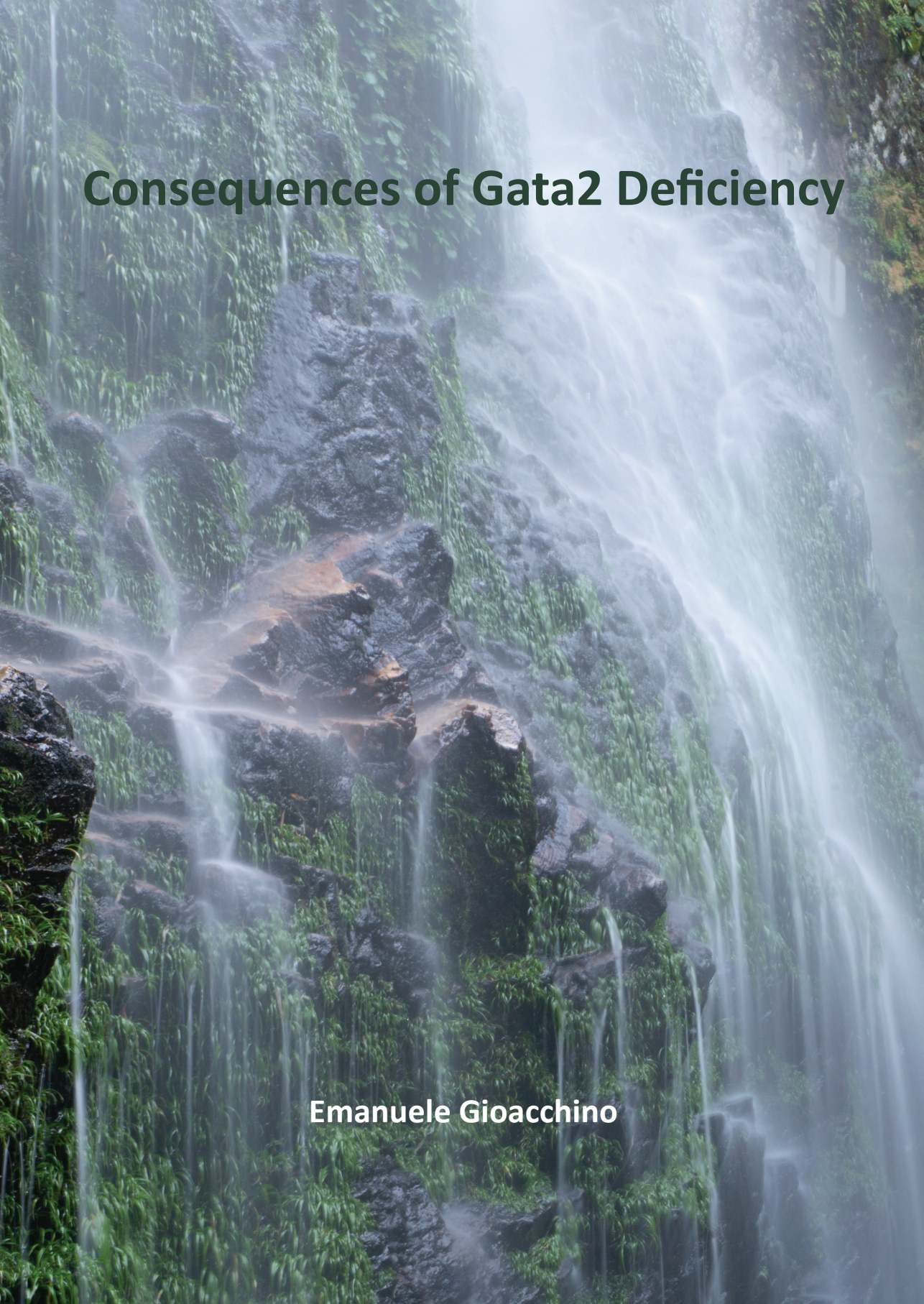
Consequences of Gata2 Deficiency



Emanuele Gioacchino

Consequences of Gata2 Deficiency

Emanuele Gioacchino

ISBN: 978-94-6416-772-6
Layout: Egied Simons
Cover: Josephine Bergmann.
Printing: Ridderprint

The work described in this thesis was performed at the Department of Hematology at the Erasmus Medical Center, Rotterdam, the Netherlands. This work has been financially supported by grants and fellowships from the Dutch Cancer Society “Koningin Wilhelmina Fonds” KWO.

Printing of this thesis was financially supported by the Erasmus University Rotterdam.

Consequences of Gata2 Deficiency

Gevolgen van Gata2 Tekort

Thesis

to obtain the degree of Doctor from the

Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Tuesday 14 September 2021 at 13.00hrs

by

Emanuele Gioacchino

born in Urbino (PU), Italy.

Erasmus University Rotterdam

DOCTORAL COMMITTEE

Promotors: Prof.dr. I.P. Touw

Other members: Prof.dr. H.R. Delwel
Prof.dr. J.N.J. Philipsen
Prof.dr. A. Bigas Salvans

Co-promoter: dr. E.M. De Pater

A nonna Adina, m capisc solo tu

CONTENTS

1

General introduction

1.1 Hematopoiesis and stem cells

Hematopoiesis is the process of blood cell production. Blood cells originate from hematopoietic stem cells (HSCs), which are characterized by their ability to self-renew and to differentiate into all blood cell types (Ogawa, 1993; Till, J.E. McCulloch, 1961). Adult HSCs reside almost exclusively in the bone marrow (BM), replicate with a low frequency, and are responsible for the production of all differentiated blood cells via intermediate cells called hematopoietic progenitor cells (HPCs) (Akashi et al., 2000; Kondo et al., 1997; Sun et al., 2014). Unlike the HSCs, HPCs are highly proliferative cells with no or limited self-renewal; they possess a more restricted lineage differentiation capacity and ensure a continuous supply of blood cells, both under homeostatic conditions and in episodes of hematological stress, such as blood cell loss or bacterial infections. The cell cycle of hematopoietic stem and progenitor cells (HSPCs) is controlled by the synergistic activity of extrinsic signals from the BM niche and hematopoietic cytokines and cell intrinsic pathways, mainly directed through transcription factors (TF) (Laurenti and Göttgens, 2018).

Mature blood cells are specialized in performing specific tasks in the organism like immune response, wound healing, oxygenation. Most blood cells have a limited life-span, lasting from a few hours (platelets) to several weeks (red blood cells). The two main blood cell categories are myeloid and lymphoid. The myeloid compartment comprises the cells of the innate immune response (neutrophils, eosinophils, and basophils), macrophages, thrombocytes, mast cells and erythrocytes. The lymphoid compartment is made up of cells responsible for the adaptive immune response, including B and T cells. In humans more than a million differentiated blood cells are produced every second. This continuous blood production requires a tight regulation. The proper control of proliferation and/or differentiation is a key aspect for a balanced and healthy blood system. This is highlighted when HSPCs struggle to differentiate into specific cell lineages or maintain a long-term self-renewal capacity, causing respectively cytopenia or leukemia (Fuchs and Chen, 2013; Seita and Weissman, 2010).

The understanding of hematopoiesis starts with the interpretation of steady state conditions. Although the mechanisms that determine how HSCs are instructed to differentiate are still not fully understood, there is consensus over a tree-like model of hematopoiesis (Akashi et al., 2000; Kondo et al., 1997). HSCs reside at the top of this model, progressively giving rise to all blood lineages, and restricting their lineage potential in a hierarchical manner. The steps of differentiation are represented by branches, where a cell restricts its potential and differentiates. This model is mostly based on immunophenotypical purification of stem and progenitor cells. Depending on the set of markers or the assay used to define progenitors, various branching patterns have been proposed leading to slightly different models of hematopoiesis. Recent advances in single-cell RNA sequencing and cellular barcoding led to a more continuous and dynamic model (Rodriguez-Fraticelli et al., 2018).

1.1.1 Identification of adult murine HSCs

The identification of HSCs has been a key aspect in hematology for various purposes, such as transplantation, re-programming technology and for a biological understanding of hematopoiesis and hematopoietic disease. Being such a small population, HSC identification was a challenging aspect in hematology. These cells can be defined both functionally and phenotypically as was pointed out by Purton and Scadden: ¨HSCs are functionally defined by their ability to sustain multi-lineage engraftment for an extended period of time upon serial transplantation into irradiated recipient mice and are immunophenotypically defined by the expression of a set markers¨ (Purton and Scadden, 2007). The transplantation of bone marrow HSC and MPPs subpopulations revealed functional differences in self-renewal, clonal lifespan and repopulation capacity in irradiated mice. Consequently, multipotent hematopoietic cells have been broadly subclassified in long-term HSC (LT-HSC), short-term HSC (ST-HSC) and multipotent progenitors (MPPs). LT-HSC can support hematopoiesis for more than 4 months in lethally irradiated recipients and maintain this competence at least in a secondary transplantation (Colvin et al., 2009; Lemischka et al., 1986; Lu et al., 2011; Morrison and Weissman, 1994; Nakauchi, Hiromitsu Sudo, Kazuhiro Ema, 2006; Oguro et al., 2013; Szilvassy et al., 1990; Yang et al., 2005).

The phenotypic definition of HSCs takes advantage of a specific combination of cell surface markers. Most methods for enriching HSCs include the exclusion of lineage positive cells (B cell, T cells, macrophages/monocytes, granulocytes and erythrocytes) and the inclusion of cells co-expressing Sca1 and c-Kit (Osawa et al., 1996; Spangrude et al., 1988; Zhao et al., 2000). The cells in this subgroup are called LSK ($Lin^{-}Sca1^{+}Kit^{+}$) and include all multipotent hematopoietic cells in the bone marrow with an HSC purity that was defined to be about 10% in steady state hematopoiesis for C57Bl/6 (Spangrude and Brooks; Spangrude et al., 1988; Uchida and Weissman, 1992). A more stringent purification method incorporates the inclusion of so-called SLAM (signaling lymphocyte activation molecule) family members, namely $CD48^{-}$ and $CD150^{+}$ (Kiel et al., 2005). The frequency of long-term reconstituting HSCs in Lin^{-} $Kit^{+}Sca^{+}CD150^{+}CD48^{-}$ (LSK SLAM) is between 20-40% and represents about 0.008% of C57BL bone marrow cells (Kent et al., 2009; Kiel et al., 2005; Yilmaz et al., 2006).

1.1.2 Characteristics and regulation of adult HSCs

The behaviour of HSCs is governed by extrinsic and intrinsic factors. For instance, during transplantation HSCs are challenged to reconstitute an irradiated recipient and consequently boost their proliferation. Contrarily, steady state HSCs are mainly quiescent, (Cabezas-Wallscheid et al., 2017; Cheshier et al., 1999; Foudi et al., 2009; Pietras et al., 2011; Wilson et al., 2008) with low mitochondrial activity (Ito et al., 2016; Vannini et al., 2016) and a preference for glycolytic metabolism over oxidative phosphorylation (Simsek et al., 2010; Takubo et al., 2013). In steady state vs transplantation setting, the proliferation difference in HSCs is due to the environment, highlighting the importance of extrinsic regulation.

Extrinsic regulation of adult steady state HSPCs is mediated by non-hematopoietic cells (like osteoblasts, endothelial cells, adipocytes, fibroblast and mesenchymal stromal cells) and differentiated hematopoietic cells like macrophages, osteoclasts (derived from monocytes and macrophages (Udagawa et al., 1990)), and lymphocytes (Hirata et al., 2018). All of these cells form a microenvironment for HSCs, also known as the HSC niche, within the bone marrow. This niche protects HSCs by limiting byproducts of aerobic metabolism, such as reactive oxygen species (ROS), supporting the prevention of DNA damage (Eliasson and Jönsson, 2010), while also maintaining and directing their self-renewal and differentiation state by releasing signals such as cytokines, chemokines and extracellular matrix molecules. Parallel to extrinsic regulation, HSC behaviour is driven intrinsically by transcription factors, signal transduction molecules, non-coding RNA and epigenetic modifiers. Transcription factors are the most researched intrinsic regulators due to their relevance with HSCs self-renewal, differentiation (for example C/EBPα for granulocyte differentiation (Zhang et al., 1997)) or apoptosis (for example P53 (Fridman and Lowe, 2003)).

1.1.3 Lineage priming and stem cell heterogeneity

The standard hematopoietic tree model assumes that HSCs are a homogeneous population with the same self-renewal and differentiation potential. However, most experiments to define HSCs were obtained from expression profiling of pooled cell subsets. Those bulk immunophenotypical analyses put together cells that are functionally often dissimilar, leading to a misinterpretation of the HSC compartment. Nowadays, phenotypic HSCs are not considered a homogeneous population anymore, yet mostly a heterogeneous lineage primed population. Already in 1997, the concept of multilineage gene expression in a single hematopoietic multipotent progenitor was recognized (Hu et al., 1997) and defined as a state called ¨promiscuous¨. This observation was based on single cell RT-PCR on a hematopoietic multipotential cell line reporting co-expression of granulocytic, monocytic and erythroblastic markers such as myeloperoxidase and beta-globin. The pattern of gene expression allowed to define a predictable functional outcome of these cells. These experiments supported the lineage priming model. Lately, a prediction model classified cells undergoing differentiation based on a gradual decrease in the number of genes that are expressed, confirming multilineage gene expression in HSCs (Gulati et al., 2020). In this model HSCs express multiple differentiation genes, representing their differentiation potential. During differentiation, one pathway gets upregulated causing the downregulation of the other differentiation pathways (Figure 1). Therefore, the expression level of lineage-restrictive transcription factors can be used to infer cell fate of HSPCs.

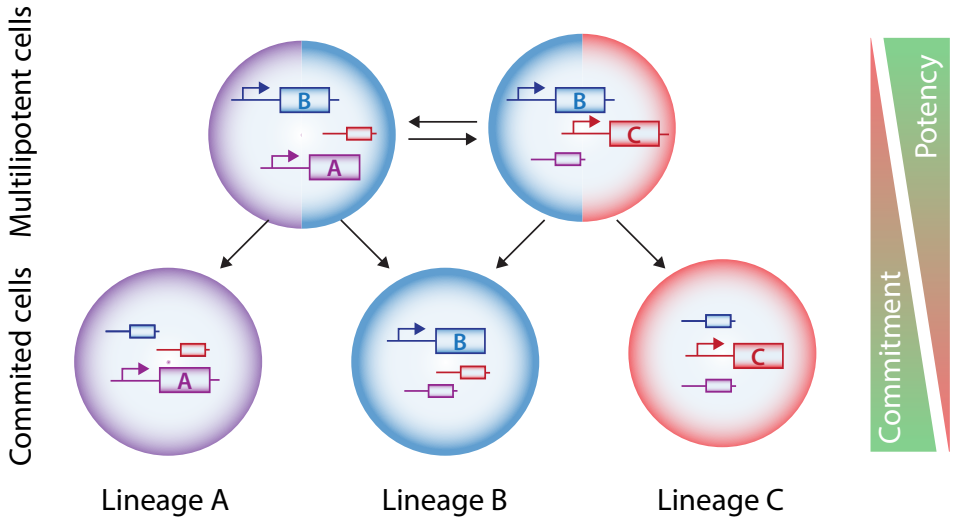


Figure 1. Multilineage priming allows HSCs to commit into different cell fates. Adapted from Nimmo et al (Nimmo et al., 2015). Multiple lineage affiliated genes are co-expressed in multipotent cells before lineage commitment. Commitment is associated with downregulation of genes associated with alternative lineages.

Even if *bona fide* HSCs and MPPs have the potential to differentiate into all hematopoietic cell types, subgroups have distinct differentiation behaviours (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015). For example, a thrombocyte-biased HSC population has been proven both transcriptionally and functionally via transplantation and *in vitro* assays (Grover et al., 2016; Rodriguez-Fraticelli et al., 2018; Sanjuan-Pla et al., 2013; Yamamoto et al., 2013). Megakaryocytes are considered to arise independently of other lineages but the reason for this atypical way of differentiation is still doubtful. Megakaryocytic priming and independence in differentiation also occurs in zebrafish (Macaulay et al., 2016). Since megakaryocytes and HSCs share common markers and niche (Cheng et al., 2020), it is still under debate whether the HSCs purification methods cause the inclusion of megakaryocytes. Besides HSCs, phenotypic MPPs are considered a heterogeneous lineage-biased population as well. MPP subgroups have been distinguished as being megakaryocytic, myeloid and lymphoid biased (Pietras et al., 2015).

The first functional evidence of a heterogeneous stem cell compartment arose from single cell transplantation experiments of phenotypic HSCs resulting in different outputs. Several single cell transplantation experiments had widely variable reconstitution kinetics and very rarely a single transplanted HSC gave rise to a balanced differentiated progeny (Dykstra et al., 2007; Morita et al., 2010; Müller-Sieburg et al., 2002; Yamamoto et al., 2013). These results are in contrast with the view of a homogeneous HSCs population and infer that every HSC possess a very specific, if not unique, self-renewal and differentiation

characteristic. Transplantation experiments test the ability of an HSC to generate other cell types when placed in a new environment, also called potency, while the direction of an HSC in an unperturbed setting represents its cell fate. Cell fate, in this context, has been studied by ex vivo barcoding and transplantation (Aiuti et al., 2013; Gerrits et al., 2010; Jordan and Lemischka, 1990; Lemischka, 1993; Lemischka et al., 1986; Mazurier et al., 2004; Shi et al., 2002; Snodgrass and Keller, 1987) and more recently by *in vivo* barcoding experiments (Rodriguez-Fraticelli et al., 2018; Yu et al., 2016), confirming the presence of HSC subpopulations in steady state hematopoiesis. In the last years, with the aid of new technologies, like single cell RNA sequencing, a molecular and functional heterogeneity in the HSC compartment has been supported by multiple research groups. The technological advances in single cell transcriptome analysis now allow for a higher sensitivity and to detect multiple genes expressed by single HSCs. In summary, transplantation, single cell RNA sequencing and barcoding experiments of putative HSCs confirmed the presence of primed HSCs (Figure 2) by revealing a heterogeneity in the expression of transcription factors and revealing hematopoiesis as a less stepwise and more continuous process (Grover et al., 2016; Karamitros et al., 2018; Moignard et al., 2013; Velten et al., 2017).

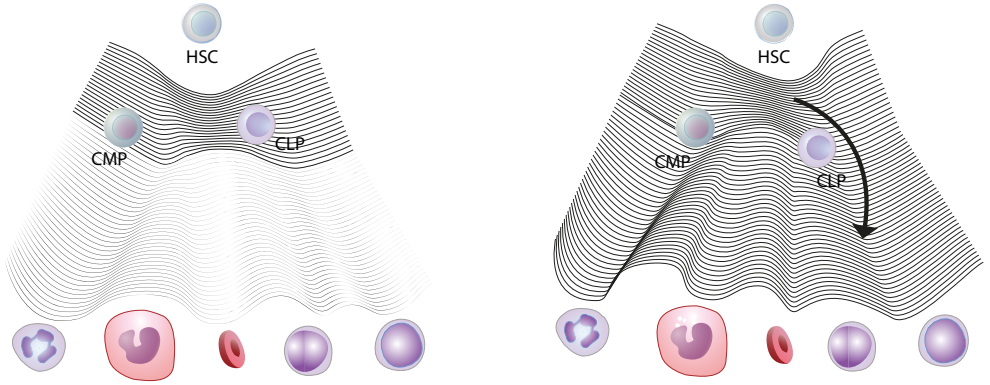


Figure 2. HSC Lineage Biases and Transcriptional Lineage Priming
Adapted from Haas et al (Haas et al., 2018). Balanced HSCs show a roughly equivalent contribution toward the production of all mature blood cells (left). Lineage biases HSCs show an unbalanced production of differentiated cells (right). Transcriptional lineage priming is an indicator of the differentiation path that multipotent cells will take.

1.2 GATA2

More than 500 transcription factors are encoded in the human genome (Hamosh et al., 2002). Six GATA factors (GATA1 to 6) have been identified in vertebrates. These proteins are pivotal transcriptional regulators involved in both physiological and pathological states. Their name is derived from their ability to bind the consensus DNA sequence (A/T) GATA(A/G) using two zinc finger protein structural motifs (Ko and Engel, 1993). While GATAs 4 to 6 have been associated to cardiac, pulmonary, pancreatic and hepatic function

(Molkentin, 2000), GATAs 1 to 3 are mostly linked to the development and-or maintenance of the hematopoietic system. Having crucial functions in the ontogeny of the hematopoietic system, each of the mouse Gata1, Gata2 or Gata3 homozygous knockout is embryonically lethal (Gata1 is hemizygous in males) (Fujiwara et al., 1996; Pandolfi et al., 1995; Tsai et al., 1994). In humans, heterozygous germline *GATA2* mutations strongly predispose to hematological diseases (Crispino and Horwitz, 2017). Co-occurring mutations are present in GATA2 heterozygous patients progressing to myelodysplastic syndromes (MDS)- acute myeloid leukemia (AML) with mutations in ASXL1 being the most common (West et al., 2014).

Germline inactivating mutations in one of the *GATA2* alleles are the primary cause of a series of disorders called GATA2 haploinsufficiency syndromes. GATA2 deficiency results in different manifestations with variable age of onset and outcomes. GATA2 deficiency is responsible for monocytopenia and mycobacterial infection (MonoMAC) syndrome and Emberger syndrome. MonoMAC is a complex phenotype involving susceptibility to various infections and cytopenias in monocytes, B-and NK cells (Vinh et al., 2010). Emberger syndrome involves lymphedema and myelodysplasia or leukemia (Hsu et al., 2011). Other GATA2 deficiencies include dendritic cell, monocyte, B and NK lymphoid (DCML) deficiency (Bigley and Collin, 2011; Dickinson et al., 2011), mild chronic neutropenia (Pasquet et al., 2013) and familial AML (Hahn et al., 2011). GATA2 haploinsufficiency has been associated with a probability to develop MDS-AML ranging from 39% at 20 years old to 80% at 40 years old (Donadieu et al., 2018). Out of all AML patients, 3.6% to 14.4% are reported to carry a *GATA2* mutation (Fasan et al., 2013; Green et al., 2013; Greif et al., 2012; Hahn et al., 2011).

In myeloid malignancies, heterozygous *GATA2* mutations are found throughout the gene, but often cluster in one of the two zinc finger domains, most likely affecting the ability of GATA2 to bind DNA and other proteins, thus causing haploinsufficiency (Collin et al., 2015). *GATA2* mutations are diverse (missense, nonsense, deletion, uniallelic or in regulatory elements) but no specific phenotype has been associated to a mutation type (except lymphedema in nonsense and deletion mutations)(Spinner et al., 2014). Mutations abrogating GATA2 (like frameshift mutations) are associated with an earlier onset of clinical presentation than less dramatic mutations (like substitution mutations), more likely to contribute in a transcriptional reduction (Dickinson et al., 2014). The WT allele in *GATA2* haploinsufficient patients appears to retain functionality (Crispino and Horwitz, 2017), excluding the need for both alleles to be non-functional for the phenotype to occur. Besides a loss in GATA2 functionality, GATA2 over-expression can result in acute myeloid leukemia (Persons et al., 1999; Tipping et al., 2009; Zhang et al., 2008). Since both GATA2 over-expression and down-regulation can cause blood phenotypes, the correct levels of GATA2 expression are essential for healthy hematopoiesis. With the goal of understanding the effect of GATA2 de-regulation in mind, in **chapter 3 and 4** we explore the effect of a zebrafish *Gata2b* and a mouse *Gata2* heterozygous knockout.

1.2.1 Gata2 and ontogeny of murine hematopoiesis

While HSCs in adult mammals reside in the bone marrow, their origin can be traced back during embryogenesis. An understanding of GATA2 function during ontogeny is particularly important because of its essential role in embryonal hematopoiesis and because the pathogenic mutations are frequently present in the germline. GATA2 possess unique roles during the generation, amplification and maintenance of embryonic HSCs (Ling et al., 2004; de Pater et al., 2013; Rodrigues et al., 2005; Tsai and Orkin, 1997). In mice the first hematopoietic cells appear at embryonic day (E)7.5 in the blood island of the yolk sac, during the so-called primitive hematopoiesis. This process is temporary and allows the developing organism to sustain growth via tissue remodelling, oxygen and metabolite distribution before the overtake of definitive hematopoiesis (Jagannathan-Bogdan and Zon, 2013). Although this process is transitory, it gives rise to permanent tissue resident macrophages (Palis et al., 1999; Tober et al., 2007). At this stage of embryonic development, Gata2 is expressed in the primitive streak and in endothelial cells of the paired dorsal aorta (Kaimakis et al., 2016; Minegishi et al., 1999; Robert-Moreno et al., 2005). A second wave of yolk-sac-derived primitive hematopoiesis causes the production of erythro-myeloid progenitor cells (EMP) by E8.5 (McGrath et al., 2015) and immune restricted progenitors (Böiers et al., 2013). These HSC-independent progenitor cells are the first hematopoietic cells with multilineage potential.

1.2.2 GATA2 in definitive hematopoiesis

The first HSCs are generated in the aorta-gonads-mesonephros (AGM) region (Dieterlen Lievre, 1975; Medvinsky and Dzierzak, 1996) and other arteries (de Bruijn et al., 2000) from specialized hemogenic endothelial cells (Ohneda et al., 1998; Zovein et al., 2008), through a conserved mechanism called endothelial to hematopoietic transition (EHT) at E10.5 (Bertrand et al., 2010; Boisset et al., 2010; Dzierzak and Bigas, 2018; Jaffredo et al., 1998; Kissa and Herbomel, 2010; Müller et al., 1994). During EHT, hemogenic endothelial cells activate a hematopoietic transcriptional program which requires the balanced expression of transcription factors, such as Runx1 and Gata2. These expression changes cause morphological alterations like the loosening of tight junctions and the rounding up of cells forming intra-aortic hematopoietic cluster cells (IAHCs). The HSPCs at this stage are marked by the co-expression of the endothelial marker CD31 and c-Kit (Sasaki et al., 2010; Yokomizo and Dzierzak, 2010). HSCs can also be found after E10.5 in vitelline and umbilical arteries, placenta, yolk sac, and embryonic head (De Bruijn et al., 2000; Gekas et al., 2005; Gordon-Keylock et al., 2013; Li et al., 2012; Medvinsky et al., 2008; Müller et al., 1994; Ottersbach and Dzierzak, 2005; Robin et al., 2011; Rybtsov et al., 2011). Gata2 is expressed during definitive hematopoiesis in the hemogenic endothelium and emerging HSCs and IAHCs at E10 (Eich et al., 2018).

Using a conditional KO, Gata2 has been deleted before the generation of HSCs using Vec-Cadherin-expressing endothelial cells (*Vec-Cre*:Gata2$^{f/f}$). This mouse model had a deficiency in intra-aortic cluster cell generation and did not generate long-term repopulating HSCs, confirming a quantitative and functional role of Gata2 in HSC generation in hemogenic endothelium. Gata2 has a dose-dependent function in the generation of HSCs. Indeed, Gata2 homozygous knock out mouse embryos are anemic and die at E10.5, not generating HSCs (Tsai et al., 1994), while heterozygous knock out embryos survive to adulthood but have more subtle changes in the hematopoietic system. Gata2 haploinsufficient embryos have significantly reduced emerging HSPCs in the AGM. This decrease in HSPCs was confirmed at E11 by a Sca-1 transgenic line (Ling et al., 2004) and at E10 by c-kit immunostaining (de Pater et al., 2013). Besides, AGM and yolk sac explants from Gata2 haploinsufficient embryos have a reduced engraftment capacity in irradiated recipients (Ling et al., 2004).

Gata2 not only controls the generation of HSCs but also their maintenance. HSPCs generated during EHT colonize the fetal liver from E11, mainly in response to Cxcl12 (Chou and Lodish, 2010). Fetal liver colonization from HSPCs is followed by their expansion, reaching a maximum of about 1000 HSC at E14.5 to 15.5 (Ema and Nakauchi, 2000; Morrison et al., 1995). Stem cells at this stage can be distinguished using SLAM markers (Kim et al., 2006). Vascular invasion in the developing bones allows the seeding of HSPCs from circulation in the BM. Gata2 has been deleted after the generation of HSCs using a conditional KO in Vav-expressing hematopoietic cells (*Vav-Cre*:Gata2$^{f/f}$). *Vav-Cre*:Gata2$^{f/f}$ mice had decreased colony formation and their phenotypic HSCs had concurrent decrease in number and increase in apoptosis. Furthermore, transplantation studies showed that LT-HSCs are absent in Vav-CRE;Gata2$^{f/f}$ embryos (de Pater et al., 2013). HSCs migrate to the BM from E17.5 and this site remains the main source of hematopoietic cells in healthy adult mice. While fetal liver HSCs are expanding, and cycle frequently, bone marrow HSCs are mostly quiescent (BM) (Christensen et al., 2004; Wilson et al., 2008). Gata2 haploinsufficient mice are viable and become adults without any apparent anomaly, including normal blood values (Rodrigues et al., 2005). However, their bone marrow has a diminished frequency of CD34- LSK, and LSK cells have an increased proliferation and a propensity for apoptosis.

While GATA2 mutations in the zinc finger domains alter binding, enhancer mutations effect its expression. Two enhancers are essential for the murine GATA2 activity, one 9.5 kb downstream (+9.5) and one 77 Kb upstream (-77) of the transcriptional start site. The deletion of the +9.5 enhancer causes lethality at E13.5, abrogating HSCs generation (Gao et al., 2013), the correspondent human enhancer is mutated in 10% of all the MonoMAC syndrome patients (Hsu et al., 2013). While a deletion of the -77 enhancer allows HSC generation but causes lethality at approximately E15.5 due to a block in myelo-erythropoiesis and HSPC expansion (Johnson et al., 2015). Although it is currently challenging to entirely weigh the various criteria regulating Gata2 dosage, balanced expression of functional GATA2 is necessary for hematopoiesis and its de-regulation is linked to leukemogenesis. Indeed, both high (Luesink et al., 2012; Menendez-Gonzalez et al., 2019) or low Gata2 expression (see haploinsufficiency) predispose to AML..

1.3 Zebrafish in hematological research

Zebrafish (*Danio rerio*) became the emerging model for hematological research in the last decades, particularly for the convenience of molecular manipulation and the possibility to visualize embryogenesis *in vivo,* due to its transparency at this stage. The external fertilization and transparency of zebrafish allows to observe embryogenesis without interferences in the growth process. These advantages, together with the available transgenic reporter lines, allowed researchers to visualize embryonic growth with unprecedented ease. For example, the first *in vivo* EHT events (Bertrand et al., 2010; Kissa and Herbomel, 2010) and the HSPCs and stromal cell interaction in the caudal hematopoietic tissue (CHT) were visualized using zebrafish (Tamplin et al., 2015). The external fertilization allows to modify the genetic material (for example using CRISPR gene editing) directly at one-cell stage. In **chapter 2** and **chapter 3** we took advantage of this opportunity to generate a gata2b knockout. Despite an evolutionary divergence, zebrafish and mammals conserved most of the molecular and physiological processes directing hematopoiesis, like the generation and amplification of HSCs. Differently from mammals, zebrafish erythrocytes and thrombocytes retain their nucleus. However, zebrafish hematopoietic cells maintain a clear resemblance to human cells in a way that all hematopoietic lineages can be distinguished (Figure 3).

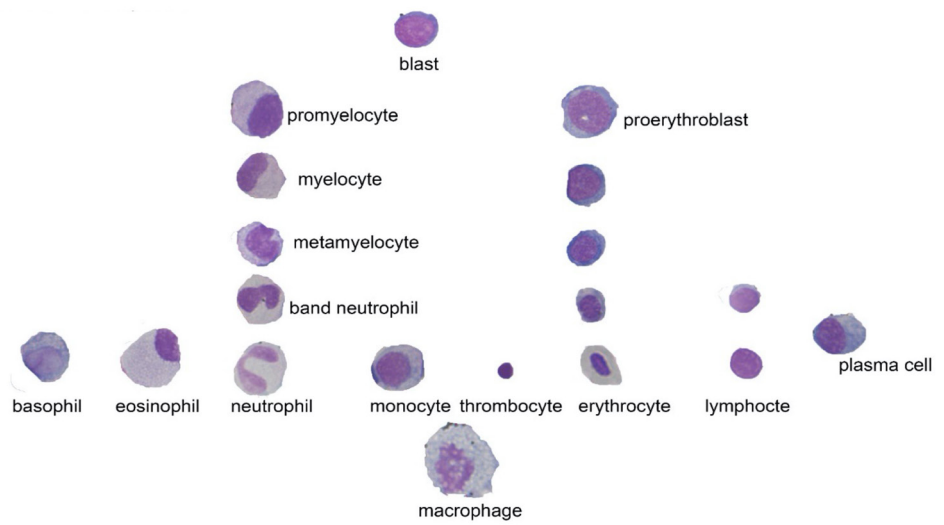


Figure 3. Hematopoietic cells in zebrafish. May-Grünwald-Giemsa stain of kidney marrow (KM) cells. Multiple differentiated cells and intermediate progenitors are recognizable.

Zebrafish belong to the infraclass of teleost, which underwent an extra whole-genome duplication compared to other vertebrates (Meyer and Schartl, 1999). While some of the duplicated genes became pseudogenes, others acquired new functions related to the original gene in a way that more than 70% of human genes have at least one zebrafish orthologue (Howe et al., 2013). The extra genome duplication in teleost caused the generation of two Gata2 orthologues in zebrafish called Gata2a and Gata2b (Gillis et al., 2009). These two paralogues genes share just 57% of sequence identity but retained a well conserved region coding for the two zinc fingers, compared to humans and mice (Butko et al., 2015)..

