The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Studying the cellular origin of HSCs in the zebrafish embryo and the role of *Gfi1* transcription factors in their formation

Maryam Jalali,

BSc, MSc, M(Res)

Thesis Submitted to The University of Nottingham for the degree of

Doctor of Philosophy

September 2016

*On hearing the news [of being awarded a Nobel Prize], a friend who knows me only too well, sent me this laconic message: '**Blood**, toil, sweat and tears always were a good mixture'.*

— Max Ferdinand Perutz
Nobel Banquet Speech (10 Dec 1962).

Abstract

In vertebrates, haematopoietic stem cells (HSCs) maintain the blood system throughout life. HSCs are believed to arise during embryogenesis from haemogenic endothelial cells (HECs) that undergo an endothelial-to-haematopoietic transition (EHT). Here, in order to trace the progeny of the embryonic ECs in zebrafish, an inducible CreER^{T2}-LoxP system was used. Following short-term induction of the Cre recombinase during early embryonic stages, Cre reporter gene expression was observed in early larval haematopoietic cells (HCs). At adult stages, PCR revealed the presence of the recombined Cre reporter gene in HCs, demonstrating that adult HCs had originated from embryonic ECs. In zebrafish, HECs of the ventral wall of the dorsal aorta (vDA) are thought to form HCs by basal epithelial-to-mesenchymal transition (bEMT), a process that depends on the transcription factor Runx1. Here, making use of the recently identified *gfi1aa*^{qmc551Gt} line, confocal microscopy showed that qmc551:GFP+ cells were HECs. While most GFP+ cells underwent bEMT, some displayed a novel type of apical departure. In *runx1* morphants, bEMT was abrogated and most GFP+ HECs remained in the vDA. Apical departure, however, was still observed in the absence of Runx1, suggesting a fundamental difference in the underlying mechanism. While *gfi1aa* expression was lost in vDA HECs of *qmc551* homozygous embryos, EHT of GFP+ HECs was completely unaffected. Up regulation of its paralogue Gfi1ab suggested functional redundancy. To study this redundancy, theCRISPR/Cas9 system was used to mutate the *gfi1ab* gene. Here, the mutant *gfi1ab* alleles *qmc552* and *qmc553* were identified. Both alleles, as well as a third allele *sh320* that was generated by a collaborator, encode truncated, most likely non-functional proteins. Initial data on *gfi1aa* and *gfi1ab* double mutant embryos showed a defect in definitive haematopoiesis. Whether HECs were affected and blocked in their ability to undergo EHT remains to be determined.

Acknowledgements

First of all I would like to thank God and the ‘Chordah Masoomeen’.

I am incredibly grateful to my supervisor and mentor Dr. Martin Gering for giving me the opportunity to be a part of his lab. I appreciate Dr. Gering's teaching approach, which includes a perfect balance of ‘challenge' and ‘support.' Thank you, Sir! I think over the years as a student in your lab I have grown as a researcher and have learnt how to be critical about not only my own but others data as well. I appreciate all your effort and support to bring me to this level and would not have been able to do this without your guidance.

I would like to thank The University of Nottingham for granting me the prestigious "Vice Chancellor's Scholarship for Research Excellence (International)" without which I would not have been able to do a PhD after my MRes.

I am extremely grateful to Dr. Tim Self and his team for their assistance with the LSM710. The high-quality images and time-lapse movies would not have been possible without Dr. Tim Self's expertise.

I would like to thank the BSU staff especially, Tom, Helen, and Mellissa for taking care of the zebrafish. It was because of your care that the zebrafish laid loads of embryos.

A big thank you to all my friends here in Nottingham Deniz, Kazhal, Sidar, Mohammad Sunny, Chris, Fatemeh, Yasmin, Jawaher, Peter and Tamil who have been there for me in good times and bad and have made my PhD experience a cherishable one.

My heart felt gratitude to Mehdi, Anjum, Laxmi, Rahul, Mahesh, Amrita and all my friends and teachers back home and around the world that have always encouraged me and prayed for me. A special mention to Aunty Iran, Nassu, Ammeh Tahseen, my Grannies,

my aunts and cousins and all my well-wishers for their encouragement and prayers. Thank you all!

Last but not the least; I would like to take this opportunity to thank my parents and my siblings, Taherreh, Faezeh, Mojtaba and Sohaila for believing in me and always encouraging me to believe in myself. I would also like to thank Saeedeh for her undying support and patience through my PhD.

Dedication

This thesis is dedicated to my parents who have always believed in my abilities and have lifted me every time I have fallen. It is your love and prayers mum and dad that have brought me this far.

List of Abbreviations

4-OHT	4-hydroxy tamoxifen
aa	Amino acid
aD	Apical departure
AGE	Agarose Gel Electrophoresis
AGM	Aorta-gonad-meshonephros
Amp	Ampicillin
bEMT	Basal epithelial-to-mesenchymal transition
BM	Bone marrow
BSA	Bovine serum albumin
CHT	Caudal hematopoietic tissue
CRISPR	Clustered, regularly interspaced, short palindromic repeats
DAB	Diaminobenzidine
DMSO	Dimethyl sulfoxide
dpf	Days post fertilization
DSB	Double stranded break
E	Days postcoitum

ECs	Endothelial cells
EDTA	Ethylenediamine tetra acetic acid
eGFP	Enhanced green fluorescent protein
EHT	Endothelial to haematopoietic transition
EMT	Epithelial-to-mesenchymal transition
EtBr	Ethidium Bromide
FCS	Foetal calf serum
gDNA	Genomic DNA
gRNA	Guide ribonucleic acid
HBSS	Hank's balanced salt solution
HDR	Homology directed repair
HECs	Haemogenic endothelial cells
hpf	Hours post fertilisation
HSC	Haematopoietic stem cell
LB (liquid or solid) medium	Luria-Bertani medium
LTR	Long term reconstitution
NEB	New England Bio Labs
NHEJ	Non-homologous end joining

NMD	Nonsense mediated decay
OCT	Optimal cutting temperature compound
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
prRBCs	Primitive red blood cells
Runx1	Runt-related transcription factor 1
SNP	Single nucleotide polymorphism
SSC	Saline-sodium citrate buffer
TAE	Tris acetate EDTA buffer
TAM	Tamoxifen
TE	Tris EDTA buffer
TS	Target site
vDA	Ventral wall of the dorsal aorta
VEGF	Vascular endothelial growth factor
WISH	Whole mount in situ hybridisation
WKM	Whole kidney marrow

Table of Contents

Abstract ..... **i**

Acknowledgements ..... **ii**

Dedication ..... **iv**

List of Abbreviations ..... **v**

Table of Contents ..... **viii**

1 Introduction **1**

1.1 Adult haematopoiesis, hematopoietic stem cells and their clinical importance 1

1.2 Embryonic haematopoiesis: A historical overview 4

1.3 Haematopoiesis in the murine embryo 6

1.3.1 The primitive wave or embryonic haematopoiesis 7

1.3.2 A prodefinitive wave of haematopoiesis gives rise to erythromyeloid progenitors 8

1.3.3 Definitive haematopoiesis in the murine embryo 10

1.3.3.1 HSCs are derived from haemogenic endothelium of the AGM 10

1.3.3.2 Primary regulatory factors involved in HC development. 11

1.4 Haematopoiesis in zebrafish 13

1.4.1 The primitive wave 14

1.4.2 In the zebrafish erythromyeloid progenitors arise during a second wave of haematopoiesis within the PBI 15

1.4.3 The haemogenic endothelium of the vDA gives rise to definitive HSCs via an endothelial-to-haematopoietic transition (EHT) 16

1.5 The structure and function of Gfi1a and Gfi1b 20

1.6 The zebrafish has three *Gfi1* paralogues 22

1.7 The *qmc551* gene trap line labels HECs of the vDA 24

1.8 *Gfi1ab* mRNA is upregulated in the vDA in the absence of *gfi1aa* 26

1.9 Lineage tracing 28

1.9.1 Migratory route of HSPCs during definitive haematopoiesis as revealed by short-term lineage tracing experiments 28

1.9.2 The Cre/LoxP system can be used to long-term lineage trace cells 31

1.10 Aims and Objectives 35

2 Materials and Methods 36

2.1 Materials 36

2.1.1 Zebrafish husbandry and maintenance of lines 36

2.1.2 Transgenic lines used in this study 36

2.2 Methods 40

2.2.1 Treatments done on embryos and adult fish to activate transgene expression 40

2.2.1.1 Induction of transgenes by 4-OHT 40

2.2.1.2 Heat shock treatment 41

2.2.2 Whole-mount in situ hybridization (WISH) of zebrafish embryos 41

2.2.2.1 Collection, fixation and storage of embryos for WISH 42

2.2.2.2 In Situ Hybridisation on zebrafish embryos 42

2.2.2.3 Probe generation 44

2.2.3 Collection of fin clip biopsies from adult zebrafish 47

2.2.4 Molecular biology techniques 48

2.2.4.1 Genomic DNA extraction 48

2.2.4.2 Genotyping after in situ hybridization 50

2.2.4.3 Primer designing 51

2.2.4.4 Polymerase Chain Reaction (PCR) 52
2.2.4.5 Agarose Gel Electrophoresis (AGE) 54
2.2.4.6 Restriction digestion of Plasmid DNA 55
2.2.4.7 Isolation of DNA fragments by gel extraction 56
2.2.4.8 Ligation of linearized DNA fragments to linearized vector 56
2.2.4.9 Cloning of DNA in E.coli 57
2.2.4.10 Isolation of plasmid DNA 57
2.2.4.11 Measuring DNA concentration 58
2.2.5 Transgene Construction 58
2.2.6 Sequencing of gel extracted DNA fragments 58
2.2.7 Adult Zebrafish WKM isolation 58
2.2.8 Preparation of blood cell suspension from isolated WKM for FAC analysis ... 59
2.2.9 Flow cytometric analysis: Fluorescent activated cell (FAC) analysis: 60
2.2.10 In vivo fluorescent analysis and imaging: 60
2.2.10.1 Preparation of embryos for confocal microscopy 61
2.2.10.2 Confocal imaging and image acquisition 61

3 An inducible CreER^{T2} recombination system demonstrates that at least a subset of adult WKM haematopoietic cells arise from flk/*kdrl*+ endothelial cells in the early embryo 62

3.1 Introduction 62
3.2 *Flk1/kdrl*-expressing endothelial cells give rise to haematopoietic cells in the caudal haematopoietic tissue 68
3.3 The inducible *kdrlCreER^{T2}/loxP* system selectively labels cells of the vasculature and haematopoietic cells of the early larvae 71

3.4 The Cre/LoxP recombination system suggests that adult blood cells arise from *flk1/kdrl*+ endothelial cells 73

3.4.1 Development of alternative *Cre-reporter* genes 76

3.5 Adult blood cells originate from embryonic endothelial cells. 78

3.6 Summary 82

4 Confocal imaging reveals qmc551:eGFP expression in HECs of the vDA that undergo both basal EMT and apical departure 83

4.1 Introduction 83

4.2 Cells that undergo EHT are arterial endothelial cells that express flk1/kdrl:eGFP and the notch reporter csl:cerulean 84

4.3 Qmc551:eGFP labels haemogenic endothelial cells of the vDA that undergo EHT/bEMT 92

4.4 Other than bEMT/EHT cells of the vDA undergo apical departure 96

4.5 aD is independent of *runx1* and occurs more frequently, beyond 40hpf in *runx1* morphant embryos. 104

4.6 Apically departing cells have a nucleus and are large enough to be intact cells 109

4.7 Summary 111

5 Novel CRISPR/Cas9-generated *gfi1ab* mutant alleles suggest defects in definitive haematopoiesis in Gfi1aa and Gfi1ab depleted embryos…112

5.1 Introduction 112

5.2 Identification of two *gfi1ab* alleles and establishment of the *qmc552* and *qmc553* mutant lines 120

5.3 Preliminary data suggests that the loss of *gfi1ab* causes a reduction in eGFP+ cells with in the CHT and thymus of *qmc551* homozygous embryos. 127

5.4 A second unidentified mutant *gfi1ab* allele segregates in the progeny of the *qmc552* founder. 136

5.5 Summary 141

6 The loss of both *gfi1a* paralogues causes a reduction in *c- myb*+ cells within the larval CHT and kidney 143

6.1 Introduction 143

6.2 A single nucleotide polymorphism (SNP) in the *gfi1ab wt* allele interfere with characterisation and identification of *sh320* mutants 150

6.3 Characterisation of double mutants 157

6.3.1 *gfi1ab* mRNA levels are unchanged in the vDA of *qmc551* single homozygous and *qmc551;sh320* double homozygous embryos 158

6.3.2 *gfi1b* expression remains restricted to prRBCs in double mutant embryos 159

6.4 Double homozygous mutant embryos display reduced *c-myb* expressing cells in the CHT, thymus and kidney 161

6.5 Summary 165

7 Discussion 167

7.1 The constitutively active Cre recombination system we used labels the progeny of *flk1/kdrl*+ cells that seed the larval CHT, thymus and kidney 167

7.2 The *flk1/kdrlCreER*^{T2} inducible system labels a subset of adult WKM blood cells showing that adult HSCs arise from *flk1/kdrl*+ cells of the early embryo alone. 169

7.3 Cells labelled and detected in the adult WKM of both the constitutive and inducible systems depends on chance and genetic drift. 171

7.4 Future prospects for lineage tracing-experiments 174

7.5 *gfi1aa*^{qmc551Gt} labels HECs giving rise to HSPCs and serves as a good tool to study zebrafish HSC emergence 175

7.6 The loss of *gfi1aa* and *gfi1ab* do not affect bEMT/EHT but cause a reduction in HSPCs within the larval CHT. 179

References 181

Appendices 195

List of Figures

Figure 1.1: Schematic depiction of the hematopoietic hierarchy. 2

Figure 1.2: Chick-quail chimera experiments that were done by Dieterlen-Lievre reveal an intra-embryonic source of all hematopoietic compartments. 5

Figure 1.3: Haematopoiesis in the mouse embryo occurs in waves. 7

Figure 1.4: Haematopoiesis takes place in three waves within the zebrafish embryo. 14

Figure 1.5: Emergence of haemangioblasts from the lateral mesoderm and genes that are expressed during the primitive wave. 16

Figure 1.6: Schematic representation of the primitive and definitive waves and the intermediates involved. 18

Figure 1.7: Definitive HSCs are marked by their close association to the vDA and by the expression of the 2 major transcription factors, *runx1* and *c-myb*. 19

Figure 1.8: Schematic representation of Gfi1 and Gfi1b and the blood hierarchy 22

Figure 1.9: Evolutionary conservation of the *Gfi1* gene family between mammals and zebrafish 24

Figure 1.10: The *(gfi1aa:eGFP)*^{Gtqmc551} labels prRBCs and ECs of the vDA. 26

Figure 1.11: Upregulation of *gfi1ab* mRNA in cells of the vDA. 27

Figure 1.12: Migratory route of HSCs/Hematopoietic progenitor cells (HSPCs) during definitive haematopoiesis. 29

Figure 1.13: Schematic representation of the Cre/LoxP recombination system in general as seen in fish. 32

Figure 3.1: eGFP and tdTom fluorescence displayed in Tg(*flk1/kdrl:egfp*)^{qmc67} and Tg(*flk1/kdrl:tdTom*)^{qmc64} transgenic embryos at 26 hpf. 63

Figure 3.2: Cre-mediated recombination of the *ef1-α:Cre* reporter transgene in endothelial cells of embryos that also carry the *flk/kdrl:cre* transgene. 65

Figure 3.3: The 4-hydroxy-tamoxifen (4-OHT)-inducible Cre*ER*^{T2}/LoxP recombination system. 67

Figure 3.4: Lack of dsRed Cre reporter gene expression in the absence of the ligand 4-OHT. 68

Figure 3.5: Double transgenic Tg(*flk1/kdrl:cre*)^{qmc101},Tg(*ef1α*:Cre reporter) embryos display dsRed expression in the endothelium and in blood cells of the CHT, the thymus and the kidney at 5 dpf. 70

Figure 3.6: *flk1/kdrl*+ endothelial cells in the early embryo give rise to blood cells that seed the embryonic CHT and early larval thymus and kidney of 4-OHT induced embryos. 72

Figure 3.7: Blood cells in the adult fish kidney marrow are derived from *flk1/kdrl*+ endothelial cells. 74

Figure 3.8: The sequence of the 1325bp and 307bp PCR fragment matches the sequence of the non-recombined and recombined *ef1α:Cre reporter* allele respectively. 75

Figure 3.9: *ef1α:Cre reporter* gene recombination detected by PCR in the kidney marrow cells isolated from *Tg(kdrl:CreER*^{T2}*)*^{qmc106}*; Tg(ef1α:Cre reporter)* double transgenic fish that were treated with 4-hydroxy tamoxifen at embryonic stages. 81

Figure 4.1: *flk1/kdrl:eGFP;csl:cerulean* labels cells of the vDA that undergo EHT giving rise to HSCs that remain in the mesenchyme below. 88

Figure 4.2: *flk1/kdrl:eGFP;csl:Cerulean* expressing cells leave the mesenchyme to join circulation via the PCV..... 91

Figure 4.3: qmc551:eGFP+ cells of the vDA undergo basal EMT to give rise to blood cells that enter the mesenchyme below. 93

Figure 4.4: qmc551:eGFP;flk1/kdrl:tdTom double positive cells in the mesenchyme join the circulation via the PCV..... 95

Figure 4.5: qmc551:eGFP;kdrl:tdTom double positive cells leave the vDA apically as well 98

Figure 4.6: qmc551 eGFP+ endothelial cells of the vDA leave apically to join the circulation via the DA..... 100

Figure 4.7: *flk1/kdrl:eGFP;csl:cer* double positive cells of the vDA give rise to apically departing circular cells that are seen to join circulation directly via the DA. 103

Figure 4.8: In qmc551; runx1 morphants bEMT of HECs of the vDA is almost completely abrogated while events of apically departing cells are still observed..... 106

Figure 4.9: The time taken for HECs to depart apically is faster than the mechanism taken for EMT. 108

Figure 4.10: Cells that undergo aD have a nucleus and are intact cells. 110

Figure 5.1: The *wt gfi1ab* gene, transcript and protein..... 115

Figure 5.2: Injecting *qmc551* homozygous embryos with gRNA1 and gRNA2 along with Cas9 causes an out of frame mutation within the *gfi1ab* gene. 119

Figure 5.3: Identification of two founders that carry mutations in *gfi1ab* in cells of their germ line. 122

Figure 5.4: Identification and establishment of *qmc552 gfi1ab* mutant line. 124

Figure 5.5: Identification and establishment of *qmc553* mutant line. 126

Figure 5.6: *gfi1ab* compound heterozygous embryos show reduced numbers of eGFP+ cells in the CHT and thymus. 129

Figure 5.7: The genotype of the M1 embryo suggests that it is *gfi1ab* compound heterozygous embryo. 131

Figure 5.8: *gfi1ab* compound heterozygous embryos do not display a *wt* gDNA *gfi1ab* fragment. 135

Figure 5.9: Oligo sets specific to each allele were designed to distinguish between the *qmc552* and *qmc553* alleles. 138

Figure 5.10: Allele specific oligo sets suggest the presence of a second mutant allele in the *qmc552* founder. 140

Figure 6.1: The *sh320* has a four base pair (4bp) insert that causes a frame shift and thus a divergent amino acid coding sequence. 145

Figure 6.2: Genotyping of the *sh320* fish was unreliable and non reproducible. 149

Figure 6.3: Genotyping of these fish revealed either a *wt* or *sh320* homozygous sequence. 151

Figure 6.4: A polymorphism present in the *gfi1ab wt* allele was identified. 155

Figure 6.5: Genotyping of the *sh320* fish was modified and corrected. 156

Figure 6.6: *gfi1ab* mRNA is unchanged in the vDA of double mutant embryos. 159

Figure 6.7: *gfi1b* mRNA is unchanged in double mutant embryos and *gata1* remains restricted to circular RBCs. 160

Figure 6.8: *c-myb* expression is significantly reduced from 3dpf within cells of the CHT. 162

Figure 6.9: Reduced *c-myb* expression in sites of definitive haematopoiesis at 5dpf. 163

Figure 7.1: Schematic representation of the genetic drift theory applied to complete or partial recombination seen in the adult WKM marrow cells. 173

Figure 7.2: Schematic representation of the mechanism of apical EMT. 177

List of Tables

Table 2.1: Buffers and Solutions used 37

Table 2.2: Incubation times in Proteinase K of embryos at different developmental stages 43

Table 2.3: List of WISH antisense probes 46

Table 2.4: Sets of primers used for PCR amplification and sequencing 51

Table 2.5: Typical PCR reaction (25µl) set up using Standard Taq polymerase and Q5 polymerase 53

Table 2.6: General PCR thermocycler steps and conditions using Standard Taq Polymerase or Q5 Polymerase 54

Table 5.1: Genotype with respect to the *gfi1ab* gene and numbers of eGFP- (*qmc551*$^{+/+}$) and eGFP+ (*qmc551/qmc551*) single embryos analysed 132

Table 6.1: Genotypes with respect to *gfi1ab* of fin clip biopsies taken from ten *qmc551* homozygous fish using oligos DB392 and DB393 152

Table 6.2: Genotypes of finclipped fish repeated after identification of the polymorphism using oligos DB751 and DB752. 157

Table 6.3: The genotype of WISH embryos reveals the phenotype seen is due to a loss of both *gfi1ab* copies along with the loss of *gfi1aa.* 164

1 Introduction

1.1 Adult haematopoiesis, hematopoietic stem cells and their clinical importance

In a healthy adult human, more than ten different types of mature blood cells are produced (Lensch, 2012; Seita and Weissman, 2010). The blood cell types produced include ~$1x10^9$ erythrocytes and ~$1x10^8$ leucocytes and a full complement of platelets and all other mature blood cells (Wang et al., 1996) (Figure 1.1). The process by which these cells are generated is known as haematopoiesis. Haematopoiesis is maintained by hematopoietic stem cells (HSCs) (Lensch, 2012; Seita and Weissman, 2010; Weissman, 2000). Residing within the adult bone marrow (BM), HSCs are immature, self-renewing multipotent progenitor cells that maintain themselves as well as give rise to cells of all blood lineages (Doulatov et al., 2012) (Figure 1.1), thereby maintaining the blood system throughout life.

On top of the hematopoietic hierarchy, HSCs are the most naïve cells capable of reconstitution of the entire blood system in haemoablated hosts (Bordignon, 2006), owing to their capacity to self-renew, thus making HSCs an extremely valuable tool in the clinic. Every year, over 50,000 HSC transplantations are performed worldwide, commonly in the treatment of congenital blood disorders (Gratwohl, 2010). Allogeneic HSC transplantation (HSCT) (using BM donated from a histocompatible or HLA-matched donor, usually a sibling) is the most commonly applied therapy (Rowe et al., 2016). Over a third of patients requiring a donor for allogeneic BM transplantation do not find a histocompatible donor (Panopoulos and Belmonte, 2012). Allogeneic stem cell transplantation carry the risk of graft rejection and graft-versus-host disease (GVHD), which remains the leading cause of morbidity in patients (Holtan et al., 2014; Slukvin, 2013).

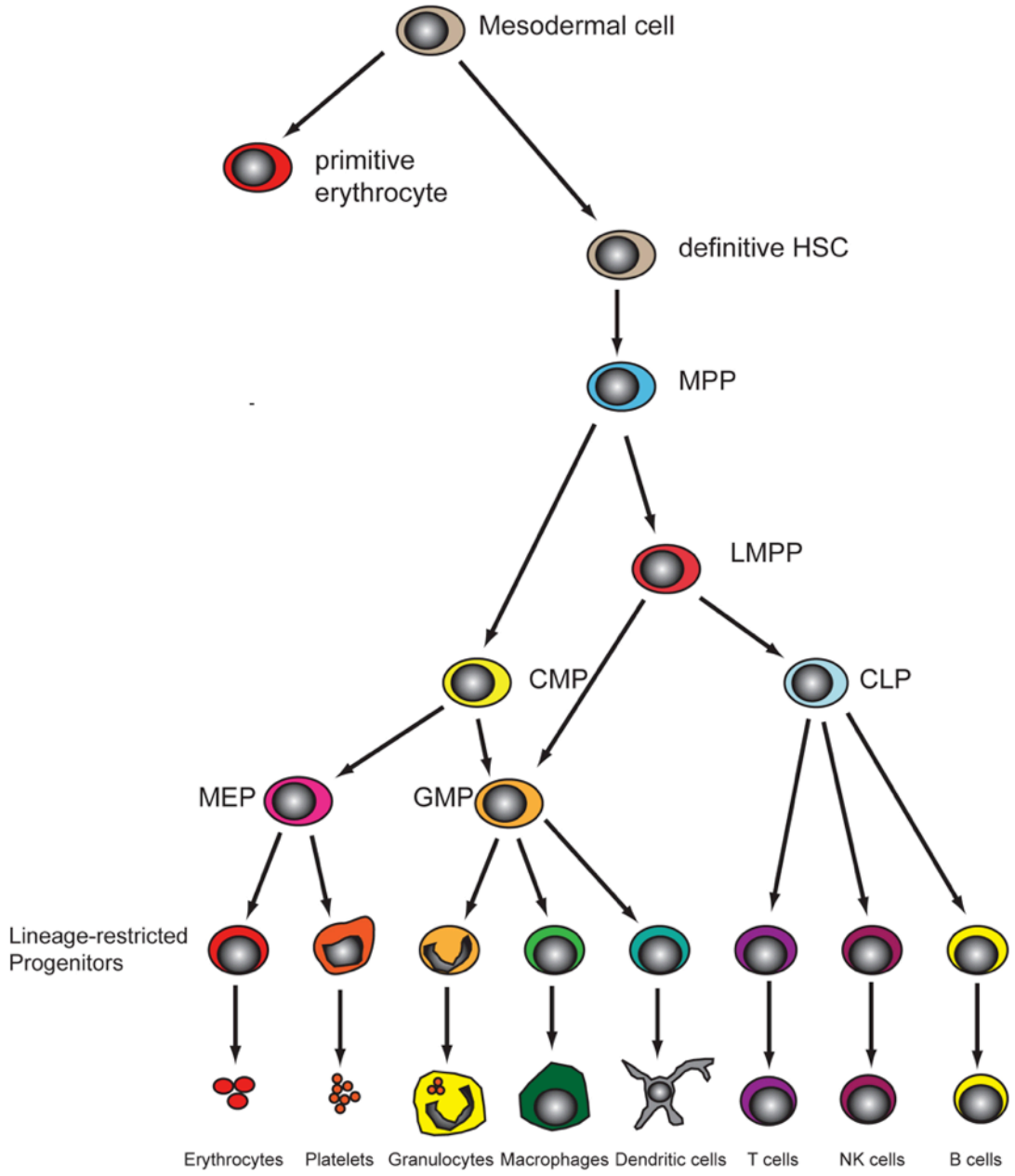


Figure 1.1: Schematic depiction of the hematopoietic hierarchy.
On top of the hierarchy is the immature, self-renewing, multipotent HSC. HSCs give rise to multipotent progenitors (MPP), which form the myeloid and lymphoid progenitor cells, i.e., Lymphoid multipotent progenitor (LMPP) and the common myeloid progenitor (CMP). LMPP can give rise to common lymphoid progenitors (CLP) and granulocyte/macrophage progenitors (GMP) and CMPs will further differentiate to form megakaryocyte/erythroid progenitors (MEP) as well as could give rise to GMPs. MEPs, GMPs and CLPs further give rise to lineage restricted progenitors that differentiate to form cells of specific blood lineages. (Image provided by Dr. Martin Gering).

Furthermore, in the absence of a histocompatible allogeneic donor, HSCT recipients rely on umbilical cord blood. Cord blood usually lacks the adequate amount of stem and

progenitor cells (HSPCs) that is required for the timely recovery and reconstitution of functional peripheral blood cells (Pineault and Abu-Khader, 2015; Rowe et al., 2016). Autologous BM transplants (using the patient's own BM) do not carry complications such as GVHD and are used dependent on whether functional, disease-free HSCs can be obtained from the patient. Many efforts are underway to engineering HSCs to use in HSCT therapy and research. The ideal engineered HSC should have self-renewal potential and should also be able to give rise to mature cells of all blood lineages for efficient oxygen transport, haemostasis and innately acquired immunity. Pluripotent stem cells (PSCs) (Murry and Keller, 2008) such as embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) (Panopoulos and Belmonte, 2012) could provide an alternative source of HSCs (Doulatov et al., 2012). Although it is possible to grow mature blood cells from ESCs, successful differentiation of PSCs into HSCs (without expressing transgenes from retroviral vectors in them) remains a challenge (Gratwohl, 2010; Walasek et al., 2012). Recent developments in protocols to differentiate pluripotent cells or reprogram more mature adult cells into HSPCs hold great promise (Ebina and Rossi, 2015; Sturgeon et al., 2013). Recently, iPSCs, along with the CRISPR/Cas9 technology-mediated gene repair of autologous cells, have augmented the progress of HSC engineering (Hendriks et al., 2016). The maintenance and expansion of isolated BM HSCs are another challenge scientists face (Lengerke and Daley, 2010). A deeper insight into the cellular origin and molecular programming of HSCs during vertebrate embryogenesis will help improve protocols aimed at deriving and maintaining HSCs from pluripotent cells *in vitro*. If this is achieved, we could produce libraries of HSCs from PSCs that could be transplanted as and when required by a patient (Lengerke and Daley, 2010).

1.2 Embryonic haematopoiesis: A historical overview

The study of haematopoiesis has a rich historical background, which involves the use of many vertebrate model organisms such as the chick, mouse and frog embryos and more recently the zebrafish embryo. All of these model organisms have their advantages for studying the ontogeny of HSCs. Data generated from each organism collated together can potentially provide a deeper understanding of the formation of HSCs during embryogenesis (reviewed by Ciau-Uitz and Patient, 2016; Ditadi et al., 2017; Dzierzak and Medvinsky, 2008; Medvinsky et al., 2011).

The yolk sac (YS) was believed to contain hematopoietic cells (HCs), but it was only until *in vitro* assays carried out by Moore and Metcalf in 1970 that the YS was experimentally demonstrated to be a site of active haematopoiesis (Moore and Metcalf, 1970). In their experiments, Moore and Metcalf showed that the YS harboured clonogenic myeloid progenitors or CFU-Cs for colony forming units-culture. CFU-Cs are progenitor cells that generate HC colonies *in vitro* and that do not have the ability to self-renew, although some can be replated. Moore and Metcalf described a second population of cells as being adult-type dHSCs and colony forming units-splenic (CFU-S). These cells have been described as immature progenitor cells that form morphologically different colonies of myeloid cells in spleens of irradiated animals following transplantation. CFU-Ss have a limited capacity of self-renew when transplanted a second time (reviewed by Medvinsky et al., 2011). Hence, with these experiments until the mid-1970s it was believed that the primary source of the adult hematopoietic system was the extra embryonic YS. Where HSCs originated and subsequently migrated to intraembryonic organs such as the liver, spleen, thymus and bone marrow (BM) (Ciau-Uitz and Patient, 2016; reviewed by Dzierzak and Medvinsky, 2008; Moore and Metcalf, 1970). In 1975 elegant chimeric experiments carried out by Dieterlen-Lievre in which quail embryos were transplanted onto the yolk sac (YS) of a chick

(Figure1.2) showed that while cells coming from the YS generated primitive red blood cells (prRBCs), they did not make a significant contribution to adult haematopoiesis (Dieterlen-Lievre, 1975). This suggested that HSCs found in the adult must have originated in the embryo itself (Dieterlen-Lievre, 1975).

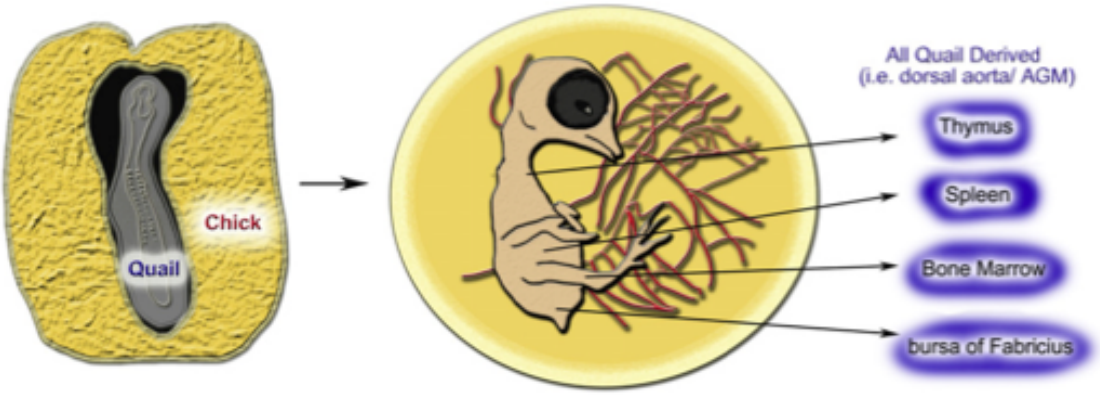


Figure 1.2: Chick-quail chimera experiments that were done by Dieterlen-Lievre reveal an intra-embryonic source of all hematopoietic compartments.
Dieterlen-Lievre was able to distinguish quail from chick cells by the morphology of the lymphocytes. Quail lymphocytes have a single thick, irregular heterochromatic mass plus some finer DNA material around the margin of the nucleus, chick lymphocytes on the other hand has numerous, finely dispersed chromocentres (Dieterlen-Lievre, 1975). (Image taken from (Zape and Zovein, 2011).

Tissue transplantation experiments carried out in *Xenopus*, also demonstrated the intraembryonic origin of HSCs (Kau and Turpen, 1983; Maéno et al., 1985). The relevance of the data generated in the avian and frog system to all vertebrates became clear only when the findings were confirmed in the murine system (reviewed by Dzierzak and Medvinsky, 2008; Medvinsky et al., 2011). A more stringent assay compared to the one carried out by Moore and Metcalf was carried out in the 1990s by Medvinsky and colleagues. In these experiments, in order to increase the sensitivity of the assay, recipient mice were irradiated to prevent the formation of endogenous CFU-S. These experiments revealed that the YS produced no CFU-S before embryonic day (E) 9.5 (Medvinsky et al.,

1993). Subsequent studies showed a lack of definitive HSCs before E11.5 (Medvinsky and Dzierzak, 1996; Müller et al., 1994). Thus these studies showed YS haematopoiesis occurs in the absence of definitive HSCs and confirmed that like in the avian and amphibian systems, mammalian haematopoiesis has an intra-embryonic origin where the aorta-gonad-mesonephros (AGM) region is able to independently give rise to HSCs (Ciau-Uitz and Patient, 2016; Dzierzak and Medvinsky, 2008; Medvinsky et al., 2011; Medvinsky and Dzierzak, 1996; Müller et al., 1994).

Over the last few years, research in the field of developmental haematopoiesis has established that during vertebrate embryogenesis, haematopoiesis occurs in successive yet overlapping waves that occur at distinct anatomical locations (Ciau-Uitz et al., 2014; Clements and Traver, 2013; Medvinsky et al., 2011). The following section will put light on haematopoiesis in the murine system.

1.3 Haematopoiesis in the murine embryo

Haematopoiesis occurs in at least three waves in the developing murine embryo. The first wave is a primitive or transient wave, which takes place in the YS blood islands. The second wave or pre-definitive wave is believed to occur within the YS but from haemogenic endothelial cells (HECs). The first two waves of haematopoiesis do not produce adult type HSCs with long-term reconstitution (LTR) ability. The third or definitive wave begins intra-embryonically when HSCs with LTR potential arise from the aorta/gonad/mesonephros (AGM) region from HECs of the ventral wall of the dorsal aorta (vDA). These cells are known to maintain haematopoiesis in the late foetus and adult BM (reviewed by (Ciau-Uitz et al., 2014) (Figure1.3).

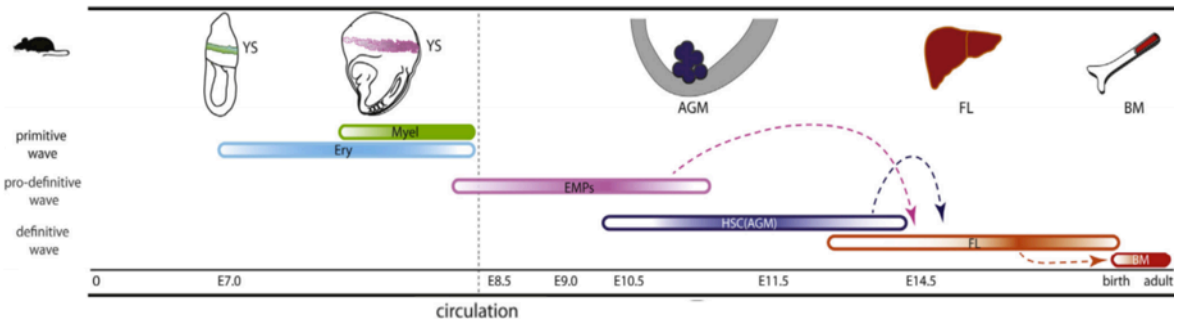


Figure 1.3: Haematopoiesis in the mouse embryo occurs in waves.
A primitive or embryonic wave, which occurs in the extra-embryonic YS from E7. This wave gives rise to a transient population of primitive myeloid (Myel) (green) and erythroid cells (Ery) (light blue). The second or prodefinitive wave is believed to also occur in the yolk sac where a transient population of EMPs (pink) originate from haemogenic endothelium. The EMPs then seed the fetal liver (FL) (pink dotted arrow). The third or definitive wave begins when HSCs with long-term reconstitution ability arise from the AGM region from a specialised population of endothelial cells known as haemogenic endothelial cells (HECs). HECs are located in the ventral floor of the dorsal aorta (DA, dark blue bar). In the mouse HSCs emerge from the lumen of the dorsal aorta in the form of clusters. Once the haematopoietic clusters emerge, HSCs enter embryonic circulation, which transports them to the FL. The FL being the main site of embryonic haematopoiesis in the mouse. But the FL is a transient site. From the FL, HSCs then migrate to the final site of haematopoiesis, the bone marrow (BM) (Image from Ciau-Uitz et al., 2014).

1.3.1 The primitive wave or embryonic haematopoiesis

In the mammalian embryo, the first wave of haematopoiesis is the primitive wave, which occurs between day 7 postcoitum (E) to E7.5 following the onset of gastrulation. The first blood cells emerge from the YS when the extra-embryonic mesoderm gateways through the posterior primitive streak to form blood islands (Ferkowicz and Yoder, 2005; Lacaud and Kouskoff, 2016; Silver and Palis, 1997). The primitive wave generates a transient population prRBCs and primitive myeloid cells (reviewed by Ciau-Uitz et al., 2014; Medvinsky et al., 2011). Cells of the primitive wave develop in the YS from bipotent progenitor cells, which are capable of giving rise to both blood and endothelial cells (ECs), a mesodermal cell, the haemangioblast (Huber et al., 2004). The YS is believed to generate erythrocytes from E7 to E9 (Palis et al., 1999). Mammalian prRBCs are large and nucleated and express both embryonic globins such as βH1 and adult globins like βmajor (Kingsley et al., 2004; Palis et al., 1999). GATA-1, a zinc finger transcription factor, plays

a main role in mediating erythroid gene expression and interacts with many proteins such as FOG1 (friend of GATA) and PU.1 (Cantor and Orkin, 2002). In the blood islands haematopoietic differentiation takes place prior to the formation of the vasculature (Medvinsky et al., 2011). prRBCs in the mammalian embryo begin to go into circulation from E8.25 to E16.5 (Brotherton et al., 1979; McGrath et al., 2003). As the embryo begins to grow and its requirement for oxygen rises, the numbers of circulating erythrocytes need to be increased. The emergence of a large number of enucleated definitive RBCs from the fetal liver at E12.5 completes this requirement of the growing embryo (reviewed by Frame et al., 2013).

At E9 the primitive wave gives rise to some primitive myeloid cells such as macrophages and megakaryocytes along with cells of the erythroid lineage (Dzierzak and Medvinsky, 2008; reviewed by McGrath et al., 2015a; 2015b; Swiers et al., 2010; Tober et al., 2007). Tissue development in the growing embryos requires cells of the primitive wave (Palis and Yoder, 2001). However, primitive cells are not enough to maintain the embryo unto adulthood. Cells of subsequent waves gradually replace the primitive cells.

1.3.2 A prodefinitive wave of haematopoiesis gives rise to erythromyeloid progenitors

The second wave of haematopoiesis which follows the primitive wave is the prodefinitive wave (Ciau-Uitz et al., 2014). Following the production of primitive cells, a transient population of erythromyeloid progenitors (EMPs) and lymphoid-primed multipotent progenitors (LMPPs) arise from HECs that surround functional blood vessels within the YS (Ciau-Uitz et al., 2014; Ditadi et al., 2017; Lensch, 2012; Palis et al., 1999; 2010; Seita and Weissman, 2010). EMPs and LMPPs of the prodefinitive wave cannot be assigned to either the primitive or the definitive wave. Hence it has been hypothesised that the

formation of these cell populations may be an independent wave which occurs within the YS (reviewed by Ditadi et al., 2017).

The EMPs produced during this second wave are $CD41^{+}$c-$kit^{+}CD16/32^{+}$ and also express adult globins along with rare B lymphoid potential (Chen et al., 2011; McGrath et al., 2015b). These cells are then believed to migrate to the fetal liver to establish myeloerythropoiesis before the arrival of definitive HSCs (Ditadi et al., 2017; Lacaud and Kouskoff, 2016; reviewed in McGrath et al., 2015b). EMPs originate in the YS was demonstrated in vascular endothelial (VE)-cadherin null mice embryos that lack circulation (Rampon and Huber, 2003) and in *Ncx1*$^{-/-}$ embryos that lack a functional heartbeat (Lux et al., 2008). In both studies very few EMPs were detected in the embryo proper, thus indicating that EMPs arise mainly in the YS. Fate-mapping experiments carried out by Gomez Perdiguero et al, identified a sequence of YS-derived, and EMP-derived haematopoiesis in the fetal liver and YS derived EMPs have the ability to produce macrophages that persist after birth (Perdiguero et al., 2015).

Just after the start of hematopoietic activity in the YS, hematopoietic activity begins in an intra-embryonic site called the para-aortic splanchnopleura (P-Sp). Explants taken from this region were cultured *in vitro*. These experiments demonstrated that the p-Sp region was capable of producing B-cell precursors and multipotent progenitors (Cumano et al., 1996).

At this stage, it might be important to mention that the direct transplantation of a large number of cells prior to E10 into an adult irradiated recipient mouse does not have the ability to produce long-term haematopoietic repopulation (reviewed by Medvinsky et al., 2011; Müller et al., 1994). Thus indicating a lack of engraftment ability, which means that at least until E10 adult type HSCs with LTR ability are absent.

1.3.3 Definitive haematopoiesis in the murine embryo

The third wave of murine embryonic haematopoiesis is the definitive wave. A crucial feature defining this wave of haematopoiesis is its ability to produce bona fide HSCs that are capable of engraftment and long-term reconstitution of all blood cell types in irradiated wild-type (wt) adult recipients (Medvinsky et al., 2011). Definitive HSCs (dHSCs) are responsible for the maintenance of both late foetal and adult haematopoiesis. It is only by E10.5 that the first LTR-HSCs are seen to emerge intra-embryonically in the aorta-gonad-mesonephros (AGM) region. This region comprises the dorsal aorta (DA), the surrounding mesonephros and the urogenital ridges. HSCs are believed to arise from ECs of the ventral wall of the dorsal aorta (vDA) (Medvinsky et al., 1993; Medvinsky and Dzierzak, 1996; Müller et al., 1994).

1.3.3.1 HSCs are derived from haemogenic endothelium of the AGM

The association of HCs with the DA is known for a long time. In the mouse DA, both the dorsal and ventral walls display an association with HC clusters. However, transplantation experiments carried out on both the dorsal and ventral sections of the DA have revealed that it is only the clusters associated with the ventral wall that contains HSCs (Ivanovs et al., 2014; Taoudi and Medvinsky, 2007). The close association of haematopoietic clusters to the endothelium of the DA implies that HSCs arise from ECs. Lineage tracing studies carried out on chick embryo DA cells in which the DA cells were labelled using either low-density lipoprotein uptake or by viral transfection provide substantial evidence that HC clusters derive from ECs (Jaffredo et al., 1998; Jaffredo et al., 2000). Early hematopoietic and ECs both in the YS blood islands, the fetal liver and the AGM region are seen to express fetal liver kinase-1 (flk1) also known as the vascular endothelial growth

factor receptor 2 (vegfr2) (reviewed by Gore et al., 2012; Kabrun et al., 1997). Conditional labelling and lineage tracing studies carried out by Chen et al. who used constitutive VE-cadherin-specific Cre-mediated labelling of ECs in the early mouse embryo showed that adult HSCs were previously positive for the endothelial marker (Chen et al., 2009). Another study by Zovein et al. using an induced VE-cadherin-specific Cre-mediated labelling of the progeny of VE-Cadherin^{+} cells also strongly suggested the endothelial origin of adult blood cells. Besides this Zovein et al. also found that the mesenchymal cells surrounding the DA do not give rise to haematopoietic cells (Zovein et al., 2008). These studies further support the idea that HSCs arise from a mesodermal cell with an EC as an intermediate. HSCs are believed to originate from the AGM region from a specialised population of ECs called haemogenic endothelial cells (HECs) of the vDA (Chen et al., 2009; Chen et al., 2011; Frame et al., 2016; Zovein et al., 2008).

At E10.5 the mouse embryo displays HSC clusters that are seen to be "budding" from the aortic floor, in a process termed, endothelial to hematopoietic transition (EHT) (Kissa and Herbomel, 2010). In the mouse in vivo imaging is not possible due to the opaque nature of the embryo. However, in vitro imaging through cultured 10.5E mouse AGM explants showed the "budding" of hematopoietic clusters into the lumen of the DA, where the clusters emerge from CD31+ ECs of the aortic floor (Boisset et al., 2010).

Several transcription factors (TFs) are involved in HSC development. The following section highlights the role of some of the TFs involved in HSC formation.

1.3.3.2 Primary regulatory factors involved in HC development.

At the genetic level, HSC development depends on some TFs. Many of which belong to the ETS and Gata families (Ciau-Uitz et al., 2014; Ciau-Uitz et al., 2010). Because of the multiple functions of these TFs in HSC development, it's hard to pinpoint which TFs are

directly involved in the emergence of HSCs from the DA. Nevertheless, the major TFs involved in HSC development include *Scl, Gata2* and *Runx1* (Ciau-Uitz et al., 2014). *Scl* and *Gata2* have been shown to play context-dependent roles in the murine embryo and are vital for both embryonic and adult blood development (Ng et al., 1994; Porcher et al., 1999; Porcher et al., 1996). *Gata2* and *Scl* null mouse embryos die before HSC emergence, thus making it difficult to study the role of these TFs in EHT (Ng et al., 1994; Tsai et al., 1994). In contrast to *Gata2* and *Scl*, *Runx1* plays almost no role in primitive haematopoiesis. Nonetheless the endothelium and the underlying mesenchyme express *Runx1*, which is a vital transcriptional regulator for definitive haematopoiesis and HSC emergence (de Bruijn et al., 2000; North et al., 2002). Transplantation experiments have shown that *Runx1* is expressed in AGM derived definitive HSCs (North et al., 2002; Nottingham et al., 2007). No definitive HSCs are derived in the mouse embryo in the absence of *Runx1* (Dzierzak and Medvinsky, 2008; North et al., 2002; Tober et al., 2013; Wang et al., 1996). Additionally, when *Runx1* was deleted specifically in endothelial cells and Vav1 expressing cells, it was seen that Runx1 is required for HSC emergence and not for HSC maintenance (Cai et al., 2011; Chen et al., 2009; North et al., 1999). Runx1 is required for the emergence of EMPs that are derived from YS haemogenic endothelium. Additionally, Ly6a-expressing haemogenic endothelium of the AGM region that gives rise to HSCs also requires *Runx1*. These findings suggest the requirement of Runx1 for the emergence of all blood lineages from haemogenic endothelium thereby it was hypothesised that Runx1 plays a role in EHT (Chen et al., 2009; Chen et al., 2011). In $Runx1^{-/-}$ mice a lack of intra-aortic clusters was observed, which suggests that Runx1 is required for EHT within the AGM region (North et al., 1999). Interestingly, the loss of haematopoietic activity in the AGM region of Runx1-depleted mice can be partially rescued by overexpressing *Runx1*. This suggests that definitive haematopoiesis occurs intra

embryonically, is localised to the AGM and is independent of yolk sac haematopoiesis (Goyama, 2004). In another study, Runx1$^{-/-}$ embryos showed that although yolk sac derived erythropoiesis was almost normal, fetal liver haematopoiesis was abrogated due to which the embryo died at around E12.5 (Okuda et al., 1996). Mucenski et al (1991) carried out experiments on *c-myb* knockout mice, and showed that although primitive erythropoiesis was unaffected, the E15 embryo develops severe anaemia. This suggested that the transcription factor *c-myb* plays a pivotal role in definitive haematopoiesis (Mucenski et al., 1991).

The HSCs in the AGM region migrate to the foetal liver, where they expand and migrate transiently to the foetal spleen, eventually to home the BM. The BM is the final niche of HSCs where they remain throughout the lifetime of the adult mammal (Cumano and Godin, 2007, reviewed by Sugiyama et al., 2011).

1.4 Haematopoiesis in zebrafish

The zebrafish (*Danio rerio*) has emerged as an ideal model organism in which to study haematopoiesis, the main reasons for this are; due to their highly conserved haematopoietic system, which occurs in sites that are anatomically similar to those of mammals (Jing and Zon, 2011), thus making it a good model to understand and study the cellular origin and molecular programming of HSCs. Zebrafish embryos can easily be exploited for genetic modification hence making it easy to carry out transgenic or loss-of-function studies. Finally, the transparent nature of its embryo makes the zebrafish extremely amenable for *in vivo* imaging.

Like in all other vertebrates, haematopoiesis in zebrafish occurs in 3 successive waves; a transient primitive wave a prodefinitve wave and a permanent or definitive wave (Ciau-Uitz et al., 2014) (Figure 1.4).

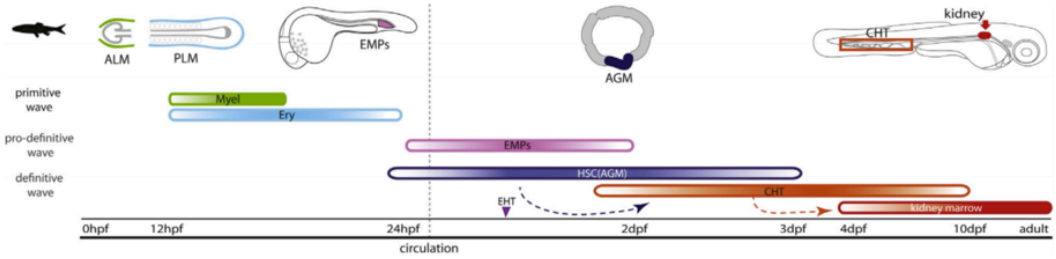


Figure 1.4: Haematopoiesis takes place in three waves within the zebrafish embryo.
The first or primitive wave in the zebrafish takes place in the anterior and posterior lateral mesoderm (ALM and PLM). This wave gives rise to erythroid (green bar) and myeloid cells (light blue bar). The second or prodefinitive wave comprises a transient population of erythromyloid progenitors (EMPs) (pink bar). EMPs in the zebrafish are thought to arise from the posterior blood islands (PBI). The third or definitive wave starts in the zebrafish when HSCs begin to arise from haemogenic endothelial cells of the ventral floor of the dorsal aorta via an endothelial to haematopoietic transition (EHT). From here the cells enter the mesenchyme below the ventral wall of the dorsal aorta and enter circulation via the posterior cardinal vein and seed the caudal haematopoietic tissue (CHT). This is a transient site where cells expand and differentiate and from here HSCs migrate to the whole kidney marrow (WKM), which is the final site of larval and adult haematopoiesis in the zebrafish (Image from (Ciau-Uitz et al., 2014).

1.4.1 The primitive wave

The primitive wave of haematopoiesis takes place in distinct anatomical locations within the zebrafish embryo: the anterior lateral mesoderm (ALM) and the posterior lateral mesoderm (PLM) (Ciau-Uitz et al., 2010) (Figure 1.4 and 1.5A). It is believed that the main purpose of the primitive wave is to produce primitive red blood cells that are required to transport oxygen, which is required for tissue formation in the developing embryo (Orkin and Zon, 2008). At early somitogenesis, a transient population of haemangioblasts (a cell capable of giving rise to both blood and ECs) arises from the ALM. In the zebrafish *scl* is expressed from the 2 to 3 somite stage (ss) in the haemangioblasts along with *lmo2, gata2* and *fli* (Liao et al., 1998). These haemangioblasts differentiate into primitive myeloid progenitors that are *pu.1*$^{+}$ and ECs that are flk1/*kdrl*$^{+}$ (Bennett et al., 2001), (Lieschke et al., 2002). *Pu.1* is expressed in the ALM at 6ss along with *scl, lmo2* and *gata2*. Later, at 10ss it is also expressed in the PLM (Figure 1.5B-E). The PLM gives rise to a second population of haemangioblasts that differentiate into erythroid and ECs (Ciau-

Uitz et al., 2010, Paik and Zon, 2010). *Gata1* an erythroid specific gene is expressed at 5ss in the PLM. Once the cells have been specified in the PLM they then migrate to the midline to form a chord of cells known as the intermediate cell mass (ICM). Unlike in birds and mammals, in the zebrafish embryo both the primitive wave as well as the definitive wave occurs intraembryonically in close proximity in the ICM (reviewed by (Clements and Traver, 2013). The ICM is comprised of prRBCs, ECs of the two major trunk vessels, the dorsal aorta (DA) and the posterior cardinal vein (PCV).

1.4.2 In the zebrafish erythromyeloid progenitors arise during a second wave of haematopoiesis within the PBI

Just like in the mouse embryo, in the zebrafish embryo a second or pro-definitive wave of haematopoiesis gives rise to a population of EMPs. In the zebrafish embryo this second wave occurs in the posterior blood islands (PBI) where EMPs are established before HSC emergence (Bertrand et al., 2007). Although the EMPs give rise to erythroid and myeloid progeny *in vitro* (Bertrand et al., 2007), there is no substantial evidence of zebrafish EMPs giving rise to erythroid cells *in vivo* (Ciau-Uitz et al., 2014).

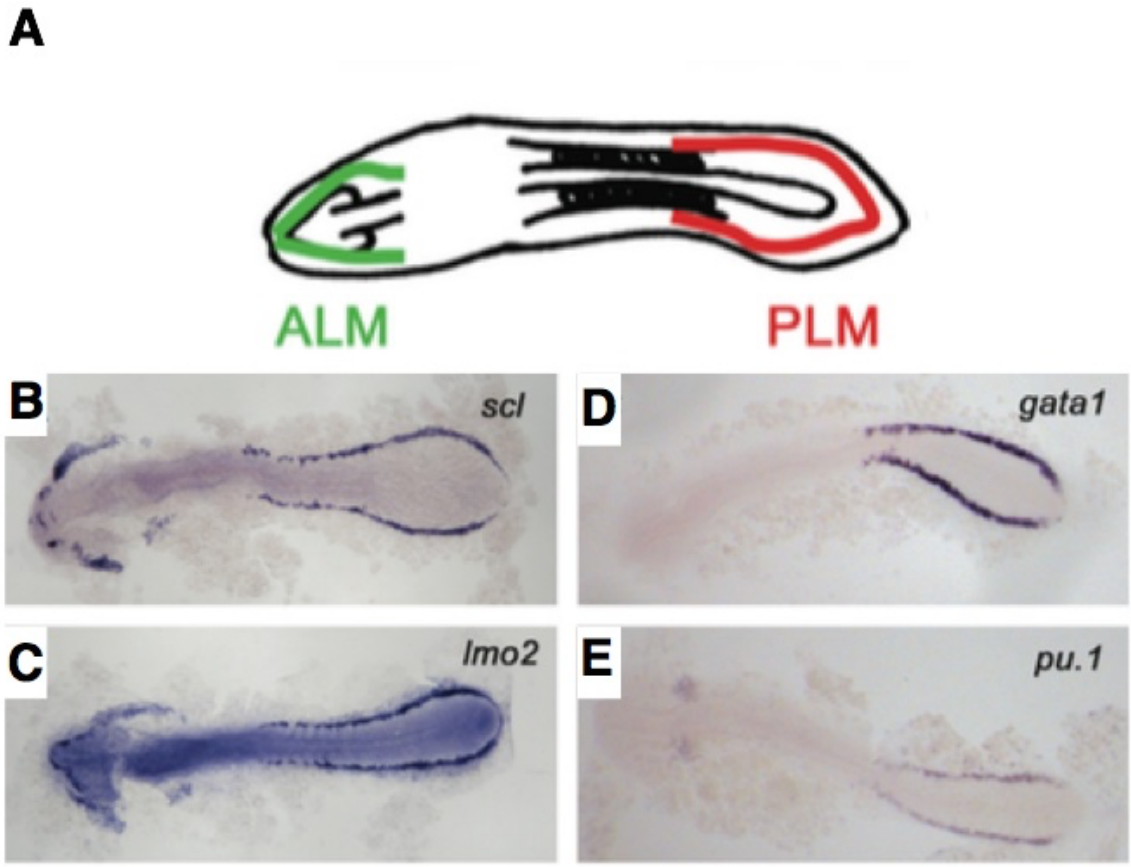


Figure 1.5: Emergence of haemangioblasts from the lateral mesoderm and genes that are expressed during the primitive wave.
A) Schematic representation of a flat mounted embryo highlighting the ALM (green) and PLM (red) from where the population of haemangioblasts arises during segmentation period. (B-C) 10ss embryos in flat mount WISH experiments showing expression of haemangioblasts markers (B) *scl* which is expressed in both the ALM and PLM and (C) *lmo2* and expression of erythroid marker (D) *gata1* and expressed mainly in the PLM (E) while myeloid marker *pu.1* expressed in both ALM and PLM. (Image taken from Paik and Zon, 2010).

1.4.3 The haemogenic endothelium of the vDA gives rise to definitive HSCs via an endothelial-to-haematopoietic transition (EHT)

The primitive erythrocytes enter circulation at around 25 to 26 hours post fertilization (hpf) as putative HSCs begin to emerge from the ventral wall of the dorsal aorta (vDA) or the AGM region from a specialised population of ECs called haemogenic endothelial cells (HECs) via basal epithelial-to-mesenchymal transition (bEMT) or more specifically an

endothelial-to hematopoietic transition (EHT) as they change from an endothelial to a hematopoietic phenotype (Bertrand et al., 2010; Kissa and Herbomel, 2010; Lam et al., 2010; Thambyrajah et al., 2016; Zhen et al., 2013) (Figure1.6). Confocal time-lapse imaging carried out on live zebrafish embryos by Kissa and Herbomel revealed bending from arterial endothelial cells to enter the mesenchyme underneath (Kissa and Herbomel, 2010). Herbomel et al (2010) described how an EC of the vDA leaves the aorta by changing morphology and squeezing out of the vessel wall, finally rounding off and entering the mesenchyme below, and the endothelial lining closes up once the cell leaves. This event occurs from 34 to 60 hpf, with events peaking between 40 to 52 hpf (Kissa and Herbomel, 2010). Once in the mesenchyme the cells enter circulation through the vein (Kissa et al., 2008). Specification of HECs integrates EC programming prior to and during gastrulation with extrinsic signals during somitogenesis (Ciau-Uitz et al., 2014; Clements and Traver, 2013).

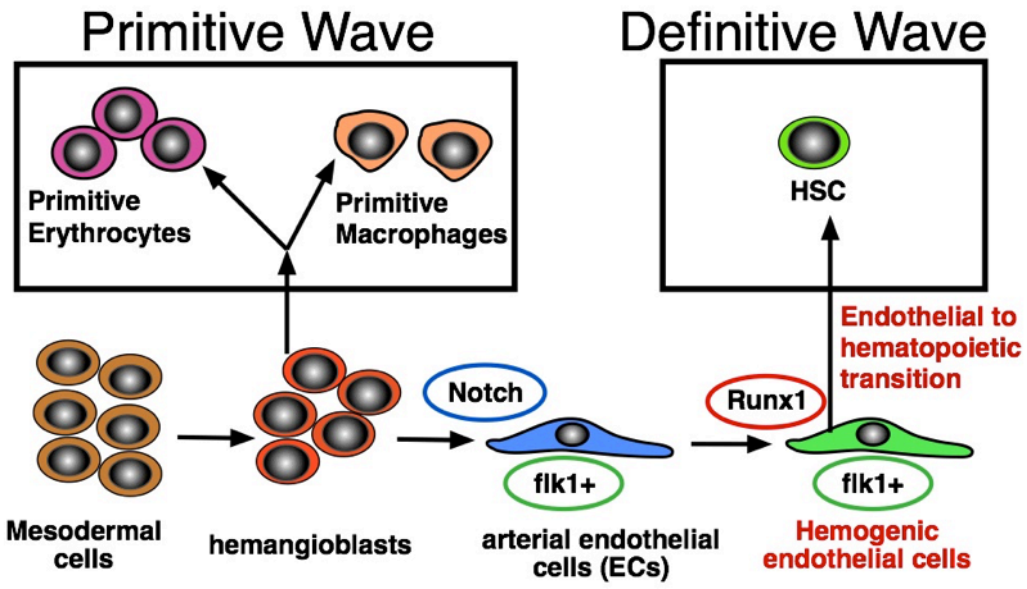


Figure 1.6: Schematic representation of the primitive and definitive waves and the intermediates involved.
The formation of the primitive and definitive waves starts from mesodermal cells, which give rise to the haemangioblasts that has a dual ability of giving rise to blood (of the primitive wave) and arterial endothelial cells that are *flk1*/*kdrl*+ and express Notch. A subpopulation of these cells become haemogenic and in the presence of *runx1* undergo EHT to give rise to HSCs.

