UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

DIA1R in zebrafish development

José Pedro Serrado Monteiro Marques

Mestrado em Biologia Evolutiva e do Desenvolvimento

Dissertação orientada por:
Ana Teresa Tavares

Maria Gabriela Rodrigues

2018

Acknowledgments

Ok, so let's do this. Disclaimer: Most likely I'll forget some people who have helped through these past months, so I apologize beforehand to them.

Throughout my thesis, I have learned that I know (almost) nothing about a lot of things. So how did I manage to do a master thesis in Developmental Biology? Well, if it wasn't for the people I'm about to mention, I certainly would have not been able to do it.

I would like to say thank you to Raquel Jacinto and Bárbara Tavares from Susana Lopes' lab (unfortunately, Bárbara is no longer with us… She went to London). They helped (more like taught) me in the *in situ* hybridization protocol for zebrafish. Both helped me through every step of the technique and gave me valuable advice to perform this laborious protocol. Without them, it would have taken me months to perfect this technique.

For qPCR, there are a lot of people to thank, and I am going to name them based on the steps necessary to do qPCR for zebrafish. Raquel Jacinto, because she gave me the protocol for the RNA extraction of zebrafish embryos. Then, for the execution of said protocol and the rest of the technique, I had the most wonderful help form Jorge Borbinha, Diogo Paramos and Ana Sofia Brandão. It was fun because we (except Sofia) were doing it for the first. It was like a throwback to the practical lessons I've had in the past. For the analysis of the data, I had the amazing help of (again) Raquel Jacinto and André Macedo. They taught me, again and again how to analyze the data to every excruciating detail.

For all the imaging I have done, I need to thank Telmo Pereira, former lab member of Jacinto's and current head of CEDOC's microscopy facility. Always very helpful and interested in solving all of the problems that occurred during image acquisition. For the image analysis, I also need to thank Gabriel Martins from IGC imaging facility because he showed me the amazing Extended depth of Field plugin and how to correct the white balance post image acquisition in Fiji. Then Telmo help me create macros in Fiji for both of the things I learned with Gabriel. If you find the images in this thesis beautiful, you can thank these two gentlemen right here.

Of course, I would like thank Antonio Jacinto for agreeing with Ana Teresa Tavares' decision of choosing me to do this project. And, obviously, Ana Teresa Tavares for the supervision. She was able to withstand a nervous and always filled with doubts student. Ana was always available for any questions regarding any issue I had throughout the project. One of the most important things in science is the ability to disregard any personal bias one might have and analyze the data for what it is and discuss science with honesty. By this I mean, instead of forcing your own views and ideas, always be humble and verify your own claims and have the ability to disregard them once proven wrong. In that regard, I think Ana is the one of the most honest scientist I have met. Whenever we would disagree on something, even interpreting results, she would first listen to my claims and then explain why they were wrong. And sometimes, if it was the right thing, she would accept my view of the data. And in all of our discussions, I never felt that I needed to withhold my opinion. Because of that it always felt that our discussions were between to scientist wanting figure out something and not between a know-it-all and quiet student. One valuable thing I learned from Ana is to be positive and optimistic (while staying critical). I also had the opportuning of presenting this work at DIA meeting in Algarve were Leonor Saúde said the same thing. "One needs to be optimistic and proud about one's work. If you are not, it means that you think you could do better. If so, then do better, because you are able to." I think I'm going to steal this quote from Leonor and claim it as my own, because it's so good, it's worth stealing…

just kidding… Or am I? (cue dramatic song). Jokes aside, I also thank Maria Gabriela Rodrigues for accepting to be my supervisor and proof-reading of the manuscript.

Of course, I have to thank my parents for their support during this long-ass journey.

Last, but most definitely not the least, I want to thank all of Jacinto's lab (and ramifications) with whom I interacted. These wonderful people are António Jacinto, Carolina Crespo, Guadalupe Cabral, Raquel Lourenço, Ricardo Costa, Ana Teresa Matias, Rita Gorgulho, Valdir Semedo, Susana Ponte, Mafalda Pinho, Diana Saraiva, Ana Farinho, Lara Carvalho, Ana Sofia Brandão, Diogo Paramos and Jorge Borbinha. Thank you all for the amazing time I spent in the lab and making me feel like I was part of the group and not just some guy who was passing by. Even when there were days in which the motivation was the lowest possible, it was always felt good to go the lab and hang around with you all. And very special thanks for Lara, Diogo, Sofia and Borbinha for meeting me in the place I cherish the most on earth. Life's good when you can share what you love with the people that matter to you.

There's so much more I could say about the people in the lab but if kept going on, the acknowledgements would be bigger than the thesis.

And now, I hope you, the reader, enjoy this piece of work.

Also, I would like to thank Checo and Geodude for correcting my objectives. I did not lose any bet for putting their names in the acknowledgments. Nope, that did not happen.

Resumo

Durante o desenvolvimento embrionário dos vertebrados, os tecidos que constituem o sistema circulatório são cruciais para a sobrevivência do embrião. O sistema circulatório fornece ao embrião oxigénio e nutrientes, e remove os resíduos metabólicos tóxicos produzidos pelos tecidos em crescimento através dos eritrócitos. Os macrófagos embrionários intervêm na remodelação de tecidos ao fagocitar possíveis agentes patogénicos e corpos apoptóticos que surgem como subproduto do desenvolvimento embrionário. Tanto a hematopoiese com a vasculogénese estão intimamente associadas. Esta observação levou à hipótese de que as células hematopoiéticas e endoteliais têm um precursor comum, o hemangioblasto. A existência deste precursor comum tem sido apoiada por estudos de diferenciação *in vitro*, bem como pela evidência de que os progenitores hematopoiéticos e endoteliais expressam vários genes em comum, como por exemplo *vegfr2*, *lmo2* e *scl*.

Durante o estudo da regulação transcricional do gene *Cerberus* na galinha, foi identificada uma região cis-regulatória que promove expressão génica nos hemangioblastos do saco vitelino. Este repórter específico de hemangioblastos (Hb-eGFP) permitiu isolar estas células via FACS e caracterizar o perfil de expressão génica das mesmas através de *microarrays*. Como esperado, os transcritos dos genes característicos de hemangioblastos foram enriquecidos na população positiva para Hb-eGFP. Além dos marcadores de hemangioblastos, esta análise levou à identificação de alguns genes não caracterizados previamente que podem estar envolvidos na diferenciação e função de hemangioblastos. Curiosamente, o segundo gene mais altamente expresso é o ortólogo do *DIA1R* humano (*Deleted in Autism 1 Related*). Os genes *DIA1R* são exclusivos dos vertebrados. Análise da sequência das proteínas DIA1R prevê que estas possuam um péptido sinal e um domínio cinático PIP49_C bastante conservado, característico da família FAM69 de proteínas semelhantes a cinases. Estas características sugerem que as proteínas DIA1R possam regular o tráfego molecular ou interferir na função de fatores segregados.

Em relação à expressão do *dia1r,* o nosso grupo realizou estudos de expressão preliminares nos embriões de galinha e ratinho. Na galinha, o *cDIA1R* (*DIA1R* da galinha) é expresso em hemangioblastos de saco vitelino, ilhotas sanguíneas e nas células endoteliais da aorta dorsal, endocárdio e vasculatura da cabeça, que são regiões com potencial hemogénico. Em particular, a expressão de *cDIA1R* parece mais elevada em células que são morfologicamente semelhantes às células do endotélio hemogénico. No cérebro do embrião de ratinho, os transcritos *mDIA1R* também são detetados no endotélio vascular e na microglia. Estes dados sugerem que o DIA1R pode ser um bom marcador e potencial regulador das linhagens hematopoiéticas intraembrionárias que derivam dos hemangioblastos do saco vitelino.

O peixe-zebra é também um excelente modelo para estudar o desenvolvimento do sangue. Assim, decidimos investigar onde o *zDIA1R* (daqui em diante *dia1r)* é expresso durante o desenvolvimento do peixe-zebra. Descobrimos que *dia1r* é detetado pela primeira vez a 12 horas-pós-fertilização (hpf) na mesoderme lateral anterior (ALM) e na mesoderme lateral posterior (PLM), tecidos estes onde os primeiros precursores de células hematopoiéticas estão localizados. Nestes estádios iniciais, o padrão de expressão de *dia1r* é semelhante ao de *scl*, um dos primeiros genes a serem expressos nas células progenitoras hematopoiéticas. Às 19,5hpf, *dia1r* e *scl* são ambos expressos no ICM, o tecido hematopoiético que deriva da PLM. Às 30hpf, o *dia1r* é expresso na aorta dorsal e no tecido hematopoiético caudal. Nestes tecidos também é expresso *lmo2*, um marcador de células hematopoiéticas estaminais/progenitoras que estão presentes nestes tecidos. Em fases de desenvolvimento posteriores (48hpf a 4dpf), a expressão *dia1r* não foi detetada.

Após a análise do padrão de expressão de *dia1r*, investigámos a sua função durante o desenvolvimento. Para tal, utilizámos um morfolino para impedir a tradução do gene, e iniciámos a criação de uma linha

mutante utilizando a tecnologia CRISPR/Cas9. Para avaliar se DIA1R tem um papel no desenvolvimento de células progenitoras hematopoiéticas, fizemos hibridação *in situ* (WISH) para os marcadores hematopoiéticos em embriões injetados com morfolinos (morfantes). Às 24hpf, os morfantes exibem uma maior expressão de *gata1*, um marcador para progenitores eritro-mielóide (EMP), e uma perda parcial da expressão de *l-plastin*, um marcador de macrófagos. Essas observações sugerem que os morfantes de *dia1r* possuem mais EMPs e menos macrófagos. Estes resultados indicam que DIA1R pode regular a proliferação e/ou a diferenciação da população de células progenitoras eritro-mielóides, promovendo a diferenciação de macrófagos.

Para a avaliação semi-quantitativa do fenótipo dos morfantes de *dia1r*, utilizámos PCR quantitativo em tempo real (RT-qPCR ou qPCR) em embriões com 19,5hpf. Com esta técnica, avaliámos a expressão de nove genes que servem de marcadores para progenitores hematopoiéticos (*scl*, *lmo2*, *c-myb*, *runx1*, *gata1* e *spi1*) e para células hematopoiéticas diferenciadas (*α-eHb*, *l-plastin* e *flk1*). Destes nove genes analisados, dois deles têm níveis de expressão significativamente diferentes: a expressão de *lmo2* foi aumentada enquanto a expressão de *l-plastin* diminuiu. Embora estes dois genes tenham sido os únicos com alterações significativas dos níveis de expressão, parece haver um enviesamento para o aumento de expressão de todos os marcadores de progenitores hematopoiéticos testados. Uma vez que as EMPs derivam de progenitores hematopoiéticos que expressam *lmo2*, a presença de mais EMPs a 24hpf (observada na análise WISH) pode ser uma consequência de uma maior população de progenitores hematopoiéticos a 19,5hpf. A diminuição de expressão de *l-plastin* a 19,5hpf pode ser devida à redução do número de macrófagos nos morfantes, o que foi também observado nas experiências de WISH. Em conclusão, os resultados de qPCR juntamente com a análise WISH dos morfantes de *dia1r* levaram-nos a suspeitar que o DIA1R pode restringir o tamanho da população de células progenitoras hematopoiéticas, seja pela regulação da sua proliferação ou pela indução da sua diferenciação para a linhagem mielóide.

Apesar da facilidade de uso de MO para estudar a função de um gene, os fenótipos induzidos por MO podem ser diferentes dos fenótipos de organismos mutantes correspondentes devido a efeitos off-target do MO ou compensação genética nos mutantes. Portanto, desenhámos uma estratégia para gerar uma linha de alelo nulo ou knockout (KO) de *dia1r* utilizando o sistema CRISPR/Cas9 (em colaboração com a Fundação Champalimaud). Para aumentar a probabilidade de criar um KO para *dia1r*, escolhemos um RNA guia (gRNA) que tem como alvo o primeiro codão do *dia1r*. O gRNA e a proteína Cas9 foram injetados em embriões no estadio de uma célula. Nesta fase, foram recolhidos alguns embriões injetados (crispants ou F0) para genotipagem e análise fenotípica preliminar. Embora a maioria dos embriões F0 sejam mosaicos para o KO de *dia1r*, estes crispants podem já ter algumas alterações fenotípicas detectáveis por WISH. Usando esta técnica, avaliámos a expressão de marcadores para progenitores hematopoiéticos (*c-myb*), progenitores eritro-mielóides (*gata1*) e macrófagos (*l-plastin*). Para além destes marcadores, também analisámos a expressão de *dia1r*. A expressão de *dia1r* pareceu estar reduzida nos crispantes em comparação com embriões não-injetados, tal como seria de esperar num mutante mosaico para *dia1r*. As nossas observações também mostraram que a expressão de *c-myb* se encontra aumentada na ICM de crispantes. Juntamente com a observação de que o *lmo2* tem maior expressão nos morfantes, os nossos resultados sugerem que a população de progenitores hematopoiéticos é maior quando se reduzem os níveis de expressão de *DIA1R*. Ao contrário do observado em embriões morfantes, os padrões de expressão de *gata1* e *l-plastin* não foram afetados em embriões crispantes. Tal pode ser devido à existência de um baixo mosaicíssimo, ausência de mutações, ou mutações *in-frame* nos embriões crispantes de *dia1r* que foram testados para a expressão de *gata1* e *l-plastin*.

De um modo geral, a nossa análise funcional indica que DIA1R pode restringir o tamanho da população de progenitores hematopoiéticos. Nós também levantamos a hipótese de que DIA1R pode controlar a proliferação de células progenitoras hematopoiéticas e/ou induzir a sua diferenciação para a linhagem mielóide.

Palavras-chave: Hematopoiese; Hemangioblastos; Desenvolvimento; Peixe-zebra; DIA1R

Abstract

During embryonic development, vasculogenesis and hematopoiesis are intimately associated. The hemangioblast is known as the common precursor cell of both endothelial and hematopoietic cells.

While studying the gene expression profile of yolk sac hemangioblasts, *CXorf36* or *deleted in autism 1 related* (*dia1r*) was identified as one of the most highly expressed genes. Previous experiments in chick and mouse embryos have shown that *dia1r* is expressed in yolk sac hemangioblasts and hemogenic endothelium, the main progenitors of primitive and definitive hematopoietic cells, respectively. This suggested that *dia1r* may have a significant role in blood development. Using the zebrafish as model system, we set out to characterize the pattern of expression and function of DIA1R throughout blood development.

We analyzed *dia1r* expression by whole-mount *in situ* hybridization (WISH) and compare it with the expression patterns of different hematopoietic markers. We found that *dia1r* expression starts to be detected at 12hpf in the anterior lateral mesoderm and posterior lateral mesoderm, the first hematopoietic tissues of the zebrafish. At 19,5hpf, *dia1r* is expressed in the intermediate cell mass (ICM), the hematopoietic tissue that derives from the PLM, whereas at 30hpf it is detected in the dorsal aorta and caudal hematopoietic tissue, where hematopoietic stem/progenitor cells are present.

To assess DIA1R function, we knocked down its translation using morpholinos, and started to generate a *dia1r* mutant line using CRISPR/Cas9 technology. For the phenotypical analysis of *dia1r* morphants, we used quantitative PCR (qPCR) to detect alterations in the expression levels of different hematopoietic markers, and WISH to analyze the distribution and size of different blood cell populations. The latter method was also used in the analysis of *dia1r* crispants (*i.e.*, F0 embryos injected with gRNA and Cas9). The qPCR analysis of *dia1r* morphant embryos showed an up-regulation of *lmo2,* a hematopoietic progenitor cell marker, and down-regulation of *l-plastin,* a macrophage marker. Concurrently, the WISH analysis of *dia1r* morphants revealed that they had more erythromyeloid progenitors (derived from the ICM) and fewer macrophages. In addition, crispant embryos had slightly more cells expressing *c-myb*, another hematopoietic progenitor cell marker.

Taken together, our results suggest that DIA1R might restrict the size of the hematopoietic progenitor cell population, either by controlling their proliferation or by promoting their differentiation into the myeloid lineage.