1.3.1 Ontogeny of zebrafish hematopoiesis

Primitive hematopoiesis in zebrafish begins in the lateral plate mesoderm (LPM). At around 15hpf the LMP forms the inner cell mass (ICM) which becomes the site of primitive hematopoiesis, generating erythroid and myeloid precursors (Detrich et al., 1995). Like in mammals, the first zebrafish definitive blood cells are derived from endothelial cells that undergo EHT. EHT occurs in the ventral side of the dorsal aorta and can be tracked in vivo using an endothelial reporter like *Tg(fli:gfp)*. During EHT, flat hemogenic endothelial cells round up, shed in the space between the dorsal aorta and posterior cardinal vein (like shown in Figure 2I of **chapter 2**) and re-enter into circulation through the posterior cardinal vein. After generation, HSPCs enter the circulation and seed the caudal hematopoietic tissue (CHT) which functions as an intermediate hematopoietic site, analogous to the fetal liver of mammals (Figure 4), to expand the HSPC pool. The CHT is a highly vascular tissue that induces expansion of HSPCs with the aid of different cytokines (Murayama et al., 2006). After amplification, HSPCs relocate to the thymus, for T-cell differentiation, and kidney starting from 4dpf (Figure 4). The zebrafish kidney marrow, similarly to the mammal's bone marrow, is the main site of definitive hematopoiesis and all hematopoietic lineages are distinguishable (Figure 3) (Davidson and Zon, 2004).

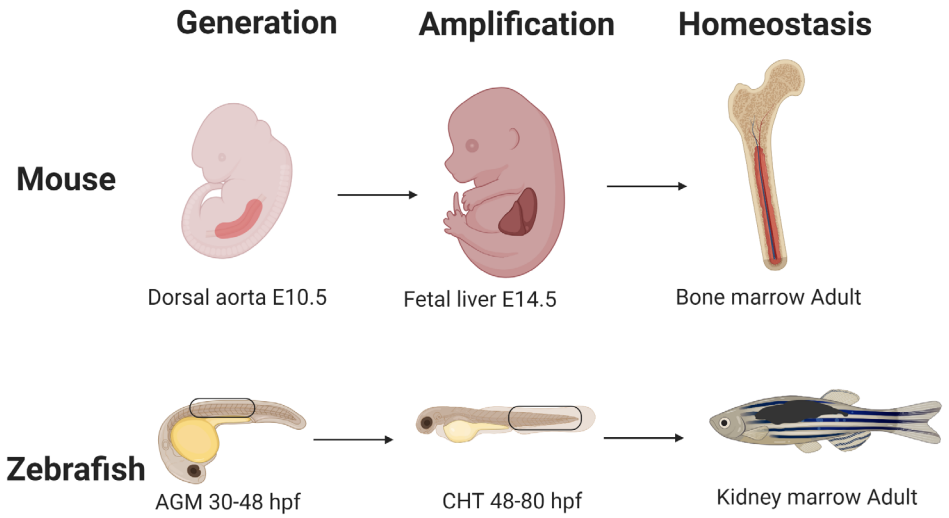


Figure 4. Hematopoietic organs are conserved in zebrafish. Principal organs and time points of hematopoiesis during mouse and zebrafish development. Created with BioRender.com

In the kidney marrow, HSPCs are marked by low levels of *Tg(CD41:GFP)* fluorescent (GFP^{low}) signal (Ma et al., 2011) or can be distinguished based on their light scatter, together with the remaining main hematopoietic populations (Figure 5) (Traver et al., 2003). In **chapter 2** and **chapter 3** we took advantage of both methods for identifying and characterizing HSPCs.

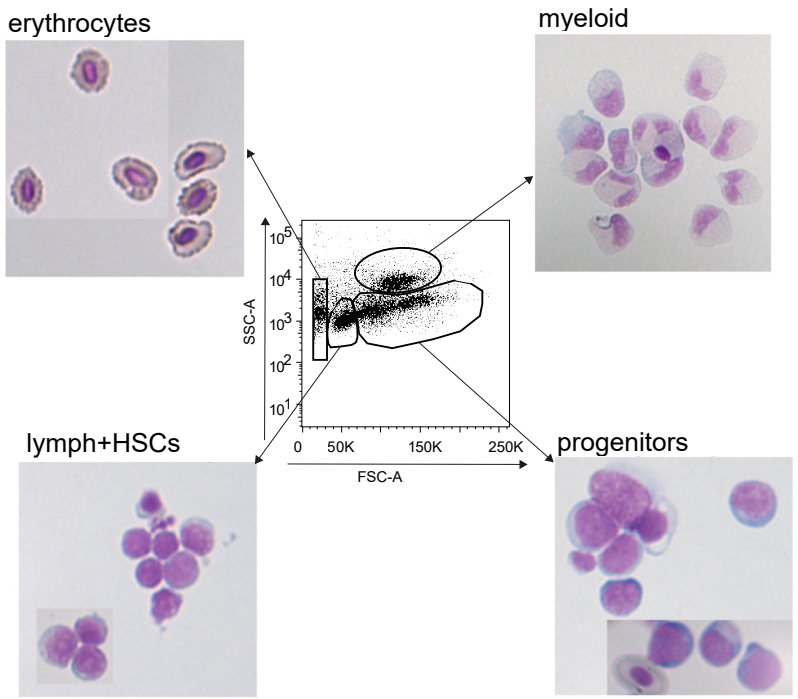


Figure 5. Hematopoietic lineage differentiation in zebrafish. Whole kidney marrow populations classification based on FSC-A/SSC-A with corresponding representative images of sorted cells after May–Grünwald-Giemsa staining. The cells within the smaller quadrants depict additional representative cells within the same gate using the same magnification.

1.3.2 Role of Gata2a and Gata2b in hematopoiesis

The expression of Gata2 factors in zebrafish hematopoiesis starts with the expression of Gata2a at around 10.3hpf in the posterior lateral plate mesoderm. Instead Gata2b starts to be expressed at 16-18hpf in endothelial cells, based on qPCR and whole-mount *in situ* hybridization data. Gata2a expression is found mostly in vascular structures like DA and PCV; while Gata2b in sites of HSCs emergence and colonization (Butko et al., 2015). In adult zebrafish, Gata2b is expressed in the thymus and kidney and, using lineage tracing, Gata2b expressing cells have been proven to give rise to most blood cells in the KM. To recapitulate, while Gata2a is expressed more broadly in vasculature, Gata2b is expressed mostly in the hematopoietic system.

In support of a functional activity in the expression sites, homozygously mutated zebrafish in *gata2a* zinc finger nuclease (ZFN) reported shunts and other circulatory problems, particularly in the formation of the dorsal aorta (Zhu et al., 2011). Dobrzycki et al., identified the i4 enhancer in *gata2a* as the equivalent of the +9.5 Gata2 enhancer in mouse. The expression of the gata2a i4 enhancer in endothelial cells was needed to drive Gata2a expression. *gata2a*-i4 enhancer homozygous mutants had lower expression of Runx1 and Gata2b in the hemogenic endothelium at 28hpf and a Notch-driven recovery of these transcription factors starting at 48hpf resulting in normal numbers of HSPCs by 5dpf (Dobrzycki et al., 2020). Gata2b morphants instead have a normal vascular development and arterial specification (Butko et al., 2015) but defects in Runx1 and C-myb expression respectively at 25 and 36hpf. Alongside, Gata2a did not rescue Gata2b morphants. These results indicate an evolutionary differentiation in the activity of Gata2a and Gata2b, and Gata2b as an uncharted transcription factor in zebrafish adult hematopoiesis.

Scope and outline

After years of human genetic analysis, substantial evidence emerged that germline mutations in *GATA2* strongly predispose to the development of MDS and AML, with a penetrance of over 80% by the age of 40 (Donadieu et al., 2018). Furthermore, germline mutated *GATA2* patients can present with a plethora of phenotypes like monocytopenia, DC-, NK and B cell-lymphopenia, neutropenia or lymphedema. In this dissertation we focus on dissecting the role of *GATA2* gene deletions in both zebrafish and mice. To this end, in **chapter 2** we generated a zebrafish *gata2b* mutant line and identified the hematopoietic consequences of the full loss of this transcription factor. Whereas in **chapter 3** we used the same *gata2b* mutant zebrafish line as chapter 2 but analysed its heterozygous loss. These models allowed a direct comparison of the effects of 3 different levels of Gata2b in the zebrafish hematopoietic system. Whereas both the heterozygous and homozygous *gata2b* mutants revealed a transcriptional hyperproliferation in HSPCs, they exhibited distinct abnormalities in kidney marrow cells. *gata2b* homozygous deletion results in lineage differentiation defects resulting in neutropenia and defective B cell differentiation, whereas heterozygous deletion of *gata2b* results in erythroid and myeloid dysplasia, both phenotypes observed in patients. Indicating that different dosages of Gata2b result in specific cellular defects.

In **chapter 4** we used a mouse model to study Gata2 haploinsufficiency. HSPCs from *Gata2* heterozygous mice (*Gata2*$^{+/-}$) had a transcriptional loss of quiescence both in embryogenesis and in adulthood. This proliferative stress, combined with the observation, in GATA2 deficient patients, of a drastic increase in MDS/AML incidence with age (Donadieu et al., 2018; Spinner et al., 2014) led us to investigate the effect of aging in *Gata2*$^{+/-}$ mice. In **chapter 4** we revealed that after aging and transplantation of *Gata2* heterozygous bone marrow recipients mice would develop leukopenia, recapitulating an aspect of the immunodeficiency reported in GATA2 patients. Finally, in **chapter 5** we summarized the most

important findings and interpreted the meaning and relevancy of the results from previous chapters. In perspective to the current available literature we evaluated the limitations of our experiments and made recommendations for future research.

REFERENCES:

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C., et al. (2013). Lentiviral hematopoietic stem cell gene therapy in patients with wiskott-aldrich syndrome. Science (80-.). *341*.

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193–197.

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

Bigley, V., and Collin, M. (2011). Dendritic cell, monocyte, B and NK lymphoid deficiency defines the lost lineages of a new GATA-2 dependent myelodysplastic syndrome. Haematologica *96*, 1081–1083.

Böiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J.C.A., Azzoni, E., Woll, P.S., Mead, A.J., Hultquist, A., Swiers, G., et al. (2013). Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. Cell Stem Cell *13*, 535–548.

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

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

De Bruijn, M.F.T.R., Speck, N.A., Peeters, M.C.E., and Dzierzak, E. (2000). Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo.

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. Dev. *142*, 1050–1061.

Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D.B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., Von Paleske, L., Renders, S., et al. (2014). Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. Cell Stem Cell *15*, 507–522.

Cabezas-Wallscheid, N., Buettner, F., Sommerkamp, P., Klimmeck, D., Ladel, L., Thalheimer, F.B., Pastor-Flores, D., Roma, L.P., Renders, S., Zeisberger, P., et al. (2017). Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. Cell *169*, 807-823.e19.

Cheng, H., Zheng, Z., and Cheng, T. (2020). New paradigms on hematopoietic stem cell differentiation. Protein Cell *11*, 34–44.

Cheshier, S.H., Morrison, S.J., Liao, X., and Weissman, I.L. (1999). In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. Proc. Natl. Acad. Sci. U. S. A. *96*, 3120–3125.

Chou, S., and Lodish, H.F. (2010). Fetal liver hepatic progenitors are supportive stromal cells for hematopoietic stem cells. Proc. Natl. Acad. Sci. U. S. A. *107*, 7799–7804.

Christensen, J.L., Wright, D.E., Wagers, A.J., and Weissman, I.L. (2004). Circulation and chemotaxis of fetal hematopoietic stem cells. PLoS Biol. *2*, E75.

Collery, R.F., and Link, B.A. (2011). Dynamic smad-mediated BMP signaling revealed through transgenic zebrafish. Dev. Dyn. *240*, 712–722.

Collin, M., Dickinson, R., and Bigley, V. (2015). Haematopoietic and immune defects associated with *GATA2* mutation. Br. J. Haematol. *169*, 173–187.

Colvin, G.A., Berz, D., Quesenberry, P.J., Papa, E., and Liu, L. (2009). Short-Term Hematopoietic Stem Cells (ST-HSC) Have Full Long-Term Capacity with Sustained but Reduced Potential Compared with LT-HSC. Blood *114*, 2550–2550.

Crispino, J.D., and Horwitz, M.S. (2017). GATA factor mutations in hematologic disease. Blood *129*, 2103–2110.

Davidson, A.J., and Zon, L.I. (2004). The “definitive” (and ’primitive’) guide to zebrafish hematopoiesis. Oncogene *23*, 7233–7246.

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*, 10713–10717.

Dickinson, R.E., Griffin, H., Bigley, V., Reynard, L.N., Hussain, R., Haniffa, M., Lakey, J.H., Rahman, T., Wang, X.N., McGovern, N., et al. (2011). Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. Blood *118*, 2656–2658.

Dickinson, R.E., Milne, P., Jardine, L., Zandi, S., Swierczek, S.I., McGovern, N., Cookson, S., Ferozepurwalla, Z., Langridge, A., Pagan, S., et al. (2014). The evolution of cellular deficiency in GATA2 mutation. Blood *123*, 863–874.

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

Dobrzycki, T., Mahony, C.B., Krecsmarik, M., Koyunlar, C., Rispoli, R., Peulen-Zink, J., Gussinklo, K., Fedlaoui, B., de Pater, E., Patient, R., et al. (2020). Deletion of a conserved Gata2 enhancer impairs haemogenic endothelium programming and adult Zebrafish haematopoiesis. Commun. Biol. *3*, 1–14.

Donadieu, J., Lamant, M., Fieschi, C., De Fontbrune, F.S., Caye, A., Ouachee, M., Beaupain, B., Bustamante, J., Poirel, H.A., Isidor, B., et al. (2018). Natural history of GATA2 deficiency in a survey of 79 French and Belgian patients. Haematologica *103*, 1278–1287.

Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S.J., Brinkman, R., and Eaves, C. (2007). Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo. Cell Stem Cell *1*, 218–229.

Dzierzak, E., and Bigas, A. (2018). Blood Development: Hematopoietic Stem Cell Dependence and Independence. Cell Stem Cell *22*, 639–651.

Eich, C., Arlt, J., Vink, C.S., Kartalaei, P.S., Kaimakis, P., Mariani, S.A., Linden, R. van der, van Cappellen, W.A., and Dzierzak, E. (2018). In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. J. Exp. Med. *215*, 233–248.

Eliasson, P., and Jönsson, J.I. (2010). The hematopoietic stem cell niche: Low in oxygen but a nice place to be. J. Cell. Physiol. *222*, 17–22.

Ema, H., and Nakauchi, H. (2000). Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. Blood *95*, 2284–2288.

Fasan, A., Eder, C., Haferlach, C., Grossmann, V., Kohlmann, A., Dicker, F., Kern, W., Haferlach, T., and Schnittger, S. (2013). GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. Leukemia *27*, 482–485.

Foudi, A., Hochedlinger, K., Van Buren, D., Schindler, J.W., Jaenisch, R., Carey, V., and Hock, H. (2009). Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. Nat. Biotechnol. *27*, 84–90.

Fridman, J.S., and Lowe, S.W. (2003). Control of apoptosis by p53. Oncogene *22*, 9030–9040.

Fuchs, E., and Chen, T. (2013). A matter of life and death: Self-renewal in stem cells. EMBO Rep. *14*, 39–48.

Fujiwara, Y., Browne, C.P., Cunniff, K., Goff, S.C., and Orkin, S.H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. Proc. Natl. Acad. Sci. U. S. A. *93*, 12355–12358.

Gao, X., Johnson, K.D., Chang, Y.I., Boyer, M.E., Dewey, C.N., Zhang, J., and Bresnick, E.H. (2013). Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. J. Exp. Med. *210*, 2833–2842.

Gekas, C., Dieterlen-Lièvre, F., Orkin, S.H., and Mikkola, H.K.A. (2005). The placenta is a niche for hematopoietic stem cells. Dev. Cell *8*, 365–375.

Gerrits, A., Dykstra, B., Kalmykowa, O.J., Klauke, K., Verovskaya, E., Broekhuis, M.J.C., De Haan, G., and Bystrykh, L. V. (2010). Cellular barcoding tool for clonal analysis in the hematopoietic system. Blood *115*, 2610–2618.

Gillis, W.Q., St John, J., Bowerman, B., and Schneider, S.Q. (2009). Whole genome duplications and expansion of the vertebrate GATA transcription factor gene family. BMC Evol. Biol. *9*, 207.

Gordon-Keylock, S., Sobiesiak, M., Rybtsov, S., Moore, K., and Medvinsky, A. (2013). Mouse extraembryonic arterial vessels harbor precursors capable of maturing into definitive HSCs. Blood *122*, 2338–2345.

Green, C.L., Tawana, K., Hills, R.K., Bödör, C., Fitzgibbon, J., Inglott, S., Ancliff, P., Burnett, A.K., Linch, D.C., and Gale, R.E. (2013). GATA2 mutations in sporadic and familial acute myeloid leukaemia patients with CEBPA mutations. Br. J. Haematol. *161*, 701–705.

Greif, P.A., Dufour, A., Konstandin, N.P., Ksienzyk, B., Zellmeier, E., Tizazu, B., Sturm, J., Benthaus, T., Herold, T., Yaghmaie, M., et al. (2012). GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. Blood *120*, 395–403.

Grover, A., Sanjuan-Pla, A., Thongjuea, S., Carrelha, J., Giustacchini, A., Gambardella, A., Macaulay, I., Mancini, E., Luis, T.C., Mead, A., et al. (2016). Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. Nat. Commun. *7*.

Gulati, G.S., Sikandar, S.S., Wesche, D.J., Manjunath, A., Bharadwaj, A., Berger, M.J., Ilagan, F., Kuo, A.H., Hsieh, R.W., Cai, S., et al. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. Science (80-.). *367*, 405–411.

Haas, S., Trumpp, A., and Milsom, M.D. (2018). Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. Cell Stem Cell *22*, 627–638.

Hahn, C.N., Chong, C.-E., Carmichael, C.L., Wilkins, E.J., Brautigan, P.J., Li, X.-C., Babic, M., Lin, M., Carmagnac, A., Lee, Y.K., et al. (2011). Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. Nat. Genet. *43*, 1012–1017.

Hamosh, A., Scott, A.F., Amberger, J., Bocchini, C., Valle, D., and McKusick, V.A. (2002). Onlined Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. *30*, 52–55.

Hirata, Y., Furuhashi, K., Ishii, H., Li, H.W., Pinho, S., Ding, L., Robson, S.C., Frenette, P.S., and Fujisaki, J. (2018). CD150 high Bone Marrow Tregs Maintain Hematopoietic Stem Cell Quiescence and Immune Privilege via Adenosine. Cell Stem Cell *22*, 445-453.e5.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. Nature *496*, 498–503.

Hsu, A.P., Sampaio, E.P., Khan, J., Calvo, K.R., Lemieux, J.E., Patel, S.Y., Frucht, D.M., Vinh, D.C., Auth, R.D., Freeman, A.F., et al. (2011). Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. Blood *118*, 2653–2655.

Hsu, A.P., Johnson, K.D., Falcone, E.L., Sanalkumar, R., Sanchez, L., Hickstein, D.D., Cuellar-Rodriguez, J., Lemieux, J.E., Zerbe, C.S., Bresnick, E.H., et al. (2013). GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. Blood *121*, 3830–3837.

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev. *11*, 774–785.

Ito, K., Turcotte, R., Cui, J., Zimmerman, S.E., Pinho, S., Mizoguchi, T., Arai, F., Runnels, J.M., Alt, C., Teruya-Feldstein, J., et al. (2016). Self-renewal of a purified Tie2+ hematopoietic stem cell population relies on mitochondrial clearance. Science (80-.). *354*, 1156–1160.

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

Jagannathan-Bogdan, M., and Zon, L.I. (2013). Hematopoiesis. Dev. *140*, 2463–2467.

Johnson, K.D., Kong, G., Gao, X., Chang, Y.I., Hewitt, K.J., Sanalkumar, R., Prathibha, R., Ranheim, E.A., Dewey, C.N., Zhang, J., et al. (2015). Cis-regulatory mechanisms governing stem and progenitor cell transitions. Sci. Adv. *1*.

Jordan, C.T., and Lemischka, I.R. (1990). Clonal and systemic analysis of long-term hematopoiesis in the mouse. Genes Dev. *4*, 220–232.

Kaimakis, P., De Pater, E., Eich, C., Kartalaei, P.S., Kauts, M.L., Vink, C.S., Van Der Linden, R., Jaegle, M., Yokomizo, T., Meijer, D., et al. (2016). Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. Blood *127*, 1426–1437.

Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., Quek, L., Otto, G., Repapi, E., Doondeea, J., et al. (2018). Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells article. Nat. Immunol. *19*, 85–97.

Kent, D.G., Copley, M.R., Benz, C., Wöhrer, S., Dykstra, B.J., Ma, E., Cheyne, J., Zhao, Y., Bowie, M.B., Zhao, Y., et al. (2009). Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. Blood *113*, 6342–6350.

Kiel, M.J., Yilmaz, Ö.H., Iwashita, T., Yilmaz, O.H., Terhorst, C., and Morrison, S.J. (2005). SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. Cell *121*, 1109–1121.

Kim, I., He, S., Yilmaz, Ö.H., Kiel, M.J., and Morrison, S.J. (2006). Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. Blood *108*, 737–744.

Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature *464*, 112–115.

Ko, L.J., and Engel, J.D. (1993). DNA-binding specificities of the GATA transcription factor family. Mol. Cell. Biol. *13*, 4011–4022.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell *91*, 661–672.

Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature *553*, 418–426.

Lemischka, I.R. (1993). Retroviral lineage studies: some principals and applications. Curr. Opin. Genet. Dev. *3*, 115–118.

Lemischka, I.R., Raulet, D.H., and Mulligan, R.C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. Cell *45*, 917–927.

Li, Z., Lan, Y., He, W., Chen, D., Wang, J., Zhou, F., Wang, Y., Sun, H., Chen, X., Xu, C., et al. (2012). Mouse embryonic head as a site for hematopoietic stem cell development. Cell Stem Cell *11*, 663–675.

Ling, K.W., Ottersbach, K., Van Hamburg, J.P., Oziemlak, A., Tsai, F.Y., Orkin, S.H., Ploemacher, R., Hendriks, R.W., and Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. J. Exp. Med. *200*, 871–882.

Lu, R., Neff, N.F., Quake, S.R., and Weissman, I.L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat. Biotechnol. *29*, 928–934.

Luesink, M., Hollink, I.H.I.M., Van Der Velden, V.H.J., Knops, R.H.J.N., Boezeman, J.B.M., De Haas, V., Trka, J., Baruchel, A., Reinhardt, D., Van Der Reijden, B.A., et al. (2012). High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. Blood *120*, 2064–2075.

Ma, D., Zhang, J., Lin, H.F., Italiano, J., and Handin, R.I. (2011). The identification and characterization of zebrafish hematopoietic stem cells. Blood *118*, 289–297.

Macaulay, I.C., Svensson, V., Labalette, C., Ferreira, L., Hamey, F., Voet, T., Teichmann, S.A., and Cvejic, A. (2016). Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. Cell Rep. *14*, 966–977.

Mazurier, F., Gan, O.I., McKenzie, J.L., Doedens, M., and Dick, J.E. (2004). Lentivector-mediated clonal tracking reveals intrinsic heterogeneity in the human hematopoietic stem cell compartment and culture-induced stem cell impairment. Blood *103*, 545–552.

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.

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

Medvinsky, A., Taoudi, S., Mendes, S., and Dzierzak, E. (2008). Analysis and manipulation of hematopoietic progenitor and stem cells from murine embryonic tissues. Curr. Protoc. Stem Cell Biol. *4*, 2A.6.1-2A.6.25.

Menendez-Gonzalez, J.B., Sinnadurai, S., Gibbs, A., Thomas, L. anne, Konstantinou, M., Garcia-Valverde, A., Boyer, M., Wang, Z., Boyd, A.S., Blair, A., et al. (2019). Inhibition of GATA2 restrains cell proliferation and enhances apoptosis and chemotherapy mediated apoptosis in human GATA2 overexpressing AML cells. Sci. Rep. *9*.

Meyer, A., and Schartl, M. (1999). Gene and genome duplications in vertebrates: The one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. *11*, 699–704.

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

Moignard, V., MacAulay, I.C., Swiers, G., Buettner, F., Schütte, J., Calero-Nieto, F.J., Kinston, S., Joshi, A., Hannah, R., Theis, F.J., et al. (2013). Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. Nat. Cell Biol. *15*, 363–372.

Molkentin, J.D. (2000). The zinc finger-containing transcription factors GATA-4, -5, and -6: Ubiquitously expressed regulators of tissue-specific gene expression. J. Biol. Chem. *275*, 38949–38952.

Morita, Y., Ema, H., and Nakauchi, H. (2010). Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. J. Exp. Med. *207*, 1173–1182.

Morrison, S.J., and Weissman, I.L. (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity *1*, 661–673.

Morrison, S.J., Hemmati, H.D., Wandycz, A.M., and Weissman, I.L. (1995). The purification and characterization of fetal liver hematopoietic stem cells. Proc. Natl. Acad. Sci. U. S. A. *92*, 10302–10306.

Müller-Sieburg, C.E., Cho, R.H., Thoman, M., Adkins, B., and Sieburg, H.B. (2002). Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. Blood *100*, 1302–1309.

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

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.

Nakauchi, Hiromitsu Sudo, Kazuhiro Ema, H. (2006). Quantitative Assessment of the Stem Cell Self-Renewal Capacity. Ann. N. Y. Acad. Sci. *938*, 18–25.

Nimmo, R.A., May, G.E., and Enver, T. (2015). Primed and ready: Understanding lineage commitment through single cell analysis. Trends Cell Biol. *25*, 459–467.

Ogawa, M. (1993). Differentiation and proliferation of hematopoietic stem cells. Blood *81*, 2844–2853.

Oguro, H., Ding, L., and Morrison, S.J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. Cell Stem Cell *13*, 102–116.

Ohneda, O., Fennie, C., Zheng, Z., Donahue, C., La, H., Villacorta, R., Cairns, B., and Lasky, L.A. (1998). Hematopoietic stem cell maintenance and differentiation are supported by embryonic aorta-gonad-mesonephros region-derived endothelium. Blood *92*, 908–919.

Osawa, M., Hanada, K.I., Hamada, H., and Nakauchi, H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34- low/negative hematopoietic stem cell. Science (80-.). *273*, 242–245.

Ottersbach, K., and Dzierzak, E. (2005). The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. Dev. Cell *8*, 377–387.

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

Pandolfi, P.P., Roth, M.E., Karis, A., Leonard, M.W., Dzierzak, E., Grosveld, F.G., Engel, J.D., and Lindenbaum, M.H. (1995). Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. Nat. Genet. *11*, 40–44.

Pasquet, M., Bellanné-Chantelot, C., Tavitian, S., Prade, N., Beaupain, B., LaRochelle, O., Petit, A., Rohrlich, P., Ferrand, C., Van Den Neste, E., et al. (2013). High frequency of GATA2 mutations in patients with mild chronic neutropenia evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. Blood *121*, 822–829.

de Pater, E., Kaimakis, P., Vink, C.S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P.S., Camper, S.A., Speck, N., and Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. J. Exp. Med. *210*, 2843–2850.

Persons, D.A., Allay, J.A., Allay, E.R., Ashmun, R.A., Orlic, D., Jane, S.M., Cunningham, J.M., and Nienhuis, A.W. (1999). Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. Blood *93*, 488–499.

Pietras, E.M., Warr, M.R., and Passegué, E. (2011). Cell cycle regulation in hematopoietic stem cells. J. Cell Biol. *195*, 709–720.

Pietras, E.M., Reynaud, D., Kang, Y.A., Carlin, D., Calero-Nieto, F.J., Leavitt, A.D., Stuart, J.A., Göttgens, B., and Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. Cell Stem Cell *17*, 35–46.

Purton, L.E., and Scadden, D.T. (2007). Limiting factors in murine hematopoietic stem cell assays. Cell Stem Cell *1*, 263–270.

Robert-Moreno, À., Espinosa, L., de la Pompa, J.L., and Bigas, A. (2005). RBPjκ-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. Development *132*, 1117–1126.

Robin, C., Ottersbach, K., Boisset, J.C., Oziemlak, A., and Dzierzak, E. (2011). CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells. Blood *117*, 5088–5091.

Rodrigues, N.P., Janzen, V., Forkert, R., Dombkowski, D.M., Boyd, A.S., Orkin, S.H., Enver, T., Vyas, P., and Scadden, D.T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. Blood *106*, 477–484.

Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M., Sun, J., Calogero, R.A., Klein, A.M., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature *553*, 212–216.

Rybtsov, S., Sobiesiak, M., Taoudi, S., Souilhol, C., Senserrich, J., Liakhovitskaia, A., Ivanovs, A., Frampton, J., Zhao, S., and Medvinsky, A. (2011). Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. J. Exp. Med. *208*, 1305–1315.

Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Jones, T.B., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232–236.

Sasaki, T., Mizuochi, C., Horio, Y., Nakao, K., Akashi, K., and Sugiyama, D. (2010). Regulation of hematopoietic cell clusters in the placental niche through SCF/Kit signaling in embryonic mouse. Development *137*, 3941–3952.

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

Shi, P.A., Hematti, P., Von Kalle, C., and Dunbar, C.E. (2002). Genetic marking as an approach to studying in vivo hematopoiesis: Progress in the non-human primate model. Oncogene *21*, 3274–3283.

Simsek, T., Kocabas, F., Zheng, J., Deberardinis, R.J., Mahmoud, A.I., Olson, E.N., Schneider, J.W., Zhang, C.C., and Sadek, H.A. (2010). The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche. Cell Stem Cell *7*, 380–390.

Snodgrass, R., and Keller, G. (1987). Clonal fluctuation within the haematopoietic system of mice reconstituted with retrovirus-infected stem cells. EMBO J. *6*, 3955–3960.

Spangrude, G.J., and Brooks, D.M. Mouse Strain Variability in the Expression of the Hematopoietic Stem Cell Antigen Ly-6A/E by Bone Marrow Cells.

Spangrude, G.J., Heimfeld, S., and Weissman, I.L. (1988). Purification and characterization of mouse hematopoietic stem cells. Science (80-.). *241*, 58–62.

Spinner, M.A., Sanchez, L.A., Hsu, A.P., Shaw, P.A., Zerbe, C.S., Calvo, K.R., Arthur, D.C., Gu, W., Gould, C.M., Brewer, C.C., et al. (2014). GATA2 deficiency: A protean disorder of hematopoiesis, lymphatics, and immunity. Blood *123*, 809–821.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

Szilvassy, S.J., Humphries, R.K., Lansdorp, P.M., Eaves, A.C., and Eaves, C.J. (1990). Quantitative assay for totipotent reconstituting hematopoietic stem cells by a competitive repopulation strategy. Proc. Natl. Acad. Sci. U. S. A. *87*, 8736–8740.

Takubo, K., Nagamatsu, G., Kobayashi, C.I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., Goda, N., Rahimi, Y., Johnson, R.S., Soga, T., et al. (2013). Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. Cell Stem Cell *12*, 49–61.

Tamplin, O.J., Durand, E.M., Carr, L.A., Childs, S.J., Hagedorn, E.J., Li, P., Yzaguirre, A.D., Speck, N.A., and Zon, L.I. (2015). Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. Cell *160*, 241–252.

Till, J.E. McCulloch, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. *14*, 213–222.

Tipping, A.J., Pina, C., Castor, A., Hong, D., Rodrigues, N.P., Lazzari, L., May, G.E., Jacobsen, S.E.W., and Enver, T. (2009). High GATA-2 expression inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. Blood *113*, 2661–2672.

Tober, J., 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.

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

Tsai, F.Y., and Orkin, S.H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood *89*, 3636–3643.

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

Uchida, N., and Weissman, I.L. (1992). Searching for hematopoietic stem cells: Evidence that Thy-l.1lo Lin− Sca-1+ Cells Are the Only Stem Cellsin C57BL/Ka-Thy-l.1 bone marrow. J. Exp. Med. *175*, 175–184.

Udagawa, N., Takahashi, N., Akatsu, T., Tanaka, H., Sasaki, T., Nishihara, T., Koga, T., Martin, T.J., and Suda, T. (1990). Origin of osteoclasts: Mature monocytes and macrophages are capable of differentiating into osteoclasts under a suitable microenvironment prepared by bone marrow-derived stromal cells. Proc. Natl. Acad. Sci. U. S. A. *87*, 7260–7264.

Vannini, N., Girotra, M., Naveiras, O., Nikitin, G., Campos, V., Giger, S., Roch, A., Auwerx, J., and Lutolf, M.P. (2016). Specification of haematopoietic stem cell fate via modulation of mitochondrial activity. Nat. Commun. *7*, 1312.

Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. *19*, 271–281.

Vinh, D.C., Patel, S.Y., Uzel, G., Anderson, V.L., Freeman, A.F., Olivier, K.N., Spalding, C., Hughes, S., Pittaluga, S., Raffeld, M., et al. (2010). Autosomal dominant and sporadic monocytopenia with susceptibility to mycobacteria, fungi, papillomaviruses, and myelodysplasia. Blood *115*, 1519–1529.

West, R.R., Hsu, A.P., Holland, S.M., Cuellar-Rodriguez, J., and Hickstein, D.D. (2014). Acquired ASXL1 mutations are common in patients with inherited GATA2 mutations and correlate with myeloid transformation. Haematologica *99*, 276–281.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R.C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C.F., Eshkind, L., Bockamp, E., et al. (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. Cell *135*, 1118–1129.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell *154*, 1112–1126.

Yang, L., Bryder, D., Adolfsson, J., Nygren, J., Månsson, R., Sigvardsson, M., and Jacobsen, S.E.W. (2005). Identification of Lin-Sca1+kit+CD34 +Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. Blood *105*, 2717–2723.

Yilmaz, Ö.H., Kiel, M.J., and Morrison, S.J. (2006). SLAM family markers are conserved among hematopoietic stem cells from old and reconstituted mice and markedly increase their purity. Blood *107*, 924–930.