The integration of the signals leads to the expression of the transcription factor *runx1*, which plays a vital role in HSC emergence (Chen et al., 2009; Gering and Patient, 2005; Kalev-Zylinska et al., 2002; North et al., 1999) (Figure 1.6). Studies on *runx1 morphant* embryos demonstrated that while *runx1* is dispensable in primitive haematopoiesis, it is required for definitive haematopoiesis as a reduction of *runx1* causes loss of *c-myb* expressing definitive progenitors (Gering and Patient, 2005; Kalev-Zylinska et al., 2002). Hence HSCs are identified by the expression of the transcriptional regulators *runx1* and *c-myb* (Figure 1.7). Further live imaging carried out on *flk1/kdrl:eGFP* trangenic embryos that were injected with a *runx1* morpholino showed that as HECs in the vDA tried to bend to undergo EHT they fragmented as they tried to round up to form an HSPC hence the EHT process was abortive (Kissa and Herbomel, 2010). This would suggest *runx1* is required for EHT. A study carried out on *runx1* mutants supports the idea of runx1 being

required for the initiation of definitive haematopoiesis but most embryos lost blood circulation and become bloodless between 8 to 12dpf and died during this bloodless phase by 11 to 20dpf. Although most of these embryos did not survive, interestingly, 20% recovered circulation and survived unto adulthood (Sood et al., 2010). Altogether these data suggest an alternative mechanism that works parallel to runx1 and that is independent of it. The loss of HSPCs and their haematopoietic progeny in the absence of *runx1* would indicate that *runx1* positive cells in the vDA are potential HSCs or their precursors (Burns, 2005; Gering and Patient, 2005; Jin et al., 2009).

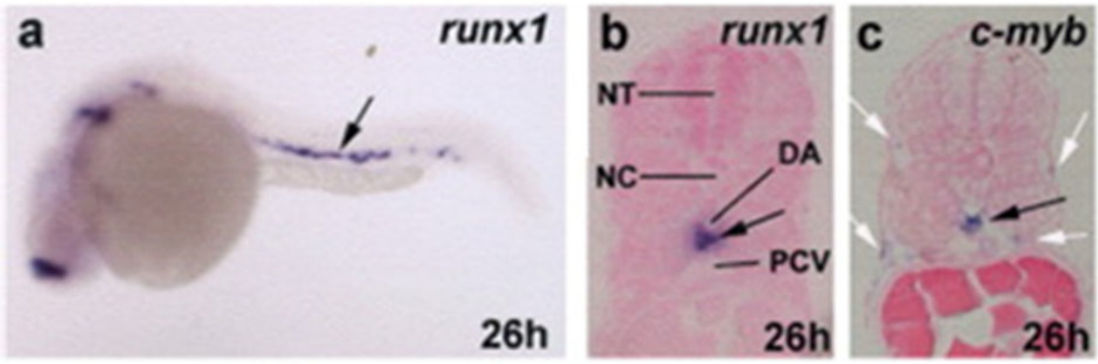


Figure 1.7: Definitive HSCs are marked by their close association to the vDA and by the expression of the 2 major transcription factors, *runx1* and *c-myb*.
In situ experiments show the expression of a) *runx1* expression in putative HSCs in the midline of a 26 hpf embryo b and c) Transverse sections show cells expressing the transcription factors *runx1* and *c-myb* and are associated with the vDA that lies dorsal to the PCV. (Image from (Gering and Patient, 2005).

Despite the similar origins of the primitive and definitive waves in the post-gastrula embryo, primitive and definitive blood cells have very different signalling requirements. Studies carried out on mutant zebrafish embryos and on zebrafish embryos treated with small molecule inhibitors that block defined signalling pathways such as Hedgehog, Vegf and Notch have shown that definitive haematopoiesis requires these signalling pathways, whereas primitive erythropoiesis is completely unaffected (Gering and Patient, 2005). Interestingly, these three signalling pathways are also required for arterial gene expression

in the DA (Lawson et al., 2001; Lawson and Weinstein, 2002). Once the HSPCs have joined circulation the bloodstream takes them to the vessel plexus. Here they leave the vessels and seed the mesenchyme, forming the caudal hematopoietic tissue (CHT). The CHT is a transient site of haematopoiesis in the zebrafish tail mesenchyme, where HSPCs occupy perivascular niches (Tamplin et al., 2015) before they migrate to the kidney and thymus which are the sites of larval and adult haematopoiesis and lymphopoiesis in the fish (Ellett and Lieschke, 2010).

Both in the murine system and in zebrafish, a number of signalling molecules and signalling sources have been implicated in HSC development and expansion (Burns, 2005; Clements et al., 2011; Gering and Patient, 2005; Nguyen et al., 2014; North et al., 2009; Wilkinson et al., 2012). One such transcription factor is Gfi1 and its closely related paralogue Gfi1b, which are transcriptional repressors known to play an overlapping role in adult HSCs and have distinct functions.

1.5 The structure and function of Gfi1a and Gfi1b

The *Growth Factor Independence 1(Gfi1)* is a transcriptional repressor and is a proto-oncogene (Gilks et al., 1993). There are two mammalian *Gfi1* paralogues: *Gfi1/Gfi1a* and *Gfi1b* (Gilks et al., 1993; Grimes et al., 1996). Both these *Gfi* genes are members of the Snail family and share high sequence homology. The Gfi1 and Gfi1b proteins are made up of 3 domains; a 20 amino acid N-terminal repressor SNAG (SNAG/Gfi1) domain, a C-terminal highly conserved DNA-binding domain with 6 C_2H_2 zinc fingers. Between the SNAG domain and the DNA binding domain is a less well characterised region (Cooney et al., 2013; Möröy et al., 2015; van der Meer et al., 2010) (Figure 1.8A and 1.9A).

During haematopoiesis many transcription factors play a role. The Gfi1 and Gfi1b play a role in the optimization of such regulatory factors (Cellot and Sauvageau, 2005; Kazanjian

et al., 2006; Möröy, 2005) (Figure 1.8B). Studies carried out on transgenic mice harbouring a reporter gene driven by *Gfi1* and *Gfi1b* promoter sequences analysed the expression patterns of both genes during haematopoiesis (Vassen et al., 2007; Yücel et al., 2004). In the mouse primitive RBC progenitors co-express both *Gfi1* and *Gfi1b* (Moignard et al., 2015) and Gfi1 or Gfi1b single knockout mice display normal primitive erythropoiesis (Hock et al., 2003; Saleque et al., 2002). However, the loss of both Gfi1 and Gfi1b results in a reduction of embryonic βH1 expression in the mouse YS, suggesting a defect in primitive erythropoiesis (Lancrin et al., 2012).

In the definitive blood system of the mouse a loss of Gfi1 alone has no effect on EHT. By contrast, Gfi1/Gfi1b double knockout mice show deficiencies in the haemogenic endothelium of both YS and the AGM region. Haemogenic endothelium derived blood cells within the YS, are unable to down-regulate endothelial genes and thus fail to go into circulation (Lancrin et al., 2012). While inside the embryo proper, although HECs of the vDA are specified, they fail to undergo EHT (Thambyrajah et al., 2015). In the murine embryo both *Gfi1* and *Gfi1b* are not only co-expressed during EHT but also in HSCs. Gfi1a and Gfi1b are expressed in HSCs and in myeloid progenitors. In HSCs, Gfi1a and Gfi1b are required to maintain both self-renewal capacity and quiescence (Khandanpour et al., 2010; Zeng et al., 2004). *Gfi1a* knockout mice develop severe inner ear defects that results in deafness and ataxic behaviour. These mice are also seen to suffer from neutropenia, which makes them susceptible to bacterial infection. Embryonic development remains normal but the new born knockout mice show growth retardation which ultimately leads to the death of the mice few months after birth (Hock et al., 2003; Karsunky et al., 2002; Wallis, 2003). Gfi1b can substitute for Gfi1a in haematopoiesis but is not sufficient for normal inner ear formation (Fiolka et al., 2006). While mice that are deplete of Gfi1b survive only unto gestation and display defects in cell types that do not co-express Gfi1

such as in definitive RBCs and megakaryocytes (Saleque et al., 2002). Conditional deletion of both *Gfi1a* and *Gfi1b* in the adult is incompatible with HSC maintenance (Khandanpour et al., 2010).

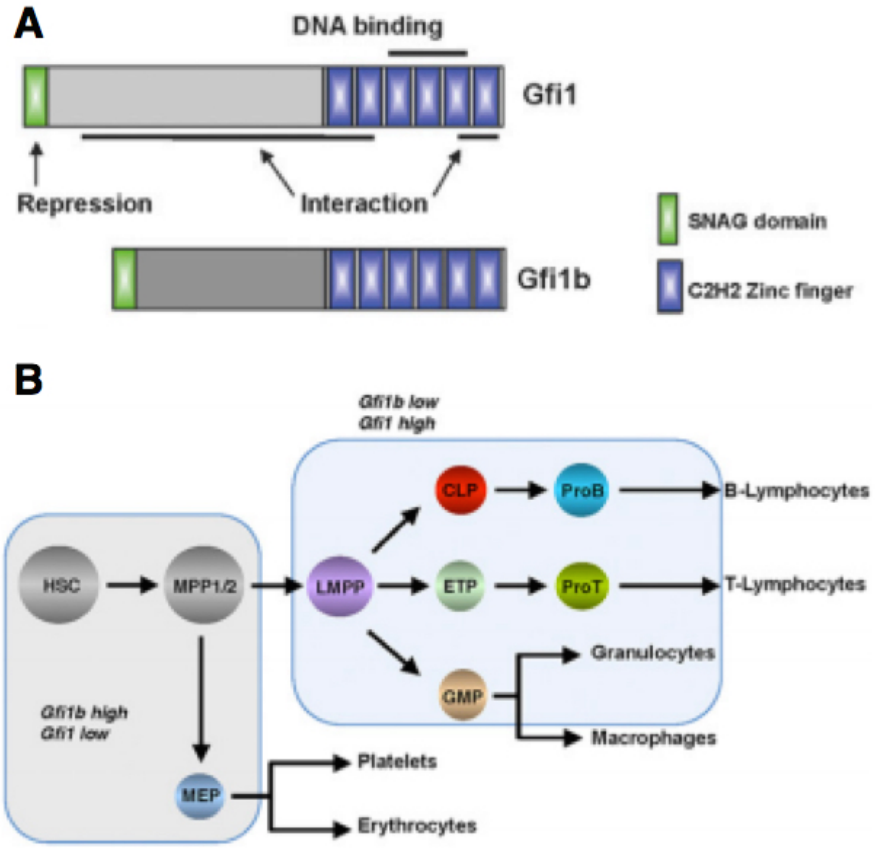


Figure 1.8: Schematic representation of Gfi1 and Gfi1b and the blood hierarchy.
A) Represents the Gfi1 and Gfi1b structure with their respective domains. B) Differentiation of HSPCs to differentiated cells. Shaded regions indicate the Gfi1 and Gfi1b expression levels. (Image from Möröy et al., 2015).

1.6 The zebrafish has three *Gfi1* paralogues

The zebrafish genome harbours three *gfi1* paralogues; *gfi1aa*, *gfi1ab* and *gfi1b*. Two of the zebrafish genes, *gfi1aa* (Wei et al., 2008) and *gfi1ab* (Dufourcq et al., 2004) are

orthologues of the mammalian *Gfi1a* gene. The two paralogues, *gfi1aa* and *gfi1ab* play a role in the inner ear, the lateral line system, the brain and in the parapineal organ (Dufourcq et al., 2004) and its role in definitive haematopoiesis is being studied in our lab (Roshana Thambyrajah, Deniz Ucanok unpublished data). A genomic sequence assembly study revealed that all three zebrafish genes consist of 6 coding exons where the first exon encodes the SNAG domain, which is conserved in fish, mouse and humans. Similarly the C-terminal zinc fingers are conserved as well. These domains are also similar in the three zebrafish paralogues (Cooney et al., 2013) (Figure 1.9A). Phylogenetic analysis showed *gfi1aa* and *gfi1ab* cluster most closely to the mammalian *Gfi1* transcription factors (Cooney et al., 2013) (Figure 1.9B) suggesting that they arose during teleost evolution through duplication of the original gene (John H Postlethwait et al., 2000; Woods et al., 2000). The third *gfi1* paralogue, *gfi1b*, is very similar to the mammalian *Gfi1b*.

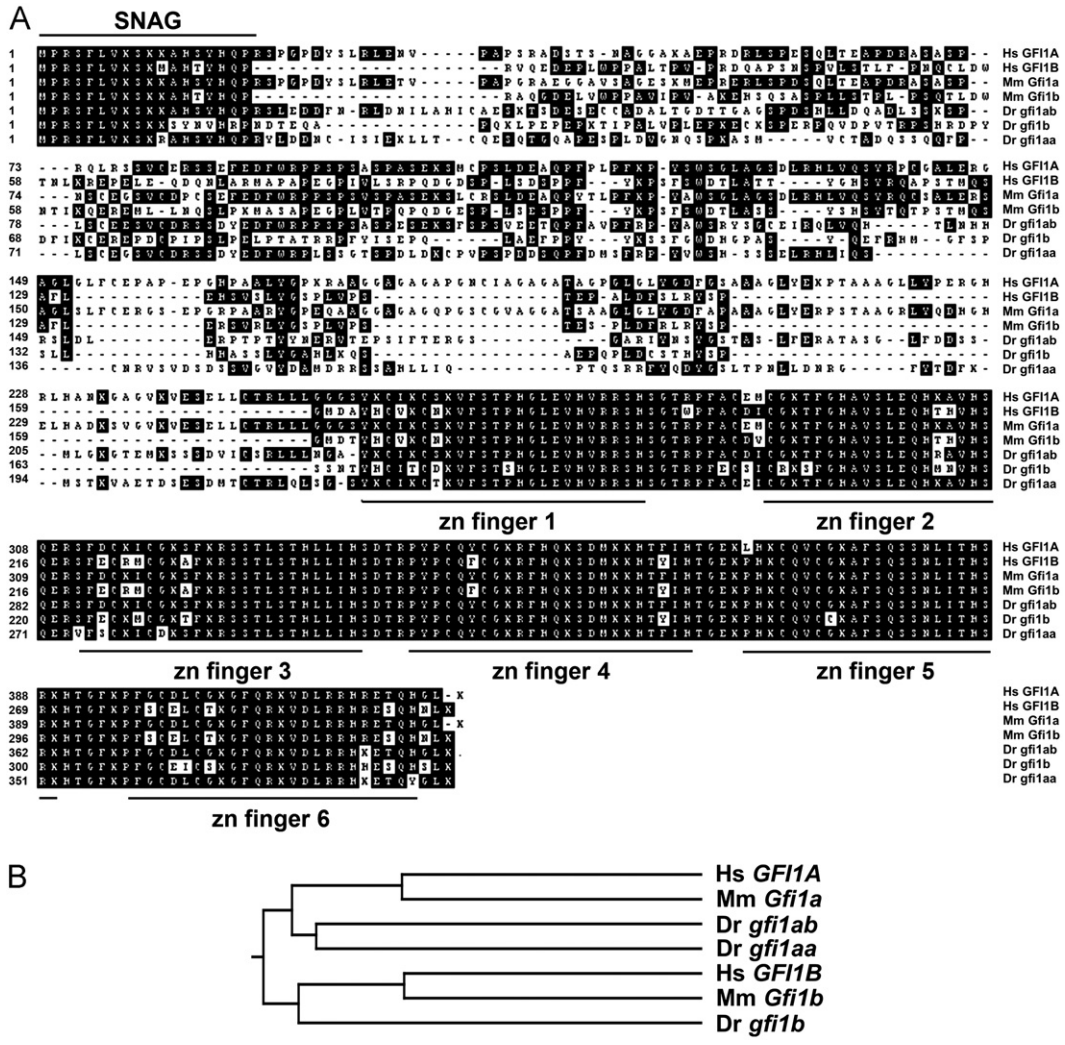


Figure 1.9: Evolutionary conservation of the *Gfi1* gene family between mammals and zebrafish.
(A) Conserved N terminal SNAG domain and C terminal 6 zinc fingers DNA binding domain amongst human mouse and fish. The middle linker domain is less conserved. (B) Phylogenetic relation of homologous Gfi factors in zebrafish, mouse and humans. Image from Cooney et al., 2013).

1.7 The *qmc551* gene trap line labels HECs of the vDA

In order to identify novel genes that are involved in vertebrate blood development during embryogenesis a transposon based gene trap approach (Kawakami, 2004) was used (Thambyrajah et al., 2016).

The recently published line *(gfi1aa:eGFP)*^{Gtqmc551} (from this point on this line will be referred to as *qmc551*) is a gene trap line that consists of an *eGFP* upstream of which is a

splice acceptor and downstream of the *eGFP* lies a polyadenylation signal. This gene trap was embedded within a Tol2 vector and injected (along with transposase mRNA) into one-cell stage wild type (wt) zebrafish embryos. Careful analysis using molecular biology techniques performed on genomic DNA (gDNA) from qmc551 embryos showed that the transposon is situated in the zebrafish *gfi1aa* gene on chromosome 2 (Thambyrajah et al., 2016). In this line, eGFP expression is observed in prRBCs as well as in cells that are associated with the vDA (Figure 1.10). *gfi1aa* expression was seen in non-haematopoietic tissue as well such as in cells of the inner ear and lateral line organ (Thambyrajah et al., 2016). In order for an EC to be part of the subset of HECs it must, a) co-express the transcription factors *runx1* and *c-myb,* (b) undergo EHT. The qmc551:eGFP cells were shown to do both of the above and additionally were also observed at the transient site of haematopoiesis, i.e. the larval CHT. Subsequently the eGFP+ cells were seen in the larval thymus and kidney. Fluorescent-activated cell sorting on the adult fish WKM revealed the presence of eGFP+ cells which mainly occupied the progenitor and lymphoid gates (Thambyrajah et al., 2016).

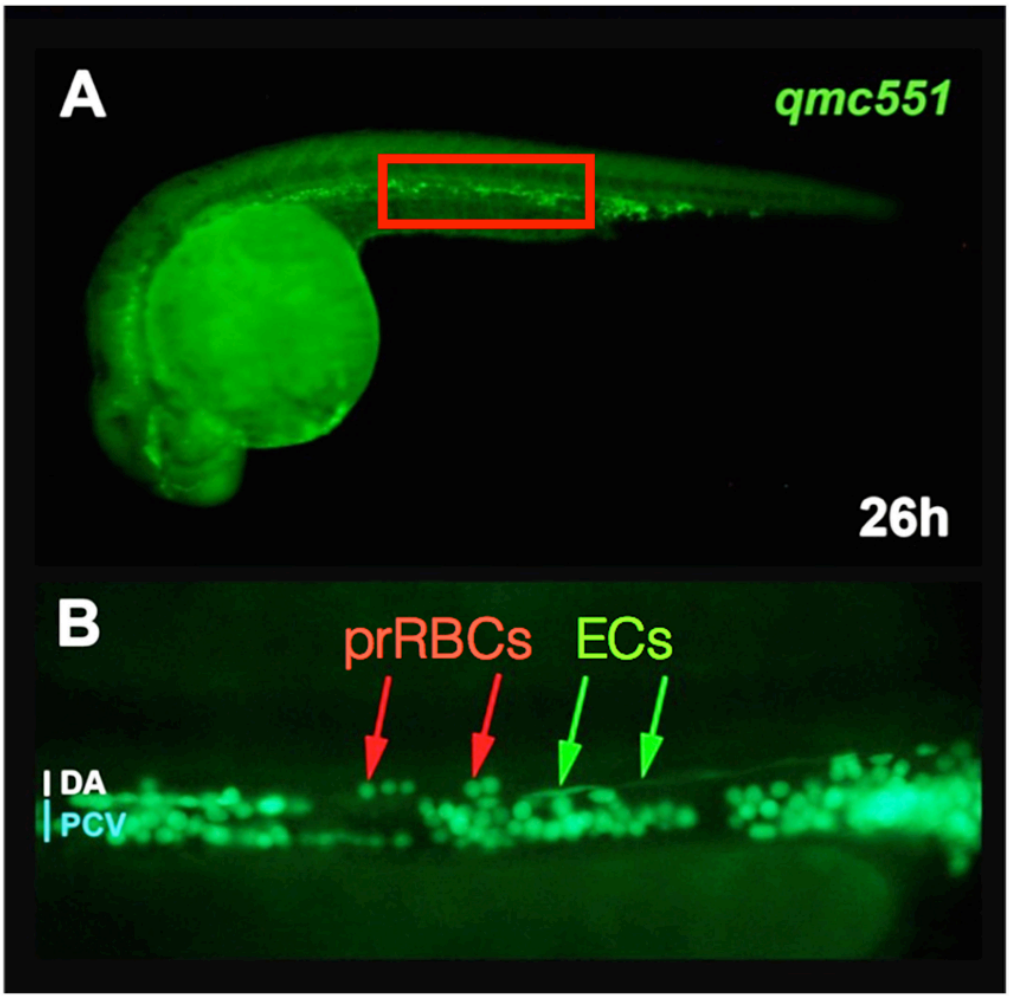


Figure 1.10: The *(gfi1aa:eGFP)*^{Gtqmc551} labels prRBCs and ECs of the vDA.
A) Highlights (red boxed region) the trunk region of a 26 hpf qmc551 embryo in which (B) close up of the vDA shows eGFP expressing spindle shaped endothelial cells (ECs). Embryo is in lateral view. (Image modified from Thambyrajah et al., 2016).

1.8 *Gfi1ab* mRNA is upregulated in the vDA in the absence of *gfi1aa*

In the *qmc551* homozygous embryos it was observed that *gfi1aa* mRNA was completely abrogated in comparison to *wt* and *qmc551* heterozygous embryos (Figure1.11Ai-Aiii). However, *gfi1ab* transcript was upregulated in cells of the vDA and slightly in the inner ear in homozygous *qmc551* embryos at 27 to 28hpf in comparison to wt and *qmc551* heterozygous embryos (Figure 1.11Bi-iii). A closer analysis of a transverse section through the trunk of a WISH s *qmc551* homozygous embryo stained for revealed the expression is

in cells of the vDA (Figure 1.11C), indicating the role of *gfi1ab* in definitive haematopoiesis (Ucanok, 2015).

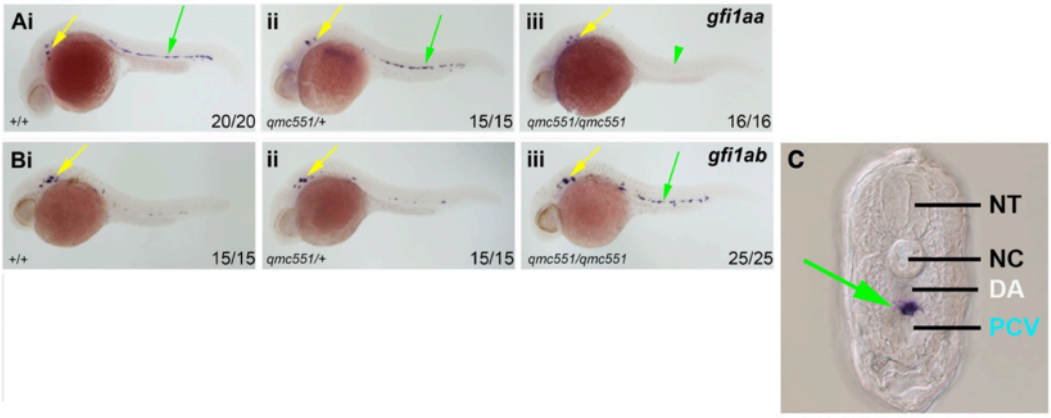


Figure 1.11: Upregulation of *gfi1ab* mRNA in cells of the vDA.
Ai-iii) WISH experiments carried out on *wt*, *qmc551* heterozygous and homozygous embryos showing a loss of *gfi1aa* mRNA in *qmc551* homozygous (Bi-iii) WISH experiments to analyse *gfi1ab* mRNA wt, *qmc551* heterozygous and homozygous embryos showing an up regulation of *gfi1ab* mRNA in *qmc551* homozygous (C) a transverse section through the trunk of a *qmc551* homozygous embryo (Biii) shows the expression is indeed in the vDA. Green arrows: HECs, yellow arrows: inner ear; green arrowhead: loss of *gfi1aa* mRNA in the vDA. (Image from Ucanok, 2015).

1.9 Lineage tracing

Fate mapping has become a major part of stem cell biology studies and helps biologists gain insights into properties of tissues. This process provides a means of studying the cellular origin of cells. Strategies of lineage tracing could involve both short-term as well as long-term tracing to identify the progeny of single cells (Kretzschmar and Watt, 2012).

1.9.1 Migratory route of HSPCs during definitive haematopoiesis as revealed by short-term lineage tracing experiments

Short-term lineage tracing experiments using a photoactivatable cell tracer dye on the developing zebrafish embryo showed that cells located in the AGM region migrate to the CHT, where they expand and differentiate before migrating to eventually seed the larval kidney and thymus (Jin et al., 2007; Murayama et al., 2006; Traver et al., 2003) (Figure 1.12I). The thymus is colonised by T cells (Lieschke, 2001; Lin, 2005; Willett et al., 1999) and seems to be the seeding and maturation site for lymphocytes (Menke et al., 2011). The HSCs from the CHT seed the kidney from 4 dpf (Ellett and Lieschke, 2010) (Figure 1.12II). Long-term lineage tracing cells from the developing embryo into adulthood would be required in order to show that cells from the embryo maintain the blood system in the adult organism.

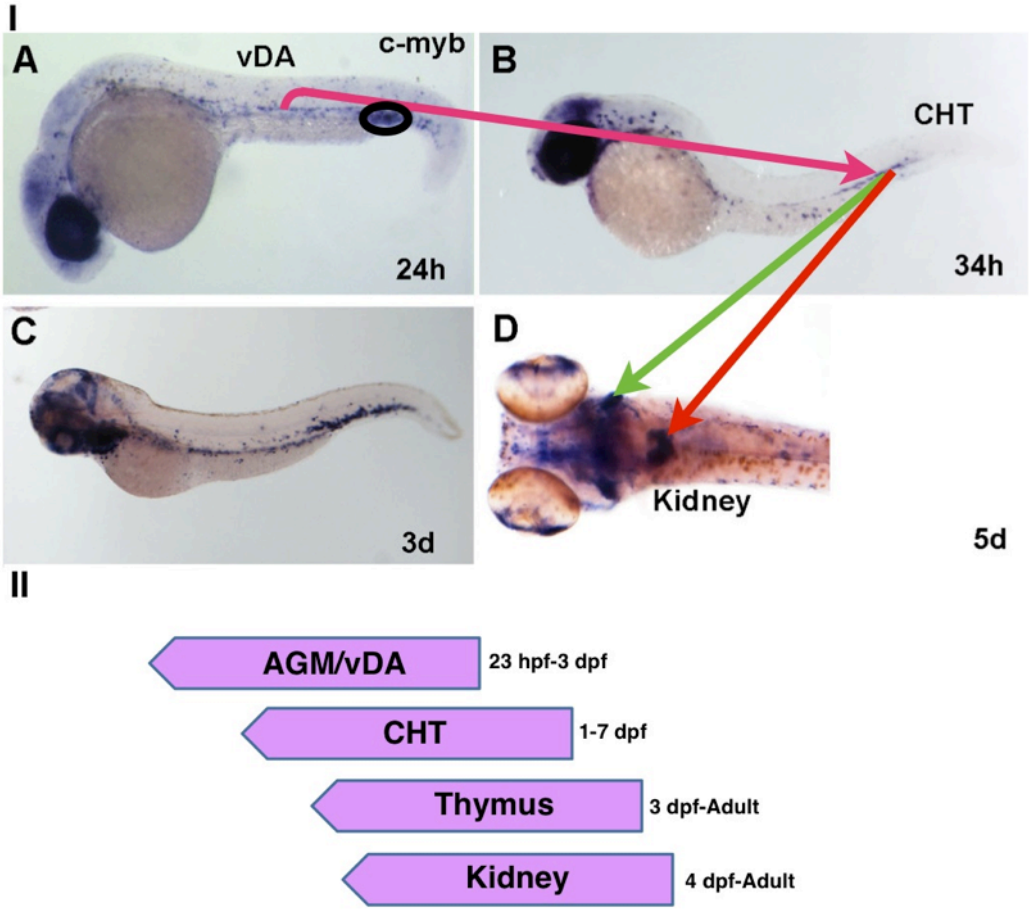


Figure 1.12: Migratory route of HSCs/Hematopoietic progenitor cells (HSPCs) during definitive haematopoiesis.
(I) *c-myb* expression pattern using WISH illustrates the migratory route of HSCs/HSPCs during zebrafish embryogenesis that have been proposed by lineage tracing experiments using photoactivatable dyes (Jin et al., 2007; Murayama et al., 2006). RNA in situ hybridisation was performed on 2hpf, 34 hpf and 5 dpf old embryos. The expression of *c-myb* in cells of the vDA that migrate to subsequent hematopoietic tissue. *c-myb* is also seen to be expressed in the branchial arches) (A) vDA at 24 hpf, also observed are a population of *c-myb* expressing cells which are most likely to be EMPs and prRBCs in the PBI (circled black) (B) the *c-myb* positive cells migrate from the vDA to the CHT (pink arrow; in situ experiments carried out on 34 hpf embryo) (C) In a 3 dpf embryo, cells in the CHT multiply and migrate to (D) the thymus (green arrow) and the kidney (red

arrow; in situ carried out on a 5 dpf embryo). II) Schematic representation of the timeline of migration of putative HSCs from the vDA to the kidney. (Figure I *c-myb in situ* images supplied by Dr. Martin Gering and was modified, Figure II from Jalali, 2012).

1.9.2 The Cre/LoxP system can be used to long-term lineage trace cells

A more elegant method to lineage trace cells from embryogenesis into adulthood is the Cre/LoxP recombination system (Nagy, 2000). The Cre/LoxP recombination system has been used for genetic recombination and efficient in carrying out Cre mediated recombination both in mouse (Georgiades et al., 2002; Zovein et al., 2008) and in zebrafish (Hans et al., 2009) where reporter gene expression was observed in a tissue specific manner.

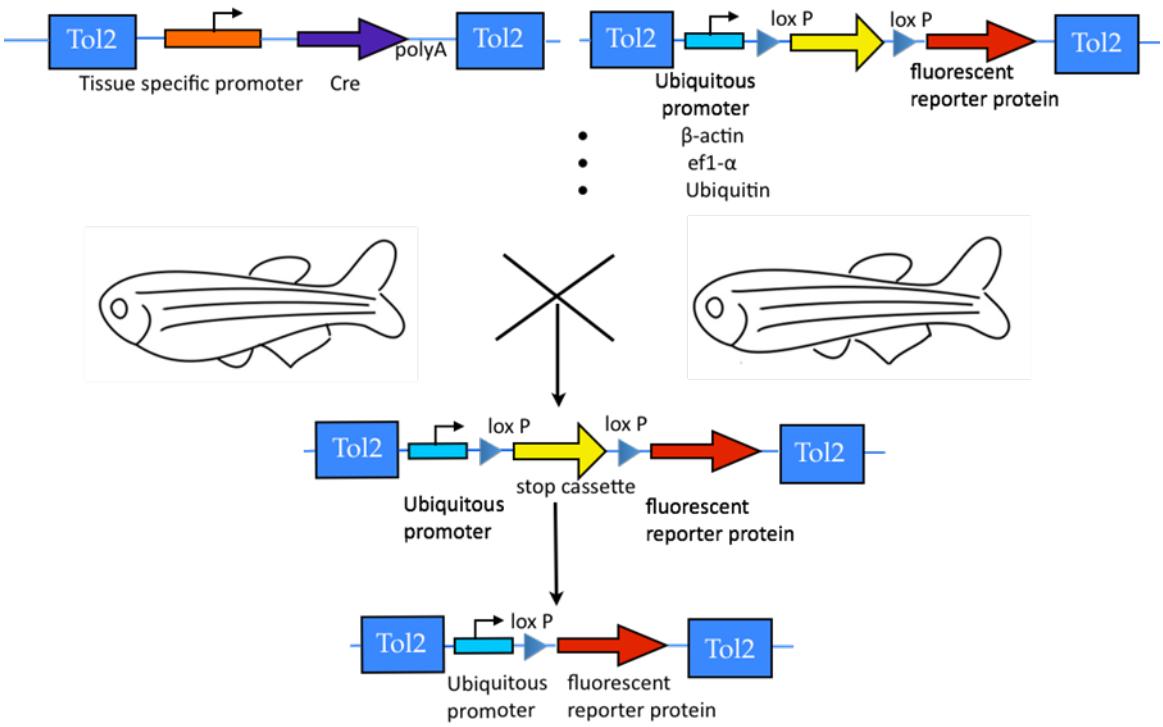


Figure 1.13: Schematic representation of the Cre/LoxP recombination system in general as seen in fish.

Making use of two transgenic fish lines; where in one line the expression of *Cre* is driven by a tissue specific promoter, depending on the tissue of interest (e.g.: *flk1/kdrl* promoter to drive *Cre* expression in endothelial cells). The second line is the reporter that carries a floxed gene usually a reporter or stop cassette; downstream of which lies a second reporter gene. The reporters are expressed under the control of a ubiquitous promoter. On crossing the two fish, embryos that carry both the Cre and the reporter transgenes would undergo recombination in cells where Cre is expressed. The floxed region is excised and the second reporter is expressed.

This system makes use of two transgenic lines, one that carries the Cre recombinase which is driven by a tissue specific promoter that would drive Cre expression in the tissue of interest, for example to drive Cre expression in endothelial cells, the endothelial cell (EC) specific promoter could be used. The second transgenic line is the reporter line (which will be referred to as the Cre reporter line throughout the text), which has a fluorescent reporter protein downstream of a ubiquitous promoter, e.g. the *β-actin* promoter (Figure 1.13). When fish carrying both transgenes are crossed with each other, the recombinase expression which is driven by the *flk1/kdrl* promoter (blood and in endothelial cells). Once expressed, Cre enters the nucleus and catalyses the deletion of the floxed region (between the two LoxP sites) to excise the "super stop cassette" or a reporter gene in *flk1/kdrl*+ cells, thus to express a reporter protein (Figure 1.13). Because the expression of the reporter protein is under the control of the ubiquitous all cells derived from *flk1/kdrl*+ ECs will be expressing the fluorescent protein from this point on.

Using the zebrafish *flk1/kdrl* promoter, usually active in endothelial cells (Choi et al., 2007; Jin, 2005), to drive Cre expression Bertrand et al (2010) showed that adult HSCs are derived from *flk1/kdrl*-positive cells, suggesting an endothelial origin of HSCs (Bertrand et al., 2010; Kissa and Herbomel, 2010). Because this study used a constitutively active Cre it is impossible to tell when the labelled HSCs arose.

In order to confirm an early embryonic origin of HSCs, a temporal control over the Cre activity is needed. For this purpose one can make use of a $CreER^{T2}$ fusion protein whose activity is inducible with tamoxifen (TAM) or its active metabolite 4-hydroxy tamoxifen (4-OHT) and, therefore, offers temporal control. Experiments using the inducible $CreER^{T2}$ system have been carried out in the mouse which have suggested that adult blood cells are derived from endothelial cells (Zovein et al., 2008). These experiments showed that haemogenic endothelial cells in the embryo, contribute to adult haematopoiesis, (Zovein et

al., 2008) however, experiments done in mouse embryos do not allow visualization of recombined cells in the early embryo. For this, these experiments can be carried out in the zebrafish to get a live overview of the recombination events.

1.10 Aims and Objectives

- Using the *flk1/kdrl* promoter usually active in ECs, to drive constitutive Cre expression, here, I aim to confirm an endothelial origin of early larval and adult haematopoietic cells.
- Once an endothelial origin of larval and adult blood cells is established, next, to determine an early embryonic origin of larval and adult blood cells, Cre expression was put under temporal control. For this purpose, a $CreER^{T2}$ fusion protein whose activity is inducible with 4-hydroxy tamoxifen (4-OHT) was used to restrict the activity to a particular time window during embryogenesis.
- As part of the lineage tracing studies, we were interested in studying cell behaviour (EHT). Using the *qmc551* line, here, I aim to show using confocal microscopy that the qmc551:eGFP cells of the vDA undergo EHT.
- *gfi1aa* expression is lost in *qmc551* homozygous embryos, but *gfi1ab* paralogue was up-regulated. Here I aim to study the effect of losing all four copies of the *gfi1a* paralogues using the *qmc551* line and three *gfi1ab* mutant lines.

2 Materials and Methods

2.1 Materials

2.1.1 Zebrafish husbandry and maintenance of lines

All zebrafish (Danio rerio) adults and embryos whether wild-type, transgenic or mutant were maintained, bred and raised at 28.5°C as per guidelines described in The Zebrafish Book (Westerfield, 2000). Fish were kept at 14 hour light/10 hour dark cycles in system water. Zebrafish are photoperiodic in nature and spawn in the mornings. Embryos are collected and kept in E3 medium containing methylene blue to avoid fungal growth. All experiments on the fish were carried in accordance to Home Office (HO) regulations and were approved by the local ethical review board under project license numbers; 40/2457, 30/3378 and 30/3356 and personal e-license number; I464EB52C.

2.1.2 Transgenic lines used in this study

Tg(flk1/kdrl:Cre)^{qmc101} *(flk1/kdrl:Cre)*, *Tg(ef1α-LoxP-eGFP-terminator-LoxP-dsRed) ef1α-Cre* reporter, *gfi1aa*^{qmc551Gt}, *Tg(kdrl/kdrl:tdTomato)*^{qmc64} *(kdrl:tdTom)*, *Tg(csl:cerulean)*^{qmc63} *(csl:cer)* an in-house generated notch signalling reporter line, which is a derivative of the *csl:venus* (Gray et al., 2013). The *csl:cerulean* transgene carries 12 Rbpj (csl) binding sites in front of a minimal rabbit β-globin promoter and *Tg(fli1a:nls-mCherry)*^{sh370} *(fli1a:nls-mCherry)* was generated in Robert Wilkinson's lab in Sheffield. The integration site of the three reporter transgenes is not known. The *Tg(kdrl:eGFP)*^{s843} *(kdrl:eGFP)* is a previously published line (Jin, 2005)

Table 2.1: Buffers and Solutions used

Buffers/ Solutions	Ingredients	Store at
Molecular biology experiments		
0.5M EDTA pH8	186.1g EDTA, NaOH (to adjust pH), distilled water	RT
Embryo lysis buffer (ELB) 200ml stock solution (no ProK)	10ml 1M Tris-HCl pH8.5, 0.4ml 0.5M EDTA pH8.0, 10 ml 10% Tween 20 (0.5%) in 200ml sterile distilled water	RT
Embryo lysis buffer (ELB) working solution ProK (10mg/ml)	ELB stock solution, 300µg/ml Proteinase K to be added fresh only at the time of use	Make up fresh
gDNA extraction buffer	1M Tris-HCl pH8.0, 0.5M EDTA pH8.0, 4M NaCl, 0.5% SDS, 200µg /ml Proteinase K	To be made fresh (Do not store)
5X base buffer	0.5g NaOH, 2ml 0.5M EDTA (pH8) bring up the volume to 100ml with sterile distilled water	RT
5X Neutralisation Solution	3.15g 1M Tris-HCl (pH8) in 100ml of Sterile distilled water	RT
50X TAE buffer	Tris Base, glacial acetic acid, 0.5M EDTA (pH8),	RT

(stock)	distilled water.	
Live embryo treatment for imaging		
0.03% PTU	0.3g U-Phenylthiourea (PTU) (Sigma, P7629) in 1L distilled water and filter sterilized.	4° C
MS222 4g/L ph7-7.5	4g Ethyl 3-aminobenzoate (Tricaine) methanesulfonate (MS222) (Sigma, A5040-100G) adjust pH with Tris-HCl pH9.5	4° C
1% Low melting point (LMP) agarose	0.5g LMP (Thermo Scientific, RO801) in 50ml E3 buffer.	RT
25mM 4-OHT stock solution	5mg 4-OHT (Sigma, H7904-5MG) dissolved in 516µl 100% EtOH to make up a stock solution. Dispense into 50µl aliquots, ideally in amber tubes to protect from light.	-20° C
WISH experiments		
4% paraformaldehyde (PFA)	Paraformaldehyde (Sigma, 15812-7), in PBS, 1M NaOH,	4° C
1X Phosphate buffered saline (PBS)	10X PBS (GIBCO, 14190250), sterile distilled water	RT
1X Phosphate	1X PBS, 0.1% Tween20	RT

buffered saline Tween20 (PBSTw)		
10mg/ml Proteinase K stock solution	Proteinase K from Tritirachium album (Sigma, P2308 100MG, 30 units/mg) in glycerol storage buffer.	-20° C
Proteinase K working solution	For In situ experiments use 0.5µg/ml of PBSTw, i.e. 1µl of stock (10mg/ml Proteinase K) in 1ml of PBSTw.	Make fresh
20X Saline Sodium Citrate (SSC) pH7	0.3M NaCl, 300mM sodium citrate	RT
2X SSC Tween20 (SSCTw)	20X SSC, 20% Tween20, sterile distilled water	RT or -20° C
0.2X SSCTw	2X SSCTw, 20% Tween20, sterile distilled water	RT or -20° C
Hybe$^{+/+}$ Buffer (pH6.0)	50% Formamide (Sigma, F9037), 5X SSC (pH7), tRNA (50mg/ml) (torula yeast, Type VI, Sigma, R6625) 50mg/ml Heparin, 9.2mM Citric acid (1M), 0.1% Tween20, sterile distilled water.	-20° C
Hybe$^{-/-}$ buffer (pH6.0)	50% Formamide (Sigma, F9037), 5X SSC (pH7), 9.2mM Citric acid, 0.1% Tween20, sterile distilled water.	-20° C
Maleic Acid Buffer	0.1M Maleic Acid, 0.15M NaCl, sterile distilled water,	4° C

(MAB) pH 7.5	adjust pH to 7.5 using NaOH pellets, autoclave	
MABTw	50 ml MAB Buffer, 250µl Tween 20 (0.1%) added when needed	RT
MAB-Block	2% Roche-Boehringer Blocking reagent TM in MAB	-20° C
BCL3 buffer (developing buffer)	1M Tris-HCl pH9.5, 5M NaCl, 1M MgCl2.6H2O, 0.1% Tween20, sterile distilled water.	RT (up to 2 weeks only)
PBSTw 20mM EDTA	PBSTw, 0.5M EDTA pH8	RT
Blood cell suspension preparation		
1X HBSS buffer	10XHBSS (GIBCO, 4065049), sterile distilled water	RT
FCS/PBS	10X PBS, FCS, sterile distilled water	-20° C

2.2 Methods

2.2.1 Treatments done on embryos and adult fish to activate transgene expression

2.2.1.1 Induction of transgenes by 4-OHT

In embryos carrying a CreERT2 expressing transgene, CreERT2 release to the nucleus was induced using 4-OHT (Sigma, H7904-5MG) (Hans et al., 2010). 25mM 4-OHT stock solution (Table 2.1) was diluted 1:10 with 100% EtOH to make up a 2.5mM working solution. 24ml of E3 medium was poured into a fish plate into which embryos were added

in a volume of ~1ml to make up the total volume in the plate to 25ml. Embryos were then treated from tail bud stage usually with 5µl (0.5µM) or 50µl (5µM) of 2.5mM 4-OHT. Control embryos were treated with the same volume of ethanol. Since 4-OHT is light sensitive, incubations were carried out in the dark at 28.5°C.

2.2.1.2 Heat shock treatment

Embryos and adult fish that carried a transgene with an hsp70 promoter were heat shocked in order to activate gene expression (Hans et al., 2010). Embryos were first raised at 28.5°C and were heat shocked on 1dpf by putting embryos into a 50ml falcon tube that contained 10ml of pre-warmed (to 37°C) E3 medium. The tube containing the embryos was incubated for 30 minutes in a water bath set to 37°C. After 30 minutes, embryos were put back into a fish plate containing E3 buffer and were incubated at 28.5°C. Generally reporter gene expression was observed with in 2 hours of the heat shock treatment. Adults were heat-shocked by pre-warming system water or E3 medium in breeding tanks to 37°C. The fish were placed in the pre-warmed water and tanks were incubated at 37°C for 8 hours with constant monitoring of the fish and the temperature of the water the fish were in.

2.2.2 Whole-mount in situ hybridization (WISH) of zebrafish embryos

Alkaline phosphatase whole mount in situ hybridization (WISH) experiments were carried out based on a previously published protocol (Broadbent and Read, 1999). WISH is a technique used to detect mRNA and thereby study the expression patterns of genes during development. In situ hybridization involves the use of an ‘antisense’ RNA probe complementary to the endogenous mRNA of the gene of interest. The antisense probe is labelled with a specific antigenic label, digoxygenin. The embryo is incubated with the labelled probe. Hybridization of the probe to the mRNA is detected using an anti-

digoxygenin antibody conjugated to alkaline phosphatase. The presence of the antibody is visualized with commercially available chromogenic substrate for alkaline phosphatase like BM purple. In order to carry out WISH experiments several steps are involved which are described in subsequent subsections.

2.2.2.1 Collection, fixation and storage of embryos for WISH

Embryos were collected at the desired time points, staging of embryos was done as described by Kimmel et al (Kimmel et al., 1995). Embryos older than 18hpf were dechorionated manually and put into 1.5ml eppendorf tubes (not more than 25 embryos per tube) before fixing with 4%PFA overnight at 4°C whereas embryos younger that 18hpf were fixed in their chorions. If embryos older than 24hpf were required for WISH experiments, they were treated with 0.003% PTU from 24hpf until they were required to be fixed, in order to ensure that the embryos did not develop pigmentation as this could obstruct the visualization of the gene expression pattern after staining with the chromogen. After overnight fixation, the 4% PFA solution was taken off and embryos were washed 3 times in PBSTw, embryos younger than 18hpf were dechorionated at this stage. Following washes with PBSTw all embryos were gradually dehydrated by taking them through 5 minute washes each with 25%, 50%, 75% (all in PBSTw) and 100% methanol. The 100% methanol was replaced and embryos were stored and preserved at -20°C at least overnight until required.

All pipetting steps were done using 500µl of solutions.

2.2.2.2 In Situ Hybridisation on zebrafish embryos

In situ experiments on zebrafish embryos was done over three days although in some cases an overnight staining was required.

DAY ONE

Embryos that were stored in methanol at -20°C were selected for an in situ and were rehydrated slowly into PBSTw, first by taking them through 5 minute washes each with 75%, 50% and 25% methanol (in PBSTw) and finally washing thrice for 5 minutes with PBSTw. Embryos from 26hpf were permeabilised by ProteinaseK digestion. A freshly prepared ProteinaseK/1ml of PBSTw solution was used. Incubation times varied according to the age of the embryos (Table 2.2).

Table 2.2: Incubation times in Proteinase K of embryos at different developmental stages

Embryo stage	Incubation times in Proteinase K
27-30 hpf	4 minutes
2dpf	17 minutes
3dpf	33 minutes
5-6 dpf	1 hour

Following permeabilisation with ProteinaseK, embryos were carefully taken through two 5-minute washes with PBSTw after which they were refixed in 4%PFA for 20 minutes. After 20 minutes the 4%PFA was discarded and embryos were then rinsed 5 times for 5mins in PBSTw at RT on a gently rocking table (Labnet GyroTwister or Grant-bio, PS-M3D). The embryos were then prepared for prehybridisation by washing once for 5

minutes with a solution of 50% PBSTw 50% Hybe+/+ solution (Table 2.1). Prehybridisation was carried out by incubating embryos in Hybe+/+ for at least 1 hour at 65°C in a heat block (Labnet International, Inc, AccuBlock ™ Digital Dry Bath) that was laid on its side. The Hybe+/+ was replaced with hybridization mix i.e. Hybe+/+ containing the probe, typically a 1:200 dilution of the probe. The tubes were incubated at 65°C overnight in a heat block laid on its side to ensure that embryos would lie along the side of the tube rather than be clumped at the base of the tube.

DAY TWO

The next day the probe was recovered and embryos were taken through a series of washes that were carried out leaving the tubes in the heat block at 65°C and with occasional agitation. Each wash was carried out for 10 minutes using different concentrations of $Hybe^{-/-}$ (Table 2.1) in 2XSSCTw, viz; 100%, 75%, 50% and 25% followed by a single 10-minute wash with 2XSSCTw and four 15 minute washes with 0.2X SSC. Next, embryos were equilibrated in a series of 5-minute washes with 25%, 50%, 75% and 100% MABTw in 0.2X SSCTw (Table 2.1) carried out at RT on a gently rocking table. After equilibration embryos were blocked with MAB-Block (Table 2.1) for at least 1hour. The MAB-Block was then replaced with anti-DIG-antibody Anti-Digoxygenin AP-Fab fragments, Roche, 11093274910), which was diluted 1:5000 in MAB-Block. Tubes were incubated at 4°C overnight on a rocking table.

2.2.2.3 Probe generation

A number of in-house and previously published WISH probes were used. In-house probes were prepared by linearizing the plasmid using appropriate restriction enzymes that cut upstream of the open reading frame of the gene of interest followed by a clean up step of the linearized DNA. Next, an invitro transcription was carried out to make the antisense

RNA probe using RNA polymerase (T3, T7 or SP6) that was DIG labelled (Roche, 11277073910). The RNA polymerase converts the DNA sequence of interest into a complementary RNA sequence. The following is a typical in vitro transcription reaction set up;

Components	Volumes
Linearized template DNA (1µg)	5µl
Sterile distilled water	7.5µl
DIG-RNA mix	2µl
5X transcription buffer	4µl
RNAase inhibitor	0.5µl
RNA polymerase (T3, T7 or SP6)	1µl
Total volume	20µl

The above reaction was incubated at 37°C for 2hours after which the DNA template was removed from the reaction by digestion with DNaseI. The incubation with DNaseI was carried out at 37°C for 15 minutes. The reaction was stopped with addition of 0.5M EDTA pH8. The RNA probe was then precipitated with 10M-ammonium acetate followed by washing of the RNA pellet with 70% ethanol and then drying the pellet finally the RNA was resuspended in RNase free water.

Table 2.3: List of WISH "antisense" Probe

Gene of interest used to make the Probe	**Plasmid (p) number from Gering Lab database, Restriction enzyme used to linearize Plasmid DNA**	**RNA polymerase used for in vitro transcription**	**Reference**
CreERT2	p301, ApaI (NEB, R0114S)	SP6	In-house
tdTomato	p230, ApaI	T7	In-house
gfi1aa	p287, SphI (NEB, R0182S)	SP6	In-house (Thambyrajah, 2011)
c-myb	p77, EcoRI (NEB, R0101S	T7	(Thompson, 1998)
rag-1	p21, HindIII (NEB,	T7	In-house
eGFP	p308, SacII (NEB,	T3	In-house (Al Khamees, 2015)

DAY THREE

Embryos were removed from 4°C and tubes were left on a gently shaking table to attain RT. Next, the antibody dilution was removed and embryos were taken through 8 X 15minute washes with MABTw, which were all done at RT on a gently rocking table. Embryos were then equilibrated 3 times by washing with developing buffer i.e. BCL3 buffer (Table 2.1) for 5minutes per wash. After the equilibration embryos were then stained with BM Purple (Roche, 11442074001). The staining solution was prepared as a 1:1 mix in BCL3 buffer. The colour was allowed to develop at RT on a gently rocking table, protected from light. Colour development times varied from an hour to staining overnight, depending on the strength of the signal. Once the embryos had stained sufficiently, the staining solution was removed and embryos were washed thrice for 5 minutes per wash with PBSTw 20mM EDTA (Table 2.1), which largely stopped the reaction. The embryos were then fixed in 4% PFA overnight at 4°C. The following day embryos were washed 3 times for 5 minutes each in PBSTw. The embryos were then slowly washed into 80% glycerol, by washing once for 5 minutes each with 30%, 50% and finally into 80% glycerol (all glycerol solutions were made in PBSTw). Embryos were stored in 80% glycerol at 4°C. Embryos were then documented and were genotyped if necessary.

2.2.3 Collection of fin clip biopsies from adult zebrafish

In order to collect fin clip biopsies, a tank containing 200ml of facility water to which approximately 6ml of 4g/L MS222 (Table 2.1) was added. The fish to be fin clipped was put into the tank in order to anesthetize it. The fish was observed for a while until it turned, belly up, indicating it was anesthetized. The fish was then carefully taken out of the tank and a fin clip biopsy was obtained using a pre-sterilized pair of scissors. The fin clip was

carefully put into a 1.5 ml eppendorf tube and the tube was left on ice for further processing, i.e. lysis buffer or ELB (Table 2.1) was added to the samples in order to obtain genomic DNA.

2.2.4 Molecular biology techniques

2.2.4.1 Genomic DNA extraction

Genomic DNA was extracted in order to genotype as per requirement either from 24-120h (I did 50h) single embryos or batches of 5 embryos, fin clip biopsies and blood cells flushed out of the adult fish whole kidney marrow. Prior to carrying out a gDNA extraction the tissue sample or embryo was first lysed followed by a gDNA extraction either by a rapid method or purified using a phenol/chloroform extraction.

2.2.4.1.1 Lysis of embryos, blood cells or fin clip biopsies

In order to extract gDNA from embryos, blood cells or fin clip biopsies it is important to digest the tissue and break open the cell membrane, for this a lysis buffer, either ELB or gDNA extraction buffer (See Table 2.1) was used, followed by an overnight incubation at 55°C. The next morning proceed to gDNA extraction by either of the following two methods (2.2.4.1.2 and 2.2.4.1.3).

2.2.4.1.2 gDNA extraction by rapid method

Following the overnight incubation in lysis buffer (2.3.1.1) the next morning tubes were vortexed briefly and centrifuged in order to bring the contents down to the bottom of the tube. Since the lysis buffer contained ProteinaseK, the ProteinaseK was inactivated by heating the tubes to 100°C for 20minutes. After 20 minutes the tubes were cooled and sterile distilled water was added to each tube followed by mixing and briefly centrifuging. The proteins and other impurities settle to the bottom of the tube whereas, although not

absolutely pure, gDNA remains suspended in the supernatant. 1µl of the gDNA containing supernatant (without disturbing the pellet) was used in a 25µl PCR reaction for genotyping done in Chapters 5 and 6.

2.2.4.1.3 gDNA extraction and purification

The rapid method for gDNA extraction described above (2.3.1.2) does not yield gDNA of prime quality and potentially contains traces of proteins, salts or other impurities, this sometimes might interfere with the genotyping. Hence, in order to obtain gDNA that is completely free of other components, a phenol chloroform extraction was carried out on samples (this method for gDNA extraction was used mainly for experiments genotyping experiments done in Chapter 3). The principle behind this organic extraction method is based on the fact that the negatively charged phosphate backbone of DNA remains soluble in an aqueous phase whereas all other components settle into the phenol layer. Briefly, following the overnight incubation of the samples at 55°C in lysis buffer, the next morning proceed to carrying out a phenol/chloroform extraction. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma, P2069) was added to each sample tubes were then vortexed (Grant bio PV-1) for 1minute and centrifuged (Fisher Scientific accuspin™ Micro or Thermo Scientific Heraeus PICO17 Centrifuge) at 13000 rpm for 5minutes. After 5 minutes, the aqueous layer (containing nucleic acids) was transferred to a new set of appropriately labelled tubes to which an equal volume of chloroform:isoamyl alcohol (24:1) (Sigma, 25666) was added and the tubes were vortexed for 1 minute and centrifuged at 13000 rpm for 5 minutes to separate the aqueous layer containing nucleic acids from the phenol layer that contained proteins and other impurities. The top aqueous layer was then carefully pipetted out and transferred to a new set of tubes and the lower organic layer was appropriately discarded following laboratory safety measures. The aqueous layer contained nucleic acids and was used to further purify DNA. DNA was

precipitated from the aqueous layer by adding 1/10th volume of sodium acetate (3M, pH5.2) plus 2.5 times volume of 100% ethanol. The tubes were left at -80°C for 30 minutes. This precipitation step in the presence of high salt reduces the amount of RNA in the DNA. After 30 minutes gently mix the tubes by inverting 5-8 times followed by centrifugation at 13000 rpm for 15 minutes. The supernatant was discarded, making sure not to disturb the pellet. The pellet was then rinsed with 70% ethanol to get rid of salts. If the pellet was disturbed, the tubes were centrifuged briefly. The 70% ethanol was discarded carefully and the pellet was left to dry completely as any reminiscence of ethanol would interfere with the enzyme Taq polymerase in the PCR. Once the pellet was completely dry, it was reconstituted in TE buffer. TE contains EDTA that is a chelating agent that gets rid of nucleases.

2.2.4.2 Genotyping after in situ hybridization

After an in situ experiment embryos are washed and stored in 80% glycerol at 4°C. Embryos should be left in 80% glycerol at least for one night before imaging. The next day, embryos were analysed, counted and imaged to document the phenotype observed. Once documented, embryos to be genotyped i.e. embryos with a phenotype (potential mutants) and without a phenotype (wt or heterozygous) were selected and individual embryos were transferred into 1.5ml eppendorf tubes. On sorting single embryos into eppendorf tubes, the embryos were then gradually washed out of 80% glycerol/PBSTw (2.2.4.2.1).

2.2.4.2.1 Washing WISH embryos out of 80% glycerol before gDNA extraction

To wash embryos out of the 80% glycerol, the embryos were taken through a series of washes, which included 5 minutes in 50% glycerol (in PBSTw) followed by 30% glycerol (in PBSTw) and finally three 5-minute washes in 1XPBSTw. The PBSTw was discarded

after the final wash and embryos were then incubated in 300mM NaCl at 65°C for 4 hours to reverse crosslinking (Gansner et al., 2008). After 4 hours of incubation in NaCl proceed towards gDNA extraction (2.2.4.2.2).

2.2.4.2.2 gDNA extraction from pre-treated WISH embryos (2.2.4.2.1)

Following the incubation in NaCl at 65°C for 4 hours, the NaCl was discarded and 1X Base buffer was added to each embryo which were then incubated at 95°C for 20 minutes. Next, the tubes were cooled and without discarding the solution neutralize with 1X neutralisation solution. The tubes were then vortexed and centrifuged at 3000 rpm for 5 minutes. The DNA is in the supernatant; 1µl was used in a 25µl PCR reaction.

2.2.4.3 Primer designing

Primers for all PCR reaction were designed using Primer 3 web version 4.0.0. http://bioinfo.ut.ee/primer3/

Table 2.4: Sets of primers used for PCR amplification and sequencing

Gering Lab Data Base (DB) Number	**Primer position**	**Primer direction**	**Primer sequence (5'-3')**
Detection of recombination fragments, that were designed within reporter constructs			
DB432	*ef1-α promoter*	Forward	TAATCCCCGCCCAGTAGAG
DB433	*dsRed*	Reverse	CCTTGGTCACCTTCAGCTTC

DB505	*Ubiquitin promoter*	Forward	TTCGGGTGAACTATCTCGTTT
DB506	*eGFP*	Reverse	CCGGTGAACAGCTCCTCG
DB573	*hsp promoter*	Forward	TCATAATGAAACAATTGCACCGA
DB574	*eGFP*	Reverse	CCCGGTGAACAGCTCCTC
Genotyping and sequencing of *gfi1ab* mutants			
DB582	Exon3	Forward	GGTCACGCGCTCATTGTAAT
DB565	Exon2	Reverse	CCTCCAAGTCTCCTCTCAGC
DB622	Intron1	Reverse	ACACACAAAACGTTTTCAGTTCT
DB751	Intron 2-3	Forward	CCTTCAGGTAAATGATTTGTGGT
DB752	Exon1	Reverse	CGCGATCTTTGGAGGATGAC
DB755	Exon3	Forward	AGATCCTCGTTCGGTGAAAA
DB757	Intron 2-3	Reverse	TTTCCATAAAGGCTACCTGTCAA
DB392	Intron1-2	Reverse	TTCGTCAAAACCACAGCTGT
DB393	Exon2	Forward	TTACGAGGACTTTTGGCGAC

2.2.4.4 Polymerase Chain Reaction (PCR)

A number of PCR reactions were set up and carried out for genotyping of various In order to carry out a PCR, depending on the number of samples a master mix was prepared, generally for most PCR reaction mixes, Taq DNA polymerase (NEB, M0273S) was used,

however for PCRs done in Chapter 5 or 6 Q5® High-Fidelity DNA polymerase (here after will be referred to as Q5 polymerase) (NEB, M0491S). Examples of a typical 25µl PCR reaction mix using Taq DNA polymerase and Q5® High-Fidelity DNA polymerases (Table 2.5).

