. 2017). This is based on evidence that stimulation of Wnt early in the differentiation protocol enhanced the generation of posterior-like mesoderm with HE potential (Kitajima et al. 2016). Consistent with

that, *in vivo* studies in *Xenopus* suggest that the location of mesodermally-derived progenitors along the anterio-posterior axis influences whether they acquire the haemogenic potential or not (Ciau-Uitz and Patient 2016). Importantly, Wnt plays a patterning role in the mesoderm of a developing embryo (Clements and Traver 2013). Wnt has also been suggested to orchestrate the decision between the primitive and definitive haematopoiesis, as only Wnt-enhanced differentiating human cells acquired T-cell potential (Sturgeon et al. 2014).

1.3. Zebrafish as a model system

Zebrafish, first proposed as a vertebrate model organism in the 1980s, has been instrumental in dissecting numerous processes underlying cell and developmental biology (Howe et al. 2013b; Avagyan and Zon 2016). Adult zebrafish and the embryos have been widely employed to model human disease, including haematological, kidney, ocular, nervous, muscle, heart and oncogenic disorders (Santoriello and Zon 2012). The zebrafish has also been established as a powerful model for *in vivo* drug discovery (Robertson et al. 2016). Moreover, sequencing and annotation of the zebrafish genome, comprising over 26,000 genes (Howe et al. 2013b), has aided deeper understanding of the functions of single genes and genetic networks. Combined with recent advances in genome engineering technologies (Varshney et al. 2015), the zebrafish has become a very powerful model for genetic studies. Importantly, a large majority of human genes, particularly those known for their involvement in human pathologies, have orthologues in zebrafish (Howe et al. 2013b; Li and Uitto 2014). Because of all the above advantages, methods to maintain and manipulate zebrafish for research have been largely standardised (Westerfield 2007), resulting in the generation of worldwide, collaborative research community that contributes with critical insights into developmental and pathological processes in vertebrates.

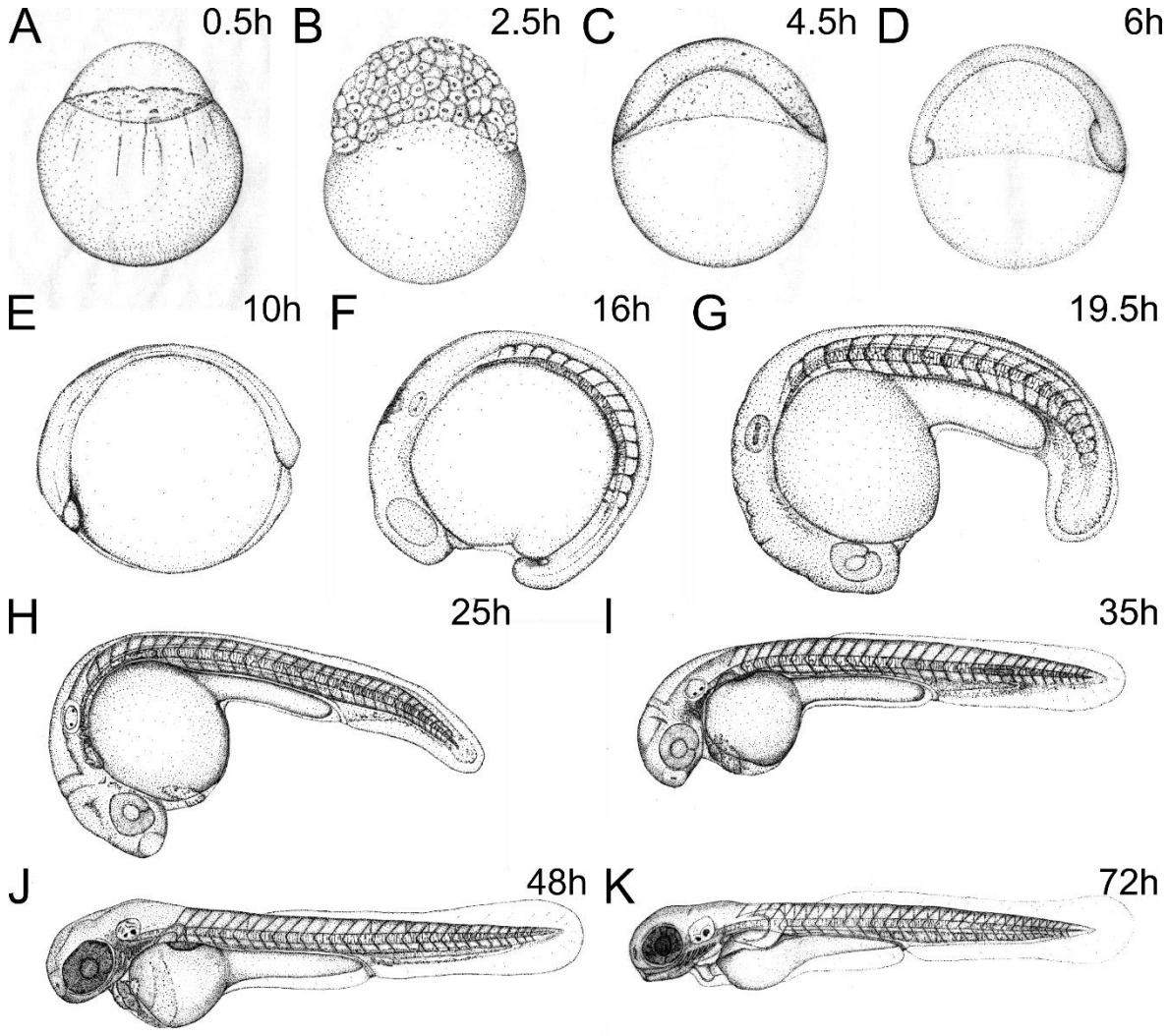


Figure 1.6. Selected stages of zebrafish embryonic development.

(A) The zebrafish zygote (up to 40min. after fertilisation) consists of yolk on the vegetal pole (bottom) and non-yolky blastodisc on the animal pole (top). (B) Through synchronous rounds of cleavage, the embryo reaches 256 cells at 2.5hpf. (C) By 4.5hpf, the embryo has initiated the epiboly, defined as the migration of blastoderm over the yolk towards the vegetal pole. (D) A 6hpf embryo has an embryonic 'shield', which defines the dorso-ventral axis. Gastrulation initiates shortly afterwards. (E) At 10hpf, the zebrafish embryo has established anterio-posterior and dorso-ventral body axes. (F) A 16hpf embryo has 14 somites. Segmentation in zebrafish generates on average 2 somites per hour. (G) By 19hpf, the tail of the embryo is clearly defined, as well as the yolk extension on the ventral side of the body. The number of somites reaches 21 at 19.5hpf. (H) At 25hpf, the heartbeat has initiated and the first blood cells start circulating. (I-K) The embryo grows and develops a streamline shape through days 2-3 of embryonic development. Most organs are generated by 72hpf. Images not to scale. A-D: Vegetal to the bottom. E-K: Anterior to the left. (Adapted from Kimmel et al. 1995)

1.3.1. Development of the zebrafish embryo

Arguably the biggest advantage of zebrafish as a model is its embryonic development. The zebrafish are capable of frequently laying hundreds of externally developing eggs, providing large amounts of material for genetic analyses. Moreover, the oviparity makes the embryos easily traceable and available for manipulation already from fertilisation. The embryos are optically transparent, making them particularly useful for live observation of developmental processes, even in a high-throughput manner (Pardo-Martin et al. 2010). They develop much more quickly than the embryos of other model vertebrate organisms, such as *Xenopus* or mouse: already at 24 hours post fertilisation (hpf) the main body axes are established and organogenesis is essentially completed by 5 days post fertilisation (dpf) (Kimmel et al. 1995).

The embryonic stage of the zebrafish is traditionally defined as the first three days since fertilisation, until the zebrafish larvae hatch from the chorion and develop a protruding mouth around 72hpf (Kimmel et al. 1995; Parichy et al. 2009). The early zebrafish egg is characterised by the presence of yolk on the vegetal pole and non-yolky blastodisc on the animal pole (**Fig. 1.6**). After the first cleavage, around 40min. post fertilisation, the blastomeres go through 15-20min. cycles of synchronous divisions, producing a mass of cells on top of the yolk (**Fig. 1.6**) (Kimmel et al. 1995). After ten cycles, the embryo enters the midblastula transition, associated with increased length of the cell cycle and initiation of zygotic transcription (Kane and Kimmel 1993). Around 4hpf, the process of epiboly initiates, characterised by migration of blastomeres, collectively termed the blastoderm, over the yolk towards the vegetal pole (**Fig. 1.6**) (Kimmel et al. 1995). When the blastoderm has covered 50% of yolk, gastrulation begins through involution of the blastoderm, followed by convergent movements forming the embryonic shield, which defines the dorso-ventral axis (Kimmel et al. 1995). By 10hpf, anterio-posterior and dorso-ventral body axes are clearly established and the germ layers are morphologically distinct (**Fig. 1.6**) (Schier and Talbot 2005). At this point, the process of

segmentation begins, generating on average two somites per hour, resulting in elongation along the anterio-posterior axis (Kimmel et al. 1995). In parallel, the neural tube forms in the dorsal part of the embryo, while the yolk extension develops on the ventral end of the trunk. A 19hpf embryo has 20 somites and a clearly defined tail (**Fig. 1.6**) (Kimmel et al. 1995). The embryo reaches the phylotypic stage at 24hpf, which is marked by the initiation of heartbeat. From this point, the retina and neural crest-derived melanophores produce pigmentation, spreading along the anterio-posterior axis (Kimmel et al. 1995). Blood circulation starts at around 26hpf. The embryo continues to grow and reaches a free-feeding stage by 5dpf.

1.3.2. Definitive haematopoiesis in the zebrafish

Despite millions of years of evolutionary distance between the zebrafish and the human, the process of making blood in the embryo is fascinatingly similar among vertebrate species (Robertson et al. 2016). Different waves of embryonic haematopoiesis exist in all vertebrates, including the zebrafish, and the genetic programmes responsible for the generation of HSCs are highly conserved (Ciau-Uitz et al. 2014). In fact, the 24hpf stage, when definitive haematopoiesis initiates in a zebrafish embryo, is characterized by the expression of evolutionarily most ancient genes (Domazet-Lošo and Tautz 2010). This, together with the above-mentioned advantages, makes zebrafish a particularly suitable model for studying definitive haematopoiesis. Moreover, zebrafish are peculiar among vertebrates in their ability to survive embryonic stages even in the absence of circulating blood (Pelster and Burggren 1996). This provides additional advantage over the mouse model, where the embryos do not survive in the absence of primitive blood. Because molecular mechanisms responsible for the generation of primitive and definitive blood largely overlap (Lancrin et al. 2009; Daniel et al. 2016), dissection of HSC-specific mechanisms in the mouse relies on tissue-specific and temporally inducible approaches, which are not necessarily required in the fish. In fact, several molecular mechanisms in definitive haematopoiesis discovered in the zebrafish have been validated in mammalian systems and

subsequently contributed to our understanding of mammalian haematopoiesis (Robertson et al. 2016).

Definitive haematopoiesis in the zebrafish is preceded by the primitive and pro-definitive waves (Ciau-Uitz et al. 2014). The earliest marker of zebrafish blood lineage is *drl*, expressed in the mesodermal progenitor from 9hpf (Mosimann et al. 2015). In the 10-somite zebrafish embryo, there are two distinct *tal1*$^{+}$ populations within the LPM, which mark the origins of blood development (**Fig. 1.7A**) (Davidson and Zon 2004). These two populations spatially separate primitive myeloid cells from primitive erythroid cells. The myeloid precursors are derived from the anterior lateral plate mesoderm (ALM) and mature in the rostral blood island (RBI) (**Fig. 1.7A**) (Gering et al. 1998; Herbomel et al. 1999). The PLM is located between the somites and pronephric duct precursors and gives rise to the DA, primitive and definitive blood and EMPs (**Fig. 1.7A**) (Gering et al. 1998; Bertrand et al. 2007; Ciau-Uitz et al. 2010a). It has been suggested that different regions of the PLM give rise to different cells, depending on their position along the anterio-posterior axis (Ciau-Uitz and Patient 2016). The most anterior cells are believed to generate vasculature, while the most posterior part of the LPM develops into the posterior blood island (PBI), the source of EMPs (**Fig. 1.7A**) (Bertrand et al. 2007). The middle part is the most mysterious, as it gives rise to the intermediate cell mass (ICM), the source of primitive erythrocytes (Detrich et al. 1995), but also to the dorsal aorta, the source of definitive HSCs (Zhong et al. 2001). To date, no fate or genetic mapping experiments have clearly separated these two populations. However, different cell populations within the PLM are believed to contribute to different tissues. They are spatially separated during the migration towards the midline, with DA precursors migrating before the haematopoietic precursors (**Fig. 1.7A**) (Gering et al. 1998; Davidson and Zon 2004). Unambiguous validation that the primitive and definitive waves are spatially separate in the PLM may require new markers of the HE

population, which are currently under investigation in our group (F. Bonkhofer, personal communication).

Around 18hpf, primitive erythrocytes are found in the ICM, the derivative of PLM (Detrich et al. 1995). After they have migrated, they remain in the midline and start circulating around 24hpf. The master regulator of erythroid cell specification in the zebrafish is *gata1* (Lyons et al. 2002), consistent with the role of Gata1 in the mouse model (Pevny et al. 1991; Pevny et al. 1995).

EMPs emerge from the posterior end of the ICM, the posterior blood island (PBI), at 24hpf (**Fig. 1.7B**) (Bertrand et al. 2007). Their differentiation into erythroid or myeloid fate is regulated by erythroid-specific Gata1 and myeloid Pu.1 (Galloway et al. 2005; Rhodes et al. 2005). The roles of Gata1 and Pu.1 vary between different blood lineages and *tif1γ* is a crucial modulator of Gata1/Pu.1 interactions (Monteiro et al. 2011). EMPs are not dependent on Notch signalling (Bertrand et al. 2010b), a feature that distinguishes them from definitive HSCs (Burns et al. 2005). The main output difference is the lymphoid potential, absent in the EMPs (Bertrand et al. 2007).

The first multilineage HSCs arise around 34hpf from the HE in the ventral wall of DA (**Fig. 1.7B**) (Kissa et al. 2008), the analogue of mammalian AGM (Davidson and Zon 2004). The endothelial-to-haematopoietic transition (EHT) has been visualised *in vivo* in live zebrafish embryos (Bertrand et al. 2010a; Kissa and Herbomel 2010; Lam et al. 2010). Kissa and Herbomel have described it as a series of morphological changes that resulted in detachment of a haematopoietic cell from the endothelial wall (Kissa and Herbomel 2010). This study used the Tg(*kdrl*:GFP) line, in which all vasculature is marked by GFP expression (Jin et al. 2005). Another study used double transgenic zebrafish to demonstrate that a subset of *kdrl*:mCherry^{+} cells expressing the *cmyb*:GFP transgene underwent the morphological transition (Bertrand et al. 2010a). The HSCs emerge from the ventral wall of DA into the sub-aortic space, located

between the DA and the posterior cardinal vein (PCV) (Kissa and Herbomel 2010). The process critically depends on *runx1*. Without *runx1*, EHT aborts and the emerging HSPCs undergo cell death (Kissa and Herbomel 2010). The budding process is also regulated by a cofactor of Runx1, Cbfβ, as shown by the analysis of *cbfb* zebrafish mutants (Bresciani et al. 2014). Moreover, on a non-cell-autonomous level, there is evidence for conserved involvement of primitive macrophages in the emergence of HCSs from the DA (Travnickova et al. 2015).

After the EHT, the HSCs enter the bloodstream through the PCV (Kissa et al. 2008) to colonise the caudal haematopoietic tissue (CHT) (**Fig. 1.7C**), the zebrafish equivalent of the mammalian foetal liver (Murayama et al. 2006). In the CHT, the HSCs become surrounded by endothelial cells in a process termed 'endothelial cuddling', which generates an HSC niche that specifies the direction of future asymmetric divisions and the fate of the daughter cells (Tamplin et al. 2015). A similar mechanism was subsequently observed in mouse foetal liver. Afterwards the HSCs migrate again with the bloodstream to colonise the kidney marrow (**Fig. 1.7C**) (Kissa et al. 2008).

The most commonly used assay to assess the output of embryonic definitive HSCs is to measure the numbers of the earliest differentiated derivatives of the DA HSCs, the T cells in the thymus, around 4-5dpf (**Fig. 1.7C**) (Murayama et al. 2006; Kissa et al. 2008). The most common assay employs *in situ* hybridization against *rag1*, a marker of T lymphocytes in the thymus (Willett et al. 1997). However, a very recent paper demonstrated that the T lymphocytes that colonise the thymus by 5dpf also originate from the HE cells in the posterior blood island before the emergence of the first HSCs in the DA (Tian et al. 2017). This will require future studies to re-evaluate the current ways to assess the presence and activity of DA HSCs.

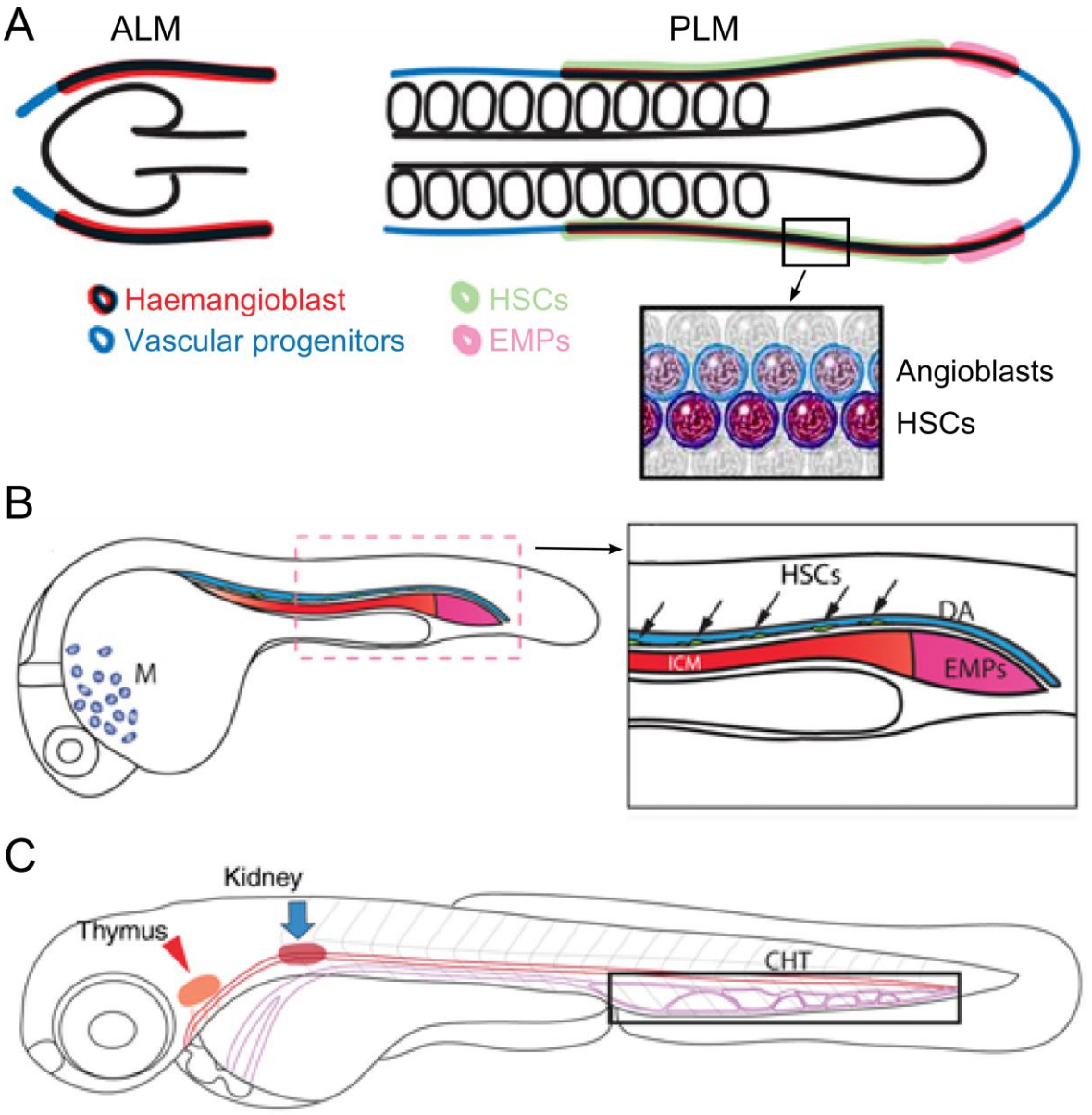


Figure 1.7. Definitive haematopoiesis in the zebrafish originates from the PLM and progresses through the CHT to kidney.

(A) At 10-somites stage, two bands of lateral plate mesoderm can be distinguished. Anterior lateral plate mesoderm (ALM) contains vascular progenitors (blue) and primitive haemangioblasts (red). Posterior lateral plate mesoderm (PLM) contains different types of progenitors depending on their position along the anterio-posterior axis. The most anterior PLM contains only vascular progenitors (blue), while its posterior end includes the precursors of pro-definitive erythromyeloid progenitors (EMPs; pink). The middle region contains the precursors of HSCs (green) and a haemangioblast population that gives rise to primitive blood (red). During the migration from the PLM to the midline, the presumptive endothelial cells (angioblasts) migrate before the HSC precursors. (B) By 24hpf, primitive macrophages (M) have derived from the ALM. The PLM gives rise to intermediate cell mass (ICM, essentially erythrocytes; red), dorsal aorta (DA; blue), EMPs (pink) and HSCs (green). The ICM contains erythroid progenitors. The HSCs (arrows) are specified in the ventral wall of the DA. (C) From 3dpf, the caudal haematopoietic tissue (CHT) is the intermediate HSC niche in zebrafish, where the HSCs proliferate and differentiate after emerging from the DA. Afterwards, they colonise the kidney (blue arrow), the zebrafish equivalent of mammalian bone marrow. The first differentiated derivatives of HSCs are T lymphocytes in the thymus (red arrowhead). (Reproduced from Davidson and Zon 2004; Monteiro et al. 2011)

1.3.3. *Gata2* paralogues in the zebrafish

Due to a whole genome duplication during the evolution of teleost fish over 300 million years ago, numerous zebrafish genes exist in the form of two paralogues (Taylor et al. 2001; Taylor et al. 2003). In the case of the *Gata* family of transcription factors, comprising six conserved vertebrate genes (*Gata1-6*), *gata1* and *gata2* have two paralogous copies in the zebrafish (Gillis et al. 2009). The *gata2* genes, *gata2a* and *gata2b*, have been suggested to undergone evolutionary subfunctionalisation from the role of the ancestral vertebrate *Gata2* gene, based on different expression patterns and functions in the embryo (Butko et al. 2015; Liu et al. 2016a). Moreover, the protein sequences of zebrafish Gata2a and Gata2b differ significantly, with the identity of 57% and similarity of 67% to each other (Butko et al. 2015). Each of them has been studied in different developmental contexts.

1.3.3.1. *Gata2a*

Shortly after the first reports suggesting the involvement of *Gata2* in the haematopoietic programming in the mouse and *Xenopus* (Kelley et al. 1994; Tsai et al. 1994), *gata2a* was identified as the zebrafish orthologue. It was first described as a marker of early haematopoietic progenitors in the mesoderm and used to follow their migration in the embryo in a series of *in situ* hybridization experiments (Detrich et al. 1995). In later studies, expression of *gata2a* was used as a readout to identify zebrafish mutants with haematopoietic defects (Stainier et al. 1995; Ransom et al. 1996). The expression pattern of *gata2a* mRNA during early somitogenesis includes endothelial and blood progenitors in the mesoderm, but it is also present in the ectoderm (Davidson and Zon 2004). Afterwards, it is expressed in all endothelial cells (**Fig. 1.8**) and in the developing central nervous system (CNS) (Yang et al. 2010; Butko et al. 2015), resembling the expression of *Gata2* in the mouse (Nardelli et al. 1999; Visel et al. 2004). Transgenic fish driven by the *gata2a* promoter expressed the fluorescent reporter in the CNS (Meng et al. 1997).

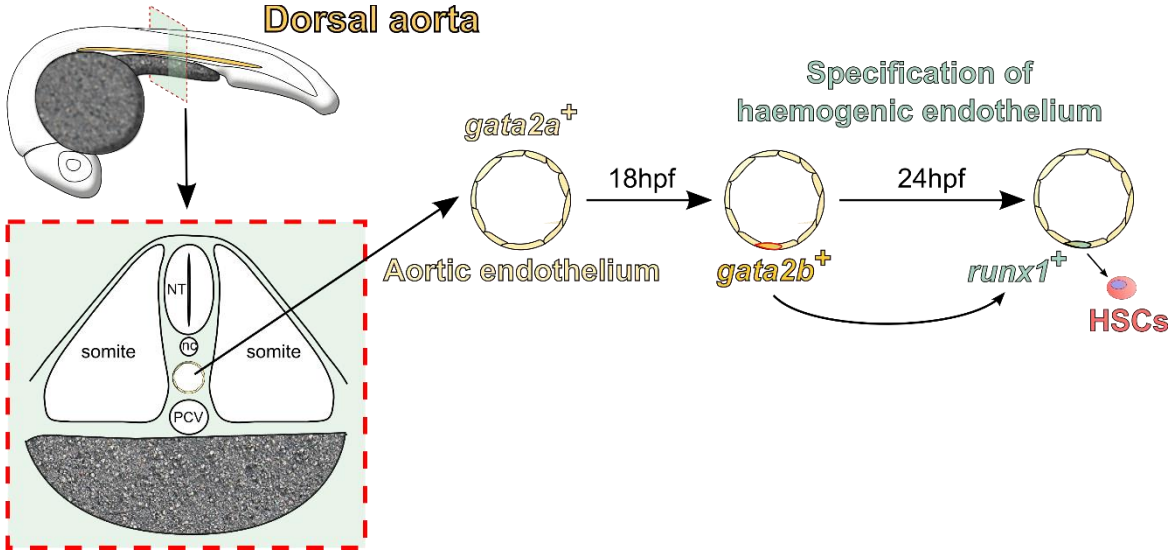


Figure 1.8. Both paralogues of zebrafish *gata2* are expressed in the DA endothelium prior to the emergence of HSCs.

Gata2a mRNA is expressed throughout the DA endothelium after the migration of PLM to the midline. Around 18hpf, the expression of *gata2b* initiates in the ventral subset of DA endothelium. This expression is required for the specification of haemogenic endothelium, activation of *runx1* from 24hpf and subsequent generation of HSCs. NT: neural tube; nc: notochord; PCV: posterior cardinal vein.

The first loss-of-function studies employing *gata2a* morpholino oligos (MOs) revealed that it was not involved in the specification of erythroid or myeloid fate during the differentiation of blood cells, suggesting its involvement at earlier stages of haematopoiesis (Galloway et al. 2005). Knockdown data from our lab showed reduced expression of endothelial and blood precursor markers in the PLM, including *fli1*, *scl*, and *gata1*, in the *gata2a* morphants (K. Tuladhar, DPhil thesis), consistent with the published role for Gata2 in the early programming of ventral mesoderm in *Xenopus laevis* (Maeno et al. 1996). However, other publications have focused on the role of *gata2a* in endothelial development. *Gata2a* morphants showed severe defects in intersomitic vessel angiogenesis at 48hpf (Fiedler et al. 2011). This effect was milder in *gata2a*^{um27} mutant embryos (characterized by a 10bp deletion in the *gata2a* protein-coding region) that showed arteriovenous shunts and reduced perfusion of the dorsal aorta at 48hpf (Zhu et al. 2011). Besides being involved in the development of

mesoderm-derived tissues, the mouse Gata2 has known functions in the CNS (Lentjes et al. 2016). The ventral spinal cord is characterized by expression of *gata2a* mRNA, which is required for the correct specification of GABAergic neurons downstream of Nkx2 family of homeobox transcription factors (Yang et al. 2010). However, a potential role for *gata2a* in definitive haematopoiesis has not been demonstrated.

1.3.3.2. *Gata2b*

Gata2b was first described for its role in the specification of haemogenic endothelium. It is characterised by a very restricted mRNA expression pattern, confined to the trunk of the embryo, more specifically the floor of the dorsal aorta from 24hpf (**Fig. 1.8**) (Butko et al. 2015). Lineage tracing experiments revealed that expression of *gata2b* marked HE and long-term definitive HSCs, while functional analysis in the MO-induced knockdown embryos revealed the requirement of *gata2b* for *runx1* expression and specification of HSCs (**Fig. 1.8**) (Butko et al. 2015). As a result, *gata2b* is currently regarded as a *bona fide* marker of HE (Gao et al. 2016). Its expression depends on Notch signalling, similarly to the mouse *Gata2* (Robert-Moreno et al. 2005; Guiu et al. 2013). The loss of *gata2b* cannot be rescued by *gata2a*, suggesting that *gata2a* does not act downstream of *gata2b* in the HE programming (Butko et al. 2015). However, possible interactions between them have not been fully investigated.

1.4. Objectives

Being arguably the most studied stem cells, the HSCs have provided important paradigms for the field of stem cell biology (Orkin and Zon 2008). The development of blood in the embryo, including the HSCs, critically relies on the orchestrated activation of genes responsible for establishing the haematopoietic fate and identity. Thorough understanding of their roles and interactions between them will likely result in valuable biological insights. Moreover, precise molecular delineation of the process of developmental haematopoiesis will allow recapitulation

of the HSC programme for clinical applications. In fact, a very recent study described *in vitro* generation of transplantable, long-term repopulating human HSCs from induced pluripotent stem cells (Sugimura et al. 2017). Importantly, this was achieved by applying lessons learnt from embryonic development about the endothelial origins of HSCs. Similarly, mouse HSCs with similar capabilities have been generated from endothelial cells (Lis et al. 2017). These reports came after many years of unsuccessful attempts to generate the HSCs (Daniel et al. 2016). Although it is too early to speak of a clinical breakthrough yet and both approaches should be optimised, these reports provide fresh hopes for future production of clinical HSCs. Deeper understanding of developmental haematopoiesis will likely play an instrumental role in this endeavour. However, further dissection of different functions of single haematopoietic genes may not be sufficient. Instead, understanding of the regulatory modes of these genes and interactions between them, including cell signalling and associated transcriptional controls, will likely provide a wealth of useful and potentially paradigm-shifting information. Thanks to recent advances in high-throughput analyses, developmental biologists are acquiring appropriate tools for figuring out the complex molecular regulatory network of the first HSCs.

Transcriptional regulation of developmental genes depends on orchestrated activity of transcription factors (*trans* regulators) and their DNA binding sites (*cis* regulatory regions). Among multiple transcription factors involved in the process of definitive haematopoiesis, *Gata2* takes a central role in integrating inputs from cell signalling networks and relaying them forward by direct regulation of numerous downstream targets. Dissection of the regulatory network involving Gata2 will on one hand further aid the recapitulation of HSC programme in the clinical setting, but also provide insights into human disorders associated with *GATA2* haploinsufficiency (Spinner et al. 2014). In this context, *cis* regulatory elements of the *Gata2* locus have been studied extensively. In the mouse, *Gata2* has served as an *in vivo* paradigm for dissecting tissue-specific expression and role of the gene determined by its enhancers, including

the critical +9.5 enhancer (Khandekar et al. 2007; Johnson et al. 2012; Gao et al. 2013). Further exploitation of this phenomenon will likely result in important practical insights for studying biological processes in development. Identifying and understanding these controls is critical, because numerous genes are involved in specification of multiple tissues at various developmental stages. Tissue-specific enhancers provide handles to dissect context-specific roles and may be the differentiating factors between, for example, non-transplantable primitive blood progenitors or long-term repopulating definitive HSCs.

Due to high degree of conservation of cellular events and molecular processes during definitive haematopoiesis, the zebrafish serves as an attractive model for studying the specification of HE and HSCs (Ciau-Uitz et al. 2014; Perlin et al. 2017). This thesis investigates the transcriptional regulatory mechanisms of *gata2* genes in the zebrafish embryo, employing high-throughput sequencing, genome engineering and low- to medium-throughput quantitative methods of tissue-specific gene expression analysis *in vivo*. The key objectives are: (1) to identify and characterize key conserved *cis* regulatory elements of *gata2a* and *gata2b* in the haematopoietic lineage; (2) to gain insights into distinct roles of *gata2a* and *gata2b* and interactions between them in the specification of HSCs; (3) to further analyse the previously reported role of *lmo4a* as a crucial negative *trans* regulator of *gata2a* expression.

Chapter 2

Materials and Methods

2.1. Solutions, oligonucleotides and molecular reagents

2.1.1. Solutions

All solutions were prepared with Milli-Q H_2O. They were stored at room temperature, 4°C or -20°C, or prepared fresh before each use (**Table 2.1**).

Table 2.1. List of solutions used in this thesis.

Name	Composition
ATAC-seq Lysis Buffer (¶)	10mM Tris-HCl, 10mM NaCl, 3mM $MgCl_2$, 0.1% IGEPAL, pH 7.4
AP Buffer for Purple (¶)	100mM Tris-HCl pH 9.5, 100mM NaCl, 50mM $MgCl_2$, 0.1% Tween-20
AP Buffer for Fast Red (¶)	100mM Tris-HCl pH 8.2, 100mM NaCl, 50mM $MgCl_2$, 0.1% Tween-20
Bleaching Solution (¶)	5% Formamide, 0.5xSSC, 10% H_2O_2
Blocking Buffer (¶)	1xPBS, 5% Goat Serum, 0.3% Triton X-100
Deyolking Buffer	116mM NaCl, 2.9mM KCl, 5mM HEPES, 1mM EDTA
E3 60x (*)	1.75% NaCl, 0.08% KCl, 0.3% $CaCl_2 \cdot 2H_2O$, 0.5% $MgCl_2 \cdot 6H_2O$, pH 7.2
EB Buffer	10mM Tris-Cl, pH 8.5 (QIAGEN)
Hanks' Solution (¶)	1xHBSS, 0.25% BSA, 10mM HEPES pH8
HBSS 10x	10x Gibco® Hanks' Balanced Salt Solution, No Calcium, No Magnesium (Life Technologies™)
HotSHOT Lysis Buffer	25mM NaOH, 0.2mM EDTA
HotSHOT Neutralisation Buffer	40mM Tris-HCl
$Hybe^{-/-}$ (*)	50% Formamide, 5xSSC, 9.2mM Citric Acid, 0.1% Tween-20
$Hybe^{+/+}$ (†)	50% Formamide, 5xSSC, 9.2mM Citric Acid, 0.1% Tween-20, 0.05% tRNA, 0.005% Heparin
LB	1% Bactotryptone, 0.5% Bactoyeast, 1% NaCl, pH 7.0
LB Agar (*)	1% Bactotryptone, 0.5% Bactoyeast, 1% NaCl, 1.5% Bactoagar, pH 7.0
MAB Block (†)	0.1M Maleic Acid, 150mM NaCl, pH 7.5, 2% Blocking Reagent™ (Roche)
MABT (¶)	0.1M Maleic Acid, 150mM NaCl, pH 7.5, 0.1% Tween-20
MS222 (*)	5g/l Ethyl-3-amino benzoate methanesulfonate salt, pH 7.0
PBS 10x	10x Gibco® Dulbecco's Phosphate-Buffered Saline, No Calcium, No Magnesium (Life Technologies™)
PBS+PFA (¶)	1xPBS, 4% paraformaldehyde
PBST	1xPBS, 0.1% Tween-20
PTU 50x (†)	0.15% 1-phenyl-2-thiourea
SOC (†)	SOC Outgrowth Medium (NEB)
SSC 20x (*)	3M NaCl, 0.3M Sodium Citrate
TAE 25x	1M Tris-acetate, 25mM EDTA
TE Buffer	10mM Tris-Cl, 1mM EDTA, pH 8.0
Trypsin + Collagenase Solution	1xHBSS, 0.05% Gibco® Trypsin+EDTA (Life Technologies™), 20mg/ml collagenase (Sigma)
X-gal/IPTG Solution (¶)	10mg/ml X-gal, 50% DMF, 500mM IPTG

(¶) – prepared fresh before each use; (*) – stored at 4°C; (†) – stored at -20°C

2.1.2. Morpholino oligonucleotides (MOs)

All MOs (**Table 2.2**) were obtained from Gene Tools (Oregon, USA). Stock solutions of 50μg/μl or 25μg/μl were prepared in H_2O and stored at room temperature. Injection mixes were made in H_2O with 0.05% Phenol Red (Sigma) and stored at -20°C. Before each use, they were incubated at 65°C for 5 minutes and cooled down to 4°C.

Table 2.2. Details of morpholino oligonucleotides used in this thesis.

Gene	Sequence (5'-3')	ng*	Reference
lmo4a	GTCAAAGGGTGCTGGAATCACTTCA	1.5	(K. Tuladhar, DPhil thesis)
gata2b	TTCACGTCCTATTGGCACACGATGC	30.0	(Butko et al. 2015)
gpr65	TAAATCGACAACTCACCATAAGTGC	4.0	(Gao et al. 2016)

*ng – injected amount in nanograms

2.1.3. Primers

All PCR and sequencing primers (**Table 2.3**) were obtained from Sigma-Aldrich, resuspended in EB Buffer to a stock concentration of 100μM and stored at -20°C. 5μM dilutions of combined primer pairs, stored at -20°C, were used for each PCR.

Table 2.3. Sequences of PCR primers used in this thesis.

Name	Sequence (5'-3')
Cas9 D10A f	AGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGT
Cas9 D10A r	ACAGAGTTGGTGCCGATGGCCAGGCCGATGCTGTACT
Cas9 D10A s1	TCGATAAGAACCTGCCCAAC
Cas9 D10A s2	GTCAGCTCGTTATACACGGT
Cas9 D10A s3	AGAACCAGACCACCCAGAAG
Cas9 D10A s4	CCAGTTCGTTTCCCTTCTGC
E1b s	GAACAATCAAGGGTCCCCAAACT
eef1a1l1 q f	GAGAAGTTCGAGAAGGAAGC
eef1a1l1 q r	CGTAGTATTTGCTGGTCTCG
gata2a i4-1.1kb f	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCAGCATCGGGATCCTATAA
gata2a i4-1.1kb r	GGGGACCACTTTGTACAAGAAAGCTGGGTCACGAATCAAACGCTTTCAG
gata2a i4-450bp f	GGGGACAAGTTTGTACAAAAAAGCAGGCTGATTGTTGAGAATGTTGTGGTGA
gata2a i4-450bp r	GGGGACCACTTTGTACAAGAAAGCTGGGTCCAGTCGAGCATGACAAACA
gata2a mut f	TGGCTAAGTGACCGTCAGAG
gata2a mut r1	TGAAACAAAACGCAGACGAC
gata2a mut r2	GGGTTTGTTGAAGACGGAAA
gata2a q f	GGACGAAAAGGAGTCCATCA
gata2a q r	GCACTCATAGCCAAGCTTCC
gata2b f1	ATGATGGATGCCCCAGCG
gata2b f2	TTTTGGCAAAGAATTCATGATGGATGCCCCAGCG
gata2b r1	TCAGCCTATAGCAGTGACTAAGC
gata2b r2	TCTGGATCATCATCGATTCAGCCTATAGCAGTGACTAAGC
hbbe1.1 q f	ATCTTCGCCAAGGCTGACTA
hbbe1.1 q r	CAGCGGCATTGTAGAGGTTT
kdrl q f	CGAACGTGAAGTGACATACGG

kdrl q r	CCCTCTACCAAACCATGTGAAA
lmo4a mut f	ACTTTGCCTCTGGATCTGCT
lmo4a mut r	TCTACATCAGTCCCGCCAAA
M13 f s	CGCCAGGGTTTTCCCAGTCACGAC
M13 r s	GTGAGCGGATAACAATTTCACACAGG
mpx q f	TGATGTTTGGTTAGGAGGTG
mpx q r	GAGCTGTTTTCTGTTTGGTG
pTALF1 s	CATCGCGCAATGCACTGAC
pTALR2 s	GGCGACGAGGTGGTCGTTGG
ubc q f	AAGAGACTCCCATACACCGC
ubc q r	ATTCTCAATGGTGTCGCTGG

f – forward primers; mut – diagnostic primers to detect genomic mutations; q – qPCR primers; r – reverse primers; s – sequencing primers

2.1.4. CRISPR/Cas9 sgRNAs

Optimised sgRNAs were designed with the CRISPR Design tool (http://crispr.mit.edu/). All primers used to make sgRNA templates (**Table 2.4**) were obtained from Sigma-Aldrich, resuspended in EB Buffer to a stock concentration of 100μM and stored at -20°C. 5μM dilutions of combined forward primer and the universal reverse primer were used for the synthesis reaction of each sgRNA template. The sgRNAs were synthesized as described previously (Bassett et al. 2013) and purified with MEGAclear™ kit (Ambion). The quality was verified on agarose gel. The sgRNA were stored at -80°C.

Table 2.4. Sequences of oligonucleotides used to synthesize sgRNA templates.

Name	Sequence (5'-3')
gata2a 1	GAAATTAATACGACTCACTATAGGG**TGACTCCATGGAAAACCCTG**GTTTTAGAGCTA GAAATAGC
gata2a 2	GAAATTAATACGACTCACTATAGGG**GGCCCATGCTTTTGCATTTT**GTTTTAGAGCTA GAAATAGC
gata2a 3	GAAATTAATACGACTCACTATAGGG**GATGTCATTCGGGCCTGCCG**GTTTTAGAGCTA GAAATAGC
gata2a 4	GAAATTAATACGACTCACTATAGGG**AGCACTATGTGTGAAGGGGC**GTTTTAGAGCTA GAAATAGC
reverse	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAAC TTGCTATTTCTAGCTCTAAAAC

Bold – sgRNA genomic target sites

2.1.5. Cas9(D10A) mRNA

For wild-type SpCas9 template, Addgene plasmid #51307 was provided by the Chen lab (Guo et al. 2014). The D10A Cas9 (Ran et al. 2013b) DNA template was generated by site-directed mutagenesis. 10μg DNA template was linearised with SacII enzyme. Cas9(D10A) mRNA was transcribed from 1μg linearised template with SP6 mMessage mMachine® kit

(Ambion) for 2h at 37°C. The DNA template was digested with DNAse I for 15min. at 37°C and the mRNA was purified with RNeasy® Micro Kit (QIAGEN). The integrity and quality of the mRNA was verified with agarose gel electrophoresis and with NanoDrop™ One (Thermo Scientific). The mRNA was stored at -80°C.

2.1.6. *Tol2* mRNA

1µg *tol2* DNA template vector (Kawakami et al. 2004) linearised with NotI was used for *in vitro* transcription with SP6 mMessage mMachine® kit (Ambion) for 3h at 37°C. The DNA template was digested with DNAse I for 15min. at 37°C and the mRNA was purified with an RNeasy® Micro Kit (QIAGEN). The quality of the mRNA was verified on an agarose gel and with a NanoDrop™ One (Thermo Scientific). Aliquots of 500ng/µl were stored at -80°C.

2.1.7. TaqMan® assays

All TaqMan® assays (**Table 2.5**) were obtained from ThermoFisher Scientific.

Table 2.5. Catalogue numbers of TaqMan® assays used in this thesis.

Gene name	TaqMan® assay catalogue number
angpt1	Dr03125214_m1
ascl1a	Dr03093273_g1
atoh8	Dr03130215_m1
brd4	Dr03431097_m1
btg2	Dr03172376_s1
cbfb	Dr03093780_m1
cbx1b	Dr03095502_m1
cdc20	Dr03133366_m1
cdk2	Dr03075917_m1
cebpd	Dr03106022_s1
cops2	Dr03114763_m1
csnk1e	Dr03439236_m1
ctnnb1	Dr03432620_m1
ctnnb2	Dr03082630_m1
dbx1a	Dr03144222_m1
dlb	Dr03124881_m1
dld	Dr03111908_m1
dll4	Dr03428646_m1
dlx2a	Dr03201055_s1
dnmt3bb.1	Dr03090840_m1

dvl2	Dr03439106_m1
ebf2	Dr03105767_m1
eef1a1l1	Dr03432748_m1
efnb2a	Dr03073975_m1
ephb2b	Dr03097160_m1
ephb4a	Dr03138273_m1
erg	Dr03109008_m1
etv2	Dr03077892_m1
etv4	Dr03138282_m1
etv5a	Dr03440241_m1
foxc1a	Dr03201313_s1
foxc1b	Dr03125188_s1
foxm1	Dr03202983_s1
foxo1a	Dr03149051_m1
fzd3a	Dr03091347_mH
fzd7a	Dr03074015_s1
gata2a	Dr03086718_m1
gata2b	Dr03140570_m1
gfi1aa	Dr03090705_m1
gfi1ab	Dr03438442_m1
gna12a	Dr03083477_m1
gsx1	Dr03077139_g1
her2	Dr03079914_s1
her4.2	Dr03160688_g1
hes6	Dr03131927_g1
hhex	Dr03079873_g1
hlx1	Dr03081830_m1
hoxa4a	Dr03080009_m1
hoxb2a	Dr03131369_m1
hoxc12b	Dr03118920_g1
hoxc13b	Dr03118921_m1
hoxc4a	Dr03086656_m1
hoxd4a	Dr03432559_m1
ikzf1	Dr03105499_m1
irf2bp2a	Dr03074782_m1
irf2bp2b	Dr03140259_m1
jun	Dr03202188_s1
kat2b	Dr03123451_m1
kdrl	Dr03432897_m1
klf17	Dr03080183_m1
lbx2	Dr03090028_m1
lef1	Dr03201167_s1
lsm12b	Dr03139893_m1
mafba	Dr03200761_s1

meox1	Dr03421220_m1
myb	Dr03432757_m1
mycla	Dr03120705_m1
nr2f1a	Dr03125009_m1
nrp1b	Dr03151877_m1
numb	Dr03091320_m1
pax6b	Dr03125136_m1
prox1a	Dr03086822_m1
psen1	Dr03131345_m1
rcor2	Dr03139615_m1
rpl13a	Dr03101115_g1
rplp0	Dr03131546_m1
runx1	Dr03074179_m1
sap30l	Dr03127103_m1
setd8a	Dr03148904_m1
smad4a	Dr03437376_m1
sox17	Dr03105677_m1
sox19a	Dr03200590_s1
sox4a	Dr03203889_s1
sox7	Dr03092284_m1
spred1	Dr03439751_m1
spry2	Dr03194153_s1
tal1	Dr03133428_m1
tbl3	Dr03096224_g1
tcf7	Dr03102668_m1
tcf7l1a	Dr03079946_m1
tgfb3	Dr03093764_m1
tmem88a	Dr03133058_m1
vangl2	Dr03201545_s1
wnt5b	Dr03150114_m1

2.1.8. *In situ* hybridization probes

10µg DNA plasmid templates coding for the antisense *in situ* hybridization probes were linearised with appropriate restriction enzymes (**Table 2.6**) and purified with QIAquick PCR Purification Kit (QIAGEN). 1µg purified linearised template was used to synthesize the riboprobes using the SP6, T3 or T7 mMessage mMachine kit (Ambion). A 20µl reaction mix contained 2µl DIG-dUTP NTP mix (Roche) and 0.5µl RNaseOUT (Invitrogen). Reactions were incubated for 2h at 37°C and the DNA template digested with DNAse I for 15min. at 37°C. 1µl 0.5M EDTA and 80µl H_2O were added and the probes were precipitated with 33µl

10M NH_4Ac and 250µl EtOH for 10min. on ice. Probe was pelleted, washed with 70% EtOH and resuspended in 50µl H_2O + 50µl $Hybe^{+/+}$. The quality was tested on an agarose gel. For *in situ* hybridization, the probes were diluted 1:100 in $Hybe^{+/+}$ and stored at -20°C.

Table 2.6. *In situ* hybridization probes used in this thesis.

Gene	Restriction site	RNA Polymerase	Reference
cmyb	EcoRI	T7	(Thompson et al. 1998)
dll4	SpeI	T7	(Monteiro et al. 2016)
fli1	XbaI	T3	(Brown et al. 2000)
foxn1	SalI	T7	(Wilkinson et al. 2009)
gata2a	EcoRI	T7	(Detrich et al. 1995)
gata2b	SacI	T7	(Monteiro et al. 2016)
gfi1aa	ClaI	T7	(Cooney et al. 2013)
gfp	SacII	SP6	This thesis
kdrl	EcoRI	T7	(Fouquet et al. 1997)
lmo4a	ApaI	SP6	R. Monteiro, unpublished
rag1	HindIII	T7	(Wilkinson et al. 2009)
runx1	HindIII	T7	(Kalev-Zylinska et al. 2002)

2.2. Maintenance and manipulation of zebrafish

2.2.1. Maintenance of zebrafish

Zebrafish (*Danio rerio*) were maintained in flowing system water at 28.5°C, conductance 450-550µS and pH 7.0±0.5 (Westerfield 2007). The daily cycle consisted of 14 hours of light and 10 hours of darkness. Fish were fed on a dry diet 3 times a day, supplemented with brine shrimp twice a day and with enriched brine shrimp 3 times a week (BMS Staff, University of Oxford). Fish suffering from infections or heart oedemas were culled according to Schedule 1 of the Animals (Scientific Procedures) Act 1986. All animal experiments were approved by the University of Oxford ethics committee.

2.2.2. Collection and staging of zebrafish embryos

Eggs were collected by natural mating and washed with system water. Embryos were grown at 24-32°C in E3 medium with methylene blue and staged according to morphological

features (Kimmel et al. 1995) corresponding to respective age in hours or days post fertilization (hpf or dpf, respectively).

2.2.3. Zebrafish lines

Published lines used in this thesis were wild type (wt^{KCL}), Tg(*kdrl*:GFP)^{s843} (Jin et al. 2005) and Tg(*kdrl*:GFP; *gata1*:DsRed)$^{s843/sd2}$ (Traver et al. 2003; Jin et al. 2005). Transgenic offspring were selected under a widefield fluorescent microscope.

2.2.4. Generation of new zebrafish lines

2.2.4.1. *Gata2a* i4 enhancer transgenic lines

One-cell zebrafish embryos were injected with 1nl of an injection mix, containing 50pg *gata2a*-i4-E1b-GFP-Tol2 construct DNA + 30pg *tol2* transposase mRNA (Kawakami et al. 2004). Transgenic founders (Tg(*gata2a*-i4:GFP)) were selected under a widefield fluorescent microscope and outbred to wt fish. Carriers of monoallelic insertions were detected by the Mendelian distribution of 50% fluorescent offspring coming from wt outcrosses. These transgenics were then inbred to homozygosity (**Fig. 2.1**).

2.2.4.2. *Gata2a* i4 enhancer mutant lines

One-cell zebrafish embryos were injected with 300pg Cas9(D10A) mRNA + 150pg each sgRNA (**Table 2.4**). Mutant founders ($gata2a^{+/\Delta i4}$) were identified by screening for the presence of a mutant PCR product in the genomic DNA of their offspring. The genomic region spanning the target sites was amplified with JumpStart™ REDTaq® ReadyMix™ PCR (see **Table 2.3** for primer sequences). Deletions were identified by the presence of a smaller PCR product on a 2% agarose gel. The mutated fragments were purified from gel, cloned into pGEM®-T Easy vector (Promega) and sequenced with M13 f/r primers. Selected $gata2a^{+/\Delta i4}$ founders were outcrossed and bred to heterozygosity and homozygosity for the selected mutant allele (**Fig. 2.2**).

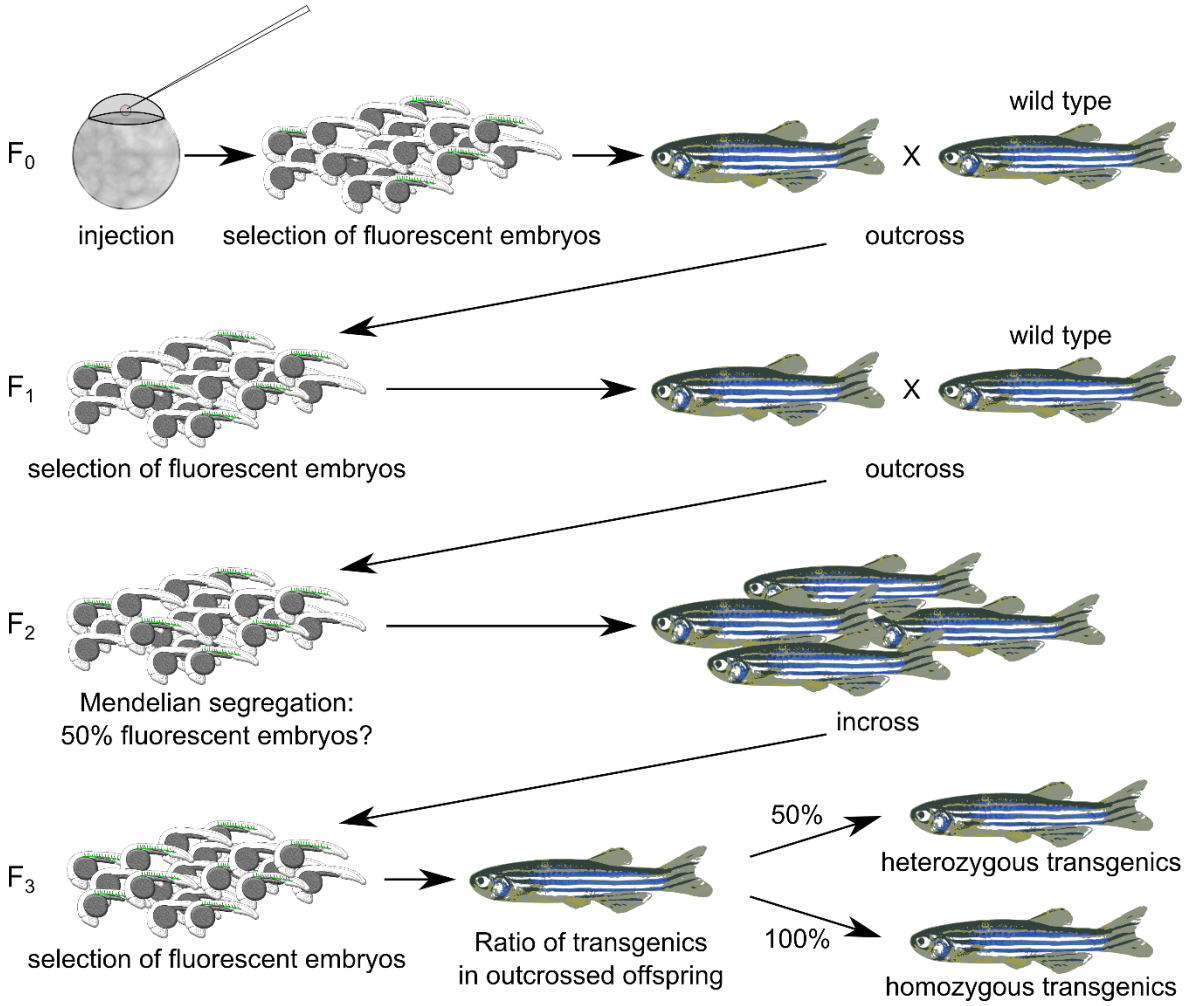


Figure 2.1. The pipeline to generate zebrafish transgenic lines.

One-cell embryos are injected with transgenesis reagents. Fluorescent embryos are selected under a fluorescent microscope. After reaching maturity, they are outcrossed and their fluorescent offspring is raised to adulthood. These F_1 fish are screened for monoallelic transgene insertions by checking for Mendelian distribution of 50% fluorescent offspring coming from outcrosses. The offspring coming from these founders are selected to grow to adulthood as monoallelic heterozygous F_2 generation. They are incrossed and their progeny is raised to maturity. The F_3 fish are tested for the ratio of transgenic embryos in their progeny after outcrossing to wild type. 50% fluorescent offspring comes from heterozygous transgenics, while 100% fluorescent offspring identifies homozygous transgenics. The cartoons are not to scale.

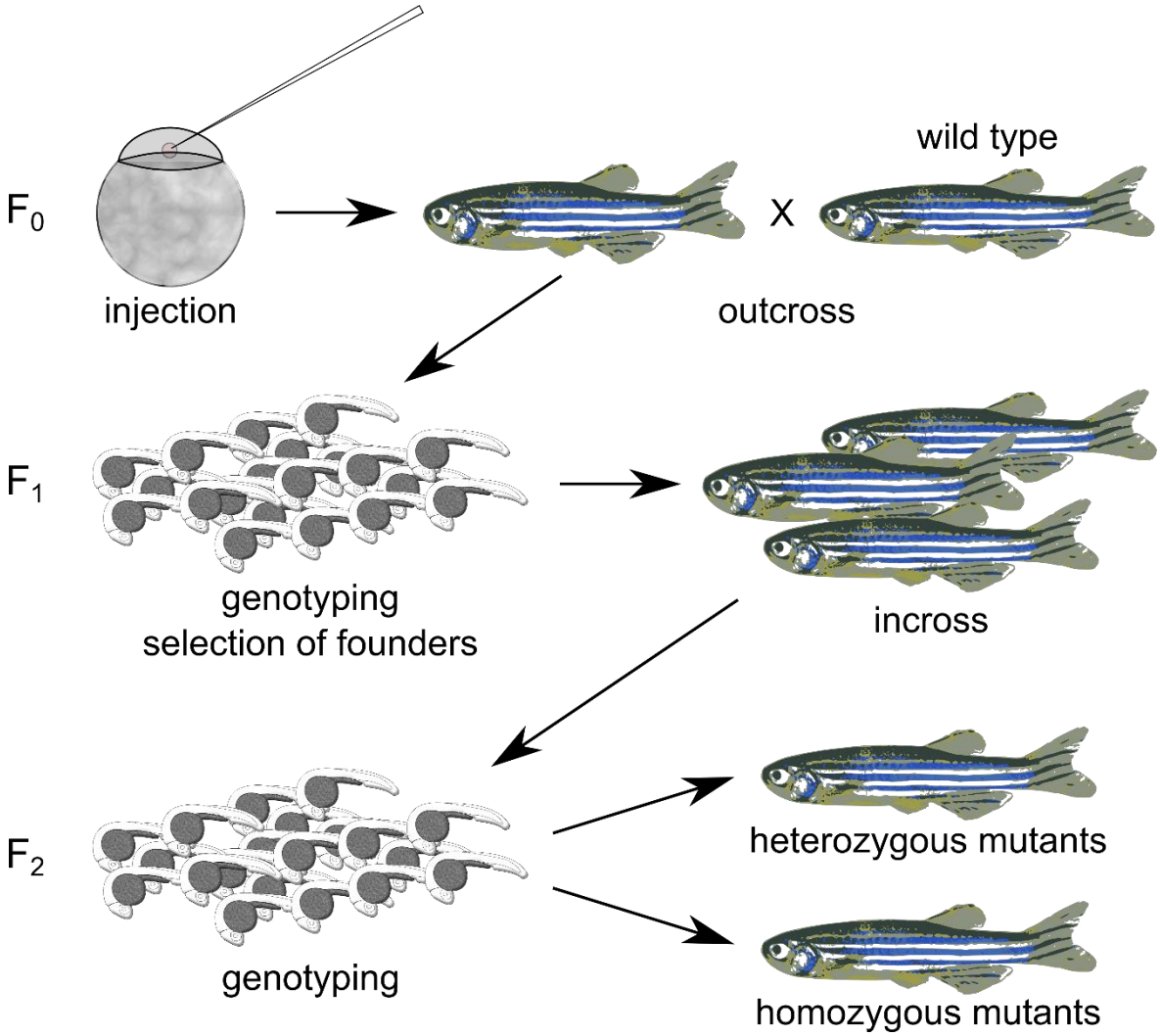


Figure 2.2. The pipeline to generate zebrafish mutant lines.

One-cell embryos are injected with CRISPR/Cas9 or TALEN reagents and raised to adulthood. After reaching maturity, they are outcrossed and their offspring screened for the presence of mutations. The progeny of identified founders is genotyped and raised to adulthood as heterozygous F_1 generation. These fish are incrossed and both heterozygous and homozygous mutants are selected by genotyping at embryonic stage. The cartoons are not to scale.

In order to generate transgenic-mutant lines, *gata2a*$^{+/\Delta i4}$ founders were crossed to Tg(*kdrl*:GFP)^{s843} or Tg(*gata2a* i4:GFP) founders.

2.2.4.3. *Lmo4a* mutant lines

One-cell zebrafish embryos were injected with 100pg left-arm + 100pg right-arm TALEN mRNAs (see section 2.3.9.3). Germline mutations in the founders (*lmo4a*$^{+/-}$) were identified using restriction fragment length polymorphism (RFLP) (Bedell et al. 2012) on the genomic DNA extracted from their offspring. The genomic region containing the target site was amplified with JumpStart™ REDTaq® ReadyMix™ PCR (see **Table 2.3** for primer sequences) and digested with SacII enzyme. Mutations were identified by the presence of an undigested PCR product on a 2% agarose gel. The undigested mutant fragments were purified from the gel, cloned into pGEM®-T Easy vector (Promega) and sequenced with M13 f/r primers. The *lmo4a*$^{+/-}$ founders were outcrossed and bred to heterozygosity and homozygosity for the selected mutant allele (**Fig. 2.2**).

2.2.4.4. Maintenance of new mutant lines

To maintain the *gata2a*$^{\Delta i4/\Delta i4}$ and *lmo4a*$^{+/-}$ lines, mutant offspring were selected by genotyping at 3-4dpf by fin clipping (Wilkinson et al. 2013). For this, the DNA was isolated using 75µl of HotSHOT Lysis Buffer and 75µl of HotSHOT Neutralization Buffer and lysis incubation time of 25-60 minutes (Truett et al. 2000). Genomic regions containing the mutated sites were amplified with Phire™ Green HotStart II PCR Master Mix using 5µl of DNA lysate in 20µl reaction volume. Mutations were identified as described above.

2.2.5. Microinjection of embryos

Embryos were aligned against a glass slide in a petri dish and injected with 1nl of the appropriate injection mix, prepared in H_2O with 0.05% Phenol Red (Sigma). DNA and RNA were injected into the cell of 1-cell stage embryos. MOs were injected into 1- to 2-cell embryos.

2.2.6. Chemical treatment of embryos

2x stocks of the chemical modulators of Wnt activity were freshly prepared in E3 and added to the equal amount of E3 containing the embryos. 16hpf embryos were treated in a 24-well plate with 0.5mM 6-bromoindirubin-3'-oxime (BIO) (Gore et al. 2011) or 10μM inhibitor of Wnt response (IWR-1) (Yin et al. 2011) diluted in DMSO. As controls, the embryos were treated in parallel with equivalent concentrations of DMSO. The embryos were incubated in the dark post-treatment.

2.2.7. Dechorionation and fixation of embryos

Embryos analysed between 18-somite and 3dpf stages were manually dechorionated with INOX #5 forceps. Embryos analysed at 4-5dpf hatched naturally. Embryos were fixed in PBS+PFA overnight at 4°C, dehydrated through increasing concentrations of EtOH in PBST and stored at -20°C in 100% EtOH. Embryos younger than the 18-somite stage were fixed in their chorions, washed with PBST, and dechorionated and stored as above.

2.2.8. Fluorescence-activated cell sorting (FACS)

20-130 embryos at the required stage were collected in Low Binding® SafeSeal® Microcentrifuge Tubes (Sorenson) and pre-homogenized by pipetting up and down in 500μl Deyolking Buffer. They were spun down for 1min. at 500g and incubated for 15min. at 30°C in Trypsin + Collagenase Solution. During that time, they were homogenized by pipetting up and down every 3min. The lysis was stopped by adding 50μl foetal bovine serum and 650μl filter-sterilized Hanks' solution. The cells were rinsed with 1ml Hanks' solution and passed through a 40μm cell strainer (Falcon®). They were resuspended in ~400μl Hanks' solution with 1:10,000 Hoechst 33258 (Molecular Probes®) and transferred to a 5ml polystyrene round bottom tube for FACS sorting. The cells were sorted on FACSAria Fusion sorter by Kevin Clark (MRC WIMM FACS Facility). The gates of GFP (488-530) and DsRed (561-582) channels were set with reference to samples derived from non-transgenic embryos. The fluorescence

readouts were compensated when necessary. For ATAC-seq library preparation, the cells were sorted into Hank's solution. For RNA isolation, the cells were sorted directly into RLT Plus buffer (QIAGEN) + 1% β-mercaptoethanol.

2.3. Preparation, manipulation and analysis of DNA

2.3.1. Transformation of competent *E. coli*

20µl of NEB 5-alpha Competent *E. coli* (High Efficiency) per transformation were thawed on ice for 10-20min. and incubated with DNA for 30 min. on ice. The cells were heat shocked for 50s at 42°C, cooled down for 1min. on ice and allowed to recover for 30-60min. shaking at 37 °C in 1ml SOC. The bacteria were centrifuged for 1min. at 13,000g and plated onto selective LB Agar plates, supplemented with 80µl of X-gal/IPTG solution for blue/white screening when required.

2.3.2. *E. coli* culture

Bacteria were cultured at 37°C as required either in shaking liquid culture (LB) or solid culture (streaking on LB Agar). As required, cultures were supplemented with 50µg/ml of Ampicillin, Kanamycin or Spectinomycin.

2.3.3. Isolation of zebrafish genomic DNA

Genomic DNA from whole clutches of embryos was extracted with the DNeasy® Blood & Tissue Kit (QIAGEN). For single live embryos, the DNA was isolated using 75µl HotSHOT Lysis Buffer, 75µl HotSHOT Neutralization Buffer and lysis incubation time of 25-60 minutes, according to the published protocol (Truett et al. 2000). For PFA-fixed embryos, genomic DNA was isolated after *in situ* hybridization using an optimised HotSHOT protocol (Dobrzycki et al. under revision – see Appendix IV). Briefly, 75µl HotSHOT Lysis Buffer was added directly to a PCR tube with a freshly-imaged embryo in <5µl 100% glycerol. The embryo was suspended

in the buffer and incubated at 95°C for 30min., then cooled to 4°C, after which 75µl HotSHOT Neutralization Buffer was added. The DNA was stored at -20°C.

2.3.4. Isolation and sequencing of plasmid DNA

Bacterial cells from liquid LB cultures were pelleted and plasmid DNA isolated with appropriate QIAGEN kit: Quicklyse® Miniprep Kit for sequencing, QIAprep® Spin Miniprep Kit for high-quality DNA for injections, or HiSpeed® Plasmid Midi Kit for higher concentration of high-quality DNA. DNA was eluted in EB and stored at -20°C. The concentration of the DNA was measured on a NanoDrop™ One spectrophotometer (Thermo Scientific). Plasmids were validated with a diagnostic restriction digest and sequenced by semi-automated Sanger sequencing (MRC WIMM Sequencing Service) with appropriate primers.

2.3.5. Polymerase chain reaction (PCR) and colony PCR

JumpStart™ REDTaq® ReadyMix™ PCR was used to amplify templates for *in situ* probes, genomic zebrafish regions for sequencing and for genotyping after *in situ* hybridization. Phire™ Green HotStart II PCR Master Mix was used for genotyping fresh embryos or fin clips. For preparation of DNA for injections or for sgRNA templates, NEB Phusion® kit was used. In all cases, manufacturer protocols were followed.

To screen bacterial colonies, individual colonies were picked and placed in 20µl H_2O. PCR was carried out as above with 3.6µl of the colony suspension. Remaining 15µl from positive colonies was used to inoculate liquid culture.

2.3.6. Agarose gel electrophoresis and gel purification

UltraPure™ Agarose (Invitrogen) was used to prepare agarose gels for purification (0.8% agarose in 0.5x TAE) or diagnostic visualisation (1-2.5%). The gels were supplemented with 0.5x GelRed™ Nucleic Acid Gel Stain (Biotium). PCR reactions containing loading dye in the Master Mix were loaded directly, other DNA samples and 1kb or 100bp DNA Ladder (NEB)

were mixed with Gel Loading Dye Purple (NEB) before loading. The gels were run at 100-200V in 0.5x TAE and imaged with trans-UV light using Bio-Rad Molecular Imager® Gel Doc™ XR+ System with Image Lab™ Software. For purification of DNA fragments, the bands were visualised through an orange filter under 470nm transillumination (Safe Imager, Invitrogen). The DNA was purified from the excised bands with the QIAEX® II Gel Extraction Kit (QIAGEN).

2.3.7. Restriction digestion, ligation and In-Fusion®

DNA was digested at for 1-12 h with NEB restriction enzymes in the provided buffers at required temperatures. Typically, 2 units of enzyme were used per 1µg of DNA. DNA fragments were ligated as required with 3:1 molar ratio of fragment to vector using T4 DNA ligase (Promega). For cloning purified PCR products into pGEM®-T Easy (Promega), 0.5µl of the vector was incubated with 3.5µl of the insert per manufacturer's instructions. In-Fusion® HD Cloning Kit (Takara) was used according to manufacturer's protocols.

2.3.8. Site-directed mutagenesis

NEB Phusion® kit was used for site-directed mutagenesis of the Addgene plasmid #51307 (Guo et al. 2014), using overlapping primers (**Table 2.3**). Standard Phusion® PCR programme was optimised to 20 cycles, annealing temperature of 60°C and 10min. extension time. The reaction was digested with DpnI for 1h at 37°C and transformed into competent *E. coli*. Isolated plasmids were verified by sequencing and stored at -20°C.

2.3.9. Construction of new plasmids

2.3.9.1. *Gata2a*-i4-GFP-Tol2 vectors

Genomic regions containing the identified 150bp-long *gata2a* i4 enhancer flanked by ±500bp (i4-1.1kb) or ±150bp (i4-450bp) were amplified from wild type zebrafish genomic DNA with NEB Phusion® polymerase (see **Table 2.3** for primer sequences) and cloned

upstream of E1b minimal promoter and GFP into a Tol2 recombination vector (Addgene plasmid #37845 from Ahituv lab (Birnbaum et al. 2012)) with Gateway® cloning technology (Life Technologies™) following the manufacturer's protocol. The resulting constructs were verified by Sanger sequencing.

2.3.9.2. *Gata2b* tissue-specific overexpression vectors

Gata2b coding region was amplified from wild type zebrafish cDNA with JumpStart™ REDTaq® ReadyMix™ PCR and cloned into pGEM®-T Easy vector (Promega). After sequencing, it was amplified with NEB Phusion® polymerase (see **Table 2.3** for primer sequences) and cloned with In-Fusion® HD Cloning Kit (Takara) into the *gata2a*-i4-GFP-Tol2 vector to replace the GFP. The resulting constructs were verified by Sanger sequencing.

2.3.9.3. *Lmo4a* TALENs

For TALENs design, Mojo Hand software was used (http://www.talendesign.org) (Neff et al. 2012). The identified target site was GGAAAGCTCCGCGGTT. The RVD-containing repeats were assembled using the Golden Gate approach in pTAL3-DDD and pTAL3-RRR vectors (Cermak et al. 2011). The resulting DNA templates were verified by sequencing and linearised with NotI enzyme. The mRNAs were transcribed from 1µg linearised template with SP6 mMessage mMachine® kit (Ambion) and purified with RNeasy® Micro Kit (QIAGEN).

2.3.9.4. DNA templates for new *in situ* hybridization probes

Gene-specific primer pairs (**Table 2.3**) were used to amplify the probe templates from wild type zebrafish cDNA with JumpStart™ REDTaq® ReadyMix™ PCR. The amplicons were gel purified and cloned into pGEM®-T Easy vector (Promega). The constructs were sequenced with M13 f/r primers.

2.3.10. ATAC-seq library preparation

Tg(*kdrl*:GFP)^{s843} embryos were dissociated for FACS at 26-27hpf to collect *kdrl*$^{+}$ and *kdrl*$^{-}$ cell populations (40,000-50,000 cells each). They were processed for ATAC library preparation using optimised standard protocol (Buenrostro et al. 2013). Briefly, after sorting into Hanks' solution, the cells were spun down at 500g at 4°C, washed with ice-cold PBS and resuspended in 50µl cold Lysis Buffer. The nuclei were pelleted for 10min. at 500g at 4°C and resuspended in the TD Buffer with Tn5 Transposase (Illumina), scaling the amounts of reagents accordingly to the number of sorted cells. The transposition reaction lasted 30min. at 37°C. The DNA was purified with PCR Purification MinElute Kit (QIAGEN). In parallel, transposase-untreated genomic DNA from *kdrl*$^{+}$ cells was purified with the DNeasy® Blood & Tissue Kit (QIAGEN). The samples were amplified with appropriate Customized Nextera primers (Buenrostro et al. 2013) in NEBNext High-Fidelity 2x PCR Master Mix (NEB). The libraries were purified with PCR Purification MinElute Kit (QIAGEN) and Agencourt AMPure XP beads (Beckmann Coulter). The quality of each library was verified using D1000 ScreenTape System (Agilent). Four biological replicas of the libraries were quantified with the KAPA Library Quantification Kit for Illumina® platforms (KAPA Biosystems). The libraries were pooled (including the Tn5-untreated control), diluted to 1ng/µl and sequenced using 75bp paired-end reads on Illumina HiSeq 4000 (Wellcome Trust Centre for Human Genetics, Oxford).

2.4. Preparation, manipulation and analysis of RNA

2.4.1. RNA extraction

To isolate RNA from whole embryos, 20-30 embryos were collected in Low Binding® SafeSeal® Microcentrifuge Tubes (Sorenson) and homogenized in 300µl TRIzol™ Reagent (Invitrogen) through a syringe and Agani™ 23Gx1" needle (Terumo®). The RNA was purified with Direct-zol™ RNA MiniPrep kit (Zymo Research). The eluate was treated for 15min. at

37°C with DNAse I and the RNA was precipitated with LiCl. The RNA was quantified and quality tested with NanoDrop™ One (Thermo Scientific).

For RNA extraction after FACS, the cells sorted into RLT Plus buffer (QIAGEN) + 1% β-mercaptoethanol were processed with the RNEasy® Micro Plus kit (QIAGEN), according to the accompanying protocol. The RNA was quantified and its quality assessed with the use of Agilent RNA 6000 Pico kit. All RNA samples were stored at -80°C.

2.4.2. RNA-seq sample preparation and sequencing

Four biological replicas of 3ng RNA isolated from *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells sorted from *lmo4a* morphants and uninjected siblings were amplified with SMARTer kit and sequenced using 75bp paired-end reads on Illumina HiSeq 4000 (Wellcome Trust Centre for Human Genetics, Oxford).

2.4.3. Reverse transcription of RNA

For SYBR® Green qRT-PCR, 150ng RNA from whole embryos or 8ng RNA from FACS-sorted cells were used for cDNA synthesis with SuperScript IV RT (Invitrogen) per manufacturer's protocol. The cDNA was then diluted 1:4-1:8 in H_2O and stored at -20°C.

For Fluidigm Biomark™ qRT-PCR, 1ng RNA from FACS-sorted cells was used for Specific Target Amplification in a 10µl reaction with the following reagents: 5µl 2xBuffer and 1.2µl enzyme mix from SuperScript III One-Step Kit (Thermo Fisher Scientific), 0.1µl SUPERase• In™ RNase Inhibitor (Ambion), 1.2µl TE buffer (Invitrogen), 2.5µl 0.2x TaqMan® assay mix. The reaction was incubated for 15min. at 50°C, for 2min. at 95°C and amplified for 20 cycles of 15s at 95°C/4min. at 60°C. The cDNA was diluted 1:5 in TE buffer and stored at -20°C.

2.4.4. Real-time quantitative PCR (RT-qPCR)

2.4.4.1. SYBR® Green qRT-PCR

3µl of the cDNA diluted in H_2O were used for technical triplicate qRT-PCR reactions of 20µl containing the Fast SYBR® Green Master Mix (Thermo Fisher Scientific) and appropriate primer pair (see **Table 2.3** for the sequences). The reactions were run on 7500 Fast Real-Time PCR System (Applied Biosystems) and the results were analysed with the accompanying software. No-template controls were run on each plate for each primer pair. Each reaction was validated with the melt curve analysis. The baseline values were calculated automatically for each reaction. The threshold values were manually set to be equal for all the reactions run on one plate, within the linear phase of exponential amplification. The relative mRNA levels in each sample were calculated by subtracting the geometric mean of Ct values for housekeeping genes *eef1a1l1* and *ubc* from the average Ct values of the technical triplicates for each gene of interest. This value (ΔCt) was then converted to a ratio relative to the housekeeping genes with the formula $2^{-\Delta Ct}$.

2.4.4.2. Fluidigm Biomark™ qRT-PCR

Diluted cDNA was used for qRT-PCR according to the Fluidigm protocol for Gene Expression with the 48.48 IFC Using Standard TaqMan® Assays (**Table 2.5**). Each sample was run in 3-6 biological replicates. The collected data were analysed with Fluidigm Real-Time PCR Analysis software (version 4.1.3). The baseline was automatically corrected using the built-in Linear Baseline Correction. The thresholds were manually adjusted for each gene to fall within the linear phase of exponential amplification, after which they were set to equal values for the housekeeping genes: *rplp0*, *rpl13a*, *cops2* (Xu et al. 2016), *lsm12b* (Hu et al. 2016) and *eef1a1l1*. The relative mRNA levels for each sample were calculated by subtracting the geometric mean of Ct values for the housekeeping genes from the Ct value for each gene of interest. This value (ΔCt) was then converted to a ratio relative to the housekeeping genes with the formula $2^{-\Delta Ct}$.

2.4.5. Whole mount *in situ* hybridization (ISH) and immunohistochemistry

In situ hybridization (ISH) was carried out according to the standard lab protocol (Jowett and Yan 1996) using digoxygenin-labeled probes (**Table 2.6**). After rehydration of fixed embryos through dilution series of EtOH in PBST, embryos older than 26hpf were treated with 10µg/ml proteinase K (Sigma) and fixed for 20min. at room temperature (RT) in PBS+PFA. The samples were hybridized overnight at 65°C, taken through a series of washes into 0.2xSSC and then into MABT. They were blocked in MAB Block for 1h at RT, incubated with Anti-digoxygenin-AP Fab Fragments (Roche) for 10h at 4°C and washed with MABT.

For chromogenic *in situ* hybridization, the embryos were stained with BM Purple AP Substrate (Roche) mixed 1:1 with AP Buffer, washed with PBST (PBS+0.05% Tween-20) and fixed in PBS+PFA. Post hybridization, embryos older than 23hpf were bleached in 5% formamide/0.5% SSC/10% H_2O_2 (Monteiro et al. 2011).

For fluorescent *in situ* hybridization, the colour was developed using SIGMAFAST Fast red/TR Naphthol (Sigma) in AP Buffer, washed with PBST and processed for immunohistochemistry by Monika Krecsmarik. The embryos were blocked in Blocking Buffer for 1h at RT. They were probed with 1:500 rabbit anti-GFP antibody (Molecular Probes®) in Blocking Buffer overnight at 4°C. Goat anti-rabbit secondary antibody coupled to AlexaFluor488 (Invitrogen) was used at 1:500 dilution for 3h at RT. Hoechst 33342 was used as a nuclear counterstain.

2.5. Imaging zebrafish embryos

2.5.1. Bright-field microscopy

After chromogenic ISH, the embryos were imaged in 100% glycerol on glass microscope slides with cavities (BRAND®) with QImaging MicroPublisher 5.0 RTV Camera and Q-Capture

Pro 7™ software (version 7.0.3). When not all embryos from a sample were imaged, representative 'high' or 'low' expression images were taken.

2.5.2. Confocal microscopy

Fluorescent images of immunostained embryos after fluorescent ISH were taken by Monika Krecsmarik on a Zeiss LSM880 confocal microscope and processed using the ZEN software (Zeiss). Cells co-expressing GFP and either *runx1* or *gata2a* in the floor of the dorsal aorta were counted manually on optical sections using the ZEN software (Zeiss).

2.5.3. Live imaging

Manually dechorionated embryos were grown in PTU to prevent pigment development. They were anaesthesized with MS222 and imaged in 3% methylcellulose on glass microscope slides with cavities (BRAND®) with the SteREO Lumar.V12 widefield fluorescent system (Zeiss) or with the ORCA-Flash 4.0 Digital Camera C11440 (Hamamatsu) attached to the Olympus MVX10 microscope, using Hamamatsu HCImageLive software (version 4.3.1.42). For higher magnification and resolution, the embryos were placed in 1% low-melting-point agarose (Sigma) on 35mm glass bottom dishes (MatTek) and imaged with the LSM780 inverted confocal microscope using the ZEN software (Zeiss).

2.6. Bioinformatic and statistical tools

2.6.1. Genomic data

Zebrafish genomic data were obtained from the Zebrafish Model Organism Database (ZFIN) (http://zfin.org/) (Howe et al. 2013a). The graphical representations were created using FancyGene v1.4 (Rambaldi and Ciccarelli 2009). Annotation of putative conserved binding sites within the i4 region and its alignment to other species were done by Sarah de Val using the BLAT tool of UCSC Genome Browser (Kent et al. 2002) and ClustalW2 (Larkin et al. 2007).

2.6.2. Quality check and mapping of high-throughput sequencing data

Raw sequenced reads were checked for base qualities, trimmed where 20% of the bases were below quality score 20, and filtered to exclude adapters using *Trimmomatic* (Version 0.32) and mapped to Zv9 reference genome (comprising 14,612 genes) (Howe et al. 2013b) by Rossella Rispoli using *STAR* with default parameters. The results were visualised using UCSC Genome Browser (http://genome-euro.ucsc.edu/) (Kent et al. 2002).

2.6.3. ATAC-seq differential peak analysis

The eight data sets were analysed with Principal Component Analysis (PCA) to identify outliers. Correlation among *kdrl*$^{+}$ and *kdrl*$^{-}$ samples was assessed with a tree map. The peaks were called for each sample using the Tn5-untreated control as input. DiffBind (EdgeR method) was used to identify differential peaks between *kdrl*$^{+}$ and *kdrl*$^{-}$ samples. The threshold for differential peaks was $p<0.05$. This analysis was done by Rossella Rispoli.

2.6.4. RNA-seq differential gene expression analysis

Counts corresponding to each gene were converted to reads per kilobase per million (RPKM) using *Subread tool*: featureCounts method (version 1.4.5-p1). The eight data sets were analysed with Principal Component Analysis (PCA) to identify outliers. EdgeR (Bioconductor package) was used for differential gene expression analysis between paired control and *lmo4a* knockdown samples, filtering out genes with >3 replicates of RPKM=0, applying generalized linear model analysis as follows: estimate 'Dispersion', fitted with 'negative binomial model' and estimate 'Generalized linear model likelihood ratio'. The threshold for differentially expressed genes was False Discovery Rate (FDR) <0.05. The above analysis was performed by Rossella Rispoli. Expressed genes (10,455) were arbitrarily defined as having average RPKM>0.5 in *lmo4a* MO-treated samples and average RPKM>0.5 in control samples. Gene symbols were converted to ZFIN ID numbers using db2db tool on bioDBnet (Mudunuri et al. 2009) or manually using ZFIN database (Howe et al. 2013a) entries for the genes of interest. The lists of ZFIN IDs were

analysed for the enrichment of gene ontology (GO) terms using the DAVID Bioinformatics Resources 6.8 (Huang et al. 2008; Huang et al. 2009).

2.6.5. *In silico* DNA analysis

Vector maps were prepared with SnapGene® (GSL Biotech LLC), which was also used to align and analyse the Sanger sequencing files (.abi). For newly generated zebrafish mutant lines, the sequences were first analysed with SnapGene®, then aligned with ClustalW2 (Larkin et al. 2007) and annotated with Jalview 2.8.1 (Waterhouse et al. 2009).

2.6.6. *In situ* hybridization image quantification

Pixel intensity values were derived as described (Dobrzycki et al. under revision – see Appendix IV). Using Fiji software (Schindelin et al. 2012), the images were inverted to negative and converted to 8-bit grayscale. A Region of Interest (ROI) containing the ISH expression signal in the dorsal aorta along the yolk sac extension was drawn manually for each embryo. Then a second ROI with the same shape and area was created in a region of the embryo that had a uniform intensity and did not contain any ISH staining. In this particular instance, this region was placed just above the notochord (**Fig. 2.3**). This area was used to define the background. A value for each region was then determined by measuring the average pixel intensity. After subtracting the value of the background region from the value of the stained region, the pixel intensity of the ISH signal was assigned to each embryo.

2.6.7. Statistical analysis

2.6.7.1. *In situ* hybridization analysis

The numbers of embryos scored as 'high' or 'low' were tested for equal distribution between morphants and uninjected controls or among wild type, heterozygous and mutant genotypes with contingency Chi-squared tests, applying Continuity Correction for 2x2 tables, using IBM® SPSS® Statistics (version 22).

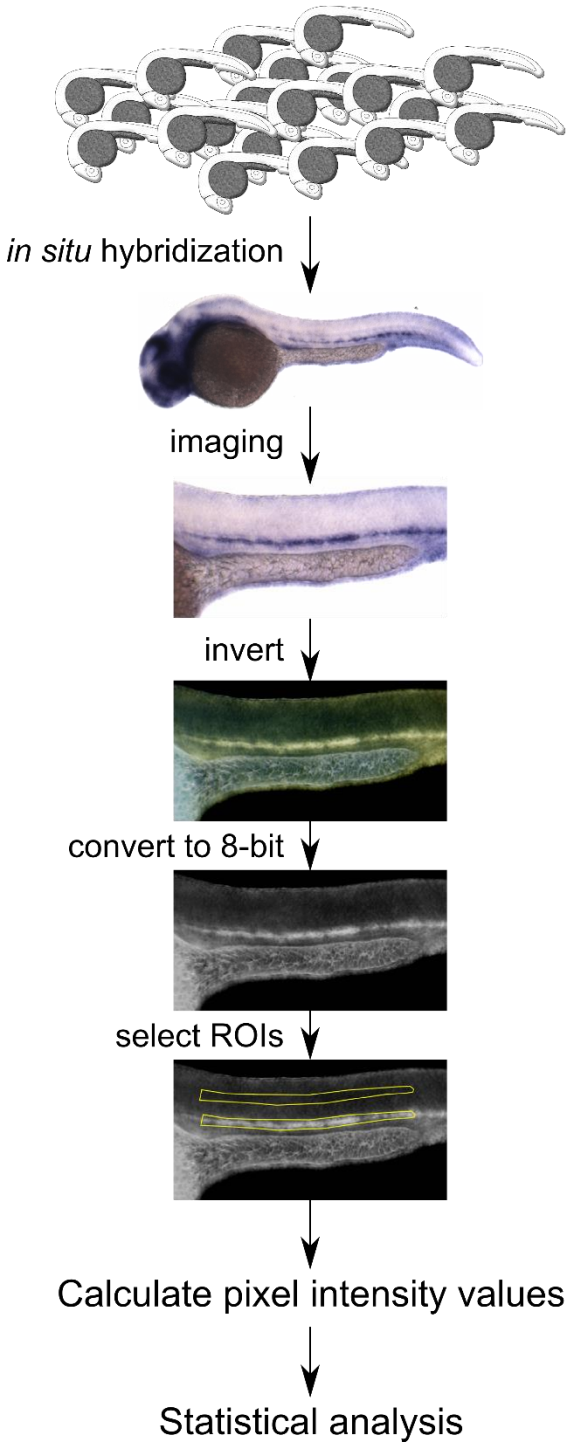


Figure 2.3. An overview of the method to quantify *in situ* hybridization (ISH) signal.

Collected embryos are probed for the measured gene with a standard ISH protocol. After imaging in 100% glycerol, the images for each embryo are inverted and converted to 8-bit greyscale. Regions of interests (ROIs) containing the ISH signal and background are manually selected and measured. The measurements, assigned to corresponding genotypes or treatments, are statistically analysed.

For digitally quantified images, the pixel intensity values were assessed for normal distribution with a Q-Q plot and transformed when necessary. Mean values (μ) of each experimental group were analysed with 2-tailed independent-samples *t*-tests or with ANOVA with 95% confidence levels, testing for the equality of variances with a Levene's test and applying the Welch correction when necessary. For ANOVA, differences between each two groups were assessed with either Tukey's post-hoc test (for equal variances) or with Games-Howell test (for unequal variances). For all these analyses, the IBM® SPSS® Statistics (version 22) package was used. The post-hoc power of the tests and required sample sizes were determined with G*Power software (version 3.0.10) (Faul et al. 2007).

2.6.7.2. qRT-PCR analysis

For SYBR® Green qRT-PCR, the ΔCt values were analysed with 2-tailed paired-samples *t*-tests with 95% confidence levels. For Fluidigm Biomark™ qRT-PCR comparing *gata2a*$^{\Delta i4/\Delta i4}$ and wt samples, the ΔCt values were analysed with 2-tailed independent-samples *t*-tests with 95% confidence levels. To compare the enrichment of gene expression in *kdrl*:GFP^{+} and *gata2a*-i4:GFP^{+} cells, the ratios of $2^{-\Delta Ct}$ values for respective populations of GFP^{+} and GFP^{-} cells were calculated $\left(\frac{2^{-\Delta Ct(GFP^{+})}}{2^{-\Delta Ct(GFP^{-})}}\right)$ and analysed using ANOVA with 95% confidence levels, testing for the equality of variances with a Levene's test and applying the Welch correction when necessary. Differences between each two groups were assessed with either Tukey's post-hoc test (for equal variances) or with Games-Howell test (for unequal variances). For all the above analyses, IBM® SPSS® Statistics (version 22) was used. For Fluidigm Biomark™ qRT-PCR on *lmo4a* morphant and control samples, the ΔCt values were analysed with 2-tailed paired-samples *t*-tests with Bonferroni correction in collaboration with Emmanouela Repapi (Computational Biology Research Group, MRC WIMM).

2.6.8. Preparation of figures

Confocal image stacks and *in situ* hybridization images were processed with Fiji software (Schindelin et al. 2012) and further prepared with GIMP 2.8.14. The graphs and heat maps representing the qRT-PCR expression levels and counts of co-expressing cells from confocal images were plotted using IBM® SPSS® Statistics (version 22). The graphs presenting individual data points, means and ±SD from the image quantification experiments were plotted using GraphPad Prism 7. The graph presenting incidence of infections and heart oedemas was plotted using GraphPad Prism 7. The graphs representing the fold change of genes belonging to each GO term in the RNA-seq experiment were plotted with GraphPad Prism 7. PCA and tree plots representing the quality of RNA-seq and ATAC-seq experiments as well as DE analysis plots were kindly provided by Rossella Rispoli. Final figures were assembled in Inkscape 0.48.