Yokomizo, T., and Dzierzak, E. (2010). Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. Development *137*, 3651–3661.

Yu, V.W.C., Yusuf, R.Z., Oki, T., Wu, J., Saez, B., Wang, X., Cook, C., Baryawno, N., Ziller, M.J., Lee, E., et al. (2016). Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. Cell *167*, 1310-1322.e17.

Zhang, D.E., Zhang, P., Wang, N.D., Hetherington, C.J., Darlington, G.J., and Tenen, D.G. (1997). Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein α-deficient mice. Proc. Natl. Acad. Sci. U. S. A. *94*, 569–574.

Zhang, S.J., Ma, L.Y., Huang, Q.H., Li, G., Gu, B.W., Gao, X.D., Shi, J.Y., Wang, Y.Y., Gao, L., Cai, X., et al. (2008). Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. Proc. Natl. Acad. Sci. U. S. A. *105*, 2076–2081.

Zhao, Y., Lin, Y., Zhan, Y., Yang, G., Louie, J., Harrison, D.E., and Anderson, W.F. (2000). Murine hematopoietic stem cell characterization and its regulation in BM transplantation. Blood *96*, 3016–3022.

Zhu, C., Smith, T., Mcnulty, J., Rayla, A.L., Lakshmanan, A., Siekmann, A.F., Buffardi, M., Meng, X., Shin, J., Padmanabhan, A., et al. (2011). Evaluation and application of modularly assembled zincfinger nucleases in zebrafish. Development *138*, 4555–4564.

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

2

Essential role for Gata2 in modulating lineage output from hematopoietic stem cells identified in zebrafish

Emanuele Gioacchino[1,7], Cansu Koyunlar[1,7], Joke Zink[1], Hans de Looper[1,2], Madelon de Jong[1], Tomasz Dobrzycki[3], Christopher B Mahony[4], Remco Hoogenboezem[1], Dennis Bosch[1], Paulette M H van Strien[1], Martin E van Royen[5], Pim J French[6], Eric Bindels[1], Kirsten J Gussinklo[1], Rui Monteiro[4], Ivo P Touw[1] and Emma de Pater[1,2]*

[1]Department of Hematology, Erasmus MC, Rotterdam, The Netherlands
[2]Cancer Genome Editing Center of Erasmus MC, Rotterdam, The Netherlands.
[3]Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK.
[4]Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham B152TT, UK.
[5]Department of Pathology, Cancer Treatment Screening Facility, Erasmus MC Optical Imaging Centre, Erasmus MC, Rotterdam, The Netherlands
[6]Department of Neurology, Cancer Treatment Screening Facility, Erasmus MC, Rotterdam, The Netherlands.
[7]Authors contributed equally

*Corresponding author: Email: e.depater@erasmusmc.nl

Key points:

1] Gata2b is required for embryonic HSPC expansion, but not HSPC generation in zebrafish

2] Gata2b plays an instructive role in the lineage output of HSPCs in zebrafish

Blood Adv (2021) 5 (13): 2687–2700.

ABSTRACT

The differentiation of hematopoietic stem cells is tightly controlled to ensure a proper balance between myeloid and lymphoid cell output. GATA2 is a pivotal hematopoietic transcription factor required for generation and maintenance of hematopoietic stem cells. GATA2 is expressed throughout development but due to early embryonic lethality in mouse, its role during adult hematopoiesis is incompletely understood. Zebrafish contains two orthologues of GATA2; Gata2a and Gata2b that are expressed in different cell types. We show that the mammalian functions of GATA2 are split between these orthologues. Gata2b deficient zebrafish have a reduction in embryonic definitive hematopoietic stem and progenitor cell (HSPC) numbers, but are viable. This allows us to uniquely study the role of GATA2 in adult hematopoiesis. *gata2b* mutants have impaired myeloid lineage differentiation. Interestingly, this defect arises not in granulocyte-monocyte progenitors, but already in HSPCs. Gata2b deficient HSPCs showed impaired progression of the myeloid transcriptional program, concomitant with increased co-expression of lymphoid genes. This results in a decrease in myeloid programmed progenitors and a relative increase in lymphoid programmed progenitors. This shift in the lineage output could function as an escape mechanism to avoid a block in lineage differentiation. These studies help to deconstruct the functions of GATA2 during hematopoiesis and show that lineage differentiation flows towards a lymphoid lineage in the absence of Gata2b.

INTRODUCTION

Hematopoietic stem cells (HSCs) have the capacity to self-renew and to generate all lineages of the hematopoietic system[1]. The HSC pool is a heterogeneous population of cells that are tightly controlled by cell-intrinsic and -extrinsic cues to maintain a balance between myeloid and lymphoid cell commitment[2-5]. It is currently under debate whether HSCs can flow between myeloid and lymphoid lineage commitment or whether the HSC pool consists of separate lymphoid biased and myeloid biased HSCs[6].

The transcription factor GATA2 has a key role in blood cell formation during mammalian embryonic development. GATA2 expression is tightly regulated during distinct stages of hematopoietic development and plays crucial roles in the specification of hemogenic endothelium (HE) and the generation and maintenance of HSCs[7-11]. A role for this transcription factor in myeloid/lymphoid commitment is supported by findings of reduced and impaired granulocyte-macrophage progenitors in *Gata2*$^{+/-}$ mice[12-14]. Conversely, retroviral mediated overexpression of *Gata2* results in enhanced self-renewal of the myeloid progenitors and a block in lymphoid differentiation[15]. Homozygous germline deletion of *Gata2* in mice results in embryonic lethality at E10, just before the generation of the first HSCs[16].

Zebrafish is an ideal *in vivo* model to study the function of GATA2 in hematopoiesis. Embryonic hematopoietic development in zebrafish is conserved with that of other vertebrates, including mammals. Like in mice, the first HSCs are generated in the dorsal aorta from hemogenic endothelial (HE) cells and are subsequently amplified in the fetal liver equivalent, the caudal hematopoietic tissue (CHT)[17-22]. The HSCs then populate the kidney marrow which is the site of adult hematopoiesis in zebrafish. In this organ all hematopoietic lineages are present[23] and hematopoietic cells morphologically resemble the corresponding human cells.

Zebrafish have two orthologues of GATA2; i.e., Gata2a and Gata2b. Previous studies have shown that *gata2b* is prominently expressed in HSPCs, whereas *gata2a* is mainly expressed in the vasculature, including the HE regulated by the conserved +9.5 enhancer previously identified in mice[24,25]. Knockdown of *gata2b* severely reduces definitive hematopoiesis during embryonic stages. Lineage tracing revealed that all definitive hematopoietic cells are derived from *gata2b* expressing cells[24], indicating that Gata2b is the predominant GATA2 orthologue required for the maintenance of hematopoietic stem cells.

In the present study, we show that Gata2b is not required for HE specification but regulates embryonic definitive HSPC expansion in the CHT. This allowed us to investigate the function of Gata2b in adult hematopoiesis and here, we demonstrate that Gata2b is necessary for balanced myeloid and lymphoid output during adulthood. Single cell transcriptome analysis revealed that Gata2b deficient HSPCs initiate an impaired myeloid gene expression program. As a result differentiation is not halted, but diverges into a lymphoid program, indicated by co-expression of lymphoid and myeloid genes within single HSPCs.

MATERIALS AND METHODS

Generation of *gata2b* mutant zebrafish

gata2b mutant zebrafish were generated using CRISPR/Cas9 targeting of exon 3. sgRNAs were designed using CHOPCHOP software and prepared according to Gagnon et al.[26] with minor adjustments.

qRT-PCR analysis

Total RNA was isolated from 6 pooled zebrafish embryos per ~~sample~~ genotype (n=6) using TRIzol Reagent (Life Technologies) and cDNA was synthesized using SuperScript III Reverse Transcriptase kit (Invitrogen). *gata2a* (FWD primer: 5'-CAAACTCCACAACGTCAACAG-3', REV primer: 5'-CCCTCACCAGATCGTTTACTC-3') and *gata2b* (FWD primer: 5'-TACACAATGTGAATCGCCCA-3', REV primer: 5'-GAAGGAGGATGGTTTGTCGT-3') expression levels were normalized to *elfa* (FWD primer: 5'-CCGCTAGCATTACCCTCC-3', REV primer: 5'-CTTCTCAGGCTGACTGTG-3') expression .

In situ hybridization (ISH) and analysis

0.003% 1-phenyl-2-thiourea (PTU) treated embryos were fixed O/N with 4% PFA in PBS containing 3% sucrose at appropriate stages and subsequently transferred to MeOH. KM smears were fixed in MeOH. ISH on embryos has been performed as previously described[27]. The *cmyb* and *runx1* probes were a kind gift from Roger Patient and quantified as described previously[28]. ISH on KM smears was performed as follows; DIG-11-UTP labelled *s100a10b* probe was incubated o/n at 68°C. Slides were blocked at RT in MABT (NaCl, Maleic Acid, 1% Tween 20)2% BSA and Sheep Serum for minimum 3 h and αDIG antibody was incubated o/n at 4°C. Staining was developed in Tris pH 9.5, $MgCl_2$, NaCl, Tween 20 with 5% PVA, NBT/BCIP at RT for two days. Cells were counterstained with Nuclear Fast Red (Sigma Aldrich) and imaged using a Leica microscope (63x magnification).

s100a10b probe synthesis

s100a10b was amplified from cDNA of adult kidney marrow (FWD primer: 5'-GAG AGC AAT GGA GAC CCT GA-3', REV primer: 5'-ACT TCT TGG CTG CTG CTT TC-3') and cloned into pCRII-TOPO. Plasmid was linearized with HindIII and antisense probe transcribed with the DIG labelling kit (Sigma-Alderich). Sense probe was used as negative control.

Transgenic lines, confocal imaging and adult KM FACS analysis

Embryos were anesthetized using tricaine (3-amino benzoic acidethylester) 160mg/L and selected for reporter expression. *Tg(fli:eGFP)*[29] and *Tg(CD41:GFP)*[30]*;Tg(flt1:RFP)*[31] embryos were imaged in 0.25% agarose with tricaine and imaged using a Leica SP5 confocal microscope pre-warmed at 28°C. *Tg(mpeg1.1:GFP)*[32], *Tg(mpx:GFP)*[33], *Tg(lck:GFP)*[34] embryos were placed in a 96 well plate (ZFplate, Hashimoto Electronic Industry Co. Ltd, Japan) and imaged using a spinning disk confocal high-throughput microscope system (Opera Phenix, Perkin Elmer) equipped with a dry 10x objective (NA 0.3). B-cell populations were analysed using *Tg(IgM:GFP)*[35] zebrafish. Adult zebrafish were euthanized, KM was isolated and dissociated by pipetting in PBS/10% FCS. 7-AAD (7-amino-actinomycin D) 0.5mg/L (BDbiosciences)or DAPI 1mg/L were used for live/dead discrimination. For embryonic proliferation assay; 25 embryos per genotype were pooled in pre-warmed PBS/10% FCS, single cell suspension was prepared by adding 1% from each collagenase (I, II and IV)(Sigma) and incubating for 45 minutes at 37°C. Proliferation was assessed after 4% PFA fixation and α-Ki67 staining for both embryonic and adult stages. The analysis was performed using FACSAriaIII (BD).

Single cell RNA sequencing

70.000 single viable cells were sorted from 2 pooled KM of female *Tg(CD41:GFP)* zebrafish and supplemented with 114-1607 CD41:GFP^{low} expressing cells. cDNA was prepared using the manufacturers protocol (10x Chromium V2) and sequenced on a Novaseq 6000 instrument (Illumina). Two WT replicates and two *gata2b*$^{-/-}$ replicates were sequenced with the following read depth; WT1: 52384 reads per cell, WT2: 43876 reads per cell, *gata2b* KO1: 53836 reads per cell and *gata2b* KO2: 46761 reads per cell. Data was analyzed using the Seurat R package[36] and detailed description is provided in the supplementary information.

Statistics

All statistical analysis was carried out in GraphPad Prism 5 (GraphPad Software). Normally distributed data were analyzed using One-way ANOVA with Tukey multiple comparison test when comparing three sample sets or a t-test when comparing two sample sets. Data with non-normal distribution were analyzed using a non-parametric Kruskal-Wallis with Dunn correction test.

RESULTS

Generation of a Gata2b deficient zebrafish line

To generate *gata2b* zebrafish mutants, we used CRISPR/Cas9 to target the third exon of the *gata2b* gene (Figure 1A). A 28 bp insertion was introduced, leading to a frameshift truncation from amino acid 185 (Figure 1B-D). qRT-PCR analysis of *gata2b* on pooled WT and *gata2b*$^{-/-}$ embryos at 30 hpf indicated that *gata2b* expression levels, a known transcriptional tartget of Gata2, was significantly reduced in mutant embryos (Figure S1A, B)[25,37]. Hereafter, we refer to this mutant as *gata2b*$^{-/-}$.

Gata2b is dispensable for the generation of hematopoietic stem cells from hemogenic endothelium

The first HSCs transdifferentiate from specialized hemogenic endothelial cells in the aorta-gonad-mesonephros (AGM) region, through a highly conserved process, known as endothelial-to-hematopoietic transition (EHT)[18,19,38-40]. In mice, *Gata2* is expressed in the endothelium, including the hemogenic endothelium (HE) of the dorsal aorta[8], and deletion of *Gata2* results in a reduction in HSC generation[7,8]. *cmyb* and *runx1* are two *bona fide* marker genes for HE at 26 hours post fertilization (hpf) in zebrafish[18,41]. We quantified *cmyb* and *runx1* expression by measuring pixel intensity of the *in situ* hybridization staining compared to background[28]. Expression of *cmyb* (Figure 1E, F) and *runx1* (Figure 1G, H) was indistinguishable between WT and Gata2b deficient embryos at 26 hpf, indicating that specification of hemogenic endothelium occurs normally in the absence of Gata2b.

Next, we examined the ability of HE to undergo EHT using *Tg(fli1a:GFP)* reporter embryos, in which GFP marks all endothelial cells, including HE[29]. Consistent with our initial results, EHT events were not significantly reduced in *gata2b*$^{-/-}$ embryos compared to WT (p = 0.077, n= 18 WT and 18 *gata2b*$^{-/-}$ embryos, Figure 1I, J and Table S1). We conclude that neither HE specification, nor HSPC generation through EHT are impaired in *gata2b*$^{-/-}$ embryos.

Gata2a is required for HE specification

GATA2 is required for the generation of HSCs in mouse[7,16], but Gata2b deficient zebrafish have intact HE and EHT (Figure 1E-J). High maternal expression of *gata2b* has been reported previously[24], and therefore residual Gata2b protein levels could possibly rescue EHT. However, maternal zygotic *gata2b*$^{-/-}$ zebrafish, that do not contain functional maternally provided *gata2b* mRNA, are viable and survive to mendelian ratios (Figure S1C, D), indicating that maternal expression of *gata2b* does not contribute to embryonic hematopoiesis. By contrast, *gata2a* is expressed in hemogenic endothelium and regulates *runx1* expression in HE[25]. Thus, we analysed *runx1* expression in *gata2a* mutants (*gata2a*$^{i4/i4}$, lacking a conserved endothelial enhancer)[25]. *runx1* expression at 26 hpf was reduced in *gata2a*$^{i4/i4}$ embryos compared to WT embryos (Figure 1K, L). This confirmed that endothelial expression of *gata2a,* but not *gata2b,* is required for the specification of hemogenic endothelium and that the different functions of mammalian GATA2 are separated between Gata2a and Gata2b in zebrafish.

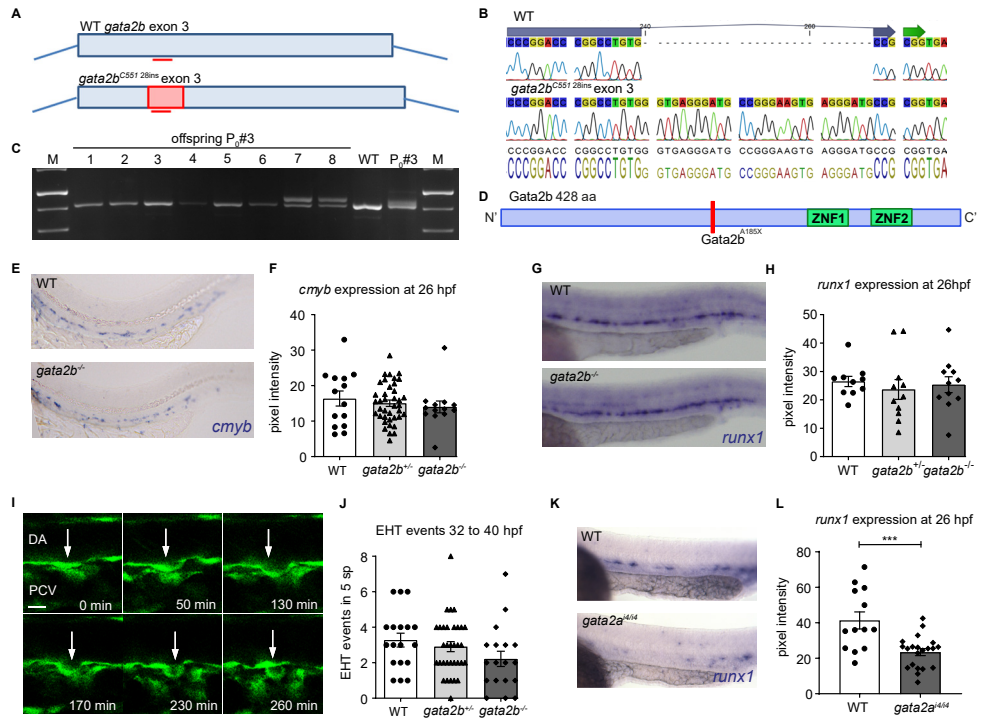


Figure 1. Newly generated Gata2b mutant does not show defects in HSPC generation.
A) Schematic representation of the CRISPR strategy targeting exon 3 of *gata2b* and the 28 nt integration in *gata2b* mutants. B) Alignment of sequencing data of WT *gata2b* exon 3, where the location of the guide is indicated in the blue arrow on top of the sequence and sequencing data from *gata2b*$^{-/-}$ DNA showing a 28 nucleotide integration. C) Gel picture showing genotyping PCR of founder 3 and the F1 with a 28 bp integration in embryo 7 and 8. D) Gata2b mutation leading to a STOP codon abrogating the protein before the DNA and protein binding znic fingers. E) Representative image of *cmyb* expression in WT and *gata2b*$^{-/-}$ embryos at 26 hpf. F) Quantitation of *cmyb* signal intensity relative to background in WT and *gata2b*$^{-/-}$ embryos at 26 hpf. G) Representative image of *runx1* expression in WT and *gata2b*$^{-/-}$ embryos at 26 hpf. H) Quantitation of *runx1* signal intensity relative to background in WT and *gata2b*$^{-/-}$ embryos at 26 hpf. I) Example of EHT event from WT *Tg(fli1a:eGFP)* transgenic zebrafish. Time indicated at the bottom right corner in minutes. Scale bar represents 10 µm. Arrow indicates endothelial cell undergoing hematopoietic transition. J) Quantitation of EHT events between 32-40 hpf in WT, *gata2b*$^{+/-}$ and *gata2b*$^{-/-}$ embryos. K) Representative example of *runx1* expression in WT and *gata2*$^{i4/i4}$ embryos at 26 hpf in the AGM region. L) Quantitation of signal intensity relative to background cells in WT and *gata2a*$^{i4/i4}$ embryos at 26 hpf, where each dot represents one embryo (41.4 ± 4.8 and 23.5 ± 2.0, n = 13 WT and n = 21 *gata2a*$^{i4/i4}$). *** = p < 0.001, error bars represent SEM. Bp = basepair, EHT = endothelial to hematopoietic transition, DA = dorsal aorta, PVC = posterior cardinal vein, sp = somite pair, hpf = hours port fertilization. Error bars represent standard error of mean (SEM).

Gata2b is required for the expansion of definitive HSPCs during the CHT amplification phase

The CHT is temporally and spatially analogous to mouse fetal liver where HSPCs undergo amplification[22]. We investigated whether the loss of Gata2b affects the number of definitive HSPCs in the CHT between 2 and 3 days post fertilization (dpf). From 44 hpf onward, definitive HSPCs are marked by co-expression of the *Tg*(*CD41:GFP*)[30] marker and the arterial *Tg(flt:RFP)*[31] marker as definitive HSPCs are derived from arteries[42]. CD41:GFP$^+$Flt:RFP$^+$ cell numbers were similar in WT and *gata2b*$^{-/-}$ embryos at 52-54 hpf (Figure 2A, B and Table S1) and 56-58 hpf (Figure 2C and Table S1). However, at 76 hpf, CD41:GFP$^+$Flt:RFP$^+$ cells were significantly reduced in the CHT in *gata2b*$^{-/-}$ embryos compared to WT (Figure 2D and Table S1).

The number of definitive HSPCs expands rapidly from 52 hpf to 76 hpf in WT embryos (6.5 fold); in *gata2b*$^{-/-}$ embryos that expansion was reduced (3.1 fold). To support our findings we investigated the expression of *cmyb,* which is a marker for proliferating HSPCs from 30 hpf[43,44]. A significant reduction in *cmyb* expression was detectable from 33 hpf onward in the AGM and CHT regions of *gata2b*$^{-/-}$ embryos compared to WT (Figure 2E-H and Table S1). This analysis detected a reduction in *cmyb* expression levels rather then quantifying HSPC numbers. However, because the number of CD41:GFP$^+$Flt:RFP$^+$ cells was not affected at 33 hpf, but *cmyb* expression was already reduced at 33 hpf, this suggests that proliferation of definitive HSPCs is affected, resulting in a reduction of definitive HSPCs at 76 hpf (Figure 2D). To test this, proliferation was assessed by flow cytometry of CD41:GFP$^+$ cells at 75 hpf in WT and *gata2b*$^{-/-}$ embryos. This analysis shows that Gata2b deficient CD41:GFP$^+$ cells have an increased proportion of cells in the G_0 phase of cell cycle explaining the reduction of HSPCs at 3 dpf (Figure 2I-K). At 5 dpf the difference in proliferation is no longer detectable although HSPC numbers are still reduced (Figure 2I and data not shown).

Single cell RNAseq identifies a lymphoid bias at the expense of myeloid lineage output in gata2b$^{-/-}$ *kidney marrow*

Because the functions of GATA2 are separated between Gata2a and Gata2b in zebrafish and Gata2b deficient zebrafish are viable, we can uniquely assess the function of Gata2b in adult hematopoiesis. To investigate the hematopoietic lineages in an unbiased manner and to assess the impact of Gata2b deficiency on the transcriptional profile of hematopoietic progenitors and differentiated cells, the progenitor population including lymphocytes from kidney marrow (KM) were isolated and processed for single-cell RNA sequencing (scRNA-seq)(Figure S2A and B). To enrich the scarce HSC population we used pooled KM from two WT and *gata2b*$^{-/-}$ *Tg(CD41:GFP)* zebrafish per sample and included all CD41:GFP^{low} expressing cells present in the kidney marrow pool as these cells were shown to contain transplantable HSCs[30] (Figure S2A and C). This resulted in a mild enrichment of phenotypic HSCs from 0.21-0.5% to 0.46 – 2.73% of CD41:GFP^{low} cells within the total progenitor population.

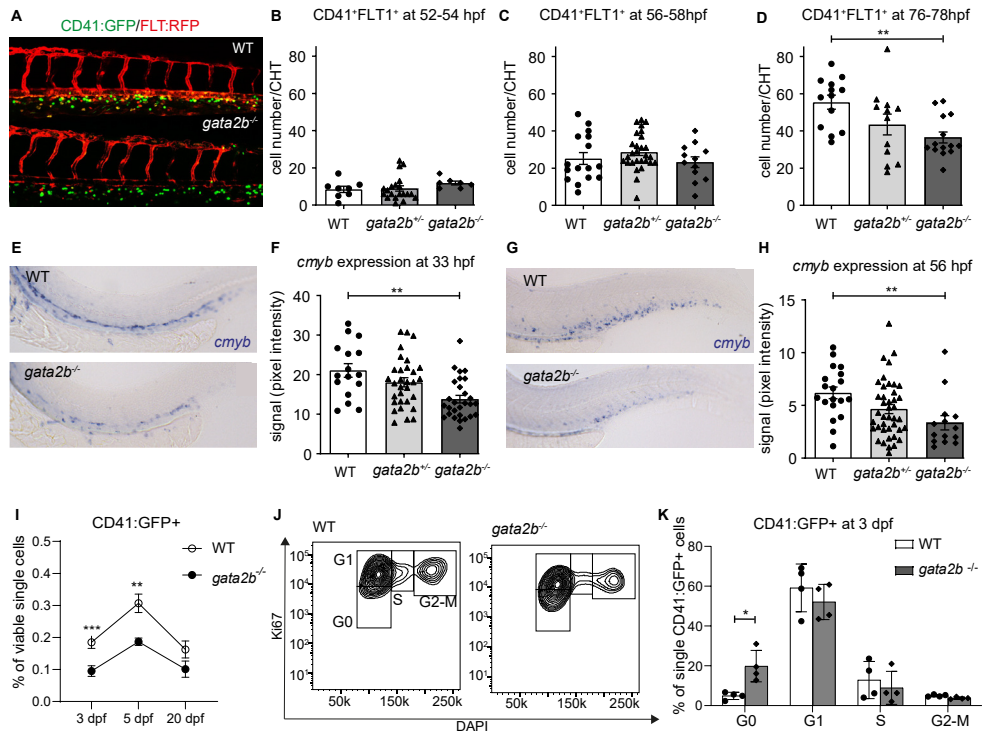


Figure 2. HSPC numbers are reduced in the CHT after 3 dpf in *gata2b*$^{-/-}$ embryos.
A) Example of *Tg(CD41:GFP); Tg(Flt:RFP)* expression in the CHT of WT and *gata2b*$^{-/-}$ embryos at 76 hpf. B-D) Quantitation of GFP$^+$RFP$^+$ cells in WT, *gata2b*$^{+/-}$ and *gata2b*$^{-/-}$ embryos at different time points where each dot represents one embryo. B) 52-54 hpf (8.50 ± 1.7 vs 11.86 ± 1.0), C) 56-58 hpf (25.19 ± 3.2 vs 23.25 ± 2.9), and D) 76-78 hpf (55.46 ± 3.8 vs 36.50 ± 3.0; P < 0.001). E) Representative example of *cmyb* expression in WT and *gata2b*$^{-/-}$ embryos at 33 hpf and F) Quantitation of signal intensity relative to background at 33 hpf where each dot represents one embryo. G) Representative example of *cmyb* expression in WT and *gata2b*$^{-/-}$ embryos at 56 hpf and H) Quantitation of signal intensity relative to background at 56 hpf where each dot represents one embryo. I) Quantitation of CD41:GFP$^+$ cell percentages by flow cytometry at 3 dpf, 5 dpf and 20 dpf. J) Cell cycle analysis by flow cytometry of Ki67 and DAPI staining of CD41:GFP$^+$ cells in WT and *gata2b*$^{-/-}$ embryos at 75 hpf. K) Bar graph representing the quantitation of cell cycle of CD41:GFP$^+$ cells in WT and *gata2b*$^{-/-}$ embryos at 75 hpf. Bars represent mean ± SEM with each dot indicating one pooled sample. hpf = hours post fertilization, dpf = days post fertilization, CHT = caudal hematopoietic tissue. * = p < 0.05, ** = p < 0.01. Error bars represent SEM.

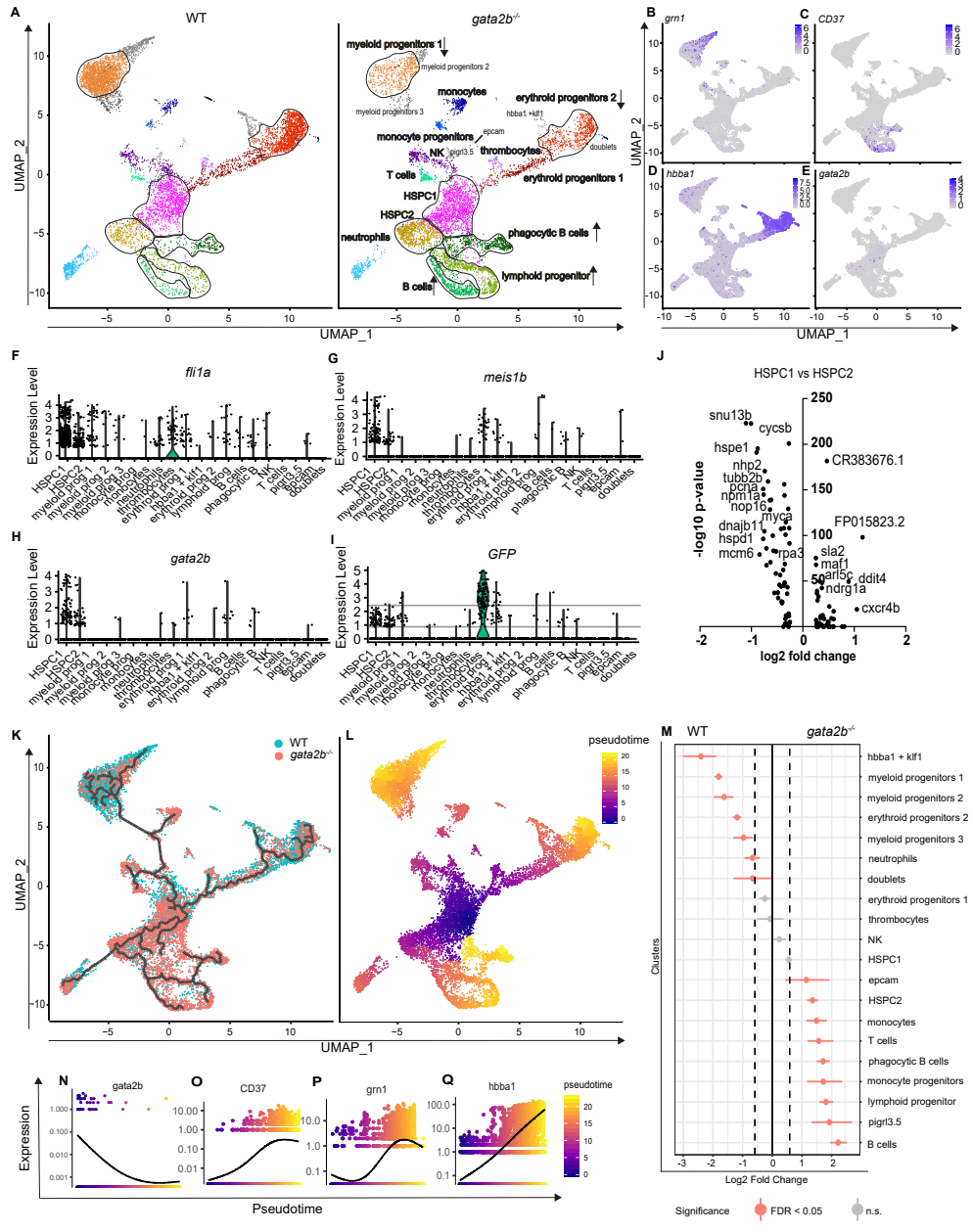
We identified 20 different cell clusters were identified using the nearest neighbor algorithm in the R Seurat package[36] (Figure 3A and S2G). Most progenitors that were sequenced expressed previously characterized differentiation markers[45-49] (Figure 3B-E and S2H). We identified 2 HSPC populations. These clusters are characterized by the robust expression of HSC genes, like *fli1a* and *meis1b*[48-50] (Figure 3F,G), *gata2b* (Figure 3H), concomitant with intermediate levels of GFP derived from the CD41:GFP transgene (Figure 3I), and low expression of differentiation markers (Figure 3B-E and S2H). Compared to HSPC2, HSPC1 exhibited a lower expression of metabolic and proliferation markers like *pcna* and

myca, suggesting that HSPC1 is more quiescent than HSPC2 (Figure 3J). Therefore, lineage trajectory analysis was started from this cluster identifying separate lineage differentiation trajectories for the erythroid-, myeloid- and lymphoid lineage (Figure 3K, L, N-Q).

Proportion analysis regarding the distribution of WT and Gata2b deficient cells between the lymphoid and myeloid lineages indicated a bias towards the lymphoid lineage in *gata2b*$^{-/-}$ cells at the expense of the myeloid lineage compared to WT (Figure 3A, K-M-and S2I and Table S2). The largest differences were observed in 3 clusters expressing high levels of *granulin 1 (grn1)*(Figure 3B) indicating that these clusters contains myeloid progenitors and were overrepresented by WT cells (Figure 3A, K, M and S2I). We defined these 3 clusters expressing myeloid specific genes with slight differences in their expression pattern as myeloid progenitors-1, -2 and -3 (Figure 3A and S2H).The *grn1* expressing cluster also showed high expression of *s100a10b,* a potential new marker for these cells (Figure S3A). Expression analysis on KM smears showed that *s100a10b* is expressed in the neutrophil lineage (Figure S3B). The B-cell clusters were overrepresented by *gata2b*$^{-/-}$ cells and showed very high expression of *immunoglobulin heavy variable 1-4* (*ighv1-4)*(Figure S2H), *CD37* (Figure 3C) and *pax5* (not shown), indicating that these were *bona fide* B-cell populations. Interestingly, we found a population of phagocytic B-cells previously identified in theleosts which express both *mpeg1.1* and B cell markers[51]. Lineage trajectory analysis indicates that these cells decend from lymphoid progenitors (Figure 3C, K, L). When pseudotime analyses was performed for WT and *gata2b*$^{-/-}$cells separately, the phagocytic B-cells did not only show a lineage differentiation trajectory from lymphoid progenitors, but also from the HSPC1 population, indicating a skewing in *gata2b*$^{-/-}$ HSPCs directly towards the lymphoid lineage (Figure S2J, K).