Table 2.5: Typical PCR reaction (25µl) set up using Standard Taq polymerase and Q5 polymerase

Components and volumes used for a PCR reaction mix using Taq polymerase		**Components and volumes used for a PCR reaction mix using Q5 polymerase**	
Components	**Volume**	**Components**	**Volume**
Sterile distilled water	15.7µl	Sterile distilled water	13.4µl
10X Standard Taq buffer	2.5µl	5X Q5 polymerase buffer	5µl
2mM dNTPs	2.5µl	2mM dNTPs	2.5µl
10mM Forward oligo	1.5µl	10mM Forward oligo	1.5µl
10mM Reverse oligo	1.5µl	10mM Reverse oligo	1.5µl
Taq polymerase	0.3µl	Q5 polymerase	0.1µl
gDNA	1µl	gDNA	1µl
Total volume	25µl	Total volume	25µl

Most PCRs were carried out in thermocyler 'Techne TC-512' or in sometimes 'Teche PRIME' was used. A typical PCR reaction thermocycle is depicted in Table 2.6

Table 2.6: General PCR thermocycle steps and conditions using Standard Taq Polymerase or Q5 Polymerase

Thermocycler Steps		Temperatures		Times	
		Taq polymerase	**Q5 polymerase**	**Taq polymerase**	**Q5 polymerase**
Initial Denaturation		94°C	96°C	2mins	2mins
Denaturation	35 cycles for Q5 polymerase and 40 cycles for Taq polymerase	94°C	96°C	15sec	10sec
Annealing		58°C	65°C	30sec	30sec
Extension		72°C	72°C	1min/kb	1min/kb
Final extension		72°C	72°C	5mins	5mins
Hold		4°C	4°C	Indefinite	Indefinite

2.2.4.5 Agarose Gel Electrophoresis (AGE)

AGE was carried out using 0.8-2% agarose (Fisher Scientific, BP160-500) gels. 0.8-1% gels were used to resolve larger fragments that were over 1kb in length and 1.5-2% gels were used to separate fragments below 1kb in length. The agarose was dissolved in 0.5X TAE (25ml 50XTAE see Table 2.1 in 1 litre distilled water) by heating the mixture in a

microwave. After the agarose powder was completely dissolved, the solution was cooled down and EtBr (0.5µg/ml) was added. EtBr intercalates into DNA in the sample thus allowing it to be visualized under UV light. After addition of EtBr, the solution was allowed to set. Once the gel had set DNA samples were mixed with 6X gel loading dye (NEB, B7024S), to facilitate loading of the DNA samples into the wells along with reference DNA ladders (100bp ladder, NEB, N3231S or 1Kb ladder, NEB, N3232S), which helped in determining the DNA fragment sizes. Fragments in the sample were resolved in an electrophoretic chamber (Advance, Mupid®-ex) at 100V for 30 minutes for analytical gels and at 50V for 1 hour in case of preparative gels. Gels were visualised and imaged using a Trans Illuminator (GelDoc-It imaging system).

2.2.4.6 Restriction digestion of Plasmid DNA

Restriction digests were typically set up as 50µl reactions for preparative digests whereas, for analysis of plasmid isolations and for screening 20µl digests were set up. Restriction enzymes (RE) and their buffers used were obtained from NEB and the optimum conditions of incubation were set as per NEB guidelines. Generally, incubation times for a restriction digestion varied from 1 hour (for high-fidelity (HF) restriction enzymes) to overnight incubations. A typical 20µl restriction digestion set up was done as follows;

Components	**Volumes**
Sterile distilled water	yµl=15.5 µl- xµl
10X BSA (not required for all RE)	2µl
10X Buffer	2µl
DNA sample (1µg)	xµl

Restriction enzyme (10U)	0.5µl
Total volume	20µl

2.2.4.7 Isolation of DNA fragments by gel extraction

After running a preparative AGE the gel was visualized under UV light, following health and safety measures. The fragment of interest was identified and was cut out of the gel using a sharp blade or scalpel, this procedure was carried out making sure to avoid long exposure of the DNA fragment to UV light, which may introduce mutations (in particular thymidine dimers) in the DNA sequence. The gel slice was then put into a pre weighed 1.5ml Eppendorf tube and was weighed again in order to determine the weight of the gel slice containing the DNA fragment of interest. The gel extraction of the DNA fragment was carried out using either commercially available gel extraction kits (Sigma, NA1111 or Macherey-Nagel NucleoSpin®, 740609) as per the guidelines specified by the manufacturer. In case, of cloning experiments extracted fragments were ligated together.

2.2.4.8 Ligation of linearized DNA fragments to linearized vector

Following linearization of vector and insert fragments by RE digestion and gel extraction of the fragments; the two fragments were ready to be ligated together and in order to do so, the T4 DNA ligase (NEB, M0202S) was used as per NEB specified guidelines and the reaction was incubated at 16°C overnight. The amount of vector and insert to be added in a 20µl ligation mix was calculated as follows;

$$\text{Amount of insert in ng} = \frac{\text{50 ng vector x size of insert}}{\text{size of vector}} \text{ x3}$$

The next day the ligation mix was used to transform into chemically competent *E.coli* cells.

2.2.4.9 Cloning of DNA in E.coli

The ligation mix prepared was subsequently used to transform 5-alpha competent E.coli (Subcloning efficiency) cells (NEB, C2988J) or for larger constructs, alternatively 5-alpha competent E.coli (High efficiency) cells (NEB, C2987H) were used. The transformation was carried out by careful addition of ligation mix to the competent *E.coli* cells. The suspension was left on ice for 15 minutes. After 15minutes the tube containing the E.coli cells and ligation mix was heat shocked at 42°C for 90 seconds. Exactly after 90 seconds the tube was placed back onto ice for 5 minutes. Recovery and phenotypic expression of the competent cells was carried out by the addition of 500µl of enriched SOC broth, followed by incubation at 37°C for 1 hour on a shaker. After 1 hour the cells were spread on LB solid medium containing Amp (100µg/ml). Plates were incubated at 37°C overnight. The next day the colonies were picked for screening and plasmid DNA isolation.

2.2.4.10 Isolation of plasmid DNA

Transformed colonies were picked and inoculated in LB broth containing Amp (100µg/ml) and the tubes were incubated at 37°C on a shaker. Plasmid DNA was isolated using culture grown for 14-16 hours using commercially available mini-prep and midi prep kits (Macherey-Nagel Nucleospin® Plasmid or NucleoBond® Xtra Midi/Maxi). Plasmid

isolation was carried out as per the instructions set by the manufacturer. Plasmids were screened for carrying the right insert by carrying out RE digestions.

2.2.4.11 Measuring DNA concentration

Concentrations of plasmid DNA and DNA fragments for sequencing were spectrophotometrically measured using NanoDrop1000 v3.2. A purity criterion for DNA was determined by the 260/280-absorbance ratio value. Where a value of absorbance ratio ~1.8 along with a clear peak (not skewed or multiple peaks) indicated a pure DNA yield.

2.2.5 Transgene Construction

Isolated plasmid DNA was injected along with Tol2 transposase mRNA into one-cell stage zebrafish embryos (Kotani et al., 2006) to generate transgenic lines. Embryos that were injected with reporter genes were observed for mosaic fluorescent reporter gene expression and were raised unto adulthood. Adult F0 fish were crossed crossed to wt fish to identify founders.

2.2.6 Sequencing of gel extracted DNA fragments

Source Biosciences Sanger Sequencing Services sequenced DNA fragments. As per the sequencing requirements 5µl of 1ng/100bp of the DNA fragment and 3.2µM concentration of the oligo to sequence the fragment was sent.

2.2.7 Adult Zebrafish WKM isolation

The method for adult fish dissection and WKM isolation was carried out as described by Gerlach et al (Gerlach et al., 2011). Briefly, the fish were selected and euthanized by treating at least for 10 minutes either with 4g/L MS222 (Table 2.1) or 4 pumps per litre of Vertak, Aqua-Sed (2-phenoxyethanol). Once the fish were completely motionless they were placed on ice until they were ready to be dissected. For the dissection, the fish was

placed on an absorbent paper towel and using a pair of sharp scissors the head of the fish was cut from under the operculum at an angle of approximately 45°. At this stage the tail was also cut and the trunk of the fish was opened up ventrally. The fish was then placed on a dissecting/wax tray. The tray was placed under a dissecting scope (Nikon SMZ645 or Nikon SMZ1000) with an external light source (Photonic PL3000) to visualize the fish anatomy clearly. The fish was pinned down to the dissection tray and the internal organs such as the gonads, intestine, swim bladder were removed. The kidney was identified which includes a head, saddle and tail region. This was carefully detached from the trunk using a pair of tweezers. For FAC analysis and sorting it is important to keep cells alive and hence the procedure was done keeping time in mind.

2.2.8 Preparation of blood cell suspension from isolated WKM for FAC analysis

Once the kidney had been isolated and it was placed in a 10mm sterile tissue culture dish (placed on ice) containing 2ml of cold 1X HBSS (see table 2.1). Next, a P1000 was used to dissociate the kidney in order to flush the blood cells out of the WKM. This was carried out for 3minutes followed by addition of 5mls of FCS/PBS (Table 2.1) into the dish. The content was mixed using a P1000 and was strained through a 40µm pore size cell strainer (Falcon, VWR, 734-0002) into 50ml falcon tubes that were kept on ice. The tube containing the flushed blood cells was centrifuged (Heraeus, Multifuge 3S-R) at 4°C at 1500 rpm for 10 minutes. The supernatant was discarded and the blood cell pellet was resuspended in 300µl of FCS/PBS (Traver et al., 2003). The suspension was then transferred in to a fresh set of 15ml tubes. All steps were carried out on ice in order to keep the cells alive. The cells were then analysed on a flow cytometer and the remaining cell suspension after the analysis was then used for genotyping.

2.2.9 Flow cytometric analysis: Fluorescent activated cell (FAC) analysis:

FAC analysis is a method carried out on a flow cytometer that uses lasers of different wavelengths to analyse cells that are fluorescently labelled. Using this method, one can sort and count fluorescently labelled cells. When cells are fluorescently labelled (for example, either with a green fluorescent protein like eGFP or a red fluorescent like dsRed or mcherry) cells are passed by a light source the fluorescent molecules are activated to a higher energy state. Cell populations can be resolved on the basis of physical properties of the cell such as size (forward scatter (FC)) and internal complexity or granularity (right angle scatter or side scatter (SC) (Brown and Wittwer, 2000). The WKM blood cell preparation (previous section/step, above) was treated with SYTOX red, which was used to exclude dead cells and debris from the analysis. The FC, SC and eGFP/dsRed/mCherry expression of cells were then analysed on a Bechman Coulter MoFlo XPD or MoFlo Astrios using the Weasel or Kaluza softwares. The excitation wavelengths for eGFP and dsRed/mCherry were 488nm and 561nm respectively the emission wave lengths were 513/26 for eGFP, 579/16nm for dsRed and 614/20 for mCherry. Sytox red on the other hand was detected at 640nm excitation and an emission at 670/30 nm.

2.2.10 In vivo fluorescent analysis and imaging:

Fluorescence analysis in embryos was done on Nikon SMZ1500 microscope that was fitted with filter sets and a mercury lamp to detect fluorescence. These microscopes were either attached to Nikon DXM1200F or Nikon Digital Sight DS-U1operated using the Nikon ACT-2U software that captured coloured images fluorescent and bright field images. A more sensitive camera was the Hamamatsu ORCA-ER C4742-80 digital camera run with IP Lab software. Although extremely sensitive it took only black and white images that could be pseudocoloured. The Hamamatsu was used with the Nikon Eclipse 80i

microscope. eGFP and dsRed expression were viewed with the help of a FITC filter set and a TRITC filter set respectively attached to a mercury light source. All embryos were anesthetised before observing and imaging using 700µl of 4g/L pH7-7.5 MS222, in order to immobilize them.

2.2.10.1 Preparation of embryos for confocal microscopy

A method to immobilise embryos was devised in which embryos were treated with a chemical cocktail of the neuromuscular blocker, Tubocurarine (20µM) (Table 2.1) along with the anaesthetic, MS222 (130µg/µl) and 0.003% PTU to prevent pigmentation in E3 medium. Embryos were incubated in the solution at 28.5°C for 1 hour to make sure the embryos are completely immobilized and anesthetized although. Once completely anesthetized but with a heartbeat and blood circulation going, the embryo was embedded onto a confocal glass bottom dish (MatTek 35mmX10mm No.0 Coverslip) using 1% low melting point agarose and overlaid with the above anaesthetic solution.

2.2.10.2 Confocal imaging and image acquisition

On the Zeiss LSM710, a C-Apochromat40x/1.2 W Korr K27 lens was used. The laser wavelengths used to excite Cerulean, GFP and dsRed/tdTom/mCherry were 405, 488 and 561 nm, respectively. Their fluorescence was detected at 454-505, 505-544 and 566-678 nm, respectively. All confocal images were analysed in Imaris (Bitplane). To determine the number blood cells in the mesenchyme and CHT, individual cells were counted in a Z-stack of confocal images. To estimate the size of a cell or a nucleus, the area of the ellipse described by the cell's or nuclei's main axes, m and n, was calculated as $A = \pi (m/2) (n/2)$. Videos exported from Imaris were annotated in iMovie. Figures were assembled using Adobe Photoshop CS6.

RESULTS

3 An inducible CreER^{T2} recombination system demonstrates that at least a subset of adult WKM haematopoietic cells arise from flk/*kdrl*+ endothelial cells in the early embryo

3.1 Introduction

Bertrand et al previously used a Cre:LoxP system to lineage-trace the progeny of ECs in zebrafish. They found expression of a recombined Cre reporter gene in adult WKM HCs. While this finding suggested an endothelial origin of adult blood cells (Bertrand et al., 2010), it remained to be established when the blood cells developed from the endothelium. Here, to test whether the blood cells arise in the embryo, a tamoxifen inducible CreER^{T2}/LoxP system was used to restrict Cre activity and, therefore, Cre reporter gene recombination to embryonic ECs.

Cis-regulatory elements of the *flk1/kdrl* gene were used to drive Cre expression in ECs. The cis-regulatory elements, an 0.8kb enhancer (-4.3, -3.5) and a 1.5kb promoter element, had previously been reported to direct expression of an e*GFP* reporter gene to ECs in transient transgenic zebrafish embryos (Choi et al., 2007). To generate transgenic lines in our lab, these elements were cloned in front of an *eGFP* (Fig 3.1Ai) and a *tdTomato* reporter gene inside a Tol2 transposon (Fig 3.1Aii). The constructs were injected with mRNA for Tol2 transposase into one-cell stage zebrafish embryos. The resulting transgenic lines *Tg(flk1/kdrl:eGFP)^{qmc67}* (Fig. 3.1Bi) and *Tg(flk1/kdrl:tdTom)^{qmc64}* (Fig. 3.1Bii) revealed green and red fluorescent protein expression in embryonic ECs, respectively. Dr. Kathryn McMahon and Dr. Roshana Thambyrajah generated the *flk1/kdrl* reporter lines qmc67 and qmc64 respectively. Unlike the stable transgenic line described

by Choi and colleagues (Choi et al., 2007), none of our lines displayed ectopic reporter expression in the brain, the eyes or the spinal cord.

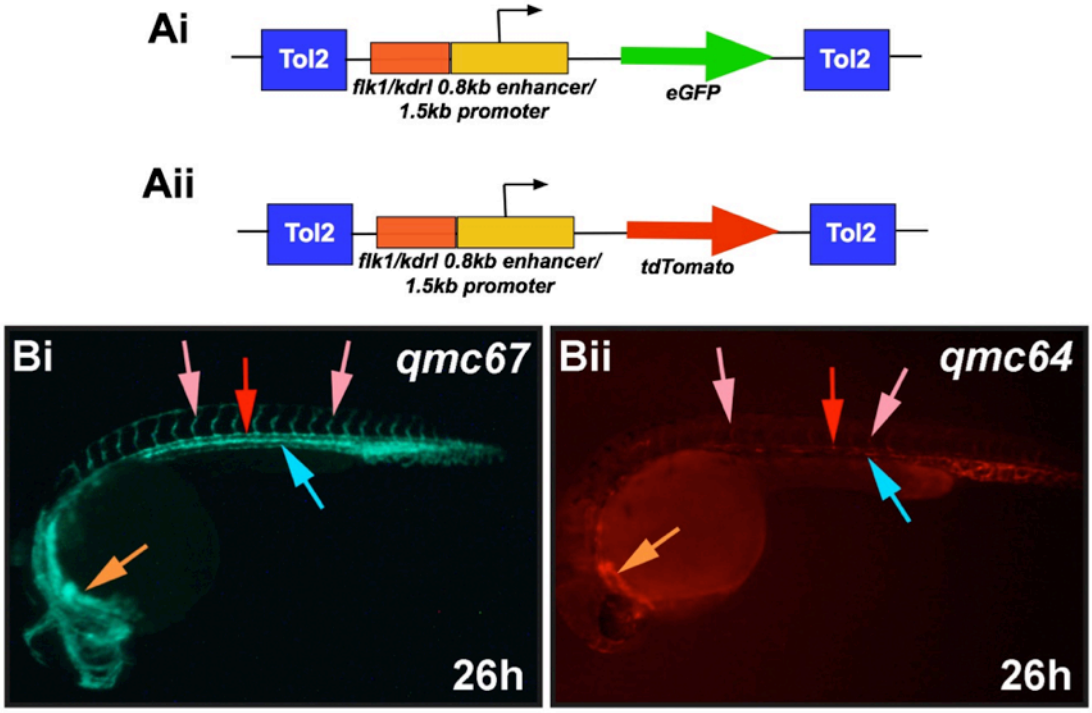


Figure 3.1: eGFP and tdTom fluorescence displayed in Tg(*flk1/kdrl:egfp*)^{qmc67} and Tg(*flk1/kdrl:tdTom*)^{qmc64} transgenic embryos at 26 hpf.
(Ai and ii) Schematic representation of the *flk1/kdrl:eGFP* and *flk1/kdrl:tdTom* transgene constructs. Both constructs contain the 0.8 kb (-4.3; -3.5 kb) enhancer and 1.5 kb promoter fragment that drives either eGFP or tdTom expression (Bi and Bii) Images of the embryos taken on a fluorescent microscope using FITC and TRITC filter sets, respectively. Arrow colour code: red arrow: dorsal aorta, blue arrow: the posterior cardinal vein, pink arrow: intersomitic vessel, orange arrow: the endocardium. (Green and red fluorescent images were taken from the Gering lab database)

Having confirmed that the cis-regulatory elements of the *flk1/kdrl* gene directed gene expression to ECs the same elements were used to drive Cre expression in ECs. For this purpose, the Cre recombinase gene was cloned downstream of the *flk1/kdrl* enhancer/promoter element on a Tol2 transposon. This transposon was inserted into the zebrafish genome generating the *Tg(flk1/kdrl.Cre)*^{qmc101} line. To visualise Cre activity in ECs and their progeny, F1 adult fish were crossed to fish that carried the Cre reporter transgene *ef1α:LoxP-eGFP-terminator-LoxP-dsRed* (Sinha et al., 2010) (from now on

referred to as the *ef1α-Cre reporter* line). This transgene carries an *ef1α* promoter that drives the expression of a floxed eGFP-reporter-Stop cassette. This cassette lies upstream of a *dsRed* reporter gene that is expressed only once Cre has recombined the loxP sites and removed the eGFP-stop cassette. In the absence of Cre, the *ef1α* promoter drives ubiquitous eGFP expression throughout the embryo. When the *flk1/kdrl:Cre* and *ef1α-Cre* reporter lines were crossed and their progeny was examined, dsRed+ spindle-shaped cells were observed in the endothelium of 4 dpf embryos that displayed ubiquitous GFP expression (Figure 3.2). In the same embryos, dsRed+ round cells were detected within the CHT, suggesting that Cre-expressing endothelial cells in the embryo had given rise to blood cells of the CHT (These experiments were performed by Andrea Haase).

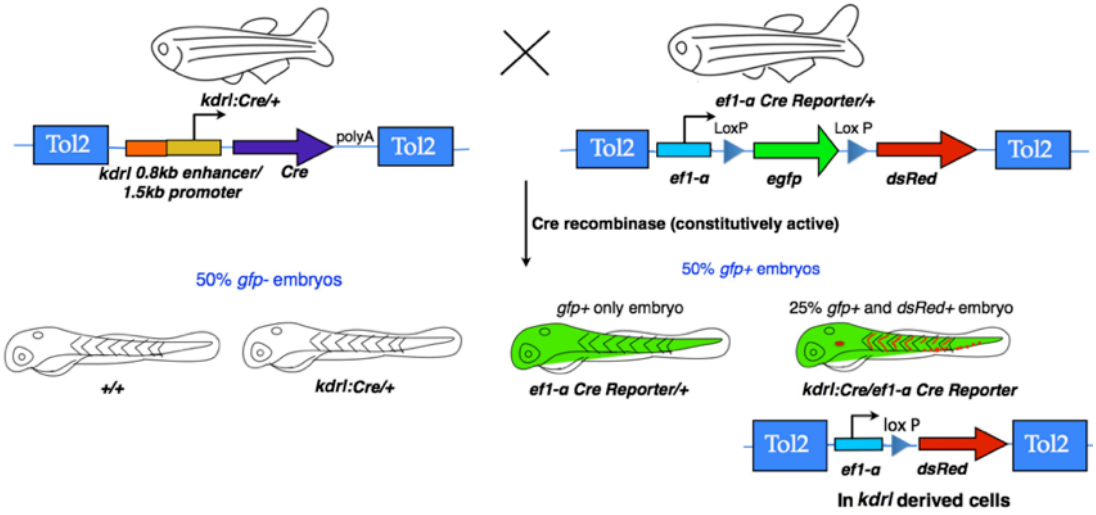


Figure 3.2: Cre-mediated recombination of the *ef1-α:Cre* reporter transgene in endothelial cells of embryos that also carry the *flk/kdrl:cre* transgene.
The *Cre* is under the control of the *kdrl* 0.8 kb enhancer and 1.5 kb promoter fragment. The line that carries this transgene was named *qmc101*. The Cre reporter line was a generous gift from Bally Cuif's lab in Paris. It expresses in which the fluorescent reporter genes under the control of a ubiquitous promoter, the *ef1-α* promoter. When heterozygous carriers of the two transgenes were crossed 50% of the embryos are eGFP- and 50% are eGFP+. Half of the eGFP+ embryos are double transgenic embryos. In the latter, Cre is expressed in flk1/kdrl+ endothelial cells. Cre recombines the *Cre* reporter gene in these cells. The *ef1-α* promoter drives expression of dsRed expression in all of the progeny of the *flk1/kdrl*+ cells.

The *flk1/kdrl* promoter and enhancer elements are expected to be active in embryonic and adult endothelial cells. To show that embryonic ECs are the source of adult haematopoietic cells Cre activity needed to be restricted to embryonic ECs. For this purpose, a second Cre driver construct was generated in which the Cre reading frame was fused to the reading frame of a modified human oestrogen receptor (ER^{T2}). In the absence of its ligand tamoxifen, ERT2 binds to the chaperone Hsp90, a protein that is located in the cytoplasm. ERT2, thus, keeps Cre in the cytoplasm and prevents it from recombining the Cre reporter gene (Figure 3.3). Exposure to tamoxifen (TAM) or its active metabolite 4-hydroxy tamoxifen (4-OHT) releases $CreER^{T2}$ from Hsp90 and allows it to enter the nucleus (Feil, 1997; Feil et al., 1996). Previous zebrafish studies had shown that TAM and 4-OHT can induce reporter gene recombination. However, recombination was faster in the presence of

the 4-OHT. Furthermore, its effective concentration (0.5μM) was one order of magnitude lower than that of TAM (5μM) (Hans et al., 2009). Thus, a 5μM 4-OHT solution was used in the following experiments.

To generate a stable *flk1/kdrl:CreER^{T2}* transgenic zebrafish line, wt embryos were injected at the one-cell stage with a plasmid that carried the transgene on a Tol2 transposon (experiments performed by Dr. Martin Gering). Transposase mRNA was co-injected with the plasmid DNA (Kikuta and Kawakami, 2009). Injected embryos were raised and the adult fish were crossed to fish that carried the *ef1α*-Cre reporter to see whether the injected fish passed the *flk1/kdrl:CreER^{T2}* transgene through their germline (Sinha et al., 2010). The progeny was exposed to 4-OHT from tailbud stage to induce recombination of the Cre reporter gene in ECs of double transgenic embryos. To achieve efficient transgene recombination in all embryonic ECs, 4-OHT was added at tailbud stage (10 hpf). Thus, 4-OHT addition occurred a few hours prior to the onset of endogenous *flk1/kdrl* expression in the lateral mesoderm (Liao et al., 1997) and over 20 hours before EHT is observed from the ventral wall of the dorsal aorta (Kissa and Herbomel, 2010). 4-OHT treatments were maintained until 5dpf.

A 4-OHT inducible *Tg(kdrl:CreER^{T2})^{qmc106}* line (from this point on will be referred to as *kdrlCreER^{T2}*) was identified and was used in subsequent experiments (Jalali, 2012).

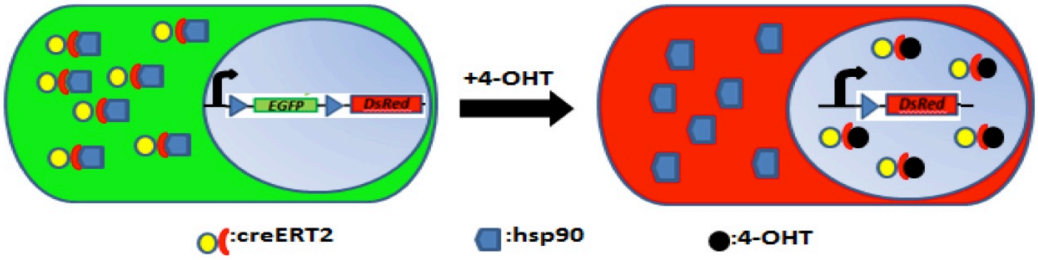


Figure 3.3: The 4-hydroxy-tamoxifen (4-OHT)-inducible Cre*ER*^{T2}/LoxP recombination system.
In the absence of the ligand 4-OHT, the *CreER*^{T2} fusion protein is held in the cytoplasm of the cell by the hsp90 protein, and the cell expresses eGFP under the control of a ubiquitous promoter. In the presence of the ligand 4-OHT, the CreER^{T2} is released form the chaperone (Hsp90) and can enter the nucleus where recombination occurs. The ubiquitous promoter drives the expression of dsRed.

To study 4-OHT-inducible reporter gene recombination in the embryo in more detail, Tg(*flk1/kdrl:CreER*^{T2})^{qmc106} F1 fish were crossed to *ef1α*:*Cre reporter* transgenic fish. Embryos were collected and exposed to 5µM 4-OHT from 10hpf. Reporter gene expression was analysed at 5 dpf under the fluorescent microscope. A control batch of embryos was exposed to the solvent, ethanol. In the control batch (n=112), ubiquitous eGFP expression was observed at 5dpf (Figure 3.4(i)) but no dsRed expression was noticed (Figure 3.4(ii)). Thus, no recombination had occurred in the absence of the ligand, confirming that the activity of *CreER*^{T2} was dependent on the presence of the ligand 4-OHT.

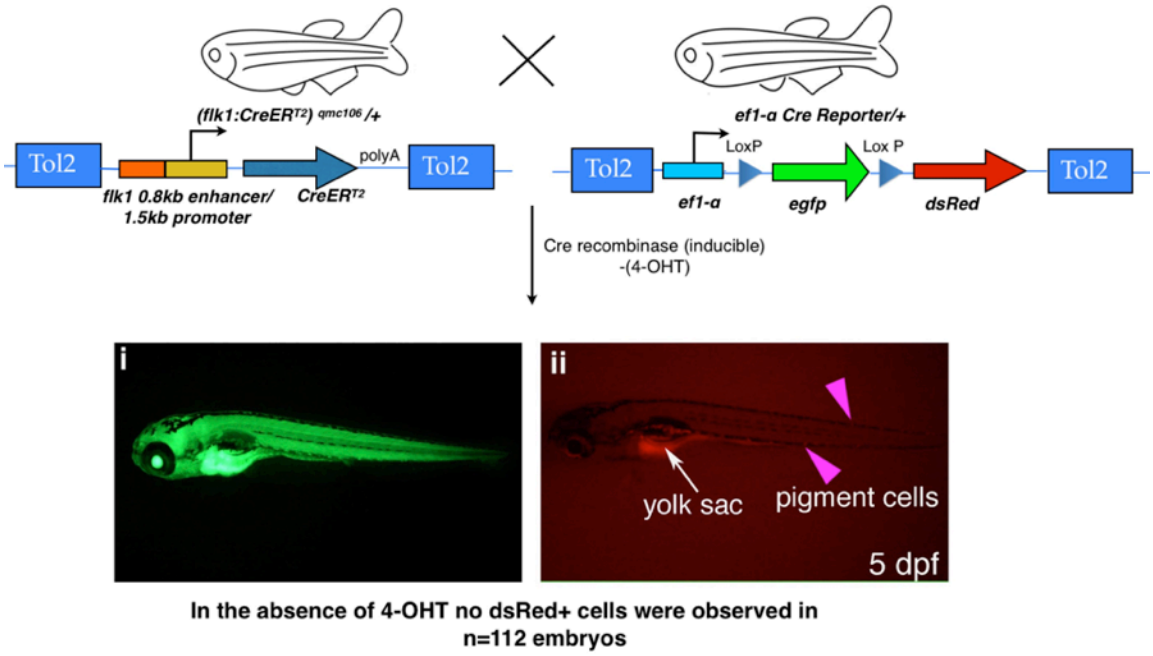


Figure 3.4: Lack of dsRed Cre reporter gene expression in the absence of the ligand 4-OHT. In the absence of 4-OHT, double transgenic Tg(*kdrl:CreERT2*)^{qmc106}; Tg(ef1α:Cre reporter) embryos express eGFP ubiquitously (i), but do not undergo recombination and hence display no dsRed+ cells (ii) when analysed on a fluorescent dissecting microscope using the FITC and TRITC filter set, respectively. Some pigment cells (pink arrowheads) and the yolk sac (white arrow) auto-fluoresce. Embryos were analysed at 5dpf.

3.2 *Flk1/kdrl*-expressing endothelial cells give rise to haematopoietic cells in the caudal haematopoietic tissue

Here, double transgenic *Tg(kdrl:Cre)*^{qmc101};Tg(*ef1α:Cre) reporter* embryos were examined beyond day 4. At day 4, these embryos displayed the expected dsRed expression in the endothelium and in blood cells of the CHT (Figure 3.5B and C). On days day 5 and 6, red fluorescence was also seen in the thymus and the kidney (Figure 3.5D, E). dsRed expression was also observed in endocardial cells of the heart (Figure 3.5D, E). By 5.5 dpf, dsRed+ cells were also detected in circulation (Supplementary Movie 3.1). Together with

the expression in the CHT, thymus and kidney, this strongly suggested that the dsRed+ cells included blood cells that had undergone Cre reporter gene recombination. Since there was no evidence of *flk1/kdrl* promoter activity in blood cells of the CHT themselves, the data suggested that the dsRed+ blood cells were derived from the *flk1/kdrl*+ embryonic ECs. It is worth noting that single transgenic *ef1α-Cre reporter* embryos did not show any dsRed expression in any tissue at any of these embryonic stages (Figure 3.5E), demonstrating that recombination of the Cre reporter was required for endothelial and haematopoietic expression of the dsRed reporter gene.

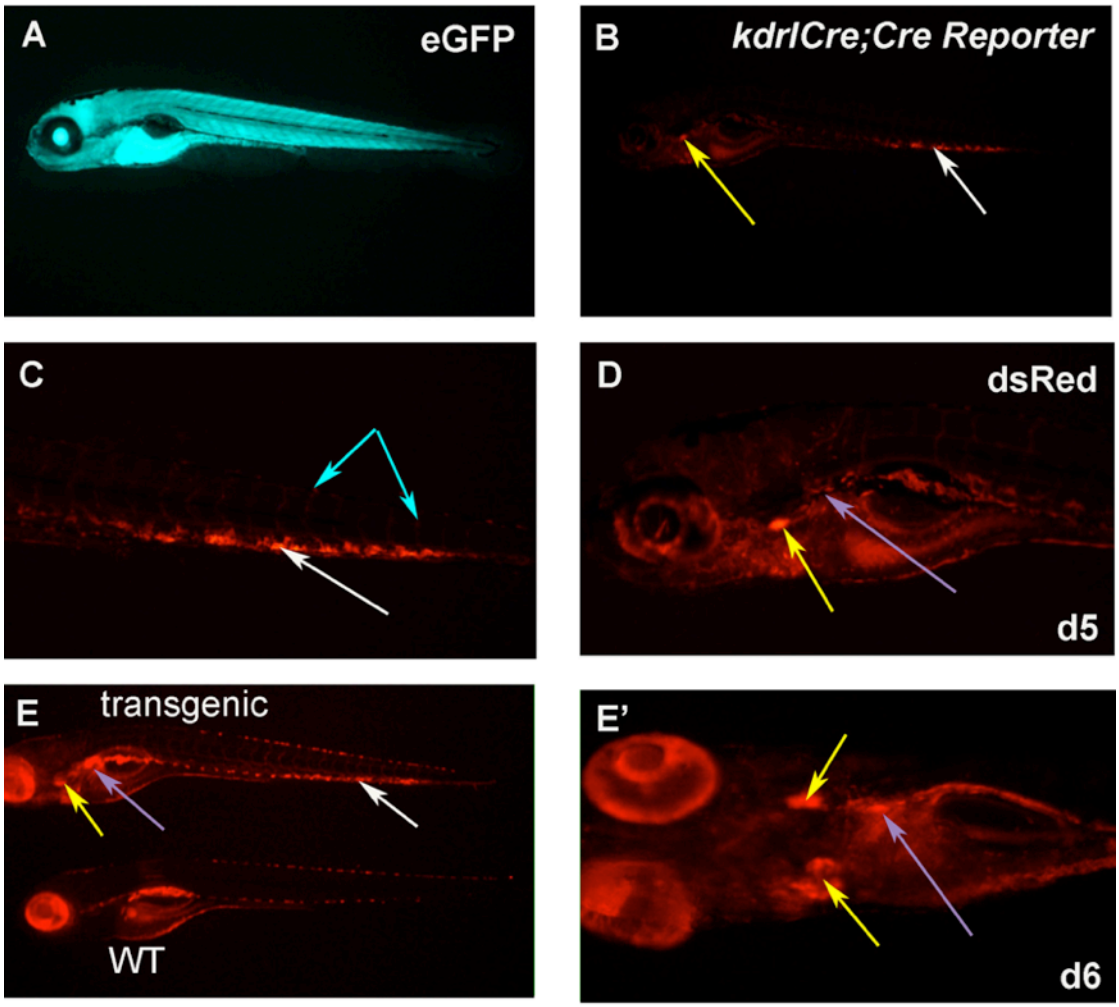


Figure 3.5: Double transgenic Tg(*flk1/kdrl:cre*)^{qmc101},Tg(*ef1α*:Cre reporter) embryos display dsRed expression in the endothelium and in blood cells of the CHT, the thymus and the kidney at 5 dpf.

(A) Ubiquitous expression of eGFP under the control of the *ef1α* promoter in a double transgenic embryo at 5 dpf. (B) dsRed fluorescence in the same embryo at 5 dpf. (C-D) Close-up images of the embryo shown in (B). (E) Red fluorescence displayed in double transgenic and wt embryos at 6 dpf. (E') shows a magnified images of the head region of the double transgenic embryo shown in (E).Please note that both embryos shown in (E) display auto-fluorescence in the pigment layer of the retina and in pigment cells throughout the body. Arrow colour code white arrow: CHT, yellow arrow: thymus, turquoise arrows: ISVs, purple arrow: kidney, blue arrow: endocardium, pink arrow: pigment cells. Arrowheads indicate the absence of cells within the control embryo, where white arrow: CHT, yellow arrow: thymus, purple arrow: kidney, blue arrow: endothelial cells.

3.3 The inducible *kdrlCreER^{T2}/loxP* system selectively labels cells of the vasculature and haematopoietic cells of the early larvae

Here, when embryos were treated with 4-OHT from 10 hp and were analysed at 5dpf under a fluorescent dissecting microscope, dsRed expression was detected in half of the embryos that showed ubiquitous eGFP expression (Figure 3.6A). dsRed-expressing cells were observed (Figure 3.6B) within the CHT, the thymus, the endocardium and in ECs of some intersomitic vessels (Figure 3.6C, D and E). dsRed expressing cells were also seen within the kidney (Figure 3.6F) of these early larvae. It is worth noting that the number of recombined cells was generally lower in 4-OHT-treated *qmc106* transgenic embryos than in *qmc101* carriers. Altogether, these data on the *Tg(kdrl:CreER^{T2})^{qmc106}, Tg(ef1α:Cre reporter)* embryos show that Cre activity can be induced by 4-OHT in endothelial cells that express CreER^{T2} under the control of the *flk1/kdrl* regulatory elements. The data support the idea that the labelled ECs give rise to dsRed-expressing HCs that seed larval sites of haematopoiesis.

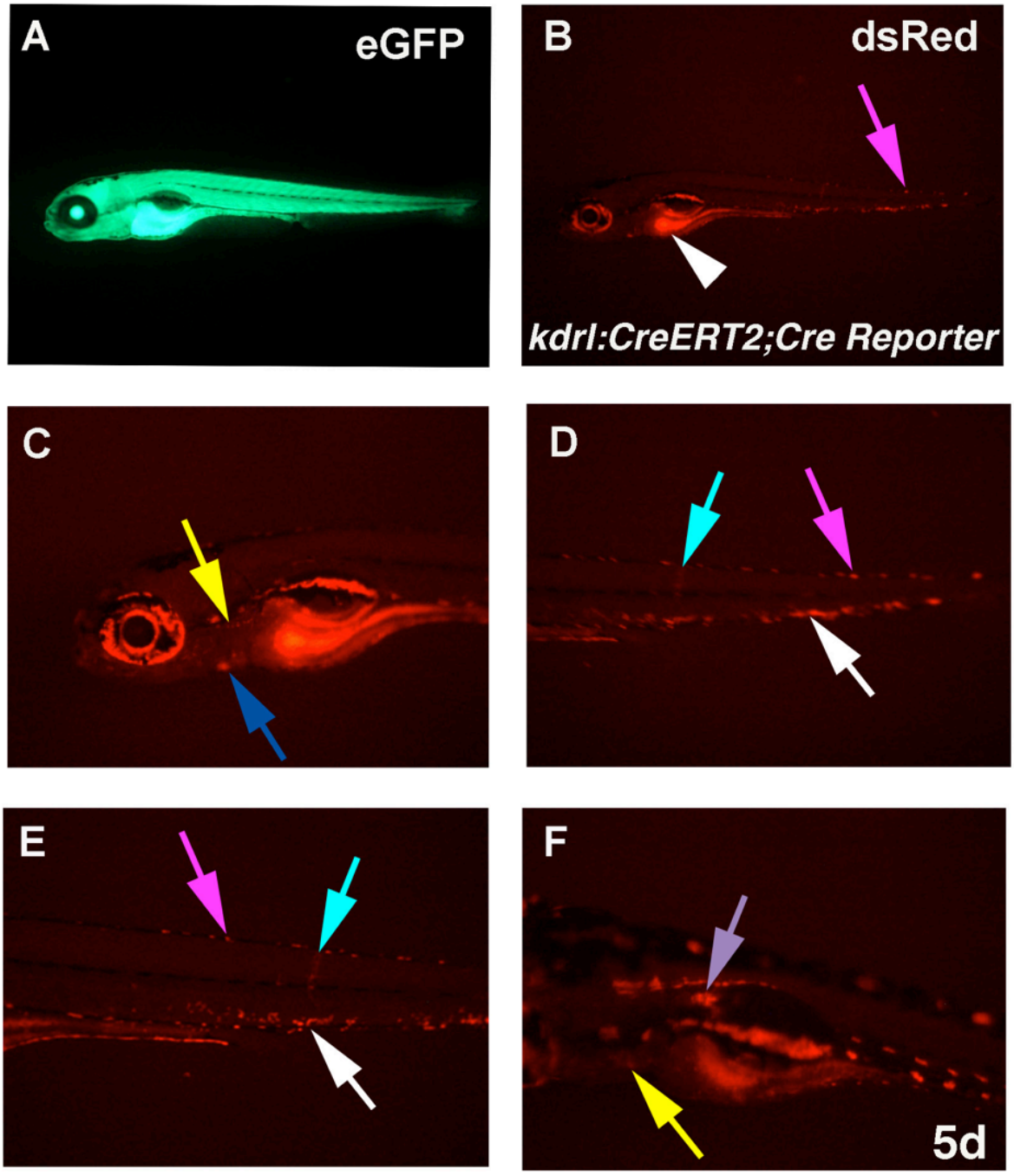


Figure 3.6: *flk1/kdrl*+ endothelial cells in the early embryo give rise to blood cells that seed the embryonic CHT and early larval thymus and kidney of 4-OHT induced embryos.
(A) eGFP is ubiquitously expressed under the control of the *ef1α* promoter. Autofluorescence was observed in the yolk sac (B) and in pigment cells. In the presence of 5μM 4-OHT from 10hpf-5dpf, recombined reporter gene expression was detected by the presence of dsRed expressing cells within the heart (C) and in the thymus (C and F), (D-E) in the CHT, intersomitic vessels and in (F) the kidney. (pink arrowhead in (B, D and E) the embryos were analysed using FITC filter set for eGFP expression and under the TRITC filter set to observe dsRed expression. Arrow colour codes; blue arrow: endocardium, yellow arrows: thymus, turquoise arrows: ISVs, white arrows: CHT, purple arrow: kidney marrow, white arrowhead: yolk sac, pink arrows: pigment cells.

3.4 The Cre/LoxP recombination system suggests that adult blood cells arise from *flk1/kdrl*+ endothelial cells

To investigate whether adult blood cells were derived from *flk1/kdrl*+ endothelial cells, double transgenic Tg(*flk1/kdrl:cre*)^{qmc101};Tg(*ef1-α:Cre reporter*) embryos with dsRed-expressing blood cells at days 4 and 5 were raised to adulthood. Adult fish were sacrificed and the whole kidney marrow (WKM) was prepared by flushing the blood cells out of the isolated kidney. Two approaches were used to examine the recombination status of the Cre reporter gene in the adult kidney cells.

In the first approach, genomic DNA was prepared from the WKM cells and PCR was carried out on gDNA. This experiment was carried out on kidney cells from control fish. These were the *ef1α*:*Cre reporter* single transgenic fish, as well as a *wt* fish. For the PCR, the primers DB432 and DB433 were designed within the *ef1α* Cre reporter transgene to amplify the region between the *ef1-α* promoter and the *dsRed* gene (Figure 3.7A). Amplification of the DNA of the non-recombined allele was expected to generate a 1325 bp fragment, while the recombined allele was meant to give a 307 bp fragment. As expected, no amplification was observed in the control PCRs (no template control and *wt* DNA). A single 1325bp fragment was amplified on gDNA obtained from *ef1α* Cre reporter single transgenic fish. By contrast, the genomic DNAs of all six double transgenic fish revealed signs of Cre reporter gene recombination. For four of them, PCR amplification yielded both the 307bp and the 1325bp fragments, suggesting partial recombination of the reporter gene (Figure 3.7Bi and Bii). Complete recombination was observed in two cases in which only the 307bp fragment was amplified (Figure 3.7C). The identities of the 1325bp and the 307bp fragments were verified by DNA sequencing (Figure 3.8A and B

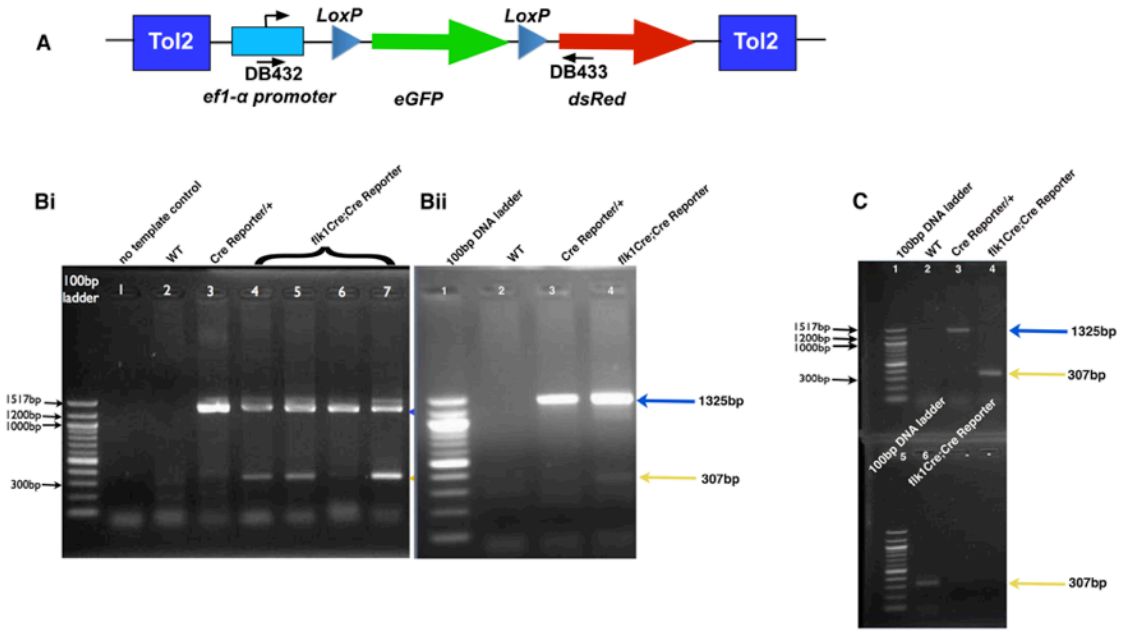


Figure 3.7: Blood cells in the adult fish kidney marrow are derived from *flk1/kdrl*+ endothelial cells.

(A) PCR primers were designed within the *ef1-α*:*Cre reporter* construct in order to detect the recombined and the non-recombined Cre reporter transgenes. The forward primer (DB432) hybridises to a 20 bp sequence in the middle of the *ef1-α* promoter, while the reverse primer (DB433) binds to a 20 bp sequence at the start of the *dsRed* gene. (Bi, Bii and C) Using primers DB432 and DB433, PCR was performed on genomic DNA isolated from cells flushed out of dissected adult kidneys. Amplification products were analysed by 1% agarose gel electrophoresis. The size of the amplification products and the DNA fragments of the 100 bp ladder are provided.

respectively), which was carried out on the sample in lane 4 (Figure 3.7 Bii).

Thus, while the data showed variable efficiency of reporter gene recombination in adult WMK cells, it was confirmed that adult WKM blood cells were derived from *flk1/kdrl*+ endothelial cells.

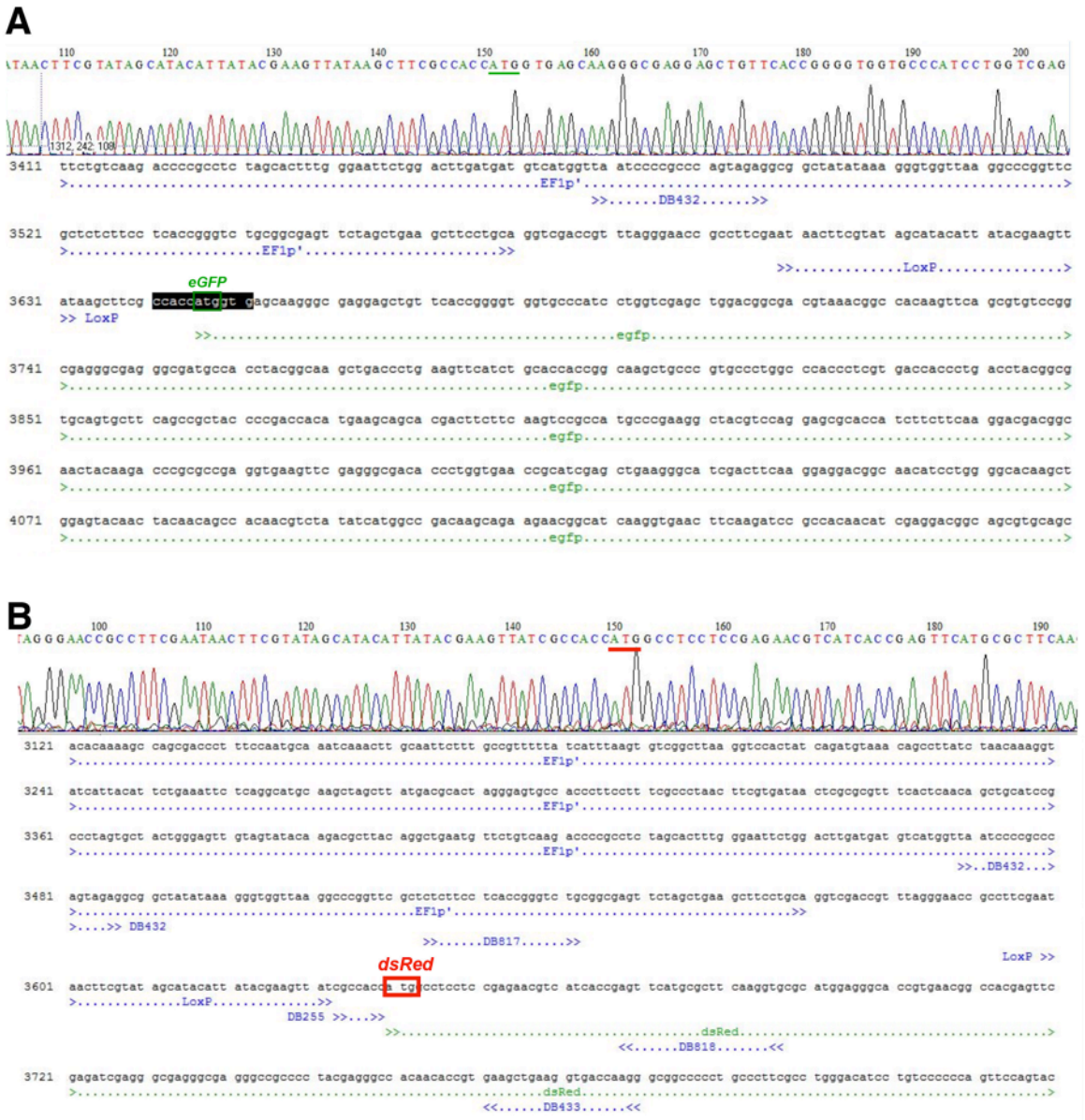


Figure 3.8: The sequence of the 1325bp and 307bp PCR fragment matches the sequence of the non-recombined and recombined *ef1α:Cre reporter* allele respectively.
(A) Primer DB432 was used to sequence the 1325 bp PCR fragment. The sequencing data were visualised as a chromatogram using BioEdit program. The sequence highlighted in the chromatogram matches the sequence of the non-recombined *ef1α:Cre reporter* allele highlighted in the sequence editing programme Clone Manager. The highlighted region starts a few base pairs 3' to the sequencing primer DB432 and includes the 3' part of the *ef1α*-promoter sequence, a *loxP* site and the *eGFP* sequence. On the BioEdit chromatogram, the ATG start codon of the *eGFP* gene is labelled. (B) Primer DB432 was used to sequence the 307 bp PCR fragment. The sequencing results are visualised in a chromatogram generated using the BioEdit software. The sequence highlighted in the chromatogram is the same sequence highlighted in the DNA editing programme Clone Manager below. The latter corresponds to the sequence of the *ef1α:Cre reporter* gene recombined by induction of $CreER^{T2}$ at the original two loxP sites. The sequence includes the 3' end of the ef1α-promoter sequence, a LoxP site and part of the dsRed reading frame. The sequence ends at the end of the reverse oligo DB433 that was used in the PCR. On the BioEdit chromatogram, the ATG start codon of the *dsRed* gene is labelled.

In a second experiment, blood cells flushed out of the adult kidney were analysed by flow cytometry. Unfortunately, only very few WKM cells displayed any fluorescence. Among the WKM cells isolated from Tg(*ef1-α:Cre reporter*) fish only about 0.74% expressed eGFP. In the double transgenic Tg(*flk1/kdrl:cre*)^{qmc101};Tg(*ef1-α:Cre reporter)* fish, red fluorescence was undetectable, yet the proportion of green fluorescent cells had dropped to 0.06%, suggesting that the majority of eGFP+ WKM cells had undergone recombination (Appendix 1). Thus, while the weak activity of the *ef1-α* promoter in adult blood cells (Hans et al., 2009) has only allowed us to analyse a minority of adult blood cells, these data together with the PCR results support the view that adult WKM blood cells are derived from *flk1/kdrl*+ ECs.

3.4.1 Development of alternative *Cre-reporter* genes

To overcome the problem with the weak *ef1-α* promoter activity, alternative Cre-reporter transgenes were developed that made use of the *ubiquitin* promoter and the *hsp70* promoter. While the former had been shown to drive high level ubiquitous reporter gene expression in the adult WKM (Mosimann et al., 2011), the latter was expected to mediate high level heat-shock-inducible gene expression.

First, the *Ubiquitin:Cre* reporter gene, an mcherry-to-eGFP *Cre* reporter gene, was generated (Appendix 2 and 3). When the *ubiquitin*:*Cre reporter* transgenic fish were established and ubiquitous mCherry expression was analysed at 4 dpf, it turned out to be fairly weak (Appendix 5). Similarly, when the transgenic fish were crossed to *Tg(kdrl:Cre)* transgenic fish, and double transgenic *Tg(kdrl:Cre; Ubiquitin:Cre reporter)* embryos were analysed at 4dpf, eGFP expression in embryonic ECs and HCs was not particularly strong. These embryos, along with wt or single transgenic Tg(*ubiquitin:Cre reporter)* animals as negative controls, were raised to adulthood. When adult WKM HCs of these fish were

analysed by flow cytometry, mCherry expression could not be detected in the single transgenic Tg(*ubiquitin:Cre reporter)* WKM cells (Appendix 6A). Furthermore, in the double transgenic *kdrl:Cre;ubiquitin-Cre reporter* animals, a mere 1.67% eGFP+ cells could be identified. Thus, while the flow cytometry data provide evidence for recombination in WKM cells the expression levels of mCherry and eGFP showed no improvement relative to the *ubiquitin:Cre reporter*. Analysis of adult HCs by PCR was inconclusive due to the large number of bands observed on the agarose gel (Appendix 6B). As the *ubiquitin:Cre reporter* transgene seemed to display weak reporter gene expression, both in the embryo and in the adult, it was replaced with the *heat-shock70* (hsp70) promoter (Thummel et al., 2005). When fish carrying the *hsp70*-Cre reporter transgene were crossed to *Tg(kdrl:Cre)* transgenic fish and the embryos obtained were heat-shocked at 37°C for 30 minutes, reporter gene expression was very strong (Appendix 7). Moreover, unlike with the other two Cre reporter lines (*ef1-α* and *Ubiquitin*) in which the expression of the second fluorescent protein gene could be observed only by 4-5dpf, with the *kdrl:Cre;hsp70 Cre reporter* transgenics, the expression of the second reporter gene (i.e, eGFP) was already detected by 3 dpf (Appendix 7) in *flk1/kdrl*-derived cells. Although the red-to-green systems seemed to give strong expression of eGFP in recombined cells in the embryo, we had problems detecting both mCherry and eGFP positive cells by flow cytometry in the adult kidney. It was therefore decided to discontinue using these *Cre reporter* lines. Subsequent experiments were carried out using the *ef1α:Cre reporter* line.

3.5 Adult blood cells originate from embryonic endothelial cells.

While the data on the Tg(*flk1/kdrl:cre*)^{qmc101}; Tg(*ef1-a:Cre reporter)* double transgenic adult fish suggested that blood cells in the adult WKM were derived from *flk1/kdrl*+ ECs, it remained uncertain whether the blood cells arose in the embryo or in the adult. To address this question, Tg(*flk1/kdrl:CreER*^{T2})^{qmc106}, Tg(*ef1α*:Cre reporter) embryos were treated 4-OHT from 10hpf-5dpf. Embryos in which dsRed expression was observed in early larval HCs (Figure 3.6) were raised to adulthood in order to analyse the adult WKM for the presence of the recombined reporter gene. To this end, PCR was carried out on gDNA prepared from adult WKM cells. WKM cells from fish derived from 4-OHT treated, eGFP negative (wt or single transgenic *flk1/kdrl:CreER*^{T2} fish) and eGFP positive only (single transgenic *ef1α Cre reporter* fish) embryos were used as negative controls. Since we did not observe recombined Cre reporter gene expression in the early larvae in the absence of the ligand, we do not expect to detect recombination in the adult HCs of fish that were treated at early embryonic stages with the ethanol solvent only. Amplification was carried out on gDNA that was isolated from cells flushed out of the adult WKM using the same primers DB432 and DB433 (Figure 3.7A) to amplify the 1325bp and/or 307bp fragments. As expected, no fragments were amplified in the negative control experiments. A single 1325bp fragment was successfully amplified on *ef1α:Cre reporter* single transgenic gDNA samples. gDNA from 10 double transgenic fish was isolated and analysed for the recombined reporter gene. The PCR revealed the presence of the recombined *ef1α:Cre reporter* gene in 3/10 fish that were analysed (Figure 3.9). The 307bp band was weak in one of the samples (Figure 3.8B lane number 8).

To verify the sequence of the 307bp and 1325bp fragments, the bands from the sample in lane 6 (Figure 3.9A) were isolated, purified and sequenced. Primer DB432 was used in the

sequencing reaction (Figure 3.7A). The sequence obtained for the 1325bp fragments confirmed that the non-recombined Cre reporter gene had been amplified and sequenced (Appendix 8). The sequence of small fragments matched the expected sequence for the recombined *ef1α:Cre reporter* gene (Appendix 8). The sequence corresponded to the 3' end of the *ef1-α* promoter sequence, the LoxP sequence and the beginning of the dsRed open reading frame (Appendix 8). The absence of the *eGFP* sequence in the 307 bp fragment confirmed successful recombination of the *ef1α:Cre reporter* gene. Thus, while the recombination efficiency may have been low, the presence of the recombined *ef1α* Cre reporter gene in adult WKM HCs suggests that the adult WKM HCs originated from embryonic ECs.