Key-words: Hematopoiesis; Hemangioblast; Development; Zebrafish; DIA1R

Index

List of Figures

Abbreviations

CDS – coding sequence

Dpf – days post-fertilization

DIA1R – deleted-in-autism 1 related

DAPI - 4',6 – diamidino-2-phenylindole

Hpf – hours post-fertilization

zDIA1R – zebrafish variant of DIA1R

cDIA1R – chick variant of DIA1R

mDIA1R – mouse variant of DIA1R

PBS - Phosphate buffered saline

GFP – Green fluorescent protein

Scl/tal-1 – stem cell leukemia/T-cell acute lymphocytic leukemia

runx1 – runt-related transcription factor 1

gata1 – GATA-binding protein 1

mpx – myeloid-specific peroxidase

c-myb - v-myb avian myeloblastosis viral oncogene homolog

csf1ra - colony stimulating factor 1 receptor a

gata2b - GATA binding protein 2b

lmo2 – LIM domain only 2

α-ehb – alpha-embryonic hemoglobin

spi1 - Spi-1 proto-oncogene

flk1/vegfr2 – fetal liver kinase 1vascular endotelial growth factor receptor 2

O/N – Overnight

CRISPR – Clustered regularly interspaced short palindromic repeats

PFA – paraformaldehyde

ef1α - elongation factor 1 α

rpl13a - ribosomal protein L13a

bp – base-pair

dNTP - deoxyribonucleotide triphosphate

min – minute(s)

h – hour(s)

KO – knockout

gRNA – guide RNA

ISH – in situ hybridization

WISH – whole-mount in situ hybridization

UTR – untranslated region

qPCR – quantitative PCR

RT-qPCR – Reverse transcriptase qPCR

Nt – nucleotide(s)

NHEJ – non-homologous end joining

HSPC – hematopoietic stem/progenitor cell

EHT – endothelial-to-hematopoietic transition

CHT – caudal hematopoietic tissue

MO – morpholino

CtMO – control morpholino

1 Introduction

1.1 Hematopoietic development in the vertebrate embryo

In the developing vertebrate embryo, few tissues are as crucial to the survival of the embryo as are hematopoietic and vascular system tissues. The hematovascular system provides oxygen and nutrients and removes toxic metabolic waste from the growing tissues, whereas embryonic-derived macrophages help tissues reshape by phagocyting apoptotic corpses and protect the embryo against pathogenic agents (Ginhoux and Guilliams, 2016; Gritz and Hirschi, 2016; Herbomel et al., 1999; Marcelo et al., 2013). In the hematopoietic system, the hematopoietic stem/progenitor cells (HSPCs) give rise to all erythroid, myeloid and lymphoid cells that will compose the immune system of the organism. Disruptions during HSPCs formation can be so severe that embryo development is no longer viable. In addition, non-lethal disruptions can lead to hematologic diseases in the adult, such as anemia and leukemia. Therefore, knowing the players of HSPC development is essential for the treatment of hematologic diseases. In the future, these treatments may take advantage of the *in* vitro generation of HSPCs from human induced pluripotent stem cells for personalized blood transplantation (Kulkeaw and Sugiyama, 2012; Robertson et al., 2016).

In the embryo, the hematopoietic system develops in two waves that are temporarily and spatially restricted and give rise to distinct hematopoietic lineages. These waves are called primitive and definitive waves. (Gritz and Hirschi, 2016; Lacaud and Kouskoff, 2017; Robertson et al., 2016).

1.1.1 Primitive wave of hematopoiesis

In the primitive wave of hematopoiesis, blood cells emerge from the blood islands present in the yolk sac. These blood islands contain the first hematopoietic and endothelial cells of the vertebrate embryo. This close developmental relationship between hematopoiesis and vasculogenesis lead to the hypothesis that these two lineages have a common progenitor, named the hemangioblast (Murray, 1932; Sabin, 1920). The existence of this cell population has been supported by *in vitro* differentiation studies (Choi et al., 1998), as well as by evidence that hematopoietic and endothelial progenitors express several genes in common, such as *vascular endothelial growth factor receptor 2* (*vegfr2* or *flk1*), *LIM domain only 2* (*lmo2*) and *T-cell acute lymphocytic leukemia* (*scl*) (reviewed in Xiong, 2008). However, recent evidence suggests that hemangioblasts may not be true bipotential precursors but rather give rise to hematopoietic cells through two types of intermediate progenitors: hemogenic angioblasts and hemogenic endothelial cells. In the blood islands, hemangioblasts give rise to embryonic erythrocytes, macrophages and megakaryocytes in the primitive wave of hematopoiesis (reviewed in Lacaud and Kouskoff, 2017). Embryonic erythrocytes are different from definitive erythrocytes because they are nucleated and predominately express embryonic forms of globulin (Kingsley et al., 2006; Palis, 2014). In contrast, primitive macrophages and megakaryocytes are functionally identical to the ones formed during the definitive wave of hematopoiesis (Tober et al., 2007).

1.1.2 Definitive wave of hematopoiesis

The definitive wave is characterized by the generation of hematopoietic progenitor cells that will give rise to all definitive blood lineages in the adult. Erythromyeloid progenitors (EMPs), the first of such progenitors, are transient hematopoietic progenitor cells that arise in the yolk sac, migrate to the fetal liver and undergo differentiation to produce definitive erythrocytes and most myeloid progenitors but not lymphoid cells (Ginhoux and Guilliams, 2016; McGrath et al., 2015). These myeloid progenitors give rise to the tissue-resident macrophages that can be found in the adult liver (Kupffer cells), central

nervous system (microglia), epidermis (Langerhans cells) and lungs (alveolar macrophages) (Gomez Perdiguero et al., 2015).

The final blood progenitors are the hematopoietic stem/progenitor cells (HSPCs). These cells arise from the hemogenic endothelium present in the ventral wall of the dorsal aorta via a process named endothelial-to-hematopoietic transition (EHT) (Ciau-Uitz et al., 2014; Monteiro et al., 2016). The HSPCs migrate to the transient hematopoietic tissues, where they will undergo massive proliferation and serve as a source of embryonic macrophages, neutrophils and monocytes. At later stages, these progenitors colonize the bone marrow (the final hematopoietic tissue) which provides a niche for HSPCs. There, they continue to proliferate and differentiate throughout adulthood into myeloid, erythroid, thromboid and lymphoid lineages, whereas the thymus will produce mature lymphoid T cells (Ciau-Uitz et al., 2014; Gore et al., 2018).

1.2 Sites and genetic regulation of zebrafish blood development

Even though zebrafish (*Danio rerio*) and humans are 400 million years apart, the development and function of embryonic blood and HSPCs is evolutionary conserved (Ciau-Uitz et al., 2014; Reischauer et al., 2016; Soza-ried et al., 2010). Zebrafish have long been used to study hematopoiesis due to their practical advantages, such as easy reproduction, large progeny, transparent embryos with external and fast development, and a fully sequenced genome. Genetic engineering in zebrafish is relatively straightforward. This enables the creation of transgenic reporter lines that, amongst other things, allow for the observation of *in vivo* blood cell dynamics throughout development using high resolution microscopy (Moro et al., 2013; Robertson et al., 2016).

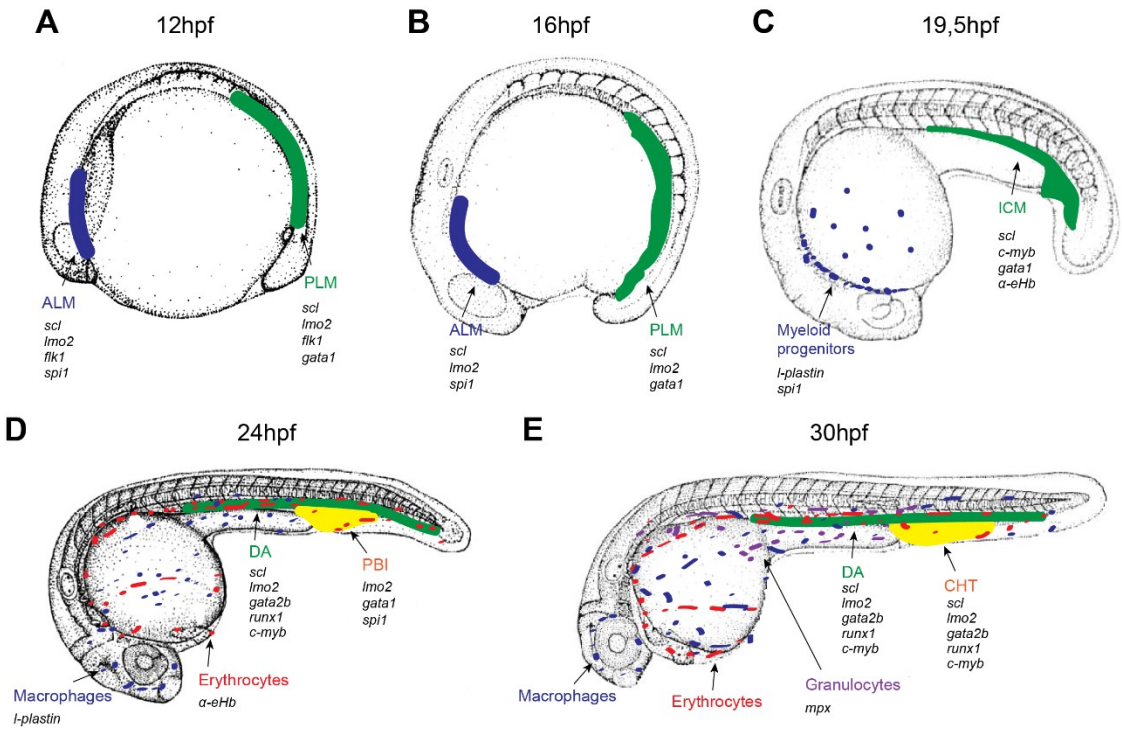


Figure 1.1 - Sites and markers of primitive and definitive hematopoiesis throughout zebrafish development
(A) Hemangioblasts arise in the anterior and posterior lateral mesoderm (ALM in blue and PLM in green, respectively). Like in other studied animals, hemangioblasts express both hematopoietic and endothelial genes. (B) At 16hpf, hemangioblast have differentiated between angioblasts and hematopoietic progenitors. (C) By 19,5hpf, hematopoietic progenitor cells start to differentiate into myeloid progenitors in ALM and primitive erythrocyte precursor cells in the intermediate cell mass (ICM). (D) At 24hpf, blood circulation starts and the first erythrocytes begin to oxygenize the developing embryo. Macrophages become functional and some migrate into different tissues to become tissue-resident macrophages (*e.g.*, microglia in the central nervous system). The ICM will give rise to the dorsal aorta (DA) and from it, the first HSPCs are born. The most posterior part of the ICM becomes the posterior blood island (PBI), where the transient EMP population is formed. (E) Between 24hpf and 30hpf, HSPCs bud of the DA and migrate to the caudal hematopoietic tissue (CHT), which originates at 30hpf from the PBI. In the CHT, HSPCs will proliferate temporally before migrating to their final hematopoietic sites (kidney marrow and thymus), where they will originate all adult blood and immune cells. Also at around 30hpf, granulocytes and macrophages arise from EMPs present in the CHT (Ciau-Uitz et al., 2014; Gore et al., 2018); Figure adapted from http://www.uoneuro.uoregon.edu/k12/Table%202.html.

Unlike birds and mammals, zebrafish do not form blood islands in the yolk sac. Instead, hemangioblasts and hematopoietic progenitors can be found in different regions of the developing embryo, as hereafter described. In the zebrafish, hemangioblasts are first specified in the anterior and posterior lateral mesoderm (ALM and PLM, respectively), the first hematopoietic tissues of the zebrafish, at the tailbud stage (10hpf) by the gene *npas4l*. This gene is responsible for the *cloche* mutant phenotype (Reischauer et al., 2016). Moreover, *npas4l* was identified as the master regulator of hemangioblast specification. Npas4l regulates the expression of *Ets variant 2* (*etv2*) and *scl*, the first genes expressed in the early endothelial and hematopoietic progenitor cells, respectively (Reischauer et al., 2016). Etv2 is an ETS transcription factor critical for vasculogenesis (Sumanas and Lin, 2006). Scl is a basic-helix-loop-helix transcription factor. *scl* and its binding partner LIM domain 2 (*lmo2*) are expressed in hemangioblasts present in the ALM and PLM. Both these genes are necessary for zebrafish hematopoiesis, given that the depletion of either *lmo2* or *scl* results in loss of primitive and definitive hematopoiesis (Dooley et al., 2005; Patterson et al., 2007).

At the 5-9 somites stage (around 12hpf), hemangioblasts present in the ALM and PLM start to have some differences in gene expression. Hemangioblasts located in the ALM express *Spi-1 proto-oncogene b* (*spi1*) while the ones in the PLM express both *spi1* and *GATA binding protein 1a* (*gata1*) (Paik and

Zon, 2010). The ETS-domain transcription factor Spi1 is a master regulator of myeloid cell development, including granulocyte and macrophage development (Scott et al., 1994), whereas the zinc-finger transcription factor Gata1 is a master regulator of erythrocyte development (Cantor and Orkin, 2002). The cross-antagonistic relationship between Gata1 and Spi1 is essential to determine the lineage commitment of hematopoietic progenitors to either the erythroid or the myeloid fate (Monteiro et al., 2011). Later, at 14-19 somites stages (around 16hpf), hemangioblasts in the ALM differentiate into myeloid progenitors. These progenitors start to express *leucocyte-specific plastin* (*l-plastin*), encoding an actin-bundling protein, and migrate beneath the cardiac tube towards the yolk-sac (Herbomel et al., 1999). In the PLM, Gata1 auto-regulates its own expression and suppresses *spi1,* leading to the specification of primitive erythroblasts (Detrich et al., 1995; Monteiro et al., 2011). These erythroblasts will later give rise to primitive erythrocytes that express embryonic forms of hemoglobulin, *α-eHB* and *β-eHb* (Kingsley et al., 2006; Ma et al., 2010; Palis, 2014). At this stage, hematopoietic progenitors and primitive erythroblasts start migrating to the midline to form the ICM, which is completed by the 20-25 somites stage (around 19,5hpf, (Detrich et al., 1995)). Also around the 20-25 somites stage, macrophage progenitors in the yolk-sac express *macrophage colony-stimulating factor receptor a* (*csf1ra*) which encodes a type III receptor tyrosine kinase (Herbomel et al., 2001; Parichy et al., 2000).

The dawn of the first heartbeat is at 24hpf (or prim-5 stage) and kickstarts the beginning of circulation. In the most posterior region of the ICM, a blood island arises. This posterior blood island (PBI) will give rise to a $gata1^{+}/lmo2^{+}$ population which are the EMPs (Bertrand et al., 2007). Like in the primitive wave, *gata1* and *spi1* cross-antagonism determines the erythroid vs. myeloid decision (Monteiro et al., 2011). The first granulocytes that arise from these EMPs can be distinguished from monocytes by the expression of *myeloid-specific peroxidase* (*mpx*), a peroxidase characteristic of neutrophil primary granules (Lieschke et al., 2001). Also around 24hpf, the ICM forms the dorsal aorta (DA) in the medial region of the zebrafish embryo. The hemogenic endothelium located at the ventral wall of the DA will give rise to definitive HSPCs *via* EHT and there are a few genes crucial to this phenomenon (Bertrand et al., 2010a; Ciau-Uitz et al., 2014). Gata binding protein 2b (Gata2b) is a transcription factor that specifies the hemogenic endothelium in the ventral wall of the DA (Butko et al., 2015). Downstream of Gata2b is runt-related transcription factor (Runx1), which is essential for the specification and EHT of HSPCs in the DA (Bresciani et al., 2014). Another important transcription factor, v-myb avian myeloblastosis viral oncogene homolog (c-Myb), acts downstream of Runx1 in the HSPCs (Burns et al., 2005) and has a central role in definitive hematopoiesis (Soza-ried et al., 2010).

In zebrafish, HSPCs will bud out of the hemogenic endothelium into the area between the DA and the posterior cardinal vein. After they emerge, HSPCs enter the circulation and journey onto the caudal hematopoietic tissue (CHT). The CHT has the same function in the zebrafish as the fetal liver in mammals, acting as a transient hematopoietic tissue where HSPCs temporally proliferate and differentiate. At the final stage of embryonic hematopoiesis, these cells will migrate to the kidney marrow, which is the zebrafish equivalent of the mammalian bone marrow (Ciau-Uitz et al., 2014; Murayama et al., 2006).

1.3 Identification of *DIA1R* as a novel hemangioblast gene

While studying the chick Cerberus transcriptional regulation (Tavares et al., 2007), Tavares et al isolated a cis-regulatory region that drives reporter gene expression specifically in yolk sac hemangioblasts. This reporter (Hb-eGFP) labels hemangioblasts since their initial migration from the primitive streak into the extraembryonic mesoderm. (Gordon-Keylock and Medvinsky, 2011; Teixeira et al., 2011; Figure 1.2). The Hb-eGFP reporter was used to isolate the hemangioblast population and characterize its gene

expression profile by microarray analysis (Teixeira et al., 2011; Figure 1.2). In addition to known hemangioblast markers, this analysis led to the identification of several previously uncharacterized genes that may be involved in hemangioblast differentiation. The second most highly differentially expressed gene was the avian ortholog of human cXorf36 or deleted in autism-1 related (DIA1R), a gene implicated in autism spectrum disorders (ASD) and X-linked mental retardation (Aziz et al., 2011a). DIA1R genes are found exclusively in vertebrates, but their function is largely unknown (Aziz et al., 2011b).