Chapter 3

Identification and characterisation of a conserved endothelial enhancer of *gata2a*

3.1. Introduction

Generation of specialised tissues during embryonic development depends on activation of genetic programmes in a localised manner. Spatial and temporal restriction of gene expression allows orchestrated differentiation of cellular diversity. Unsurprisingly, whole-transcriptome profiling in different tissues during development revealed highly dynamic gene expression patterns. Examples include the development of liver (Si-Tayeb et al. 2010), brain (Kang et al. 2011), spleen (Li et al. 2017), or, beyond animal biology, tomato fruit (Pattison et al. 2015). Moreover, single-cell sequencing of very early mouse embryos revealed gene expression changes responsible for generation of different cell types present already soon after fertilisation (Guo et al. 2010). Another study profiled whole-transcriptome expression data across 60 developmental stages of zebrafish (from embryos to ageing fish) (Domazet-Lošo and Tautz 2010), revealing the highly dynamic nature of gene expression.

Such molecular patterns largely depend on differential transcriptional activation of lineage-specific genes. In general, the transcription initiates in response to the action of *cis* and *trans* regulators. The latter mainly refer to transcription factors, while regulation in *cis* encompasses DNA regulatory elements, such as promoters or enhancers. Enhancers have recently gained a lot of attention, although the term 'enhancer' was already coined decades ago to describe DNA sequences located several hundreds of base pairs away from the promoter region that acted in *cis* to increase the levels of transcription from coding regions (Banerji et al. 1981). A more updated definition describes enhancers as regulatory DNA elements that bring together DNA-bound transcription factors to initiate gene expression. Through DNA looping, transcription factors bound to enhancers are brought towards promoter regions, which can be far away, to initiate expression of target genes (Levine 2010). In the pre-genomic era, identification and validation of enhancers was very tedious and difficult. More recently developed high-throughput methods have largely increased the speed of this process, allowing

much better understanding of the role of the distal regulatory regions. It is currently believed that throughout evolution, the appearance of long-range enhancers allowed the metazoans to develop complex genetic controls (Levine 2010). Moreover, differences in complexity have likely arisen from modifications to regulatory mechanisms rather than increased number of genes (Jones et al. 2012; Dailey 2015). By activating the same genes in various developmental contexts, enhancers provide a mode to achieve correct, precise spatial and temporal expression patterns of developmental genes (Visel et al. 2009; Levine 2010).

Examples of individual developmental enhancers, critical for correct localized gene expression during development, have been described in the frog *Xenopus laevis*, sea squirt *Ciona intestinalis*, fruit fly *Drosophila melanogaster* or nematode *Caenorhabditis elegans* (Levine 2010). There are also well defined examples of developmental genes regulated by several enhancers, such as *eve* in the fly or *Hoxb1* in the mouse (Levine 2010). One of the most striking examples of changes in regulatory DNA having impact on morphology is the evolutionary loss of pelvic fins in sticklebacks due to changes in *cis*-regulatory regions of the *Pitx1* gene (Shapiro et al. 2004; Chan et al. 2010).

Transcriptional activity at enhancer regions depends on dynamic remodelling of chromatin structure, which allows or prohibits binding of transcription factors (Mellor 2005). Consequently, one of the defining features of active DNA regulatory elements is chromatin accessibility (Encode Consortium 2012). Moreover, active enhancer regions are characterised by epigenetic chromatin marks. In fact, developmental enhancers have a specific chromatin signature, including such histone modifications as H3K4me1 and H3K27Ac (Creyghton et al. 2010; Rada-Iglesias et al. 2011), identifiable experimentally by chromatin immunoprecipitation (ChIP). Together with other features, such as binding of P300 protein (Visel et al. 2009; Blow et al. 2010), they are used as a proxy for defining enhancer elements. However, whether they are a cause or a consequence of the enhancer activity remains to be resolved (Dailey 2015). In

any case, assaying for these marks has allowed prediction of hundreds of thousands of human enhancers (Dailey 2015). More recent opinions suggest this number to be in the order of millions (Lopes et al. 2016).

A recent study employed ChIP combined with high-throughput sequencing (ChIP-seq) to investigate changes in acetylation of H3K27 across seven developmental stages in the forebrain, heart and liver of a developing mouse embryo (Nord et al. 2013). With this, the authors aimed to determine how much difference there was among the usage of different enhancers throughout different developmental stages in embryonic tissues. The analysis, comprising nearly 90,000 potential developmental enhancers, revealed a striking difference between proximal (promoter-associated) and distal (non-promoter-associated) H3K27Ac$^+$ chromatin regions. While proximal active chromatin loci were commonly identified across all tissues, distal regulatory regions exhibited mostly tissue-specific histone acetylation. Interestingly, activity of the enhancers appeared to change progressively over time and varied between tissues. In fact, most enhancers associated with tissue-specific genes possessed active marks during only a subset of examined developmental stages. Nearly half of all distal enhancers were restricted in their activity to only one or two time points. In further validation experiments, LacZ reporter constructs were used to test the activity of 32 putative enhancers in mouse embryos *in vivo*. 70% of them remained in agreement with H3K27Ac-predicted activities. Collectively, these data strongly demonstrate that distal enhancers play crucial roles in driving spatial and temporal specificity of gene expression during development. A similar analysis done on an *in vitro* model of human haematopoietic and erythroid development confirmed this dynamic nature and the critical role of distal *cis* regulatory elements in determining cell fate (Huang et al. 2016). Therefore, because the enhancer landscape changes extensively and rapidly during development, understanding the distal regulatory regions is necessary to dissect the mechanisms orchestrating embryonic patterning. Moreover, thanks to their high specificity, the

distal enhancers provide a handle to study temporally and spatially restricted expression of developmental genes.

In the developing haematopoietic system, the first enhancer to be linked to the generation of HSPCs was the mouse 3' enhancer of *Scl/Tal1* (Sánchez et al. 1999; Göttgens et al. 2002). Its sequence was then used to map and validate further enhancers of haematopoietic genes, such as *Fli-1* (Donaldson et al. 2005). However, only the discovery of an intronic *Runx1* enhancer paved the way for employing *cis* regulatory elements to their full potential. This highly conserved region, termed +23.5 enhancer, was identified computationally by sequence homology independently by two groups (Nottingham et al. 2007; Ng et al. 2010). It was further characterised by open chromatin in a DNAseI hypersensitivity assay (Nottingham et al. 2007) and by methylation of H3K4 (Ng et al. 2010). During the *in vivo* validation, the identified 531bp fragment drove LacZ activity at haematopoietic sites, but not in other endothelial or mesenchymal cells that express *Runx1* (Nottingham et al. 2007). Subsequently generated transgenic mice expressing GFP under the control of this enhancer (23GFP) confirmed this highly specific expression pattern in the subset of embryonic endothelium (Bee et al. 2010; Ng et al. 2010), also providing a tool to characterise $23GFP^+$ cell populations with FACS. Profiling of these cells revealed their haematopoietic nature, supporting the new role of the 23GFP transgene as a handle to isolate highly enriched haemogenic populations from mouse embryos (Ng et al. 2010; Swiers et al. 2013a). Moreover, single cell analyses revealed a unique profile of $23GFP^+$ endothelium, suggesting that it formed a separate population during the maturation of HE and HSC specification (Swiers et al. 2013a). This was the first example of using a transgenic reporter driven by an enhancer to precisely trace the origins of HSCs and dissect their stepwise development. Strikingly, although no orthologous enhancer was identified in the zebrafish *runx1* locus, the mouse +23.5 529bp region was sufficient to drive fluorescent transgene expression in zebrafish embryos at haematopoietic sites (Ng et al. 2010), suggesting a highly conserved nature of the *cis* controls during HSC programming.

In the case of *Gata2*, several putative enhancers were identified in its locus (Grass et al. 2006), but only one was found critical for the *Gata2* activity in the HE (Gao et al. 2013; Sanalkumar et al. 2014). This enhancer, located within intron 4, was termed the VE enhancer (Khandekar et al. 2007) or +9.5 element (Johnson et al. 2012; Gao et al. 2013). It was found to be essential for expression of haematopoietic genes, including *Runx1*, and the generation of HSPCs in the mouse embryo (Gao et al. 2013). In transgenic assays, this region drove expression of LacZ in all endothelial cells of early mouse embryos (Khandekar et al. 2007). Furthermore, specific deletion of *Gata2* within its domain of expression impaired HSPC generation, strongly supporting a role of this element in HE programming (Lim et al. 2012). The importance of this *cis* element led to a suggestion for existence of an 'HSC cistrome', a genome-wide ensemble of similar DNA elements collectively required for the genetic specification and differentiation of HSCs (Gao et al. 2013; Hewitt et al. 2015). Moreover, clinical studies demonstrated that mutations in this region in humans confer increased susceptibility to cancer or immunodeficiencies (Johnson et al. 2012). However, further evolutionary conservation of this well-defined enhancer has not been investigated. We hypothesised that a homologous region in zebrafish could provide a useful tool to monitor the haematopoietic function of *gata2* in live embryos.

This chapter describes the identification of a conserved distal regulatory region within the locus of zebrafish *gata2a*, which we termed the i4 enhancer. This region is characterized by chromatin accessibility specifically in the endothelial cells of the embryo before the generation of HSCs. Presented functional analysis reveals that this enhancer drives expression of a GFP transgene in embryos in the precursors of HSCs. Its activity is dynamic and transgenic *gata2a*-i4:GFP fish provide a useful novel tool to monitor the haematopoietic activity of *gata2a*.

3.2. A highly conserved region within the zebrafish *gata2a* locus is accessible to transposase specifically in the endothelium

Distal regulatory elements can be identified with various methods. Besides the already described ChIP-based assays, several approaches, such as DNAse-seq (Song and Crawford 2010), MNase-seq (Valouev et al. 2011) or FAIRE-seq (Simon et al. 2012), rely on detecting the open chromatin present at the active enhancer regions. However, these protocols contain multiple loss-prone steps. Consequently, they are limited by requiring very high amounts of starting material, making them practically non-applicable in *in vivo* embryonic systems without *ex vivo* expansion. The recently developed Assay for Transposase-Accessible Chromatin (ATAC-seq) (Buenrostro et al. 2013) allows mapping of open chromatin regions with reduced amount of starting material, using even as few as 500 cells. It was first done in a B-cell line, but also on patient-derived T-cells in a clinical setting. From a single blood draw, $CD4^+$ and $CD8^+$ T-cells were separately analysed (Buenrostro et al. 2013), confirming FACS-sorted cells as a useful source for ATAC-seq profiling.

To identify open chromatin regions specific to endothelial cells in zebrafish, we employed the Tg(*kdrl*:GFP) transgenic line, in which GFP marks all endothelium (Jin et al. 2005). We used FACS to sort *kdrl*:GFP^+ and *kdrl*:GFP^- cells from 26hpf embryos for generation of ATAC-seq libraries (**Fig. 3.1A**). Principal Component Analysis of DNA sequencing reads generated from 2 *kdrl*:GFP^+ replicas and 4 *kdrl*:GFP^- replicas revealed strong differences between the two cell populations (**Fig. 3.1B**). A correlation analysis further supported this observation, validating the high quality of the sequenced libraries (**Fig. 3.1C**). For further analysis, we focused on the loci of both zebrafish *gata2* homologues, expecting to identify open chromatin regions specific to the *kdrl*:GFP^+ population that would be responsible for their endothelial expression.

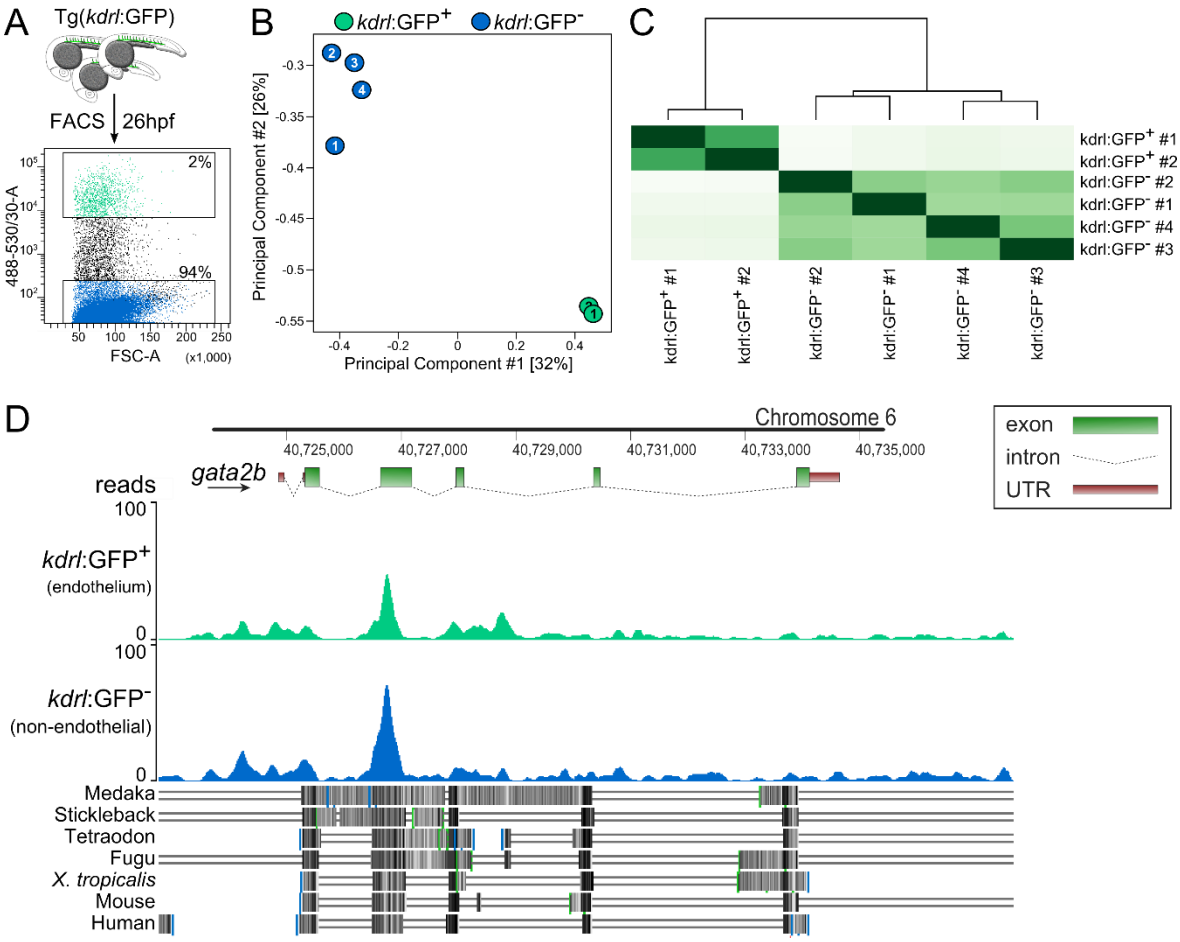


Figure 3.1. Analysis of *gata2b* locus with ATAC-seq does not reveal endothelial enhancers.

(A) *Kdrl*:GFP+ (sea green) and *kdrl*:GFP- (blue) cells were FACS-sorted from 26hpf embryos and used for preparation of ATAC-seq libraries. (B) Principal Component Analysis of 2 *kdrl*:GFP+ replicas and 4 *kdrl*:GFP- replicas, showing strong separation of the two cell populations. The same numbers inside the circles denote replicas coming from the same FACS. (C) Clustering of the ATAC-seq samples showing high correlation between *kdrl*:GFP+ replicas and among *kdrl*:GFP- replicas. The same numbers represent replicas coming from the same FACS. (D) Stacked means of two biological replicates (#1 and #2) mapped to Zv9 genome. Differential peak analysis did not identify any chromatin regions in the locus of *gata2b* that were significantly more open in the *kdrl*:GFP+ population. Data for panels B-D were provided by Rossella Rispoli.

Notably, we observed several peaks overlapping the exons of *gata2a* and *gata2b* (**Fig. 3.1, 3.2**). A non-biased analysis of enhancer regions with ChIP-seq against H3K4me1, H3K27Ac and P300 revealed that 7% of all sequencing peaks overlapped coding exons (Birnbaum et al. 2012). However, they were demonstrated to frequently regulate neighbouring genes, rather than the genes they were located in. Although these regions are potentially interesting and worth investigating in future studies, we decided to focus on the peaks outside of the coding regions.

Differential peak analysis identified four differentially open sites within the 20kb genomic region spanning the *gata2a* locus (**Fig. 3.2**). Three of them were located within the introns. All these putative endothelial enhancers showed a high level of sequence conservation across other fish species, but only the ~1kb region within the fourth intron (termed i4) was conserved across other vertebrates, including mouse and human. In fact, this region is homologous to the crucial Gata2 enhancer identified previously in the mouse (Khandekar et al. 2007; Gao et al. 2013) and human (Johnson et al. 2012).

Surprisingly, no differential peaks were identified in the *gata2b* locus (**Fig. 3.1**), although expression of *gata2b* is highly restricted to haemogenic endothelium (Butko et al. 2015). However, it is possible that other endothelial cells contributing to the whole *kdrl*$^{+}$ population masked the identification of haemogenic enhancers. Because of the lack of identified *gata2b* putative enhancers, we focused on the *gata2a*-i4 region in our further analysis.

3.3. The *gata2a*-i4 enhancer contains conserved putative transcription factor binding sites

One characteristic of enhancers is the presence of specific sequences (motifs) recognized by transcription factors from different families (Levine 2010). Such modules of transcription factor binding sites are believed to allow synergistic binding to determine enhancer-driven

transcription (Istrail and Davidson 2005). To further investigate the sequence of the identified *gata2a*-i4 site as a potential enhancer, transcription factor binding sites were computationally mapped. The highly conserved 150bp-long element contains several putative transcription factor motifs showing particularly high level of conservation across vertebrates, including consensus target sites for the Gata and Ets families (**Fig. 3.2**). These families of transcription factors play important roles in developmental haematopoiesis (Wozniak et al. 2007; Ciau-Uitz et al. 2013b) and their binding motifs are often found in close association (Liu et al. 2015). Moreover, Gata and Ets binding sites were previously shown to be critical for the function of the *Runx1* +23.5 enhancer (Nottingham et al. 2007).

Previous studies on the orthologous *Gata2* enhancer in mouse have revealed that the Ets sites are only responsible for fine-tuning the expression levels, while the E-box region is essential for the *Gata2* activity in embryonic endothelium (Khandekar et al. 2007). Although the positioning of the E-box site relative to the adjacent GATA site differs in zebrafish and mammals (**Fig. 3.2**), the necessary spacer distance of ~9bp between the two *cis* sites (Wozniak et al. 2008) is conserved across the vertebrate genomes analysed. Moreover, for developmental enhancers, a similar but not identical arrangement of target binding sites within the enhancer regions may have arisen due to convergent evolution to produce the same function (Levine 2010). Therefore, this site may be a crucial target for transcriptional complexes containing an E-box-binding factor and a GATA family member not only in the mouse (Sanalkumar et al. 2014), but also in the zebrafish.

In summary, the intronic enhancer (i4) identified in the zebrafish *gata2a* locus contains highly conserved binding sites for key haematopoietic transcription factors, suggesting that genetic regulation of *gata2a* expression in the HE is conserved across vertebrates.

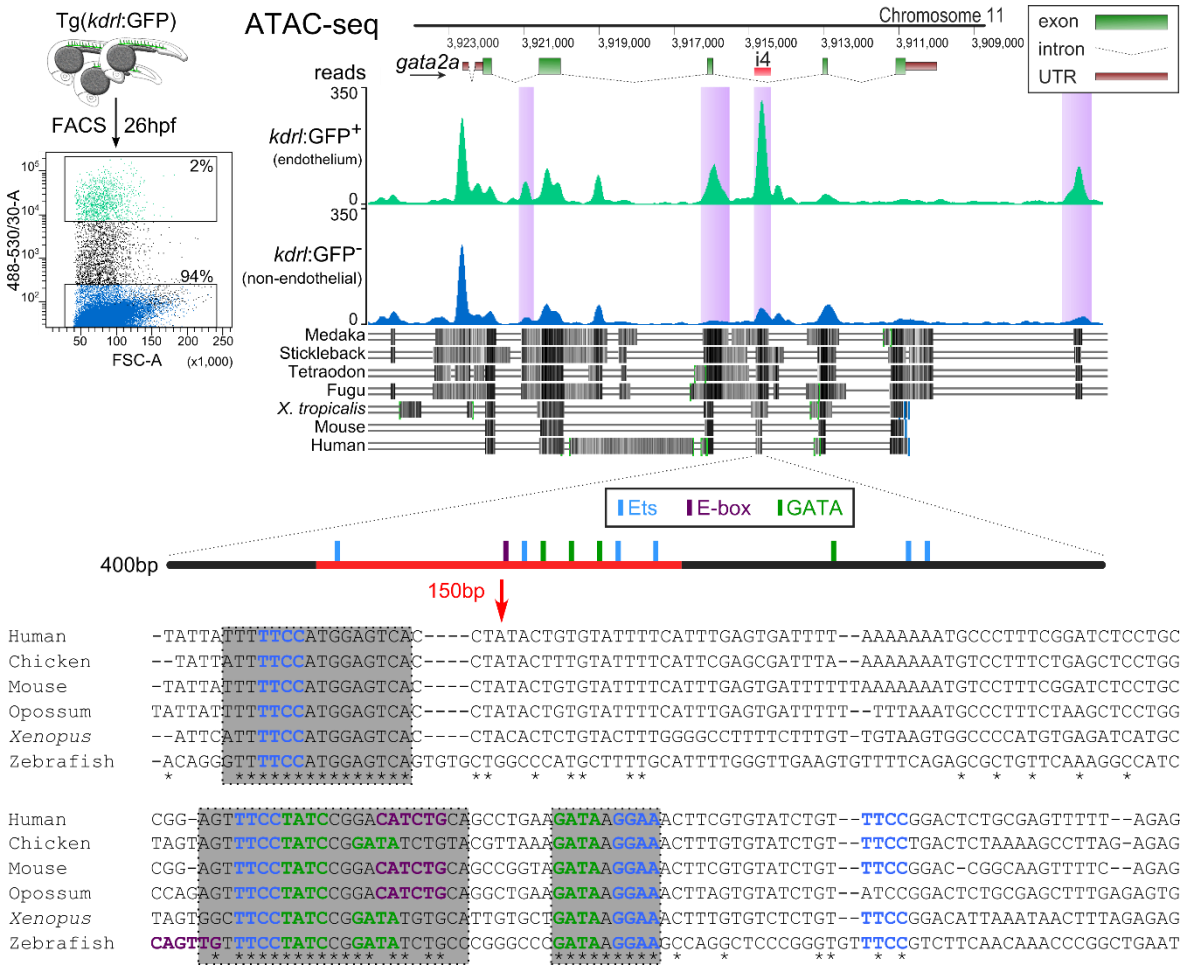


Figure 3.2. Identification of a conserved endothelial enhancer of zebrafish *gata2a*.

Kdrl:GFP$^+$ (sea green) and *kdrl*:GFP$^-$ (blue) cells were FACS-sorted from 26hpf embryos and used for preparation of ATAC-seq libraries. The image of the mapped reads presents stacked means of two biological replicates. Differential peak analysis (by Rossella Rispoli) identified four chromatin regions (purple shading) in the locus of *gata2a* that are significantly more open in the *kdrl*:GFP$^+$ population (p<0.0001). A region in the fourth intron (i4) is conserved throughout vertebrates. The highly conserved 150bp region (red) contains putative transcription factor binding sites (mapped computationally using BLAT tool of UCSC Genome Browser by Sarah de Val). Light blue: Ets binding sites; purple: E-box binding sites; green: GATA binding sites; grey shading: regions of high conservaton; asterisks: conserved residues.

3.4. The *gata2a*-i4 enhancer drives GFP expression in haematopoietic precursors

3.4.1. The *gata2a*-i4:GFP transgenic zebrafish embryos are characterised by GFP expression in the PLM and in the dorsal aorta

Not all accessible chromatin regions represent an active element, therefore identification of enhancers requires their validation in transgenic assays (Dailey 2015). To investigate the activity of the *gata2a*-i4 enhancer *in vivo*, the conserved genomic 150bp region (**Fig. 3.2**), together with flanking ±500bp (*gata2a*-i4-1.1kb:GFP) or ±150bp (*gata2a*-i4-450bp:GFP) was cloned into a reporter E1b:GFP construct (see Chapter 2.3.9.1) and integrated into the zebrafish genome using Tol2 transgenesis (Kawakami et al. 2004). The longer construct, in addition to the conserved sequence, included the whole of the small ATAC-seq peak immediately downstream of the *gata2a*-i4 region (**Fig. 3.2**), while the shorter one only included part of it. Transgenic zebrafish lines were established as described in Chapter 2.2.4.1 (**Fig. 2.1**). The earliest activity of the enhancer was observed at 14-somite stage (14ss), when *gfp* mRNA was detected in the PLM (**Fig. 3.3A-E**). After 22hpf, the reporter signal was pan-endothelial (**Fig. 3.3F-K**, **Fig. 3.5A-B**), in agreement with the activity of the mouse +9.5 enhancer (Khandekar et al. 2007). We also observed strong GFP fluorescence in the heart (**Fig. 3.3F**, **Fig. 3.5A**). Around 27hpf, higher intensities of GFP fluorescence and higher levels of *gfp* mRNA were visible in the floor of the DA, the source of the first HSCs (Swiers et al. 2013b) (**Fig. 3.3H-K**, **Fig. 3.5A-B**). At these stages, the expression of GFP protein and *gfp* mRNA overlapped completely in the endothelial cells of the DA, as demonstrated by confocal imaging of fluorescently labelled 25hpf embryos (**Fig. 3.3N**). While the protein was still visible in the vasculature around 3dpf, it was likely carried over from earlier stages, since the *gfp* mRNA was not detectable anymore (**Fig. 3.3L-M**). Overall, these data confirm the presence of an active element and suggest that the i4 enhancer is active *in vivo* at the correct time window and tissue to regulate definitive

haematopoiesis. Importantly, an analogous pattern was observed in mouse embryos driving LacZ under the VE enhancer (Khandekar et al. 2007), suggesting functional conservation of the *gata2a*-i4 enhancer across vertebrates.

3.4.2. The *gata2a*-i4-GFP$^+$ cells include *gata2a*$^+$ and *runx1*$^+$ haematopoietic precursors

To characterize the identity of the *gata2a*-i4-1.1kb:GFP$^+$ cells, the transgenic embryos were simultaneously probed for *gata2a* mRNA and immunostained for GFP protein at 30hpf (**Fig. 3.4A**). Since the mouse +9.5 enhancer is required for *Gata2* expression in the developing HSPCs (Gao et al. 2013), we were not surprised to find a large overlap between *gata2a*$^+$ and GFP$^+$ cells (**Fig. 3.4B**). Interestingly, a proportion of GFP$^+$ cells did not express *gata2a* mRNA (**Fig. 3.4A**), suggesting that the activity of the enhancer was not completely indicative of *gata2a* expression. This could suggest that some cells require activity of other endothelial enhancers (see **Fig. 3.1**) to trigger transcription of *gata2a*. Other possible explanations could be based on different post-transcriptional stabilities and controls of *gfp* and *gata2a* mRNAs. For instance, in *gfi1aa*^{amc551} transgenic fish, expressing GFP under *gfi1aa* promoter, GFP protein and mRNA were more stable than *gfi1aa* mRNA (Thambyrajah et al. 2016). Moreover, amounts of Gata2 protein are very dynamic due to high proteasomic turnover (Lurie et al. 2008) and we speculate that related mechanisms may regulate the turnover of *gata2a* mRNA, resulting in incomplete overlap between *gata2a* mRNA and GFP in Tg(*gata2a*-i4-1.1kb:GFP) zebrafish embryos. Importantly, the GFP signal was absent in *gata2a*-expressing neural cells (**Fig. 3.4A**).

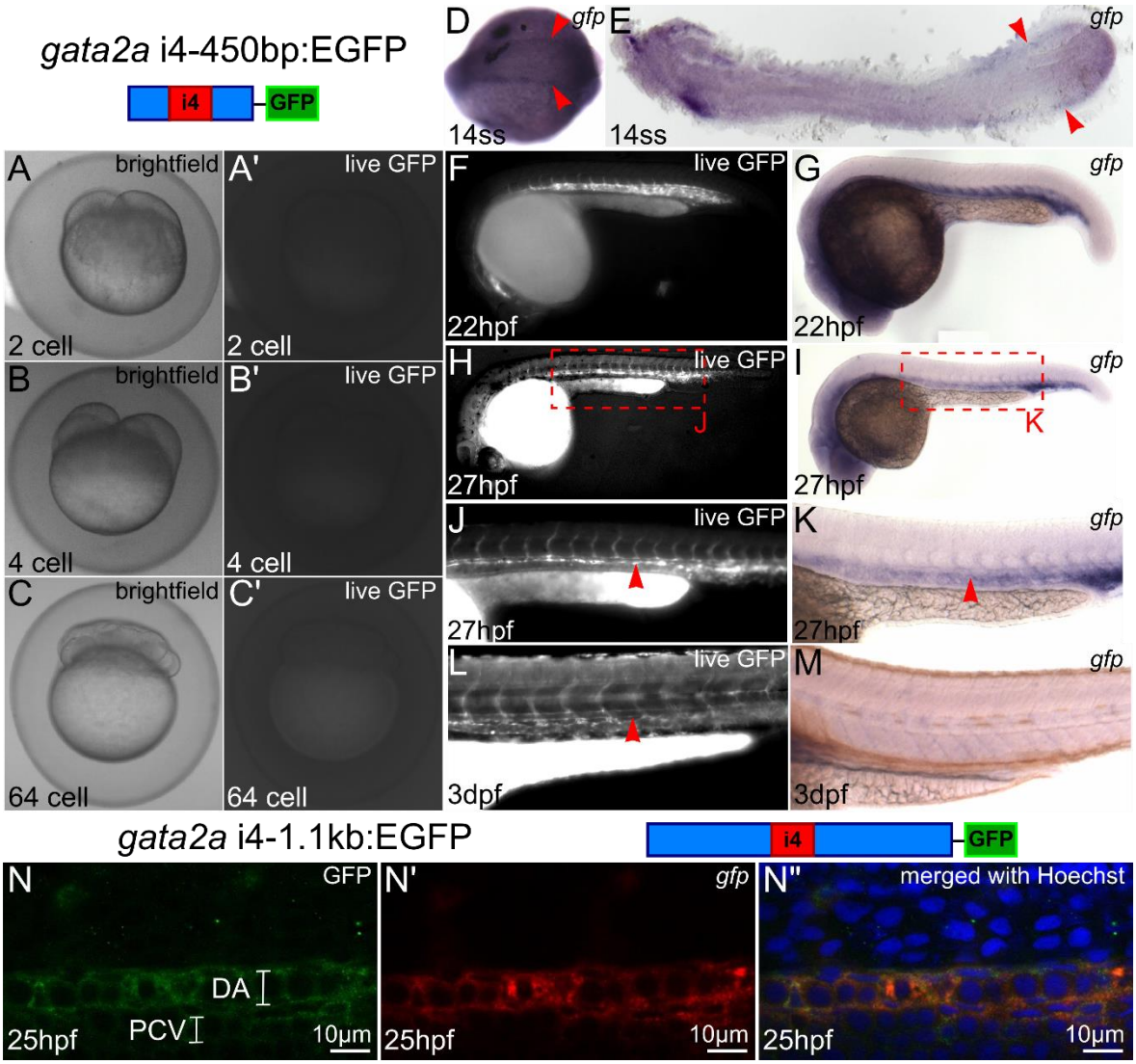


Figure 3.3. ***Gata2a*-i4 enhancer drives GFP expression in PLM, DA and endothelial cells.**

(A-C) Brightfield and widefield fluorescent images of live 2-cell (A-A'), 4-cell (B-B') and 64-cell (C-C') Tg(*gata2a*-i4-450bp:GFP) zebrafish embryo, showing no GFP expression until 64 cell stage. (D) *In situ* hybridization of a 14-somite Tg(*gata2a*-i4-450bp:GFP) embryo shows *gfp* mRNA expression in the posterior lateral plate mesoderm (arrowheads), visible over high background staining in the yolk. (E) Flat mount of the embryo from panel D, showing *gfp* mRNA expression in the posterior lateral plate mesoderm (arrowheads). (F-G) Fluorescent image of a live Tg(*gata2a*-i4-450bp:GFP) embryo and *in situ* hybridization image of the same embryo at 22hpf reveal activity of the i4 enhancer in the forming dorsal aorta and the intermediate cell mass. GFP is also present in the heart. (H-K) At 27hpf, GFP protein is detected with fluorescent microscopy in the endothelial cells and is particularly strong in the dorsal aorta (arrowhead). *In situ* hybridization reveals high degree of overlap between *gfp* RNA and protein expression. Panels J and K show magnified trunks of embryos in panels H and K as indicated. (L-M) By 3dpf, GFP protein persists in the dorsal aorta (arrowhead) and other endothelial cells, but *gfp* mRNA is undetectable by *in situ* hybridization. (N-N') Confocal images (provided by Monika Krecsmarik) of a trunk fragment of a Tg(*gata2a*-i4-1.1kb:GFP) embryo immunostained with anti-GFP antibody (N) and probed for *gfp* mRNA (N') at 25hpf. (N") Merged images from panels N-N' with Hoechst nuclear staining in blue, showing complete overlap of GFP protein and mRNA.

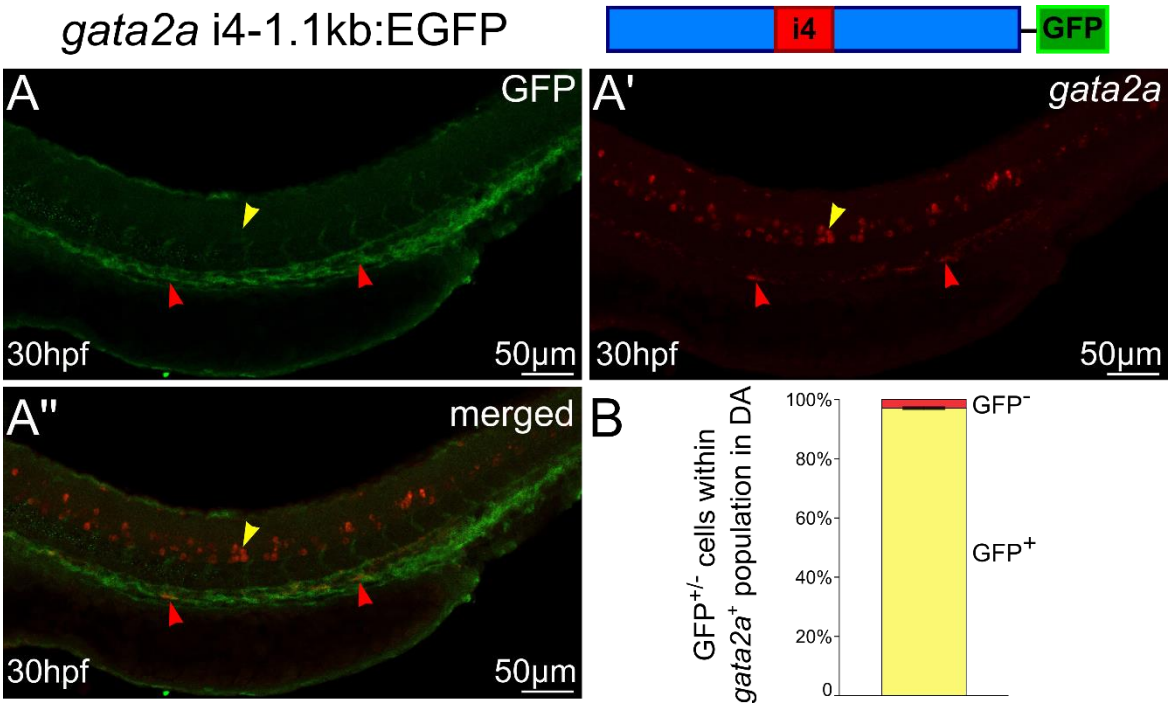


Figure 3.4. *Gata2a*-expressing endothelial cells in DA are GFP-positive in *gata2a*-i4:GFP transgenic embryos.

(A-A') Confocal images (provided by Monika Krecsmarik) of the trunk of a Tg(*gata2a*-i4-1.1kb:GFP) embryo immunostained with anti-GFP antibody (A) and probed for *gata2a* mRNA (A') at 30hpf. (A") Merged images from panels A-A', showing *gata2a*-expressing cells in the DA that are also GFP$^+$ (red arrowheads). *Gata2a*-expressing neural cells (yellow arrowhead) are GFP$^-$. (B) Counting of the *gata2a*$^+$ cells in the DA (by Monika Krecsmarik) represented in panels A'-A" in 9 embryos shows that >95% of *gata2a*$^+$ cells in the DA are also GFP$^+$. N=2. Error bars: ±SD.

Analogous analysis investigating co-expression of GFP driven by the i4 enhancer and *runx1* mRNA in the DA revealed that over 90% of *runx1*$^+$ cells at 25hpf were also GFP$^+$ (**Fig. 3.5**). *Runx1* is one of the earliest markers of developing HSPCs and a *bona fide* HE marker (Burns et al. 2002; Kalev-Zylinska et al. 2002; Gering and Patient 2005). Therefore, we conclude that the GFP expression under the *gata2a*-i4 enhancer marks the majority of the haemogenic endothelial population.

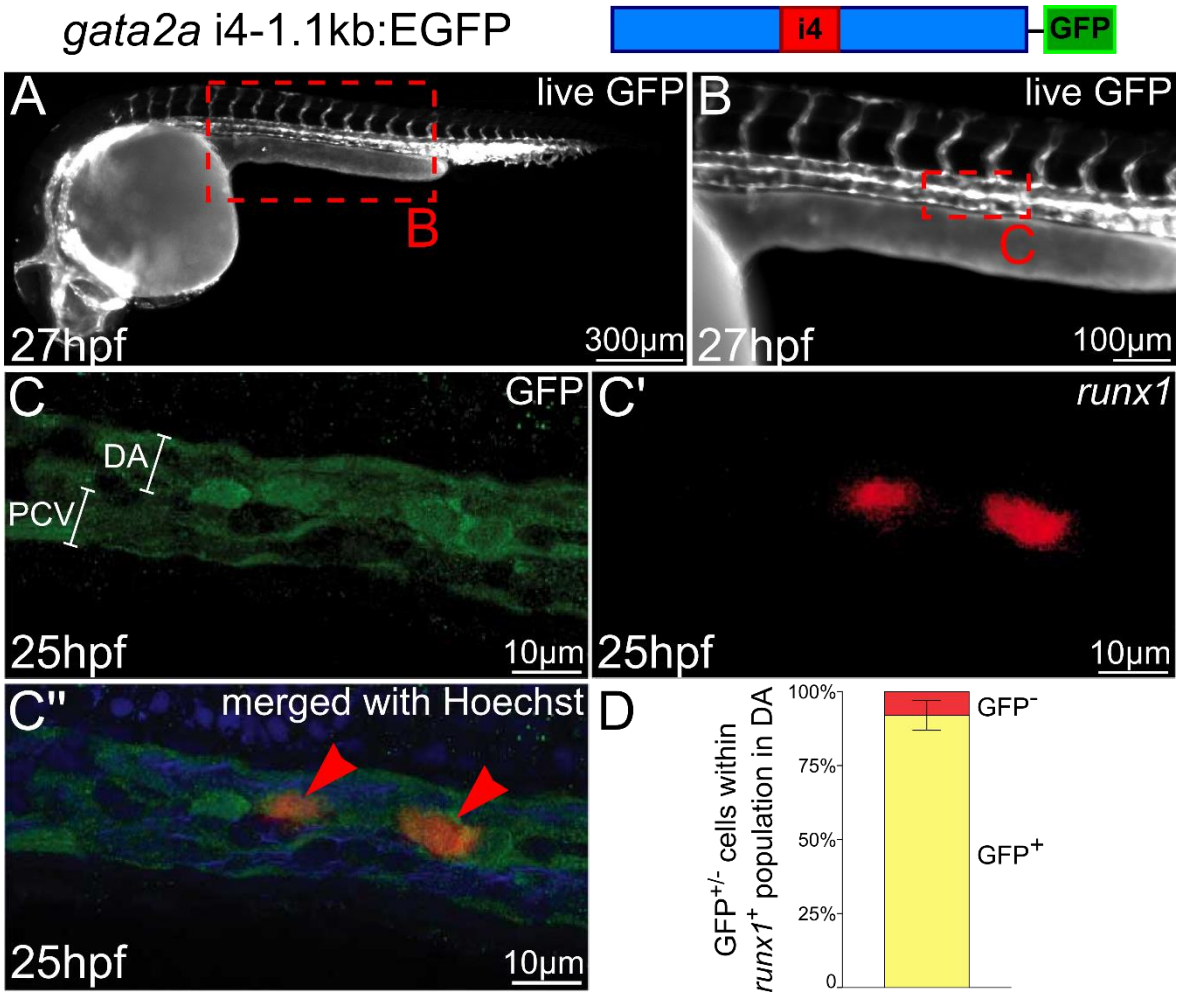


Figure 3.5. GFP-positive cells in *gata2a*-i4:GFP transgenic fish include *runx1*⁺ cells.

(A) Widefield fluorescent image of a live Tg(*gata2a*-i4-1.1kb:GFP) zebrafish embryo at 27hpf, demonstrating GFP fluorescence in the endothelial cells and in the heart. (B) Higher magnification image of the trunk of the embryo from panel A. (C-C') Confocal images (provided by Monika Krecsmarik) of the dorsal aorta (DA) and posterior cardinal vein (PCV) of a Tg(*gata2a*-i4-1.1kb:GFP) embryo immunostained with anti-GFP antibody (C) and probed for *runx1* mRNA (C') at 25hpf. See panel B for approximate position within the embryo. (C") Merged images from panels C-C', also showing Hoechst nuclear staining in blue. (D) Counting of the *runx1*⁺ cells (by Monika Krecsmarik) represented in panels C'-C" in 25 embryos shows that >90% of *runx1*⁺ cells are also GFP⁺. N=3. Error bars: ±SD.

3.4.3. The *gata2a*-i4-GFP$^+$ cells express endothelial and haematopoietic markers

To construct a genetic profile of the *gata2a*-i4-1.1kb:GFP$^+$ and *gata2a*-i4-450bp:GFP$^+$ cells, we used FACS to sort them from 29hpf transgenic embryos, together with their GFP$^-$ counterparts (**Fig. 3.6**). Cells sorted from the Tg(*gata2a*-i4-1.1kb:GFP) fish displayed stronger levels of GFP fluorescence than those from Tg(*gata2a*-i4-450bp:GFP) fish (**Fig. 3.6**). Moreover, the *gata2a*-i4-1.1kb:GFP$^+$ cells clustered into three distinct populations (K. Clark, WIMM FACS facility), one of them potentially representing heart cells, while *gata2a*-i4-450bp:GFP$^+$ cells did not display any apparent clustering in FACS plots (**Fig. 3.6**). Tg(*kdrl*:GFP) embryos, in which GFP marks all endothelial cells (Jin et al. 2005), were used for reference. We used the Fluidigm Biomark™ qRT-PCR system to measure mRNA levels of several haematopoietic, endothelial and cell signalling genes in GFP$^+$ and GFP$^-$ cells. We then analysed the enrichment for these markers in the GFP$^+$ population by comparing the ratios of expression in GFP$^+$ over GFP$^-$ cells among the three transgenic lines. For the large majority of haematopoietic genes investigated (including *runx1*, *gata2b* and *myb*), used as markers of the haemogenic population within the endothelium, there was no significant difference in enrichment among all three lines (**Fig. 3.6**). Expression of *gata2a* was more enriched in *kdrl*:GFP$^+$ and *gata2a*-i4-1.1kb:GFP$^+$ cells than in the *gata2a*-i4-450bp:GFP$^+$ cells, suggesting that the larger genomic region around the 150bp *gata2a*-i4 enhancer may be active in more *gata2a*-expressing cells within the embryonic endothelium. While most of the endothelial-specific genes were, unsurprisingly, more enriched in the *kdrl*:GFP$^+$ cells, there was very little difference in genetic profiles between the two *gata2a*-i4 transgenic lines (**Fig. 3.6**). A similar pattern was observed for genes involved in Notch and Wnt signalling (**Fig. 3.6**). In summary, the genetic profile of both *gata2a*-i4-1.1kb:GFP$^+$ and *gata2a*-i4-450bp:GFP$^+$ cells is very similar, suggesting they can be used interchangeably for studying haematopoietic processes in development. These cells also include the haemogenic

population within the endothelium, as demonstrated by the enrichment for HE markers similar to the Tg(*kdrl*:GFP) line.

3.5. Activity of the *gata2a*-i4 enhancer increases upon *gpr65* knockdown

Previous work identified a G-protein-coupled receptor, GPR65, involved in establishing repressive chromatin at the +9.5 enhancer of *Gata2* in mouse (Gao et al. 2016). In addition, knockdown of zebrafish *gpr65* resulted in increased expression of *runx1* at 36hpf, which was explained by increased *gata2b* expression at 36hpf (Gao et al. 2016). However, these results were unconvincing, as they relied on non-quantitated *in situ* hybridization. Moreover, the analysis in zebrafish was only done at 36hpf, almost 12h after the initiation of *runx1* expression in the DA (Kalev-Zylinska et al. 2002) and 6h after the first EHT events (Kissa and Herbomel 2010). Since the *gata2a*-i4 enhancer is likely to be the zebrafish equivalent of the +9.5 mouse region, we hypothesized that *gpr65* acted on the i4 enhancer at earlier stages than 36hpf. To address this, we injected Tg(*gata2a*-i4-450bp:GFP) with *gpr65* MO. Images of 16 *gpr65* morphants and 15 control (uninjected) siblings probed at 29hpf for *gfp* mRNA (**Fig. 3.7A-B**) were digitally quantified (see **Fig. 2.3**). We found a small, but significant increase of approximately 15% in pixel intensity levels of the *gfp in situ* staining in the DA of *gpr65* knockdown embryos, compared to non-injected controls (μ_{wt}=30.1, μ_{MO}=35; t=2.81, d.f.=29, p<0.01) (**Fig. 3.7C**). The pixel intensity values measured from wild type and morphant embryos showed relatively low dispersion, with coefficients of variation of 11% and 16%, respectively. This suggests that the expression of *gfp* mRNA in the Tg(*gata2a*-i4-450bp:GFP) embryos may be useful as a sensor for *in vivo* activity of the enhancer. The power of the above *t* test to detect the difference at 0.05 level was 89%.

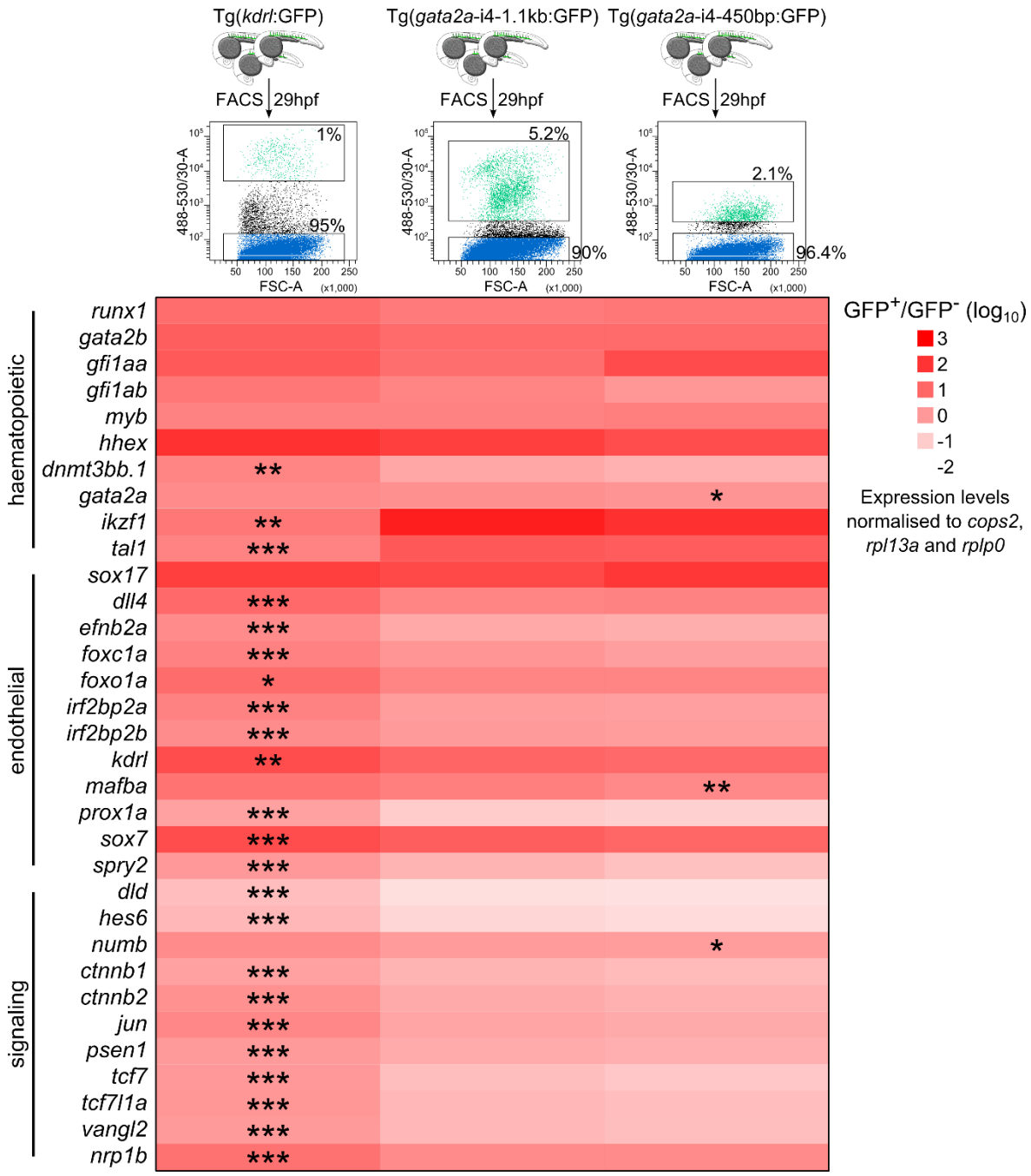


Figure 3.6. GFP⁺ cells in Tg(*gata2a*-i4-1.1kb:GFP) and Tg(*gata2a*-i4-450bp:GFP) transgenic fish express similar levels of haematopoietic and endothelial genes compared to GFP⁻ cells.

GFP⁺ and GFP⁻ cells were FACS-sorted from Tg(*kdrl*:GFP) (N=3), Tg(*gata2a*-i4-1.1kb:GFP) (N=4) and Tg(*gata2a*-i4-450bp:GFP) (N=4) zebrafish embryos at 29hpf. Expression levels of selected genes were measured with Fluidigm Biomark™ system and normalised to housekeeping genes *cops2*, *rpl13a* and *rplp0*. Ratio of expression levels in GFP⁺ over GFP⁻ cells was calculated for each gene. The heat map represents $\log_{10}$ values of these ratios. Several key haematopoietic regulators show no difference in enrichment of expression in GFP+ cells across the three transgenic lines. Most endothelial genes and genes involved in Wnt, Notch and Neuropilin signaling have statistically different enrichment (asterisks) only in the Tg(*kdrl*:GFP)⁺ cells, while there is little difference between Tg(*gata2a*-i4-1.1kb:GFP) and Tg(*gata2a*-i4-450bp:GFP) lines. Asterisks indicate significant changes (detected with ANOVA) in expression enrichment levels compared to the two other populations: *p<0.05, **p<0.01, ***p<0.001.

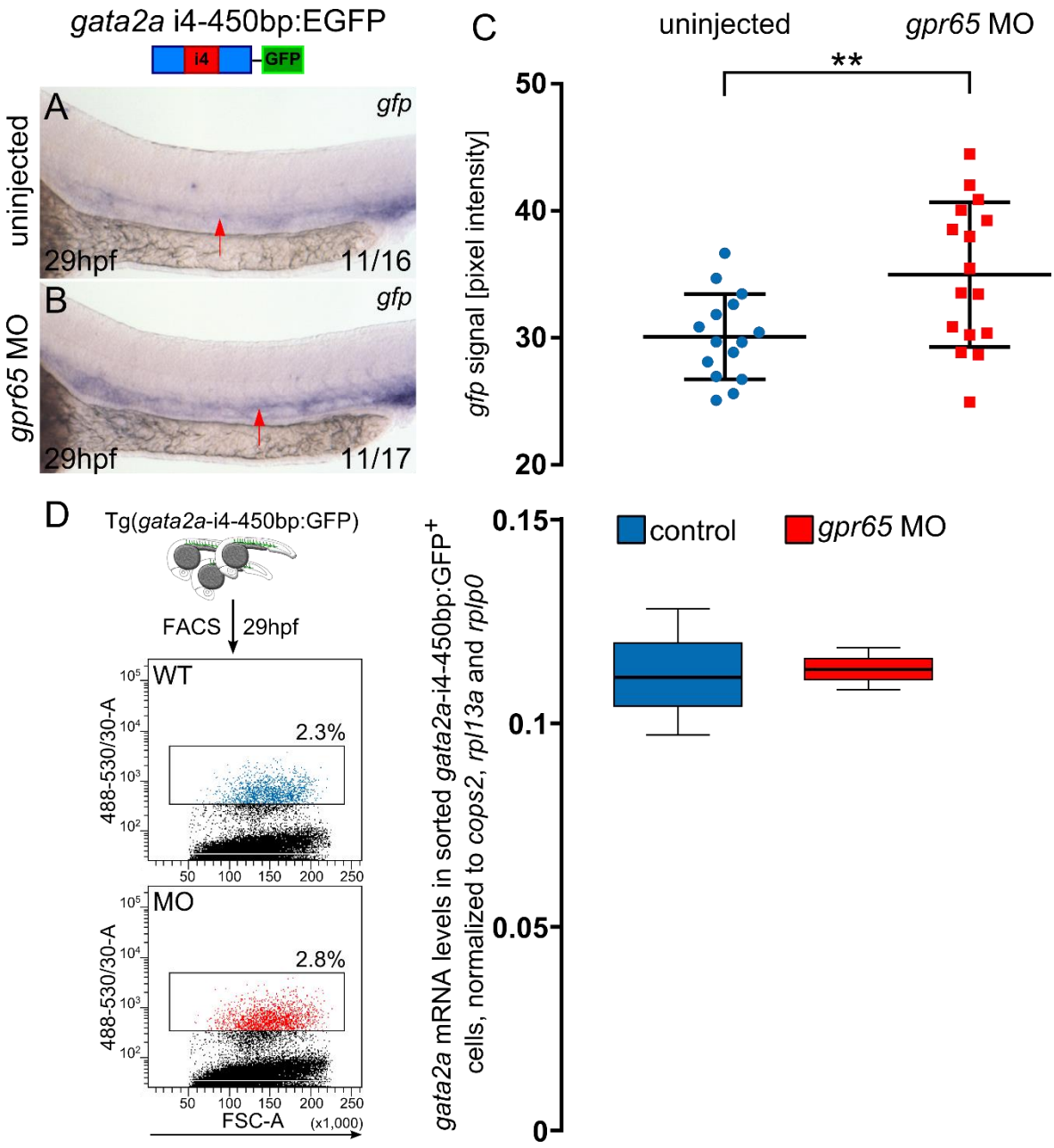


Figure 3.7. Expression of *gfp* mRNA in Tg(*gata2a*-i4-450bp:GFP) transgenic fish depends on the expression of *gpr65*.

(A) A representative image (11/16) of uninjected Tg(*gata2a*-i4-450bp:GFP) embryos aged 29hpf after *in situ* hybridization against *gfp*, showing expression in the dorsal aorta (arrow). (B) A representative image (11/17) of *gpr65* MO-injected Tg(*gata2a*-i4-450bp:GFP) embryos aged 29hpf probed for *gfp* mRNA, showing increased expression in the dorsal aorta compared to uninjected controls (arrow). (C) Quantification of the *gfp* mRNA signal from uninjected (n=15) and *gpr65* morphant (n=16) embryos shows significant increase in *gfp* pixel intensity upon *gpr65* knockdown in Tg(*gata2a*-i4-450bp:GFP) embryos at 29hpf. N=1. Error bars: mean±SD. $^{**}p<0.01$. (D) *Gpr65* morphants have more *gata2a*-i4-450bp:GFP^{+} cells, as demonstrated by FACS. There is no significant difference in *gata2a* mRNA levels in the sorted GFP^{+} cells between uninjected controls and *gpr65* MO-injected 29hpf embryos, measured by qRT-PCR. N=3.

Surprisingly, knockdown of *gpr65* did not result in increased *gata2a* mRNA levels in sorted GFP$^+$ cells from Tg(*gata2a*-i4-450bp:GFP) embryos, measured by qRT-PCR (**Fig. 3.7D**). However, the applied relative quantification method relies on normalisation to reference housekeeping genes (Livak and Schmittgen 2001), effectively measuring mRNA levels per cell. Therefore, the lack of difference can be explained by the fact that *gpr65* MO-injected embryos had on average 20% more GFP$^+$ cells than their uninjected siblings (2.8% total sorted cells compared to 2.3%; N=3) (**Fig. 3.7D**). This suggests that the observed increase in the activity of the i4 enhancer in response to *gpr65* knockdown likely results from a higher number of cells expressing the transgene rather than higher expression levels in the same cell number. This is in agreement with the proposed role of GPR65 in establishing silencing chromatin at the +9.5 region (Gao et al. 2016), where the lack of the protein would result in more cells with the *Gata2* enhancer accessible for transcriptional activation. Consequently, there are likely more cells expressing *gata2a* in *gpr65* morphants compared to uninjected controls, resulting in higher global levels of *gata2a* expression in the endothelium. This hypothesis would need to be validated with qRT-PCR for *gata2a* mRNA in sorted *kdrl*:GFP$^+$ cells from *gpr65* morphants and control siblings. A complementary analysis clarifying the link between *gpr65* and the *gata2a*-i4 enhancer would involve injecting *gpr65* MO into *gata2a*$^{\Delta i4/\Delta i4}$ mutants (see Chapter 4).

3.6. Discussion

3.6.1. Overview

Developmental enhancers are not only responsible for specific patterns of gene expression, but also provide a handle to mark subsets of cells expressing the gene of interest. In most cases, their activity is much more restricted than the embryo-wide expression of the gene they regulate (Nord et al. 2013). Previously, this phenomenon has been employed for example to isolate haemogenic endothelial cells from mouse embryos expressing GFP under the control

of the *Runx1* +23.5 enhancer (Swiers et al. 2013a). While numerous enhancers can be identified computationally by sequence homology (Levine 2010), most of these studies do not include fish in their analyses (Khandekar et al. 2007; Nottingham et al. 2007). Conservation of noncoding regions down to fish would suggest even stronger evolutionary importance of the enhancers in question. However, it would also provide novel tools to study tissue-specific gene expression *in vivo*, particularly useful in the transparent, externally developing zebrafish embryo. In this chapter, we have employed molecular and transgenic tools to identify and characterise a conserved enhancer of the zebrafish *gata2a* gene.

3.6.2. Genomic locus of *gata2a* but not *gata2b* contains a highly conserved intronic enhancer previously characterized in mice

ATAC-seq is a recently developed method to identify open chromatin regions, indicative of *cis* regulatory regions (Encode Consortium 2012), from limited cell populations (Buenrostro et al. 2013). Here we used it to compare chromatin accessibility profiles in FACS-sorted *kdrl*:GFP^{+} cells against GFP^{-} cells. The *kdrl*:GFP transgene marks all the endothelial cells (Jin et al. 2005), and we aimed to identify transposase-accessible regions specific to the endothelial population that contains the precursor of HSCs (Lacaud and Kouskoff 2017). The selected time point (26hpf) was chosen to fall shortly after the initiation of *runx1* expression in the DA (Kalev-Zylinska et al. 2002) and well before the first EHT events (Kissa and Herbomel 2010). By this approach, our analysis included all the early precursors of HSCs, but was not restricted to only haemogenic endothelium. With this, we also hoped to identify *cis* regulatory regions that would confer expression of genes important for HSC generation via non-cell-autonomous mechanisms, for example through signalling pathways within the endothelium (Gritz and Hirschi 2016).

We were surprised to find no differentially endothelial-specific enhancers of *gata2b*. *Gata2b* expression is restricted in the embryo to the ventral wall of the dorsal aorta (Butko et al.

2015) and it was conceivable that such expression pattern would be defined by tissue-specific *cis* regulators. In fact, recently published ATAC-seq data on sorted nuclei from 24hpf *fli1a*:GFP zebrafish embryos suggest the presence of an endothelial-specific enhancer within the *gata2b* locus (Quillien et al. 2017). However, we could not detect it in our analysis. It is likely that the sorting strategy, including also non-haemogenic endothelium, diluted the enrichment for open chromatin in the *gata2b* locus. Moreover, the tissue specificity of *gata2b* expression could be largely instructed by its promoter region. While the accessibility of the enhancers varies widely when comparing different cellular materials, promoters tend to be marked by open chromatin across different cell populations and tissues (Dailey 2015). This results in little differences in promoter accessibility detected with methods like ATAC-seq, making it more difficult to ascribe tissue specificity to the promoter regions. Moreover, *gata2b* expression in haemogenic endothelium could be controlled by a more distal enhancer, assigned to a different neighbouring gene in the computational analysis. Indeed, a recent analysis in *Drosophila melanogaster*, matching *in vivo* enhancer activities to gene expression patterns, suggested that 12-21% of developmental enhancers were not responsible for driving transcription of the most closely located genes (Kvon et al. 2014).

In contrast, we found four differentially enriched ATAC-seq peaks in the endothelial population within the *gata2a* locus. The expression of *gata2a* is much more widespread than *gata2b*, including all endothelial cells, but also the ectoderm and the developing central nervous system (CNS) (Davidson and Zon 2004; Yang et al. 2010; Butko et al. 2015), much more resembling the expression of *Gata2* in the mouse (Nardelli et al. 1999; Visel et al. 2004). Indeed, the mouse Gata2 has known multiple functions not only in the developing haematopoietic system, but also in the CNS (Vicente et al. 2012; Lentjes et al. 2016). Therefore, we expected multiple tissue-specific *cis* controls to exist within the *Gata2* locus. We were not surprised to find several potential enhancers responsible for the endothelial expression of *gata2a*. They all

exhibited sequence conservation across other fish species, but only one region, within the fourth intron (subsequently termed i4), was conserved across other vertebrates, including the human. Detailed sequence analysis confirmed the presence of multiple conserved units of transcription factor binding sites within the i4 enhancer. In fact, identification of evolutionarily conserved non-coding sequences or searching for clusters of transcription factor motifs is a widely accepted method of computational prediction of developmental enhancers (Levine 2010). Here we used such approach to support the experimental ATAC-seq data. In fact, the i4 enhancer is orthologous to the previously described mouse *Gata2* enhancer, termed the VE enhancer (Khandekar et al. 2007) or +9.5 enhancer (Gao et al. 2013). This enhancer proved instrumental in dissecting the function of *Gata2* in definitive and adult haematopoiesis (Lim et al. 2012; Gao et al. 2013). The homologous region in human was also correlated with susceptibility to blood-related disorders (Johnson et al. 2012). Here we demonstrate for the first time the conservation of this region all the way to zebrafish, suggesting it has played an essential role throughout the evolution of the haemato-vascular system.

3.6.3. The *gata2a*-i4 enhancer drives GFP expression in haematopoietic cell populations

The mouse VE *Gata2* enhancer was found to be active in all endothelial cells in the mouse *in vivo* transgenic assays (Khandekar et al. 2007). We tested whether this tissue-specific domain was conserved for the zebrafish *gata2a*-i4 enhancer using a GFP transgene. Indeed, we observed GFP fluorescence in all the endothelial cells of zebrafish embryos at the stages of HE programming. We could also trace the origins of these cells back to the PLM, the source of precursors of endothelium and HSCs (Davidson and Zon 2004; Medvinsky et al. 2011), by assaying for *gfp* expression at 14ss. Importantly, we made use of the imaging power provided by the transparent zebrafish embryo to quantify the spatial co-expression of the transgene with endogenous *gata2a* and the crucial marker of haemogenic endothelial progenitors, *runx1* (Burns

et al. 2002; Kalev-Zylinska et al. 2002; Gering and Patient 2005), in intact embryos. This analysis has demonstrated that while the i4 enhancer is active in other endothelial cells, its activity domain includes most of the haematopoietic precursors within the HE. Intriguingly, the enhancer-driven GFP expression did not overlap completely with *gata2a* mRNA expression, suggesting that other endothelial enhancers may be required for the full domain of activity of *gata2a*. Future analyses of other chromatin accessible regions will provide further insights into *cis* regulation of *gata2a* in the haemogenic endothelium. Our ATAC-seq analysis already identified potential candidates. In addition, a recent study analysed the ATAC-seq profile in the nuclei of *fli1a*:GFP^{+} endothelium at 24hpf and also identified the enrichment for the *gata2a*-i4 enhancer (Quillien et al. 2017). This data set could serve as an additional source of useful information.

The generated Tg(*gata2a*-i4-1.1kb:GFP) and Tg(*gata2a*-i4-450bp:GFP) transgenic lines also allowed us to profile the transcriptomes of the GFP^{+} cells. Although the *gata2a*-i4-1.1kb:GFP^{+} cells displayed stronger GFP fluorescence and clustered into distinct populations, we found little difference between the two lines in enrichment for haematopoietic and endothelial markers. Future analyses will determine whether the longer version of the transgenic enhancer construct marks additional cells to those labelled by the shorter version. Potentially, the two lines may serve as tools to distinguish HE cells undergoing a burst of *gata2a* expression during the HSC specification. Interestingly, the enrichment for the expression of HE markers in both *gata2a*-i4:GFP^{+} populations was very similar to the enrichment observed in *kdrl*:GFP^{+} cells, further confirming that the activity domain of the i4 enhancer includes the haemogenic endothelial population, as well as the rest of the endothelium. Similar analysis of genes involved in conserved signalling pathways, including Notch and Wnt, also revealed little difference between the two new transgenic lines. Both Notch and Wnt (Bigas et al. 2013; Butko et al. 2016; Richter et al. 2017) have known functions in specifying the HE within endothelium

and expression of the corresponding genes in these endothelial fractions was not surprising. However, all these genes were more enriched in the *kdrl*:GFP$^+$ population, compared to the *gata2a*-i4:GFP$^+$ cells. Possible explanations for this observation include more restrictive sorting strategy of *kdrl*:GFP$^+$ cells, resulting in 'purer' endothelial populations. The *gata2a*-i4:GFP$^+$ sorted populations may include other cells, resulting in more diluted expression of the genes involved in Notch and Wnt signalling. These cells, characterization of which is beyond the scope of this thesis, likely include heart cells, marked by strong *gata2a*-i4:GFP fluorescent signal.

Very importantly, we did not observe GFP expression in the developing CNS, which is characterized by high levels of *gata2a* mRNA (Yang et al. 2010). In fact, previously generated transgenic fish driven by the *gata2a* promoter expressed the transgenic reporter in the CNS (Meng et al. 1997). Therefore, our functional *in vivo* analysis provides a new example of achieving tissue and cell specificity by using an enhancer fused to a minimal promoter to drive a transgene. As observed previously for *Runx1* transgenic mice, the expression domain marked by the +23.5 enhancer was much more restricted than the expression domain of *Runx1* mRNA (Nottingham et al. 2007). We propose the *gata2a*-i4 enhancer as a new tool to monitor the endothelial and haematopoietic activity of *gata2a* in zebrafish.

3.6.4. The *gata2a*-i4:GFP transgenic line provides a new tool to assay the activity of the i4 enhancer *in vivo*

The described activity domain of *gata2a*-i4-GFP includes a large subset of endothelial cells, closely resembling endothelial populations marked in other transgenic lines, such as Tg(*kdrl*:GFP) (Jin et al. 2005). Therefore, it is unlikely to serve as a novel tool to isolate enriched haemogenic endothelial populations, as done previously in *Runx1*-23:GFP transgenic mice (Ng et al. 2010; Swiers et al. 2013a). However, here we utilize the enhancer transgenic to quantify its activity *in vivo* in response to a molecular perturbation. Firstly, increased expression of *gfp* under the control of the *gata2a*-i4 enhancer under MO-mediated knockdown of *gpr65* suggests an

update to the currently proposed regulation of *Gata2* by *Gpr65* (Gao et al. 2016), where zebrafish *gata2a* rather than *gata2b* would be the direct target of *gpr65*. Secondly, this experiment demonstrates that the expression of *gata2a*-i4:*gfp* is dynamic and can be used to monitor expression changes during the development of HSCs. To our knowledge, this is the first instance of quantifying the expression of an enhancer transgene in zebrafish as a proxy for tissue-specific gene expression.

Chapter 4

The role of *gata2a* in the programming of the haemogenic endothelium revealed by genomic deletion of its endothelial enhancer

4.1. Introduction

Recent years have brought the development of genome engineering technologies, such as zinc-finger nucleases (ZFNs) (Bibikova et al. 2002; Morton et al. 2006), transcription activator-like effector nucleases (TALENs) (Miller et al. 2011; Zhang et al. 2011a) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 (Jinek et al. 2012). These technologies have enabled researchers to generate an unprecedented range of mutant models by precisely targeting genomic loci with high efficiency. Each of the tools has proven effective in zebrafish (Doyon et al. 2008; Meng et al. 2008; Huang et al. 2011; Hwang et al. 2013). This resulted in rapid development of genome editing in the zebrafish field and more than 100 genes have been effectively targeted in a site-directed manner (Varshney et al. 2015). Before these new technologies became available for reverse genetics, morpholino oligonucleotides (MOs) were used for gene knock down studies in zebrafish for over 15 years (Nasevicius and Ekker 2000). However, discrepancies between the phenotypes induced by genetic mutations and MOs have been exposed (Kok et al. 2015; Place and Smith 2017), resulting in intense debates over the validity of each approach. Some genetic mutants, supported by MOs as a secondary tool, have successfully been utilised to reveal novel molecular mechanisms (Gore et al. 2016). On the contrary, other mutants exhibited genomic compensatory mechanisms otherwise absent in MO-treated zebrafish (Rossi et al. 2015), a phenomenon observed also in other model species (El-Brolosy and Stainier 2017). Notably, all the zebrafish gene knockouts described to date have targeted coding DNA and not *cis* regulatory regions.

Generation of tissue-specific mutants by deleting endogenous enhancer regions has been demonstrated in mouse by deleting the +9.5 enhancer of *Gata2* (Johnson et al. 2012; Gao et al. 2013). This approach resulted in depletion of *Gata2* mRNA in the endothelium after the formation of vessels, allowing dissection of distinct roles of *Gata2* by circumventing earlier endothelial phenotypes. Another example of studying the effect of a loss of an enhancer on

embryonic development involved deleting the locus control region of β-globin (Bender et al. 2000). More recently, an 18kb distal upstream enhancer element from the *GATA2* locus was targeted in human cell lines with TALENs and CRISPR/Cas9 (Gröschel et al. 2014). Removal of the enhancer in its native locus resulted in reduced *GATA2* levels. Translocation of this element in AML inv(3)/t(3;3) patients and in MUTZ-3 AML cells resulted in overexpression of EVI1. This phenotype could be reversed after deletion of the enhancer, demonstrating the importance of this *cis* element for expression of the genes falling under its regulation (Gröschel et al. 2014). Another study demonstrated that CRISPR/Cas9-mediated removal of a 12kb enhancer of *BCL11A* resulted in the loss of *BCL11A* expression in human erythroid cell lines (Canver et al. 2015). Similar results from analogous *in vitro* approaches were reported for enhancers of *SOX2* (Li et al. 2014b; Zhou et al. 2014) and *TAL1* (Mansour et al. 2014). Although there are still considerable limitations to functional validation of enhancers by their genomic deletion (Lopes et al. 2016), the above evidence strongly suggests that targeting *cis* regulatory elements may provide a useful tool to generate mutant animal models with tissue-specific depletion of target gene expression.

This chapter describes CRISPR/Cas9-mediated deletion (*gata2a*$^{\Delta i4}$) of the *gata2a*-i4 enhancer, described in detail in Chapter 3. Deletion of this region results in downregulation of *gata2a* expression specifically in the endothelial cells prior to the specification of the HSC programme in the HE. This tissue-specific molecular phenotype is also associated with decreased expression of other haematopoietic genes, including *gata2b* and *runx1*, without affecting endothelial development. The results presented here demonstrate that Gata2a is an upstream regulator of its paralogue, *gata2b*, and is required during definitive haematopoiesis. Moreover, the evolutionary subfunctionalisation of two zebrafish *gata2* paralogues provides a handle to dissect distinct roles of *Gata2* in haematopoiesis.

4.2. Generation and detection of CRISPR/Cas9-induced deletions of *gata2a*-i4 enhancer

Applying genome engineering tools to target non-coding DNA regions, such as enhancers, requires special considerations. While even short indels within protein-coding regions often result in frameshifts and nonsense mutations, such modifications are unlikely to be sufficient to disrupt enhancers (Lopes et al. 2016). Therefore, a single sgRNA/Cas9 complex targeting one position within the enhancer would likely be insufficient to influence gene expression. A more promising solution lies within targeting two guides to the borders of the enhancer region, as done previously *in vitro* with TALENs and CRISPR/Cas9 to delete an 18kb distal commonly translocated segment (CTS) from the *GATA2* locus (Gröschel et al. 2014).

Following this strategy, we identified potential sgRNA target sites flanking the 150bp highly conserved, active region within intron 4 of the *gata2a* locus (see **Fig. 3.2**). To reduce potential off-target effects of CRISPR/Cas9, we utilized the D10A 'nickase' version of Cas9 nuclease (Mali et al. 2013; Ran et al. 2013b), together with two pairs of sgRNAs flanking the enhancer (**Fig. 4.1A**). Mutant carriers were isolated as described in the Methods chapter (**Fig. 2.2**). Sequencing of the isolated mutant alleles revealed two deletions of 215bp (Δ78-292) and 231bp (Δ73-303) (**Fig. 4.1B**). Both deletions included the highly conserved E-box, Ets and GATA transcription factor binding sites. The Δ73-303 allele was selected for further experiments under the term *gata2a*$^{\Delta i4}$. To unambiguously genotype wild types, heterozygotes and homozygous mutants, we designed a strategy consisting of two PCR primer pairs (**Fig. 4.1**). One primer pair flanked the whole region, producing a 600bp wild type band and 369bp mutant band. In the second primer pair, one of the primers was designed to bind within the deleted region, only giving a 367bp band in the presence of the wild type allele (**Fig. 4.1C**).

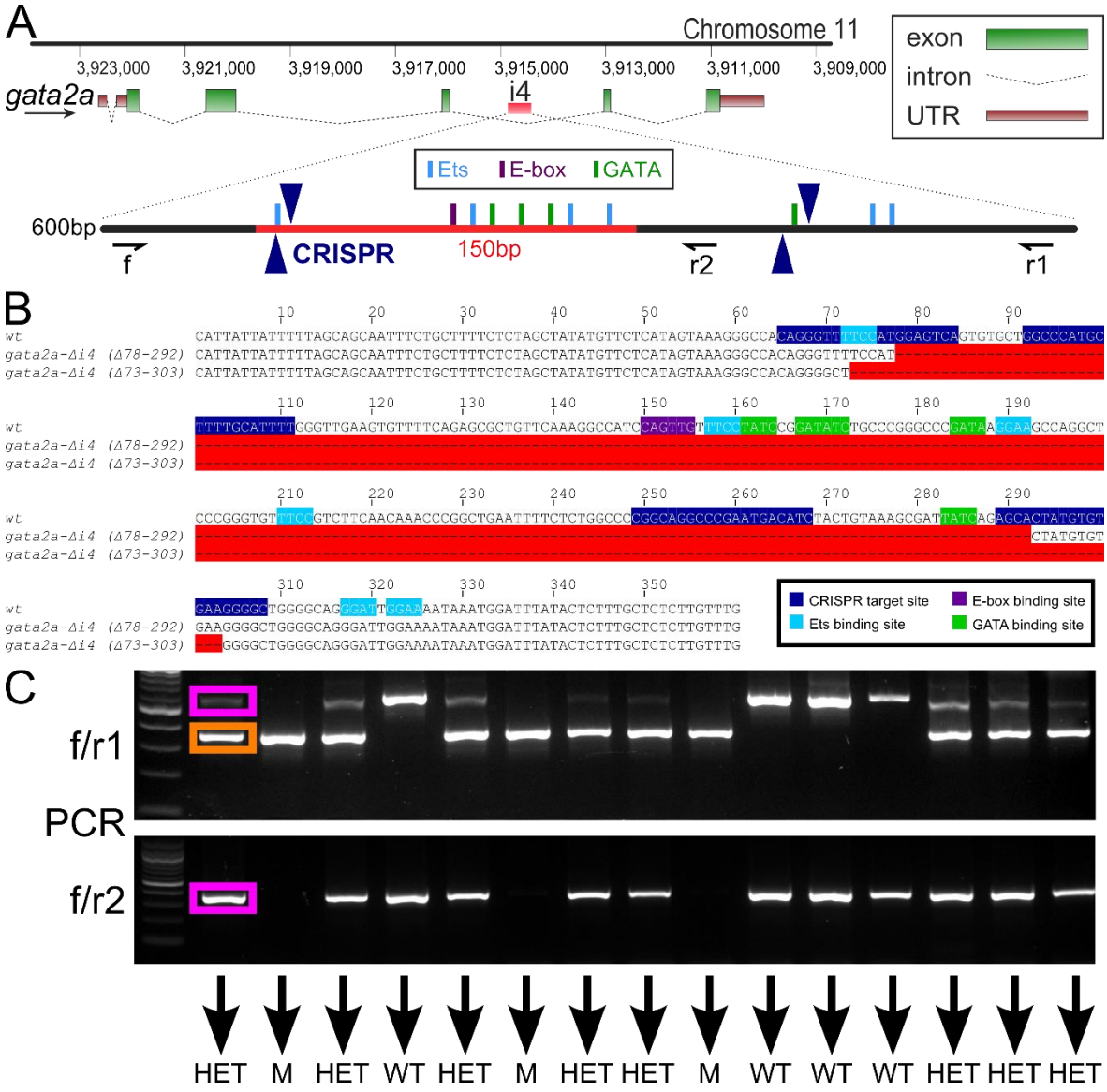


Figure 4.1. Generation of *gata2a*$^{\Delta i4}$ zebrafish mutants with CRISPR/Cas9.

(A) Two pairs of CRISPR sgRNAs (dark blue) were designed to flank the highly conserved 150bp region within the *gata2a*-i4 enhancer. Approximate binding positions of diagnostic PCR primers f, r1 and r2 are indicated. Light blue: Ets binding sites; purple: E-box binding sites; green: GATA binding sites (mapped by Sarah de Val). (B) Two isolated mutant alleles (Δ78-292 and Δ73-303) carry deletions (red gaps) including the majority of highly conserved transcription factor binding sites. (C) Genotyping of embryos from the incross of two *gata2a*$^{+/\Delta i4}$ mutants with PCR relies on the identification of bands specific to wild type alleles (pink) and the mutant-specific band (orange). Primers f, r1 and r2 are the same as in panel A. The first lane from the left in each gel: 100bp marker. WT: wild type; HET: heterozygote; M: mutant.

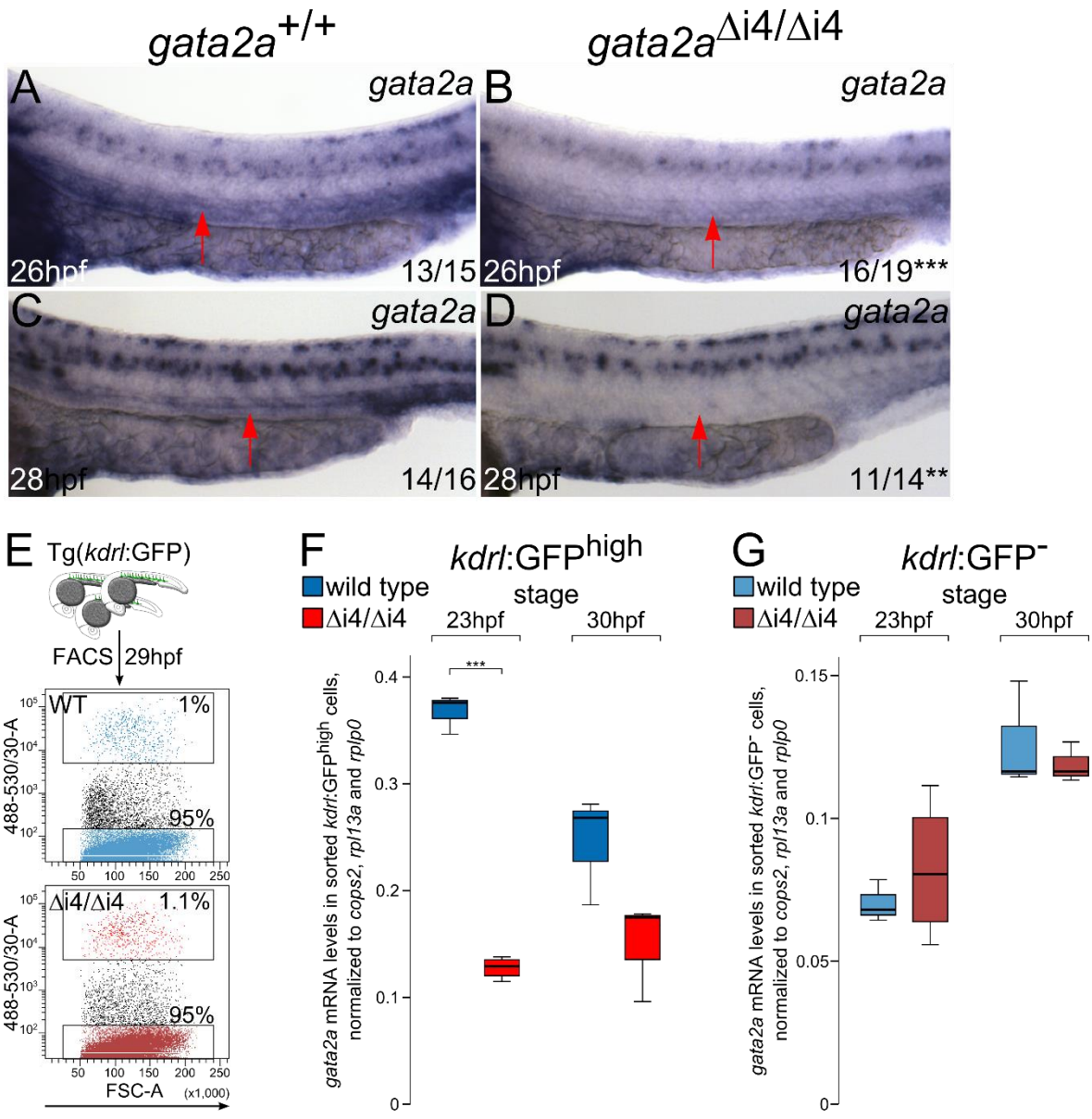


Figure 4.2. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type, express reduced levels of *gata2a* mRNA specifically in the endothelium.

(A-D) A significant majority of *gata2a*$^{\Delta i4/\Delta i4}$ mutants have reduced levels of *gata2a* mRNA in the dorsal aorta (arrows) at 26hpf and at 28hpf, compared to wild type siblings, as detected with *in situ* hybridization. The expression in the neural tube appears unaffected. $^{**}p<0.01$; $^{***}p<0.001$. (E) *Kdrl*:GFP^{high} and *kdrl*:GFP^{-} cells were sorted from non-mutant (WT, blue) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) embryos carrying the Tg(*kdrl*:GFP) transgene. (F-G) QRT-PCR on RNA isolated from the sorted *kdrl*:GFP^{high} or *kdrl*:GFP^{-} cells (panel E) shows decreased levels of *gata2a* mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf and at 30hpf, compared to wild type. There is no difference in *gata2a* mRNA levels in non-endothelial cells between wt and *gata2a*$^{\Delta i4/\Delta i4}$ mutants. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples. Note different scales of expression levels. $^{***}p<0.001$.

4.3. Deletion of the *gata2a*-i4 enhancer results in decreased expression of *gata2a* in the embryonic endothelium

Deletion of a single enhancer, even when it displays a very high degree of tissue-specificity, is not guaranteed to affect expression levels of the target gene. Some *cis* elements may be redundant and only deletion of a few of them may result in reduced transcriptional activity (Lopes et al. 2016). However, previously reported deletion of only the +9.5 enhancer of *Gata2* was sufficient to deplete *Gata2* mRNA in the developing endothelium (Johnson et al. 2012; Gao et al. 2013). Therefore, we wanted to investigate whether the deletion of the *gata2a*-i4 enhancer affected levels of *gata2a* mRNA in the endothelium.

In situ hybridization analysis of *gata2a*$^{\Delta i4/\Delta i4}$ mutants in comparison to their wild type siblings revealed decreased levels of *gata2a* mRNA in the DA region in the majority of homozygous mutant embryos at 26hpf (X^2=14.177, d.f.=1, p<0.001) (**Fig. 4.2A-B**). Similarly, 28hpf embryos phenotypically scored as having 'low' levels of *gata2a* expression in the DA were significantly overrepresented within the mutant siblings (X^2=10.720, d.f.=1, p<0.01) (**Fig. 4.2C-D**). The *gata2a* expression in the neural tube appeared unaffected in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants (**Fig. 4.2A-D**).

Because adult homozygous *gata2a*$^{\Delta i4/\Delta i4}$ fish were viable and fertile, they were crossed to Tg(*kdrl*:GFP) transgenics to generate Tg(*kdrl*:GFP); *gata2a*$^{\Delta i4/\Delta i4}$ embryos. These fish, along with non-mutant Tg(*kdrl*:GFP), were used for FACS-mediated isolation of *kdrl*:GFP^{+} and *kdrl*:GFP^{-} cells to quantitatively compare mRNA expression levels of *gata2a* in the endothelial and non-endothelial cells of wild type and *gata2a*$^{\Delta i4/\Delta i4}$ embryos (**Fig. 4.2E**). Fluidigm Biomark™ qRT-PCR analysis revealed significantly decreased levels of *gata2a* mRNA in the *kdrl*:GFP^{+} endothelial population of 23hpf mutant embryos, compared to wild types (t=20.026, d.f.=5, p<0.001) (**Fig. 4.2F**). At 30hpf this decrease was not statistically significant at the p=0.05 level.

However, this was likely due to a decrease in expression of *gata2a* in the endothelial cells of wild type embryos that appears to occur normally during development, while the mutant expression levels remained low (**Fig. 4.2F**). Importantly, there was no difference in *gata2a* expression levels in the non-endothelial population (*kdrl*:GFP^{-} cells) between wild types and *gata2a*$^{\Delta i4/\Delta i4}$ mutants at either 23hpf or 30hpf (**Fig. 4.2G**). Altogether, these data suggest that genomic deletion of the *gata2a*-i4 enhancer is sufficient to reduce expression of *gata2a* specifically in the endothelial tissue. This is in agreement with mouse data, where deletion of the *Gata2* +9.5 enhancer reduced *Gata2* expression in the endothelium, but not in the brain (Johnson et al. 2012).

4.4. Loss of the *gata2a*-i4 enhancer does not affect endothelial development

The previously described *gata2a*$^{um27/um27}$ mutant, characterized by a 10bp deletion within the protein-coding region, showed arteriovenous shunts and reduced perfusion of dorsal aorta, detectable around 48hpf (Zhu et al. 2011). Therefore, we investigated whether the *gata2a*$^{\Delta i4/\Delta i4}$ mutants had a similar phenotype. Moreover, since the *gata2a*-i4 enhancer was active throughout the endothelial tissue (see **Fig. 3.3**), we expected an endothelial phenotype in the mutants. Surprisingly, endothelial development, assessed by expression of *kdrl* mRNA at 28hpf and Tg(*kdrl*:GFP) transgene at 48hpf, was unaffected in the *gata2a*$^{\Delta i4/\Delta i4}$ embryos (**Fig. 4.3A-F**). Moreover, the *kdrl* expression levels were not reduced in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants (**Fig. 4.3G**), contrary to the reduction in *kdrl* mRNA in *gata2a*$^{um27/um27}$ fish (Zhu et al. 2011). Expression levels of other endothelial markers, such as *dld* or *dll4*, measured with qRT-PCR, were also unchanged in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants, compared to wild type (**Fig. 4.3H-I**). Therefore, the *gata2a*$^{\Delta i4/\Delta i4}$ deletion appears to circumvent the endothelial phenotype of *gata2a*^{um27} mutants, characterised by *gata2a* loss-of-function in all the cells of the embryo. Because the DA endothelium is specified normally, this also confirms that the observed reduction in *gata2a* expression, described in

Chapter 4.3, results primarily from reduced transcription of *gata2a* mRNA, rather than reflecting a secondary phenotype resulting from an earlier loss of endothelial cells that normally express *gata2a*.

4.5. *Gata2a*$^{\Delta i4/\Delta i4}$ mutants exhibit defects in the expression of haematopoietic genes

4.5.1. Deletion of the *gata2a*-i4 enhancer results in reduced expression of *runx1* at early stages of HE specification

The endothelial phenotype of *gata2a* mutants has been previously described (Zhu et al. 2011). Whether it plays a role in HSC development has not been addressed to date, mainly because of existing evidence suggesting that *gata2b* is the zebrafish *Gata2* ortholog responsible for HSC specification (Butko et al. 2015). To investigate whether *gata2a* plays a role in HSC development, we compared the expression of *runx1*, the key marker of definitive haematopoiesis in zebrafish (Burns et al. 2002; Kalev-Zylinska et al. 2002; Gering and Patient 2005), in wild type and *gata2a*$^{\Delta i4/\Delta i4}$ embryos.