Figure 3. Single cell analysis reveals that *gata2b*$^{-/-}$ kidney marrow cells are overrepresented in lymphoid lineage clusters and reduced in erythroid and myeloid lineage clusters compared to WT.
A) Split UMAP of WT and *gata2b*$^{-/-}$ cells with cluster indication of enriched (arrow up) or reduced (arrow down) cell clusters in *gata2b*$^{-/-}$ cells. B-E) Pooled WT and *gata2b*$^{-/-}$ UMAP feature analysis with gradual gene expression in shades of blue. Expression pattern of B) *granulin1* (*grn1),* C) *cluster of differentiation 37* (*CD37),* D) *hemoglobin, beta adult (hbba1),* E) *GATA binding protein 2b* (*gata2b).* F-I) Violin plots representing the expression levels of genes within the different clusters and each dot represents expression in one cell. F) *fli-1 proto-oncogene* (*fli1a),* G) *meis homeobox 1b* (*meis1b),* H) *GATA binding protein 2b* (*gata2b),* I) *green fluorescent protein* (*GFP)*, indicating CD41:GFP^{low} cells. J) Volcano plot comparing HSPC1 vs HSPC2. At the left of the Y axis there are genes in HSPC1 with an average logarithmic fold change less than -0.25 and to the right are genes with a logarithmic fold change higher than 0.25 compared to HSPC2. K) Lineage differentiation trajectory depicted on UMAP with WT cells in blue and *gata2b*$^{-/-}$ cells in pink. L) Pseudotime analysis assuming HSPC1 as a starting point. M) quantitation of proportions of distribution between WT and *gata2b*$^{-/-}$ cells in the different clusters. Significant differences are indicated in pink. N-Q) pseudotime analysis of gene expression in lineage trajectory analysis of N) *gata2b,* O) *cluster of differentiation 37* (*CD37)*, P) *granulin1* (*grn1)* ans Q) *hemoglobin, beta adult (hbba1)*. UMAP; uniform manifold approximation projection.

Figure 3



Lack of Gata2b leads to reduced neutrophil numbers and increased lymphoid progenitors in adult kidney marrow

Because scRNA-seq analysis showed a major switch in lineage differentiation, we asked whether hematopoietic differentiation was affected in the adult *gata2b*$^{-/-}$ KM using scatter profile-, transgenic marker- and morphological analysis[30,32,33,52]. While *gata2b*$^{-/-}$ embryos did not show signs of altered lineage differentiation up to 5 dpf (Figure S4A-F), scatter profiles of adult *gata2b*$^{-/-}$ zebrafish KM showed a significant reduction in the myeloid population (Figure 4A, B and Table 1) and a relative increase in the scatter population containing HSPCs and lymphoid cells at 4 months post fertilization (mpf) and onward (Figure 4A, C and Table 1). This skewing in the population frequencies persisted with age (Figure 4B, C). To further address how the myeloid lineage was affected by the loss of Gata2b, *Tg(mpx:GFP)* expression, specifically marking neutrophils[33,53] and *Tg(mpeg1.1:GFP),* marking monocytes and phagocytic B-cells[32,54] was assessed. No significant difference was observed in mpeg:GFP^{+} cells between WT and *gata2b*$^{-/-}$ KM (Figure S4O-Q). *gata2b*$^{-/-}$ zebrafish showed a severe reduction in mpx:GFP^{+} neutrophils in the kidney marrow at 4 mpf (Figure 4D-F and Table 1). Sorted mpx:GFP^{+} cells from these zebrafish showed that the remaining *gata2b*$^{-/-}$ mpx:GFP^{+} cells did not reach WT levels of GFP, had a more immature neutrophil morphology and a block at the promyelocyte stage (Figure 4G, H and S3E), indicating that Gata2b is required for terminal neutrophil differentiation. This could be a result of the reduction in myeloid progenitors in the single cell data (Figure 3A and M).

Because GATA2 is also required for HSC maintenance in mice[7,12,13], we asked whether Gata2b deficiency resulted in a block in HSPC differentiation and thus an accumulation of HSPCs. In zebrafish, CD41:GFP^{low} expression marks the HSPC population most stringently[30]. Although CD41:GFP^{+} cell numbers and percentages were reduced during embryonic development (Figure 2D, I), at 20 dpf these percentages normalize resulting in comparable numbers of CD41:GFP^{low} cells during adulthood (Figure 2I and 4I-K) indicating that the accumulation of the population containing lymphoid cells and HSPCs in *gata2b*$^{-/-}$ KM is not due to a differentiation block in HSPCs, but due to an increase in lymphoid cells.

Tg(IgM:GFP), marking B-cells[35] and *Tg(lck:GFP)*, marking T-cells[34] were used to asses lymphoid differentiation. We did not find an increase in the lck:GFP^{+} population (Figure S4M, N and Table 1). However, in *Tg(IgM:GFP)* zebrafish we identified several populations of IgM:GFP^{+} cells with a significant increase in immature IgM:GFP^{+} cells (IgM:GFP3 fraction)(Figure 4L, N and S4G-M). We could classify the different IgM:GFP^{+} populations as lymphoblastic cells, lymphocytes, plasmacells and phagocytic B-cells[51,55](Figure 4O-S). In particular phagocytic B-cells were increased in *gata2b*$^{-/-}$ KM compared to WT, but mature plasmacells were significantly reduced (Figure 4S, $P<0.01$ and $P<0.001$). Although the majority of the cells in the lymphoid and HSPC population was not marked by known lymphoid lineage markers IgM:GFP or lck:GFP, we could still detect a significant increase in immature B-cells, confirming the increase in lymphoid output in KM in *gata2b*$^{-/-}$ zebrafish compared to WT.

Gene expression analysis reveals different HSPC populations in zebrafish

Next, we explored the molecular origin of the increase in lymphoid lineage output observed in Gata2b deficient zebrafish. Previous studies show that blocked neutrophil differentiation results in a shift towards monocytic lineage differentiation[56]. Our data shows indeed an increase in monocyte progenitors and monocytes (Figure 3M). However, we also detect a shift towards the lymphoid lineage, indicating that Gata2b is required for lineage programming in more immature progenitors. First, we tested if the lymphoid lineage bias was detectable in the HSPC clusters we identified as HSPC1 and HSPC2 (Figure 3A).

HSPC1 makes up 18 percent of the total analyzed kidney marrow population. Because HSCs are a rare population of cells based on transplantation studies, we hypothesized that HSPC1 also contains other progenitor cells. We subclustered HSPC1 to subdivide these progenitors (Figure 5A,B). In this way, clustering is not based on gene expression differences found in comparison to more committed cells, but only based on gene expression differences within the HSPC1 population. Eight subclusters were identified based on differential gene expression analysis (Figure S5A). These subclusters were classified as a quiescent subcluster with very low gene expression, an HSC subcluster with expression of *meis1b* and *fli1a*, three myeloid subclusters, one lymphoid subcluster, a proliferative subcluster and an undefined subcluster (Figure 5B, E, F and S5A). Interestingly, proportion analysis of WT and *gata2b*$^{-/-}$ cells showed that Gata2b deficient cells almost entirely lost the quiescent subcluster and gained a proliferative subcluster, and myeloid subcluster 3 (Figure 5C, D). A differential expression analysis of the whole HSPC1 cluster revealed a downregulation of myeloid genes like *s100a10b*, *grn1*, *csf3b* and *cepba* in *gata2b*$^{-/-}$ HSPC1 (Figure 5G) but not a clear upregulation of lymphoid genes. When the same comparison was done in HSPC2 cells (Figure S6A, B), we detected a larger reduction in the myeloid gene expression program in the entire HSPC2 cluster (Figure S6C, D) and found that *gata2b*$^{-/-}$ cells expressed higher levels of lymphoid genes like *ikzf1, fcer1gl, ighv1-4, ccr9a* and *xbp1* (Figure S6E-J). Psuedotime analysis showed the differentiation trajectories within the HSPC1 cluster when started from the quiescent subcluster, containing most CD41:GFP expressing cells (Figure 3E, F and H, I). When inferring gene expression in psuedotime analysis of HSPC1, we found that *gata2b* and *meis1* expression are highest in the quiescent population and decrease during differentiation (Figure 5J-L). Interestingly, *gata2a* expression does not overlap with *gata2b* indicating that *gata2a* does not compensate for the loss of *gata2b* in Gata2b deficient HSPCs (Figure 5K, M). As HSPCs become more mature, they first upregulate proliferation markers like *pcna* and *mki67* both highly expressed in the *gata2b*$^{-/-}$ unique subcluster (Figure 5N, O). Assessing proliferation by flow cytometry of CD41:GFP^{low} cells in WT and *gata2b*$^{-/-}$ KM, we found that *gata2b*$^{-/-}$ CD41:GFP^{low} cells have increased numbers of cells in S phase, indicating that *gata2b*$^{-/-}$ zebrafish HSPCs are more proliferative (Figure 5P-R). Together these data show that the absence of Gata2b leads to transcriptional changes in the HSPC compartment concomitant with a shift in lineage output from the myeloid lineage towards the lymphoid lineage.

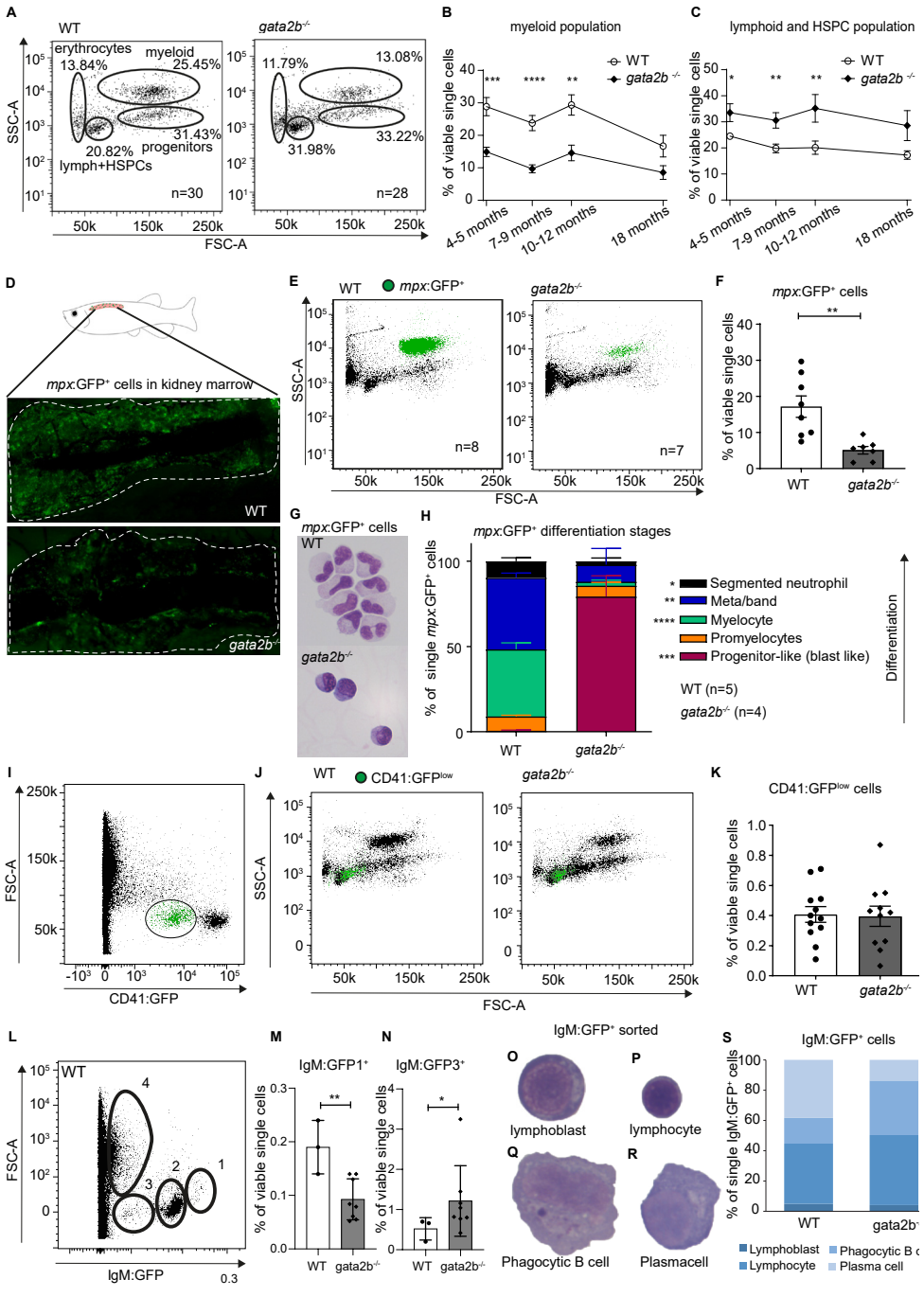


Figure 4. Gata2b deficiency results in decreased myeloid differentiation in adult zebrafish kidney marrow.
A) Gating strategy of FACS analysis of whole kidney marrow of WT and *gata2b*$^{-/-}$ zebrafish. Percentages represent the average of all zebrafish analyzed per genotype. B-C) Quantitation as percentages of single viable cells over time of B) myeloid and C) lymphoid and HSPC populations. D) Representative example of *Tg(mpx:GFP)* expression in WT and *gata2b*$^{-/-}$ zebrafish kidney marrow by fluorescence microscopy. E) Forward and side scatter profile of *Tg(mpx:GFP)* expression in WT and *gata2b*$^{-/-}$ zebrafish kidney marrow in green. F) Quantitation of *Tg(mpx:GFP)*$^{+}$ cells expressed as percentage in single viable cells. Each dot represents kidney marrow analysis of one zebrafish. G) Representative figure of sorted *Tg(mpx:GFP)*$^{+}$ cells from WT and *gata2b*$^{-/-}$ zebrafish kidney marrow after MGG staining. H) Quantification of *Tg(mpx:GFP)*$^{+}$ cells from WT and *gata2b*$^{-/-}$ zebrafish kidney marrow based on the differentiation phenotype using MGG staining. I) Gating strategy for CD41:GFP^{low} expressing cells in total kidney marrow in green. J) Forward- and side-scatter plot of WT and *gata2b*$^{-/-}$ kidney marrow cells and CD41:GFP^{low} expressing cells in green. K) Quantification of the frequency of CD41:GFP^{low} cells in single live cells of total kidney marrow. Each dot represents kidney marrow analysis of one zebrafish. L) FSC/GFP scatter profile of *Tg(IgM:GFP)* WT KM. M) Quantitation of Gating 1 of *Tg(IgM:GFP)* WT and *gata2b*$^{-/-}$ KM as percentage of single viable cells. Each dot represents kidney marrow analysis of one zebrafish. N) Quantitation of Gating 3 of *Tg(IgM:GFP)* WT and *gata2b*$^{-/-}$ KM as percentage of single viable cells. Each dot represents kidney marrow analysis of one zebrafish. O-R) representative image of sorted IgM:GFP^{+} cells indicating O) lymphoplastic cell, P) lymphocyte, Q) phagocytic B cell and R) plasmacell. S) quantitation of sorted IgM:GFP^{+} cells per genotype. SSC = side scatter, FSC = forward scatter, KM = kidney marrow. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Error bars represent SEM.

Differential gene expression analysis reveals decreased myeloid marker expression in gata2b$^{-/-}$ *HSPCs and aberrant co-expression of myeloid and lymphoid genes.*

Overall, the expression of myeloid genes in *gata2b*$^{-/-}$ HSPC1 is reduced, but the percentage of *gata2b*$^{-/-}$ HSPC1s with detectable expression of myeloid genes such as *grn1* was increased (Figure 5G and 6A). This apparent contradiction was clarified by an overall transcript upregulation (Figure 6A, *actinb1* expression), indicative of a loss of quiescence. It is known that HSPCs can co-express myeloid and lymphoid genes before lineage decision[57,58]. While WT cells had a clear dichotomy in expression of myeloid and lymphoid genes, *gata2b*$^{-/-}$ HSPCs had a higher fraction of cells co-expressing lymphoid and myeloid genes (Figure 6B-E). For example, increased co-expression of a phagocytic B-cell marker, *igic1s1*, could be detected in *gata2b*$^{-/-}$ HSPC1 cells together with the myeloid marker *cebpb* (Figure 6B). This result suggests that the loss of Gata2b does not halt HSPC differentiation but re-directs this towards another lineage. Interestingly, when we infer psuedotime analysis of only WT and only *gata2b*$^{-/-}$ cells, this is exactly as we find. In *gata2b*$^{-/-}$ cells, phagocytic B cells can be formed from both lymphoid progenitors, as well as HSPC1 cells as opposed to WT phagocytic B cells (Figure S2J, K). Based on this data we conclude that the lymphoid bias in *gata2b*$^{-/-}$ zebrafish kidney marrow initiated in the most immature HSPC population. This is due to a failure to elicit proper expression of the myeloid differentiation program and concomitant upregulation of the lymphoid program, that redirects HSPCs towards a lymphoid fate.

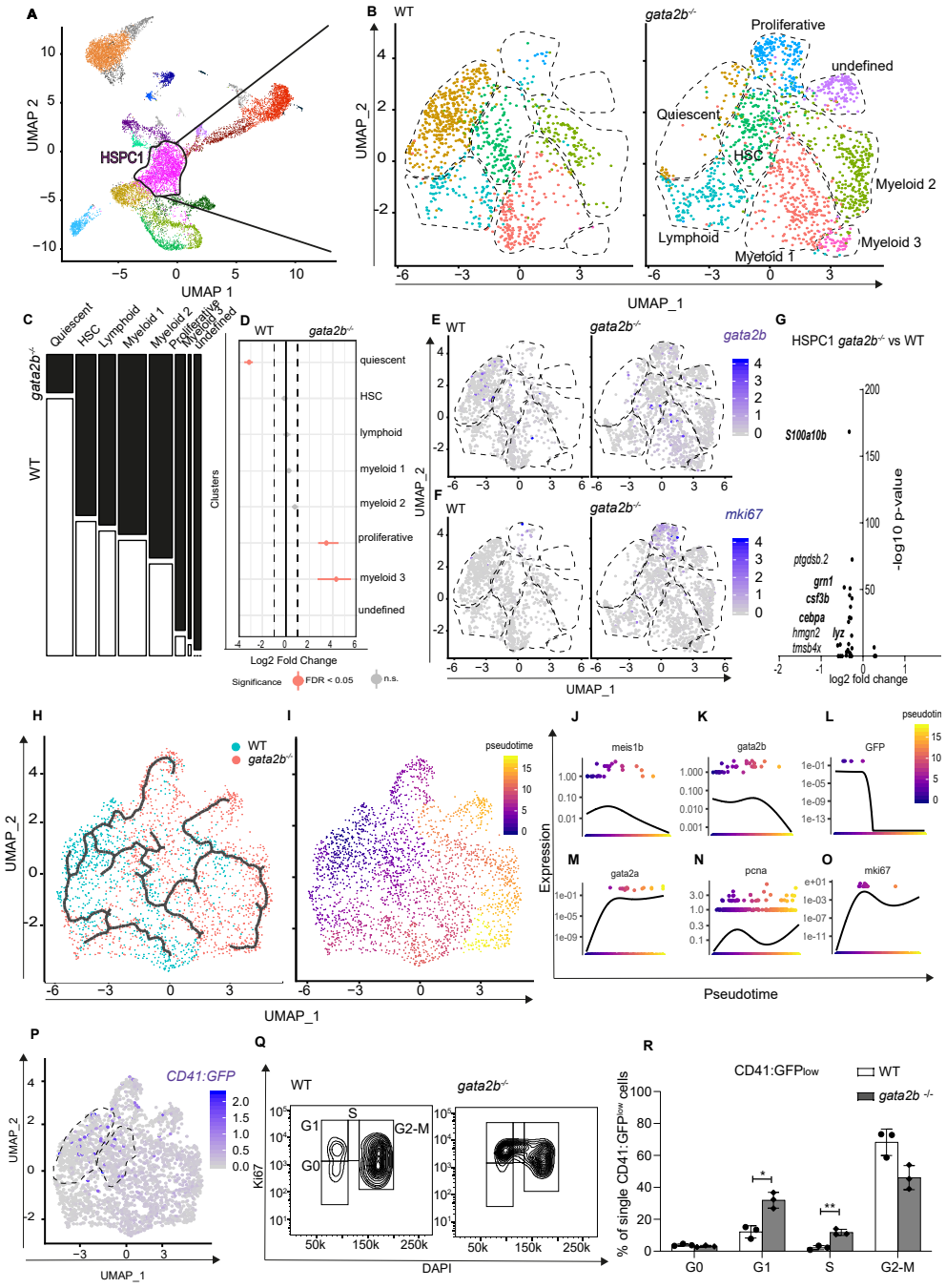


Figure 5. Subclustering of HSPC1 demonstrates the loss of a *gata2b* expressing quiescent subcluster and the appearance of a proliferative subcluster.

A) Cluster selection for subclustering. B) Reclustering of the HSPC1 population split between WT and *gata2b*$^{-/-}$cells. C) Genotype distribution of each of the clusters with WT cells in white and *gata2b*$^{-/-}$cells in black. D) quantitation of proportions of distribution between WT and *gata2b*$^{-/-}$ cells in the different clusters. Significant differences are indicated in pink. E-F) WT and *gata2b*$^{-/-}$ feature analysis with gradual gene expression in shades of blue within HSPC1 cells of D) *gata2b* E) *mki67*. G) Volcano plot comparing HSPC1 *gata2b*$^{-/-}$ vs WT. At the left of the Y axis gene expression in *gata2b*$^{-/-}$ HSPC1s with an average logarithmic fold change less than -0.25 and to the right gene expression with a logarithmic fold change higher than 0.25 compared to WT HSPC1s. Each dot represents a gene. H) Lineage differentiation trajectory depicted on UMAP with WT cells in blue and *gata2b*$^{-/-}$ cells in pink. I) Pseudotime analysis assuming the quiescent population as starting point. J-O) gene expression analysis on pseudotime analysis with J) *meis1b*, K) *gata2b* L) *GFP,* M) *gata2a,* N) *pcna*, O) *ki67*. P) WT and *gata2b*$^{-/-}$ feature analysis with gradual gene expression of *GFP* in shades of blue within HSPC1 cells. Dotted circles indicate the quiescent and HSC subcluster. Q) Cell cycle analysis by flow cytometry of Ki67 and DAPI staining of CD41:GFP^{low} cells in adult WT and *gata2b*$^{-/-}$ KM cells. R) Bar graph representing the quantitation of cell cycle of CD41:GFP^{low} cells in adult WT and *gata2b*$^{-/-}$ KM cells. Bars represent mean ± SEM, each dot indicates analysis from one zebrafish. UMAP; uniform manifold approximation projection.

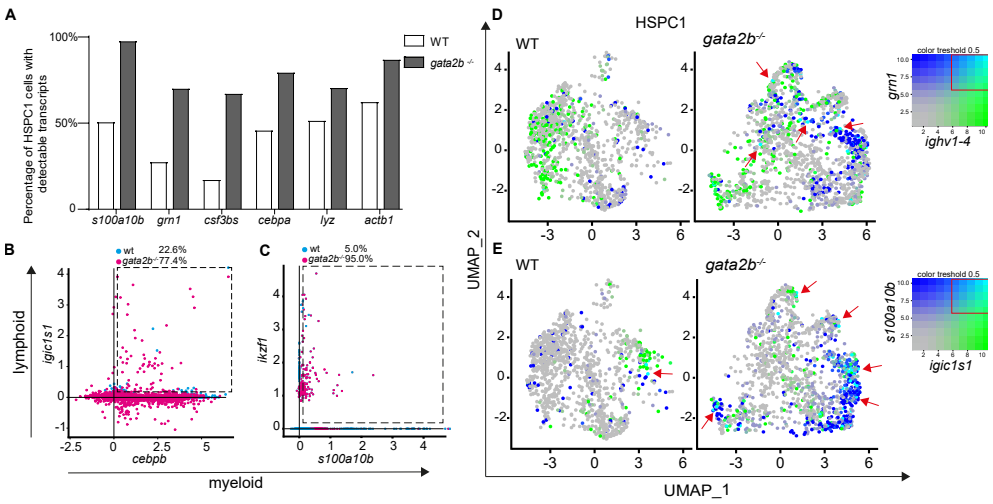


Figure 6. Gata2b deficient HSPC1 show co-expression of myeloid and lymphoid gene expression programs
A) Bar graph representing the percentage of WT HSPC1 cells in white and *gata2b*$^{-/-}$ HSPC1 cells in grey with at least one read of *s100a10b*, *grn1*, *colony stimulating factor 3b (*csf3bs), *CCAAT enhancer binding protein alpha (cebpa), lysozyme (lyz)* or *actin beta1 (actb1),* indicating more cells with detectable myeloid gene expression in Gata2b deficient HSPC1s. B) Co-expression analysis of the lymphoid gene *igic1s1* with the myeloid gene *cebpb* and C) the lymphoid gene *IKAROS Family Zinc Finger 1* (*ikzf1*) with the myeloid gene *s100a10b*. Values represent percentages of WT and *gata2b*$^{-/-}$ HSPC1 cells co-expressing myeloid and lymphoid genes (within the dashed box). D) WT and *gata2b*$^{-/-}$ feature analysis representing co-expression analysis of the lymphoid gene *Immunoglobulin heavy variable 1-4* (*ighv1-4*) with the myeloid gene *grn1* and E) the phagocytic B cell marker immunoglobulin light iota constant 1-s1 (*igic1s1*) with the myeloid gene *s100a10b* with myeloid genes in blue and lymphoid genes in green. Co-expression of the myeloid and lymphoid genes is represented in turquoise indicated by red arrows in the WT and *gata2b*$^{-/-}$ feature analysis. Coloring threshold set in quantiles, min.cutoff= q10, max.cutoff= q90.

DISCUSSION

In this study, we showed that the function of mammalian GATA2 in zebrafish is split between Gata2a and Gata2b. Gata2a is required for HE specification upstream of Gata2b. Gata2b is not vital for embryonic generation of HSPCs, but supports their expansion in the caudal hematopoietic tissue. However, during adulthood, Gata2b is required for the quiescent HSPC population and in its absence HSPCs are more proliferative. In addition, Gata2b deficient kidney marrow from adult zebrafish showed a lymphoid bias at the expense of the myeloid lineage based on scatter profiles and transgenic marker analyses. Single cell transcriptome analysis showed that the stem and progenitor cells were the origin of the increased lymphoid lineage output in *gata2b*$^{-/-}$ kidney marrow cells, due to a failure to increase myeloid gene expression to sufficient levels and a subsequent co-expression of both myeloid and lymphoid genes in *gata2b*$^{-/-}$ HSPCs. These data establish that Gata2b is vital for maintaining the myeloid differentiation program while restricting lymphoid differentiation.

The molecular mechanism controlling lineage commitment has long been thought to be regulated by stochastic variations in the levels of transcription factors, and progenitors are committed to a lineage choice[59]. However, later reports suggested that some transcription factors have a reinforcing activity for terminal differentiation and propose that microenvironmental or upstream regulators are decisive for lineage commitment[60]. This would suggest that when these reinforcing factors are removed, cells can redirect their lineage. Our results are consistent with Gata2b being required for stemness of HSCs. Single cell transcriptome analysis showed a unique cluster of Gata2b deficient cells with upregulation of genes related to proliferation, suggestive of a role for Gata2b in cell cycle adaptation. The quiescent subcluster was almost entirely lost and the the CD41:GFP^{low} population showed increased proliferation. Loss of quiescence in HSPCs then increases the expression of commitment genes resulting in cells co-expressing lymphoid and myeloid lineage markers as detected in *gata2b*$^{-/-}$ HSPCs and Gata2b is therefore an essential cell-intrinsic regulator of lineage output in HSPCs.

In mouse, GATA2 is also required for the maintenance of HSCs after they are generated[7]. During embryonic hematopoiesis the number and percentage of HSPCs is reduced due to reduced proliferation, but during adult stages Gata2b deficient HSCs as marked by CD41:GFP^{low} expression are not reduced and proliferation is increased, probably responsible for the normalization in HSC numbers (Figure 4I-K and 5Q,R). We do not find upregulation of *gata2a* in these cells as a rescue mechanism (Figure 5M). Single cell transcriptome analysis identifies several HSPC populations with unique transcriptional signatures. Interestingly, the CD41:GFP^{low} expressing cells were scattered among different HSPC populations. Transplantation data suggest that only a minority of these cells are *bona fide* HSCs[30,61]. Because zebrafish are outbred, limiting dilution transplantation studies result in a gross underestimation of actual HSC numbers. This indicates that further research could provide

us with a more stringent marker for HSCs in zebrafish. In Gata2b deficient HSPC1s, the quiescent HSPC population is absent (Figure 5B-D). This could represent the true quiescent HSC population. Interestingly, this does not affect survival of the zebrafish.

Not all myeloid lineage differentiation was abrogated in Gata2b deficient zebrafish and few intact neutrophils remained present. Also, the monocyte progenitor and -cluster, marked by *mpeg1.1* were present in Gata2b deficient KM. Previous studies found that if neutrophil development is blocked, myeloid differentiation progresses towards to monocytic lineage[56]. Besides an increase in monocytic progenitors, we also detect a redirection of lineage differentiation at a much earlier state leading to increases in B-cell populations (Figure 3M, S2J, K). This indicates that in Gata2b deficient HSPCs, a reprogramming occurs both in immature cells to delineate lineage differentiation towards the lymphoid lineage, but also in the myeloid lineage to redirect the lineage to monocytes, again indicating separate functions for Gata2 in lineage differentiation. Interestingly, the number and percentage of plasmacells was reduced (Figure 4S). Together with the severe neutropenia, this is very similar to patients sufferening from MonoMAC syndrome, which is characterized by neutropenia, monocytopenia, DC- and B-cell lymphopenia[62-64]. These syndromes are caused by haploinsufficiency of the GATA2 transcription factor. Despite the severe neutropenia, no infections were observed in Gata2b deficient zebrafish probably due to the SPF conditions of the animal facility.

In conclusion, we find that Gata2b is required for proliferation of the HSPC pool in the CHT and is vital for myeloid lineage differentiation in the adult, both in the HSPC compartment and for terminal differentiation. Loss of Gata2b consequently induces a differentiation diversion towards the lymphoid lineage.

Acknowledgements

We thank members of the de Pater, Touw, Raaijmakers and Schneider-Kramann labs for helpful discussions. We thank Dr Cupedo and Prof. Dr. Dzierzak (University of Edinburgh) for careful reading of the manuscript. We thank the Experimental Animal Facility of Erasmus MC for animal husbandry and the Erasmus Optical Imaging Center for confocal microscopy services. This research is supported by the European Hematology Association (junior non clinical research fellowship)(EdP), the Dutch Cancer Foundation KWF/Alpe d'HuZes (SK10321)(EdP), the British Heart Foundation (BHF IBSR Fellowship FS/13/50/30436 to RM, CBM), the Wellcome Trust (PhD Scholarship #WT102345/Z/13/Z. to TD), the Daniel den Hoed Foundation for support of the Cancer Genome Editing Center (IT) and the Josephine Nefkens Foundation for purchase of the Chromium 10x (IT). Graphical abstract is created with BioRender.com.

Authorship contributions

EdP, EG and CK conceived the study; EG, CK, HdL, JZ, DB, TD, CBM, PvS, MvR, and EB performed experiments; EG, CK, MdJ, RH, CBM, RM, KG and EdP analysed results; RM, PF and IT provided resources and EG, CK and EdP wrote the manuscript and IT revised the manuscript.