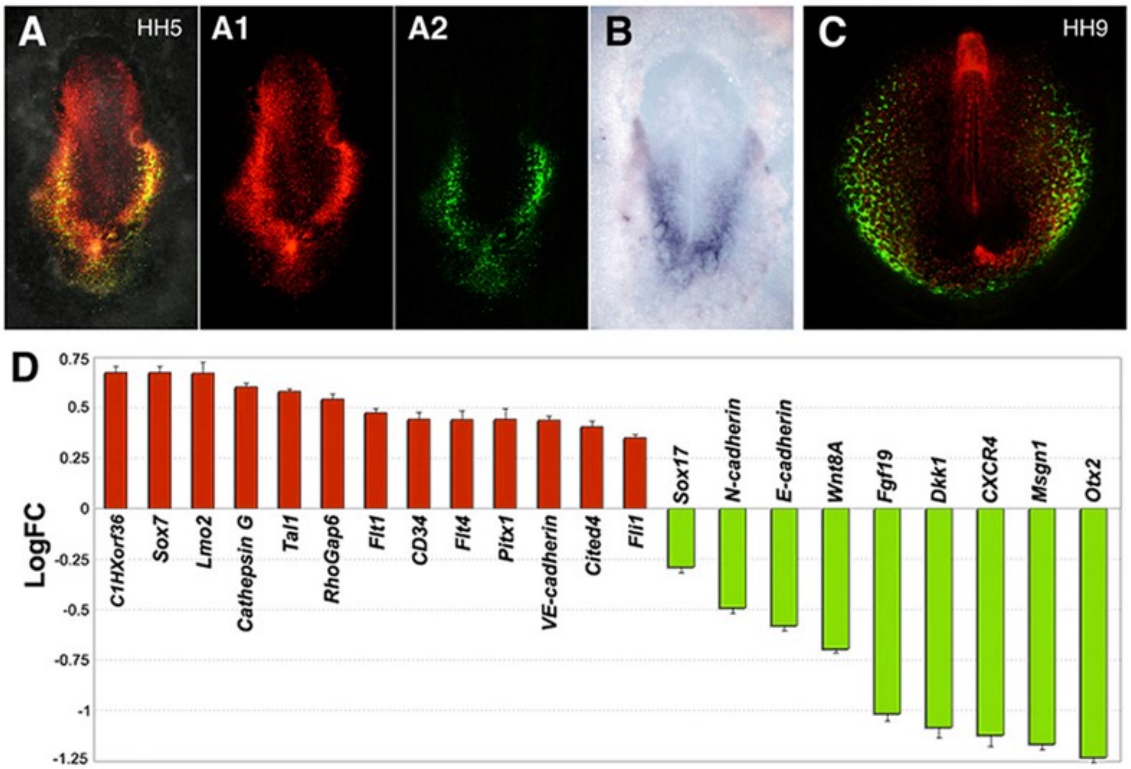


Figure 1.2 - Hb-eGFP reporter expression and microarray analysis
Hb-eGFP reporter is expressed in the posterior extraembryonic population of hemangioblasts in the chicken embryo at HH5 (Hamburger and Hamilton, 1951) (A). This is corroborated by *vegfr2* expression in the same cells (B). At later stages, cells expressing eGFP aggregate into blood islands (C). To validate the reporter and to characterize the gene expression profile of hemangioblasts, the Hb-eGFP+ (A2) and RFP+ (A1) cell populations were isolated and their expression profiles were compared. As expected, the microarray data shows enrichment of known hemangioblast markers (Scl/tal1, Lmo2, Fli1) and down-regulation of markers for other cell types (Sox17 – mesoderm, Otx2 – neuroectoderm) in the Hb-eGFP-positive population. cDIA1R (or C1HXorf36) is one of the most highly expressed genes in hemangioblasts (D). Adapted from Teixeira et al., 2011.

Based on sequence analysis, DIA1R proteins were predicted to have a signal peptide and a highly conserved PIP49_C protein-kinase domain, characteristic of the FAM69 family of kinase-like proteins (Aziz et al., 2011a; Dudkiewicz et al., 2013). However, it does not have a conserved active site, thus indicating that DIA1R may be either an atypical kinase or a pseudo-kinase. These features suggest that DIA1R proteins may regulate molecular traffic or interfere with the function of secreted factors (Aziz et al., 2011a).

Concerning gene expression, our group performed preliminary expression studies in the chick and mouse embryos. The analysis of *cDIA1R* expression during chick development revealed that this gene is expressed in yolk sac hemangioblasts, blood islands and endothelial cells of the dorsal aorta,

endocardium and head vasculature (Figure 3), which are regions known to have hemogenic potential (Swiers et al., 2013). In particular, *cDIA1R* expression appears to be higher in cells that are morphologically similar to hemogenic endothelial cells (Figure 3 D') (Serrado Marques et al., 2018). In the brain of the mouse embryo, *mDIA1R* transcripts are also detected in the vascular endothelium and microglia (Figure 3 G, H, respectively). Consistently with our observations, transcripts of *DIA1R* orthologs have been detected in hematopoietic progenitors and endothelial cells of zebrafish embryos (Cannon et al., 2013; *cc058* gene), in endothelial and microglial cells from embryonic and adult mouse brains (Gay et al., 2013; *4930578C19Rik* gene) and in human endothelial cells (Bhasin et al., 2010). Taken together, these findings suggest that DIA1R may be a good marker and potential regulator of the intraembryonic hemogenic lineages that derive from yolk sac hemangioblasts.

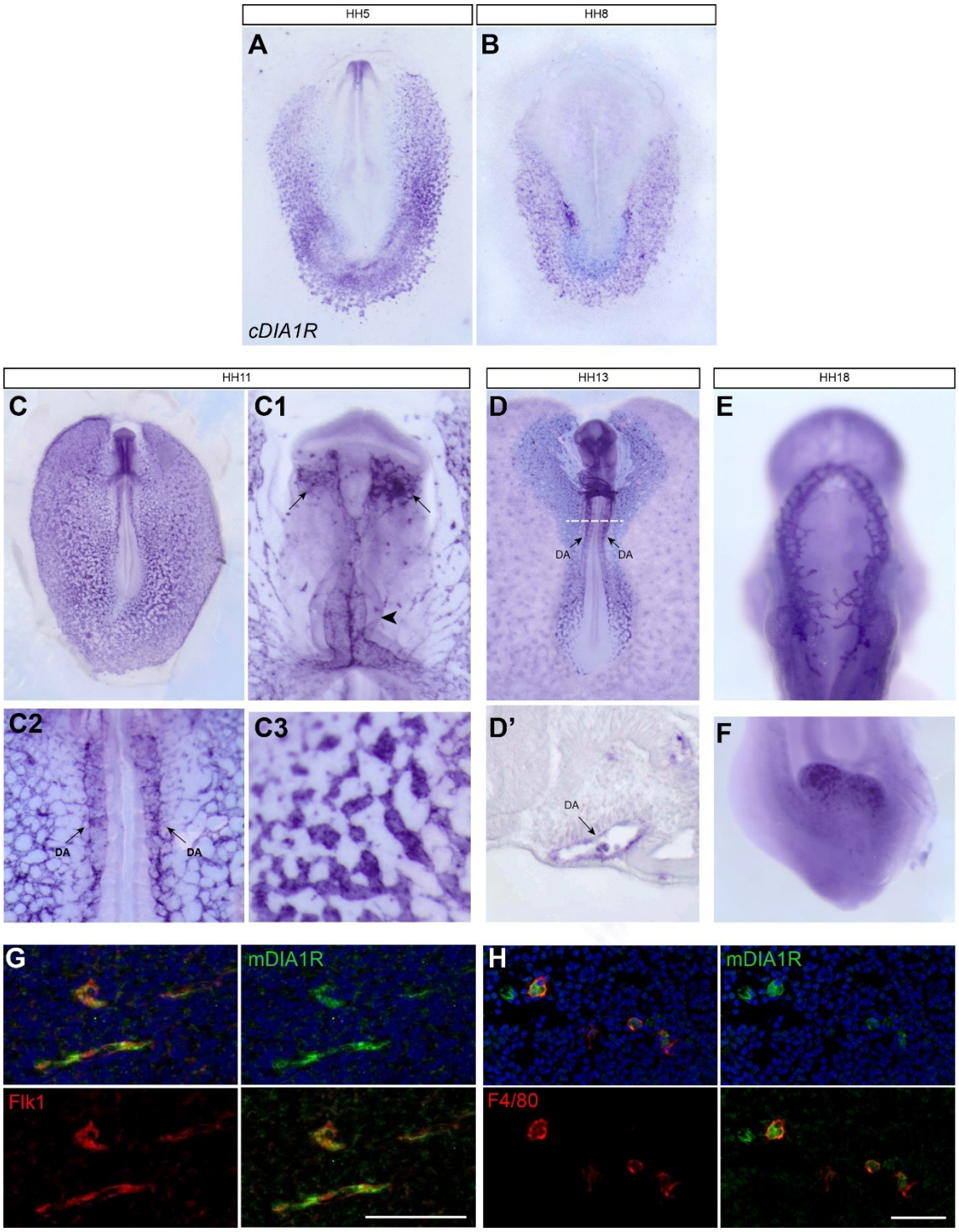


Figure 1.3 - DIA1R expression in chick and mouse embryos.
In the chick embryo, *cDIA1R* expression starts to be detected at HH5 (A) in the extraembryonic mesoderm that will form the yolk sac blood islands (B, C, C3). At HH11, *cDIA1R* is also expressed in the developing head vasculature (arrows in C1), endocardium (arrowhead in C1) and dorsal aorta (DA) (C2). At later stages, *cDIA1R* is detected in the blood vessels of the embryo, with high intensity in the dorsal aorta (D, D'), head vasculature (E) and allantois (F). In the mouse neuroepithelium, *mDIA1R* is detected both in the neurovasculature (G) and in microglial cells (H). Chick images are adapted from Serrado Marques et al., 2018. Mouse images are unpublished data. Scale bar – 50 µm.

1.4 Objectives

Previous work in the lab has found that *dia1r* is expressed in hemangioblasts and hemogenic endothelium of the chick and mouse embryos. However, neither the expression pattern nor the function of the gene have not been significantly addressed in zebrafish, an excellent model for studying blood development (Gore et al., 2018). With that in mind, the goals of the project were to:

1) Characterize the pattern of expression of *zDIA1R* (henceforth *dia1r*) throughout zebrafish development; and

2) Decipher a possible role for *dia1r* in hematopoietic development.

2 Methods

2.1 Zebrafish husbandry and embryo collection

Zebrafish wild-type lines used in this project were maintained in a re-circulating system with a light cycle of 14 hours (h) of light and 10h of darkness per day, with water temperature at 28ºC. Embryos used in experiments were grown at 25ºC, 28ºC or 33ºC in Embryo Medium and collected at different developmental stages. These stages were shield (6hpf), bud (10hpf), 3-somite (11hpf), 6-somite (12hpf), 10-somite (14hpf), 14-somite (16hpf), 21-somite (19.5hpf), 26-somite (22hpf), prim-5 (24hpf), prim-15 (30hpf), long-pec (48hpf), protruding-mouth (3dpf) and day 4 larvae (4dpf). Even though 3dpf and 4dpf zebrafish are considered to be larvae (Westerfield, 2007), we will name them "embryos" from here on for the sake of clarity and conciseness.

2.2 Morpholino injection

DIA1R antisense morpholino and 5bp-mismatch control-morpholino (GeneTools) were used to assess DIA1R function (Supplementary table 3). MO or CtMO injections were performed in embryos of the same posture. To determine the ideal working concentration, 0.75, 1, 1.5 and 2ng of morpholino diluted in Danieu medium were injected in embryos at one-cell stage using glass capillaries, a pressure injector (PV820 Pneumatic PicoPump; hold pressure = 3psi; eject pressure = 20psi) and a Nikon SMZ745 stereoscope. The volume injected was 1.4nL. Once they reached the appropriate developmental stage, embryos were either fixed in 4%PFA for whole-mount *in situ* hybridization or dissociated and homogenized in Lysis Buffer (from the RNeasy Mini kit) for RNA extraction.

2.3 Whole-mount *in situ* hybridization (WISH)

In order to determine where DIA1R is expressed, zebrafish embryos were processed for whole-mount *in situ* hybridization (ISH) at the development stages defined above (subchapter 2.1), as described hereafter. The WISH method employed was the Thisse's protocol (Thisse and Thisse, 2008) with a few modifications, as described hereafter.

2.3.1 RNA probe generation

In order to generate the DIA1R riboprobe, we started by amplifying and cloning two sequences of the zebrafish cDNA: *zDIA1R* cannon (Cannon et al., 2013a), which spans the entire coding sequence (1272bp) and *zDIA1R* short, which spans a shorter sequence in the 3' region (511bp). Using total RNA from 24hpf zebrafish embryos, we amplified the cDNA with the Transcriptor High Fidelity cDNA Synthesis kit (Roche), following the available protocol. To make the template-primer mix, random hexamer primers were used. The primers used for the amplification of *zDIA1R* cannon and *zDIA1R* short probes are in Supplementary table 1.

For each probe, PCR mix was prepared with 2µL of cDNA solution, 10X Taq Buffer (containing 20mM MgCl2), 5µL of deoxyribonucleotide triphosphate (dNTP) mix (with 2mM each), 1.0µM of each primer, and 0.2µL of Taq DNA Polymerase (Thermo Fisher Scientific), in a final reaction volume of 25µL. Briefly, reactions started with an initial denaturation step at 95°C for 1 minute (min), followed by 20 cycles of denaturation at 95°C for 30 sec, annealing at 72°C for 1min, and extension at 72°C for 1min, and a final extension step at 72°C for 10min. To increase the amount of amplified product, a second PCR was performed using the same conditions but 2µL of the first PCR reaction instead of the cDNA.

The PCR products were detected by electrophoresis on a 1% agarose gel, and purified with QIAquick gel extraction kit from QIAGEN.

The vector used for cloning and transcribing the *dia1r* probes was pGEM-T easy vector by Promega. We followed Promega's protocol for the DNA ligation and transformation procedures, using in-house DH5α competent cells. pGEM-T easy vector has two polymerase promoters (T7 and SP6) which allows for transcription of sense and anti-sense RNA probes, regardless of the orientation of the DNA fragment.

In addition to *dia1r*, we used riboprobes for several markers of different hematopoietic differentiation stages: *cMyb*, *runx1*, *lmo2*, *gata1*, *gata2b*, *mpx*, *l-plastin* and *csf1ra* (Supplementary table 2). The transcription vectors for these markers were given to us by Rita Fior (Champalimaud Foundation), which she received from Rui Monteiro (University of Oxford; who gave us his consent). The provided DNA samples were amplified using the same competent cells, in order to obtain the required amount of plasmid DNA for probe synthesis.

After successful transformation, DNA preparation was performed using QIAprep Spin Miniprep Kit by QIAGEN, and DNA concentration was measured on NanoDrop 2000 (Thermo Scientifc). To verify the orientation in which *zDIA1R* cannon and *zDIA1R* short cDNAs were insert, we performed double digestion with *EcoR*V and *Pst*I restriction enzymes. *EcoR*V restriction site is near the 3' end of both *dia1r* probes (insert) and *Pst*I restriction site is near the beginning of the SP6 promoter (vector). After digestion and gel electrophoresis, we got a short band (around 60bp) for both probes. This tells us that SP6 promoter will give us the anti-sense probe and T7 promoter will transcribe the sense (or control) probe.

For the transcription of the RNA probes, it is essential to use RNase-free material. In brief, vectors containing the cDNA for the probes were linearized with appropriate restriction enzymes, and transcribed with an RNA polymerase that gives us the anti-sense probes (and *dia1r* sense probe as control), as listed in Supplementary table 2. The nucleotides used were labeled with digoxygenin (DIG). The RNA presence was confirmed by electrophoresis on a 1% agarose gel, and its concentration measured on NanoDrop 2000.

2.3.2 Embryo fixation

Upon reaching the desired developmental stage, zebrafish embryos were dechorionated, fixed overnight in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) at 4°C, dehydrated through a series of methanol/PBS solutions (25%, 50%, 75% and 100% methanol) and stored at -20°C for a minimum of 2h (embryos can be stored for months in 100% methanol at 20°C). For the WISH protocol, fixed embryos were rehydrated and washed twice in PBS plus 0.1% Tween 20 (PBT) for 5min each. When necessary, pigmentation was removed by bleaching the embryos with a 3% H2O2/1% KOH solution (this process can take up to 1h, depending on embryo's age).

2.3.3 Day 1 - Permeabilization and hybridization of the embryos

On the first day, embryos were washed twice in PBT for 5min and digested with proteinase K for the duration specified in the protocol, except for the older ones (*i.e.*, 40min for 3dpf larvae). After digestion, embryos were immediately re-fixed in 4%PFA for 20min, washed five times in PBT for 5 min each and pre-hybridized in pre-hybridization solution (50%formamide, 5x SSC, 0.1%Tween 20, 9.2mM citric acid to pH6; preHYB-MIX) at 70°C for 2h. Then, embryos were incubated overnight at 70°C in hybridization solution (preHYB-MIX plus 50 µg/mL heparin and 50 µg/mL *Torula* RNA) containing 1.5 ng/µL for the *dia1r* probes and 1 ng/µL for the others.