Quantitative *in situ* hybridization analysis (see **Fig. 2.3**) with the use of ANOVA revealed decreased expression of *runx1* in *gata2a*$^{\Delta i4/\Delta i4}$ embryos compared to wild type siblings at 24hpf (μ_{wt}=17.3, μ_{mut}=6.4; F=11.735, d.f.=2, 77; p<0.001), 26hpf (μ_{wt}=24.3, μ_{mut}=13.4; F=4.549, d.f.=2, 60; p<0.05), 28hpf (μ_{wt}=34.8, μ_{mut}=25.3; F=4.956, d.f.=2, 58; p=0.01) and 30hpf (μ_{wt}=23.7, μ_{mut}=11.6; F=7.077, d.f.=2, 56; p<0.01) (**Fig. 4.4A-H**). Further qRT-PCR analysis in sorted *kdrl*:GFP^{+} and *kdrl*:GFP^{-} cells (**Fig. 4.4I**) revealed significantly decreased levels of *runx1* mRNA in the *kdrl*:GFP^{+} endothelial population of 23hpf mutant embryos, compared to wild type (t=2.585, d.f.=5, p<0.05) (**Fig. 4.4J**). At 30hpf, the *runx1* levels were higher in the mutants

than in wild type, but this difference was not statistically significant. This discrepancy between *in situ* hybridization and qRT-PCR data may be partly explained by the fact that the *kdrl*:GFP^{+} population likely includes the *runx1*-expressing EMPs (Bertrand et al. 2007), which are not included in the quantification of *in situ* hybridization. In the *kdrl*:GFP^{-} non-endothelial population, *runx1* levels were slightly increased in *gata2a*$^{\Delta i4/\Delta i4}$ embryos compared to wild type at 23hpf (t=2.682, d.f.=5, p<0.05) and 30hpf (t=3.871, d.f.=4, p<0.05) (**Fig. 4.4K**). Because *gata2a* levels were not affected in these cells (**Fig. 4.2G**), there may be a non-cell autonomous effect on *runx1* in the nervous system, where *runx1* is normally expressed. A hypothetical explanation may involve signals from *gata2a*-deficient endothelial cells that would stimulate non-endothelial expression of *runx1*. On the other hand, *gata2a*$^{\Delta i4/\Delta i4}$ embryos may have undetectable reduction in *gata2a* levels in the heart, where the *gata2a*-i4 enhancer drove GFP expression. In this scenario, the observed increase in *runx1* levels could originate from the heart tissue. However, further investigation is beyond the scope of this thesis. Unsurprisingly, levels of *dnmt3bb.1*, a downstream target of *runx1* (Gore et al. 2016), were significantly lower in the *kdrl*:GFP^{+} fraction of the mutants at 23hpf (t=2.997, d.f.=5, p<0.05), compared to wild type (**Supplementary Fig. 4.1**). At 30hpf we could still detect decreased *dnmt3bb.1* expression by qRT-PCR, although it was not significant at the p=0.05 level (**Supplementary Fig. 4.1**). We observed no differences in *dnmt3bb.1* expression in the *kdrl*:GFP^{-} population (**Supplementary Fig. 4.1**).

Taken together, these data suggest that deletion of the *gata2a*-i4 enhancer results in the reduction of *runx1* expression specifically in the endothelium at early stages of HE programming, suggesting a previously unreported role of *gata2a* in definitive haematopoiesis. Moreover, this observation correlates well with published data, where +9.5$^{-/-}$ mouse AGM explants showed lower levels of *Runx1* expression compared to +9.5$^{+/+}$ controls (Gao et al. 2013), further supporting the critical evolutionary role of the intronic enhancer of *Gata2* in HSC specification.

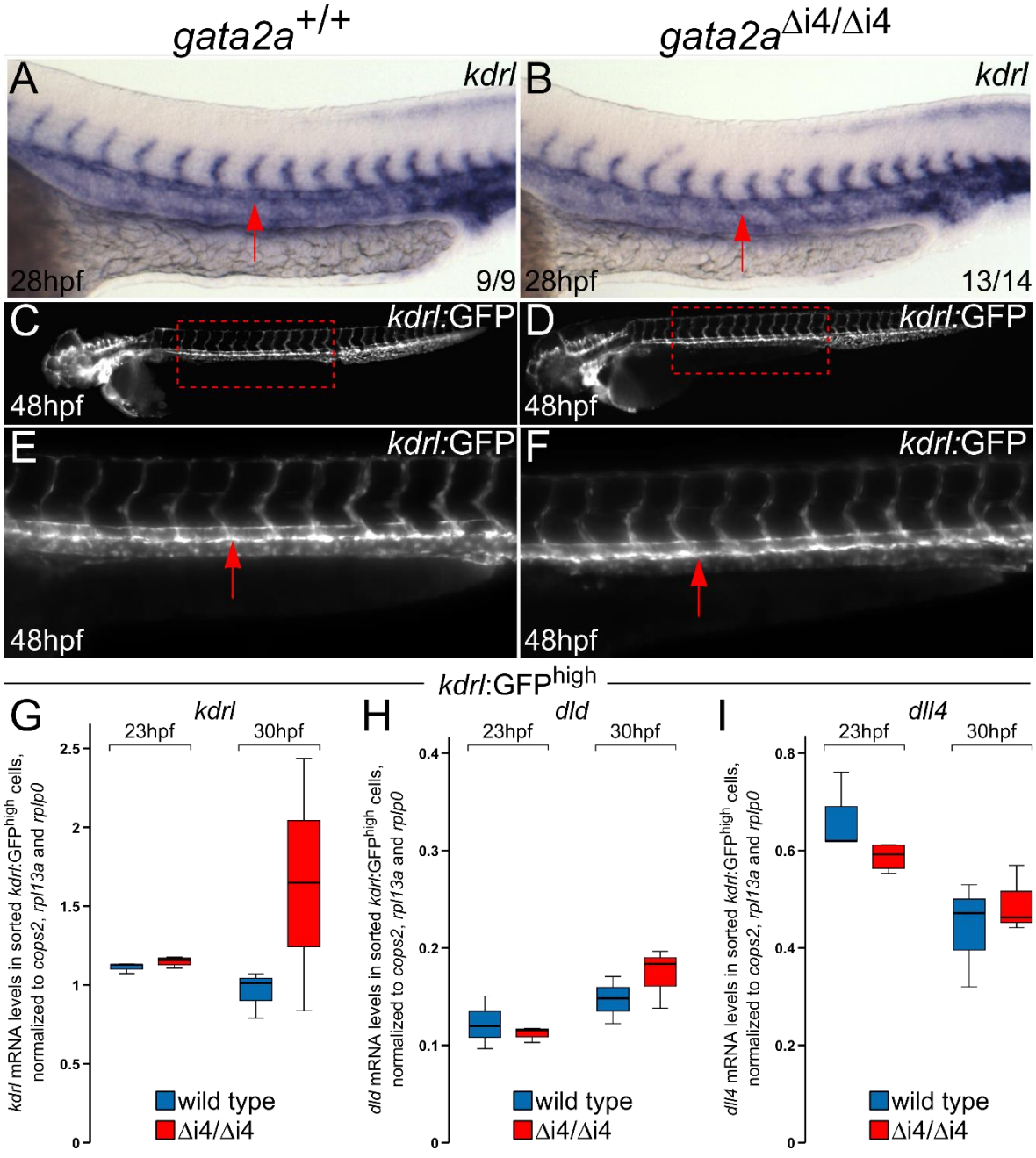


Figure 4.3. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants exhibit no defects in endothelial development.

(A-B) *In situ* hybridization for the endothelial marker *kdrl* at 28hpf reveals no difference between *gata2a*$^{\Delta i4/\Delta i4}$ mutants and wt siblings. The dorsal aorta (arrows) appears unaffected. (C-D) Live images of 48hpf Tg(*kdrl*:GFP) and Tg(*kdrl*:GFP); *gata2a*$^{\Delta i4/\Delta i4}$ demonstrate normal endothelial development of the mutants. (E-F) Zoomed images of the embryos in panels C-D, within the outlined areas. The endothelium of the dorsal aorta (arrows) appears normal in the *gata2a*$^{\Delta i4/\Delta i4}$ embryos. (G-I) QRT-PCR in sorted *kdrl*:GFP^{high} cells from wt (blue) and *gata2a*$^{\Delta i4/\Delta i4}$ mutants (red) shows no differences in the levels of *kdrl*, *dld* and *dll4* mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf and 30hpf, compared to wild type. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples.

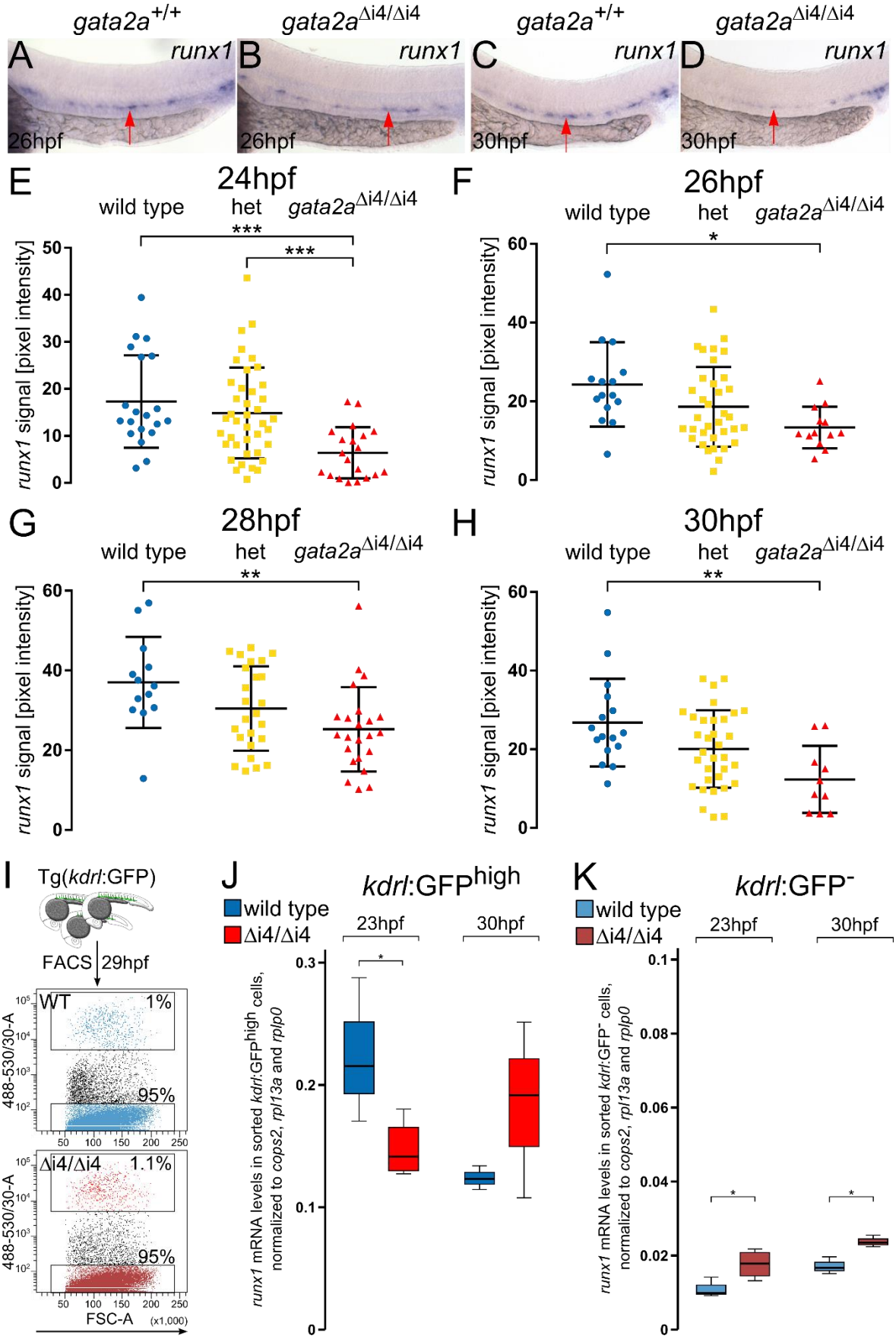

gata2a+/+
gata2aΔi4/Δi4
gata2a+/+
gata2aΔi4/Δi4
A
runx1
26hpf
B
runx1
26hpf
C
runx1
30hpf
D
runx1
30hpf
E
24hpf
wild type
het
gata2aΔi4/Δi4
***
***
runx1 signal [pixel intensity]
F
26hpf
wild type
het
gata2aΔi4/Δi4
*
runx1 signal [pixel intensity]
G
28hpf
wild type
het
gata2aΔi4/Δi4
**
runx1 signal [pixel intensity]
H
30hpf
wild type
het
gata2aΔi4/Δi4
**
runx1 signal [pixel intensity]
I
Tg(kdrl:GFP)
FACS 29hpf
WT
1%
95%
Δi4/Δi4
1.1%
95%
488-530/30-A
FSC-A (x1,000)
J
kdrl:GFPhigh
wild type
Δi4/Δi4
23hpf
30hpf
runx1 mRNA levels in sorted kdrl:GFPhigh cells, normalized to cops2, rpl13a and rplp0
K
kdrl:GFP-
wild type
Δi4/Δi4
23hpf
30hpf
runx1 mRNA levels in sorted kdrl:GFP- cells, normalized to cops2, rpl13a and rplp0

Figure 4.4. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type embryos, express lower levels of *runx1* in the dorsal aorta during the specification of HE.

(A-D) Representative lateral images of the trunks of *runx1*-probed wild type and *gata2a*$^{\Delta i4/\Delta i4}$ embryos at 26hpf and 30hpf, showing the expression of *runx1* mRNA in the dorsal aorta (arrows). (E-H) Quantification of the *runx1* mRNA signal, detected by *in situ* hybridization, from wild type (blue), heterozygous *gata2a*$^{+/\Delta i4}$ (het, yellow) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) siblings at 24-30hpf shows significant decrease in *runx1* pixel intensity in the DA in the homozygous mutants compared to wt at all stages. Error bars: mean±SD. *p<0.05; **p<0.01; ***p<0.001. (I) *Kdrl*:GFP^{high} and *kdrl*:GFP^{-} cells were sorted from non-mutant (WT, blue) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) embryos carrying the Tg(*kdrl*:GFP) transgene. (J-K) QRT-PCR on RNA isolated from the sorted *kdrl*:GFP^{high} or *kdrl*:GFP^{-} cells (panel E) shows decreased levels of *runx1* mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf but not at 30hpf, compared to wild type. In non-endothelial cells, *gata2a*$^{\Delta i4/\Delta i4}$ mutants express higher levels of *runx1* mRNA than wt controls. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples. Note different scales of expression levels. *p<0.05.

4.5.2. Deletion of the *gata2a*-i4 enhancer results in reduced expression of *gata2b* at early stages of HE specification

Previous investigation into the potential relationship between the two zebrafish *gata2* paralogues showed that *gata2a* was not sufficient to rescue the haematopoietic defect induced by *gata2b* MO, excluding the scenario of *gata2a* acting downstream of *gata2b* during the specification of HSCs (Butko et al. 2015). Thus, we tested whether *gata2a* could act upstream of *gata2b* by measuring expression levels of *gata2b* in *gata2a*$^{\Delta i4/\Delta i4}$ embryos.

Quantitative analysis of *in situ* hybridization demonstrated significantly decreased expression of *gata2b* in *gata2a*$^{\Delta i4/\Delta i4}$ embryos compared to wild type siblings at 22hpf (μ_{wt}=35.8, μ_{mut}=24.3; t=3.908, d.f.=47, p<0.001), 26hpf (μ_{wt}=42.6, μ_{mut}=25.1; t=5.361, d.f.=34, p<0.001), 28hpf (μ_{wt}=39, μ_{mut}=30.1; F=5.05, d.f.=2, 54; p=0.01) and 30hpf (μ_{wt}=32.6, μ_{mut}=26; F=3.34, d.f.=2, 80; p<0.05), but not at 24hpf, where the observed decrease in pixel intensity of about 20% (μ_{wt}=27.8, μ_{mut}=21.8) was not statistically significant at the p=0.05 level (**Fig. 4.5A-I**). Fluidigm Biomark™ qRT-PCR analysis in sorted *kdrl*:GFP^{+} cells (see **Fig. 4.2E**) further confirmed significantly decreased levels of *gata2b* mRNA in the endothelial population of 23hpf mutant embryos, compared to wild type (t=3.334, d.f.=5, p<0.05) (**Fig. 4.5J**). At 30hpf, the mutants showed higher levels of *gata2b* mRNA than wild type controls, but this difference was not statistically significant (**Fig. 4.5J**). Again, as in the case of *runx1*, this discrepancy may be caused by the EMPs (Bertrand et al. 2007), although there is no documented evidence for *gata2b* expression in this population. Alternatively, this recovery of *gata2b* expression by 30hpf suggests that other regulators of *gata2b* might compensate for the lack of endothelial *gata2a* in *gata2a*$^{\Delta i4/\Delta i4}$ mutants. In summary, these data suggest that *gata2a* is responsible for early expression of *gata2b* in the DA to programme the HE, further supporting a role for *gata2a* in definitive haematopoiesis.

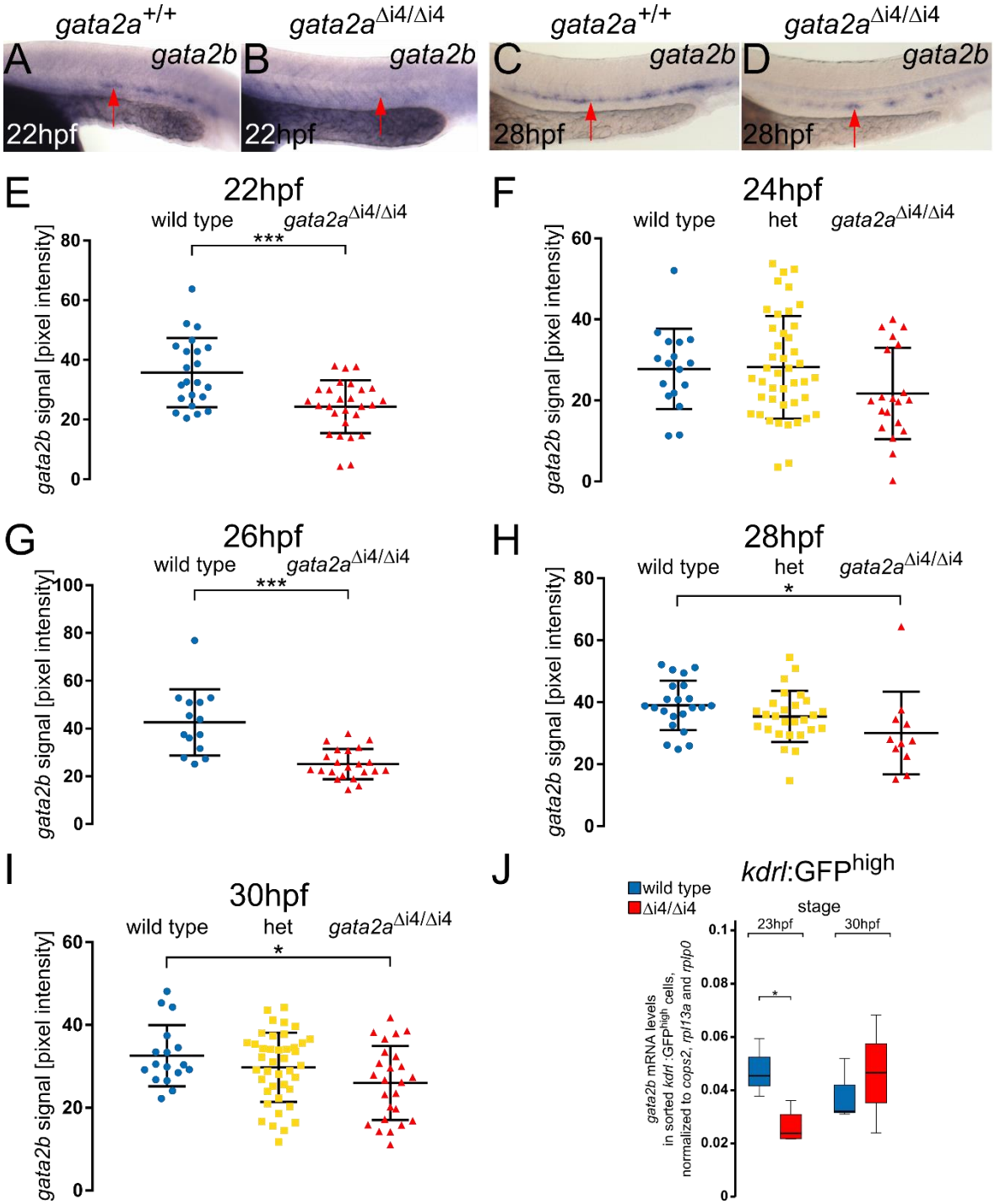


Figure 4.5. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type embryos, express lower levels of *gata2b* in the dorsal aorta during the specification of HE.

(A-D) Representative lateral images of trunks of *gata2b*-probed wild type and *gata2a*$^{\Delta i4/\Delta i4}$ embryos at 22hpf and 28hpf, showing the expression of *gata2b* mRNA in the dorsal aorta (arrows). (E-I) Quantification of the *gata2b* mRNA signal, detected by *in situ* hybridization, from wild type (blue), heterozygous *gata2a*$^{+/\Delta i4}$ (het; yellow) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) siblings at 22-30hpf shows significant decrease in *gata2b* pixel intensity in the DA in the homozygous mutants compared to wt at nearly all stages. Error bars: mean±SD. *p<0.05; ***p<0.001. (J) QRT-PCR in sorted *kdrl*:GFP^{high} cells demonstrates decreased levels of *gata2b* mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf but not at 30hpf, compared to wild type. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples. *p<0.05.

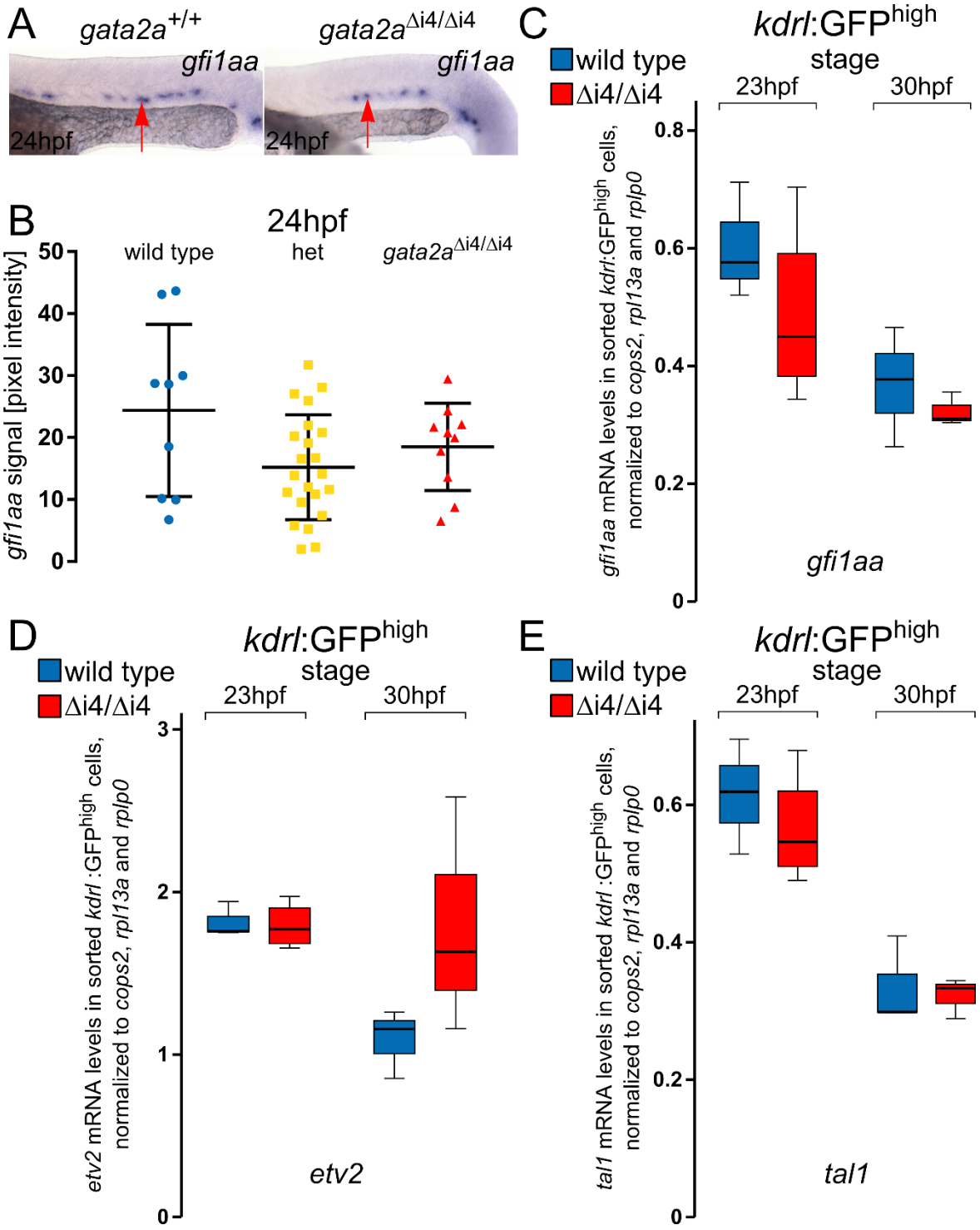


Figure 4.6. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants do not express significantly reduced levels of *gfi1aa*, *etv2* or *tal1*, compared to wild type embryos.

(A) Representative lateral images of the trunks of *gfi1aa*-probed wild type and *gata2a*$^{\Delta i4/\Delta i4}$ embryos at 24hpf, showing the expression of *gfi1aa* mRNA in the dorsal aorta (arrows). (B) Quantification of the *gfi1aa* mRNA signal, detected by *in situ* hybridization, from wild type (blue), heterozygous *gata2a*$^{+/\Delta i4}$ (het; yellow) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) siblings at 24hpf shows a small, but insignificant decrease in *gfi1aa* pixel intensity in the DA in the homozygous mutants compared to wt. Error bars: mean±SD. (C-E) QRT-PCR in sorted *kdrl*:GFP^{high} cells shows no significant changes in the levels of *gfi1aa*, *etv2* or *tal1* mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf or 30hpf, compared to wild type. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples.

4.5.3. Deletion of the *gata2a*-i4 enhancer does not result in complete loss of HE specification

Although *runx1* and *gata2b* are regarded as critical HE markers (Swiers et al. 2010; Butko et al. 2015), we investigated whether other genes expressed in HE during HSC programming were also affected in *gata2a*$^{\Delta i4/\Delta i4}$ mutants.

Gfi1aa, originally identified as a regulator of primitive haematopoiesis (Wei et al. 2008), is expressed in the ventral wall of the DA, the primary site of definitive haematopoiesis (Cooney et al. 2013) (**Fig. 4.6A**). In fact, at 24-36hpf expression of *gfi1aa* can be used as a marker of the haemogenic endothelial tissue independently of *runx1* expression (Cooney et al. 2013). More recently, a gene trap line, *qmc551*, expressing GFP under the endogenously regulated promoter of *gfi1aa*, confirmed the expression of *gfi1aa* in the haemogenic endothelial cells (Thambyrajah et al. 2016). This expression depends on Vegfa and Notch signalling, but not on *runx1* (Thambyrajah et al. 2016). Our quantification of *in situ* hybridization showed no significant difference in the expression of *gfi1aa* in *gata2a*$^{\Delta i4/\Delta i4}$ embryos (n=10) compared to wild type siblings (n=9) at 24hpf (μ_{wt}=24.4, μ_{mut}=18.5) (**Fig. 4.6B**). Further qRT-PCR analysis in sorted *kdrl*:GFP^{+} cells (see **Fig. 4.2E**) confirmed this observation, demonstrating no significant effect on *gfi1aa* expression induced by *gata2a*$^{\Delta i4/\Delta i4}$ homozygous mutation (**Fig. 4.6C**). This suggests that *gata2a* does not regulate the expression of *gfi1aa* in the HE. In addition, it confirms that the HE is still preserved in *gata2a*$^{\Delta i4/\Delta i4}$ embryos and that the reduced expression of *gata2b* and *runx1* that we observed did not result from the loss of the tissue normally expressing these genes.

We also measured the mRNA levels of *tal1*, a conserved gene critical for developmental haematopoiesis (Shivdasani et al. 1995; Patterson et al. 2007) and its upstream regulator *etv2* (Ren et al. 2010; Ciau-Uitz et al. 2013a) using qRT-PCR in sorted *kdrl*:GFP^{+} cells (see **Fig. 4.2E**). We observed no significant changes in endothelial expression of either of the genes in the endothelial *gata2a*$^{\Delta i4/\Delta i4}$ mutants compared to wild type (**Fig 4.6D-E**), suggesting that *gata2a*

is not required for the expression of *etv2* or *tal1* in HE. The former is not surprising, as *Etv2* has been previously shown to be upstream of *Gata2* in the mouse and *Xenopus* (Ciau-Uitz et al. 2013a; Liu et al. 2015). However, their relationship is context-dependent and both transcription factors work together during embryonic haematopoiesis (Ciau-Uitz et al. 2013a; Shi et al. 2014). As for *tal1*, it has been positioned downstream of *gata2* in the genetic cascade responsible for primitive and definitive haematopoiesis in zebrafish PLM (Patterson et al. 2005). However, despite several pieces of evidence for cooperative action of the two proteins (Nottingham et al. 2007; Pimanda et al. 2007b; Tripic et al. 2009; Wu et al. 2014; Org et al. 2015), direct reciprocal transcriptional regulation has not been clearly demonstrated, confounding the efforts to position *Tal1* and *Gata2* in the genetic hierarchy. In fact, Tal1 binds to the *Gata2* +9.5 enhancer in the mouse, suggesting it acts upstream of *Gata2* (Sanalkumar et al. 2014). However, an inverse relationship has also been described, where Gata2 bound to the 3' *Tal1* enhancer (Göttgens et al. 2002). Our data demonstrate that zebrafish *tal1* is not regulated by endothelial *gata2a*. Because *gata2b* is reduced in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants, these data also argue against *tal1* regulation by *gata2b*. Transcriptional regulation of these genes is likely to be complex, context-dependent in different cell types and possibly independent of each other to a certain extent.

4.6. Lack of endothelial *gata2a* does not affect the activity of the *gata2a*-i4 enhancer, while overexpression of *gata2b* increases the activity of the enhancer

Previous work on the +9.5 *Gata2* enhancer in mouse demonstrated that *Gata2* could bind to its intronic enhancer, suggesting an autoregulatory loop (Li et al. 2011; Sanalkumar et al. 2014). In a follow-up study, *Gata2* was found to bind to other enhancers with a similar arrangement of transcription factor binding sites. Genes under the control of these enhancers

were regulated by *Gata2* (Hewitt et al. 2015). However, autoregulatory transcriptional effects of *Gata2* binding to its own enhancer have not been demonstrated or quantified. Therefore, we investigated whether modulation of *Gata2* orthologues in the zebrafish affected the activity of the *gata2a*-i4 enhancer.

Surprisingly, the activity of the *gata2a*-i4 enhancer, measured by the expression of *gfp* in Tg(*gata2a*-i4-450bp:GFP) fish (see Chapter 3) at 28hpf, did not differ between wild type and *gata2a*$^{\Delta i4/\Delta i4}$ siblings (μ_{wt}=25.6, μ_{mut}=24.3) (**Fig. 4.7A**). Because *gata2a* mRNA is severely reduced in the endothelium at this stage (**Fig. 4.2F**), this observation suggests that Gata2a does not regulate its own expression via the i4 enhancer at this developmental stage. Therefore, we tested if modulation of *gata2b* affected the activity of the *gata2a*-i4:GFP transgene. Knockdown of *gata2b* with a MO resulted in a slight, non-significant decrease in *gfp* pixel intensity compared to uninjected embryos (μ_{wt}=28.4, μ_{MO}=25) (**Fig. 4.7B**). This was not surprising, since the expression of *gata2b* is restricted to the haemogenic subset of the DA endothelium (Butko et al. 2015) and its depletion was unlikely to affect the pan-endothelial activity of the *gata2a*-i4 enhancer. However, ectopic overexpression of *gata2b* under the i4 enhancer by injecting the *gata2a*-i4-450bp:*gata2b* construct resulted in a significant increase in the *gfp* pixel intensity, compared to uninjected controls (μ_{wt}=28.4, μ_{OE}=25.1; F=10.257, d.f.=2, 68; p<0.01) (**Fig. 4.7B**). This suggests that expression of *gata2b* in the activity domain of the *gata2a*-i4 enhancer, which includes endothelial and haematopoietic cells, is sufficient to increase *gata2a*-i4 enhancer activity. This positions Gata2b as a potential regulator of *gata2a* through the i4 enhancer, which may involve binding to one of the annotated Gata motifs in the i4 enhancer (see **Fig. 3.2**).

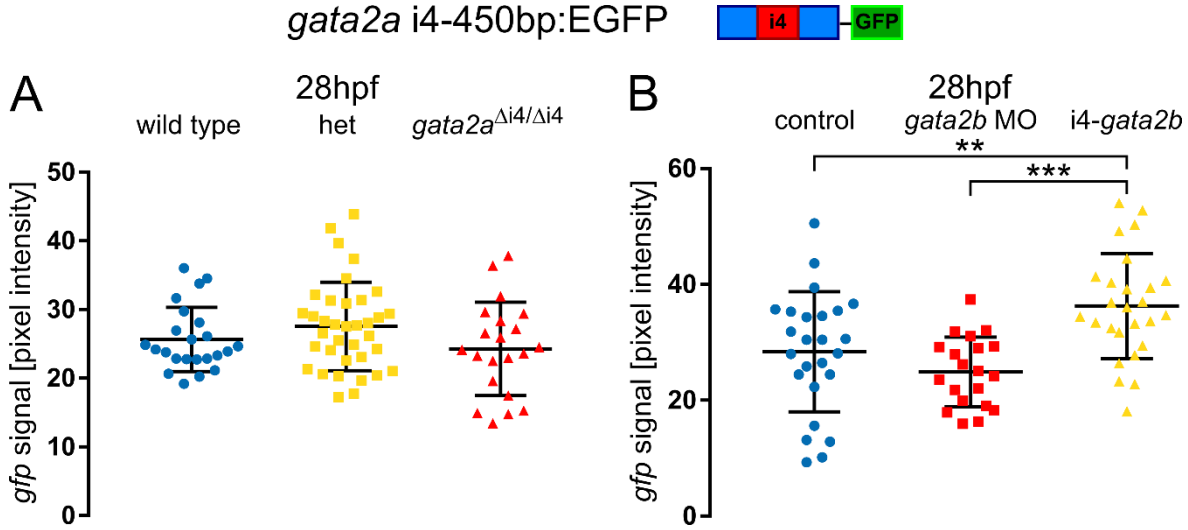


Figure 4.7. Expression of *gata2a*-i4:*gfp* is not affected by the loss of endothelial *gata2a*, but increases upon ectopic overexpression of *gata2b*.

(A) Quantification of the *gfp* mRNA signal, detected by *in situ* hybridization, from non-mutant Tg(*gata2a*-i4-450bp:GFP) (wild type; blue), their transgenic siblings heterozygous for *gata2a*$^{\Delta i4}$ mutation (het; yellow) and *gata2a*$^{\Delta i4/\Delta i4}$ homozygous siblings (red) at 28hpf shows no significant differences in *gfp* pixel intensity in the DA depending on the levels of endothelial *gata2a* expression. Error bars: mean±SD. (B) Quantification of the *gfp* mRNA signal, detected by *in situ* hybridization, from Tg(*gata2a*-i4-450bp:GFP) (control; blue) and their transgenic siblings injected with *gata2b* MO (red) or with *gata2a*-i4-450bp:*gata2b* construct (yellow) demonstrates a slight, but non-significant decrease in *gfp* pixel intensity in the DA in response to *gata2b* knockdown. The *gfp* pixel intensity is significantly increased upon overexpression of *gata2b* under the control of the *gata2a*-i4 enhancer. N=1. Error bars: mean±SD. **p<0.01; ***p<0.001.

4.7. *Gata2a*$^{\Delta i4/\Delta i4}$ mutants recover from the early haematopoietic phenotype

4.7.1. Expression of late markers of embryonic HSC activity is unaffected in *gata2a*$^{\Delta i4/\Delta i4}$ mutants

Deletion of the *Gata2* +9.5 enhancer in mouse dramatically reduces the colony forming and long-term repopulation potential of the embryonic AGM region (Gao et al. 2013). Similarly, heterozygous deletions of the homologous region in human patients were correlated with immunodeficiencies and MonoMAC syndrome (Johnson et al. 2012; Hsu et al. 2013; Spinner et al. 2014), suggesting a long term haematopoietic defect. Therefore, we tested whether

expression of markers of haematopoietic activity in the embryo was affected after the period of HE specification.

Quantitative *in situ* hybridization analysis of *runx1* mRNA in the dorsal aorta at 48hpf showed no significant difference in pixel intensity levels between *gata2a*$^{\Delta i4/\Delta i4}$ mutants (n=27) and wild type controls (n=19) (μ_{wt}=33.1, μ_{mut}=37.5) (**Fig. 4.8A-B**). In fact, the qRT-PCR data from sorted *kdrl*:GFP^{+} cells suggested recovery of *runx1* expression already at 30hpf (**Fig. 4.4H**). These data suggest that the decrease of *runx1* expression at early stages of HE programming in *gata2a*$^{\Delta i4/\Delta i4}$ embryos is transient and recovers by 2dpf.

Expression of *cmyb* in the CHT at around 4dpf can be used to assess the output of HSCs from the DA (Murayama et al. 2006). *In situ* hybridization analysis revealed no difference in *cmyb* signal between *gata2a*$^{\Delta i4/\Delta i4}$ and wild type fish at 4dpf (**Fig. 4.8C-D**), suggesting normal embryonic HSC activity by day 4 in the mutants.

Rag1 marks T lymphocytes in the thymus (Willett et al. 1997), which originate from the DA HSCs (Murayama et al. 2006; Kissa et al. 2008). We assessed thymic *rag1* expression by *in situ* hybridization to compare the T-cell output of embryonic HSCs in *gata2a*$^{\Delta i4/\Delta i4}$ and wild type embryos at 4-5dpf. While at 4dpf around half of the *gata2a*$^{\Delta i4/\Delta i4}$ embryos showed a reduction in the *rag1* signal compared to wild type (**Fig. 4.8E-F**), this effect was absent at 5dpf (**Fig. 4.8G-H**), further suggesting that embryonic haematopoiesis in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants is largely unaffected after 2dpf.

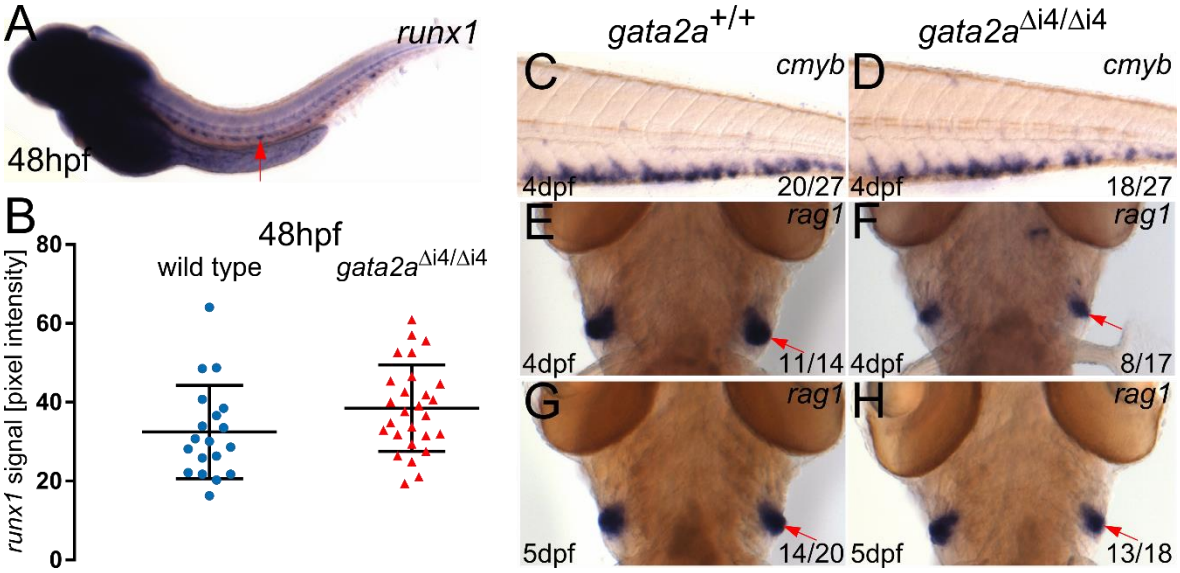


Figure 4.8. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants recover the expression of markers of definitive haematopoiesis.

(A) Representative images of a *runx1*-probed wild type embryo at 48hpf, demonstrating the expression of *runx1* mRNA in the dorsal aorta (arrow). (B) Quantification of the *runx1* mRNA signal, detected by *in situ* hybridization, from wild type (blue) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) siblings at 48hpf shows no significant difference in *runx1* pixel intensity in the DA between the homozygous mutants and wt. Error bars: mean±SD. (C-D) *In situ* hybridization images of the CHT regions, showing no difference in *cmyb* mRNA between wild type and *gata2a*$^{\Delta i4/\Delta i4}$ siblings at 4dpf. (E-H) *In situ* hybridization images of the ventral side of the head, showing a slight decrease (relative to wild type) in *rag1* mRNA in the thymus (arrows) in approximately half of the homozygous mutant embryos at 4dpf. This effect is absent at 5dpf.

4.7.2. The haematopoietic recovery of *gata2a*$^{\Delta i4/\Delta i4}$ mutants is likely mediated by *gata2b* and may involve Wnt signalling pathway

Although the haematopoietic recovery of *gata2a*$^{\Delta i4/\Delta i4}$ embryos disagrees with data from mouse +9.5$^{-/-}$ mutants, we hypothesized that the recovery could be due to the presence of two *Gata2* orthologues in zebrafish, where *gata2a* and *gata2b* would behave as if they were independent *gata2* 'alleles'. *Gata2b* has been shown to be crucial for HE and HSC specification (Butko et al. 2015) and our qRT-PCR data from sorted *kdrl*:GFP^{+} cells suggested recovery of *gata2b* expression in *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 30hpf (**Fig. 4.5H**). Therefore, we investigated whether forced ectopic expression of *gata2b* was sufficient to speed up the haematopoietic recovery of *gata2a*$^{\Delta i4/\Delta i4}$ embryos. We injected wild type and mutant embryos with a construct driving expression of *gata2b* under the control of the *gata2a*-i4-450bp enhancer and measured

runx1 expression at 28hpf in the DA by quantitative *in situ* hybridization. As expected, we observed a significant decrease in *runx1* pixel intensity in *gata2a*$^{\Delta i4/\Delta i4}$ embryos compared to wild type (μ_{wt}=38.8, μ_{mut}=17.9; F=47.373, d.f.=3, 89; p<0.001) (**Fig. 4.9**). Ectopic expression of *gata2b* under the *gata2a*-i4 enhancer significantly increased the *runx1* pixel intensity levels in both wild type (μ_{wt}=38.8, $\mu_{wt+gata2b}$=53.4; p<0.01) and mutants (μ_{mut}=17.9, $\mu_{mut+gata2b}$=33.2; p<0.001) (**Fig. 4.9**). Importantly, it was sufficient to bring the *runx1* expression levels in the mutants up to the levels detected in uninjected wild type fish (μ_{wt}=38.8, $\mu_{mut+gata2b}$=33.2; n.s.) (**Fig. 4.9**), suggesting that expression of *gata2b* was capable of speeding up the haematopoietic recovery, marked by *runx1* expression, in *gata2a*$^{\Delta i4/\Delta i4}$ mutants. Thus, pan-endothelial *gata2b* is able to recover the definitive haematopoietic programme even in the absence of endothelial *gata2a*.

Detailed dissection of the molecular mechanism of this recovery remains beyond the scope of this thesis, but preliminary qRT-PCR data suggest that Wnt signalling pathway may be involved. Wnt signals are important at multiple stages of HSC specification (Richter et al. 2017) and recent data suggest that HE must respond to Wnt cues to generate HSCs (Grainger et al. 2016). Our analysis of mRNA levels in sorted *kdrl*:GFP^{+} cells from *gata2a*$^{\Delta i4/\Delta i4}$ mutants and wild type controls (see **Fig. 4.2E**) showed a significant decrease in the mutants in the expression of members of the Wnt signalling pathway at 23hpf, including *ctnnb2* (t=2.837, d.f.=5, p<0.05), *fzd3a* (t=4.179, d.f.=5, p<0.01), *tcf7* (t=4.271, d.f.=5, p<0.01) and *wnt5b* (t=2.837, d.f.=5, p<0.05) (**Fig. 4.10**). Importantly, expression of all these genes recovered to wild type levels by 30hpf, indicating a potential role for the Wnt pathway in the haematopoietic recovery of *gata2a*$^{\Delta i4/\Delta i4}$ mutants. In contrast, we did not observe changes in endothelial expression of selected genes from the Notch (*dld*, *dll4*, *hes6*) or Tgfβ (*tgfb3*) signalling pathways (**Fig. 4.11**), which are also important for HSC specification (Butko et al. 2016; Monteiro et al. 2016). This does not exclude the possibility that these pathways participate in the recovery, especially because *gata2b* is regulated by Notch (Butko et al. 2015), but these hypotheses would have to be verified by further experiments.

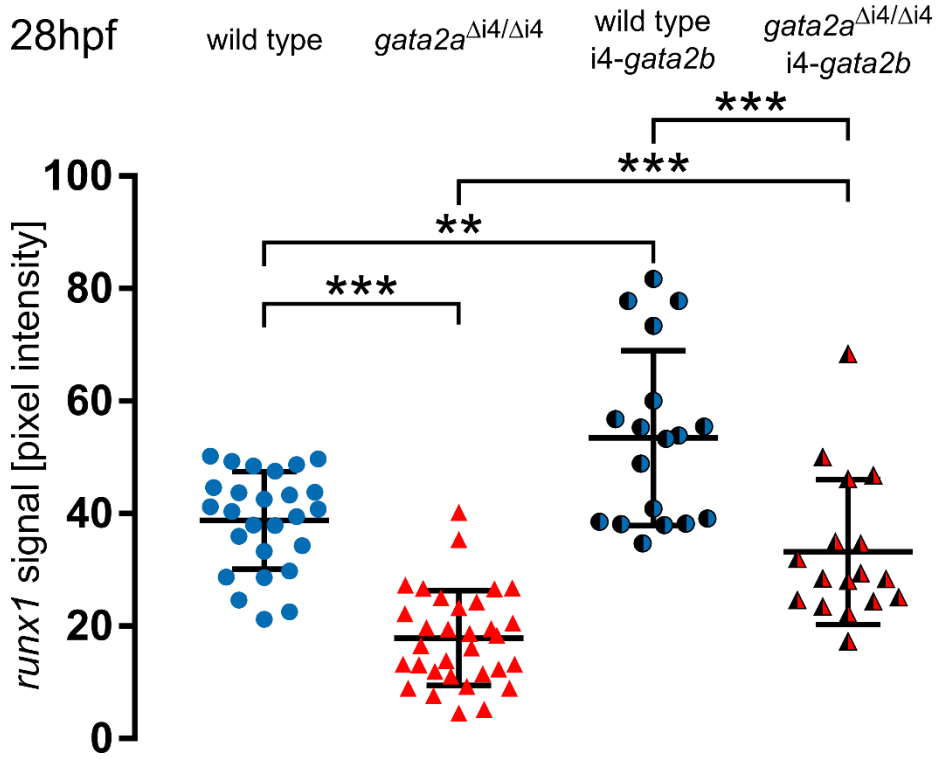


Figure 4.9. Endothelial overexpression of *gata2b* rescues the early *runx1* phenotype of *gata2a*$^{\Delta i4/\Delta i4}$ mutants.

Quantification of the *runx1* mRNA signal, detected by *in situ* hybridization, from 28hpf wild type embryos (blue), *gata2a*$^{\Delta i4/\Delta i4}$ mutants (red) and their siblings injected with *gata2a*-i4-450bp:*gata2b* construct (shaded blue and red). Ectopic expression of *gata2b* significantly increases the *runx1* pixel intensity in the DA in wt and in mutants, rescuing the latter to the non-injected wt expression levels. Error bars: mean±SD. **p<0.01; ***p<0.001.

4.7.3. Adult *gata2a*$^{\Delta i4/\Delta i4}$ mutants suffer from infections and heart oedemas

Although the expression of haematopoietic markers recovers in the *gata2a*$^{\Delta i4/\Delta i4}$ embryos, these mutants may have long term effects on adult haematopoiesis. In fact, we observed a high incidence of infections and heart oedemas in the adult fish, with over 25% suffering from one of these defects by 6 months of age, compared to <1% of wild type fish (**Fig. 4.12**). While the heart oedemas may point to an additional function of the *gata2a*-i4 enhancer in the heart or the lymphatic system, infections may suggest immune defects, as observed in human patients with heterozygous mutations in the homologous enhancer region of *GATA2* (Johnson et al. 2012). Further experiments, employing FACS of whole kidney marrow from adult fish (Traver et al. 2003), will provide better understanding of the long term effects of *gata2a*$^{\Delta i4}$ deletion.

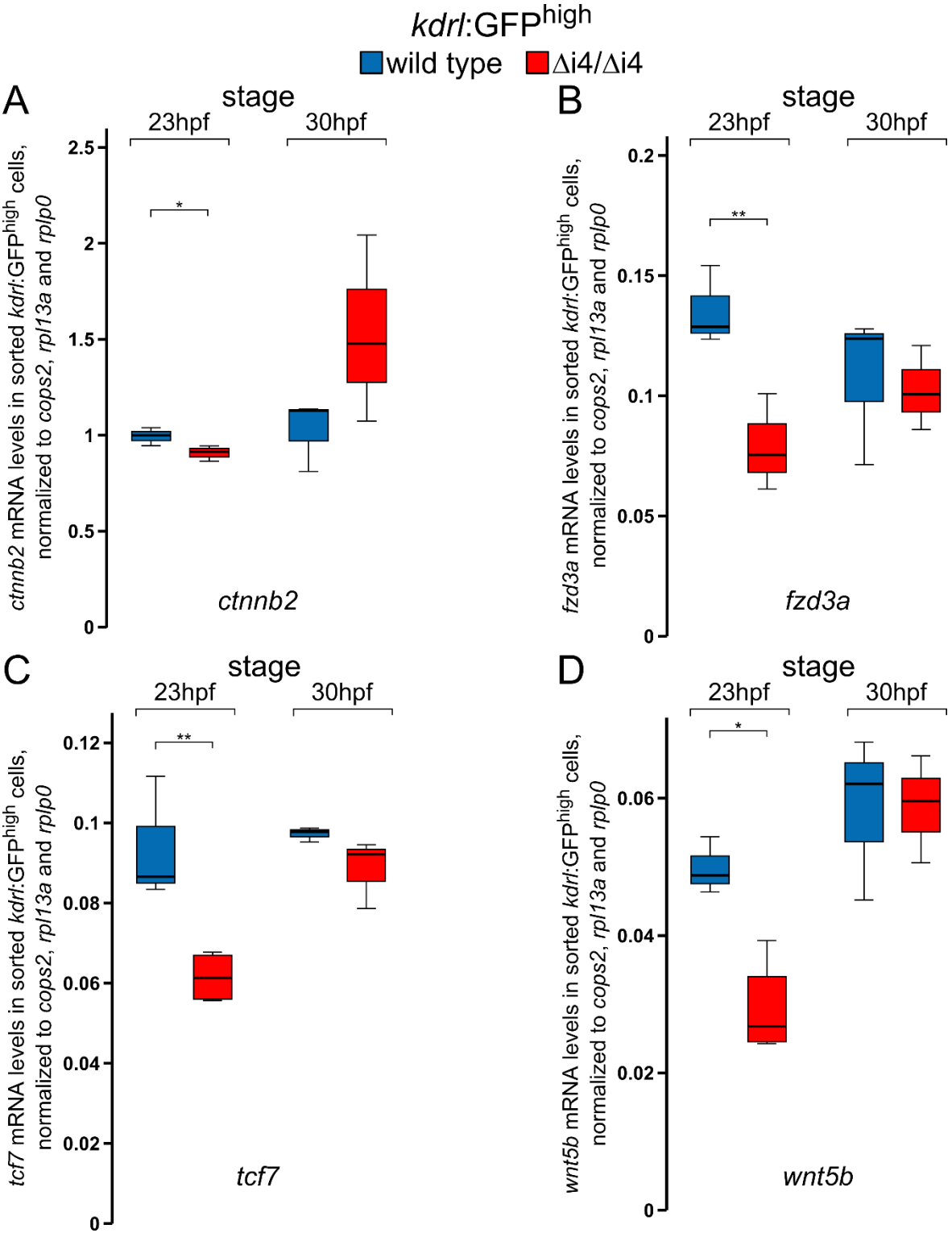


Figure 4.10. Endothelial expression of Wnt signalling genes recovers in *gata2a*$^{\Delta i4/\Delta i4}$ mutants.

QRT-PCR in sorted *kdrl*:GFP^{high} cells from wt (blue) and *gata2a*$^{\Delta i4/\Delta i4}$ mutants (red) shows significant decrease in the levels of *ctnnb2* (A), *fzd3a* (B), *tcf7* (C) and *wnt5b* (D) mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf, compared to wild type. This effect is absent at 30hpf. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples. *p<0.05; **p<0.01.

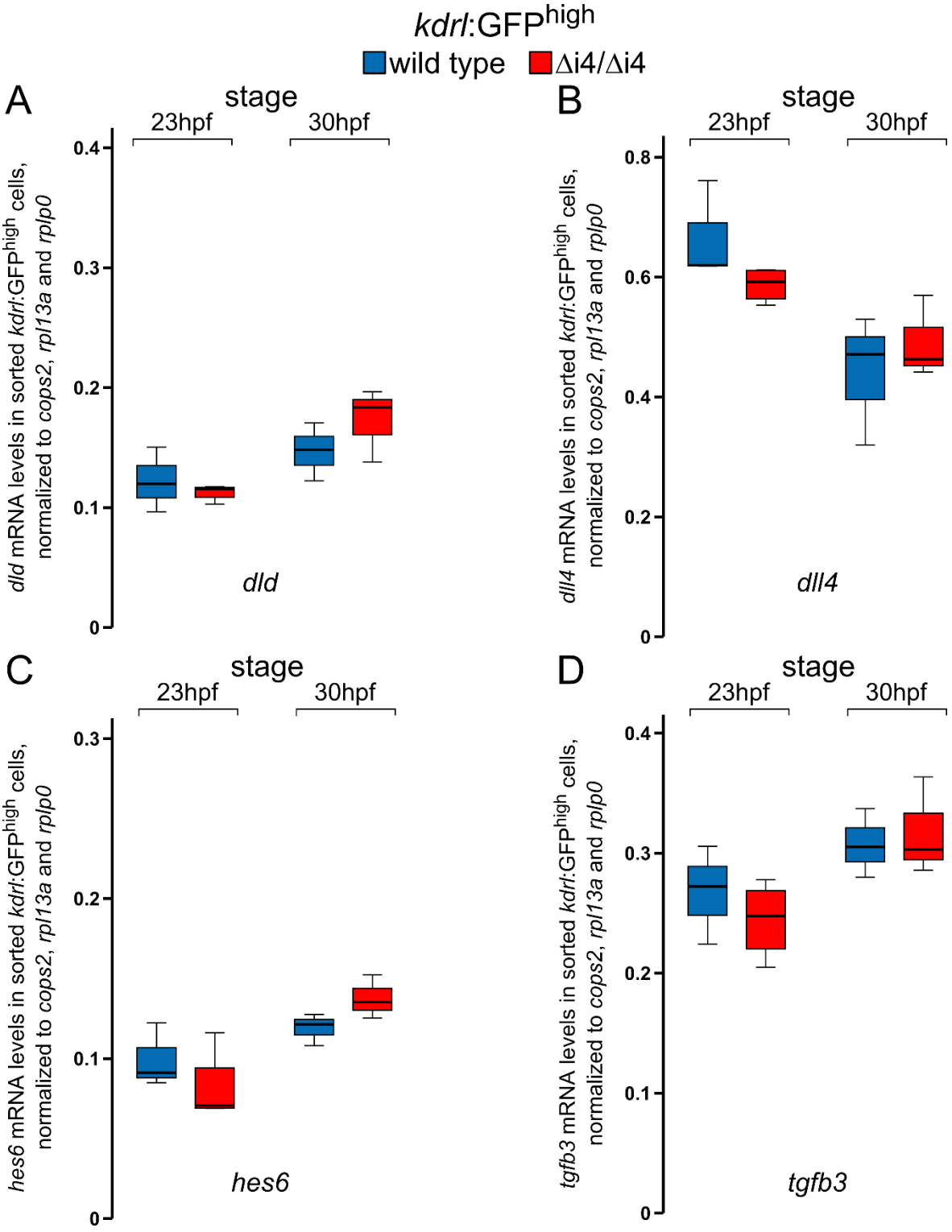


Figure 4.11. Endothelial expression of selected Notch and Tgfβ signalling genes is not affected in *gata2a*$^{\Delta i4/\Delta i4}$ mutants.

QRT-PCR in sorted *kdrl*:GFP^{high} cells from wt (blue) and *gata2a*$^{\Delta i4/\Delta i4}$ mutants (red) shows no significant changes in the levels of *dld* (A), *dll4* (B), *hes6* (C) and *tgfb3* (D) mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf and 30hpf, compared to wild type. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples.

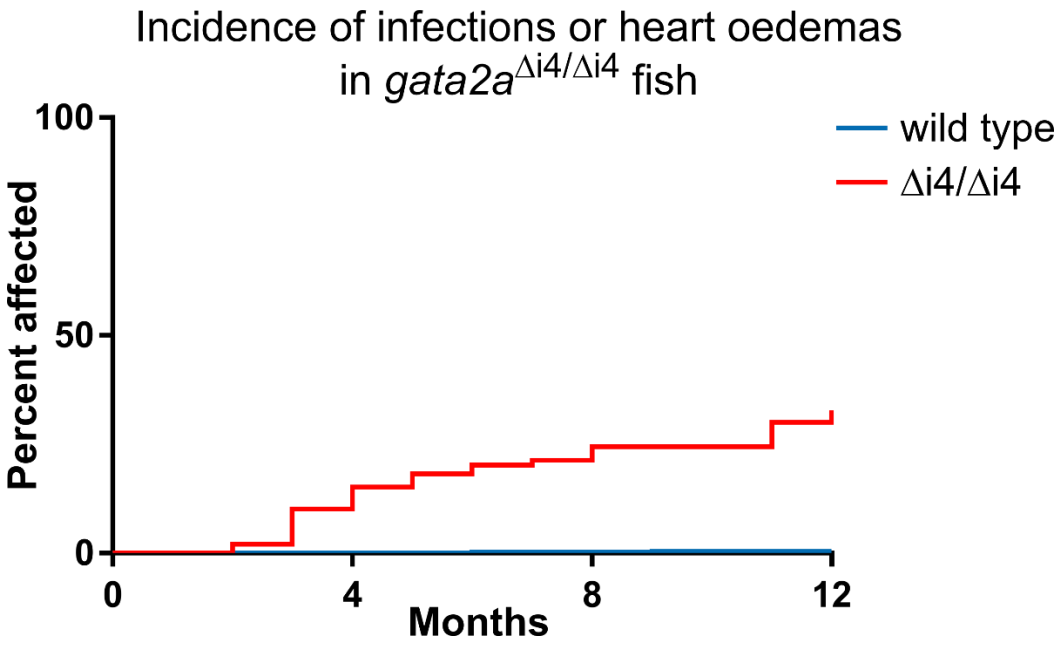


Figure 4.12. *Gata2a*$^{\Delta i4/\Delta i4}$ adult fish suffer from infections and heart oedemas.

Over 25% of *gata2a*$^{\Delta i4/\Delta i4}$ mutants (red) catch infections or suffer from heart oedemas by 6 months. Only around 65% survive infection- or oedema-free for more than 12 months. <1% of wild type fish (blue) exhibit such defects. The graph does not include deaths by other causes.

4.8. Discussion

4.8.1. Overview

Recent development of genome editing tools has provided new, efficient ways of generating targeted gene knockouts for reverse genetic studies in various model organisms. However, results from genetic mutants have challenged previous, widely used knockdown studies by providing contradictory results (Kok et al. 2015; El-Brolosy and Stainier 2017; Ledford 2017). Therefore, there is a need for better understanding of the molecular effects of genomic mutations and their differences from gene knockdowns. Moreover, most studies published to date describe genomic mutations in protein coding regions of targeted genes. The alternative of generating gene knockouts by targeting *cis*-regulatory elements is still in its infancy.

There are very few examples reported to date and, to our knowledge, none in the zebrafish model. However, ongoing advances in high-throughput identification of enhancers (Dailey 2015) provide more and more putative control regions for targeting with genome editing tools. This approach may be particularly useful for genetic studies during embryonic development. Because the activity of developmental enhancers is usually much more restricted than the embryo-wide expression of the genes under their control (Nord et al. 2013), deletion of tissue-specific enhancers could aid generation of tissue-specific mutants, without the need for dual systems, such as Cre-lox. However, such an approach, still in its early days, requires a number of special considerations, such as the size of the intended deletion and potential redundancy of enhancers (Lopes et al. 2016).

Arguably the best evidence in favour of targeting developmental enhancers comes from the mouse model, where deletion of the +9.5 enhancer of *Gata2* resulted in tissue-specific depletion of *Gata2* mRNA in the endothelium, which caused a defect in definitive haematopoiesis (Johnson et al. 2012; Gao et al. 2013). However, there are very few other publications reporting the effects of endogenous enhancer deletions and most of them come from *in vitro* studies. To date, there are no such data from zebrafish, although many enhancers are conserved among vertebrates, at least at the sequence level (Levine 2010). Successful enhancer deletion studies in zebrafish would provide novel tools to study effects of tissue-specific gene knockouts or knockdowns in the zebrafish embryo, a very powerful model for cell and developmental biology, especially in the field of haematopoiesis (Santoriello and Zon 2012; Robertson et al. 2016). In this chapter, we have employed CRISPR/Cas9 to delete a conserved intronic enhancer of the zebrafish *gata2a* gene, described in detail in Chapter 3. We have used the generated mutants to study the role of *gata2a* in definitive haematopoiesis in zebrafish.

4.8.2. A conserved intronic enhancer is required for the endothelial expression of *gata2a* but its deletion does not affect endothelial development

Because of the modular and complex organisation of gene enhancers, deletion of whole regions of putative enhancer activity is the most likely approach to affect gene expression (Lopes et al. 2016). This, however, requires the simultaneous use of two gene editing reagents, each of them targeting a flank of the enhancer. In the case of CRISPR/Cas9, the editing efficiency of paired guides is much lower than the efficiency of single targeting (Canver et al. 2014). However, because of the relatively short generation time of zebrafish and the ability to generate stable mutant lines from a single germline mutation, transient editing efficiency is not a major concern.

We defined the boundaries of the *gata2a*-i4 enhancer based on sequence conservation and ATAC-seq profile (see **Fig. 3.2**). We targeted them with pairs of sgRNAs and D10A Cas9 'nickase' (Mali et al. 2013; Ran et al. 2013b) to reduce potential off-target effects. We isolated two stable mutant lines, both of which carried deletions that included all the highly conserved motifs. In further gene expression analysis, homozygous mutant embryos carrying a 231bp deletion ($gata2a^{\Delta i4/\Delta i4}$) showed reduction in *gata2a* mRNA expression specifically in the endothelial tissue at the stages of HE specification. This agrees with the mouse data, where deletion of the homologous +9.5 enhancer resulted in depletion of *Gata2* specifically in the endothelium and developing haematopoietic lineages, but not in the nervous system (Johnson et al. 2012). This also positions the *gata2a*-i4 enhancer as critical for the endothelial expression of *gata2a*, similar to the mouse +9.5 enhancer, shown to be absolutely required for the haematopoietic expression and function of *Gata2* (Snow et al. 2011; Sanalkumar et al. 2014). Although many *cis* elements may have redundant functions in regulating transcriptional activity (Lopes et al. 2016), our data on the *gata2a*-i4 enhancer suggest that its absolute requirement is conserved across vertebrates.

The previously described *gata2a*^{um27} mutant showed defects in DA morphogenesis, detectable around 48hpf (Zhu et al. 2011). Moreover, mouse +9.5$^{-/-}$ mutants also displayed minor endothelial defects around day E13.5, occurring after the earlier HSPC defect (Johnson et al. 2012). Therefore, we investigated whether the *gata2a*$^{\Delta i4/\Delta i4}$ mutants suffered from disruptive endothelial development by observing *kdrl*-expressing cells *in vivo* or by *in situ* hybridization. To our surprise, we noted no endothelial defects until 48hpf. Although the *gata2a*-i4 enhancer is active throughout the endothelial population and its deletion reduces the expression of *gata2a* in the endothelium, it is possible that the specification of endothelial integrity by *gata2a* occurs independently of the i4 enhancer. We speculate that the defect observed in *gata2a*^{um27} mutants results from an earlier requirement for *gata2a*, before it becomes dependent on i4 enhancer activity. Importantly, the lack of severe endothelial defects allowed us to grow the *gata2a*$^{\Delta i4/\Delta i4}$ fish to adult stages, which facilitated collection of samples for qRT-PCR analyses from homozygous mutants.

Taken together, these data demonstrate that deletion of a conserved *gata2a* endothelial enhancer can achieve a tissue-specific decrease in *gata2a* expression. To our knowledge, this is the first instance when deleting a zebrafish enhancer affects expression of the gene it regulates, leading the way for other studies to achieve tissue-specific gene knockout or knockdown via enhancer deletion or mutation.

4.8.3. Functional analysis of the *gata2a*-i4 reveals non-redundant roles for Gata2a and Gata2b in HE programming

Whole genome duplication during evolution resulted in the generation of multiple paralogues in teleost fish, including the Gata family of transcription factors (Gillis et al. 2009). A recent publication has identified *gata2b* as the critical Notch-regulated marker of HE, required for *runx1* expression in the DA and generation of HSCs (Butko et al. 2015). Because specification of HSPCs is only one of the described roles of Gata2 (Spinner et al. 2014),

subfunctionalisation of the two zebrafish *Gata2* paralogues has been suggested, assigning the haematopoietic role to *gata2b* and the other roles, including endothelial and nervous system development, to *gata2a* (Zhu et al. 2011; Butko et al. 2015). Therefore, and also due to a lack of appropriate gene knockdown or knockout tools, a potential haemogenic function for *gata2a* has not been investigated. Because the mouse *Gata2* $+9.5^{-/-}$ mutants displayed defects in HSPC specification (Gao et al. 2013), we investigated whether zebrafish *gata2a*$^{\Delta i4/\Delta i4}$ mutants displayed a similar phenotype.

Indeed, we observed reduced mRNA expression of *runx1*, the key *bona fide* HE marker (Burns et al. 2002; Kalev-Zylinska et al. 2002; Gering and Patient 2005), in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants in the endothelium at early stages of HE specification and before the first EHT events, which initiate around 30hpf (Kissa and Herbomel 2010). This was also associated with a reduction in endothelial *gata2b* mRNA levels at the same stages, suggesting that *gata2a* is required for expression of *gata2b*, which in turn regulates the expression of *runx1* (Butko et al. 2015) (**Fig. 4.13**). Future live imaging experiments will determine whether the reduction in *gata2b* and *runx1* mRNA levels in the mutants *gata2a*$^{\Delta i4/\Delta i4}$ results in reduced number of EHT events.

In contrast to *runx1* and *gata2b*, we did not observe any reduction in the endothelial expression of other selected HE markers. These included *gfi1aa*, a *runx1*-independent marker of HE not required for HSC specification (Cooney et al. 2013; Thambyrajah et al. 2016), suggesting that the early specification of the pre-HSC structure of HE was not affected in the *gata2a*$^{\Delta i4/\Delta i4}$ embryos. Expression of *etv2*, an upstream regulator of *Gata2* (Ciau-Uitz et al. 2013a; Liu et al. 2015), was unaffected as well, suggesting that this genetic hierarchy is conserved in zebrafish. *Tal1*, a downstream target of *etv2* in the HE specification programme (Ren et al. 2010; Ciau-Uitz et al. 2013a), was also unaffected. Although both *Tal1* and *Gata2* are required for haematopoietic development (Pimanda et al. 2007b), their relationship in zebrafish has not been studied beyond the PLM stage, where *gata2* regulates *tal1* expression (Patterson et al. 2005). These analyses are

further confounded not only by the presence of two *gata2* paralogues, but also by two zebrafish-specific *tal1* isoforms that are required for the HSC specification programme (Zhen et al. 2013). Our results, in which we did not distinguish between the two isoforms, demonstrate that after 23hpf, expression of *tal1* in the DA is not regulated by endothelial *gata2a* or *gata2b*. It is possible that these genes are regulated in parallel, independently of each other.

Because studies on the mouse +9.5 *Gata2* enhancer suggested autoregulation of *Gata2* through this enhancer (Li et al. 2011; Sanalkumar et al. 2014), we also investigated whether modulation of the two *gata2* zebrafish paralogues affected the expression of the transgene driven by the *gata2a*-i4 enhancer, described in Chapter 3. While neither endothelial-specific *gata2a*$^{\Delta i4/\Delta i4}$ knockdown nor systemic *gata2b* MO-mediated knockdown significantly affected the *gata2a*-i4:*gfp* activity at 28hpf, forced expression of *gata2b* under the control of the *gata2a*-i4 enhancer resulted in increased levels of *gfp* mRNA in the DA. This suggests that Gata2b has the potential to modulate the activity of the *gata2a*-i4 enhancer, possibly via direct binding to the conserved Gata motifs. The lack of a major effect in the *gata2b* morphants likely results from the very restricted expression domain of *gata2b* (Butko et al. 2015), affecting only few of the cells normally expressing *gata2a*-i4:*gfp*. These data do not exclude the direct binding of Gata2a to its enhancer at earlier stages of HE specification, but suggest that such binding might only have a transient function. Another possibility involves another Gata factor, Gata3, expressed in the mouse AGM at E11.5 in various cell types, including endothelial cells (Fitch et al. 2012). Analysis of *Gata3* $^{-/-}$ mutants revealed a requirement for Gata3 during the specification of HSPCs (Fitch et al. 2012). In fact, our RNA-seq data (see Chapter 5.3) demonstrate that zebrafish *gata3* is also expressed in the endothelium at 28hpf (see Appendix II), suggesting that it may be involved in regulating the *gata2a*-i4 enhancer. This poses a hypothetical model in which a 'triad' of sequentially acting Gata factors, comprising Gata2a, Gata2b and Gata3, may be involved in regulating the specification of HE in zebrafish. This model is particularly interesting in the light of the genome-

wide 'Gata switches' in mouse and human models. These are defined as *cis*-regulatory regions where Gata1 displaces Gata2 to alter the expression of target genes (Huang et al. 2016). Moreover, Gata2 directly regulates *Gata3* during dendritic cell differentiation in the mouse (Onodera et al. 2016) and it is possible that similar mechanisms are recycled in multiple developmental contexts.

Taken together, these data reveal a previously unreported role for *gata2a* upstream of *gata2b* and *runx1* in early HE specification (**Fig. 4.13**). This genetic cascade appears independent of *etv2* and *tal1*, providing better understanding of the molecular events responsible for HE programming.

4.8.4. Recovery of definitive haematopoiesis in *gata2a*$^{\Delta i4/\Delta i4}$ mutants likely results from the activity of *gata2b*

The expression of *runx1* is widely used as a marker for HSPC specification and reduced expression in the DA is commonly associated with reduced haematopoietic potential (de Bruijn and Dzierzak 2017). However, it has been demonstrated before that, even in the absence of functional *runx1*, zebrafish are able to recover definitive and adult haematopoiesis (Sood et al. 2013). Therefore, we investigated whether a similar recovery mechanism was present in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants, especially in the light of recently described *Gata2*-independent haematopoietic progenitor cells (HPCs) in the mouse (Kaimakis et al. 2016). Although those HPCs differed from *Gata2*$^{+}$ HSCs in their genetic programme and output, they maintained a limited blood-producing potential even in the absence of *Gata2*, demonstrating that not all blood cells absolutely depend on the activity of *Gata2* (Kaimakis et al. 2016).

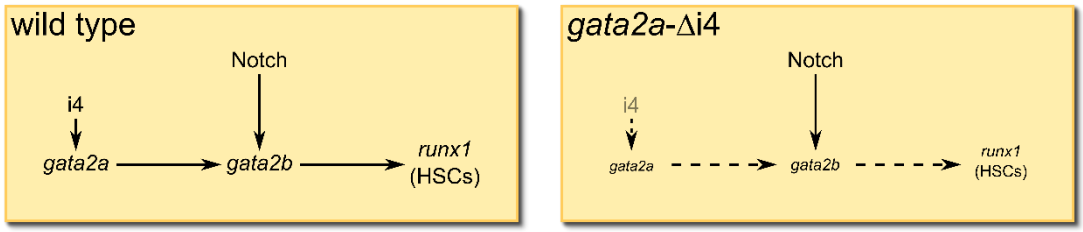


Figure 4.13. Zebrafish *gata2a* regulates the early HE programme through *gata2b*.

Endothelial expression of *gata2a* is regulated through the i4 enhancer. Its activity is required for the expression of *gata2b* in the haemogenic endothelium, which in turn activates *runx1* expression and specifies the HSC programme in the HE. In the absence of endothelial *gata2a* upon the deletion of the *gata2a*-i4 enhancer, expression levels of *gata2b* and *runx1* are initially reduced. However, regulation of *gata2b* by Notch is not affected in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants, which might contribute to the recovery of definitive haematopoiesis in the mutants.

We observed signs of recovered endothelial expression of *runx1* and *gata2b* at 30hpf by qRT-PCR. Although this effect was not apparent in quantitative *in situ* hybridization, this discrepancy likely resulted from expression in EMPs, included in the qRT-PCR but not in the *in situ* hybridization. However, quantitative *in situ* hybridization against *runx1* in the DA at 48hpf demonstrated full recovery of *runx1* expression in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants. This was further associated with normal expression of *cmyb* in the CHT at 4dpf, used as an indicator of the presence of HSCs originating from the DA (Murayama et al. 2006). Probing for the expression of *rag1* in the thymus (Willett et al. 1997), which marks T cells, the first differentiated derivatives of DA HSCs (Murayama et al. 2006), showed a slight decrease at 4dpf, but not at 5dpf. This suggests that although there may be a delay in the generation and functioning of HSCs in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants, the zebrafish have apparently normal definitive haematopoietic programme.

Interestingly, mouse +9.5$^{-/-}$ mutants suffered from long term haematopoietic defects (Gao et al. 2013). We hypothesized that this discrepancy could be due to the presence of two homologues of *Gata2* in the zebrafish (Gillis et al. 2009), despite the fact that Gata2a and Gata2b

are only about 50% identical at the protein level (Butko et al. 2015). Indeed, forced expression of *gata2b* under the *gata2a*-i4 enhancer rescued the expression of *runx1* in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants to wild type levels. A complementary experiment to further validate this hypothesis would involve depletion of *gata2b* in the *gata2a*$^{\Delta i4/\Delta i4}$ mutants, either with MO (Butko et al. 2015) or by using *gata2b* mutants (E. de Pater, personal communication). Lack of recovery in these embryos would confirm the requirement for *gata2b* to rescue haematopoiesis in the fish lacking endothelial *gata2a*. Another Gata factor involved in the rescue could be Gata3, previously demonstrated to be required for HSPC specification (Fitch et al. 2012). Future projects will investigate this hypothesis.

Gata2 levels in the HE are controlled by Notch signalling in the mouse embryo (Robert-Moreno et al. 2005) and this mechanism is conserved in the case of zebrafish *gata2b* (Butko et al. 2015). Therefore, we speculated that the recovery of *gata2b* expression and the subsequent haemogenic programme in *gata2a*$^{\Delta i4/\Delta i4}$ mutants could be due to the activity of Notch signalling. Indeed, we did not observe changes in expression in members of the Notch signalling pathway, *dld*, *dll4*, *hes6*, normally expressed in the endothelium. However, validation of this hypothesis would also require profiling of *notch1*, the key Notch receptor in the HE (Robert-Moreno et al. 2005; Kim et al. 2014; Butko et al. 2015). A complementary future experiment would involve timed blocking of Notch signalling in *gata2a*$^{\Delta i4/\Delta i4}$ mutants. Lack of recovery in these embryos would strongly suggest for the role of Notch-*gata2b* axis in the haematopoietic rescue of the fish lacking endothelial *gata2a*.

In contrast to Notch, we observed reduced expression of genes involved in the Wnt signalling pathway: *ctnnb2*, *fzd3a*, *tcf7* and *wnt5b* at 23hpf in the *gata2a*$^{\Delta i4/\Delta i4}$ embryos, which recovered by 30hpf to wild type levels. Wnt signalling is required for the generation of HSPCs (Richter et al. 2017) and there have been suggested links between Wnt and Notch pathways in the somites (Clements et al. 2011; Genthe and Clements 2017). However, the possibility of

Wnt-mediated haematopoietic rescue of *gata2a*$^{\Delta i4/\Delta i4}$ mutants is hypothetical and based on preliminary data. In particular, there are no currently reported direct links between Wnt and *Gata2* expression.

Although the *gata2a*$^{\Delta i4/\Delta i4}$ embryos recovered the expression of HSC markers after 2dpf, we observed a high incidence of infections and heart oedemas in the adults. FACS analysis of whole kidney marrow in *runx1*$^{W84X/W84X}$ mutants, which apparently recovered definitive haematopoiesis despite the lack of *runx1*, showed a decrease in the total number of blood cells in the kidney, especially myeloid and progenitor cells (Sood et al. 2010). This suggests lineage-specific differentiation defects, which may also be present in the *gata2a*$^{\Delta i4/\Delta i4}$ adult fish. This idea is supported by studies from the mouse system, where different subsets of HE, depending on their molecular profile, have been suggested to give rise to different blood populations. For instance, *Sox17* expression in *in vitro* model of HE was required for the generation of T lymphocytes, but not erythroid or myeloid cells (Clarke et al. 2013). Another example includes *Cbfβ*, expression of which distinguished EMPs from HSCs at the level of HE (Chen et al. 2011). We speculate that *gata2a* may be required for the specification of a lineage-specific HE subset. Future experiments, involving FACS of whole kidney marrow, will test this hypothesis and potentially provide novel insights into the role of *gata2a*, regulated by the conserved i4 enhancer, in specifying the adult blood lineages.

Chapter 5

The role of *lmo4a* in the specification of HSCs

5.1. Introduction

The activity of developmental enhancers depends on binding of transcription factors (TFs) to the DNA. Through accompanying DNA looping, the TFs bound to these enhancer elements are brought together and towards promoter regions to initiate expression of target genes (Levine 2010). However, not every enhancer can be bound by every transcription factor. Enhancers are composed of modules of transcription factor binding sites (motifs) that determine their function (Istrail and Davidson 2005). Through enhancers specific TFs achieve synergy to trigger tissue-specific transcription of target genes (Dailey 2015). Because the activity of enhancers depends on combinatorial expression and binding of multiple transcription factors, they only become active when several defined TFs are present in the same cell (Spitz and Furlong 2012). This observation suggests a model where enhancer regions are bound by multi-protein complexes of transcription factors.

One of the best characterised examples of such regulatory complexes involved in the development of the haematopoietic system is the pentameric complex, consisting of Ldb1, Lmo2, E47, Gata1 and Tal1 (Wadman et al. 1997; El Omari et al. 2013). This complex recognizes adjacent E-box and GATA motifs through the DNA-binding domains of Gata1 and Tal1 (Wadman et al. 1997) and has been shown to induce the erythroid differentiation programme in *Xenopus* (Mead et al. 2001). A variation of this complex, composed of Ldb1, Lmo2, Tal1, an Ets factor and Gata2, has been proposed to bind to a *runx1* enhancer and regulate its transcription in the definitive haematopoietic lineage, based on *in vitro* ChIP data from mouse 416B myeloid progenitor cells (Nottingham et al. 2007). A similar complex may recognize E-box-(8bp)-GATA composite elements (Wozniak et al. 2008) to regulate the expression of *Gata2* through the +9.5 enhancer (Sanalkumar et al. 2014). Recently, such composite elements have been found more widespread throughout the genome, suggesting more global regulation driven by Gata2-containing protein complexes in the development of

HSPCs (Hewitt et al. 2015). Moreover, Tal1-Gata complexes have also been shown to recognize less stringent motifs of CTG-(7-9bp)-Gata in erythroid differentiation, where the CTG represents half of the E-box consensus site (Kassouf et al. 2010; Soler et al. 2010).

In the above complexes, recognition of the DNA motifs is mediated by the DNA-binding domains of Gata and E-box factors. Ldb (Lim domain binding) and Lmo (Lim domain only) proteins, integral components of these transcription factor machineries, do not have DNA-binding capacities and their role is to establish a scaffold integrating the DNA-binding factors (Sang et al. 2014). Proteins from the Lmo family are conserved from the nematode *C. elegans* to the human and four of them have been identified to date: Lmo1-4 (Sang et al. 2014). The function of Lmo1 has mostly been assigned to adult stages through its involvement in T-cell acute lymphoblastic leukaemia, while Lmo3 primarily functions in the developing and adult nervous system (Sang et al. 2014). Lmo2, on the other hand, has been widely studied for its role in developmental haematopoiesis. *Lmo2* knockout mice are embryonic lethal due to defective development of blood cells (Warren et al. 1994) and the loss of *Lmo2* results in the loss of haematopoietic repopulation capacity in adult mice (Yamada et al. 1998). It is also required for embryonic haematopoiesis in zebrafish (Patterson et al. 2007).

Lmo4 is structurally the most distant member of the Lmo protein family and its expression pattern is more ubiquitous than those of other Lmos (Sang et al. 2014). It has mostly been studied in the context of breast cancer (Montañez-Wiscovich et al. 2009; Montañez-Wiscovich et al. 2010). Several protein partners of Lmo4 have been identified, supporting its role as a scaffold protein in transcription factor complexes (Sang et al. 2014). In particular, it was identified as a binding partner of Ldb1 (Grutz et al. 1998; Meier et al. 2006), an important regulator of definitive haematopoiesis (Li et al. 2011) that binds to the +9.5 *Gata2* enhancer in the mouse (Sanalkumar et al. 2014). Moreover, previous collaborative work of our group identified other protein partners of Lmo4, comprising a few negative regulators of transcription:

Hdac7, Bcor, Ncor or Ep400 (K. Tuladhar, DPhil thesis; C. Andrieu-Soler, unpublished data). Further functional analyses suggested a role for an Lmo4-nucleated protein complex in negative regulation of gene expression in the developing haematopoietic lineage (K. Tuladhar, DPhil thesis). In the zebrafish, because of the whole genome duplication event during the evolution of the teleosts (Taylor et al. 2003), there are two *lmo4* homologues: *lmo4a* and *lmo4b*. There are currently very few publications studying the roles of these genes. For instance, *lmo4b* has been associated with forebrain development (McCollum et al. 2007), while the expression pattern of *lmo4a* includes somitic mesoderm, the cardiovascular system including the DA, branchial arches, limb buds and nervous system structures (Lane et al. 2002). In addition, MO-mediated knockdown of *lmo4a* results in the reduction of *runx1* expression in the dorsal aorta (Meier et al. 2006). Work in our lab analysed both paralogues in more detail, revealing 93.4% homology between the two proteins (K. Tuladhar, DPhil thesis). Lmo4b is more closely related to the human Lmo4 than Lmo4a and the expression pattern of *lmo4b* resembled more closely the expression pattern of mouse and frog *Lmo4* (K. Tuladhar, DPhil thesis). Further functional analyses suggested subfunctionalisation of the zebrafish *lmo4* paralogues, with *lmo4b* being required for the specification of multiple tissues before 26hpf, while *lmo4a* was found important for the specification of HE and HSCs (K. Tuladhar, DPhil thesis). Its molecular function was assigned to negative fine-tuning of the levels of *gata2a* expression, but mechanistic details of this regulation have not been resolved.

This chapter presents the analysis of the effects of *lmo4a* MO-mediated knockdown on the transcriptome of endothelial cells, including the HE, during HE specification stages. We confirm previous observations that *lmo4a* is a negative regulator of *gata2a*. This regulation may be direct through the *gata2a*-i4 enhancer. We also identify genes potentially upstream of *gata2a* affected by the *lmo4a* knockdown. We suggest that the Wnt signalling pathway may be mediating the regulation between *lmo4a* and *gata2a*. We also present preliminary data from

TALEN-generated *lmo4a* knockout mutants confirming the upregulation of *gata2a* upon *lmo4a* loss of function.

5.2. Characterisation of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cell population for the transcriptomic profiling of endothelium at the stage of HE specification

Expression of the Vegf receptor *kdrl* is one of the earliest markers of endothelial cells in the embryo (Jin et al. 2005). In the Tg(*kdrl*:GFP) transgenic line, GFP marks all endothelium (Jin et al. 2005) and appears more tissue-specific than, for instance, Tg(*fli1*:GFP), where GFP expression was also strong in primitive erythrocytes (Lawson and Weinstein 2002). Therefore, we decided to use the Tg(*kdrl*:GFP) embryos to identify the most appropriate cell population for transcriptomic profiling of endothelial cells, including the HE, at 28hpf. FACS of Tg(*kdrl*:GFP) transgenic embryos revealed three distinct populations along the GFP intensity axis (**Fig. 5.1A**), which we termed *kdrl*:GFP^{-}, *kdrl*:GFP^{low} and *kdrl*:GFP^{high} (K. Clark, WIMM FACS facility). We tested whether the low levels of GFP fluorescence could be attributed to GFP protein being carried over from earlier embryonic stages by measuring the expression of endogenous *kdrl* mRNA in these populations. Indeed, the *kdrl* mRNA expression levels in the *kdrl*:GFP^{low} population were much lower than in the *kdrl*:GFP^{high} cells and comparable to those in *kdrl*:GFP^{-} population (**Fig. 5.1B**), suggesting that only high levels of GFP fluorescence marked true, *kdrl*-expressing, endothelial cells. To further support this observation, qRT-PCR data revealed strong enrichment for other endothelial markers, including HE markers, in the *kdrl*:GFP^{high} over *kdrl*:GFP^{low} cells (R. Monteiro, unpublished data). To our surprise, we also observed expression of *hbbe1.1* and *mpx* mRNA in the *kdrl*:GFP^{high} cells (data not shown), suggesting the presence of cells of erythroid (Quinkertz and Campos-Ortega 1999) or myeloid

(Bennett et al. 2001) fate, respectively. Therefore, we hypothesized that these cells could be sorted out by negatively selecting for the expression of the *gata1*:DsRed transgene that marks erythroid cells and their precursors (Traver et al. 2003). Indeed, qRT-PCR analysis of six distinct cell populations distinguished by the levels of GFP and DsRed fluorescence in Tg(*kdrl*:GFP; *gata1*:DsRed) embryos revealed that the *gata1*$^{+}$ populations expressed *hbbe1.1* and *mpx*, even when they were marked with high levels of *kdrl*:GFP fluorescence (**Fig. 5.1B**). We hypothesized that the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells could be EMPs, previously characterized by the expression of the *gata1*-driven transgene (Bertrand et al. 2007). Further transcriptomic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} and *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells with Fluidigm Biomark™ showed that the cells without *gata1*:DsRed fluorescence expressed significantly higher levels of endothelial markers, compared to *gata1*:DsRed^{+} cells (**Fig. 5.1C**). By contrast, the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells had significantly higher levels of mRNA expression of markers of haematopoietic cells (**Fig. 5.1C**). This likely results from the fact that there were on average five times as many *kdrl*:GFP^{high}; *gata1*:DsRed^{-} as *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells (**Fig. 5.1A**), resulting in more diluted representation of cells with haemogenic fate within the *kdrl*:GFP^{high}; *gata1*:DsRed^{-} population. However, expression of *cmyb*, previously described as one of the defining markers of mouse EMPs (Hoeffel et al. 2015), was particularly strongly enriched in the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells. In summary, these data suggest that the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells may represent the EMP population, whereas the *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cell population is highly appropriate for isolation of endothelial cells, including the HE, at 28hpf.

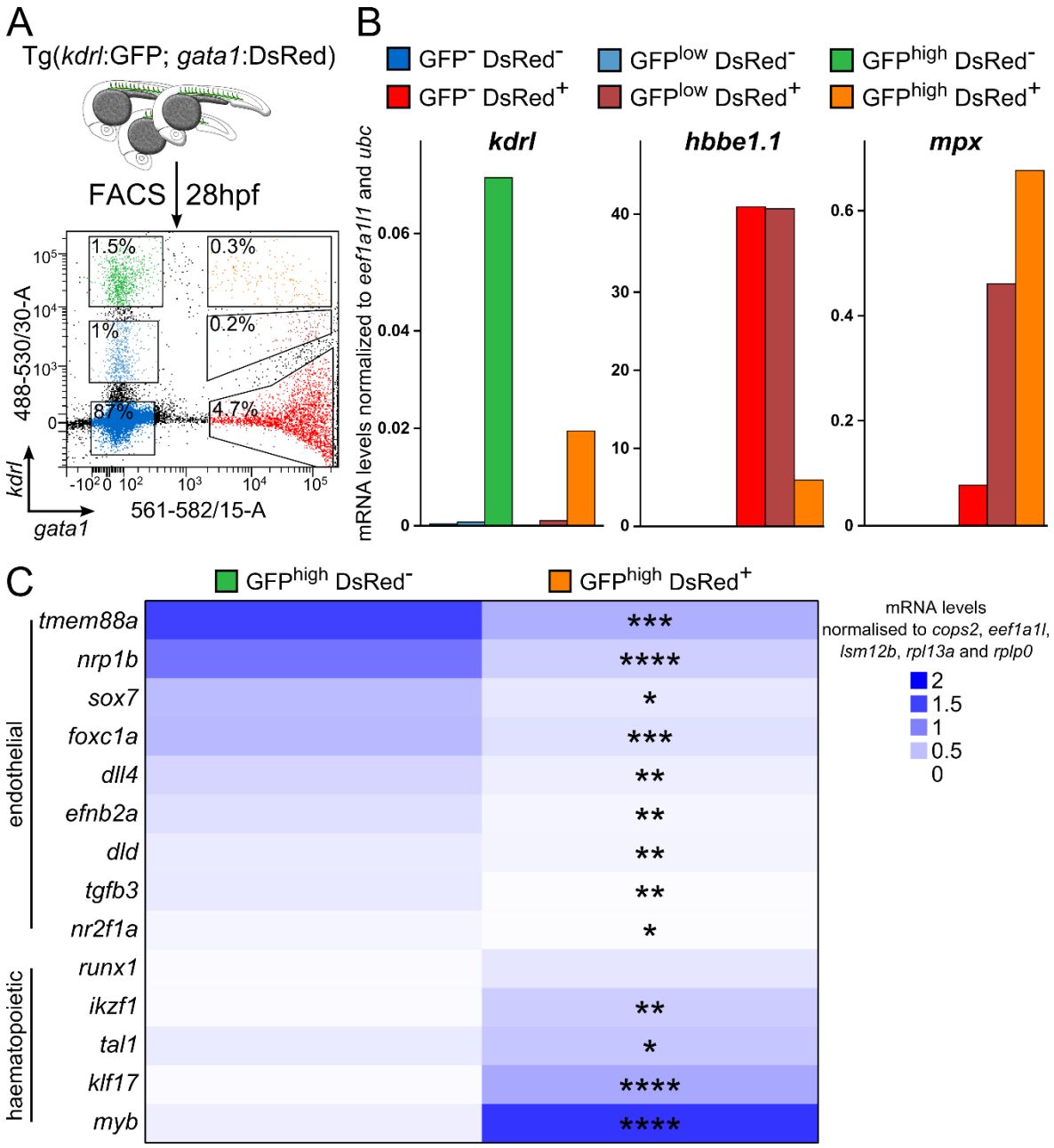


Figure 5.1. Genetic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells reveals their endothelial identity.