Disclosures

The authors declare no conflicts of interests

REFERENCES

1. Sawai CM, Babovic S, Upadhaya S, et al. Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity*. 2016;45(3):597-609.
2. Yamamoto R, Morita Y, Ooehara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*. 2013;154(5):1112-1126.
3. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013;502(7470):232-236.
4. Muller-Sieburg CE, Cho RH, Karlsson L, Huang JF, Sieburg HB. Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood*. 2004;103(11):4111-4118.
5. Dykstra B, Kent D, Bowie M, et al. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell*. 2007;1(2):218-229.
6. Watcham S, Kucinski I, Gottgens B. New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood*. 2019;133(13):1415-1426.
7. de Pater E, Kaimakis P, Vink CS, et al. Gata2 is required for HSC generation and survival. *J Exp Med*. 2013;210(13):2843-2850.
8. Gao X, Johnson KD, Chang YI, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med*. 2013;210(13):2833-2842.
9. Snow JW, Trowbridge JJ, Fujiwara T, et al. A single cis element maintains repression of the key developmental regulator Gata2. *PLoS Genet*. 2010;6(9):e1001103.
10. Mehta C, Johnson KD, Gao X, et al. Integrating Enhancer Mechanisms to Establish a Hierarchical Blood Development Program. *Cell Rep*. 2017;20(12):2966-2979.
11. Johnson KD, Kong G, Gao X, et al. Cis-regulatory mechanisms governing stem and progenitor cell transitions. *Sci Adv*. 2015;1(8):e1500503.
12. Ling KW, Ottersbach K, van Hamburg JP, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med*. 2004;200(7):871-882.
13. Rodrigues NP, Janzen V, Forkert R, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*. 2005;106(2):477-484.
14. Rodrigues NP, Boyd AS, Fugazza C, et al. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood*. 2008;112(13):4862-4873.
15. Nandakumar SK, Johnson K, Throm SL, Pestina TI, Neale G, Persons DA. Low-level GATA2 overexpression promotes myeloid progenitor self-renewal and blocks lymphoid differentiation in mice. *Exp Hematol*. 2015;43(7):565-577 e561-510.
16. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371(6494):221-226.
17. Bertrand JY, Cisson JL, Stachura DL, Traver D. Notch signaling distinguishes 2 waves of definitive hematopoiesis in the zebrafish embryo. *Blood*. 2010;115(14):2777-2783.
18. Bertrand JY, Chi NC, Santoso B, Teng S, Stainier DY, Traver D. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*. 2010;464(7285):108-111.
19. Kissa K, Herbomel P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature*. 2010;464(7285):112-115.
20. Warga RM, Kane DA, Ho RK. Fate mapping embryonic blood in zebrafish: multi- and unipotential lineages are segregated at gastrulation. *Dev Cell*. 2009;16(5):744-755.
21. Ciau-Uitz A, Monteiro R, Kirmizitas A, Patient R. Developmental hematopoiesis: ontogeny, genetic programming and conservation. *Exp Hematol*. 2014;42(8):669-683.
22. Tamplin OJ, Durand EM, Carr LA, et al. Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell*. 2015;160(1-2):241-252.
23. Davidson AJ, Zon LI. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene*. 2004;23(43):7233-7246.
24. Butko E, Distel M, Pouget C, et al. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development*. 2015;142(6):1050-1061.
25. Dobrzycki T, Mahony CB, Krecsmarik M, et al. Deletion of a conserved Gata2 enhancer impairs haemogenic endothelium programming and adult Zebrafish haematopoiesis. *Commun Biol*. 2020;3(1):71.
26. Gagnon JA, Valen E, Thyme SB, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One*. 2014;9(5):e98186.
27. Chocron S, Verhoeven MC, Rentzsch F, Hammerschmidt M, Bakkers J. Zebrafish Bmp4 regulates left-right asymmetry at two distinct developmental time points. *Dev Biol*. 2007;305(2):577-588.
28. Dobrzycki T, Krecsmarik M, Bonkhofer F, Patient R, Monteiro R. An optimised pipeline for parallel image-based quantification of gene expression and genotyping after in situ hybridisation. *Biol Open*. 2018;7(4).
29. Lawson ND, Weinstein BM. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol*. 2002;248(2):307-318.
30. Ma D, Zhang J, Lin HF, Italiano J, Handin RI. The identification and characterization of zebrafish hematopoietic stem cells. *Blood*. 2011;118(2):289-297.
31. Bussmann J, Bos FL, Urasaki A, Kawakami K, Duckers HJ, Schulte-Merker S. Arteries provide essential guidance cues for lymphatic endothelial cells in the zebrafish trunk. *Development*. 2010;137(16):2653-2657.
32. Ellett F, Pase L, Hayman JW, Andrianopoulos A, Lieschke GJ. mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood*. 2011;117(4):e49-56.
33. Renshaw SA, Loynes CA, Trushell DM, Elworthy S, Ingham PW, Whyte MK. A transgenic zebrafish model of neutrophilic inflammation. *Blood*. 2006;108(13):3976-3978.
34. Langenau DM, Ferrando AA, Traver D, et al. In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. *Proc Natl Acad Sci U S A*. 2004;101(19):7369-7374.
35. Page DM, Wittamer V, Bertrand JY, et al. An evolutionarily conserved program of B-cell development and activation in zebrafish. *Blood*. 2013;122(8):e1-11.
36. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411-420.
37. Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc Natl Acad Sci U S A*. 2003;100(15):8811-8816.
38. Medvinsky A, Dzierzak E. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*. 1996;86(6):897-906.
39. Jaffredo T, Gautier R, Eichmann A, Dieterlen-Lièvre F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development*. 1998;125(22):4575-4583.
40. Boisset JC, van Cappellen W, Andrieu-Soler C, Galjart N, Dzierzak E, Robin C. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*. 2010;464(7285):116-120.
41. Bertrand JY, Kim AD, Teng S, Traver D. CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development*. 2008;135(10):1853-1862.

42. Bonkhofer F, Rispoli R, Pinheiro P, et al. Blood stem cell-forming haemogenic endothelium in zebrafish derives from arterial endothelium. *Nat Commun*. 2019;10(1):3577.
43. Bein K, Husain M, Ware JA, Mucenski ML, Rosenberg RD, Simons M. c-Myb function in fibroblasts. *J Cell Physiol*. 1997;173(3):319-326.
44. North TE, Goessling W, Walkley CR, et al. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature*. 2007;447(7147):1007-1011.
45. Carmona SJ, Teichmann SA, Ferreira L, et al. Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome Res*. 2017;27(3):451-461.
46. Danilova N, Bussmann J, Jekosch K, Steiner LA. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. *Nat Immunol*. 2005;6(3):295-302.
47. Kortum AN, Rodriguez-Nunez I, Yang J, et al. Differential expression and ligand binding indicate alternative functions for zebrafish polymeric immunoglobulin receptor (pIgR) and a family of pIgR-like (PIGRL) proteins. *Immunogenetics*. 2014;66(4):267-279.
48. Macaulay IC, Svensson V, Labalette C, et al. Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. *Cell Rep*. 2016;14(4):966-977.
49. Tang Q, Iyer S, Lobbardi R, et al. Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. *J Exp Med*. 2017;214(10):2875-2887.
50. Athanasiadis EI, Botthof JG, Andres H, Ferreira L, Lio P, Cvejic A. Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat Commun*. 2017;8(1):2045.
51. Li J, Barreda DR, Zhang YA, et al. B lymphocytes from early vertebrates have potent phagocytic and microbicidal abilities. *Nat Immunol*. 2006;7(10):1116-1124.
52. Traver D, Paw BH, Poss KD, Penberthy WT, Lin S, Zon LI. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol*. 2003;4(12):1238-1246.
53. Bennett CM, Kanki JP, Rhodes J, et al. Myelopoiesis in the zebrafish, Danio rerio. *Blood*. 2001;98(3):643-651.
54. Ferrero G, Gomez E, Lyer S, et al. The macrophage-expressed gene (mpeg) 1 identifies a subpopulation of B cells in the adult zebrafish. *J Leukoc Biol*. 2020;107(3):431-443.
55. Moss LD, Monette MM, Jaso-Friedmann L, et al. Identification of phagocytic cells, NK-like cytotoxic cell activity and the production of cellular exudates in the coelomic cavity of adult zebrafish. *Dev Comp Immunol*. 2009;33(10):1077-1087.
56. Olsson A, Venkatasubramanian M, Chaudhri VK, et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature*. 2016;537(7622):698-702.
57. Drissen R, Buza-Vidas N, Woll P, et al. Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat Immunol*. 2016;17(6):666-676.
58. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*. 2017;19(4):271-281.
59. Graf T, Enver T. Forcing cells to change lineages. *Nature*. 2009;462(7273):587-594.
60. Hoppe PS, Schwarzfischer M, Loeffler D, et al. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature*. 2016;535(7611):299-302.
61. Lin HF, Traver D, Zhu H, et al. Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood*. 2005;106(12):3803-3810.
62. Vinh DC, Patel SY, Uzel G, et al. Autosomal dominant and sporadic monocytopenia with susceptibility to mycobacteria, fungi, papillomaviruses, and myelodysplasia. *Blood*. 2010;115(8):1519-1529.
63. Hsu AP, Sampaio EP, Khan J, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*. 2011;118(10):2653-2655.

64. Hahn CN, Chong CE, Carmichael CL, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet*. 2011;43(10):1012-1017.

	WT	***gata2b*$^{-/-}$**
Scatter analysis in freq. of single viable cells	**N = 26**	**N = 23**
Erythrocytes in KM (3 to 12 months)	25.4 ± 2.8	25.9 ± 3.0
Progenitors in KM (3 to 12 months)	26.8 ± 2.0	24.4 ± 1.9
Lymphocytes and HSCs in KM (3 to 12 months)	18.1 ± 1.1	29.0 ± 1.9****
Myeloid cells in KM (3 to 12 months)	20.3 ± 1.4	10.3 ± 0.9****
Reporter expression in freq. of single viable cells		
Tg(*mpx*:GFP) (GFP+ in single live cells)	17.2 ± 3.0 (n=8)	5.1 ± 1.0 (n=7)**
Tg(*mpeg*:GFP)^{gl22Tg} (GFP+ in single live cells)	3.5 ± 1.9 (n=4)	2.1 ± 0.7(n=6)
Tg(*LCK*:GFP) (GFP+ in single live cells)	3.1 ± 0.8 (n =4)	3.4 ± 0.6 (n=4)
Tg(CD41 :GFP) (GFP low in single live cells)	0.4 ± 0.1 (n=9)	0.4 ± 0.1 (n=8)

Table 1. Adult hematopoietic cell quantitations
The data are mean± SEM. KM; kidney marrow, n= number of zebrafish used in analysis. *P< 0.05, **P< 0.01, ***P<0.001, ****P<0.0001. If data are normally distributed we used One-way ANOVA with Tukey post-test. If data are not normally distributed we used Kruskal-Wallis with Dunn's post-test.

SUPPLEMENTARY METHODS

Single cell transcriptome analysis of the progenitor compartment of WT and gata2b$^{-/-}$ zebrafish allows for unbiased lineage investigation

Kidney marrow from two 5 mpf female *Tg(CD41:GFP)* WT or *gata2b*$^{-/-}$ zebrafish were pooled and sorted and experiment was performed in replicate (Figure S2A-C). In total 70.000 cells from the gate in panel B were sorted and between 214 and 1607 CD41:GFP^{low} cells were added to this pool in PBS/10%FSC/2%BSA/2% carp serum. This resulted in final CD41:GFP^{low} percentage of 0.45 – 2.73%. The sorting strategy also included the CD41:GFP^{high} expressing cells which were previously identified as thrombocytes [30] (Figure S2C) so for every CD41:GFP^{low} cell, 1.9 CD41:GFP^{high} cells were present in our final population. 7630 cells in WT1, 4033 in WT2, 3675 in *gata2b*$^{-/-}$ 1 and 5229 in *gata2b*$^{-/-}$ 2 were obtained after quality control (Figure S2D-F), with a read depth of approximately 50.000 reads per cell. Gross differences in cell numbers between WT and *gata2b*$^{-/-}$ cells may influence cluster identification when a nearest-neighbor algorithm is used. Therefore, all replicates were randomly down-sampled to 3675 to match each other. First the replicates were aligned using anchor based integration and then the WT and *gata2b*$^{-/-}$ samples were aligned using the same method to correctly identify the clusters, avoiding batch specific differences. We could identify 20 different cell clusters using the R Seurat package[36](Figure S2G).

Cluster identification

Using the FindMarkers function in Seurat, differentially expressed genes were identified compared to the other clusters. Subsequently the functions FeaturePlot and VlnPlot were used to analyze gene expression patterns between clusters to test validity and exclusivity of individual clusters. Finally, marker analysis was visualized using the DoHeatmap function. Importantly, known Gata2 target genes like *cebpb* (p = 0,0002) and *alas2* (p = 4.77E-53) were found significantly downregulated in gata2b mutants.

Erythroid cells, thrombocytes, neutrophils, monocytes and their progenitors were identified by comparing our data to known single cell expression analysis and known lineage markers[45,48-50,65]. Canonical lineage differentiation markers are generally low expressing transcription factors and are poorly amplified by droplet based single cell sequencing methods, therefore we have presented the lineage differentiation using a combination of canonical lineage differentiation markers and high differentially expressed genes between clusters as identified by the FindMarkers function of Seurat and presented in a heatmap (Figure S2H). In short, cells of the erythroid lineage are known to express hemoglobins like *hbba1. itga2b* was used as a marker for thrombocytes in the form of *Tg(CD41:GFP)*, *mpeg1.1* for monocytes and *lysozyme (lyz)* for neutrophils (Figure S2H). Furthermore, the differentiated populations were devoid of expression of proliferation genes like *myca* and *pcna,* indeed suggesting that these are differentiated cells and not a progenitor compartment (Figure S2H). Interestingly, one population expressed markers like *CCAAT enhancer binding protein alpha* (*cebpa)* and *granulin1 (grn1)* (Figure 3B) which are expressed in the monocyte lineage. This population also showed very high expression levels of *s100a10b* (Figure S3A). *In situ* hybridization confirmed that macrophages expressed high levels of *s100a10b* (Figure S3A, B).

The sorting strategy for the kidney progenitor population also contains cells of the lymphoid lineage. A lymphoid progenitor population expressing *rag1* was found. Furthermore, this population also expressed the *rag1* homologue *topoisomerase 2a (top2a)* (Figure S2H), probably as an alternative mechanism in V(D)J recombination[66], and high expression of proliferation markers *myca* and *pcna* suggestive of the lymphoid progenitors. T cells were marked by the expression of *tox, il2rb* and *dusp2*, NK cells were marked by *nkl.3, nkl.4* and *ccl33.3* and two distinct populations of B-cells were marked by expression of *CD37, CD79a* and *pax5* (Figure S2H, Figure 3C). Interestingly, the second B-cell population showed high levels of the immunoglobulins *ighv1-4* and *ighz*[46,65] but did not express the proliferative marker *pcna*, indicating this is a more mature or activated population of B-cells (Figure S2H). Analysis of IgM:GFP transgenic zebrafish [35] showed indeed the presence of several B-cell populations (Figure 4L-S, Figure S4G-L).

One cluster was marked by high expression of *pigrl3.5* which encodes a polymeric Ig receptor[47]. The *pigrl3.5* expressing cells are in close association in the UMAP with the NK population suggesting this is a lymphoid population (Figure S2G, H).

A small cluster with expression of *epithelial cell adhesion molecule (epcam)* was detected (Figure S2H). We hypothesize that this may be niche cells, but since this population is very small and our sorting strategy was not meant to obtain the niche cells, we have not further investigated this population.

Proportional difference between WT and *gata2b*$^{-/-}$ cells in clusters is calculated using scProportionTest package (Figure 3M, 5D and S6D)[67]

HSPC identification

The most immature population would be an hematopoietic stem or progenitor cell (HSPC). These cell types are marked by little expression of lineage markers and expression of proliferation markers. In zebrafish *fli1a* and *meis1b* are typical HSC markers[48-50]. The HSPC1 and HSPC2 populations meet these criteria by expressing low levels of lineage markers (Figure 3B-D). These populations showed high expression levels of proliferation markers like *myca, pcna* and *mki67* (Figure 3J, Figure S2H) and some cells with high levels of the stem cell markers *fli1a* and *meis1b* (Figure 3F, G). Putative HSC, represented by CD41:GFP^{low} sorted cells, are present in both the HSPC1 and HSPC2 clusters (Figure 3I). As expected, the GFP high expressing cells were thrombocytes. *gata2b* gene expression was enriched in the

HSPC1 cluster compared to the other clusters (Figure 3H), indicating that this population would be most affected by the loss of Gata2b. Also, highest expression of *fli1a* and *meis1b* were found in HSPC1 (Figure 3F, G) indicating that this cluster contains some HSCs.

Lineage trajectory analysis

After defining the clusters we performed trajectory and pseudotime analysis using Monocle 3 on the integrated data set and HSPC1 cluster, both carrying Seurat embeddings[68]. Monocle 3 uses an algorithm to learn the differentiation trajectory according to the gene expression of each cell. Once the trajectory graph was learned (Figure 3K and 5H), we used get_earliest_principal_node function and chose the "root" to produce the pseudotime graph. For whole data set HSPC1 cluster was chosen as a starting point and for HSPC1 cluster quiescent subcluster was chosen to generate the pseudotime graphs (Figure 3L, 5I). We used plot_genes_in_pseudotime function to identify the gene expression in pseudotime and found that *gata2b* expression is the highest in the most immature HSPCs and decreases with differentiation (Figure 3N). We also showed that when HSPC1 cluster is chosen as a starting point, the expression of lineage specific genes such as *CD37*, *grn1* and *hbba1* were increased in pseudotime (Figure 3O-Q) which shows that HSPC1 cluster is indeed the most immature cluster within the whole dataset. Similarly, in HSPC1 cluster, when quiescent subcluster was chosen as a starting point, we found that the expression of stem cell marker *meis1b*, also *gata2b* and *CD41:GFP* were decreased in pseudotime, confirming the differentiation trajectory for these cells (Figure 5J-L).

Table S1.

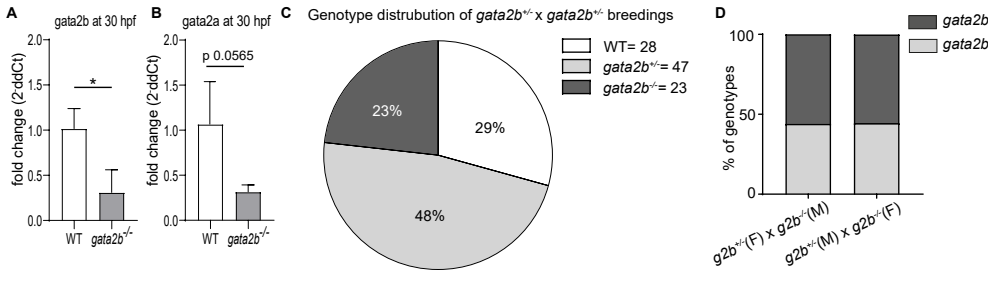
	WT	***gata2b*$^{+/-}$**	***gata2b*$^{-/-}$**
EHT events			
AGM (32 to 40hpf)	3.3 ± 0.4 (n=18)	2.9 ± 0.3 (n=33)	2.2 ± 0.4 (n=18)
CD41:GFP^{+}Flt1:RFP^{+} cells			
CHT (52-54hpf)	8.5 ± 1.7 (n=8)	9.0 ± 1.5 (n=20)	11.9 ± 1.0 (n=7)
CHT (58-60hpf)	25.2 ± 3.2 (n=16)	28.7 ± 1.9 (n=28)	23.3 ± 2.9 (n=12)
CHT (76-58hpf)	55.5 ± 3.8 (n=13)	43.4 ±5.6 (n=12)	36.5 ± 3.0 (n=14)**
***cmyb* expression intensity**			
AGM (24 hpf)	33.0 ± 6.3 (n=14)	28.5 ± 4.5 (n=39)	30.6 ± 2.6 (n=13)
AGM (33 hpf)	32.9 ± 10.9 (n=16)	30.9 ± 7.9 (n=31)	28.5 ± 6.5(n=28)**
CHT (56 hpf)	10.5 ± 2.5 (n=19)	12.8 ± 0.5 (n=41)	4.2 ± 1.1 (n=14)**
CHT (76 hpf)	21.3 ± 7.4 (n=19)	20.8 ± 1.125 (n=43)**	16.142 ± 1.950 (n=15)*
***Runx1* expression intensity**			
AGM (26 hpf)	24.1 ± 9.6 (n= 10)	23.7 ± 11.6 (n= 11)	25.4 ± 9.3 (n=11)
AGM (36 hpf)	30.0 ± 7.1 (n=18)	23.1 ± 8.1 (n=20)	25.3 ± 6.8 (n=11)
mpeg:GFP^{+} cells			
CHT (54hpf)	513.9±23.67 (n=18)	564.9±23.51 (n=26)	545.9±33.27 (n=14)
mpx:GFP^{+} cells			
CHT (75hpf)	227.5±14.40 (n=11)	221.1±8.63 (n=26)	235.0±15.49 (n=6)
LCK:GFP^{+} pixel number			
Thymus (5dpf)	2093±186.2 (n=9)	2449±165.5 (n=18)	2354±206.8 (n=17)

The data are mean ± SEM. n= number of zebrafish embryos used in analysis. AGM; aorta-gonad-mesonephros region, CHT; caudal hematopoietic tissue. *P< 0.05, **P< 0.01. If data were normally distributed we used One-way ANOVA with Tukey post-test. If data were not normally distributed we used Kruskal-Wallis with Dunn's post-test.

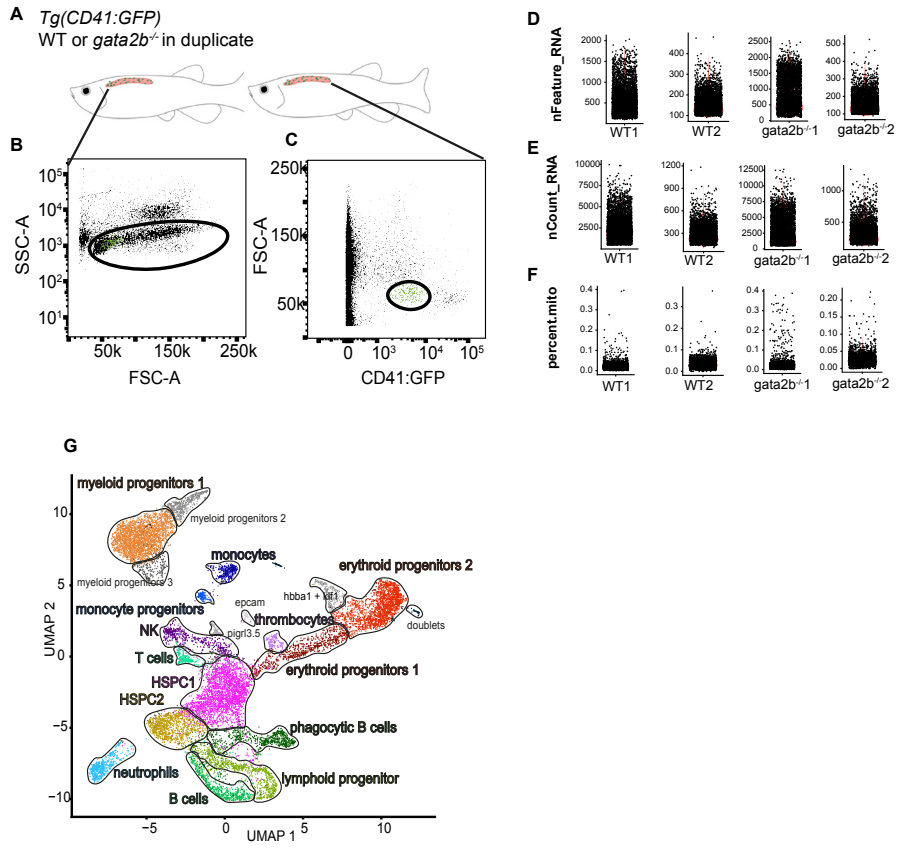
Table S2. Cell numbers within the different populations in single cell analysis

	HSPC1	HSPC2	Myeloid progenitors 1	Myeloid progenitors 2	Myeloid progenitors 3	Monocyte progenitor	Monocytes	Neutrophils	Trombocytes	Ery progenitors	hbba1 + klf1	Erythrocytes	Lymphoid progenitors	B-cells	Phagocytic B-cells	Nk cells	T cells	pigrl3.5	epcam	doublets
WT 2	588	188	1110	300	146	33	1	216	39	331	6	301	105	44	58	164	21	5	3	16
WT 1	498	110	822	31	109	59	30	142	55	112	168	100	888	65	149	109	34	20	21	45
gata2b$^{-/-}$2	846	102	410	61	71	119	41	55	19	202	9	148	267	201	238	99	60	20	9	14
gata2b$^{-/-}$1	572	570	77	34	44	110	49	144	60	127	20	361	332	248	362	185	84	64	38	20
Total	2504	970	2419	426	370	121	321	557	173	772	203	1818	792	558	807	557	199	109	71	95

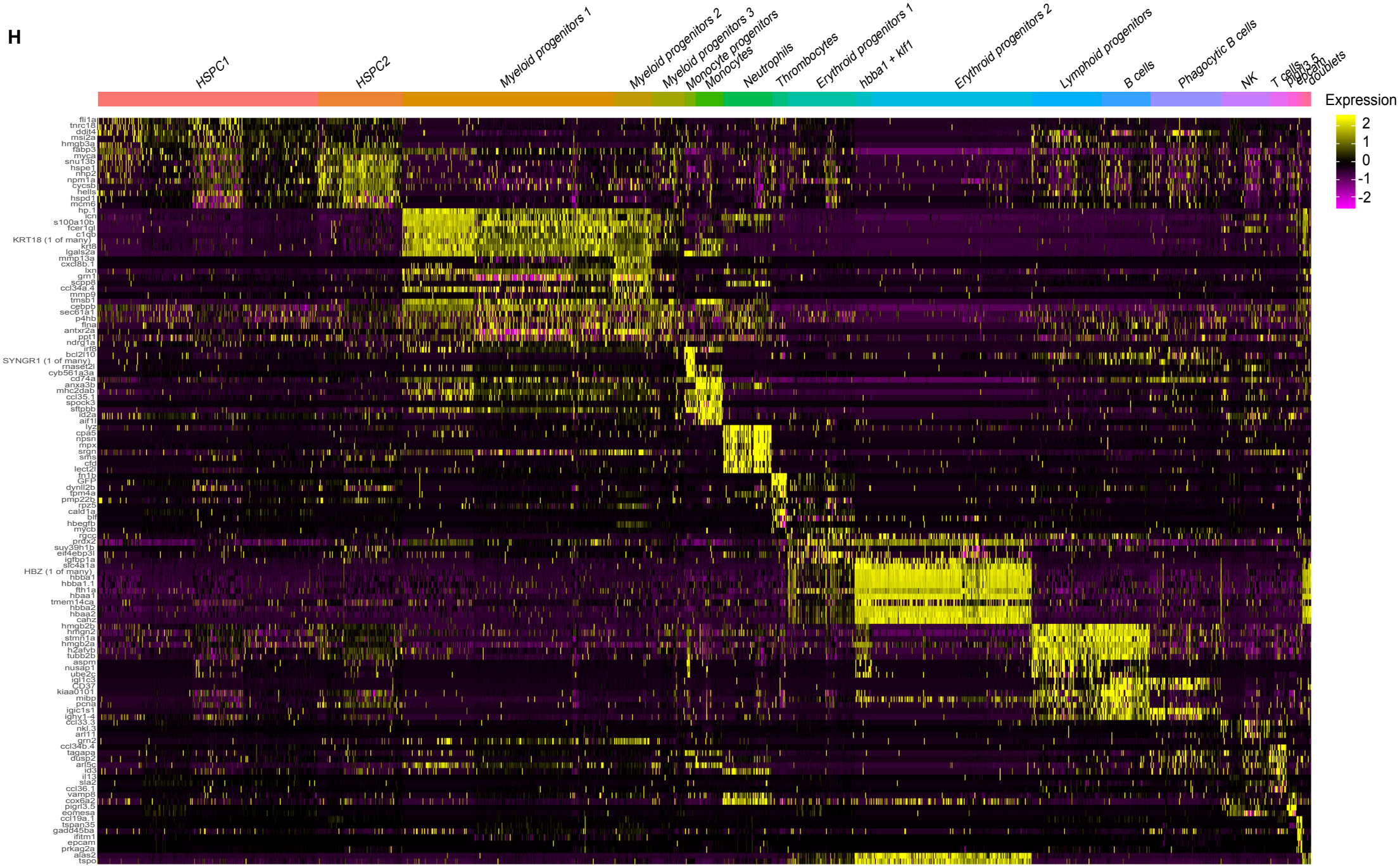
SUPPLEMENTARY FIGURES

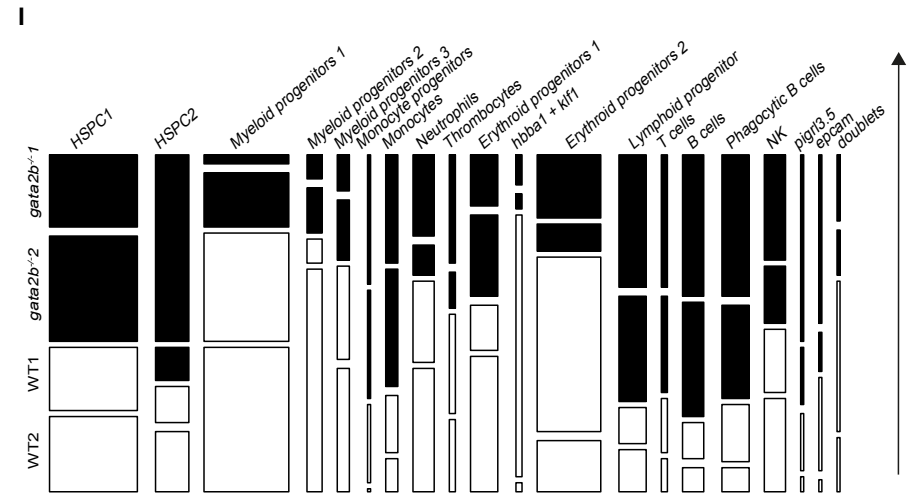


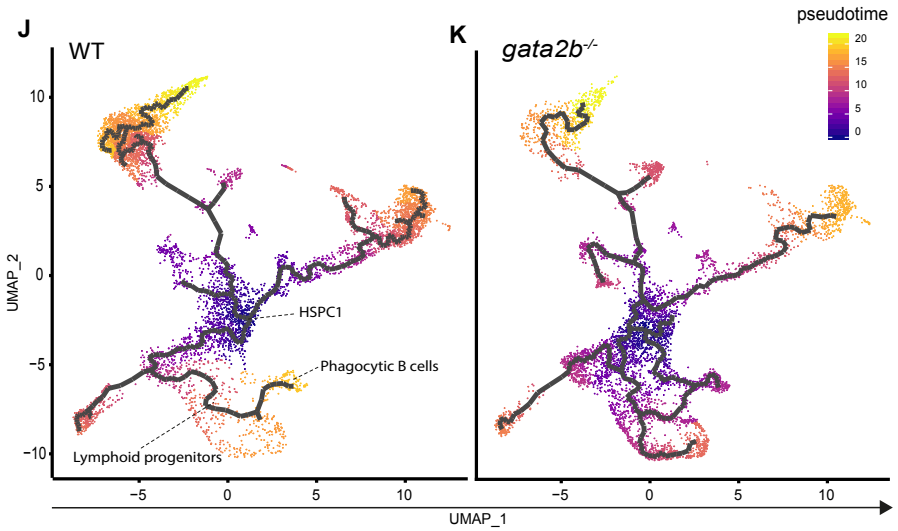
Supplementary Figure 1. Maternal contribution of *gata2b* does not affect *gata2b*$^{-/-}$ survival.
A) qRT-PCR of *gata2b* references against *elfa* on pooled WT and *gata2b*$^{-/-}$ embryos at 30 hpf (n = 3) indicating that *gata2b* expression levels are significantly reduced in mutant embryos (p = 0.0218). B) qRT-PCR of *gata2a* references against *elfa* on pooled WT and *gata2b*$^{-/-}$ embryos at 30 hpf (n = 3) shows reduced *gata2a* expression levels are in mutant embryos (p = 0.0565). C) Genotype distribution of matings between *gata2b*$^{+/-}$ and *gata2b*$^{+/-}$ zebrafish. D) Genotype distribution of matings between *gata2b*$^{+/-}$ female and *gata2b*$^{-/-}$ zebrafish males or *gata2b*$^{+/-}$ male and *gata2b*$^{-/-}$ zebrafish females. *; P<0.05.



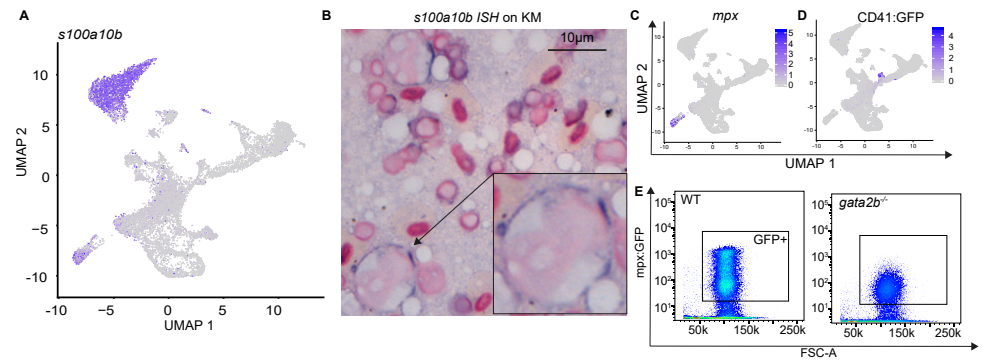
Supplementary Figure 2. Single cell RNA sequencing reveals several progenitor populations
A-C) Experimental strategy to obtain single cells for RNA sequencing. B) FACS plot indicating the progenitor population which was sorted and supplemented with the remaining C) CD41:GFP^{low} expressing cells from the kidney marrow pool of cells. D-E-F) Quality control parameters for each sample G) UMAP showing cluster analysis on aggregated data set of both WT and *gata2b*$^{-/-}$ cells indicating 20 different clusters with different colors. H) Heatmap showing top10 marker genes for each cluster calculated in an unbiased way. I) Genotype distribution of each of the clusters, area of the bars indicate the cell numbers in each cluster, white = WT, black = *gata2b*$^{-/-}$. Each replicate is depicted in a separate bar. J) Differentiation trajectory and pseudotime calculated only for WT cells. K) Differentiation trajectory and pseudotime calculated only for *gata2b*$^{-/-}$ cells.