2.3.4 Day 2 - Washes and incubation with anti-DIG antibody alkaline phosphatase

On the second day, riboprobes were recovered and embryos were rinsed with preHYB-MIX at 70°C, followed by a series of hot washes (70°C) in 75% preHYB-MIX / 25% 2x SSC, 50% preHYB-MIX / 50% 2x SSC, 25% preHYB-MIX / 75% 2x SSC and 2x SSC for 15min each. Embryos were then washed with 0,2x SSC twice for 30 min and went through a series of 0.2x SSC/PBT washes (75%/25%, 50%/50%, 25%/75% respectively) and PBT for 10min each at room temperature (RT). Embryos were incubated in blocking solution (PBT plus 100mg of bovine albumin serum (BSA) and 2% goat serum (GS)) for at least 1h in a horizontal rotator and incubated overnight (O/N) at 4°C in blocking solution with alkaline phosphatase-coupled anti-Digoxigenin antibody (Roche) at a 1:5000 concentration.

2.3.5 Day 3 - Washes and staining

On the third day, embryos were briefly rinsed in PBT, washed in PBT six times, 15 to 60min each and three times in staining buffer (also known as NTMT; 100mM NaCl, 100mM TrisHCl pH 9.5, 50mM MgCl2, 0.1% Tween 20) for 5min each. Finally, embryos were incubated in staining solution (NTMT plus NBT/BCIP (referino)) in the dark at RT for 30min to 20h.

2.4 RNA isolation and real-time quantitative PCR (RT-qPCR)

To evaluate more precisely the effect of *dia1r* knockdown on hemangioblast lineage commitment, the expression of nine genes known to be markers of different hematopoietic differentiation stages was analyzed by RT-qPCR. The primers used for qPCR amplification can be found in Supplementary table 4.

Total RNA was isolated from a pool of 50 wild-type or morpholino-injected embryos at 21-somite stage using the RNeasy Mini Kit (QIAGEN), and cDNA was synthetized from 1µg of RNA using Transcriptor High Fidelity cDNA Synthesis Kit (Roche). RT-qPCR was performed using FastStart SYBR Green Master Mix (Roche) on a LightCycler 96 system (Roche). Each 10 ul reaction had 5µL of SYBR Green (stock was 2x concentrated), 3µL (1µM) of primers and 2µL (10ng) of cDNA. The RT-qPCR conditions are described in Supplementary table 5.

SYBR Green is fluorescent cyanine dye that binds to double-stranded DNA. The fluorescence signal will serve as a proxy for the amount of DNA. The level above which the fluorescent signal is detected is known as "cycle threshold" or Ct. During the PCR exponential phase, the fluorescent signal doubles with each cycle as does the target DNA template (amplicon). Since the Ct is logarithmically proportional to the amount of DNA, we can compare the expression levels of different genes (Schefe et al., 2006). To normalize the expression of each gene of interest, we subtract its Ct with the Ct of a housekeeping (HK) gene in the same sample. This is known as the ΔCt method and permits comparison of gene expression among different samples. In our case, we used the Ct average of two HK genes: elongation factor 1 α (*ef1α*) and ribosomal protein L13a (*rpl13a*) (Tavares et al., 2017).

Five different wild-type lines were analyzed in triplicate for morpholino-injected (MO), control morpholino-injected (CtMO) and non-injected (wild-type) samples. To inference statistically significant differences between expression levels, we compared MO values against wild-type values, MO values against CtMO values and CtMO against wild-type values using the t-student test in Microsoft Excel 2016. Figure 3.3 E shows the MO values relative to the CtMO values.

2.5 CRISPR/Cas9 strategy

dia1r mutant zebrafish were generated using the CRISPR/Cas9 system in collaboration with Joana F. Monteiro and Ana C. Certal, at the Fish Platform of Champalimaud Foundation. This technology was used to generate targeted indel mutations in *dia1r* based on well-established protocols (Gagnon et al., 2014; Vejnar et al., 2016). The guide RNAs (gRNA) were designed using CRISPRscan (http://www.crisprscan.org/; Moreno-Mateos et al., 2015). From the resulting list of possible gRNAs with high scores, we selected one that targets the start codon of *dia1r*, thus raising the probability of creating a knockout (KO) mutation, with the sequence 5' - TGGCAGGGATATGGTCAG - 3'.

To verify if the wild-type zebrafish strain used has any allelic variants that differs from the reference genome sequence (Ensembl databases), we amplified and sequenced the gRNA target region from a sample of genomic DNA provided by the Champalimaud Fish Platform. The primers used in this PCR amplification were also used for genotyping the CRISPR mutants (Supplementary table 6).

One cell-stage wild-type zebrafish embryos (Tuebingen strain) were microinjected with the gRNA and Cas9 protein. In zebrafish, most germline-transmitted mutations happen very early after injection (Varshney et al., 2015). By delivering Cas9 protein instead of Cas9 mRNA we expect to speed up the onset of CRISPR/Cas9-mediated genome editing upon injection, increasing the germline mutagenesis rate (Gagnon et al., 2014). After injection, a pool of F0 embryos was collected and genotyped by PCR amplification of the targeted mutation region. Using this methodology, we obtained some crispant embryos for a preliminary study, and a first generation of mutants (e.g., F0 with indel mutations) is currently being grown. WISH was performed on crispant embryos to check for hematopoietic gene expression differences. In the near future, F1 heterozygous founders will be confirmed by genotyping as above, and allele identity will be assessed by Sanger sequencing.

2.6 Imaging methods

Images of WISH stained embryos were taken using a widefield microscope Zeiss Z2 (5x, 10x and 20x objectives), Axiocam 506 Mono camera and ZEN 2012 PRO software. To optimize image contrast, we used a white sheet of paper as background and an external light source to lighten the embryos from above. Z-stack images of the embryos were acquired to have all the focus points in a single file. Then, these z-stacks files were processed with Extended Depth of Field, an executable plugin for Fiji (http://bigwww.epfl.ch/demo/edf/), to have all the regions of interest focused in a single image. Finally, to correct the white color in the WISH images, the white background script was used (https://github.com/pmascalchi/ImageJ_Auto-white-balance-correction/releases/tag/v1.0).

All images were processed with Fiji, Adobe Photoshop CS5 and Adobe Illustrator CC.

3 Results

3.1 Generation of riboprobes for zebrafish *dia1r* and hematopoietic markers

To investigate where *dia1r* is expressed in the zebrafish embryo, we created riboprobes to label *dia1r* RNA to perform whole-mount in situ hybridization (WISH). For the amplification of the riboprobe cDNA templates, we used two different pairs of primers (Supplementary table 1), one pair yielding a fragment with 1276bp (Cannon et al., 2013; henceforth *zDIA1R* cannon; whole CDS) and the other yielding a fragment with 511bp (henceforth *zDIA1R* short; Figure 3.1). Both pairs only differ in their forward primer. Even though *zDIA1R* cannon was already proven to work (Cannon et al., 2013a), we generated a shorter probe to see which one was the best at labelling *dia1r* mRNA. After the insertion of the amplicon in a pGEM-T easy vector, we proceeded to transcribe *in vitro* the two *dia1r* probes, as well as riboprobes for several markers of different hematopoietic differentiation stages - *c-myb*, *gata1*, *runx1*, *gata2b*, *lmo2*, *scl*, *l-plastin*, *csf1ra* and *mpx* (Figure 3.1; reviewed in Ciau-Uitz et al., 2014; Robertson et al., 2016). These markers were used for comparison with *dia1r* expression pattern, as well as for the evaluation of potential hematopoietic phenotypes of *dia1r* morphants and crispants. *Gata2b, runx1, scl* and *lmo2* label hematopoietic stem cells precursors (HSPCs) present in the DA (Barbieri et al., 2016; Monteiro et al., 2016).

When we were testing the various probes in WISH of zebrafish embryos, *scl* and *lmo2* probes worked well the first time, while *gata2b* and *runx1* needed optimization (data not shown). Since these four genes are expressed in the same cell populations (Barbieri et al., 2016; Monteiro et al., 2016), we decided not to use *gata2b* and *runx1* probes. *zDIA1R* short and *zDIA1R* cannon probes were tested in embryos at 16hpf, 24hpf and 30hpf. Since the *zDIA1R* short probe did not give preliminary results (data not shown), we only used the *zDIA1R* cannon probe to study *dia1r* developmental expression (hereafter *zDIA1R* cannon probe = *dia1r* probe).

Even though Cannon and co-workers (Cannon et al., 2013a) have performed WISH for *dia1r*, they did not describe in detail its expression pattern at various stages of development. Since our work is the first to characterize *dia1r* expression in great detail by WISH, a sense probe for the gene was used as a control for the anti-sense probe. The sense probe has the same sequence as the mRNA of the gene so it should not hybridize with it. The absence of labeling when using the sense probe indicates that the signal detected with the anti-sense probe is due to specific binding to its target. In the first WISH performed on zebrafish embryos, we used and compared the labeling of *dia1r* sense and anti-sense probes. As seen in Supplementary figure 1, the sense probe is not detected, thus indicating that our *dia1r* anti-sense probe is specifically labelling *dia1r* mRNA. The embryos hybridized with the sense probe were left developing for an extended period of time, which explains the dark color and high background observed due to substrate precipitation.

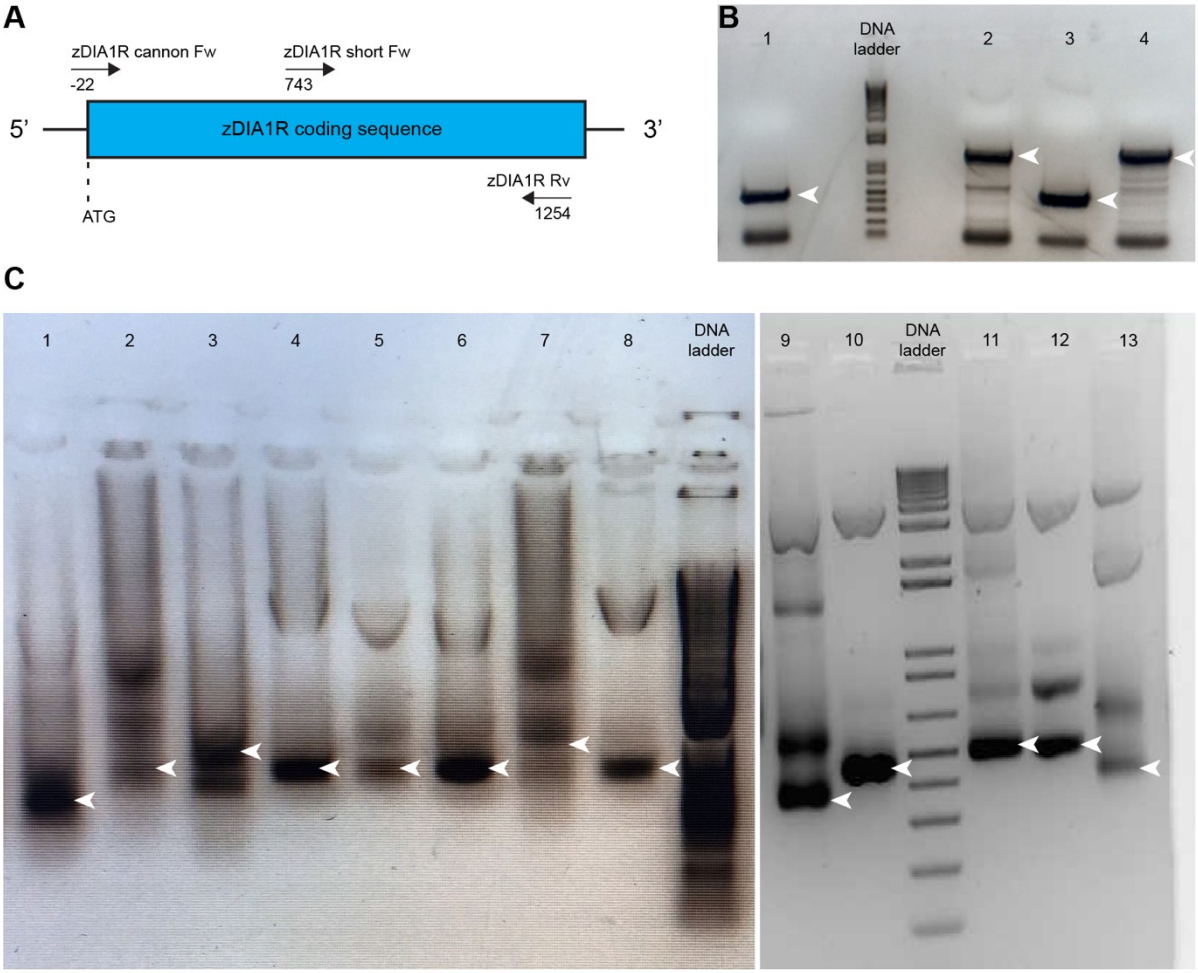


Figure 3.1 - Riboprobes of dia1r and hematopoietic markers
(A) Schematic representation of the primers used for the generation of zDIA1R riboprobes. (B) Electrophoresis gel of PCR amplification of zDIA1R probe templates from cDNA of 24hpf and 48hpf zebrafish embryos. Arrowheads indicate PCR products used for pGEM-T easy cloning. (1) *zDIA1R* short probe amplicon (511bp) from 24hpf cDNA; (2) *zDIA1R* cannon probe amplicon (1276bp) from 24hpf cDNA; (3) *zDIA1R* short probe amplicon from 48hpf cDNA; (4) *zDIA1R* cannon probe amplicon from 48hpf cDNA. (C) Electrophoresis gel of *in vitro* transcription of riboprobes. Arrowheads indicate the riboprobes. (1) *zDIA1R* short anti-sense; (2) *zDIA1R* cannon sense; (3) *zDIA1R* cannon anti-sense; (4) *c-myb*; (5) *gata1*; (6) *runx1*; (7) *gata2b*; (8) *lmo2*; (9) *zDIA1R* short sense; (10) *scl*; (11) *l-plastin*; (12) *csf1ra*; (13) *mpx*. DNA ladder used was 1kb plus from Thermo Scientific.

3.2 Zebrafish *dia1r* pattern of expression

Using the previously mentioned riboprobes, we set out to investigate where *dia1r* is expressed during zebrafish development. *dia1r* is first detected at 12hpf in the anterior lateral mesoderm (ALM) and the posterior lateral mesoderm (PLM), which are the sites where the early precursors of hematopoietic cells are located in the zebrafish (Reischauer et al., 2016). *dia1r* continues to be expressed in the same regions at 16hpf. At these early stages, *dia1r* pattern of expression is similar to that of *scl,* an early marker for hematopoietic progenitor cells (Figure 3.1 A-H). At 19,5hpf, *dia1r* and *scl* are both expressed in the ICM (Figure expression). At 30hpf, *dia1r* is expressed in these same regions and its pattern of expression is very much alike to that of *lmo2*, another marker of HSPCs (Figure 3.1 K-L). At later developmental stages (48hpf to 4dpf) *dia1r* expression could not be detected. WISH for additional hematopoietic markers (Supplementary figure 1) revealed that *dia1r* is not expressed in myeloid cells that derive from the primitive wave of hematopoiesis (Supplementary figure 1 C, E and F). Altogether, these results

allow us to say that *dia1r* is being expressed in regions known to be the birthplace of all hematopoietic stem/progenitor cells.