(A) Six cell populations were sorted from 28hpf Tg(*kdrl*:GFP; *gata1*:DsRed) zebrafish embryos using FACS: *kdrl*:GFP^{-}; *gata1*:DsRed^{-} (blue), *kdrl*:GFP^{low}; *gata1*:DsRed^{-} (light blue), *kdrl*:GFP^{high}; *gata1*:DsRed^{-} (green), *kdrl*:GFP^{-}; *gata1*:DsRed^{+} (red), *kdrl*:GFP^{low}; *gata1*:DsRed^{+} (purple) and *kdrl*:GFP^{high}; *gata1*:DsRed^{+} (orange). (B) QRT-PCR on RNA isolated from these cell populations reveals that the fractions with high *kdrl*:GFP levels express much higher levels of endogenous *kdrl* mRNA than other cells. Expression of *hbbe1.1* is only detectable in *gata1*:DsRed^{+} fractions. Expression of *mpx* mRNA is highest in the *kdrl*:GFP^{+}; *gata1*:DsRed^{+} (double positive) cells. N=1. (C) Comparison of gene expression profiles between *kdrl*:GFP^{high}; *gata1*:DsRed^{-} and *kdrl*:GFP^{high}; *gata1*:DsRed^{+} shows enrichment for endothelial markers in the *kdrl*:GFP^{high}; *gata1*:DsRed^{-} population and enrichment for genes expressed in haematopoietic progenitors in the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells. N=3. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

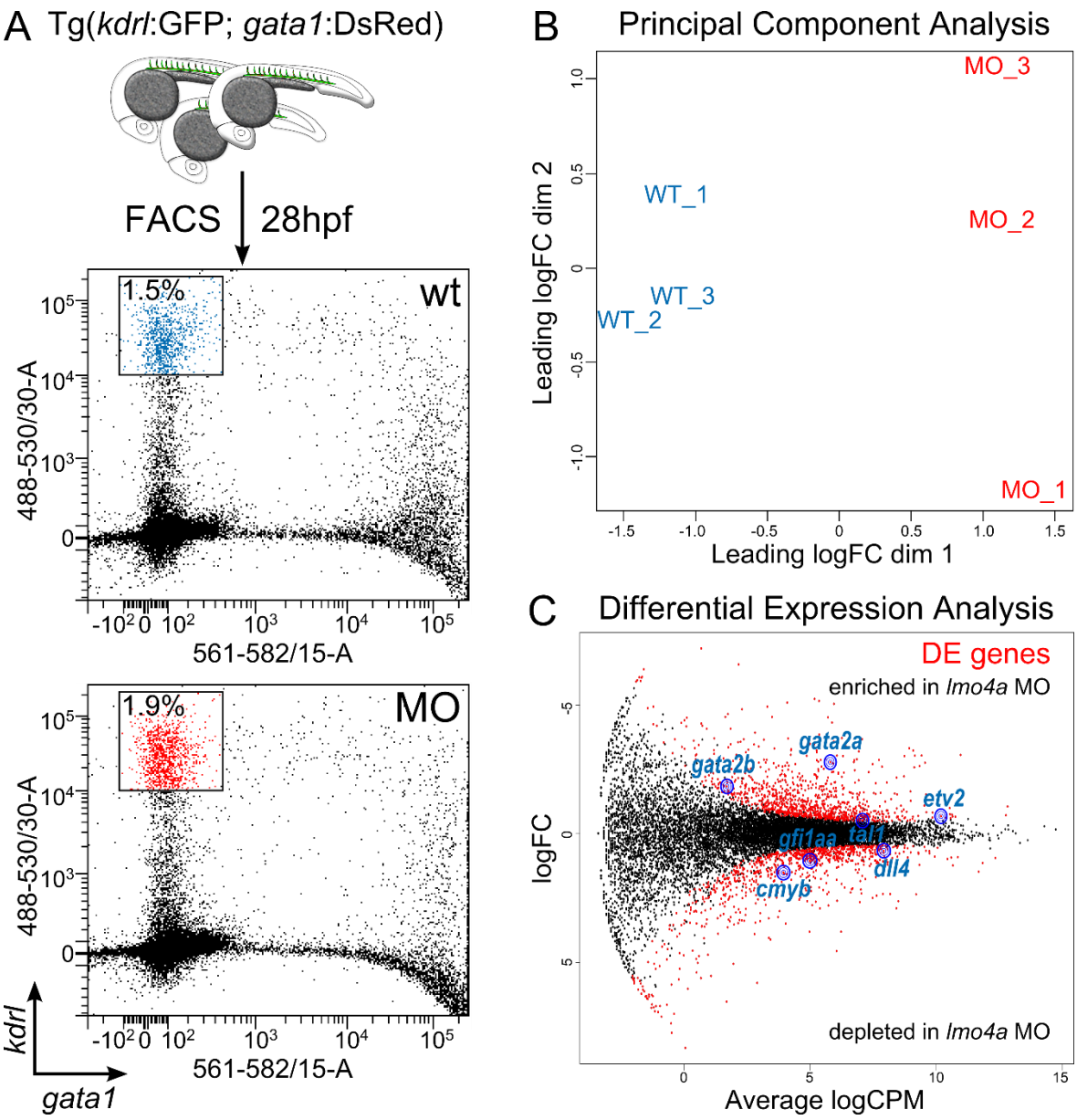


Figure 5.2. Transcriptomic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells with RNA-seq reveals multiple differentially expressed genes upon *lmo4a* knockdown.

(A) *Kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells were sorted from 28hpf wild type (wt; blue) and *lmo4a* MO-injected embryos (MO; red). (B) Principal Component Analysis of RNA sequencing reads (by Rossella Rispoli) demonstrates clear separation of the 3 wt and 3 MO replicas. (C) Differential Expression (DE) Analysis of RNA sequencing reads from wt and *lmo4a* MO samples (by Rossella Rispoli) reveals over 1600 genes (red dots) enriched (top part) or depleted (bottom part) in the morphants, compared to wt. Selected genes known to be involved in haematopoietic development are marked in blue. FC: fold change; CPM: counts per million.

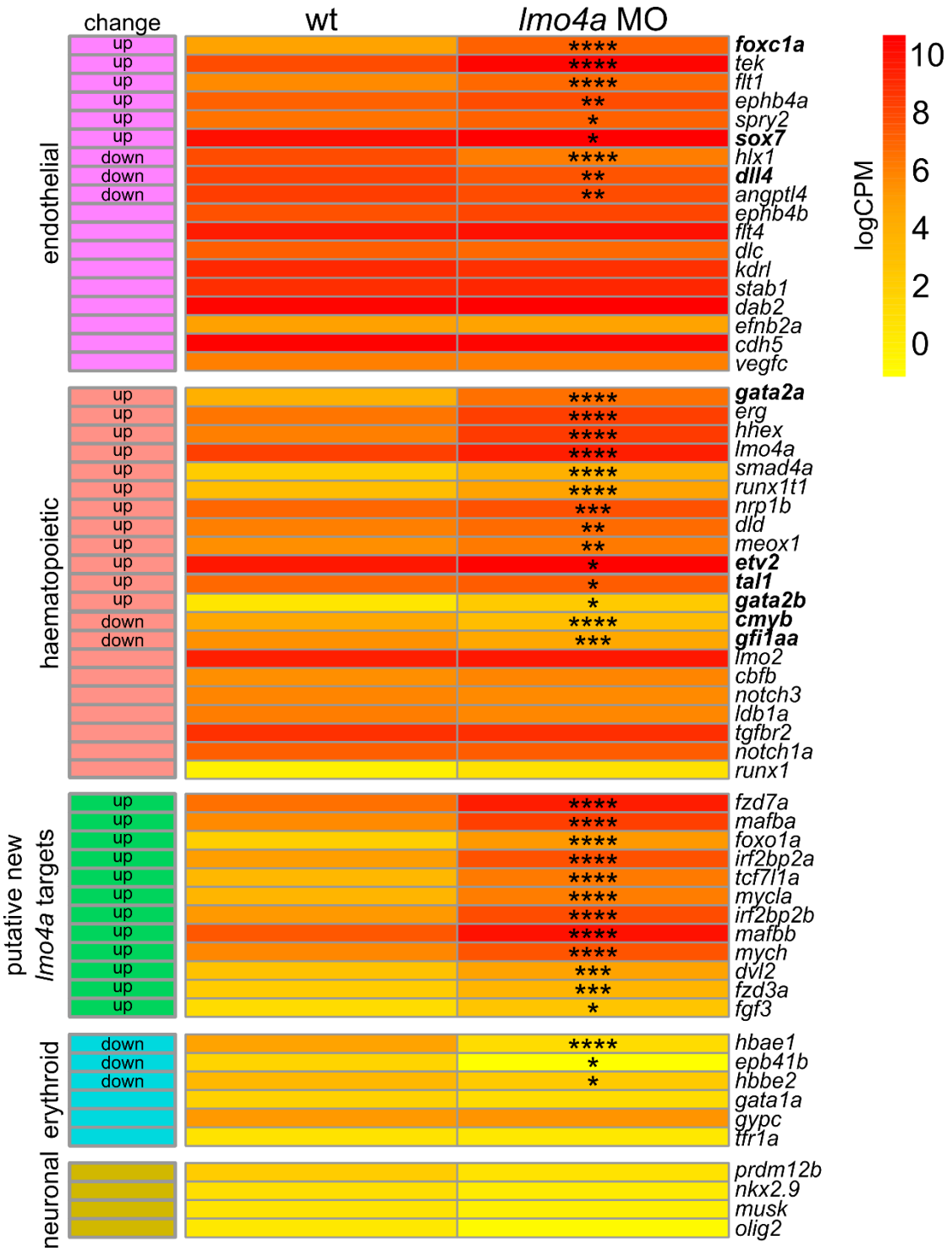


Figure 5.3. Transcriptomic profiling of *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells with RNA-seq confirms their endothelial identity and reveals differentially expressed genes upon *lmo4a* knockdown.

Expression profiling of selected genes in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from 28hpf wild type embryos (wt) and *lmo4a* morphants demonstrates that genes associated with endothelial (violet) or haematopoietic (pink) fate are expressed at higher levels in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} than erythroid (cyan) or neuronal (yellow) genes. Expression of multiple markers of endothelial or haematopoietic cell fate changes significantly (up: upregulation; down: downregulation) upon *lmo4a* knockdown compared to uninjected controls. In addition, novel putative *lmo4a* targets enriched in the morphants have been identified (green). Gene names in bold correspond to the genes discussed in Figures 5.2 and 5.4-5.8. N=3. *FDR<0.05, **FDR<0.01, ***FDR<0.001, ****FDR<0.0001. Figure provided by Rossella Rispoli.

5.3. RNA-seq in *lmo4a* morphants reveals changes in expression of known haemogenic endothelial genes

Lmo4a is expressed in the dorsal aorta at 24hpf (Lane et al. 2002; K. Tuladhar, DPhil thesis), suggesting that it could play a role in the programming of HSCs. Previous studies on the function of *lmo4a* in the specification of HE have been limited to the analysis of the expression of candidate genes (Meier et al. 2006; K. Tuladhar, DPhil thesis). To characterize genome-wide molecular effects of *lmo4a* knockdown in the endothelium, we conducted an RNA sequencing (RNA-seq) experiment on *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells sorted at 28hpf from *lmo4a* MO-injected embryos and uninjected control siblings (**Fig. 5.2A**). Principal Component Analysis showed good separation of the RNA sequencing reads from wild type and morphant samples, confirming the high quality of the data (**Fig. 5.2B**). Differential Expression analysis revealed 1606 genes affected by the knockdown of *lmo4a*, comprising 907 genes with elevated mRNA levels in the endothelium in *lmo4a* morphants and 699 genes downregulated in *lmo4a* morphants (**Fig. 5.2C**). These included known genes involved in HE programming. We observed upregulation of *gata2a*, *gata2b*, *etv2* and *tal1* and downregulation of *dll4*, *gfi1aa* and *cmyb* upon *lmo4a* knockdown (**Fig. 5.2C, Fig. 5.3**). Appendix II contains the list of all identified differentially expressed genes.

Further analysis of the RNA-seq data revealed more endothelial and haematopoietic genes affected by *lmo4a* knockdown (**Fig. 5.3**). For instance, we observed an increase in mRNA levels of *foxc1a* and *sox7*, early regulators of endothelial and HE specification (Costa et al. 2012b; Veldman and Lin 2012), in *lmo4a* morphants compared to wild types (**Fig. 5.3**). On the other hand, we found no changes in expression of *kdrl*, *dab2*, *flt4*, *efnb2a* or *ephb4b*, suggesting correct endothelial development and arterio-venous specification in *lmo4a* morphants (**Fig. 5.3**), in agreement with previous observations (K. Tuladhar, DPhil thesis). The analysis also revealed several putative novel targets of *lmo4a*, not previously associated with haematopoietic

development (**Fig. 5.3**). We selected 84 of the differentially expressed genes for further validation with Fluidigm Biomark™ qRT-PCR, based on their known involvement in transcriptional regulation, cell signalling or haematopoietic processes (**Supplementary Table 5.1**). Some of these targets will be discussed in detail in the following subchapters. Importantly, the expression analysis confirmed the endothelial identity of the *kdrl*:GFP^{high}; *gata1*:DsRed^{-} fraction, by showing high mRNA levels of endothelial genes and low levels of erythroid or neuronal genes (**Fig. 5.3**).

5.4. Expression of *gata2a* is regulated by Lmo4a

5.4.1. *Gata2a* mRNA levels are elevated in *lmo4a* morphants

Previous work in our lab demonstrated upregulation of *gata2a* upon *lmo4a* knockdown by *in situ* hybridization and qRT-PCR in sorted *kdrl*:GFP^{+} cells at 24-30hpf (K. Tuladhar, DPhil thesis). The RNA-seq experiment in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells at 28hpf confirmed this observation, showing significant enrichment of *gata2a* mRNA in *lmo4a* morphants compared to wild types (**Fig. 5.4A**). Further Fluidigm Biomark™ qRT-PCR analysis in the same cell population (**Fig. 5.4B**) showed a minor increase in *gata2a* levels both at 22hpf and 28hpf upon *lmo4a* knockdown (**Fig. 5.4C**). The observed differences were not statistically significant, likely due to poor quality of the *gata2a* Taqman® assay (see Table 2.5), which produced robust PCR amplification in negative control samples (data not shown). However, the increase in *gata2a* mRNA levels was confirmed as statistically significant in a series of SYBR® Green qRT-PCR experiments in 28hpf embryos (**Supplementary Fig. 5.1**). QRT-PCR analysis in the *kdrl*:GFP^{-}; *gata1*:DsRed^{-} cell population from the same embryos showed no difference in *gata2a* levels (**Fig. 5.4D**), suggesting a tissue-specific effect of *lmo4a* knockdown. *In situ* hybridization analysis at 22hpf did not reveal any visible differences between the morphants and controls in the endothelium, mainly due to high levels of *gata2a* expression in the ICM, obscuring the

expression in the developing DA (**Fig. 5.5A-B**). However, the majority of *lmo4a* MO-injected embryos showed a clear increase in *gata2a* mRNA levels in the dorsal aorta at 26hpf (X^2=8.032, d.f.=1, p<0.01) and 28hpf (X^2=12.884, d.f.=1, p<0.001) (**Fig. 5.5C-F**).

5.4.2. *Lmo4a* mutants exhibit increased levels of *gata2a* expression

Recent reports have questioned the specificity of the MO-induced phenotypes (Kok et al. 2015; Place and Smith 2017). Therefore, we have employed TALENs to generate a loss-of-function mutant of *lmo4a* (**Supplementary Fig. 5.2**). We have isolated and established a mutant line (see **Fig. 2.2**) carrying a 5bp deletion (Δ25-29) upstream of the LIM domains, the main functional domains of Lmo proteins (Deane et al. 2004; Sang et al. 2014). Based on the sequence analysis, we predicted this deletion to be a nonsense mutation, completely changing the predicted protein sequence and resulting in a premature stop codon upstream of the last exon-exon junction, potentially resulting in nonsense-mediated decay of the *lmo4a* mRNA (Maquat 2004). Homozygous *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ fish were viable and fertile, so we could analyse the effect of *lmo4a* knockout in maternal zygotic mutants to exclude any potential maternal mRNA or protein activity. We observed increased *gata2a* expression by *in situ* hybridization in the majority of maternal zygotic homozygous *lmo4a* mutants, compared to wild type controls (X^2=7.489, d.f.=1, p<0.01) (**Fig. 5.5G-H**). Therefore, we conclude that the genetic loss-of-function mutants phenocopy the effect on *gata2a* mRNA expression observed in *lmo4a* morphants.

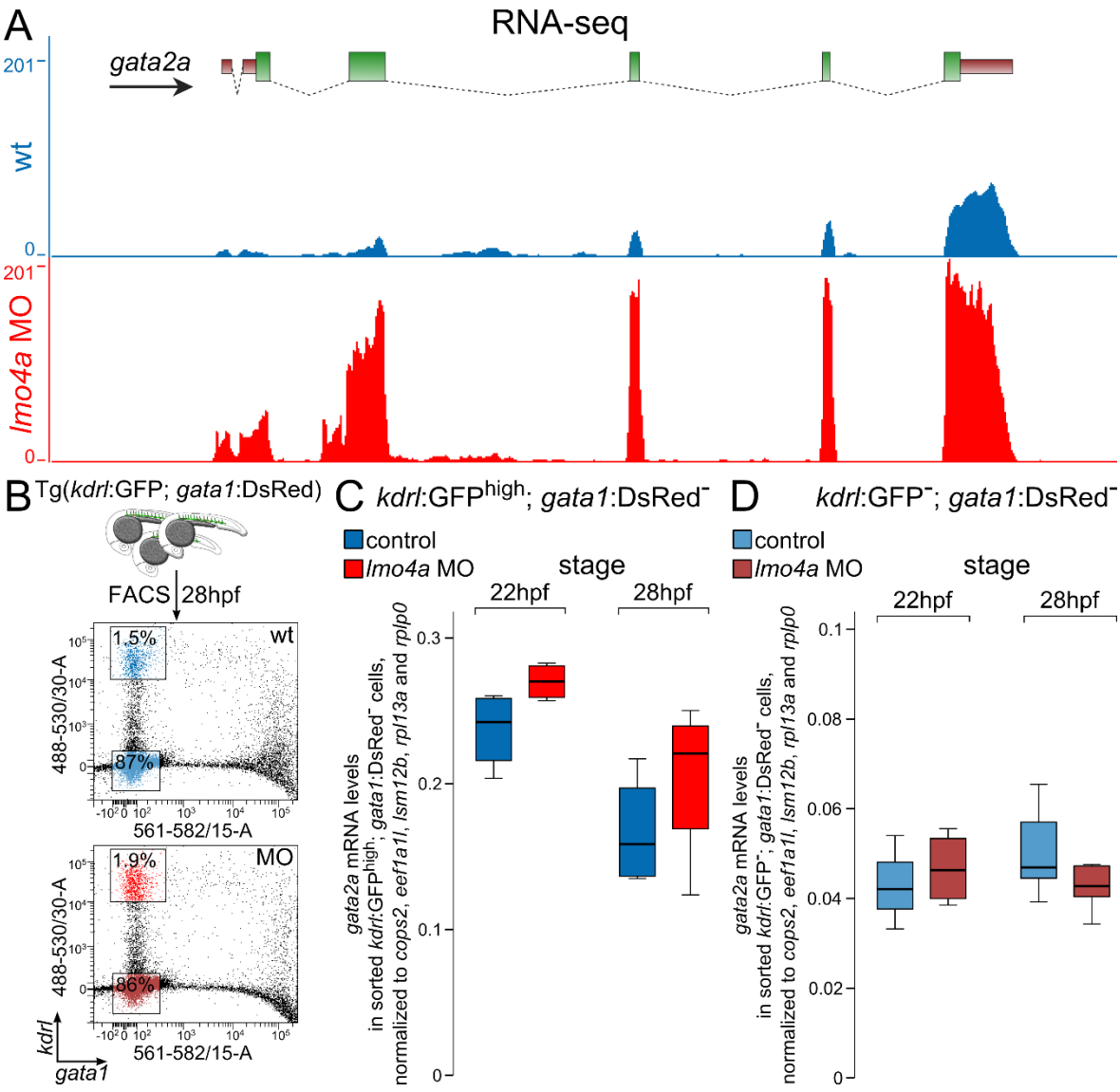


Figure 5.4. Expression of *gata2a* increases in the endothelium of *lmo4a* morphants.

(A) RNA sequencing reads mapped to *gata2a* locus in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from 28hpf control embryos (wt; blue) and *lmo4a* morphants (red) demonstrates enrichment of *gata2a* mRNA levels upon *lmo4a* knockdown. (B) *Kdrl*:GFP^{high}; *gata1*:DsRed^{-} and *kdrl*:GFP^{-}; *gata1*:DsRed^{-} cells were sorted from wild type (wt; blue) and *lmo4a* MO-injected embryos (MO; red) at 22hpf and 28hpf. The plots are representative of 28hpf embryos. (C-D) QRT-PCR on RNA isolated from the sorted cell populations (panel B) shows non-significantly increased levels of *gata2a* mRNA in the endothelium of *lmo4a* morphants at 22hpf and at 28hpf, compared to wild type. There is no difference in *gata2a* mRNA levels in double negative cells between wt and *lmo4a* morphants. N=4 for 22hpf, N=6 for 28hpf. Note different scales of expression levels between the different cell populations.

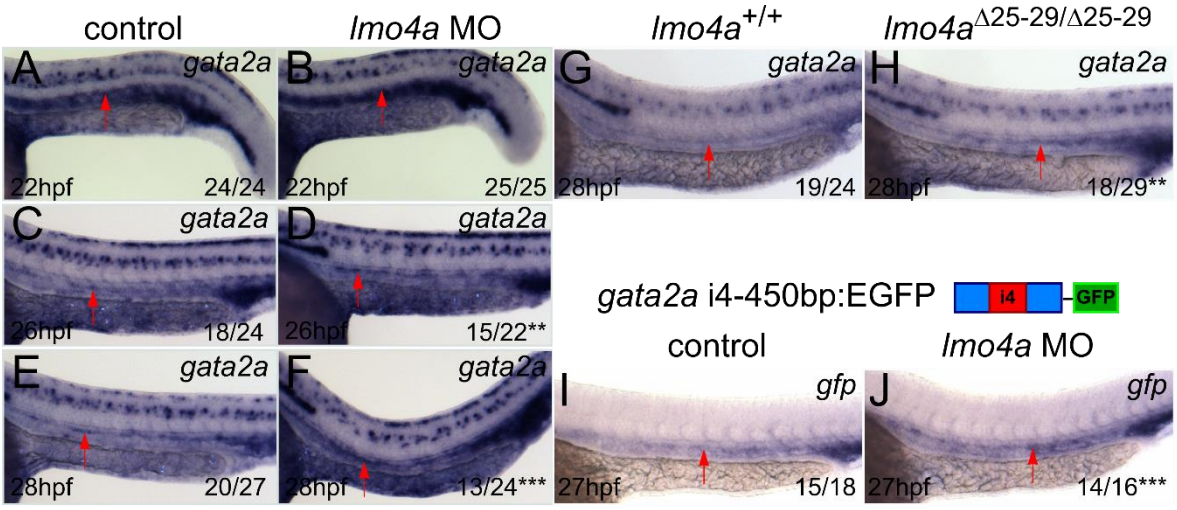


Figure 5.5. Expression of *gata2a* mRNA increases in the endothelium of *lmo4a* morphants and mutants after 26hpf, which is also associated with increased *gata2a*-i4 enhancer activity.

(A-F) Expression of *gata2a* mRNA in the dorsal aorta (arrows) appears unchanged in *lmo4a* morphants compared to uninjected siblings at 22hpf, but is increased at 26hpf and 28hpf in a significant majority of embryos, as measured with *in situ* hybridization. (G-H) Majority of *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ maternal homozygous mutants have increased levels of *gata2a* mRNA in the dorsal aorta (arrows), compared to wild type. (I-J) Expression of *gfp* driven by the *gata2a*-i4-450bp enhancer element increases in the dorsal aorta (arrows) in the majority of *lmo4a* morphants, compared to uninjected controls. $^{**}p<0.01$, $^{***}p<0.001$.

5.4.3. The activity of the *gata2a*-i4 enhancer increases upon *lmo4a* knockdown

ChIP data from our collaborators showed Lmo4 binding to the *Gata2* +9.5 enhancer in mouse erythroleukaemia (MEL) cells (C. Andrieu-Soler, unpublished data). In addition, the mouse +9.5 enhancer is bound by Ldb1 (Sanalkumar et al. 2014), a protein partner of Lmo4 in differentiated MEL cells (Meier et al. 2006). Therefore, we hypothesized that *lmo4a* could regulate *gata2a* through the zebrafish equivalent of +9.5, the *gata2a*-i4 enhancer. Injection of *lmo4a* MO resulted in increased *gfp* expression in Tg(*gata2a*-i4-450bp:GFP) fish at 27hpf, compared to wild type (X^2=14.285, d.f.=1, $p<0.001$) (**Fig. 5.5I-J**). The pixel intensity levels were not digitally quantified due to distorted morphology of most *lmo4a* morphants. This observation suggests that the knockdown of *lmo4a* might affect the expression of *gata2a* through upregulated activity of the *gata2a*-i4 enhancer. Analysis of *gata2a*$^{\Delta i4/\Delta i4}$ mutants injected with *lmo4a* MO will

further verify this hypothesis. However, the observed effect could result from increased proliferation and higher number of *gata2a*-i4:GFP^{+} cells upon *lmo4a* knockdown rather than changes on the transcriptional level. Flow cytometry analyses in the *lmo4a* morphants and control siblings, as done in *gpr65* morphants (see Chapter 3.5), should resolve this uncertainty.

5.5. Knockdown of *lmo4a* affects the expression of several genes involved in HE and HSC programming

5.5.1. Expression of *gata2b* mRNA is not majorly affected in *lmo4a* morphants

In the previous chapter, we provided evidence for a model in which *gata2a* acted upstream of *gata2b*, a crucial marker of the haemogenic endothelium in the zebrafish (Butko et al. 2015). Our RNA-seq analysis revealed an increased expression of *gata2b* in the *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells in *lmo4a* morphants at 28hpf (**Fig. 5.2C, 5.3**), so we wanted to verify this observation with other methods. We observed strong foci of *gata2b* mRNA expression in the dorsal aorta of *lmo4a* morphants at 22-23hpf, which were absent in control embryos (**Fig. 5.6A-D**). This effect was not visible at 24hpf, when no difference in *gata2b* expression was apparent between the morphants and wild types (**Fig. 5.6E-F**). Quantitation of *in situ* hybridization (see **Fig. 2.3**) against *gata2b* at 30hpf showed a slight but non-significant increase in pixel intensity in the *lmo4a* morphants compared to uninjected controls (μ_{wt}=36, μ_{MO}=41.8) (**Fig. 5.6G**). We further analysed the expression of *gata2b* with Fluidigm Biomark™ qRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells at 22hpf and 28hpf (**Fig. 5.6H**). Because *gata2b* expression in *kdrl*:GFP^{-}; *gata1*:DsRed^{-} population was very low (data not shown), we used the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} fraction, possibly representing the EMPs (see Chapter 5.2), as an internal control (**Fig. 5.6H**). Contrary to the *in situ* hybridization and RNA-seq data, we observed a slight and

non-significant decrease in *gata2b* mRNA levels upon *lmo4a* knockdown in the endothelial fraction at both 22hpf and 28hpf (**Fig. 5.6I**). The most likely explanation for these data is that *gata2b* expression is not affected by *lmo4a* knockdown and the observed slight variations result from natural fluctuations of mRNA levels in the developing embryos. There was no difference in *gata2b* expression in the *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells (**Fig. 5.6J**).

5.5.2. Selected early HE markers are upregulated upon *lmo4a* knockdown

Differential Expression analysis of the RNA-seq data (**Fig. 5.2C**) revealed significant upregulation of *foxc1a*, *etv2*, *tal1* and *sox7* (**Fig. 5.3**), all of which are involved in HE and blood programming (Shivdasani et al. 1995; Patterson et al. 2007; Ren et al. 2010; Costa et al. 2012b; Veldman and Lin 2012; Ciau-Uitz et al. 2013a; Liu et al. 2015). We further tested their expression with qRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells (see **Fig. 5.2A**) at 22hpf and 28hpf. We observed increased levels of all four genes at both stages in *lmo4a* morphants compared to wild types (**Fig. 5.7A-D**). Upregulation of *foxc1a* and *tal1* was significant at 28hpf ($p<0.05$ and $p<0.01$, respectively), while *etv2* and *sox7* were significantly upregulated at both stages ($p<0.01$ and $p<0.05$, respectively) (**Fig. 5.7A-D**). These data suggest that the knockdown of *lmo4a* may already affect genes higher in the genetic hierarchy than *gata2a*. *Etv2* was previously shown to be upstream of *Gata2* in the mouse and *Xenopus* (Ciau-Uitz et al. 2013a; Liu et al. 2015) and the data presented in Chapter 4 suggest conservation of this relationship in the zebrafish. *Foxc1a* acts upstream of *etv2* in zebrafish mesodermal progenitors (Veldman and Lin 2012). Therefore, we suggest a model in which *lmo4a*-induced upregulation of *foxc1a* in the endothelium causes secondary upregulation of *etv2* and its downstream targets, *gata2a* and *tal1* (Ren et al. 2010; Ciau-Uitz et al. 2013a; Liu et al. 2015), placing *lmo4a* as a global regulator of the expression of HE-specifying genes. *Sox7* was recently shown to be expressed in the HE prior to the emergence of blood progenitors *in vitro* and its downregulation was necessary to allow haematopoietic differentiation (Gandillet et al. 2009; Costa et al. 2012b). Its upregulation in *lmo4a* morphants

could be the reason for decreased expression of *cmyb*, the marker of HSCs (**Fig. 5.2C, 5.3**). Surprisingly, digital quantification of the *sox7 in situ* hybridization signal (**Fig. 2.3**) in *lmo4a* MO-injected embryos and uninjected controls showed significant downregulation in *sox7* pixel intensity in the morphants (μ_{wt}=74.5, μ_{MO}=62.6; t=5.519, d.f.=47, p<0.001) (**Fig. 5.7E**). At 28hpf, there was no significant difference between the *lmo4a* morphants and the controls (μ_{wt}=53.7, μ_{MO}=55.9) (**Fig. 5.7F**). This discrepancy could result from a combination of several reasons, the first being distorted morphology of *lmo4a* morphants. Another factor could be stronger expression of *sox7* in the roof of the dorsal aorta than in its floor, potentially affecting the precision of the quantification, which was originally designed for the signal in the floor of the dorsal aorta. Defining the most appropriate region of interest for the signal quantification is further confounded by the expression of *sox7* in PBI vessels. Moreover, the background signal calculation could be affected by the fact that *sox7* is expressed in the intersomitic vessels, suggesting that the *in situ* quantification method may not work well for this gene.

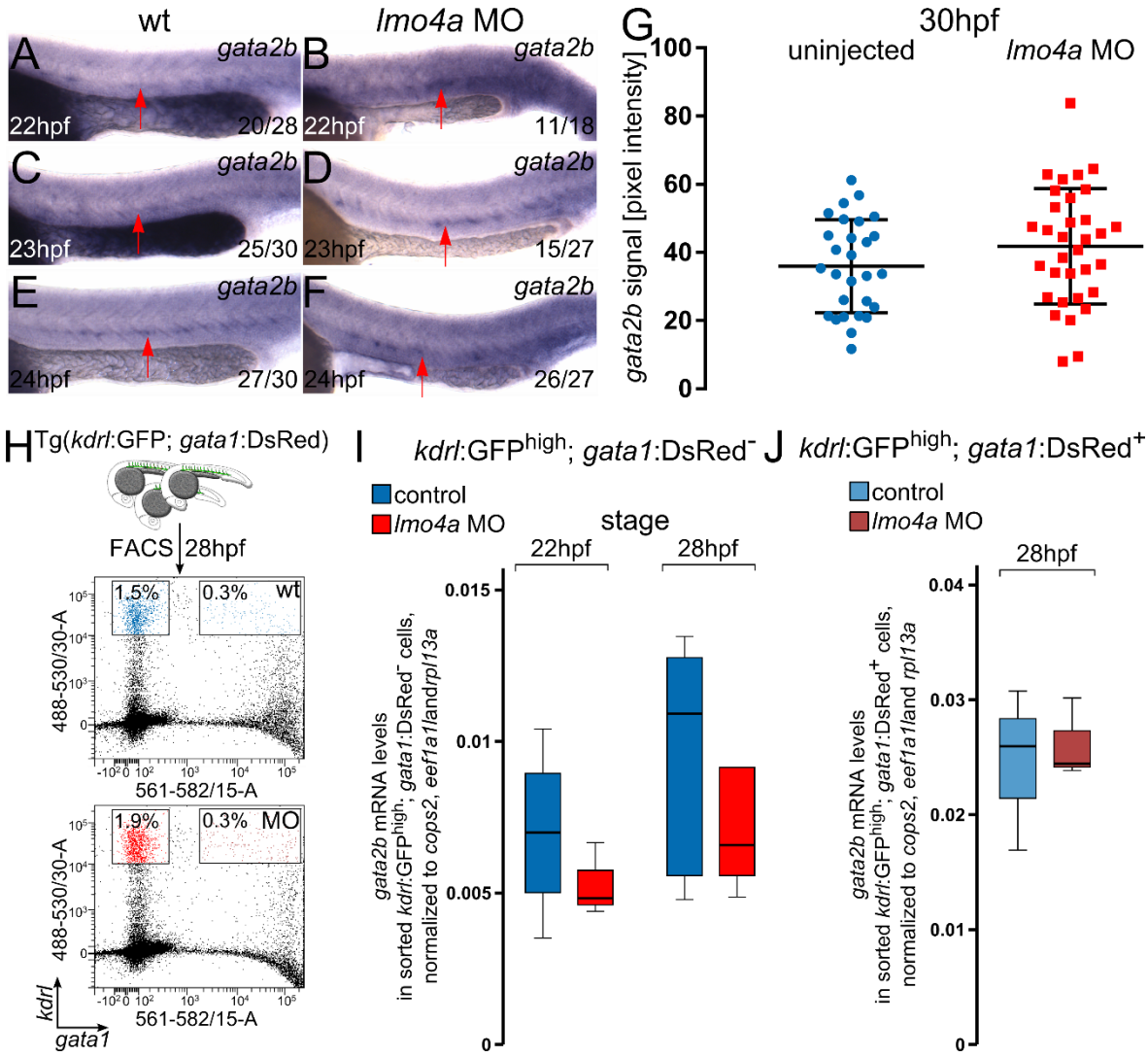


Figure 5.6. Expression of *gata2b* mRNA is not affected by *lmo4a* knockdown.

(A-F) Around half of *lmo4a* morphants have stronger foci of *gata2b* mRNA expression in the dorsal aorta (arrows) at 22-23hpf compared to control siblings, but not at 24hpf, measured with *in situ* hybridization. (G) Quantitation of the *in situ* hybridization *gata2b* signal in 30hpf uninjected embryos (blue) and *lmo4a* morphants (red) shows slight, but not statistically significant increase in *gata2b* mRNA expression. Error bars: mean±SD. (H) *Kdrl*:GFP^{high}; *gata1*:DsRed^{-} and *kdrl*:GFP^{high}; *gata1*:DsRed^{+} cells were sorted from wild type (wt; blue) and *lmo4a* MO-injected embryos (MO; red) at 22hpf and 28hpf. The plots are representative of 28hpf embryos. (I-J) QRT-PCR on RNA isolated from the sorted cell populations (panel H) shows slight yet insignificant decrease in *gata2b* mRNA levels in the endothelium of *lmo4a* morphants at 22hpf and at 28hpf, compared to wild type. There is no difference in *gata2b* mRNA levels in double positive cells between control siblings and *lmo4a* morphants. N=4 for 22hpf, N=6 for 28hpf in panel I, N=3 for panel J. Note different scales of expression levels.

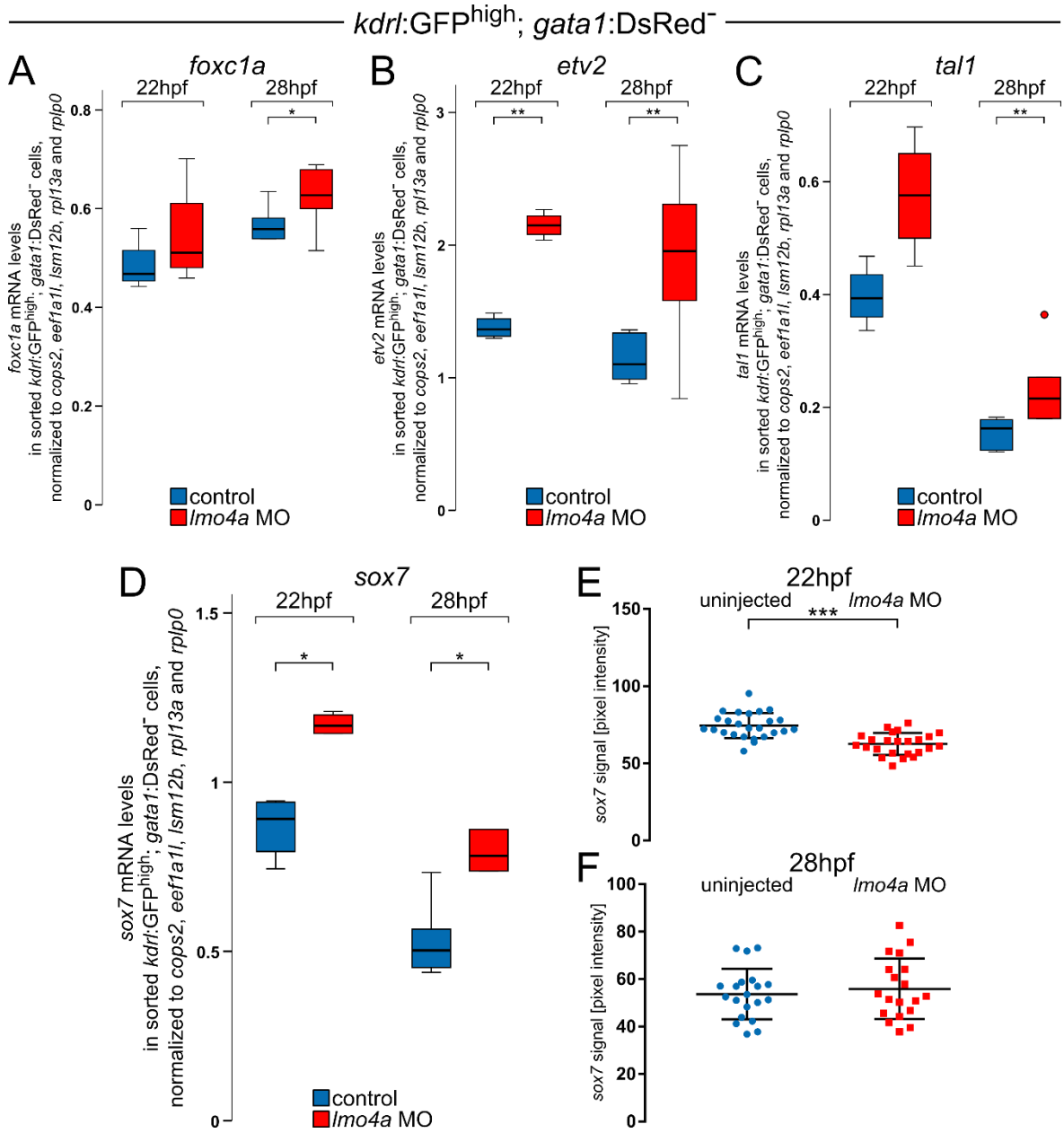


Figure 5.7. Expression of selected genes required for HE specification increases upon *lmo4a* knockdown.

QRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from wild type (blue) and *lmo4a* morphants (red) shows significant increase in the levels of *foxc1a* (A), *etv2* (B), *tal1* (C) and *sox7* (D) mRNA in the endothelium of *lmo4a* morphants at 28hpf, compared to wild type. This increase is also significant at 22hpf in the case of *etv2* (B) and *sox7* (D). N=4 for 22hpf, N=6 for 28hpf. *p<0.05, **p<0.01. (E-F) Quantification of the *sox7* mRNA signal, detected by *in situ* hybridization, from uninjected siblings (blue) and *lmo4a* morphants (red) at 22hpf shows a significant reduction in *sox7* pixel intensity in the DA of *lmo4a* morphants. At 28hpf there is no significant difference. Error bars: mean±SD. ***p<0.001.

5.5.3. *Lmo4a* morphants express reduced levels of *gfi1aa* and *cmyb*

Previous work in our lab suggested that *lmo4a*-induced upregulation of *gata2a* resulted in the loss of HSC specification, which could be rescued by inhibition of Gata activity (K. Tuladhar, DPhil thesis). *In vitro* studies suggested that overexpression of *Gata2* indeed negatively affected HSC generation and functionality (Persons et al. 1999; Tipping et al. 2009). A couple of reports from *in vivo* models also described molecular mechanisms maintaining balanced *Gata2* expression to allow the generation of functional HSPCs (Guiu et al. 2013; Gao et al. 2016). Our RNA-seq experiment in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} at 28hpf showed significant downregulation in mRNA levels of *dll4*, *gfi1aa* and *cmyb* (**Fig. 5.2C, 5.3**), all associated with HE identity and haematopoietic development (Thompson et al. 1998; Emambokus et al. 2003; Duarte et al. 2004; Hayashi and Kume 2008; Butko et al. 2016; Thambyrajah et al. 2016). We further investigated these effects with whole-mount *in situ* hybridization and Fluidigm Biomark™ qRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells (**Fig. 5.8**). Expression of *dll4* in the dorsal aorta appeared slightly decreased in the *lmo4a* morphants by *in situ* hybridization at 30hpf (**Fig. 5.8A-B**), while qRT-PCR revealed a slight, non-significant decrease upon *lmo4a* knockdown at 22hpf and no difference between the two samples at 28hpf (**Fig. 5.8C**). On the other hand, the majority of *lmo4a* MO-injected embryos showed a clear reduction in *gfi1aa* mRNA expression in the dorsal aorta at 30hpf (X^2=11.733, d.f.=1, p<0.01) (**Fig. 5.8D-E**). This was further confirmed by qRT-PCR analysis, showing significant reduction in *gfi1aa* mRNA levels in the endothelium of *lmo4a* morphants at 28hpf compared to wild types (p<0.05) (**Fig. 5.8F**). We made very similar observations for *cmyb*, with reduced *in situ* signal in *lmo4a* morphants at 30hpf (X^2=9.42, d.f.=1, p<0.01) (**Fig. 5.8G-H**) and significantly reduced mRNA levels, measured with qRT-PCR at 28hpf (p<0.001) (**Fig. 5.8I**). These data suggest that knockdown of *lmo4a* results in deficient HE and HSC specification.

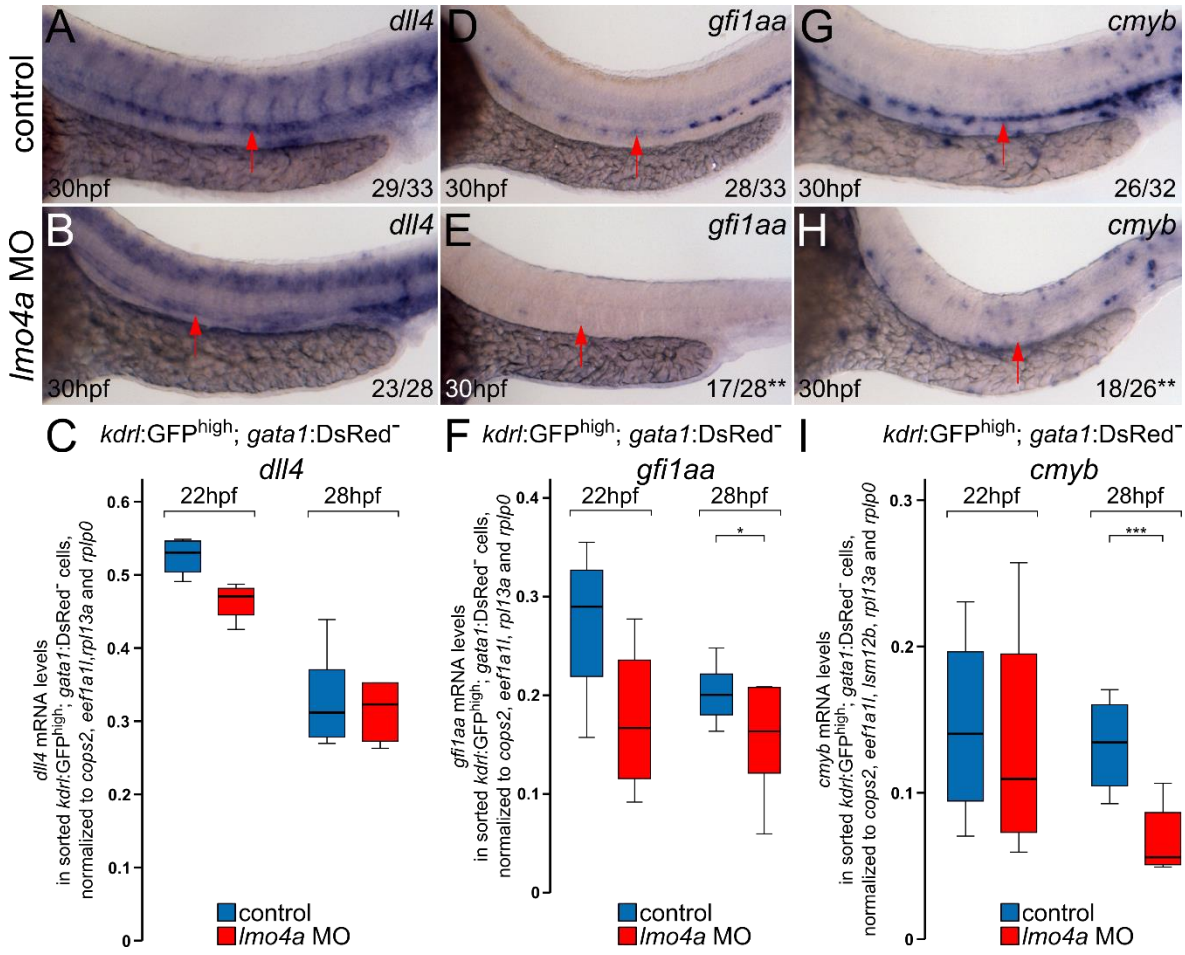


Figure 5.8. Expression of *gfi1aa* and *cmyb* is reduced in *lmo4a* morphants.

(A-B) *Lmo4a* morphants do not express *dll4* in the intersomitic vessels, but the *dll4* expression in the dorsal aorta (arrows) is not affected at 30hpf, compared to wild type, as measured by *in situ* hybridization. (C) QRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from uninjected siblings (blue) and *lmo4a* morphants (red) shows slight reduction in *dll4* mRNA levels upon *lmo4a* knockdown at 22hpf, but not at 28hpf. (D-E) *In situ* hybridization shows reduced expression of *gfi1aa* in the dorsal aorta (arrows) upon *lmo4a* knockdown at 30hpf, compared to controls. (F) QRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from control embryos (blue) and *lmo4a* morphants (red) shows reduced *gfi1aa* mRNA levels in *lmo4a* morphants at 22hpf and 28hpf. (G-H) At 30hpf, mRNA expression of *cmyb*, probed with *in situ* hybridization, decreases upon *lmo4a* knockown, compared to wild type. (I) QRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from wild type (blue) and *lmo4a* knockdown (red) embryos shows a reduction in *cmyb* mRNA levels in *lmo4a* morphants at 28hpf, but not 22hpf. N=4 for 22hpf, N=6 for 28hpf. *p<0.05, **p<0.01, ***p<0.001.

5.6. Specification of the main embryonic vessels is not affected in the *lmo4a* morphants

Previous work in our lab showed that while *lmo4a* morphants showed defects in the angiogenesis of intersomitic vessels (ISVs), vasculogenesis of the dorsal aorta was unaffected, compared to uninjected controls (K. Tuladhar, DPhil thesis). In agreement with that, we confirmed the ISV defect by *in situ* hybridization against *fli1* and *kdrl*, markers of embryonic endothelium (Lawson and Weinstein 2002; Jin et al. 2005). *Lmo4a* morphants analysed at 30hpf exhibited complete loss of ISVs, but marker expression in the dorsal aorta was unaffected compared to control embryos (**Fig. 5.9A-D**). Consistent with that, injection of *lmo4a* MO into Tg(*kdrl*:GFP) embryos resulted in the loss of ISVs, but not the dorsal aorta (**Fig. 5.9E-F**). Furthermore, we observed no differences in the mRNA expression of *efnb2a*, an artery-specific marker, and its venous equivalent, *ephb4a* (Wang et al. 1998; Gerety et al. 1999) between *lmo4a* morphants and wild types in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells at 22hpf or 28hpf, suggesting that arteriovenous specification was not affected (**Fig. 5.9G-H**). We also used qRT-PCR to measure the expression of *erg*, an endothelial marker required for the specification of HSCs (Baltzinger et al. 1999; Loughran et al. 2008; Wilson et al. 2010). We observed no difference in expression between *lmo4a* morphants and wild types (**Fig. 5.9I**), further supporting the observation that vasculogenesis, occurring before the onset of HE specification and ISV angiogenesis, was not affected by *lmo4a* knockdown.

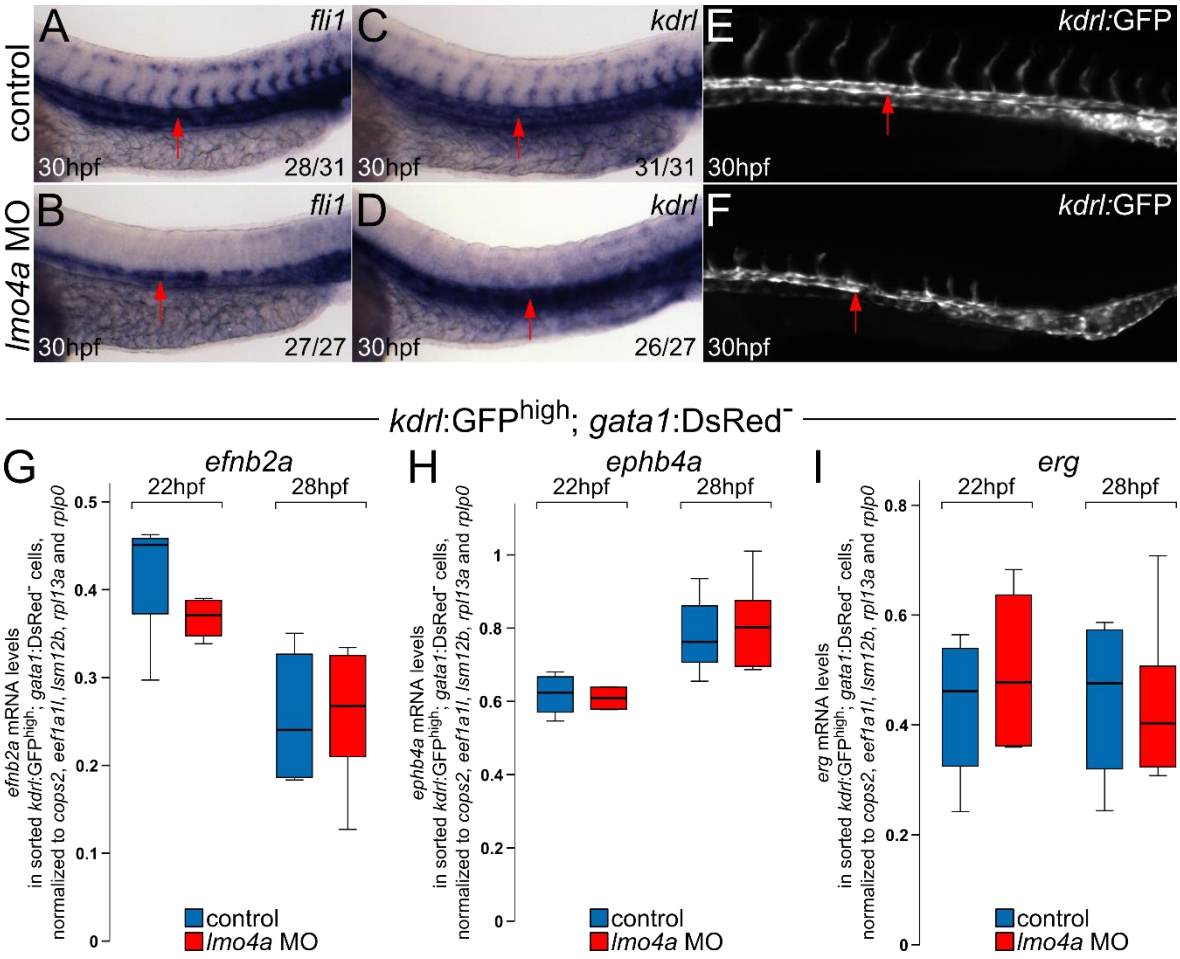


Figure 5.9. Specification of the dorsal aorta endothelium is not affected in *lmo4a* morphants.

(A-D) *In situ* hybridization for the endothelial markers *fli1* and *kdrl* at 30hpf reveals the loss of intersomitic vessels in *lmo4a* morphants. The dorsal aorta (arrows) appears unaffected. (E-F) Live images of 30hpf Tg(*kdrl*:GFP) control embryos and *lmo4a* morphants demonstrate normal dorsal aorta development (arrows) of the morphants. (G-I) QRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells shows no significant changes in the levels of *efnb2a*, *ephb4a* or *erg* mRNA in the endothelium of *lmo4a* morphants (red) at 22hpf or 28hpf, compared to uninjected siblings (blue). N=4 for 22hpf, N=6 for 28hpf.

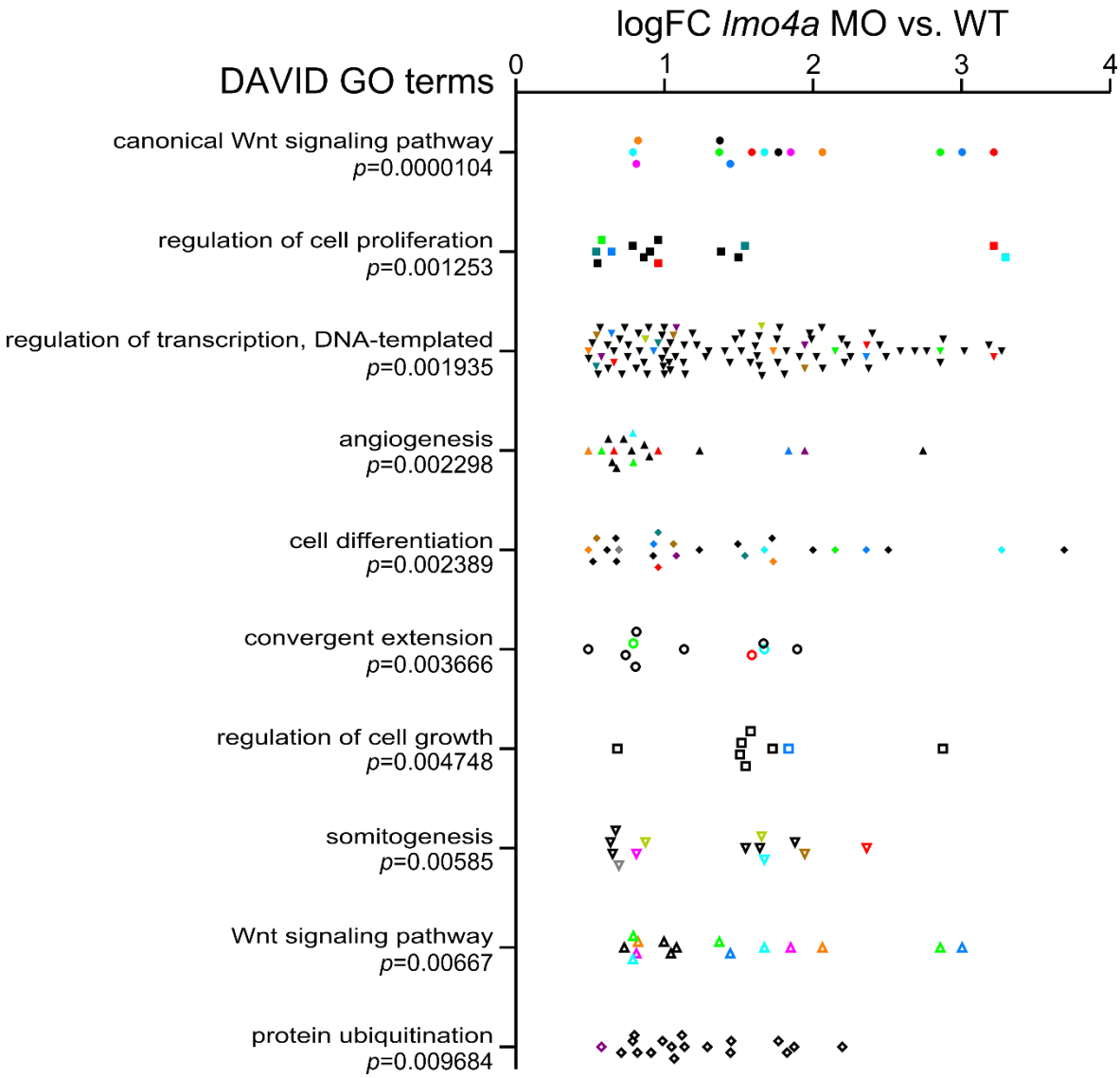


Figure 5.10. Gene Ontology term analysis suggests a role of *lmo4a* in regulating Wnt signalling.

Gene Ontology (GO) term analysis of 905 genes differentially upregulated in *lmo4a* morphants over uninjected controls in RNA-seq experiment demonstrates significant enrichment for genes involved in Wnt signalling pathway. Transcription factors are also among the most enriched GO terms. Other terms involve important developmental processes, including cell proliferation and differentiation or angiogenesis. Single points represent single genes within each GO term. Same colours mark the same genes present in multiple GO terms. FC: fold change. Top 10 GO terms (p<0.01) are presented.

5.7. Wnt signalling likely acts downstream of *lmo4a* and upstream of *gata2a*

5.7.1. Expression of Wnt signalling genes is increased in *lmo4a* morphants

Gene Ontology (GO) term analysis of lists of differentially expressed genes from high-throughput sequencing experiments can reveal potential regulatory patterns, otherwise difficult to notice with candidate gene approaches (Huang et al. 2008; Huang et al. 2009). In our GO term analysis of the RNA-seq data (see Chapter 5.3), we focused on the 907 genes upregulated upon *lmo4a* knockdown (Appendix 1), because previous structural and functional analyses strongly supported the involvement of Lmo4a in a negative *trans* regulatory complex, including known negative transcriptional regulators, such as Hdac7 or Bcor (K. Tuladhar, DPhil thesis). We were not surprised to find the enrichment for genes classified into terms such as cell proliferation, cell growth or differentiation (**Fig. 5.10**), which in the context of the developing embryo represent developmental processes, confirming an important role for *lmo4a* in tissue patterning. We also observed enrichment for DNA-binding transcription factors (**Fig. 5.10**), potentially placing *lmo4a* at a high level in gene regulatory networks orchestrating developmental processes in the endothelium. Intriguingly, we found GO terms associated with the Wnt signalling pathway to be among the most enriched (**Fig. 5.10**). Many of these genes were in fact classified into more than one of the enriched GO terms (indicated by the same colours in **Fig. 5.10**). We further analysed the expression of the identified genes upregulated in RNA-seq involved in Wnt signalling with Fluidigm Biomark™ qRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells at 22hpf or 28hpf (see **Supplementary Table 5.1**). We found very little difference in the mRNA levels at 22hpf, but at 28hpf we observed increased expression in *lmo4a* morphants of a few genes, including *ctnnb1* (n.s.), *jun* ($p<0.01$), *fzd7a* (n.s.), *lef1* ($p<0.01$), *csnk1e* ($p<0.001$) or *tcf7* ($p<0.05$) (**Fig. 5.11**). The Wnt signalling pathway is known to be involved in

the specification of HE and HSPCs (Richter et al. 2017), so we decided to further investigate a potential *lmo4a*-Wnt link.

5.7.2. Inhibition of Wnt signalling reduces *gata2a* expression in *lmo4a* morphants

Expression of *Gata2* in the developing haematopoietic system is regulated by inputs from Notch (Robert-Moreno et al. 2005) and BMP (Kirmizitas et al. 2017) signalling pathways. In the zebrafish, Notch signals regulate *gata2b*, but not *gata2a* (Butko et al. 2015). However, potential regulation of *Gata2* by Wnt has not been investigated to date.

We hypothesized that Wnt signalling could mediate the *lmo4a*-induced upregulation of *gata2a*. To test this hypothesis, we made use of available chemical modulators of Wnt activity. IWR-1 is a selective small molecule inhibitor of Wnt signalling (Chen et al. 2009a), previously used to treat zebrafish embryos after 9hpf to study swimbladder development (Yin et al. 2011) or endothelial specification (Hübner et al. 2017). BIO is a selective inhibitor of GSK-3β and its activity results in enhanced Wnt signalling (Meijer et al. 2003). Treatment of zebrafish embryos from 16hpf with 500nM BIO resulted in ISV angiogenesis defects, but the DA specification was unaffected (Gore et al. 2011). We first investigated whether expression of *gata2a* mRNA, assayed with *in situ* hybdridization at 30hpf, was affected by the chemical modulation of Wnt signalling, introduced from 16hpf. While treatment with BIO had no apparent effect compared to DMSO-treated controls, exposure to IWR-1 resulted in decreased *gata2a* expression in around half of the treated embryos (**Fig. 5.12A-C**). Furthermore, treating *lmo4a* morphants with IWR-1 rescued the overexpression of *gata2a* down to wild type levels (**Fig. 5.12D-E**), suggesting that Wnt signalling acts downstream of *lmo4a* in regulating the expression of *gata2a*.

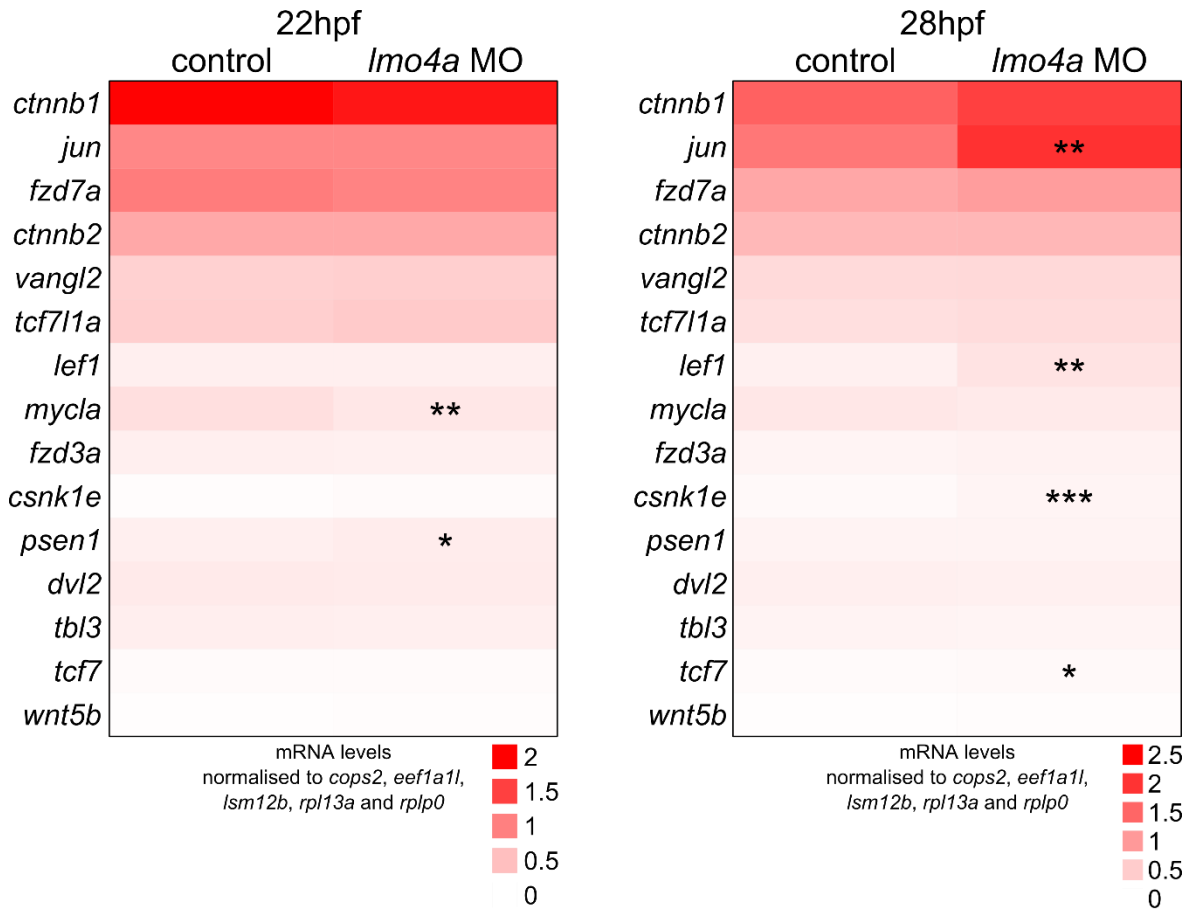


Figure 5.11. Expression of few Wnt signalling genes is increased in *lmo4a* morphants.

QRT-PCR measuring mRNA expression of genes involved in Wnt signalling in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from control embryos and *lmo4a* morphants shows little difference between the endothelium of *lmo4a* morphants and uninjected siblings at 22hpf. The only significant differences are downregulation of *mycla* and upregulation of *psen1*. At 28hpf, many Wnt signalling genes are expressed at higher levels upon *lmo4a* knockdown, including significantly upregulated *jun*, *lef1*, *csnk1e* and *tcf7*. Other differences are not statistically significant. N=4 for 22hpf, N=6 for 28hpf. *p<0.05, **p<0.01, ***p<0.001.

5.7.3. Chemical modulation of Wnt signalling affects the activity of the *gata2a*-i4 enhancer

The highly conserved *gata2a*-i4 enhancer contains binding motifs for Gata, Ets and E-box transcription factors (see **Fig. 3.2**). However, we investigated whether other regions of the enhancer could be targeted by the effectors of Wnt signalling. Indeed, we found a putative Tcf binding site in the 5' region of the enhancer (R. Monteiro, unpublished data). To test whether the activity of the *gata2a*-i4 enhancer could be regulated by Wnt, we quantified the *gfp in situ*

hybridization signal (see **Fig. 2.3**) in Tg(*gata2a*-i4-450bp:GFP) and Tg(*gata2a*-i4-1.1kb:GFP) embryos treated with IWR-1 or BIO. We found significantly decreased pixel intensity in the Tg(*gata2a*-i4-450bp:GFP) embryos treated with IWR-1, while the embryos treated with BIO showed an increase in *gfp* mRNA signal (μ_{DMSO}=31.3, $\mu_{IWR\text{-}1}$=21.7, μ_{BIO}=41.1; F=30.905, d.f.=2, 75; p<0.001) (**Fig. 5.12F**). Treatment of Tg(*gata2a*-i4-1.1kb:GFP) showed a similar trend (μ_{DMSO}=43.9, $\mu_{IWR\text{-}1}$=40.8, μ_{BIO}=48.7), but the differences were not significant at the p=0.05 level (**Fig. 5.12G**). Together these data suggest that Wnt signalling could mediate the effect of *lmo4a* knockdown on *gata2a* expression via the *gata2a*-i4 enhancer. However, the observed effects could result from the fact that Wnt is a know modulator of cell proliferation (Loh et al. 2016). The differences may arise from larger number of cells expressing *gata2a* or *gfp*. Flow cytometry of *gata2a*-i4:GFP^{+} cells in *lmo4a* morphants and control siblings would address this possibility.

5.8. Loss of *lmo4a* function – differences between morphants and mutants

5.8.1. Expression of *runx1* is reduced in *lmo4a* morphants

As mentioned in Chapter 5.5.3, upregulation of *gata2a* upon *lmo4a* knockdown resulted in the loss of HSC programming (K. Tuladhar, DPhil thesis). Although we already observed decreased expression of *cmyb* in *lmo4a* morphants, we also measured the expression of *runx1*, one of the earliest *bona fide* HE markers (Burns et al. 2002; Kalev-Zylinska et al. 2002; Gering and Patient 2005). Consistently with previous reports (Meier et al. 2006), we observed decreased levels of *runx1* mRNA in the majority of *lmo4a* morphants compared to uninjected controls (X^2=9.547, d.f.=1, p<0.01) (**Fig. 5.13A-B**). We also observed a slight, yet non-significant decrease in *runx1* levels in *lmo4a* morphants compared to wild types, measured with qRT-PCR in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells at 28hpf, but not at 22hpf (**Fig. 5.13C**). These data suggest that *lmo4a* is required for *runx1* expression in the HE.

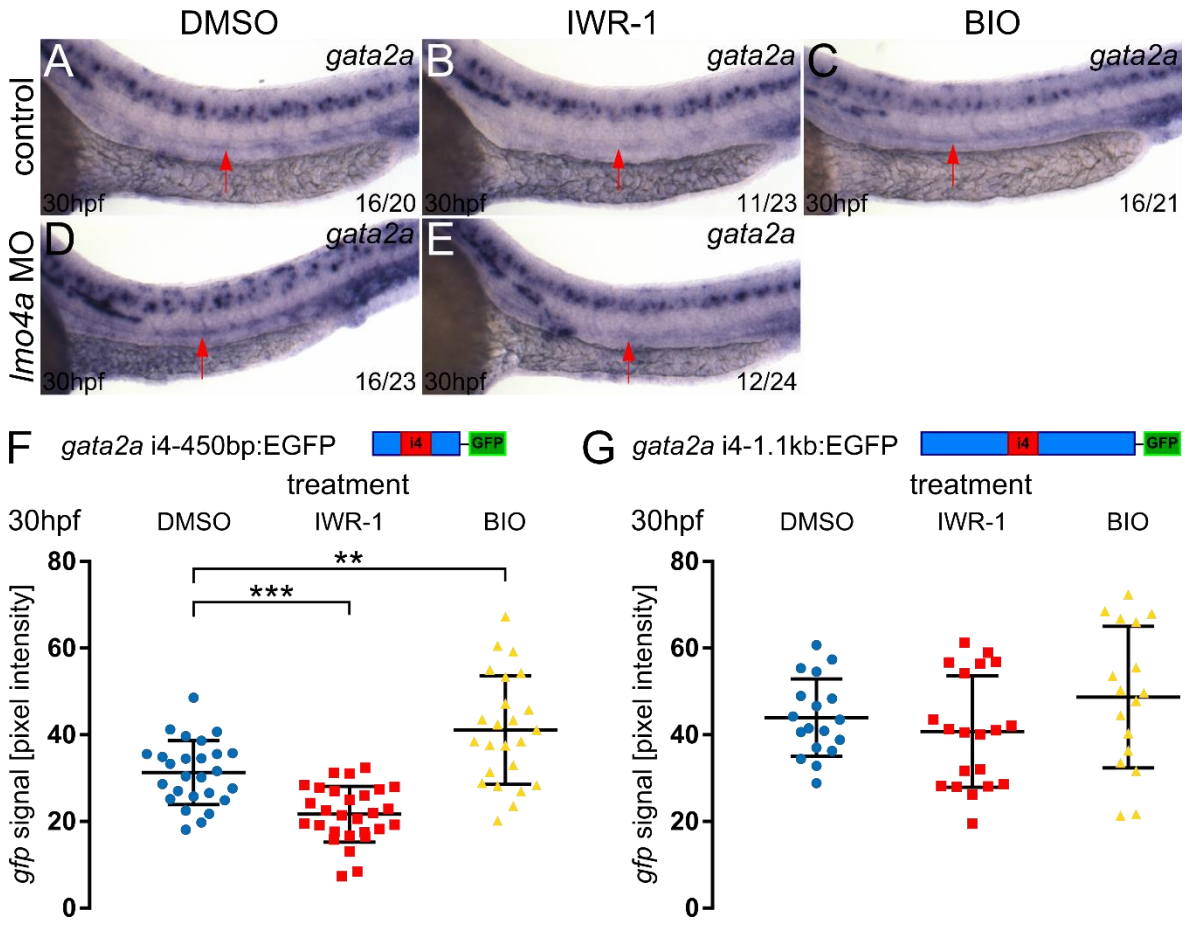


Figure 5.12. Chemical inhibition of Wnt signalling rescues *gata2a* overexpression in *lmo4a* morphants, potentially involving the i4 enhancer.

(A-C) *In situ* hybridization shows decreased *gata2a* mRNA expression at 30hpf in around 50% of embryos in the dorsal aorta (arrows) upon Wnt inhibition with IWR-1, compared to DMSO-treated control. Treatment with Wnt agonist BIO does not affect *gata2a* expression at 30hpf. (D-E) Inhibition of Wnt signalling with IWR-1 rescues increased *gata2a* mRNA levels in *lmo4a* morphants down to control levels. (F-G) Quantification of the *gfp* mRNA signal, detected by *in situ* hybridization, control Tg(*gata2a*-i4-450bp:GFP) (DMSO; blue) and their transgenic siblings treated with IWR-1 (red) or BIO (yellow) at 30hpf shows that *gfp* pixel intensity in the DA is positively affected by changes in Wnt signalling. Same analysis of Tg(*gata2a*-i4-1.1kb:GFP) embryos demonstrates a similar trend, but the differences are not significant. Error bars: mean±SD. **p<0.01; ***p<0.001.

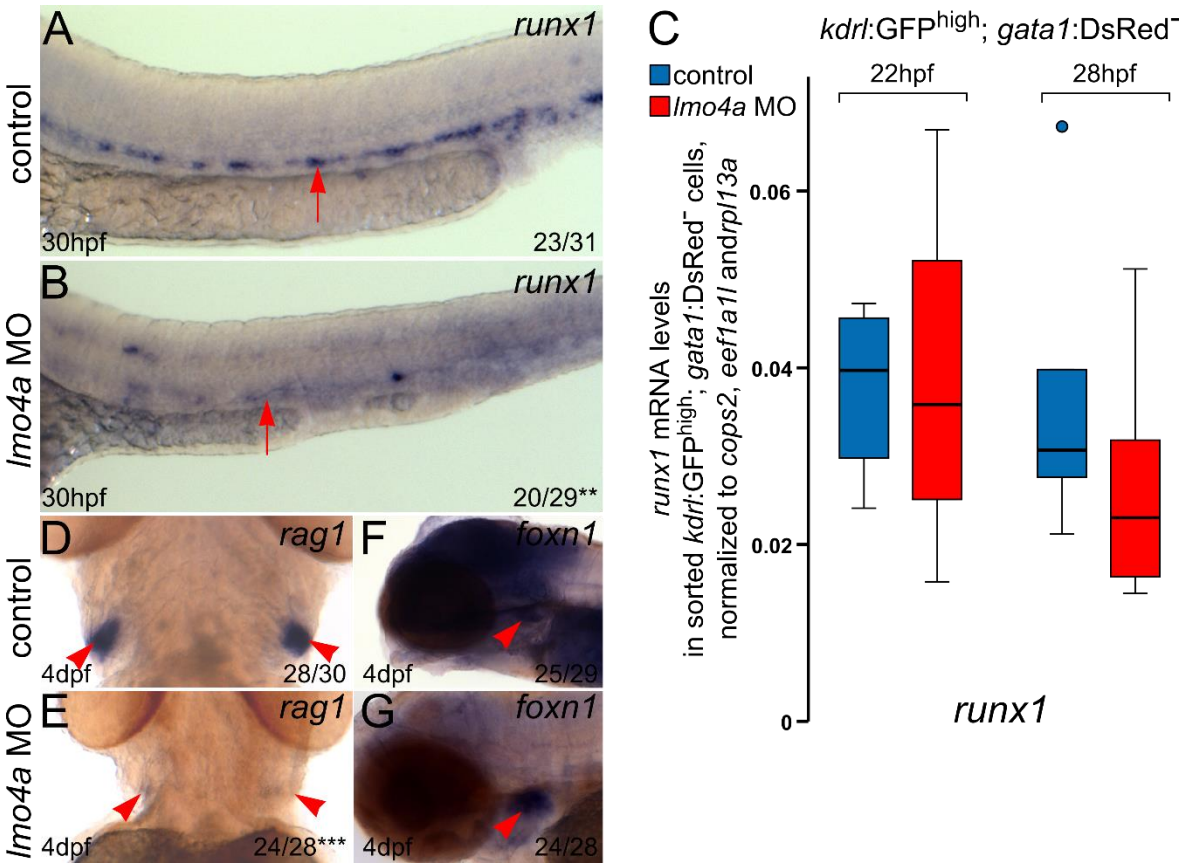


Figure 5.13. Expression of *runx1* and *rag1* is reduced upon *lmo4a* knockdown.

(A-B) Majority of *lmo4a* morphants show reduced expression of *runx1* in the dorsal aorta (arrows) at 30hpf, as detected with *in situ* hybridization. (C) QRT-PCR in sorted *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells from uninjected (blue) and *lmo4a* knockdown (red) embryos shows a reduction in *runx1* mRNA levels in *lmo4a* morphants at 28hpf, but not 22hpf. N=4 for 22hpf, N=6 for 28hpf. (D-E) *Rag1* expression in the thymi (arrowheads) is reduced in the majority of *lmo4a* morphants at 4dpf, compared to uninjected controls, as measured with *in situ* hybridization. (F-G) Expression of *foxn1* in the thymi (arrowheads) is not affected by *lmo4a* knockdown. $^{**}p<0.01$, $^{***}p<0.001$.

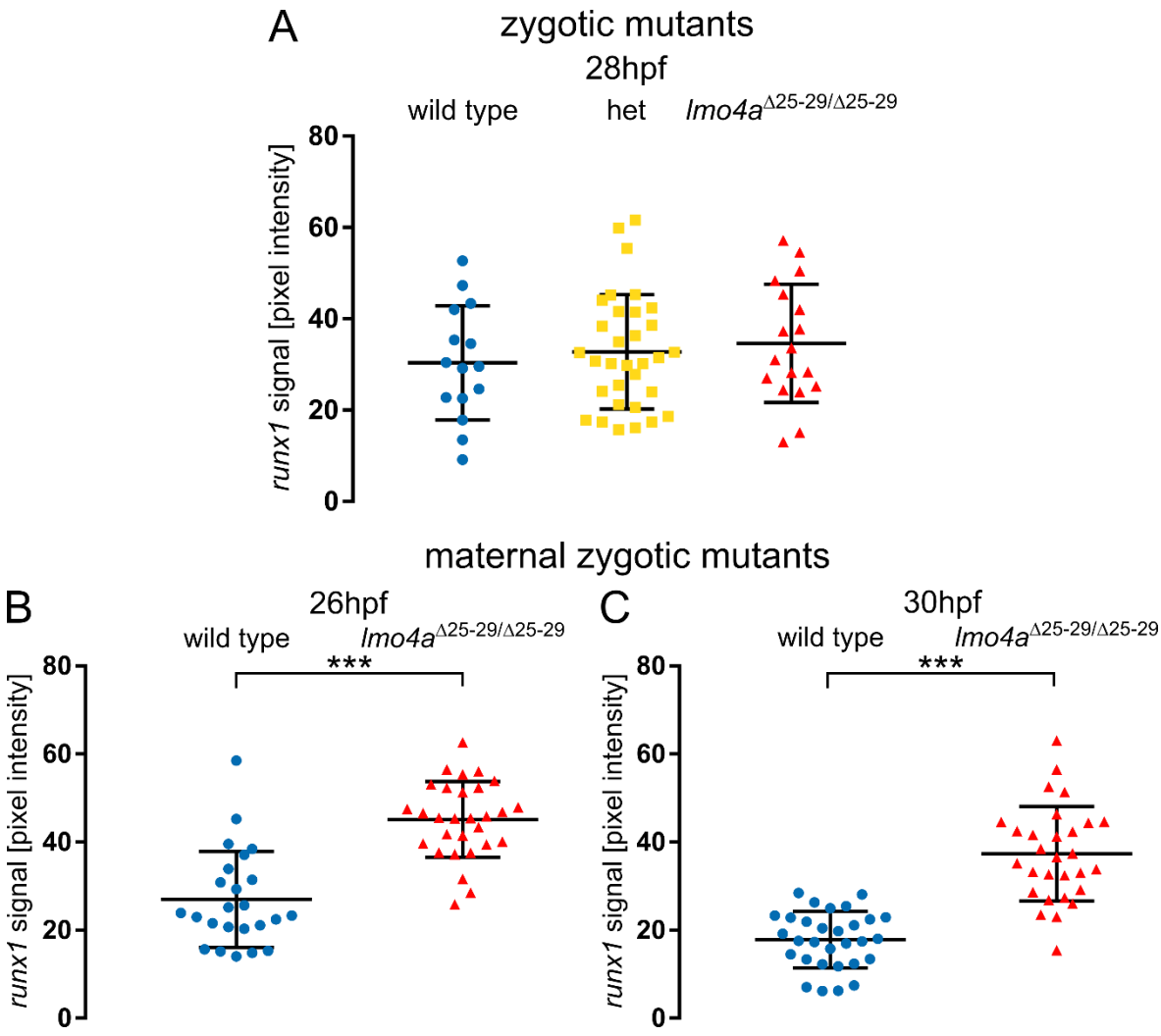


Figure 5.14. Expression of *runx1* is increased in the dorsal aorta of *lmo4a* maternal homozygous mutants.

(A) Quantification of the *runx1* mRNA signal, detected by *in situ* hybridization, from 28hpf wild type (blue), heterozygous *lmo4a*$^{+/-}$ (het; yellow) and *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ mutant (red) embryos from one clutch shows no significant difference in *runx1* pixel intensity among the different genotypes. (B-C) Quantification of the *runx1 in situ* hybridization signal from wild type (blue) and maternal zygotic *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ mutants (red) reveals significant increase in *runx1* mRNA expression upon *lmo4a* knockout compared to wild type. Error bars: mean±SD. ***p<0.001.

To analyse the early output of the HSCs generated from the dorsal aorta HE in the *lmo4a* morphants, we used *in situ* hybridization to probe for *rag1*, a marker of T lymphocytes in the thymus (Willett et al. 1997) that originate from the DA HSCs (Murayama et al. 2006). We found reduced expression of thymic *rag1* in most *lmo4a* morphants compared to control embryos (X^2=33.464, d.f.=1, p<0.001) (**Fig. 5.13D-E**). Importantly, expression of *foxn1*, marking thymic epithelium (Ma et al. 2012), was not affected (**Fig. 5.13F-G**), suggesting a loss of *rag1*$^+$ T-cells rather than a loss of thymus in the *lmo4a* morphants. Together, these data provide evidence that knockdown of *lmo4a* negatively affects the generation and activity of the first HSCs in the embryo.

5.8.2. Expression of *runx1* increases in *lmo4a* mutants

In the light of recent reports casting doubts on the specificity of MO-induced phenotypes (Kok et al. 2015; Place and Smith 2017), we analysed expression of *runx1* in TALEN-generated *lmo4a* mutants, described in Chapter 5.4.2. Quantitative *in situ* hybridization (**Fig. 2.3**) against *runx1* mRNA in a clutch of embryos coming from an incross of two *lmo4a*$^{+/-}$ heterozygotes showed no significant effect on *runx1* expression in *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ mutants at 28hpf (μ_{wt}=30.3, μ_{mut}=34.6) (**Fig. 5.14A**). However, when maternal zygotic *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ (obtained from a cross of two adult homozygotes) were compared to wild type embryos, we observed a significant increase in *runx1* pixel intensity in *lmo4a* mutants at 26hpf (μ_{wt}=26.9, μ_{mut}=45.1; t=6.778, d.f.=51, p<0.001) and 30hpf (μ_{wt}=17.8, μ_{mut}=37.4; t=8.581, d.f.=57, p<0.001) (**Fig. 5.14B-C**). This is contrary to the data from *lmo4a* morphants, despite the consistency of *gata2a* phenotype between the *lmo4a* morphants and mutants (see Chapter 5.4.2). This discrepancy suggests that there may be compensatory mechanisms rescuing *runx1* expression in *lmo4a* mutants that are absent in the morphants (El-Brolosy and Stainier 2017). However, it is also possible that the decreased *runx1* expression in *lmo4a* morphants results from off-target effects of the MO or due to technical differences during *in situ* hybridization. Interestingly, the effect observed in *lmo4a* mutants agrees

with a report from the mouse, where elevated expression of *Gata2* resulted in an upregulation of *Runx1* (Guiu et al. 2013). Importantly, this, in turn, resulted in the loss of functional HSPCs (Guiu et al. 2013). Further experiments will analyse the expression of *cmyb* in the CHT of 4dpf *lmo4a* mutants to assess the early output of DA HSCs and explain the discrepancy.

5.9. Discussion

5.9.1. Overview

Lmo (Lim domain only) proteins are a class of transcription factors involved in important developmental processes, including neural development and haematopoiesis (Sang et al. 2014). Although they do not bind to DNA, they play crucial roles in formation of DNA-binding regulatory complexes, such as the well characterized pentameric complex, consisting of Lmo2, Ldb1, E47, Gata1 and Tal1 (Wadman et al. 1997). While Lmo2 has a relatively well defined role in developmental haematopoiesis (Warren et al. 1994; Yamada et al. 1998; Patterson et al. 2007), initial studies have also suggested a role for Lmo4 in this process (Meier et al. 2006; K. Tuladhar, DPhil thesis). In this chapter, we have further investigated the role of *lmo4a* in regulating the genetic networks involved in HE programming.

5.9.2. Lmo4a is a negative regulator of *gata2a* expression in the endothelium

Lmo4 is a conserved scaffold protein, required for forming transcription factor complexes in various contexts (Sang et al. 2014). In the zebrafish, there are two *lmo4* homologues, *lmo4a* and *lmo4b*. Analysis of both paralogues has suggested evolutionary subfunctionalisation, assigning a cell-autonomous role in the specification of HE and HSCs to *lmo4a* but not *lmo4b* (K. Tuladhar, DPhil thesis). In particular, knockdown of *lmo4a* was found to result in upregulation of *gata2a* expression in the endothelium at the HE specification stages

(K. Tuladhar, DPhil thesis). We confirmed this observation with RNA-seq on FACS-sorted endothelial cells, defined as expressing high levels of *kdrl*:GFP and no *gata1*:DsRed (*kdrl*:GFP^{high}; *gata1*:DsRed^{-}). Further validation with *in situ* hybridization showed upregulated levels of *gata2a* in the dorsal aorta of *lmo4a* morphants compared to uninjected controls at 26-28hpf, consistent with previous data (K. Tuladhar, DPhil thesis). We also observed increased levels of *gata2a* mRNA upon *lmo4a* knockdown measured by qRT-PCR in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells at 22hpf and 28hpf. Previous qRT-PCR experiments, done in less stringently sorted *kdrl*:GFP^{+} cells, showed a significant increase in *gata2a* expression in *lmo4a* morphants at 24hpf and 30hpf (K. Tuladhar, DPhil thesis). Importantly, we verified the increase in *gata2a* expression by *in situ* hybridization in *lmo4a* mutants, generated for the purpose of this thesis using TALENs. In the light of recent discrepancies between morphant and mutant phenotypes in zebrafish (Kok et al. 2015), genomic mutants are becoming more and more critical in reverse genetic studies. We believe our *lmo4a* mutant will provide a useful tool to the zebrafish community, especially because homozygous mutant adults are viable and fertile.