H
HSPC1
HSPC2
Myeloid progenitors 1
Myeloid progenitors 2
Myeloid progenitors 3
Monocyte progenitors
Monocytes
Neutrophils
Thrombocytes
Erythroid progenitors 1
hbba1 + klf1
Erythroid progenitors 2
Lymphoid progenitors
B cells
Phagocytic B cells
NK
T cells
pdgfra 5
pdgfcain
doublets
Expression
2
1
0
-1
-2
fli1a
tnrc18
ddit4
msi2a
hmgb3a
fabp3
myca
snu13b
hspe1
nhp2
npm1a
cycsb
hells
hspd1
mcm6
hp.1
icn
s100a10b
fcer1gl
c1qb
KRT18 (1 of many)
krt8
lgals2a
mmp13a
cxcl8b.1
lxn
grn1
scpp8
ccl34a.4
mmp9
tmsb1
cebpb
sec61a1
p4hb
flna
antxr2a
ppt1
ndrg1a
irf8
bcl2l10
SYNGR1 (1 of many)
rnaset2l
cyb561a3a
cd74a
anxa3b
mhc2dab
ccl35.1
spock3
sftpbb
id2a
aif1l
lyz
cpa5
npsn
mpx
srgn
srhs
cfd
lect2l
fn1b
GFP
dynll2b
tpm4a
pmp22b
rpz5
cald1a
blf
hbegfb
mycb
rgcc
prdx2
suv39h1b
eif4ebp3l
igfbp1a
slc4a1a
HBZ (1 of many)
hbba1
hbba1.1
fth1a
hbaa1
tmem14ca
hbba2
hbaa2
cahz
hmgb2b
hmgn2
stmn1a
hmgb2a
h2afvb
tubb2b
aspm
nusap1
ube2c
igl1c3
CD37
kiaa0101
mibp
pcna
igic1s1
ighv1-4
ccl33.3
nkl.3
arl11
grn2
ccl34b.4
tagapa
dusp2
arl5c
id3
il13
sla2
ccl36.1
vamp8
cox6a2
pigrl3.5
eomesa
ccl19a.1
tspan35
gad045ba
ifitm1
epcam
prkag2a
alas2
tspo

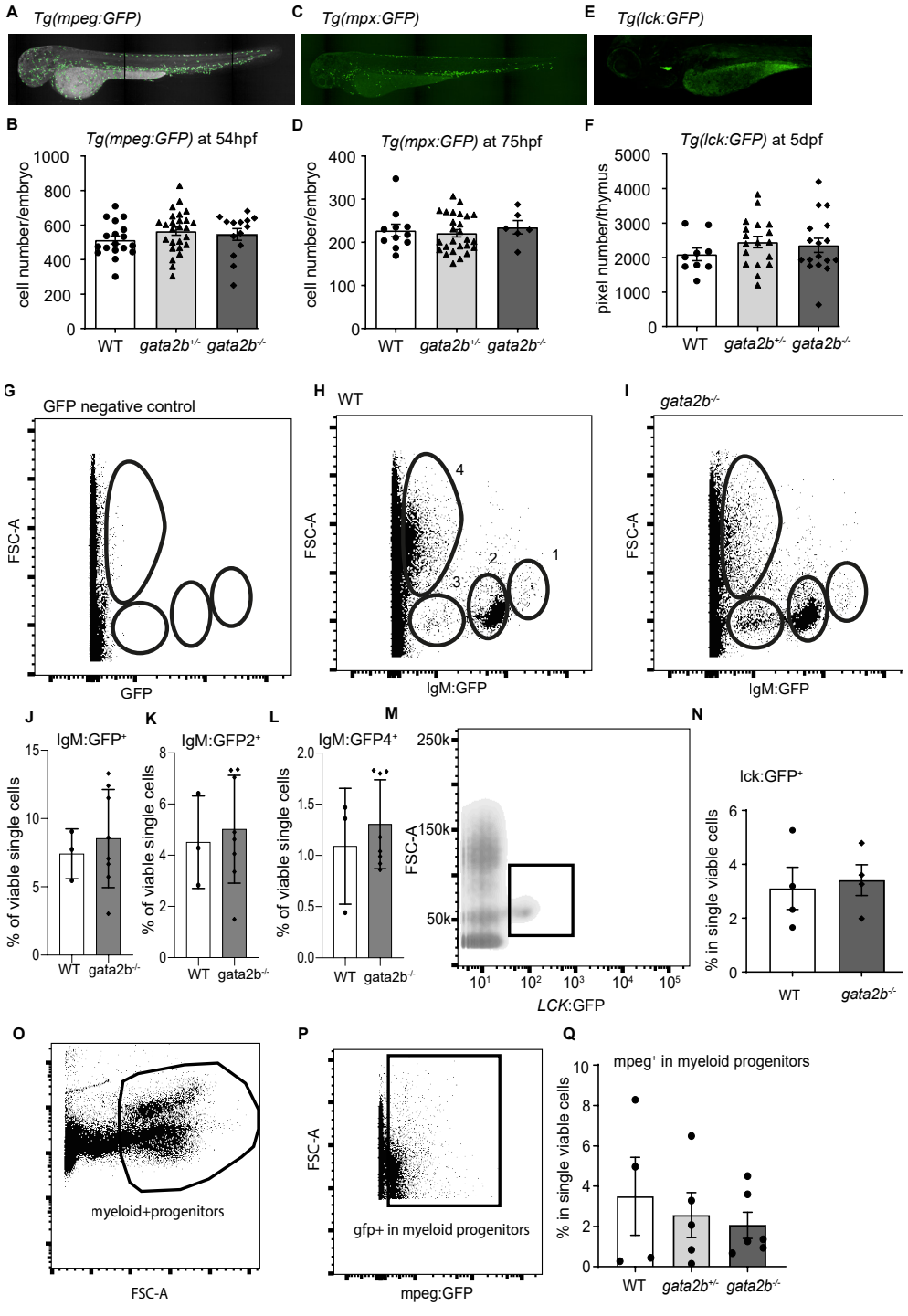




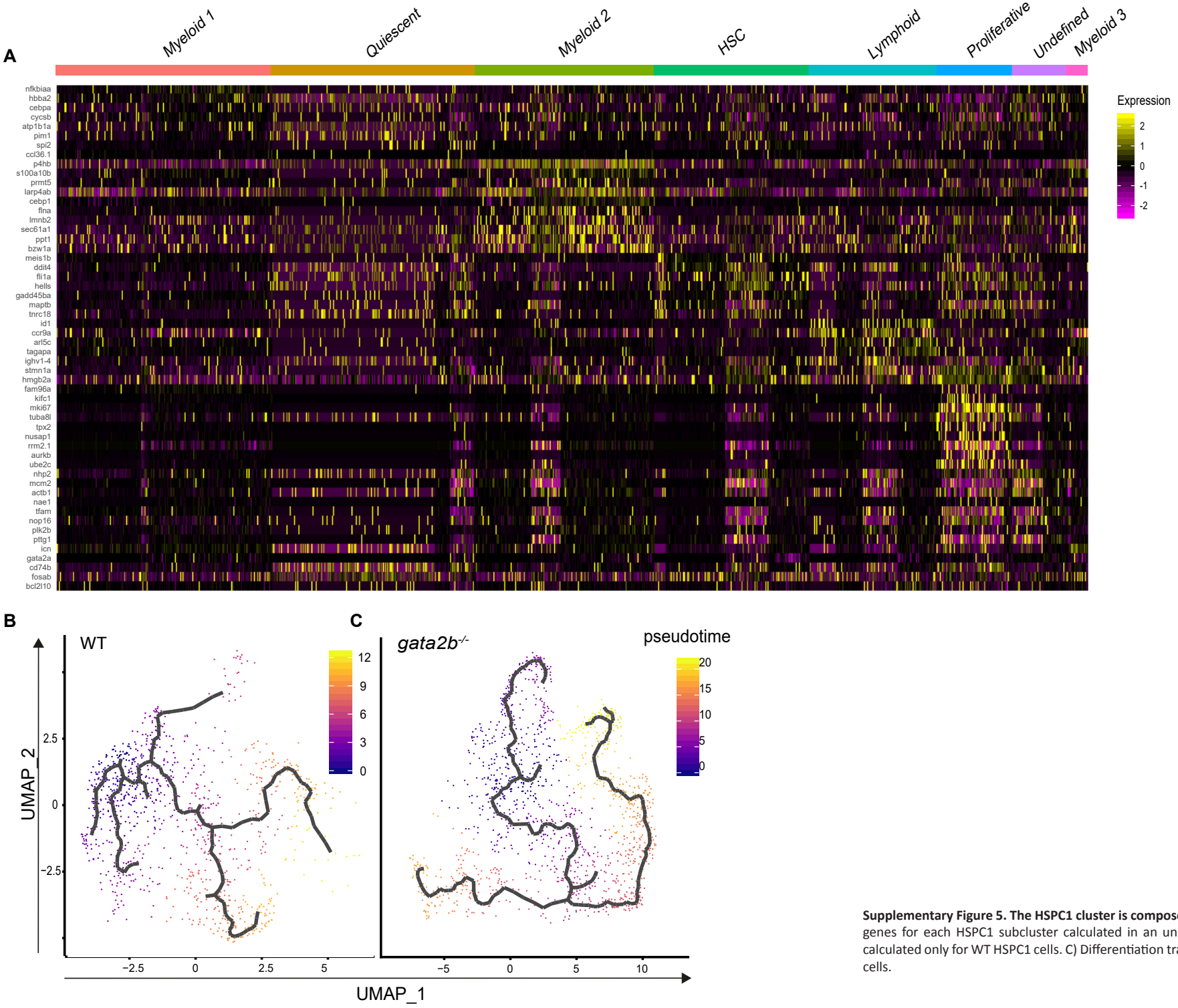
Supplementary Figure 2. Single cell RNA sequencing reveals several progenitor populations
A-C) Experimental strategy to obtain single cells for RNA sequencing. B) FACS plot indicating the progenitor population which was sorted and supplemented with the remaining C) CD41:GFP^{low} expressing cells from the kidney marrow pool of cells. D-E-F) Quality control parameters for each sample G) UMAP showing cluster analysis on aggregated data set of both WT and *gata2b*$^{-/-}$ cells indicating 20 different clusters with different colors. H) Heatmap showing top10 marker genes for each cluster calculated in an unbiased way. I) Genotype distribution of each of the clusters, area of the bars indicate the cell numbers in each cluster, white = WT, black = *gata2b*$^{-/-}$. Each replicate is depicted in a separate bar. J) Differentiation trajectory and pseudotime calculated only for WT cells. K) Differentiation trajectory and pseudotime calculated only for *gata2b*$^{-/-}$ cells.



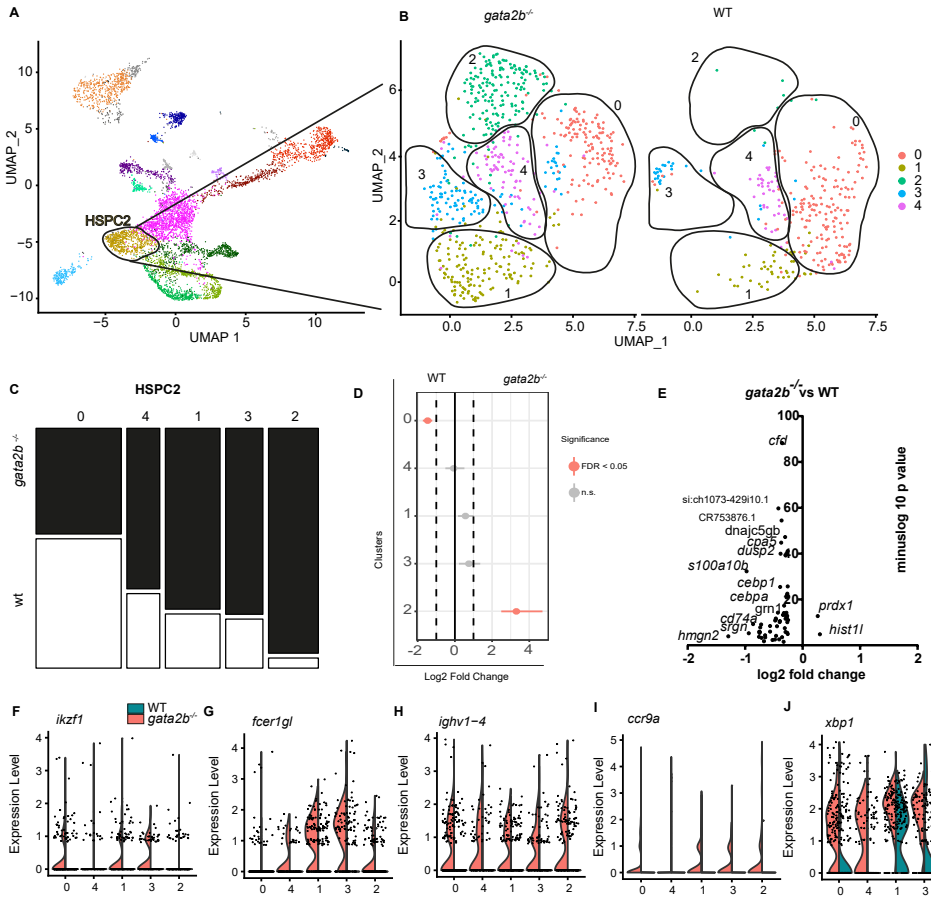
Supplementary Figure 3. *s100a10b* is expressed in the neutrophil lineage. A) Feature analysis with gradual gene expression in shades of blue of *s100a10b*. B) WT representative images of *s100a10b in situ* hybridization on zebrafish kidney marrow smears. A banded neutrophil is indicated by the arrow and shown enlarged in the corner. Scale bar indicates 10 µm. UMAP = Uniform manifold approximation and Projection. C) Feature analysis showing *mpx* expression in the neutrophil cluster. D) Feature analysis showing high *CD41:*GFP expression in the thrombocytes cluster and low *CD41:*GFP expression in the HSPC1 and HSPC2 clusters. E) Gating strategy for *Tg(mpx:GFP)* in WT and *Gata2b*$^{-/-}$ KM.



Supplementary Figure 4. Differentiation markers are not altered in *gata2b*$^{-/-}$ embryos.
A) Representative picture of *Tg(mpeg:GFP)* embryo at 54 hpf. B) Quantitation of *mpeg*:GFP^{+} cells in WT, *gata2b*$^{+/-}$ and *gata2b*$^{-/-}$ embryos at 54 hpf. C) Representative picture of *Tg(mpx:GFP)* embryos at 75 hpf. D) Quantitation of *mpx*:GFP^{+} cells in WT, *gata2b*$^{+/-}$ and *gata2b*$^{-/-}$ embryos. E) Representative picture of *Tg(lck:GFP)* embryos at 5 dpf F) Lck:GFP^{+} area represented as pixel number in WT, *gata2b*$^{+/-}$ and *gata2b*$^{-/-}$ embryos. Each dot represent one embryo, see Table S1 for exact cell numbers and numbers of embryos analyzed. G) Negative GFP control. H) WT *Tg(IgM:GFP)* zebrafish KM with 4 GFP populations gated with clear differences in size (FSC) or GFP positivity, I) Similar gating strategy for *Gata2b*$^{-/-}$ *Tg(IgM:GFP)* KM. J-L) Quantitation of total IgM:GFP^{+}, IgM:GFP2^{+} and IgM:GFP4^{+} populations for WT and *Gata2b*$^{-/-}$ KM cells as percentage of single viable cells. M) Gating strategy for *Tg(lck:GFP)* WT and *Gata2b*$^{-/-}$ KM. N) Quantitation of GFP^{+} cells in *Tg(lck:GFP)* WT and *Gata2b*$^{-/-}$ KM as percentage of single viable cells. O-P) Gating strategy of *Tg(mpeg1.1:GFP)* WT and *Gata2b*$^{-/-}$ KM. Q) Quantitation of GFP^{+} cells in *Tg(mpeg1.1:GFP)* WT, *gata2b*$^{+/-}$ and *Gata2b*$^{-/-}$ KM as percentage of single viable cells. Error bars represent SE



Supplementary Figure 5. The HSPC1 cluster is composed of multiple HSPC subtypes. A) Heatmap showing marker genes for each HSPC1 subcluster calculated in an unbiased way. B) Differentiation trajectory and pseudotime calculated only for WT HSPC1 cells. C) Differentiation trajectory and pseudotime calculated only for *gata2b*$^{-/-}$ HSPC1 cells.



Supplementary Figure 6. HSPC2 shows reduced myeloid differentiation and increased in lymphoid differentiation. A) Selection of HSPC2 for further analysis in B) subclusters (0-4) of *gata2b*$^{-/-}$ cells on the left and WT cells on the right. C) Proportion analysis of the different subclusters, indicating unequal distribution of WT and *gata2b*$^{-/-}$ cells in the individual subclusters. D) Pointrange plot showing the difference between proportion of WT and *gata2b*$^{-/-}$ cells for each HSPC2 subcluster calculated by permutation test. If FDR < 0.05, point is colored in pink and if not in grey. E) Volcanoplot showing differential gene expression analysis between WT and *gata2b*$^{-/-}$ cells showing robust downregulation of myeloid genes in *gata2b*$^{-/-}$ cells. F-J) violin plots of individual lymphoid gene expression within subclusters of HSPCs with WT cells in green and *gata2b*$^{-/-}$ cells in pink of F) *ikzf2*, G) *fcer1gl*, H) *ighv1-4*, I) *ccr9a* and J) *xpb1*.

SUPPLEMENTARY REFERENCES

1. Sawai CM, Babovic S, Upadhaya S, et al. Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity*. 2016;45(3):597-609.
2. Yamamoto R, Morita Y, Ooehara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*. 2013;154(5):1112-1126.
3. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013;502(7470):232-236.
4. Muller-Sieburg CE, Cho RH, Karlsson L, Huang JF, Sieburg HB. Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood*. 2004;103(11):4111-4118.
5. Dykstra B, Kent D, Bowie M, et al. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell*. 2007;1(2):218-229.
6. Watcham S, Kucinski I, Gottgens B. New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood*. 2019;133(13):1415-1426.
7. de Pater E, Kaimakis P, Vink CS, et al. Gata2 is required for HSC generation and survival. *J Exp Med*. 2013;210(13):2843-2850.
8. Gao X, Johnson KD, Chang YI, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med*. 2013;210(13):2833-2842.
9. Snow JW, Trowbridge JJ, Fujiwara T, et al. A single cis element maintains repression of the key developmental regulator Gata2. *PLoS Genet*. 2010;6(9):e1001103.
10. Mehta C, Johnson KD, Gao X, et al. Integrating Enhancer Mechanisms to Establish a Hierarchical Blood Development Program. *Cell Rep*. 2017;20(12):2966-2979.
11. Johnson KD, Kong G, Gao X, et al. Cis-regulatory mechanisms governing stem and progenitor cell transitions. *Sci Adv*. 2015;1(8):e1500503.
12. Ling KW, Ottersbach K, van Hamburg JP, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med*. 2004;200(7):871-882.
13. Rodrigues NP, Janzen V, Forkert R, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*. 2005;106(2):477-484.
14. Rodrigues NP, Boyd AS, Fugazza C, et al. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood*. 2008;112(13):4862-4873.
15. Nandakumar SK, Johnson K, Throm SL, Pestina TI, Neale G, Persons DA. Low-level GATA2 overexpression promotes myeloid progenitor self-renewal and blocks lymphoid differentiation in mice. *Exp Hematol*. 2015;43(7):565-577 e561-510.
16. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371(6494):221-226.
17. Bertrand JY, Cisson JL, Stachura DL, Traver D. Notch signaling distinguishes 2 waves of definitive hematopoiesis in the zebrafish embryo. *Blood*. 2010;115(14):2777-2783.
18. Bertrand JY, Chi NC, Santoso B, Teng S, Stainier DY, Traver D. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*. 2010;464(7285):108-111.
19. Kissa K, Herbomel P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature*. 2010;464(7285):112-115.
20. Warga RM, Kane DA, Ho RK. Fate mapping embryonic blood in zebrafish: multi- and unipotential lineages are segregated at gastrulation. *Dev Cell*. 2009;16(5):744-755.

21. Ciau-Uitz A, Monteiro R, Kirmizitas A, Patient R. Developmental hematopoiesis: ontogeny, genetic programming and conservation. *Exp Hematol*. 2014;42(8):669-683.
22. Tamplin OJ, Durand EM, Carr LA, et al. Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell*. 2015;160(1-2):241-252.
23. Davidson AJ, Zon LI. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene*. 2004;23(43):7233-7246.
24. Butko E, Distel M, Pouget C, et al. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development*. 2015;142(6):1050-1061.
25. Dobrzycki T, Mahony CB, Krecsmarik M, et al. Deletion of a conserved Gata2 enhancer impairs haemogenic endothelium programming and adult Zebrafish haematopoiesis. *Commun Biol*. 2020;3(1):71.
26. Gagnon JA, Valen E, Thyme SB, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One*. 2014;9(5):e98186.
27. Chocron S, Verhoeven MC, Rentzsch F, Hammerschmidt M, Bakkers J. Zebrafish Bmp4 regulates left-right asymmetry at two distinct developmental time points. *Dev Biol*. 2007;305(2):577-588.
28. Dobrzycki T, Krecsmarik M, Bonkhofer F, Patient R, Monteiro R. An optimised pipeline for parallel image-based quantification of gene expression and genotyping after in situ hybridisation. *Biol Open*. 2018;7(4).
29. Lawson ND, Weinstein BM. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol*. 2002;248(2):307-318.
30. Ma D, Zhang J, Lin HF, Italiano J, Handin RI. The identification and characterization of zebrafish hematopoietic stem cells. *Blood*. 2011;118(2):289-297.
31. Bussmann J, Bos FL, Urasaki A, Kawakami K, Duckers HJ, Schulte-Merker S. Arteries provide essential guidance cues for lymphatic endothelial cells in the zebrafish trunk. *Development*. 2010;137(16):2653-2657.
32. Ellett F, Pase L, Hayman JW, Andrianopoulos A, Lieschke GJ. mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood*. 2011;117(4):e49-56.
33. Renshaw SA, Loynes CA, Trushell DM, Elworthy S, Ingham PW, Whyte MK. A transgenic zebrafish model of neutrophilic inflammation. *Blood*. 2006;108(13):3976-3978.
34. Langenau DM, Ferrando AA, Traver D, et al. In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. *Proc Natl Acad Sci U S A*. 2004;101(19):7369-7374.
35. Page DM, Wittamer V, Bertrand JY, et al. An evolutionarily conserved program of B-cell development and activation in zebrafish. *Blood*. 2013;122(8):e1-11.
36. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411-420.
37. Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc Natl Acad Sci U S A*. 2003;100(15):8811-8816.
38. Medvinsky A, Dzierzak E. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*. 1996;86(6):897-906.
39. Jaffredo T, Gautier R, Eichmann A, Dieterlen-Lièvre F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development*. 1998;125(22):4575-4583.
40. Boisset JC, van Cappellen W, Andrieu-Soler C, Galjart N, Dzierzak E, Robin C. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*. 2010;464(7285):116-120.
41. Bertrand JY, Kim AD, Teng S, Traver D. CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development*. 2008;135(10):1853-1862.
42. Bonkhofer F, Rispoli R, Pinheiro P, et al. Blood stem cell-forming haemogenic endothelium in zebrafish derives from arterial endothelium. *Nat Commun*. 2019;10(1):3577.

43. Bein K, Husain M, Ware JA, Mucenski ML, Rosenberg RD, Simons M. c-Myb function in fibroblasts. *J Cell Physiol*. 1997;173(3):319-326.
44. North TE, Goessling W, Walkley CR, et al. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature*. 2007;447(7147):1007-1011.
45. Carmona SJ, Teichmann SA, Ferreira L, et al. Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome Res*. 2017;27(3):451-461.
46. Danilova N, Bussmann J, Jekosch K, Steiner LA. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. *Nat Immunol*. 2005;6(3):295-302.
47. Kortum AN, Rodriguez-Nunez I, Yang J, et al. Differential expression and ligand binding indicate alternative functions for zebrafish polymeric immunoglobulin receptor (pIgR) and a family of pIgR-like (PIGRL) proteins. *Immunogenetics*. 2014;66(4):267-279.
48. Macaulay IC, Svensson V, Labalette C, et al. Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. *Cell Rep*. 2016;14(4):966-977.
49. Tang Q, Iyer S, Lobbardi R, et al. Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. *J Exp Med*. 2017;214(10):2875-2887.
50. Athanasiadis EI, Botthof JG, Andres H, Ferreira L, Lio P, Cvejic A. Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat Commun*. 2017;8(1):2045.
51. Li J, Barreda DR, Zhang YA, et al. B lymphocytes from early vertebrates have potent phagocytic and microbicidal abilities. *Nat Immunol*. 2006;7(10):1116-1124.
52. Traver D, Paw BH, Poss KD, Penberthy WT, Lin S, Zon LI. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol*. 2003;4(12):1238-1246.
53. Bennett CM, Kanki JP, Rhodes J, et al. Myelopoiesis in the zebrafish, Danio rerio. *Blood*. 2001;98(3):643-651.
54. Ferrero G, Gomez E, Lyer S, et al. The macrophage-expressed gene (mpeg) 1 identifies a subpopulation of B cells in the adult zebrafish. *J Leukoc Biol*. 2020;107(3):431-443.
55. Moss LD, Monette MM, Jaso-Friedmann L, et al. Identification of phagocytic cells, NK-like cytotoxic cell activity and the production of cellular exudates in the coelomic cavity of adult zebrafish. *Dev Comp Immunol*. 2009;33(10):1077-1087.
56. Olsson A, Venkatasubramanian M, Chaudhri VK, et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature*. 2016;537(7622):698-702.
57. Drissen R, Buza-Vidas N, Woll P, et al. Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat Immunol*. 2016;17(6):666-676.
58. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*. 2017;19(4):271-281.
59. Graf T, Enver T. Forcing cells to change lineages. *Nature*. 2009;462(7273):587-594.
60. Hoppe PS, Schwarzfischer M, Loeffler D, et al. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature*. 2016;535(7611):299-302.
61. Lin HF, Traver D, Zhu H, et al. Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood*. 2005;106(12):3803-3810.
62. Vinh DC, Patel SY, Uzel G, et al. Autosomal dominant and sporadic monocytopenia with susceptibility to mycobacteria, fungi, papillomaviruses, and myelodysplasia. *Blood*. 2010;115(8):1519-1529.
63. Hsu AP, Sampaio EP, Khan J, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*. 2011;118(10):2653-2655.

64. Hahn CN, Chong CE, Carmichael CL, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet*. 2011;43(10):1012-1017.

65. Moore JC, Tang Q, Yordan NT, et al. Single-cell imaging of normal and malignant cell engraftment into optically clear prkdc-null SCID zebrafish. *J Exp Med*. 2016;213(12):2575-2589.

66. Riou JF, Grondard L, Petitgenet O, Abitbol M, Lavelle F. Altered topoisomerase I activity and recombination activating gene expression in a human leukemia cell line resistant to doxorubicin. *Biochem Pharmacol*. 1993;46(5):851-861.

67. Miller SA, Policastro RA, Sriramkumar S, et al. LSD1 promotes secretory cell specification to drive BRAF mutant colorectal cancer. *bioRxiv*. 2020:2020.2009.2025.313536.

68. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496-502.

3

Zebrafish gata2b haploinsufficiency results in erythroid-dysplasia

Emanuele Gioacchino, Cansu Koyunlar, Joke Zink, Hans de Looper, Remco Hoogenboezem, Eric Bindels, Ivo Touw and Emma de Pater

Manuscript Submitted

ABSTRACT

The expression level of the transcription factor GATA2 is pivotal in hematopoiesis. Germline GATA2 haploinsufficiency manifests with immunodeficiency, bone marrow failure and a predisposition to myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML). The mechanisms driving GATA2 related phenotypes are still unexplained. To explore the role of GATA2 haploinsufficiency on the hematopoietic system we use a zebrafish model with a mutation in the orthologue of *GATA2, Gata2b*. Morphological analysis reveals that *gata2b*$^{+/-}$ kidney marrow (KM) shows erythroid dysplasia. To gain insight in the molecular origin of the dysplasia we used both bulk and single cell RNA sequencing. It unveiled that *gata2b*$^{+/-}$ HSPCs have a downregulation of cytoskeletal transcripts and aberrant proliferative signatures and ribosome biogenesis, possibly contributing to the erythroid dysplasia in *gata2b*$^{+/-}$ fish.

INTRODUCTION

GATA2 has a leading role in the generation and maintenance of the hematopoietic system (Gao et al., 2013; de Pater et al., 2013; Tsai et al., 1994). In humans, the clinical features of GATA2 deficiency are broad and include immunodeficiency, pulmonary, vascular/lymphatic dysfunctions and a strong propensity to develop MDS and AML (Donadieu et al., 2018; Spinner et al., 2014). GATA2 deficiency has been linked to the pathogenesis of Emberger syndrome (Ostergaard et al., 2011), MonoMAC (Hsu et al., 2011), dendritic cell, monocyte, and B and natural killer lymphoid deficiency (DCML) (Dickinson et al., 2011), and familiar forms of AML (Hahn et al., 2011). GATA2 haploinsufficiency has a variable impact on disease phenotype and age of onset, but presents with a high penetrance (Donadieu et al., 2018; Spinner et al., 2014). Surprisingly, there is no clear correlation between the occurrence of GATA2 mutations and severity of clinical symptoms, even between family members who share the same mutations (Mutsaers et al., 2013; Wang et al., 2015). In mice, GATA2 has an essential regulatory activity in the hematopoietic stem cell (HSC) compartment and has a function at several stages of HSC generation and maintenance. GATA2 null mice are lethal at embryonic day (E) 10.5 (Tsai et al., 1994), but *Gata2* heterozygous (*Gata2*$^{+/-}$) mice survive to adulthood with normal blood values. However, *Gata2*$^{+/-}$ mice possess a reduced phenotypic HSC compartment in their bone marrow (BM) (Rodrigues et al., 2005) and develop a blood phenotype after aging and BM transplantation (Gioacchino et al 2021). Whereas mouse models emerged as a precious source for identifying possible functions of GATA2, the mechanism causing the different aspects of GATA2 deficiency syndromes is still undiscovered. To better understand the biology of GATA2 and identify potential disease mechanisms, zebrafish serves as an attractive alternative model. Zebrafish has the advantage of having 2 *GATA2* orthologues, *gata2a* and *gata2b*. Gata2a is expressed mostly in vasculature (Butko et al., 2015) and is required for an appropriate hemogenic endothelium programming, upstream of Gata2b (Dobrzycki et al., 2020; Gioacchino et al 2021). Gata2b is expressed in HSPCs (Butko et al., 2015) and its homozygous deletion (*gata2b*$^{-/-}$) causes a lack of myeloid differentiation and a lymphoid bias (Gioacchino et al 2021). To evaluate the transcriptional and phenotypic consequences of reduced Gata2b expression, we assessed hematopoietic cell development in *gata2b* heterozygous zebrafish (*gata2b*$^{+/-}$). We hypothesized that, with half a Gata2b dose, zebrafish would show reduced myeloid differentiation, although to a lesser extent than when knocking out *gata2b* completely. Instead, we observed an unprecedented erythroid dysplasia in the kidney marrow. We therefore conclude that *gata2b* heterozygous knockout zebrafish develop a distinct phenotype in comparison to the homozygous knockout. These results, besides recapitulating morphological abnormalities occurring in GATA2 patients, highlight a concentration-dependent action of gata2b in zebrafish hematopoiesis.

MATERIAL AND METHODS

Generation and genotyping of Gata2b heterozygous zebrafish

We generated *gata2b* heterozygote zebrafish using CRISPR/Cas9 as previously described (Gioacchino et al 2021), and maintained them under standard conditions. Genotyping of the 28bp insertion in the exon 3 of *gata2b* gene was accomplished by isolating genomic DNA from fin biopsy, followed by polymerase chain reaction (PCR), using forward primer 5' CTGTCGATGACGCAACACTG 3' and reverse primer 5' TGTCGTCATGTTTCCGAGCA 3', giving a product of 363 bp. PCR had a denaturing temperature of 95°C for 15", annealing 54°C for 15" and extension at 72°C for 30" for 35 cycles.

Zebrafish embryos were kept at 28,5°C on a 14h/10h light-dark cycle in HEPES-buffered E3 medium. Zebrafish were anesthetized using tricaine and euthanized by ice-water. Animal studies were approved by the animal Welfare/Ethics Committee in accordance to the legislation in the Netherlands.

Kidney marrow isolation and analysis

Adult zebrafish were euthanized and the kidney marrow was removed mechanically using tweezers and dissociated by pipetting in phosphate buffered saline (PBS)/10% fetal calf serum (FCS) to obtain a single-cell suspension. 7-AAD (7-amino-actinomycin D) (Stem –kit reagents) 0.5mg/L or DAPI 1mg/L was used for live/dead discrimination. The sorting and analysis were performed using FACS AriaIII (BD Bioscences). When sorting, the cells were sorted in non-stick cooled micro tubes (Ambion) containing 10% FCS in PBS.

May-Grünwald-Giemsa stain of KM smears

Adult zebrafish kidney marrow was removed mechanically, using tweezers. Kidney marrow smears were obtained by wiping the tissue on slides (SuperFrost Plus, Thermo Scientific) in duplicates. The smears were subsequently fixed for one minute in 100% MeOH before staining five minutes in May Grünwald solution (diluted 1:1 in phosphate buffer) and ten minutes in Giemsa solution (diluted 1:20 in phosphate buffer) followed by a last rinsing step in tap water.

Analysis was performed by counting 200-500 hematopoietic cells of each kidney marrow smear; excluding mature erythrocytes and thrombocytes. Cells were categorized as: blast, myelocyte, neutrophil, eosinophil, lymphocyte or erythroblast. Furthermore, if dysplasia was observed within a specific lineage, the percentage of dysplastic cells within that lineage was determined by additional counting of at least 50 cells within that specific lineage.

Single cell RNA sequencing

kidney marrow cells were isolated (as describes in bone marrow isolation) and $7x10^4$ viable cells were sorted from 2 pooled males *Tg(CD41:GFP, runx1:RFP)* of wild type (WT) or *gata2b*$^{+/-}$ zebrafish at 1 year of age. $7x10^4$ single viable cells were sorted from kidney marrows from one female of WT or *gata2b*$^{+/-}$ zebrafish between 18-20 months of age. cDNA was prepared using the manufacturers protocol (10x Chromium V3) and sequenced on a Novaseq 6000 instrument (Illumina). After sample processing and upstream analysis, gathered information on 18147 cells for WT and 10849 cells for *gata2b*$^{+/-}$. Considering the difference in cell number between samples, we normalized for this factor and assessed difference between the 2 conditions. Two WT replicates had a read depth of 24135 and 28043 reads per cell and the two *gata2b*$^{+/-}$ replicates were sequenced with a read depth of 21790 and 27885. Data was analyzed using the Seurat R package.