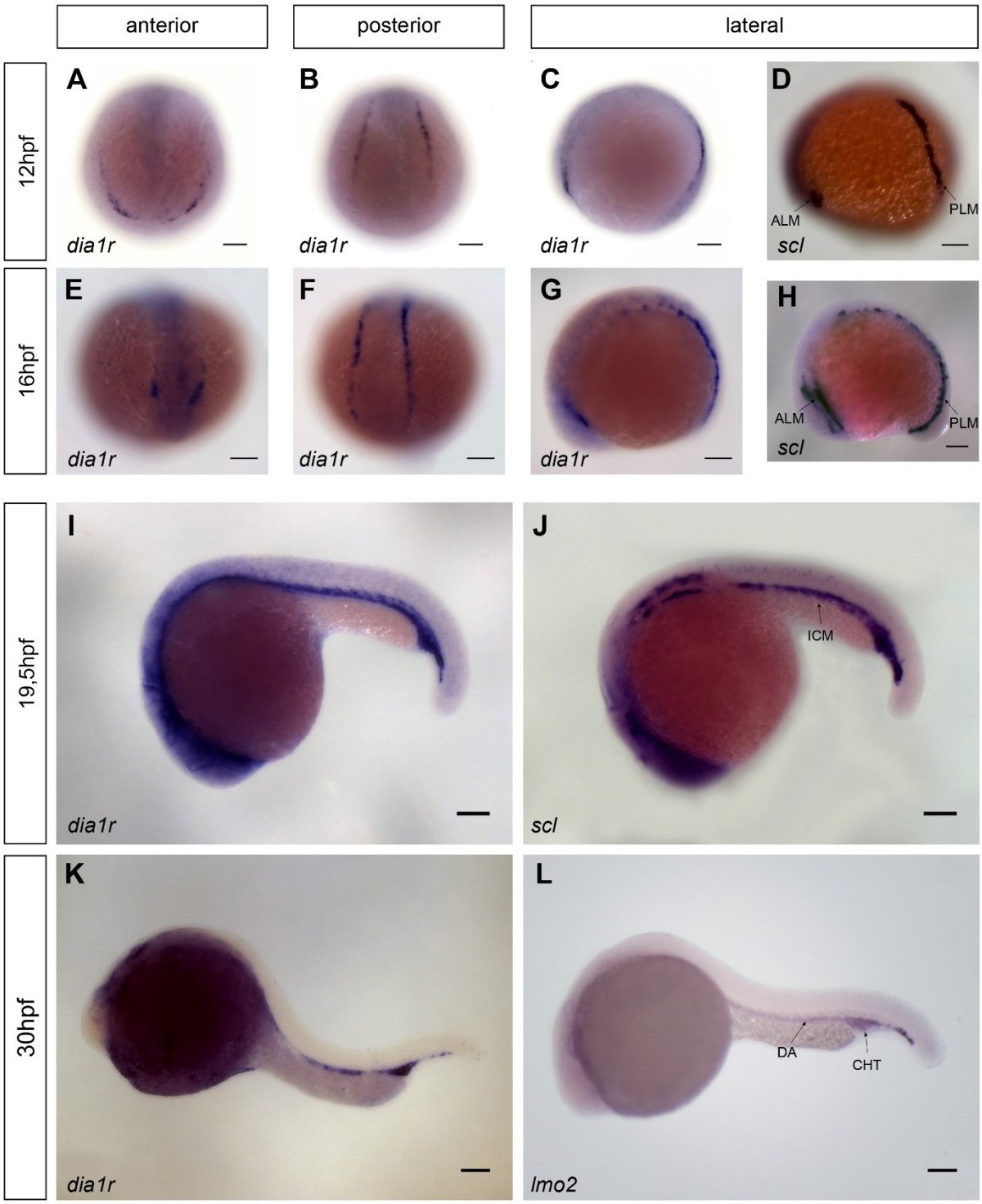


Figure 3.2 - dia1r expression throughout zebrafish development.
WISH of *dia1r* reveals its expression in hematopoietic precursor sites. (A-D) WISH of 12hpf embryos for *dia1r* (A-C) and *scl* (D). *dia1r* is expressed in the ALM and PLM, where we also detect *scl,* one of the earliest genes expressed in hematopoietic precursor cells. (E-H) WISH of 16hpf embryos for *dia1r* (E-G) and *scl* (H). At 16hpf, *dia1r* continues to be expressed in the same sites as before. (I-J) WISH of 19,5hpf embryos for *dia1r* (I) and *scl* (J). At this stage *dia1r* is present in the ICM and PBI, both of which are cell populations that derivate from the PLM. (K-L) WISH of 30hpf embryos for *dia1r* (K) and *lmo2* (L). *dia1r* is expressed in the DA and CHT, the final hematopoietic progenitor sites in the embryos. This expression pattern is

similar to that of *lmo2*, a marker for hematopoietic stem/progenitor cells. ALM – anterior lateral mesoderm; PLM – posterior lateral mesoderm; ICM – intermediate cell mass; PBI – posterior blood island; DA – dorsal aorta; CHT – caudal hematopoietic tissue. Scale bar = 100 μm.

3.3 Analysis of DIA1R function in zebrafish development

After the analysis of the expression pattern of zebrafish *dia1r*, we aimed at investigating its function during development. We first established a *dia1r* knockdown strategy using morpholinos (MOs). MOs are synthetic oligonucleotides with about 25 subunits that are similar to DNA and RNA oligonucleotides, except they have a morpholine ring instead of a ribose ring. MOs interfere with gene expression by either inhibiting translation or preventing RNA splicing. The morpholine ring allows MOs to hybridize with complementary mRNA and DNA while being resistant to nucleases, which makes them more stable than other knockdown techniques (Eisen and Smith, 2008). In our case, we used a 25bp MO that blocks *dia1r* translation (table). MO injection delays embryo development (Stainier et al., 2017; Yajima et al., 2012). To account for it, we used a 5bp mismatch control morpholino (CtMO; Supplementary table 3).

3.3.1 DIA1R knockdown by morpholino injection

In a first study, we used *Fli-GFP:Gata-RFP* (*fli-gata*) transgenic zebrafish for MO-mediated *dia1r* knockdown, which allow for easy screening of phenotypes in the endothelium and blood cells. Previous work in the lab showed that *fli-gata* morphant (MO-injected) embryos seemed to have more *gata1*-labeled cells, suggesting that they could have more erythrocyte precursor cells. However, since these preliminary experiments were performed at another institute (IGC), we decided to repeat the assay at CEDOC in order to further confirm the results. *fli-gata* embryos were injected with 2ng of MO and CtMO. We observed the same phenotype as before, but many *fli-gata* morphants were deformed at 2dpf (87%) and they all died at 3dpf (data not shown). In this assay, all control morphants (*i.e.*, injected with CtMO) were similar to the non-injected embryos (NI) and all lived past 3dpf, indicating that the concentration of the experimental MO should be optimized.

The optimization of MO concentration was performed in wild-type embryos. CtMO concentration was always 2ng/1,4nL, since it was shown to work properly in the *fli-gata* morphants experiments. In the first assay, embryos were injected with 1ng of MO, 2ng of MO and CtMO (Assay 1; Supplementary figure 2). All control morphants and more than half of MO injected embryos died the day after injection. Subsequent assays were performed with lower concentrations (0,5, 0,75 and 1,5ng per 1,4nL; Assays 2 and 3; Supplementary figure 2). In these trials, the mortality rate was reduced to an average of 30%. However, similar numbers were observed in the non-injected group. This was probably because of a problem with the air conditioner in the institute's fish facility during the period when these two assays were performed, which increased the stress of the zebrafish and led to bad postures. Once the facility problem was solved, fish postures improved, and we decided to test once again a higher MO concentration. The final two assays were performed with injections of 2ng MO, and mortality rates in all groups were now lower than before (Assays 4 and 5; Supplementary figure 2). Therefore, we used 2ng/1,4nL as the optimal MO concentration.

For the WISH analysis of *dia1r* morphants, embryos were injected with 2ng of CtMO/MO and fixed at 16hpf, 24hpf and 30hpf. By performing WISH for the hematopoietic markers in morphant embryos we can assess if DIA1R has a role in the development of hematopoietic progenitor cells. At 16hpf, *scl* expression is not affected in morphant embryos, suggesting that DIA1R may act downstream of Scl (Supplementary figure 3). At 24hpf, morphant embryos display a higher expression of *gata1*, which is a marker for both erythromyeloid progenitor (EMP) (Bertrand et al., 2007) and hematopoietic progenitor cells committed to the erythroid lineage (Belele et al., 2009), and a partial loss of *l-plastin* expression, a

macrophage marker (Warga et al., 2016) (Figure morphants). These observations suggest that DIA1R morphants have more EMPs and more erythrocyte progenitor cells and less macrophages. At 30hpf, the expression of the granulocyte marker *mpx* is not altered (Supplementary figure 3; Lieschke et al., 2001), thus indicating that the granulocyte population is not affected in *dia1r* morphants. These first results start to suggest that DIA1R regulates the proliferation and/or differentiation of the erythromyeloid progenitor cell population, restricting the erythroid lineage and promoting primitive macrophage differentiation (Herbomel et al., 1999; Monteiro et al., 2011).

By observing morphant embryos processed by WISH, we can identify the tissues in which gene expression is modified. However, WISH analysis is mainly qualitative, meaning that the interpretation of morphant phenotypes is subject to some observational bias. For the evaluation of *dia1r* morphants phenotype in a semi-quantitative manner, we used real-time quantitative PCR (RT-qPCR or qPCR). With this technique, we assessed the expression of nine genes that serve as markers for hematopoietic stemness (*scl*, *lmo2*, *c-myb, runx1, gata1* and *spi1*) and differentiation (*α-eHb, l-plastin* and *flk1*) (Jin, 2005; Ma et al., 2010; Robertson et al., 2016). We decided to perform this analysis in 19,5hpf morphant embryos because this stage enabled us to test both early hematopoietic precursors and differentiation markers. For the normalization of gene expression, we used two housekeeping genes (*ef-1α* and *rpl13a*) to better buffer up possible variation due to experimental procedures.

Of the nine genes analyzed, two of them have significantly different expression levels: *lmo2* expression was increased whereas *l-plastin* expression was decreased (Figure 3.3, E). While only these two genes were altered, there appears to be a bias for the up-regulation of all hematopoietic progenitor markers tested (Figure 3.3, E). Since EMPs derive from *lmo2*-expressing hematopoietic progenitors, the presence of more EMPs at 24hpf (observed in WISH analysis) can be a consequence of a larger population of hematopoietic progenitors at 19,5hpf. The down-regulation of *l-plastin* at 19,5hpf could be due to existing less macrophages in *dia1r* morphants, which was indeed observed in our WISH experiments. *Gata1* expression was not significantly altered in our qPCR essay, suggesting that erythroid progenitors are not affected in *dia1r* morphants and the higher expression observed in WISH experiments could be due to more EMPs.

In conclusion, the qPCR results together with the WISH analysis of DIA1R morphants lead us to think that DIA1R may restrict the size of the hematopoietic progenitor cell population (of both HSPCs and EMPs), either by regulating their proliferation or by promoting their differentiation toward the myeloid lineage.

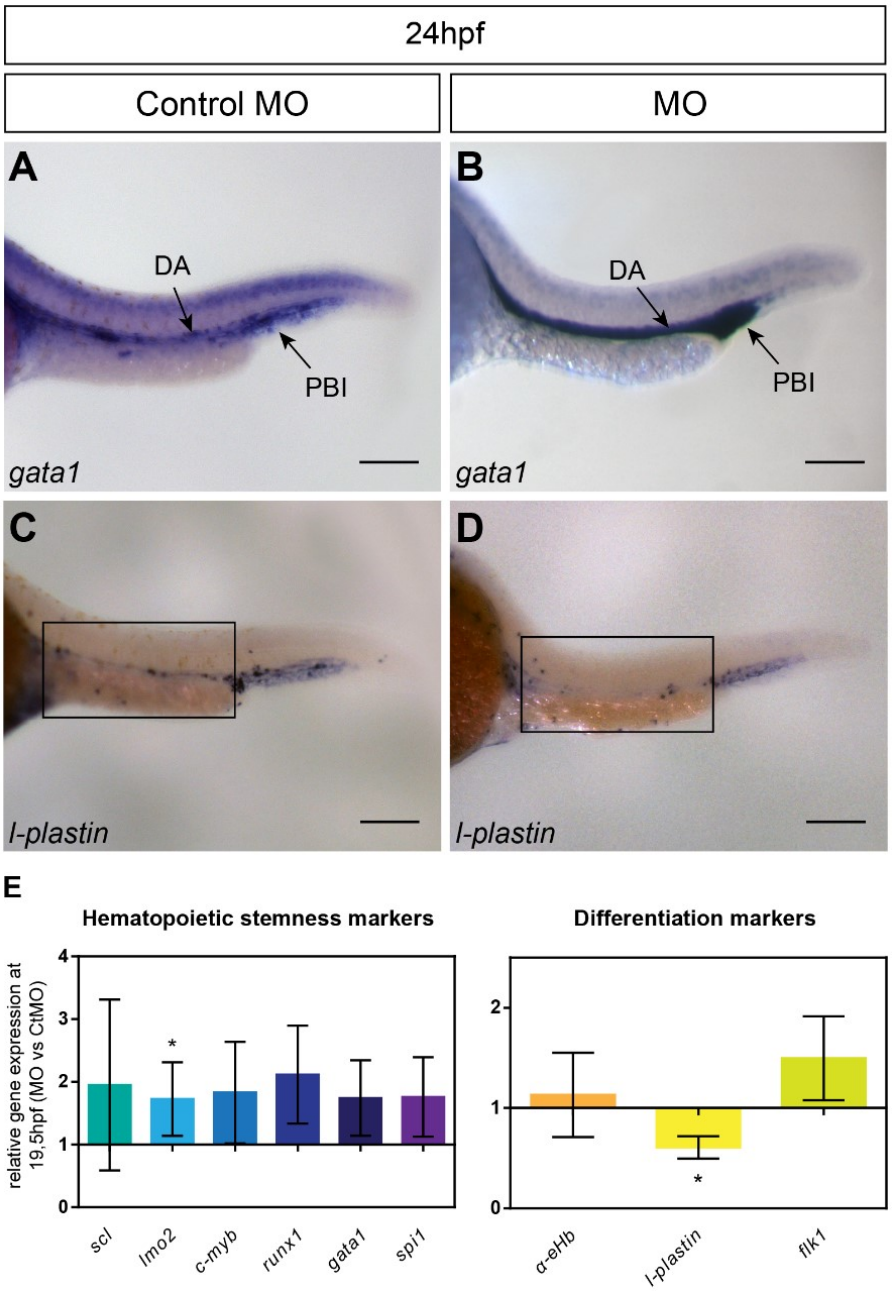


Figure 3.3 - Effect of dia1r knockdown via morpholino injection.
(A-B) Expression of *gata1* in control (A) and *dia1r* (B) morphants. (C-D) Expression of *l-plastin* in control (C) and *dia1r* (D) morphants. (E) Expression of hematopoietic stemness markers (*scl, lmo2, c-myb, gata1, runx1, spi1*) and differentiation markers (*α-eHb, l-plastin, flk1*) by RT-qPCR in 19,5hpf *dia1r* morphants. *lmo2* has a significantly higher expression and *l-plastin* has a significantly lower expression in *dia1r* morphant embryos. Although not statistically significant, there appears to be a bias toward the up-regulation of hematopoietic progenitor markers. Results are shown as fold change of *dia1r* morphant values relative to control morphant values (MO vs CtMO), both normalized to endogenous housekeeping gene expression. Scale bar = 140 µm.

3.3.2 DIA1R mutants by CRISPR/Cas9 technology

Despite the ease of use of MO knockdown, MO-induced phenotypes can be different from those of the corresponding mutants due to MO off-target effects (Kok et al., 2015) or genetic compensation in mutant embryos (Rossi et al., 2015). Therefore, we designed a strategy to generate a *dia1r* null-allele mutant or knockout (KO) line using the CRISPR/Cas9 system (in collaboration with the Fish Platform of the Champalimaud Foundation). To maximize the chances of creating a KO for *dia1r*, we chose a guide

RNA (gRNA) that targets the start codon of *dia1r* (Figure 3.4 A; gRNA selected with CRISPRscan; see methods). The gRNA and Cas9 protein were injected into one-cell stage wild-type embryos, followed by PCR-based genotyping of a sample of injected embryos (F0 generation) using primers that span the targeted genomic region (Supplementary table 6; Figure 3.4A). If the mutation is successful, F0 embryos will have the *dia1r* locus mutated with insertions or deletions (indels), caused by non-homologous end joining repair mechanism. Indeed, in the genotyping PCR of 12 samples (two injected embryos in each pool) we observed smears in five of them (Figure 3.4 B), which are evidence of the amplification of DNA fragments with different sizes. There was no amplification in the other samples, which could be due to problems in the DNA extraction from these samples. Overall, the PCR shows that various indels are present in the *dia1r* locus in these samples, thus indicating that the *dia1r* gene was successfully mutated.

Although most F0 embryos will be mosaic for *dia1r* mutations (Jao et al., 2013), these CRISPR knockout embryos (crispants) may already have some phenotypic change that can be evaluated by WISH. As these are mosaic embryos, a huge change in the expression of genes regulated by DIA1R it is not expected. Even so, a detectable mis-regulation of gene expression in crispants will suggest that the mutation strategy was successful, and F0 embryos may have a different phenotype from wild-type zebrafish. Therefore, in order to verify the potential phenotype caused by DIA1R KO, we analyzed the expression of hematopoietic markers in crispant embryos (19,5hpf).

Since we had a small number of fixed embryos, we only tested one marker for hematopoietic progenitors (*c-myb*), one for erythromyeloid progenitors (*gata1*) and one for macrophages (*l-plastin*). In addition to these markers, we also analyzed the expression of *dia1r*, which appeared to be slightly lower in crispant embryos in comparison with wild-type embryos, as expected in a mosaic mutant (Figure 3.4 C). Our observations also showed that *c-myb* expression was slightly up-regulated in the ICM of crispant embryos. Together with the observation that *lmo2* is up-regulated in *dia1r* morphants, our results suggest that the hematopoietic progenitor population may be increased when DIA1R levels or activity are lower. Unlike in morphant embryos, *gata1* and *l-plastin* expression patterns were not affected in crispant embryos (Figure 3.4 C). This may be due to the existence of low mosaicism, absence of mutations or in-frame mutations in those *dia1r* crispant embryos that were tested for the expression of *gata1* and *l-plastin*.

Altogether, the morphant and crispant analyses indicate that DIA1R may be restricting the size of hematopoietic progenitor population. We hypothesize that DIA1R may act either by controlling the proliferation of hematopoietic progenitor cells and/or by inducing their differentiation. In addition, our results suggest that DIA1R may indeed promote the differentiation of hematopoietic progenitor cells into primitive macrophages.