We also made use of the Tg(*gata2a*-i4-450bp:GFP) transgenic line, described in detail in Chapter 3, to assess the effect of *lmo4a* knockdown on the activity of the *gata2a*-i4 enhancer. We found increased *gfp* mRNA levels in *lmo4a* morphants compared to wild type, suggesting that Lmo4a could act directly through the i4 enhancer in regulating *gata2a* expression. This is further supported by ChIP data from MEL cells, where Lmo4 bound to the mouse equivalent of *gata2a*-i4, the *Gata2* +9.5 enhancer (C. Andrieu-Soler, unpublished data). Furthermore, the +9.5 enhancer is bound by Ldb1 (Sanalkumar et al. 2014), a known protein partner of Lmo4 (Grutz et al. 1998; Meier et al. 2006). It is conceivable to suggest a complex nucleated around Lmo4a and Ldb1 binding to the *gata2a*-i4 enhancer in the developing zebrafish endothelium. In fact, this enhancer contains a composite E-box-(11bp)-Gata motif (Wozniak et al. 2008), recognized by Lmo2-Ldb1-Gata1-Tal1-containing protein complexes (Kassouf et al. 2010). Lmo2 and

Lmo4 differ in sequence and structure and to date there is no clear evidence that they can be interchangeable in protein complexes. However, they are highly homologous within the LIM domains, the main functional units of Lmo proteins (Deane et al. 2004; Sang et al. 2014). We hypothesize that alternative complexes compete for binding to the *gata2a*-i4 enhancer, one involving Lmo2 and the other one involving Lmo4a. The former would positively affect the expression of *gata2a*, while the latter would repress *gata2a* transcription. This idea is supported by Co-IP experiments in MEL cells, which identified multiple negative regulators of transcription as binding partners of Lmo4 (Meier et al. 2006; K. Tuladhar, DPhil thesis). The concept of different transcriptional output depending on different protein partners in DNA-binding complexes is not entirely new. Formation of a complex with Tal1 and Lmo2 conferred the activating, but not repressive, role of Gata1 (Tripic et al. 2009). In addition, Gata1 functions differently when complexed with Lmo2 or Eto2 (Fujiwara et al. 2010). Therefore, we speculate that Lmo2 and Lmo4a could compete for the DNA-binding transcription factors responsible for recognizing the *gata2a*-i4 enhancer. However, this hypothesis would have to be verified with ChIP experiments and reciprocal functional analyses on the activity of each of the two Lmo proteins in the absence of the other one.

5.9.3. Wnt signalling is a potential mediator of the *lmo4a-gata2a* regulation

Our RNA-seq experiment revealed nearly 700 genes downregulated upon *lmo4a* knockdown and this data set will be investigated in future projects. In this thesis, because of previous evidence for the existence of an Lmo4-mediated negative regulatory complex (K. Tuladhar, DPhil thesis), we focused on the genes upregulated upon *lmo4a* knockdown in the RNA-seq experiment. The GO term analysis revealed a significant enrichment for the genes involved in Wnt signalling, a known player in the development of HSPCs (Richter et al. 2017). We observed upregulation of a few Wnt signalling genes in the endothelium of *lmo4a* morphants by qRT-PCR, suggesting an intriguing global regulation of genes involved at different stages of

the signalling cascade by Lmo4a. Although this effect was transcriptional, another effect on the protein level cannot be excluded. A link between a cytosolic Lmo protein and Wnt signalling has been uncovered recently, whereby Lmo2 modulated Wnt signalling by binding to Dvl1/2 in solid tumours (Liu et al. 2016b). Therefore, in addition to modulating the transcription of some Wnt signalling genes, we speculate that Lmo4a could also have another function by binding to cytosolic members of the Wnt cascade. However, investigation of this hypothesis was beyond the scope of this thesis.

To test the idea that the increase in Wnt signalling in *lmo4a* morphants could mediate the elevated expression of *gata2a* mRNA, we treated *lmo4a* MO-injected zebrafish embryos with IWR-1, a known antagonist of Wnt signalling (Chen et al. 2009a). This treatment reduced the expression of *gata2a* back to control levels, positioning Wnt downstream of *lmo4a* in regulating *gata2a*.

We further speculated whether Wnt could act through the *gata2a*-i4 enhancer. We found a Tcf binding site within the enhancer, suggesting a putative direct regulation by Wnt (Cavallo et al. 1998; Daniels and Weis 2005). This site was mapped computationally and would have to be validated by ChIP or by CRISPR-mediated deletion. It would also be interesting to check sequence conservation of this site in the mouse +9.5 enhancer, which is bound by Lmo4 in MEL cells (C. Andrieu-Soler, unpublished data). Future projects could investigate potential direct interaction between Lmo4 and Tcf in regulating the conserved *Gata2* endothelial enhancer.

In a functional assay, treatment of Tg(*gata2a*-i4-1.1kb:GFP) and Tg(*gata2a*-i4-450bp:GFP) fish with IWR-1 and BIO, a known agonist of Wnt signalling (Meijer et al. 2003), revealed dynamic changes in *gfp* mRNA expression in response to modulation of Wnt: inhibition of Wnt decreased the activity of the enhancer, while its activation increased the enhancer activity. Intriguingly, expression of *gata2a* itself did not increase upon treatment with BIO, suggesting

that the i4 enhancer only regulates the expression of *gata2a* up to a certain level. There may be other fine-tuning mechanisms preventing *gata2a* overexpression in certain contexts, suggesting the importance of balanced levels of *gata2a* in the endothelium. Nevertheless, these data suggest that Wnt may be a mediator of the *lmo4a-gata2a* regulation. While previous reports demonstrated regulation of *Gata2* expression in the haematopoietic system by Notch (Robert-Moreno et al. 2005) and BMP (Kirmizitas et al. 2017), this is, to our knowledge, the first evidence linking Wnt to *Gata2*. Further experiments would quantify the expression of *gata2a* and verify this effect in *lmo4a* mutants. Future projects will also investigate whether the effect of modulated Wnt is relayed on the transcriptional level or whether it increases the levels of *gata2a* by stimulating cell proliferation.

5.9.4. A genetic network responsible for precise regulation of *gata2a* expression is required for HSC specification

Besides the above evidence for direct or Wnt-dependent regulation of *gata2a* by Lmo4a through the i4 enhancer, our RNA-seq analysis also revealed upregulation of other transcription factors, potentially acting upstream of *gata2a*. One of them is *Etv2*, a conserved upstream regulator of *Gata2* (Ciau-Uitz et al. 2013a; Liu et al. 2015). *Etv2* is an Ets transcription factor and it could regulate *gata2a* by direct binding to the Ets motifs within the i4 enhancer. Moreover, the data presented in Chapter 4 demonstrated that expression of *etv2* was not affected by the loss of *gata2a*, arguing against the possibility of *gata2a* acting upstream of *etv2*. Interestingly, we also observed upregulation of *foxc1a* upon *lmo4a* knockdown. *Foxc1a* was previously shown to regulate *etv2* in zebrafish angioblasts (Veldman and Lin 2012) and a similar mechanism may be in place in the endothelium of the dorsal aorta. This points to a potential *lmo4a-foxc1a-etv2-gata2a* gene regulatory cascade, which would have to be verified with epistasis experiments. Notably, this cascade could also involve *tal1*, a downstream target of *etv2* (Ren et al. 2010; Ciau-Uitz et al. 2013a) and a potential protein partner of Gata2a (Nottingham et al. 2007; Pimanda et al. 2007b;

Tripic et al. 2009; Wu et al. 2014; Org et al. 2015), which was also upregulated in *lmo4a* morphants. Such a position for *lmo4a* at the top of a transcription factor regulatory cascade is further supported by the Gene Ontology (GO) term analysis, which listed DNA-binding transcription factors among the most enriched within all the genes upregulated upon *lmo4a* knockdown. However, this *lmo4a*-dependent network likely does not involve *gata2b*. Although we observed upregulation of its expression in RNA-seq, further *in situ* hybridization and qRT-PCR experiments could not unambiguously verify the effect of *lmo4a* knockdown on *gata2b* expression. In the previous chapter, we demonstrated that endothelial Gata2a was necessary for early *gata2b* expression. The data presented here suggests that upregulation of *gata2a* induced by *lmo4a* knockdown is not sufficient to cause upregulation of *gata2b*. Gata2b is critically required in the zebrafish for the specification of the HE (Butko et al. 2015) so it is likely that the Gata2b-dependent mechanisms of HE programming are not affected in *lmo4a* morphants.

A few *in vitro* and *in vivo* studies suggested that too much *Gata2* negatively affected HSPC functionality and balanced *Gata2* expression was required to allow the generation of functional HSPCs (Persons et al. 1999; Tipping et al. 2009; Guiu et al. 2013; Gao et al. 2016). Consistent with that, we observed reduced endothelial expression of *runx1* and *cmyb*, the markers of developing HSCs (Thompson et al. 1998; Burns et al. 2002; Kalev-Zylinska et al. 2002; Gering and Patient 2005) in *lmo4a* morphants compared to wild types. We also observed a reduction in *gfi1aa*, a marker of HE (Cooney et al. 2013; Thambyrajah et al. 2016), and reduced *rag1* expression in the thymus (Willett et al. 1997; Murayama et al. 2006), suggesting that knockdown of *lmo4a* indeed affected HE and HSC specification. We speculate that this haematopoietic defect could be due to increased expression of *sox7*, potentially acting downstream of *gata2a*. Downregulation of *sox7* expression is necessary to allow haematopoietic differentiation (Gandillet et al. 2009; Costa et al. 2012b) and the loss of HSCs could result from maintained Sox7 activity. Importantly, generation of the dorsal aorta, the source of first embryonic HSCs,

was not affected in *lmo4a* morphants, consistent with previous observations (K. Tuladhar, DPhil thesis).

The *lmo4a*-induced loss of HSCs can be rescued by chemical inhibition of Gata activity (K. Tuladhar, DPhil thesis), strongly suggesting the importance of *gata2a* in mediating the haematopoietic phenotype. This also further supports the model in which balanced, permissive levels of *Gata2* in the endothelium are required for the generation of HSPCs (Guiu et al. 2013), although the precise mechanism is still unknown. Since Gata2 is a DNA-binding transcription factor involved in multiple complexes, its overexpression could result in ectopic binding or exhaustion of its partners from other complexes.

5.9.5. Compensatory mechanisms potentially rescue *runx1* expression in *lmo4a* mutants

Because of the known discrepancies between MO-induced and genomic mutant phenotypes (Kok et al. 2015; Place and Smith 2017), we measured the expression of *runx1* in the *lmo4a* homozygous mutants. We observed no difference in the embryos coming from the incross of *lmo4a*$^{+/-}$ heterozygotes, explainable by the activity of maternally-deposited *lmo4a* mRNA (Lane et al. 2002). However, when we compared maternal zygotic homozygous *lmo4a* mutants to wild type embryos, we observed, surprisingly, a significant increase in *runx1* signal in the mutants at 26hpf and 30hpf. This is in agreement with mouse data, where an increase in *Gata2* was associated with an increase in *Runx1* expression (Guiu et al. 2013). Nevertheless, there is a discrepancy between zebrafish *lmo4a* morphants and mutants. It is possible that the *lmo4a* mutants activate compensatory mechanisms, more and more widely reported in the zebrafish field (El-Brolosy and Stainier 2017). Therefore, our mutants may contribute to improving the understanding of these mechanisms in the future. Potential compensation could come from the closest relative of *lmo4a*, *lmo4b*, or even from *lmo2*. However, this would need verifying with expression analysis in the *lmo4a* mutants and potentially generation of double

mutants. On the other hand, the observed discrepancy between mutants and morphants may arise from off-target effects on *runx1* expression in the *lmo4a* MO-injected embryos. In this instance, the newly generated *lmo4a* mutants will provide an invaluable tool to understand the role of Lmo4a in the HE specification and HSC programming.

Conclusions

Haematopoietic stem cells (HSCs) are arguably the best model for stem cell biology (Orkin and Zon 2008). Generating HSCs *in vitro* holds a great potential, both biological and clinical (Daniel et al. 2016). However, the future success of these efforts likely depends on the thorough understanding of the molecular and cellular processes governing the *de novo* generation of HSCs during embryonic development. Although numerous mechanisms have been uncovered, recent views point to the importance of more global understanding of the molecular networks orchestrating definitive haematopoiesis. Integration of gene regulatory networks, *cis*-regulatory elements and intercellular signalling cues will likely provide a better overview of the intriguing nuances that distinguish the multilineage, long-term engrafting HSCs from other blood progenitor populations. We have explored these avenues in the context of the conserved transcription factor Gata2.

Gata2 is critically required for the generation and maintenance of embryonic HSCs (Tsai et al. 1994; de Pater et al. 2013). Mutations in the *GATA2* locus have been associated with human haematological disorders (Spinner et al. 2014). Expression of *Gata2* in the haemogenic endothelium (HE) and its subsequent role in the specification of the HSCs depends on its intronic enhancer, identified in the mouse and human (Khandekar et al. 2007; Johnson et al. 2012; Gao et al. 2013). We have investigated the conservation of this enhancer in other vertebrates and identified a homologous region in the intron 4 of zebrafish *gata2a* locus (i4) (Chapter 3). This region is characterised by open chromatin specifically in the endothelial cells of the developing embryo. We have established transgenic lines driving pan-endothelial expression of GFP under the control of this enhancer, including the HE, thus demonstrating a high level of functional evolutionary conservation of the i4 region. We further demonstrated that the reporter expression of *gfp* in these lines is highly dynamic and responsive to perturbations affecting *gata2* expression *in vivo*. Thus, we created a new tool to quantify tissue-specific activity of *cis*-regulatory elements *in vivo*.

Furthermore, we have generated a library of open chromatin sites in the endothelial cell population by ATAC-seq (Chapter 3). Future analysis of these genome-wide data will provide a wealth of information that we can explore to identify novel regulatory elements involved in the HE programming and HSC formation.

We employed CRISPR/Cas9 to delete the i4 enhancer from the *gata2a* locus, describing the first instance when deletion of an endogenous *cis* element in the zebrafish directly affects tissue-specific expression of its regulated gene (Chapter 4). The *gata2a*$^{\Delta i4/\Delta i4}$ mutation caused an endothelial-specific loss of *gata2a* expression during the HE programming stages, which allowed us to characterise the previously unexplored function of *gata2a* in definitive haematopoiesis. Gata2a acts upstream of its paralogue, *gata2b*, at the early stages of HE programming, but becomes redundant by 2dpf and *gata2a*$^{\Delta i4/\Delta i4}$ mutants eventually recover from the early defects in HE programming and are able to grow to adulthood. Future explanation of the haematopoietic recovery will improve our understanding of how different molecular inputs feed into the programming of HE and generation of HSCs. Moreover, our results suggest that the duplication of *gata2* genes in the zebrafish (Gillis et al. 2009) likely resulted in haematopoietic subfunctionalisation of the paralogues, providing an opportunity to unpick the different roles of *Gata2* in the multi-step process of definitive haematopoiesis. For example, the *gata2a*$^{\Delta i4/\Delta i4}$ mutants are prone to infections (Chapter 4) and recent preliminary data from our lab (M. Krecsmarik, R Monteiro, personal communication) suggest that the kidney marrow is hypocellular. These phenotypes are typically found in humans suffering from MonoMAC syndrome, caused by *Gata2* haploinsufficiency (Hsu et al. 2013). This suggests that the zebrafish *gata2a* and *gata2b* paralogues may, in the adult bone marrow, function as *gata2* 'alleles'. Therefore, we hope to explore the use of our *gata2a*$^{\Delta i4/\Delta i4}$ mutants as an animal model for human diseases linked to *Gata2* haploinsufficiency.

In this thesis, we also explored *trans* regulatory mechanisms regulating *gata2a* expression in the endothelium. Previous work in our lab identified *lmo4a* as a crucial negative regulator of *gata2a* endothelial expression (K. Tuladhar, DPhil thesis) and we further investigated this relationship. Our evidence suggests that Lmo4a is a more global negative regulator acting upstream of a transcriptional network that programmes the HE (Chapter 5). This network potentially includes Wnt signalling, considered to be among the most important cues distinguishing primitive from definitive haematopoiesis (Ditadi et al. 2017). Intermediate levels of Gata2 are required for the correct specification of the HSPCs (Guiu et al. 2013; K. Tuladhar, DPhil thesis). We provide evidence for a novel Lmo4a-Wnt-Gata2a cascade, potentially acting through the i4 enhancer, establishing intermediate levels of *gata2a* expression in the HE. We also used TALENs to generate a new *lmo4a* mutant, which phenocopied the MO-induced upregulation of *gata2a*. However, we observed a discrepancy in the expression of *runx1* between mutants and morphants; the mutants did not show decreased *runx1* expression in the HE. Of note, there is a second zebrafish paralogue of *Lmo4*, *lmo4b*, that could potentially explain the mutant phenotype. Future analysis will likely provide new insights into the incompletely understood mechanism of genetic compensation in zebrafish, a phenomenon that has been widely observed in genetic knockouts, but not knockdowns (El-Brolosy and Stainier 2017).

In summary, this thesis provides new insights into the molecular networks establishing the definitive haematopoietic programme in the embryo. Most of the biological insights are directly relevant to the zebrafish, but high-degree conservation of the haematopoietic development will likely help understanding the process of haematopoiesis in other vertebrate systems (Robertson et al. 2016). We also provide new technical approaches, describing *in vivo* quantification of enhancer activity using transgenic zebrafish lines. Moreover, we describe the first deletion of an endogenous enhancer in the zebrafish, resulting in tissue-specific loss of target gene expression. We expect to set a precedent in the field so that more endogenous enhancers will be successfully targeted in the future. We believe that this work will likely

contribute to the understanding of the effects of genomic mutations, aiding the generation of zebrafish mutants in the future.

References

Amaya E. 2013. The hemangioblast: a state of competence. *Blood* **122**: 3853-3854.
Andersson ER, Sandberg R, Lendahl U. 2011. Notch signaling: simplicity in design, versatility in function. *Development* **138**: 3593-3612.
Avagyan S, Zon LI. 2016. Fish to Learn: Insights into Blood Development and Blood Disorders from Zebrafish Hematopoiesis. *Hum Gene Ther* **27**: 287-294.
Ayllón V, Bueno C, Ramos-Mejía V, Navarro-Montero O, Prieto C, Real PJ, Romero T, García-León MJ, Toribio ML, Bigas A, Menendez P. 2015. The Notch ligand DLL4 specifically marks human hematoendothelial progenitors and regulates their hematopoietic fate. *Leukemia* **29**: 1741-1753.
Baltzinger M, Mager-Heckel AM, Remy P. 1999. Xl erg: expression pattern and overexpression during development plead for a role in endothelial cell differentiation. *Dev Dyn* **216**: 420-433.
Banerji J, Rusconi S, Schaffner W. 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**: 299-308.
Bassett AR, Tibbit C, Ponting CP, Liu JL. 2013. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. *Cell Rep* **4**: 220-228.
Batta K, Florkowska M, Kouskoff V, Lacaud G. 2014. Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *Cell Rep* **9**: 1871-1884.
Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG, 2nd, Tan W, Penheiter SG, Ma AC, Leung AY, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC. 2012. In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**: 114-118.
Bee T, Swiers G, Muroi S, Pozner A, Nottingham W, Santos AC, Li P-SS, Taniuchi I, de Bruijn MF. 2010. Nonredundant roles for Runx1 alternative promoters reflect their activity at discrete stages of developmental hematopoiesis. *Blood* **115**: 3042-3050.
Bender MA, Bulger M, Close J, Groudine M. 2000. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. *Mol Cell* **5**: 387-393.
Bennett CM, Kanki JP, Rhodes J, Liu TX, Paw BH, Kieran MW, Langenau DM, Delahaye-Brown A, Zon LI, Fleming MD, Look AT. 2001. Myelopoiesis in the zebrafish, Danio rerio. *Blood* **98**: 643-651.
Bertrand JY, Chi NC, Santoso B, Teng S, Stainier DY, Traver D. 2010a. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* **464**: 108-111.
Bertrand JY, Cisson JL, Stachura DL, Traver D. 2010b. Notch signaling distinguishes 2 waves of definitive hematopoiesis in the zebrafish embryo. *Blood* **115**: 2777-2783.
Bertrand JY, Kim AD, Violette EP, Stachura DL, Cisson JL, Traver D. 2007. Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. *Development* **134**: 4147-4156.
Bertwistle D, Walmsley ME, Read EM, Pizzey JA, Patient RK. 1996. GATA factors and the origins of adult and embryonic blood in Xenopus: responses to retinoic acid. *Mech Dev* **57**: 199-214.
Bibikova M, Golic M, Golic KG, Carroll D. 2002. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**: 1169-1175.
Bigas A, Guiu J, Gama-Norton L. 2013. Notch and Wnt signaling in the emergence of hematopoietic stem cells. *Blood Cells Mol Dis* **51**: 264-270.
Birnbaum RY, Clowney EJ, Agamy O, Kim MJ, Zhao J, Yamanaka T, Pappalardo Z, Clarke SL, Wenger AM, Nguyen L, Gurrieri F, Everman DB, Schwartz CE, Birk OS, Bejerano G, Lomvardas S, Ahituv N. 2012. Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res* **22**: 1059-1068.
Blow MJ, McCulley DJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Bristow J, Ren B, Black BL, Rubin EM, Visel A, Pennacchio LA. 2010. ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* **42**: 806-810.
Boisset JC, Clapes T, Klaus A, Papazian N, Onderwater J, Mommaas-Kienhuis M, Cupedo T, Robin C. 2015. Progressive maturation toward hematopoietic stem cells in the mouse embryo aorta. *Blood* **125**: 465-469.
Boisset JC, van Cappellen W, Andrieu-Soler C, Galjart N, Dzierzak E, Robin C. 2010. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* **464**: 116-120.
Bos FL, Hawkins JS, Zovein AC. 2015. Single-cell resolution of morphological changes in hemogenic endothelium. *Development* **142**: 2719-2724.

Bresciani E, Carrington B, Wincovitch S, Jones M, Gore AV, Weinstein BM, Sood R, Liu PP. 2014. CBFβ and RUNX1 are required at 2 different steps during the development of hematopoietic stem cells in zebrafish. *Blood* **124**: 70-78.

Brown LA, Rodaway AR, Schilling TF, Jowett T, Ingham PW, Patient RK, Sharrocks AD. 2000. Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. *Mech Dev* **90**: 237-252.

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**: 1213-1218.

Burns CE, DeBlasio T, Zhou Y, Zhang J, Zon L, Nimer SD. 2002. Isolation and characterization of runxa and runxb, zebrafish members of the runt family of transcriptional regulators. *Exp Hematol* **30**: 1381-1389.

Burns CE, Traver D, Mayhall E, Shepard JL, Zon LI. 2005. Hematopoietic stem cell fate is established by the Notch-Runx pathway. *Genes Dev* **19**: 2331-2342.

Butko E, Distel M, Pouget C, Weijts B, Kobayashi I, Ng K, Mosimann C, Poulain FE, McPherson A, Ni CW, Stachura DL, Del Cid N, Espin-Palazon R, Lawson ND, Dorsky R, Clements WK, Traver D. 2015. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development* **142**: 1050-1061.

Butko E, Pouget C, Traver D. 2016. Complex regulation of HSC emergence by the Notch signaling pathway. *Dev Biol* **409**: 129-138.

Cai Z, de Bruijn M, Ma X, Dortland B, Luteijn T, Downing RJ, Dzierzak E. 2000. Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. *Immunity* **13**: 423-431.

Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, Maeda T, Paw BH, Orkin SH. 2014. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J Biol Chem* **289**: 21312-21324.

Canver MC, Smith EC, Sher F, Pinello L, Sanjana NE, Shalem O, Chen DD, Schupp PG, Vinjamur DS, Garcia SP, Luc S, Kurita R, Nakamura Y, Fujiwara Y, Maeda T, Yuan G-CC, Zhang F, Orkin SH, Bauer DE. 2015. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**: 192-197.

Cavallo RA, Cox RT, Moline MM, Roose J, Polevoy GA, Clevers H, Peifer M, Bejsovec A. 1998. Drosophila Tcf and Groucho interact to repress Wingless signalling activity. *Nature* **395**: 604-608.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**: e82.

Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, Myers RM, Petrov D, Jónsson B, Schluter D, Bell MA, Kingsley DM. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**: 302-305.

Chanda B, Ditadi A, Iscove NN, Keller G. 2013. Retinoic Acid Signaling Is Essential for Embryonic Hematopoietic Stem Cell Development. *Cell* **155**: 215-227.

Charbord P, Pouget C, Binder H, Dumont F, Stik G, Levy P, Allain F, Marchal C, Richter J, Uzan B, Pflumio F, Letourneur F, Wirth H, Dzierzak E, Traver D, Jaffredo T, Durand C. 2014. A systems biology approach for defining the molecular framework of the hematopoietic stem cell niche. *Cell stem cell* **15**: 376-391.

Chen B, Dodge ME, Tang W, Lu J, Ma Z, Fan C-W, Wei S, Hao W, Kilgore J, Williams NS, Roth MG, Amatruda JF, Chen C, Lum L. 2009a. Small molecule–mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nat Chem Biol* **5**: 100-107.

Chen MJ, Li Y, De Obaldia ME, Yang Q, Yzaguirre AD, Yamada-Inagawa T, Vink CS, Bhandoola A, Dzierzak E, Speck NA. 2011. Erythroid/myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells. *Cell Stem Cell* **9**: 541-552.

Chen MJ, Yokomizo T, Zeigler BM, Dzierzak E, Speck NA. 2009b. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* **457**: 887-891.

Chen XD, Turpen JB. 1995. Intraembryonic origin of hepatic hematopoiesis in Xenopus laevis. *J Immunol* **154**: 2557-2567.

Choi K, Kennedy M, Kazarov A, Papadimitriou JC, Keller G. 1998. A common precursor for hematopoietic and endothelial cells. *Development* **125**: 725-732.

Ciau-Uitz A, Liu F, Patient R. 2010a. Genetic control of hematopoietic development in Xenopus and zebrafish. *Int J Dev Biol* **54**: 1139-1149.

Ciau-Uitz A, Monteiro R, Kirmizitas A, Patient R. 2014. Developmental hematopoiesis: ontogeny, genetic programming and conservation. *Exp Hematol* **42**: 669-683.

Ciau-Uitz A, Patient R. 2016. The embryonic origins and genetic programming of emerging haematopoietic stem cells. *FEBS Lett* **590**: 4002-4015.

Ciau-Uitz A, Pinheiro P, Gupta R, Enver T, Patient R. 2010b. Tel1/ETV6 specifies blood stem cells through the agency of VEGF signaling. *Dev Cell* **18**: 569-578.

Ciau-Uitz A, Pinheiro P, Kirmizitas A, Zuo J, Patient R. 2013a. VEGFA-dependent and -independent pathways synergise to drive Scl expression and initiate programming of the blood stem cell lineage in Xenopus. *Development* **140**: 2632-2642.

Ciau-Uitz A, Walmsley M, Patient R. 2000. Distinct Origins of Adult and Embryonic Blood in Xenopus. *Cell* **102**: 787-796.

Ciau-Uitz A, Wang L, Patient R, Liu F. 2013b. ETS transcription factors in hematopoietic stem cell development. *Blood Cells Mol Dis* **51**: 248-255.

Clarke RL, Yzaguirre AD, Yashiro-Ohtani Y, Bondue A, Blanpain C, Pear WS, Speck NA, Keller G. 2013. The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol* **15**: 502-510.

Clements WK, Kim AD, Ong KG, Moore JC, Lawson ND, Traver D. 2011. A somitic Wnt16/Notch pathway specifies haematopoietic stem cells. *Nature* **474**: 220-224.

Clements WK, Traver D. 2013. Signalling pathways that control vertebrate haematopoietic stem cell specification. *Nat Rev Immunol* **13**: 336-348.

Cooney JD, Hildick-Smith GJ, Shafizadeh E, McBride PF, Carroll KJ, Anderson H, Shaw GC, Tamplin OJ, Branco DS, Dalton AJ, Shah DI, Wong C, Gallagher PG, Zon LI, North TE, Paw BH. 2013. Teleost growth factor independence (gfi) genes differentially regulate successive waves of hematopoiesis. *Dev Biol* **373**: 431-441.

Corada M, Orsenigo F, Morini MF, Pitulescu ME, Bhat G, Nyqvist D, Breviario F, Conti V, Briot A, Iruela-Arispe ML, Adams RH, Dejana E. 2013. Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nat Commun* **4**: 2609.

Coşkun S, Chao H, Vasavada H, Heydari K, Gonzales N, Zhou X, de Crombrugghe B, Hirschi KK. 2014. Development of the fetal bone marrow niche and regulation of HSC quiescence and homing ability by emerging osteolineage cells. *Cell Rep* **9**: 581-590.

Costa G, Kouskoff V, Lacaud G. 2012a. Origin of blood cells and HSC production in the embryo. *Trends Immunol* **33**: 215-223.

Costa G, Mazan A, Gandillet A, Pearson S, Lacaud G, Kouskoff V. 2012b. SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. *Development* **139**: 1587-1598.

Coultas L, Nieuwenhuis E, Anderson GA, Cabezas J, Nagy A, Henkelman RM, Hui C-CC, Rossant J. 2010. Hedgehog regulates distinct vascular patterning events through VEGF-dependent and -independent mechanisms. *Blood* **116**: 653-660.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**: 21931-21936.

Dailey L. 2015. High throughput technologies for the functional discovery of mammalian enhancers: New approaches for understanding transcriptional regulatory network dynamics. *Genomics* **106**: 151-158.

Daniel MG, Pereira C-F, Lemischka IR, Moore KA. 2016. Making a Hematopoietic Stem Cell. *Trends Cell Biol* **26**: 202-214.

Daniels DL, Weis WI. 2005. Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol* **12**: 364-371.

Davidson AJ, Zon LI. 2004. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene* **23**: 7233-7246.

de Bruijn M, Dzierzak E. 2017. Runx transcription factors in the development and function of the definitive hematopoietic system. *Blood* **129**: 2061-2069.

de Bruijn MF, Ma X, Robin C, Ottersbach K, Sanchez M-JJ, Dzierzak E. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* **16**: 673-683.

de Bruijn MF, Speck NA, Peeters MC, Dzierzak E. 2000. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *Embo J* **19**: 2465-2474.

de Pater E, Kaimakis P, Vink CS, Yokomizo T, Yamada-Inagawa T, van der Linden R, Kartalaei PS, Camper SA, Speck N, Dzierzak E. 2013. Gata2 is required for HSC generation and survival. *J Exp Med* **210**: 2843-2850.

Deane JE, Ryan DP, Sunde M, Maher MJ, Guss MJ, Visvader JE, Matthews JM. 2004. Tandem LIM domains provide synergistic binding in the LMO4:Ldb1 complex. *Embo J* **23**: 3589-3598.

Detrich HW, Kieran MW, Chan FY, Barone LM, Yee K, Rundstadler JA, Pratt S, Ransom D, Zon LI. 1995. Intraembryonic hematopoietic cell migration during vertebrate development. *Proc Natl Acad Sci U S A* **92**: 10713-10717.

Dieterlen-Lièvre F, Martin C. 1981. Diffuse intraembryonic hemopoiesis in normal and chimeric avian development. *Dev Biol* **88**: 180-191.

Ditadi A, Sturgeon CM, Keller G. 2017. A view of human haematopoietic development from the Petri dish. *Nat Rev Mol Cell Biol* **18**: 56-67.

Ditadi A, Sturgeon CM, Tober J, Awong G, Kennedy M, Yzaguirre AD, Azzola L, Ng ES, Stanley EG, French DL, Cheng X, Gadue P, Speck NA, Elefanty AG, Keller G. 2015. Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat Cell Biol* **17**: 580-591.

Domazet-Lošo T, Tautz D. 2010. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* **468**: 815-818.

Donaldson IJ, Chapman M, Kinston S, Landry JR, Knezevic K, Piltz S, Buckley N, Green AR, Göttgens B. 2005. Genome-wide identification of cis-regulatory sequences controlling blood and endothelial development. *Hum Mol Genet* **14**: 595-601.

Dooley KA, Davidson AJ, Zon LI. 2005. Zebrafish scl functions independently in hematopoietic and endothelial development. *Dev Biol* **277**: 522-536.

Doulatov S, Vo LT, Chou SS, Kim PG, Arora N, Li H, Hadland BK, Bernstein ID, Collins JJ, Zon LI, Daley GQ. 2013. Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. *Cell stem cell* **13**: 459-470.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL. 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol* **26**: 702-708.

Drake CJ, Brandt SJ, Trusk TC, Little CD. 1997. TAL1/SCL is expressed in endothelial progenitor cells/angioblasts and defines a dorsal-to-ventral gradient of vasculogenesis. *Dev Biol* **192**: 17-30.

Duarte A, Hirashima M, Benedito R, Trindade A, Diniz P, Bekman E, Costa L, Henrique D, Rossant J. 2004. Dosage-sensitive requirement for mouse Dll4 in artery development. *Genes Dev* **18**: 2474-2478.

Durand C, Robin C, Bollerot K, Baron MH, Ottersbach K, Dzierzak E. 2007. Embryonic stromal clones reveal developmental regulators of definitive hematopoietic stem cells. *Proc Natl Acad Sci U S A* **104**: 20838-20843.

Dzierzak E, Speck NA. 2008. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat Immunol* **9**: 129-136.

Eilken HM, Nishikawa S-I, Schroeder T. 2009. Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457**: 896-900.

El-Brolosy MA, Stainier DY. 2017. Genetic compensation: A phenomenon in search of mechanisms. *PLoS genetics* **13**: e1006780.

El Omari K, Hoosdally SJ, Tuladhar K, Karia D, Hall-Ponsele E, Platonova O, Vyas P, Patient R, Porcher C, Mancini EJ. 2013. Structural basis for LMO2-driven recruitment of the SCL:E47bHLH heterodimer to hematopoietic-specific transcriptional targets. *Cell Rep* **4**: 135-147.

Elcheva I, Brok-Volchanskaya V, Kumar A, Liu P, Lee JH, Tong L, Vodyanik M, Swanson S, Stewart R, Kyba M, Yakubov E, Cooke J, Thomson JA, Slukvin I. 2014. Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. *Nat Commun* **5**: 4372.

Elefanty AG, Begley CG, Hartley L, Papaevangeliou B, Robb L. 1999. SCL expression in the mouse embryo detected with a targeted lacZ reporter gene demonstrates its localization to hematopoietic, vascular, and neural tissues. *Blood* **94**: 3754-3763.

Emambokus N, Vegiopoulos A, Harman B, Jenkinson E, Anderson G, Frampton J. 2003. Progression through key stages of haemopoiesis is dependent on distinct threshold levels of c-Myb. *Embo J* **22**: 4478-4488.

Encode Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.

Espín-Palazón R, Stachura DL, Campbell CA, García-Moreno D, Del Cid N, Kim AD, Candel S, Meseguer J, Mulero V, Traver D. 2014. Proinflammatory signaling regulates hematopoietic stem cell emergence. *Cell* **159**: 1070-1085.

Faul F, Erdfelder E, Lang A-G, Buchner A. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* **39**: 175-191.

Ferkowicz MJ, Starr M, Xie X, Li W, Johnson SA, Shelley WC, Morrison PR, Yoder MC. 2003. CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development* **130**: 4393-4403.

Ferkowicz MJ, Yoder MC. 2005. Blood island formation: longstanding observations and modern interpretations. *Exp Hematol* **33**: 1041-1047.

Fiedler J, Jazbutyte V, Kirchmaier BC, Gupta SK, Lorenzen J, Hartmann D, Galuppo P, Kneitz S, Pena J, Sohn-Lee C, Loyer X, Soutschek J, Brand T, Tuschl T, Heineke J, Martin U, Schulte-Merker S, Ertl G, Engelhardt S, Bauersachs J, Thum T. 2011. MicroRNA-24 Regulates Vascularity After Myocardial Infarction. *Circulation* **124**: 720-730.

Fitch SR, Kimber GM, Wilson NK, Parker A, Mirshekar-Syahkal B, Göttgens B, Medvinsky A, Dzierzak E, Ottersbach K. 2012. Signaling from the Sympathetic Nervous System Regulates Hematopoietic Stem Cell Emergence during Embryogenesis. *Cell Stem Cell* **11**: 554-566.

Fouquet B, Weinstein BM, Serluca FC, Fishman MC. 1997. Vessel patterning in the embryo of the zebrafish: guidance by notochord. *Dev Biol* **183**: 37-48.

Frame JM, Fegan KH, Conway SJ, McGrath KE, Palis J. 2016. Definitive Hematopoiesis in the Yolk Sac Emerges from Wnt-Responsive Hemogenic Endothelium Independently of Circulation and Arterial Identity. *Stem Cells* **34**: 431-444.

Frame JM, McGrath KE, Palis J. 2013. Erythro-myeloid progenitors: "Definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells Mol Dis* **51**: 220-225.

Fujiwara T, Lee HY, Sanalkumar R, Bresnick EH. 2010. Building multifunctionality into a complex containing master regulators of hematopoiesis. *Proc Natl Acad Sci U S A* **107**: 20429-20434.

Galloway JL, Wingert RA, Thisse C, Thisse B, Zon LI. 2005. Loss of Gata1 but Not Gata2 Converts Erythropoiesis to Myelopoiesis in Zebrafish Embryos. *Dev Cell* **8**: 109-116.

Gama-Norton L, Ferrando E, Ruiz-Herguido C, Liu Z, Liu Z, Guiu J, Islam AB, Lee S-UU, Yan M, Guidos CJ, López-Bigas N, Maeda T, Espinosa L, Kopan R, Bigas A. 2015. Notch signal strength controls cell fate in the haemogenic endothelium. *Nat Commun* **6**: 8510.

Gandillet A, Serrano AG, Pearson S, Lie-A-Ling M, Lacaud G, Kouskoff V. 2009. Sox7-sustained expression alters the balance between proliferation and differentiation of hematopoietic progenitors at the onset of blood specification. *Blood* **114**: 4813-4822.

Gao X, Johnson KD, Chang YI, Boyer ME, Dewey CN, Zhang J, Bresnick EH. 2013. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med* **210**: 2833-2842.

Gao X, Wu T, Johnson KD, Lahvic JL, Ranheim EA, Zon LI, Bresnick EH. 2016. GATA Factor-G-Protein-Coupled Receptor Circuit Suppresses Hematopoiesis. *Stem cell reports* **6**: 368-382.

Garcia-Porrero JA, Manaia A, Jimeno J, Lasky LL, Dieterlen-Lièvre F, Godin IE. 1998. Antigenic profiles of endothelial and hemopoietic lineages in murine intraembryonic hemogenic sites. *Dev Comp Immunol* **22**: 303-319.

Garriock RJ, Mikawa T. 2011. Early arterial differentiation and patterning in the avian embryo model. *Semin Cell Dev Biol* **22**: 985-992.

Garza A, Sinha A, Bowman TV. 2017. Concise Review: Hematopoietic Stem Cell Origins: Lessons from Embryogenesis for Improving Regenerative Medicine. *STEM CELLS Translational Medicine* **6**: 60-67.

Genthe JR, Clements WK. 2017. R-spondin 1 is required for specification of hematopoietic stem cells through Wnt16 and Vegfa signaling pathways. *Development* **144**: 590-600.

Gerety SS, Wang HU, Chen ZF, Anderson DJ. 1999. Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development. *Mol Cell* **4**: 403-414.

Gering M, Patient R. 2005. Hedgehog signaling is required for adult blood stem cell formation in zebrafish embryos. *Dev Cell* **8**: 389-400.

Gering M, Rodaway AR, Göttgens B, Patient RK, Green AR. 1998. The SCL gene specifies haemangioblast development from early mesoderm. *Embo J* **17**: 4029-4045.

Gering M, Yamada Y, Rabbitts TH, Patient RK. 2003. Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. *Development* **130**: 6187-6199.

Gillis WQ, St John J, Bowerman B, Schneider SQ. 2009. Whole genome duplications and expansion of the vertebrate GATA transcription factor gene family. *BMC Evol Biol* **9**: 207.

Goessling W, North TE, Loewer S, Lord AM, Lee S, Stoick-Cooper CL, Weidinger G, Puder M, Daley GQ, Moon RT, Zon LI. 2009. Genetic interaction of PGE2 and Wnt signaling regulates developmental specification of stem cells and regeneration. *Cell* **136**: 1136-1147.

Gomez Perdiguero E, Klapproth K, Schulz C, Busch K, Azzoni E, Crozet L, Garner H, Trouillet C, de Bruijn MF, Geissmann F, Rodewald H-RR. 2015. Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* **518**: 547-551.

Gore AV, Athans B, Iben JR, Johnson K, Russanova V, Castranova D, Pham VN, Butler MG, Williams-Simons L, Nichols JT, Bresciani E, Feldman B, Kimmel CB, Liu PP, Weinstein BM. 2016. Epigenetic regulation of hematopoiesis by DNA methylation. *eLife* **5**: e11813.

Gore AV, Swift MR, Cha YR, Lo B, McKinney MC, Li W, Castranova D, Davis A, Mukouyama Y-sS, Weinstein BM. 2011. Rspo1/Wnt signaling promotes angiogenesis via Vegfc/Vegfr3. *Development* **138**: 4875-4886.

Gori JL, Butler JM, Chan Y-YY, Chandrasekaran D, Poulos MG, Ginsberg M, Nolan DJ, Elemento O, Wood BL, Adair JE, Rafii S, Kiem H-PP. 2015. Vascular niche promotes hematopoietic multipotent progenitor formation from pluripotent stem cells. *J Clin Invest* **125**: 1243-1254.

Göttgens B, Nastos A, Kinston S, Piltz S, Delabesse EC, Stanley M, Sanchez M-JJ, Ciau-Uitz A, Patient R, Green AR. 2002. Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *EMBO J* **21**: 3039-3050.

Grainger S, Richter J, Palazón RE, Pouget C, Lonquich B, Wirth S, Grassme KS, Herzog W, Swift MR, Weinstein BM, Traver D, Willert K. 2016. Wnt9a Is Required for the Aortic Amplification of Nascent Hematopoietic Stem Cells. *Cell Rep* **17**: 1595-1606.

Grass JA, Jing H, Kim SI, Martowicz ML, Pal S, Blobel GA, Bresnick EH. 2006. Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Mol Cell Biol* **26**: 7056-7067.

Grego-Bessa J, Luna-Zurita L, del Monte G, Bolós V, Melgar P, Arandilla A, Garratt AN, Zang H, Mukouyama Y-SS, Chen H, Shou W, Ballestar E, Esteller M, Rojas A, Pérez-Pomares JMM, de la Pompa JLL. 2007. Notch signaling is essential for ventricular chamber development. *Dev Cell* **12**: 415-429.

Gritz E, Hirschi KK. 2016. Specification and function of hemogenic endothelium during embryogenesis. *Cell Mol Life Sci* **73**: 1547-1567.

Gröschel S, Sanders MA, Hoogenboezem R, de Wit E, Bouwman BAM, Erpelinck C, van der Velden VHJ, Havermans M, Avellino R, van Lom K, Rombouts EJ, van Duin M, Döhner K, Beverloo HB, Bradner JE, Döhner H, Löwenberg B, Valk PJM, Bindels EMJ, de Laat W, Delwel R. 2014. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* **157**: 369-381.

Grutz G, Forster A, Rabbitts TH. 1998. Identification of the LMO4 gene encoding an interaction partner of the LIM-binding protein LDB1/NLI1: a candidate for displacement by LMO proteins in T cell acute leukaemia. *Oncogene* **17**: 2799-2803.

Guiu J, Shimizu R, D'Altri T, Fraser ST, Hatakeyama J, Bresnick EH, Kageyama R, Dzierzak E, Yamamoto M, Espinosa L, Bigas A. 2013. Hes repressors are essential regulators of hematopoietic stem cell development downstream of Notch signaling. *J Exp Med* **210**: 71-84.

Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. 2010. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* **18**: 675-685.

Guo X, Zhang T, Hu Z, Zhang Y, Shi Z, Wang Q, Cui Y, Wang F, Zhao H, Chen Y. 2014. Efficient RNA/Cas9-mediated genome editing in Xenopus tropicalis. *Development* **141**: 707-714.

Hadland BK, Huppert SS, Kanungo J, Xue Y, Jiang R, Gridley T, Conlon RA, Cheng AM, Kopan R, Longmore GD. 2004. A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. *Blood* **104**: 3097-3105.

Hadland BK, Varnum-Finney B, Poulos MG, Moon RT, Butler JM, Rafii S, Bernstein ID. 2015. Endothelium and NOTCH specify and amplify aorta-gonad-mesonephros-derived hematopoietic stem cells. *J Clin Invest* **125**: 2032-2045.

Hayashi H, Kume T. 2008. Foxc transcription factors directly regulate Dll4 and Hey2 expression by interacting with the VEGF-Notch signaling pathways in endothelial cells. *PLoS One* **3**: e2401.

He Q, Zhang C, Wang L, Zhang P, Ma D, Lv J, Liu F. 2015. Inflammatory signaling regulates hematopoietic stem and progenitor cell emergence in vertebrates. *Blood* **125**: 1098-1106.

Herbomel P, Thisse B, Thisse C. 1999. Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development* **126**: 3735-3745.

Hewitt KJ, Kim DH, Devadas P, Prathibha R, Zuo C, Sanalkumar R, Johnson KD, Kang Y-AA, Kim J-SS, Dewey CN, Keles S, Bresnick EH. 2015. Hematopoietic Signaling Mechanism Revealed from a Stem/Progenitor Cell Cistrome. *Mol Cell* **59**: 62-74.

Hirai H, Ogawa M, Suzuki N, Yamamoto M, Breier G, Mazda O, Imanishi J, Nishikawa S-I. 2003. Hemogenic and nonhemogenic endothelium can be distinguished by the activity of fetal liver kinase (Flk)-1 promoter/enhancer during mouse embryogenesis. *Blood* **101**: 886-893.

Hoeffel G, Chen J, Lavin Y, Low D, Almeida FF, See P, Beaudin AE, Lum J, Low I, Forsberg EC, Poidinger M, Zolezzi F, Larbi A, Ng LG, Chan JK, Greter M, Becher B, Samokhvalov IM, Merad M, Ginhoux F. 2015. C-Myb(+) erythro-myeloid progenitor-derived fetal monocytes give rise to adult tissue-resident macrophages. *Immunity* **42**: 665-678.

Howe DG, Bradford YM, Conlin T, Eagle AE, Fashena D, Frazer K, Knight J, Mani P, Martin R, Moxon SA, Paddock H, Pich C, Ramachandran S, Ruef BJ, Ruzicka L, Schaper K, Shao X, Singer A, Sprunger B, Van Slyke CE, Westerfield M. 2013a. ZFIN, the Zebrafish Model Organism Database: increased support for mutants and transgenics. *Nucleic Acids Res* **41**: 60.

Howe K Clark MD Torroja CF Torrance J Berthelot C Muffato M Collins JE Humphray S McLaren K Matthews L McLaren S Sealy I Caccamo M Churcher C Scott C Barrett JC Koch R Rauch G-J White S Chow W Kilian B Quintais LT Guerra-Assunção JA Zhou Y Gu Y Yen J Vogel J-H Eyre T Redmond S Banerjee R Chi J Fu B Langley E Maguire SF Laird GK Lloyd D Kenyon E Donaldson S Sehra H Almeida-King J Loveland J Trevanion S Jones M Quail M Willey D Hunt A Burton J Sims S McLay K Plumb B Davis J Clee C Oliver K Clark R Riddle C Elliot D Eliott D Threadgold G Harden G Ware D Begum S Mortimore B Mortimer B Kerry G Heath P Phillimore B Tracey A Corby N Dunn M Johnson C Wood J Clark S Pelan S Griffiths G Smith M Glithero R Howden P Barker N Lloyd C Stevens C Harley J Holt K Panagiotidis G Lovell J Beasley H Henderson C Gordon D Auger K Wright D Collins J Raisen C Dyer L Leung K Robertson L Ambridge K Leongamornlert D McGuire S Gilderthorp R Griffiths C Manthravadi D Nichol S Barker G Whitehead S Kay M Brown J Murnane C Gray E Humphries M Sycamore N Barker D Saunders D Wallis J Babbage A Hammond S Mashreghi-Mohammadi M Barr L Martin S Wray P Ellington A Matthews N Ellwood M Woodmansey R Clark G Cooper JD Cooper J Tromans A Grafham D Skuce C Pandian R Andrews R Harrison E Kimberley A Garnett J Fosker N Hall R Garner P Kelly D Bird C Palmer S Gehring I Berger A Dooley CM Ersan-Ürün Z Eser C Geiger H Geisler M Karotki L Kirn A Konantz J Konantz M Oberländer M Rudolph-Geiger S Teucke M Lanz C Raddatz G Osoegawa K Zhu B Rapp A Widaa S Langford C Yang F Schuster SC Carter NP Harrow J Ning Z Herrero J Searle SMJ Enright A Geisler R Plasterk RHA Lee C Westerfield M de Jong PJ Zon LI Postlethwait JH Nüsslein-Volhard C Hubbard TJP Crollius H Rogers J, Stemple DL. 2013b. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**: 498-503.

Hsu AP, Johnson KD, Falcone LE, Sanalkumar R, Sanchez L, Hickstein DD, Cuellar-Rodriguez J, Lemieux JE, Zerbe CS, Bresnick EH, Holland SM. 2013. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. *Blood* **121**: 3830-3837.
Hu Y, Xie S, Yao J. 2016. Identification of Novel Reference Genes Suitable for qRT-PCR Normalization with Respect to the Zebrafish Developmental Stage. *PLoS One* **11**: e0149277.
Huang D, Sherman BT, Lempicki RA. 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44-57.
-. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**: 1-13.
Huang J, Liu X, Li D, Shao Z, Cao H, Zhang Y, Trompouki E, Bowman TV, Zon LI, Yuan G-C, Orkin SH, Xu J. 2016. Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Dev Cell* **36**: 9-23.
Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. 2011. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol* **29**: 699-700.
Huber TL, Kouskoff V, Fehling HJ, Palis J, Keller G. 2004. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* **432**: 625-630.
Hübner K, Grassme KS, Rao J, Wenke NK, Zimmer CL, Korte L, Mu Ller K, Sumanas S, Greber B, Herzog W. 2017. Wnt Signaling Positively Regulates Endothelial Cell Fate Specification in the Fli1a-Positive Progenitor Population via Lef1. *Dev Biol* **430**: 142-155.
Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**: 227-229.
Irion S, Clarke RL, Luche H, Kim I, Morrison SJ, Fehling H-JJ, Keller GM. 2010. Temporal specification of blood progenitors from mouse embryonic stem cells and induced pluripotent stem cells. *Development* **137**: 2829-2839.
Istrail S, Davidson EH. 2005. Logic functions of the genomic cis-regulatory code. *Proc Natl Acad Sci U S A* **102**: 4954-4959.
Jaffredo T, Gautier R, Eichmann A, Dieterlen-Lièvre F. 1998. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125**: 4575-4583.
Jang IH, Lu Y-FF, Zhao L, Wenzel PL, Kume T, Datta SM, Arora N, Guiu J, Lagha M, Kim PG, Do EK, Kim JH, Schlaeger TM, Zon LI, Bigas A, Burns CE, Daley GQ. 2015. Notch1 acts via Foxc2 to promote definitive hematopoiesis via effects on hemogenic endothelium. *Blood* **125**: 1418-1426.
Jin S-W, Beis D, Mitchell T, Chen J-N, Stainier DYR. 2005. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**: 5199-5209.
Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816-821.
Johnson KD, Hsu AP, Ryu M-JJ, Wang J, Gao X, Boyer ME, Liu Y, Lee Y, Calvo KR, Keles S, Zhang J, Holland SM, Bresnick EH. 2012. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J Clin Invest* **122**: 3692-3704.
Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, Birney E, Searle S, Schmutz J, Grimwood J, Dickson MC, Myers RM, Miller CT, Summers BR, Knecht AK, Brady SD, Zhang H, Pollen AA, Howes T, Amemiya C, Team B, Baldwin J, Bloom T, Jaffe DB, Nicol R, Wilkinson J, Lander ES, Di Palma F, Lindblad-Toh K, Kingsley DM. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**: 55-61.
Jowett T, Yan YL. 1996. Double fluorescent in situ hybridization to zebrafish embryos. *Trends Genet* **12**: 387-389.
Kaimakis P, de Pater E, Eich C, Kartalaei P, Kauts M-L, Vink CS, van der Linden R, Jaegle M, Yokomizo T, Meijer D, Dzierzak E. 2016. Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. *Blood* **127**: 1426-1437.
Kalev-Zylinska ML, Horsfield JA, Flores MV, Postlethwait JH, Vitas MR, Baas AM, Crosier PS, Crosier KE. 2002. Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development* **129**: 2015-2030.
Kallianpur AR, Jordan JE, Brandt SJ. 1994. The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. *Blood* **83**: 1200-1208.

Kane DA, Kimmel CB. 1993. The zebrafish midblastula transition. *Development* **119**: 447-456.
Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AMMM, Pletikos M, Meyer KA, Sedmak G, Guennel T, Shin Y, Johnson MB, Krsnik Z, Mayer S, Fertuzinhos S, Umlauf S, Lisgo SN, Vortmeyer A, Weinberger DR, Mane S, Hyde TM, Huttner A, Reimers M, Kleinman JE, Sestan N. 2011. Spatio-temporal transcriptome of the human brain. *Nature* **478**: 483-489.
Kanz D, Konantz M, Alghisi E, North TE, Lengerke C. 2016. Endothelial-to-hematopoietic transition: Notch-ing vessels into blood. *Ann N Y Acad Sci* **1370**: 97-108.
Kartalaei PS, Yamada-Inagawa T, Vink CS, de Pater E, van der Linden R, Marks-Bluth J, van der Sloot A, van den Hout M, Yokomizo T, van Schaick-Solernó ML, Delwel R, Pimanda JE, van Ijcken WF, Dzierzak E. 2015. Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *J Exp Med* **212**: 93-106.
Kassouf MT, Hughes JR, Taylor S, McGowan SJ, Soneji S, Green AL, Vyas P, Porcher C. 2010. Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. *Genome Res* **20**: 1064-1083.
Kataoka H, Hayashi M, Nakagawa R, Tanaka Y, Izumi N, Nishikawa S, Jakt ML, Tarui H, Nishikawa S-I. 2011. Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRα+ primitive mesoderm. *Blood* **118**: 6975-6986.
Kawakami K, Takeda H, Kawakami N, Kobayashi M, Matsuda N, Mishina M. 2004. A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev Cell* **7**: 133-144.
Kelley C, Yee K, Harland R, Zon LI. 1994. Ventral expression of GATA-1 and GATA-2 in the Xenopus embryo defines induction of hematopoietic mesoderm. *Dev Biol* **165**: 193-205.
Kennedy M, D'Souza SL, Lynch-Kattman M, Schwantz S, Keller G. 2007. Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. *Blood* **109**: 2679-2687.
Kent JW, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, and Haussler D. 2002. The Human Genome Browser at UCSC. *Genome Res* **12**: 996-1006.
Khandekar M, Brandt W, Zhou Y, Dagenais S, Glover TW, Suzuki N, Shimizu R, Yamamoto M, Lim KC, Engel JD. 2007. A Gata2 intronic enhancer confers its pan-endothelia-specific regulation. *Development* **134**: 1703-1712.
Kim AD, Melick CH, Clements WK, Stachura DL, Distel M, Panakova D, MacRae C, Mork LA, Crump JG, Traver D. 2014. Discrete Notch signaling requirements in the specification of hematopoietic stem cells. *EMBO J* **33**: 2363-2373.
Kim PG, Albacker CE, Lu Y-fF, Jang I-hH, Lim Y, Heffner GC, Arora N, Bowman TV, Lin MI, Lensch MW, De Los Angeles A, Zon LI, Loewer S, Daley GQ. 2013. Signaling axis involving Hedgehog, Notch, and Scl promotes the embryonic endothelial-to-hematopoietic transition. *Proc Natl Acad Sci U S A* **110**: E141-E150.
Kim PG, Canver MC, Rhee C, Ross SJ, Harriss JV, Tu CH, Orkin SH, Tucker HO, Daley GQ. 2016. Interferon-α signaling promotes embryonic HSC maturation. *Blood* **128**: 204-216.
Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**: 253-310.
Kirmizitas A, Meiklejohn S, Ciau-Uitz A, Stephenson R, Patient R. 2017. Dissecting BMP signaling input into the gene regulatory networks driving specification of the blood stem cell lineage. *Proc Natl Acad Sci U S A* **114**: 5814-5821.
Kissa K, Herbomel P. 2010. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* **464**: 112-115.
Kissa K, Murayama E, Zapata A, Cortes A, Perret E, Machu C, Herbomel P. 2008. Live imaging of emerging hematopoietic stem cells and early thymus colonization. *Blood* **111**: 1147-1156.
Kitajima K, Nakajima M, Kanokoda M, Kyba M, Dandapat A, Tolar J, Saito MK, Toyoda M, Umezawa A, Hara T. 2016. GSK3β inhibition activates the CDX/HOX pathway and promotes hemogenic endothelial progenitor differentiation from human pluripotent stem cells. *Exp Hematol* **44**: 68-74.
Kobayashi I, Kobayashi-Sun J, Kim AD, Pouget C, Fujita N, Suda T, Traver D. 2014. Jam1a-Jam2a interactions regulate haematopoietic stem cell fate through Notch signalling. *Nature* **512**: 319-323.

Kok FO, Shin M, Ni CW, Gupta A, Grosse AS, van Impel A, Kirchmaier BC, Peterson-Maduro J, Kourkoulis G, Male I, DeSantis DF, Sheppard-Tindell S, Ebarasi L, Betsholtz C, Schulte-Merker S, Wolfe SA, Lawson ND. 2015. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev Cell* **32**: 97-108.

Krebs LT, Shutter JR, Tanigaki K, Honjo T, Stark KL, Gridley T. 2004. Haploinsufficient lethality and formation of arteriovenous malformations in Notch pathway mutants. *Genes Dev* **18**: 2469-2473.

Krebs LT, Xue Y, Norton CR, Shutter JR, Maguire M, Sundberg JP, Gallahan D, Closson V, Kitajewski J, Callahan R, Smith GH, Stark KL, Gridley T. 2000. Notch signaling is essential for vascular morphogenesis in mice. *Genes Dev* **14**: 1343-1352.

Kumano K, Chiba S, Kunisato A, Sata M, Saito T, Nakagami-Yamaguchi E, Yamaguchi T, Masuda S, Shimizu K, Takahashi T, Ogawa S, Hamada Y, Hirai H. 2003. Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. *Immunity* **18**: 699-711.

Kumaravelu P, Hook L, Morrison AM, Ure J, Zhao S, Zuyev S, Ansell J, Medvinsky A. 2002. Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development* **129**: 4891-4899.

Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, Dickson BJ, Stark A. 2014. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature* **512**: 91-95.

Labastie MC, Cortés F, Roméo PH, Dulac C, Péault B. 1998. Molecular identity of hematopoietic precursor cells emerging in the human embryo. *Blood* **92**: 3624-3635.

Lacaud G, Kouskoff V. 2017. Hemangioblast, hemogenic endothelium, and primitive versus definitive hematopoiesis. *Exp Hematol* **49**: 19-24.

Lam EY, Hall CJ, Crosier PS, Crosier KE, Flores MV. 2010. Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. *Blood* **116**: 909-914.

Lancrin C, Mazan M, Stefanska M, Patel R, Lichtinger M, Costa G, Vargel Ö, Wilson NK, Möröy T, Bonifer C, Göttgens B, Kouskoff V, Lacaud G. 2012. GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood* **120**: 314-322.

Lancrin C, Sroczynska P, Stephenson C, Allen T, Kouskoff V, Lacaud G. 2009. The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* **457**: 892-895.

Landry J-RR, Kinston S, Knezevic K, de Bruijn MF, Wilson N, Nottingham WT, Peitz M, Edenhofer F, Pimanda JE, Ottersbach K, Göttgens B. 2008. Runx genes are direct targets of Scl/Tal1 in the yolk sac and fetal liver. *Blood* **111**: 3005-3014.

Lane M, Runko AP, Roy NM, Sagerström CG. 2002. Dynamic expression and regulation by Fgf8 and Pou2 of the zebrafish LIM-only gene, lmo4. *Mech Dev* **119**: S185–S189.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Lawson ND, Scheer N, Pham VN, Kim CH, Chitnis AB, Campos-Ortega JA, Weinstein BM. 2001. Notch signaling is required for arterial-venous differentiation during embryonic vascular development. *Development* **128**: 3675-3683.

Lawson ND, Vogel AM, Weinstein BM. 2002. sonic hedgehog and vascular endothelial growth factor act upstream of the Notch pathway during arterial endothelial differentiation. *Dev Cell* **3**: 127-136.

Lawson ND, Weinstein BM. 2002. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol* **248**: 307-318.

Ledford H. 2017. CRISPR studies muddy results of older gene research. *Nature*.

Lee D, Park C, Lee H, Lugus JJ, Kim SH, Arentson E, Chung YS, Gomez G, Kyba M, Lin S, Janknecht R, Lim D-SS, Choi K. 2008. ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell stem cell* **2**: 497-507.

Lee Y, Manegold JE, Kim AD, Pouget C, Stachura DL, Clements WK, Traver D. 2014. FGF signalling specifies haematopoietic stem cells through its regulation of somitic Notch signalling. *Nat Commun* **5**: 5583.

Lentjes MH, Niessen HE, Akiyama Y, de Bruïne AP, Melotte V, van Engeland M. 2016. The emerging role of GATA transcription factors in development and disease. *Expert Rev Mol Med* **18**: e3.

Leung A, Ciau-Uitz A, Pinheiro P, Monteiro R, Zuo J, Vyas P, Patient R, Porcher C. 2013. Uncoupling VEGFA functions in arteriogenesis and hematopoietic stem cell specification. *Dev Cell* **24**: 144-158.

Levine M. 2010. Transcriptional enhancers in animal development and evolution. *Curr Biol* **20**: R754-R763.

Li C, Lan Y, Schwartz-Orbach L, Korol E, Tahiliani M, Evans T, Goll MG. 2015. Overlapping Requirements for Tet2 and Tet3 in Normal Development and Hematopoietic Stem Cell Emergence. *Cell Rep* **12**: 1133-1143.

Li G, Zhao Y, Wang J, Liu B, Sun X, Guo S, Feng J. 2017. Transcriptome profiling of developing spleen tissue and discovery of immune-related genes in grass carp (Ctenopharyngodon idella). *Fish Shellfish Immunol* **60**: 400-410.

Li L, Jothi R, Cui K, Lee JY, Cohen T, Gorivodsky M, Tzchori I, Zhao Y, Hayes SM, Bresnick EH, Zhao K, Westphal H, Love PE. 2011. Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nat Immunol* **12**: 129-136.

Li Q, Uitto J. 2014. Zebrafish as a model system to study skin biology and pathology. *J Invest Dermatol* **134**: 1-6.

Li Y, Esain V, Teng L, Xu J, Kwan W, Frost IM, Yzaguirre AD, Cai X, Cortes M, Maijenburg MW, Tober J, Dzierzak E, Orkin SH, Tan K, North TE, Speck NA. 2014a. Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev* **28**: 2597-2612.

Li Y, Rivera CM, Ishii H, Jin F, Selvaraj S, Lee AY, Dixon JR, Ren B. 2014b. CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. *PLoS One* **9**: e114485.

Liakhovitskaia A, Rybtsov S, Smith T, Batsivari A, Rybtsova N, Rode C, de Bruijn M, Buchholz F, Gordon-Keylock S, Zhao S, Medvinsky A. 2014. Runx1 is required for progression of CD41+ embryonic precursors into HSCs but not prior to this. *Development* **141**: 3319-3323.

Liao EC, Paw BH, Oates AC, Pratt SJ, Postlethwait JH, Zon LI. 1998. SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish. *Genes Dev* **12**: 621-626.

Lie-A-Ling M, Marinopoulou E, Li Y, Patel R, Stefanska M, Bonifer C, Miller C, Kouskoff V, Lacaud G. 2014. RUNX1 positively regulates a cell adhesion and migration program in murine hemogenic endothelium prior to blood emergence. *Blood* **124**: e11-e20.

Lim KC, Hosoya T, Brandt W, Ku CJ, Hosoya-Ohmura S, Camper SA, Yamamoto M, Engel JD. 2012. Conditional Gata2 inactivation results in HSC loss and lymphatic mispatterning. *J Clin Invest* **122**: 3705-3717.

Limbourg FP, Takeshita K, Radtke F, Bronson RT, Chin MT, Liao JK. 2005. Essential role of endothelial Notch1 in angiogenesis. *Circulation* **111**: 1826-1832.

Ling K-WW, Ottersbach K, van Hamburg JP, Oziemlak A, Tsai F-YY, Orkin SH, Ploemacher R, Hendriks RW, Dzierzak E. 2004. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med* **200**: 871-882.

Lis R, Karrasch CC, Poulos MG, Kunar B, Redmond D, Duran JG, Badwe CR, Schachterle W, Ginsberg M, Xiang J, Tabrizi A, Shido K, Rosenwaks Z, Elemento O, Speck NA, Butler JM, Scandura JM, Rafii S. 2017. Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature* **545**: 439-445.

Liu F, Li D, Yu Y, Kang I, Cha MJ, Kim J, Park C, Watson DK, Wang T, Choi K. 2015. Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. *EMBO reports* **16**: 654-669.

Liu F, Patient R. 2008. Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. *Circ Res* **103**: 1147-1154.

Liu F, Walmsley M, Rodaway A, Patient R. 2008. Fli1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol* **18**: 1234-1240.

Liu J, Jiang J, Wang Z, He Y, Zhang Q. 2016a. Origin and evolution of GATA2a and GATA2b in teleosts: insights from tongue sole, Cynoglossus semilaevis. *PeerJ* **4**: e1790.

Liu Y, Huang D, Wang Z, Wu C, Zhang Z, Wang D, Li Z, Zhu T, Yang S, Sun W. 2016b. LMO2 attenuates tumor growth by targeting the Wnt signaling pathway in breast and colorectal cancer. *Sci Rep* **6**: 36050.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods* **25**: 402-408.

Lizama CO, Hawkins JS, Schmitt CE, Bos FL, Zape JP, Cautivo KM, Pinto H, Rhyner AM, Yu H, Donohoe ME, Wythe JD, Zovein AC. 2015. Repression of arterial genes in hemogenic endothelium is sufficient for haematopoietic fate acquisition. *Nat Commun* **6**: 7739.

Logan CY, Nusse R. 2004. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol* **20**: 781-810.

Loh KM, van Amerongen R, Nusse R. 2016. Generating Cellular Diversity and Spatial Form: Wnt Signaling and the Evolution of Multicellular Animals. *Dev Cell* **38**: 643-655.

Lopes R, Korkmaz G, Agami R. 2016. Applying CRISPR-Cas9 tools to identify and characterize transcriptional enhancers. *Nat Rev Mol Cell Biol* **17**: 597-604.

Loughran SJ, Kruse EA, Hacking DF, de Graaf CA, Hyland CD, Willson TA, Henley KJ, Ellis S, Voss AK, Metcalf D, Hilton DJ, Alexander WS, Kile BT. 2008. The transcription factor Erg is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells. *Nat Immunol* **9**: 810-819.

Lugus JJ, Chung YS, Mills JC, Kim SI, Grass J, Kyba M, Doherty JM, Bresnick EH, Choi K. 2007. GATA2 functions at multiple steps in hemangioblast development and differentiation. *Development* **134**: 393-405.

Lugus JJ, Park C, Ma YD, Choi K. 2009. Both primitive and definitive blood cells are derived from Flk-1+ mesoderm. *Blood* **113**: 563-566.

Luis TC, Naber BA, Fibbe WE, van Dongen JJ, Staal FJ. 2010. Wnt3a nonredundantly controls hematopoietic stem cell function and its deficiency results in complete absence of canonical Wnt signaling. *Blood* **116**: 496-497.

Lurie LJ, Boyer ME, Grass JA, Bresnick EH. 2008. Differential GATA factor stabilities: implications for chromatin occupancy by structurally similar transcription factors. *Biochemistry* **47**: 859-869.

Lyons SE, Lawson ND, Lei L, Bennett PE, Weinstein BM, Liu PP. 2002. A nonsense mutation in zebrafish gata1 causes the bloodless phenotype in vlad tepes. *Proc Natl Acad Sci U S A* **99**: 5454-5459.

Ma D, Wang L, Wang S, Gao Y, Wei Y, Liu F. 2012. Foxn1 maintains thymic epithelial cells to support T-cell development via mcm2 in zebrafish. *Proceedings of the National Academy of Sciences* **109**: 21040-21045.

Maeno M, Mead PE, Kelley C, Xu R-H, Kung H-F, Suzuki A, Ueno N, Zon LI. 1996. The role of BMP-4 and GATA-2 in the induction and differentiation of hematopoietic mesoderm in Xenopus laevis. *Blood* **88**: 1965-1972.

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM. 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**: 833-838.

Mandal L, Banerjee U, Hartenstein V. 2004. Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. *Nat Genet* **36**: 1019-1023.

Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, Etchin J, Lawton L, Sallan SE, Silverman LB, Loh ML, Hunger SP, Sanda T, Young RA, Look AT. 2014. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **346**: 1373-1377.

Maquat LE. 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol* **5**: 89-99.

Marcelo KL, Sills TM, Coskun S, Vasavada H, Sanglikar S, Goldie LC, Hirschi KK. 2013. Hemogenic endothelial cell specification requires c-Kit, Notch signaling, and p27-mediated cell-cycle control. *Dev Cell* **27**: 504-515.

Marshall CJ, Kinnon C, Thrasher AJ. 2000. Polarized expression of bone morphogenetic protein-4 in the human aorta-gonad-mesonephros region. *Blood* **96**: 1591-1593.

McCollum CW, Amin SR, Pauerstein P, Lane M. 2007. A zebrafish LMO4 ortholog limits the size of the forebrain and eyes through negative regulation of six3b and rx3. *Dev Biol* **309**: 373-385.

McGrath KE, Frame JM, Fegan KH, Bowen JR, Conway SJ, Catherman SC, Kingsley PD, Koniski AD, Palis J. 2015. Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Rep* **11**: 1892-1904.

Mead PE, Deconinck AE, Huber TL, Orkin SH, Zon LI. 2001. Primitive erythropoiesis in the Xenopus embryo: the synergistic role of LMO-2, SCL and GATA-binding proteins. *Development* **128**: 2301-2308.

Medvinsky A, Dzierzak E. 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* **86**: 897-906.

Medvinsky A, Rybtsov S, Taoudi S. 2011. Embryonic origin of the adult hematopoietic system: advances and questions. *Development* **138**: 1017-1031.

Meier N, Krpic S, Rodriguez P, Strouboulis J, Monti M, Krijgsveld J, Gering M, Patient R, Hostert A, Grosveld F. 2006. Novel binding partners of Ldb1 are required for haematopoietic development. *Development* **133**: 4913-4923.

Meijer L, Skaltsounis A-L, Magiatis P, Polychronopoulos P, Knockaert M, Leost M, Ryan XP, Vonica C, Brivanlou A, Dajani R, Crovace C, Tarricone C, Musacchio A, Roe MS, Pearl L, Greengard P. 2003. GSK-3-selective inhibitors derived from Tyrian purple indirubins. *Chem Biol* **10**: 1255-1266.

Mellor J. 2005. The dynamics of chromatin remodeling at promoters. *Mol Cell* **19**: 147-157.

Mendelson A, Frenette PS. 2014. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. *Nat Med* **20**: 833-846.

Meng A, Tang H, Ong BA, Farrell MJ, Lin S. 1997. Promoter analysis in living zebrafish embryos identifies a cis-acting motif required for neuronal expression of GATA-2. *Proc Natl Acad Sci U S A* **94**: 6267-6272.

Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. 2008. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol* **26**: 695-701.

Mikkola HK, Fujiwara Y, Schlaeger TM, Traver D, Orkin SH. 2003. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* **101**: 508-516.

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ. 2011. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**: 143-148.

Minegishi N, Ohta J, Yamagiwa H, Suzuki N, Kawauchi S, Zhou Y, Takahashi S, Hayashi N, Engel JD, Yamamoto M. 1999. The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood* **93**: 4196-4207.

Montañez-Wiscovich ME, Seachrist DD, Landis MD, Visvader J, Andersen B, Keri RA. 2009. LMO4 is an essential mediator of ErbB2/HER2/Neu-induced breast cancer cell cycle progression. *Oncogene* **28**: 3608-3618.

Montañez-Wiscovich ME, Shelton MD, Seachrist DD, Lozada KL, Johnson E, Miedler JD, Abdul-Karim FW, Visvader JE, Keri RA. 2010. Aberrant expression of LMO4 induces centrosome amplification and mitotic spindle abnormalities in breast cancer cells. *J Pathol* **222**: 271-281.

Monteiro R, Pinheiro P, Joseph N, Peterkin T, Koth J, Repapi E, Bonkhofer F, Kirmizitas A, Patient R. 2016. Transforming Growth Factor β Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells. *Dev Cell* **38**: 1-13.

Monteiro R, Pouget C, Patient R. 2011. The gata1/pu.1 lineage fate paradigm varies between blood populations and is modulated by tif1γ. *EMBO J* **30**: 1093-1103.

Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, Figueroa ME, Vasanthakumar A, Patel J, Zhao X, Perna F, Pandey S, Madzo J, Song C, Dai Q, He C, Ibrahim S, Beran M, Zavadil J, Nimer SD, Melnick A, Godley LA, Aifantis I, Levine RL. 2011. Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation. *Cancer Cell* **20**: 11-24.

Morton J, Davis MW, Jorgensen EM, Carroll D. 2006. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. *Proc Natl Acad Sci U S A* **103**: 16370-16375.

Mosimann C, Panáková D, Werdich AA, Musso G, Burger A, Lawson KL, Carr LA, Nevis KR, Sabeh MK, Zhou Y, Davidson AJ, DiBiase A, Burns CE, Burns CG, MacRae CA, Zon LI. 2015. Chamber identity programs drive early functional partitioning of the heart. *Nat Commun* **6**: 8146.

Mucenski ML, McLain K, Kier AB, Swerdlow SH, Schreiner CM, Miller TA, Pietryga DW, Scott WJ, Potter SS. 1991. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65**: 677-689.

Mudunuri U, Che A, Yi M, Stephens RM. 2009. bioDBnet: the biological database network. *Bioinformatics* **25**: 555-556.

Murayama E, Kissa K, Zapata A, Mordelet E, Briolat V, Lin H-F, Handin RI, Herbomel P. 2006. Tracing Hematopoietic Precursor Migration to Successive Hematopoietic Organs during Zebrafish Development. *Immunity* **25**: 963-975.

Nakagawa M, Ichikawa M, Kumano K, Goyama S, Kawazu M, Asai T, Ogawa S, Kurokawa M, Chiba S. 2006. AML1/Runx1 rescues Notch1-null mutation-induced deficiency of para-aortic splanchnopleural hematopoiesis. *Blood* **108**: 3329-3334.

Nardelli J, Thiesson D, Fujiwara Y, Tsai F-Y, Orkin SH. 1999. Expression and Genetic Interaction of Transcription Factors GATA-2 and GATA-3 during Development of the Mouse Central Nervous System. *Dev Biol* **210**: 305-321.

Nasevicius A, Ekker SC. 2000. Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* **26**: 216-220.

Neff KL, Argue DP, Ma AC, Lee HB, Clark KJ, Ekker SC. 2012. Mojo Hand, a TALEN design tool for genome editing applications. *BMC Bioinformatics* **14**: 1-7.

Ng CE, Yokomizo T, Yamashita N, Cirovic B, Jin H, Wen Z, Ito Y, Osato M. 2010. A Runx1 intronic enhancer marks hemogenic endothelial cells and hematopoietic stem cells. *Stem Cells* **28**: 1869-1881.

Nguyen PD, Hollway GE, Sonntag C, Miles LB, Hall TE, Berger S, Fernandez KJ, Gurevich DB, Cole NJ, Alaei S, Ramialison M, Sutherland RL, Polo JM, Lieschke GJ, Currie PD. 2014. Haematopoietic stem cell induction by somite-derived endothelial cells controlled by meox1. *Nature* **512**: 314-318.

Nishikawa SI, Nishikawa S, Hirashima M, Matsuyoshi N, Kodama H. 1998. Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. *Development* **125**: 1747-1757.

Nord AS, Blow MJ, Attanasio C, Akiyama JA, Holt A, Hosseini R, Phouanenavong S, Plajzer-Frick I, Shoukry M, Afzal V, Rubenstein JL, Rubin EM, Pennacchio LA, Visel A. 2013. Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**: 1521-1531.

North T, Gu TL, Stacy T, Wang Q, Howard L, Binder M, Marín-Padilla M, Speck NA. 1999. Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126**: 2563-2575.

North TE, de Bruijn MF, Stacy T, Talebian L, Lind E, Robin C, Binder M, Dzierzak E, Speck NA. 2002. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16**: 661-672.

Nottingham WT, Jarratt A, Burgess M, Speck CL, Cheng JF, Prabhakar S, Rubin EM, Li PS, Sloane-Stanley J, Kong ASJ, de Bruijn MF. 2007. Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. *Blood* **110**: 4188-4197.

Okuda T, van Deursen J, Hiebert SW, Grosveld G, Downing JR. 1996. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84**: 321-330.

Olsson A-KK, Dimberg A, Kreuger J, Claesson-Welsh L. 2006. VEGF receptor signalling - in control of vascular function. *Nat Rev Mol Cell Biol* **7**: 359-371.

Onodera K, Fujiwara T, Onishi Y, Itoh-Nakadai A, Okitsu Y, Fukuhara N, Ishizawa K, Shimizu R, Yamamoto M, Harigae H. 2016. GATA2 regulates dendritic cell differentiation. *Blood* **128**: 508-518.

Org T, Duan D, Ferrari R, Montel-Hagen A, Van Handel B, Kerenyi MA, Sasidharan R, Rubbi L, Fujiwara Y, Pellegrini M, Orkin SH, Kurdistani SK, Mikkola HK. 2015. Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. *EMBO J* **34**: 759-777.

Orkin SH, Zon LI. 2008. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell* **132**: 631-644.

Padron-Barthe L, Temino S, Villa del Campo C, Carramolino L, Isern J, Torres M. 2014. Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo. *Blood* **124**: 2523-2532.

Palis J. 2008. Ontogeny of erythropoiesis. *Curr Opin Hematol* **15**: 155-161.
Palis J, Robertson S, Kennedy M, Wall C, Keller G. 1999. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126**: 5073-5084.
Pardo-Martin C, Chang T-YY, Koo BK, Gilleland CL, Wasserman SC, Yanik MF. 2010. High-throughput in vivo vertebrate screening. *Nature methods* **7**: 634-636.
Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE. 2009. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev Dyn* **238**: 2975-3015.
Patterson LJ, Gering M, Eckfeldt CE, Green AR, Verfaillie CM, Ekker SC, Patient R. 2007. The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish. *Blood* **109**: 2389-2398.
Patterson LJ, Gering M, Patient R. 2005. Scl is required for dorsal aorta as well as blood formation in zebrafish embryos. *Blood* **105**: 3502-3511.
Pattison RJ, Csukasi F, Zheng Y, Fei Z, van der Knaap E, Catalá C. 2015. Comprehensive Tissue-Specific Transcriptome Analysis Reveals Distinct Regulatory Programs during Early Tomato Fruit Development. *Plant Physiol* **168**: 1684-1701.
Peeters M, Ottersbach K, Bollerot K, Orelio C, de Bruijn M, Wijgerde M, Dzierzak E. 2009. Ventral embryonic tissues and Hedgehog proteins induce early AGM hematopoietic stem cell development. *Development* **136**: 2613-2621.
Pelster B, Burggren WW. 1996. Disruption of hemoglobin oxygen transport does not impact oxygen-dependent physiological processes in developing embryos of zebra fish (Danio rerio). *Circ Res* **79**: 358-362.
Pereira C-FF, Chang B, Qiu J, Niu X, Papatsenko D, Hendry CE, Clark NR, Nomura-Kitabayashi A, Kovacic JC, Ma'ayan A, Schaniel C, Lemischka IR, Moore K. 2013. Induction of a hemogenic program in mouse fibroblasts. *Cell stem cell* **13**: 205-218.
Perlin JR, Robertson AL, Zon LI. 2017. Efforts to enhance blood stem cell engraftment: Recent insights from zebrafish hematopoiesis. *J Exp Med* **214**: 2817-2827.
Persons DA, Allay JA, Allay ER, Ashmun RA, Orlic D, Jane SM, Cunningham JM, Nienhuis AW. 1999. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood* **93**: 488-499.
Pevny L, Lin CS, D'Agati V, Simon MC, Orkin SH, Costantini F. 1995. Development of hematopoietic cells lacking transcription factor GATA-1. *Development* **121**: 163-172.
Pevny L, Simon MC, Robertson E, Klein WH, Tsai SF, D'Agati V, Orkin SH, Costantini F. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**: 257-260.
Pimanda JE, Donaldson IJ, de Bruijn MF, Kinston S, Knezevic K, Huckle L, Piltz S, Landry J-RR, Green AR, Tannahill D, Göttgens B. 2007a. The SCL transcriptional network and BMP signaling pathway interact to regulate RUNX1 activity. *Proc Natl Acad Sci U S A* **104**: 840-845.
Pimanda JE, Ottersbach K, Knezevic K, Kinston S, Chan WY, Wilson NK, Landry JR, Wood AD, Kolb-Kokocinski A, Green AR, Tannahill D, Lacaud G, Kouskoff V, Gottgens B. 2007b. Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc Natl Acad Sci U S A* **104**: 17692-17697.
Place ES, Smith JC. 2017. Zebrafish atoh8 mutants do not recapitulate morpholino phenotypes. *PLoS One* **12**: e0171143.
Poole TJ, Coffin JD. 1988. Developmental angiogenesis: quail embryonic vasculature. *Scanning Microsc* **2**: 443-448.
-. 1989. Vasculogenesis and angiogenesis: two distinct morphogenetic mechanisms establish embryonic vascular pattern. *J Exp Zool* **251**: 224-231.
Porcher C, Chagraoui H, Kristiansen MS. 2017. SCL/TAL1: a multifaceted regulator from blood development to disease. *Blood* **129**: 2051-2060.
Porcher C, Swat W, Rockwell K, Fujiwara Y, Alt FW, Orkin SH. 1996. The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* **86**: 47-57.
Pouget C, Peterkin T, Simoes FC, Lee Y, Traver D, Patient R. 2014. FGF signalling restricts haematopoietic stem cell specification via modulation of the BMP pathway. *Nat Commun* **5**: 5588.

Qian F, Zhen F, Xu J, Huang M, Li W, Wen Z. 2007. Distinct Functions for Different scl Isoforms in Zebrafish Primitive and Definitive Hematopoiesis. *PLoS Biol* **5**: e132.

Quillien A, Abdalla M, Yu J, Ou J, Zhu LJ, Lawson ND. 2017. Robust Identification of Developmentally Active Endothelial Enhancers in Zebrafish Using FANS-Assisted ATAC-Seq. *Cell Rep* **20**: 709-720.

Quinkertz A, Campos-Ortega JA. 1999. A new beta-globin gene from the zebrafish, betaE1, and its pattern of transcription during embryogenesis. *Dev Genes Evol* **209**: 126-131.

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**: 279-283.

Rambaldi D, Ciccarelli FD. 2009. FancyGene: dynamic visualization of gene structures and protein domain architectures on genomic loci. *Bioinformatics* **25**: 2281-2282.

Ran D, Shia W-JJ, Lo M-CC, Fan J-BB, Knorr DA, Ferrell PI, Ye Z, Yan M, Cheng L, Kaufman DS, Zhang D-EE. 2013a. RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. *Blood* **121**: 2882-2890.

Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F. 2013b. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**: 1380-1389.

Ransom DG, Haffter P, Odenthal J, Brownlie A, Vogelsang E, Kelsh RN, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Mullins MC, Nüsslein-Volhard C. 1996. Characterization of zebrafish mutants with defects in embryonic hematopoiesis. *Development* **123**: 311-319.

Real PJ, Ligero G, Ayllon V, Ramos-Mejia V, Bueno C, Gutierrez-Aranda I, Navarro-Montero O, Lako M, Menendez P. 2012. SCL/TAL1 regulates hematopoietic specification from human embryonic stem cells. *Mol Ther* **20**: 1443-1453.

Reischauer S, Stone OA, Villasenor A, Chi N, Jin S-WW, Martin M, Lee MT, Fukuda N, Marass M, Witty A, Fiddes I, Kuo T, Chung W-SS, Salek S, Lerrigo R, Alsiö J, Luo S, Tworus D, Augustine SM, Mucenieks S, Nystedt B, Giraldez AJ, Schroth GP, Andersson O, Stainier DY. 2016. Cloche is a bHLH-PAS transcription factor that drives haemato-vascular specification. *Nature* **535**: 294-298.

Ren X, Gomez GA, Zhang B, Lin S. 2010. Scl isoforms act downstream of etsrp to specify angioblasts and definitive hematopoietic stem cells. *Blood* **115**: 5338-5346.

Rhodes J, Hagen A, Hsu K, Deng M, Liu TX, Look AT, Kanki JP. 2005. Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev Cell* **8**: 97-108.

Richard C, Drevon C, Canto PY, Villain G, Bollerot K, Lempereur A, Teillet MA, Vincent C, Rossello Castillo C, Torres M, Piwarzyk E, Speck NA, Souyri M, Jaffredo T. 2013. Endothelio-mesenchymal interaction controls runx1 expression and modulates the notch pathway to initiate aortic hematopoiesis. *Dev Cell* **24**: 600-611.

Richter J, Traver D, Willert K. 2017. The role of Wnt signaling in hematopoietic stem cell development. *Crit Rev Biochem Mol Biol* **52**: 414-424.

Robb L, Elwood NJ, Elefanty AG, Köntgen F, Li R, Barnett LD, Begley CG. 1996. The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse. *Embo J* **15**: 4123-4129.

Robert-Moreno A, Espinosa L, de la Pompa JLL, Bigas A. 2005. RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development* **132**: 1117-1126.

Robert-Moreno A, Guiu J, Ruiz-Herguido C, López ME, Inglés-Esteve J, Riera L, Tipping A, Enver T, Dzierzak E, Gridley T, Espinosa L, Bigas A. 2008. Impaired embryonic haematopoiesis yet normal arterial development in the absence of the Notch ligand Jagged1. *Embo J* **27**: 1886-1895.

Robertson AL, Avagyan S, Gansner JM, Zon LI. 2016. Understanding the regulation of vertebrate hematopoiesis and blood disorders – big lessons from a small fish. *FEBS Lett* **590**: 4016-4033.

Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, Stainier DYR. 2015. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**: 230-233.

Rowlinson JM, Gering M. 2010. Hey2 acts upstream of Notch in hematopoietic stem cell specification in zebrafish embryos. *Blood* **116**: 2046-2056.

Ruiz-Herguido C, Guiu J, D'Altri T, Inglés-Esteve J, Dzierzak E, Espinosa L, Bigas A. 2012. Hematopoietic stem cell development requires transient Wnt/β-catenin activity. *J Exp Med* **209**: 1457-1468.

Rybtsov S, Ivanovs A, Zhao S, Medvinsky A. 2016. Concealed expansion of immature precursors underpins acute burst of adult HSC activity in foetal liver. *Development* **143**: 1284-1289.

Salanga MC, Meadows SM, Myers CT, Krieg PA. 2010. ETS family protein ETV2 is required for initiation of the endothelial lineage but not the hematopoietic lineage in the Xenopus embryo. *Dev Dyn* **239**: 1178-1187.

Sanalkumar R, Johnson KD, Gao X, Boyer ME, Chang YI, Hewitt KJ, Zhang J, Bresnick EH. 2014. Mechanism governing a stem cell-generating cis-regulatory element. *Proc Natl Acad Sci U S A* **111**: E1091-1100.

Sánchez M, Göttgens B, Sinclair AM, Stanley M, Begley CG, Hunter S, Green AR. 1999. An SCL 3' enhancer targets developing endothelium together with embryonic and adult haematopoietic progenitors. *Development* **126**: 3891-3904.

Sang M, Ma L, Sang M, Zhou X, Gao W, Geng C. 2014. LIM-domain-only proteins: multifunctional nuclear transcription coregulators that interacts with diverse proteins. *Mol Biol Rep* **41**: 1067-1073.

Santoriello C, Zon LI. 2012. Hooked! Modeling human disease in zebrafish. *J Clin Invest* **122**: 2337-2343.

Sawamiphak S, Kontarakis Z, Stainier DY. 2014. Interferon gamma signaling positively regulates hematopoietic stem cell emergence. *Dev Cell* **31**: 640-653.

Schier AF, Talbot WS. 2005. Molecular genetics of axis formation in zebrafish. *Annu Rev Genet* **39**: 561-613.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-YY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**: 676-682.

Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, Schluter D, Kingsley DM. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**: 717-723.

Shi X, Richard J, Zirbes KM, Gong W, Lin G, Kyba M, Thomson JA, Koyano-Nakagawa N, Garry DJ. 2014. Cooperative interaction of Etv2 and Gata2 regulates the development of endothelial and hematopoietic lineages. *Dev Biol* **389**: 208-218.

Shivdasani RA, Mayer EL, Orkin SH. 1995. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**: 432-434.

Si-Tayeb K, Lemaigre FPP, Duncan SA. 2010. Organogenesis and development of the liver. *Dev Cell* **18**: 175-189.

Silver L, Palis J. 1997. Initiation of murine embryonic erythropoiesis: a spatial analysis. *Blood* **89**: 1154-1164.

Simon JM, Giresi PG, Davis IJ, Lieb JD. 2012. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat Protoc* **7**: 256-267.

Snow JW, Trowbridge JJ, Johnson KD, Fujiwara T, Emambokus NE, Grass JA, Orkin SH, Bresnick EH. 2011. Context-dependent function of "GATA switch" sites in vivo. *Blood* **117**: 4769-4772.

Soler E, Andrieu-Soler C, de Boer E, Bryne JC, Thongjuea S, Stadhouders R, Palstra R-JJ, Stevens M, Kockx C, van Ijcken W, Hou J, Steinhoff C, Rijkers E, Lenhard B, Grosveld F. 2010. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev* **24**: 277-289.

Song L, Crawford GE. 2010. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor Protocols*.

Sood R, Carrington B, Bishop K, Jones M, Rissone A, Candotti F, Chandrasekharappa SC, Liu P. 2013. Efficient methods for targeted mutagenesis in zebrafish using zinc-finger nucleases: data from targeting of nine genes using CompoZr or CoDA ZFNs. *PLoS One* **8**: e57239.

Sood R, English MA, Belele CL, Jin H, Bishop K, Haskins R, McKinney MC, Chahal J, Weinstein BM, Wen Z, Liu PP. 2010. Development of multilineage adult hematopoiesis in the zebrafish with a runx1 truncation mutation. *Blood* **115**: 2806-2809.

Spinner MA, Sanchez LA, Hsu AP, Shaw PA, Zerbe CS, Calvo KR, Arthur DC, Gu W, Gould CM, Brewer CC, Cowen EW, Freeman AF, Olivier KN, Uzel G, Zelazny AM, Daub JR, Spalding CD, Claypool RJ, Giri NK, Alter BP, Mace EM, Orange JS, Cuellar-Rodriguez J, Hickstein DD,

Holland SM. 2014. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. *Blood* **123**: 809-821.