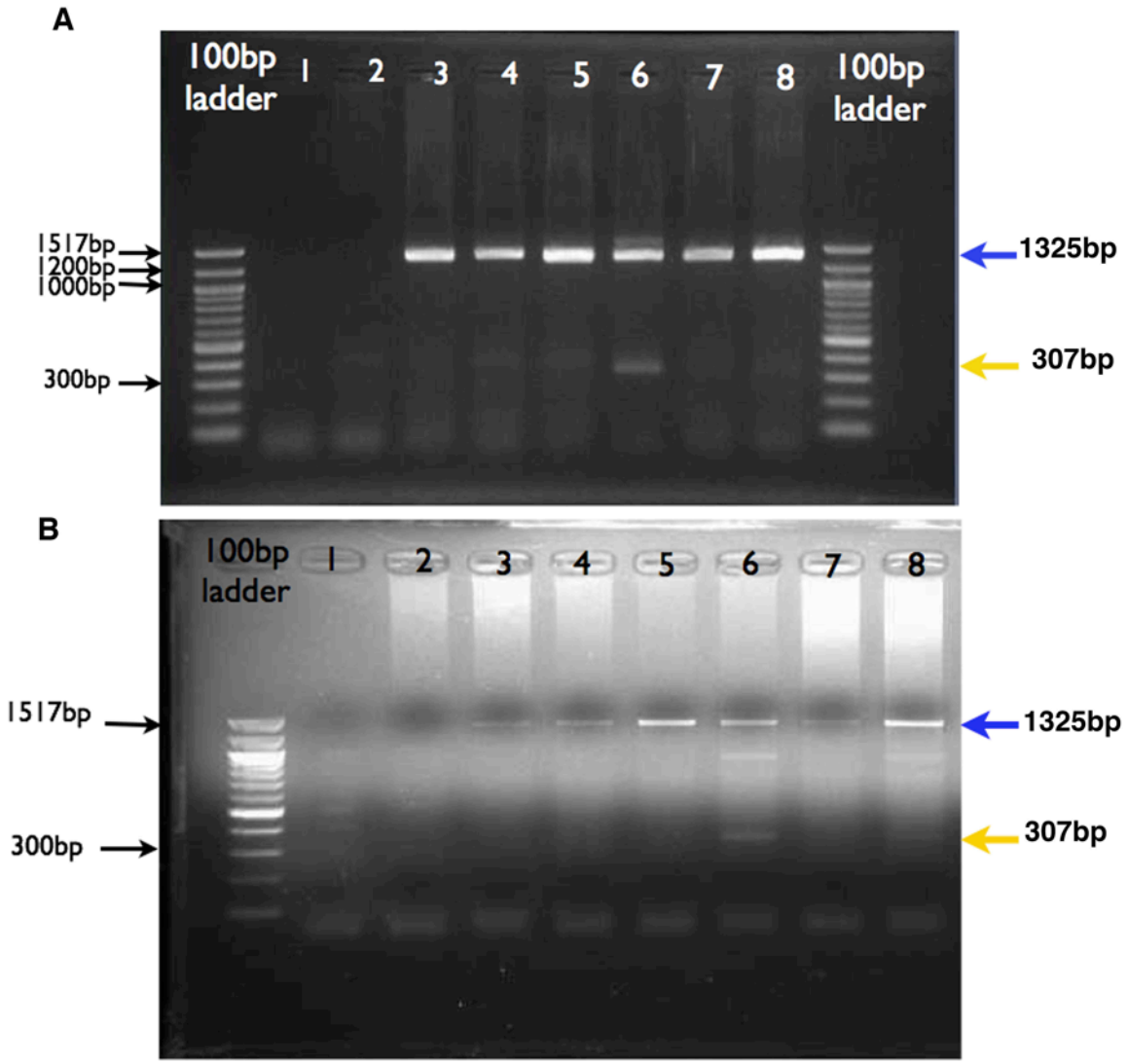

A
100bp ladder
1
2
3
4
5
6
7
8
100bp ladder
1517bp
1200bp
1000bp
300bp
1325bp
307bp
B
100bp ladder
1
2
3
4
5
6
7
8
1517bp
300bp
1325bp
307bp

Figure 3.9: *ef1α:Cre reporter* gene recombination detected by PCR in the kidney marrow cells isolated from *Tg(kdrl:CreER^{T2})^{qmc106}; Tg(ef1α:Cre reporter)* double transgenic fish that were treated with 4-hydroxy tamoxifen at embryonic stages.
Agarose gel electrophoresis carried out on PCR products (using the same pair of oligonucleotides shown in (Figure 3.7A) obtained from the genomic DNA of 10 *Tg(kdrl:CreER^{T2}; ef1α:Cre reporter)* fish. (A) Has a 100bp DNA ladder in the first and last lanes the well labelled 1 contained a water/no template control, well labelled 2 contained the PCR product obtained from gDNA of a wt fish, well numbered 3 contained the PCR product obtained from gDNA of an *ef1α Cre reporter* fish that displays the unrecombined fragment (indicated by the blue arrow). Lanes 4 to 8 contained PCR products obtained from gDNA of *Tg(kdrl:CreER^{T2}; ef1α Cre reporter)* fish, all of the lanes contain the 1325bp band. We see a second band i.e. at approximately the 307bp (recombined fragment) position (ochre arrow) in lane number 6. Wells 4,5,7 and 8 do not show a band at the 307bp position. (B) Contained a 100bp DNA ladder in the first well the well labelled 1 contained a water/no template control, well labelled 2 contained the PCR product obtained from gDNA of a wt fish, well numbered 3 contained the PCR product obtained from gDNA of an ef1α Cre Reporter fish where we see the unrecombined fragment (indicated by the blue arrow). Lanes 4 to 8 contained PCR products obtained from gDNA of *Tg(kdrl:CreER^{T2}; ef1α Cre reporter)* fish, all of the lanes contain the 1325bp band. We see a second band i.e. at approximately the 307bp (recombined fragment) position (ochre arrow) in lane number 6 and a weak band in this position in Lane number 8. Wells 4,5 and 7 do not show a band at the 307bp position, indicating the absence of recombination in these samples. Genomic DNA for all samples was obtained from blood cells flushed out of the adult fish kidney marrow.

3.6 Summary

- Here we show in 4-5dpf embryos, which express a constitutive Cre that the progeny of *kdrl*+ ECs do indeed contribute to blood cells in the CHT and in the larval kidney. These cells were still detectable in the adult WKM.
- The inducible CreER^{T2} system allowed us to label *kdrl*+ cells and its progeny, both in the early larvae and in the adult WKM. Moreover, we were also able to show the presence of recombined reporter gene expression both in early larval and adult HCs. These results strongly indicate that larval and adult HCs are derived from embryonic ECs.

4 Confocal imaging reveals qmc551:eGFP expression in HECs of the vDA that undergo both basal EMT and apical departure

4.1 Introduction

As part of the lineage tracing experiments single cell resolution (1-2µm) time-lapse confocal imaging was carried out to image and study cell behaviour in the vDA of the early zebrafish embryo. Initially, cell behaviour was investigated using a double transgenic line, *Tg(flk1/kdrl:eGFP)*^{s843}*;(csl:cerulean)*^{qmc63}. In *flk1/kdrl:eGFP;csl:cerulean* double transgenic embryos *flk1/kdrl:eGFP)*^{s843} is expressed in both arterial and venous ECs. The second transgene, *Tg(csl:cerulean)*^{qmc63} expresses cerulean in all cells that receive a Notch signal, which includes arterial ECs of the vDA. This double transgenic line was utilised to (a) establish the correct experimental conditions for the time-lapse in our lab and (b) once the right conditions for the time-lapse were achieved, the established method was used to explore more into the *qmc551* line at the cellular level. The double transgenic line *Tg(flk1/kdrl:eGFP)*^{s843}*;(csl:cerulean)*^{qmc63} was used as a control to confirm cellular events observed in the *qmc551* line.

So far it was shown qmc551:eGFP is expressed in spindle-shaped ECs of the vDA and round prRBCs cells (Figure 1.10) (Thambyrajah et al., 2016). If the eGFP+ spindle-shaped cells of the vDA are HECs, then they should potentially undergo EHT (Kissa and Herbomel, 2010) or bEMT. If this would be the case, then the *qmc551* line would be a valuable tool that could be used to study cellular and molecular events of HECs and their progeny during early vertebrate development. Using the *qmc551* line and other transgenic lines we have discovered a rare but novel form of EMT, which differentially relies on the

transcription factor *runx1,* but with a mechanism similar to the budding EHT mechanism seen in the mouse (Boisset et al., 2010).

4.2 Cells that undergo EHT are arterial endothelial cells that express flk1/kdrl:eGFP and the notch reporter csl:cerulean

Studies have shown that blood flow is required for the emergence of HSCs in the early embryo (Adamo et al., 2009; reviewed in Clements and Traver, 2013; North et al., 2009). The anaesthetic (Tricaine) slows down blood circulation at higher concentrations. Thus using a lower concentration of the anaesthetic would allow blood flow within the embryo to remain relatively normal. Using lower levels of the anaesthetic, the *in vivo* imaging of the vDA was carried out under conditions of comparatively normal blood flow. EHT or bEMT is known to occur from 34 to 60hpf, with the optimum number of events occurring from 40-52hpf (Kissa and Herbomel, 2010). Thus it was decided to begin imaging the vDA of a double transgenic *Tg(flk1/kdrl:eGFP)*^{s843}*;Tg(csl:cerulean)*^{qmc63} embryo from 31 hpf. For imaging, an embryo at this time point it was important to (a) immobilise it but still (b) have blood circulation to remain relatively normal and (c) allowing for the embryo to be able to grow out during the imaging. Hence keeping in mind these points we came up with a method in which we treated embryos with a chemical cocktail, which contained a neuromuscular blocker, tubocurarine (20μM) along with the anaesthetic, Tricaine (130μg/μl) in E3 medium. The use of this chemical cocktail ensured the twitching of the embryo was minimised. Embryos were then embedded in 1% low melting point agarose, allowing the embryo to grow during the imaging. Finally, the embryo was overlaid with the above chemical cocktail, which ensured that the embryo remained anaesthetized and immobile during the entire duration of the imaging. Embryos seemed to have blood circulation, and a heart beat at 3 dpf when the experiment was terminated. Kissa and Herbomel used a 160 μg/μl concentration of the anaesthetic, Tricaine, (Kissa and

Herbomel, 2010) in their experiments which could have had a possible effect on blood flow and thereby in the process of EHT. Once the embryo was immobilised and ready to image, next, the settings on the confocal microscope were determined. Kissa and Herbomel (2010) took z-stacks, where they imaged every 5-6 minutes over a period of 108 hours (Kissa and Herbomel, 2010). Imaging every 5 to 6 minutes might be a long interval, which may cause us to lose out on capturing important events (EHT). Hence we reduced our settings to image every 3 minutes and imaged over a period of 72 hours. However, reducing the time intervals introduced the risk of too much exposure to the laser, which caused bleaching of the fluorescent proteins. To avoid bleaching, we decided not to take z-stacks but to instead image a single 1-2µm optical sagittal section at 5 to 6 different positions going laterally through the embryo from anterior to posterior. Movement of the vessel tissue due to pulsating blood flow made early imaging of the embryo difficult. As the embryo grew over the duration of the imaging, there was a tendency of it going out of focus of the region we set out to image. The data generated using the double transgenic line, *Tg(flk1/kdrl:eGFP;csl:cer)* confirmed the EHT event shown by Kissa and Herbomel and Bertrand et al. (Bertrand et al., 2010; Kissa and Herbomel, 2010). More precisely, a double positive ($eGFP^+$ + cer^+) spindle-shaped cell was observed that was part of the endothelium lining the floor of the DA (Figure 4.1(i-iii)), which slowly bends, creating a gap in the endothelium above (Figure 4.1(iv-xiv)). The cell then rounds up as it ultimately ingresses into the mesenchyme below and the gap that was created then closes above the newly formed HC (Figure 4.1(v-xx)) and (Supplementary Movie 4.1) (Thambyrajah et al., 2016) published Movie 4 (encircled in red). The process took about 3.5 hours. Apart from confirming the EHT event described by the other labs we were also able to show that the cells that underwent the transition were not only *flk1/kdrl*+ ECs but these cells are notch

expressing arterial ECs of the vDA giving more confidence in the data generated that the observed cells are most likely to be HECs.

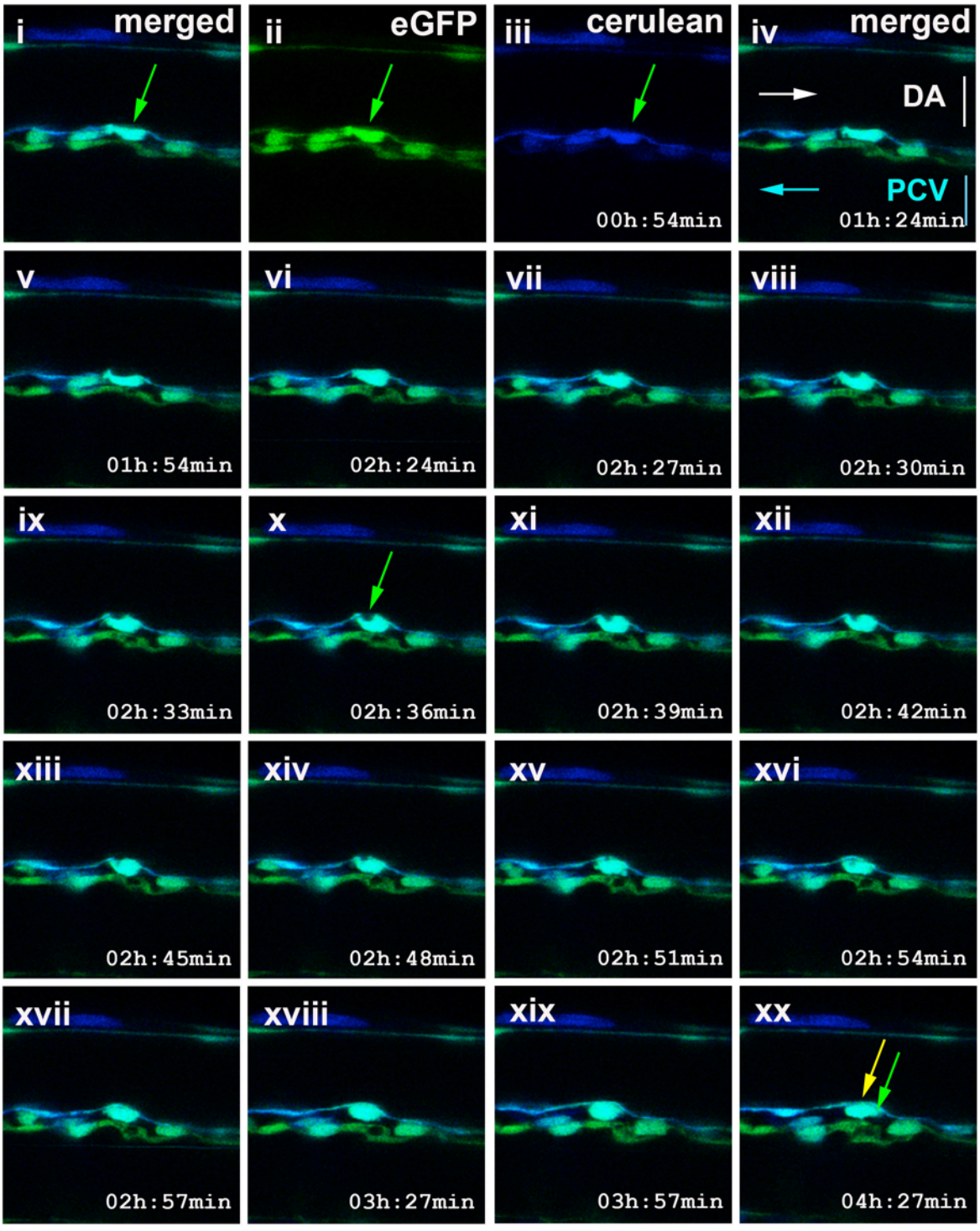

i merged
ii eGFP
iii cerulean
iv merged
DA
PCV
00h:54min
01h:24min
v
vi
vii
viii
01h:54min
02h:24min
02h:27min
02h:30min
ix
x
xi
xii
02h:33min
02h:36min
02h:39min
02h:42min
xiii
xiv
xv
xvi
02h:45min
02h:48min
02h:51min
02h:54min
xvii
xviii
xix
xx
02h:57min
03h:27min
03h:57min
04h:27min

Figure 4.1: *flk1/kdrl:eGFP;csl:cerulean* labels cells of the vDA that undergo EHT that gives rise to HSCs that remain in the mesenchyme below.

The above image series was taken from a time-lapse confocal movie of the trunk region of a double transgenic embryo starting from 48 hpf in which the embryo was placed laterally with anterior facing left and dorsal side up. Optical sections were 1.5μm thick and images were taken at 3-minute intervals. Endothelial cells of the DA co-express both reporter transgenes (i-iv) an elongated spindle shaped cell that expresses both (i) eGFP and cerulean as seen in individual channels (ii) shows the cell is green (eGFP+) and (iii) blue (cerulean+) (iv) highlights the DA and the PCV and shows the direction of blood circulation in the these two major vessels. (v-xiv) shows the cell as it bends and transitions to enter into the mesenchyme as a gap is created in the endothelium above. (xv-xx) As the cell is completely part of the mesenchyme the endothelium above it closes up (xx) (yellow arrow). This event was seen to take up to 3.5 hours. Green arrows: the cell that is undergoing the transition, white arrow: direction of blood flow in the DA, turquoise arrow: direction of blood flow in the PCV.

Once in the mesenchyme, the cells seem to remain there for varying amounts of time as seen from the variable intensity of green and blue fluorescence before they joined circulation via the PCV. Cells that had only just undergone bEMT seemed to express both eGFP and cerulean to a relatively high intensity as compared to cells that stayed in the mesenchyme. The expression of either eGFP or cerulean appeared variably reduced, indicating their longer stay within the mesenchyme. As cells exit the mesenchyme to join circulation via the vein, they seem to have residual blue or green intensity. In this particular example, the cell (in Figure 4.2, Supplementary Movie 4.2) appears to have residual eGFP expression. The weak eGFP expression indicates that the cell may have stayed within the mesenchyme for a long time before squeezing out via the dorsal wall of the vein to join the circulation. This event was observed to take place over a period of 36 minutes (in Figure 4.2, Supplementary Movie 4.2). Cells were seen entering the vein via different routes, where one cell was captured emerging from the dorsal wall of the vein to join circulation, and another was seen to crawl around the vein before entering the vein through the ventral wall of the vein. Several other cells were seen, trapped within a venous endothelial pocket, more like the "endothelial cuddling" in the CHT as described before (Tamplin et al., 2015). The cells are then seen to find their way out of the pocket and eventually get released into the vein to join the circulation ((Thambyrajah et al., 2016) published Movie 4).

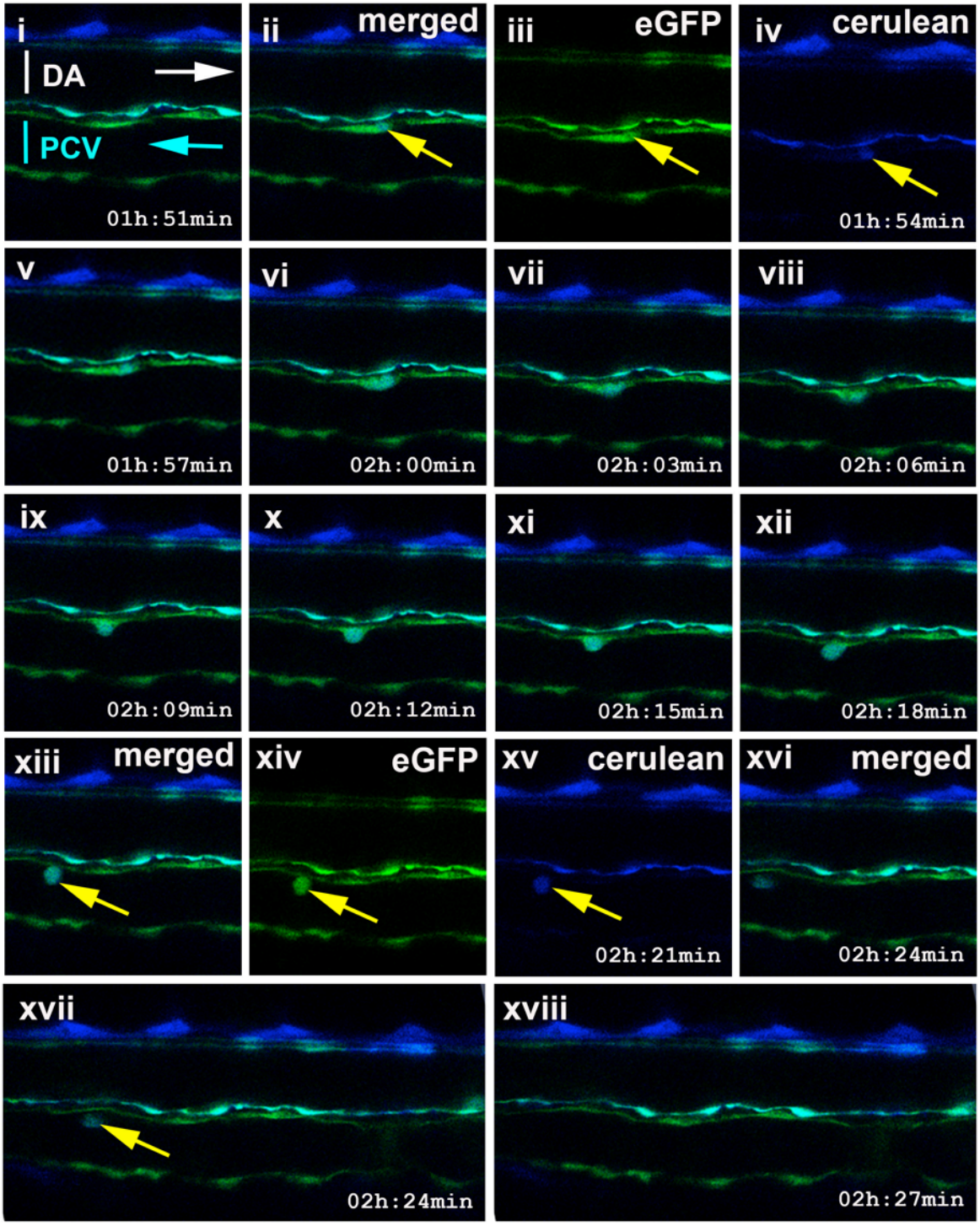

i
DA
PCV
01h:51min
ii merged
iii eGFP
iv cerulean
01h:54min
v
01h:57min
vi
02h:00min
vii
02h:03min
viii
02h:06min
ix
02h:09min
x
02h:12min
xi
02h:15min
xii
02h:18min
xiii merged
xiv eGFP
xv cerulean
02h:21min
xvi merged
02h:24min
xvii
02h:24min
xviii
02h:27min

Figure 4.2: *flk1/kdrl:eGFP;csl:Cerulean* expressing cells leave the mesenchyme to join circulation via the PCV.

The above image series was taken from a time-lapse confocal movie taken of the trunk region of a double transgenic embryo in which the embryo was placed laterally with anterior facing left and dorsal side up. The time-lapse was started from 48 hpf. Optical sections were 1.5µm thick and images were taken at 3-minute intervals. Endothelial cells of the DA and cells derived from ECs of the vDA through an EHT co-express both reporter transgenes. (i) shows the DA that co-expresses both reporter transgenes. The white arrow indicates the direction of blood flow in the DA. eGFP is expressed only in the PCV endothelial cells. The turquoise arrow indicates the direction of blood flow in the PCV. (ii) A cell in the mesenchyme is seen to co-express both (iii) eGFP and (iv) cerulean (v-xii) the cell is seen to squeeze its way out of the dorsal wall of the vein and (xiii) exits, still co-expressing both transgenes as seen in (xiv) the cell is eGFP+ and (xv) Cer+ and finally (xvi-xvii) to join the circulation via the PCV and (xviii) can no longer be seen. Yellow arrows: double positive cell in the mesenchyme that enters the PCV to join the circulation.

4.3 Qmc551:eGFP labels haemogenic endothelial cells of the vDA that undergo EHT/bEMT

Once we had established the experimental conditions for the time-lapse and had captured EHT events using the *Tg(flk1/kdrl:eGFP;csl:cerulean)* double transgenic embryos, we next wanted to study more about the behaviour of the eGFP+ cells that make up the vDA in the *qmc551* embryos.

For an EC to be characterised as being part of the specialised population with HEC, they should have the ability to undergo an EHT or bEMT. Hence, to show that the *gfi1aa*^{qmc551G} HECs that undergo bEMT, time-lapse experiments were carried out on 31 hpf *qmc551:eGFP;(flk1/kdrl:tdTomato)*^{qmc64} double transgenic embryos. As we know, *flk1/kdrl* is expressed in all ECs of both arteries and veins. Thus as the *Tg(flk1/kdrl:tTtom)*^{qmc64} line labels ECs of the DA it also outlines the PCV. A *qmc551:eGFP;(flk1/kdrl:tdTomato)*^{qmc64} double positive, i.e. a yellow spindle-shaped cell was seen lining the floor of the DA (Figure 4.3(i-iv) supplementary movie 4.3). The cell was seen to slowly bend creating a gap in the endothelium above as it gradually rounds up to enter the mesenchyme below (Figure 4.3 v-ix). Once in the mesenchyme, the gap in the endothelium closes above the newly formed HSPC. Throughout the event, the cell expresses eGFP and tdTom (Figure 4.3 x-xvi). This event was seen to occur over about 5 hours into the time-lapse experiment.

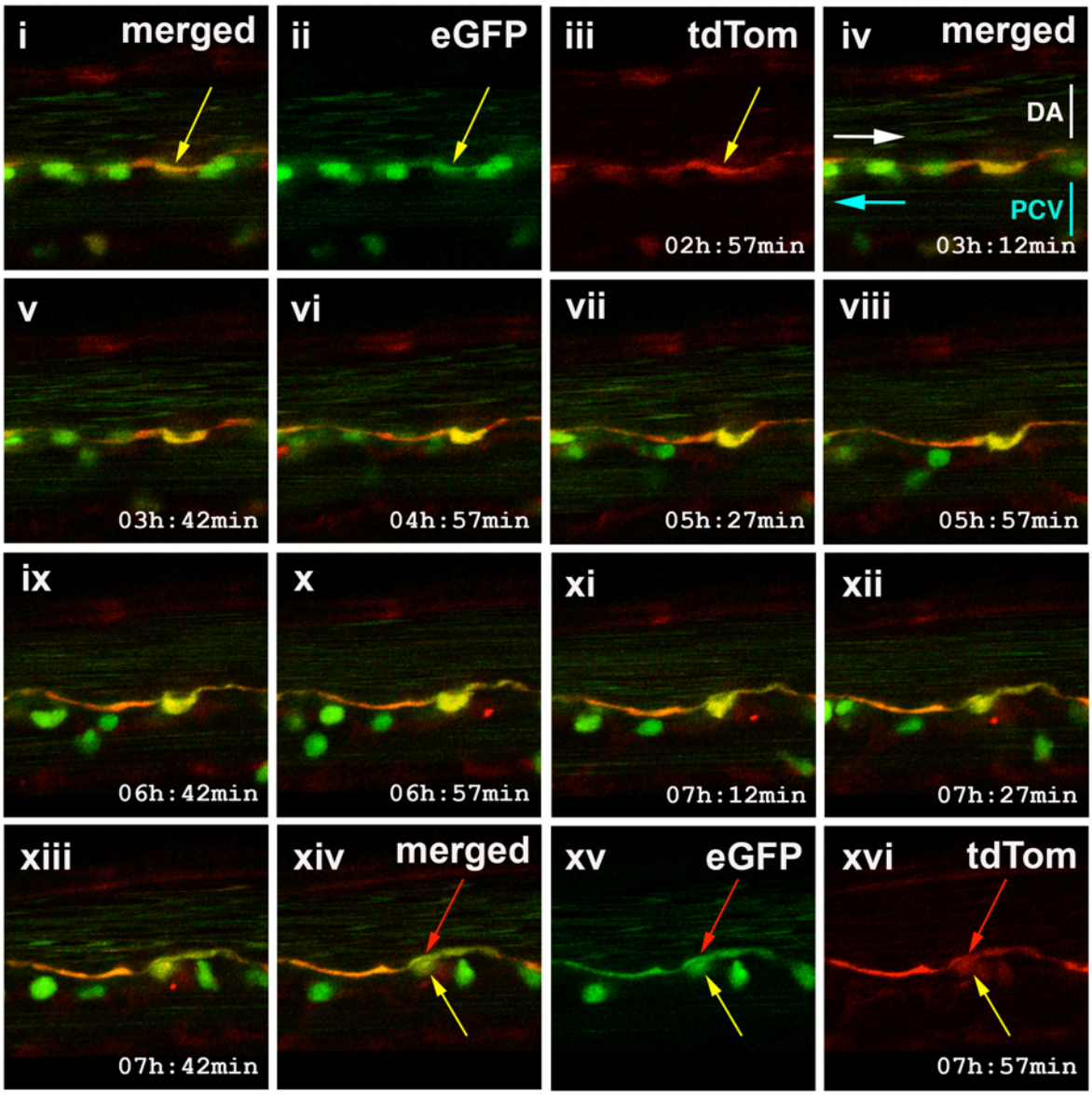


Figure 4.3: qmc551:eGFP+ cells of the vDA undergo basal EMT to give rise to blood cells that enter the mesenchyme below.
The above image series was taken from a time-lapse confocal movie performed on the vDA region of a 31 hpf *qmc551:eGFP;flk1/kdrl:tdTom* double transgenic embryo. The embryo was placed laterally with anterior facing left and dorsal side up. Optical sections were 2.1µm thick and images were taken at 3-minute intervals. (i-iv) highlights an elongated cell of the vDA, which co-expresses both eGFP and tdTom (v-xii) the cell slowly bends and ingresses into the mesenchyme below (xiii-xvi) the cell is in the mesenchyme completely while the endothelium above it closes up and the cell continues to co-express both eGFP and tdTom. This process seems to take about 5 hours. Yellow arrow: double positive HEC in the vDA that undergoes EHT and the derived mesenchymal cell, red arrow: endothelium that closes above the hEC derived mesenchymal, cell, white arrow: direction of blood flow in the DA, turquoise arrow: direction of blood flow in the PCV.

Once cells had undergone EHT/bEMT, they then remained within the mesenchyme for variable amounts of time, as the colour intensity of the cells suggests. Some cells seem to

express residual amounts of eGFP while others appeared to be expressing it more strongly. These cells then made their way from the mesenchyme, to join the circulation via the PCV (Figure 4.4, Supplementary Movie 4.4). We have recently published another example of a cell entering circulation via the PCV using a *qmc551:eGFP;(flk1/kdrl:tdTom)*^{qmc64} double transgenic embryo (Thambyrajah et al., 2016).

These events of bEMT and cells joining the circulation via the PCV were seen in both double transgenic *flk1/kdrl:eGFP;csl:cerulean* and *qmc551:eGFP;flk1/kdrl:tdTom* embryos. The bEMT observed in the two double transgenic lines occurred from around 40 to 49 hpf which fits with the peak of EHT described by Kissa and Herbomel i.e. 40 to 52 hpf (Kissa and Herbomel, 2010). Hence these data from the two double transgenic lines show that with both lines we are labelling and observing arterial ECs that give rise to mesenchymal cells that join circulation via the vein. Cells of the vDA undergoing bEMT strongly suggest that the cells are HECs giving rise to HSPCs. We have also been able to show that with the *qmc551* line we are indeed labelling HECs that give rise to HCs that transiently occupy the subaortic space, and eventually join the circulation via the PCV, during this entire process the cells maintained eGFP expression. Thus the data obtained on the *qmc551* line strongly suggest that qmc551:eGFP expression marks HECs of the vDA prior to, during and after bEMT/EHT (Thambyrajah et al., 2016)

During the process of collecting data and observing the bEMT/EHT in the two transgenic lines described above, we came across a new observation, an event no one had described so far in the literature. The rest of this chapter deals with getting more insight into this novel and interesting process.

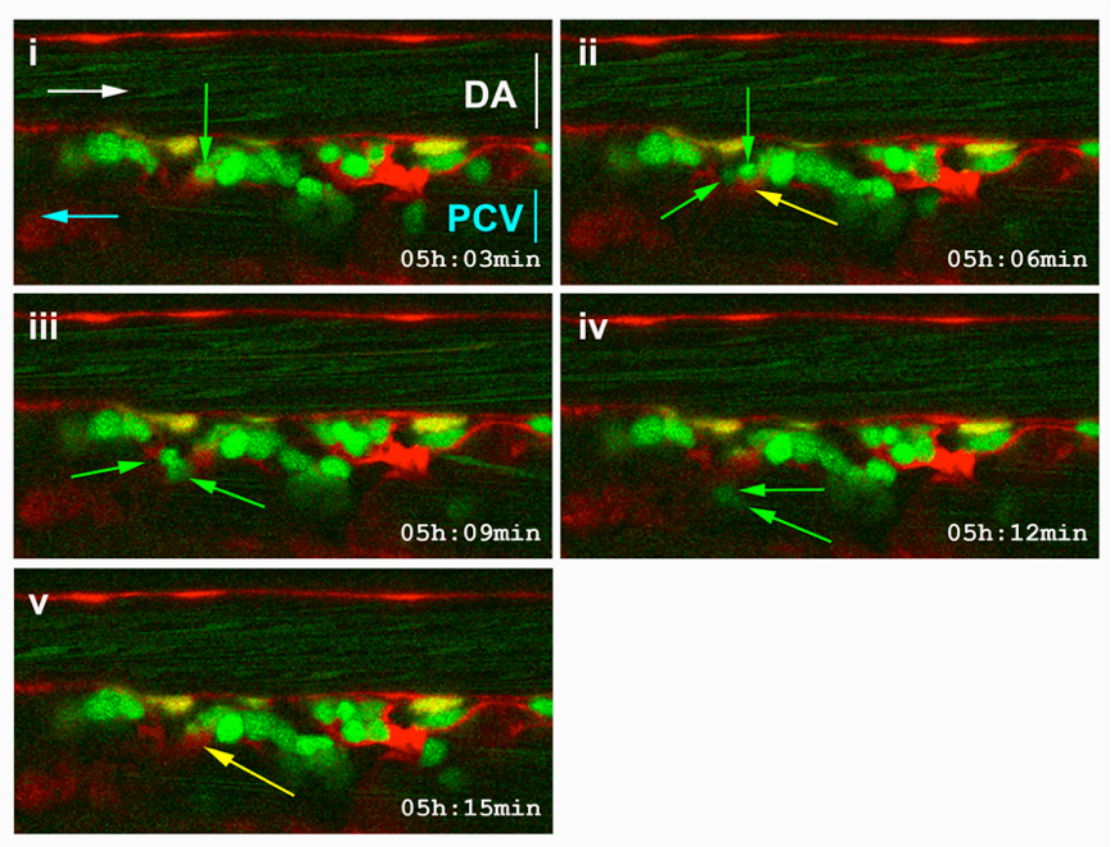


Figure 4.4: qmc551:eGFP;flk1/kdrl:tdTom double positive cells in the mesenchyme join the circulation via the PCV.

The above image series was taken from a time-lapse confocal movie taken of the trunk region of a double transgenic embryo. The embryo was placed laterally with anterior facing left and dorsal side up and the time-lapse experiment was started from 44 hpf. Optical sections were 1.5µm thick and images were taken every 3-minutes. Mesenchymal cells derived from the vDA through an EHT co-express both reporter transgenes. (i-v) demonstrates 2 cells, with varying eGFP intensity, leaving the mesenchyme and entering the vein to join circulation. Green arrows: The 2 mesenchymal cells that joined circulation via the PCV, yellow arrow: dorsal wall of the vein, white arrow: direction of blood flow in the DA, turquoise arrow: direction of blood flow in the PCV.

4.4 Other than bEMT/EHT cells of the vDA undergo apical departure

As we analysed the data obtained from the double transgenic *qmc551:eGFP;flk1/kdrl:tdTom* embryos to confirm that the double positive cells of the vDA do indeed undergo a bEMT/EHT and once within the mesenchyme these double positive cells also join circulation via the vein, we came across another interesting and novel observation. We demonstrated for the first time, in *qmc551:eGFP;flk/kdrl:tdTom* double transgenic embryos that cells emerge from the vDA enter circulation via the DA directly, a process we coined as, apical endothelial to mesenchymal transition (aEMT) or apical departure (aD). As observed from the time-lapse movies, this mechanism seemed more like a budding process rather than a gradual bending and ingression into the mesenchyme below. Here, a double positive cell within the vDA is seen to give rise to another round cell which seems to be associated with another twin cell (Figure 4.5, Movie 4.5 and (Thambyrajah et al., 2016). This event was seen to occur at an earlier time point, before 40hpf, i.e. before the onset of bEMT/EHT.

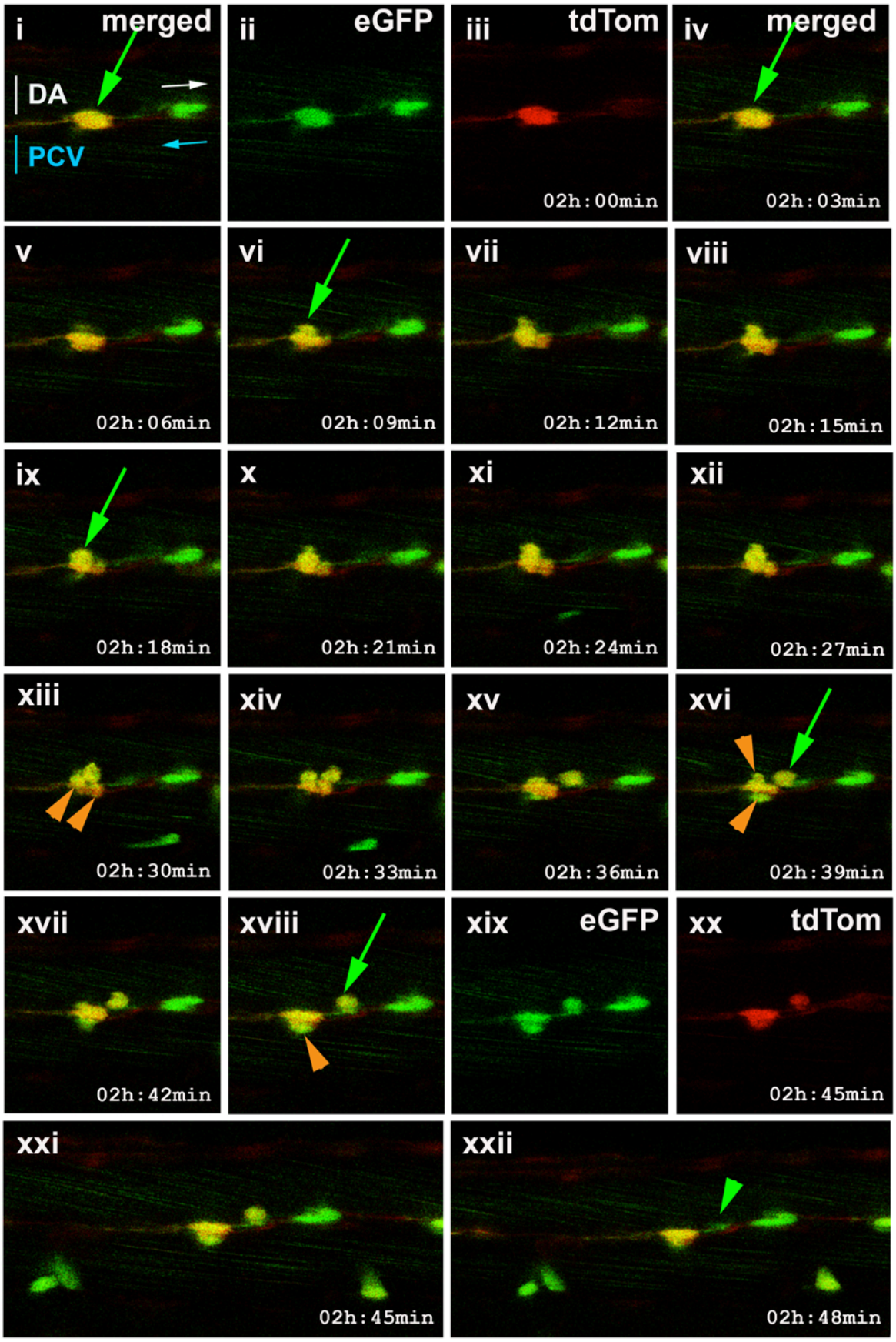

i merged
DA
PCV
ii eGFP
iii tdTom
02h:00min
iv merged
02h:03min
v
02h:06min
vi
02h:09min
vii
02h:12min
viii
02h:15min
ix
02h:18min
x
02h:21min
xi
02h:24min
xii
02h:27min
xiii
02h:30min
xiv
02h:33min
xv
02h:36min
xvi
02h:39min
xvii
02h:42min
xviii
xix eGFP
xx tdTom
02h:45min
xxi
02h:45min
xxii
02h:48min

Figure 4.5: qmc551:eGFP;kdrl:tdTom double positive cells leave the vDA apically as well.

The above image series was taken from a time-lapse confocal movie taken of the trunk region of a double transgenic embryo. The embryo was placed laterally with anterior facing left and dorsal side up. The time-lapse experiment was started from 33 hpf. Optical sections were 2.1μm thick and images were taken every 3-minutes. (i-v) At an early stage in development, an eGFP/Tomato-double positive cell in the ventral wall of the dorsal aorta of a qmc551; kdrl/kdrl:tomato embryo (vi-xii) gives rise to a cell that enters the lumen of the DA, along with this cell are 2 more cells (orange arrow heads) (xiii-xxi) the first cell migrates on the ventral wall and then (xxii) disappears into the circulation (green arrow head). Green arrow: double positive cell within the vDA and its derived cell that joins circulation via the DA, white arrow: direction of blood flow in the DA, turquoise arrow: direction of blood flow in the PCV.

To confirm this novel finding, we repeated the experiment on another *qmc551* embryo, this time without the *Tg(flk1/kdrl:tdTom)* in the background. Here, we were able to capture the event again, where an eGFP+ EC of the vDA was seen (Figure 4.6 i-iv) from which a round cell buds off above (Figure 4.6 v-xii) while another cell is seen below it (orange arrowhead in Figure 4.6 ix). The cell above detaches from the endothelium i.e. undergoes aD just around 40 hpf (Figure 4.6 xiii). The round cell then remains on the vDA and slowly migrates, eventually to join the circulation (Supplementary Movie 4.6).

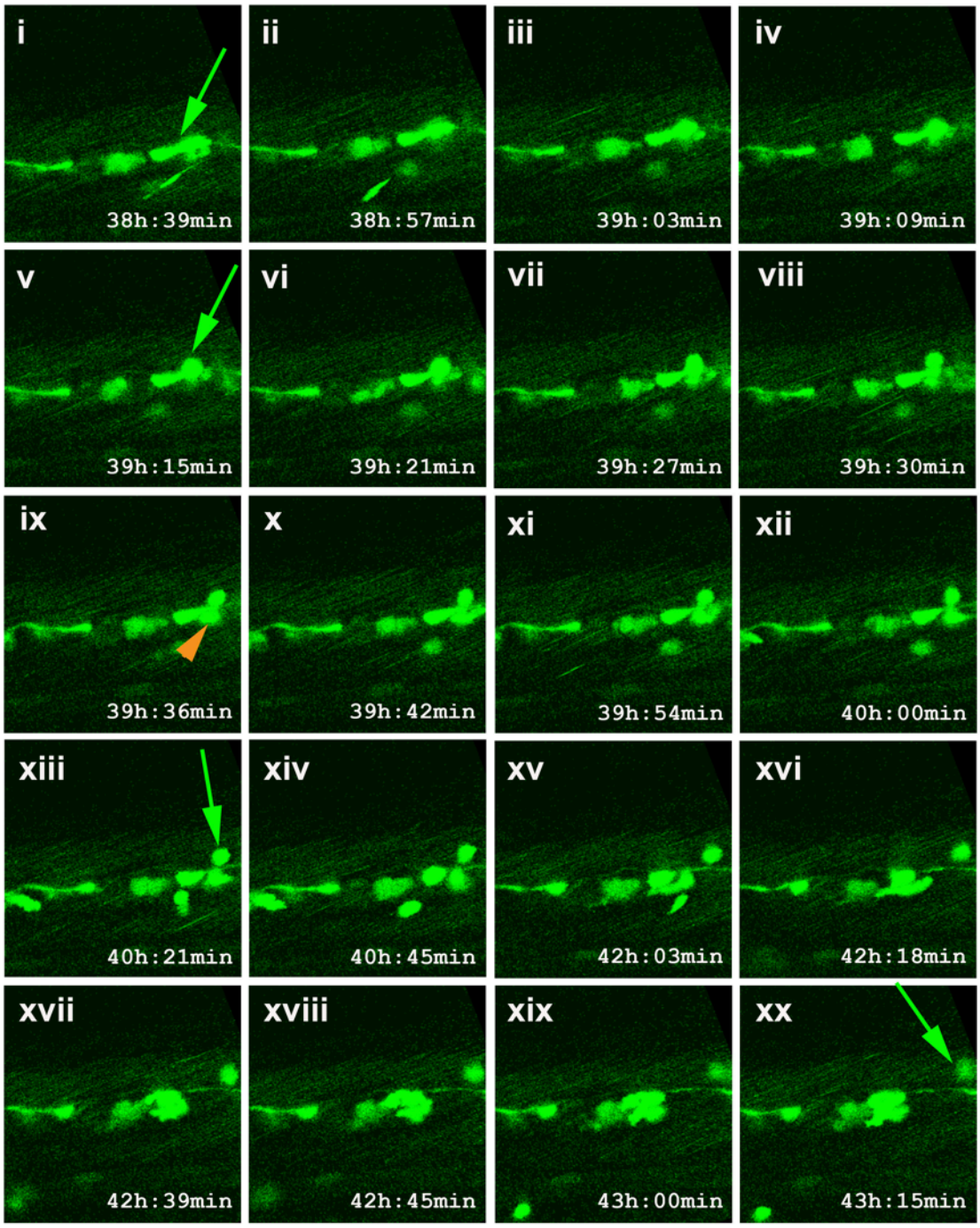


Figure 4.6: qmc551 eGFP+ endothelial cells of the vDA leave apically to join the circulation via the DA.

The above image series was taken from a time-lapse confocal movie taken of the trunk region of a qmc551 transgenic embryo. The embryo was placed laterally with anterior facing left and dorsal side up. The time-lapse experiment was started from 33 hpf. Optical sections were 1.5µm thick and images were taken every 3-minutes. (i-iv) An elongated cell of the vDA that expresses qmc551:eGFP buds off slowly at one end as it (v-xii) gives rise to a circular cell that seems to also be associated to another circular cell below (orange arrow head) and (xiii-xviii) remains over the endothelium of the vDA for a while before it (xix-xx) gradually seems to disappear into the circulation via the DA. Green arrow: hEC that gives rise to an HC, which joins the circulation via the DA. Timings are indicated on the right hand corner of each image.

Since the *qmc551* line is a gene trap line, we thought this event might be an artefact of the gene trap. To further investigate this possibility, we decided to image a *flk1/kdrl:eGFP;csl:cerulean* double transgenic embryo as a control. An apically departing cell was observed in the *flk1/kdrl:eGFP;csl:cerulean* double transgenic embryo as well. The event occurred by 35hpf and the apically departing cell was seen to be associated with another budding cell below the vDA (Figure 4.7i-xviii, Supplementary Movie 4.7).

All of the examples of apically departing cells we observed occurred in the early embryo just before or around 40hpf. This event has been observed only at earlier time points (34-40 hpf), probably just before the onset of bEMT/EHT. The fact that cells of the vDA undergo aD at earlier time points and bEMT later may indicate the presence of another factor that directs the cells to move either downwards towards the mesenchyme (at later time points), or earlier to move upwards into the DA. Along with this, we noticed that the budding of these cells from an EC of the vDA was always associated with another round cell below or alongside the cell that departed apically (indicated by orange arrowheads in (Figures 4.5 xiii, xvi and xviii, 4.6ix and 4.7v-vi). The presence of the twin cell suggests a sort of asymmetric cell division as described previously (Tamplin et al., 2015) where a cell divides, and one cell stays while the other cell goes into circulation.

Our data show that aD occurs in the early embryo, before 40hpf. Overall, our results suggest that although rare (observed ten times) in comparison to the bEMT/EHT, apically departing cells are indeed derived from ECs of the vDA as labelled by the flk1/kdrl in both the transgenes used. These cells are arterial cells that express Notch, and finally, these cells express qmc551:eGFP, strongly suggesting that it is the HECs of the vDA that undergo both bEMT as well as aEMT. The mechanism by which cells of the vDA depart apically was clearly distinct from the mechanism described for EHT or the bEMT, which involve bending of the HECs into the mesenchyme below. aD seemed more like a budding process.

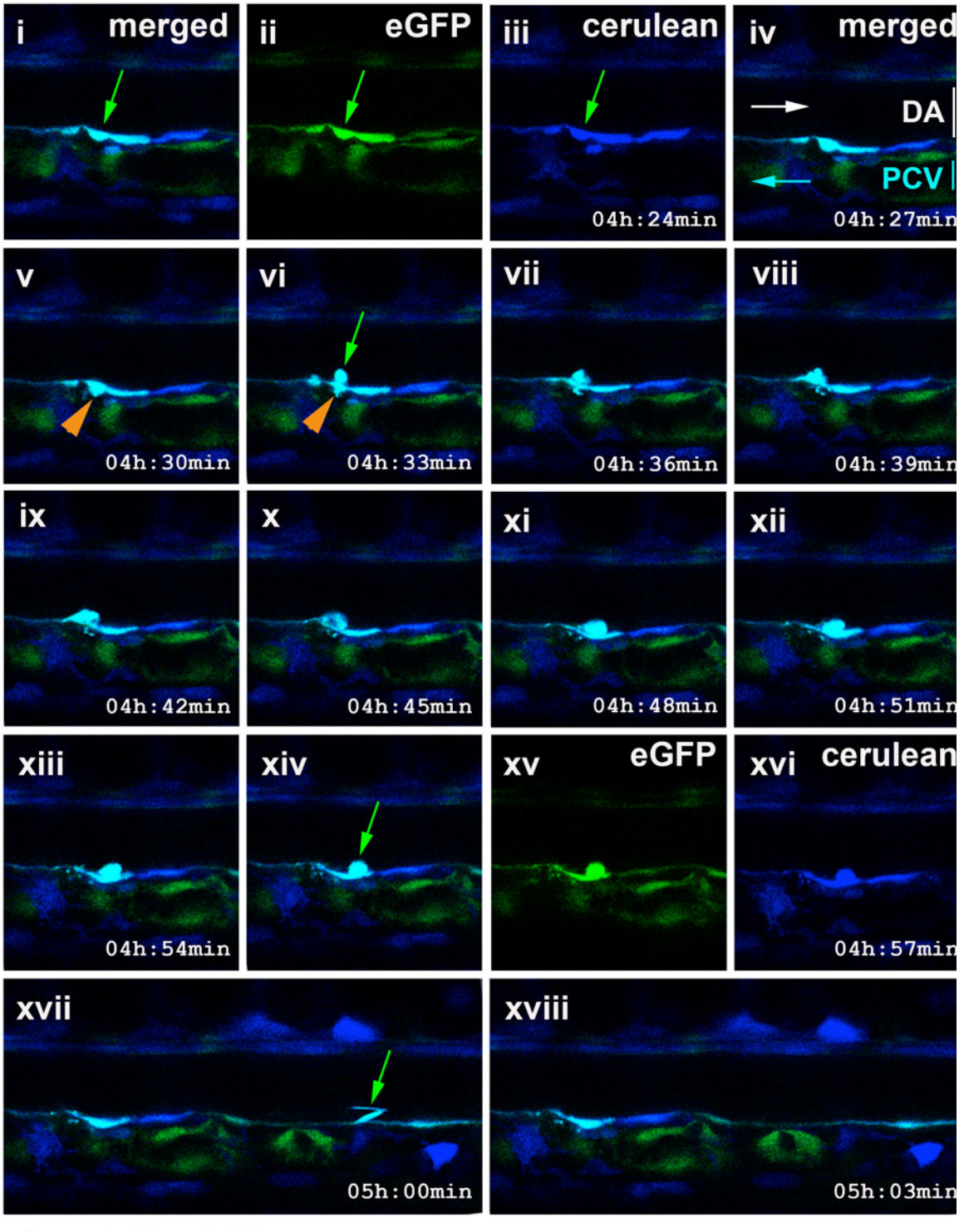

i merged
ii eGFP
iii cerulean
iv merged
DA
PCV
04h:24min
04h:27min
v
vi
vii
viii
04h:30min
04h:33min
04h:36min
04h:39min
ix
x
xi
xii
04h:42min
04h:45min
04h:48min
04h:51min
xiii
xiv
xv eGFP
xvi cerulean
04h:54min
04h:57min
xvii
xviii
05h:00min
05h:03min

Figure 4.7: *flk1/kdrl:eGFP;csl:cer* double positive cells of the vDA give rise to apically departing circular cells that are seen to join circulation directly via the DA.
The above image series was taken from a time-lapse confocal movie taken of the trunk region of a double transgenic embryo in which the embryo was placed laterally with anterior facing left and dorsal side up. The time-lapse was started from 30 hpf. Optical sections were 1.5µm thick and images were taken at 3-minute intervals. (i-iv) A double positive elongated cell is seen in the vDA, i.e. that expresses both (i) eGFP and (ii) cerulean which slowly (v) develops a small node below (orange arrow head) and (vi) and a budding cell begins to appear above as the node (orange arrowhead) continues to be visible (vii-xvi) the budding cell begins to grow as it continues to express both eGFP and cerulean, and (xvii) rolling over the vDA to finally (xviii) disappear into circulation. Green arrow: HEC that gives rise to an HC, which joins the circulation via the DA, white arrow: direction of blood flow in the DA, turquoise arrow: direction of blood flow in the PCV.

4.5 aD is independent of *runx1* and occurs more frequently, beyond 40hpf in *runx1* morphant embryos.

Previous time-lapse studies in zebrafish have highlighted the requirement of *runx1* in EHT and have described how in the *runx1* morphant embryos cells of the vDA attempt to undergo EHT but disintegrate or "explode" (Kissa and Herbomel, 2010; Zhen et al., 2013). What remains to be answered is what happens to the remaining HECs of the vDA and are we still able to see cells leave the vDA apically? Recently we have shown that qmc551:eGFP expression is independent of runx1 (Thambyrajah et al., 2016). For this purpose, Deniz Ucanok injected 2-4 cell stage qmc551:eGFP embryos with 3ng of *runx1* splice morpholino (Gering and Patient, 2005). The endothelium of the vDA and caudal artery (CA) remained eGFP+, suggesting that HECs are retained in the *runx1* morphants. However, there is a dramatic reduction of qmc551:eGFP^{+} HCs in the mesenchyme below the vDA (Figure 4.8 Ai-ii) as compared to an uninjected control embryo. Next, I carried out confocal imaging on morphant and control embryos. Z-stacks taken through the CHT of 50hpf *runx1* morphants revealed a dramatic reduction of qmc551:eGFP+ HCs as compared to uninjected control embryos (Figure 4.8 Aiii-iv). Next, to study cellular behaviour in the vDA, I carried out time-lapse imaging on the *qmc551* transgenic *runx1* morphant embryos. The confocal time lapse imaging data revealed that most HECs remained in the vDA. bEMT or EHT in these morphant embryos was almost entirely abrogated (observed only once). Instead, we observed many individual HECs that rounded upwards into the lumen of the DA, forming protrusions, and then budding off and vigorously moving and disappearing into circulation via the DA (observed thirteen times) (Figure 4.8 Bi-viii, Supplementary Movie 4.8 shows two cells). Here as well, apically departing cells were seen to be associated with small vesicles that were observed in the

mesenchyme below. Apically departing HECs were seen beyond 40hpf in the *runx1* morphants, but not in wt embryos (Section 4.4 Figures 4.5-4.7).

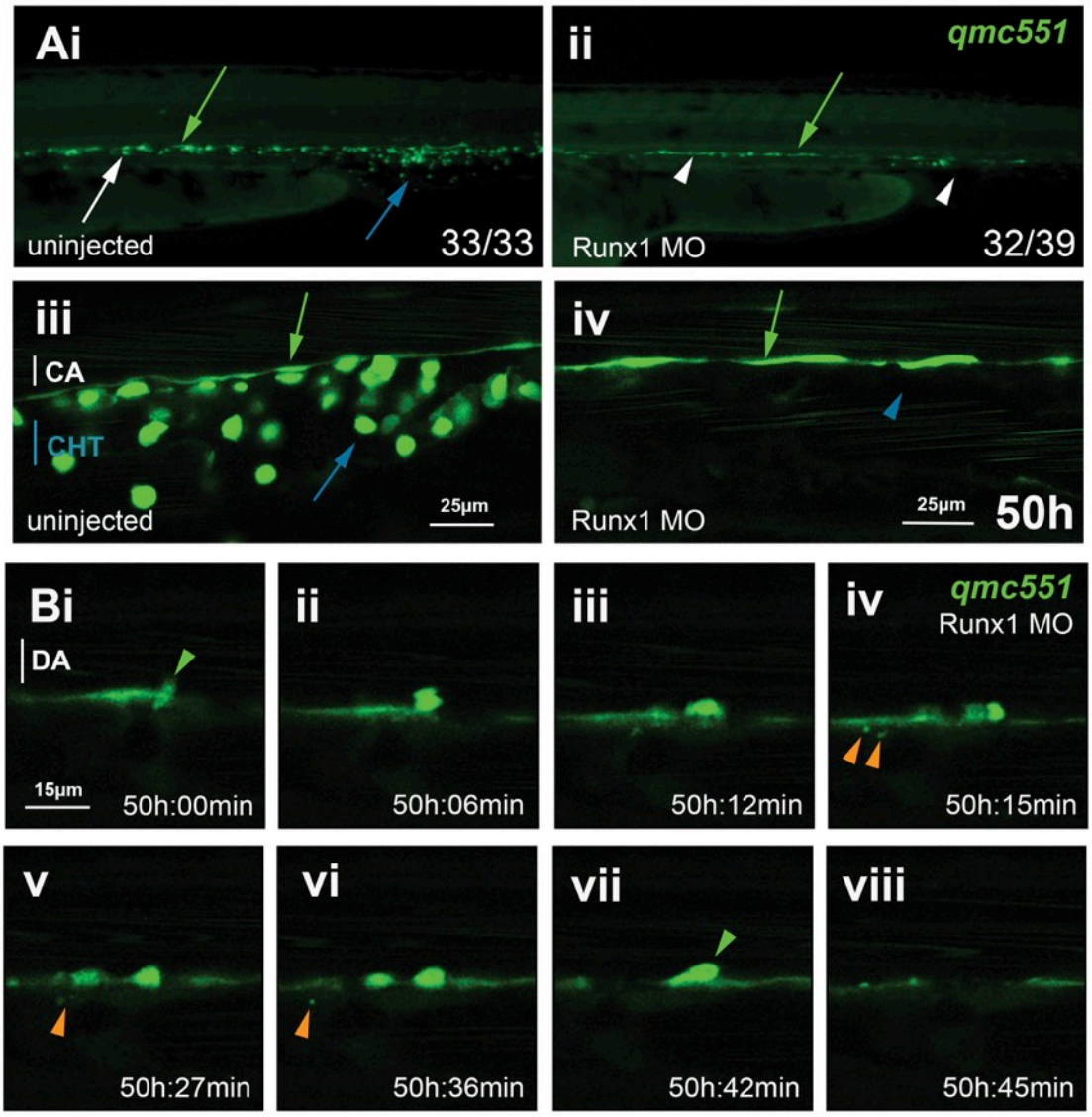


Figure 4.8: In qmc551; runx1 morphants bEMT of HECs of the vDA is almost completely abrogated while events of apically departing cells are still observed.
(Ai-ii) Images were taken on a dissecting fluorescent microscope and images (Aiii-iv and Bi-viii) are confocal images. In all images the embryos positioned with anterior to the left and dorsal side up. Ai-ii) Lateral views of the trunk and tail show HECs of the vDA that seem normal in both uninjected qmc551 and the *qmc551;runx1 morphant* embryos. Mesenchymal cells that underwent bEMT and HSPCs in the CHT in (Ai), while a reduction in mesenchymal cells and HSPCs in the CHT in the morphant (Aii). Aiii-iv) Z-stacks taken through the CHTs (2µm thick optical slices) of the same embryos shows the HECs of the CA and HSCPs in the CHT (Aiii) while in the morphant embryo (Aiv) HSPCs are almost completely abrogated in the CHT. (Bi-viii) Confocal time lapse series (2µm thick optical slices) through the trunk of a qmc551 runx1 morphant embryo shows an apically departing intra-aortic cell and mini vesicles green arrows: HECs, white arrow: Mesenchymal cells that underwent bEMT, blue arrow: HSPCs in the CHT white arrowhead: reduction in mesenchymal cells that underwent bEMT, blue arrowhead: abrogated HSPCs in the CHT, green arrowheads: apically departing intra-aortic cell, orange arrowheads: mini vesicles.

These data showed that although bEMT completely depends on *runx1*, aD of cells into the subaortic space is independent of the transcription factor *runx1*.

Besides the fact that aD differentially relies on *runx1*, another interesting observation with the aD cells, both in *wt* and *morphant* embryos was the time taken for these events to complete. It was observed that the mechanism of aD proceeded significantly faster (around 1hour) in comparison to the time taken for cells to undergo bEMT, which took approximately 5 hours to complete (Figure 4.9).

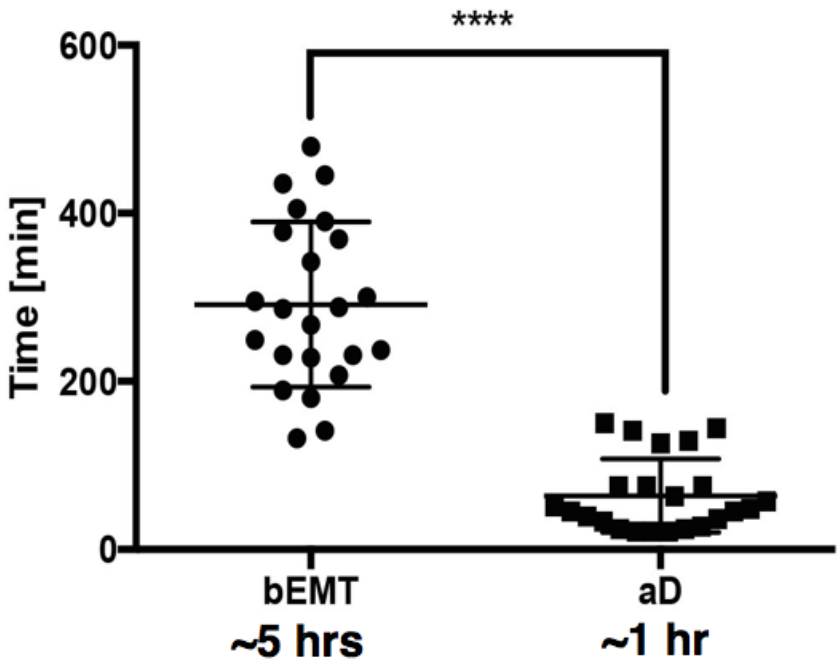


Figure 4.9: The time taken for HECs to depart apically is faster than the mechanism taken for EMT.