RNA extraction and RNA quality control

Total sample RNA isolation was performed according to the standard protocol of RNA isolation with Trizol and GenElute LPA (Sigma). Quality and quantity of the total RNA was checked on a 2100 Bioanalyzer (Agilent) using the Agilent RNA 6000 Pico Kit.

RNA Sequencing and gene set enrichment analysis (GSEA)

cDNA was prepared using the SMARTer Ultra Low RNA kit (Clontech) for Illumina Sequencing. The gene expression values were measured as FPKM (Fragments per kilobase of exon per million fragments mapped). Fragment counts were determined per gene with HTSeq-count, utilizing the strict intersection option, and subsequently used for differential expression analysis using the DESeq2 package, with standard parameters, in the R environment.

Statistics

Statistical analysis was carried out in GraphPad Prism 8 (GraphPad Software). Unless otherwise specified, data were analyzed using unpaired, 2 tailed Student's t-test. Statistical significance was defined as $p<0.05$. Graphs are means ± SEM and the number of replicates is indicated in the figure legend.

RESULTS

gata2b$^{+/-}$ kidney marrow differentiation remains intact over time

A previously generated *gata2b* mutant zebrafish line was used (Gioacchino et al 2021). Here, we test the consequences of gata2b heterozygosity (*gata2b*$^{+/-}$) in the zebrafish hematopoietic system. Since *gata2b* homozygous mutant (*gata2b*$^{-/-}$) zebrafish had alterations in the lymphoid and myeloid lineage, we analyzed *gata2b*$^{+/-}$ KM by flow cytometry to discriminate an eventual kidney marrow lineage bias (Boatman et al., 2013; Traver et al., 2003) (Fig. 1A). *gata2b*$^{+/-}$ zebrafish had no significant difference in the distribution of either myeloid, erythroid, progenitors or HSPCs and lymphoid populations compared to wild type (WT) (Fig. 1A). To further specify differentiated hematopoietic cells more stringently, we bred *gata2b*$^{+/-}$ zebrafish to various transgenic zebrafish lines, as reporters for differentiated cell types (Supp. Fig. 1). We used *Tg(mpx:GFP)* for neutrophils (Supp. Fig. 1A-B), *Tg(lck:GFP)* for T cells (Supp. Fig. 1C-D), *Tg(Igm:GFP)* for B cells (Supp. Fig. 1E-F), *Tg(mpeg:GFP* in cells of the myeloid gate) for macrophages (Supp. Fig. 1G-H), and *Tg(mpeg:GFP* in cells of the lymphoid gate) as a supplementary control for B cells quantification (Supp. Fig. 1G-I). No significant differences in mature hematopoietic subsets were observed between WT and *gata2b*$^{+/-}$ zebrafish. Since GATA2 haploinsufficiency manifestations might require longer periods of time to become evident (Donadieu et al., 2018; Spinner et al., 2014), we tested the effect of *gata2b* heterozygosity after aging zebrafish up to 18 months. No significant differences between the proportions of cell populations were detected over time in *gata2b*$^{+/-}$ kidney marrow compared to WT (Fig. 1B-E). However, the erythroid population increased, indiscriminately of genotype, after 12 months of age ($P<0.001$) (Fig. 1D). While the remaining myeloid, lymphoid+HSC and progenitor populations did not vary as dramatically (Fig. 1B-C-E). In summary, based on FSC-A vs SSC-A and transgenic reporters we conclude that the major differentiation lineages are not impaired in *gata2b*$^{+/-}$ KM. Additionally, we note a previously unreported erythroid bias in aged zebrafish.

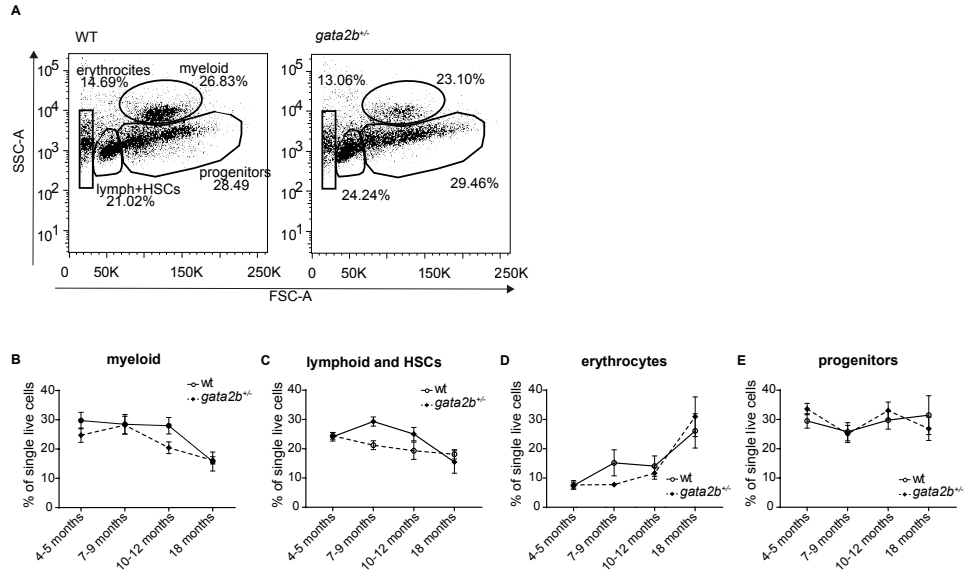


Fig. 1: *gata2b*$^{+/-}$ kidney marrow does not have alteration in differentiation over time
A) Gating strategy of FACS analysis of whole kidney marrow of WT and *gata2b*$^{+/-}$ zebrafish. Percentages represent the average of all zebrafish analyzed per genotype. Quantitation as percentages of the different cell populations in single viable cells of B) myeloid C) lymphoid and HSCs D) erythrocytes E) progenitors in WT and *gata2b*$^{+/-}$ zebrafish kidney marrow over time. Mean ± SEM.

gata2b$^{+/-}$ kidney marrow shows erythroid-myeloid dysplasia

GATA2 haploinsufficiency does not always manifests with cytopenias (Donadieu et al., 2018; Spinner et al., 2014). Therefore, regardless of a conserved overall differentiation process in *gata2b*$^{+/-}$ KM, we morphologically assessed whole KM smears as myelodysplasia is the most common form of leukemia upon *GATA2* haploinsufficiency. MGG staining of KM smear quantification (excluding differentiated erythrocytes) revealed that *gata2b*$^{+/-}$ kidney marrow had no overall differences in lineage differentiation. However, a significant decrease in eosinophils was observed (Fig. 2A), representing less than 5% of the total KM cells. Morphological analysis further showed that, while WT zebrafish had predominantly normal kidney marrow cells, all *gata2b*$^{+/-}$ samples had a considerable fraction of dysplastic cells in the erythroid lineage. On average 0.5% of WT erythroid cells showed dysplastic features, compared to 9.9% of *gata2b*$^{+/-}$ erythroid cells (Fig. 2B), the latter representing 4.5% of the total kidney marrow population of *gata2b*$^{+/-}$. Furthermore, we found clear myeloid lineage dysplasia in 25% of *gata2b*$^{+/-}$ zebrafish (Fig. 2C). In these samples, 30% of myeloid cells were dysplastic compared to the WT average of 0.3%. While the myeloid dysplasia was mostly represented by multi-lobated nuclei, the erythroid abnormalities were ranging from nuclear deformities and double nuclei to irregular cytoplasm or an almost complete lack of cytoplasm (Fig. 2C-D). The remaining cell types did not seem to be affected morphologically by Gata2b haploinsufficiency. These results indicate that *gata2b* heterozygosity causes haploinsufficiency and induces dysplasia, predominantly in erythroid progenitors.

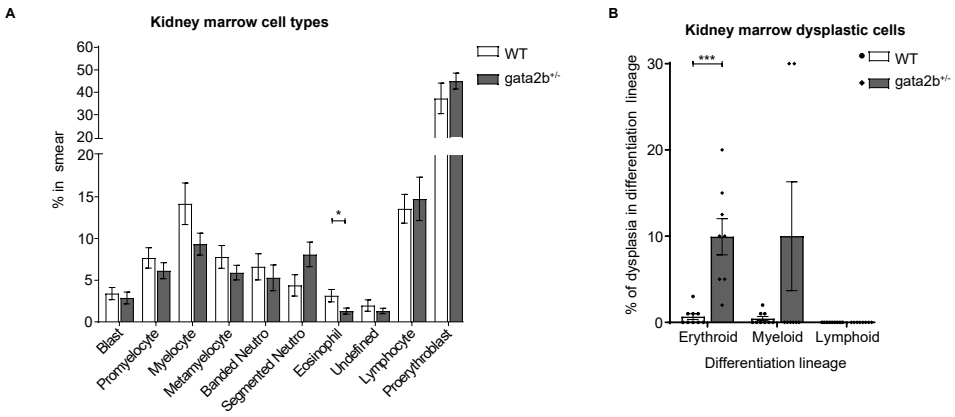


Fig. 2: gata2b$^{+/-}$ kidney marrow shows erythroid dysplasia
A) Frequency of kidney marrow (KM) cells in smears of WT (n=9) and *gata2b*$^{+/-}$ (n=9) zebrafish. Differentiated erythrocytes were excluded from quantification. B) Frequency of dysplastic cells of the erythroid, myeloid and lymphoid lineage in KM smears of WT (n=9) and *gata2b*$^{+/-}$ (n=9) zebrafish. Representative pictures of **kidney marrow smears after** May-Grünwald-Giemsa staining of C) WT KM smears and D) *gata2b*$^{+/-}$ KM smears. 1) Binucleated erythroblast; 2) Blebbing in cytoplasm of proerythroblast; 3) Binucleated promyelocyte; 4) Lobed nucleus and micronucleus in erythroblast; 5) Multinucleated promyelocyte; 6) Irregular cytoplasm in erythroid precursor; 7) Lobed nucleus in erythrocytes; 8) Blebbing in cytoplasm of blast. * P value<0.5, ***P value<0.001. Data represents mean ± Standard error of the mean.

Single cell transcriptome analysis reveals unique subpopulations in WT and Gata2b$^{+/-}$ KM indicative of a cell maturation arrest.

Next, we sought to characterize the dysplastic cell population in *gata2b*$^{+/-}$ KM. Through flow cytometry, we sorted 4 cell populations based on light scatter (Fig 1A) and observed dysplastic cells in the progenitor and lymphoid + HSPCs gate of *gata2b*$^{+/-}$, indicating that dysplastic cells could be viably sorted (Fig. 2D, panel 7 and 8). However, flow cytometry did not identify a uniquely separated population of dysplastic cells, possibly because of their heterogeneity in shape. To identify the transcriptional characteristics of *gata2b*$^{+/-}$ and to assess small subpopulations (possibly dysplastic), otherwise masked in bulk RNA sequencing (Shalek et al., 2014; Villani et al., 2017), we performed single cell RNA sequencing (scRNA-seq). We sorted lymphoid and HSPC and progenitor (Fig. 3A) because they included the dysplastic cells and HSPCs. These cells were split in 19 clusters (Fig 3B) using the nearest neighbor algorithm in the R Seurat package (Butler et al., 2018). Each cluster was classified based on differentially expressed genes (Suppl. Fig. 2) and known differentiation markers. In short, cells of the erythroid lineage are known to express hemoglobins like *hbba1. itga2b* was used as a marker for thrombocytes and *mpeg1.1* for monocytes and *lysozyme (lyz)* for neutrophils. Furthermore, the differentiated populations were devoid of expression of proliferation genes like *mki67* and *pcna* indeed suggesting that these are differentiated cells and not a progenitor compartment (Fig 3C). The populations expressing markers like *CCAAT enhancer binding protein alpha* (*cebpa)* and *granulin1 (grn1)* were identified as myeloid progenitors 1 and 2. T cells were marked by expression of *tox, il2rb* and *dusp2* expression, NK cells were marked by *nkl.3, nkl.4* and *ccl33.3* and B cells were marked by expression of *CD37, CD79a*, *pax5* and high levels of the immunoglobulin *ighv1-4*. Corroborating a conserved lymphoid and myeloid differentiation program, the cluster proportion analysis indicated an overall similar distribution of cells within clusters between *gata2b*$^{+/-}$ and WT (Fig. 4A).

Even though clusters were evenly distributed, dysplastic cells were almost exclusively observed in *gata2b*$^{+/-}$ KM (Fig. 2B). This discrepancy lead us to hypothesize that the clustering did not separate dysplastic cells in one cluster but that they were included within erythroid and myeloid clusters. For a more precise subdivision of small populations, we subclustered erythroid and myeloid clusters (including their progenitors) and identified 5 significantly overrepresented subclusters in *gata2b*$^{+/-}$ KM (Suppl Fig. 3). These subclusters represented 5.6% of the total sequenced *gata2b*$^{+/-}$ cells, a fraction comparable to the amount of dysplastic cells observed in *gata2b*$^{+/-}$ KM. Additionally, these subclusters had downregulation of tubulin transcripts (not shown), suggesting a loss of cytoskeletal structure, a characteristic of dysplasia. However, these subclusters were not characterized by a unique transcriptome compared to other subclusters, indicating either that *gata2b*$^{+/-}$ dysplastic cells are transcriptionally undistinguishable or that cell clustering is a suboptimal method for their identification.

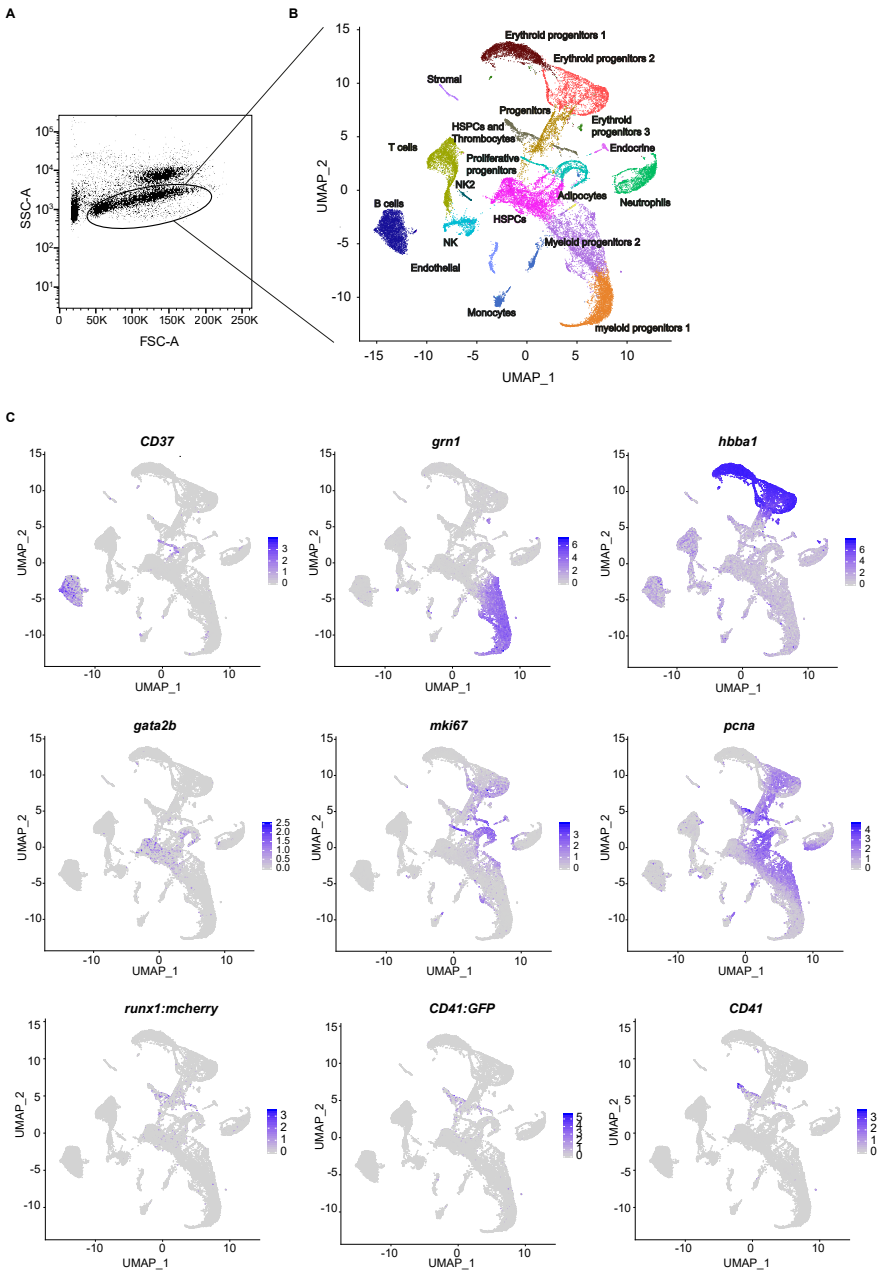


Fig. 3: Single cell analysis reveals multiple populations of differentiated and progenitor cells
A) Kidney marrow light scatter sorting strategy for scRNA sequencing analysis and B) UMAP clustering showing clusters in different colors with respective attributed nomenclature. D)Pooled WT and *gata2b*$^{+/-}$ UMAP feature analysis with gradual gene expression in shades of blue. Expression pattern of *cluster of differentiation 37* (*CD37),* granulin1 (*grn1),* hemoglobin beta adult 1 (*hbba1*), *GATA binding protein 2b* (*gata2b),* marker of proliferation Ki-67(*mKi67*), proliferating cell nuclear antigen (*pcna*), *mCherry* representing mCherry positive cells of the Runx:mCherry transgene, *GFP,* representing GFP expression from the CD41:GFP transgene and CD41, integrin alpha 2b (*itga2b*). FDR=False discovery rate. FD=Fold difference.

Further investigation of *gata2b*$^{+/-}$ cluster cells revealed that "myeloid progenitors 1" and "erythroid progenitors 2" clusters contain underrepresented subclusters (Fig 4B and C). The loss of these cells suggests that *gata2b*$^{+/-}$ cell maturation might be impaired, a phenomenon often found in dysplasia (Nováková et al., 2016). In line with the hypothesis of a block of differentiation, the *gata2b*$^{+/-}$ "erythroid progenitor 2" cluster, while having higher signatures of DNA replication, had a downregulation of transcripts necessary for cell mitosis (Fig. 4D). A block in the G2 phase of cell cycle can explain the origin of multi-lobated nuclei and other nuclear abnormalities observed in *gata2b*$^{+/-}$ dysplastic cells. Since cell divisions are crucial for the differentiation program, Gata2b haploinsufficiency could cause a block in cell cycle progression in a fraction of erythroid and myeloid cells, causing their loss in differentiation. Overall, in our single cell transcriptome analysis, dysplastic cells do not form a separate cluster but are likely scattered in the myeloid and erythroid progenitor clusters and possibly originate as a result of proliferative alterations causing a differentiation block.

gata2b$^{+/-}$ HSPCs have a high nucleic acid metabolism and downregulated cytoskeletal and protein ubiquitination transcripts

Multiple observations point to Gata2b affecting HSPC behavior. In *gata2b*$^{+/-}$ KM, dysplasia was observed in both myeloid and erythroid cells, suggesting an aberrancy derived from a common progenitor. Additionally, Gata2b expression was present mostly in the HSPC cluster (Fig 3C) and Gata2b knockout was shown to affect HSPCs in zebrafish (Gioacchino et al., 2021). Unfortunately, the HSPC population in the single cell analysis showed great heterogeneity with scattered HSPCs expressing differentiation markers like *ighv1-4* indicative of B cell differentiation, *Nkl.2* indicative of NK-differentiation and *icn* indicative of myeloid differentiation (Suppl Fig. 2). Therefore, we further purified the HSC population using CD41:GFP^{low} expressing cells (Fig. 5A), as these were shown to contain transplantable HSCs (Lin et al., 2005; Ma et al., 2011). Although numerically unaffected by *gata2b* heterozygosity (Fig. 5B), gene expression was markedly different in *gata2b*$^{+/-}$ CD41:GFP^{low} cells compared to WT HSPCs. Gene set enrichment (GSEA) and KEGG analysis revealed 71 upregulated and 80 downregulated gene sets (P<0.05). In *gata2b*$^{+/-}$ CD41:GFP^{low} cells, upregulated gene sets were related to DNA replication and ribosome biogenesis. Simultaneously, in *gata2b*$^{+/-}$ HSPCs, downregulated gene sets were related to cytoskeleton and protein degradation (Fig. 5C-D). These data suggest that *gata2b*$^{+/-}$ HSPCs have an aberrant expression of genes controlling cell structure and replication, which is possibly at the source of the dysplasia observed in the kidney marrow of *gata2b*$^{+/-}$ zebrafish.

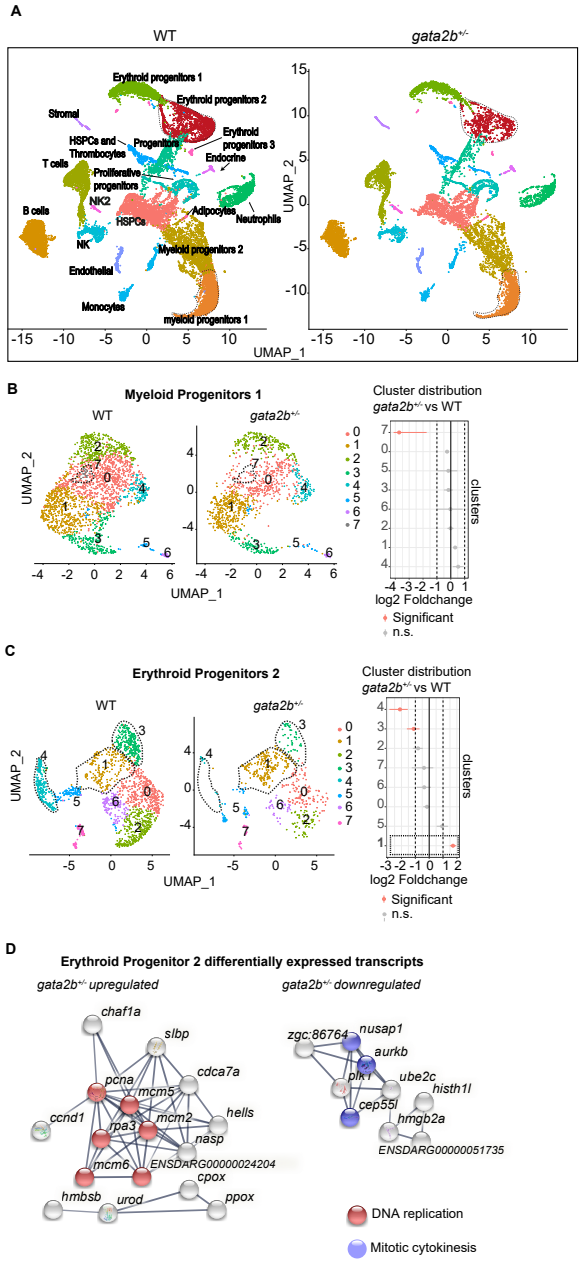


Fig. 4: *gata2b*$^{+/-}$ contains differentially-represented subclusters in myeloid and erythroid progenitors
A) Split UMAP representing cells in the clusters (with different colors) separating WT and *gata2b*$^{+/-}$ cells. B) Split UMAP representing the cell distribution and subclusters in the Myeloid progenitors 1 and C) Erythroid progenitors 2 (distinguishable by different colors) in WT and *gata2b*$^{+/-}$ with respective quantification of distribution between genotype. Significantly differentially distributed subclusters in pink. FDR<0.05 & Log2 fold change >1. D) STRING network of upregulated and downregulated transcripts in *gata2b*$^{+/-}$ "erythroid progenitors 2". Only networks with more than 2 interactions were represented. Highlighted in red DNA replication genes from KEGG pathways and highlighted in blue mitotic cytokinesis genes form Biological processes (gene Ontology). Fold change >0.05 & adjusted P value <0.05.

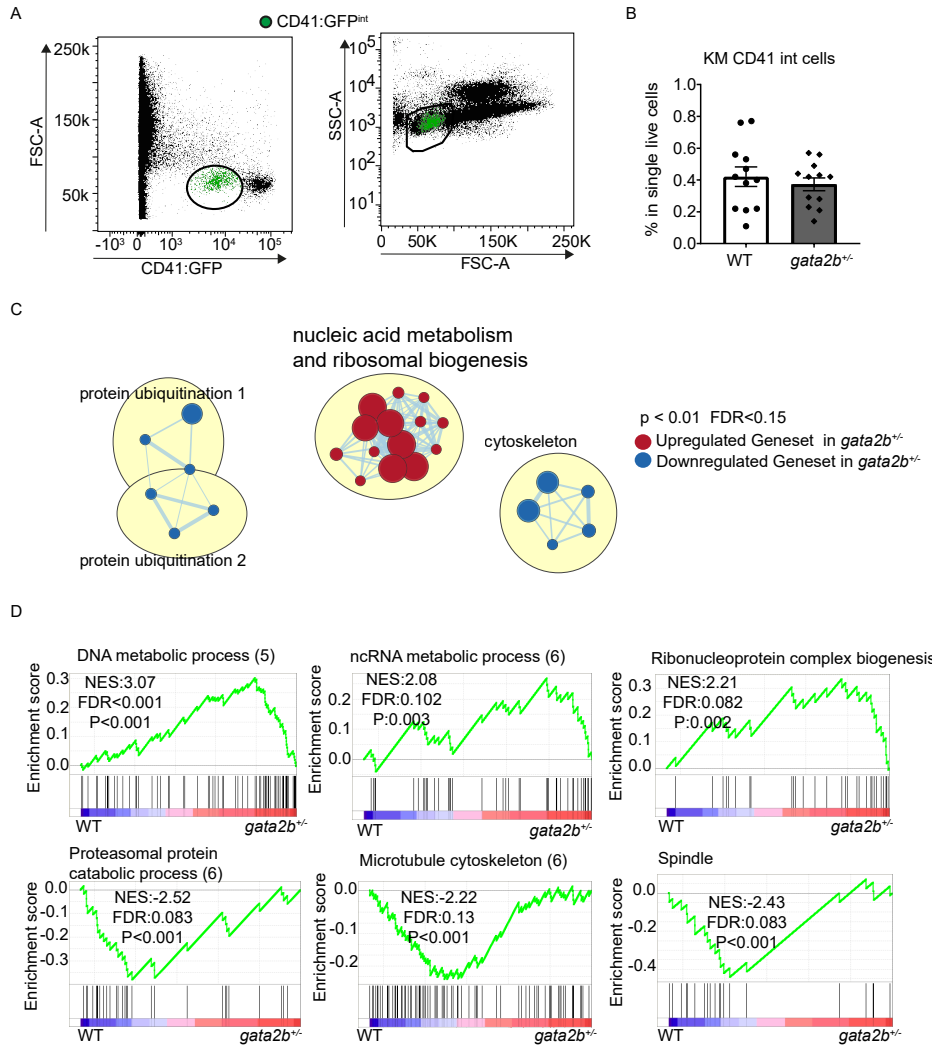


Fig. 5: *gata2b*$^{+/-}$ HSPCs have a high nucleic acid metabolism and downregulated cytoskeletal and protein ubiquitination transcripts
A) Representative figure of identification of CD41:GFP^{int} population (in green) and distribution in the FSC-A SSC-A kidney marrow population. B) Quantification of the percentage of CD41:GFP^{int} cells in WT and *gata2b*$^{+/-}$ kidney marrow single live cells. C) Cytoscape 3.8.2 enrichment map depicting significantly ($p < 0.01$ and FDR<0.15) upregulated and downregulated gene sets. The clusters in yellow are generated using the AutoAnnotation app of Cytoscape 3.8.2, the clusters names were manually edited. D) Representative gene set enrichments plots showing the profile of the enrichment score (ES) and positions of gene set members on the rank ordered list.

DISCUSSION

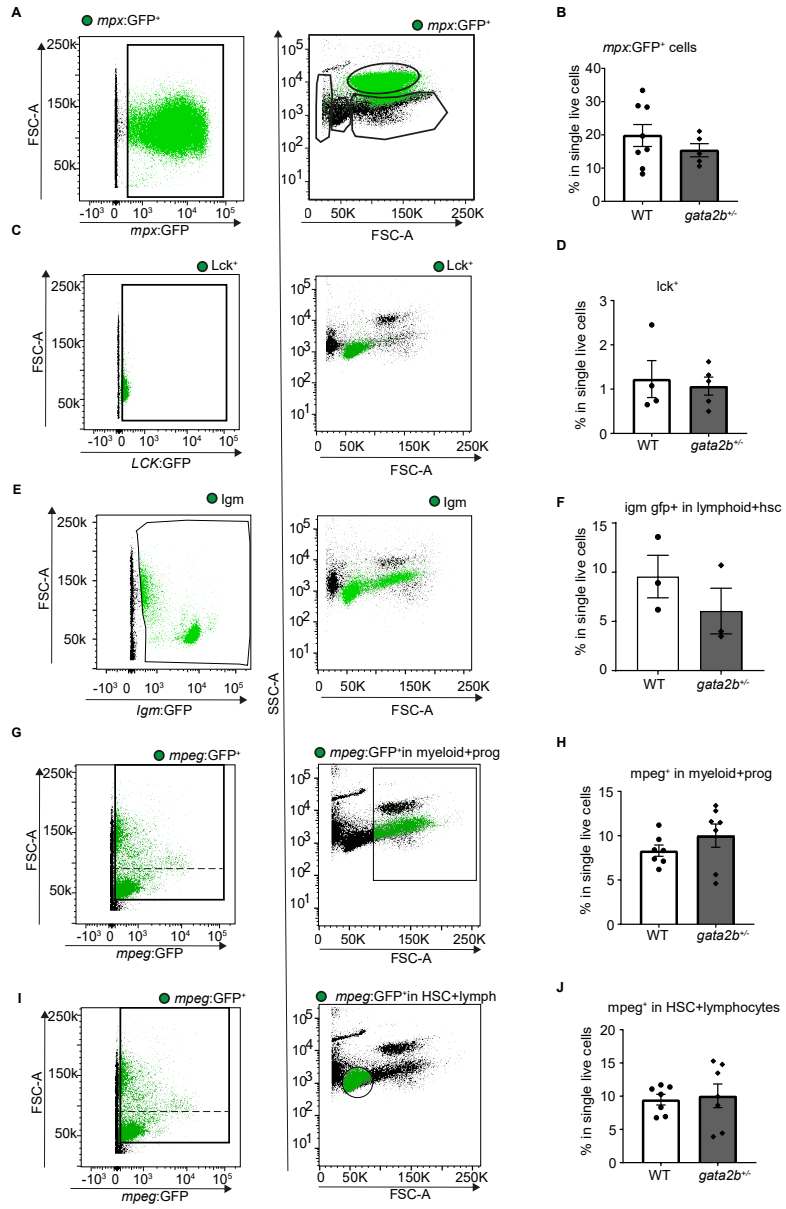
In humans, a balanced GATA2 expression is essential for proper hematopoiesis. Consequently, more than 90% of GATA2 mutant patients develop a life-threatening hematologic disease by 40 years of age (Donadieu et al., 2018). While the consequences of GATA2 mutations became obvious in the last decades, the regulation of GATA2 activity, and the relationship with human bone marrow failure remain incompletely understood.

Here, we used transgenic reporters, morphological phenotyping and RNA sequencing to analyze *gata2b* heterozygous zebrafish. We reveal that, while major differentiation lineages remain intact, *gata2b* heterozygosity causes dysplasia in erythroid and myeloid progenitors of zebrafish kidney marrow. Single cell RNAseq analysis does not identify a single population identifiable as dysplastic and dysplastic cell types are likely scattered through erythroid and myeloid clusters. In the future, the isolation of single dysplastic cells could help us define their transcriptome, meanwhile a tubulin reporter zebrafish (Goldman et al., 2001) could reveal whether the loss of tubulin is an attribute of *gata2b*$^{+/-}$ dysplasia.

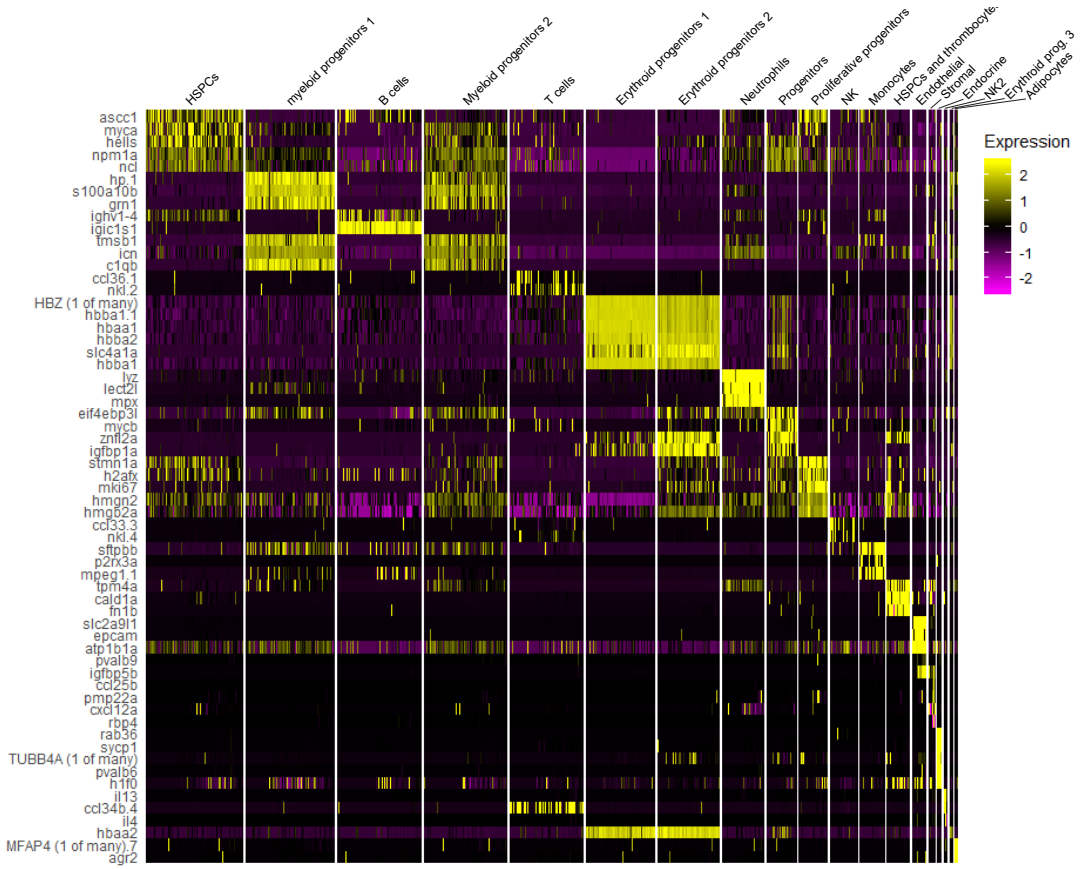
However, after isolating the most primitive HSPCs (CD41 intermediate), *gata2b*$^{+/-}$ hyperproliferation is evident and transcriptionally resembles the loss of quiescence observed in Gata2b knockout (*gata2b*$^{-/-}$) (Gioacchino et al 2021).