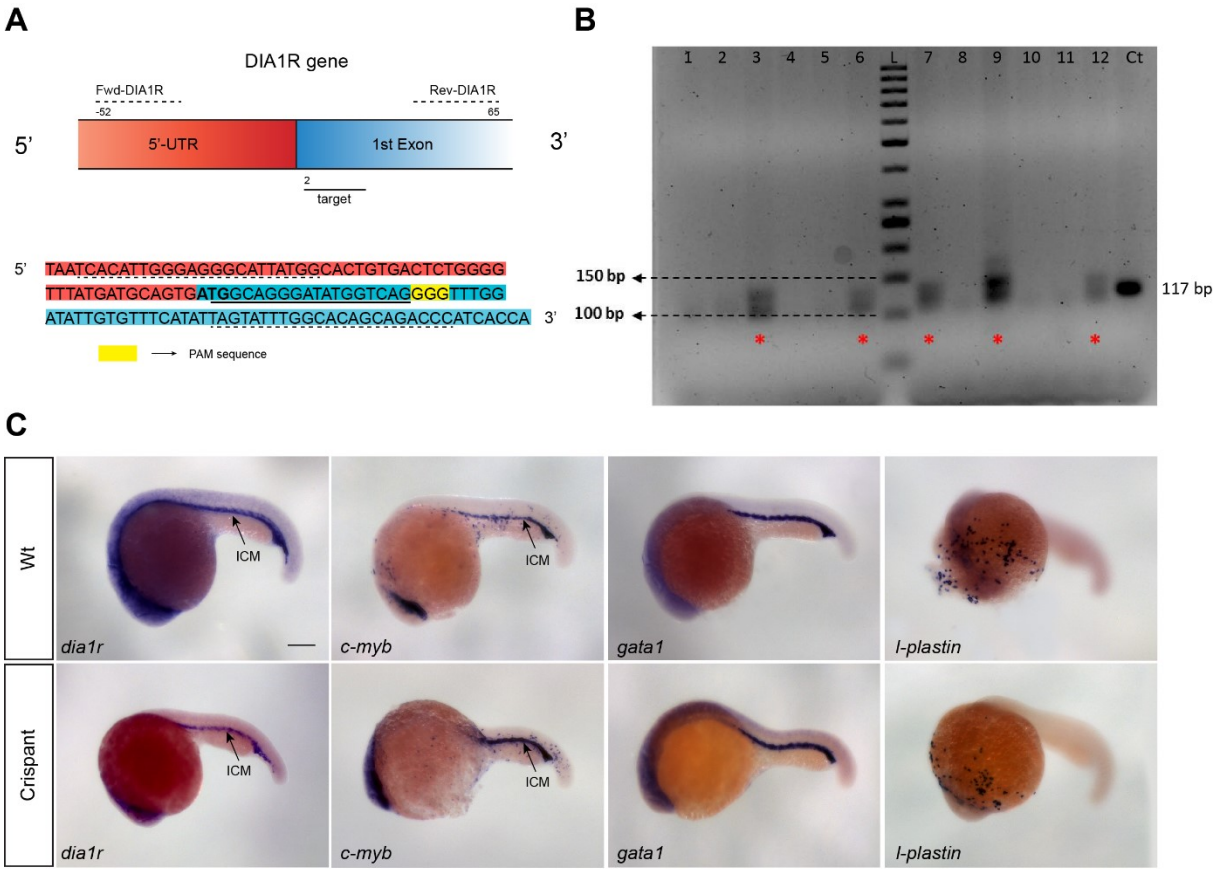


Figure 3.4 - CRISPR strategy for DIA1R KO zebrafish and crispant analysis.
(A) Schematic representation of the KO strategy using the CRISPR/Cas9 technology. The dashed lines represent the primers used for sequencing and genotyping, whereas the straight line represents the gRNA target sequence. PAM sequence is depicted in yellow. (B) Electrophoresis gel of the genotyping PCR of 24 hpf embryos injected with gRNA and Cas9 protein (two embryos per sample/lane). Five samples (3, 6, 7, 9 and 12) have smears, indicating the occurrence of indels in the locus of interest. (C) WISH of 19,5 hpf DIA1R crispant embryos for *dia1r*, *c-myb, gata1* and *l-plastin*. In comparison with Wt, the ICM of crispants embryos appears to have fewer cells expressing *dia1r* and more cells expressing *c-myb* (arrows). *gata1* and *l-plastin* expression is not affected. Ct – control (Wt embryos); Wt – wild-type; L – 1kb ladder. Scale bar = 140 µm.

4 Discussion

4.1 *dia1r* is expressed in hematopoietic precursors cells

We found out that *dia1r* is expressed in regions known to have hematopoietic potential in the zebrafish embryo. Namely, *dia1r* is expressed in the ALM and PLM at 12hpf and 16hpf, in the ICM at 19,5hpf, and in the DA and the CHT at 30hpf. We did not detect *dia1r* expression at 48hpf. However, we did not perform WISH in embryos at stages between 30hpf and 48hpf. During this period, HSPCs are still budding out of the DA and migrating to the CHT (Bertrand et al., 2010a; Ciau-Uitz et al., 2014). Therefore, it is likely that *dia1r* may continue to be expressed during this time window in the same regions detected at 30hpf.

In addition to hematopoietic progenitor cells, there are also angioblasts in the PLM and cardiac progenitors in the ALM (Gore et al., 2012; Hogan and Schulte-Merker, 2017; Peterkin et al., 2009). Also in the DA and CHT, both endothelial and hematopoietic cells are present, as well as other cell types (Murayama et al., 2006). In order to be certain of which cell populations express *dia1r*, we can perform co-localization experiments. One of these is two color whole-mount RNA *in situ* hybridization, in which we use two probes (one for *dia1r* and another for a marker gene of hematopoietic progenitor cells) labelled with different antigens (digoxygenin or fluorescein) and detected with two different subtracts (NBT/BCIP and Fast Red) (Bennett et al., 2001). Another method is to perform FISH for *dia1r* and immunohistochemistry for GFP in a transgenic line with labelled hematopoietic cells (*e.g.*, CD41-GFP) or differentiated cells from the blood lineage (*e.g.*, MPEG-GFP). Previous work in the lab determined that the available antibodies against human DIA1R do not work in zebrafish. In order to perform immunohistochemistry for zebrafish DIA1R, we can either create a specific antibody or generate a transgenic line with DIA1R fused with a Myc or a FLAG tag. Using the later approach, we can detect the tag by immunohistochemistry and use it as a proxy for the localization of DIA1R.

4.2 DIA1R restricts the size of hematopoietic progenitor cells

To determine the function of *dia1r* in blood development, we knocked it down via morpholino injection, and evaluated *dia1r* morphant phenotypes by WISH and qPCR of hematopoietic marker genes. WISH analysis of 24hpf *dia1r* morphants in comparison with controls showed that *gata1*, a EMP marker, had higher expression, and *l-plastin,* a macrophage marker, was expressed in fewer cells. Moreover, qPCR analysis of 19,5hpf *dia1r* morphants unveiled an up-regulation of *lmo2* expression*,* a hematopoietic progenitor cell marker, and down-regulation of *l-plastin* expression. These results suggest that DIA1R constrains the size of hematopoietic progenitor cell population, either by controlling their proliferation or inducing their differentiation. In addition to morphant experiments, we also analyzed the phenotype of a few crispant embryos at 19,5hpf. The WISH analysis of these embryos showed that *dia1r* expression was slightly reduced in crispants, indicating that our CRISPR strategy may be successful in creating a mutant line for *dia1r.* Of the three other probes used (*gata1, c-myb* and *l-plastin*), only *c-myb* expression had a slightly higher expression in crispant embryos, while *l-plastin* and *gata1* were not altered. These observations did not corroborate the alterations in primitive macrophages (*l-plastin* expression) and EMPs (*gata1* expression) observed in morphants. However, since crispants are F0 embryos, we do not know what kind of mutation(s) occurred nor how many cells (if any) were mutated in these individuals. Therefore, the existence of *gata1* and *l-plastin* phenotypes should to be further investigated in *dia1r* mutant embryos.

Although our results point to a role for DIA1R in hematopoietic development, additional experiments are required to confirm this hypothesis. First off, we need a bigger sample number of morphant embryos for the WISH and qPCR analyses, and the expression of the same genes should be evaluated by both methods. The qPCR data showed a differential expression for *lmo2*, but since we did not performed WISH for *lmo2* in *dia1r* morphants, we cannot say with certainty that *lmo2* up-regulation is a result of the presence of more hematopoietic progenitor cells. Moreover, the qPCR was performed using 19,5hpf embryos, and we did not manage to analyze embryos at this stage by WISH. Therefore, it is crucial to obtain WISH and qPCR data on *dia1r* morphants at the same developmental stages in order to show which cell types are being altered (WISH) and which genes are being altered, with semi-quantitative data (qPCR). We also need to increase the sample size for qPCR analysis in order to increase the statistical power of this analysis, and to optimize the protocol procedures for minimizing sample variation.

Secondly, it is essential to verify the phenotype observed in morphants embryos. Morpholinos (MO) are prone to off-targets effects, which may not be revealed by using 5bp-mismatch control morpholinos. Moreover, morpholino phenotypes are known to be more severe than mutation phenotypes due to genetic compensation of a knocked-out gene, hypomorphic nature of the mutant allele or phenotypic rescue of zygotic mutants by maternal mRNAs (Rossi et al., 2015; Schulte-Merker and Stainier, 2014; Stainier et al., 2017) Various experiments can be performed to validate a MO-phenotype. For example, one could use different morpholinos (ATG and splice blocking) and perform rescue experiments by co-injecting MO and *dia1r* mRNA. Nonetheless, a MO-phenotype should always be validated by comparing it to a mutant. In addition, the best way to verify MO specificity is to inject it and assess its effect in a mutant of that same gene (Stainier et al., 2017).

As mentioned above, our *dia1r* morphant results showed a reduction in the number of macrophages and indicated the presence of more EMPs (suggested by the higher expression of *gata1* observed by WISH and up-regulation of *lmo2* in the qPCR essay*;* Figure morphants). At 24hpf, *gata1* is also expressed in erythroid precursor cells that give rise to primitive erythrocytes. Therefore, the up-regulation of *gata1* in morphant embryos could also indicate that DIA1R plays a role in cross-antagonism between *gata1* and *spi1* that determines the erythroid versus myeloid fate (Monteiro et al., 2011). If DIA1R acts on this decision, and not on the differentiation decision in hematopoietic precursor cells, then we would expect not only a decrease in *l-plastin* (like we observed), but also a decrease in *spi1* expression. However, our qPCR data tends towards an up-regulation of *spi1* when *dia1r* is knocked down, indicating that DIA1R might not act on this cross-antagonism.

Finally, while we propose that DIA1R may promote the differentiation of hematopoietic progenitor cells (both HSPCs and EMPs), our observations focused only on macrophage differentiation. To confirm our hypothesis, we will need to investigate if other cell types of the blood lineage are affected (erythrocytes, lymphocytes and thrombocytes). We can verify DIA1R role in the differentiation of these cell types either qualitatively by WISH of *dia1r* morphant embryos or semi-quantitatively by qPCR for differentiation marker genes of all hematopoietic cell types. In addition, we could inject the morpholino in transgenic reporter lines for the different cell types and quantify the number of fluorescent cells by flow cytometry. These transgenic lines are summarized in a recent review by Gore et al (Gore et al., 2018).

4.3 Future approaches to investigate the regulation of *dia1r* expression

There are two major ways of identifying the factors that regulate the expression of a particular gene. One is through targeted approaches, in which the potential regulating factors are known, and the other is through approaches that identify new factors that might control gene expression.

4.3.1 Targeted approaches

Decades of research in hematopoiesis and its genetic control has yielded much information about the genes that control the development of hematopoietic cells. By knowing which transcription factors (and the signaling pathways that control their activity) specify the multiple hematopoietic lineages that arise throughout development, we can accurately speculate which genes may regulate *dia1r* expression. A relative quick way to identify them would be to analyze *dia1r* expression by WISH in mutants or morphants of those putative regulatory factors.

In 2016, the gene responsible for the *cloche* mutant was finally identified (Reischauer et al., 2016). It is named *npas4l* and encodes a basic helix-loop-helix PAS transcription factor. *Npas4l* expression precedes that of *scl* and *etv2,* the earliest hemangioblast genes, and it acts as a master regulator for the hemato-vascular development. It stands to reason that Npas4l might also regulate *dia1r* expression at early stages, but probably indirectly through scl and etv2, whose expression precedes that of *dia1r*. Scl and Lmo2 are both necessary for the formation of primitive hematopoietic cells (Patterson et al., 2007), whereas Etv2 is critical for vasculogenesis (Sumanas and Lin, 2006). Since *dia1r* is expressed in hematopoietic progenitor cells and *dia1r* knockdown seems to affect only hematopoietic populations, it is more likely that Scl and Lmo2 regulate its expression than Etv2.

At later stages, *dia1r* expression is detected in the DA and CHT, suggesting that it could be expressed in HSPCs and EMPs. DIA1R presence in these cells may depend on the transcription factors that characterize or specify these populations. HSPCs are specified in the hemogenic endothelium by Runx1. *Runx1* is induced by Notch1a/b signaling in these cells (Kim et al., 2014), which distinguishes HSPCs from EMPs (Bertrand et al., 2010b). Downstream of Runx1 is *c-Myb*, which is critical for the survival of HSPCs (Soza-ried et al., 2010). Among these three potential regulators, *dia1r* expression in HSPCs is most likely to be controlled by Notch1a/b or Runx1 than by c-Myb, given that *c-myb* expression is altered in *dia1r* crispants (Figure 3.4).

Concerning EMPs, this hematopoietic cell population is independent from Notch and can be characterized by the co-expression *gata1* and *lmo2* (Bertrand et al., 2007). If DIA1R expression is found to be controlled by Lmo2 in primitive hematopoiesis, then it might also be controlled by this factor in EMPs.

Human *DIA1R* (or *cXorf36*) transcripts were shown to be reduced in a microarray analysis of ERG small interfering RNA (siRNA)-treated HUVECs (Yuan et al., 2009; NCBI GEO profiles), thus indicating that ERG may activate *DIA1R* expression in hemogenic endothelial cells. ERG is an endothelial transcription factor essential for definitive hematopoiesis in mice (Loughran et al., 2008). In zebrafish, both *erg* and *dia1r* transcripts are first detected at 12hpf. In this model organism, *erg* expression is *scl* dependent, but it does not seem to have a role in blood development (Ellett et al., 2009; Liu and Patient, 2008). While morpholino knockdown of *erg* in zebrafish leads to hemorrhage, the same effect is not observed in *erg* mutant mice. Given this contradictory function in zebrafish and mice, *erg* might not be a regulator of *dia1r* expression in zebrafish. Nonetheless, to verify if *dia1r* acts downstream of Erg in zebrafish, we could analyze *dia1r* expression in ERG morphants and mutants.

4.3.2 Novel approach to identify new potential regulators

The approach described hereafter is really powerful, in the sense that it enables us to identify new transcriptional regulators of our locus of interest. However, it is expensive, complex and time-consuming, and is best employed when there is little to no data about the potential regulators of expression of our gene of interest.

Chromatin affinity purification with mass spectrometry (ChAP-MS) is a technique that enables the identification of proteins and histone posttranslational modifications (hPTMs) at a single genomic locus. This tool allows the isolation and enrichment of a target genomic locus through affinity-based purification, and the investigation of hPTMs and associated proteins through mass spectrometry (Byrum et al., 2012). Byrum *et al.* talk about the difficulties in creating the DNA affinity handle with homologous recombination in mammalian cell lines (and likely also in other vertebrate cell lines). With current technologies, like CRISPR, it becomes plausible to have a direct adaptation of this technique.

For zebrafish *dia1r* specifically, this technique is a bit farfetched. First, we need to identify the cis-regulatory region of *zDIA1R*. Then, if that region is bigger than 1000bp, we might need to perform multiple ChAP-MS. We would need to culture different zebrafish cell types in which we knew that *dia1r* was repressed or activated. And finally, we would need to have a control situation similar to the one described by Byrum *et al.* in which we would culture zebrafish cell lines in isotopically heavy lysine. Nonetheless, in the near future this technique might be properly develop for vertebrate systems (http://grantome.com/grant/NIH/R01-GM106024-02) and so, more feasible to be used for the identification of novel regulators of *dia1r* expression.

4.4 DIA1R as a valuable marker for the maintenance of the HSPC *in vitro*

It is difficult to generate HSPCs *in vitro*. This emphasizes the importance of the HSPC niche that exits in the adult hematopoietic sites, be it the bone marrow in mouse and human or the kidney marrow in zebrafish, in the maintenance of HSPCs *in vivo.* Although there have been successful attempts to maintain HSPCs in culture with long-term mouse bone marrow stromal cultures (Dexter and Testa, 1976; Dexter et al., 1977) and zebrafish kidney stromal cell lines (Stachura et al., 2009), there is still a need to improve their culture conditions and make them more straightforward.