Spitz F, Furlong EE. 2012. Transcription factors: from enhancer binding to developmental control. *Nature reviews Genetics* **13**: 613-626.

Stachura DL, Svoboda O, Campbell CA, Espín-Palazón R, Lau RP, Zon LI, Bartůněk P, Traver D. 2013. The zebrafish granulocyte colony-stimulating factors (Gcsfs): 2 paralogous cytokines and their roles in hematopoietic development and maintenance. *Blood* **122**: 3918-3928.

Stainier DY, Weinstein BM, Detrich HW, Zon LI, Fishman MC. 1995. Cloche, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. *Development* **121**: 3141-3150.

Sturgeon CM, Ditadi A, Awong G, Kennedy M, Keller G. 2014. Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. *Nat Biotechnol* **32**: 554-561.

Sugimura R, Jha D, Han A, Soria-Valles C, da Rocha E, Lu Y-F, Goettel JA, Serrao E, Rowe GR, Malleshaiah M, Wong I, Sousa P, Zhu TN, Ditadi A, Keller G, Engelman AN, Snapper SB, Doulatov S, Daley GQ. 2017. Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature* **545**: 432-438.

Sugiyama D, Ogawa M, Hirose I, Jaffredo T, Arai K-i, Tsuji K. 2003. Erythropoiesis from acetyl LDL incorporating endothelial cells at the preliver stage. *Blood* **101**: 4733-4738.

Sumanas S, Lin S. 2006. Ets1-related protein is a key regulator of vasculogenesis in zebrafish. *PLoS Biol* **4**: e10.

Swiers G, Baumann C, O'Rourke J, Giannoulatou E, Taylor S, Joshi A, Moignard V, Pina C, Bee T, Kokkaliaris KD, Yoshimoto M, Yoder MC, Frampton J, Schroeder T, Enver T, Göttgens B, de Bruijn MFTR. 2013a. Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat Commun* **4**: 2924.

Swiers G, de Bruijn M, Speck NA. 2010. Hematopoietic stem cell emergence in the conceptus and the role of Runx1. *Int J Dev Biol* **54**: 1151-1163.

Swiers G, Rode C, Azzoni E, de Bruijn MF. 2013b. A short history of hemogenic endothelium. *Blood Cells Mol Dis* **51**: 206-212.

Takada R, Satomi Y, Kurata T, Ueno N, Norioka S, Kondoh H, Takao T, Takada S. 2006. Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. *Dev Cell* **11**: 791-801.

Tamplin OJ, Durand EM, Carr LA, Childs SJ, Hagedorn EJ, Li P, Yzaguirre AD, Speck NA, Zon LI. 2015. Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell* **160**: 241-252.

Tanaka Y, Joshi A, Wilson NK, Kinston S, Nishikawa S, Göttgens B. 2012. The transcriptional programme controlled by Runx1 during early embryonic blood development. *Dev Biol* **366**: 404-419.

Tang Y, Bai H, Urs S, Wang Z, Liaw L. 2013. Notch1 activation in embryonic VE-cadherin populations selectively blocks hematopoietic stem cell generation and fetal liver hematopoiesis. *Transgenic Res* **22**: 403-410.

Tavian M, Coulombel L, Luton D, Clemente HS, Dieterlen-Lièvre F, Péault B. 1996. Aorta-associated CD34+ hematopoietic cells in the early human embryo. *Blood* **87**: 67-72.

Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y. 2003. Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res* **13**: 382-390.

Taylor JS, Van de Peer Y, Braasch I, Meyer A. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos Trans R Soc Lond B Biol Sci* **356**: 1661-1679.

Thambyrajah R, Ucanok D, Jalali M, Hough Y, Wilkinson RN, McMahon K, Moore C, Gering M. 2016. A gene trap transposon eliminates haematopoietic expression of zebrafish Gfi1aa, but does not interfere with haematopoiesis. *Dev Biol* **417**: 25-39.

Thompson MA, Ransom DG, Pratt SJ, MacLennan H, Kieran MW, Detrich HW, Vail B, Huber TL, Paw B, Brownlie AJ, Oates AC, Fritz A, Gates MA, Amores A, Bahary N, Talbot WS, Her H, Beier DR, Postlethwait JH, Zon LI. 1998. The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Dev Biol* **197**: 248-269.

Tian Y, Xu J, Feng S, He S, Zhao S, Zhu L, Jin W, Dai Y, Luo L, Qu JY, Wen Z. 2017. The first wave of T lymphopoiesis in zebrafish arises from aorta endothelium independent of hematopoietic stem cells. *J Exp Med* **214**: 3347-3360.

Tijssen MR, Cvejic A, Joshi A, Hannah RL, Ferreira R, Forrai A, Bellissimo DC, Oram SH, Smethurst PA, Wilson NK, Wang X, Ottersbach K, Stemple DL, Green AR, Ouwehand WH, Göttgens B. 2011. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev Cell* **20**: 597-609.

Tipping AJ, Pina C, Castor A, Hong D, Rodrigues NP, Lazzari L, May GE, Jacobsen SE, Enver T. 2009. High GATA-2 expression inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. *Blood* **113**: 2661-2672.

Tober J, Yzaguirre AD, Piwarzyk E, Speck NA. 2013. Distinct temporal requirements for Runx1 in hematopoietic progenitors and stem cells. *Development* **140**: 3765-3776.

Traver D, Paw BH, Poss KD, Penberthy WT, Lin S, Zon LI. 2003. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol* **4**: 1238-1246.

Travnickova J, Tran Chau V, Julien E, Mateos-Langerak J, Gonzalez C, Lelièvre E, Lutfalla G, Tavian M, Kissa K. 2015. Primitive macrophages control HSPC mobilization and definitive haematopoiesis. *Nat Commun* **6**: 6227.

Tripic T, Deng W, Cheng Y, Zhang Y, Vakoc CR, Gregory GD, Hardison RC, Blobel GA. 2009. SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. *Blood* **113**: 2191-2201.

Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, Warman ML. 2000. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* **29**: 52-54.

Tsai FY, Keller G, Kuo FC, Weiss M, Chen J, Rosenblatt M, Alt FW, Orkin SH. 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* **371**: 221-226.

Tsai FY, Orkin SH. 1997. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood* **89**: 3636-3643.

Turpen JB, Kelley CM, Mead PE, Zon LI. 1997. Bipotential primitive-definitive hematopoietic progenitors in the vertebrate embryo. *Immunity* **7**: 325-334.

Ueno H, Weissman IL. 2006. Clonal analysis of mouse development reveals a polyclonal origin for yolk sac blood islands. *Dev Cell* **11**: 519-533.

Uyttendaele H, Marazzi G, Wu G, Yan Q, Sassoon D, Kitajewski J. 1996. Notch4/int-3, a mammary proto-oncogene, is an endothelial cell-specific mammalian Notch gene. *Development* **122**: 2251-2259.

Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A. 2011. Determinants of nucleosome organization in primary human cells. *Nature* **474**: 516-520.

Van Handel B, Montel-Hagen A, Sasidharan R, Nakano H, Ferrari R, Boogerd CJ, Schredelseker J, Wang Y, Hunter S, Org T, Zhou J, Li X, Pellegrini M, Chen J-NN, Orkin SH, Kurdistani SK, Evans SM, Nakano A, Mikkola HK. 2012. Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell* **150**: 590-605.

Varshney GK, Sood R, Burgess SM. 2015. Understanding and Editing the Zebrafish Genome. *Adv Genet* **92**: 1-52.

Veldman MB, Lin S. 2012. Etsrp/Etv2 is directly regulated by Foxc1a/b in the zebrafish angioblast. *Circ Res* **110**: 220-229.

Vicente C, Conchillo A, Garcia-Sanchez MA, Odero MD. 2012. The role of the GATA2 transcription factor in normal and malignant hematopoiesis. *Crit Rev Oncol Hematol* **82**: 1-17.

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**: 854-858.

Visel A, Thaller C, Eichele G. 2004. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. *Nucleic Acids Res* **32**: D552–D556.

Vogeli KM, Jin S-W, Martin GR, Stainier DYR. 2006. A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. *Nature* **443**: 337-339.

Wadman IA, Osada H, Grutz GG, Agulnick AD, Westphal H, Forster A, Rabbitts TH. 1997. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J* **16**: 3145-3157.

Walmsley M, Ciau-Uitz A, Patient R. 2002. Adult and embryonic blood and endothelium derive from distinct precursor populations which are differentially programmed by BMP in Xenopus. *Development* **129**: 5683-5695.

Wang HU, Chen ZF, Anderson DJ. 1998. Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. *Cell* **93**: 741-753.

Wang Q, Stacy T, Binder M, Marin-Padilla M, Sharpe AH, Speck NA. 1996. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci U S A* **93**: 3444-3449.

Wang S, Yin J, Chen D, Nie F, Song X, Fei C, Miao H, Jing C, Ma W, Wang L, Xie S, Li C, Zeng R, Pan W, Hao X, Li L. 2013. Small-molecule modulation of Wnt signaling via modulating the Axin-LRP5/6 interaction. *Nat Chem Biol* **9**: 579-585.

Wareing S, Eliades A, Lacaud G, Kouskoff V. 2012. ETV2 expression marks blood and endothelium precursors, including hemogenic endothelium, at the onset of blood development. *Dev Dyn* **241**: 1454-1464.

Warren AJ, Colledge WH, Carlton MB, Evans MJ, Smith AJ, Rabbitts TH. 1994. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell* **78**: 45-57.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189-1191.

Wei W, Wen L, Huang P, Zhang Z, Chen Y, Xiao A, Huang H, Zhu Z, Zhang B, Lin S. 2008. Gfi1.1 regulates hematopoietic lineage differentiation during zebrafish embryogenesis. *Cell Res* **18**: 677-685.

Weiss A, Attisano L. 2013. The TGFbeta superfamily signaling pathway. *Wiley Interdiscip Rev Dev Biol* **2**: 47-63.

Westerfield M. 2007. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio).* Univ of Oregon Press, Eugene, OR.

Wilkinson RN, Elworthy S, Ingham PW, van Eeden FJM. 2013. A method for high-throughput PCR-based genotyping of larval zebrafish tail biopsies. *Biotechniques* **55**: 314-316.

Wilkinson RN, Koudijs MJ, Patient RK, Ingham PW, Schulte-Merker S, van Eeden FJ. 2012. Hedgehog signaling via a calcitonin receptor-like receptor can induce arterial differentiation independently of VEGF signaling in zebrafish. *Blood* **120**: 477-488.

Wilkinson RN, Pouget C, Gering M, Russell AJ, Davies SG, Kimelman D, Patient R. 2009. Hedgehog and Bmp polarize hematopoietic stem cell emergence in the zebrafish dorsal aorta. *Dev Cell* **16**: 909-916.

Willett CE, Zapata AG, Hopkins N, Steiner LA. 1997. Expression of zebrafish rag genes during early development identifies the thymus. *Dev Biol* **182**: 331-341.

Wilson NK, Calero-Nieto FJ, Ferreira R, Göttgens B. 2011. Transcriptional regulation of haematopoietic transcription factors. *Stem Cell Res Ther* **2**: 6.

Wilson NK, Foster SD, Wang X, Knezevic K, Schutte J, Kaimakis P, Chilarska PM, Kinston S, Ouwehand WH, Dzierzak E, Pimanda JE, de Bruijn MF, Gottgens B. 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**: 532-544.

Wozniak RJ, Boyer ME, Grass JA, Lee Y, Bresnick EH. 2007. Context-dependent GATA factor function: combinatorial requirements for transcriptional control in hematopoietic and endothelial cells. *J Biol Chem* **282**: 14665-14674.

Wozniak RJ, Keles S, Lugus JJ, Young KH, Boyer ME, Tran TM, Choi K, Bresnick EH. 2008. Molecular hallmarks of endogenous chromatin complexes containing master regulators of hematopoiesis. *Mol Cell Biol* **28**: 6681-6694.

Wu W, Morrissey CS, Keller CA, Mishra T, Pimkin M, Blobel GA, Weiss MJ, Hardison RC. 2014. Dynamic shifts in occupancy by TAL1 are guided by GATA factors and drive large-scale reprogramming of gene expression during hematopoiesis. *Genome Res* **24**: 1945-1962.

Xu H, Li C, Zeng Q, Agrawal I, Zhu X, Gong Z. 2016. Genome-wide identification of suitable zebrafish Danio rerio reference genes for normalization of gene expression data by RT-qPCR. *J Fish Biol.*

Yamada Y, Warren AJ, Dobson C, Forster A, Pannell R, Rabbitts TH. 1998. The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis. *Proc Natl Acad Sci U S A* **95**: 3890-3895.

Yang L, Rastegar S, Strähle U. 2010. Regulatory interactions specifying Kolmer-Agduhr interneurons. *Development* **137**: 2713-2722.

Yin A, Korzh S, Winata CL, Korzh V, Gong Z. 2011. Wnt Signaling Is Required for Early Development of Zebrafish Swimbladder. *PLoS One* **6**: e18431.

Yokomizo T, Dzierzak E. 2010. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development* **137**: 3651-3661.

Yokomizo T, Ng CE, Osato M, Dzierzak E. 2011. Three-dimensional imaging of whole midgestation murine embryos shows an intravascular localization for all hematopoietic clusters. *Blood* **117**: 6132-6134.

Yokomizo T, Ogawa M, Osato M, Kanno T, Yoshida H, Fujimoto T, Fraser S, Nishikawa S, Okada H, Satake M, Noda T, Nishikawa S, Ito Y. 2001. Requirement of Runx1/AML1/PEBP2alphaB for the generation of haematopoietic cells from endothelial cells. *Genes Cells* **6**: 13-23.

Yokomizo T, Takahashi S, Mochizuki N, Kuroha T, Ema M, Wakamatsu A, Shimizu R, Ohneda O, Osato M, Okada H, Komori T, Ogawa M, Nishikawa S-I, Ito Y, Yamamoto M. 2007. Characterization of GATA-1(+) hemangioblastic cells in the mouse embryo. *Embo J* **26**: 184-196.

Yoshimoto M, Montecino-Rodriguez E, Ferkowicz MJ, Porayette P, Shelley WC, Conway SJ, Dorshkind K, Yoder MC. 2011. Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. *Proc Natl Acad Sci U S A* **108**: 1468-1473.

Yoshimoto M, Porayette P, Glosson NL, Conway SJ, Carlesso N, Cardoso AA, Kaplan MH, Yoder MC. 2012. Autonomous murine T-cell progenitor production in the extra-embryonic yolk sac before HSC emergence. *Blood* **119**: 5706-5714.

Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. 2011a. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29**: 149-153.

Zhang P, He Q, Chen D, Liu W, Wang L, Zhang C, Ma D, Li W, Liu B, Liu F. 2015. G protein-coupled receptor 183 facilitates endothelial-to-hematopoietic transition via Notch1 inhibition. *Cell Res* **25**: 1093-1107.

Zhang Y, Jin H, Li L, Qin FX, Wen Z. 2011b. cMyb regulates hematopoietic stem/progenitor cell mobilization during zebrafish hematopoiesis. *Blood* **118**: 4093-4101.

Zhen F, Lan Y, Yan B, Zhang W, Wen Z. 2013. Hemogenic endothelium specification and hematopoietic stem cell maintenance employ distinct Scl isoforms. *Development* **140**: 3977-3985.

Zhong TP, Childs S, Leu JP, Fishman MC. 2001. Gridlock signalling pathway fashions the first embryonic artery. *Nature* **414**: 216-220.

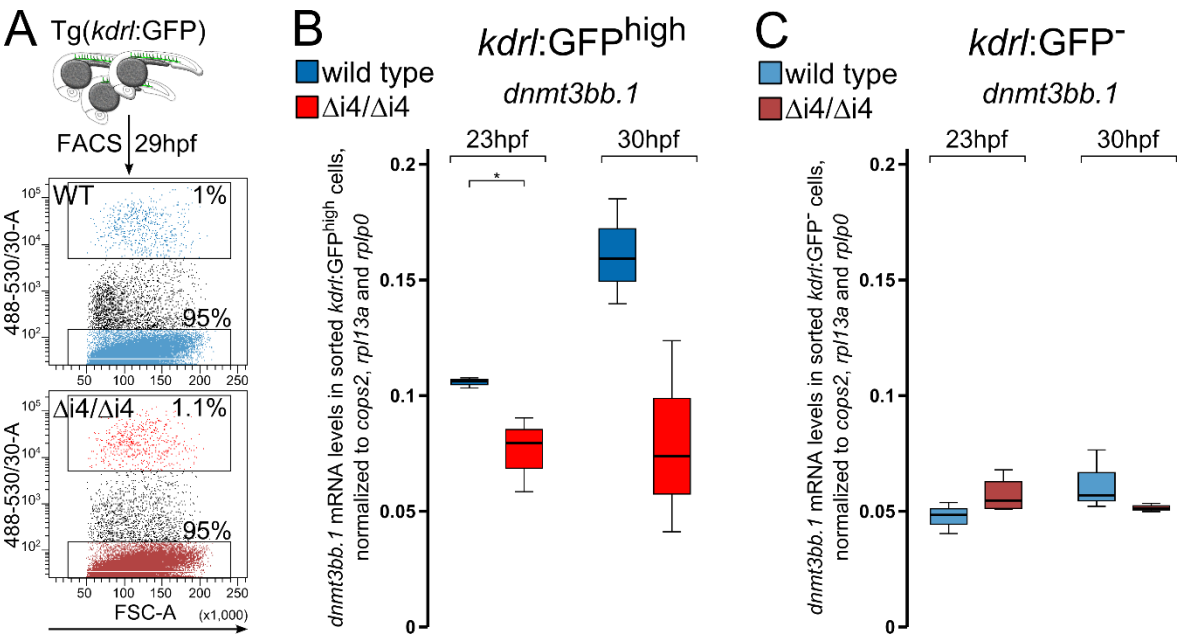
Zhou HY, Katsman Y, Dhaliwal NK, Davidson S, Macpherson NN, Sakthidevi M, Collura F, Mitchell JA. 2014. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. *Genes Dev* **28**: 2699-2711.

Zhu C, Smith T, McNulty J, Rayla AL, Lakshmanan A, Siekmann AF, Buffardi M, Meng X, Shin J, Padmanabhan A, Cifuentes D, Giraldez AJ, Look AT, Epstein JA, Lawson ND, Wolfe SA. 2011. Evaluation and application of modularly assembled zinc-finger nucleases in zebrafish. *Development* **138**: 4555-4564.

Zovein AC, Hofmann JJ, Lynch M, French WJ, Turlo KA, Yang Y, Becker MS, Zanetta L, Dejana E, Gasson JC, Tallquist MD, Iruela-Arispe LM. 2008. Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. *Cell Stem Cell* **3**: 625-636.

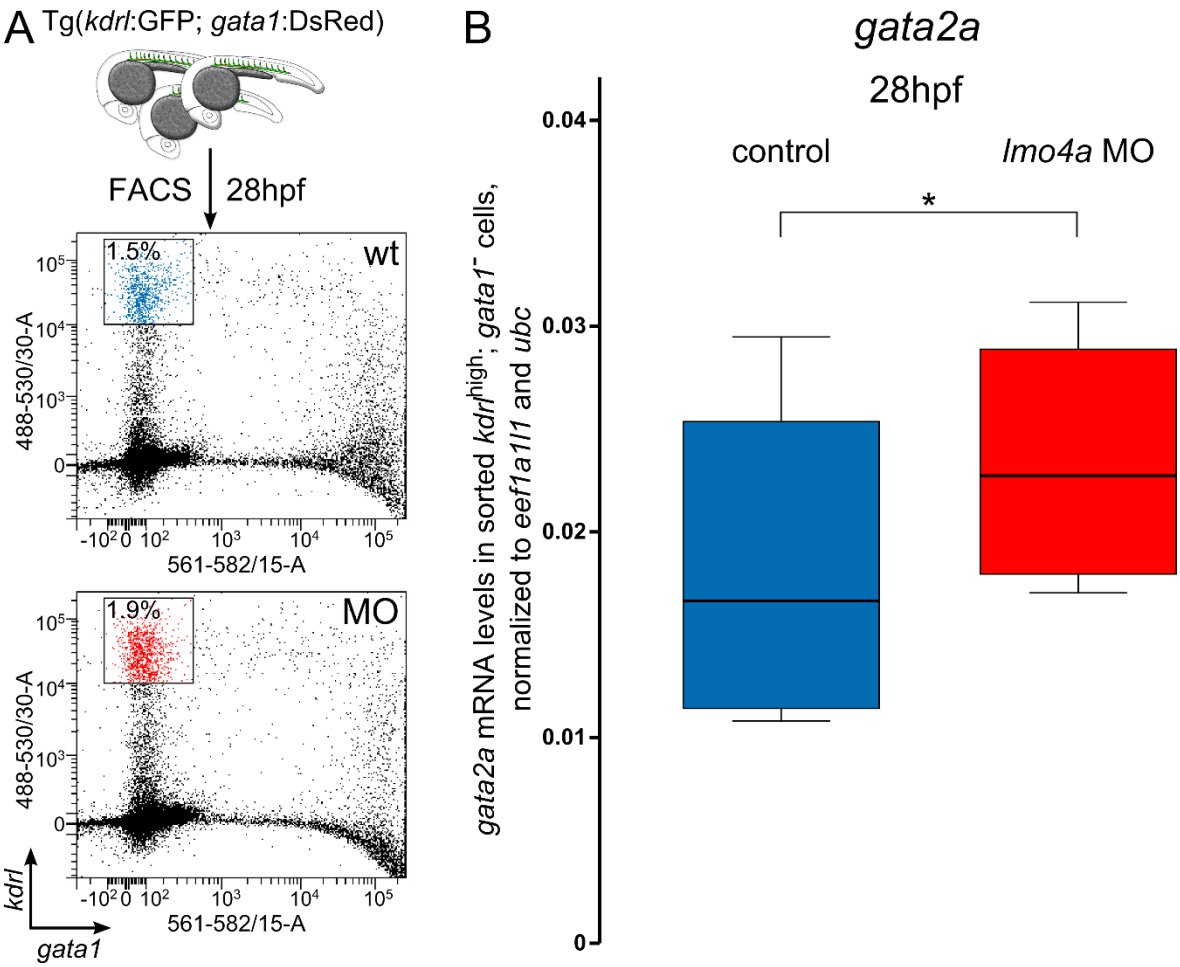
Zovein AC, Turlo KA, Ponec RM, Lynch MR, Chen KC, Hofmann JJ, Cox TC, Gasson JC, Iruela-Arispe ML. 2010. Vascular remodeling of the vitelline artery initiates extravascular emergence of hematopoietic clusters. *Blood* **116**: 3435-3444.

Appendix I: Supplementary Figures



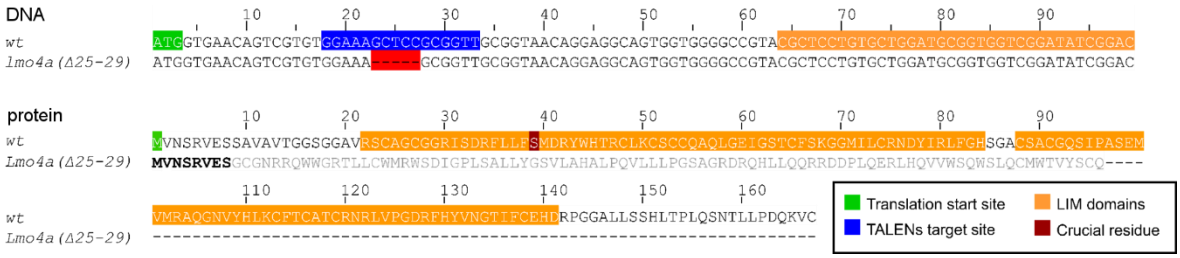
Supplementary Figure 4.1. *Gata2a*$^{\Delta i4/\Delta i4}$ zebrafish mutants, compared to wild type embryos, express lower levels of *dnmt3bb.1* in the dorsal aorta during the specification of HE.

(A) *Kdrl*:GFP^{high} and *kdrl*:GFP^{-} cells were sorted from non-mutant (WT, blue) and *gata2a*$^{\Delta i4/\Delta i4}$ (red) embryos carrying the Tg(*kdrl*:GFP) transgene. (B-C) QRT-PCR on RNA isolated from the sorted *kdrl*:GFP^{high} or *kdrl*:GFP^{-} cells (panel E) shows decreased levels of *dnmt3bb.1* mRNA in the endothelium of *gata2a*$^{\Delta i4/\Delta i4}$ mutants at 23hpf and 30hpf, compared to wild type. In non-endothelial cells, there is no significant difference between *gata2a*$^{\Delta i4/\Delta i4}$ mutants and wt controls. N=4 for *gata2a*$^{\Delta i4/\Delta i4}$ at 23hpf, N=3 for other samples. *p<0.05.



Supplementary Figure 5.1. Expression of *gata2a* increases in the endothelium of *lmo4a* morphants.

(A) *Kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells were sorted from wild type (wt; blue) and *lmo4a* MO-injected embryos (MO; red) at 28hpf. (B) QRT-PCR on RNA isolated from the sorted cell populations (panel A) shows significantly increased levels of *gata2a* mRNA in the endothelium of *lmo4a* morphants at 28hpf, compared to wild type. N=4. *p<0.05.



Supplementary Figure 5.2. Generation of *lmo4a*$^{\Delta 25\text{-}29}$ zebrafish mutants with TALENs.

TALENs were designed to a region (blue) ~20bp downstream of *lmo4a* translation start site (green). Isolated mutant alleles carry 5bp deletions (Δ25-29) (red gaps) upstream of the conserved LIM domains (orange), resulting in a frameshift after S8. Resulting mutant protein is predicted to lack the LIM domains, including the crucial S39 (purple).

Appendix II: Differentially expressed genes identified in the *lmo4a* RNA-seq experiment (related to Chapter 5.3).

Genes upregulated in *lmo4a* morphants

GeneID	logFC	PValue	FDR	ZFIN ID
serpinb1l3	-5.21554	3.543E-27	1.328E-24	ZDB-GENE-030131-7059
pigm	-5.07582	1.022E-33	5.744E-31	ZDB-GENE-040426-1393
ghitm	-4.30337	5.447E-69	7.959E-65	ZDB-GENE-040426-1537
tmem11	-3.83343	7.399E-60	1.544E-56	ZDB-GENE-050417-37
sf3a2	-3.69716	8.525E-49	1.132E-45	ZDB-GENE-040426-897
lect1	-3.68531	8.527E-25	2.709E-22	ZDB-GENE-010713-1
mcm7	-3.65676	1.598E-43	1.556E-40	ZDB-GENE-020419-27
pigc	-3.63300	9.704E-22	2.363E-19	ZDB-GENE-030131-2714
ubiad1	-3.47292	5.714E-37	4.394E-34	ZDB-GENE-030131-3205
pggt1b	-3.46618	3.905E-29	1.678E-26	ZDB-GENE-050913-85
atxn3	-3.44850	1.491E-18	2.690E-16	ZDB-GENE-040426-983
hyi	-3.37084	3.088E-20	6.182E-18	ZDB-GENE-040426-1273
cdkn1ba	-3.29146	5.116E-24	1.557E-21	ZDB-GENE-040812-3
foxo1a	-3.26539	1.288E-36	8.960E-34	ZDB-GENE-061013-59
clcn4	-3.24817	3.343E-14	3.908E-12	ZDB-GENE-061013-353
ppp1r21	-3.22531	1.479E-34	9.007E-32	ZDB-GENE-041114-8
cldn5a	-3.21634	1.651E-14	2.062E-12	ZDB-GENE-040426-2442
jun	-3.21329	2.635E-66	9.627E-63	ZDB-GENE-030131-7859
irf2bp2b	-3.17866	4.477E-13	4.332E-11	ZDB-GENE-040426-2574
ap4b1	-3.14208	2.950E-08	1.286E-06	ZDB-GENE-040426-1284
creg1	-3.04487	1.621E-21	3.821E-19	ZDB-GENE-070112-2042
dync1h1	-3.01629	1.208E-16	1.877E-14	ZDB-GENE-030131-7050
mafbb	-3.01422	8.667E-12	6.958E-10	ZDB-GENE-010605-4
lhfpl2a	-3.00925	2.526E-23	7.383E-21	ZDB-GENE-040426-1839
fam53b	-3.00020	1.858E-43	1.696E-40	ZDB-GENE-040426-2495
fzd7a	-2.99938	1.646E-51	2.673E-48	ZDB-GENE-990415-223
irf2bp2a	-2.87284	3.361E-34	1.965E-31	ZDB-GENE-030131-7809
zgc:85866	-2.87128	2.328E-30	1.063E-27	ZDB-GENE-040426-3
zgc:64022	-2.86518	3.408E-10	2.184E-08	ZDB-GENE-040426-1340
tmem39a	-2.85682	7.942E-17	1.248E-14	ZDB-GENE-030131-1928
wwtr1	-2.85457	1.501E-19	2.848E-17	ZDB-GENE-051101-1
tcf7l1a	-2.85396	2.756E-25	8.950E-23	ZDB-GENE-980605-30
hhex	-2.85325	3.347E-15	4.658E-13	ZDB-GENE-980526-299
nkain1	-2.83614	3.911E-32	1.905E-29	ZDB-GENE-040426-1472
ipmkb	-2.80844	9.580E-08	3.674E-06	ZDB-GENE-061215-5
zgc:165608	-2.77999	3.374E-12	2.866E-10	ZDB-GENE-070615-36
alas1	-2.76400	4.317E-12	3.625E-10	ZDB-GENE-001229-2
gata2a	-2.76253	5.313E-35	3.376E-32	ZDB-GENE-980526-260
trim8a	-2.74541	2.277E-15	3.262E-13	ZDB-GENE-130107-1
tek	-2.73735	1.782E-18	3.176E-16	ZDB-GENE-990415-56
dusp7	-2.73687	5.475E-12	4.520E-10	ZDB-GENE-030131-5457
hspb8	-2.72227	1.260E-13	1.334E-11	ZDB-GENE-030131-2480
ciao1	-2.71048	1.322E-29	5.855E-27	ZDB-GENE-040426-839
mafba	-2.68210	1.433E-50	2.093E-47	ZDB-GENE-980526-515
isg15	-2.62616	2.086E-12	1.803E-10	ZDB-GENE-021211-1
pcsk7	-2.62434	1.283E-12	1.143E-10	ZDB-GENE-030131-7293

tnfaip6	-2.62010	7.319E-22	1.844E-19	ZDB-GENE-060421-2654
h1f0	-2.60924	1.395E-23	4.161E-21	ZDB-GENE-030131-337
rab3gap1	-2.59032	1.358E-13	1.428E-11	ZDB-GENE-061027-53
arid3a	-2.58586	5.938E-33	3.213E-30	ZDB-GENE-140106-97
nr2f1a	-2.58514	1.899E-46	2.312E-43	ZDB-GENE-980526-115
tbc1d20	-2.57493	8.667E-11	5.990E-09	ZDB-GENE-050626-61
psmd3	-2.55960	3.051E-23	8.742E-21	ZDB-GENE-040426-1444
prg4b	-2.55316	2.577E-14	3.112E-12	ZDB-GENE-041014-324
zgc:113944	-2.52496	3.149E-08	1.349E-06	ZDB-GENE-050913-155
slc16a2	-2.51407	1.089E-08	5.235E-07	ZDB-GENE-120117-1
emc6	-2.50565	9.536E-11	6.542E-09	ZDB-GENE-030131-9164
tdrd5	-2.50364	2.585E-09	1.399E-07	ZDB-GENE-050208-502
slc27a1b	-2.49858	1.122E-03	1.290E-02	ZDB-GENE-061013-672
zgc:162310	-2.49248	7.166E-07	2.252E-05	ZDB-GENE-070410-69
mycl1a	-2.48985	4.812E-21	1.082E-18	ZDB-GENE-040426-2439
wdr91	-2.48344	1.119E-13	1.202E-11	ZDB-GENE-050113-2
znrf1	-2.45959	4.510E-09	2.362E-07	ZDB-GENE-040426-1905
tp53inp1	-2.44704	9.179E-38	7.451E-35	ZDB-GENE-031018-3
eps8	-2.43433	1.026E-10	7.002E-09	ZDB-GENE-040426-871
limch1b	-2.41796	1.947E-24	6.054E-22	ZDB-GENE-030131-9210
zgc:56039	-2.39680	6.258E-07	2.014E-05	ZDB-GENE-030131-7832
hbp1	-2.39026	7.568E-15	9.963E-13	ZDB-GENE-050522-414
nr2f6a	-2.37304	9.258E-29	3.758E-26	ZDB-GENE-030131-4362
dnm2b	-2.37187	4.294E-22	1.141E-19	ZDB-GENE-040426-2371
foxc1a	-2.35867	1.240E-32	6.246E-30	ZDB-GENE-010302-1
atoh8	-2.35649	8.497E-12	6.859E-10	ZDB-GENE-061215-7
zgc:92481	-2.33371	1.270E-05	2.859E-04	ZDB-GENE-040912-181
cox15	-2.32643	2.011E-20	4.260E-18	ZDB-GENE-040426-955
zgc:92789	-2.32447	2.132E-06	5.866E-05	ZDB-GENE-040718-280
pdcd6ip	-2.30532	2.109E-26	7.269E-24	ZDB-GENE-040426-2678
zgc:77439	-2.30177	7.517E-08	2.976E-06	ZDB-GENE-031010-24
zgc:110091	-2.28978	2.064E-21	4.788E-19	ZDB-GENE-050417-34
snx27b	-2.27818	1.395E-17	2.290E-15	ZDB-GENE-040426-1529
mrpl20	-2.27038	1.931E-04	2.973E-03	ZDB-GENE-030131-5299
tsc22d1	-2.25055	6.990E-17	1.110E-14	ZDB-GENE-030131-7785
ppm1h	-2.22927	5.491E-09	2.825E-07	ZDB-GENE-061027-190
nelfa	-2.22805	2.712E-17	4.355E-15	ZDB-GENE-030131-6400
thap11	-2.22743	2.107E-08	9.623E-07	ZDB-GENE-040426-2121
azin1a	-2.21684	6.519E-18	1.108E-15	ZDB-GENE-050417-20
slc43a2a	-2.21531	6.299E-16	9.297E-14	ZDB-GENE-040426-964
carhsp1	-2.21166	2.421E-06	6.613E-05	ZDB-GENE-040426-1479
cdr2l	-2.20008	3.564E-12	3.010E-10	ZDB-GENE-030131-8171
socs3b	-2.19541	6.745E-08	2.700E-06	ZDB-GENE-040426-2528
nkx2.3	-2.19116	4.270E-07	1.418E-05	ZDB-GENE-990415-178
pink1	-2.18224	2.677E-04	3.892E-03	ZDB-GENE-041212-53
parp8	-2.17110	5.351E-09	2.763E-07	ZDB-GENE-060929-432
ncl1	-2.16634	1.777E-13	1.828E-11	ZDB-GENE-030131-8788
btg2	-2.16312	1.519E-16	2.313E-14	ZDB-GENE-000210-15

bloc1s4	-2.15125	4.063E-08	1.711E-06	ZDB-GENE-040801-215
foxl1	-2.14834	1.049E-14	1.332E-12	ZDB-GENE-040426-1181
gsk3aa	-2.14622	1.587E-06	4.555E-05	ZDB-GENE-060503-796
kcmf1	-2.13889	9.911E-16	1.434E-13	ZDB-GENE-040426-1383
mpp2b	-2.12410	1.731E-07	6.293E-06	ZDB-GENE-040704-70
gde1	-2.09102	2.337E-07	8.247E-06	ZDB-GENE-040426-881
tmem65	-2.08524	7.870E-08	3.100E-06	ZDB-GENE-081022-99
btd	-2.08276	1.567E-03	1.695E-02	ZDB-GENE-060825-186
klc1b	-2.07437	1.234E-12	1.106E-10	ZDB-GENE-030131-7439
cbx7a	-2.07312	2.594E-15	3.680E-13	ZDB-GENE-050417-400
mri1	-2.06790	1.134E-05	2.599E-04	ZDB-GENE-080204-109
zgc:153012	-2.06304	8.048E-06	1.918E-04	ZDB-GENE-061103-559
dvl2	-2.06236	5.829E-06	1.453E-04	ZDB-GENE-041118-20
mych	-2.05836	4.586E-10	2.852E-08	ZDB-GENE-030219-51
rdh14b	-2.02494	1.963E-07	7.014E-06	ZDB-GENE-030131-6605
diras1a	-2.02301	2.159E-05	4.474E-04	ZDB-GENE-030131-5788
si:dkey-222b8.1	-2.02167	8.320E-07	2.559E-05	ZDB-GENE-060503-248
atp6v0cb	-2.02130	2.552E-06	6.945E-05	ZDB-GENE-030131-4127
ahcyl2	-2.00002	2.249E-13	2.298E-11	ZDB-GENE-040115-5
fgf3	-1.99816	1.707E-03	1.825E-02	ZDB-GENE-980526-178
hps3	-1.99757	4.384E-10	2.749E-08	ZDB-GENE-061110-115
sap30l	-1.98775	4.402E-22	1.149E-19	ZDB-GENE-030131-5287
slc30a4	-1.98534	2.964E-14	3.550E-12	ZDB-GENE-040426-1408
hn1a	-1.98491	3.010E-03	2.770E-02	ZDB-GENE-070410-142
cebpd	-1.97735	4.140E-23	1.163E-20	ZDB-GENE-020111-4
ghdc	-1.97687	4.217E-04	5.790E-03	ZDB-GENE-050208-165
rom1b	-1.97412	2.092E-04	3.177E-03	ZDB-GENE-040426-1073
slc16a9a	-1.97103	1.503E-10	1.003E-08	ZDB-GENE-040426-1364
gpd1c	-1.97005	2.139E-26	7.269E-24	ZDB-GENE-040426-2576
sp4	-1.96378	1.595E-05	3.453E-04	ZDB-GENE-031202-3
dennd2db	-1.96264	1.358E-27	5.221E-25	ZDB-GENE-041114-34
rnf150a	-1.95028	5.101E-11	3.672E-09	ZDB-GENE-060213-1
faxdc2	-1.94476	6.751E-05	1.187E-03	ZDB-GENE-040426-2907
erg	-1.94321	2.061E-26	7.269E-24	ZDB-GENE-041212-40
dmrt2b	-1.94284	2.811E-07	9.694E-06	ZDB-GENE-080813-1
rnf139	-1.93731	1.702E-09	9.387E-08	ZDB-GENE-080401-4
ftr14	-1.93728	5.013E-06	1.274E-04	ZDB-GENE-060512-201
edn2	-1.93025	2.130E-21	4.863E-19	ZDB-GENE-060503-774
cherp	-1.92787	2.401E-17	3.899E-15	ZDB-GENE-040407-1
slc9a3.2	-1.92707	5.183E-03	4.264E-02	ZDB-GENE-060503-640
bri3	-1.91545	3.897E-18	6.779E-16	ZDB-GENE-030131-3038
irf2bpl	-1.91435	4.453E-13	4.332E-11	ZDB-GENE-040426-1968
smad4a	-1.90875	4.371E-09	2.297E-07	ZDB-GENE-990603-8
zgc:153247	-1.90755	3.823E-06	1.001E-04	ZDB-GENE-060825-214
snx16	-1.90488	9.133E-08	3.521E-06	ZDB-GENE-040426-1832
si:ch211-216l23.2	-1.90448	6.907E-15	9.259E-13	ZDB-GENE-030131-3806
ttc8	-1.89286	5.992E-03	4.779E-02	ZDB-GENE-030131-8846
tbl3	-1.88860	1.683E-08	7.831E-07	ZDB-GENE-041114-104

bmp1b	-1.87958	1.791E-03	1.891E-02	ZDB-GENE-060818-2
pdlim7	-1.87772	1.426E-12	1.263E-10	ZDB-GENE-040426-2092
arpc5la	-1.87390	4.792E-08	2.012E-06	ZDB-GENE-040426-1453
mkrn1	-1.87277	1.260E-09	7.025E-08	ZDB-GENE-020213-1
usp12a	-1.86939	8.581E-06	2.036E-04	ZDB-GENE-060228-3
noa1	-1.85928	2.767E-06	7.488E-05	ZDB-GENE-080220-18
abcc5	-1.85860	2.460E-19	4.549E-17	ZDB-GENE-050517-17
mrpl38	-1.85652	8.011E-08	3.130E-06	ZDB-GENE-040426-2373
zgc:66427	-1.85160	7.348E-12	5.965E-10	ZDB-GENE-040426-1913
fzd3a	-1.84887	1.654E-05	3.546E-04	ZDB-GENE-990415-225
col9a2	-1.83510	1.625E-06	4.629E-05	ZDB-GENE-030131-9786
si:ch211-212o1.2	-1.83492	4.535E-04	6.175E-03	ZDB-GENE-041111-310
igfbp2a	-1.83467	5.162E-10	3.104E-08	ZDB-GENE-000125-12
plekha1	-1.83057	2.277E-10	1.485E-08	ZDB-GENE-040426-2780
bves	-1.82793	1.355E-16	2.085E-14	ZDB-GENE-040624-11
klhl8	-1.82425	3.063E-08	1.324E-06	ZDB-GENE-070209-226
zgc:63470	-1.82387	1.268E-05	2.859E-04	ZDB-GENE-040426-1126
gata2b	-1.82049	1.619E-03	1.739E-02	ZDB-GENE-040718-440
rab14	-1.81840	1.141E-09	6.486E-08	ZDB-GENE-030826-20
gopc	-1.81466	1.002E-05	2.335E-04	ZDB-GENE-030131-5086
zgc:64190	-1.81346	2.000E-03	2.048E-02	ZDB-GENE-030131-7568
kat2b	-1.80725	3.415E-05	6.706E-04	ZDB-GENE-060503-207
ubald2	-1.77872	5.839E-07	1.892E-05	ZDB-GENE-050522-185
tbx3a	-1.77590	1.844E-05	3.895E-04	ZDB-GENE-070209-80
twf1a	-1.77230	1.348E-06	3.956E-05	ZDB-GENE-030131-4402
zgc:113364	-1.77150	1.737E-03	1.849E-02	ZDB-GENE-050410-5
ankrd54	-1.76860	2.898E-13	2.880E-11	ZDB-GENE-040718-318
psen1	-1.76738	2.789E-10	1.803E-08	ZDB-GENE-991119-4
fbxl14b	-1.76705	1.724E-06	4.882E-05	ZDB-GENE-030131-3556
irx5b	-1.76301	1.180E-07	4.442E-06	ZDB-GENE-040628-5
b3gnt5a	-1.76218	2.874E-11	2.165E-09	ZDB-GENE-030131-8470
pou2f1a	-1.75650	1.136E-03	1.301E-02	ZDB-GENE-990415-185
nptnb	-1.75420	2.097E-03	2.124E-02	ZDB-GENE-040426-1821
bxdc2	-1.74801	7.403E-05	1.286E-03	ZDB-GENE-060518-1
zgc:65811	-1.74580	2.546E-03	2.439E-02	ZDB-GENE-040426-1428
babam1	-1.74181	1.747E-08	8.103E-07	ZDB-GENE-040808-6
pappab	-1.74106	2.803E-03	2.630E-02	ZDB-GENE-060526-71
pcmtd2	-1.73731	2.842E-04	4.092E-03	ZDB-GENE-060312-14
zgc:77838	-1.73670	3.429E-04	4.831E-03	ZDB-GENE-040426-1889
slc30a1a	-1.73595	9.518E-13	8.802E-11	ZDB-GENE-040426-1840
zgc:63863	-1.73342	4.368E-05	8.267E-04	ZDB-GENE-040426-1278
ets2	-1.73245	1.964E-08	9.023E-07	ZDB-GENE-050522-552
igfbp3	-1.72816	1.220E-09	6.881E-08	ZDB-GENE-040412-1
evla	-1.72811	5.965E-08	2.435E-06	ZDB-GENE-000607-65
zgc:195173	-1.72553	2.409E-11	1.843E-09	ZDB-GENE-081022-190
nog1	-1.72461	5.814E-03	4.678E-02	ZDB-GENE-991206-8
abhd17c	-1.72206	1.248E-08	5.939E-07	ZDB-GENE-040426-714
ddrgk1	-1.72124	1.859E-06	5.213E-05	ZDB-GENE-040426-1050

rnps1	-1.71980	6.216E-05	1.116E-03	ZDB-GENE-030131-5228
zgc:158357	-1.71759	2.190E-06	6.015E-05	ZDB-GENE-070112-1172
wipi1	-1.70682	6.626E-09	3.339E-07	ZDB-GENE-040426-1404
nudt8	-1.70528	7.334E-09	3.645E-07	ZDB-GENE-061013-219
slc3a2b	-1.69840	5.880E-11	4.171E-09	ZDB-GENE-040122-2
colgalt1	-1.69494	2.913E-08	1.278E-06	ZDB-GENE-070720-6
ago4	-1.69486	2.261E-08	1.026E-06	ZDB-GENE-030131-5357
zgc:158270	-1.68400	6.099E-04	7.901E-03	ZDB-GENE-070112-1442
vangl2	-1.67307	1.450E-06	4.214E-05	ZDB-GENE-020507-3
ptpn12	-1.66828	8.116E-06	1.931E-04	ZDB-GENE-040426-1441
hipk2	-1.66626	3.351E-07	1.141E-05	ZDB-GENE-031125-4
camk4	-1.66481	1.026E-03	1.200E-02	ZDB-GENE-050417-76
zgc:92287	-1.66298	4.497E-05	8.467E-04	ZDB-GENE-040718-350
cxcr3.2	-1.65884	1.930E-10	1.264E-08	ZDB-GENE-041114-186
med18	-1.65707	8.056E-04	9.909E-03	ZDB-GENE-040426-1276
hes6	-1.65431	1.608E-07	5.904E-06	ZDB-GENE-030828-5
pcdh10b	-1.64203	1.666E-07	6.101E-06	ZDB-GENE-030721-4
xpot	-1.64140	5.778E-08	2.378E-06	ZDB-GENE-041210-4
rasl11b	-1.64081	2.079E-09	1.142E-07	ZDB-GENE-040426-793
maf	-1.63831	8.521E-18	1.431E-15	ZDB-GENE-010605-2
adam10b	-1.63626	1.728E-04	2.694E-03	ZDB-GENE-071115-1
irf6	-1.63573	6.711E-04	8.497E-03	ZDB-GENE-040426-1137
tbc1d16	-1.62735	1.914E-10	1.260E-08	ZDB-GENE-041010-179
ybx1	-1.62322	2.756E-13	2.758E-11	ZDB-GENE-000629-3
zgc:158320	-1.61942	1.524E-05	3.354E-04	ZDB-GENE-061201-33
pbx1b	-1.61124	2.388E-03	2.320E-02	ZDB-GENE-070424-11
cpsf6	-1.60350	1.713E-13	1.775E-11	ZDB-GENE-030131-5277
ifit5	-1.60310	2.969E-03	2.749E-02	ZDB-GENE-120406-2
lats1	-1.60205	9.486E-04	1.134E-02	ZDB-GENE-050523-2
igsf8	-1.59888	3.758E-05	7.283E-04	ZDB-GENE-040426-2444
pdpk1a	-1.59835	2.063E-03	2.102E-02	ZDB-GENE-061013-109
h2afy	-1.59825	2.032E-04	3.109E-03	ZDB-GENE-060421-4796
sorbs3	-1.59746	6.395E-05	1.139E-03	ZDB-GENE-050327-1
zgc:65772	-1.59432	1.557E-05	3.396E-04	ZDB-GENE-040426-1415
zgc:56304	-1.59166	4.887E-14	5.535E-12	ZDB-GENE-040426-2698
klhl26	-1.58829	9.718E-04	1.156E-02	ZDB-GENE-030131-5975
ctnnb2	-1.58818	7.810E-20	1.542E-17	ZDB-GENE-040426-2575
prkacab	-1.58558	1.524E-07	5.609E-06	ZDB-GENE-040801-208
gbp1	-1.58075	1.048E-08	5.070E-07	ZDB-GENE-040718-32
cyr61	-1.58044	6.997E-12	5.712E-10	ZDB-GENE-060404-5
glcci1	-1.57964	1.134E-03	1.299E-02	ZDB-GENE-031010-41
runx1t1	-1.57849	4.930E-09	2.571E-07	ZDB-GENE-060929-164
pdlim5b	-1.57665	1.113E-08	5.332E-07	ZDB-GENE-040426-908
c1qbp	-1.57661	7.359E-09	3.645E-07	ZDB-GENE-050417-408
ctc1	-1.57169	1.101E-09	6.308E-08	ZDB-GENE-061103-271
snrpb	-1.56974	4.613E-04	6.253E-03	ZDB-GENE-040426-1819
cyp24a1	-1.56404	4.906E-04	6.582E-03	ZDB-GENE-060825-1
mapk1	-1.56146	5.838E-06	1.453E-04	ZDB-GENE-030722-2

si:ch73-56p18.4	-1.56088	6.410E-06	1.574E-04	ZDB-GENE-121127-1
mtdha	-1.56030	8.358E-21	1.823E-18	ZDB-GENE-030616-615
gadd45g	-1.55614	1.876E-03	1.956E-02	ZDB-GENE-040426-1882
plk3	-1.54656	9.535E-20	1.858E-17	ZDB-GENE-030131-6134
htra1b	-1.54624	9.936E-04	1.175E-02	ZDB-GENE-080219-7
larp6	-1.54583	4.729E-10	2.904E-08	ZDB-GENE-030131-6403
shisa3	-1.54415	3.023E-08	1.311E-06	ZDB-GENE-070112-2022
syngr3a	-1.54281	3.469E-04	4.870E-03	ZDB-GENE-050522-541
abl2	-1.54164	1.401E-04	2.254E-03	ZDB-GENE-020809-2
spata4	-1.54093	2.095E-03	2.124E-02	ZDB-GENE-050417-153
pdhx	-1.53880	6.336E-04	8.132E-03	ZDB-GENE-040426-1539
sox4a	-1.53753	7.183E-11	5.046E-09	ZDB-GENE-030131-8290
zgc:158619	-1.53450	2.434E-06	6.635E-05	ZDB-GENE-070209-116
dyrk1b	-1.53284	1.879E-12	1.644E-10	ZDB-GENE-060503-499
samsn1a	-1.53046	1.015E-08	4.928E-07	ZDB-GENE-030131-8639
sepp1a	-1.52749	1.680E-10	1.111E-08	ZDB-GENE-030311-1
gna12a	-1.52198	1.214E-12	1.095E-10	ZDB-GENE-050221-1
cyr61l1	-1.51880	2.609E-07	9.077E-06	ZDB-GENE-050417-136
irf2bp1	-1.51847	1.645E-04	2.585E-03	ZDB-GENE-050420-389
batf3	-1.51793	1.022E-03	1.197E-02	ZDB-GENE-041014-356
zgc:92249	-1.51644	1.039E-03	1.211E-02	ZDB-GENE-040718-368
mntb	-1.51501	3.519E-03	3.134E-02	ZDB-GENE-070820-8
ctgfa	-1.50842	1.015E-19	1.951E-17	ZDB-GENE-030131-102
spata2	-1.50585	1.380E-05	3.069E-04	ZDB-GENE-070822-5
dctn2	-1.50403	2.956E-08	1.286E-06	ZDB-GENE-040426-1279
hnrnpaba	-1.50393	1.666E-08	7.778E-07	ZDB-GENE-030131-185
src	-1.49904	2.070E-06	5.740E-05	ZDB-GENE-030131-3809
lygl1	-1.49900	8.879E-04	1.077E-02	ZDB-GENE-040718-461
kiaa0907	-1.49875	1.994E-04	3.058E-03	ZDB-GENE-030131-405
cyhr1	-1.49853	4.515E-12	3.770E-10	ZDB-GENE-060929-720
plxnd1	-1.49653	5.067E-04	6.755E-03	ZDB-GENE-040426-1828
qkia	-1.49434	6.767E-12	5.555E-10	ZDB-GENE-990415-230
anxa11b	-1.49233	7.096E-15	9.426E-13	ZDB-GENE-030707-5
qpctla	-1.48755	1.430E-09	7.943E-08	ZDB-GENE-030131-1441
zgc:158659	-1.48624	9.799E-07	2.952E-05	ZDB-GENE-070112-492
bnip4	-1.48334	9.445E-14	1.022E-11	ZDB-GENE-051113-212
arl6ip5b	-1.47778	9.367E-13	8.774E-11	ZDB-GENE-040718-223
l3mbtl3	-1.47634	3.669E-03	3.246E-02	ZDB-GENE-040724-127
srsf1b	-1.47455	1.418E-06	4.137E-05	ZDB-GENE-040426-1467
cx35.4	-1.47408	6.800E-10	3.990E-08	ZDB-GENE-050417-174
zgc:77391	-1.47100	5.041E-08	2.111E-06	ZDB-GENE-040426-1830
amh	-1.47088	6.596E-07	2.100E-05	ZDB-GENE-050706-192
slit2	-1.46752	4.635E-10	2.870E-08	ZDB-GENE-010306-3
lrig3	-1.46703	4.866E-05	9.047E-04	ZDB-GENE-130530-866
itch	-1.46524	7.269E-06	1.744E-04	ZDB-GENE-120919-4
mob2b	-1.45687	1.494E-09	8.267E-08	ZDB-GENE-060616-319
srebf1	-1.45537	4.582E-04	6.227E-03	ZDB-GENE-090812-3
march8	-1.44943	1.064E-06	3.172E-05	ZDB-GENE-080204-8

mylipa	-1.44568	4.042E-06	1.051E-04	ZDB-GENE-030131-7831
apip	-1.44443	9.531E-08	3.665E-06	ZDB-GENE-040912-128
tcf7	-1.44387	4.219E-04	5.790E-03	ZDB-GENE-050222-4
mta3	-1.44096	1.542E-03	1.678E-02	ZDB-GENE-030131-6485
zgc:174863	-1.42702	2.674E-05	5.397E-04	ZDB-GENE-080204-87
ranbp3b	-1.42696	1.083E-03	1.252E-02	ZDB-GENE-040426-1627
xpo4	-1.42263	1.623E-06	4.629E-05	ZDB-GENE-030131-3062
gtpbp3	-1.42024	1.372E-03	1.521E-02	ZDB-GENE-050522-335
pus7	-1.41588	1.368E-04	2.213E-03	ZDB-GENE-060620-1
gsk3ab	-1.41275	6.796E-14	7.466E-12	ZDB-GENE-990714-3
camlg	-1.41155	1.438E-06	4.184E-05	ZDB-GENE-040426-2407
adam17a	-1.40923	5.505E-14	6.094E-12	ZDB-GENE-030131-2862
shrprbck1r	-1.40848	4.850E-06	1.239E-04	ZDB-GENE-080220-52
zgc:154116	-1.40809	6.480E-06	1.589E-04	ZDB-GENE-061013-84
znf503	-1.40765	1.160E-05	2.648E-04	ZDB-GENE-031113-5
psen2	-1.40572	1.480E-12	1.303E-10	ZDB-GENE-000330-9
atrn	-1.40248	5.361E-06	1.353E-04	ZDB-GENE-090312-21
dyrk1aa	-1.38959	1.254E-03	1.413E-02	ZDB-GENE-050302-29
itm2bb	-1.38928	3.099E-13	3.059E-11	ZDB-GENE-040426-2139
cxcl-c5c	-1.38116	8.514E-05	1.448E-03	ZDB-GENE-111004-2
aaas	-1.37631	1.907E-03	1.979E-02	ZDB-GENE-030912-4
lrp6	-1.37347	3.717E-04	5.173E-03	ZDB-GENE-090821-1
nipa1	-1.37242	6.044E-03	4.803E-02	ZDB-GENE-041010-164
dvl3a	-1.37033	1.784E-07	6.437E-06	ZDB-GENE-010319-1
nptna	-1.36788	5.829E-08	2.393E-06	ZDB-GENE-030804-7
gch1	-1.36680	7.588E-05	1.314E-03	ZDB-GENE-070720-5
rab1bb	-1.36547	1.686E-07	6.157E-06	ZDB-GENE-040625-133
mtrf1	-1.36477	2.565E-03	2.450E-02	ZDB-GENE-050220-8
lonrf1l	-1.35991	2.115E-06	5.842E-05	ZDB-GENE-081104-397
stx17	-1.35926	6.188E-03	4.879E-02	ZDB-GENE-041114-164
btbd2a	-1.35777	3.349E-03	3.004E-02	ZDB-GENE-030829-3
scaper	-1.35774	2.092E-03	2.124E-02	ZDB-GENE-060303-5
zbtb14	-1.35189	2.207E-04	3.325E-03	ZDB-GENE-040426-2946
scp2a	-1.35166	4.851E-05	9.029E-04	ZDB-GENE-040426-1846
spred1	-1.34985	1.700E-05	3.626E-04	ZDB-GENE-040426-2338
b2m	-1.34298	6.283E-08	2.543E-06	ZDB-GENE-980526-88
tjp2b	-1.33318	5.743E-06	1.437E-04	ZDB-GENE-040718-58
si:dkey-25e12.3	-1.32877	4.140E-03	3.548E-02	ZDB-GENE-041001-132
sesn1	-1.32607	3.751E-14	4.350E-12	ZDB-GENE-040718-408
sesn2	-1.32442	2.278E-05	4.675E-04	ZDB-GENE-070108-1
rngtt	-1.32330	1.608E-06	4.608E-05	ZDB-GENE-040426-2087
cyp3c1	-1.32220	1.933E-11	1.494E-09	ZDB-GENE-030131-3060
zgc:56409	-1.31641	1.197E-04	1.966E-03	ZDB-GENE-040426-1019
sqstm1	-1.30929	9.640E-15	1.236E-12	ZDB-GENE-040426-2204
bag2	-1.30765	9.400E-09	4.578E-07	ZDB-GENE-040426-1399
nit1	-1.30716	3.925E-03	3.414E-02	ZDB-GENE-040912-65
khdrbs1a	-1.30714	2.224E-04	3.347E-03	ZDB-GENE-000210-25
gbp	-1.30560	1.540E-05	3.376E-04	ZDB-GENE-990714-7

mapkapk3	-1.30422	1.292E-04	2.098E-03	ZDB-GENE-061215-124
cremb	-1.29944	6.331E-07	2.033E-05	ZDB-GENE-050417-150
wdsub1	-1.29037	3.145E-03	2.868E-02	ZDB-GENE-061110-97
tchp	-1.28519	1.172E-05	2.668E-04	ZDB-GENE-060421-3368
aebp2	-1.27785	2.262E-04	3.393E-03	ZDB-GENE-040426-2885
spsb3a	-1.27741	5.724E-04	7.505E-03	ZDB-GENE-030131-4182
zgc:153383	-1.27692	2.603E-03	2.481E-02	ZDB-GENE-060825-178
itsn2a	-1.27661	7.376E-04	9.181E-03	ZDB-GENE-041014-84
skor1a	-1.27399	1.739E-07	6.306E-06	ZDB-GENE-080722-5
phlda1	-1.27163	1.917E-04	2.958E-03	ZDB-GENE-041010-103
cox4i2	-1.26993	5.837E-10	3.453E-08	ZDB-GENE-040426-1775
ispd	-1.26837	5.906E-03	4.736E-02	ZDB-GENE-061110-16
slc7a6	-1.26794	1.292E-07	4.817E-06	ZDB-GENE-000607-21
LOC100002393	-1.26720	6.542E-04	8.312E-03	ZDB-GENE-091027-2
galnt2	-1.26165	9.016E-08	3.485E-06	ZDB-GENE-041111-110
lmo4a	-1.26159	9.437E-13	8.783E-11	ZDB-GENE-010702-1
bysl	-1.26139	7.933E-08	3.108E-06	ZDB-GENE-040426-1287
vrk2	-1.25873	2.293E-03	2.262E-02	ZDB-GENE-040426-1046
minal	-1.25583	6.012E-04	7.816E-03	ZDB-GENE-040426-1283
nkd3	-1.25467	1.208E-04	1.981E-03	ZDB-GENE-060503-63
map2k1	-1.25279	2.707E-08	1.195E-06	ZDB-GENE-040426-2759
cry1a	-1.24955	3.631E-06	9.543E-05	ZDB-GENE-010426-2
si:dkey-188i13.11	-1.24462	2.909E-05	5.815E-04	ZDB-GENE-090313-209
zgc:136895	-1.24075	3.227E-03	2.932E-02	ZDB-GENE-060421-4694
per2	-1.23977	1.036E-04	1.733E-03	ZDB-GENE-011220-2
plcg1	-1.23867	6.930E-09	3.480E-07	ZDB-GENE-030421-3
btbd6a	-1.23673	1.351E-05	3.016E-04	ZDB-GENE-080218-26
nupr1	-1.23368	2.188E-07	7.741E-06	ZDB-GENE-030131-4653
ccdc90b	-1.23321	6.725E-04	8.506E-03	ZDB-GENE-040724-162
syncripl	-1.23226	1.251E-05	2.826E-04	ZDB-GENE-030131-3104
scrn2	-1.22903	4.527E-05	8.514E-04	ZDB-GENE-040724-75
vapa	-1.22818	1.164E-07	4.396E-06	ZDB-GENE-030131-6540
stm	-1.22789	3.879E-07	1.306E-05	ZDB-GENE-031112-4
fblim1	-1.22756	1.223E-13	1.304E-11	ZDB-GENE-061013-662
atat1	-1.22599	6.113E-03	4.836E-02	ZDB-GENE-040426-2120
ankef1a	-1.22535	1.855E-03	1.945E-02	ZDB-GENE-041001-196
aamp	-1.22367	2.564E-04	3.751E-03	ZDB-GENE-040426-2370
tada1	-1.21918	6.069E-03	4.813E-02	ZDB-GENE-040426-58
dnajc28	-1.21898	4.095E-03	3.524E-02	ZDB-GENE-050417-130
nfasca	-1.21678	6.215E-03	4.893E-02	ZDB-GENE-080229-6
wdr26b	-1.21531	1.420E-07	5.240E-06	ZDB-GENE-041014-277
lix1l	-1.21112	3.992E-05	7.695E-04	ZDB-GENE-060503-777
galk2	-1.21082	9.403E-04	1.127E-02	ZDB-GENE-041114-143
lrrc15	-1.20113	3.188E-10	2.052E-08	ZDB-GENE-031113-8
msrb2	-1.20057	2.124E-03	2.145E-02	ZDB-GENE-040426-2304
tp53bp2a	-1.19751	8.249E-07	2.545E-05	ZDB-GENE-040516-8
sema6d	-1.19580	1.340E-05	2.999E-04	ZDB-GENE-040426-1933
utp18	-1.19043	2.507E-04	3.679E-03	ZDB-GENE-060929-668

srebf2	-1.18980	7.642E-08	3.018E-06	ZDB-GENE-070410-8
dlst	-1.18679	8.831E-07	2.705E-05	ZDB-GENE-030326-1
tnfsf10l3	-1.18602	7.106E-05	1.239E-03	ZDB-GENE-070606-1
qdpra	-1.17590	5.042E-05	9.326E-04	ZDB-GENE-070705-197
cep97	-1.17497	6.579E-05	1.159E-03	ZDB-GENE-031030-11
flt1	-1.17085	4.776E-11	3.455E-09	ZDB-GENE-050407-1
heca	-1.16909	4.363E-04	5.964E-03	ZDB-GENE-041001-173
fam126a	-1.16777	2.205E-05	4.557E-04	ZDB-GENE-040426-1806
gpcpd1	-1.16335	4.826E-05	8.995E-04	ZDB-GENE-060503-472
wbp1la	-1.16281	5.528E-06	1.390E-04	ZDB-GENE-030131-1961
st6gal2a	-1.16095	1.185E-04	1.950E-03	ZDB-GENE-060322-4
naa30	-1.15897	7.010E-07	2.212E-05	ZDB-GENE-081022-73
mea1	-1.15598	4.813E-06	1.232E-04	ZDB-GENE-070424-117
prkar2aa	-1.15067	1.346E-07	4.991E-06	ZDB-GENE-030131-1007
fam110b	-1.14971	4.046E-03	3.488E-02	ZDB-GENE-050626-70
arid1aa	-1.14770	1.268E-03	1.427E-02	ZDB-GENE-080220-35
irf9	-1.14102	4.582E-07	1.508E-05	ZDB-GENE-040912-165
chmp1a	-1.13783	3.039E-05	6.025E-04	ZDB-GENE-040426-1474
unc5c	-1.13679	1.083E-04	1.802E-03	ZDB-GENE-030131-9196
brap	-1.13664	5.219E-04	6.924E-03	ZDB-GENE-040718-168
brd4	-1.13580	9.039E-07	2.763E-05	ZDB-GENE-030131-267
cutc	-1.13559	1.566E-05	3.405E-04	ZDB-GENE-041010-144
zgc:175248	-1.13375	2.483E-04	3.657E-03	ZDB-GENE-080204-93
appa	-1.13319	2.746E-06	7.444E-05	ZDB-GENE-000616-13
cthrc1a	-1.13204	2.753E-08	1.212E-06	ZDB-GENE-040808-22
si:dkey-97m3.1	-1.13153	1.229E-03	1.389E-02	ZDB-GENE-060503-367
hoxb2a	-1.12954	4.296E-05	8.164E-04	ZDB-GENE-990415-103
numb	-1.12765	9.171E-06	2.162E-04	ZDB-GENE-060422-1
zgc:101562	-1.12756	4.390E-03	3.721E-02	ZDB-GENE-041010-74
ms4a17a.11	-1.12742	7.632E-09	3.755E-07	ZDB-GENE-050417-157
ankrd13c	-1.12580	1.539E-04	2.453E-03	ZDB-GENE-030131-3892
rilpl1	-1.12515	9.445E-04	1.131E-02	ZDB-GENE-061201-51
vim	-1.12296	1.005E-05	2.339E-04	ZDB-GENE-011212-3
ccdc3	-1.11961	3.273E-03	2.958E-02	ZDB-GENE-041210-326
keap1b	-1.11926	1.968E-03	2.022E-02	ZDB-GENE-080508-1
lsm14aa	-1.11729	1.366E-03	1.517E-02	ZDB-GENE-040426-1567
mtmr7b	-1.11682	1.308E-06	3.845E-05	ZDB-GENE-050913-66
cdkal1	-1.11288	3.843E-03	3.356E-02	ZDB-GENE-040426-1443
cmbl	-1.10820	1.576E-03	1.699E-02	ZDB-GENE-071004-21
reps1	-1.10777	1.430E-05	3.166E-04	ZDB-GENE-030131-9536
map7d1b	-1.10061	3.421E-07	1.162E-05	ZDB-GENE-080215-23
parp3	-1.09944	1.763E-05	3.734E-04	ZDB-GENE-030131-7435
zgc:64106	-1.09788	9.347E-07	2.834E-05	ZDB-GENE-040426-1370
gnsb	-1.09377	2.267E-03	2.242E-02	ZDB-GENE-030131-5846
tmed7	-1.09302	2.078E-08	9.517E-07	ZDB-GENE-040426-2570
trim44	-1.09264	4.671E-03	3.914E-02	ZDB-GENE-050522-37
si:ch211-199o1.2	-1.09257	5.379E-06	1.355E-04	ZDB-GENE-060503-867
zgc:85936	-1.09251	1.183E-04	1.950E-03	ZDB-GENE-040426-2450

slc26a11	-1.08906	2.930E-06	7.870E-05	ZDB-GENE-030131-5349
si:ch211-234h8.7	-1.08762	2.065E-04	3.153E-03	ZDB-GENE-061009-16
wif1	-1.08262	1.852E-06	5.204E-05	ZDB-GENE-990712-17
foxa1	-1.08215	6.729E-04	8.506E-03	ZDB-GENE-990415-78
apoc1l	-1.07797	6.087E-06	1.510E-04	ZDB-GENE-030131-1074
fam129ba	-1.07630	1.372E-07	5.074E-06	ZDB-GENE-070112-1222
nr1d2a	-1.07609	1.961E-04	3.016E-03	ZDB-GENE-040504-1
tmem88a	-1.07334	4.237E-09	2.235E-07	ZDB-GENE-040426-1808
ucp2	-1.07293	1.171E-09	6.629E-08	ZDB-GENE-990708-8
socs3a	-1.06762	3.395E-06	8.988E-05	ZDB-GENE-030131-7349
npr3	-1.06690	6.246E-07	2.014E-05	ZDB-GENE-060531-91
atp13a	-1.06565	1.107E-07	4.192E-06	ZDB-GENE-040426-2804
foxj1a	-1.06203	2.962E-03	2.746E-02	ZDB-GENE-060929-1178
rab12	-1.06108	3.181E-05	6.272E-04	ZDB-GENE-050522-555
zdhhc5a	-1.06102	3.870E-05	7.481E-04	ZDB-GENE-090312-92
adrm1	-1.05781	4.843E-04	6.516E-03	ZDB-GENE-040426-905
smek2	-1.05605	1.825E-04	2.838E-03	ZDB-GENE-030131-8142
ppap2c	-1.05566	3.561E-06	9.376E-05	ZDB-GENE-030131-7425
fam192a	-1.05390	9.743E-05	1.640E-03	ZDB-GENE-030131-6692
zgc:175284	-1.05282	1.523E-03	1.660E-02	ZDB-GENE-080219-49
ifi30	-1.05041	2.638E-07	9.155E-06	ZDB-GENE-030131-8447
zdhhc14	-1.05005	4.917E-04	6.585E-03	ZDB-GENE-040724-21
zgc:63907	-1.04962	2.417E-05	4.925E-04	ZDB-GENE-040426-1292
csnk1g1	-1.04542	7.228E-04	9.035E-03	ZDB-GENE-041212-63
h1fx	-1.04488	1.230E-08	5.872E-07	ZDB-GENE-030131-1228
zgc:174888	-1.04444	8.937E-09	4.382E-07	ZDB-GENE-030328-33
il1rapl1a	-1.04434	2.319E-05	4.752E-04	ZDB-GENE-061013-249
helz	-1.04096	3.653E-04	5.107E-03	ZDB-GENE-040426-2419
ddit3	-1.04061	2.378E-08	1.063E-06	ZDB-GENE-070410-90
krt8	-1.04050	9.085E-06	2.148E-04	ZDB-GENE-030411-5
atxn2	-1.03441	1.082E-03	1.251E-02	ZDB-GENE-060526-217
atf3	-1.03422	9.665E-08	3.697E-06	ZDB-GENE-040426-728
zgc:123258	-1.03405	4.739E-03	3.955E-02	ZDB-GENE-051120-84
zgc:136439	-1.03266	5.919E-08	2.423E-06	ZDB-GENE-060421-2858
ehd2b	-1.03203	5.745E-08	2.371E-06	ZDB-GENE-030131-9347
khdrbs1b	-1.02838	2.232E-04	3.356E-03	ZDB-GENE-040426-2344
clybl	-1.02212	4.752E-05	8.891E-04	ZDB-GENE-030516-6
cyp2ad3	-1.02081	5.053E-05	9.335E-04	ZDB-GENE-041001-156
adipor2	-1.02010	6.712E-08	2.694E-06	ZDB-GENE-041210-60
terfa	-1.01762	3.122E-03	2.856E-02	ZDB-GENE-020419-38
si:dkey-69o16.5	-1.01596	4.661E-05	8.731E-04	ZDB-GENE-081104-432
mid1ip1b	-1.01339	2.829E-05	5.678E-04	ZDB-GENE-040426-1720
lxn	-1.01289	1.607E-05	3.474E-04	ZDB-GENE-041212-18
tle2	-1.01036	4.378E-04	5.979E-03	ZDB-GENE-030131-2826
ythdf1	-1.00984	1.464E-03	1.607E-02	ZDB-GENE-040426-2786
entpd8	-1.00538	5.210E-05	9.576E-04	ZDB-GENE-040724-142
ammecr1	-1.00437	6.935E-06	1.686E-04	ZDB-GENE-040426-1533
pmaip1	-1.00366	6.269E-04	8.078E-03	ZDB-GENE-070119-3

si:ch211-71m22.1	-1.00334	9.753E-08	3.721E-06	ZDB-GENE-070705-193
calcoco1a	-1.00207	1.361E-03	1.513E-02	ZDB-GENE-030131-6911
ccdc104	-1.00190	1.824E-03	1.919E-02	ZDB-GENE-060421-8162
mst1	-1.00156	4.010E-05	7.719E-04	ZDB-GENE-020806-3
prrx1a	-1.00111	2.610E-05	5.275E-04	ZDB-GENE-020905-5
col18a1	-0.99990	4.000E-07	1.344E-05	ZDB-GENE-030516-3
setd8a	-0.99988	3.025E-04	4.316E-03	ZDB-GENE-060825-37
ralbp1	-0.99979	1.281E-05	2.879E-04	ZDB-GENE-040426-2098
csnk1e	-0.99889	8.308E-05	1.422E-03	ZDB-GENE-030131-7873
rap2b	-0.99631	3.773E-04	5.240E-03	ZDB-GENE-040426-2818
zgc:92140	-0.99597	1.813E-06	5.105E-05	ZDB-GENE-040912-161
zgc:158364	-0.99498	1.152E-03	1.317E-02	ZDB-GENE-090609-1
slc1a4	-0.99363	6.271E-03	4.926E-02	ZDB-GENE-030616-566
bcl7ba	-0.99337	1.615E-06	4.619E-05	ZDB-GENE-040426-2193
zmiz1a	-0.99125	8.320E-04	1.017E-02	ZDB-GENE-080220-27
anxa13l	-0.99067	3.516E-06	9.285E-05	ZDB-GENE-050522-310
mylipb	-0.98862	4.309E-06	1.111E-04	ZDB-GENE-061027-67
snap29	-0.98855	3.466E-04	4.870E-03	ZDB-GENE-041111-226
ppm1db	-0.98854	2.794E-04	4.034E-03	ZDB-GENE-041114-27
rbbp5	-0.98842	3.009E-03	2.770E-02	ZDB-GENE-040426-886
meis4.1a	-0.98557	3.807E-04	5.268E-03	ZDB-GENE-020122-3
gbx2	-0.98520	1.512E-08	7.083E-07	ZDB-GENE-020509-2
tmem234	-0.98238	2.859E-05	5.730E-04	ZDB-GENE-090313-326
histh1l	-0.98130	1.757E-05	3.732E-04	ZDB-GENE-050417-145
npdc1a	-0.98049	1.920E-05	4.031E-04	ZDB-GENE-031204-31
tpbgl	-0.97953	4.398E-03	3.725E-02	ZDB-GENE-030827-4
akt2l	-0.97643	4.101E-07	1.365E-05	ZDB-GENE-040121-5
cdc37	-0.97575	6.437E-07	2.054E-05	ZDB-GENE-040426-940
ankrd39	-0.97470	5.501E-03	4.478E-02	ZDB-GENE-081104-153
lyrm5b	-0.97452	1.551E-04	2.466E-03	ZDB-GENE-070424-20
arfrp1	-0.97420	2.362E-03	2.304E-02	ZDB-GENE-050227-14
rnft1	-0.97248	4.339E-07	1.438E-05	ZDB-GENE-040426-2324
mapk11	-0.97156	3.860E-05	7.471E-04	ZDB-GENE-040625-75
slc18a2	-0.96811	8.811E-05	1.497E-03	ZDB-GENE-080514-1
zgc:101744	-0.96328	4.793E-04	6.461E-03	ZDB-GENE-050320-111
foxc1b	-0.95978	7.633E-04	9.468E-03	ZDB-GENE-010302-2
ptk2aa	-0.95927	8.965E-05	1.516E-03	ZDB-GENE-031114-5
bcl6a	-0.95881	8.918E-04	1.080E-02	ZDB-GENE-040426-1696
tox2	-0.95851	1.729E-03	1.843E-02	ZDB-GENE-040426-2095
zgc:113363	-0.95829	3.368E-05	6.633E-04	ZDB-GENE-050327-7
crtac1a	-0.95823	3.669E-03	3.246E-02	ZDB-GENE-030131-9704
s100v2	-0.95669	1.517E-04	2.422E-03	ZDB-GENE-030131-8909
cyp2v1	-0.95216	1.352E-05	3.016E-04	ZDB-GENE-040912-139
pdpk1b	-0.95198	9.651E-07	2.914E-05	ZDB-GENE-040426-1820
mettl9	-0.95017	5.905E-05	1.065E-03	ZDB-GENE-061013-752
si:ch211-51e12.7	-0.94997	2.595E-04	3.791E-03	ZDB-GENE-030131-8279
clul1	-0.94950	2.588E-06	7.029E-05	ZDB-GENE-031118-57
wasla	-0.94672	3.574E-05	6.991E-04	ZDB-GENE-070209-220

ablim1a	-0.94612	3.166E-05	6.251E-04	ZDB-GENE-080219-41
cfp	-0.94372	2.533E-03	2.428E-02	ZDB-GENE-030131-9247
ssh2b	-0.94295	5.322E-05	9.753E-04	ZDB-GENE-081205-4
si:dkey-204f11.64	-0.94061	1.528E-06	4.420E-05	ZDB-GENE-030131-8332
pgrmc2	-0.93477	7.873E-05	1.360E-03	ZDB-GENE-040426-2102
abhd2a	-0.93444	6.457E-05	1.148E-03	ZDB-GENE-040426-784
rnf11a	-0.93373	2.299E-03	2.265E-02	ZDB-GENE-040426-1277
ldlrad3	-0.93033	1.556E-04	2.471E-03	ZDB-GENE-030131-1533
foxi2	-0.92915	7.891E-05	1.361E-03	ZDB-GENE-031126-2
zgc:113314	-0.92819	3.300E-03	2.971E-02	ZDB-GENE-050913-19
sacm1la	-0.92816	9.665E-04	1.151E-02	ZDB-GENE-060503-122
srsf6b	-0.92698	2.307E-08	1.041E-06	ZDB-GENE-041219-1
slc22a7b	-0.92692	1.332E-05	2.986E-04	ZDB-GENE-040426-1311
dhrs1	-0.92688	2.518E-04	3.690E-03	ZDB-GENE-030616-591
spast	-0.92628	2.091E-04	3.177E-03	ZDB-GENE-040426-2331
serpina10a	-0.92535	1.696E-04	2.662E-03	ZDB-GENE-041014-246
rab32a	-0.92452	1.832E-03	1.924E-02	ZDB-GENE-031006-10
arsk	-0.92056	2.382E-03	2.316E-02	ZDB-GENE-060929-332
cers5	-0.91823	1.086E-06	3.233E-05	ZDB-GENE-030131-1740
aamdc	-0.91718	6.565E-05	1.159E-03	ZDB-GENE-050522-64
calm1a	-0.91602	6.876E-08	2.745E-06	ZDB-GENE-030131-8308
arih1	-0.91212	3.617E-03	3.205E-02	ZDB-GENE-030131-5213
zgc:154075	-0.90919	8.922E-05	1.511E-03	ZDB-GENE-060929-910
mpv17l2	-0.90679	3.368E-03	3.015E-02	ZDB-GENE-040718-306
snx11	-0.90620	1.922E-07	6.882E-06	ZDB-GENE-041111-253
furina	-0.90610	2.823E-03	2.644E-02	ZDB-GENE-040901-1
jak2a	-0.90442	1.944E-03	2.005E-02	ZDB-GENE-980526-481
optn	-0.90283	6.755E-04	8.531E-03	ZDB-GENE-030131-8103
ism1	-0.90063	2.599E-04	3.794E-03	ZDB-GENE-050523-3
abcg2a	-0.90063	9.306E-07	2.827E-05	ZDB-GENE-050517-35
cx30.3	-0.90027	2.588E-07	9.068E-06	ZDB-GENE-040406-1
ivns1abpa	-0.89879	2.819E-03	2.642E-02	ZDB-GENE-031222-2
rbm4.2	-0.89875	1.001E-04	1.679E-03	ZDB-GENE-030131-3019
skia	-0.89738	8.840E-04	1.073E-02	ZDB-GENE-990715-9
lbx2	-0.89554	2.424E-03	2.344E-02	ZDB-GENE-001206-2
galm	-0.89308	2.118E-04	3.213E-03	ZDB-GENE-040718-66
fbxl2	-0.89182	2.450E-03	2.366E-02	ZDB-GENE-030925-12
irak4	-0.89172	1.405E-04	2.259E-03	ZDB-GENE-040426-738
si:ch211-239i4.3	-0.88936	2.320E-04	3.463E-03	ZDB-GENE-081107-30
smarcd1	-0.88660	5.500E-05	1.002E-03	ZDB-GENE-030131-1835
nfybb	-0.88530	2.478E-04	3.657E-03	ZDB-GENE-050306-25
zgc:154093	-0.88381	2.974E-05	5.929E-04	ZDB-GENE-061103-517
tmem66	-0.88332	2.023E-04	3.099E-03	ZDB-GENE-060929-208
zgc:110340	-0.88173	1.027E-07	3.897E-06	ZDB-GENE-050320-36
hapln3	-0.88075	7.319E-08	2.906E-06	ZDB-GENE-040426-2089
pvrl3b	-0.88026	7.033E-06	1.704E-04	ZDB-GENE-050327-14
blzf1	-0.87958	1.345E-04	2.179E-03	ZDB-GENE-010724-1
rnf103	-0.87847	2.499E-04	3.674E-03	ZDB-GENE-040912-171

casp9	-0.87795	5.523E-04	7.264E-03	ZDB-GENE-030825-5
prnprs3	-0.87478	6.369E-07	2.041E-05	ZDB-GENE-041221-3
mef2aa	-0.87441	4.432E-06	1.140E-04	ZDB-GENE-990415-163
thap4	-0.87067	6.817E-04	8.602E-03	ZDB-GENE-070912-5
serpinb1	-0.86917	1.709E-04	2.670E-03	ZDB-GENE-040718-400
midn	-0.86844	1.270E-07	4.748E-06	ZDB-GENE-030131-2433
sat1b	-0.86679	5.752E-03	4.646E-02	ZDB-GENE-041001-39
ifrd1	-0.86614	7.067E-07	2.225E-05	ZDB-GENE-030131-6132
arid3b	-0.86611	2.353E-07	8.284E-06	ZDB-GENE-070209-182
nfe2l2a	-0.86459	1.109E-05	2.552E-04	ZDB-GENE-030723-2
s100t	-0.86452	2.134E-03	2.149E-02	ZDB-GENE-050320-61
tnfrsf1a	-0.86314	1.231E-04	2.012E-03	ZDB-GENE-040426-2252
zgc:92873	-0.86248	1.845E-04	2.862E-03	ZDB-GENE-040718-182
snx25	-0.86198	1.855E-04	2.874E-03	ZDB-GENE-121105-11
kcnk6	-0.86061	8.997E-04	1.088E-02	ZDB-GENE-050309-213
ccdc102a	-0.86056	6.339E-04	8.132E-03	ZDB-GENE-030131-4044
zfyve27	-0.85763	5.822E-05	1.053E-03	ZDB-GENE-061013-502
si:ch211-214p16.2	-0.85646	1.647E-05	3.545E-04	ZDB-GENE-030616-607
pxmp2	-0.85607	2.274E-04	3.404E-03	ZDB-GENE-040912-184
zgc:92052	-0.85457	7.922E-05	1.365E-03	ZDB-GENE-040718-81
bnip3lb	-0.85399	5.409E-05	9.867E-04	ZDB-GENE-040325-1
nenf	-0.85371	6.266E-06	1.544E-04	ZDB-GENE-050320-129
si:dkey-206d17.12	-0.85343	6.062E-04	7.859E-03	ZDB-GENE-081031-82
mthfd1l	-0.85258	1.673E-03	1.792E-02	ZDB-GENE-041001-133
mfge8b	-0.85225	4.982E-05	9.250E-04	ZDB-GENE-070209-137
smarce1	-0.84883	4.232E-04	5.801E-03	ZDB-GENE-030131-967
rdh10a	-0.84835	1.546E-03	1.680E-02	ZDB-GENE-070112-2242
tmem120a	-0.84102	5.390E-04	7.121E-03	ZDB-GENE-070424-21
slc26a5	-0.83932	1.009E-05	2.345E-04	ZDB-GENE-030131-1566
adprh	-0.83929	1.974E-03	2.025E-02	ZDB-GENE-040801-190
ubl7a	-0.83603	5.727E-04	7.505E-03	ZDB-GENE-050417-285
atf4b2	-0.83524	5.637E-06	1.415E-04	ZDB-GENE-070928-23
pon2	-0.83487	1.766E-04	2.751E-03	ZDB-GENE-030131-7116
pfkfb1	-0.83361	1.862E-03	1.947E-02	ZDB-GENE-030131-5664
fam101a	-0.83304	5.369E-03	4.393E-02	ZDB-GENE-030616-80
si:dkey-28b4.7	-0.83094	1.546E-05	3.376E-04	ZDB-GENE-050506-118
zbtb37	-0.83043	4.888E-03	4.060E-02	ZDB-GENE-050208-621
ewsr1a	-0.82836	3.515E-04	4.924E-03	ZDB-GENE-030131-2317
srsf2a	-0.82812	6.997E-04	8.775E-03	ZDB-GENE-040426-2706
mon1a	-0.82787	5.040E-03	4.164E-02	ZDB-GENE-070615-6
sfpq	-0.82744	1.655E-05	3.546E-04	ZDB-GENE-040426-2452
tob1a	-0.82534	7.155E-06	1.722E-04	ZDB-GENE-031030-4
slc25a44b	-0.82350	6.915E-04	8.696E-03	ZDB-GENE-050320-89
zgc:77849	-0.82305	5.533E-07	1.797E-05	ZDB-GENE-030131-1107
si:ch211-105j21.7	-0.82287	4.062E-03	3.500E-02	ZDB-GENE-110408-16
ppp2r2d	-0.82262	1.126E-03	1.293E-02	ZDB-GENE-040426-2086
stub1	-0.82078	1.644E-04	2.585E-03	ZDB-GENE-030131-2963
zgc:112432	-0.82057	3.369E-04	4.761E-03	ZDB-GENE-050417-304

flrt2	-0.81827	2.469E-03	2.377E-02	ZDB-GENE-070705-267
tmem59l	-0.81814	9.947E-05	1.673E-03	ZDB-GENE-040426-2651
map1lc3a	-0.81786	3.467E-05	6.800E-04	ZDB-GENE-040426-2232
si:dkey-286j15.3	-0.81782	4.134E-03	3.545E-02	ZDB-GENE-060503-635
rps2	-0.81740	7.004E-04	8.777E-03	ZDB-GENE-040426-2454
prrg4	-0.81641	6.349E-04	8.138E-03	ZDB-GENE-050417-112
ephb4a	-0.81364	2.334E-04	3.476E-03	ZDB-GENE-990415-62
otud5b	-0.81351	1.601E-03	1.721E-02	ZDB-GENE-040801-50
lef1	-0.81302	5.367E-05	9.816E-04	ZDB-GENE-990714-26
sema5a	-0.81262	3.809E-03	3.335E-02	ZDB-GENE-030131-8969
creb3l3l	-0.81257	2.493E-05	5.066E-04	ZDB-GENE-040426-2942
pik3ip1	-0.81049	5.944E-04	7.734E-03	ZDB-GENE-031030-14
optc	-0.80992	5.860E-04	7.646E-03	ZDB-GENE-030131-1798
dirc2	-0.80830	3.394E-05	6.674E-04	ZDB-GENE-040912-167
ctnnb1	-0.80732	3.688E-03	3.256E-02	ZDB-GENE-980526-362
sult2st1	-0.80176	8.419E-06	2.000E-04	ZDB-GENE-030219-114
siae	-0.80118	2.196E-03	2.196E-02	ZDB-GENE-040426-1032
rab40b	-0.80104	1.441E-03	1.588E-02	ZDB-GENE-040718-257
ube2h	-0.80102	1.113E-04	1.847E-03	ZDB-GENE-030616-67
dgkaa	-0.79841	1.156E-04	1.908E-03	ZDB-GENE-060616-305
taz	-0.79813	2.705E-03	2.553E-02	ZDB-GENE-030131-684
hps1	-0.79426	2.747E-03	2.588E-02	ZDB-GENE-051113-152
pi4k2a	-0.79384	3.043E-03	2.792E-02	ZDB-GENE-040426-2675
proza	-0.79318	1.277E-03	1.434E-02	ZDB-GENE-061013-403
wnt5b	-0.79298	4.629E-03	3.885E-02	ZDB-GENE-980526-87
arhgdig	-0.79260	9.953E-07	2.988E-05	ZDB-GENE-040426-1493
mcm4	-0.79191	6.887E-05	1.205E-03	ZDB-GENE-030131-9544
fzd6	-0.79074	9.998E-04	1.180E-02	ZDB-GENE-040426-1451
fbxl4	-0.79041	3.485E-03	3.109E-02	ZDB-GENE-041114-187
sav1	-0.78956	4.540E-05	8.527E-04	ZDB-GENE-040912-28
rnf8	-0.78952	3.020E-03	2.773E-02	ZDB-GENE-040426-849
nlk1	-0.78910	1.588E-03	1.712E-02	ZDB-GENE-040701-1
prox1b	-0.78731	2.357E-05	4.816E-04	ZDB-GENE-050107-3
si:ch73-142c19.1	-0.78658	1.522E-03	1.659E-02	ZDB-GENE-090311-15
ampd3b	-0.78488	2.763E-03	2.600E-02	ZDB-GENE-030131-5929
lmcd1	-0.78275	1.958E-03	2.018E-02	ZDB-GENE-040426-1067
ctsz	-0.78099	2.211E-05	4.564E-04	ZDB-GENE-041010-139
slc39a1	-0.78038	2.874E-03	2.680E-02	ZDB-GENE-030131-9917
zgc:77304	-0.77958	2.265E-03	2.242E-02	ZDB-GENE-030131-6724
agbl5	-0.77919	6.834E-04	8.616E-03	ZDB-GENE-040822-29
ms4a17a.9	-0.77838	1.591E-05	3.450E-04	ZDB-GENE-050320-79
col1a2	-0.77480	1.070E-03	1.242E-02	ZDB-GENE-030131-8415
smap1	-0.77418	1.040E-05	2.409E-04	ZDB-GENE-060920-2
kctd12.2	-0.77393	1.652E-05	3.546E-04	ZDB-GENE-060117-4
zgc:92360	-0.77292	1.861E-03	1.947E-02	ZDB-GENE-040718-470
pla2g7	-0.77058	1.721E-03	1.835E-02	ZDB-GENE-040426-2248
ppp1r3b	-0.76858	1.554E-03	1.685E-02	ZDB-GENE-030131-5496
zgc:113263	-0.76708	2.339E-03	2.292E-02	ZDB-GENE-050306-34