The times required for apical and basal departure were significantly different (Student t-test: $p<0.0001$). As the bEMT mechanism was seen to take an average of 5 hours to complete, aD proceeds much faster, ~1hour.

4.6 Apically departing cells have a nucleus and are large enough to be intact cells

Next, to investigate whether aD cells have a nucleus and whether they are large enough to be intact cells *qmc551:eGFP;fli1:nls-mCherry* double transgenic embryos were used. Confocal imaging of *runx1* morphant and uninjected control embryos that expressed a nuclear-localised mCherry in HECs showed that apically departing cells were nucleated (Figure 4.10A and B), demonstrating that these intra-aortic cells are intact cells.

Next, the sizes of the intra-aortic cells as they left the vDA and just before they joined circulation via the DA of both *runx1* morphants and the control embryos was measured (Figure 4.10A, B and Ci) (same cell as documented in Figure 4.7xiv). We also measured cells that had just undergone bEMT and had just entered the mesenchyme below the vDA (Figure 4.10Cii). Additionally, nuclei of HSPCs in the CHT were also measured (Figure 4.10Ciii). Comparing the cell sizes of the intra-aortic cells to the cells that had undergone bEMT and those that were in the CHT revealed that some of the intra-aortic cells had dimensions similar to HEC-derived mesenchymal cells or HSPCs in the CHT. Others were slightly smaller but were at least the size of a nucleus. Intra-aortic cells in *runx1* morphants were observed to be slightly smaller (Figure 4.10D). Since the cells that were measured were taken from single slices confocal images, there was a possibility that cells were being clipped, thus making them appear smaller.

Altogether these data demonstrate that HECs of the vDA undergo two types of transition, bEMT and an aD, which differentially rely on the transcription factor *runx1*. Interestingly, cells that undergo aD bud directly into the lumen of the DA from where they join circulation, just like HECs in the mammalian DA (Boisset et al., 2010).

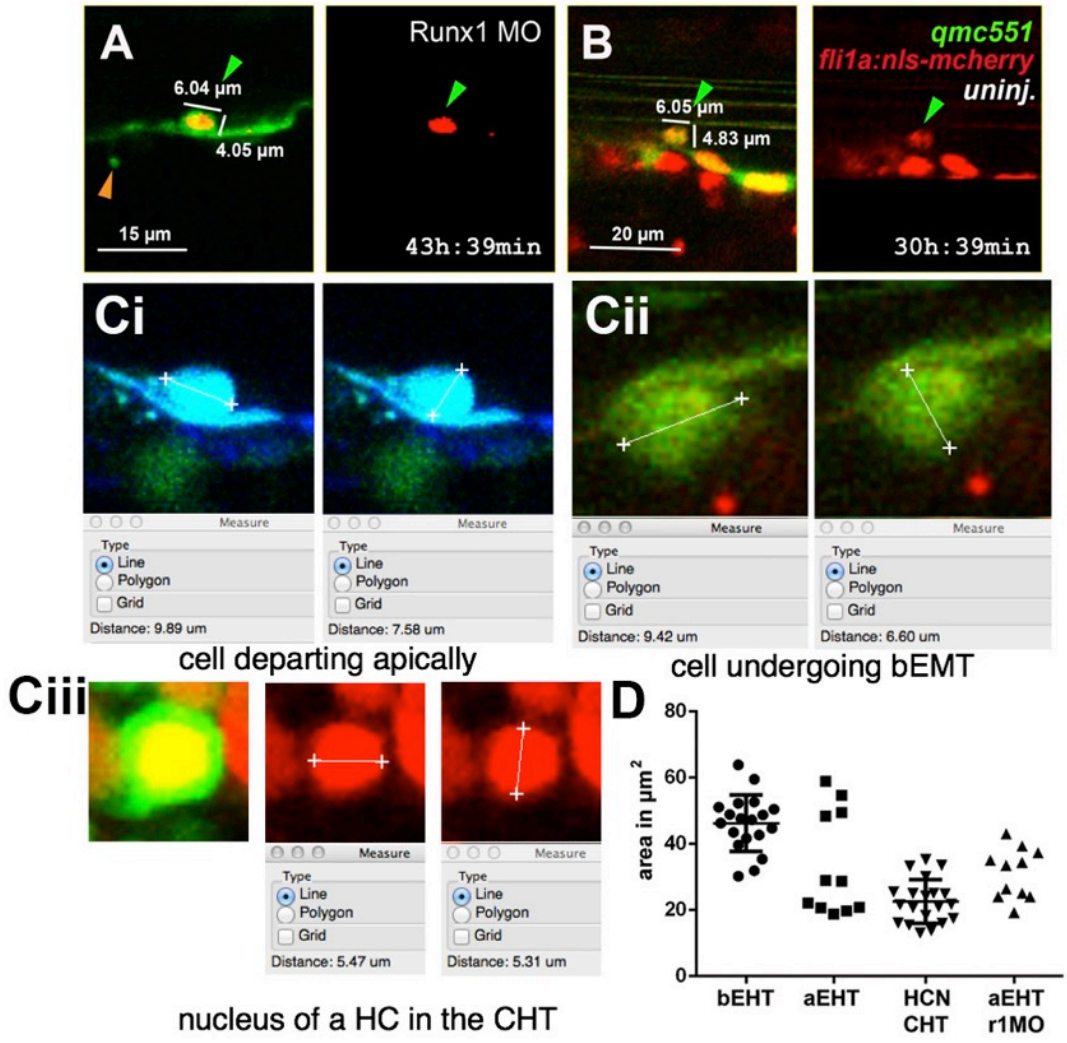


Figure 4.10: Cells that undergo aD have a nucleus and are intact cells.
(A-B) confocal images (1.5µm thick optical sections) through the DA of a qmc551; fli1a-nls.mCherry (A) runx1 morphant embryo and (B) an uninjected control embryo displaying eGFP-mCherry merged and mCherry only displaying aD in apically departing cells have a nucleus. (Ci-iii) measuring axes 'm' and 'n' of cells in Imaris (Bitplane) (Ci) intra-aortic cell in the *flk1/kdrl:eGFP;csl:cer* transgenic embryo (Cii) an HEC derived mesenchymal cell from a qmc551:eGFP;kdrl:tom embryo and measuring the axes 'm' and 'n' of the (Ciii) nucleus of an HSPC in the CHT of a *qmc551:eGFP;fli1nls:mCherry* transgenic embryo. (D) Comparison of cell sizes and nuclei sizes. Scatter plots are provided with mean and SD shown where size distribution is normal.

4.7 Summary

- We show that cells that undergo EHT are not only kdrl+ but also express the notch reporter.
- The *gfi1aa*^{qmc551Gt} line labels HECs of the vDA that undergo EHT and that also leave the vDA apically.
- Single-cell resolution confocal microscopy reveals a second and novel mode of EHT other than basal EMT, i.e. apical departure (aD) or aEMT of cells.
- The mechanism of aD is distinct from the bEMT process. Where aD is more a budding process rather than the bending of HECs, which is the bEMT mechanism.
- bEMT and aD differentially rely on the transcription factor *runx1*.
- aD is seen to occur only up to 40hpf in control embryos where as in runx1 morphants this process occurs even beyond 40hpf.
- The process of aD is significantly faster than bEMT, i.e. takes about 1hour where as bEMT takes an average of 5 hours to complete.
- aD cells have nuclei and some are as big as cells with in the mesenchyme or CHT or are at least the size of a nucleus.
- The aD of cells is reminiscent of EHT seen in the mouse.

5 Novel CRISPR/Cas9-generated *gfi1ab* mutant alleles suggest defects in definitive haematopoiesis in Gfi1aa and Gfi1ab depleted embryos.

5.1 Introduction

The *qmc551* homozygous fish survive up to adulthood and are fertile. In the absence of Gfi1aa expression its paralogue Gfi1ab is up regulated in the vDA (Ucanok, 2015). This strongly suggests that (a) Gfi1aa represses *gfi1ab* expression in the vDA and (b) Gfi1ab functionally replaces Gfi1aa during the onset of definitive haematopoiesis and possibly also during its maintenance. To search for further evidence for functional redundancy and to learn more about the function of the gfi1a genes in HEC and HSC formation we decided to generate a *gfi1ab* mutant using the CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR associated (Cas)) technology. We generated the mutant allele in the qmc551 line in order to be able to generate double mutants that could then be analyzed by live imaging. The idea was to use the qmc551:eGFP reporter to visualize HECs and their ability to undergo EHT in a single and double mutant background.

In order to generate *gfi1ab* mutants, the CRISPR/Cas9 system was employed. The system has evolved as bacterial and archaeal adaptive immune systems to protect these organisms from invading viruses and plasmids (Bhaya et al., 2011; Wiedenheft et al., 2012). Invading nucleic acids are destroyed by Cas9 nucleases that find their targets with the help of CRISPR RNAs. (Wiedenheft et al., 2012). The information of invading bacteriophage or plasmid DNA is saved in short DNA sequences called “protospacers” that are incorporated into the genome of the bacteria. These short DNA sequences serve as a sort of immune system “memory”. The sequences are expressed as short guide RNAs that help Cas9 find

its targets (Gupta and Musunuru, 2014; Wiedenheft et al., 2012). This system can be exploited as a genetic tool to carry out RNA-programmable genome editing (Jinek et al., 2012). One of the simplest CRISPR/Cas systems employed for genome-editing tool for heterologous expression of a CRISPR/Cas system is a TypeII CRISPR system from *Streptococcus pyogenes*; the system comprises a Cas9 endonuclease along with two small RNAs, namely a CRISPR RNA (crRNA) and a trans-activating crRNA (tracr RNA) complex, which are sufficient to drive RNA-guided cleavage of the invading nucleic acid (Barrangou et al., 2007; Brouns et al., 2008; Deltcheva et al., 2011). Instead of two separate RNAs a chimeric RNA (single guide RNA) can be used to guide the Cas9 to the Target Sites (TSs). TSs are characterised by a 20 bp sequence that is followed by an NGG nucleotide sequence also known as the protospacer-adjacent motif (PAM) (Cong et al., 2013; reviewed by Gupta and Musunuru, 2014; Jinek et al., 2012). Guided by the single guide RNA, which has incorporated into it the sequence complementary to the TS, the Cas9 nuclease generates a double stranded break (DSB) at the TS. The DSB is subsequently repaired by one of two different mechanisms in the cells. The most commonly used mechanism is non-homologous end joining (NHEJ). NHEJ is an error-prone mechanism, and it is this feature of the mechanism that can be exploited to efficiently introduce insertion/deletion (indel) mutations of various lengths, which can lead to a frame-shift in the coding sequence of a gene and to the introduction of a premature stop codon. This will cause the generation of a truncated, possibly non-functional protein. A premature stop codon may furthermore stimulate nonsense-mediated decay (NMD) of the mRNA, leading to a drastic reduction in protein abundance. The alternative repair mechanism is homology-directed repair (HDR). It requires the presence of a homologous sequence that can be used as a repair template. This mechanism can be used to insert

specific mutations through homologous recombination with the broken target locus (Bedell et al., 2012; Gupta and Musunuru, 2014; Sander and Joung, 2014).

Recent studies carried out in the zebrafish have successfully used the CRISPR/Cas9 system to introduce site-specific DSBs in several genomic loci (Chang et al., 2013; Hwang et al., 2013; Jao et al., 2013). Dr. Chris Moore, a post doc in our lab used the CRISPR/Cas9 system to target the zebrafish *gfi1ab* gene. The wt *gfi1ab* gene lies on chromosome 6 of the *Danio rerio* (zebrafish) genome and consists of six coding exons (Ensembl:ENSDART00000065139.4) (Figure 5.1 Ai). They encode a protein (UniProtKB F1QKD4-1) of 396 amino acids (aa), with two well-conserved protein domains (Dufourcq et al., 2004). One of them is the 20aa SNAG domain that mediates transcriptional repressor activity and is encoded by the 3' end of exon 1 (Figure 5.1 Aii). The other domain is the C-terminal DNA binding domain that is encoded by the 3' end of exon3 and exons4 to 6 and consists of six C-terminal C_2H_2 zinc fingers. The SNAG domain and the DNA binding domains are separated by a less conserved linker domain (Cooney et al., 2013) (Figure 5.1 Aii).

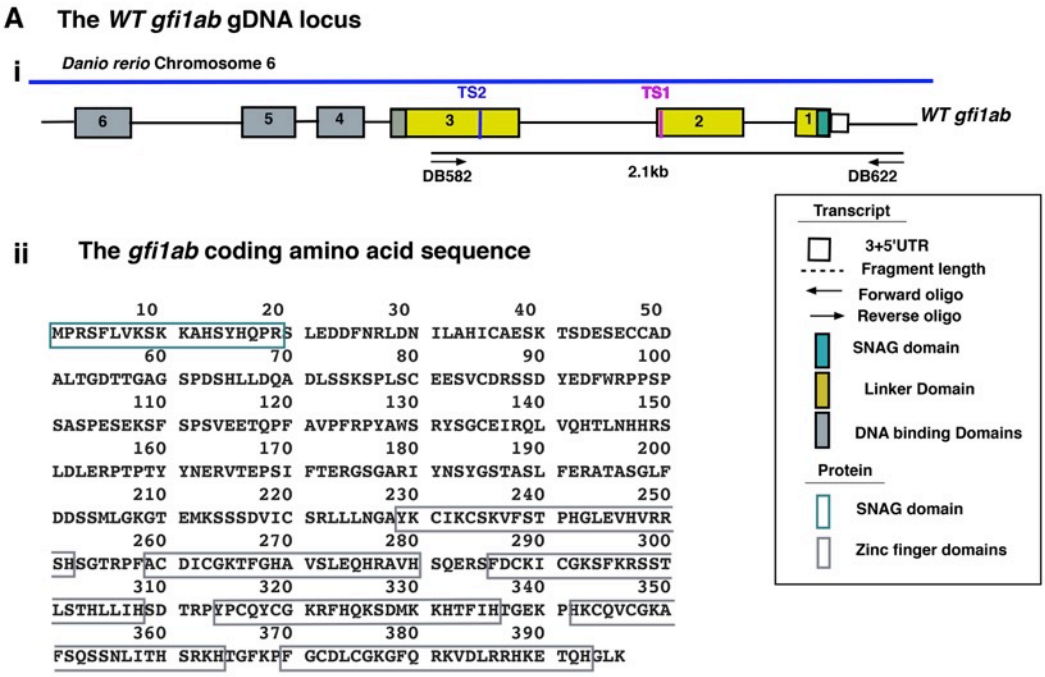


Figure 5.1: The *wt gfi1ab* gene, transcript and protein.
Ai) The *wt gfi1ab* gDNA locus that is located in chromosome 6 of the *Danio rerio* genome. The gene is made up of 6 coding exons. The target sites were selected in the linker domain (mustard coloured), i.e in exons 2 and 3 at the end of exon2 is TS1 (represented with a pink bar) and within exon3 (represented with a blue bar). ii) A schematic representation of the *gfi1ab* gene highlighting the conserved domains; the SNAG domain and the DNA binding domains. Target sites are highlighted in exons 2 and 3 and primers were designed outside of the target sites, DB622 and DB582 that would amplify a 2.1kb fragment in the wt gfi1ab gene iii) shows the translated gfi1ab protein sequence (UniProtKB F1QKD4-1), highlighted in turquoise is the SNAG domain and the regions within grey boxes highlights the 6 zinc (Zn) finger domains.

In order to generate a mutation in *gfi1ab* that would lead to the production of a C-terminally truncated Gfi1ab protein that lacks the DNA binding domain Chris searched exonic DNA sequences for suitable TSs using the ZiFiT Targeter programme (Version 4.2 (http://zifit.partners.org/ZiFiT/ChoiceMenu.aspx) and (Sander et al., 2010; 2007). The programme identifies TSs based on the principle that the sgRNA:Cas9 system can target any sequence of the form 5'-GG-N18-NGG-3'. Based on this criterion the most promising TSs were located at the 3' end of exon2 TS1 (5' GGTGATGCTGATGGTGATGGTGG 3') and in the middle of exon3 TS2 (5' GGTACTCGGGGTGTGAAATCAGG 3') (Figure 5.1 Ai) i.e. in DNA sequences that encode the linker domain of Gfi1ab (Cooney et al., 2013),

(Figure 5.1 Aii). Frameshift mutations induced by gRNA/Cas9 at these TSs would lead to the loss of the protein's DNA binding domain and are thus predicted to seriously affect the function of the protein.

Using customized oligonucleotide linkers, the TS sequences were inserted into the gRNA expression vector pDR274 that harbours a T7 promoter (Hwang et al., 2013). The plasmid was linearised and a gRNA run-off transcript generated in vitro using T7 RNA polymerase. The Cas9 mRNA was transcribed in vitro from linearised plasmid pCS2-nCas9n, which encodes a zebrafish codon-optimized Cas9 with C- and N-terminal nuclear localization signals (Jao et al., 2013). The gRNA and Cas9 mRNA were then coinjected into one-cell stage zebrafish embryos. Injecting both gRNAs with Cas9 allowed us to test both guide RNAs in a single experiment and offered the possibility to recognize genomic deletions simply by PCR amplification.

Initially, Dr. Chris Moore injected the two gRNAs along with the Cas9 mRNA into *qmc551* homozygous embryos. If the *gfi1ab* locus was targeted efficiently one could study the loss of Gfi1ab expression in the context of a Gfi1aa-depleted embryo. Furthermore, eGFP expression in cells of the vDA and HSPCs would highlight any defects that occur during definitive haematopoiesis. When the injected embryos were analyzed at 3dpf it was observed that they displayed a reduced number of eGFP+ cells in the CHT (Figure 5.2 (Av) green arrow head) and in the thymus (Figure 5.2 (Avi) green arrow head) relative to uninjected control siblings (Figure 5.2 (Aii and iii) green arrows). To confirm that the abnormal phenotype correlated with a mutation in *gfi1ab*, single embryos were genotyped. First, gDNA was isolated from four uninjected and four injected embryos (U1-U4 and I1-I4 respectively) (Figure 5.1B). Then, oligos DB582 and DB622 were used to PCR amplify the genomic region that contains both TSs. The expected wt PCR fragment is 2.1kb in size. Successful targeting of both TSs and deletion of the intervening 1kb sequence during

NHEJ would result in the generation of a 1.1kb fragment in the genomic PCR (Figure 5.2 C). For all uninjected embryos, wt 2.1kb fragments were amplified. On gDNA from injected embryos that displayed the CHT phenotype, fragments of different sizes were amplified, including a 1.1kb fragment (Figure 5.2B). This variation may be caused by the NHEJ or HDR mechanisms that join the two ends of the DNA strand by adding or deleting nucleotides. To confirm the sequences of the amplified bands, Chris purified and sequenced both the 2.1kb fragment from the control (U2) and the 1.1kb fragment obtained from an injected (mutant) embryo (I4) using the same oligos that were used for the PCR amplification, i.e. DB582 and DB622.

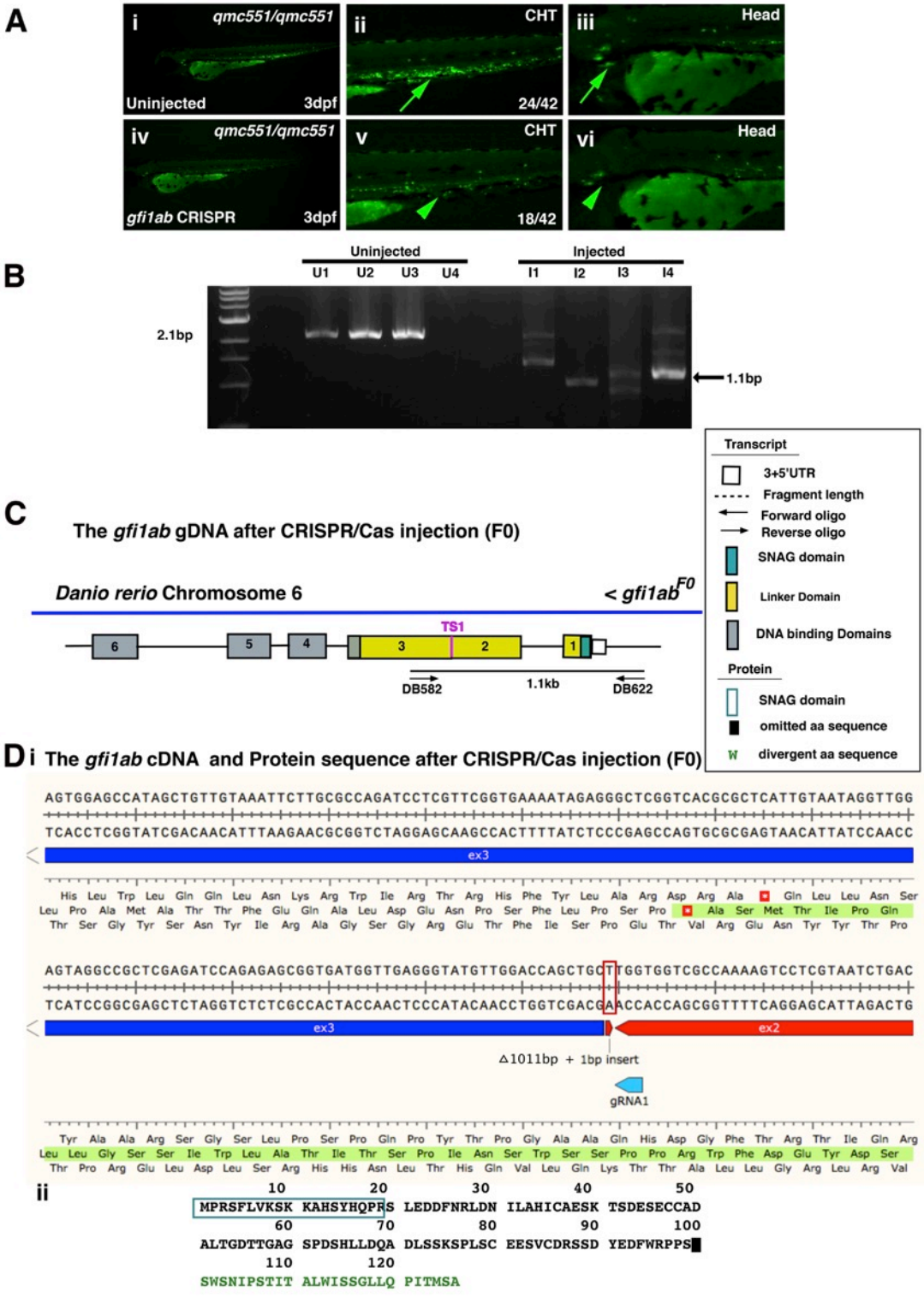

A
i qmc551/qmc551
Uninjected 3dpf
ii CHT
24/42
iii Head
iv qmc551/qmc551
gfi1ab CRISPR 3dpf
v CHT
18/42
vi Head
B
Uninjected
U1 U2 U3 U4
Injected
I1 I2 I3 I4
2.1bp
1.1bp
C
The gfi1ab gDNA after CRISPR/Cas injection (F0)
Danio rerio Chromosome 6
< gfi1ab F0
TS1
6 5 4 3 2 1
DB582
1.1kb
DB622
Transcript
3+5'UTR
Fragment length
Forward oligo
Reverse oligo
SNAG domain
Linker Domain
DNA binding Domains
Protein
SNAG domain
omitted aa sequence
W divergent aa sequence
D i The gfi1ab cDNA and Protein sequence after CRISPR/Cas injection (F0)
AGTGGAGCCATAGCTGTTGTAAATTCTTGCGCCAGATCCTCGTTCGGTGAAAATAGAGGGCTCGGTCACGCGCTCATTGTAATAGGTTGG
TCACCTCGGTATCGACAACATTTAAGAACGCGGTCTAGGAGCAAGCCACTTTTATCTCCCGAGCCAGTGCGCGAGTAACATTATCCAACC
ex3
His Leu Trp Leu Gln Gln Leu Asn Lys Arg Trp Ile Arg Thr Arg His Phe Tyr Leu Ala Arg Asp Arg Ala Gln Leu Leu Asn Ser
Leu Pro Ala Met Ala Thr Thr Phe Glu Gln Ala Leu Asp Glu Asn Pro Ser Phe Leu Pro Ser Pro Ala Ser Met Thr Ile Pro Gln
Thr Ser Gly Tyr Ser Asn Tyr Ile Arg Ala Gly Ser Gly Arg Glu Thr Phe Ile Ser Pro Glu Thr Val Arg Glu Asn Tyr Tyr Thr Pro
AGTAGGCCGCTCGAGATCCAGAGAGCGGTGATGGTTGAGGGTATGTTGGACCAGCTGCTTGGTGGTCGCCAAAAGTCCTCGTAATCTGAC
TCATCCGGCGAGCTCTAGGTCTCTCGCCACTACCAACTCCCATACAACCTGGTCGACGAACCACCAGCGGTTTTCAGGAGCATTAGACTG
ex3
ex2
Δ1011bp + 1bp insert
gRNA1
Tyr Ala Ala Arg Ser Gly Ser Leu Pro Ser Pro Gln Pro Tyr Thr Pro Gly Ala Ala Gln His Asp Gly Phe Thr Arg Thr Ile Gln Arg
Leu Leu Gly Ser Ser Ile Trp Leu Ala Thr Ile Thr Ser Pro Ile Asn Ser Trp Ser Ser Pro Pro Arg Trp Phe Asp Glu Tyr Asp Ser
Thr Pro Arg Glu Leu Asp Leu Ser Arg His His Asn Leu Thr His Gln Val Leu Gln Lys Thr Thr Ala Leu Leu Gly Arg Leu Arg Val
ii
10 20 30 40 50
MPRSFLVKSK KAHSYHQPRS LEDDFNRLDN ILAHICAESK TSDESECCAD
60 70 80 90 100
ALTGDTTGAG SPDSHLLDQA DLSSKSPLSC EESVCDRSSD YEDFWRPPS
110 120
SWSNIPSTIT ALWISSGLLQ PITMSA

Figure 5.2: Injecting *qmc551* homozygous embryos with gRNA1 and gRNA2 along with Cas9 causes a deletion and thereby an out of frame mutation within the *gfi1ab* gene.
Ai-iii) is a control/uninjected *qmc551/qmc551* 3dpf embryo that displayed eGFP+ cells in the (ii) CHT and (iii) thymus (head region) respectively (indicated by the green arrows in image ii and iii), iv-vi) was a *qmc551/qmc551* embryo injected with the CRISPR/Cas9 that display reduced eGFP+ cells in the (v) CHT and (vi) thymus which is located in the head region (indicated by green arrowheads in image v and vi) at 3dpf (B) Genomic DNA from individual uninjected, U1, U2, U3 and U4 (the PCR for U4 failed, hence no band was visible for U4) and injected embryos I1, I2, I3 and I4 were amplified and displayed a 2.1kb band and 1.1kb band respectively on a 1% agarose gel C) schematically depicts the deletion in the F0 *qmc551* homozygous embryos that were injected with the two gRNAs and the Cas9 mRNA, which deletes almost 1kb between the two target sites (TS1 pink bar and TS2 blue bar) while the repair mechanism causes 1bp insertion. Using the same oligos used to amplify the wt fragment between the two TSs, i.e. oligos DB622 and DB582, in the mutant embryo amplify a 1.1kb fragment. Di) Shows the gfi1ab cDNA sequence of the injected F0 mutant embryo, in which there was, a 1bp insertion (red box), which causes a frameshift and thus a 1101bp deletion. Dii) displays the amino acid coding sequence in the mutant, where the SNAG domain is retained and amino acid 95 (95aa) which is Tryptophan (W) up to 99aa in the mutant sequence is unchanged. From then onwards the amino acid sequence changes in the mutant and is replaced by a divergent aa sequence from 125aa (Images A and B were provided by Dr. Chris Moore).

The sequencing data from the injected (F0) embryos showed that both fragments represented *gfi1ab* sequences and that there is indeed a 1011bp deletion and a 1bp insertion between the two TSs in the DNA isolated from one of the injected embryos (Figure 5.2 C and Di). This mutation leads to a frame shift in the *gfi1ab* coding sequence. The mutant allele is predicted to encode a protein that shares the first 99aa with wt Gfi1ab. The following divergent sequence is encoded by the alternative reading frame (Figure 5.2 Di-ii). The predicted loss the DNA binding domain suggests that the truncated Gfi1ab protein will not be able to fulfil any of those roles that require its direct interaction with DNA. The presence of a frame shift mutation in *gfi1ab* suggests that the reduction in the number of eGFP+ cells observed in the sites of haematopoiesis is unlikely to be caused by an off-target effect of the CRISPR/Cas9 system. Instead, it suggests that Gfi1aa and Gfi1ab play a redundant role in definitive haematopoiesis. The severity of the phenotype made it unlikely that double mutant larvae would not survive to adulthood. For that reason, Chris, subsequently, injected the gRNA and cas9 mRNA into *qmc551* heterozygous embryos. None of these embryos showed an obvious phenotype. These embryos could be raised to sexual maturity.

5.2 Identification of two *gfi1ab* alleles and establishment of the *qmc552* and *qmc553* mutant lines

At 4-5 months of age, founders were to be identified that carried a *gfi1ab* mutation in the germline (at this stage I took over this project). For this purpose, potential founders were crossed to wt fish. From each successful cross, ten batches (denoted as B1-B10 in Figure 5.3A and B) of randomly selected 5 embryos were collected and gDNA was extracted. The gDNA was used as a template to amplify the sequence including the TSs (using the same oligos DB622 and DB582 used by Chris) (Figure 5.1 Aii). Based on the previous genotyping of the injected F0 embryos, we expected to find not only the wt 2.1kb

fragment, but also a smaller DNA fragment. Of the ten fish that were screened two carriers of a *gfi1ab* mutation were identified. One was ♂3 and the other was ♀1 (Figure 5.3). PCR analysis of the gDNA isolated from the embryo batches derived from ♂3 revealed the presence of a 1.1kb DNA fragment. This fragment could be detected in all batches, suggesting that at least every fifth embryo carried this deletion mutation i.e. B1-B10 (Figure 5.3A). Thus, the transmission rate was at least 20%. In the case of the ♀1 samples, 8/10 batches displayed the 1.1kb mutant band along with the 2.1kb wt fragment (B1-B6, B8 and B10 Figure 5.3B). The other two (B7 and B9 Figure 5.3B) displayed the 2.1kb band only. Thus, the transmission rate should be at least 16%.

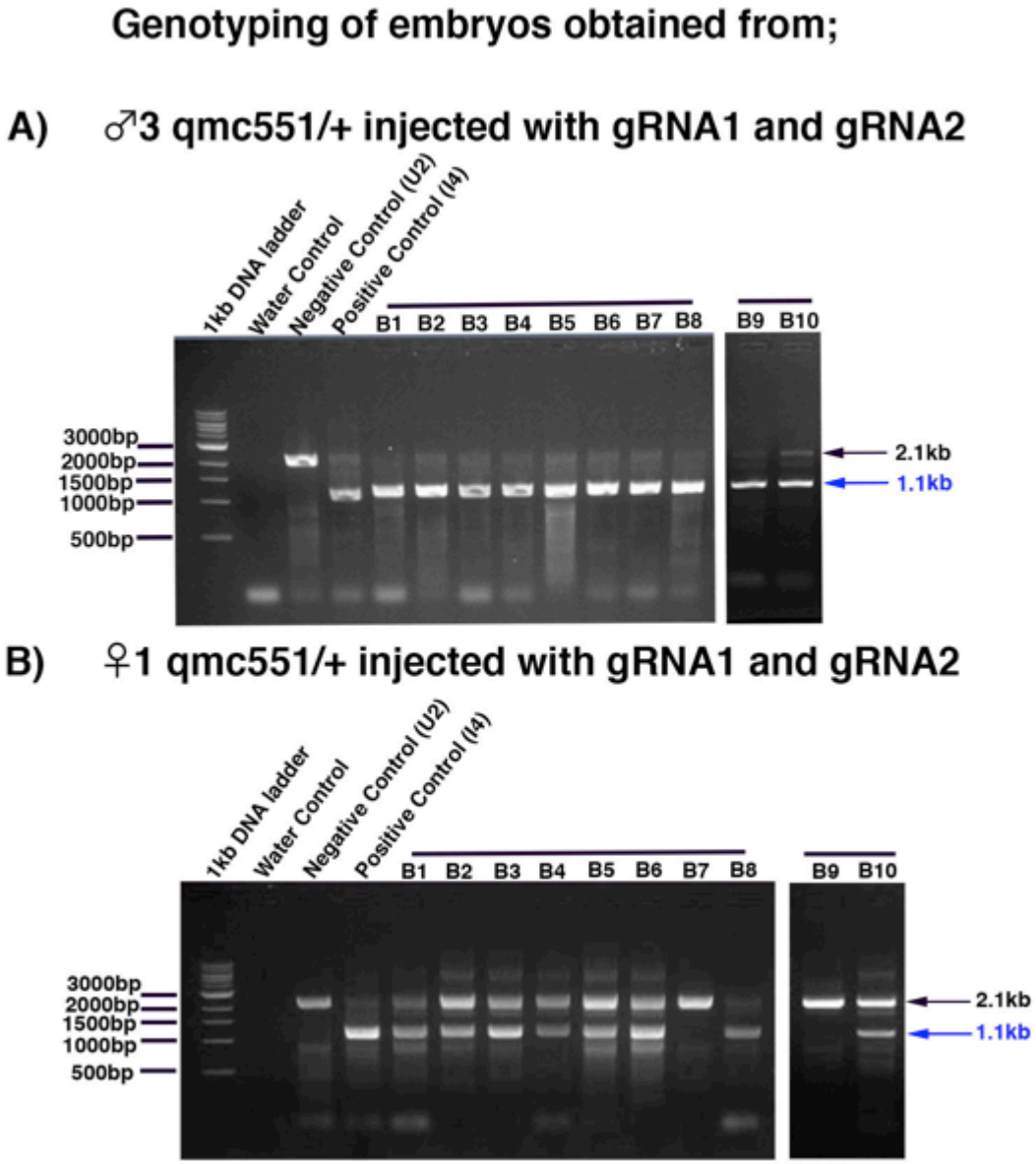


Figure 5.3: Identification of two founders that carry mutations in *gfi1ab* in cells of their germ line.
A) embryos obtained from a cross between ♂3 and a wt fish were genotyped in 10 batches (B1-B10) of 5 embryos each. The amplified DNA samples were run on a 1% agarose gel along with water, negative and positive controls to separate the fragments. A 1kb ladder was also run in parallel. The water control was clean and showed no band. The negative control was gDNA amplified from an uninjected embryo (U2) (see Figure 6.1C) whereas the positive control was gDNA amplified from a *qmc551* homozygous injected embryo (I4) of the preliminary experiments (see Figure 5.2C). The negative control displayed the wt fragment of 2.1bp, whereas the positive control displayed both the 2.1kb as well as the mutant 1.1kb band. The rest of the lanes contain the samples i.e. 10 batches of 5 embryos denoted as B1-B10. All of the 10 batches in this case displayed the 1.1kb band as well as the 2.1kb wt fragment. B) Embryos obtained from ♀1 were similarly genotyped in 10 batches (B1-B10) of 5 embryos each here 8 out of 10 batches displayed both the wt 2.1kb band along with the mutant 1.1kb band, however B7 and B9 displayed the 2.1kb fragment but not the 1.1kb band.

Next, the 1.1kb PCR fragments were isolated from the gel and sequenced to (a) verify that the small fragments represented gfi1ab sequences, (b) determine the exact size of the deletion/insertion and (c) predict the consequences for the gfi1ab reading frame.

Sanger sequencing analysis of the ♂3 ~1.1kb fragment from batch B7 verified that the amplified DNA was derived from the gfi1ab locus. It revealed a 928bp deletion that started in intron 2 and extended beyond TS2. The deleted sequence was replaced by a 3bp insertion. In addition, there was a 6bp deletion in TS1 (Figure 5.4B). The 928 bp deletion eliminates the splice acceptor at the intron 2-exon 3 junction. It is, therefore, likely that processing of the primary transcript leads to the omission of exon 3 sequences as exon 2 sequences are spliced to exon 4 (Figure 5.4 B and C). The splicing of exon2 to exon 4 causes a frame mutation and the generation of a premature stop codon (Figure 5.4C). The protein product retains the first 104 amino acids (excluding the proline and serine residues encoded by the six nucleotides that were deleted in TS1. The protein, thus, contains the SNAG domain, but it lacks its DNA binding domain. Instead, it carries a C-terminal, amino acid sequence that is encoded by an alternative reading frame. The loss of the DNA binding domain probably means that the truncated protein is unable to fulfil those functions that require DNA binding (Figure 5.3D). This new mutant allele of *gfi1ab* was called *qmc552*. The founder ♂3 was crossed to a wt female. The progeny established the *gfi1ab*^{qmc552} line.

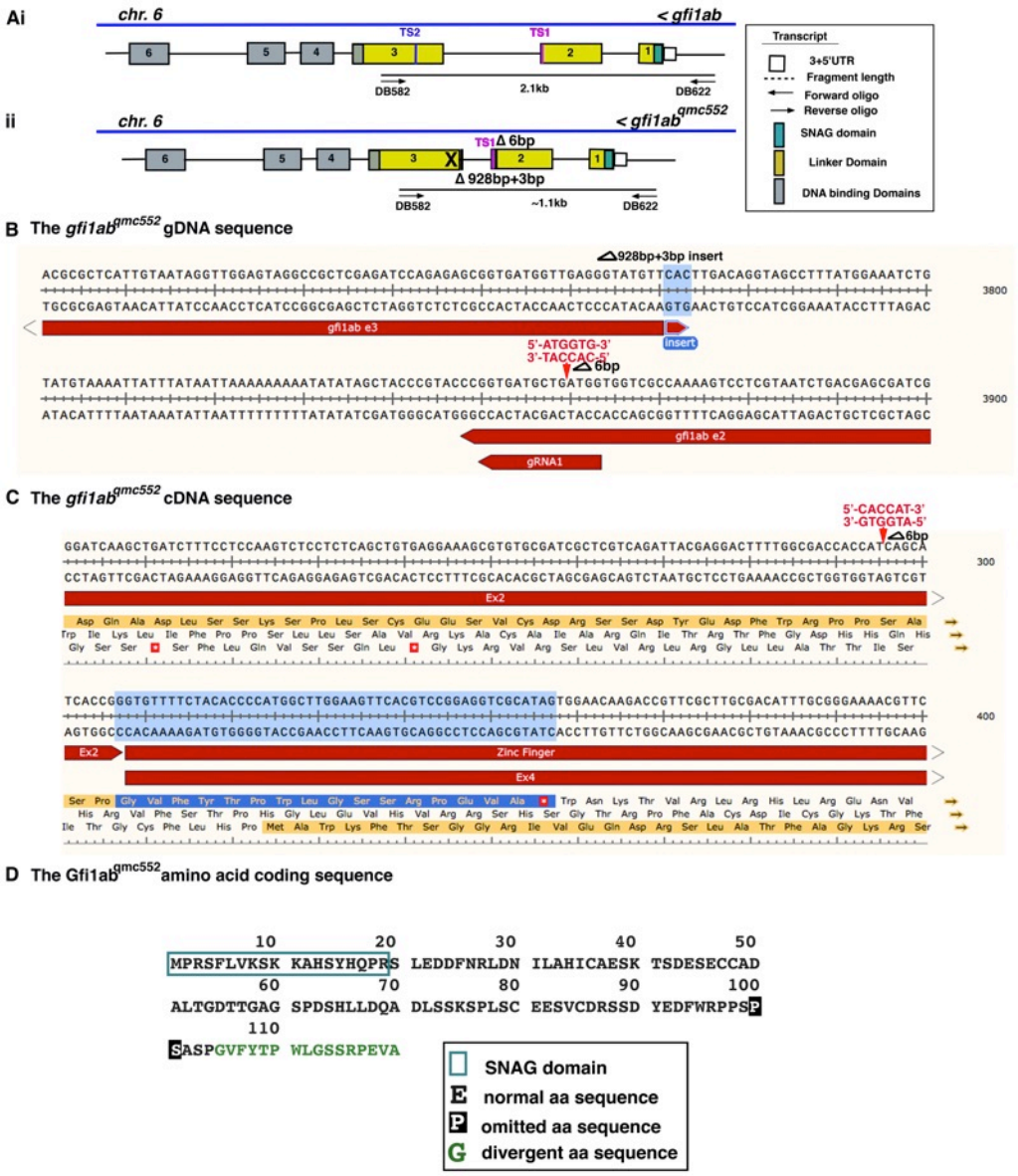


Figure 5.4: Identification and establishment of *qmc552 gfi1ab* mutant line.
Ai) The wt gfi1ab gene is present on chromosome 6 of the zebrafish genome, with six coding exons. TS1 and TS2 were designed in exons2 and 3. (ii) Schematic representation of the *qmc552* locus displaying the 1.1kb fragment that was amplified using oligos DB582 and DB622. The mutant has a 928bp deletion in intron 2-3 and a 3bp insertion. There was also a 6bp deletion in exon2 B) Sequence analysis using SnapGene of the 1.1kb mutant fragment revealed a 6bp deletion at TS1 in exon2, a 928bp deletion and a 3bp insertion at the 3' end of exon 3. Exon3 seems to be deleted completely. C) Shows the *qmc552* cDNA sequence, which shows the 6bp deletion at the cut site that leads to the loss of exon3. The reading frame and the divergent amino acid sequence are highlighted in blue. D) displays the amino acid (aa) coding sequence of qmc552. The SNAG domain is retained and the amino acid sequence remains unchanged up to amino acid 99 (99aa) in the *qmc552* mutant, from then on two amino acids are deleted and the sequence is replaced by a divergent amino acid sequence.

We next purified and sequenced the ~1.1kb fragment from batch B3 of ♀1. The sequencing data was analysed and found to correspond to the *gfi1ab* locus. It carried a 987 bp deletion that started in TS1 and ended in TS2 (Figure 5.5 Ai-ii and B), suggesting that Cas9 had cut the DNA in both sites and that the resulting DSBs were repaired by NHEJ. The 987bp deletion fused exons 2 and 3 and resulted in a frame shift and the generation of a premature stop codon (Figure 5.5B). In the mutant allele, the first 99 codons are maintained. The 6bp deletion in TS1 eliminates a proline and a serine codon (Figure 5.5 B and C). The next three codons encode amino acids 102-105 of the wt Gfi1ab. The frame shift added a divergent sequence of 30 amino acids that is encoded by an alternative reading frame (Figure 5.5C). The truncated protein lacks the DNA binding domain, suggesting that it is not able to fulfil all its normal functions. This new *gfi1ab* allele was named *gfi1ab*^{qmc553}. The female founder of this line was crossed to a *wt* male to establish a new line.

Embryos obtained from the *qmc552* and *qmc553* outcrossed to *wt* fish were sorted for eGFP expression. Half of them also carry the mutant *gfi1ab* allele.

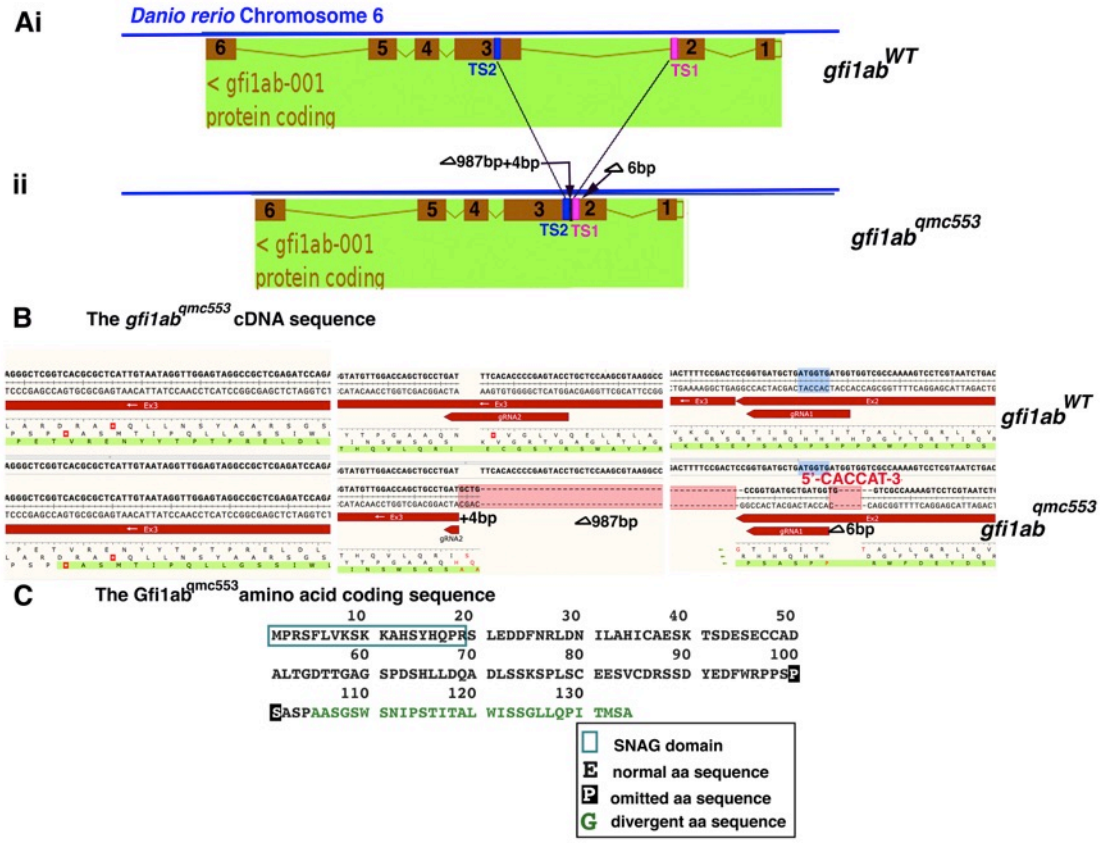


Figure 5.5: Identification and establishment of *qmc553* mutant line.
Ai-ii) A schematic representation of the *wt gfi1ab* gene, which lies in chromosome 6 of the zebrafish genome and has 6 coding exons. After injection of the gRNAs/Cas9 mRNAs, which cut at both TSs and the NHEJ repair mechanism, causes indels at the TSs, 4bp insertion and a 6bp deletion, that leads to the loss of the intron between exons2 and 3, a 987bp deletion (the image B displays only part of the 987bp big deletion). Bi) A sequence comparison between the *wt* and *qmc553* alleles at the site of deletion displays the indels at the TSs and the big deletion that may have been caused by the HDR repair mechanism. C) The coding aa sequence of the wt and qmc553 is the same until 99aa (displayed from 91aa in B) after 99aa there are omitted amino acid sequences followed by an Alanine, Serine and Proline (ASP), which is the same as the wt sequence. From this point on the amino acid sequence is replaced by a divergent amino acid sequence in qmc553.

5.3 Preliminary data suggests that the loss of *gfi1ab* causes a reduction in eGFP+ cells with in the CHT and thymus of *qmc551* homozygous embryos.

While these embryos were raised, I tried to get preliminary data on the phenotype of the *gfi1aa/gfi1ab* double mutant embryos. I crossed the *qmc552* male and the *qmc553* female founders with each other to generate qmc551 homozygous embryos that are compound heterozygous carriers for the two *gfi1ab* mutations.

Embryos were collected and sorted for eGFP- *(gfi1aa$^{+/+}$)*, eGFP+ (low) (*gfi1aa$^{qmc551/+}$*) and eGFP+ (high) (*gfi1aa$^{qmc551/qmc551}$*) on late 1 dpf. Among the eGFP+ (high) *qmc551* homozygous embryos, we expect to have embryos that also carried both *qmc552* and *qmc553* alleles. Based on the phenotype of the CRISPR/Cas9-injected *qmc551 homozygous* F0 fish (Figure 5.2Av-vi), we expected these *qmc551 homozygous;qmc552/qmc553* embryos to display defects in sites of larval and adult definitive haematopoiesis (Figure 5.6). To explore this possibility, 13 eGFP+ (high) *qmc551* homozygous embryos (Figure5.6 A-A') were collected from a cross of the *qmc552* and *qmc553* founders. Under the fluorescent microscope, one, embryo M1, displayed a phenotype during embryonic development. While this embryo showed normal GFP expression in the lateral line organ and in the inner ear, the number of eGFP+ cells was reduced in the CHT and thymus (Figure5.6 B-B') at 3dpf, suggesting a haematopoietic phenotype similar to that of the CRISPR/Cas9 injected *qmc551* homozygous F0 embryos.

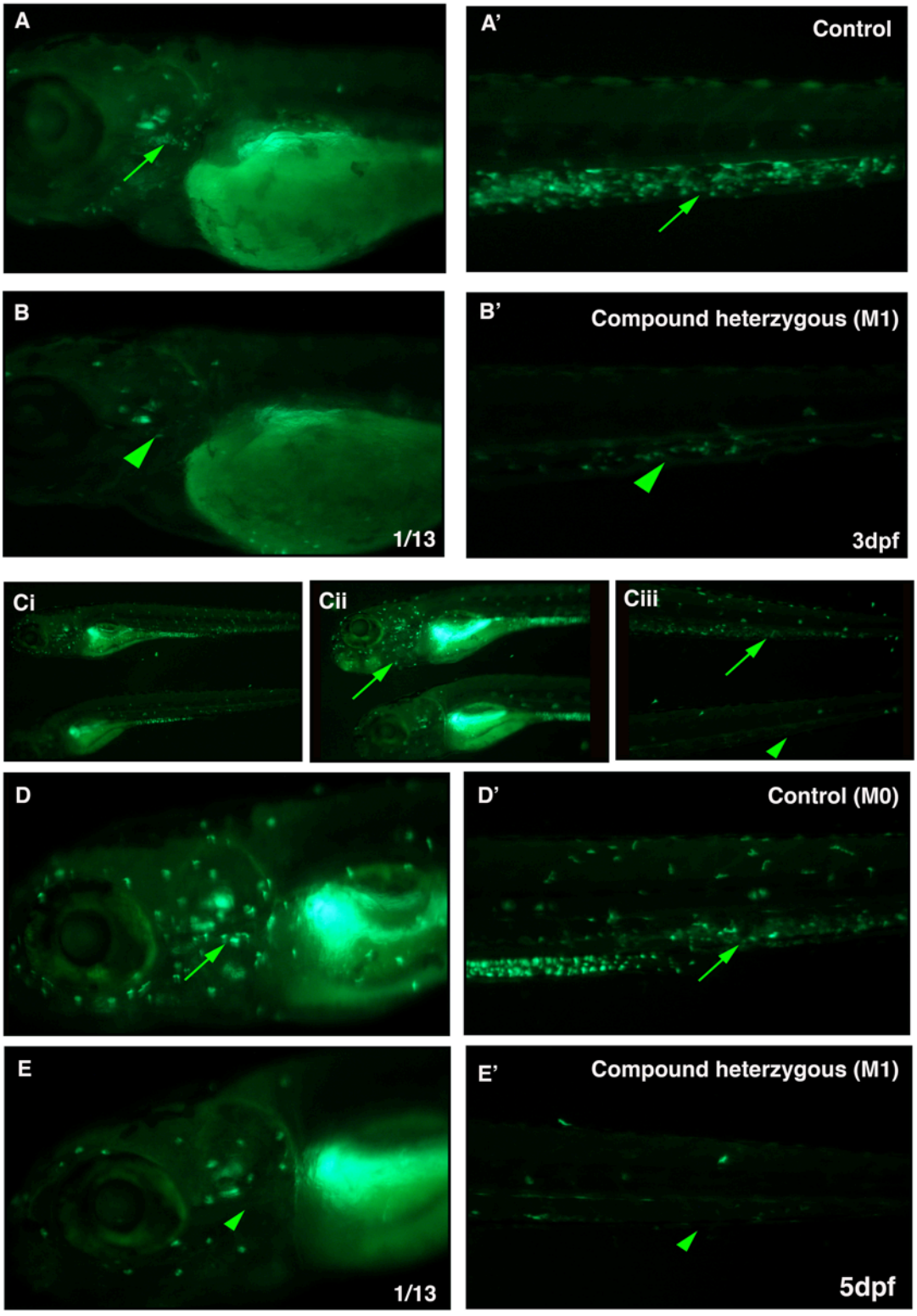

A
A'
Control
B
1/13
B'
Compound heterzygous (M1)
3dpf
Ci
Cii
Ciii
D
D'
Control (M0)
E
1/13
E'
Compound heterzygous (M1)
5dpf

Figure 5.6: *gfi1ab* compound heterozygous embryos show reduced numbers of eGFP+ cells in the CHT and thymus.
(A-A') Fluorescent images of a control embryo on 3dpf displays large number of eGFP+ cells in the CHT and thymus (B-B') in comparison to a *gfi1ab* potential compound heterozygous (M1) embryo that had significant reduction in eGFP+ cells with in these sites of definitive haematopoiesis. (Ci-iii) Fluorescent images taken on 5dpf of the embryo M1 with reduced number of eGFP+ were imaged alongside a control M0 embryo, which had a high number of eGFP+ cells. (D-D') Displays the control M0 embryo that shows more eGFP+ in the thymus and CHT, whereas (E-E') the M1 embryo had almost no staining in the thymus and CHT. (eGFP+ cells in the control embryo M0 are indicated with a green arrow and loss or reduction of eGFP+ cells in the M1 embryo is indicated with green arrowheads).

In comparison to another eGFP high embryo, embryo M0 (Figure 5.6 D-D') the M1 embryo remained to display a significant reduction in eGFP+ cells within the thymus and CHT at 5dpf (Figure 5.6 E-E'). Besides the $eGFP^+$ cells within the sites of larval and adult haematopoiesis, there were many wandering cells visible throughout the head and tail region of the M0 embryo, in comparison to M1. These cells are thought be immature leucocytes (Thambyrajah et al., 2016).

In order to further confirm that the phenotype seen in the M1 embryo correlated with a loss of both wt gfi1ab copies we carried out a series of PCR experiments to genotype the embryo. For this purpose, gDNA was isolated from individual eGFP high embryos (n=13, which includes the embryo with a phenotype) as well as individual $eGFP^-$ negative siblings (n=14). We analysed this number of embryos in order to get an idea of the distribution of the wt and the two mutant *gfi1ab* alleles amongst their progeny of the two founders. First, I performed PCR amplification with oligos DB565 and DB582 to help me distinguish between the *wt, qmc552* and *qmc553* alleles (Figure 5.7Ai-iii). With these oligos we expect a 1,173bp, 241bp and 184bp fragments with the *wt, qmc552* and *qmc553* alleles, respectively (Figure 5.7Ai-iii). Although the presence of numerous non-specific PCR products complicates the interpretation of the gel, it nevertheless allowed us to distinguish allele-specific DNA bands. The genotyping showed that nine eGFP-negative embryos carried only the *wt* allele of *gfi1ab*. Three GFP-negative embryos appeared to be heterozygous for the *qmc552* allele (embryos 3-1.5, 3-1.6 and 3-1.11) while two embryos (3-1.2 and 3-1.9) displayed only the *qmc553*-specific PCR fragment (Figure 5.7B, Table 5.1). Among the eGFP+ (high) embryos, seven appeared to be *gfi1ab wt*, five were heterozygous for *qmc552* (embryos M0, M4, M5, M8 and M11) and one embryo, embryo M1, displayed only the 184bp *qmc553* fragment (Figure 5.7B). Just like in the case of the two GFP-negative *qmc553* carriers, a second *gfi1ab* allele could not be amplified.

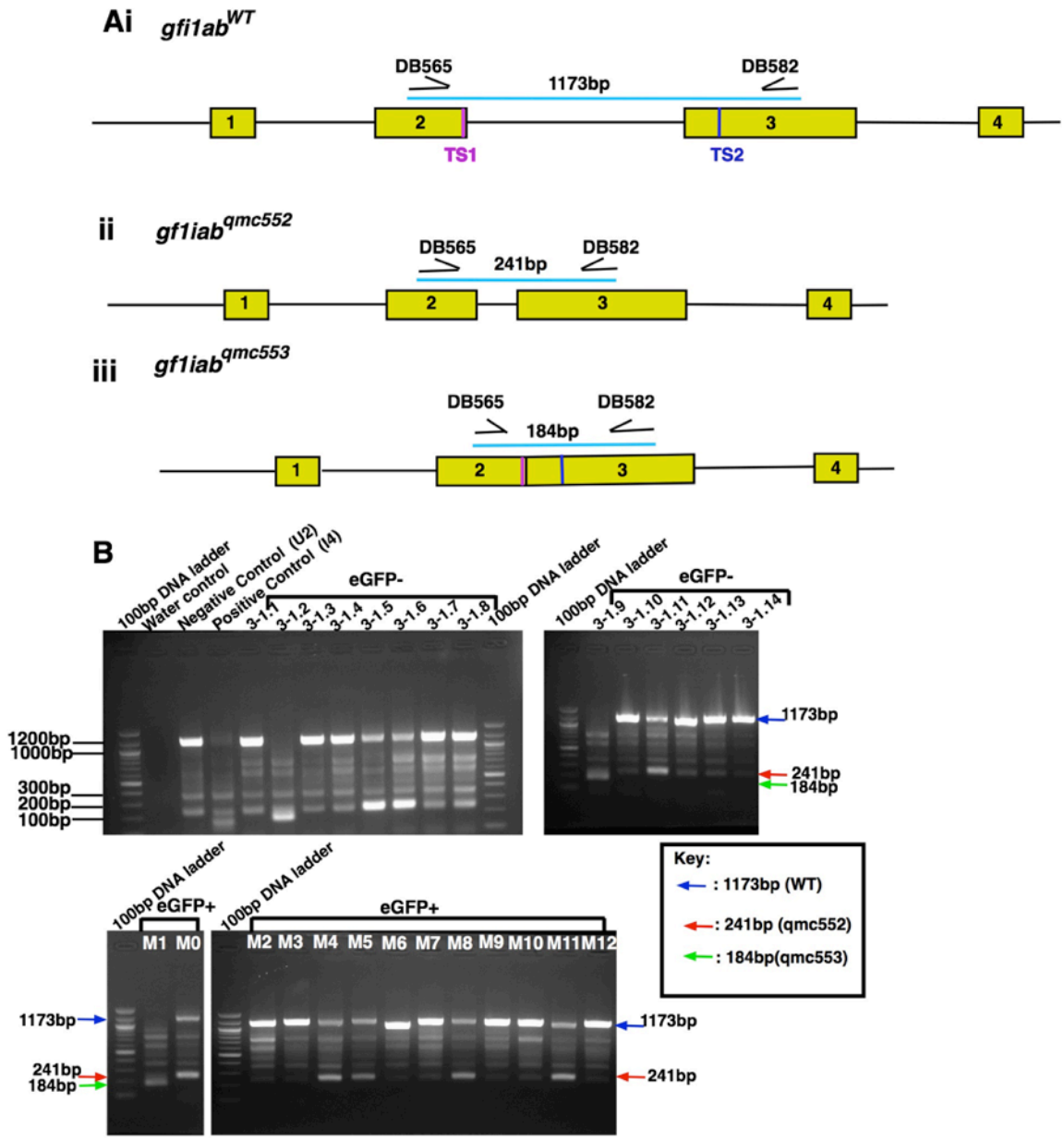


Figure 5.7: The genotype of the M1 embryo suggests that it is *gfi1ab* compound heterozygous embryo.

(Ai-iii) Displays the *gfi1ab*$^{+/+}$, *gfi1ab*^{qmc552} and *gfi1ab*^{qmc553} loci and the binding sites of oligos DB565 and DB582. The expected fragment sizes using these oligos is 1173bp, 241bp and 184bp fragments with the *wt, qmc552* and *qmc553* alleles respectively. B) 1.5% agarose gel displays PCR amplified gDNA from individual eGFP- and eGFP+ (high) embryos using oligos DB565 and DB582. The eGFP^{-} embryos were numbered as 3-1.1-3-1.14. Where embryos 3-1.2 and 3-1.9 seem to display only a 184bp *qmc553* fragment, while 3-1.5, 3-1.6 and 3-1.11 seem to display the 1173bp *wt* and 241bp *qmc552* fragments. We think the other embryos in this eGFP^{-} batch may be *wt* for the *gfi1ab*. Among the eGFP^{+} (high) batch of embryos the M1 embryo does not display a *wt* 1173bp fragment but only displayed a 184bp *qmc553* fragment. Embryo M0, M4, M5, M8 and M11 display both 1173bp wt and the 241bp *qmc552* fragments. We think the rest of the embryos were wt for the *gfi1ab* allele. Control included, a water control that was clear, a negative control, which was PCR amplified gDNA from F0 uninjected (U2) and the positive control was PCR amplified gDNA from F0 injected (I4) (Figure 5.2B).