It is also important to know which genes regulate the stemness of HSPCs in order to either maintain their stem cell state or induce their differentiation into specific cell types of interest (*e.g.*, macrophages). Since *dia1r* is expressed in hematopoietic precursor cells, it could be used as a marker for the isolation of these precursor cells in zebrafish. Moreover, if DIA1R has the function we propose in this report, it could be crucial for controlling HSPC differentiation. Namely, suppressing this gene can be useful for the maintenance of HSPC *in vitro*, whereas overexpressing it could be used to induce their differentiation into the myeloid lineage. In conclusion, DIA1R is a viable candidate modulator of hematopoietic stemness and myeloid differentiation.

5 Bibliography

Aziz, A., Harrop, S. P. and Bishop, N. E. (2011a). DIA1R is an X-linked gene related to Deleted In Autism-1. *PLoS One* **6**,.

Aziz, A., Harrop, S. P. and Bishop, N. E. (2011b). Characterization of the deleted in autism 1 protein family: Implications for studying cognitive disorders. *PLoS One* **6**,.

Barbieri, E., Deflorian, G., Pezzimenti, F., Valli, D., Saia, M., Meani, N., Gruszka, A. M. and Alcalay, M. (2016). Nucleophosmin leukemogenic mutant activates Wnt signaling during zebrafish development. *Oncotarget* **7**, 55302–55312.

Belele, C. L., English, M. A., Chahal, J., Burnetti, A., Finckbeiner, S. M., Gibney, G., Kirby, M., Sood, R. and Liu, P. P. (2009). Differential requirement for Gata1 DNA binding and transactivation between primitive and definitive stages of hematopoiesis in zebrafish. *Blood* **114**, 5162–5172.

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

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

Bertrand, J. Y., Cisson, J. L., Stachura, D. L. and Traver, D. (2010b). Notch signaling distinguishes 2 waves of definitive hematopoiesis in the zebrafish embryo. *Blood* **115**, 2777–2783.

Bhasin, M., Yuan, L., Keskin, D. B., Otu, H. H., Libermann, T. A. and Oettgen, P. (2010). Bioinformatic identification and characterization of human endothelial cell-restricted genes. *BMC Genomics* **11**, 342.

Bresciani, E., Carrington, B., Wincovitch, S., Jones, M., Gore, A. V, Brant, M., Sood, R., Liu, P. P., Bresciani, E., Carrington, B., et al. (2014). CBFβ and RUNX1 are required at 2 different steps during the development of hematopoietic stem cells in zebrafish. *Blood* **124**, 70–78.

Burns, C. E., Traver, D., Mayhall, E., Shepard, J. L. and Zon, L. I. (2005). Hematopoietic stem cell fate is established by the Notch – Runx pathway. *Genes Dev.* **19**, 2331–2342.

Butko, E., Distel, M., Pouget, C., Weijts, B., Kobayashi, I., Ng, K., Mosimann, C., Poulain, 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.

Byrum, S. D., Raman, A., Taverna, S. D. and Tackett, A. J. (2012). ChAP-MS: A method for identification of proteins and histone posttranslational modifications at a single genomic locus. *Cell Rep.* **2**, 198–205.

Cannon, J. E., Place, E. S., Eve, A. M. J., Bradshaw, C. R., Sesay, A., Morrell, N. W. and Smith, J. C. (2013a). Global analysis of the haematopoietic and endothelial transcriptome during zebrafish development. *Mech. Dev.* **130**, 122–131.

Cannon, J. E., Place, E. S., Eve, A. M. J., Bradshaw, C. R., Sesay, A., Morrell, N. W. and Smith, J. C. (2013b). Global analysis of the haematopoietic and endothelial transcriptome during zebrafish development. *Mech. Dev.* **130**, 122–131.

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

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C. and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. *Development* **125**, 725–32.

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

Detrich, H. W., Kieran, M. W., Chan, F. Y., Barone, L. M., Yee, K., Rundstadler, J. A., Pratt, S., Ransom, D. and Zon, L. I. (1995). Intraembryonic hematopoietic cell migration during vertebrate development. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 10713–7.

Dexter, T. M. and Testa, N. G. (1976). Differentiation and proliferation of hemopoietic cells in culture. *Methods Cell Biol.* **14**, 387–405.

Dexter, T. M., Allen, T. D. and Lajtha, L. G. (1977). Conditions controlling the proliferation of haemopoietic stem cells in vitro. *J. Cell. Physiol.* **91**, 335–344.

Dooley, K. A., Davidson, A. J. and Zon, L. I. (2005). Zebrafish scl functions independently in hematopoietic and endothelial development. *Dev. Biol.* **277**, 522–536.

Dudkiewicz, M., Lenart, A. and Pawłowski, K. (2013). A Novel Predicted Calcium-Regulated Kinase Family Implicated in Neurological Disorders. *PLoS One* **8**, 1–10.

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

Ellett, F., Kile, B. T. and Lieschke, G. J. (2009). The role of the ETS factor erg in zebrafish vasculogenesis. *Mech. Dev.* **126**, 220–229.

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C. and Schier, A. F. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**, 5–12.

Gay, L., Miller, M. R., Ventura, P. B., Devasthali, V., Vue, Z., Thompson, H. L., Temple, S., Zong, H., Cleary, M. D., Stankunas, K., et al. (2013). Mouse TU tagging: A chemical/genetic intersectional method for purifying cell type-specific nascent RNA. *Genes Dev.* **27**, 98–115.

Gering, M., Rodaway, A. R., Göttgens, B., Patient, R. K. and Green, A. R. (1998). The SCL gene specifies haemangioblast development from early mesoderm. *EMBO J.* **17**, 4029–45.

Ginhoux, F. and Guilliams, M. (2016). Tissue-Resident Macrophage Ontogeny and Homeostasis. *Immunity* **44**, 439–449.

Gomez Perdiguero, E., 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.

Gordon-Keylock, S. and Medvinsky, A. (2011). Endothelio-hematopoietic relationship: getting closer to the beginnings. *BMC Biol.* **9**, 88.

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

Gore, A. V., Pillay, L. M., Venero Galanternik, M. and Weinstein, B. M. (2018). The zebrafish: A

fintastic model for hematopoietic development and disease. *Wiley Interdiscip. Rev. Dev. Biol.* e312.

Gritz, E. and Hirschi, K. K. (2016). Specification and function of hemogenic endothelium during embryogenesis. *Cell. Mol. Life Sci.* **73**, 1547–1567.

Hamburger, V. and Hamilton, H. L. (1951). A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49–92.

Herbomel, P., Thisse, B. and Thisse, C. (1999). Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development* **126**, 3735–45.

Herbomel, P., Thisse, B. and Thisse, C. (2001). Zebrafish Early Macrophages Colonize Cephalic Mesenchyme and Developing Brain, Retina, and Epidermis through a M-CSF Receptor-Dependent Invasive Process. *Dev. Biol.* **238**, 274–288.

Hogan, B. M. and Schulte-Merker, S. (2017). How to Plumb a Pisces: Understanding Vascular Development and Disease Using Zebrafish Embryos. *Dev. Cell* **42**, 567–583.

Hu, B., Zhang, W., Feng, X., Ji, W., Xie, X. and Xiao, W. (2014). Zebrafish eaf1 suppresses foxo3b expression to modulate transcriptional activity of gata1 and spi1 in primitive hematopoiesis. *Dev. Biol.* **388**, 81–93.

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.* **110**, 13904–13909.

Jin, S.-W. (2005). Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**, 5199–5209.

Kim, A. D., Melick, C. H., Clements, W. K., Stachura, D. L., Distel, M., Panáková, D., Macrae, C., Mork, L. A., Crump, J. G., Traver, D., et al. (2014). Discrete Notch signaling requirements in the specification of hematopoietic stem cells. **33**, 2363–2374.

Kingsley, P. D., Malik, J., Emerson, R. L., Bushnell, T. P., McGrath, K. E., Bloedorn, L. A., Bulger, M. and Palis, J. (2006). "Maturational" globin switching in primary primitive erythroid cells. *Blood* **107**, 1665–1672.

Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., vanImpel, A., Kirchmaier, B. C., Peterson-Maduro, J., Kourkoulis, G., Male, I., et al. (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev. Cell* **32**, 97–108.

Kulkeaw, K. and Sugiyama, D. (2012). Zebrafish erythropoiesis and the utility of fish as models of anemia. 1–11.

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

Lieschke, G. J., Oates, A. C., Crowhurst, M. O., Ward, A. C. and Layton, J. E. (2001). Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. *Blood* **98**, 3087–3096.

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

Loughran, S. J., Kruse, E. A., Hacking, D. F., de Graaf, C. A., Hyland, C. D., Willson, T. A., Henley, K. J., Ellis, S., Voss, A. K., Metcalf, D., et al. (2008). The transcription factor Erg is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells. *Nat. Immunol.* **9**, 810–819.

Ma, A. C. H., Chung, M. I. S., Liang, R. and Leung, A. Y. H. (2010). A DEAB-sensitive aldehyde

dehydrogenase regulates hematopoietic stem and progenitor cells development during primitive hematopoiesis in zebrafish embryos. *Leukemia* **24**, 2090–2099.

Marcelo, K. L., Goldie, L. C. and Hirschi, K. K. (2013). Regulation of endothelial cell differentiation and specification. *Circ. Res.* **112**, 1272–1287.

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. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Rep.* **11**, 1892–1904.

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

Monteiro, R., Pinheiro, P., Joseph, N., Peterkin, T., Koth, J., Repapi, E., Bonkhofer, F., Kirmizitas, A. and Patient, R. (2016). Transforming Growth Factor β Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells. *Dev. Cell* **38**, 358–370.

Moreno-Mateos, M. A., Vejnar, C. E., Beaudoin, J. D., Fernandez, J. P., Mis, E. K., Khokha, M. K. and Giraldez, A. J. (2015). CRISPRscan: Designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* **12**, 982–988.

Moro, E., Vettori, A., Porazzi, P., Schiavone, M., Rampazzo, E., Casari, A., Ek, O., Facchinello, N., Astone, M., Zancan, I., et al. (2013). Generation and application of signaling pathway reporter lines in zebrafish. *Mol. Genet. Genomics* **288**, 231–242.

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.

Murray, P. D. F. (1932). The Development in vitro of the Blood of the Early Chick Embryo. *Proc. R. Soc. B Biol. Sci.* **111**, 497–521.

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

Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. *Front. Physiol.* **5 JAN**, 1–9.

Parichy, D. M., Ransom, D. G., Paw, B., Zon, L. I. and Johnson, S. L. (2000). An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, Danio rerio. *Development* **127**, 3031–3044.

Patterson, L. J., Gering, M., Eckfeldt, C. E., Green, A. R., Verfaillie, C. M., Ekker, S. C. and Patient, R. (2007). The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish. *Blood* **109**, 2389–2398.

Peterkin, T., Gibson, A. and Patient, R. (2009). Common genetic control of haemangioblast and cardiac development in zebrafish. *Development* **136**, 1465–1474.

Reischauer, S., Stone, O. A., Villasenor, A., Chi, N., Jin, S. W., Martin, M., Lee, M. T., Fukuda, N., Marass, M., Witty, A., et al. (2016). Cloche is a bHLH-PAS transcription factor that drives haemato-vascular specification. *Nature* **535**, 294–298.

Robertson, A. L., Avagyan, S., Gansner, J. M. and Zon, L. I. (2016). Understanding the regulation of vertebrate hematopoiesis and blood disorders: big lessons from a small fish. *FEBS Lett.* 1–18.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M. and Stainier, D. Y. R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**, 230–233.

Sabin, F. R. (1920). Studies on the origin of blood vessels and of red corpuscles as seen in the living blastoderm of the chick during the second day of incubation. In *Contributions to Embryology*, .

Schefe, J. H., Lehmann, K. E., Buschmann, I. R., Unger, T. and Funke-Kaiser, H. (2006). Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's C T difference" formula. *J. Mol. Med.* **84**, 901–910.

Schulte-Merker, S. and Stainier, D. Y. R. (2014). Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. *Development* **141**, 3103–3104.

Scott, E. W., Simon, M. C., Anastasi, J. and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–7.

Serrado Marques, J., Teixeira, V., Jacinto, A. and Tavares, A. (2018). Identification of Novel Hemangioblast Genes in the Early Chick Embryo. *Cells* **7**, 9.

Shen, L. J., Chen, F. Y., Zhang, Y., Cao, L. F., Kuang, Y., Zhong, M., Wang, T. and Zhong, H. (2013). MYCN Transgenic Zebrafish Model with the Characterization of Acute Myeloid Leukemia and Altered Hematopoiesis. *PLoS One* **8**, 1–12.

Soza-ried, C., Hess, I., Netuschil, N., Schorpp, M. and Boehm, T. (2010). Essential role of c-myb in definitive hematopoiesis is evolutionarily conserved. *Pnas* **107**, 17304–17308.

Stachura, D. L., Reyes, J. R., Bartunek, P., Paw, B. H., Zon, L. I. and Traver, D. (2009). Zebrafish kidney stromal cell lines support multilineage hematopoiesis. *Blood* **114**, 279–289.

Stainier, D. Y. R., Raz, E., Lawson, N. D., Ekker, S. C., Burdine, R. D., Eisen, J. S., Ingham, P. W., Schulte-Merker, S., Yelon, D., Weinstein, B. M., et al. (2017). Guidelines for morpholino use in zebrafish. *PLoS Genet.* **13**, 6–10.

Sumanas, S. and Lin, S. (2006). Ets1-related protein is a key regulator of vasculogenesis in zebrafish. *PLoS Biol.* **4**, 0060–0069.

Swiers, G., Rode, C., Azzoni, E. and De Bruijn, M. F. T. R. (2013). A short history of hemogenic endothelium. *Blood Cells, Mol. Dis.* **51**, 206–212.

Tavares, A. T., Andrade, S., Silva, A. C. and Belo, J. A. (2007). Cerberus is a feedback inhibitor of Nodal asymmetric signaling in the chick embryo. *Development* **134**, 2051–2060.

Tavares, B., Jacinto, R., Sampaio, P., Pestana, S., Pinto, A., Vaz, A., Roxo-Rosa, M., Gardner, R., Lopes, T., Schilling, B., et al. (2017). Notch/Her12 signalling modulates, motile/immotile cilia ratio downstream of Foxj1a in zebrafish left-right organizer. *Elife* **6**, e25165.

Teixeira, V., Arede, N., Gardner, R., Rodríguez-León, J. and Tavares, A. T. (2011). Targeting the hemangioblast with a novel cell type-specific enhancer. *BMC Dev. Biol.* **11**, 76.

Thisse, C. and Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69.

Thompson, M. a., Ransom, D. G., Pratt, S. J., MacLennan, H., Kieran, M. W., Detrich, H. W., Vail, B., Huber, T. L., Paw, B., Brownlie, A. J., et al. (1998). The cloche and spadetail Genes Differentially Affect Hematopoiesis and Vasculogenesis. *Dev. Biol.* **197**, 248–269.

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.

Varshney, G. K., Pei, W., LaFave, M. C., Idol, J., Xu, L., Gallardo, V., Carrington, B., Bishop, K., Jones, M., Li, M., et al. (2015). High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res.* **25**, 1030–42.

Vejnar, C. E., Moreno-Mateos, M. A., Cifuentes, D., Bazzini, A. A. and Giraldez, A. J. (2016). Optimized CRISPR-Cas9 system for genome editing in zebrafish. *Cold Spring Harb. Protoc.* **2016**, 856–870.

Veldman, M. B. and Lin, S. (2012). Etsrp/Etv2 is directly regulated by Foxc1a/b in the zebrafish angioblast. *Circ. Res.* **110**, 220–229.

Warga, R. M., Wicklund, A., Webster, S. E. and Kane, D. A. (2016). Progressive loss of RacGAP1/ogre activity has sequential effects on cytokinesis and zebrafish development. *Dev. Biol.* **418**, 307–322.

Westerfield, M. (2007). *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio), 5th Edition*.

Xiong, J. W. (2008). Molecular and developmental biology of the hemangioblast. *Dev. Dyn.* **237**, 1218–1231.

Yajima, Y., Vestergaard, M. C. and Takagi, M. (2012). Damage in brain development by morpholino knockdown of zebrafish dax1. *J. Biosci. Bioeng.* **113**, 683–688.

Yuan, L., Nikolova-Krstevski, V., Zhan, Y., Kondo, M., Bhasin, M., Varghese, L., Yano, K., Carman, C. V., Aird, W. C. and Oettgen, P. (2009). Antiinflammatory effects of the ETS factor ERG in endothelial cells are mediated through transcriptional repression of the interleukin-8 gene. *Circ. Res.* **104**, 1049–1057.

Zou, J., Li, W. Q., Li, Q., Li, X. Q., Zhang, J. T., Liu, G. Q., Chen, J., Qiu, X. X., Tian, F. J., Wang, Z. Z., et al. (2011). Two functional MicroRNA-126s repress a novel target gene p21-activated kinase 1 to regulate vascular integrity in zebrafish. *Circ. Res.* **108**, 201–209.