arrdc2	-0.76699	3.693E-05	7.175E-04	ZDB-GENE-030131-7146
si:dkeyp-69e1.8	-0.76632	2.970E-03	2.749E-02	ZDB-GENE-070912-676
lemd3	-0.76576	2.361E-04	3.514E-03	ZDB-GENE-030131-490
spata5l1	-0.76459	5.711E-03	4.618E-02	ZDB-GENE-060929-204
dlat	-0.76399	1.575E-04	2.499E-03	ZDB-GENE-030131-2921
cry4	-0.76378	2.749E-04	3.978E-03	ZDB-GENE-010426-7
glyctk	-0.76299	3.241E-03	2.938E-02	ZDB-GENE-061013-308
akap8l	-0.76226	8.134E-04	9.988E-03	ZDB-GENE-040426-2303
mto1	-0.76213	1.553E-03	1.685E-02	ZDB-GENE-070209-253
atf6b	-0.76031	1.298E-03	1.455E-02	ZDB-GENE-061103-166
gstt1b	-0.75924	2.273E-04	3.404E-03	ZDB-GENE-040426-1491
si:dkey-15j16.2	-0.75899	5.704E-05	1.034E-03	ZDB-GENE-030131-5900
eif4ebp1	-0.75642	6.175E-06	1.527E-04	ZDB-GENE-030131-2332
tbrg1	-0.75529	3.352E-03	3.005E-02	ZDB-GENE-080220-62
si:ch211-175l6.9	-0.75064	4.087E-03	3.520E-02	ZDB-GENE-060503-768
acy1	-0.74989	3.250E-03	2.942E-02	ZDB-GENE-040426-755
dclre1c	-0.74958	1.938E-03	2.001E-02	ZDB-GENE-041210-72
ewsr1b	-0.74700	5.157E-03	4.253E-02	ZDB-GENE-030131-1600
fastk	-0.74426	1.715E-03	1.830E-02	ZDB-GENE-070912-680
spred2b	-0.74312	9.829E-04	1.166E-02	ZDB-GENE-040426-2342
zgc:162544	-0.74277	4.091E-04	5.645E-03	ZDB-GENE-030131-9020
capn5a	-0.74139	2.066E-03	2.104E-02	ZDB-GENE-061103-214
ryk	-0.74124	2.825E-03	2.644E-02	ZDB-GENE-070209-277
ebag9	-0.74094	2.468E-04	3.646E-03	ZDB-GENE-040426-1093
plekhb2	-0.74054	1.489E-04	2.386E-03	ZDB-GENE-040426-754
zgc:92066	-0.74016	1.020E-03	1.197E-02	ZDB-GENE-040718-72
unc119.1	-0.73808	3.794E-04	5.260E-03	ZDB-GENE-030219-130
smad9	-0.73560	3.720E-03	3.273E-02	ZDB-GENE-031014-1
dact2	-0.73233	6.949E-05	1.213E-03	ZDB-GENE-030131-9975
serinc1	-0.73138	1.860E-05	3.917E-04	ZDB-GENE-030131-4415
edn1	-0.73092	1.078E-03	1.249E-02	ZDB-GENE-000920-1
nrp1b	-0.72864	2.081E-05	4.337E-04	ZDB-GENE-040611-1
wwc3	-0.72843	5.492E-04	7.231E-03	ZDB-GENE-070209-229
zgc:112315	-0.72834	1.925E-03	1.993E-02	ZDB-GENE-050522-174
mospd1	-0.72535	8.020E-04	9.873E-03	ZDB-GENE-040426-2691
ptprma	-0.72502	1.496E-03	1.637E-02	ZDB-GENE-031113-24
zgc:113121	-0.72498	2.024E-03	2.069E-02	ZDB-GENE-050227-20
chmp2bb	-0.72434	1.414E-04	2.271E-03	ZDB-GENE-040426-2539
exosc3	-0.72335	5.038E-04	6.730E-03	ZDB-GENE-050706-140
grasp	-0.72307	8.125E-05	1.393E-03	ZDB-GENE-060312-30
nid1b	-0.72302	1.642E-04	2.585E-03	ZDB-GENE-070802-3
id1	-0.72277	6.357E-04	8.141E-03	ZDB-GENE-990415-96
dcaf5	-0.72185	1.641E-03	1.761E-02	ZDB-GENE-060929-808
pik3c3	-0.72165	2.904E-04	4.172E-03	ZDB-GENE-050410-13
iars	-0.71850	2.814E-04	4.059E-03	ZDB-GENE-030131-6325
si:dkey-260j18.2	-0.71725	3.719E-03	3.273E-02	ZDB-GENE-030131-3956
meox1	-0.71602	6.038E-04	7.843E-03	ZDB-GENE-040718-149
zgc:158328	-0.71499	4.441E-03	3.747E-02	ZDB-GENE-070112-1202

tmem198b	-0.71419	1.156E-03	1.319E-02	ZDB-GENE-050809-123
snx27a	-0.71337	1.333E-03	1.486E-02	ZDB-GENE-060503-371
zgc:174637	-0.71315	2.178E-03	2.181E-02	ZDB-GENE-071004-70
rnf14	-0.71232	6.490E-04	8.261E-03	ZDB-GENE-040625-65
rhd	-0.71089	3.279E-03	2.959E-02	ZDB-GENE-051213-1
ubxn4	-0.71041	1.421E-03	1.568E-02	ZDB-GENE-040426-1981
pi4kb	-0.70971	3.720E-03	3.273E-02	ZDB-GENE-061013-96
il13ra1	-0.70776	3.195E-04	4.536E-03	ZDB-GENE-060503-103
zgc:92308	-0.70713	6.086E-05	1.095E-03	ZDB-GENE-040912-142
tpd52l2a	-0.70373	2.674E-03	2.531E-02	ZDB-GENE-030131-7047
pdgfab	-0.70295	2.068E-04	3.154E-03	ZDB-GENE-060929-124
znf704	-0.70242	1.793E-03	1.891E-02	ZDB-GENE-060503-775
gatad2ab	-0.70199	6.354E-03	4.973E-02	ZDB-GENE-040426-2372
rpl19	-0.70018	2.575E-03	2.456E-02	ZDB-GENE-040426-2290
tsku	-0.69991	2.929E-03	2.717E-02	ZDB-GENE-030131-4683
tbc1d17	-0.69892	3.730E-03	3.277E-02	ZDB-GENE-050522-79
cecr1a	-0.69846	1.156E-03	1.319E-02	ZDB-GENE-030902-4
dld	-0.69622	8.300E-05	1.422E-03	ZDB-GENE-990415-47
ephb3a	-0.69617	2.327E-03	2.284E-02	ZDB-GENE-990415-60
zgc:165573	-0.69498	2.126E-03	2.145E-02	ZDB-GENE-070615-27
tango2	-0.69014	4.226E-03	3.605E-02	ZDB-GENE-040808-42
aimp2	-0.68806	1.447E-03	1.591E-02	ZDB-GENE-040426-2652
zgc:101583	-0.68759	5.063E-03	4.180E-02	ZDB-GENE-041212-75
igfbp7	-0.68660	5.106E-04	6.802E-03	ZDB-GENE-040426-2423
vps8	-0.68631	2.306E-03	2.269E-02	ZDB-GENE-060503-348
zgc:73238	-0.68328	4.690E-04	6.340E-03	ZDB-GENE-040426-1726
fbxo7	-0.68257	6.284E-03	4.929E-02	ZDB-GENE-050913-68
cald1b	-0.68046	3.946E-03	3.424E-02	ZDB-GENE-090313-229
zfpm2b	-0.68025	4.529E-03	3.812E-02	ZDB-GENE-060130-5
rnf111	-0.67945	2.854E-03	2.663E-02	ZDB-GENE-070808-2
ubl7b	-0.67865	5.911E-04	7.698E-03	ZDB-GENE-040426-1334
si:dkey-165n16.1	-0.67793	3.135E-03	2.861E-02	ZDB-GENE-100922-224
ppdpfb	-0.67577	1.715E-04	2.678E-03	ZDB-GENE-030131-8247
tmem206	-0.67563	1.598E-03	1.720E-02	ZDB-GENE-040426-1173
igf2b	-0.67428	1.701E-04	2.667E-03	ZDB-GENE-030131-2935
dkk3	-0.67426	5.201E-03	4.276E-02	ZDB-GENE-110519-1
baiap2a	-0.67402	6.956E-04	8.740E-03	ZDB-GENE-060421-2756
fuca2	-0.67387	2.607E-04	3.801E-03	ZDB-GENE-040822-39
snx22	-0.67327	2.574E-03	2.456E-02	ZDB-GENE-060825-154
zgc:110182	-0.67312	1.847E-03	1.939E-02	ZDB-GENE-050417-381
bmper	-0.67169	2.154E-03	2.166E-02	ZDB-GENE-030219-146
abi2b	-0.67159	1.503E-03	1.643E-02	ZDB-GENE-060519-5
riok3	-0.67097	1.590E-04	2.512E-03	ZDB-GENE-040801-133
dbnlb	-0.67074	3.203E-03	2.914E-02	ZDB-GENE-050522-293
dhrs3a	-0.66980	3.674E-03	3.248E-02	ZDB-GENE-040801-217
col4a5	-0.66801	3.987E-03	3.449E-02	ZDB-GENE-030131-2281
gar1	-0.66776	5.463E-03	4.455E-02	ZDB-GENE-040426-893
fam43a	-0.66770	1.896E-04	2.932E-03	ZDB-GENE-030131-6098

gro1	-0.66275	1.491E-03	1.635E-02	ZDB-GENE-990415-85
etv2	-0.66229	1.324E-03	1.477E-02	ZDB-GENE-050622-14
pon3	-0.66139	2.644E-03	2.510E-02	ZDB-GENE-050522-353
net1	-0.66050	5.471E-03	4.458E-02	ZDB-GENE-041121-18
zgc:112084	-0.65991	7.325E-04	9.141E-03	ZDB-GENE-050417-283
tmem180	-0.65934	2.123E-03	2.145E-02	ZDB-GENE-041212-83
ubap1	-0.65846	8.419E-04	1.029E-02	ZDB-GENE-040426-950
ppp1r12c	-0.65553	5.729E-03	4.630E-02	ZDB-GENE-010724-11
atpif1b	-0.65491	1.925E-04	2.968E-03	ZDB-GENE-050506-113
gadd45ba	-0.65433	7.370E-05	1.282E-03	ZDB-GENE-040426-1971
ap1s3b	-0.65411	5.376E-05	9.819E-04	ZDB-GENE-040912-62
slc43a1b	-0.65362	3.957E-03	3.429E-02	ZDB-GENE-070209-173
zfp36l1a	-0.65258	2.095E-03	2.124E-02	ZDB-GENE-030131-9860
crlf3	-0.65248	1.283E-04	2.086E-03	ZDB-GENE-050417-354
rab5c	-0.65178	1.485E-04	2.381E-03	ZDB-GENE-031118-30
cybrd1	-0.65055	2.762E-04	3.992E-03	ZDB-GENE-050522-365
tfg	-0.64860	3.293E-03	2.966E-02	ZDB-GENE-030131-8410
junbb	-0.64754	5.493E-04	7.231E-03	ZDB-GENE-040426-2666
rcbtb1	-0.64608	4.220E-04	5.790E-03	ZDB-GENE-030131-7951
lsm14ab	-0.64599	1.445E-03	1.591E-02	ZDB-GENE-040426-2462
cpn1	-0.64281	1.875E-03	1.956E-02	ZDB-GENE-030131-9116
dgcr2	-0.64274	4.123E-03	3.539E-02	ZDB-GENE-040718-404
dcaf11	-0.64228	1.101E-03	1.271E-02	ZDB-GENE-050809-116
tbk1	-0.64207	2.230E-03	2.215E-02	ZDB-GENE-060512-359
cap1	-0.64025	1.081E-03	1.251E-02	ZDB-GENE-030131-6459
sas	-0.63867	4.120E-04	5.680E-03	ZDB-GENE-030131-9960
zgc:101723	-0.63788	2.428E-04	3.595E-03	ZDB-GENE-040927-25
hnrpkl	-0.63738	8.340E-05	1.425E-03	ZDB-GENE-040426-1923
snx14	-0.63603	1.179E-03	1.340E-02	ZDB-GENE-040724-144
wdr13	-0.63535	2.038E-03	2.080E-02	ZDB-GENE-050522-274
slc39a9	-0.63480	2.252E-03	2.234E-02	ZDB-GENE-060314-3
ptena	-0.63461	3.114E-03	2.851E-02	ZDB-GENE-030131-3776
prdx1	-0.63197	5.876E-04	7.660E-03	ZDB-GENE-050320-35
rsrc2	-0.63194	1.656E-03	1.775E-02	ZDB-GENE-040426-2457
zgc:55512	-0.63113	3.289E-04	4.653E-03	ZDB-GENE-030131-4195
pttg1ipb	-0.62927	2.473E-03	2.377E-02	ZDB-GENE-041010-10
sh3glb1b	-0.62796	2.135E-03	2.149E-02	ZDB-GENE-030131-6851
matn4	-0.62750	2.969E-04	4.253E-03	ZDB-GENE-050208-64
grapa	-0.62638	7.850E-04	9.729E-03	ZDB-GENE-050522-347
ntn1a	-0.62535	3.166E-03	2.886E-02	ZDB-GENE-990415-169
rpz2	-0.62524	3.558E-03	3.160E-02	ZDB-GENE-030131-9071
atp6v1c1a	-0.62508	4.237E-04	5.802E-03	ZDB-GENE-030616-612
commd1	-0.62499	5.343E-03	4.377E-02	ZDB-GENE-060512-329
mef2d	-0.62311	7.071E-04	8.854E-03	ZDB-GENE-990415-164
dcxr	-0.62155	1.471E-03	1.613E-02	ZDB-GENE-030131-7002
mxi1	-0.62100	3.971E-03	3.438E-02	ZDB-GENE-990415-151
zgc:109934	-0.61785	6.476E-04	8.257E-03	ZDB-GENE-050522-428
stambpa	-0.61779	2.114E-03	2.138E-02	ZDB-GENE-040426-1551

lfng	-0.61676	2.161E-04	3.266E-03	ZDB-GENE-980605-16
sepw1	-0.61665	6.190E-04	7.998E-03	ZDB-GENE-030410-5
grnb	-0.61638	1.130E-03	1.297E-02	ZDB-GENE-030131-7393
ephx2	-0.61552	5.924E-03	4.738E-02	ZDB-GENE-041212-70
ugt5b2	-0.61413	2.697E-03	2.550E-02	ZDB-GENE-100406-4
arrdc3a	-0.61409	2.900E-03	2.699E-02	ZDB-GENE-030131-2913
marcksl1b	-0.61170	9.515E-04	1.137E-02	ZDB-GENE-040426-2315
sox7	-0.60668	1.971E-03	2.024E-02	ZDB-GENE-040109-4
trappc6b	-0.60486	6.145E-03	4.854E-02	ZDB-GENE-041010-122
chmp5b	-0.60447	2.719E-04	3.945E-03	ZDB-GENE-040426-1374
rtcb	-0.60431	8.653E-04	1.054E-02	ZDB-GENE-040426-2096
zgc:162228	-0.60359	9.658E-04	1.151E-02	ZDB-GENE-070410-43
mpp1	-0.60286	2.333E-03	2.288E-02	ZDB-GENE-031113-4
si:ch211-154o6.6	-0.60186	5.180E-03	4.264E-02	ZDB-GENE-141216-19
gnpda1	-0.60090	4.304E-03	3.667E-02	ZDB-GENE-050417-417
gsr	-0.59993	1.865E-03	1.950E-02	ZDB-GENE-050522-116
sh3glb2b	-0.59791	6.630E-04	8.410E-03	ZDB-GENE-040426-833
tpcn2	-0.59710	2.101E-03	2.127E-02	ZDB-GENE-061103-202
wbp2nl	-0.59567	5.026E-03	4.158E-02	ZDB-GENE-030131-4012
capgb	-0.59333	9.838E-04	1.166E-02	ZDB-GENE-030131-8541
marcksl1a	-0.59255	2.341E-03	2.292E-02	ZDB-GENE-040426-2146
phactr4b	-0.59157	1.810E-03	1.907E-02	ZDB-GENE-040426-1131
hps5	-0.58889	2.406E-03	2.332E-02	ZDB-GENE-070410-80
eif4g2a	-0.58828	3.862E-04	5.339E-03	ZDB-GENE-030131-2173
rabep2	-0.58545	2.459E-03	2.369E-02	ZDB-GENE-070112-572
si:dkey-222f8.3	-0.58313	3.349E-03	3.004E-02	ZDB-GENE-030131-3658
spry2	-0.58093	1.269E-03	1.427E-02	ZDB-GENE-030131-7038
stk11	-0.58085	2.373E-03	2.310E-02	ZDB-GENE-050417-383
si:ch211-206k20.5	-0.58050	5.003E-03	4.142E-02	ZDB-GENE-030131-2912
zgc:113184	-0.57997	3.728E-03	3.277E-02	ZDB-GENE-050522-9
zbtb16a	-0.57816	6.252E-04	8.064E-03	ZDB-GENE-030131-1989
abca1b	-0.57573	1.513E-03	1.651E-02	ZDB-GENE-030131-9826
keap1a	-0.57350	2.743E-03	2.586E-02	ZDB-GENE-030131-556
nrros	-0.57267	6.009E-03	4.788E-02	ZDB-GENE-040808-36
efna1b	-0.57182	1.035E-03	1.209E-02	ZDB-GENE-041007-5
zgc:92046	-0.57042	1.397E-03	1.546E-02	ZDB-GENE-040718-85
herpud2	-0.57013	1.889E-03	1.965E-02	ZDB-GENE-040426-868
pbx2	-0.56950	1.306E-03	1.460E-02	ZDB-GENE-000405-5
ttc13	-0.56615	6.166E-03	4.868E-02	ZDB-GENE-080722-25
znf710a	-0.56381	2.038E-03	2.080E-02	ZDB-GENE-030131-5559
cotl1	-0.56370	2.223E-03	2.211E-02	ZDB-GENE-030131-8325
smox	-0.56170	2.982E-03	2.754E-02	ZDB-GENE-031201-3
ak4	-0.55838	2.086E-03	2.119E-02	ZDB-GENE-040426-2505
ppifb	-0.55771	5.214E-03	4.283E-02	ZDB-GENE-051030-126
irx3b	-0.55641	3.768E-03	3.309E-02	ZDB-GENE-030131-525
zgc:123105	-0.55605	2.477E-03	2.379E-02	ZDB-GENE-051113-272
tgfb3	-0.55219	3.780E-03	3.313E-02	ZDB-GENE-030723-4
phc1	-0.55076	2.134E-03	2.149E-02	ZDB-GENE-081022-39

etv4	-0.54627	3.710E-03	3.270E-02	ZDB-GENE-990415-71
junba	-0.54438	6.484E-04	8.261E-03	ZDB-GENE-040426-2172
slc35f6	-0.54437	2.257E-03	2.238E-02	ZDB-GENE-040718-297
ei24	-0.54193	5.487E-03	4.469E-02	ZDB-GENE-050417-457
si:dkey-177p2.6	-0.54068	6.191E-03	4.879E-02	ZDB-GENE-030131-5926
gadd45ab	-0.53497	3.430E-03	3.064E-02	ZDB-GENE-040704-60
exoc1	-0.53478	2.342E-03	2.292E-02	ZDB-GENE-030131-1057
b3galnt2	-0.53476	6.282E-03	4.929E-02	ZDB-GENE-050522-358
hig1	-0.53394	1.570E-03	1.696E-02	ZDB-GENE-030826-15
dctn3	-0.53196	1.900E-03	1.975E-02	ZDB-GENE-040704-65
stxbp2	-0.53127	4.420E-03	3.740E-02	ZDB-GENE-040426-2482
kifap3a	-0.53117	5.363E-03	4.390E-02	ZDB-GENE-040912-74
col8a2	-0.52932	1.231E-03	1.390E-02	ZDB-GENE-060503-488
psmb6	-0.52663	3.072E-03	2.816E-02	ZDB-GENE-990415-216
tmem134	-0.52220	3.909E-03	3.402E-02	ZDB-GENE-040426-1345
yipf3	-0.52164	5.528E-03	4.498E-02	ZDB-GENE-030131-6443
znf207a	-0.52091	1.851E-03	1.942E-02	ZDB-GENE-040426-2273
dse	-0.52082	4.424E-03	3.741E-02	ZDB-GENE-041001-202
smad6a	-0.51771	4.386E-03	3.719E-02	ZDB-GENE-011015-1
anxa13	-0.51600	4.429E-03	3.743E-02	ZDB-GENE-010406-1
psmg1	-0.51414	4.122E-03	3.539E-02	ZDB-GENE-040718-469
map1lc3b	-0.51380	2.967E-03	2.749E-02	ZDB-GENE-030131-1145
sirt2	-0.51239	2.271E-03	2.242E-02	ZDB-GENE-030131-1028
atp6v1d	-0.51105	5.040E-03	4.164E-02	ZDB-GENE-040426-727
odc1	-0.51079	3.132E-03	2.860E-02	ZDB-GENE-010816-1
il17rd	-0.50536	4.174E-03	3.568E-02	ZDB-GENE-020320-5
akr1a1b	-0.50494	3.778E-03	3.313E-02	ZDB-GENE-050417-118
si:ch73-281f12.4	-0.50368	5.587E-03	4.535E-02	ZDB-GENE-130530-682
tspan4b	-0.50346	3.241E-03	2.938E-02	ZDB-GENE-070410-127
gabarapa	-0.50129	6.377E-03	4.985E-02	ZDB-GENE-030131-5174
si:ch211-191d15.2	-0.49752	5.168E-03	4.257E-02	ZDB-GENE-040724-113
kpna4	-0.49449	2.389E-03	2.320E-02	ZDB-GENE-030131-5584
cited4b	-0.49293	2.128E-03	2.146E-02	ZDB-GENE-030425-5
tal1	-0.49156	3.533E-03	3.142E-02	ZDB-GENE-980526-501
swap70b	-0.48986	4.751E-03	3.962E-02	ZDB-GENE-030131-3587
acsbg2	-0.48933	4.377E-03	3.716E-02	ZDB-GENE-030131-7099
papola	-0.48743	6.235E-03	4.906E-02	ZDB-GENE-030131-3507
ube2d3	-0.48500	2.832E-03	2.650E-02	ZDB-GENE-030131-551
exoc7	-0.48495	6.053E-03	4.807E-02	ZDB-GENE-040426-2752
frem1b	-0.48165	2.606E-03	2.482E-02	ZDB-GENE-050208-783
clta	-0.47736	5.336E-03	4.373E-02	ZDB-GENE-040426-1986
amdhd2	-0.47076	4.978E-03	4.126E-02	ZDB-GENE-040426-1871
march6	-0.46865	6.365E-03	4.978E-02	ZDB-GENE-070912-530
rhbdf1	-0.45711	6.033E-03	4.801E-02	ZDB-GENE-040704-75
sept8a	-0.44994	5.772E-03	4.657E-02	ZDB-GENE-030131-309

Genes downregulated in *lmo4a* morphants

GeneID	logFC	PValue	FDR	ZFIN ID
pvalb4	5.80881	2.074E-06	5.740E-05	ZDB-GENE-040625-48
cryba1b	5.06758	4.308E-03	3.669E-02	ZDB-GENE-040718-326
pvalb1	4.57264	2.968E-36	1.971E-33	ZDB-GENE-040625-57
myhz1.2	4.13299	2.034E-14	2.497E-12	ZDB-GENE-070705-73
and2	4.05870	1.121E-12	1.024E-10	ZDB-GENE-041105-2
and1	3.82491	3.637E-13	3.567E-11	ZDB-GENE-030131-9105
myhz1.1	3.82096	5.395E-18	9.274E-16	ZDB-GENE-000322-5
cyp4t8	3.54601	2.070E-04	3.154E-03	ZDB-GENE-031219-3
myhz2	3.50895	2.714E-20	5.507E-18	ZDB-GENE-020604-1
si:ch211-251b21.1	3.48896	9.157E-04	1.104E-02	ZDB-GENE-060809-5
crybb1	3.46602	1.022E-03	1.197E-02	ZDB-GENE-010813-1
dct	3.39810	1.886E-08	8.721E-07	ZDB-GENE-000508-1
pgam2	3.25103	5.748E-11	4.097E-09	ZDB-GENE-040116-6
zgc:86709	3.17374	4.944E-09	2.571E-07	ZDB-GENE-040625-35
mylpfb	3.10001	6.078E-08	2.467E-06	ZDB-GENE-040912-115
hspb9	3.07113	2.775E-11	2.112E-09	ZDB-GENE-080214-6
hbae1	3.02095	2.065E-14	2.515E-12	ZDB-GENE-980526-80
murcb	2.93331	1.137E-09	6.486E-08	ZDB-GENE-041212-87
icn2	2.91318	2.015E-07	7.165E-06	ZDB-GENE-070822-9
atp2a1	2.90955	3.905E-11	2.867E-09	ZDB-GENE-020905-1
LOC573653	2.80408	3.958E-14	4.553E-12	ZDB-GENE-131120-128
smyhc1	2.78888	6.373E-09	3.233E-07	ZDB-GENE-030131-271
pmela	2.77673	3.737E-15	5.151E-13	ZDB-GENE-030131-9818
mdh1b	2.71753	5.157E-14	5.752E-12	ZDB-GENE-060929-384
acta1b	2.67727	2.594E-20	5.338E-18	ZDB-GENE-030131-55
tnnt3a	2.65524	2.457E-20	5.130E-18	ZDB-GENE-000322-3
mylpfa	2.63424	1.639E-11	1.274E-09	ZDB-GENE-990712-15
tnni2a.4	2.62037	1.919E-14	2.376E-12	ZDB-GENE-040625-119
evpla	2.58259	3.533E-05	6.920E-04	ZDB-GENE-030829-19
srl	2.57277	4.730E-10	2.904E-08	ZDB-GENE-061110-91
tyrp1b	2.56660	7.485E-06	1.790E-04	ZDB-GENE-040718-249
foxn4	2.52320	1.675E-06	4.753E-05	ZDB-GENE-990415-277
ckma	2.50406	8.484E-15	1.107E-12	ZDB-GENE-980526-109
casq2	2.49858	5.161E-10	3.104E-08	ZDB-GENE-010724-12
wu:fc46h12	2.49302	3.008E-05	5.979E-04	ZDB-GENE-030131-3611
actn3a	2.48603	2.584E-05	5.237E-04	ZDB-GENE-000329-9
dhrs7cb	2.45852	4.903E-07	1.606E-05	ZDB-GENE-050522-226
chrna1	2.45183	4.543E-06	1.167E-04	ZDB-GENE-980526-137
tgm2a	2.43114	5.098E-05	9.388E-04	ZDB-GENE-040912-78
tmod4	2.42887	1.228E-06	3.640E-05	ZDB-GENE-991019-9
tpma	2.42146	1.130E-12	1.025E-10	ZDB-GENE-990415-269
her4.2	2.40524	4.589E-14	5.239E-12	ZDB-GENE-060815-1
hhatla	2.40426	3.336E-07	1.139E-05	ZDB-GENE-040426-1865
gsx1	2.39157	2.313E-03	2.275E-02	ZDB-GENE-041008-135
aox3	2.38480	3.538E-03	3.144E-02	ZDB-GENE-001205-2
actc1b	2.37108	1.048E-12	9.631E-11	ZDB-GENE-000322-1

zgc:66286	2.36363	2.946E-11	2.196E-09	ZDB-GENE-040426-1593
tmem38a	2.35860	2.677E-07	9.271E-06	ZDB-GENE-040426-1888
ckmb	2.35824	7.527E-13	7.096E-11	ZDB-GENE-040426-2128
cyt1l	2.35319	4.194E-10	2.641E-08	ZDB-GENE-061026-4
sfrp5	2.35103	2.799E-11	2.119E-09	ZDB-GENE-011108-2
mylz3	2.33967	5.336E-13	5.096E-11	ZDB-GENE-000322-6
pax6b	2.30525	6.057E-06	1.505E-04	ZDB-GENE-001031-1
tnnt3b	2.30211	6.634E-15	8.975E-13	ZDB-GENE-030520-2
zgc:92061	2.29867	5.389E-08	2.237E-06	ZDB-GENE-040718-78
pygma	2.29553	2.653E-13	2.673E-11	ZDB-GENE-050522-482
coe2	2.28510	2.521E-13	2.558E-11	ZDB-GENE-990715-11
aldoab	2.28125	1.180E-21	2.828E-19	ZDB-GENE-040426-2299
smyhc2	2.26239	3.129E-03	2.859E-02	ZDB-GENE-070822-12
zgc:85869	2.25650	6.078E-08	2.467E-06	ZDB-GENE-040426-2321
p4ha1b	2.24316	5.006E-16	7.464E-14	ZDB-GENE-030131-4089
stmn2b	2.22319	4.156E-05	7.949E-04	ZDB-GENE-050522-207
elavl4	2.22000	2.273E-05	4.675E-04	ZDB-GENE-990415-246
zgc:165571	2.21885	1.108E-03	1.276E-02	ZDB-GENE-030131-7249
krtt1c19e	2.19958	3.095E-08	1.330E-06	ZDB-GENE-050506-95
pitx2	2.19928	1.258E-04	2.049E-03	ZDB-GENE-990714-27
cav3	2.16345	2.080E-03	2.117E-02	ZDB-GENE-050522-426
dlx2a	2.14630	3.078E-05	6.095E-04	ZDB-GENE-980526-212
myl1	2.13569	1.916E-08	8.831E-07	ZDB-GENE-030131-8109
desma	2.13345	5.556E-10	3.314E-08	ZDB-GENE-980526-221
dpysl5a	2.12991	4.680E-06	1.200E-04	ZDB-GENE-030131-3136
tmem45b	2.11158	2.780E-03	2.614E-02	ZDB-GENE-040426-1904
chrna5	2.10714	9.856E-06	2.301E-04	ZDB-GENE-050417-440
mat1a	2.10673	7.011E-08	2.792E-06	ZDB-GENE-030131-6127
acta1a	2.10397	9.123E-21	1.960E-18	ZDB-GENE-050417-267
gch2	2.10150	2.628E-04	3.828E-03	ZDB-GENE-001205-3
zgc:193505	2.09820	1.197E-03	1.359E-02	ZDB-GENE-030131-7103
tnnc2	2.09189	8.674E-10	5.030E-08	ZDB-GENE-000322-2
ldb3a	2.08345	1.964E-03	2.022E-02	ZDB-GENE-040121-6
nr2e1	2.04888	1.775E-03	1.878E-02	ZDB-GENE-040801-127
ankrd37	1.99747	3.520E-06	9.285E-05	ZDB-GENE-070308-1
cabp1b	1.98499	4.943E-07	1.616E-05	ZDB-GENE-040718-111
apobec2a	1.98076	6.313E-05	1.128E-03	ZDB-GENE-050306-15
cap2	1.97138	5.830E-04	7.612E-03	ZDB-GENE-040426-1758
anxa1c	1.96841	9.525E-04	1.137E-02	ZDB-GENE-030131-5274
dbx1a	1.95852	1.761E-05	3.734E-04	ZDB-GENE-000128-8
fez1	1.95804	2.164E-04	3.266E-03	ZDB-GENE-040426-2723
wu:fb17f05	1.94089	4.549E-03	3.827E-02	ZDB-GENE-030131-221
ptrfb	1.93703	2.241E-11	1.724E-09	ZDB-GENE-030131-5509
zgc:162267	1.92139	1.217E-04	1.994E-03	ZDB-GENE-070410-53
dbnla	1.92010	5.916E-03	4.738E-02	ZDB-GENE-040704-42
tnc	1.89575	7.374E-07	2.307E-05	ZDB-GENE-980526-104
stac3	1.89213	1.868E-03	1.951E-02	ZDB-GENE-040801-248
si:ch211-243g18.2	1.88579	7.474E-06	1.790E-04	ZDB-GENE-030131-2830

hmp19	1.88429	2.301E-04	3.439E-03	ZDB-GENE-030131-3694
col8a1a	1.87187	2.325E-04	3.467E-03	ZDB-GENE-070912-274
aqp3a	1.86807	9.360E-09	4.574E-07	ZDB-GENE-040426-2826
hoxc13b	1.86797	2.029E-05	4.242E-04	ZDB-GENE-000822-5
plp1a	1.86591	4.267E-05	8.118E-04	ZDB-GENE-001202-1
znf207b	1.85350	2.728E-25	8.950E-23	ZDB-GENE-030131-7377
zgc:110333	1.84633	9.084E-04	1.097E-02	ZDB-GENE-050417-331
tmsb	1.83639	2.974E-06	7.959E-05	ZDB-GENE-050307-5
cox6b1	1.83573	1.051E-06	3.140E-05	ZDB-GENE-040718-448
mybpc2b	1.83390	1.181E-03	1.341E-02	ZDB-GENE-050320-80
ank1a	1.82709	5.703E-03	4.614E-02	ZDB-GENE-041010-44
rx1	1.82623	3.395E-04	4.793E-03	ZDB-GENE-990415-236
neurod4	1.81857	1.964E-05	4.111E-04	ZDB-GENE-030730-1
cryba4	1.81433	6.049E-04	7.850E-03	ZDB-GENE-050522-143
elavl3	1.81116	1.333E-07	4.955E-06	ZDB-GENE-980526-76
elovl7a	1.80789	7.963E-04	9.836E-03	ZDB-GENE-030131-6149
slc45a2	1.78550	1.367E-03	1.517E-02	ZDB-GENE-050208-97
gpm6aa	1.78306	9.018E-05	1.522E-03	ZDB-GENE-030710-7
hlx1	1.77940	6.457E-21	1.430E-18	ZDB-GENE-030131-5304
tuba1l	1.77260	4.851E-15	6.624E-13	ZDB-GENE-061114-1
spata6l	1.77066	2.655E-03	2.517E-02	ZDB-GENE-040426-1369
ap1m3	1.76735	7.894E-04	9.776E-03	ZDB-GENE-040801-159
tnni2b.1	1.74509	3.540E-08	1.504E-06	ZDB-GENE-050417-49
jph2	1.74143	4.709E-03	3.938E-02	ZDB-GENE-030131-9848
egln3	1.74044	3.678E-03	3.249E-02	ZDB-GENE-040426-2541
tnnc1b	1.73958	9.960E-07	2.988E-05	ZDB-GENE-040625-62
gstm3	1.73662	3.088E-04	4.403E-03	ZDB-GENE-050309-24
klhl40a	1.73491	5.790E-04	7.567E-03	ZDB-GENE-050916-1
sp9	1.73246	7.972E-04	9.838E-03	ZDB-GENE-040426-2313
zgc:110283	1.72575	5.961E-03	4.765E-02	ZDB-GENE-050417-72
capn9	1.72495	1.117E-04	1.851E-03	ZDB-GENE-010724-2
cyt1	1.70173	8.032E-07	2.492E-05	ZDB-GENE-991008-6
LOC100149066	1.69586	2.121E-05	4.409E-04	-
klhl31	1.69192	2.386E-04	3.542E-03	ZDB-GENE-030131-2052
tnw	1.69180	1.342E-11	1.054E-09	ZDB-GENE-990415-262
tnnt1	1.68760	3.082E-08	1.329E-06	ZDB-GENE-080723-27
blf	1.67922	3.313E-11	2.445E-09	ZDB-GENE-050721-1
ahcyl1	1.67553	2.560E-03	2.450E-02	ZDB-GENE-021206-7
llgl1	1.67312	4.107E-09	2.174E-07	ZDB-GENE-060825-69
pah	1.67233	2.551E-03	2.442E-02	ZDB-GENE-031006-2
crabp1b	1.67223	6.724E-06	1.646E-04	ZDB-GENE-040624-3
mab21l2	1.67189	4.191E-05	7.995E-04	ZDB-GENE-011101-3
angptl7	1.66851	4.638E-04	6.275E-03	ZDB-GENE-041010-175
actc1a	1.65916	2.158E-12	1.855E-10	ZDB-GENE-040520-4
vipr2	1.65900	1.552E-03	1.684E-02	ZDB-GENE-010406-4
zgc:92533	1.63388	2.148E-09	1.175E-07	ZDB-GENE-040801-181
myl10	1.63329	3.649E-07	1.234E-05	ZDB-GENE-050417-421
zgc:154054	1.61580	1.574E-03	1.699E-02	ZDB-GENE-061103-283

fitm1	1.61572	3.325E-03	2.990E-02	ZDB-GENE-050306-28
vasna	1.60917	1.567E-03	1.695E-02	ZDB-GENE-050522-43
si:busm1-118j2.5	1.60838	7.710E-07	2.402E-05	ZDB-GENE-030131-7722
anxa1b	1.60319	4.366E-03	3.709E-02	ZDB-GENE-030131-6554
slc38a4	1.59650	3.412E-09	1.813E-07	ZDB-GENE-041010-14
quo	1.59453	5.063E-04	6.755E-03	ZDB-GENE-030131-9944
sepw2b	1.59406	1.762E-03	1.868E-02	ZDB-GENE-030428-2
myoc	1.59251	4.107E-06	1.065E-04	ZDB-GENE-050425-2
ptgdsb	1.59010	5.488E-08	2.272E-06	ZDB-GENE-030131-8436
klf17	1.58128	8.696E-08	3.379E-06	ZDB-GENE-010129-1
LOC100148329	1.56995	7.102E-10	4.151E-08	ZDB-GENE-060815-1
atp1b1b	1.56771	3.405E-04	4.802E-03	ZDB-GENE-001127-4
zgc:165555	1.56026	4.430E-05	8.352E-04	ZDB-GENE-070620-17
stmn1b	1.55779	5.037E-05	9.326E-04	ZDB-GENE-050417-397
scinla	1.55450	5.699E-05	1.034E-03	ZDB-GENE-030131-2005
si:ch211-191i18.2	1.55433	3.147E-14	3.738E-12	ZDB-GENE-060503-769
cyp2y3	1.55300	9.793E-04	1.162E-02	ZDB-GENE-050522-490
ctnnd2b	1.54841	1.374E-03	1.522E-02	ZDB-GENE-070912-532
pitx1	1.54494	6.881E-06	1.676E-04	ZDB-GENE-060421-3623
zgc:101853	1.54045	5.769E-04	7.547E-03	ZDB-GENE-041212-13
cldne	1.53560	5.376E-08	2.237E-06	ZDB-GENE-010328-5
zgc:158296	1.53374	4.471E-03	3.768E-02	ZDB-GENE-070112-1272
aspm	1.53116	3.697E-18	6.508E-16	ZDB-GENE-050208-620
cmyb	1.52670	2.608E-07	9.077E-06	ZDB-GENE-991110-14
vcanb	1.52515	1.354E-05	3.016E-04	ZDB-GENE-030131-2185
zic4	1.52192	3.183E-03	2.898E-02	ZDB-GENE-040622-4
myom1a	1.51623	1.851E-05	3.902E-04	ZDB-GENE-030131-2856
mdkb	1.50896	1.540E-06	4.439E-05	ZDB-GENE-010131-6
thbs4b	1.50749	2.330E-09	1.270E-07	ZDB-GENE-020708-4
serpinf1	1.50650	9.905E-10	5.698E-08	ZDB-GENE-040912-2
cygb1	1.50415	1.510E-05	3.327E-04	ZDB-GENE-020513-1
mgst1.2	1.50257	9.295E-04	1.117E-02	ZDB-GENE-040704-59
zgc:92518	1.49257	6.258E-05	1.122E-03	ZDB-GENE-040718-53
kera	1.48970	1.569E-05	3.407E-04	ZDB-GENE-041210-165
krt4	1.48929	6.374E-05	1.137E-03	ZDB-GENE-000607-83
plce1	1.48307	9.696E-06	2.270E-04	ZDB-GENE-061212-4
hsp90aa1.1	1.47913	8.026E-07	2.492E-05	ZDB-GENE-990415-94
zgc:171416	1.47452	1.092E-05	2.516E-04	ZDB-GENE-041111-230
pcdh18a	1.47409	1.984E-04	3.048E-03	ZDB-GENE-000607-1
diaph3	1.47292	2.653E-03	2.517E-02	ZDB-GENE-040801-144
tnni1al	1.46686	7.979E-05	1.373E-03	ZDB-GENE-040625-81
tnnt2d	1.46140	1.888E-03	1.965E-02	ZDB-GENE-050626-97
tk1	1.45983	6.904E-04	8.690E-03	ZDB-GENE-030131-5801
unc45b	1.45841	6.299E-05	1.127E-03	ZDB-GENE-020919-3
mdm4	1.45330	3.086E-11	2.289E-09	ZDB-GENE-030131-6872
ascl1a	1.45287	1.745E-05	3.712E-04	ZDB-GENE-980526-90
ppl	1.44927	5.683E-05	1.033E-03	ZDB-GENE-041008-196
dnph1	1.44327	2.859E-06	7.694E-05	ZDB-GENE-081107-5

ccne2	1.44286	4.673E-11	3.408E-09	ZDB-GENE-030131-9689
zgc:194989	1.44191	9.559E-06	2.245E-04	ZDB-GENE-080723-3
hoxd4a	1.42153	2.701E-03	2.552E-02	ZDB-GENE-980526-214
parvb	1.41754	1.937E-03	2.001E-02	ZDB-GENE-030131-4411
anln	1.41616	1.406E-11	1.099E-09	ZDB-GENE-030131-4486
lipg	1.41527	3.468E-07	1.176E-05	ZDB-GENE-040426-699
cdc20	1.40765	9.641E-15	1.236E-12	ZDB-GENE-040426-2044
tpm2	1.40004	3.393E-06	8.988E-05	ZDB-GENE-040625-114
txlnbb	1.39793	1.056E-03	1.231E-02	ZDB-GENE-070424-77
si:ch211-106g4.2	1.38070	3.207E-04	4.550E-03	ZDB-GENE-081105-78
actn3b	1.38044	4.839E-03	4.029E-02	ZDB-GENE-030131-2040
hmmr	1.37676	1.458E-08	6.850E-07	ZDB-GENE-030131-731
tox	1.37193	2.708E-08	1.195E-06	ZDB-GENE-070912-181
tfap2a	1.36514	3.937E-03	3.421E-02	ZDB-GENE-011212-6
si:ch211-69g19.2	1.36193	1.832E-16	2.759E-14	ZDB-GENE-030804-13
zgc:153629	1.34626	2.055E-05	4.289E-04	ZDB-GENE-061027-116
zgc:112234	1.34268	1.732E-03	1.845E-02	ZDB-GENE-050522-106
hoxc12b	1.33954	9.463E-04	1.132E-02	ZDB-GENE-000329-17
her15.1	1.33714	2.964E-06	7.948E-05	ZDB-GENE-030707-2
aldob	1.33579	1.961E-12	1.706E-10	ZDB-GENE-030131-383
asf1ba	1.33157	1.095E-03	1.265E-02	ZDB-GENE-031118-197
fmoda	1.32819	1.114E-04	1.847E-03	ZDB-GENE-041008-23
foxm1	1.32549	9.793E-06	2.290E-04	ZDB-GENE-040426-1275
ckbb	1.31942	1.073E-08	5.174E-07	ZDB-GENE-020103-2
hapln1a	1.30995	3.865E-06	1.009E-04	ZDB-GENE-050302-175
zic2b	1.30647	1.225E-09	6.884E-08	ZDB-GENE-030219-92
cdh7	1.30038	5.678E-03	4.599E-02	ZDB-GENE-061019-3
adrb3a	1.29535	6.298E-05	1.127E-03	ZDB-GENE-080917-21
rbfox1	1.28892	6.276E-03	4.928E-02	ZDB-GENE-040927-11
atp1a2a	1.28577	3.175E-03	2.892E-02	ZDB-GENE-001212-6
top2a	1.26949	3.335E-14	3.908E-12	ZDB-GENE-030131-2453
plk4	1.26942	4.621E-07	1.517E-05	ZDB-GENE-030619-14
hist1h4l	1.26755	7.433E-11	5.196E-09	ZDB-GENE-070927-10
krcp	1.25428	5.417E-10	3.244E-08	ZDB-GENE-030131-2126
her15.2	1.25379	9.847E-08	3.747E-06	ZDB-GENE-070627-1
hmgb2a	1.24813	9.102E-10	5.257E-08	ZDB-GENE-051120-126
slc29a2	1.24249	2.281E-04	3.412E-03	ZDB-GENE-050220-15
mcm5	1.24220	4.621E-12	3.836E-10	ZDB-GENE-021209-1
ndc80	1.24174	5.052E-09	2.618E-07	ZDB-GENE-030131-904
dck	1.23561	5.878E-05	1.062E-03	ZDB-GENE-041024-3
dlb	1.23444	1.545E-05	3.376E-04	ZDB-GENE-980526-114
capn8	1.23380	5.695E-06	1.427E-04	ZDB-GENE-050522-279
nedd9	1.23243	4.336E-04	5.933E-03	ZDB-GENE-061013-318
slc4a11	1.23234	5.749E-04	7.527E-03	ZDB-GENE-090313-183
arl4ca	1.23194	5.578E-03	4.532E-02	ZDB-GENE-050417-265
kpna2	1.22796	8.238E-11	5.732E-09	ZDB-GENE-040718-22
si:dkeyp-35b8.5	1.22483	1.157E-11	9.140E-10	ZDB-GENE-110404-2
LOC325449	1.22231	1.470E-10	9.895E-09	ZDB-GENE-050522-252

rasgrp3	1.22114	2.781E-09	1.500E-07	ZDB-GENE-070424-82
cdca7a	1.21988	1.155E-11	9.140E-10	ZDB-GENE-050417-29
sox19a	1.21065	6.538E-05	1.157E-03	ZDB-GENE-980526-102
pvalb2	1.20086	5.186E-11	3.714E-09	ZDB-GENE-000322-4
arhgef7a	1.20052	1.016E-05	2.355E-04	ZDB-GENE-060322-11
calr3b	1.19251	8.355E-08	3.256E-06	ZDB-GENE-030131-9907
si:ch211-63o20.7	1.19173	1.388E-06	4.058E-05	ZDB-GENE-041001-194
zgc:113255	1.18855	1.004E-03	1.183E-02	ZDB-GENE-050327-45
stmn1a	1.18698	2.893E-11	2.168E-09	ZDB-GENE-031006-14
si:dkey-11j8.1	1.18355	2.810E-06	7.574E-05	ZDB-GENE-030616-151
hells	1.18135	1.616E-10	1.073E-08	ZDB-GENE-030131-9923
atad5b	1.17923	9.157E-06	2.162E-04	ZDB-GENE-030131-338
hoxc4a	1.17715	1.273E-03	1.431E-02	ZDB-GENE-990415-112
nexn	1.17428	8.099E-04	9.953E-03	ZDB-GENE-041114-92
g2e3	1.17167	7.445E-09	3.675E-07	ZDB-GENE-030616-400
cenpe	1.17083	6.453E-08	2.605E-06	ZDB-GENE-060929-860
slc29a1a	1.16637	1.661E-06	4.721E-05	ZDB-GENE-050913-138
gapdh	1.15995	1.368E-06	4.007E-05	ZDB-GENE-030115-1
hmga1a	1.15602	1.193E-07	4.481E-06	ZDB-GENE-010502-1
tspan2	1.15389	5.008E-05	9.287E-04	ZDB-GENE-050522-511
haus4	1.15225	7.296E-07	2.288E-05	ZDB-GENE-040718-122
gfap	1.14948	4.318E-03	3.674E-02	ZDB-GENE-990914-3
mcm6	1.14375	4.688E-11	3.408E-09	ZDB-GENE-030909-6
msi1	1.14290	1.212E-05	2.746E-04	ZDB-GENE-050320-86
ncaph2	1.14273	7.075E-06	1.712E-04	ZDB-GENE-041210-172
cep55l	1.14240	1.173E-03	1.335E-02	ZDB-GENE-050522-222
pcolcea	1.13927	1.391E-03	1.540E-02	ZDB-GENE-041014-370
sfrp1a	1.13534	1.873E-04	2.899E-03	ZDB-GENE-040310-5
stil	1.12945	7.334E-04	9.144E-03	ZDB-GENE-020419-39
mki67	1.12517	8.690E-11	5.990E-09	ZDB-GENE-030131-9771
wdhd1	1.11932	2.246E-08	1.022E-06	ZDB-GENE-030131-1665
itgb4	1.11750	5.217E-03	4.283E-02	ZDB-GENE-030131-7209
fkbp11	1.11261	6.963E-04	8.740E-03	ZDB-GENE-030616-357
dlx5a	1.11149	4.617E-03	3.878E-02	ZDB-GENE-990415-49
smc2	1.10913	3.597E-10	2.285E-08	ZDB-GENE-030131-105
kif11	1.10308	3.690E-10	2.334E-08	ZDB-GENE-020426-1
mcm3	1.09968	7.299E-09	3.640E-07	ZDB-GENE-020419-4
zgc:174160	1.09937	5.963E-07	1.928E-05	ZDB-GENE-030131-6252
kazald2	1.09594	3.723E-04	5.176E-03	ZDB-GENE-070928-14
dut	1.09278	1.088E-10	7.396E-09	ZDB-GENE-041010-93
fam188a	1.09199	1.494E-05	3.298E-04	ZDB-GENE-061110-13
nusap1	1.08846	6.580E-09	3.327E-07	ZDB-GENE-030827-5
cep128	1.08751	5.119E-04	6.813E-03	ZDB-GENE-070410-34
fam60al	1.08438	6.292E-09	3.203E-07	ZDB-GENE-031114-2
acrc	1.08361	1.411E-05	3.130E-04	ZDB-GENE-050320-153
fam169aa	1.08095	8.905E-05	1.509E-03	ZDB-GENE-060825-117
rgs4	1.07874	2.435E-05	4.955E-04	ZDB-GENE-030131-9839
dctd	1.07692	1.533E-06	4.428E-05	ZDB-GENE-050417-116

aldh1a2	1.07266	5.670E-03	4.595E-02	ZDB-GENE-011010-3
dnmt1	1.07191	1.197E-07	4.484E-06	ZDB-GENE-990714-15
gfi1aa	1.06855	6.848E-06	1.670E-04	ZDB-GENE-050522-534
ccnb3	1.06797	1.889E-07	6.782E-06	ZDB-GENE-060929-684
pold1	1.06617	3.201E-09	1.713E-07	ZDB-GENE-060429-1
dlgap5	1.06377	3.237E-09	1.726E-07	ZDB-GENE-040912-160
appl2	1.06230	5.580E-03	4.532E-02	ZDB-GENE-081016-2
rbfox1l	1.06063	2.913E-03	2.706E-02	ZDB-GENE-040923-2
aurka	1.06028	5.702E-09	2.924E-07	ZDB-GENE-040801-161
cdk1	1.05775	5.624E-10	3.341E-08	ZDB-GENE-010320-1
rrm1	1.05641	1.486E-10	9.958E-09	ZDB-GENE-990415-247
ttk	1.05623	3.194E-09	1.713E-07	ZDB-GENE-030123-1
sass6	1.04854	1.044E-04	1.742E-03	ZDB-GENE-040426-2784
arhgap11a	1.04773	1.888E-06	5.286E-05	ZDB-GENE-040426-2562
agr2	1.04596	1.301E-03	1.457E-02	ZDB-GENE-050417-214
kcnj2a	1.04403	2.190E-05	4.532E-04	ZDB-GENE-120319-1
ube2c	1.04403	2.477E-09	1.345E-07	ZDB-GENE-051030-48
ephb2b	1.04273	2.316E-08	1.041E-06	ZDB-GENE-050522-415
dhfr	1.04164	8.959E-08	3.472E-06	ZDB-GENE-010406-5
tuba1b	1.03910	2.951E-08	1.286E-06	ZDB-GENE-030822-1
prdm1a	1.03414	1.811E-07	6.517E-06	ZDB-GENE-030131-2193
traip	1.03171	4.401E-04	6.004E-03	ZDB-GENE-040801-30
zgc:162618	1.02840	1.298E-05	2.914E-04	ZDB-GENE-030131-4476
znf185	1.02463	6.503E-04	8.270E-03	ZDB-GENE-061103-355
tubb5	1.02425	8.391E-07	2.576E-05	ZDB-GENE-031110-4
zgc:163030	1.02403	1.580E-04	2.502E-03	ZDB-GENE-060810-77
kcnk5b	1.02233	1.858E-03	1.946E-02	ZDB-GENE-040426-1297
pbk	1.02118	3.104E-07	1.065E-05	ZDB-GENE-030523-2
si:dkey-57a22.11	1.02024	7.164E-04	8.963E-03	ZDB-GENE-081104-423
prkcbb	1.01853	2.682E-05	5.405E-04	ZDB-GENE-040426-1178
pou6f1	1.01765	1.503E-03	1.643E-02	ZDB-GENE-980526-530
ptgr1	1.01570	3.239E-08	1.384E-06	ZDB-GENE-041212-80
setd8b	1.01002	4.111E-06	1.065E-04	ZDB-GENE-071004-17
rfc1	1.00671	5.814E-06	1.452E-04	ZDB-GENE-070410-99
zgc:153631	1.00555	2.267E-03	2.242E-02	ZDB-GENE-060825-208
ncaph	1.00248	2.126E-06	5.861E-05	ZDB-GENE-070112-652
fanci	1.00045	1.545E-05	3.376E-04	ZDB-GENE-030131-2905
rnaseh2a	0.99775	6.402E-07	2.047E-05	ZDB-GENE-040426-976
suv39h1b	0.99419	3.324E-06	8.831E-05	ZDB-GENE-030131-5105
sort1b	0.99381	4.170E-03	3.568E-02	ZDB-GENE-030131-7447
fen1	0.99293	1.280E-08	6.074E-07	ZDB-GENE-031112-11
esco2	0.98550	1.769E-07	6.399E-06	ZDB-GENE-050913-156
zgc:55733	0.98396	4.535E-07	1.496E-05	ZDB-GENE-040426-791
ccnb2	0.98294	1.563E-05	3.405E-04	ZDB-GENE-030429-12
tuba8l2	0.98192	1.135E-05	2.599E-04	ZDB-GENE-040426-1646
pif1	0.98148	3.607E-05	7.038E-04	ZDB-GENE-030131-4928
pole	0.98040	2.545E-04	3.726E-03	ZDB-GENE-070705-557
prc1b	0.97884	9.503E-07	2.875E-05	ZDB-GENE-040426-777

zgc:153910	0.97587	1.007E-04	1.687E-03	ZDB-GENE-061013-398
cbx1b	0.97006	2.503E-04	3.676E-03	ZDB-GENE-040625-68
zgc:153405	0.96839	5.214E-06	1.321E-04	ZDB-GENE-060825-170
tacc3	0.96790	2.041E-07	7.238E-06	ZDB-GENE-050522-327
banp	0.96708	2.340E-05	4.788E-04	ZDB-GENE-050522-54
sc:d189	0.96650	3.619E-05	7.042E-04	ZDB-GENE-080303-11
zgc:86764	0.96603	1.085E-05	2.505E-04	ZDB-GENE-040718-435
sema4e	0.96336	2.808E-05	5.644E-04	ZDB-GENE-990715-7
prkrira	0.96248	6.981E-06	1.694E-04	ZDB-GENE-040426-2045
trip13	0.96019	1.924E-05	4.032E-04	ZDB-GENE-040426-1488
pim1	0.95983	2.924E-04	4.197E-03	ZDB-GENE-061013-139
extl3	0.95900	5.390E-03	4.405E-02	ZDB-GENE-041124-2
atf7ip	0.95806	2.689E-08	1.194E-06	ZDB-GENE-061103-178
spry4	0.95769	4.080E-07	1.361E-05	ZDB-GENE-010803-2
adck5	0.95748	2.202E-03	2.198E-02	ZDB-GENE-081104-149
fbxo5	0.95674	4.112E-05	7.895E-04	ZDB-GENE-030131-4027
rpl3	0.95621	2.550E-05	5.176E-04	ZDB-GENE-030131-1291
acat2	0.95565	1.395E-04	2.250E-03	ZDB-GENE-990714-22
haus6	0.95561	1.582E-04	2.502E-03	ZDB-GENE-030131-5517
wu:fb15g10	0.95489	6.885E-05	1.205E-03	ZDB-GENE-030131-8575
prickle2	0.95473	6.579E-05	1.159E-03	ZDB-GENE-030724-6
mettl13	0.95411	8.316E-04	1.017E-02	ZDB-GENE-060929-60
msh2	0.95340	4.171E-06	1.079E-04	ZDB-GENE-040426-2932
zgc:113984	0.95109	3.421E-03	3.059E-02	ZDB-GENE-050626-82
mfsd2aa	0.94535	3.506E-03	3.125E-02	ZDB-GENE-041114-166
hmgb2b	0.94464	1.243E-06	3.677E-05	ZDB-GENE-040912-122
meis3	0.94423	5.062E-05	9.340E-04	ZDB-GENE-010406-2
hat1	0.94364	5.296E-06	1.339E-04	ZDB-GENE-040912-99
fam83d	0.94283	2.981E-05	5.935E-04	ZDB-GENE-041111-264
plk1	0.94082	3.328E-08	1.418E-06	ZDB-GENE-021115-7
pkp3a	0.93956	2.082E-03	2.117E-02	ZDB-GENE-060512-114
ncapd2	0.93831	3.632E-08	1.538E-06	ZDB-GENE-050506-59
dnmt4	0.93487	6.894E-05	1.205E-03	ZDB-GENE-050314-1
depdc1a	0.93212	3.581E-05	6.996E-04	ZDB-GENE-030131-9827
si:dkeyp-117h8.4	0.93142	1.550E-03	1.684E-02	ZDB-GENE-030804-4
cenpo	0.92834	1.944E-03	2.005E-02	ZDB-GENE-070928-16
ankrd10a	0.92668	1.362E-03	1.514E-02	ZDB-GENE-040426-700
rpl10a	0.92514	3.001E-06	8.015E-05	ZDB-GENE-030131-2025
nrf1	0.92399	6.481E-05	1.148E-03	ZDB-GENE-001221-1
zgc:66479	0.91772	5.983E-03	4.778E-02	ZDB-GENE-031030-10
katnb1	0.91648	1.147E-05	2.624E-04	ZDB-GENE-040426-1954
ccna2	0.91578	3.733E-08	1.576E-06	ZDB-GENE-020418-1
si:ch211-255a21.1	0.91492	4.108E-05	7.895E-04	ZDB-GENE-030131-6690
si:ch211-244o22.2	0.91321	1.483E-05	3.278E-04	ZDB-GENE-030131-3271
haus5	0.91014	1.225E-04	2.004E-03	ZDB-GENE-041114-150
hmgb1b	0.90967	6.661E-08	2.681E-06	ZDB-GENE-030131-8480
pcna	0.90954	1.273E-06	3.749E-05	ZDB-GENE-000210-8
acsl1a	0.90907	6.826E-05	1.199E-03	ZDB-GENE-050809-115

hirip3	0.90613	5.323E-05	9.753E-04	ZDB-GENE-030131-5753
tnfaip8l1	0.90474	1.142E-06	3.391E-05	ZDB-GENE-040426-695
caprin1b	0.90412	2.561E-08	1.141E-06	ZDB-GENE-040426-2020
parp1	0.90303	4.011E-07	1.344E-05	ZDB-GENE-030131-3955
rrbp1b	0.90279	5.526E-05	1.005E-03	ZDB-GENE-030429-36
cdca8	0.89765	1.252E-06	3.696E-05	ZDB-GENE-041114-171
racgap1	0.89691	1.716E-05	3.655E-04	ZDB-GENE-030131-1917
mfsd2ab	0.89317	1.010E-04	1.690E-03	ZDB-GENE-040801-89
tfap2c	0.89104	2.843E-03	2.657E-02	ZDB-GENE-041212-61
dpf3	0.89025	3.017E-03	2.772E-02	ZDB-GENE-041014-190
ska3	0.88790	1.251E-04	2.040E-03	ZDB-GENE-030131-2997
primpol	0.88741	3.247E-04	4.602E-03	ZDB-GENE-051113-100
actn1	0.88679	2.255E-05	4.647E-04	ZDB-GENE-081113-4
hand2	0.88573	3.505E-04	4.915E-03	ZDB-GENE-000511-1
nasp	0.88526	2.280E-08	1.032E-06	ZDB-GENE-030131-5511
pold2	0.88241	8.057E-05	1.385E-03	ZDB-GENE-030616-157
ercc6l	0.88176	3.693E-04	5.148E-03	ZDB-GENE-060531-56
tgm2b	0.88125	3.857E-03	3.365E-02	ZDB-GENE-030131-2576
mlh1	0.88102	6.461E-04	8.245E-03	ZDB-GENE-040426-1600
her2	0.88072	2.205E-03	2.198E-02	ZDB-GENE-980526-274
kif22	0.87938	3.280E-07	1.122E-05	ZDB-GENE-080204-34
btg3	0.87757	3.020E-04	4.314E-03	ZDB-GENE-031113-19
zgc:110540	0.87267	2.120E-04	3.213E-03	ZDB-GENE-050327-77
zgc:55461	0.87250	6.310E-06	1.552E-04	ZDB-GENE-040426-2879
p4ha2	0.87031	4.585E-04	6.227E-03	ZDB-GENE-030827-1
spc24	0.86877	6.171E-05	1.109E-03	ZDB-GENE-050522-456
orc5	0.86722	9.946E-04	1.175E-02	ZDB-GENE-040426-2275
armc1l	0.85913	1.127E-05	2.589E-04	ZDB-GENE-030131-9819
gtse1	0.85802	1.193E-04	1.960E-03	ZDB-GENE-050522-493
ska1	0.85709	4.829E-04	6.503E-03	ZDB-GENE-040718-334
atic	0.85673	1.323E-03	1.477E-02	ZDB-GENE-011212-4
knstrn	0.85607	3.337E-03	2.997E-02	ZDB-GENE-081022-187
unga	0.85561	4.604E-04	6.246E-03	ZDB-GENE-040426-900
dscc1	0.85468	1.580E-04	2.502E-03	ZDB-GENE-040912-35
cdk2	0.85379	1.918E-06	5.358E-05	ZDB-GENE-040426-2741
dnaja1	0.85276	3.708E-03	3.270E-02	ZDB-GENE-030131-2642
zwilch	0.85262	5.294E-04	7.007E-03	ZDB-GENE-050420-101
rad51	0.85233	2.595E-05	5.251E-04	ZDB-GENE-040426-2286
ankrd12	0.85144	6.582E-05	1.159E-03	ZDB-GENE-040426-846
hmgn3	0.85051	1.546E-05	3.376E-04	ZDB-GENE-041114-180
zgc:173506	0.84798	2.687E-04	3.903E-03	ZDB-GENE-030131-7120
mad2l1	0.84770	2.234E-06	6.125E-05	ZDB-GENE-030515-3
pacsin1b	0.84550	5.888E-03	4.729E-02	ZDB-GENE-050913-35
ccnb1	0.84405	1.580E-06	4.545E-05	ZDB-GENE-000406-10
zgc:114060	0.84292	1.417E-03	1.565E-02	ZDB-GENE-050913-92
cdc6	0.84267	2.126E-05	4.413E-04	ZDB-GENE-050506-131
lin9	0.84235	5.783E-03	4.661E-02	ZDB-GENE-030131-1747
prox1a	0.84036	3.002E-03	2.765E-02	ZDB-GENE-980526-397

egr1	0.84005	2.202E-03	2.198E-02	ZDB-GENE-980526-320
pogza	0.83990	3.933E-03	3.419E-02	ZDB-GENE-040914-76
gins2	0.83956	2.737E-04	3.963E-03	ZDB-GENE-050419-19
snapc4	0.83428	1.110E-03	1.278E-02	ZDB-GENE-030131-5794
zgc:66125	0.83400	3.032E-05	6.019E-04	ZDB-GENE-040426-1545
nuf2	0.83252	2.440E-04	3.609E-03	ZDB-GENE-040426-1121
oxct1a	0.83064	4.937E-04	6.606E-03	ZDB-GENE-040723-3
lypd6	0.83056	1.907E-03	1.979E-02	ZDB-GENE-040912-117
kif23	0.82858	1.999E-06	5.552E-05	ZDB-GENE-991019-4
ssrp1a	0.82829	3.708E-07	1.251E-05	ZDB-GENE-031118-9
kif2c	0.82715	4.776E-05	8.913E-04	ZDB-GENE-070912-298
med16	0.82480	1.534E-04	2.447E-03	ZDB-GENE-090306-4
fance	0.82297	6.258E-03	4.921E-02	ZDB-GENE-060510-5
etv5a	0.82090	2.879E-05	5.763E-04	ZDB-GENE-030131-9572
fanca	0.81606	2.316E-03	2.276E-02	ZDB-GENE-060508-1
dtl	0.81560	5.129E-04	6.820E-03	ZDB-GENE-020419-34
kif14	0.81555	1.361E-03	1.513E-02	ZDB-GENE-030131-5759
whsc1	0.81524	1.208E-05	2.740E-04	ZDB-GENE-030131-2581
raver1	0.81416	2.369E-03	2.309E-02	ZDB-GENE-050522-99
dnajc9	0.81368	7.914E-04	9.792E-03	ZDB-GENE-040718-130
wu:fc13c02	0.81291	1.221E-03	1.381E-02	ZDB-GENE-090313-189
msh6	0.81285	1.671E-05	3.569E-04	ZDB-GENE-020905-3
nat8l	0.81230	4.984E-03	4.129E-02	ZDB-GENE-030729-4
ankrd10b	0.80980	8.531E-04	1.041E-02	ZDB-GENE-061013-49
prim1	0.80901	4.187E-04	5.761E-03	ZDB-GENE-990603-6
smarca5	0.80846	9.178E-07	2.800E-05	ZDB-GENE-021125-1
snrpa1	0.80829	9.294E-06	2.187E-04	ZDB-GENE-041114-82
hs3st3b1a	0.80711	3.830E-03	3.349E-02	ZDB-GENE-070202-4
melk	0.80596	6.319E-04	8.128E-03	ZDB-GENE-990603-5
rbb4	0.80378	6.755E-07	2.146E-05	ZDB-GENE-030131-848
usp1	0.80360	7.217E-06	1.734E-04	ZDB-GENE-030131-726
aldh4a1	0.79918	1.137E-03	1.301E-02	ZDB-GENE-040426-1179
abi1b	0.79849	5.222E-04	6.924E-03	ZDB-GENE-060929-1182
nme2b.1	0.79734	1.899E-05	3.992E-04	ZDB-GENE-000210-32
rtkn2a	0.79639	1.144E-04	1.892E-03	ZDB-GENE-041010-225
dnmt3b	0.79573	2.389E-05	4.875E-04	ZDB-GENE-050314-4
ckap5	0.79314	7.143E-06	1.722E-04	ZDB-GENE-051120-174
zgc:101640	0.79310	2.347E-03	2.295E-02	ZDB-GENE-041114-158
tpi1b	0.79297	7.558E-06	1.805E-04	ZDB-GENE-020416-4
rad54l	0.79265	1.022E-03	1.197E-02	ZDB-GENE-040426-968
znfl2a	0.79246	2.998E-03	2.765E-02	ZDB-GENE-030828-7
dlg1	0.79081	3.820E-03	3.342E-02	ZDB-GENE-010724-8
orc4	0.79069	1.826E-04	2.838E-03	ZDB-GENE-040426-2224
zgc:92411	0.78900	5.326E-05	9.753E-04	ZDB-GENE-040808-52
marcksb	0.78855	9.304E-07	2.827E-05	ZDB-GENE-030131-1921
tut1	0.78722	7.950E-04	9.828E-03	ZDB-GENE-050706-68
sf3b3	0.78610	1.093E-04	1.818E-03	ZDB-GENE-040426-2901
ppa1b	0.78512	1.184E-04	1.950E-03	ZDB-GENE-050417-375

slc1a3a	0.78489	1.206E-03	1.369E-02	ZDB-GENE-030131-2159
zgc:112332	0.78422	2.936E-04	4.210E-03	ZDB-GENE-050522-387
cdh2	0.78094	2.138E-04	3.237E-03	ZDB-GENE-990415-171
ppig	0.78075	1.633E-05	3.519E-04	ZDB-GENE-040426-2822
dbf4	0.77936	2.397E-03	2.324E-02	ZDB-GENE-091020-13
seta	0.77923	1.217E-05	2.752E-04	ZDB-GENE-030131-2221
smc3	0.77788	2.261E-06	6.188E-05	ZDB-GENE-030131-3196
cd2bp2	0.77711	1.044E-04	1.742E-03	ZDB-GENE-040426-925
hs2st1a	0.77695	2.379E-03	2.315E-02	ZDB-GENE-070112-2312
ghra	0.77399	4.401E-03	3.726E-02	ZDB-GENE-070509-1
ahcy	0.77240	3.993E-06	1.040E-04	ZDB-GENE-031219-6
jam2a	0.77014	1.776E-05	3.755E-04	ZDB-GENE-031204-3
dido1	0.76910	2.939E-05	5.866E-04	ZDB-GENE-030131-6117
lbr	0.76669	4.237E-06	1.094E-04	ZDB-GENE-030804-11
tagln3b	0.76664	2.350E-03	2.297E-02	ZDB-GENE-030131-4199
orc1	0.76659	6.890E-04	8.679E-03	ZDB-GENE-030131-6960
cep135	0.76371	1.287E-03	1.444E-02	ZDB-GENE-041210-325
pyroxd2	0.76318	9.226E-04	1.110E-02	ZDB-GENE-050506-147
zgc:101116	0.76238	1.297E-03	1.454E-02	ZDB-GENE-040801-78
tnpo2	0.75990	1.276E-04	2.076E-03	ZDB-GENE-040822-16
ppm1g	0.75921	3.022E-06	8.057E-05	ZDB-GENE-030425-4
ncapg	0.75881	4.153E-05	7.949E-04	ZDB-GENE-041111-282
mybl2	0.75830	3.012E-04	4.310E-03	ZDB-GENE-041007-1
cdc14b	0.75629	1.207E-05	2.740E-04	ZDB-GENE-040426-820
asf1bb	0.75555	5.965E-05	1.075E-03	ZDB-GENE-040426-2594
dnmt3	0.75514	3.622E-04	5.070E-03	ZDB-GENE-990712-11
cdc7	0.75366	1.789E-03	1.890E-02	ZDB-GENE-041114-113
cenpa	0.75352	1.059E-03	1.233E-02	-
exo1	0.75327	3.842E-03	3.356E-02	ZDB-GENE-040426-2828
zgc:174906	0.74989	3.021E-03	2.773E-02	ZDB-GENE-030131-6125
carm1	0.74982	1.912E-04	2.954E-03	ZDB-GENE-040724-77
cenpk	0.74892	3.290E-03	2.966E-02	ZDB-GENE-090313-204
gmnn	0.74879	2.353E-03	2.299E-02	ZDB-GENE-030429-30
hsd17b12a	0.74842	6.042E-03	4.803E-02	ZDB-GENE-030131-5628
pdia3	0.74838	5.435E-04	7.168E-03	ZDB-GENE-031002-9
hoxd12a	0.74571	6.403E-04	8.178E-03	ZDB-GENE-990415-118
rif1	0.74521	6.416E-05	1.142E-03	ZDB-GENE-030131-3406
rad18	0.74218	5.766E-03	4.655E-02	ZDB-GENE-040426-745
cdca5	0.74078	2.481E-04	3.657E-03	ZDB-GENE-030131-235
nucks1b	0.73964	1.707E-04	2.670E-03	ZDB-GENE-030131-4544
ect2	0.73798	1.965E-03	2.022E-02	ZDB-GENE-040831-2
kalrnb	0.73686	4.606E-05	8.641E-04	ZDB-GENE-060421-7244
tmpob	0.73607	7.485E-05	1.298E-03	ZDB-GENE-050522-8
tprb	0.73438	4.131E-05	7.922E-04	ZDB-GENE-030131-6410
ilf3a	0.73401	1.399E-04	2.254E-03	ZDB-GENE-030131-3151
chtf18	0.73356	8.731E-04	1.060E-02	ZDB-GENE-050522-508
calr	0.73317	4.772E-05	8.913E-04	ZDB-GENE-030131-4042
tmpoa	0.73219	3.614E-05	7.041E-04	ZDB-GENE-030131-6003

naprt1	0.73114	4.635E-03	3.888E-02	ZDB-GENE-040426-1897
smc1al	0.72993	1.047E-05	2.420E-04	ZDB-GENE-040426-57
gart	0.72846	3.146E-04	4.481E-03	ZDB-GENE-000616-14
etaa1	0.72744	5.161E-03	4.254E-02	ZDB-GENE-080204-12
impdh1b	0.72686	9.967E-05	1.674E-03	ZDB-GENE-030219-206
chek1	0.72364	2.472E-03	2.377E-02	ZDB-GENE-040426-896
arpp19b	0.72310	1.593E-04	2.513E-03	ZDB-GENE-040426-930
aldh7a1	0.72257	1.384E-04	2.235E-03	ZDB-GENE-030131-6129
pbxip1b	0.72066	4.189E-03	3.579E-02	ZDB-GENE-070112-2032
zgc:101819	0.71796	7.311E-04	9.131E-03	ZDB-GENE-030131-740
agpat2	0.71712	1.766E-03	1.872E-02	ZDB-GENE-061103-541
ergic3	0.71575	9.888E-04	1.170E-02	ZDB-GENE-040426-795
nap1l1	0.71377	1.167E-05	2.660E-04	ZDB-GENE-030516-2
paxip1	0.71230	8.006E-04	9.872E-03	ZDB-GENE-041010-214
alcamb	0.71193	2.084E-04	3.172E-03	ZDB-GENE-030131-1768
rhogd	0.71064	2.189E-03	2.191E-02	ZDB-GENE-040426-1504
gli3	0.71063	1.831E-03	1.924E-02	ZDB-GENE-041111-162
dek	0.70701	2.187E-04	3.298E-03	ZDB-GENE-060512-191
mastl	0.70286	1.975E-03	2.025E-02	ZDB-GENE-040801-128
ogt.1	0.69815	1.879E-03	1.958E-02	ZDB-GENE-030131-9631
klf2b	0.69762	2.498E-03	2.398E-02	ZDB-GENE-011109-2
ing5b	0.69741	2.695E-03	2.549E-02	ZDB-GENE-030616-462
rpa3	0.69594	3.432E-04	4.831E-03	ZDB-GENE-040426-977
e2f7	0.69559	2.162E-03	2.173E-02	ZDB-GENE-030131-3527
zgc:171223	0.69105	2.306E-03	2.269E-02	ZDB-GENE-070928-7
snrnp200	0.69047	1.641E-04	2.585E-03	ZDB-GENE-081105-64
mcmbp	0.69037	1.037E-03	1.210E-02	ZDB-GENE-030131-9676
h2afvb	0.69000	1.236E-04	2.018E-03	ZDB-GENE-050506-24
hoxb1a	0.68917	1.425E-03	1.571E-02	ZDB-GENE-990415-101
ak1	0.68782	4.328E-03	3.681E-02	ZDB-GENE-040822-37
birc5a	0.68554	1.145E-03	1.309E-02	ZDB-GENE-030826-1
smc4	0.68481	8.397E-05	1.433E-03	ZDB-GENE-020419-21
anp32e	0.68295	4.317E-05	8.182E-04	ZDB-GENE-040426-2448
mgaa	0.67988	2.177E-03	2.181E-02	ZDB-GENE-030603-1
dll4	0.67786	2.837E-04	4.089E-03	ZDB-GENE-041014-73
chtf8	0.67632	4.103E-03	3.528E-02	ZDB-GENE-050419-148
men1	0.67524	5.905E-03	4.736E-02	ZDB-GENE-991110-12
rps29	0.67513	1.370E-04	2.215E-03	ZDB-GENE-040622-5
aqr	0.67407	3.956E-03	3.429E-02	ZDB-GENE-040426-1189
nploc4	0.67198	9.724E-04	1.156E-02	ZDB-GENE-030131-5658
prkcbp1l	0.67070	3.804E-04	5.268E-03	ZDB-GENE-030131-2199
pgam1a	0.67026	6.028E-03	4.800E-02	ZDB-GENE-030131-1827
fabp3	0.66688	1.145E-04	1.892E-03	ZDB-GENE-020318-2
zgc:162740	0.66671	4.599E-03	3.866E-02	ZDB-GENE-030131-6219
tp63	0.66583	1.248E-03	1.407E-02	ZDB-GENE-030819-1
e2f4	0.66449	1.010E-03	1.189E-02	ZDB-GENE-040426-2773
agps	0.66167	5.303E-03	4.349E-02	ZDB-GENE-031118-14
zpr1	0.66151	4.125E-03	3.539E-02	ZDB-GENE-040426-2110

pop7	0.66128	1.505E-03	1.644E-02	ZDB-GENE-041010-88
chaf1b	0.65838	1.779E-03	1.881E-02	ZDB-GENE-040426-1947
nol6	0.65810	8.728E-04	1.060E-02	ZDB-GENE-030131-6294
asph	0.65620	2.662E-03	2.522E-02	ZDB-GENE-031112-5
cse1l	0.65269	2.881E-04	4.144E-03	ZDB-GENE-990603-1
tpm4b	0.64950	2.801E-03	2.630E-02	ZDB-GENE-030131-2964
nup107	0.64813	1.305E-04	2.117E-03	ZDB-GENE-041210-245
icn	0.64636	1.700E-03	1.819E-02	ZDB-GENE-030131-8599
ticrr	0.64495	4.105E-03	3.528E-02	ZDB-GENE-040728-1
hoxa4a	0.64009	6.263E-03	4.923E-02	ZDB-GENE-000823-4
tshz3b	0.63846	2.975E-03	2.750E-02	ZDB-GENE-100319-2
ipo9	0.63844	4.875E-04	6.553E-03	ZDB-GENE-040426-2953
kpnb3	0.63792	1.068E-03	1.242E-02	ZDB-GENE-030424-2
shcbp1	0.63791	1.633E-03	1.753E-02	ZDB-GENE-030131-5865
cep57l1	0.63565	5.924E-03	4.738E-02	ZDB-GENE-040808-27
rfc4	0.63376	8.527E-04	1.041E-02	ZDB-GENE-040824-3
cbx2	0.63233	1.008E-03	1.186E-02	ZDB-GENE-030131-5502
zgc:158604	0.63217	2.912E-03	2.706E-02	ZDB-GENE-070112-882
entpd1	0.63214	4.165E-04	5.736E-03	ZDB-GENE-040801-58
bora	0.63203	5.548E-03	4.511E-02	ZDB-GENE-041114-165
prpf8	0.63114	1.705E-04	2.670E-03	ZDB-GENE-030131-577
lrpprc	0.63100	1.107E-03	1.276E-02	ZDB-GENE-030131-7887
slbp	0.62964	2.051E-03	2.091E-02	ZDB-GENE-030131-9686
zgc:92139	0.62546	3.276E-04	4.639E-03	ZDB-GENE-040718-47
safb	0.62469	1.926E-03	1.993E-02	ZDB-GENE-030131-698
slc25a22	0.62246	2.226E-03	2.212E-02	ZDB-GENE-040426-2745
zgc:110216	0.62118	3.109E-03	2.848E-02	ZDB-GENE-050320-42
spryd4	0.62099	6.199E-03	4.883E-02	ZDB-GENE-040718-180
angptl4	0.62042	5.326E-04	7.043E-03	ZDB-GENE-041111-222
slc38a5b	0.61764	6.336E-04	8.132E-03	ZDB-GENE-040718-395
mdh1aa	0.61645	2.973E-03	2.750E-02	ZDB-GENE-040204-1
igf2bp3	0.61599	1.305E-03	1.460E-02	ZDB-GENE-000308-1
skiv2l2	0.61429	6.666E-04	8.448E-03	ZDB-GENE-040426-2854
aifm1	0.61428	1.918E-03	1.989E-02	ZDB-GENE-030826-11
polr3glb	0.61322	5.400E-03	4.409E-02	ZDB-GENE-041010-32
pinx1	0.61295	1.771E-03	1.875E-02	ZDB-GENE-030515-2
nipblb	0.61155	2.493E-04	3.669E-03	ZDB-GENE-030131-6070
spc25	0.61039	4.216E-03	3.601E-02	ZDB-GENE-060810-13
rpa1	0.60978	1.494E-04	2.391E-03	ZDB-GENE-030912-3
usp33	0.60952	2.849E-03	2.660E-02	ZDB-GENE-040426-2323
prp19	0.60739	2.247E-04	3.375E-03	ZDB-GENE-030131-263
hadhab	0.60718	1.318E-03	1.472E-02	ZDB-GENE-041111-204
noc2l	0.60702	2.788E-03	2.620E-02	ZDB-GENE-000607-75
nup155	0.60696	2.456E-03	2.369E-02	ZDB-GENE-040426-711
her12	0.60486	9.363E-04	1.123E-02	ZDB-GENE-040824-5
crnkl1	0.60322	7.358E-04	9.166E-03	ZDB-GENE-040426-694
naa15b	0.60194	2.357E-03	2.301E-02	ZDB-GENE-030131-6031
tfam	0.59945	1.213E-03	1.373E-02	ZDB-GENE-061013-552

polr2h	0.59943	4.867E-03	4.048E-02	ZDB-GENE-050417-66
atrx	0.59559	3.782E-04	5.248E-03	ZDB-GENE-030912-11
arid4a	0.59464	6.393E-04	8.173E-03	ZDB-GENE-050809-105
tra2b	0.59402	3.256E-03	2.946E-02	ZDB-GENE-040426-1094
ncapg2	0.59347	5.649E-03	4.583E-02	ZDB-GENE-070410-100
foxp4	0.58629	2.929E-03	2.717E-02	ZDB-GENE-101207-1
ddx39b	0.58183	4.915E-04	6.585E-03	ZDB-GENE-040426-1259
hnrpdl	0.58057	2.734E-03	2.579E-02	ZDB-GENE-040426-2717
zgc:91897	0.57880	6.094E-03	4.824E-02	ZDB-GENE-040711-1
sf3b5	0.57547	2.637E-03	2.505E-02	ZDB-GENE-040718-181
gata3	0.57511	1.072E-03	1.243E-02	ZDB-GENE-990415-82
galnt7	0.57509	4.733E-03	3.952E-02	ZDB-GENE-050522-284
ints7	0.57440	6.057E-03	4.808E-02	ZDB-GENE-021220-6
pane1	0.57137	4.687E-03	3.923E-02	ZDB-GENE-040611-4
nutf2l	0.57079	2.204E-03	2.198E-02	ZDB-GENE-020416-1
rpa2	0.56988	5.409E-04	7.140E-03	ZDB-GENE-010131-3
map2k6	0.56967	1.034E-03	1.209E-02	ZDB-GENE-010202-3
zw10	0.56795	6.339E-03	4.969E-02	ZDB-GENE-040426-800
cbx3a	0.56421	5.264E-04	6.974E-03	ZDB-GENE-030131-1158
ran	0.56417	2.757E-03	2.596E-02	ZDB-GENE-990415-88
cdc25	0.55762	4.365E-03	3.709E-02	ZDB-GENE-000330-5
tubg1	0.55728	2.636E-03	2.505E-02	ZDB-GENE-040426-836
wsb1	0.55616	4.383E-03	3.719E-02	ZDB-GENE-030131-1946
zgc:114188	0.55607	8.015E-04	9.873E-03	ZDB-GENE-050320-15
adh5	0.55590	8.597E-04	1.048E-02	ZDB-GENE-011003-1
si:dkey-67c22.2	0.55418	6.252E-04	8.064E-03	ZDB-GENE-030131-5845
sephs1	0.55254	1.168E-03	1.331E-02	ZDB-GENE-030131-3670
supt16h	0.55036	6.138E-03	4.851E-02	ZDB-GENE-031118-96
si:ch73-46j18.5	0.54927	3.586E-03	3.182E-02	ZDB-GENE-030131-7892
hnrnpul1	0.54865	1.401E-03	1.549E-02	ZDB-GENE-040426-2432
si:ch211-114c12.2	0.54694	4.023E-03	3.472E-02	ZDB-GENE-030131-8522
idh2	0.54647	6.568E-04	8.339E-03	ZDB-GENE-031118-95
immt	0.54618	1.537E-03	1.673E-02	ZDB-GENE-030131-5417
cks1b	0.54590	5.662E-03	4.591E-02	ZDB-GENE-040801-112
acaa2	0.54456	4.435E-03	3.744E-02	ZDB-GENE-040426-1996
nucks1a	0.54082	9.301E-04	1.117E-02	ZDB-GENE-040912-175
rcor2	0.54055	2.173E-03	2.181E-02	ZDB-GENE-040426-1812
ncalda	0.53786	2.270E-03	2.242E-02	ZDB-GENE-080220-28
chd4a	0.53679	1.967E-03	2.022E-02	ZDB-GENE-041111-187
setdb1a	0.53671	4.041E-03	3.486E-02	ZDB-GENE-030131-2421
dnajb1b	0.53431	4.680E-03	3.919E-02	ZDB-GENE-030131-5455
hdgfrp2	0.53283	3.278E-03	2.959E-02	ZDB-GENE-040426-2104
rcc2	0.53184	2.234E-03	2.218E-02	ZDB-GENE-040426-2213
ddx19	0.53031	2.806E-03	2.632E-02	ZDB-GENE-020419-30
brd3a	0.52353	1.674E-03	1.792E-02	ZDB-GENE-030131-6141
h3f3a	0.52218	3.247E-03	2.941E-02	ZDB-GENE-040426-1928
ezh2	0.52079	4.978E-03	4.126E-02	ZDB-GENE-041111-259
mbd3b	0.51429	4.173E-03	3.568E-02	ZDB-GENE-030131-9792

ruvbl1	0.51329	4.523E-03	3.809E-02	ZDB-GENE-030109-2
prdx4	0.51217	6.036E-03	4.801E-02	ZDB-GENE-030131-1096
eif4a3	0.51059	1.748E-03	1.857E-02	ZDB-GENE-040426-915
ctcf	0.51011	3.126E-03	2.858E-02	ZDB-GENE-040624-5
ktn1	0.50758	3.773E-03	3.311E-02	ZDB-GENE-030408-4
tfdp1a	0.50682	4.719E-03	3.944E-02	ZDB-GENE-040426-2883
zgc:92107	0.50579	3.434E-03	3.065E-02	ZDB-GENE-041114-4
tuba1l2	0.50532	4.610E-03	3.873E-02	ZDB-GENE-090507-4
sec61al2	0.50064	3.855E-03	3.365E-02	ZDB-GENE-021016-2
rtn4a	0.50016	1.928E-03	1.993E-02	ZDB-GENE-030710-1
tex10	0.49838	6.071E-03	4.813E-02	ZDB-GENE-040426-2827
nup205	0.49781	5.996E-03	4.780E-02	ZDB-GENE-040715-8
fkbp9	0.49529	3.898E-03	3.394E-02	ZDB-GENE-040801-23
puf60b	0.49466	2.327E-03	2.284E-02	ZDB-GENE-040625-116
ppp6r3	0.49025	2.988E-03	2.758E-02	ZDB-GENE-030131-3251
uba2	0.48870	3.282E-03	2.961E-02	ZDB-GENE-040426-2681
sap30bp	0.48841	6.349E-03	4.973E-02	ZDB-GENE-030131-457
egfl6	0.48768	5.987E-03	4.778E-02	ZDB-GENE-040718-157
thoc2	0.48462	3.523E-03	3.135E-02	ZDB-GENE-030616-54
sae1	0.48202	5.805E-03	4.673E-02	ZDB-GENE-040625-21
emp2	0.48042	3.993E-03	3.450E-02	ZDB-GENE-040822-24
srrm1	0.47981	2.370E-03	2.309E-02	ZDB-GENE-040426-2789
pleca	0.47893	3.321E-03	2.988E-02	ZDB-GENE-030131-8984
atp5h	0.47017	4.299E-03	3.665E-02	ZDB-GENE-040426-1658
cpsf5	0.46591	5.456E-03	4.451E-02	ZDB-GENE-040426-1316
cbx5	0.46128	5.910E-03	4.737E-02	ZDB-GENE-030131-5553
rbb4l	0.45896	5.304E-03	4.349E-02	ZDB-GENE-030131-848
banf1	0.45818	4.435E-03	3.744E-02	ZDB-GENE-030131-6657
uqcrc1	0.45639	4.942E-03	4.101E-02	ZDB-GENE-040426-1792

Appendix III: Supplementary Table 5.1

Supplementary Table 5.1. Expression levels in *kdrl*:GFP^{high}; *gata1*:DsRed^{-} cells of selected genes identified as differentially expressed upon *lmo4a* morphants in the RNA-seq experiment, measured with Fluidigm Biomark™ qRT-PCR at 22hpf and 28hpf.

gene	stage	cell population	sample	normalized mRNA levels
cdk2	22hpf	kdrl+	wt	.09719301
			MO	.05762175
		neg	wt	.09498036
			MO	.06789121
	28hpf	kdrl+	wt	.08205283
			MO	.05942641
		double++	wt	.04878083
			MO	.03388502
		neg	wt	.09520103
			MO	.07066272
cttnb2	22hpf	kdrl+	wt	.68482627
			MO	.68459846
		neg	wt	.35410035
			MO	.38567632
	28hpf	kdrl+	wt	.69133216
			MO	.70129769
		double++	wt	.37309595
			MO	.43646125
		neg	wt	.41709661
			MO	.43154850
dld	22hpf	kdrl+	wt	.14350686
			MO	.19286572
		neg	wt	.33893165
			MO	.21468533
	28hpf	kdrl+	wt	.15511713
			MO	.28094030
		double++	wt	.08670278
			MO	.18246482
		neg	wt	.46899930
			MO	.33408529
efnb2a	22hpf	kdrl+	wt	.41544781
			MO	.36763413
		neg	wt	.14118292
			MO	.15453869
	28hpf	kdrl+	wt	.25457431
			MO	.25523364
		double++	wt	.08031958
			MO	.08728428
		neg	wt	.12297084
			MO	.12944556

erg	22hpf	kdrl+	wt	.43214307
			MO	.49923592
		neg	wt	.00601185
			MO	.00391236
	28hpf	kdrl+	wt	.44578007
			MO	.44199976
		double++	wt	.24277390
			MO	.23252910
		neg	wt	.00700443
			MO	.00614148
etv2	22hpf	kdrl+	wt	1.37815827
			MO	2.14966101
		neg	wt	.04983012
			MO	.04478933
	28hpf	kdrl+	wt	1.14017983
			MO	1.89803227
		double++	wt	.48456735
			MO	.76060452
		neg	wt	.02029001
			MO	.02248230
fzd3a	22hpf	kdrl+	wt	.12844068
			MO	.12089206
		neg	wt	.39779533
			MO	.37148556
	28hpf	kdrl+	wt	.11171895
			MO	.12585519
		double++	wt	.03349371
			MO	.03020559
		neg	wt	.45285654
			MO	.39826019
fzd7a	22hpf	kdrl+	wt	1.02384167
			MO	.97446298
		neg	wt	.79147953
			MO	.78986685
	28hpf	kdrl+	wt	.86000990
			MO	.95848717
		double++	wt	.22955483
			MO	.25427804
		neg	wt	.77253051
			MO	.75572470
gata2a	22hpf	kdrl+	wt	.23714944
			MO	.26996354
		neg	wt	.04312105
			MO	.04667795
	28hpf	kdrl+	wt	.16708073