Table 5.1: Genotype with respect to the *gfi1ab* gene and numbers of eGFP- (*qmc551*$^{+/+}$) and eGFP+ (*qmc551/qmc551*) single embryos analysed

eGFP^{-} (*gfi1aa*$^{+/+}$) n=14	
Genotype (for *gfi1ab*)	Numbers
wt	9/14
qmc552 (heterozygous)	3/14
qmc553 only	2/14
eGFP^{+} (*qmc551/qmc551*) n=13	
wt	7/13
qmc552 (heterozygous)	5/13
qmc553 only	1/13

In order to confirm that embryos e3-1.2, e3-1.9 and M1 embryo did not carry a *wt gfi1ab* allele; we employed another set of oligos. PCR amplification with oligos DB622 and DB582 results in the generation of 2159bp (~2.1kb), 1226bp (~1.2kb) and 1171bp (~1.1kb) fragments with the *wt*, *qmc552* and *qmc553* alleles, respectively (Figure 5.8 Ai-Aiii). While the mutant cannot reliably be resolved on the agarose gel, the experiments can tell us whether a *wt* allele is present in the gDNA of the embryos. The PCR analysis (Figure 5.8B) confirmed the genotyping result for the GFP-negative *qmc552* heterozygous embryo e3-1.5 and showed again that the gDNAs of embryos e3-1.2, e3-1.9 and M1 (Lanes 7-9) displayed only what appeared to be a ~1.1kb *qmc553* fragment, but not the ~2.1kb wt fragment. Whether the lack of a second *gfi1ab* allele in the *qmc553* mutant

embryos (a) suggests the presence of an additional undetectable mutant allele in the germline of the *qmc552* founder, (b) is caused by insufficient resolution of the *qmc552* and *qmc553* bands on the gel or (c) reflects unexplained preferential amplification of the *qmc553* allele over the *wt* and *qmc552* alleles in both sets of PCR reactions (please note that oligo DB582 was used in both reactions) remained unclear. The data gave us a rough indication of mutant allele transmission (~30% for *qmc552* and ~11% for *qmc553*, Table 5.1) and showed that the GFP high embryo M1 with the haematopoietic phenotype could indeed be a *qmc551* homozygous, *qmc553* compound heterozygous embryo.

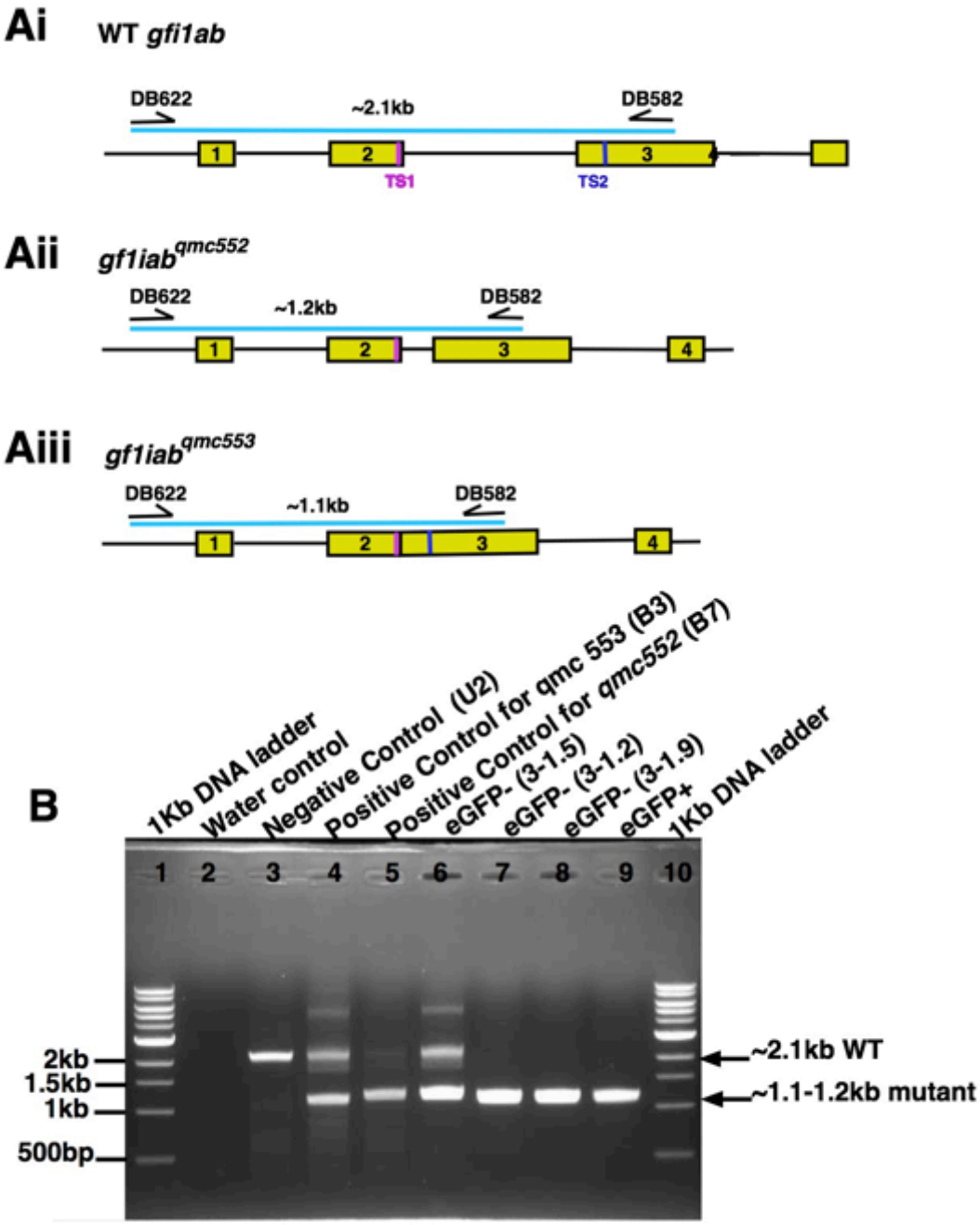

Ai
WT gfi1ab
DB622
~2.1kb
DB582
1
2
3
TS1
TS2
Aii
gf1iab qmc552
DB622
~1.2kb
DB582
1
2
3
4
Aiii
gf1iab qmc553
DB622
~1.1kb
DB582
1
2
3
4
B
1Kb DNA ladder
Water control
Negative Control (U2)
Positive Control for qmc 553 (B3)
Positive Control for qmc552 (B7)
eGFP- (3-1.5)
eGFP- (3-1.2)
eGFP- (3-1.9)
eGFP+
1Kb DNA ladder
1 2 3 4 5 6 7 8 9 10
2kb
1.5kb
1kb
500bp
~2.1kb WT
~1.1-1.2kb mutant

Figure 5.8: *gfi1ab* compound heterozygous embryos do not display a *wt* gDNA *gfi1ab* fragment.
(Ai-Aiii) Schematically highlights the region of the *gfi1ab* locus where the TSs were designed and the oligos DB622 and DB582 lay. The expected fragment sizes using these oligos is ~2.1kb, ~1.2kb and ~1.1kb fragments with the *wt, qmc552* and *qmc553* alleles respectively. B) Is a 1% agarose gel samples were applied, in which Lane1 contained the 1kb DNA ladder (NEB) highlighting the 500bp, 1kb, 1.5kb and 2kb fragments, LanM0 was a water control that was clear, Lane3 contained a negative control, which was PCR amplified gDNA uninjected control (U2) (Figure 5.2B) and displayed a ~2.1kb wt gDNA *gfi1ab* fragment, Lanes 4-5 contained positive controls for *qmc553* and *qmc552*, B3 and B7 respectively (Figure 5.3 A and B respectively), which displayed ~1.1kb and ~1.2kb respectively. These lanes also displayed the ~2.1kb wt fragment, although weak in Lane5 (B7). Lane 6 was a PCR amplified gDNA from an eGFP- *qmc552* heterozygous embryo, which displayed both the *wt* (~2.1kb) fragment as well as the ~1.2kb *qmc552* fragment. Lanes 7-8 were PCR amplified gDNA from the two eGFP- embryos expected compound heterozygous that displayed no *wt* ~2.1kb fragment but that displayed only a ~1.1-1.2kb fragments. Lane9 was the PCR amplified gDNA from the eGFP+ embryo expected compound heterozygous (documented in Figure 5.6 B-B', (Ci-iii and E-E') that displayed no wt ~2.1kb fragment but that displayed only a ~1.1-1.2kb fragment. Lane 10 contained was the 1kb DNA ladder (NEB).

5.4 A second unidentified mutant *gfi1ab* allele segregates in the progeny of the *qmc552* founder.

In order to further try and resolve the two mutant alleles in the compound heterozygous mutant samples, we used oligo sets specific to each allele. The *wt* oligo set (oligo set 'a') DB757 and DB619 hybridizes to sequences within intron 2 of the *gfi1ab wt* locus and allow amplification of a 795bp fragment from wt gDNA (Figure 5.9Ai,Bi). As the sequence that oligo DB619 binds to is not present within the *qmc552* or *qmc553* alleles, this oligo set 'a' should not amplify DNA sequences of the *qmc552* and *qmc553* alleles. Since the *qmc552* allele, but not *qmc553*, retains part of the intron 2 sequences, oligo DB757 was designed within this region. In combination with the exon 3 oligo DB755 (DB755 and DB757 are 'b' set of oligos) it allows specific amplification of a 130bp *qmc552* fragment, as well as the generation of a 1 kb *wt* band (Figure 5.9Aii,Bi). For the amplification of the *qmc553* fragment we used DB565 and DB582 oligos ('c' set of oligos), which bind upstream and downstream of the TSs in exon 2 and exon 3, respectively, and is identical to the set used in the first PCR reactions (Figure 5.7). The size of the expected PCR product is 184bp, 241bp and 1.1kb for the *qmc553*, *qmc552* and *wt* alleles of *gfi1ab*, respectively (Figure 5.9Aiii,Bi). Initially, short extension times were used in the PCR reactions with oligo sets b and c in the hope to prevent the amplification of the large 1 kb wt fragments, but experience has shown that wt PCR fragment were amplified even under these conditions (Figure 5.9Bii).

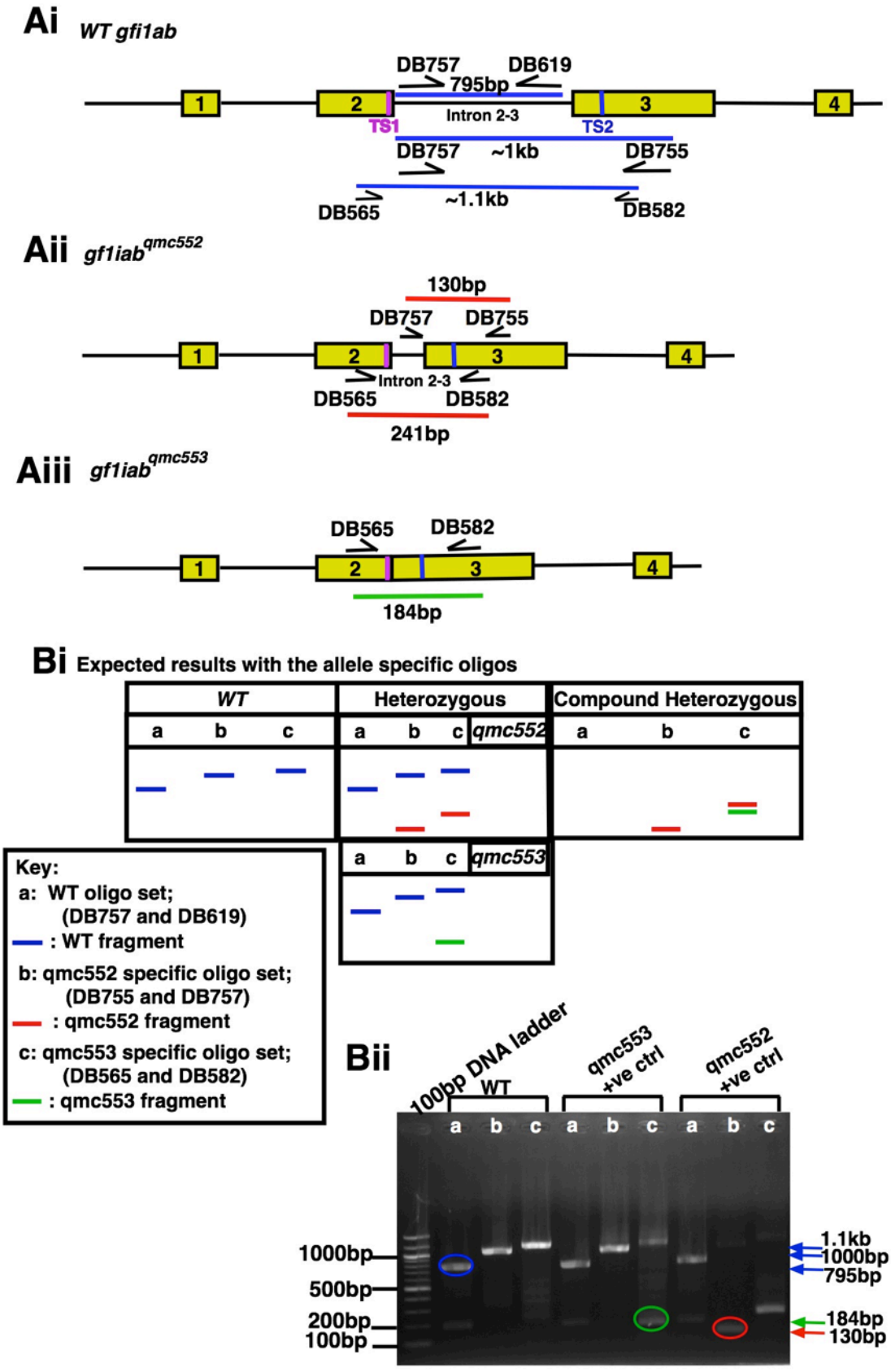

Ai WT gfi1ab
DB757
DB619
795bp
1
2
3
4
Intron 2-3
TS1
TS2
DB757
~1kb
DB755
DB565
~1.1kb
DB582
Aii gf1iab qmc552
130bp
DB757
DB755
1
2
3
4
Intron 2-3
DB565
DB582
241bp
Aiii gf1iab qmc553
DB565
DB582
1
2
3
4
184bp
Bi Expected results with the allele specific oligos
WT
Heterozygous
Compound Heterozygous
a b c
a b c qmc552
a b c
a b c qmc553
Key:
a: WT oligo set;
(DB757 and DB619)
: WT fragment
b: qmc552 specific oligo set;
(DB755 and DB757)
: qmc552 fragment
c: qmc553 specific oligo set;
(DB565 and DB582)
: qmc553 fragment
Bii
100bp DNA ladder
WT
qmc553 +ve ctrl
qmc552 +ve ctrl
a b c a b c a b c
1000bp
500bp
200bp
100bp
1.1kb
1000bp
795bp
184bp
130bp

Figure 5.9: Oligo sets specific to each allele were designed to distinguish between the *qmc552* and *qmc553* alleles.
(A) Schematically highlights the region of the *gfi1ab* locus (not all of the gene) where the TSs was designed and the oligos specific to each allele lay as well as common oligos have also been displayed. Ai) Depicts the *wt gfi1ab* locus where TS1 and TS2 were designed. Oligos DB619 and DB757 (set 'a' oligos) were selected to amplify a 795bp fragment from wt gfi1ab gDNA but this pair together does not amplify a fragment in *qmc552* or *qmc553*. Depicted are also the oligos specific to *qmc552* and *qmc553* that amplify bands in the *gfi1ab wt* allele, DB755 and DB757 amplify a ~1kb fragment and DB565 and DB582 amplify a ~1.1kb fragment. Aii) Highlights the $gfi1ab^{qmc552}$ locus at the gRNAs/Cas9 cut site. The oligos specific to the *qmc552* allele were DB755 and DB757 (set 'b' oligos) giving a 130bp fragment and *qmc552* allele. Oligos DB565 and DB582 also amplify a 241bp fragment in this region. Aiii) Highlights the $gfi1ab^{qmc553}$ locus at the gRNAs/Cas9 cut site. The oligos specific to the *qmc553* allele were DB565 and DB582 (c set 'c' oligos, this oligo set was also used in previously, see Figure 5.7. Here, however the PCR conditions were changed to reduce the number of non-specific bands, as seen in Figure 5.7) that amplify an 184bp fragment on gDNA. Bi) Is a schematic representation of the expected results on a gel using specific oligos. Bii) Is a representative example of a 1.5% agarose gel on which gDNA from each a *wt*, a positive control for *qmc553* (B3 Figure 5.3B) and a positive control for *qmc552* (B7 Figure 5.3A) were amplified using all 3 sets of oligos, 'a', 'b' and 'c'. Colour codes: blue represents *wt* fragments, red represents *qmc552* fragments and green represents *qmc553* fragments.

Genomic DNAs from several embryos were re-analyzed using the allele-specific oligo sets (Figure 5.10). Embryos 3-1.5 and M0 that the previous PCRs had revealed to be heterozygous for the *qmc552* allele of *gfi1ab* (see Figure 5.7) displayed all the expected *wt* and *qmc552* bands in the PCRs with oligo sets a, b and c. On the gDNA of the eGFP- embryo 3-1.2 and the eGFP+ (high) embryo M1 (Figure5.7), only oligo set c gave a convincing PCR product that corresponded to the 184bp *qmc553* specific fragment. There was a weak *wt* band in the PCR on the gDNA of embryo 3-1.2 with the wt-specific oligo set a (in lane 2), but the presence of a *wt gfi1ab* allele was not confirmed with oligo sets b and c (Figure 5.10). Thus, these data confirm the presence of the *qmc553* allele and the absence of *wt* or *qmc552* allele in the gDNA of the GFP-negative embryo 3-1.2 and the eGFP-high qmc551 homozygous embryo M1. What has happened to the second *gfi1ab* allele, i.e. the one inherited from the *qmc552* founder, remains to be investigated. The genotyping data on the embryos 3-1.2 and M1 seem to suggest that this allele of *gfi1ab* cannot be detected with our oligos sets, suggesting that a second mutant *gfi1ab* allele is segregating in the progeny of the *qmc552* founder. It appears that the haematopoietic phenotype observed in eGFP-high *qmc551* homozygous embryo M1 correlate indeed with the loss of all *wt gfi1ab* gene copies in the embryo.

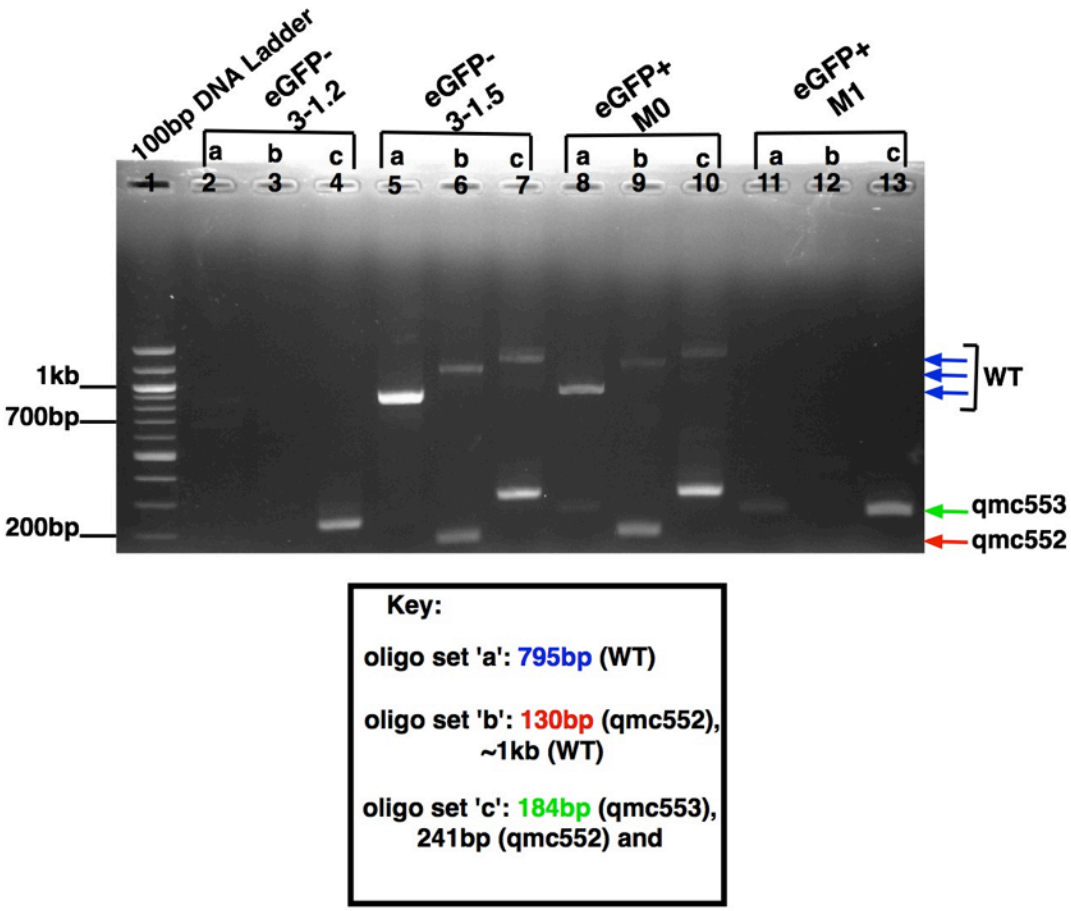


Figure 5.10: Allele specific oligo sets suggest the presence of a second mutant allele in the *qmc552* founder.
A 1.5% gel was run to resolve the fragments that were amplified from gDNA. Lane1 shows a 100bp DNA ladder (NEB) 200bp, 700bp and 1kb positions have been highlighted. Lane M0-4 contained amplified gDNA obtained from one eGFP$^-$ compound heterozygous embryo (3-1.2, Figure 5.7) using all 3 sets of oligos a, b and c respectively. Lane 2 with set (a) oligos displayed a weak wt band (795bp) one eGFP$^-$ *qmc552* heterozygous embryo (3-1.5, Figure 5.7) was PCR amplified using the 3 sets of oligos a, b and c and gDNA from eGFP$^+$ M0 and compound heterozygous embryo M1 were also amplified using the 3 sets of oligos a, b and c. Lanes 2,5,8 and 11 were samples amplified using oligos DB619 and DB757 (a), Lanes 3,6,9 and 12 were samples amplified using oligos DB755 and DB757 (b) and Lanes 4,7, 10 and 13were samples amplified using oligos DB565 and DB582 (c).

Since this sample displayed neither the *wt gfi1ab* fragment nor the *qmc552* fragment, one would have to assume the presence of a yet uncharacterized *gfi1ab* mutant allele.

5.5 Summary

- The *qmc551* homozygous embryos survive up to adulthood and are fertile. It was found that *gfi1ab* is upregulated in the vDA in the absence of *gfi1aa* expression, thus suggesting that *gfi1aa* and *gfi1ab* play redundant roles during definitive haematopoiesis (Ucanok, 2015).
- To investigate this functional redundancy, mutant alleles of gfi1ab were generated using the CRISPR/Cas technology. Injection of the CRISPR guide RNAs against target sites 1 and 2 in exons 2 and 3 of *gfi1ab* together with Cas9 mRNA into one-cell stage *qmc551* homozygous zebrafish embryos suggested that the knockout of the gene *gfi1ab* in *gfi1aa* mutants results in a specific phenotype in definitive haematopoiesis (Chris Moore, personal communication).
- *qmc551* heterozygous embryos were injected with the gfi1ab gRNAs and Cas9 mRNA and were raised to adulthood (Chris Moore, personal communication).
- Among the injected fish, two founders were identified; ♂3 and ♀1, which carried the $gfi1ab^{qmc552}$ and $gfi1ab^{qmc553}$ mutant alleles, respectively.
- Sequence analysis revealed that the *qmc552* allele carries a 928 bp deletion that encompasses a large part of intron 2, as well as the 5' end of exon 3. It is assumed that in the absence of the splice acceptor at the 5'end of exon 3 RNA processing will lead to the splicing of exon 2 to exon 4 in the mRNA that will result in a frameshift and the generation of a premature stop codon.
- Sequence analysis of the *qmc553* allele revealed a deletion of 987 bp that starts in TS1 and ends in TS2, effectively fusing exon 2 and 3 sequences. The mutation

shifts the *gfi1ab* reading frame and leads to the introduction of a premature stop codon.

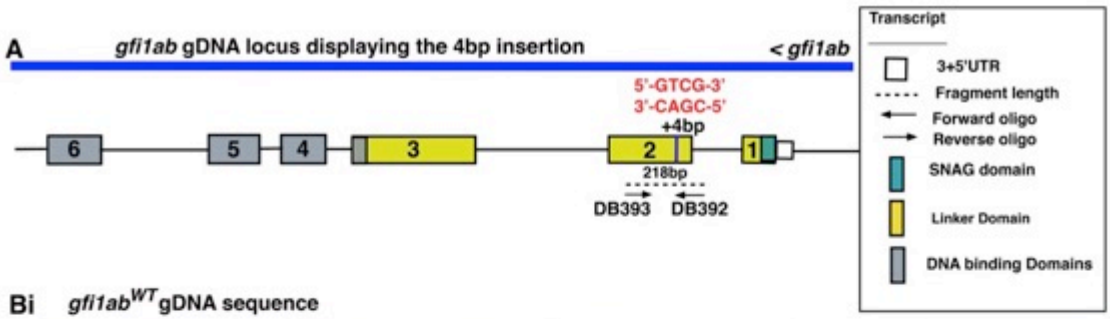
- Both alleles, *qmc552* and *qmc553* encode truncated proteins that retain the SNAG repressor domain, but lack the DNA binding domain. These truncated proteins are likely to be unable to fulfil the roles of Gfi1ab that require active DNA binding.
- The two founders were crossed to *wt* fish to generate embryos that establish new mutant lines.
- The two founders were crossed to each other to generate compound heterozygous fish that are deficient for Gfi1aa and Gfi1ab. Among the GFP-high *qmc551* homozygous progeny, one embryo was identified that displayed a reduced number of qmc551:eGFP-positive definitive blood cells in the CHT.
- Genotyping revealed that it carried a *qmc553* allele, but no second *gfi1ab* allele. The lack of a second allele suggests that the observed phenotype does correlate with a loss of Gfi1aa and Gfi1ab expression. It also indicates that a second mutant *gfi1ab* allele segregates in the progeny of the *qmc552* founder. The exact nature of this second mutant allele requires further investigation.

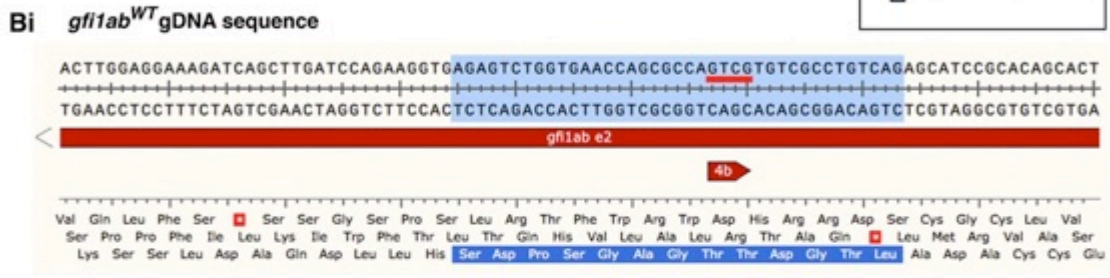
6 The loss of both *gfi1a* paralogues causes a reduction in *c-myb*+ cells within the larval CHT and kidney

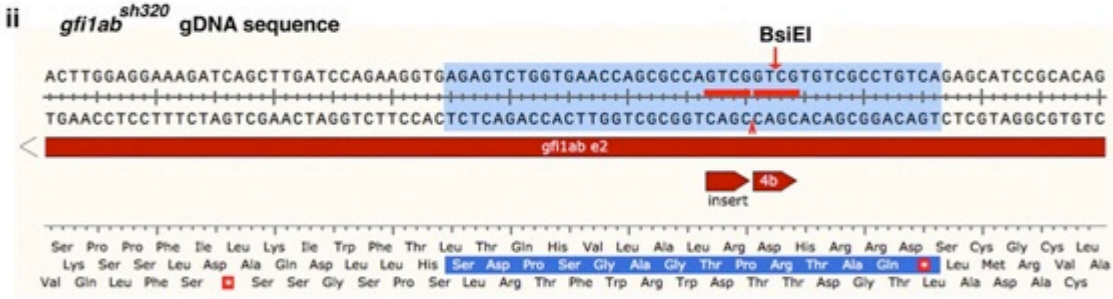
6.1 Introduction

In addition to the qmc552 and qmc553 *gfi1ab* mutant lines, Dr. Robert Wilkinson's lab in Sheffield identified, another *gfi1ab* mutant allele called *sh320*. This line was crossed to the qmc551 line with the view to generate double mutants that would be characterised.

The *sh320* allele was created by targeting a sequence in exon 2 of the *gfi1ab* gene using zinc finger nucleases (Urnov et al., 2010). The mutation carried by the *sh320* allele is a duplication of the 4bp sequence 5'-GTCG-3' in exon 2 of the *gfi1ab* locus. (Figure 6.1A and B). The 4bp duplication leads to a frame shift in the *gfi1ab* coding sequence and the possible production of a truncated protein product. The mutant protein is predicted to retain the N-terminal SNAG domain, but to lack the six-zinc finger DNA binding domain. They are replaced by a divergent amino acid sequence (Figure 6.1 Ci-ii). While the truncated protein may retain the ability to interact with other proteins via the SNAG domain, the absence of the DNA binding domain is likely to interfere with many of the protein's normal functions.







Ci Gfi1ab^WT protein coding sequence

```
        10        20        30        40        50
MPRSFLVKSK KAHSYHQPRS LEDDFNRLDN ILAHICAESK TSDESECCAD
        60        70        80        90       100
ALTGDTTGAG SPDSHLLDQA DLSSKSPLSC EESVCDRSSD YEDFWRPPSP
       110       120       130       140       150
SASPESEKSF SPSVEETQPF AVPFRPYAWS RYSGCEIRQL VQHTLNHHRS
       160       170       180       190       200
LDLERPTPTY YNERVTEPSI FTERGSGARI YNSYGSTASL FERATASGLF
       210       220       230       240       250
DDSSMLGKGT EMKSSSDVIC SRLLLNGAYK CIKCSKVFST PHGLEVHVRR
       260       270       280       290       300
SHSGTRPFAC DICGKTFGHA VSLEQHRAVH SQERSFDCKI CGKSFKRSST
       310       320       330       340       350
LSTHLLIHSD TRPYPCQYCG KRFHQKSDMK KHTFIHTGEK PHKCQVCGKA
       360       370       380       390
FSQSSNLITH SRKHTGFKPF GCDLCGKGFQ RKVDLRRHKE TQHGLK
```

ii Gfi1ab^sh320 protein coding sequence

```
        10        20        30        40        50
MPRSFLVKSK KAHSYHQPRS LEDDFNRLDN ILAHICAESK TSDESECCAD
        60        70
ALTGDTTDWR WFTRLSPSGS S
```

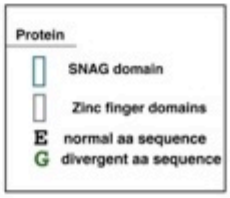
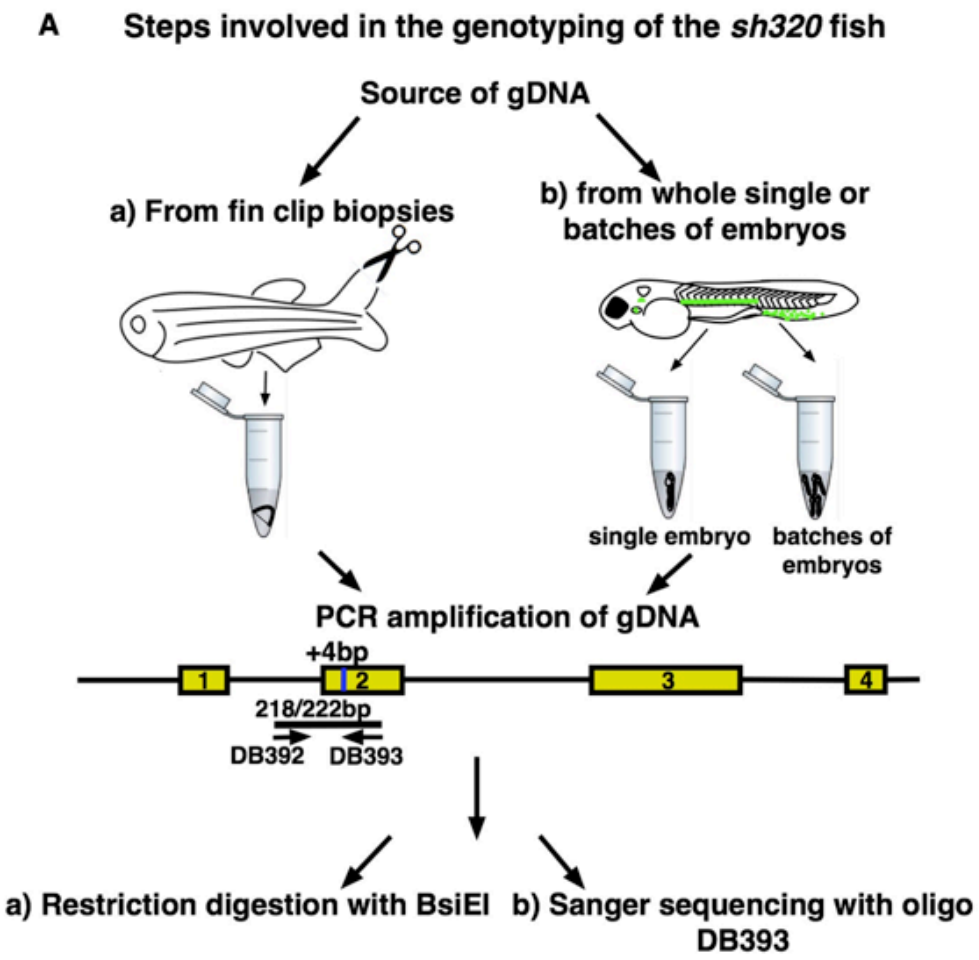


Figure 6.1: The *sh320* has a four base pair (4bp) insert that causes a frame shift and thus a divergent amino acid coding sequence.
(A) Schematically displays the *gfi1ab* transcript that shows the 6 coding exons, the SNAG domain is marked at the 5' end of exon1. Following the SNAG domain but continuing in exon1 is the linker domain. The linker domain ends just before the 3' end of exon3. The DNA binding domains are encoded by the 3' end of exon3 continue through from exons 4 to 6. The *sh320* mutant has a 4bp insertion/duplication, which is close to the 5' end of exon2 was characterised and identified by PCR amplification using oligos DB392 and DB393. Bi) shows the *gfi1ab WT* and ii) *sh320* mutant gDNA sequences, where in the mutant the 4 bp duplication causes a frame shift and thus causes a premature stop. The 4bp insertion also introduces a BsiEI restriction site. Ci) Displays the amino acid coding sequence of the WT Gfi1ab sequence that marks the 20aa SNAG domain and the 6 zinc finger domains. ii) in the *sh320* mutant the SNAG domain is conserved but the aa sequence is replaced by a divergent aa sequence from 58aa and the DNA binding domain is lost.

Double heterozygous *qmc551/+; sh320/+* fish survived to adulthood (Rob Wilkinson and Deniz Ucanok personal communication), whether the double homozygous fish were viable too was unclear. In order to generate the double homozygous embryos we incrossed the double heterozygous fish to get embryos that included double homozygous carriers.

To identify the *sh320* carriers in the subsequent generations it was decided to make use of the BsiEI site that was fortuitously introduced by the 4bp mutation. Fin clip biopsies of adult fish, single embryos or batches of embryos were used to prepare the gDNA that served as the template for PCR amplification. In the PCR reaction, oligos DB392 and DB393 were used to amplify a 218bp and 222bp fragment (Figure 6.2A) in wt and mutant DNA, respectively, that included the sequence duplication in the *sh320* allele. The PCR product was subsequently digested with BsiEI, which cuts the *sh320* allele in half and leaves the wt allele uncut. The amplified gDNA could also be Sanger sequenced using either oligos DB392 or DB393 (Figure 6.2A), which could further distinguish between $gfi1ab^{+/+}$, $gfi1ab^{sh320/+}$ and $gfi1ab^{sh320/sh320}$ carriers and confirm the genotypes of the fish analysed. A previous PhD student in the lab, Deniz Ucanok performed these genotyping experiments to identify $gfi1ab^{+/+}$, $gfi1ab^{sh320/+}$ and $gfi1ab^{sh320/sh320}$ carriers. However, the genotyping did not seem to work efficiently. When on the same gDNA sample a PCR and restriction digestion was repeated, occasionally, the data were inconsistent, suggesting that the experiment was not reproducible. Likewise when on the same gDNA sample an amplification and restriction digestion assay was done and she tried to confirm the results using PCR amplification and sequencing, there were in consistencies, a heterozygous pattern on the gel turned out to be homozygous as per the sequencing data obtained (Figure 6.2B) Therefore this would suggest that either the (a) oligos used to amplify the 218bp fragment was preferentially amplifying the mutant allele or (b) the restriction digest was an

incomplete digest, which could be due to the restriction enzyme, BsiEI not cutting efficiently. Due to time constraints these questions could not be addressed further.


A
Steps involved in the genotyping of the sh320 fish
Source of gDNA
a) From fin clip biopsies
b) from whole single or batches of embryos
single embryo
batches of embryos
PCR amplification of gDNA
+4bp
1
2
3
4
218/222bp
DB392
DB393
a) Restriction digestion with BsiEI
b) Sanger sequencing with oligo DB393

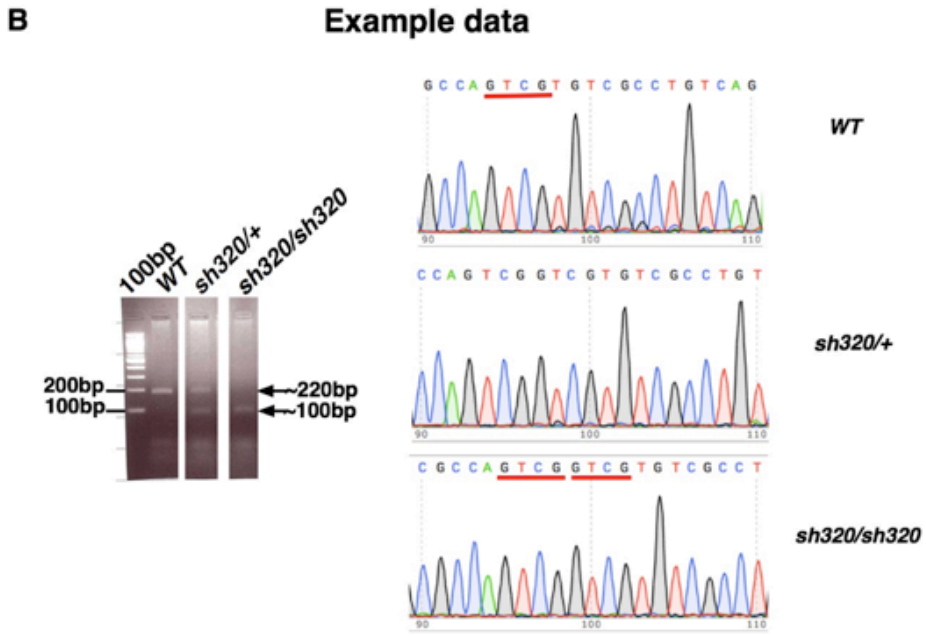

B
Example data
100bp
WT
sh320/+
sh320/sh320
200bp
100bp
~220bp
~100bp
WT
sh320/+
sh320/sh320

Figure 6.2: Genotyping of the *sh320* fish was unreliable and non reproducible.
Schematic representation of the steps involved in the initial genotyping of the *sh320* carriers. (A) Depicts the sources of gDNA, which were either from fin clips, single embryos or batches of embryos. The gDNA obtained was then PCR amplified using DB392 and DB393 followed by a BsiEI restriction digestion and agarose gel electrophoresis or Sanger sequencing (B) shows representative data on gDNA that was either PCR amplified and digested or PCR amplified and sequenced using oligo DB393 (Representative data images of gel and sequences was provided by Deniz Ucanok).

6.2 A single nucleotide polymorphism (SNP) in the *gfi1ab wt* allele interfere with characterisation and identification of *sh320* mutants

We revisited the genotyping of the *sh320* allele to resolve the issues. Embryos derived from an incross between double heterozygous *qmc551;sh320* carriers were selected for eGFP high expression and were then raised to adulthood. Fin clip biopsies were prepared from ten individual adult fish and gDNA was extracted. Using oligos DB392 and DB393, the 218/222bp fragment was amplified (Figure 6.3Ai). We did not carry out the restriction digestion, as they seemed to be unreliable. Instead, a 1.3% agarose gel was run to visualize the ~220bp fragment (Figure 6.3Ai). The fragment was subsequently purified and sequenced (Figure 6.3Bi-ii) using DB393. The sequencing data revealed that 3/10 samples, FC2, FC4 and FC10, contained a single "GTCG" nucleotide sequence and displayed a wt sequence pattern (Figure 6.3Bi) and Table 6.1. By contrast, 7/10 fish, FC1, FC3, FC5, FC6, FC7, FC8, FC9 and FC10 contained the duplicated 4bp (GTCG) sequence (Table 6.1) and (Figure 6.3Bii). An *sh320* heterozygous sequence was not found. Assuming Mendelian segregation of alleles, which is expected to be a 25 (*wt*): 50 (*hets*): 25(*homs*), this outcome of 30(*wt*): 70(*homs*) seemed surprising and unexpected.

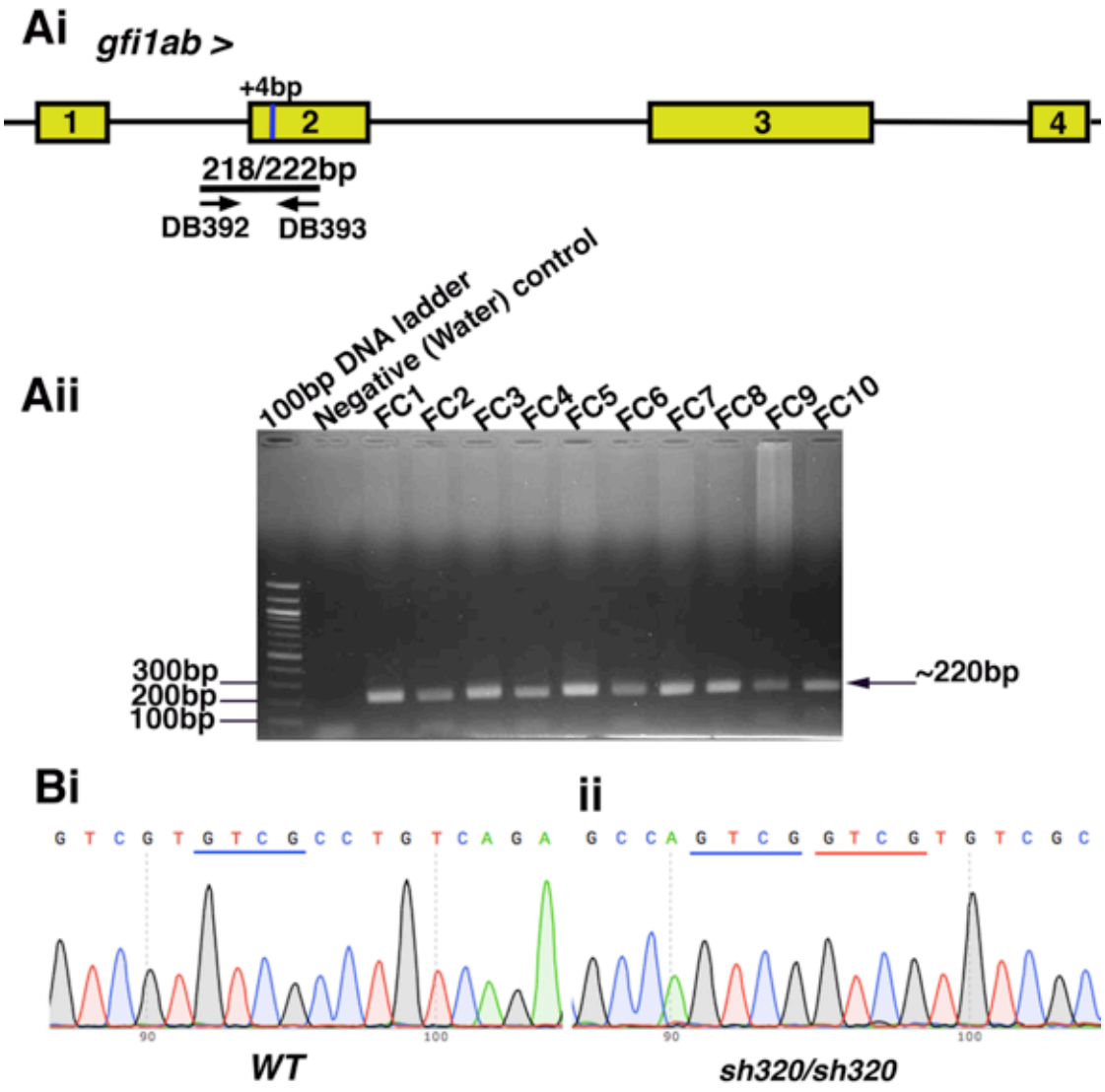


Figure 6.3: Genotyping of these fish revealed either a *wt* or *sh320* homozygous sequence.
(Ai) A schematic representation of the *gfi1ab* locus displaying the insertion/duplication within the *sh320* mutant line. The oligos DB392 and DB393 were used to PCR amplify the region around the insertion to obtain a 218bp (~220bp) fragment. (Aii) A 1.3% agarose gel displays a 100bp DNA ladder in the first lane with the 100bp, 200bp and 300bp marks highlighted. The second lane contains a negative water control. The samples FC1-FC10 all display the ~220bp amplified fragment. (Bi-ii) Subsequent sequencing data showed the presence of either a (i) *wt* sequence pattern, which contains a single "GTCG" or a (ii) *sh320/sh320* sequence pattern was observed, which contains the "GTCG" as well as a duplication of it (highlighted in red in the sequence). (Image of chromatograms taken from SnapGene Viewer version 3.2)

Table 6.1: Genotypes with respect to *gfi1ab* of fin clip biopsies taken from ten *qmc551* homozygous fish using oligos DB392 and DB393

Fin Clip (FC) numbers	*gfi1ab* Genotype
FC1	*sh320/sh320*
FC2	+/+
FC3	*sh320/sh320*
FC4	+/+
FC5	*sh320/sh320*
FC6	*sh320/sh320*
FC7	*sh320/sh320*
FC8	*sh320/sh320*
FC9	*sh320/sh320*
FC10	+/+

To verify the genotyping data, we decided to cross three of the identified *sh320* homozygous fish to *wt* fish to see whether all of the progeny were *sh320* heterozygous. Initially, three embryos from the cross between FC7 crossed to *wt* were randomly selected. The gDNA obtained from the individual embryos was amplified using oligos DB392 and DB393 to obtain a ~220bp fragment. The fragment was cleaned up and sequenced. The sequencing revealed that one out of the three embryos displayed a *wt* sequence (data not shown). Thus, our data strongly suggested that the PCR on the gDNA from the fin clips

had failed to amplify the *wt* allele. This may be due to a polymorphism in the sequence bound by one of the oligos used to PCR amplify the region around the duplication.

In order to check whether a polymorphism did indeed interfere with the genotyping, we designed 2 oligos DB752 and DB751 upstream and downstream of DB392 and DB393 respectively (Figure 6.4Ai). The new set of oligos was used to amplify a ~690bp fragment, initially just on gDNA from FCs 2, 7 and 8. The PCR products were separated on a 1.5% agarose gel (Figure 6.4Ai-ii). The ~690bp fragment was subsequently purified and sequenced. Initially, the sequencing was carried out using oligos DB392, DB393 and DB752. However, only the sequencing reaction with DB393 returned readable sequences. Close inspection of these sequences showed, FC7 and FC8 gDNA contained both *wt* and *sh320* mutant sequences, suggesting that the fish were *sh320* heterozygous. By contrast, FC2 gDNA turned out to be *wt* (Figure 6.4Bi-iii). This *wt* sequence also revealed a single nucleotide polymorphism (SNP) downstream of the 'GTCG' sequence that was located within the sequence bound by oligo DB392. A "thymine (T)" was replaced by an "adenine (A)" (Figure 6.4Ci-iii). Due to the SNP in the *wt* sequence, oligo DB392 would preferentially bind to the *sh320* mutant allele, if present in the same DNA sample, biasing DNA amplification in favour of the *sh320* allele. Thus, a heterozygous DNA sample would, depending on the PCR conditions, return a homozygous genotyping result.

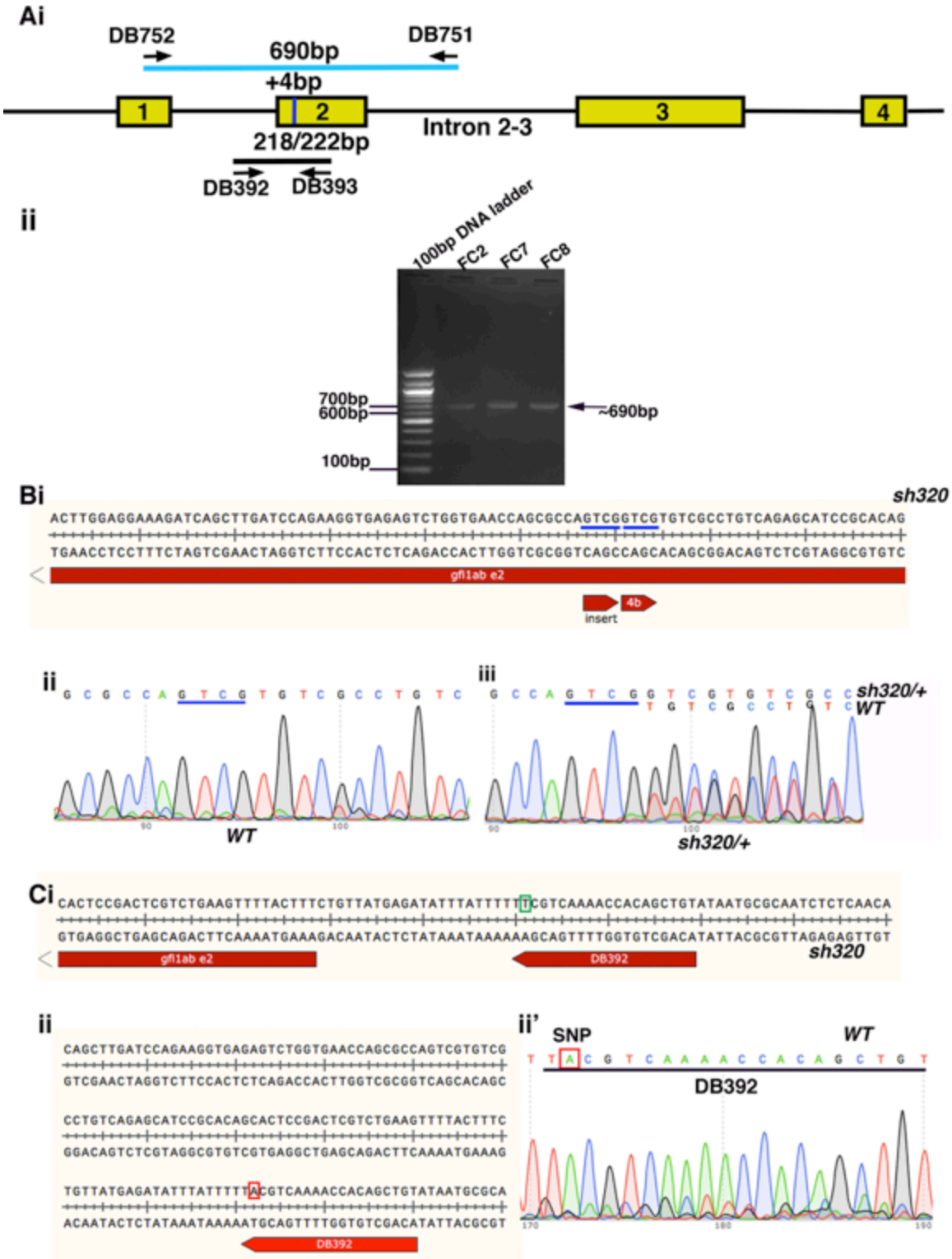

Ai
DB752
690bp
DB751
+4bp
1
2
3
4
Intron 2-3
218/222bp
DB392
DB393
ii
100bp DNA ladder
FC2
FC7
FC8
700bp
600bp
100bp
~690bp
Bi
sh320
ACTTGGAGGAAAGATCAGCTTGATCCAGAAGGTGAGAGTCTGGTGAACCAGCGCCAGTCGGTCGTGTCGCCTGTCAGAGCATCCGCACAG
TGAACCTCCTTTCTAGTCGAACTAGGTCTTCCACTCTCAGACCACTTGGTCGCGGTCAGCCAGCACAGCGGACAGTCTCGTAGGCGTGTC
gfi1ab e2
4b
insert
ii
WT
iii
sh320/+
WT
sh320/+
Ci
CACTCCGACTCGTCTGAAGTTTTACTTTCTGTTATGAGATATTTATTTTTTCGTCAAAACCACAGCTGTATAATGCGCAATCTCTCAACA
GTGAGGCTGAGCAGACTTCAAAATGAAAGACAATACTCTATAAATAAAAAAGCAGTTTTGGTGTCGACATATTACGCGTTAGAGAGTTGT
gfi1ab e2
DB392
sh320
ii
CAGCTTGATCCAGAAGGTGAGAGTCTGGTGAACCAGCGCCAGTCGTGTCG
GTCGAACTAGGTCTTCCACTCTCAGACCACTTGGTCGCGGTCAGCACAGC
CCTGTCAGAGCATCCGCACAGCACTCCGACTCGTCTGAAGTTTTACTTTC
GGACAGTCTCGTAGGCGTGTCGTGAGGCTGAGCAGACTTCAAAATGAAAG
TGTTATGAGATATTTATTTTTACGTCAAAACCACAGCTGTATAATGCGCA
ACAATACTCTATAAATAAAAATGCAGTTTTGGTGTCGACATATTACGCGT
DB392
ii'
SNP
WT
DB392

Figure 6.4: A polymorphism present in the *gfi1ab wt* allele was identified.
(Ai) Schematically represents the *gfi1ab* locus in which the 4bp insertion/duplication is situated within the *sh320* mutant allele. Oligos DB392 and DB393 were initially used to amplify a ~220bp fragment around the 4bp insertion. A new set of oligos DB751 and DB752 were designed outside of DB392 and DB393 to amplify fragment a 690bp long. (ii) gDNA amplified from FC2, 7 and 8 to visualize a 690bp fragment on a 1.5% agarose gel along side a 100bp DNA ladder (NEB). Bi) Displays the sh320 gDNA sequence highlighting the "GTCG" and its duplication (underlined in blue) ii) Sanger sequencing displaying a *gfi1ab wt* sequence pattern, chromatogram image taken on Snap Gene Viewer version 3.2, highlights the GTCG (underlined in blue). ii) Sanger sequencing displaying a heterozygous *sh320* sequence pattern, chromatogram image taken on Snap Gene Viewer version 3.2, highlights the GTCG (underlined in blue), after which the sequence runs into double peaks that represent both *wt* and *sh320* mutant nucleotide sequences. Ci) shows the *sh320* gDNA sequence that highlights (in green) a 'T' in oligo DB392. Cii-ii') Highlights (in red) the polymorphism in the *wt* (ii) gDNA sequence 'A' instead of a 'T' and (ii') The Sanger sequence chromatogram highlights the polymorphism 'A' (in red) within the DB392 sequence.

Next, we repeated the genotyping on the remaining fin clip biopsies using oligos DB751 and DB752 for the PCR amplification and oligo DB393 was used for the sequencing (Figure 6.5). Unfortunately, the PCR on gDNA from FC9 had failed; and the sequencing could not be performed. DNA sequencing on the gDNAs of the other fin clip biopsies revealed that FCs 1, 3, 5 and 6 contained both *wt* and *sh320* mutant sequences, suggesting that the fish were *sh320* heterozygous. By contrast, FCs 4 and 10 gDNA turned out to be *wt* (Table 6.2). Thus, altogether, our results revealed that all *sh320* carriers were heterozygous, suggesting that homozygous *qmc551* fish that were also homozygous for *sh320* were not viable under the conditions, which they were raised.

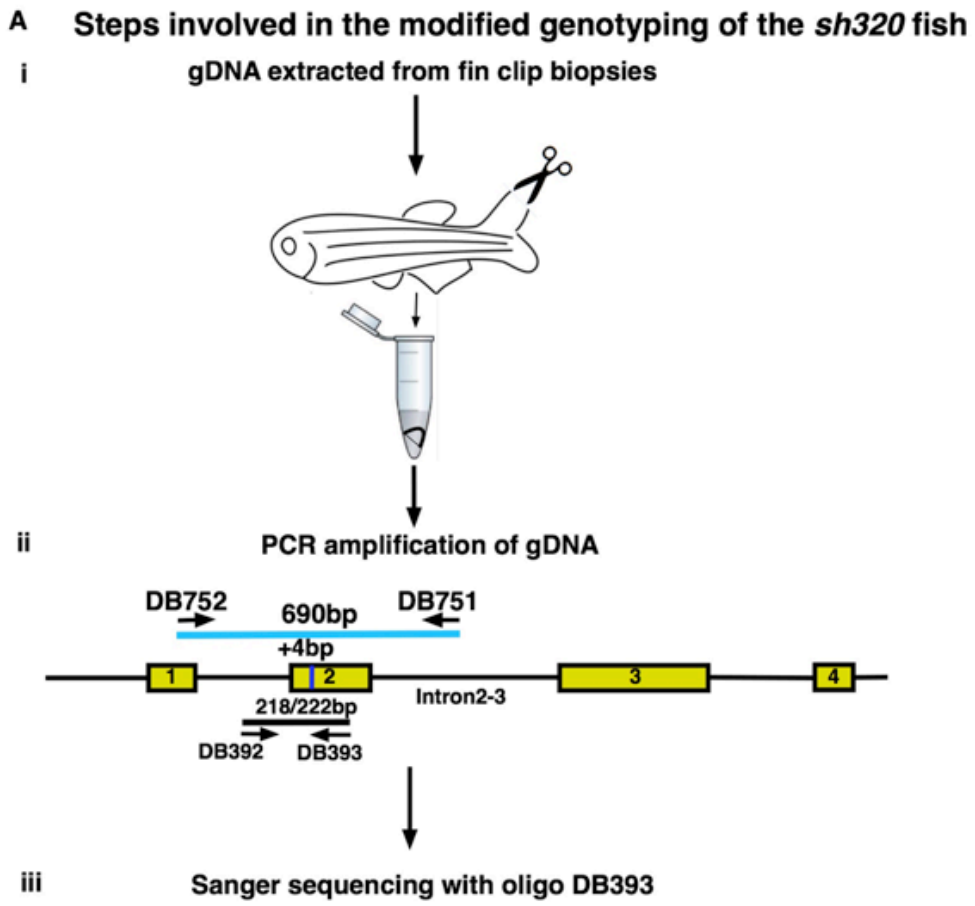


Figure 6.5: Genotyping of the *sh320* fish was modified and corrected.
Schematic representation of the steps involved in the genotyping of the *sh320* carriers. Ai-iii) Depicts gDNA extraction from fin clips. The gDNA obtained was then PCR amplified using DB751 and DB752 outside of oligos DB392 and DB393 to amplify a ~690bp fragment, which was subsequently sequenced using oligo DB393.

Table 6.2: Genotypes of finclipped fish repeated after identification of the polymorphism using oligos DB751 and DB752.

FC numbers	*gfi1ab* genotype
FC1	*sh320*/+
FC2	+/+
FC3	*sh320*/+
FC4	+/+
FC5	*sh320*/+
FC6	*sh320*/+
FC7	*sh320*/+
FC8	*sh320*/+
FC9	PCR failed
FC10	+/+

6.3 Characterisation of double mutants

Next, the six identified *qmc551 homozygous; sh320 heterozygous* fish were used in incrosses to generate embryos that are *qmc551/qmc551* as well as *sh320/sh320*. Because the parents were homozygous for *qmc551*, all of their progeny was eGFP+ high, demonstrating that the embryos were homozygous for the *qmc551Gt* allele of *gfi1aa*. Among these embryos, according to Mendelian genetics, we would expect 25% to carry 2 *wt gfi1ab* alleles, half to be *sh320* heterozygous and another quarter to be *sh320*

homozygous carriers. These embryos were then used to characterise a phenotype that may be caused due to the loss of both *gfi1a* paralogues. Under the microscope, all embryos of the batch displayed the same normal phenotype. I, therefore, performed a number of WISH experiments to learn more about haematopoiesis in these embryos.