6 Supplementary Information

6.1 Appendix A - Supplementary tables

Supplementary table 1 - zDIA1R probe amplification primers

Primers' name	Sequence
cannon Fw	5' - CTCTGGGGTTTATGATGCAGTGATGG - 3'
cannon Rv	5' - CTAGAACTTGTCACTGTAGAGACATTCCG - 3'
short Fw	5' - AGATCTCGCCTACCAGCTTCT - 3'
short Rv	5' - CTAGAACTTGTCACTGTAGAGACATTCCG - 3'

Supplementary table 2 – List of riboprobes used for whole-mount *in situ* hybridization

Probe	Vector	Antibiotic	Anti-sense Linearization	Anti-sense Polymerase	Reference
zDIA1R cannon	pGEMT-Easy	Amp	*Sac*I	SP6	(Cannon et al., 2013b)
zDIA1R short	pGEMT-Easy	Amp	*Spe*I	SP6	This work
cMyb	pBK-CMV	Kan	*EcoR*I	T7	(Thompson et al 1998)
gata1	pBluescript SK(-)	Amp	*Xba*I	T7	(Detrich et al., 1995)
l-plastin	pBSK+	Amp	*Not*I	T7	(Herbomel et al. 1999)
runx1	pBluescript SK(+)	Amp	*Hind*III	T7	Unpublished
mpx	pBK-CMV	Kan	*EcoR*I	T7	(Shen et al., 2013)
gata2b	pGEMT-Easy	Amp	*Sac*I	T7	Unpublished
lmo2	pBK-CMV	Kan	*EcoR*I	T7	(Thompson et al 1998)
csf1ra	pGEMT-Easy	Amp	*Nco*I	SP6	Unpublished
Scl	pZE62	Amp	*EcoR*I	T7	(Gering et al., 1998)

Supplementary table 3 – Morpholinos used in *dia1r* knockdown

	Type	Sequence
zDIA1R-ATG	Blocks translation	5' - CCCCTGACCATATCCCTGCCATCAC - 3'
zDIA1R-ATG control	5bp mismatch control morpholino	5' - CCCCTCAGCATATCGCTCCGATCAC - 3'

Supplementary table 4 –List of primers used in qPCR

Gene	Primers	Reference
α-ehb	Fw 5'- TGCTCTCTCCAGGATGTTGA - 3' Rv 5'- TCACAGTCTTGCCGTGTTTC - 3'	(Ma et al., 2010)
c-myb	Fw 5'- TTTCTACCGAATCGAACAGATG - 3' Rv 5'- CAATCACCCGTTGGTCTTCT - 3'	(Ma et al., 2010)
gata1	Fw 5'- AAGATGGGACAGGCCACTAC - 3' Rv 5'- TGCTGACAATCAGCCTCTTTT - 3'	(Ma et al., 2010)
spi1	Fw 5'- GGGCAGTTTTAACCAAAGATCA - 3' Rv 5'- CCCAAGAGTGATCGTTCTGAC - 3'	(Ma et al., 2010)
runx1	Fw 5'- CGTCTTCACAAACCCTCCTC - 3' Rv 5'- TACTGCTTCATCCGGCTTCT - 3'	(Ma et al., 2010)
scl	Fw 5'- GGAGATGCGGAACAGTATGG - 3' Rv 5'- GAAGGCACCGTTCACATTCT - 3'	(Veldman and Lin, 2012)
lmo2	Fw 5'- AAATGAGGAGCCGGTGGAT - 3' Rv 5'- GCTCGATGGCCTTCAGAAA - 3'	(Bertrand et al., 2008)
l-plastin	Fw 5'- ACTCACACAGGCGGGAAAAGG - 3' Rv 5'- GCCATCAGACACGCAAACACC - 3'	(Hu et al., 2014)
flk1	Fw 5'- CTGGTGGAGAGGCTAGGAGA - 3' Rv 5'- TGATCGGGATGTAGTGCTTTC - 3'	(Zou et al., 2011)
ef1α	Fw 5'-CCTGGGAGTGAAACAGCTG - 3' Rv 5'-GCCTCCAGCATGTTGTCAC - 3'	made by Ana Brandão
rpl-13a	Fw 5'- TGACAAGAGAAAGCGCATGGTT - 3' Rv 5'- GCCTGGTACTTCCAGCCAACTT - 3'	made by Raquel Jacinto

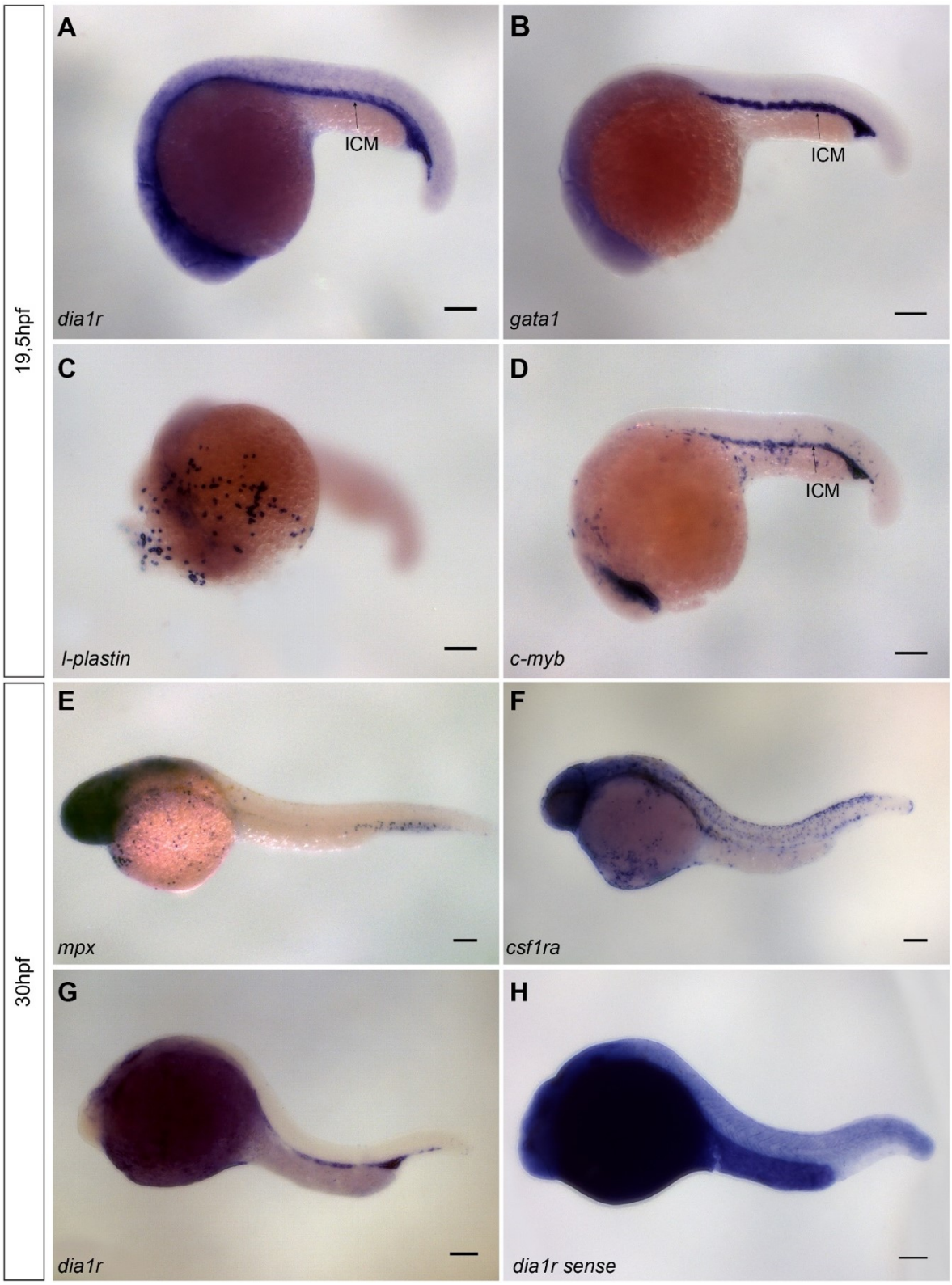
Supplementary table 5 – qPCR conditions

Program	Temperature	Duration	Fluorescence acquisition
Preincubation	95°C	600 seconds	None
3-step-amplification	95°C - Denaturation 60°C - Annealing 72°C - Extension	10 seconds 10 seconds 20 seconds	None None Single
Melting curve	95°C 65°C 97°C	10 seconds 60 seconds 1 second	None None Continuous
Cooling	37°C	30 seconds	None

Supplementary table 6 – CRISPR primers for genotyping and sequencing

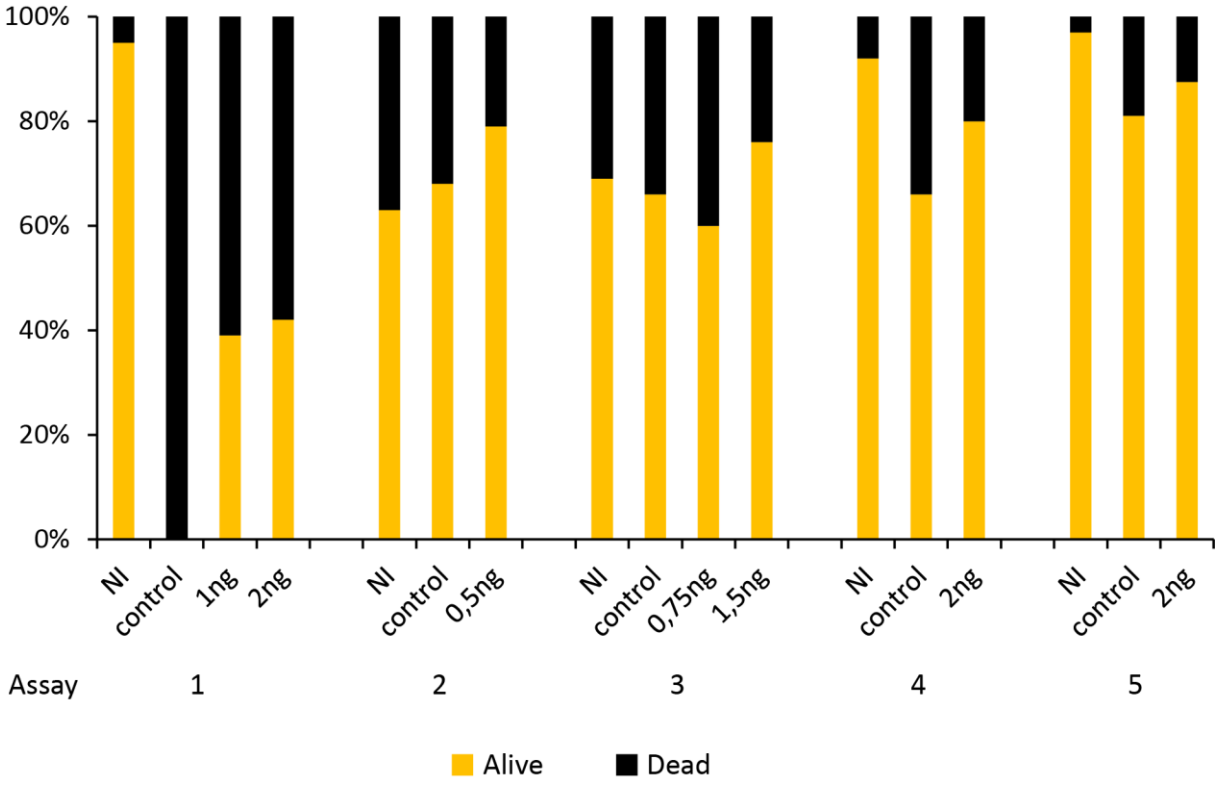
Primer name	Sequence
Fwd-DIA1R	5' - TCACATTGGGAGGGCATTATGG - 3'
Rev-DIA1R	5' - GGGTCTGCTGTGCCAAATACTA - 3'

6.2 Appendix B – Supplementary figures

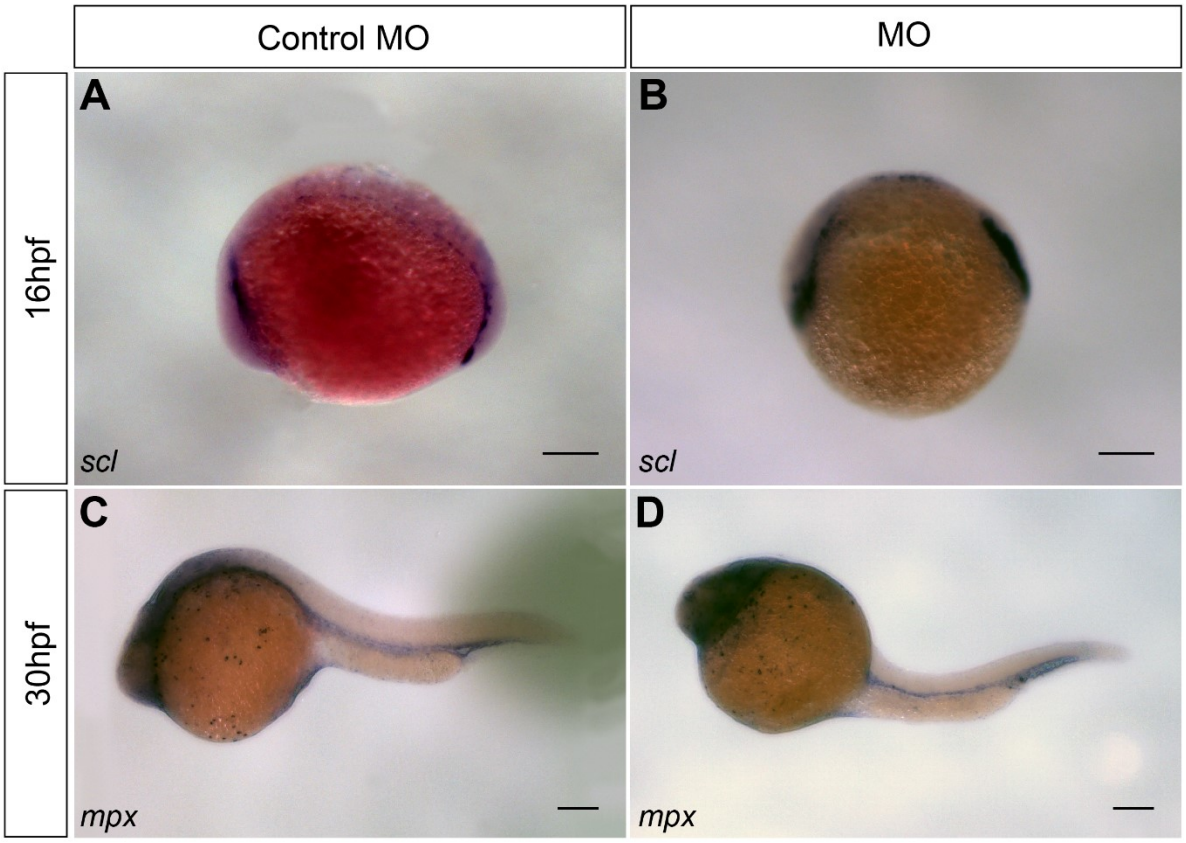


Supplementary figure 1 –Expression by WISH of *dia1r* in comparison with other hematopoietic markers and with *dia1r* control probe. (A-D) Expression of *dia1r* (A), *gata1* (B), *l-plastin* (C) and *c-myb* (D) in 19,5hpf embryos. *gata1* is expressed in the ICM and *l-plastin* is expressed in macrophages. At this stage *c-myb* labels both in the ICM and myeloid progenitor cells. (E-F) Expression of *mpx* (E), *csf1ra* (F), *dia1r* (G) and *dia1r* sense probe (H) in 30hpf embryos. *mpx* labels neutrophils and *csf1ra* is expressed in both macrophages and pigment cells. *dia1r* expression overlaps with *gata1* (B) and *c-myb* (D) in the

ICM (arrows), but it does not overlap with *l-plastin* (C), *csf1ra* (F) or *mpx* (E). These observations suggest that *dia1r* is expressed in HSPCs but not in more differentiated myeloid cells. The *dia1r* sense probe does not label any structures, thus indicating that the signal detected by the *dia1r* anti-sense probe (*zDIA1R* cannon) is specific to *dia1r* mRNA. ICM – intermediate cell mass. Scale bar = 100 µm.



Supplementary figure 2 – Optimization of MO concentration. Graphic representation of zebrafish mortality rate assessed at 24hpf. Assays are organized in a chronological manner. Embryos were either non-injected (NI) or injected with 1,4nL of *dia1r* 5-mismatch control MO (control; 2ng) or *dia1r* MO at different concentrations. These experiments revealed an optimal MO concentration of 2ng/1,4nL. *n* = 80-100 per condition in each assay.



Supplementary figure 3 – WISH of unaffected genes in *dia1r* morphants embryos. Embryos were injected with *dia1r* control morpholino (Control MO) or *dia1r* morpholino (MO). (A-B) Expression of *scl* in control (A) and *dia1r* (B) morphants. (C-D) Expression of *mpx* in control (C) and *dia1r* (D) morphants. The expression of both *scl* and *mpx* is similar in control and *dia1r* morphants. Scale bar = 140 µm.