			MO	.20401085
		double++	wt	.07719211
			MO	.07976240
		neg	wt	.05000793
			MO	.04249855
her2	22hpf	kdrl+	wt	.23188931
			MO	.14404444
		neg	wt	.48697247
			MO	.39913556
	28hpf	kdrl+	wt	.08925223
			MO	.09679281
		double++	wt	.05102027
			MO	.04389082
		neg	wt	.43765938
			MO	.43029988
hhex	22hpf	kdrl+	wt	.03321575
			MO	.04455577
		neg	wt	.00253562
			MO	.00154753
	28hpf	kdrl+	wt	.03263720
			MO	.04017012
		double++	wt	.01491722
			MO	.01912836
		neg	wt	.00299825
			MO	.00309329
hoxa4a	22hpf	kdrl+	wt	.07415402
			MO	.08061492
		neg	wt	.13398649
			MO	.12369734
	28hpf	kdrl+	wt	.12669839
			MO	.11411517
		double++	wt	.02417739
			MO	.02659046
		neg	wt	.15359877
			MO	.15200621
ikzf1	22hpf	kdrl+	wt	.05067816
			MO	.04256884
		neg	wt	.00447837
			MO	.00371359
	28hpf	kdrl+	wt	.03069109
			MO	.02176853
		double++	wt	.39742320
			MO	.55675964
		neg	wt	.00367912
			MO	.00386200

irf2bp2a	22hpf	kdrl+	wt	.31806234
			MO	.37836614
		neg	wt	.10294789
			MO	.11365025
	28hpf	kdrl+	wt	.37843167
			MO	.43372090
		double++	wt	.24362031
			MO	.27435267
		neg	wt	.12015374
			MO	.12507191
kat2b	22hpf	kdrl+	wt	.12095644
			MO	.11977382
		neg	wt	.03681342
			MO	.02855068
	28hpf	kdrl+	wt	.11909738
			MO	.15537696
		double++	wt	.09939318
			MO	.10934767
		neg	wt	.05144461
			MO	.04176013
kdrl	22hpf	kdrl+	wt	.97645819
			MO	1.08438309
		neg	wt	.01680581
			MO	.01476943
	28hpf	kdrl+	wt	.95347370
			MO	1.05328987
		double++	wt	.48324885
			MO	.61503576
		neg	wt	.02033258
			MO	.02392026
mycla	22hpf	kdrl+	wt	.25422729
			MO	.18864321
		neg	wt	.25316105
			MO	.18551559
	28hpf	kdrl+	wt	.23654043
			MO	.20800094
		double++	wt	.11173520
			MO	.12106054
		neg	wt	.28333469
			MO	.23759246
nr2f1a	22hpf	kdrl+	wt	.05583221
			MO	.05544140
		neg	wt	.04932438
			MO	.03383235
	28hpf	kdrl+	wt	.07688905

			MO	.09879097
		double++	wt	.02657190
			MO	.02701146
		neg	wt	.07828175
			MO	.05638307
numb	22hpf	kdrl+	wt	.05050948
			MO	.04999902
		neg	wt	.01088005
			MO	.00970911
	28hpf	kdrl+	wt	.04809855
			MO	.05598432
		double++	wt	.03361896
			MO	.04359746
		neg	wt	.01545595
			MO	.01618596
sox7	22hpf	kdrl+	wt	.86798685
			MO	1.17138469
		neg	wt	.01462872
			MO	.01180214
	28hpf	kdrl+	wt	.53271010
			MO	.80624829
		double++	wt	.18732937
			MO	.33907619
		neg	wt	.01381170
			MO	.01594886
tcf7l1a	22hpf	kdrl+	wt	.37496496
			MO	.41295791
		neg	wt	.23499299
			MO	.21613369
	28hpf	kdrl+	wt	.31300611
			MO	.34735672
		double++	wt	.13207916
			MO	.14633042
		neg	wt	.25681354
			MO	.24094389
tgfb3	22hpf	kdrl+	wt	.24812620
			MO	.19264567
		neg	wt	.07077362
			MO	.06283640
	28hpf	kdrl+	wt	.17614061
			MO	.20139613
		double++	wt	.02571965
			MO	.02362559
		neg	wt	.13505022
			MO	.13251898

ascl1a	22hpf	kdrl+	wt	.01304848
			MO	.01207314
		neg	wt	.12763095
			MO	.11468488
	28hpf	kdrl+	wt	.02409183
			MO	.02159077
		double++	wt	.00792307
			MO	.00567812
		neg	wt	.24406630
			MO	.22774148
cbx1b	22hpf	kdrl+	wt	.06950944
			MO	.04372507
		neg	wt	.08286049
			MO	.05492304
	28hpf	kdrl+	wt	.06549633
			MO	.05848895
		double++	wt	.03400133
			MO	.02349983
		neg	wt	.11580415
			MO	.07216144
cebpd	22hpf	kdrl+	wt	.00886480
			MO	.02802371
		neg	wt	.02011884
			MO	.03765186
	28hpf	kdrl+	wt	.02158561
			MO	.05887942
		double++	wt	.00278174
			MO	.01386261
		neg	wt	.01882115
			MO	.02802167
dlb	22hpf	kdrl+	wt	.01007651
			MO	.00775053
		neg	wt	.11034597
			MO	.09445833
	28hpf	kdrl+	wt	.02643807
			MO	.02449909
		double++	wt	.00655934
			MO	.00412019
		neg	wt	.20936604
			MO	.17506707
ebf2	22hpf	kdrl+	wt	.00496203
			MO	.00398819
		neg	wt	.05556680
			MO	.04208892
	28hpf	kdrl+	wt	.01285649

			MO	.01293429
		double++	wt	.00317811
			MO	.00128538
		neg	wt	.10210204
			MO	.07016540
ephb2b	22hpf	kdrl+	wt	.02275792
			MO	.01864768
		neg	wt	.08365583
			MO	.07483816
	28hpf	kdrl+	wt	.02997835
			MO	.02508588
		double++	wt	.01829055
			MO	.01067041
		neg	wt	.11723909
			MO	.08641798
gfi1aa	22hpf	kdrl+	wt	.27304752
			MO	.17576682
		neg	wt	.01101266
			MO	.02006251
	28hpf	kdrl+	wt	.20243337
			MO	.15409710
		double++	wt	.50085149
			MO	.41060917
		neg	wt	.01159348
			MO	.02461947
gfi1ab	22hpf	kdrl+	wt	.01115877
			MO	.00854691
		neg	wt	.00318529
			MO	.00483503
	28hpf	kdrl+	wt	.00657848
			MO	.00474818
		double++	wt	.02219343
			MO	.02116466
		neg	wt	.00378858
			MO	.00327586
gna12a	22hpf	kdrl+	wt	.06832308
			MO	.07355897
		neg	wt	.01691850
			MO	.02083915
	28hpf	kdrl+	wt	.08199401
			MO	.08458559
		double++	wt	.03003045
			MO	.03633950
		neg	wt	.01860154
			MO	.01804011

gsx1	22hpf	kdrl+	wt	.01520178
			MO	.01892858
		neg	wt	.03260916
			MO	.02729476
	28hpf	kdrl+	wt	.00690162
			MO	.00665668
		double++	wt	.00260125
			MO	.00104386
		neg	wt	.05024994
			MO	.04137104
hlx1	22hpf	kdrl+	wt	.05946253
			MO	.02264624
		neg	wt	.00270263
			MO	.00185523
	28hpf	kdrl+	wt	.10958172
			MO	.04627512
		double++	wt	.05492494
			MO	.06690336
		neg	wt	.00720836
			MO	.00401907
hoxc13b	22hpf	kdrl+	wt	.00052995
			MO	.00067385
		neg	wt	.01099530
			MO	.00899848
	28hpf	kdrl+	wt	.00466815
			MO	.00326479
		double++	wt	.00068853
			MO	.00017328
		neg	wt	.00765073
			MO	.00801322
hoxc4a	22hpf	kdrl+	wt	.02411085
			MO	.02706137
		neg	wt	.10505593
			MO	.10360864
	28hpf	kdrl+	wt	.05061399
			MO	.04610107
		double++	wt	.00971503
			MO	.00476832
		neg	wt	.19858191
			MO	.17505499
klf17	22hpf	kdrl+	wt	.10865465
			MO	.09200477
		neg	wt	.02891932
			MO	.04733444
	28hpf	kdrl+	wt	.04176259

			MO	.02615482
		double++	wt	.68472424
			MO	.55060389
		neg	wt	.01301383
			MO	.01618687
lbx2	22hpf	kdrl+	wt	.00814747
			MO	.00852463
		neg	wt	.01986834
			MO	.01635549
	28hpf	kdrl+	wt	.01266709
			MO	.02054308
		double++	wt	.00208151
			MO	.00328309
		neg	wt	.02175764
			MO	.02102399
pax6b	22hpf	kdrl+	wt	.00575086
			MO	.00546804
		neg	wt	.06058996
			MO	.04662354
	28hpf	kdrl+	wt	.00771045
			MO	.01161516
		double++	wt	.00271390
			MO	.00148707
		neg	wt	.06205810
			MO	.05403211
prox1a	22hpf	kdrl+	wt	.02312620
			MO	.02049724
		neg	wt	.05519800
			MO	.05670486
	28hpf	kdrl+	wt	.07449492
			MO	.07204746
		double++	wt	.01782763
			MO	.02336734
		neg	wt	.10803285
			MO	.09857885
tbl3	22hpf	kdrl+	wt	.13354464
			MO	.12274465
		neg	wt	.23275965
			MO	.23562554
	28hpf	kdrl+	wt	.11950346
			MO	.10879611
		double++	wt	.15842884
			MO	.13534866
		neg	wt	.17795054
			MO	.21623241

tcf7	22hpf	kdrl+	wt	.04167000
			MO	.03860817
		neg	wt	.10664165
			MO	.11113817
	28hpf	kdrl+	wt	.04514801
			MO	.06000949
		double++	wt	.01666761
			MO	.01546459
		neg	wt	.05285576
			MO	.06080558
hoxc12b	22hpf	kdrl+	wt	.00167276
			MO	.00095589
		neg	wt	.04677588
			MO	.03264281
	28hpf	kdrl+	wt	.03506025
			MO	.01862337
		double++	wt	.00440968
			MO	.00179065
		neg	wt	.02607826
			MO	.02921717
runx1	22hpf	kdrl+	wt	.03770876
			MO	.03860618
		neg	wt	.00659574
			MO	.00417878
	28hpf	kdrl+	wt	.03621078
			MO	.02664858
		double++	wt	.19602741
			MO	.24004401
		neg	wt	.00909490
			MO	.00949304
brd4	22hpf	kdrl+	wt	.70835681
			MO	.74434733
		neg	wt	.46862551
			MO	.48254836
	28hpf	kdrl+	wt	.71893134
			MO	.76364912
		double++	wt	.60561973
			MO	.52688338
		neg	wt	.66023834
			MO	.59888590
btg2	22hpf	kdrl+	wt	1.27945588
			MO	1.94913420
		neg	wt	1.36936158
			MO	1.77671245
	28hpf	kdrl+	wt	1.14220379

			MO	1.94885545
		double++	wt	.91099314
			MO	1.13445168
		neg	wt	1.24230533
			MO	1.31823735
cdc20	22hpf	kdrl+	wt	.21332935
			MO	.12973675
		neg	wt	.24729992
			MO	.16650018
	28hpf	kdrl+	wt	.27086283
			MO	.14504013
		double++	wt	.24614071
			MO	.20525527
		neg	wt	.31560622
			MO	.22012332
ctnnb1	22hpf	kdrl+	wt	1.98822658
			MO	1.83088268
		neg	wt	1.88240036
			MO	1.85522984
	28hpf	kdrl+	wt	1.54977688
			MO	1.86988820
		double++	wt	.87810332
			MO	.86486542
		neg	wt	1.77051908
			MO	1.84473642
dvl2	22hpf	kdrl+	wt	.16594858
			MO	.15925206
		neg	wt	.20176689
			MO	.18780343
	28hpf	kdrl+	wt	.16144697
			MO	.14370252
		double++	wt	.06437868
			MO	.06054493
		neg	wt	.25144099
			MO	.19588384
ephb4a	22hpf	kdrl+	wt	.61884132
			MO	.60897551
		neg	wt	.21789619
			MO	.20366799
	28hpf	kdrl+	wt	.78064353
			MO	.81226579
		double++	wt	.45145794
			MO	.48218642
		neg	wt	.24624654
			MO	.21736331

etv4	22hpf	kdrl+	wt	.20408877
			MO	.20239559
		neg	wt	.08402023
			MO	.07687284
	28hpf	kdrl+	wt	.15871798
			MO	.15395762
		double++	wt	.02485127
			MO	.02665585
		neg	wt	.08105398
			MO	.07726976
etv5a	22hpf	kdrl+	wt	.22319641
			MO	.18565954
		neg	wt	.29347109
			MO	.29789045
	28hpf	kdrl+	wt	.28831661
			MO	.20088979
		double++	wt	.09193392
			MO	.08603196
		neg	wt	.24735726
			MO	.21850811
foxc1a	22hpf	kdrl+	wt	.48448401
			MO	.54552604
		neg	wt	.09785753
			MO	.09566768
	28hpf	kdrl+	wt	.54646803
			MO	.62307292
		double++	wt	.24092431
			MO	.25214275
		neg	wt	.16322094
			MO	.10376881
foxc1b	22hpf	kdrl+	wt	.09746424
			MO	.10516529
		neg	wt	.14652837
			MO	.17364884
	28hpf	kdrl+	wt	.14127010
			MO	.13739565
		double++	wt	.04907285
			MO	.05351064
		neg	wt	.20219898
			MO	.17245712
foxm1	22hpf	kdrl+	wt	.03214304
			MO	.02649962
		neg	wt	.06503930
			MO	.04366199
	28hpf	kdrl+	wt	.03737543

			MO	.01815568
		double++	wt	.02684834
			MO	.02079894
		neg	wt	.08417121
			MO	.04482315
her4.2	22hpf	kdrl+	wt	.23224178
			MO	.11733785
		neg	wt	1.54938821
			MO	1.06228613
	28hpf	kdrl+	wt	.30623306
			MO	.16836344
		double++	wt	.13758858
			MO	.09339772
		neg	wt	2.87672423
			MO	1.64591728
hes6	22hpf	kdrl+	wt	.04523172
			MO	.05339650
		neg	wt	.19030505
			MO	.14588102
	28hpf	kdrl+	wt	.07061001
			MO	.04593055
		double++	wt	.05115026
			MO	.03443289
		neg	wt	.26767996
			MO	.13105627
jun	22hpf	kdrl+	wt	.94377690
			MO	.94390967
		neg	wt	.34172532
			MO	.49816863
	28hpf	kdrl+	wt	1.33122233
			MO	2.02190340
		double++	wt	1.07068024
			MO	1.29438226
		neg	wt	.51736649
			MO	.71218016
mafba	22hpf	kdrl+	wt	.82490866
			MO	1.02270410
		neg	wt	.10722309
			MO	.11503519
	28hpf	kdrl+	wt	.95006011
			MO	1.42542449
		double++	wt	.48191437
			MO	.83843609
		neg	wt	.11634451
			MO	.09304511

myb	22hpf	kdrl+	wt	.14535139
			MO	.13388288
		neg	wt	.11461475
			MO	.13614279
	28hpf	kdrl+	wt	.13276973
			MO	.06752149
		double++	wt	1.59370356
			MO	1.53310429
		neg	wt	.16287433
			MO	.11628759
nrp1b	22hpf	kdrl+	wt	1.86563003
			MO	1.65509132
		neg	wt	.07024134
			MO	.05870639
	28hpf	kdrl+	wt	1.09978119
			MO	1.20959422
		double++	wt	.37657472
			MO	.41087287
		neg	wt	.14183786
			MO	.09051835
psen1	22hpf	kdrl+	wt	.12360061
			MO	.15293142
		neg	wt	.13967761
			MO	.11719448
	28hpf	kdrl+	wt	.11771035
			MO	.12139678
		double++	wt	.11240409
			MO	.11141548
		neg	wt	.13548945
			MO	.08717683
rcor2	22hpf	kdrl+	wt	.34992885
			MO	.34191209
		neg	wt	.36537344
			MO	.27756234
	28hpf	kdrl+	wt	.26941538
			MO	.26725058
		double++	wt	.26245627
			MO	.19962850
		neg	wt	.39508238
			MO	.30235127
sap30l	22hpf	kdrl+	wt	.59124366
			MO	.48865488
		neg	wt	.39656604
			MO	.36430598
	28hpf	kdrl+	wt	.54353054

			MO	.53928511
		double++	wt	.45569160
			MO	.40046158
		neg	wt	.44180164
			MO	.39409284
setd8a	22hpf	kdrl+	wt	.28171863
			MO	.33695706
		neg	wt	.22894844
			MO	.24416393
	28hpf	kdrl+	wt	.19487236
			MO	.23697663
		double++	wt	.28311479
			MO	.30768146
		neg	wt	.24614845
			MO	.26822233
sox19a	22hpf	kdrl+	wt	.03339227
			MO	.04230187
		neg	wt	.32594730
			MO	.26679064
	28hpf	kdrl+	wt	.03633499
			MO	.01316823
		double++	wt	.01707937
			MO	.00820763
		neg	wt	.33693694
			MO	.24639025
sox4a	22hpf	kdrl+	wt	.33502601
			MO	.40190458
		neg	wt	.15944372
			MO	.20626387
	28hpf	kdrl+	wt	.39849868
			MO	.47145778
		double++	wt	.17300759
			MO	.16613720
		neg	wt	.28422884
			MO	.20678570
spry2	22hpf	kdrl+	wt	.07789921
			MO	.09774837
		neg	wt	.07741793
			MO	.06452301
	28hpf	kdrl+	wt	.08559318
			MO	.07657739
		double++	wt	.04128592
			MO	.03786902
		neg	wt	.08745607
			MO	.05924732

tal1	22hpf	kdrl+	wt	.39753844
			MO	.57473802
		neg	wt	.04869134
			MO	.04761896
	28hpf	kdrl+	wt	.15512333
			MO	.23470482
		double++	wt	.45408799
			MO	.32361752
		neg	wt	.05110147
			MO	.02798170
tmem88a	22hpf	kdrl+	wt	2.04969830
			MO	2.32725193
		neg	wt	.04146007
			MO	.04538070
	28hpf	kdrl+	wt	1.49792378
			MO	2.30036485
		double++	wt	.62534628
			MO	.85846940
		neg	wt	.04472701
			MO	.02797581
vangl2	22hpf	kdrl+	wt	.36188485
			MO	.37942196
		neg	wt	.31562534
			MO	.27715332
	28hpf	kdrl+	wt	.36475374
			MO	.37463296
		double++	wt	.17488581
			MO	.17228298
		neg	wt	.36274575
			MO	.27508234
cbfb	22hpf	kdrl+	wt	.12486460
			MO	.15084250
		neg	wt	.01582811
			MO	.02268621
	28hpf	kdrl+	wt	.07546985
			MO	.10270370
		double++	wt	.11765397
			MO	.14122788
		neg	wt	.03438847
			MO	.02188149
csnk1e	22hpf	kdrl+	wt	.02705254
			MO	.04163201
		neg	wt	.02809610
			MO	.02651821
	28hpf	kdrl+	wt	.06062759

			MO	.10669609
		double++	wt	.05242526
			MO	.09203426
		neg	wt	.05257754
			MO	.03863247
dll4	22hpf	kdrl+	wt	.52512090
			MO	.46344481
		neg	wt	.01770338
			MO	.01846910
	28hpf	kdrl+	wt	.33011183
			MO	.33867355
		double++	wt	.13539650
			MO	.15344614
		neg	wt	.02854171
			MO	.02042474
foxo1a	22hpf	kdrl+	wt	.07107223
			MO	.09256423
		neg	wt	.00555136
			MO	.00588789
	28hpf	kdrl+	wt	.07462831
			MO	.07727491
		double++	wt	.02903392
			MO	.03873558
		neg	wt	.01250207
			MO	.00687185
hoxd4a	22hpf	kdrl+	wt	.03255445
			MO	.02525121
		neg	wt	.16433228
			MO	.16294519
	28hpf	kdrl+	wt	.05229112
			MO	.04224558
		double++	wt	.02812416
			MO	.02774432
		neg	wt	.18808657
			MO	.16325592
lef1	22hpf	kdrl+	wt	.12424367
			MO	.12855554
		neg	wt	.08026764
			MO	.08858806
	28hpf	kdrl+	wt	.14615522
			MO	.27093628
		double++	wt	.04597719
			MO	.05509444
		neg	wt	.09556673
			MO	.09742284

meox1	22hpf	kdrl+	wt	.02397244
			MO	.04327143
		neg	wt	.02953532
			MO	.02554426
	28hpf	kdrl+	wt	.01866383
			MO	.02581027
		double++	wt	.00516236
			MO	.00628545
		neg	wt	.02338237
			MO	.02204389
smad4a	22hpf	kdrl+	wt	.07997981
			MO	.07972524
		neg	wt	.06208751
			MO	.06022870
	28hpf	kdrl+	wt	.07426912
			MO	.08073884
		double++	wt	.03825557
			MO	.03549145
		neg	wt	.06439962
			MO	.05669148
spred1	22hpf	kdrl+	wt	.02512036
			MO	.02269222
		neg	wt	.01854670
			MO	.01469553
	28hpf	kdrl+	wt	.03093394
			MO	.03030841
		double++	wt	.00614901
			MO	.01007810
		neg	wt	.02454998
			MO	.02155778
wnt5b	22hpf	kdrl+	wt	.01820923
			MO	.02291014
		neg	wt	.02732691
			MO	.02327225
	28hpf	kdrl+	wt	.02407717
			MO	.02632252
		double++	wt	.00762624
			MO	.00778821
		neg	wt	.03146978
			MO	.02514865
atoh8	22hpf	kdrl+	wt	.00305299
			MO	.00323601
		neg	wt	.00317329
			MO	.00312950
	28hpf	kdrl+	wt	.01184525

			MO	.01180793
		double++	wt	.00485201
			MO	.00446161
		neg	wt	.00573436
			MO	.00577571
dbx1a	22hpf	kdrl+	wt	.00055194
			MO	.00074721
		neg	wt	.01860296
			MO	.01239978
	28hpf	kdrl+	wt	.00338533
			MO	.00105633
		double++	wt	.00134595
			MO	.00046124
		neg	wt	.03243272
			MO	.01613025
dlx2a	22hpf	kdrl+	wt	.00213131
			MO	.00132608
		neg	wt	.04536718
			MO	.04785040
	28hpf	kdrl+	wt	.00496068
			MO	.00140830
		double++	wt	.00173169
			MO	.00046280
		neg	wt	.07047550
			MO	.05027737
gata2b	22hpf	kdrl+	wt	.00697960
			MO	.00517748
		neg	wt	.00037304
			MO	.00041518
	28hpf	kdrl+	wt	.00973845
			MO	.00793057
		double++	wt	.02454043
			MO	.02614793
		neg	wt	.00263341
			MO	.00129716
hoxb2a	22hpf	kdrl+	wt	.00730667
			MO	.00660912
		neg	wt	.00685295
			MO	.00600146
	28hpf	kdrl+	wt	.00802213
			MO	.00570267
		double++	wt	.00225894
			MO	.00155197
		neg	wt	.01034307
			MO	.00851884

Appendix IV: Manuscript submitted for peer-reviewed publication describing the method to quantitatively analyse *in situ* hybridization in zebrafish embryos.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Title: **An optimized pipeline for parallel image-based quantification of gene expression and genotyping after *in situ* hybridization**

Running title: **Optimized ISH quantification**

Tomasz Dobrzycki[+1], Monika Krecsmarik[+1,2], Florian Bonkhofer[1], Roger Patient[1,2] and Rui Monteiro*[1,2]

1. MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford, OX3 9DS, United Kingdom
2. BHF Centre of Research Excellence, Oxford, United Kingdom

[+]These authors contributed equally.

*Contact details:

MRC Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford, OX3 9DS, United Kingdom

Email: rui.monteiro@imm.ox.ac.uk

Key words: zebrafish, mutant, bias, genotyping, image quantification

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

SUMMARY STATEMENT

Our optimized protocol to genotype zebrafish mutant embryos after *in situ* hybridization and digitally quantify the *in situ* signal will help to standardize existing experimental designs and methods of analysis.

ABSTRACT

Advances in genome engineering have resulted in the generation of numerous zebrafish mutant lines. A commonly used method to assess gene expression in the mutants is *in situ* hybridization (ISH). Because the embryos can be distinguished by genotype after ISH, comparing gene expression between wild type and mutant siblings can be done blinded and in parallel. Such experimental design reduces the technical variation between samples and minimises the risk of bias. This approach, however, requires an efficient method of genomic DNA extraction from post-ISH fixed zebrafish samples to ascribe phenotype to genotype. Here we describe a method to obtain PCR-quality DNA from 95-100% of zebrafish embryos, suitable for genotyping after ISH. In addition, we provide an image analysis protocol for quantifying gene expression of ISH-probed embryos, adaptable to analyse different expression patterns. Finally, we show that intensity-based image analysis enables accurate representation of the variability of gene expression detected by ISH and that it correlates well with quantitative methods like qRT-PCR. By combining genotyping after ISH and computer-based image analysis, we have established a high-confidence, unbiased methodology to assign gene expression levels to specific genotypes, and applied it to the analysis of molecular phenotypes of newly generated *lmo4a* mutants.

INTRODUCTION

The emergence of genome engineering technologies, including zinc-finger nucleases (ZFNs)(Doyon et al., 2008; Meng et al., 2008), transcription activator-like effector nucleases (TALENs)(Huang et al., 2011) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 (Hwang et al., 2013), has enabled zebrafish researchers to generate a wide range of mutant lines by precisely targeting genomic loci with high efficiency (Varshney et al., 2015). These methods provide alternatives to morpholino oligonucleotides (MOs), which have been used for gene knock down studies for over 15 years (Nasevicius and Ekker, 2000). In the light of recent concerns regarding the reliability of phenotypes induced by MOs (Kok et al., 2015), targeted gene knockouts have become a powerful tool to verify the transient MO phenotypes (Rossi et al., 2015) or to study genetic mutants, using MOs as a secondary tool (Gore et al., 2016). Low-cost protocols for DNA isolation (Meeker et

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

al., 2007) and for genotyping of zebrafish embryos (Wilkinson et al., 2013) have allowed faster generation and analysis of new mutant lines.

One of the most widely used methods to analyse molecular phenotypes during embryonic development is *in situ* hybridization (ISH). This technique is used to detect spatial expression patterns and tissue-specific changes in mRNA levels. Relevant protocols with several extensions have been standardised (Westerfield, 2007) and numerous validated probes are curated on the ZFIN database (Howe et al., 2013). Using ISH to analyse MO phenotypes requires processing of the MO-treated and control samples separately, which may result in an inherent bias and an increased risk of technical variation. Lack of reported measures to reduce the risk of bias has recently been exposed in a meta-analysis of *in vivo* animal studies (Macleod et al., 2015). Genomic mutants offer a way to overcome these issues. Wild type and mutant embryos can be processed in one sample, assessed phenotypically in a blinded manner and then distinguished based on their genotype. This way, the technical variation between samples is minimised and phenotypic assessment is largely bias-free. A few publications have reported approaches to genotype mutant lines after ISH, using either proteinase K treatment (Gore et al., 2016; Zhu et al., 2011) or commercial kits (Bresciani et al., 2014) for subsequent DNA extraction. However, the efficiency of these DNA extraction and genotyping methods has not been adequately demonstrated.

Reporting of phenotypes has usually been confined to one representative ISH image per condition, limiting the ability to quantitatively represent the variability of the phenotype. Approaches to score the expression levels as ‘high’, ‘medium’ or ‘low’ by eye (Blaser et al., 2017; Gao et al., 2016; Genthe and Clements, 2017; Peterkin et al., 2007; Place and Smith, 2017) are subjective and limited in how accurately they represent the effect of a knockout or a knockdown. Furthermore, visual scoring is inherently prone to poor reproducibility and low sensitivity. A more accurate way to quantitatively evaluate the change in gene expression levels involves counting the cells that contain the ISH signal (Espín-Palazón et al., 2014). However, this method is difficult to apply to compact anatomical structures and cell boundaries are hard to distinguish after ISH. Intensity-based image analysis using a selected region of interest (ROI) provides an objective alternative to visual scoring. Fan and colleagues have recently described a method using the ImageJ software to quantify ISH signal intensity in mouse embryos, where measurements were taken along a straight line drawn across the developing forelimb (Fan et al., 2015). Wen et al. have adapted this technique for use in the zebrafish to quantify gene expression levels in the mesoderm during early embryonic development (Wen et al., 2017). This approach can be further optimised to quantify gene expression after ISH in other regions and at other stages of the developing embryo.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Here we provide an optimised protocol for fast, inexpensive and highly efficient isolation of DNA from fixed zebrafish embryos for PCR and genotyping after ISH. This approach, based on the Hot Sodium Hydroxide and Tris (HotSHOT) method (Truett et al., 2000), is extremely reliable and allowed successful genotyping in 95-100% of the embryos. In addition, we propose a detailed step-by-step guide for gene expression intensity measurements based on modifications to the previously described quantitation methods (Fan et al., 2015; Wen et al., 2017). This pipeline (Fig. 1A) provides a tool for high-confidence, bias-free reporting of molecular phenotypes using standard ISH. We demonstrate its usefulness by assessing expression changes in *gpr65* morphants and *runx1* and *lmo4a* mutants.

RESULTS

HotSHOT genomic DNA extraction is suitable for PCR amplification in ISH-probed embryos

Our aim was to establish a fast and cost-effective way to extract DNA from individual embryos after ISH. First, we optimized the published HotSHOT protocol (Truett et al., 2000) for raw genomic DNA extraction, followed by JumpStart™ REDTaq® PCR. By testing amplicons designed for four different loci, we found that a 5-minute incubation time in the lysis buffer was sufficient to extract DNA to amplify fragments ranging from 144 to 600 base pairs (bp) by PCR (Fig. 1B). Increasing incubation time did not increase DNA yield (Fig. 1B). Next, we tested the 5-minute DNA extraction method on embryos at different stages by amplifying a 300bp fragment from the *runx1* locus. We successfully extracted DNA from fish ranging from 14hpf to 5dpf that had been stored for up to 2 years in glycerol at room temperature (Fig. 1C; Table 1). Incubation for less than 30 min was on average more efficient (allowing successful genotyping of more embryos) than incubation times longer than 65 min (Table 1). When comparing two different commercially available PCR master mixes, we found the JumpStart™ REDTaq® ReadyMix™ to be on average more efficient than Phire™ Green HotStart II PCR Master Mix (data not shown), likely due to differences in the optimal salt concentrations for the enzymes in each PCR mix.

Digital quantification of the ISH signal in the dorsal aorta reveals a significant decrease of dnmt3bb.1 mRNA levels in runx1$^{W84X/W84X}$ mutants and a significant increase of runx1 expression in gpr65 morphants

To demonstrate the usefulness of our genotyping protocol in a known mutant, we imaged 130 embryos from the incross of *runx1*$^{+/W84X}$ heterozygotes (Jin et al., 2009), fixed at 33hpf and probed

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

for *dnmt3bb.1* mRNA, a known downstream target of *runx1* within the haemogenic endothelium (Gore et al., 2016) (Fig. 2A). Scoring of the images as 'high', 'medium' or 'low' showed a Mendelian 1:2:1 distribution of phenotypes (Fig. S1A). We then genotyped all the imaged embryos with 100% efficiency using the above protocol and RFLP (Fig. 2B). The observed Mendelian distribution of phenotypes, resulting from the first phenotypic assessment, did not entirely correspond to the respective genotypes. While the 'low' phenotype was significantly overrepresented in the homozygous mutant group (X^2=95.3, d.f.=4; p<0.001), there was no difference in the distribution of 'high'- and 'medium'-expressing embryos among wild type and heterozygous fish (X^2=1.35, d.f.=1; p>0.2) (Fig. S1B). To implement a more quantitative assessment, we used Fiji (Schindelin et al., 2012) to digitally quantify pixel intensities of each embryo in the dorsal aorta region along the yolk sac extension prior to genotyping. This approach relied on allocating two separate ROIs to each image: one containing the staining (yellow ROI, Fig. 1A) and another containing an equal area of the embryo without any staining (green ROI, Fig. 1A). By subtracting the background value from the staining value, a number was assigned to each embryo. After genotyping, we compared the signal intensity values in wild type, heterozygous and mutant embryos. As expected, the *runx1* mutant embryos showed a statistically significant reduction by approximately 50% of *dnmt3bb.1* signal compared to wild types or heterozygotes (μ_{wt}=54, μ_{het}=50.1, μ_{mut}=26.3; F=132.97, d.f.=2, 69.2; p<0.001). In contrast, there was no significant difference in signal intensity between wild types and heterozygotes (μ_{wt}=54, μ_{het}=50.1; p>0.3) (Fig. 2C). We found that the signal intensity values in all groups were very dispersed, with high coefficients of variation (24%, 22% and 21% for wild type, heterozygote and mutant groups, respectively).

We applied the same analysis method to the images of 16 *gpr65* morphants and 16 control (uninjected) siblings probed at 29hpf for *runx1* mRNA, previously shown to be downstream of *gpr65* (Gao et al., 2016). We found a significant increase of approximately 25% in pixel intensity levels of the *runx1* probe staining in *gpr65* morphants, compared to non-injected controls (μ_{wt}=29, μ_{gpr65}=37.2; t=2.38, d.f.=30; p<0.05) (Fig. 3). The values of the wild type embryos and the morphants showed dispersion with the coefficients of variation of 34% and 26%, respectively.

Digital quantification of runx1 ISH signal in the dorsal aorta reveals a discrepancy between the phenotype in lmo4a$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ mutants and previously described lmo4a morphants

We applied the ISH quantification and genotyping methods to analyse the phenotype of *lmo4a* mutants. We used TALENs to generate a 5bp deletion in the coding region of the *lmo4a* gene (*lmo4a*$^{\Delta 25\text{-}29}$) (Fig. 4A). We then incrossed *lmo4a* heterozygous parents and quantified the expression of *runx1* in 28hpf embryos prior to genotyping them with 100% efficiency using RFLP (Fig. 4B).

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

MO-mediated knockdown of *lmo4a* was previously shown to result in reduced expression of *runx1* in the dorsal aorta at 30hpf (Meier et al., 2006). Surprisingly, we found no significant differences in *runx1* pixel intensity at 28hpf among wild type, heterozygous and homozygous siblings (μ_{wt}=30.3, μ_{het}=32.8, μ_{mut}=34.6; F=0.509, d.f.=2, 62; p>0.6) (Fig. 4C,D). The coefficients of variation of the values of the wild types, heterozygotes and homozygous mutants were 41%, 38% and 37%, respectively.

DISCUSSION

The HotSHOT method of genomic DNA isolation, originally designed for mouse ear notch samples (Truett et al., 2000), offers a fast and cost-effective way to genotype animals. While it has been used to isolate genomic DNA from PFA-fixed zebrafish samples before (Cooney et al., 2013; Meeker et al., 2007), its efficiency in this setting had not been reported. In consequence, the method has not been widely adopted in the community and many research groups rely on time-consuming and more expensive DNA extraction methods involving proteinase K treatment (Gore et al., 2016) or commercially available kits (Bresciani et al., 2014; Sood et al., 2013). Here we report that PCR-quality DNA can be rapidly extracted from 95-100% of fixed zebrafish embryos aged from 14hpf to 5dpf with an optimised HotSHOT protocol after ISH. This DNA can be subsequently used to genotype the samples with simple, inexpensive PCR followed by a digest to detect restriction enzyme sites disrupted by the mutation, as done previously for fresh tissue (Hruscha et al., 2013). To facilitate this, mutations can be designed to target restriction enzyme recognition sites. In fact, an online TALEN and CRISPR/Cas9 design tool Mojo Hand (Neff et al., 2012) readily provides restriction enzyme sites targeted by a desired mutation. While newly emerging alternatives to standard PCR and RFLP may provide a higher speed of genotyping (D'Agostino et al., 2016; Lee et al., 2016), the protocol described here for genotyping after ISH is attractive due to its high efficiency and demonstrated robustness in our hands.

Genotyping zebrafish embryos after ISH is important because it allows processing of mutant and wild type embryos in one batch, therefore limiting the technical variation between samples. In addition, expression levels of the target gene can be assessed in a non-biased way, because the embryos can be distinguished by their genotype only after phenotypic assessment. This is a powerful way to control for unconscious bias, a serious issue in *in vivo* animal research (Macleod et al., 2015). Assessing mRNA levels in post-ISH embryos has been performed visually, either by scoring the phenotypes into discrete groups (Blaser et al., 2017; Gao et al., 2016; Genthe and Clements, 2017) or by cell counting (Espín-Palazón et al., 2014). However, these approaches are prone to subjectivity and poor reproducibility. Furthermore, visual scoring can be difficult to carry out and interpret due to expression level differences between individuals of the same genotype. Indeed, we show here

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

that pixel intensities of the ISH signal in wild type embryos probed for the transcription factor *runx1* show high dispersion in wild type embryos with over 25% and up to 40% coefficient of variation (Figs 3B, 4D). These results indicate that the interpretation of phenotypes based purely on expected Mendelian distribution from heterozygous incrosses might be misleading. As we demonstrate, visual scoring of embryos from a heterozygous *runx1*$^{+/W84X}$ incross into 'high', 'medium' and 'low' groups based on *dnmt3bb.1* expression levels gives a phenotypic Mendelian distribution of 1:2:1, which could suggest a haploinsufficiency effect on the regulation of *dnmt3bb.1* expression. Genotyping of these embryos revealed that the vast majority of 'low'-expressing ones were indeed genetically homozygous mutant. However, both 'high'- and 'medium'-expressing embryos were distributed similarly across wild type and heterozygous fish, disproving the haploinsufficiency hypothesis. As a possible explanation for this discrepancy, we found that the signal intensity values in all three genotypes were highly dispersed, with coefficients of variation over 20%. Therefore, each ISH experiment done on embryos from a heterozygous incross should be followed by genotyping to avoid misleading conclusions due to the variability of the ISH signal intensities in embryos of the same genotype.

We would argue that the use of digital image analysis on ISH-probed samples is critical for objective, statistical demonstration of changes in expression levels. Here we describe an imaging protocol based on previous studies (Fan et al., 2015; Wen et al., 2017) to measure gene expression intensity in the trunk region of 1-2 day old zebrafish embryos. We show that the average ISH staining intensity for *dnmt3bb.1* mRNA is significantly decreased in *runx1* mutants compared to wild type siblings, in agreement with previously reported qRT-PCR quantitation of *dnmt3bb.1* levels in whole embryos (Gore et al., 2016). Thus, our method is robust and should be adopted instead of less reliable visual scoring methods. It could also be used as an alternative to qRT-PCR experiments where these require larger numbers of animals and are prone to errors due to a limited number of highly reliable internal controls (Xu et al., 2016). Our quantification method addresses all of these limitations. Furthermore, it presents a way to measure changes in expression levels in a very tissue-specific manner, which is useful in the case of genes with multiple developmental roles. We believe it will be particularly helpful for studying other genes with expression patterns that are spatially restricted, such as *gata2b*, a haematopoietic gene expressed in the ventral wall of the dorsal aorta (Butko et al., 2015). An alternative modification to our method could involve subtracting the average intensity of another stained region from the intensity of the selected ROI – for instance, using the *runx1* or *dnmt3bb.1* signal in the head as an internal control. However, expression levels vary widely between different tissues and there is a risk of saturating the signal, reducing the dynamic range used for the

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

comparisons. Therefore, we believe that using an unstained region with the same area as a background measurement provides a more reliable way to quantify the ISH signal in each embryo.

We also propose a way to quantitatively represent variation in gene expression levels without relying on subjective and biased scoring. For instance, we could replicate the previously reported increase in *runx1* expression in *gpr65* morphants (Gao et al., 2016) with our method, but we represented it in a more objective quantitative way, also allowing powerful statistical analysis. In fact, we achieved 63% power to detect a 25% increase in pixel intensity at the $p<0.05$ level for sample sizes as small as 16 for each condition. This method of analysis also allows precise calculations of required sample sizes to achieve given power. In the presented example, 80% power would require 24 MO-injected embryos and 24 uninjected controls. Such calculations are essential in animal research (Dell et al., 2002), but they are notoriously unreported (Macleod et al., 2015). In fact, recently published updated guidelines for the use of MOs in zebrafish encourage statistical analysis of phenotypes and advocate the use of blinded assessment (Stainier et al., 2017) and both points can be addressed with our method. In addition, there is scope to automate the phenotypical analysis using this method. We have generated a batch conversion Fiji macro that converts all images in a given file directory (see Supplementary Detailed Protocol) ready for manual ROI selection. Future optimisations may involve automation of intensity measurements of the ISH images (Chen et al., 2011).

We have further applied the described method to analyse the molecular phenotype of an unpublished *lmo4a* mutant. In contrast to previous reports showing decreased *runx1* expression in *lmo4a* morphants (Meier et al., 2006), we found no difference in *runx1* pixel intensity levels between wild types, heterozygotes and *lmo4a* homozygous mutants. This discrepancy may result from compensatory mechanisms activated in mutants but not morphants, a phenomenon now frequently reported in zebrafish (El-Brolosy and Stainier, 2017). Potential compensation could come from the *lmo4b* paralogue or from *lmo2*, another gene from the *lmo* family that is expressed in the dorsal aorta and required for the specification of the haematopoietic programme (Patterson et al., 2007).

METHODS

Maintenance of zebrafish and morpholino oligonucleotide injections

All animal experiments were approved by the local ethics committee. Wild type and *runx1*^{W84X} (Jin et al., 2009) and *lmo4a*$^{\Delta 25\text{-}29}$ mutant zebrafish (*Danio rerio* Hamilton) were maintained and bred according to standard procedures (Westerfield, 2007). Embryos were collected by natural mating and staged according to morphological features (Kimmel et al., 1995) corresponding to respective

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

age in hours or days post fertilisation (hpf or dpf, respectively). For *gpr65* knockdown, wild type one-cell stage embryos were injected with 4ng of GPR65_SP MO (Gao et al., 2016).

Generation of lmo4a mutants

For TALENs design, Mojo Hand software was used (http://www.talendesign.org)(Neff et al., 2012). The identified target site was GGAAAGCTCCGCGGTT. The RVD-containing repeats were assembled using the Golden Gate approach in pTAL3-DDD and pTAL3-RRR vectors (Cermak et al., 2011). The resulting DNA templates were verified by sequencing and linearised with NotI enzyme. The mRNAs were transcribed from 1µg linearised template with SP6 mMessage mMachine® kit (Ambion) and purified with RNeasy® Micro Kit (QIAGEN).

One-cell zebrafish embryos were injected with 100pg left-arm + 100pg right-arm TALEN mRNAs. Germline mutations in the founders (*lmo4a*$^{+/-}$) were identified using restriction fragment length polymorphism (RFLP)(Bedell et al., 2012) (see primer sequences below) with SacII enzyme on the genomic DNA extracted from their offspring. Mutations were identified by the presence of an undigested PCR product on a 2% agarose gel. The undigested mutant fragments were purified from the gel, cloned into pGEM®-T Easy vector (Promega) and sequenced. The *lmo4a*$^{+/-}$ founders were outcrossed to generate heterozygous carriers, identified with fin-clipping and genotyping at 3dpf as described previously (Wilkinson et al., 2013).

Whole-mount in situ hybridization

ISH was carried out according to the standard lab protocol (Jowett and Yan, 1996) using digoxygenin-labelled *dnmt3bb.1* (Gore et al., 2016) and *runx1* (Kalev-Zylinska et al., 2002) probes. Post hybridization, the embryos were bleached in 5% formamide/0.5% SSC/10% H_2O_2 (Monteiro et al., 2011) and imaged in 100% glycerol with QImaging MicroPublisher 5.0 RTV Camera and Q-Capture Pro 7™ software (version 7.0.3), using the same exposure, magnification and illumination settings for each embryo.

DNA extraction and genotyping

Genomic DNA was isolated from 4% paraformaldehyde(PFA)-fixed embryos using the original HotSHOT protocol (Truett et al., 2000) (Fig. 1a). Briefly, 40-75 µl of lysis buffer (25 mM NaOH, 0.2 mM EDTA) was added directly to a PCR tube with a freshly-imaged embryo in <5 µl 100% glycerol. To test the efficiency of the DNA extraction, embryos were suspended in the buffer and incubated at 95°C for 5-120 minutes, then cooled to 4°C, after which an equal volume of neutralisation buffer (40-75 µl 40 mM Tris-HCl) was added (see detailed Supplementary protocol). Genomic regions containing the mutated sites in the *runx1* locus were amplified with JumpStart™ REDTaq®

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

ReadyMix™ PCR or with Phire™ Green HotStart II PCR Master Mix according to manufacturer protocols, using 5 µl of DNA lysate in a 20 µl reaction volume and the following primer sequences: 5'-GCTCTGGTGGGCAAACTG-3' and 5'-CATGTGTTTGGACTGTGGGG-3' for *runx1*; and 5'-ACTTTGCCTCTGGATCTGCT-3' and 5'-TCTACATCAGTCCCGCCAAA-3' for *lmo4a*. The presence of *runx1* and *lmo4a* mutations were verified by restriction fragment length polymorphism (RFLP)(Hruscha et al., 2013) with HaeII or SacII enzymes respectively on a 2% agarose gel, using 100bp DNA Ladder as a reference. The sequences of additional primers used were: 5'-GCACCACAGTGGACATTGAT-3' and 5'-GTTGTAGAGGGCCAGCACTT-3' for the *tgfβ3* locus; 5'-CATTAATGCGAGGGATACGG-3' and 5'-AAAAGAGCCACGGTAGGTGA-3' for the *tgfβr2* locus; and 5'-TGGCTAAGTGACCGTCAGAG-3' and 5'-TGAAACAAAACGCAGACGAC-3' for the *gata2a* locus.

Digital image analysis

Using Fiji software (Schindelin et al., 2012), the images were inverted to negative and converted to 8-bit grayscale. A Region of Interest (ROI) containing the ISH expression signal in the dorsal aorta along the yolk sac extension was drawn manually for each embryo. Then a second ROI with the same shape and area was created in a region of the embryo that has a uniform intensity and does not contain any ISH staining. In this particular instance, this region was placed just above the notochord (Fig. 1A). This area was used to define the background. A value for each region was then determined by measuring the average pixel intensity. After subtracting the value of the background region from the value of the stained region, the pixel intensity of the ISH signal was assigned to each embryo (Fig. 1A and detailed Supplementary protocol).

Statistical analysis

The numbers of embryos scored as 'high', 'medium' or 'low' were tested for equal distribution among wild type, heterozygous and mutant genotypes with contingency Chi-squared tests, applying Continuity Correction for 2x2 tables. For digitally analysed images, the pixel intensity values were assessed for normal distribution with Q-Q plots and transformed with sqrt function if necessary. Mean values (μ) of each experimental group were analysed with 2-tailed independent-samples *t*-test (for *gpr65* MO experiment) or with ANOVA (for mutant experiments) with 95% confidence levels, testing for the equality of variances with a Levene's test and applying the Welch correction when necessary. For ANOVA, differences between each two groups were assessed with either Tukey's post-hoc test (for equal variances) or with Games-Howell test (for unequal variances). For all these analyses the IBM® SPSS® Statistics (version 22) package was used. The degree of variability in each sample was assessed by calculating the coefficients of variation, defined as $\frac{s_{(x)}}{\mu}$, with $s_{(x)}$ being the standard deviation. The post-hoc power of the tests and required sample sizes were determined

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

with G*Power software (version 3.0.10)(Faul et al., 2007). The graphs presenting individual data points, means and ±s.d. were plotted using GraphPad Prism 7.

ACKNOWLEDGEMENTS

We thank Dr Dominic Waithe from the Wolfson Imaging Centre for advice on developing the image quantification method. We thank Rossella Rispoli for help with generating the batch conversion Fiji macro. We thank staff members of Biomedical Services at the John Radcliffe Hospital for monitoring and feeding zebrafish.

COMPETING INTERESTS

No competing interests declared.

FUNDING

T.D. was funded by a Wellcome Trust Chromosome and Developmental Biology PhD Scholarship (#WT102345/Z/13/Z). R.M. and M.K. were funded by the British Heart Foundation (BHF IBSR Fellowship FS/13/50/30436). R.K.P. and F.B. were funded by the Medical Research Council. R.M. acknowledges support from the BHF Centre of Research Excellence (RE/13/1/30181), Oxford.

DATA AVAILABILITY

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

REFERENCES

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug, R. G., 2nd, Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y., et al. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**, 114-118.

Blaser, B. W., Moore, J. L., Hagedorn, E. J., Li, B., Riquelme, R., Lichtig, A., Yang, S., Zhou, Y., Tamplin, O. J., Binder, V., et al. (2017). CXCR1 remodels the vascular niche to promote hematopoietic stem and progenitor cell engraftment. *J. Exp. Med.* **214**, 1011-1027.

Bresciani, E., Carrington, B., Wincovitch, S., Jones, M., Gore, A. V., Weinstein, B. M., Sood, R. and Liu, P. P. (2014). CBFβ and RUNX1 are required at 2 different steps during the development of hematopoietic stem cells in zebrafish. *Blood* **124**, 70-78.

Butko, E., Distel, M., Pouget, C., Weijts, B., Kobayashi, I., Ng, K., Mosimann, C., Poulain, F. E., McPherson, A., Ni, C. W., et al. (2015). Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development* **142**, 1050-1061.

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J. and Voytas, D. F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82.

Chen, S., Zhu, Y., Xia, W., Xia, S. and Xu, X. (2011). Automated analysis of zebrafish images for phenotypic changes in drug discovery. *J. Neurosci. Methods* **200**, 229-236.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Cooney, J. D., Hildick-Smith, G. J., Shafizadeh, E., McBride, P. F., Carroll, K. J., Anderson, H., Shaw, G. C., Tamplin, O. J., Branco, D. S., Dalton, A. J., et al. (2013). Teleost growth factor independence (gfi) genes differentially regulate successive waves of hematopoiesis. *Dev. Biol.* **373**, 431-441.

D'Agostino, Y., Locascio, A., Ristoratore, F., Sordino, P., Spagnuolo, A., Borra, M. and D'Aniello, S. (2016). A Rapid and Cheap Methodology for CRISPR/Cas9 Zebrafish Mutant Screening. *Mol. Biotechnol.* **58**, 73-78.

Dell, R. B., Holleran, S. and Ramakrishnan, R. (2002). Sample size determination. *Ilar Journal* **43**, 207-213.

Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., Amora, R., Hocking, T. D., Zhang, L., Rebar, E. J., et al. (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.* **26**, 702-708.

El-Brolosy, M. A. and Stainier, D. Y. (2017). Genetic compensation: A phenomenon in search of mechanisms. *PLoS genetics* **13**.

Espín-Palazón, R., Stachura, D. L., Campbell, C. A., García-Moreno, D., Del Cid, N., Kim, A. D., Candel, S., Meseguer, J., Mulero, V. and Traver, D. (2014). Proinflammatory signaling regulates hematopoietic stem cell emergence. *Cell* **159**, 1070-1085.

Fan, Y., Richelme, S., Avazeri, E., Audebert, S., Helmbacher, F., Dono, R. and Maina, F. (2015). Tissue-specific gain of RTK signalling uncovers selective cell vulnerability during embryogenesis. *PLoS Genet* **11**, e1005533.

Faul, F., Erdfelder, E., Lang, A.-G. and Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175-191.

Gao, X., Wu, T., Johnson, K. D., Lahvic, J. L., Ranheim, E. A., Zon, L. I. and Bresnick, E. H. (2016). GATA Factor-G-Protein-Coupled Receptor Circuit Suppresses Hematopoiesis. *Stem cell reports* **6**, 368-382.

Genthe, J. R. and Clements, W. K. (2017). R-spondin 1 is required for specification of hematopoietic stem cells through Wnt16 and Vegfa signaling pathways. *Development* **144**, 590-600.

Gore, A. V., Athans, B., Iben, J. R., Johnson, K., Russanova, V., Castranova, D., Pham, V. N., Butler, M. G., Williams-Simons, L., Nichols, J. T., et al. (2016). Epigenetic regulation of hematopoiesis by DNA methylation. *eLife* **5**.

Howe, D. G., Bradford, Y. M., Conlin, T., Eagle, A. E., Fashena, D., Frazer, K., Knight, J., Mani, P., Martin, R., Moxon, S. A., et al. (2013). ZFIN, the Zebrafish Model Organism Database: increased support for mutants and transgenics. *Nucleic Acids Res.* **41**, 60.

Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C. and Schmid, B. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982-4987.

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S. and Zhang, B. (2011). Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.* **29**, 699-700.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J. R. and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227-229.

Jin, H., Sood, R., Xu, J., Zhen, F., English, M. A., Liu, P. P. and Wen, Z. (2009). Definitive hematopoietic stem/progenitor cells manifest distinct differentiation output in the zebrafish VDA and PBI. *Development* **136**, 647-654.

Jowett, T. and Yan, Y. L. (1996). Double fluorescent in situ hybridization to zebrafish embryos. *Trends Genet.* **12**, 387-389.

Kalev-Zylinska, M. L., Horsfield, J. A., Flores, M. V., Postlethwait, J. H., Vitas, M. R., Baas, A. M., Crosier, P. S. and Crosier, K. E. (2002). Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development* **129**, 2015-2030.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253-310.

Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., van Impel, A., Kirchmaier, B. C., Peterson-Maduro, J., Kourkoulis, G., Male, I., et al. (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev. Cell* **32**, 97-108.

Lee, H. B., Schwab, T. L., Koleilat, A., Ata, H., Daby, C. L., Cervera, R., McNulty, M. S., Bostwick, H. S. and Clark, K. J. (2016). Allele-Specific Quantitative PCR for Accurate, Rapid, and Cost-Effective Genotyping. *Hum. Gene Ther.* **27**, 425-435.

Macleod, M. R., McLean, A., Kyriakopoulou, A., Serghiou, S., de Wilde, A., Sherratt, N., Hirst, T., Hemblade, R., Bahor, Z., Nunes-Fonseca, C., et al. (2015). Risk of Bias in Reports of In Vivo Research: A Focus for Improvement. *PLoS Biol.* **13**, e1002273.

Meeker, N., Hutchinson, S., Ho, L. and Trede, N. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *Biotechniques* **43**, 610-614.

Meier, N., Krpic, S., Rodriguez, P., Strouboulis, J., Monti, M., Krijgsveld, J., Gering, M., Patient, R., Hostert, A. and Grosveld, F. (2006). Novel binding partners of Ldb1 are required for haematopoietic development. *Development* **133**, 4913-4923.

Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D. and Wolfe, S. A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.* **26**, 695-701.

Monteiro, R., Pouget, C. and Patient, R. (2011). The gata1/pu.1 lineage fate paradigm varies between blood populations and is modulated by tif1γ. *EMBO J.* **30**, 1093-1103.

Nasevicius, A. and Ekker, S. C. (2000). Effective targeted gene 'knockdown' in zebrafish. *Nat. Genet.* **26**, 216-220.

Neff, K. L., Argue, D. P., Ma, A. C., Lee, H. B., Clark, K. J. and Ekker, S. C. (2012). Mojo Hand, a TALEN design tool for genome editing applications. *BMC Bioinformatics* **14**, 1-7.

Patterson, L. J., Gering, M., Eckfeldt, C. E., Green, A. R., Verfaillie, C. M., Ekker, S. C. and Patient, R. (2007). The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish. *Blood* **109**, 2389-2398.

Peterkin, T., Gibson, A. and Patient, R. (2007). Redundancy and evolution of GATA factor requirements in development of the myocardium. *Dev. Biol.* **311**, 623-635.

Place, E. S. and Smith, J. C. (2017). Zebrafish atoh8 mutants do not recapitulate morpholino phenotypes. *PLoS One* **12**, e0171143.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M. and Stainier, D. Y. R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**, 230-233.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682.

Sood, R., Carrington, B., Bishop, K., Jones, M., Rissone, A., Candotti, F., Chandrasekharappa, S. C. and Liu, P. (2013). Efficient methods for targeted mutagenesis in zebrafish using zinc-finger nucleases: data from targeting of nine genes using CompoZr or CoDA ZFNs. *PLoS One* **8**, e57239.

Stainier, D. Y. R., Raz, E., Lawson, N. D., Ekker, S. C., Burdine, R. D., Eisen, J. S., Ingham, P. W., Schulte-Merker, S., Yelon, D., Weinstein, B. M., et al. (2017). Guidelines for morpholino use in zebrafish. *PLOS Genetics* **13**.

Truett, G. E., Heeger, P., Mynatt, R. L., Truett, A. A., Walker, J. A. and Warman, M. L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* **29**, 52-54.

Varshney, G. K., Sood, R. and Burgess, S. M. (2015). Understanding and Editing the Zebrafish Genome. *Adv. Genet.* **92**, 1-52.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Wen, B., Yuan, H., Liu, X., Wang, H., Chen, S., Chen, Z., de The, H., Zhou, J. and Zhu, J. (2017). GATA5 SUMOylation is indispensable for zebrafish cardiac development. *Biochim. Biophys. Acta*.

Westerfield, M. (2007). *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio).* (5th edn). Eugene, OR: Univ of Oregon Press.

Wilkinson, R. N., Elworthy, S., Ingham, P. W. and van Eeden, F. J. M. (2013). A method for high-throughput PCR-based genotyping of larval zebrafish tail biopsies. *Biotechniques* **55**.

Xu, H., Li, C., Zeng, Q., Agrawal, I., Zhu, X. and Gong, Z. (2016). Genome-wide identification of suitable zebrafish Danio rerio reference genes for normalization of gene expression data by RT-qPCR. *J. Fish Biol.*

Zhu, C., Smith, T., McNulty, J., Rayla, A. L., Lakshmanan, A., Siekmann, A. F., Buffardi, M., Meng, X., Shin, J., Padmanabhan, A., et al. (2011). Evaluation and application of modularly assembled zinc-finger nucleases in zebrafish. *Development* **138**, 4555-4564.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Figure 1. Overview of the method to extract DNA for genotyping zebrafish mutants after ISH and measure the mRNA levels. A) Embryos collected from an incross of fish heterozygous for a mutant allele are probed for the measured gene with a standard ISH protocol. After imaging in 100% glycerol, genomic DNA is extracted using the HotSHOT protocol by adding the lysis buffer directly to the embryo in a 0.2ml PCR tube, followed by a ≥5 min incubation at 95°C. This DNA is used for genotyping of the embryos by PCR and restriction fragment length polymorphism (RFLP). In parallel, the images for each embryo are inverted and converted to 8-bit greyscale. ROIs of identical shape and size containing the ISH signal (yellow) and background (green) are manually selected and measured. The measurements, assigned to corresponding genotypes, are statistically analysed. B) DNA fragments amplified from single fixed and ISH-probed embryos after incubation in HotSHOT lysis buffer for 5, 15, 30, 45, 60 and 120 min followed by JumpStart™ REDTaq® ReadyMix™ PCR. Primer pairs designed for the *gata2a*, *runx1*, *tgfbr2b* and *tgfb3* produced 600, 300, 176 and 144 bp fragments, respectively. Three wild type embryos were used for each time point. First lane from the left: 100bp DNA ladder. C) DNA fragments amplified from single fixed and ISH-probed wild-type embryos after DNA extraction in HotSHOT lysis buffer for 5 min using a primer pair to amplify a 300bp fragment of the *runx1* locus with JumpStart™ REDTaq® ReadyMix™ PCR. First lane from the left: 100bp DNA ladder. The lanes after the ladder represent embryos that are 14h, 1, 2, 3 and 4 days old and have been stored in glycerol at room temperature for 2 years.

Figure 2. *Runx1* mutants have significantly reduced levels of *dnmt3bb.1* mRNA detected by ISH. A) A representative image of ISH for *dnmt3bb.1* in a 33hpf wild type embryo, showing the expression in the dorsal aorta. B) 2% agarose gel showing representative genotypes of wild type (WT), heterozygous (HET) and mutant (MUT) *runx1* embryos, distinguished by RFLP. Yellow: wild type 214bp + 124bp bands, pink: 338bp mutant band. First lane from the left: 100bp DNA ladder. C) Pixel intensity values of *dnmt3bb.1* mRNA in *runx1* mutant embryos (N=36) are significantly decreased compared to wild types (N=32) and heterozygotes (N=62). The coefficients of variation are 24%, 22% and 21% for wild type, heterozygote and mutant groups, respectively. The bars represent mean ± s.d. ***p<0.001 (Games-Howell post-hoc test).

Figure 3. *Gpr65* morphants have significantly increased levels of *runx1* mRNA detected by ISH. A) A representative image of ISH for *runx1* in a 29hpf wild type embryo, showing the expression in the dorsal aorta. B) Pixel intensity values of *runx1* mRNA in in *gpr65* MO-injected embryos (N=16) are significantly higher than in uninjected control siblings (N=16). The coefficients of variation are 34% and 26% for wild type and morphant groups, respectively. The bars represent mean ± s.d. *p<0.05 (*t* test). The power of the *t* test to detect the difference at 0.05 level was 63%.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Figure 4. *Runx1* levels detected by ISH are not affected in *lmo4a* mutants. A) TALENs were designed to a region (blue) ~20bp downstream of the *lmo4a* translation start site (green). Isolated mutant alleles carry 5bp deletions (Δ25-29) (red gaps) upstream of the conserved LIM domains (orange), resulting in a frameshift after S8. The resulting mutant protein is predicted to lack the LIM domains, including the crucial S39 (purple). B) 2% agarose gel showing representative genotypes of wild type (WT), heterozygous (HET) and mutant (MUT) *lmo4a* embryos, distinguished by RFLP. Yellow: wild type 258bp + 87bp bands, pink: 340bp mutant band. First lane from the left: 100bp DNA ladder. C) A representative image of ISH for *runx1* in a 28hpf embryo, showing the expression in the dorsal aorta. D) Quantification of the *runx1* mRNA signal, detected by ISH, from 28hpf wild type (N=15), heterozygous *lmo4a*$^{+/-}$ (het) (N=34) and *lmo4a*$^{\Delta 25\text{-}29/\Delta 25\text{-}29}$ mutant (N=18) embryos from one clutch shows no significant difference in *runx1* pixel intensity among the different genotypes. The coefficients of variation are 41%, 38% and 37% for wild type, heterozygote and mutant groups, respectively. The bars represent mean ± s.d.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Table 1. The efficiency of the HotSHOT protocol for post-ISH DNA extraction from zebrafish embryos at different stages depending on the lysis incubation time.

Incubation time / Age	≤30 min	≥65 min
22-26hpf	99.3% (N=9)	96.4% (N=1)
28-30hpf	99.3% (N=7)	94.2% (N=8)
33-48hpf	100% (N=2)	100% (N=2)
4-5dpf	96.8% (N=3)	88.7% (N=2)

N: number of independent experiments

Fig. 1

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

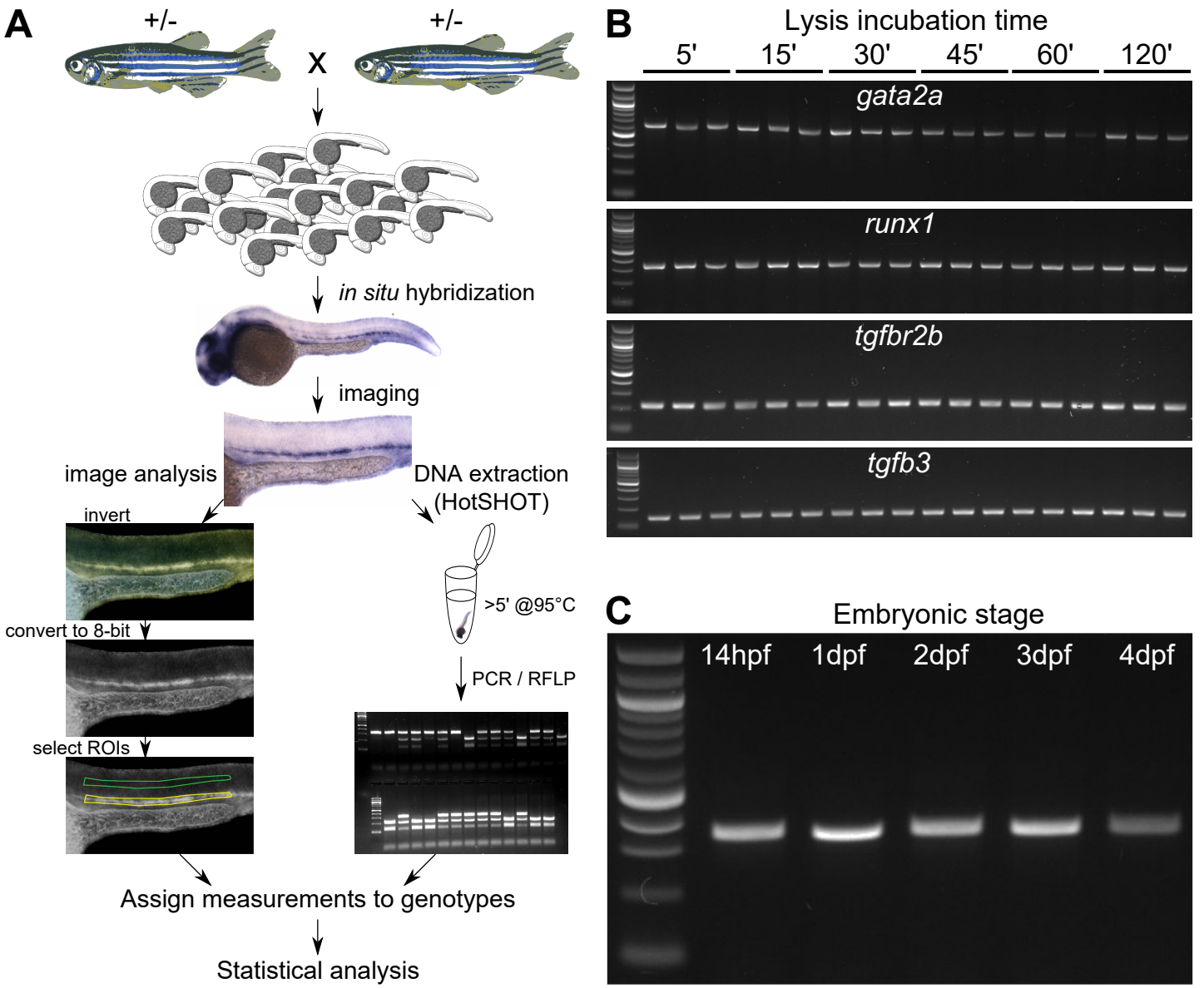
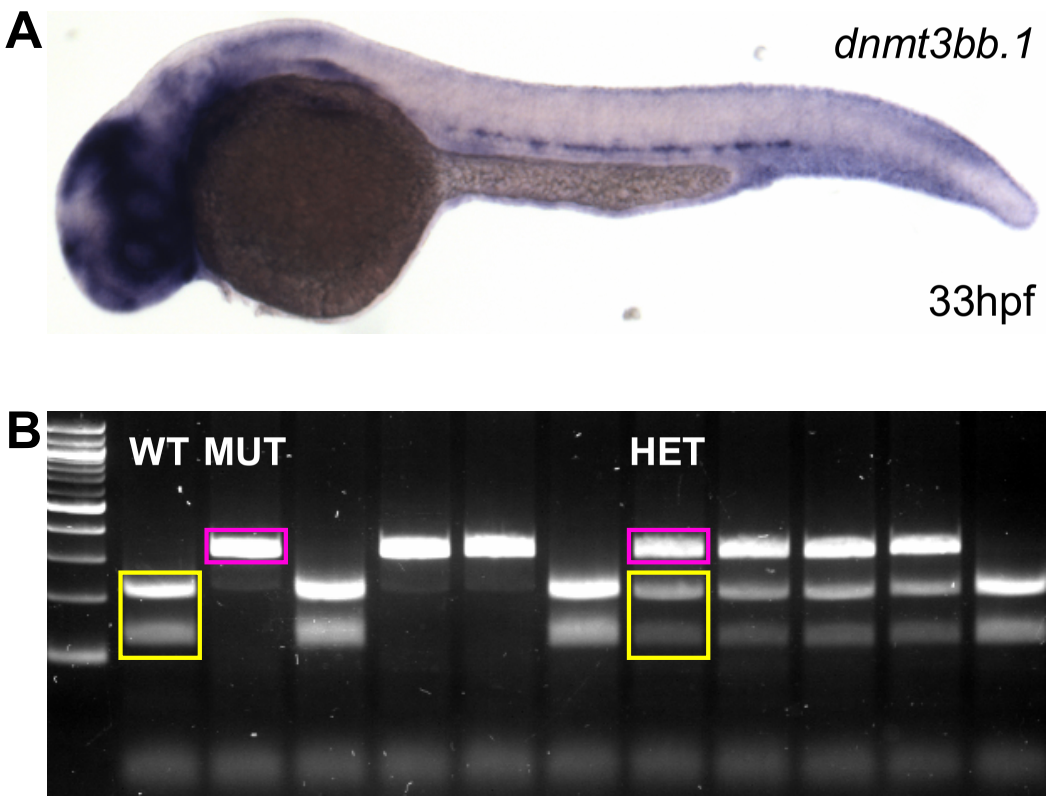


Fig. 2


A
dnmt3bb.1
33hpf
B
WT MUT
HET

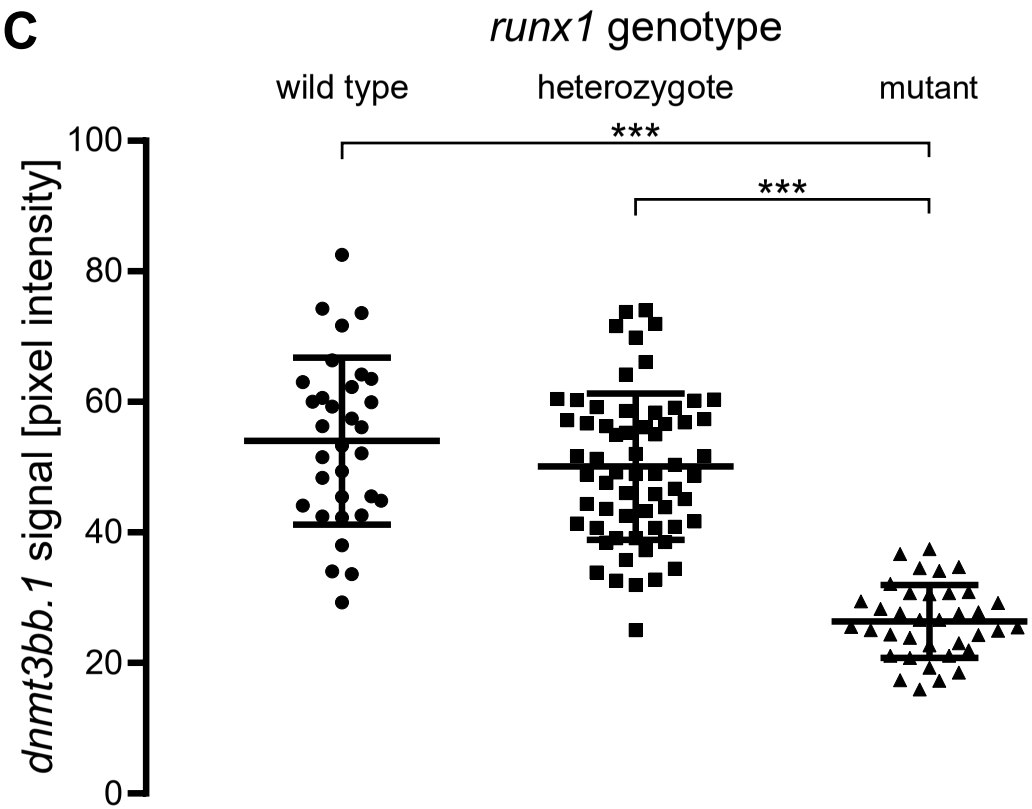

C
runx1 genotype
wild type
heterozygote
mutant
***
***
dnmt3bb.1 signal [pixel intensity]
100
80
60
40
20
0

Fig. 3

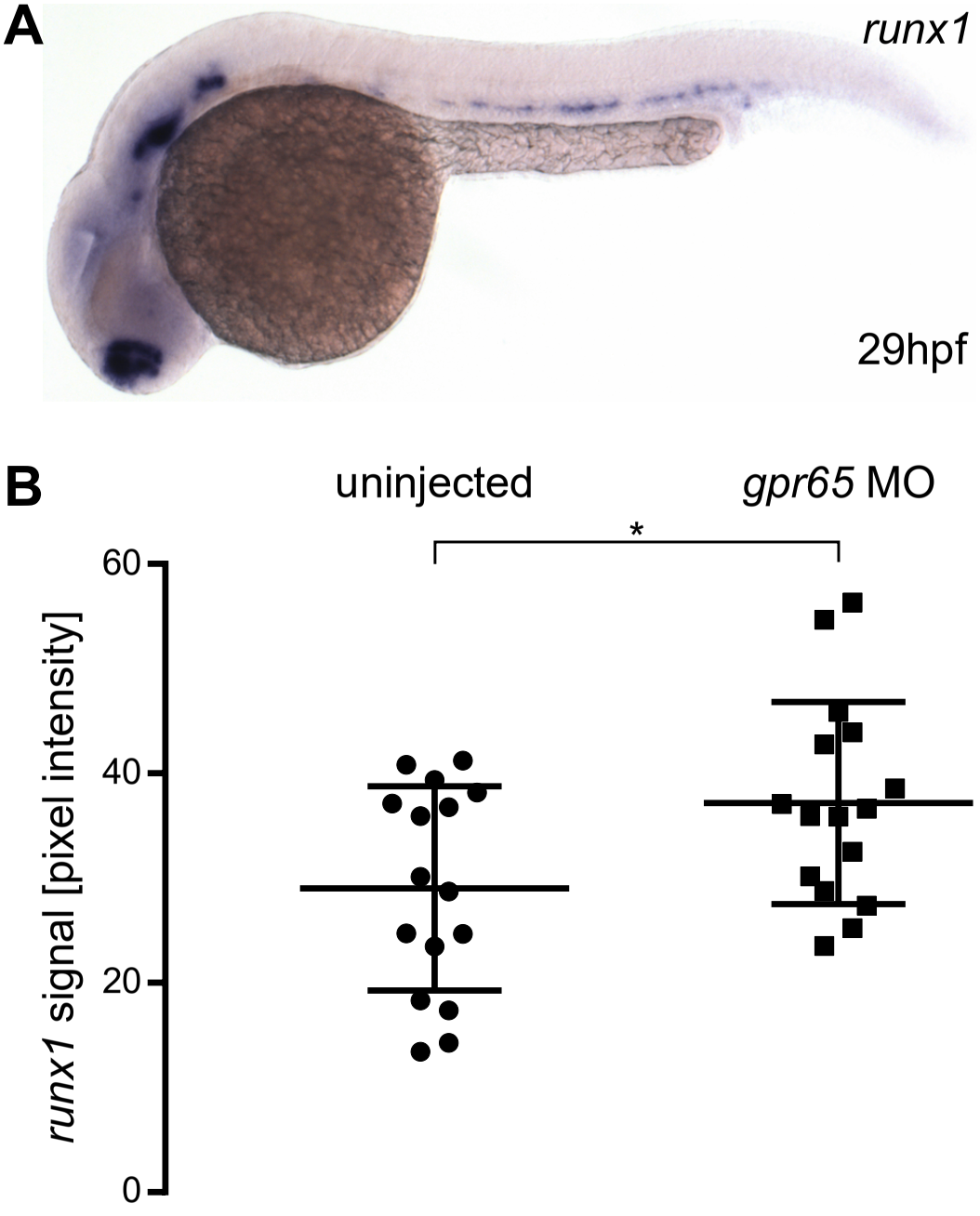
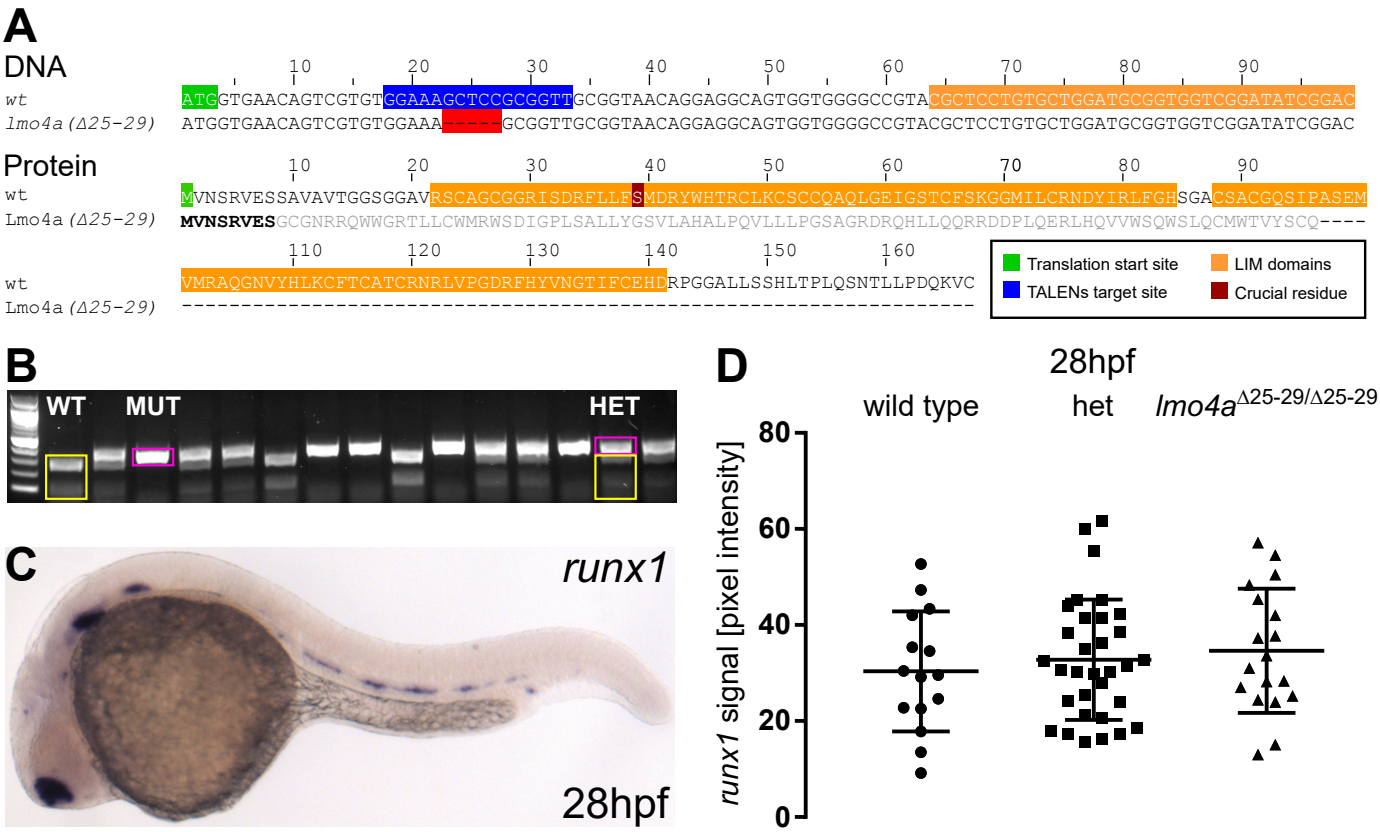

A
runx1
29hpf
B
uninjected
gpr65 MO
*
runx1 signal [pixel intensity]
60
40
20
0

Fig. 4

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.



bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

SUPPLEMENTARY DETAILED PROTOCOL

DNA EXTRACTION AND IMAGE PROCESSING PROTOCOL

for

An optimized pipeline for parallel image-based quantification of gene expression and genotyping after *in situ* hybridization

Tomasz Dobrzycki, Monika Krecsmarik, Florian Bonkhofer, Roger Patient and Rui Monteiro

REAGENTS AND EQUIPMENT

- Stereomicroscope equipped with a digital camera and imaging software
- Fiji Software
- 0.2 ml PCR tubes
- glycerol
- forceps *or* P20 pipette and tips
- single- or multi-channel P200 pipette and tips
- BRAND® glass microscope slide with cavity for imaging embryos: L × W 76 mm × 26 mm, thickness 1.2-1.5 mm, with 1 concavity (BR475505, Sigma)
- HotSHOT alkaline lysis buffer (25 mM NaOH, 0.2 mM disodium EDTA, pH 12)
- HotSHOT neutralization buffer (Tris HCl 40 mM, pH 5)

PROCEDURES

Imaging

1. Prepare a glass slide with a drop of 100% glycerol.
2. Place the first ISH-probed embryo from 80% glycerol onto the slide and take a picture of the embryo with the appropriate orientation and magnification for the ISH staining.
3. Transfer the embryo either with forceps or a P20 pipette into a fresh 0.2 ml PCR tube with a minimum amount of glycerol.
4. Continue taking pictures of the remaining embryos using the same illumination, magnification and exposure settings.

DNA extraction

5. Add 40 µl of HotSHOT alkaline lysis buffer to the PCR tubes (from step **3**.) and make sure the embryos are in the solution. ** Using up to 75 µl does not affect the efficiency of the protocol.*
6. Incubate the tubes at 95°C for at least 5 minutes. ** Incubation longer than 65 min might reduce the success rate of the amplification of a PCR product.*
7. Cool the tubes to 4°C before adding an equal volume (40 µl) of HotSHOT neutralization buffer. ** Overnight incubation at 4°C after the neutralization improves PCR efficiency*. Spin down the samples before each use.
8. Use 5 µl of DNA lysate for PCR to genotype the embryos.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Image analysis

9. For image processing, open the image in Fiji and invert the image with "Edit▸Invert". Then change the image type to 8-bit ("Image▸Type▸8-bit"). ** We have generated a Fiji macro that performs these actions simultaneously on all images in a directory.*
10. Using the polygon selection tool, draw your region of interest (ROI) manually on the image around the region containing the signal.
11. Press "t" to open the ROI manager. Use the ROI manager's Measure command to measure the intensity of the ROI. Copy the "mean" value from the Results window to Excel or equivalent calculation software.
12. Move the same ROI, ensuring the same size and shape as the original region, to a region of the zebrafish not containing any staining. Repeat step **11** to measure the background.
13. To obtain the mean pixel intensity of the ISH signal, subtract in Excel the mean intensity value of the background region from that of the stained region for each embryo.

Statistical analysis

14. Assign each intensity value to a genotype (from step **8.**). ** If scoring phenotypes by eye (eg. 'high', 'medium' or 'low'), assign the phenotypes to genotypes and analyse with contingency chi-squared test.*
15. Plot all the values on one Q-Q plot to identify any deviations from normal distribution. ** Normal distribution can also be verified with the Kolmogorov-Smirnov test or the Shapiro-Wilk test. However, for large sample sizes there is a high risk of false positives in these tests.*
16. If there are strong deviations from normal distribution, transform all the values (using e.g. ln or sqrt functions) to make sure they are normally distributed before proceeding.
17. Analyse the differences between the values (transformed if necessary) assigned to each genotype (wt vs heterozygote vs mutant) with 2-tailed ANOVA with 95% confidence levels, accounting for the equality of variances with a Levene's test and Welch correction. For pairwise comparisons between each pair of genotypes, use Tukey's (equal variances) or Games-Howell (unequal variances) post-hoc test.
18. Plot the untransformed values (from step **14**.) as dot plots for the best representation of the results.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

Fig. S1

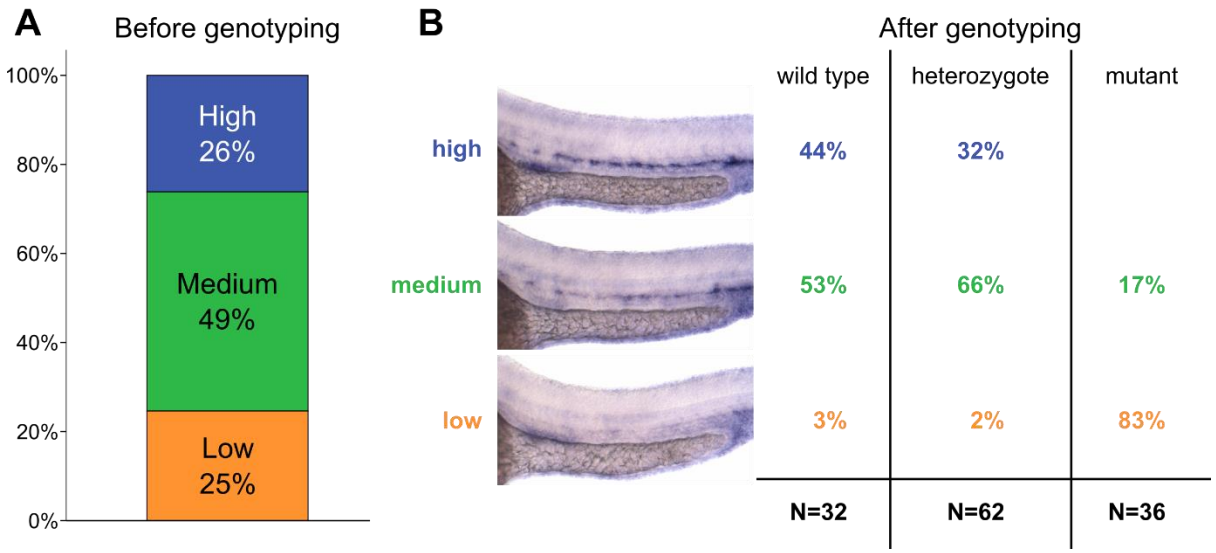


Figure S1. Phenotypic scoring of *dnmt3bb.1*-probed embryos does not fully correspond to their *runx1* genotype. A) The phenotypes of 130 33hpf *dnmt3bb.1*-probed embryos coming from the incross of *runx1* heterozygotes scored as 'high' (N=34), 'medium' (N=64) and 'low' (N=32) follow a Mendelian distribution. B) Distribution of the phenotypes according to the genotype. Images show a representative phenotype scored as 'high', 'medium' or 'low'. Percentages show the occurrence of the phenotypes for the given genotype.

bioRxiv preprint first posted online Jun. 13, 2017; doi: http://dx.doi.org/10.1101/149591. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.

```
// Recursively converts images to 8-bit/invert starting in a user-specified
directory.
macro "conv"{
dir1 = getDirectory("Choose Source Directory ");
list = getFileList(dir1);
dirD = getDirectory("Choose Destination Directory ");
counter = 1;
setBatchMode(true);
for (i=0; i<list.length; i++)
    {
      path = dir1+list[i];
      if (endsWith(path, "tif")){
            print(path);
            run("TIFF Virtual Stack...", "open=path");
            run("8-bit");
            run("Invert");
            counter_converted=IJ.pad(counter,2);
            print(counter_converted);
            saveAs("Tiff", dirD+"conv"+counter_converted);
      close();
      i = i + 1;
      call("java.lang.System.gc");
      }
   counter = counter+1;
   i = i-1;
  }
}
```

Appendix V: Haematopoiesis rap

Performed at FameLab regional final on 08.03.2017, at the WIMM Day on 24.03.2017 and at the DPhil viva on 30.11.2017.

Yo,
Remember when you cut yourself, or fell on your chin?
You would have seen blood flowing through your skin
And probably, this made you feel sad
To lose something as important as blood.
But why is blood important? It does quite a bit,
It distributes the nutrients after you eat,
It collects the waste to be removed with piss
And keeps the body balanced – or in homeostasis.
But wait! The blood's not only a transporter,
In fact, it contains more than salt and water,
There are also cells, of three main types:
Erythrocytes, leukocytes and thrombocytes…
No need for sighs,
Their names may not be nice
But we also call them differently – let's start with leukocytes
Or white blood cells – they're like the police
That protect our body from disease
So the medics – in their community
Call them the cells that provide immunity.
Now on to thrombocytes, or platelets for short,
When you have a wound they form the clot.
Erythrocytes are red blood cells (or RBC),
They transport oxygen, to be used for energy
Production, from lungs to everywhere
And collect the CO_2 to exhale with the air…
Now - would you be amazed
If I told you that all these cells come from the same place?
Because there is one cell, that does the job quite well,
Its name, hard to spell – haematopoietic stem cell.
It's really powerful, we have quite a few of them
And each one can make a whole new blood system.
So these blood stem cells can be worth more than gold
And they first turn up in embryo that is six weeks old,
In its main blood vessel – the dorsal aorta.
This story is quite long, but I'll try to make it shorter:
The walls of this aorta are made of cells that are flat,
Some receive a signal and become a little fat,
They enter the bloodstream, undergoing transformation
To acquire long-term blood-making potential.
They first go to liver to divide like mad
And when increased in numbers they go back to blood
And from there, they move straight as an arrow
To their final place, called the bone marrow
And until the end of life – this is where they stay
To make a trillion (!) new blood cells every day!
And this process, on which I'm writing my thesis,
We call so beautifully – haematopoiesis.

I know that was a lot and maybe by September
You'll forget most of it, I just want you to remember:
next time you bleed, there's no need to cry,
Thanks to these blood stem cells you're gonna be Okay!