6.3.1 *gfi1ab* mRNA levels are unchanged in the vDA of *qmc551* single homozygous and *qmc551;sh320* double homozygous embryos

The stability of mRNA can be compromised by the presence of a premature stop codon. Such a transcript may be recognized and degraded by the nonsense-mediated decay (NMD) pathway (Kervestin and Jacobson, 2012; Lykke-Andersen and Jensen, 2015; Saveanu and Jacquier, 2016). A typical target for the NMD pathway has a stop codon positioned more than 50-55 nucleotides upstream of the last exon-exon junction or has a long 3' untranslated region (UTR) (Lykke-Andersen and Jensen, 2015). The mRNA encoded by the *sh320* allele meets the first criterion in that it carries a premature stop codon approximately 700 bp upstream of the last exon-exon junction. To test whether the *sh320* mRNA is a target for NMD, *gfi1ab* mRNA levels were examined in *qmc551* single and *qmc551;sh320* double homozygous embryos. At 27-28 hpf, single *qmc551* homozygous embryos display high levels of *gfi1ab* mRNA in the vDA (Ucanok, 2015). If the presence of a premature stop codon induced NMD, one might be able to detect a reduction of *gfi1ab* transcript levels by WISH in double homozygous embryos. Embryos used in the WISH experiments were derived from an incross of *qmc551* homozygous, *sh320* heterozygous fish. The WISH experiments revealed that *gfi1ab* mRNA levels was very similar in the vDA of all 17 embryos (Figure 6.6 Ai-iii). While this semi-quantitative assay suggests that the *gfi1ab*^{sh320} mRNA is not subject to the NMD pathway it remains possible that NMD-mediated degradation is balanced by increased *gfi1ab* transcription in the double mutant embryos.

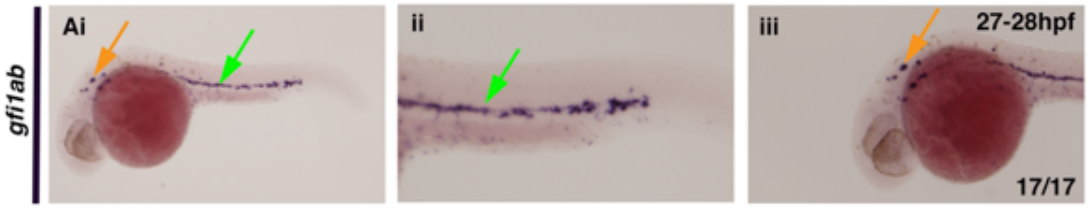


Figure 6.6: *gfi1ab* mRNA is unchanged in the vDA of double mutant embryos.
(A) 27-28hpf embryos obtained from an in cross between qmc551 homozygous; sh320/+ were fixed and WISH was carried out to detect gfi1ab mRNA which is detected in the inner ear (orange arrows) and in endothelial cells of the vDA (green arrows). The embryo is presented in lateral view, with dorsal facing up and anterior of the embryo facing left.

6.3.2 *gfi1b* expression remains restricted to prRBCs in double mutant embryos

In the absence of Gfi1aa in *qmc551* homozygous embryos, *gfi1ab* expression is upregulated in the vDA (Ucanok, 2015). In the same embryos the third paralogue, *gfi1b* expression is restricted to prRBCs and is not observed in the ECs of the vDA (Ucanok, 2015). In order to investigate whether *gfi1b* expression is upregulated in the absence of both *gfi1aa* and *gfi1ab* transcription factors at 27 to 28hpf *gfi1b* WISH experiments were performed on embryos derived from an incross of *qmc551 homozygous;sh320 heterozygous* fish. In most embryos, *gfi1b* mRNA was seen in spherical cells around the YS and PBI only (Figure 6.7 Ai-iii). In some of the embryos *gfi1b* expressing cells were also seen in the trunk region, but these cells were spherical and seemed to be in circulation (Figure 6.7 Bi-ii). A transverse section through the trunk of the same embryo confirmed the presence a round prRBC within the DA. There was no apparent expression in spindle-shaped ECs of the vDA (Figure 6.7 Biii), suggesting that a loss of Gfi1aa and Gfi1ab does not lead to an up-regulation of Gfi1b in the vDA.

The *gfi1b* WISH experiment suggested that prRBCs develop normally in Gfi1aa and Gfi1ab-depleted embryos. To confirm this finding, expression of the erythroid specific

marker gata1 was investigated. *Gata1* expression was detected by WISH at 20hpf when prRBCs are in the ICM (Figure 6.7 Ci-ii). There was no change in *gata1* expression within embryos of the batch. Altogether, this would suggest that primitive erythropoiesis is compatible with depletion of Gfi1aa and Gfi1ab.

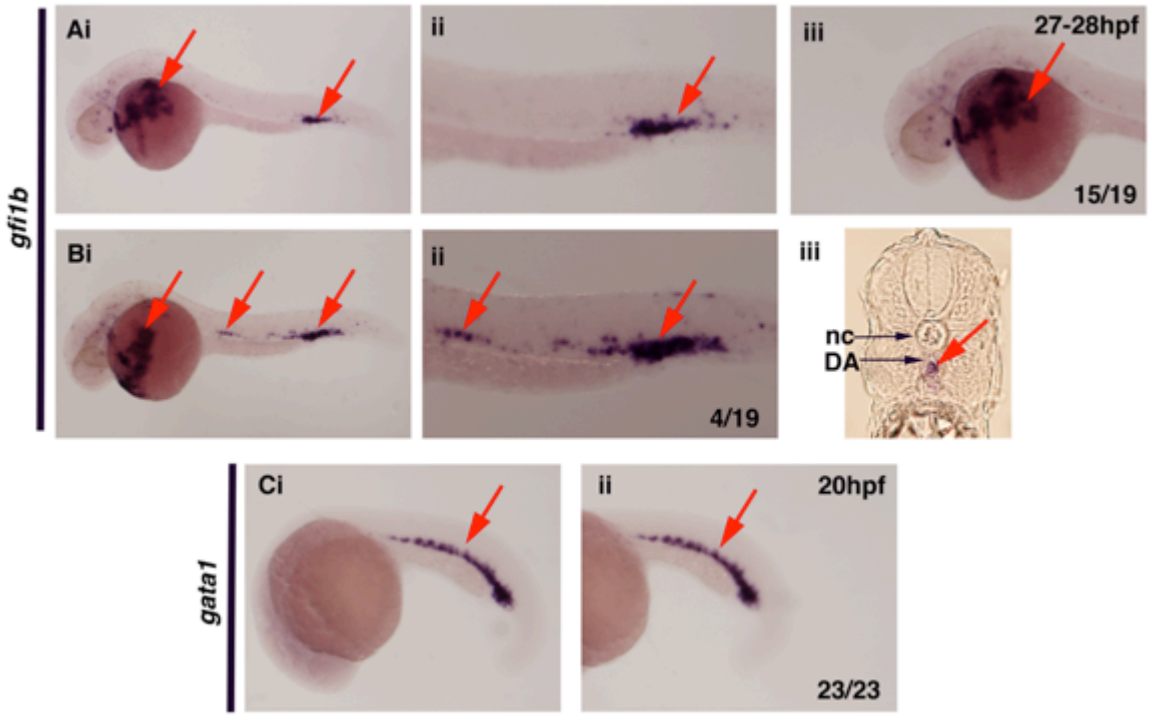


Figure 6.7: *gfi1b* mRNA is unchanged in double mutant embryos and *gata1* remains restricted to circular RBCs.
(A) 27 to 28hpf embryos obtained from an in cross between *qmc551* homozygous; *sh320/+* were fixed and WISH was carried out to detect (Ai-iii and Bi-iii) *gfi1b* mRNA (Bi-ii) displays circular cells with in the DA region (Ciii) a transverse section through the highlighted region displays the presence of a circular prRBC in the DA. (Ci-ii) 20hpf embryos obtained from an in cross between *qmc551* homozygous; *sh320/+* were fixed and *gata1* expression was detected by WISH. Embryos are presented in lateral view, with dorsal facing up and anterior of the embryo facing left. Orange arrows: inner ear, green arrows: endothelial cells of the vDA, red arrows: circular prRBCS, nc: notochord, DA: dorsal aorta.

6.4 Double homozygous mutant embryos display reduced *c-myb* expressing cells in the CHT, thymus and kidney

We then wanted to investigate whether the loss of both Gfi1aa and Gfi1ab proteins had a consequence on definitive haematopoiesis. For this purpose, embryos were collected and fixed at 2dpf, 3dpf and 5 dpf respectively and examined for *c-myb* expression. By 5dpf HSPCs from the CHT migrate to the larval kidney and T-lymphocytes progenitors seed the larval thymus (Paik and Zon, 2010), respectively. WISH analysis at 2dpf showed that the majority of embryos (15/19) displayed high level *c-myb* expression in the CHT and in non-hematopoietic tissue such as in the branchial arches (Figure 6.8Ai-ii). However four embryos (4/19) displayed a slight reduction of *c-myb* expression within the CHT, despite having normal staining in the region of the branchial arches (Figure 6.8Aiii-iv). At 3 dpf, more *c-myb*$^{+}$ cells were observed in the tail tissue as displayed by *c-myb* mRNA expression in 14/18 embryos (Figure 6.8Bi-ii). *c-myb* expression is significantly reduced in 4/18 embryos (Figure 6.8Biii-iv).

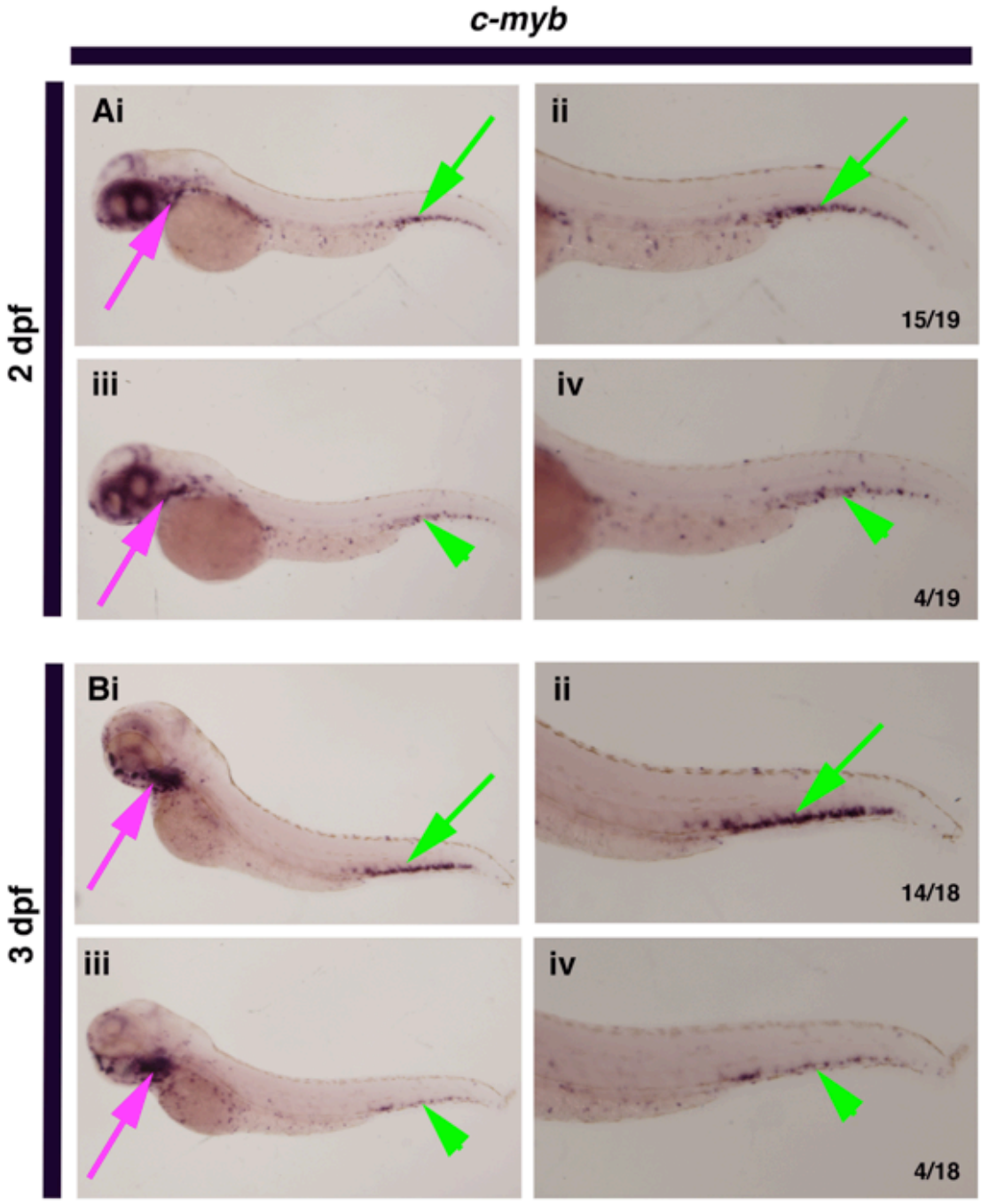


Figure 6.8: *c-myb* expression is significantly reduced from 3dpf within cells of the CHT. (A-B) *c-myb* mRNA expression at 2dpf and 3dpf detected by WISH in a mix of *qmc551 homozygous;sh320/+* or *gfi1ab*$^{+/+}$ or *sh320* homozygous (A) 2dpf embryos (i-ii) highlights cells within the CHT (iii-iv) shows a slight reduction in cells within the CHT. (B) 3dpf embryos (i-ii) Shows an embryo with a significant number of *c-myb* expressing cells in the CHT (iii-iv) an embryo that displays a significant reduction in cells of the CHT. All Embryos are displayed in lateral view dorsal side facing up and anterior facing left. Pink arrows: Branchial arches, green arrows: HSPCs in CHT, green arrowheads: reduction in expression in HSPCs within the CHT.

At 5dpf, HSPCs c-myb mRNA was detected by WISH within cells of the CHT, thymus and kidney, as well as in non-hematopoietic tissue such as in the branchial arches and in the retina (Figure 6.9i-iii). Although non-hematopoietic expression of *c-myb* mRNA is retained, 4/13 embryos showed a significant reduction in *c-myb*$^{+}$ cells within the three sites of haematopoiesis, i.e. in the CHT, thymus and kidney (Figure 6.9 iv-vi). These results would strongly suggest that a loss of both *gfi1a* paralogues causes a defect during definitive haematopoiesis.

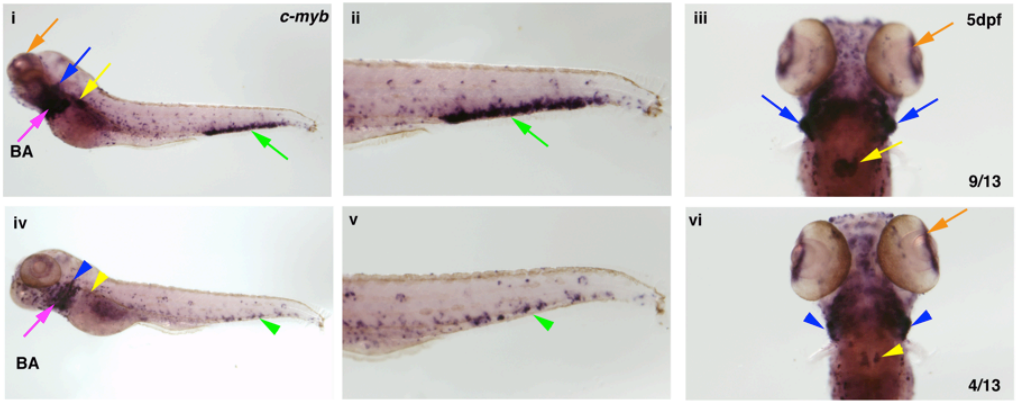


Figure 6.9: Reduced *c-myb* expression in sites of definitive haematopoiesis at 5dpf.
(i-vi) 5dpf embryos express *c-myb* mRNA as displayed by WISH experiments (i-iii) An embryo displaying *c-myb* mRNA expression in haematopoietic and non-haematopoietic tissues (iv-vi) a potential mutant, with significant reduction in *c-myb* expressing cells within haematopoietic tissue. (i-ii and iv-v) embryos are presented in lateral view, with dorsal facing up and anterior of the embryo facing left. (iii and vi) embryos presented in dorsal view. Pink arrows (BA): branchial arches, orange arrows: retina, green arrows: HSPCs in the CHT, blue arrows: thymus, yellow arrows: larval kidney, green arrows heads: reduction of HSPCs in the CHT, blue arrows heads: reduction of lymphocytes in the thymus, yellow arrowheads: reduction in HSPCs within the larval kidney.

To investigate whether the 5dpf embryos with the reduced levels of *c-myb* expression were indeed double homozygous mutants, we isolated their gDNA for *gfi1ab* genotyping. The same was done on 9 of the 5dpf embryos that displayed normal *c-myb* expression.

Table 6.3: The genotype of WISH embryos reveals the phenotype seen is indeed due to a loss of both *gfi1ab* copies along with the loss of *gfi1aa*.

Embryo selected as per phenotype displayed by WISH	***gfi1ab* genotype**
Embryos with normal c-myb expression	
Embryo 1	+/+
Embryo 2	+/+
Embryo 3	*sh320*/+
Embryo 4	+/+
Embryo 5	+/+
Embryo 6	*sh320*/+
Embryo 7	*sh320*/+
Embryo 8	*sh320*/+
Embryo 9	*sh320*/+
Embryos with a phenotype	
Embryo 1	*sh320/sh320*
Embryo 2	*sh320/sh320*
Embryo 3	*sh320/sh320*
Embryo 4	*sh320/sh320*

The genotyping results (Table 6.3) confirmed that the loss of *c-myb* expression correlates with the embryos being *sh320* homozygous. These data suggest that the loss of both *gfi1a* paralogues leads to a defect in definitive haematopoiesis. Interestingly, *sh320* heterozygous embryos displayed normal *c-myb* expression, suggesting that a single copy of one of the *gfi1a* paralogues is sufficient to maintain normal *c-myb* expression during definitive haematopoiesis.

6.5 Summary

- A *gfi1ab* mutant line, *gfi1ab*^{sh320} was established in Robert Wilkinson's Lab in Sheffield in order to investigate the effect of losing the gfi1a paralogues in qmc551 homozygous embryos. The *sh320* mutant allele carries a 4bp "GTCG" insertion/duplication at the 5' end of exon 2.
- The identification of the *sh320* mutation involved PCR amplification of gDNA using oligos DB392 and DB393 to amplify a 218/222bp fragment that needed to be subsequently cut using restriction enzyme BsiEI. This type of genotyping proved unreliable, but it remained unclear whether this was due to problems with PCR amplification or BsiE1 restriction.
- To resolve the genotyping issue two new oligos DB752 and DB751 were designed upstream and downstream of DB392 and DB393. Sequencing of the PCR products with oligo DB393 proved reliable for genotyping and revealed that an SNP segregating in the progeny of our wt fish had affected the binding of oligo DB392, biased the PCR amplification in favour of mutant allele and rendered previous genotyping efforts unreliable.
- Using the new method for identification of the *sh320* mutation we were thus able to correctly identify *sh320* heterozygous fish that were also *qmc551* homozygous.

These fish were incrossed to generate double mutant embryos that were then characterised.

- The expression of *gfi1ab* mRNA in cells of the vDA and inner ear remained unchanged in double homozygous embryos. *gfi1b* expression was also tested but does not seem to be upregulated in the vDA in the absence of the *gfi1a* paralogues. Erythroid specific *gata1* expression remains unchanged at 20hpf suggesting that primitive erythropoiesis is compatible with depletion of Gfi1aa and Gfi1ab.
- A significant reduction in *c-myb*$^{+}$ cells in the CHT, thymus and kidney of a quarter of embryos was observed at 5dpf. Genotyping these embryos confirmed the loss of *c-myb* expression correlates with embryos being *sh320* homozygous. These data suggest that a loss of both *gfi1a* paralogues causes a defect during definitive haematopoiesis.

7 Discussion

7.1 The constitutively active Cre recombination system we used labels the progeny of *flk1/kdrl*+ cells that seed the larval CHT, thymus and kidney

I used a constitutively active Cre/LoxP recombination system, similar to Bertrand et al to show successful reporter gene expression in the developing 5dpf embryo. The embryos displayed dsRed expressing cells in the CHT, thymi, intersomitic vessels, in the endocardium and in cells of the kidney. The presence of labelled cells within the CHT and within the larval thymus and kidney (the sites of larval and adult lymphopoiesis and haematopoiesis respectively) (Ellett and Lieschke, 2010), supported the findings that HSPCs are derived from *flk1*/*kdrl*+ cells (Bertrand et al., 2010). Bertrand et al. did not show the presence of reporter expression in the embryo. This was probably due to the fact that based on a conference abstract (6th European Zebrafish, Genetics and Development meeting held in Rome in 2009, Program and Abstracts p92), their *β-actin-switch* reporter line also shows dsRed express reporter expression in prRBCs of the embryo. This may suggest that the larger *flk1/kdrl* promoter fragment (7kb) drives expression earlier than our *flk1/kdrl* 0.8/1.5kb enhancer/promoter fragment in the posterior lateral mesoderm and in erythrocyte progenitors. We did not observe fluorescence in any circulating prRBCs suggesting that the *ef1α* promoter is not active in these cells. Since we did not see any circulating eGFP+ cells at 4dpf this would suggest that the *ef1α* promoter is not a good promoter to study prRBC activity. This property of the *ef1α* promoter worked to our advantage, as we did not have interfering prRBCs while imaging. The green-to-red reporter we used had an advantage over the reporter used by Bertrand and colleagues, where our reporter line allowed us to visualize and analyse cells before recombination as they

expressed the eGFP and subsequently analyse dsRed cells after a successful recombination event.

The fluorescent reporter protein dsRed was first detected on late 4 dpf, hence although recombination must have occurred in the early developing embryo, when the *flk1/kdrl* is switched on at around 8 to10ss (Choi et al., 2007). This could be firstly due to the *dsRed* mRNA which needs to be translated in order to make a sufficient amount of the fluorescent protein that can be visualised under the microscope. Therefore even though recombination takes place much earlier on in the developing embryo, dsRed expression is only seen after 4 dpf (Drew E. MSci, 2010). Secondly, dsRed has a very slow and incomplete maturation period (Baird et al., 2000) which further delays when the recombined cells can be detected.

Confocal imaging carried out on the CHT of 5 dpf embryos showed dsRed expressing elongated cells as well as round cells (Appendix 8 Bi and iii), indicating that the *flk1/kdrl* promoter was active in these cells, which thereby undergo recombination hence expressing the second reporter gene, i.e. *dsRed*. The presence of dsRed expressing spindle shaped ECs suggests that not all elongated cells undergo bEMT or EHT and not all of the EC cells are haemogenic in nature, which fits with the idea in the mouse, that only a subset of *flk1/kdrl*+ ECs are HECs (Jaffredo et al., 1998; Nishikawa et al., 1998; North et al., 2002). The confocal imaging also revealed the presence of round cells that were green and not red (i.e. not coming from flk1/kdrl+ cells) (Appendix 8 Bi and ii), this suggests that not all blood cells are derived from a *flk1/kdrl* expressing cell or these cells were mesenchymal cells that were not haematopoietic in nature.

The larvae analysed were then raised and analysed to adulthood and the adult WKM was analysed for the presence of recombination by FAC analysis (Appendix 1). Bertrand et al found 96% myeloid cells of the adult fish WKM showed reporter gene expression.

Admittedly, we had problems detecting reporter gene expression using FAC analysis. We found that the relative percentage of eGFP expressing cells flushed out of the WKM of a single transgenic *ef1-α* Cre reporter fish (expressed only eGFP) to be 0.45% (Appendix1 C) while a double transgenic displayed 0.06% eGFP expressing cells and 0.74% dsRed expressing cells (Appendix 1 D), which indicated that in the double transgenics almost all cells had undergone recombination in the adult fish WKM. These low percentages reflected the low activity of the *ef1α* promoter in adult cells, which has been reported to be most active in proliferating cells during gastrulation and mid-somitogenesis (Collins et al., 2010; Hans et al., 2009; Thummel et al., 2006). Unfortunately, these results were not reproducible hence we decided to detect recombination by genotyping the cells flushed out of the adult fish WKM. Here we observed 2/6 fish had undergone complete recombination while 4/6 should partial recombination showing that this system is stochastic (This will be discussed later in section 7.3).

7.2 The *flk1/kdrlCreER^{T2}* inducible system labels a subset of adult WKM blood cells showing that adult HSCs arise from *flk1/kdrl*+ cells of the early embryo alone.

Bertrand et al. had claimed using the constitutive system that cells found in the adult fish WKM at 6 months were derived from the early embryo alone (Bertrand et al., 2010), but in fact, without a temporal control on the activity of the Cre recombinase this claim cannot be made. After establishing a tight, non-leaky inducible system, which offered spatial control on Cre activity again under the control of the *flk1/kdrl* promoter (Choi et al., 2007) but this time additionally under temporal control with the administration of the ligand 4-OHT. 4-OHT was administered keeping in mind the commencement of the *flk1/kdrl* promoter activity, which begins at early somitogenesis and that drives the expression of the Cre recombinase. Ligand administration was also carried out keeping in mind the time of EHT

(Kissa and Herbomel, 2010). Using this system I was able to show the presence of recombined, dsRed expressing cells in the larval CHT, thymus and kidney, the sites of transitory haematopoiesis and larval and adult sites of lymphopoiesis and haematopoiesis respectively in the fish (Ellett and Lieschke, 2010). The CHTs of these fish were analysed using confocal microscopy, which revealed the presence of dsRed expressing cells within the tail tissue (Appendix 8 Ci and iii). This confirmed the generation of HSPCs from *flk1/kdrl*+ cells that were labelled in the early embryo between 0-5dpf. Experiments using the inducible CreER^{T2} system, carried out in the mouse have also suggested that adult blood cells are derived from ECs (Zovein et al., 2008). These experiments showed that HECs in the embryo contribute to adult haematopoiesis as (Chen et al., 2009; 2011; Zovein et al., 2008). However, experiments done in the mouse embryos do not allow visualization of recombined cells in the early embryo. Due to the transparent nature of the zebrafish embryo we can image and visualize live cells that have carried out recombination and that end up in the CHT of the early larvae. These embryos were then raised and adult fish WKM were analysed for the presence of recombined cells by PCR. We were thus able to show in at least 2/10 fish that a subset of blood cells of the WKM of the adult fish had undergone recombination. The adult bone marrow was found to contain anywhere from 2-24% of reporter labelled cells and it was reported that this might be due to TAM dose variation (Zovein et al., 2008). Here, PCR results in the inducible *CreER^{T2}* experiments revealed low recombination efficiency in the adult, which matched relatively low numbers of dsRed+ HCs in the early larvae (Figure 3.6E Appendix 8 Ci and iii). Poor recombination efficiency could either be due to inefficient 4-OHT treatments. Alternatively, it may indicate that the *flk1/kdrl* 0.8kb enhancer/1.5kb promoter elements (Choi et al., 2007) did not drive *CreER^{T2}* expression very efficiently as suggested by WISH to detect *CreER^{T2}* mRNA (Appendix 9). Nevertheless, we were able to demonstrate a

lineage relationship between *flk1/kdrl* derived early embryonic and adult HSCs in the zebrafish and that the cells found in the adult WKM were indeed derived from embryonic *flk1/kdrl*+ cells.

The system still has a few caveats: while it is possible that all blood cells are derived from *flk1/kdrl*+ ECs that underwent EHT/bEMT there is a possibility that blood cells express *flk1/kdrl* directly, without going through the EHT/bEMT; or even arising from primitive progenitors. These questions still remain unanswered.

7.3 Cells labelled and detected in the adult WKM of both constitutive and inducible systems depends on chance and genetic drift.

With the constitutive and inducible systems we were able to label *flk1/kdrl*+ cells and their progeny which we were able to detect by the expression of dsRed within the larval CHT, thymus and kidney and finally within the adult WKM. We found that the efficiency of recombination from both systems differed and that the number of dsRed labelled cells were more in both the larval CHT (Figures 3.5C and 3.6D and E and Appendix 8 panels B and C) and adult kidney (Figures 3.7B and C and 3.9) of fish that expressed *Cre* constitutively under the *flk1/kdrl* promoter. Once cells from the CHT seed the adult WKM the number of cells seeding the WKM niche depends on the number of cells that left the CHT. Since more cells were labelled in the CHTs of the early larvae that expressed a *Cre* constitutively, we expect more cells to seed the adult WKM here. In contrast, the induced larvae had fewer dsRed expressing cells in the CHT to begin with. Once these cells migrate to the WKM they divide. This could be briefly explained by one of the proposed theories (Snippert and Clevers, 2011), which could be explained in terms of HSCs in the WKM. HSCs within the WKM would divide through symmetrical cell division. The result of symmetrical division would be the formation of two daughter HSCs (Shahriyari and Komarova, 2013). After cell division, the fate of the daughter cells is determined

independently and stochastically such that two HSCs, two differentiated cells or one HSC and one differentiated cell could result (asymmetrical cell division) (Snippert and Clevers, 2011). Following cell division the theory of genetic drift could be applied to what we see regarding complete and partial recombination. Hence applying this theory to our results, in the 4/6 cases where partial recombination was observed, this would depend on the time point at which the WKM was analysed. If the WKM was analysed at an early time point (say time point 1 or 3 from Figure 7.1), then we would see be able to amplify both unrecombined and recombined fragments. If we analyse the WKM at a later time point (time point 5 Figure 7.1) i.e. when cells get fixed, then according to the genetic drift theory (Masel, 2011) we were able to amplify only the recombined fragment. Thus indicating that the cells in the WKM of those 2/6 fish have been fixed at having all dsRed+ cells. In comparison, the WKM cells of the induced fish displayed partial recombination in at least 2/10 fish indicating that the analysis was carried out at a point when the cells were not fixed. In the other fish, we had lost all dsRed-expressing cells because of which the recombined fragment was undetectable. This could be because we started off with a small number of labelled cells which may have divided in a way that the fixed population was only eGFP+ (Figure 7.1), hence we could detect only the unrecombined fragment in the gDNA of those fish.

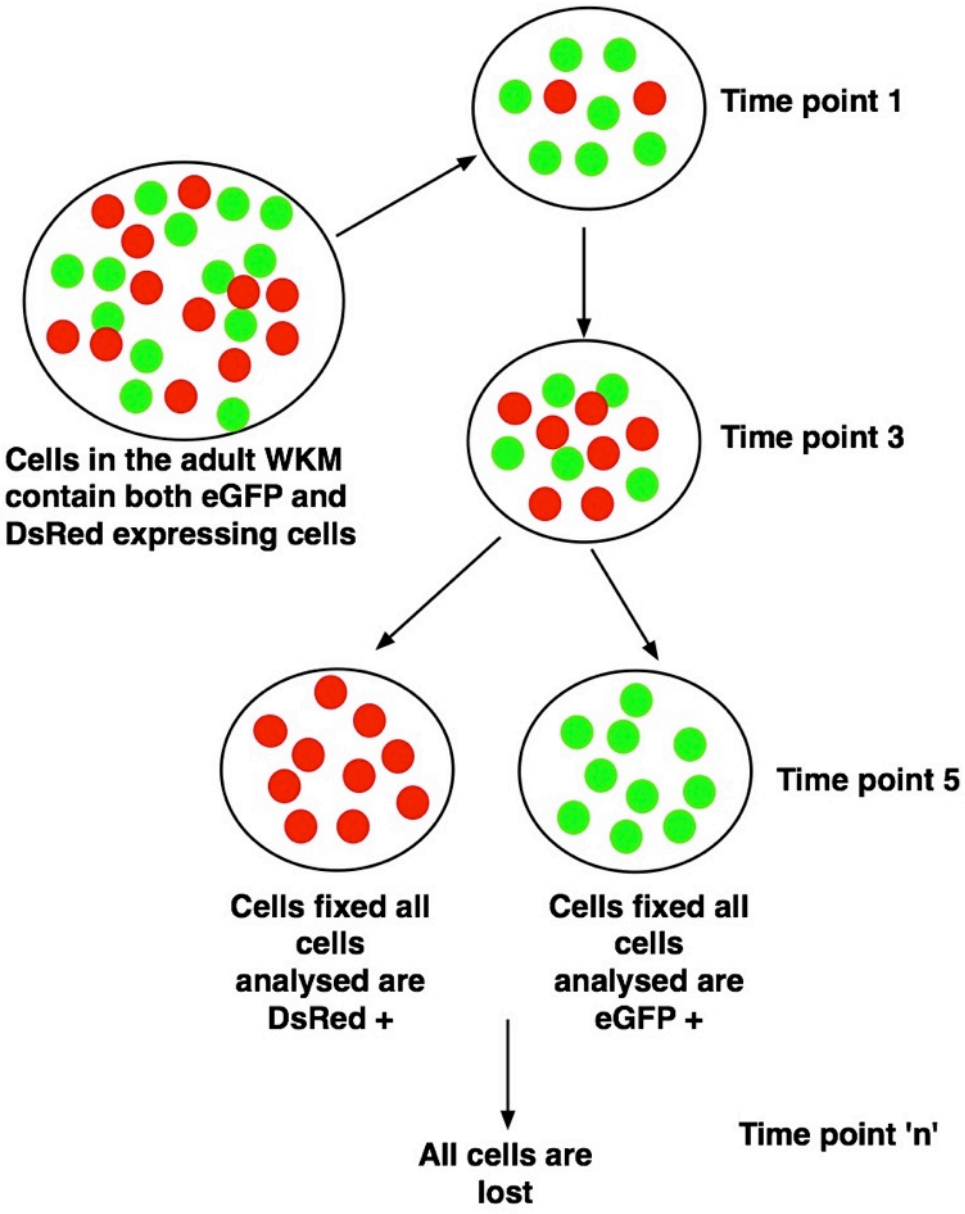


Figure 7.1: Schematic representation of the genetic drift theory applied to complete or partial recombination seen in the adult WKM marrow cells.
After cells in the WKM have undergone either symmetrical or asymmetrical cell division, depending on a given time point, say Time point 1 if cells are analysed here, we get a number of eGFP and dsRed expressing cells by the subsequent time points the divisions again depending on chance we get either a more number of green or red cells. At a certain time point say Time point 5 and the type of cell divisions we could get either all green or all red cells which get fixed when we analyse them. If we analyse them before they are fixed we will have a presence of both bands, for example at Time point 1 or at Time point 3. Beyond a certain time point say 'n' we may have lost all labelled cells.

Altogether we have shown that blood HSPCs arise from *flk1/kdrl*+ cells. A subset of labelled cells found in the adult WKM came from a cohort of cells that were labelled during a finite time window (0 to 5dpf) in the embryo. We also saw that the constitutive Cre system labelled more cells in the embryo as well as in the adult suggesting that the *flk1/kdrl* promoter might be active during later stages in development due to which the Cre recombinase carries out recombination in *flk1/kdrl*+ cells. A study carried out by Pelosi et al. on human cord blood haematopoietic progenitors that were activated using mouse hepatocyte medium, differentiate into adherent EC precursors, which have haemogenic property, capable of giving rise to blood cells. This study highlights the possibility of haemato-endothelial precursors arising *de novo* in human post-natal life, i.e. arising later during development (Pelosi et al., 2012).

7.4 Future prospects for lineage tracing-experiments

From both the constitutive and inducible systems, cells in the CHTs double transgenic embryos could be enumerated and embryos grown to adulthood separately depending on the numbers of dsRed expressing cells present. Once the embryos are adults, blood cells flushed out of the kidney marrow can be analysed by FACS and percentages obtained could be compared to the numbers counted in the CHT of the early larvae. This could help correlate the number of dsRed+ cells counted in the early larval CHT the relative percentages of recombined reporter labelled cells found later in the adult. To quantify developing HSCs in the zebrafish single-cell and clonal analysis of HSCs was carried out by Henninger et al. Making use of the inducible $CreER^{T2}$ system in developing cd41:eGFP+ HSCs allowed them to study the contribution of targeted cells to adult blood. They also made use of the multicolour *Zebrabow* system with *draculin:*$CreER^{T2}$, which is active in early haematopoiesis to label HSCs in the early embryo with a unique colour.

Each HSC and its progeny is barcoded with a unique colour, thus allowing them to quantify the number of HSCs that arise during development (Henninger et al., 2016).

In our experiments the Cre reporter could be modified by replacing the *ef1α* promoter with a more HSC specific promoter/enhancer element such as *c-myb* to drive reporter gene expression in cells of the blood lineage. Another possibility would be to use the *qmc551* line and using the latest CRISPR/Cas9 technology to introduce a *Cre* between the splice acceptor and eGFP, to drive *Cre* activity in HECs and be able to visualize where Cre is expressed with the help of the eGFP in the *qmc551* line. And use a reporter line to trace the progeny of HECs.

7.5 *gfi1aa*^{qmc551Gt} labels HECs giving rise to HSPCs and serves as a good tool to study zebrafish HSC emergence

Previously, using a *flk1/kdrl:eGFP* transgenic line it was shown that ECs of the vDA undergo a bending and transition into blood cells that enter the mesenchyme underneath the DA in a process termed EHT (Bertrand et al., 2010; Kissa and Herbomel, 2010). Our data to show that these cells express the Notch reporter csl:cerulean, which confirms that these cells are arterial in nature.

The zebrafish *gfi1aa*^{qmc551Gt} isolated by Roshana Thambyrajah expresses an eGFP reporter inside and outside of the haematopoietic system. Within the blood system, eGFP was detected in prRBCs and HECs of the vDA and in HSPCs of the larval and adult sites of haematopoiesis. The expression seen in the *qmc551* embryos is wider than the expression pattern observed in a previously published *gfi1aa* enhancer trap line which was restricted to prRBCs (Wei et al., 2008). Using single slice confocal time-lapse imaging, I have shown for the first time that zebrafish HECs can undergo both apical and basal EMT (aD and bEMT) (Thambyrajah et al., 2016). In the zebrafish, bEMT/EHT was demonstrated in live

confocal time-lapse studies, which were carried out on embryos that carried endothelial specific reporter transgenes. In these studies, embryos co-expressed with the EC specific and blood specific transgenes that indicated the hematopoietic nature of the cells during or after EHT (Bertrand et al., 2010; Kissa and Herbomel, 2010; Kissa et al., 2008; Lam et al., 2010; Zhen et al., 2013).

The mechanisms of aD and bEMT as observed seemed to be distinct. aD was observed to be a sort of "budding" mechanism in comparison to the previously described EHT/bEMT mechanism which involves the "bending" of the EC of the vDA, which rounds up and enters the mesenchyme below (Kissa and Herbomel, 2010). Along with the apically departing cells there always seemed to be a protrusion towards the underlying mesenchyme. Whether this protrusion originated from the cell that departed apically or from another cell could not be resolved (Figures 4.5-4.7). It also remains to be seen whether the process entails an asymmetrical cell division, with one daughter cell entering the lumen of the DA and the other remaining in the endothelium (Tamplin and Zon, 2010). The production of a few mini vesicles was frequently observed associated below the cell that departed apically. Generation of these mini vesicles have previously been reported in Runx1 morphants (Kissa and Herbomel, 2010; Zhen et al., 2013) (Figure 7.2). We assume that in the morphant embryos these may be the failure to efficiently reprogram the cells in the absence of *Runx1* (Lancrin et al., 2012; Liakhovitskaia et al., 2014) leads to the maintenance of even stronger local interactions and the shedding of more mini vesicles. We speculate that these mini vesicles maybe extracellular vesicles such as micro vesicles or exosomes that are known to be involved in the transfer of genetic information in the form of mRNA, miRNA and proteins to recipient cells (Lee et al., 2012).

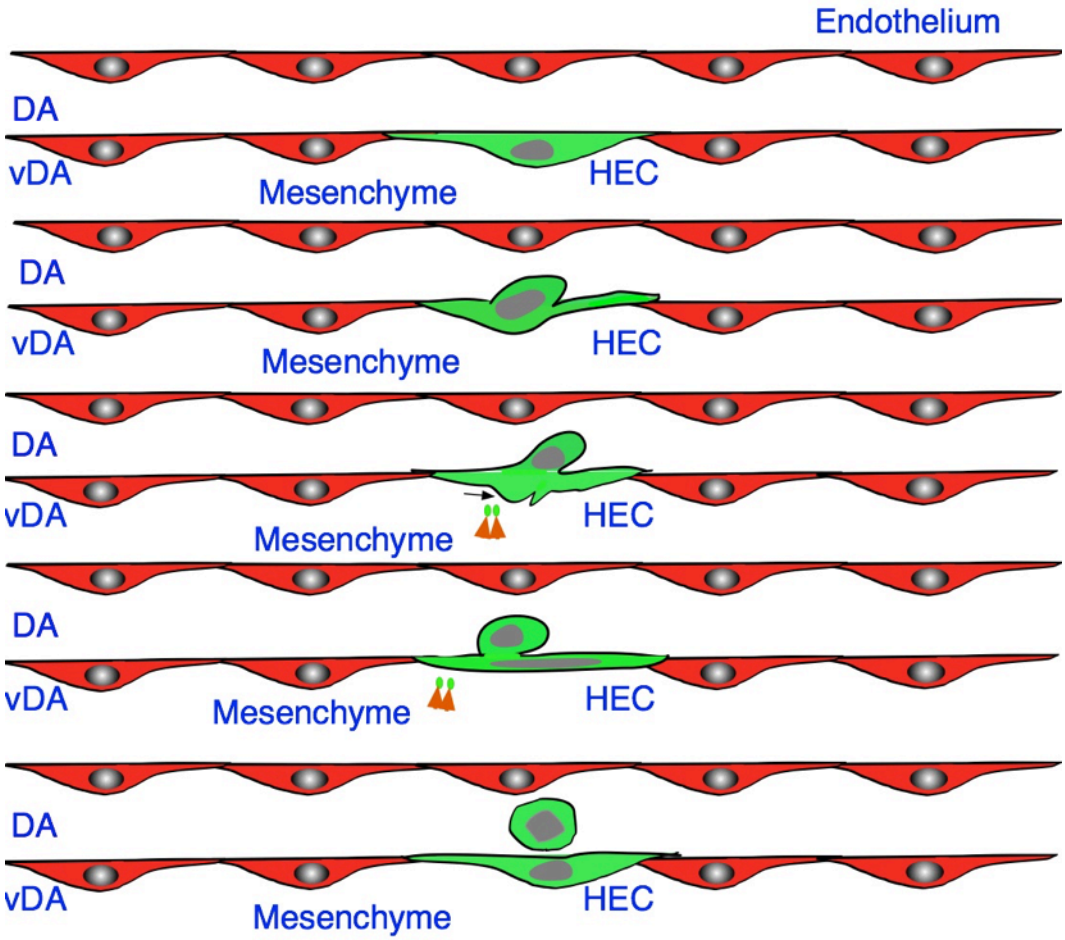


Figure 7.2: Schematic representation of the mechanism of apical EMT.
An HEC within the vDA slowly bulges upward or apically to form a bud which forms protrusions below in the mesenchyme and also associated to these protrusions are mini vesicles. The bud slowly detaches from the endothelium while the mini vesicles are released in the mesenchyme below. The cell enters the lumen of the DA and joins circulation. Green cell: is a hEC in the vDA, black arrow: protrusions, orange arrowheads: mini vesicles.

Besides the bEMT and aD differing in mechanism, they also differ in time scale. The average time taken for a cell to undergo apical departure was found to be 63 minutes, which is similar to the time frame taken for mouse HC clusters to bud into the lumen of the DA (approximately 1hour and 15 minutes) (Boisset et al., 2010). We have also shown that these apically departing cells have a nucleus and their sizes would suggest that these are indeed intact cells.

Altogether, the early expression of eGFP in HECs in *qmc551* embryos and its up regulation upon EHT makes it a good tool in comparison to using just EC specific transgenic embryos to visualize cell behaviour in the earliest events. The *qmc551* line gives us an exceptional early view on HSC emergence and cell behaviour in the zebrafish embryo. The recently published *gata2b:GAL4;UAS:GFP* double transgenic line seems to give a similarly early reporter expression (Butko et al., 2015). This line and the *qmc551* line could be good tools to visualize *in vivo* live imaging of early HSC emergence.

The apical departure of cells is a rare event (observed 10 times) that occurs only before 40hpf in *wt* embryos. In Runx1-morphant embryos where bEMT is blocked aD is more frequent and is observed beyond 40hpf. In zebrafish runx1 mutants it was reported that while most of the embryos die between 11 to 20 dpf, 20% of the embryos regained blood circulation by 15 to 21dpf and were seen to survive upto adulthood (Sood et al., 2010). Irrespective of the mechanism, aD may at least in part explain why those 20% runx1 mutants recover in the zebrafish whose development is clearly less dependent on haematopoiesis than the mammalian embryo (Hess et al., 2013).

The fate of apically departing cells is unclear, and it would be important to trace them. As they arise from the vDA we assume that they are blood progenitors that seed the larval CHT. Using light sheet microscopy, it may be possible to follow the cells. However, given

the low frequency of aEMT, the effort and time required to trace them appeared disproportionally high to the probably low contribution that these cells make to adult haematopoiesis.

7.6 The loss of *gfi1aa* and *gfi1ab* do not affect bEMT/EHT but cause a reduction in HSPCs within the larval CHT.

The zebrafish genome has three *gfi1* paralogues (Cooney et al., 2013; Dufourcq et al., 2004; Wei et al., 2008) *gfi1aa, gfi1ab* and *gfi1b*. Initially when discovered, the gfi1ab paralogue was reported to be expressed within cells of the the inner ear, the lateral line system, the brain and in the parapineal organ and hence was not considered as a gene of interest in haematopoiesis (Cooney et al., 2013; Dufourcq et al., 2004). In our lab using the *qmc551* line it was found that in the absence of *gfi1aa* in *qmc551* homozygous embryos, the expression of *gfi1b* is retained in prRBCs (Thambyrajah et al., 2016) and at 27-28 hpf there was an up regulation of *gfi1ab* transcript in the vDA (Ucanok, 2015), thus maintaining normal primitive and definitive haematopoiesis in the qmc551 homozygous fish. Here, using the *gfi1ab* mutants and the *qmc551* line to generate double mutants (that have lost both *gfi1aa* and *gfi1ab*), we show in these embryos that there was no obvious effect on the qmc551:eGFP cells of the vDA and mesenchyme below, i.e EHT seemed to be normal but in our preliminary data on the *qmc552* and *qmc553* compound heterozygous embryo M1 there was a reduction in eGFP expressing cells within the larval CHT and thymus on 5dpf (Figure 5.6). The double mutant *qmc551/qmc551;sh320/sh320* were analysed for *c-myb* mRNA and it was observed that there was reduced number of *c-myb* expressing cells in the larval CHT, thymus and kidney at 5 dpf (Figure 6.9). There was already a slight reduction of *c-myb* expression at 2dpf with in the CHT (Figure 6.8). The reduced *c-myb* expression would suggest that although the cells of the HECs of the vDA may undergo EHT they are unable to expand and differentiate into lineages as they seed

the larval CHT, thymus and kidney. In the mouse there are two *Gfi1* paralogues, *Gfi1* and *Gfi1b*. A phylogenetic analysis showed that the zebrafish *gfi1aa* and *gfi1ab* cluster most closely and are orthologues of the mammalian *Gfi1* gene. This suggests that they arose from the duplication of the single original *gfi1a* gene. The presence of two gene copies suggests possible functional redundancy. However, the divergence of their gene expression patterns indicates that the roles of the original *gfi1a* gene may have been split between *gfi1aa* and *gfi1ab* in the fish. By contrast, *Gfi1b* in the mouse clusters more closely to the *gfi1b* in the fish (Cooney et al., 2013). Mice that have lost *Gfi1b* have abnormal erythropoiesis and megakaryopoiesis and thus do not survive at gestation (Saleque et al., 2002). Only double knockout mice display a defect in HEC, where it has been reported that HECs are specified but fail to undergo EHT (Thambyrajah et al., 2015).

So this suggests that *Gfi1* and *Gfi1b* are (a) redundant in the mouse, the same as *gfi1aa/ab* in the fish and (b) likely regulating each other's expression, just like *gfi1aa/ab* in the fish. Hence in the fish we might need to knock all three paralogues, which would include the *gfi1b* along with *gfi1aa* and *gfi1ab* in order to generate triple mutants. These embryos can then be analysed to study the effect of the loss of gfi1a/b proteins on HECs and to see whether they undergo EHT or not.

References

Adamo, L., Naveiras, O., Wenzel, P.L., McKinney-Freeman, S., Mack, P.J., Gracia-Sancho, J., Suchy-Dicey, A., Yoshimoto, M., Lensch, M.W., Yoder, M.C., Garcia-Cardena, G., Daley, G.Q. (2009). Biomechanical forces promote embryonic haematopoiesis. Nature *459*, 1131–1135.

Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (2000). Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. Proc. Natl. Acad. Sci. U.S.A. *97*, 11984–11989.

Barrangou, R., Fremaux, C., Deveau, H., and Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315,* 1709-1712.

Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug, R.G., II, Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y.H., et al. (2012). In vivo genome editing using a high-efficiency TALEN system. Nature *490*, 114–118.

Bennett, C.M., Kanki, J.P., Rhodes, J., Liu, T.X., Paw, B.H., Kieran, M.W., Langenau, D.M., Delahaye-Brown, A., Zon, L.I., Fleming, M.D., et al. (2001). Myelopoiesis in the zebrafish, Danio rerio. Blood *98*, 643–651.

Bertrand, J.Y., Kim, A.D., Violette, E.P., Stachura, D.L., Cisson, J.L., and Traver, D. (2007). Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. Development *134*, 4147–4156.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y.R., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. Nature *464*, 108–111.

Bhaya, D., Davison, M., and Barrangou, R. (2011). CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. Annu. Rev. Genet. *45*, 273–297.

Boisset, J.-C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature *464*, 116–120.

Bordignon, C. (2006). Stem-cell therapies for blood diseases. Nature *441*, 1100–1102.

Broadbent, J., and Read, E.M. (1999). Wholemount in situ hybridization of Xenopus and zebrafish embryos. Methods Mol. Biol. *127*, 57–67.

Brotherton, T.W., Chui, D.H., Gauldie, J., and Patterson, M. (1979). Hemoglobin ontogeny during normal mouse fetal development. Proc. Natl. Acad. Sci. U.S.a. *76*, 2853–2857.

Brouns, S., Jore, M.M., Lundgren, M., and Westra, E.R. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. ….

Brown, M., and Wittwer, C. (2000). Flow cytometry: principles and clinical applications in

hematology. Clin. Chem. *46*, 1221–1229.

Burns, C.E. (2005). Hematopoietic stem cell fate is established by the Notch-Runx pathway. Genes & Development *19*, 2331–2342.

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.

Cai, X., Gaudet, J.J., Mangan, J.K., Chen, M.J., De Obaldia, M.E., Oo, Z., Ernst, P., and Speck, N.A. (2011). Runx1 Loss Minimally Impacts Long-Term Hematopoietic Stem Cells. PloS One *6*, e28430.

Cantor, A.B.A., and Orkin, S.H.S. (2002). Transcriptional regulation of erythropoiesis: an affair involving multiple partners. Oncogene *21*, 3368–3376.

Cellot, S., and Sauvageau, G. (2005). Gfi-1: another piece in the HSC puzzle. Trends Immunol. *26*, 68–71.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos. Cell Res *23*, 465–472.

Chen, M.J., Li, Y., De Obaldia, M.E., Yang, Q., Yzaguirre, A.D., Yamada-Inagawa, T., Vink, C.S., Bhandoola, A., Dzierzak, E., and Speck, N.A. (2011). Erythroid/Myeloid Progenitorsand Hematopoietic Stem Cells Originate from Distinct Populations of Endothelial Cells. Stem Cell *9*, 541–552.

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887–891.

Choi, J., Dong, L., Ahn, J., Dao, D., Hammerschmidt, M., and Chen, J.-N. (2007). FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. Developmental Biology *304*, 735–744.

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

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

Ciau-Uitz, A., Monteiro, R., Kirmizitas, A., and Patient, R. (2014). Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. Exp. Hematol. *42*, 669–683.

Clements, W.K., and Traver, D. (2013). Signalling pathways that control vertebrate haematopoietic stem cell specification. Nature Publishing Group *13*, 336–348.

Clements, W.K., Kim, A.D., Ong, K.G., Moore, J.C., Lawson, N.D., and Traver, D. (2011). A somitic Wnt16/Notch pathway specifies haematopoietic stem cells. Nature *474*, 220–224.

Collins, R.T., Linker, C., and Lewis, J. (2010). mAZe: a tool for mosaic analysis of gene function in zebrafish. Nature Publishing Group *7*, 219–223.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science *339*, 819–823.

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). Developmental Biology. Developmental Biology *373*, 431–441.

Cumano, A., Dieterlen-Lievre, F., and Godin, I. (1996). Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura. Cell *86*, 907–916.

Cumano, A., and Godin, I. (2007). Ontogeny of the Hematopoietic System. Annu. Rev. Immunol. *25*, 745–785.

de Bruijn, M.F.M., Speck, N.A.N., Peeters, M.C.M., and Dzierzak, E.E. (2000). Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. Embo J *19*, 2465–2474.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

Dieterlen-Lievre, F. (1975). On the origin of haemopoietic stem cells in the avian embryo: an experimental approach. J Embryol Exp Morphol *33*, 607–619.

Ditadi, A., Sturgeon, C.M., and Keller, G. (2017). A view of human haematopoietic development from the Petri dish. Nature Publishing Group *18*, 56–67.

Doulatov, S., Notta, F., Laurenti, E., and Dick, J.E. (2012). Hematopoiesis: A Human Perspective. Stem Cell *10*, 120–136.

Dufourcq, P., Rastegar, S., Strähle, U., and Blader, P. (2004). Parapineal specific expression of gfi1 in the zebrafish epithalamus. Gene Expression Patterns *4*, 53–57.

Dzierzak, E., and Medvinsky, A. (2008). The discovery of a source of adult hematopoietic cells in the embryo. Development *135*, 2343–2346.

Ebina, W., and Rossi, D.J. (2015). Transcription factor-mediated reprogramming toward hematopoietic stem cells. The EMBO Journal *34*, 694–709.

Ellett, F., and Lieschke, G.J. (2010). Zebrafish as a model for vertebrate hematopoiesis. Current Opinion in Pharmacology *10*, 563–570.

Feil, R.E.A. (1997). Regulation of Cre Recombinase Activity by Mutated Estrogen Receptor Ligand-Binding Domains. Biochem Biophys Res Commun *237,* 1–6.

Feil, R., Brocard, J., Mascrez, B., LeMeur, M., Metzger, D., and Chambon, P. (1996). Ligand-activated site-specific recombination in mice. Proc. Natl. Acad. Sci. U.S.A. *93*,

10887–10890.

Ferkowicz, M.J., and Yoder, M.C. (2005). Blood island formation: longstanding observations and modern interpretations. Exp. Hematol. *33*, 1041–1047.

Fiolka, K., Hertzano, R., Vassen, L., Zeng, H., Hermesh, O., Avraham, K.B., Dührsen, U., and Möröy, T. (2006). Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. Nature Publishing Group *7*, 326–333.

Frame, J.M., Fegan, K.H., Conway, S.J., McGrath, K.E., and 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, J.M., McGrath, K.E., and Palis, J. (2013). Erythro-myeloid progenitors: "Definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. Blood Cells, Molecules, and Diseases *51*, 220–225.

Gansner, J.M., Madsen, E.C., Mecham, R.P., and Gitlin, J.D. (2008). Essential role for fibrillin-2 in zebrafish notochord and vascular morphogenesis. Dev Dyn *237*, 2844–2861.

Georgiades, P., Ogilvy, S., Duval, H.L.N., Licence, D.R., Charnock-Jones, D.S., Smith, S.K., and Print, C.G. (2002). vavCre Transgenic mice: A tool for mutagenesis in hematopoietic and endothelial lineages. Genesis *34*, 251–256.

Gering, M., and Patient, R. (2005). Hedgehog Signaling Is Required for Adult Blood Stem Cell Formation in Zebrafish Embryos. Developmental Cell *8*, 389–400.

Gerlach, G.F., Schrader, L.N., and Wingert, R.A. (2011). Dissection of the Adult Zebrafish Kidney. JoVE 54, e2839.

Gilks, C.B., Bear, S.E., and Grimes, H.L., Tsichlis P. N., (1993). Progression of interleukin-2 (IL-2)-dependent rat T cell lymphoma lines to IL-2-independent growth following activation of a gene (Gfi-1) encoding a novel zinc finger protein. Molecular and Cellular biology *13,* 1759-1768.

Gore, A.V., Monzo, K., Cha, Y.R., Pan, W., and Weinstein, B.M. (2012). Vascular Development in the Zebrafish. Cold Spring Harbor Perspectives in Medicine *2*, a006684–a006684.

Goyama, S. (2004). The transcriptionally active form of AML1 is required for hematopoietic rescue of the AML1-deficient embryonic para-aortic splanchnopleural (P-Sp) region. Blood *104*, 3558–3564.

Gratwohl, A. (2010). Hematopoietic Stem Cell TransplantationA Global Perspective. Jama *303*, 1617.

Gray, C., Bratt, D., Lees, J., daCosta, M., Plant, K., Watson, O.J., Solaymani-Kohal, S., Tazzyman, S., Serbanovic-Canic, J., Crossman, D.C., et al. (2013). Loss of function of parathyroid hormone receptor 1 induces Notch-dependent aortic defects during zebrafish vascular development. Arteriosclerosis, Thrombosis, and Vascular Biology *33*, 1257–1263.

Grimes, H.L., Chan, T.O., Zweidler-McKay, P.A., Tong, B., and Tsichlis, P.N. (1996). The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol. Cell. Biol. *16*, 6263–6272.

Gupta, R.M., and Musunuru, K. (2014). Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. J. Clin. Invest. *124*, 4154–4161.

Hans, S., Freudenreich, D., Geffarth, M., Kaslin, J., Machate, A., and Brand, M. (2010). Generation of a non-leaky heat shock-inducible Cre line for conditional Cre/lox strategies in zebrafish. Dev Dyn *240*, 108–115.

Hans, S., Kaslin, J., Freudenreich, D., and Brand, M. (2009). Temporally-controlled site-specific recombination in zebrafish. PloS One *4*, e4640.

Hendriks, W.T., Warren, C.R., and Cowan, C.A. (2016). Protocol Review. Stem Cell *18*, 53–65.

Henninger, J., Santoso, B., Hans, S., Durand, E., Moore, J., Mosimann, C., Brand, M., Traver, D., and Zon, L. (2016). Clonal fate mapping quantifies the number of haematopoietic stem cells that arise during development. Nature Cell Biology *19*, 17–27.

Hess, I., Iwanami, N., Schorpp, M., and Boehm, T. (2013). Zebrafish model for allogeneic hematopoietic cell transplantation not requiring preconditioning. Proc. Natl. Acad. Sci. U.S.A. *110*, 4327–4332.

Hock, H., Hamblen, M.J., Rooke, H.M., and Traver, D. (2003). Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. Immunity *18,* 109-120.

Holtan, S.G., Pasquini, M., and Weisdorf, D.J. (2014). Acute graft-versus-host disease: a bench-to-bedside update. Blood *124*, 363–373.

Huber, T.L., Kouskoff, V., Joerg Fehling, H., Palis, J., and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature *432*, 625–630.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Brief communications. Nat. Biotechnol. *31*, 227–229.

Ivanovs, A., Rybtsov, S., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2014). Stem Cell Reports. Stem Cell Reports *2*, 449–456.

Jaffredo, T., Gautier, R., Eichmann, A., and Dieterlen-Lievre, F. (1998). Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. Development *125*, 4575–4583.

Jaffredo, T., Gautier, R., Brajeul, V., and Dieterlen-Lievre, F. (2000). Tracing the Progeny of the Aortic Hemangioblast in the Avian Embryo. Developmental Biology *224*, 204–214.

Jalali, M. (2012). Using the Cre/loxP recombination system to study the role of the PI3K/AKT signalling transduction pathway and the embryonic cellular origin of adult haematopoietic stem cells in the zebrafish. M(Res) Thesis. The University of Nottingham.

Jao, L.-E., Wente, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl. Acad. Sci. U.S.A. *110*, 13904–13909.

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*, 1397–1397.

Jin, H., Xu, J., and Wen, Z. (2007). Migratory path of definitive hematopoietic stem/progenitor cells during zebrafish development. Blood *109*, 5208–5214.

Jin, S.W. (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, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science *337*, 816–821.

Jing, L., and Zon, L.I. (2011). Zebrafish as a model for normal and malignant hematopoiesis. Disease Models & Mechanisms *4*, 433–438.

John H Postlethwait, Woods, I.G., Ngo-Hazelett, P., Yi-Lin Yan, P.D.K., Chu, F., Huang, H., Hill-Force, A., and Talbot, W.S. (2000). Zebrafish Comparative Genomicsand the Origins of Vertebrate ChromosomesJohn H. Postlethwait,1,3 Ian G. Woods,2 Phuong Ngo-Hazelett,1 Yi-Lin Yan,1Peter D. Kelly,2 Felicia Chu,2 Hui Huang,2 Alicia Hill-Force,1 and William S. Talbot2. 1–14.

Kabrun, N.N., Bühring, H.J.H., Choi, K.K., Ullrich, A.A., Risau, W.W., and Keller, G.G. (1997). Flk-1 expression defines a population of early embryonic hematopoietic precursors. Development *124*, 2039–2048.

Kalev-Zylinska, M.L., Horsfield, J.A., Flores, M.V.C., 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.

Karsunky, H., Zeng, H., Schmidt, T., Zevnik, B., Kluge, R., Schmid, K.W., Dührsen, U., and Möröy, T. (2002). Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. Nat. Genet. *30*, 295–300.

Kau, C.L., and Turpen, J.B. (1983). Dual contribution of embryonic ventral blood island and dorsal lateral plate mesoderm during ontogeny of hemopoietic cells in Xenopus laevis. J. Immunol. *131*, 2262–2266.

Kawakami, K. (2004). A Transposon-Mediated Gene Trap Technique Approach Identifies DevelopmentallyRegulated Genes in Zebrafish. Developmental Cell. *7,* 1–12.

Kazanjian, A., Gross, E.A., and Grimes, H.L. (2006). The growth factor independence-1 transcription factor: New functions and new insights. Critical Reviews in Oncology/Hematology *59*, 85–97.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature

translational termination. Nature Publishing Group *13*, 700–712.

Khandanpour, C., Sharif-Askari, E., Vassen, L., Gaudreau, M.-C., Zhu, J., Paul, W.E., Okayama, T., Kosan, C., and Möröy, T. (2010). Evidence that growth factor independence 1b regulates dormancy and peripheral blood mobilization of hematopoietic stem cells. Blood *116*, 5149–5161.

Kikuta, H., and Kawakami, K. (2009). Transient and Stable Transgenesis Using Tol2 Transposon Vectors. In Methods in Molecular Biology, (Totowa, NJ: Humana Press), pp. 69–84.

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.

Kingsley, P.D., Malik, J., Fantauzzo, K.A., and Palis, J. (2004). Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis. Blood *104*, 19–25.

Kissa, K., and 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., Cortés, A., Perret, E., Machu, C., and Herbomel, P. (2008). Live imaging of emerging hematopoietic stem cells and early thymus colonization. Blood *111*, 1147–1156.

Kotani, T., Nagayoshi, S., Urasaki, A., and Kawakami, K. (2006). Transposon-mediated gene trapping in zebrafish. Methods *39*, 199–206.

Kretzschmar, K., and Watt, F.M. (2012). Lineage Tracing. Cell *148*, 33–45.

Lacaud, G., and Kouskoff, V. (2016). Accepted Manuscript. Exp. Hematol. 1–16.

Lam, E.Y., Hall, C.J., Crosier, P.S., Crosier, K.E., and Flores, M.V. (2010). Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. Blood *116*.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N.K., Moroy, T., Bonifer, C., et al. (2012). GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. Blood *120*, 314–322.

Lawson, N.D., Scheer, N., Pham, V.N., Kim, C.H., Chitnis, A.B., Campos-Ortega, J.A., and Weinstein, B.M. (2001). Notch signaling is required for arterial-venous differentiation during embryonic vascular development. Development *128*, 3675–3683.

Lawson, N.D., and Weinstein, B.M. (2002). Arteries and veins: making a difference with zebrafish. Nat Rev Genet *3*, 674–682.

Lee, Y., EL Andaloussi, S., and Wood, M.J.A. (2012). Exosomes and microvesicles: extracellular vesicles for genetic information transfer and gene therapy. Human Molecular Genetics *21*, R125–R134.

Lengerke, C., and Daley, G.Q. (2010). Autologous blood cell therapies from pluripotent

stem cells. Blood Reviews *24*, 27–37.

Lensch, M.W. (2012). An Evolving Model of Hematopoietic Stem Cell Functional Identity. Stem Cell Rev and Rep *8*, 551–560.

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

Liao, E.C., Paw, B.H., Oates, A.C., Pratt, S.J., Postlethwait, J.H., and Zon, L.I. (1998). SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish. Genes & Development *12*, 621–626.

Liao, W., Bisgrove, B.W., Sawyer, H., Hug, B., Bell, B., Peters, K., Grunwald, D.J., and Stainier, D.Y. (1997). The zebrafish gene cloche acts upstream of a flk-1 homologue to regulate endothelial cell differentiation. Development *124*, 381–389.

Lieschke, G.J. (2001). Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. Blood *98*, 3087–3096.

Lieschke, G.J., Oates, A.C., Paw, B.H., Thompson, M.A., Hall, N.E., Ward, A.C., Ho, R.K., Zon, L.I., and Layton, J.E. (2002). Zebrafish SPI-1 (PU.1) Marks a Site of Myeloid Development Independent of Primitive Erythropoiesis: Implications for Axial Patterning. Developmental Biology *246*, 274–295.

Lin, H.F. (2005). Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. Blood *106*, 3803–3810.

Lux, C.T., Yoshimoto, M., McGrath, K., Conway, S.J., Palis, J., and Yoder, M.C. (2008). All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac. Blood *111*, 3435–3438.

Lykke-Andersen, S., and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nature Publishing Group *16*, 665–677.

Maéno, M., Tochinai, S., and Katagiri, C. (1985). Differential participation of ventral and dorsolateral mesoderms in the hemopoiesis of Xenopus, as revealed in diploid-triploid or interspecific chimeras. Developmental Biology *110*, 503–508.

Masel, J. (2011). Genetic drift. Curbio *21*, R837–R838.

McGrath, K.E., Frame, J.M., and Palis, J. (2015a). Seminars in Immunology. Seminars in Immunology *27*, 379–387.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015b). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Reports *11*, 1892–1904.

McGrath, K.E., Koniski, A.D., Malik, J., and Palis, J. (2003). Circulation is established in a stepwise pattern in the mammalian embryo. Blood *101*, 1669–1676.

Medvinsky, A.L., Samoylina, N.L., Müller, A.M., and Dzierzak, E.A. (1993). An early pre-liver intraembryonic source of CFU-S in the developing mouse. Nature *364,* 64-66.

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

Medvinsky, A., and Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. Cell *86*, 897–906.

Menke, A.L., Spitsbergen, J.M., Wolterbeek, A.P.M., and Woutersen, R.A. (2011). Normal Anatomy and Histology of the Adult Zebrafish. Toxicologic Pathology *39*, 759–775.

Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A.J., Tanaka, Y., Wilkinson, A.C., Buettner, F., Macaulay, I.C., Jawaid, W., Diamanti, E., et al. (2015). Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nat. Biotechnol. *33*, 269–276.

Moore, M.A.S., and Metcalf, D. (1970). Ontogeny of the Haemopoietic System: Yolk Sac Origin of In Vivo and In Vitro Colony Forming Cells in the Developing Mouse Embryo*. British Journal of Haematology *18*, 279–296.

Mosimann, C., Kaufman, C.K., Li, P., Pugach, E.K., Tamplin, O.J., and Zon, L.I. (2011). Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. Development *138*.

Möröy, T. (2005). The zinc finger transcription factor Growth factor independence 1 (Gfi1). The International Journal of Biochemistry & Cell Biology *37*, 541–546.

Möröy, T., Vassen, L., Wilkes, B., and Khandanpour, C. (2015). From cytopenia to leukemia: the role of Gfi1 and Gfi1b in blood formation. Blood *126*, 2561–2569.

Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J., and Potter, S.S. (1991). A Functional C-Myb Gene Is Required For Normal Murine Fetal Hepatic Hematopoiesis. Cell *65*, 677–689.

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

Murry, C.E., and Keller, G. (2008). Differentiation of Embryonic Stem Cells to Clinically Relevant Populations: Lessons from Embryonic Development. Cell *132*, 661–680.

Müller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzak, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291–301.

Nagy, A. (2000). Cre recombinase: the universal reagent for genome tailoring. Genesis *26*, 99–109.

Ng, Y.K., George, K.M., Engel, J.D., and Linzer, D.I. (1994). GATA factor activity is required for the trophoblast-specific transcriptional regulation of the mouse placental lactogen I gene. Development *120*, 3257–3266.

Nguyen, P.D., Hollway, G.E., Sonntag, C., Miles, L.B., Hall, T.E., Berger, S., Fernandez, K.J., Gurevich, D.B., Cole, N.J., Alaei, S., et al. (2014). Haematopoietic stem cell induction by somite-derived endothelial cells controlled by meox1. Nature *512*, 314–318.

Nishikawa, S.I., Nishikawa, S., Kawamoto, H., Yoshida, H., Kizumoto, M., Kataoka, H., and Katsura, Y. (1998). In vitro generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. Immunity *8*, 761–769.

North, T.E., de Bruijin, M., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E., and Speck, N.A. (2002). Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. Immunity *16*, 661–672.

North, T.T., Gu, T.L.T., Stacy, T.T., Wang, Q.Q., Howard, L.L., Binder, M.M., Marín-Padilla, M.M., and Speck, N.A.N. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. Development *126*, 2563–2575.

North, T.E., Goessling, W., Peeters, M., Li, P., Ceol, C., Lord, A.M., Weber, G.J., Harris, J., Cutting, C.C., Huang, P., et al. (2009). Hematopoietic Stem Cell Development Is Dependent on Blood Flow. Cell *137*, 736–748.

Nottingham, W.T., Jarratt, A., Burgess, M., Speck, C.L., Cheng, J.-F., Prabhakar, S., Rubin, E.M., Li, P.-S., Sloane-Stanley, J., Kong-a-San, J., et al. (2007). Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood *110*, 4188–4197.

Okuda, T., vanDeursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321–330.

Orkin, S.H.S., and Zon, L.I.L. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. Cell *132*, 14–14.

Paik, E.J., and Zon, L.I. (2010). Hematopoietic development in the zebrafish. Int. J. Dev. Biol. *54*, 1127–1137.

Palis, J., and Yoder, M.C. (2001). Yolk-sac hematopoiesis: the first blood cells of mouse and man. Exp. Hematol. *29*, 927–936.

Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development *126*, 5073–5084.

Palis, J., Malik, J., McGrath, K.E., and Kingsley, P.D. (2010). Primitive erythropoiesis in the mammalian embryo. Int. J. Dev. Biol. *54*, 1011–1018.

Panopoulos, A.D., and Belmonte, J.C.I. (2012). Induced pluripotent stem cells in clinical hematology. Current Opinion in Hematology *19*, 256–260.

Pelosi, E., Castelli, G., Martin-Padura, I., Bordoni, V., Santoro, S., Conigliaro, A., Cerio, A.M., De Santis Puzzonia, M., Marighetti, P., Biffoni, M., et al. (2012). Human Haemato-Endothelial Precursors: Cord Blood CD34+ Cells Produce Haemogenic Endothelium. PloS One *7*, e51109.

Perdiguero, E.G., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., de Bruijn, M.F., Geissmann, F., et al. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature *518*, 547–551.

Pineault, N., and Abu-Khader, A. (2015). Advances in umbilical cord blood stem cell expansion and clinical translation. Exp. Hematol. *43*, 498–513.

Porcher, C., Liao, E.C., Fujiwara, Y., Zon, L.I., and Orkin, S.H. (1999). Specification of hematopoietic and vascular development by the bHLH transcription factor SCL without direct DNA binding. Development *126*, 4603–4615.

Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F.W., and Orkin, S.H. (1996). The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. Cell *86*, 47–57.

Rampon, C., and Huber, P. (2003). Multilineage hematopoietic progenitor activity generated autonomously in the mouse yolk sac: analysis using angiogenesis-defective embryos. Int. J. Dev. Biol. *47*, 273–280.

Rowe, R.G., Mandelbaum, J., Zon, L.I., and Daley, G.Q. (2016). Engineering Hematopoietic Stem Cells: Lessons from Development. Stem Cell *18*, 707–720.

Saleque, S., Cameron, S., and Orkin, S.H. (2002). The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. Genes & Development *16*, 301–306.

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D. (2010). ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Research *38*, W462–W468.

Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., and Dobbs, D. (2007). Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. Nucleic Acids Research *35*, W599–W605.

Sander, J.D., and Joung, J.K. (2014). CriSPr-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. *32*, 347–355.

Saveanu, C., and Jacquier, A. (2016). How cells kill a “killer” messenger. Elife *5,* 1-3.

Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. Wiley Interdiscip Rev Syst Biol Med *2*, 640–653.

Shahriyari, L., and Komarova, N.L. (2013). Symmetric vs. asymmetric stem cell divisions: an adaptation against cancer? PLoS ONE *8*, e76195.

Silver, L., and Palis, J. (1997). Initiation of Murine Embryonic Erythropoiesis: A Spatial Analysis. Blood *89*, 1154–1164.

Sinha, D.K., Neveu, P., Gagey, N., Aujard, I., Le Saux, T., Rampon, C., Gauron, C., Kawakami, K., Leucht, C., Bally-Cuif, L., et al. (2010). Photoactivation of the CreER T2 recombinase for conditional site-specific recombination with high spatiotemporal resolution. Zebrafish *7*, 199–204.

Slukvin, I.I. (2013). Hematopoietic specification from human pluripotent stem cells: current advances and challenges toward de novo generation of hematopoietic stem cells. Blood *122*, 4035–4046.

Snippert, H.J., and Clevers, H. (2011). Tracking adult stem cells. Nature Publishing Group *12*, 113–122.

Sood, R., English, M.A., Belele, C.L., Jin, H., Bishop, K., Haskins, R., McKinney, M.C., Chahal, J., Weinstein, B.M., Wen, Z., et al. (2010). Development of multilineage adult hematopoiesis in the zebrafish with a runx1 truncation mutation. Blood *115*, 2806–2809.

Sturgeon, C.M., Ditadi, A., Clarke, R.L., and Keller, G. (2013). Defining the path to hematopoietic stem cells. Nat. Biotechnol. *31*, 416–418.

Sugiyama, D., Inoue-Yokoo, T., Fraser, S.T., Kulkeaw, K., Mizuochi, C., and Horio, Y. (2011). Embryonic Regulation of the Mouse Hematopoietic Niche. The Scientific World Journal *11*, 1770–1780.

Swiers, G., De Bruijn, M., and Speck, N.A. (2010). Hematopoietic stem cell emergence in the conceptus and the role of Runx1. Int. J. Dev. Biol. *54*, 1151–1163.

Tamplin, O.J., and Zon, L.I. (2010). Blood flow: metalloproteases cut loose in primitive erythrocytes. Curr Biol *20*.

Tamplin, O.J., Durand, E.M., Carr, L.A., Childs, S.J., Hagedorn, E.J., Li, P., Yzaguirre, A.D., Speck, N.A., and Zon, L.I. (2015). Hematopoietic Stem Cell Arrival Triggers Dynamic Remodeling of the Perivascular Niche. Cell *160*, 241–252.

Taoudi, S., and Medvinsky, A. (2007). Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. Proc. Natl. Acad. Sci. U.S.A. *104*, 9399–9403.

Thambyrajah, R., Ucanok, D., Jalali, M., Hough, Y., Wilkinson, R. N., McMahon, K., Moore, C., and Gering, M. (2016). A gene trap transposon eliminates haematopoietic expression of zebrafish *Gfi1aa*, but does not interfere with haematopoiesis. Developmental Biology *417*, 25-39.

Thambyrajah, R. (2011). A Gene Trap Screen Reveals the Expression of the Transcription Factor *Gfi1.1* in Haemogenic Endothelial Cells of the Zebrafish Embryo. Ph.D. Thesis, The University of Nottingham.

Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Möröy, T., et al. (2015). GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. Nature Cell Biology *18*, 21–32.

Thompson, M.A., Ransom, D. G., Pratt, S. J., MacLennan, H., Kieran, M. W., Detrich III, W., Vail, B., Huber, T. L., Paw, B., Oates, A. C., Brownlie, A. J., Fritz, A., Gates, M., Bahary, N., Talbot, H., Beier, D. R., Postlethwait, J. H., and Leonard I Zon (1998). The cloche and spadetail Genes Differentially Affect Hematopoiesis and Vasculogenesis. Developmental Biology *197,* 248-269.

Thummel, R., Burket, C.T., and Hyde, D.R. (2006). Two Different Transgenes to Study Gene Silencing and Re-Expression During Zebrafish Caudal Fin and Retinal Regeneration. The Scientific World Journal *6*, 65–81.

Thummel, R., Burket, C.T., Brewer, J.L., Sarras, M.P., Li, L., Perry, M., McDermott, J.P., Sauer, B., Hyde, D.R., and Godwin, A.R. (2005). Cre-mediated site-specific recombination in zebrafish embryos. Dev Dyn *233*, 1366–1377.

Tober, J., Koniski, A., McGrath, K.E., Vemishetti, R., Emerson, R., de Mesy-Bentley, K.K.L., Waugh, R., and Palis, J. (2007). The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. Blood *109*, 1433–1441.

Tober, J., Yzaguirre, A.D., Piwarzyk, E., and Speck, N.A. (2013). Distinct temporal requirements for Runx1 in hematopoietic progenitors and stem cells. Development *140*, 3765–3776.

Traver, D., Paw, B.H., Poss, K.D., Penberthy, W.T., Lin, S., and Zon, L.I. (2003). Transplantation and *in vivo* imaging of multilineage engraftment in zebrafish bloodless mutants. Nat Immunol *4,* 1238-1246.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature *371*, 221–226.

Ucanok, D. (2015). Zebrafish Gfi1 paralogs play redundant roles in primitive and definitive haematopoietic development. Ph.D. Thesis, The University of Nottingham.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. Nat Rev Genet *11*, 636–646.

van der Meer, L.T., Jansen, J.H., and van der Reijden, B.A. (2010). Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia *24*, 1834–1843.

Vassen, L., Okayama, T., and Moroy, T. (2007). Gfi1b:green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. Blood *109*, 2356–2364.

Walasek, M.A., van Os, R., and de Haan, G. (2012). Hematopoietic stem cell expansion: challenges and opportunities. Annals of the New York Academy of Sciences *1266*, 138–150.

Wallis, D. (2003). The zinc finger transcription factor Gfi1, implicated in lymphomagenesis, is required for inner ear hair cell differentiation and survival. Development *130*, 221–232.

Wang, Q., Stacy, T., and Binder, M., Marin-Padilla, M., Sharpe, A. H., and Speck N. A. (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 *97,* 248-269.

Wei, W., Wen, L., Huang, P., Zhang, Z., Chen, Y., Xiao, A., Huang, H., Zhu, Z., Zhang,

B., and Lin, S. (2008). Gfi1.1 regulates hematopoietic lineage differentiation during zebrafish embryogenesis. Cell Res *18*, 677–685.

Weissman, I.L. (2000). Stem cells: units of development, units of regeneration, and units in evolution. Cell *100,* 157-168.

Westerfield, M. (2000). The Zebrafish Book.

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331–338.

Wilkinson, R.N., Koudijs, M.J., Patient, R.K., Ingham, P.W., Schulte-Merker, S., and van Eeden, F.J.M. (2012). Hedgehog signaling via a calcitonin receptor-like receptor can induce arterial differentiation independently of VEGF signaling in zebrafish. Blood *120*, 477–488.

Willett, C.E.C., Cortes, A.A., Zuasti, A.A., and Zapata, A.G.A. (1999). Early hematopoiesis and developing lymphoid organs in the zebrafish. Dev Dyn *214*, 323–336.

Woods, I.G., Kelly, P.D., Chu, F., Ngo-Hazelett, P., Yan, Y.L., Huang, H., Postlethwait, J.H., and Talbot, W.S. (2000). A comparative map of the zebrafish genome. Genome Res. *10*, 1903–1914.

Yücel, R., Kosan, C., Heyd, F., and Möröy, T. (2004). Gfi1:Green Fluorescent Protein Knock-in Mutant Reveals Differential Expression and Autoregulation of the Growth Factor Independence 1 (Gfi1) Gene during Lymphocyte Development. J. Biol. Chem. *279*, 40906–40917.

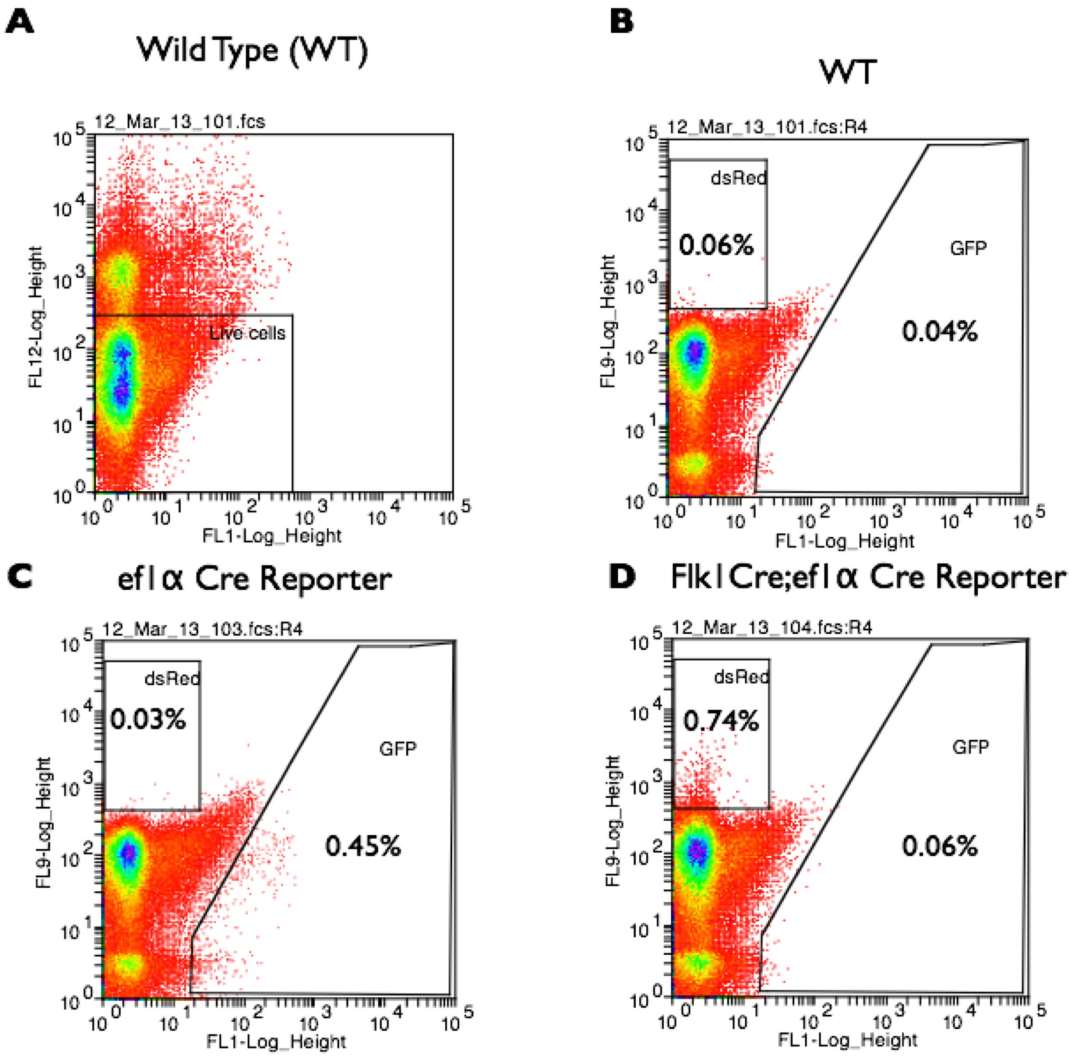
Zape, J.P., and Zovein, A.C. (2011). Hemogenic endothelium: Origins, regulation, and implications for vascular biology. Seminars in Cell and Developmental Biology *22*, 1036–1047.

Zeng, H., Yücel, R., Kosan, C., Klein-Hitpass, L., and Möröy, T. (2004). Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. Embo J *23*, 4116–4125.

Zhen, F., Lan, Y., Yan, B., Zhang, W., and Wen, Z. (2013). Hemogenic endothelium specification and hematopoietic stem cell maintenance employ distinct Scl isoforms. Development *140*, 3977–3985.

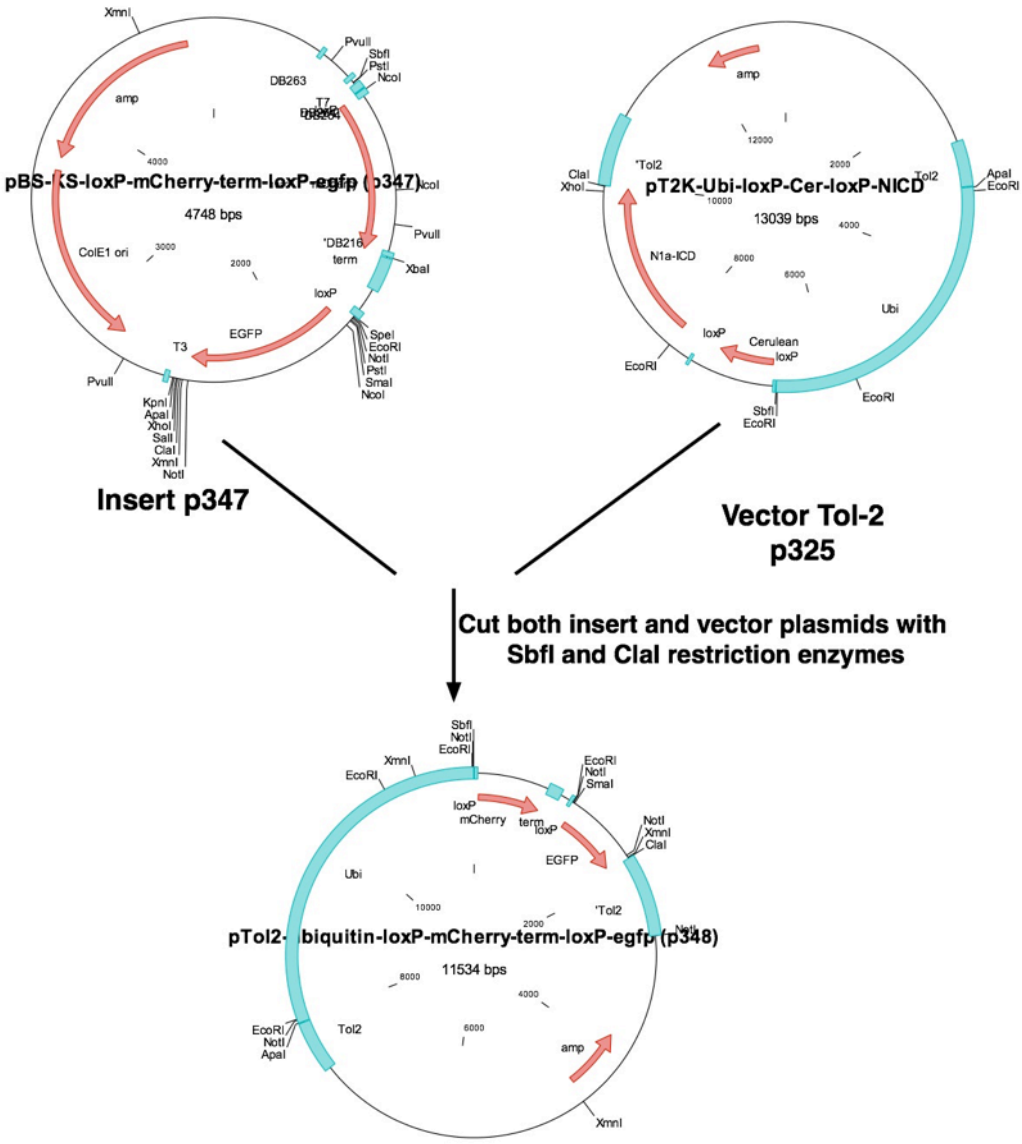
Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., et al. (2008). Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. Cell Stem Cell *3*, 625–636.

Appendices



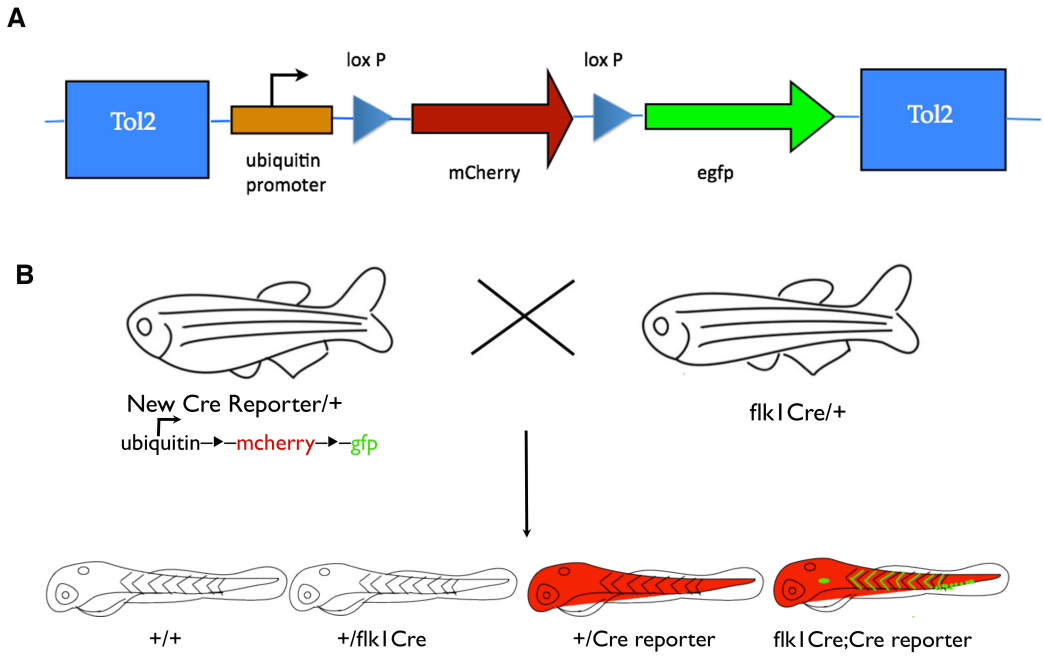
Appendix 1: FACS analysis shows recombination that occurred in blood cells flushed out of the adult kidney of *Tg(flk1/kdrl:Cre; ef1α-Cre Reporter)* transgenic fish.

(A) shows cells from the WT kidney that were stained with sytox red which stains dead cells and debris. Cells stained with sytox are detected in the far-red wavelength. In order to include only the live cells in our analysis, we used a gate with a threshold of just above 102. B) From the selected live wt cells a mere 0.04% of eGFP and 0.06% dsRed, most likely to be auto-fluorescing cells. C) shows cells isolated from blood cells flushed out of the kidney of a single transgenic *ef1α-Cre reporter* fish shows 0.45% eGFP+ cells and 0.03% (negligible) dsRed cells D) FAC analysis on blood cells flushed out of the kidney of a double transgenic *Tg(flk1/kdrlCre;ef1α-Cre reporter)* fish shows 0.74% dsRed+ cells and 0.06% eGFP+ cells, indicating almost complete recombination. In all panels; FL-1 on the x-axis detects eGFP+ cells at 507nm using the 488nm laser, FL-9 on the y-axis in panels B, C and D detects dsRed+ cells at a wavelength of 592nm detected by a 561nm laser and FL-12 on the y-axis in panel (A) uses the 640nm laser to detect sytox positive cells in the far red wavelength at 658nm.



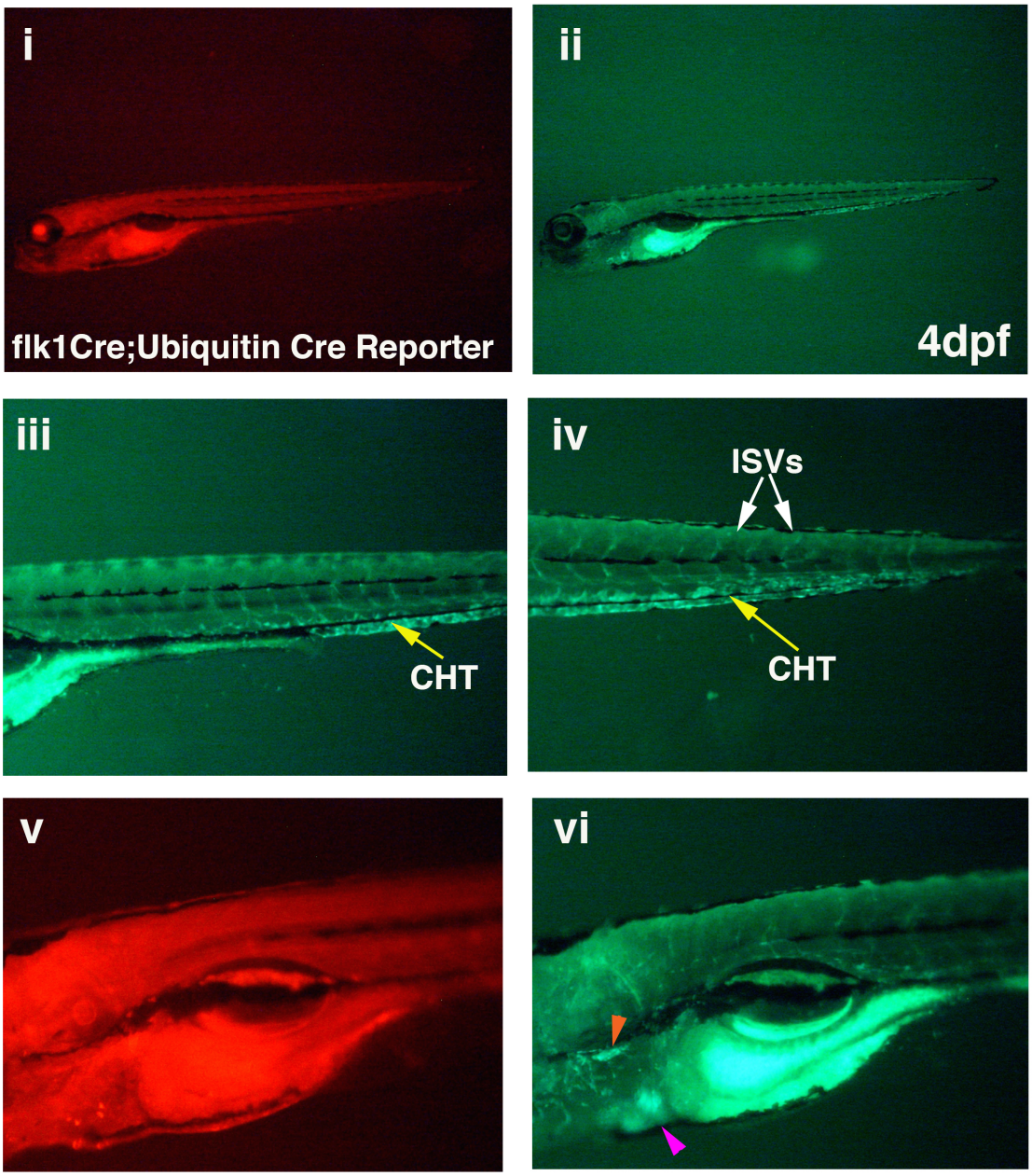
Appendix 2: Generation of the red-to-green Ubiquitin-Cre Reporter construct.

The red-to-green Cre reporter was p347 was in the plasmid blue script and the vector tol2 was in a NICD construct p325 made by Mohammad Al Khamees. Both the p347 and p325 were cut using the restriction enzymes SbfI and ClaI. The ubiquitin promoter in the Tol2 vector was ligated to the red-to-green construct to give the Ubiquitin Cre Reporter construct which was then subsequently injected along with transposase mRNA in one-cell stage zebrafish embryos.



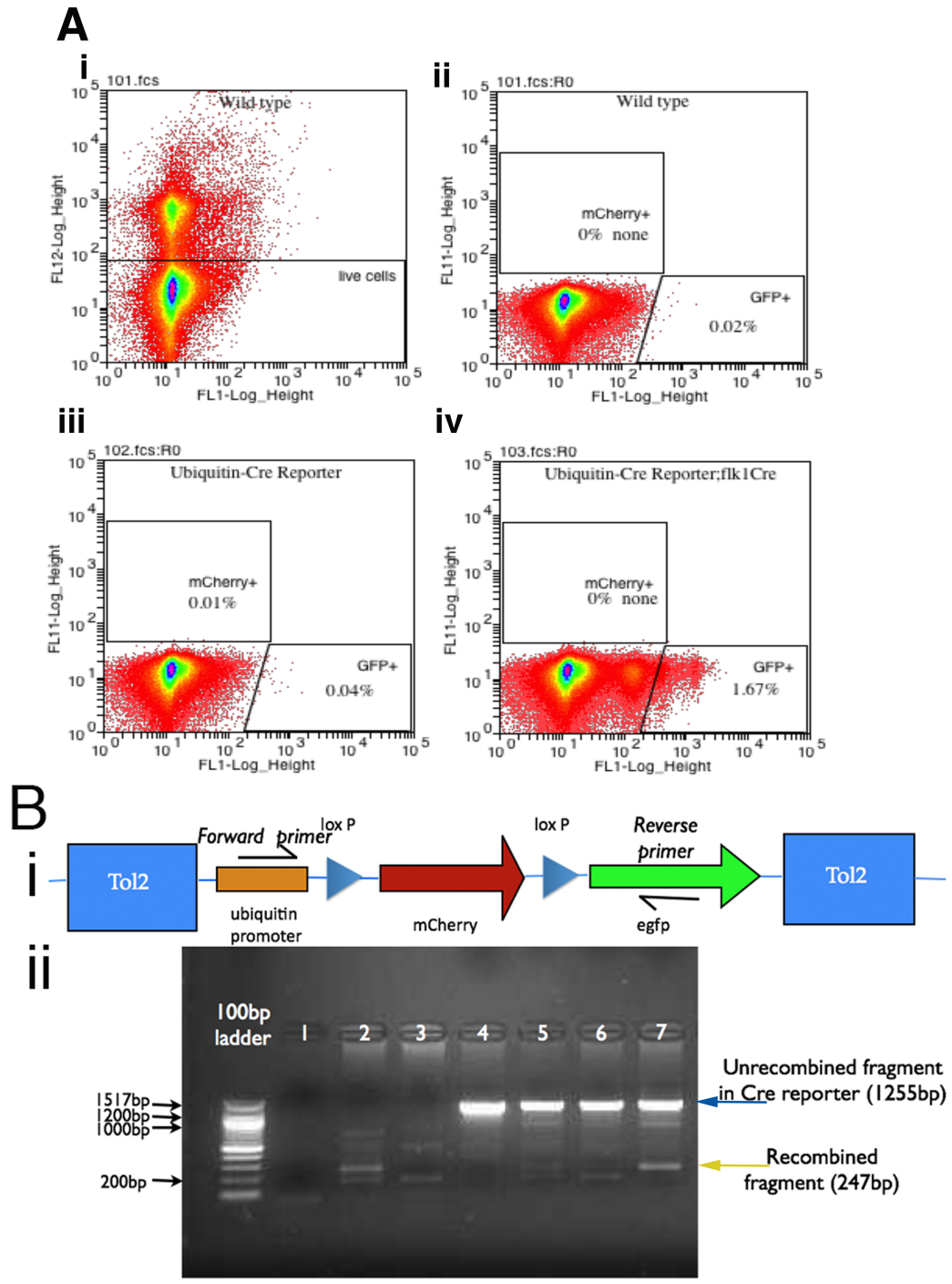
Appendix 3: Schematic representation of the constitutively active Cre recombination system using a new red-to-green reporter.

(A) I developed a new red-to-green (mcherry-to-eGFP) *Cre reporter* construct in which a red fluorescent gene, mcherry lies within the 2 LoxP sites and is expressed in the absence of recombination. A second reporter gene, i.e *eGFP* lies downstream of the *mcherry*. In the event of recombination, the floxed region is deleted and eGFP is expressed in *flk1/kdrl* derived cells. The fluorescent reporter protein expression is under the control of a *ubiquitous* promoter, the *ubiquitin* promoter. (B) When a *flk1/kdrl:Cre/+* fish was crossed to a *ubiquitin Cre Reporter/+* fish, we expect half of the embryos to be mcherry- and half to be mcherry+. Half of the mcherry+ embryos would express eGFP on 4-5dpf in flk1derived cells after recombination. We expect to see eGFP expression in cells within the CHT, in intersomitic vessels, the thymus and in the endocardium. The genotype of the different embryos obtained has been depicted under each embryo shown.



Appendix 4: The Ubiquitin promoter seems to be a weak ubiquitous promoter as revealed by low levels of fluorescent gene expression in the embryo.

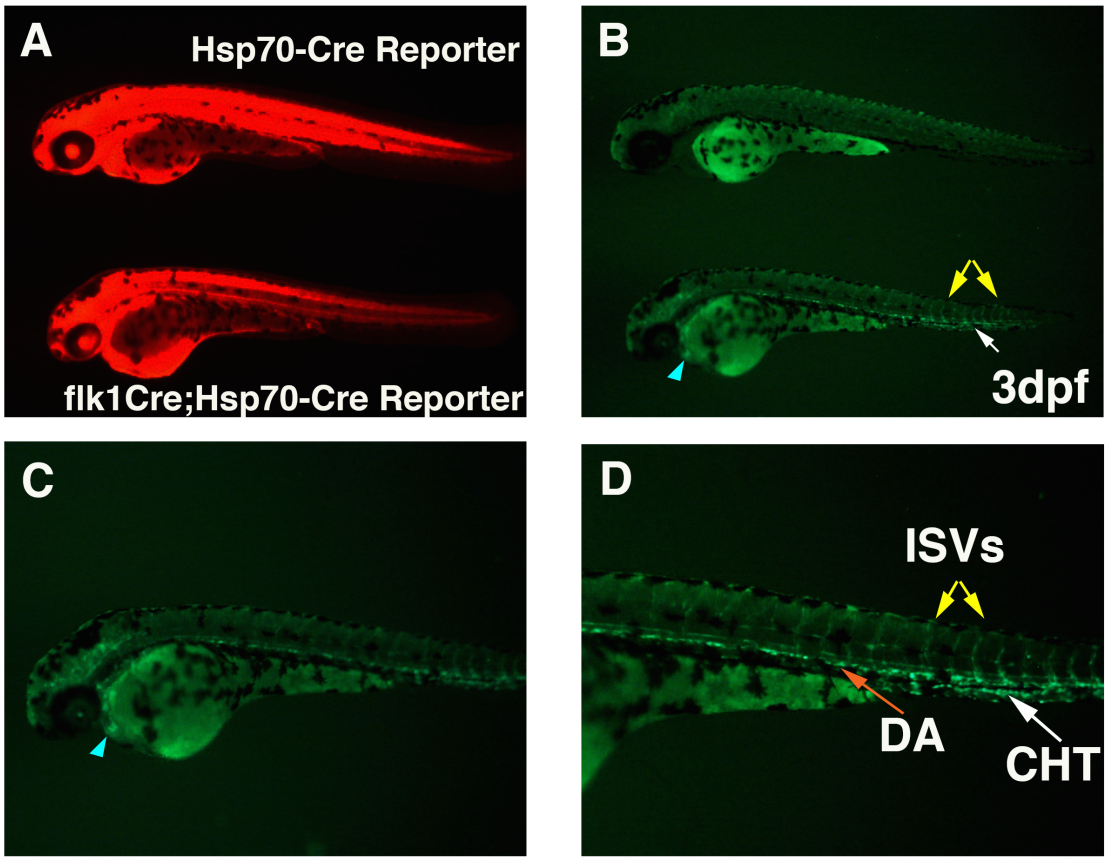
The image depicts a 4-5dpf Tg(flk1/kdrl:Cre;Ubiquitin-Cre reporter) embryo. (i) shows relatively weak ubiquitous expression of mcherry, which is driven by the ubiquitin promoter. mcherry is expressed in the absence of recombination. (ii) After recombination, flk1/kdrl derived cells express eGFP. (iii-iv) eGFP expression is seen in the CHT (yellow arrow) and in intersomitic vessels (ISVs) (white arrows) (v-vi) show eGFP expression in the thymus (orange arrowhead) and in the endocardium/heart (pink arrowhead).



Appendix 5: **The *ubiquitin* promoter seems to be a stronger ubiquitous promoter in the adult compared to the *ef1α* promoter.**

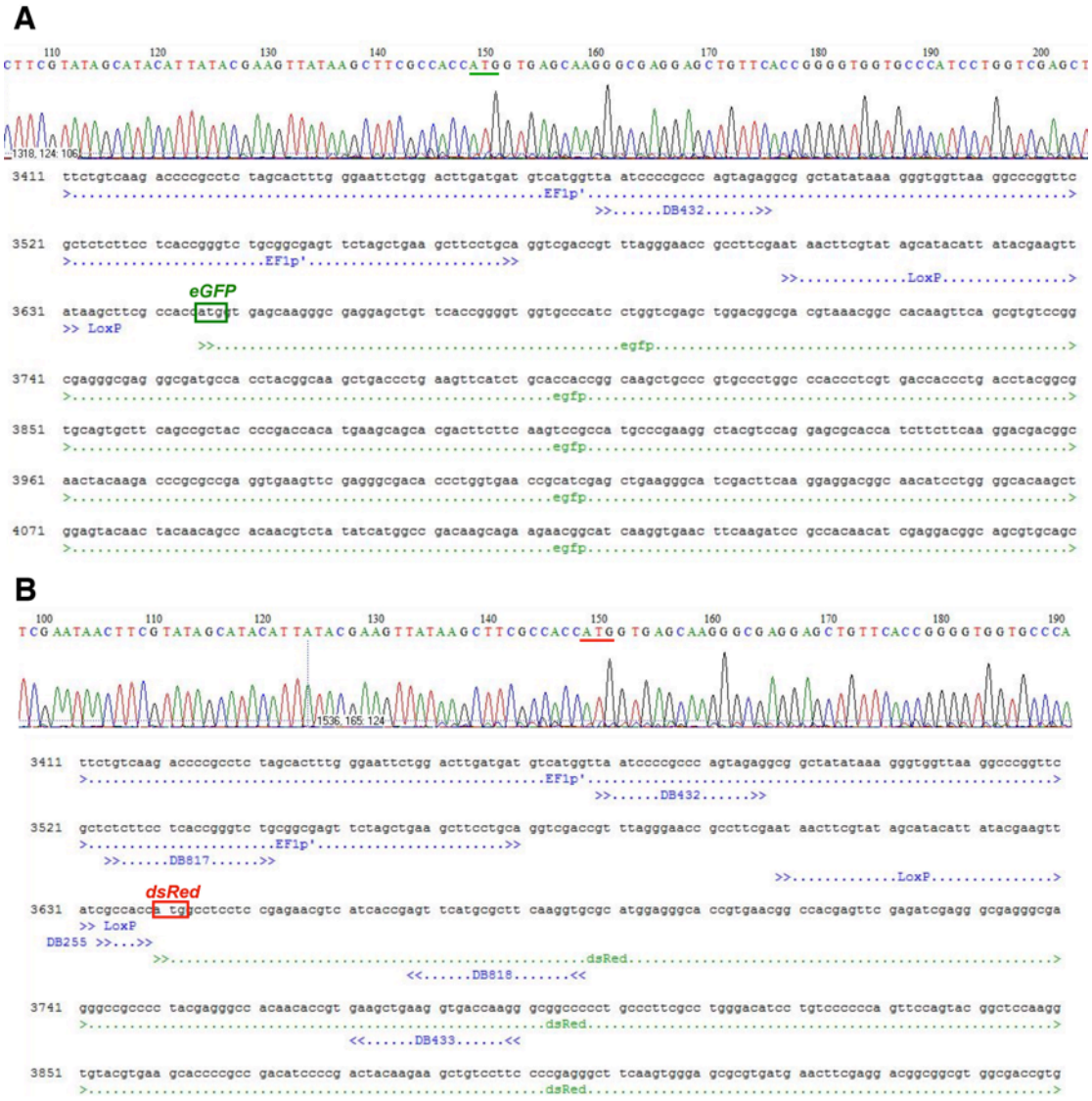
A) FACS analysis shows recombination that occurred in blood cells flushed out of the adult kidney of *Tg(flk1/kdrl:Cre;Ubiquitin Cre Reporter)* transgenic fish (Ai) shows cells from the *wt* kidney that were stained with sytox red which stains dead cells and debris. Cells stained with sytox are detected in the far-red wavelength. In order to include only the live cells in our analysis, we used a gate with a threshold of just above 102.Aii) From the selected live wt cells a mere 0.02% of eGFP and 0% mcherry, most likely to be auto-fluorescing cells. Aiii) shows cells isolated from blood cells flushed out of the kidney of a *ubiquitin-Cre reporter* fish that showed 0.01% mcherry+ cells (negligible) and 0.04% (negligible) eGFP+ cells Aiv) FAC analysis on blood cells flushed out of

the kidney of a double transgenic *Tg(flk1:Cre; Ubiquitin Cre reporter)* fish shows 0% mcherry+ cells and 1.67% eGFP+ cells, indicating that the FAC sorter continues to have problems detecting red fluorescence. In all panels; FL-1 on the x-axis detects eGFP+ cells at 507nm using the 488nm laser, FL-11 on the y-axis in panels Aii, Aiii and Aiv detects mcherry+ cells at a wavelength of 610nm detected by a 561nm laser and FL-12 on the y-axis in panel (Ai) uses the 640nm laser to detect sytox positive cells in the far red wavelength at 658nm. (B) An agarose gel electrophoresis was carried out on PCR products obtained from gDNA. The PCR was carried out using primers depicted in (Bi), the forward primer was in the Ubiquitin promoter and the reverse primer was situated within the *eGFP* gene. (Bii) The agarose gel electrophoresis shows a 100bp DNA ladder in the first well the well labeled 1 contained a water/no template control, well labeled 2 contained the PCR product obtained from gDNA of a wt fish and shows many non-specific bands, well numbered 3 contained the PCR product obtained from gDNA of an mcherry- sibling fish whose genotype could either be +/+ or *flk1/kdrl:Cre/*+, hence this would be an additional negative control. The well labeled 4 contained the PCR product obtained from gDNA of an Ubiquitin Cre Reporter fish the band (indicated by the blue arrow) seems to be at the 1255bp position which is the size of the unrecombined fragment, this sample does not show a band at the 247bp (recombined fragment) position. Lanes 5 to 7 contained PCR products obtained from gDNA of *flk1/kdrl:Cre; Ubiquitin Cre reporter* fish, all of the lanes contain the 1255bp band (unrecombined fragment). We see a strong second band i.e. at approximately the 247bp (recombined fragment) position (ochre arrow) in lane number 7. Wells labelled 5 and 6 show a weak band at the 247bp position, and some non-specific bands. Genomic DNA for all samples was obtained from blood cells flushed out of the adult fish kidney marrow.



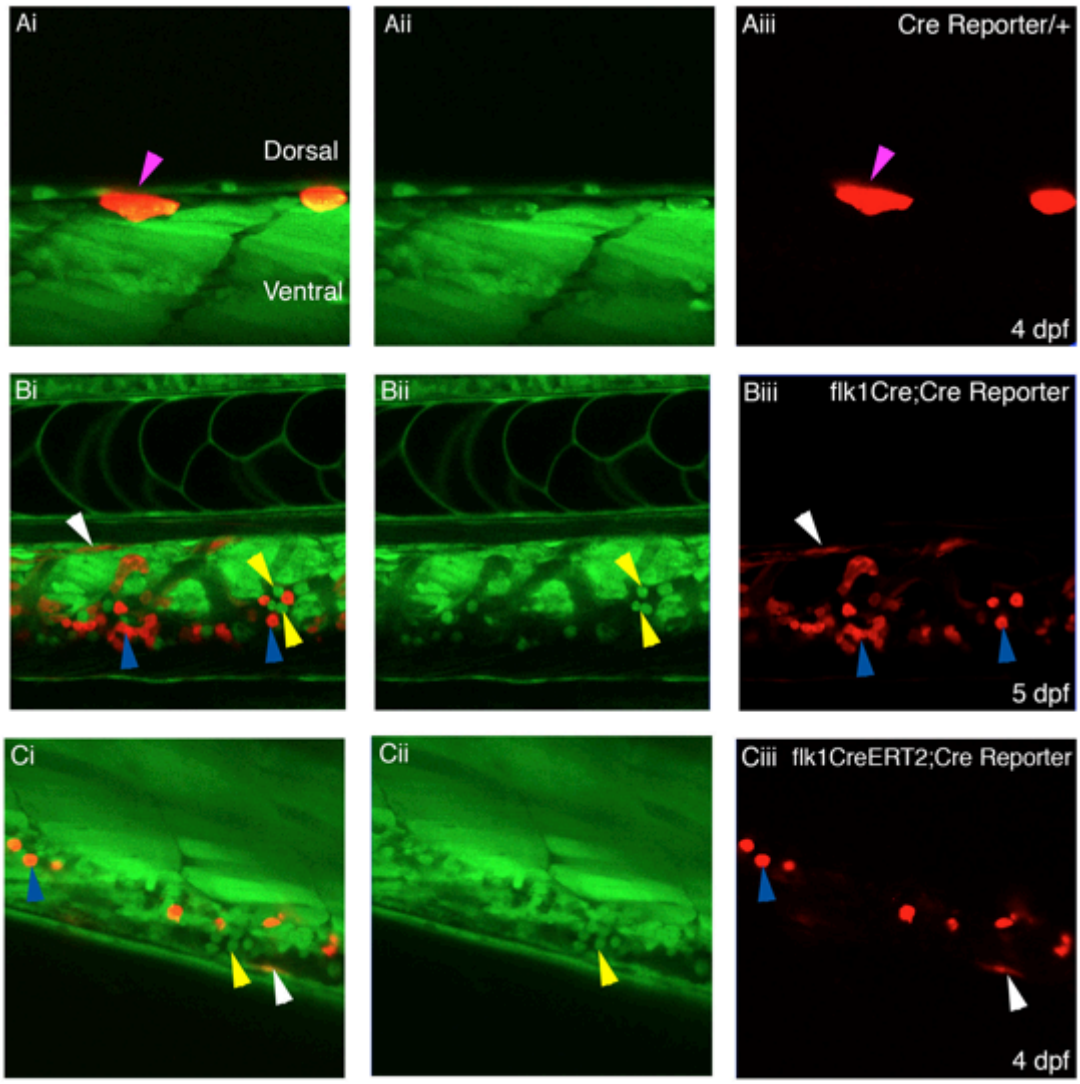
Appendix 6: The *hsp70* promoter seems to be a promising ubiquitous promoter.

When a *Tg(hsp70-Cre reporter)* fish was crossed to a *flk1/kdrl:Cre* fish, embryos were heat shocked in order to observe fluorescent gene expression. (A) depicts 3dpf embryos with ubiquitous mcherry expression, which could be observed after heat shock. (B) shows the embryos as seen under the green channel where eGFP expression is seen in *flk1/kdrl* derived cells in the (C) endocardium (turquoise arrowhead in B and C) and (D) in the dorsal aorta (DA) (orange arrow), the CHT (white arrows in B and D) and in intersomitic vessels (ISVs) (yellow arrows in B and D).



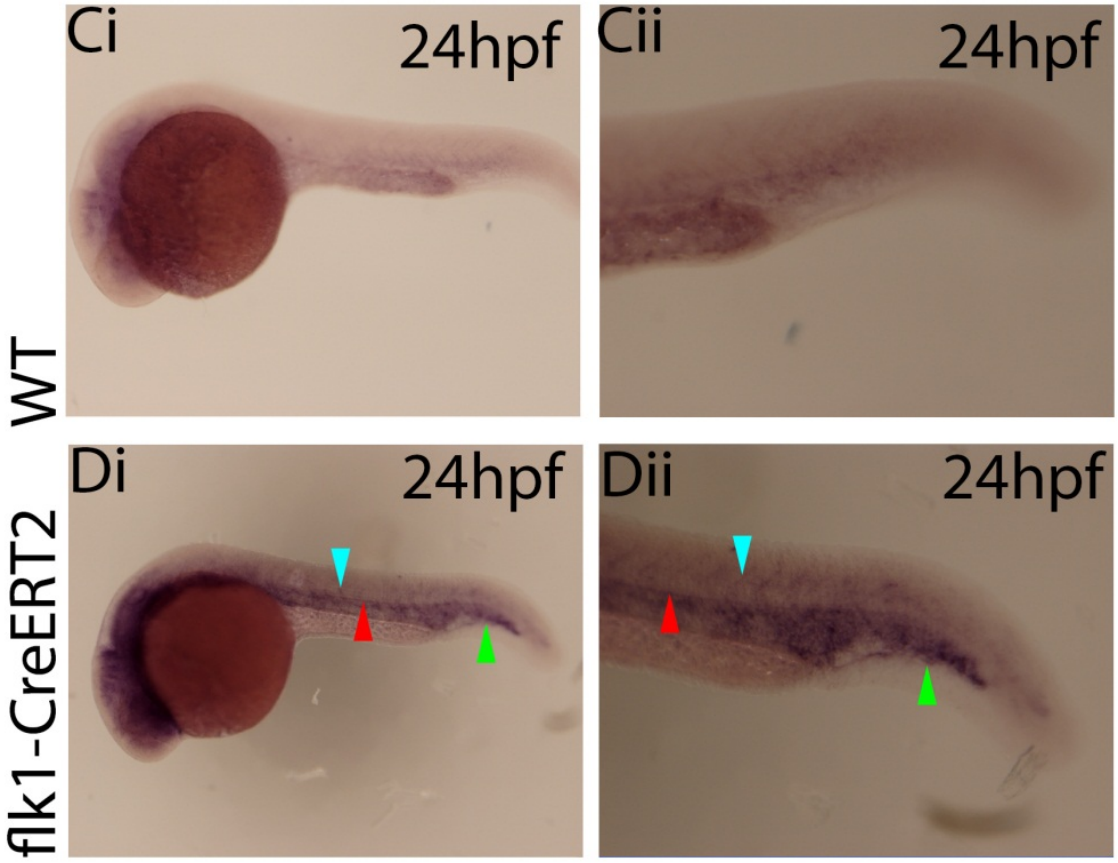
Appendix 7: The sequence of the 1325bp and 307bp PCR fragment matches the sequence of the non-recombined and recombined *ef1α:CreER^{T2}* reporter allele respectively.

(A) Primer DB432 was used to sequence the 1325 bp PCR fragment. The sequencing data were visualised as a chromatogram using BioEdit program. The sequence highlighted in the chromatogram matches the sequence of the non-recombined *ef1α:CreER^{T2}* reporter allele highlighted in the sequence editing programmed Clone Manager. The highlighted region starts a few base pairs 3' to the sequencing primer DB432 and includes the 3' part of the *ef1α*-promoter sequence, a loxP site and the eGFP sequence. On the BioEdit chromatogram, the ATG start codon of the *eGFP* gene is labelled. (B) Primer DB432 was used to sequence the 307 bp PCR fragment. The sequencing results are visualised in a chromatogram generated using the BioEdit software. The sequence highlighted in the chromatogram is the same sequence highlighted in the DNA editing programme Clone Manager below. The latter corresponds to the sequence of the *ef1α:CreER^{T2}* reporter gene recombined by Cre at the original two loxP sites. The sequence includes the 3' end of the *ef1α*-promoter sequence, a LoxP site and part of the *dsRed* reading frame. The sequence ends at the end of the reverse oligo DB433 that was used in the PCR. On the BioEdit chromatogram, the ATG start codon of the dsRed gene is labeled.



Appendix 8: Confocal imaging carried out on the CHT of 4-5 dpf embryos demonstrate, lower recombination efficiency in induced embryos compared to embryos that expressed Cre constitutively.

All images are displayed as merged, eGFP only and dsRed only (Panel A) shows a section of the tail region of an *Tg(ef1-Cre Reporter)* 4dpf embryo, i.e. that expresses eGFP only. (Ai and Aiii) red auto-fluorescing pigment cells were observed on the dorsal surface of the tail. (Panel B) depicts a section of the CHT of a 5dpf *Tg(flk1:Cre;ef1-Cre Reporter)* double transgenic embryo, (Bi and Biii) showing dsRed+, spindle shaped endothelial cells and dsRed+ circular blood cells were observed. (Bi and Bii) eGFP+ circular mesenchymal cells were observed in the CHT. (Panel C) displays a section of the CHT of a 4dpf *Tg(flk1:CreER^{T2}; ef1α Cre Reporter)* induced embryo that shows (Ci and Ciii) a small population of dsRed+ circular cells and (Ci and Ciii) a spindle shaped cell and (Ci and Cii) some eGFP+ circular cells. Images Ai, Bi and Ci are merged images of the green and red channels, Aii, Bii and Cii display the image as seen in the green channel alone and Aiii, Biii and Ciii show images as viewed in the red channel only. Optical sections were in (A) 1.1µm, (B) 2.1µm and (C) 2µm. Arrowhead colour codes: white arrowheads: dsRed+ spindle shaped ECs, yellow arrowheads: eGFP+ circular mesenchymal cells, blue arrowheads: dsRed+ circular blood cells, pink arrowheads: pigment cells. All embryos faced to the left with dorsal side up.



Appendix 9: The *flk1/kdrl* 0.8kb enhancer 1.5kb promoter fragment drives weak *CreER^{T2}* mRNA expression at 24hpf.

(Ci-ii) wt embryo shows no *CreER^{T2}* mRNA expression (Di-ii) At 24hpf the double transgenic embryo displays weak *CreER^{T2}* mRNA expression in the DA (red arrowhead), intersomitic vessels (turquoise arrow head) and cells in the CHT (green arrow head) (Image from Dara Nayer Nouri, MSci 2014).