Interestingly, whereas in *gata2b*$^{-/-}$ there is an abrogation of the myeloid lineage and a bias to lymphoid differentiation (Gioacchino et al 2021), *gata2b*$^{+/-}$ did not simply result in an intermediate phenotype between WT and *gata2b*$^{-/-}$. Instead, it developed a unique dysplasia, not observed in *gata2b*$^{-/-}$. The differences in the homozygous and heterozygous *gata2b* knockout phenotype support a role of gene dosage in the severity of the GATA2 deficiency phenotype, possibly explaining the disease heterogeneity between patients. Since both erythroid and myeloid dysplasia can be observed in GATA2 patients (Donadieu et al., 2018), we propose that the dysplastic cells in *gata2b*$^{+/-}$ represent an aspect of the clinical syndromes associated with *GATA2* heterozygosity. It remains to be established how *gata2b*$^{+/-}$ HSPCs would respond to insults such as infections or severe bleeding.

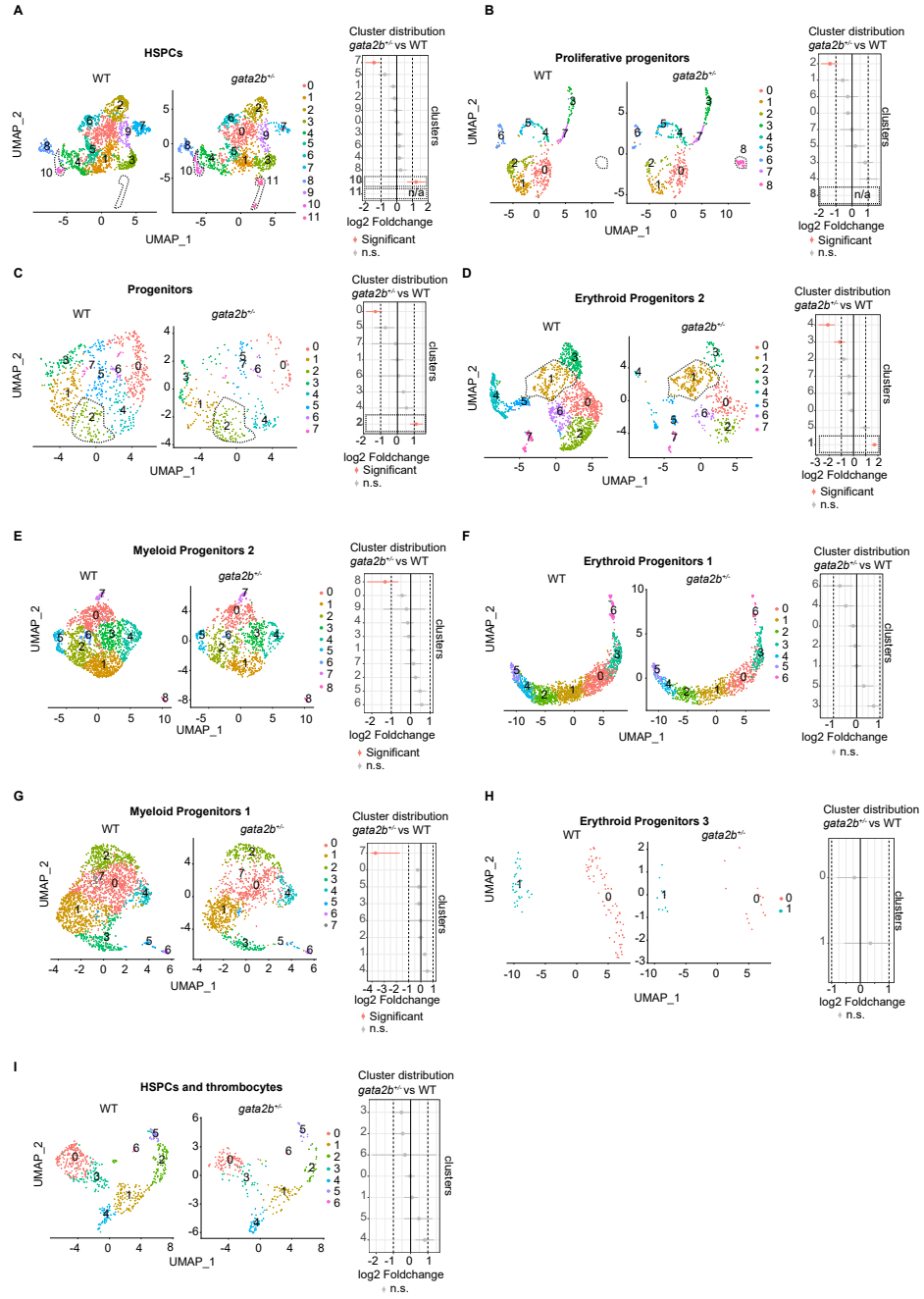
In conclusion, while retaining the major differentiation lineages intact, *gata2b*$^{+/-}$ zebrafish develop a hyperproliferative stem and progenitor compartment which leads to the generation of erythroid and myeloid dysplastic cells. Taken together, our model provides insights into the consequences of Gata2b dosage in normal development, revealing transcriptional networks affected by Gata2b haploinsufficiency in zebrafish.



Supplementary Fig. 1: Transgenic zebrafish reporter lines show no overall differentiation difference in *gata2b*$^{+/-}$
A) *mpx* positive cells, depicted in green, with distribution in FSC-A SSC-A graph. Followed by B) Quantification of the percentage of mpx:GFP+ cells in WT and *gata2b*$^{+/-}$ kidney marrow single live cells. C) *Lck* positive cells, depicted in green, with distribution in FSC-A SSC-A graph. Followed by D) Quantification of the percentage of lck:GFP+ cells in WT and *gata2b*$^{+/-}$ kidney marrow single live cells. E) *Igm* positive cells, depicted in green, with distribution in FSC-A SSC-A graph. Followed by F) Quantification of the percentage of IgM:GFP+ cells in WT and *gata2b*$^{+/-}$ kidney marrow single live cells. G) *mpeg* positive cells, depicted in green, mark monocytes (top) and phagocytic B-cells (bottom). Followed by H-I) Quantification of the percentage of mpeg:GFP+ cells in WT and *gata2b*$^{+/-}$ kidney marrow single live cells. Data represents mean ± Standard error of the mean.



Supplementary Fig. 2: Heatmap with top 5 transcripts per cluster
Heatmap representing the expression level of the top 5 expressed transcripts per cluster. Transcripts highly expressed in multiple clusters were not repeated in the list (like hemoglobin in erythroid clusters). Transcripts identified only by their chromosome location were not included.



Supplementary Fig. 3: *gata2b*$^{+/-}$ cells have a different distribution of subclusters compared to WT
Split UMAP representing the cell distribution in the various subclusters (distinguishable by different colors) in WT and *gata2b*$^{+/-}$ for A) HSPCs B) Proliferative progenitors C) Progenitors D) Erythroid progenitors 2 E) Myeloid progenitors 2, F) Erythroid progenitors 1, G) Myeloid progenitors 1, H) Erythroid progenitors 3, I) HSPCs and thrombocytes. With respective quantification of distribution between genotypes. Significantly differentially distributed clusters in pink. FDR<0.05 & Log2 fold change >1.

REFERENCES

Boatman, S., Barrett, F., Satishchandran, S., Jing, L., Shestopalov, I., and Zon, L.I. (2013). Assaying hematopoiesis using zebrafish. Blood Cells, Mol. Dis. *51*, 271–276.

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. Dev. *142*, 1050–1061.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Dickinson, R.E., Griffin, H., Bigley, V., Reynard, L.N., Hussain, R., Haniffa, M., Lakey, J.H., Rahman, T., Wang, X.N., McGovern, N., et al. (2011). Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. Blood *118*, 2656–2658.

Dobrzycki, T., Mahony, C.B., Krecsmarik, M., Koyunlar, C., Rispoli, R., Peulen-Zink, J., Gussinklo, K., Fedlaoui, B., de Pater, E., Patient, R., et al. (2020). Deletion of a conserved Gata2 enhancer impairs haemogenic endothelium programming and adult Zebrafish haematopoiesis. Commun. Biol. *3*, 1–14.

Donadieu, J., Lamant, M., Fieschi, C., De Fontbrune, F.S., Caye, A., Ouachee, M., Beaupain, B., Bustamante, J., Poirel, H.A., Isidor, B., et al. (2018). Natural history of GATA2 deficiency in a survey of 79 French and Belgian patients. Haematologica *103*, 1278–1287.

Gao, X., Johnson, K.D., Chang, Y.I., Boyer, M.E., Dewey, C.N., Zhang, J., and Bresnick, E.H. (2013). Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. J. Exp. Med. *210*, 2833–2842.

Goldman, D., Hankin, M., Li, Z., Dai, X., and Ding, J. (2001). Transgenic zebrafish for studying nervous system development and regeneration. Transgenic Res. *10*, 21–33.

Hahn, C.N., Chong, C.E., Carmichael, C.L., Wilkins, E.J., Brautigan, P.J., Li, X.C., Babic, M., Lin, M., Carmagnac, A., Lee, Y.K., et al. (2011). Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. Nat. Genet. *43*, 1012–1019.

Hsu, A.P., Sampaio, E.P., Khan, J., Calvo, K.R., Lemieux, J.E., Patel, S.Y., Frucht, D.M., Vinh, D.C., Auth, R.D., Freeman, A.F., et al. (2011). Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. Blood *118*, 2653–2655.

Kordasti, S.Y., Afzali, B., Lim, Z., Ingram, W., Hayden, J., Barber, L., Matthews, K., Chelliah, R., Guinn, B., Lombardi, G., et al. (2009). IL-17-producing CD4+ T cells, pro-inflammatory cytokines and apoptosis are increased in low risk myelodysplastic syndrome. Br. J. Haematol. *145*, 64–72.

Lin, H.F., Traver, D., Zhu, H., Dooley, K., Paw, B.H., Zon, L.I., and Handin, R.I. (2005). Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. Blood *106*, 3803–3810.

Ma, D., Zhang, J., Lin, H.F., Italiano, J., and Handin, R.I. (2011). The identification and characterization of zebrafish hematopoietic stem cells. Blood *118*, 289–297.

Mutsaers, P.G.N.J., Van De Loosdrecht, A.A., Tawana, K., Bödör, C., Fitzgibbon, J., and Menko, F.H. (2013). Highly variable clinical manifestations in a large family with a novel GATA2 mutation. Leukemia *27*, 2247–2248.

Ostergaard, P., Simpson, M.A., Connell, F.C., Steward, C.G., Brice, G., Woollard, W.J., Dafou, D., Kilo, T., Smithson, S., Lunt, P., et al. (2011). Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). Nat. Genet. *43*, 929–931.

Ou, R., Huang, J., Shen, H., Liu, Z., Zhu, Y., Zhong, Q., Zheng, L., Yao, M., She, Y., Zhou, S., et al. (2017). Transcriptome analysis of CD34+ cells from myelodysplastic syndrome patients. Leuk. Res. *62*, 40–50.

de Pater, E., Kaimakis, P., Vink, C.S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P.S., Camper, S.A., Speck, N., and Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. J. Exp. Med. *210*, 2843–2850.

Rinker, E.B., Dueber, J.C., Qualtieri, J., Tedesco, J., Erdogan, B., Bosompem, A., and Kim, A.S. (2016). Differential expression of ribosomal proteins in myelodysplastic syndromes. J. Clin. Pathol. *69*, 176–180.

Rodrigues, N.P., Janzen, V., Forkert, R., Dombkowski, D.M., Boyd, A.S., Orkin, S.H., Enver, T., Vyas, P., and Scadden, D.T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. Blood *106*, 477–484.

Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature *510*, 363–369.

Spinner, M.A., Sanchez, L.A., Hsu, A.P., Shaw, P.A., Zerbe, C.S., Calvo, K.R., Arthur, D.C., Gu, W., Gould, C.M., Brewer, C.C., et al. (2014). GATA2 deficiency: A protean disorder of hematopoiesis, lymphatics, and immunity. Blood *123*, 809–821.

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

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

Villani, A.C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science (80-.). *356*.

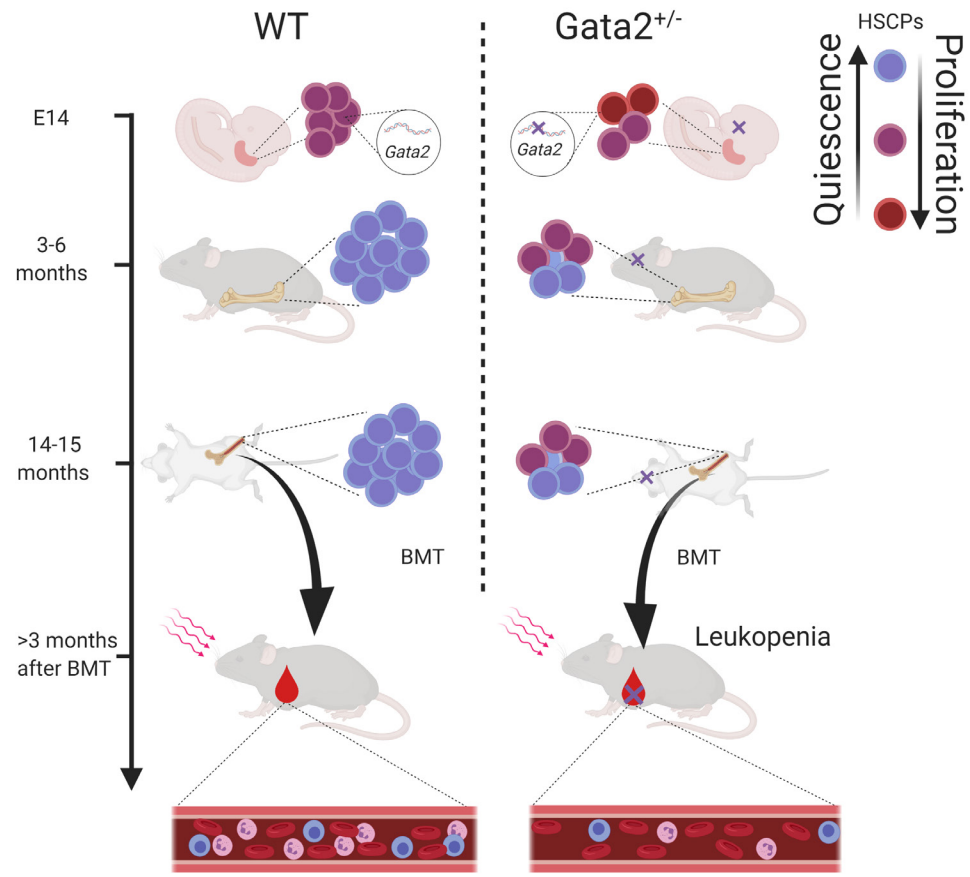
Wang, X., Muramatsu, H., Okuno, Y., Sakaguchi, H., Yoshida, K., Kawashima, N., Xu, Y., Shiraishi, Y., Chiba, K., Tanaka, H., et al. (2015). GATA2 and secondary mutations in familial myelodysplastic syndromes and pediatric myeloid malignancies. Haematologica *100*, e398–e401.

4

GATA2 haploinsufficiency reduces fitness of aged hematopoietic stem cells

Emanuele Gioacchino, Cansu Koyunlar, Hans de Looper, Joke Peulen, Mariëtte Ter Borg, Remco Hoogenboezem, Dennis Bosch, Paulette van Strien, Kirsten J Gussinklo, Eric Bindels, Elaine Dzierzak, Ivo Touw and Emma de Pater.

Manuscript In Submission



Created with BioRender.com

ABSTRACT

GATA2 is a transcription factor with key functions in both embryonic and adult hematopoiesis. Patients with germline *GATA2* mutations are predisposed to develop a broad range of pathological phenotypes, ranging from immunodeficiencies to acute myeloid leukemia. This phenotype heterogeneity, associated with hundreds of possible mutations and variable penetrance, caused GATA2 pathogenic contribution to be enigmatic. We show here that *Gata2*$^{+/-}$ hematopoietic stem and progenitor cells (HSPCs) lose quiescence from embryonic day (E)14 to 16 months. Despite this HSC phenotype, *Gata2*$^{+/-}$ animals did not develop hematological anomalies. To challenge the HSCs, aged WT and *Gata2*$^{+/-}$ HSCs were transplanted into lethally irradiated recipients. Mice transplanted with aged *Gata2*$^{+/-}$ BM cells developed leukopenia, recapitulating the immunodeficiency reported in GATA2 patients and a measure for the reduction in fitness of *Gata2*$^{+/-}$ HSCs. Our results highlight the role of GATA2 in maintaining HSCs homeostatic quiescence and fitness and suggest that the combination of GATA2 haploinsufficiency, ageing and replicative stress are necessary and sufficient to cause a GATA2 deficiency manifestation.

INTRODUCTION

Hematopoietic stem cells (HSCs) are characterized by their quiescence state (Nakamura-Ishizu et al., 2014; Orford & Scadden, 2008) and by their ability to generate all hematopoietic cells via multipotency and self-renewal (Ogawa, 1993; Seita & Weissman, 2010). These properties allow HSCs to sustain the hematopoietic system throughout life. Due to the vast and continuous demand of differentiated blood cells, HSCs fitness is critical for the balanced production of cells of the hematopoietic system. Regulation of transcription, driven by transcription factors like GATA2, is key in the control of HSCs and provides a crucial level of control against immune diseases (Bhagwat & Vakoc, 2015).

Human *GATA2* germline mutations are the cause of the bone marrow failure syndrome comprised of a series of disorders with variable penetrance assembled under the GATA2 deficiency syndromes (Rachel Emma Dickinson et al., 2011; Hsu et al., 2011; Ostergaard et al., 2011). 73% of patients suffer from various cytopenias at diagnosis (49% monocytopenia, 39% neutrophenia and consistent B-cell lymphopenia) 15% of patients suffer from lymphedema and more than 80% develop MDS or leukemia (Donadieu et al., 2018). Hundreds of pathogenic *GATA2* mutations have been identified (Collin et al., 2015; Wlodarski et al., 2016, 2017). Roughly 90% of these are either missense mutations within zinc finger (ZF) 2 or truncating mutations preceding ZF2 rendering one copy of the protein nonfunctional (Wlodarski et al., 2017). Therefore, *GATA2* mutations are considered to cause GATA2 haploinsufficiency (Cortés-Lavaud et al., 2015). However, the different clinical phenotypes are not strictly linked to specific mutation types and it is currently unclear how GATA2 haploinsufficiency contributes to the different aspects of GATA2 deficiency syndromes (Hahn et al., 2011; Spinner et al., 2014). Interestingly, patients carrying *GATA2* germline mutations have much greater risk of developing clinical symptoms and MDS in older age (Donadieu et al., 2018; Spinner et al., 2014): 38% chance to remain asymptomatic by age 20 and 8% by age 40, and the risk of developing MDS goes from 6% at 10 years, to 81% at age of 40 (Donadieu et al., 2018). This indicates that the hematopoietic system of patients carrying germline *GATA2* mutations accumulate molecular changes that lead to the progression to leukemia.

The *Gata2* gene is required for both the embryonic generation and maintenance of HSCs in adulthood (de Pater et al., 2013) and is well conserved between humans and mice (Mazumder et al., 2016). During mammalian hematopoietic development, definitive HSCs originate in the aorta-gonad-mesonephrons (AGM) region (Dzierzak & Speck, 2008; Ivanovs et al., 2011; Medvinsky & Dzierzak, 1996) via endothelial to hematopoietic transition (EHT). HSCs then migrate to the fetal liver where they amplify before populating the bone marrow around birth (E18-21 in mouse) and reside throughout adulthood. *Gata2* germline knock-out mice (*Gata2*$^{-/-}$) are embryonically lethal at embryonic day (E) 10, just before the generation of the first HSCs, due to hematopoietic failure (Tsai et al., 1994). Zebrafish studies

showed the involvement of Gata2b, the zebrafish orthologue of GATA2, to function in hematopoietic lineage differentiation within the HSC compartment (Gioacchino et al 2021). *Gata2* heterozygous knock-out mice (*Gata2*$^{+/-}$) are viable, breed normally and have normal hematologic profiles; despite showing a reduction in hematopoietic stem and progenitor cells (HSPCs) during both embryogenesis (de Pater et al., 2013; Ling et al., 2004; Tsai et al., 1994) and adulthood (Guo et al., 2013; Rodrigues et al., 2005).

Because GATA2 plays pivotal roles in the hematopoietic system throughout development and because the incidence of MDS/AML increases with age in patients with GATA2 deficiency syndrome, we hypothesize that molecular changes in embryonic HSCs eventually lead to the bone marrow failure syndrome observed in patients carrying germline GATA2 mutations. Therefore, we assessed the hematopoietic system from embryonic hematopoiesis throughout adulthood and after aging. The transplantation of aged HSCs resulted in pancytopenia and reduced HSC fitness in *Gata2*$^{+/-}$ mice compared to WT. Molecular assessment of the HSC compartment at various time-points revealed that, throughout development, HSCs lose quiescence and are under proliferative stress. This abnormalities progress into a reduction in HSCs fitness and manifests as a bone marrow failure syndrome, similar to patients carrying germline *GATA2* mutations.

MATERIALS AND METHODS

Mouse maintenance

Gata2 heterozygous mice have been previously described (Tsai et al., 1994). Animal studies were approved by the Animal Welfare/Ethics Committee of the EDC in accordance with legislation in the Netherlands.

Bone marrow sampling

Mouse bone marrow (BM) was obtained by flushing femurs and tibias using a 25G (BD Microlance) needle with Phosphate buffer solution (PBS) supplemented with 5 IU/mL penicillin, 5 µg/mL streptomycin, and 10% fetal calf serum (FCS). The bone marrow cell suspension was filtered using a 30-micron nylon mesh, Ficoll treated and stored on ice until use. An aliquot of the BM cell suspension was diluted with phosphate-buffered saline (PBS) and loaded into a hemocytometer to count nucleated cells.

Peripheral blood count

Blood was sampled from cheek puncture and collected in a Microtainer Brand Tube with EDTA (ethylenediaminetetraacetic acid;). Samples were analyzed on a Hemavet 850 (Drew scientific).

Colony-forming unit assay

MethoCult GF M3434 (Stem Cell Technologies, Vancouver, BC, Canada) was defrosted overnight at 4°C and used according to manufacturer's instructions to enumerate colony-forming units (CFU). For optimal colony growth, 100 freshly sorted LSK SLAM cells were plated in 1,1ml of Methocult placed in 10mm style Falcon petri dish (Corning Incorporated) and colonies scored after 10 days of culture at 37°C with 5% CO2. After colony count, MethoCult was removed from the plates using PBS with 10% FCS at 37°C, cells counted and 10^4 cells re-plated for 1st,2nd and 3rd re-plating.

Growth of primitive erythroid progenitor cells (BFU-E), granulocyte-macrophage progenitor cells (CFU-GM, CFU-G and CFU-M), and multi-potential granulocyte, erythroid, macrophage, megakaryocyte progenitor cells (CFU-GEMM) were scored using an inverted microscope.

LSK SLAM staining

BM cells were labeled with lineage antibodies:
Rat Anti Mouse CD45R/B220 Clone RA3-6B2 (cat 553089), PE Rat Anti-Mouse CD3 Molecular Complex Clone 17-A2 (Cat 555275), Rat Anti-CD11b Clone M1/70 (Cat 553311), Rat Anti-Mouse TER-119 (Cat 553673), Rat Anti-Mouse Ly-6G and Ly-6C (Gr1) Clone RB6-8C5 (Cat 553128) (all from BD Bioscience). To identify murine hematopoietic stem and progenitor

cells, BM cells were co-stained with APC Rat Anti Mouse CD117 Clone:2B8 Cat:553356 (BD Bioscience), BB700 Rat anti Mouse Ly-6A (sca1) Clone D7 Cat 742089 (BD Bioscience), Alexa Fluor 700 anti-mouse CD48 Cat 103426 (BioLegend), Rat anti-mouse CD150 Cat 115914 (BioLegend). All LSK SLAM FACS antibodies incubations were performed in PBS+10% FCS for 20 min on ice. Fetal liver LSK SLAM analyses were performed according to (I. Kim et al., 2006).

Cell proliferation and apoptosis assays and γH2AX

Apoptosis was analyzed using FITC Annexin V Apoptosis Detection Kit I (BD Biosciences) according to the recommendation of the manufacturer. To detect death from apoptotic cells, Annexin-V was used together with 4,6-diamidino-2-phenylindole (DAPI) (Molecular Probes).

For Cell cycle analysis$1x10^6$ freshly isolated BM cells, stained with surface markers, were fixed while vortexing using 2% PFA and incubated for 1 hour at 4 degrees in the dark. After washing, cells were permeabilized re-suspending the cell pellet in PBS/0.2% triton with Dapi 1:500 and incubated overnight at 4°C in the dark. After a washing step, cells were stained using Ki67 FITC (Thermofisher, catalog #11-5698-82) 1:25 in PBS 10% FCS with DAPI 1:500 and incubated for 2 hours at 4 degrees in the dark vortexing every 30 minutes. After washing, cells were resuspended in PBS with DAPI 1:500 and analyzed (Szade et al., 2016). γH2AX levels were assessed in cells fixed and permeabilized with Cytofix/Cytoperm Fixation/ Permeabilization Solution Kit (BD Biosciences) by incubating cells with Alexa Fluor 488 mouse anti-γH2AX (N1-431, Cat 560445, BD Biosciences), diluted in 1X Perm/Wash buffer (BD Biosciences) 1:200. All FACS events were recorded using a BD LSR II Flow Cytometer or a BD FACSAria III and analyzed with FlowJo 7.6.5 software (Tree Star). Cells were sorted with a BD FACSAria III.

RNA extraction and RNA quality control

Total sample RNA isolation was performed according to the standard protocol of RNA isolation with Trizol and GenElute LPA (Sigma). Quality and quantity of the total RNA was checked on a 2100 Bioanalyzer (Agilent) using the Agilent RNA 6000 Pico Kit.

RNA Sequencing and gene set enrichment analysis (GSEA)

cDNA was prepared using SMARTer procedure with SMARTer Ultra Low RNA kit (Clontech) for Illumina Sequencing. SMARTer Ultra Low RNA Kit (Clonetech) for Illumina Sequencing was used for cDNA based. The gene expression values were measured as FPKM (Fragments per kilobase of exon per million fragments mapped). Fragment counts were determined per gene with HTSeq-count, utilizing the strict intersection option, and subsequently used for differential expression analysis using the DESeq2 package, with standard parameters, in the R environment. Gene expression is measured as FPKM (Fragment per kilobase of exon per million fragments mapped). GSEA (Gene Set Enrichment Analysis) was performed on the FPKM values using the curated C2 collection of gene sets.

Serial bone marrow transplantation

$3x10^6$ freshly isolated nucleated bone marrow $CD45.2^+$ cells were transplanted via tail vein injection into lethally irradiated (10.5Gy) 6-8 weeks old $CD45.1^+$ recipient mice. The donor cell chimerism was determined in peripheral blood every month after transplantation. Blood was lysed from red blood cells and assessed using Gr1, CD11b, CD3, CD19 (eBioD3). For secondary transplant, $3x10^6$ bone marrow cells were collected from recipients and transplanted into lethally irradiated $CD45.1^+$ recipient mice.

Cytoscape analysis

Enrichment Map and AutoAnnotate applications were used on Cytoscape 3.8.2 (Shannon et al., 2003) to obtain GeneSet clusters. Some of the cluster names have been manually edited from the original autoannotation for a better definition. False discovery rate cutoff =0,25, P value cutoff=0,01, Edge cutoff=0,4, Node cutoff Q value =0,05. Only clusters with 5 nodes for E14 and clusters with 10 nodes for adult and aged samples were presented.

Statistics

Data are presented as mean ± SEM. All statistical analysis was carried out in GraphPad Prism 8.0.1 (GraphPad Software Inc., San Diego, CA). Normally distributed data were analyzed using an unpaired t test, not normally distributed data were analyzed using Mann-Whitney test. A p value less than 0.05 was considered significant.

RESULTS

The HSPC compartment of Gata2$^{+/-}$ embryos is hyperproliferative

Because GATA2 haploinsufficiency syndromes are innate genetic disorders and the associated phenotypes worsen with age, we hypothesize that, during embryogenesis, it is possible to detect early changes supporting malignancy. Therefore, we investigated the transcriptome of phenotypic HSCs, expressing Lin$^-$Sca1$^+$cKit$^+$CD48$^-$CD150$^+$Mac-1$^+$ (LSK SLAM), at embryonic day (E) 14. At this time point, HSCs frequently divide and are greatly expanded (Ema & Nakauchi, 2000; I. Kim et al., 2006; Kumaravelu et al., 2002). Our results show that GATA2 haploinsufficient mouse (*Gata2*$^{+/-}$) embryos possess a reduced fetal liver (FL) HSPC compartment compared to wild type (WT) at E14 (Fig. 1a). The HSPC reduction is not caused by a slower amplification in *Gata2*$^{+/-}$, as it is was demonstrated not to be delayed (Ling et al., 2004 and unpublished results). Hierarchical clustering by principal component (PC) analysis of the RNA-Seq data revealed an evident separation of the *Gata2*$^{+/-}$ cells by the first principal component (PC_1) (Fig. 1b). To determine the correlation between WT and *Gata2*$^{+/-}$ transcriptome we performed Gene Set Enrichment Analysis using curated gene sets. We selected significant differentially expressed genes (p<0.01 and FDR<0.25) and performed Cytoscape analysis (Otasek et al., 2019; Shannon et al., 2003). Gene-sets (circles) are connected (lines) if a gene is shared between gene sets. In this way, similar gene sets are grouped, creating a network, facilitating the interpretation of the results. While translation gene sets were downregulated, cell cycle and aberrant proliferation gene sets resulted the main group of gene sets upregulated in *Gata2*$^{+/-}$ E14 LSK SLAM cells compared to WT (Fig. 1c). For example, gene sets like “DNA replication”, “Myc targets UP”, “Cell cycle retinoblastoma (RB1) targets’’, and “Dividing vs Normal cells in chronic myeloid leukemia (CML)” were upregulated and “Peptide chain elongation” was downregulated in GATA2 haploinsufficient cells (Fig. 1d and Supplementary Fig.1a). Overall, *Gata2*$^{+/-}$ phenotypic HSCs are quantitatively reduced but transcriptionally more proliferative at E14, indicating that the reduction of HSC numbers in *Gata2*$^{+/-}$ E14 FL is not caused by a reduction in proliferation.

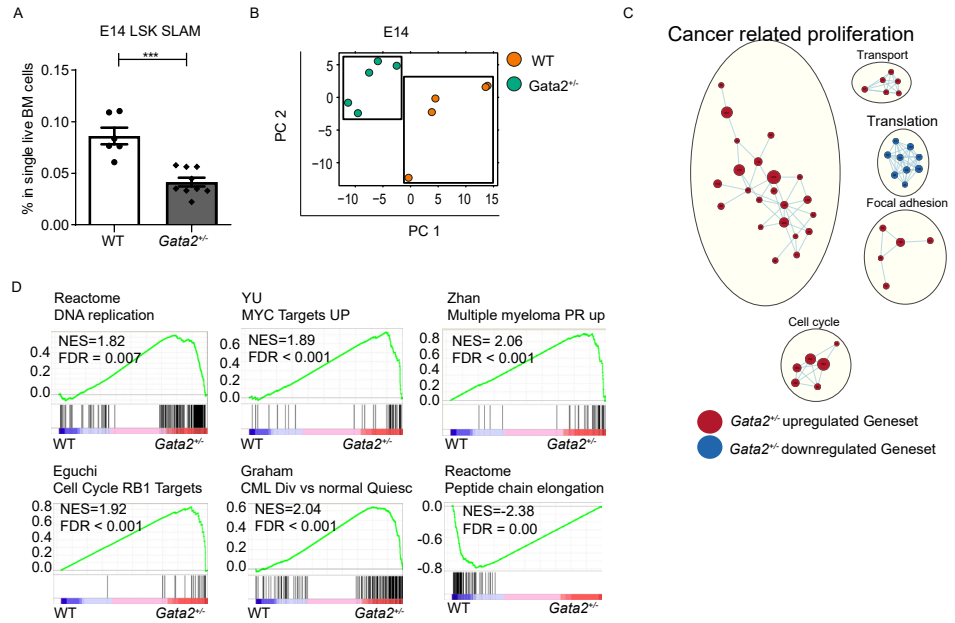


Figure 1. *Gata2*$^{+/-}$ embryonic HSPCs are diminished but transcriptionally more proliferative than WT A) Percentage of Lin$^-$Sca1$^+$cKit$^+$CD48$^-$CD150$^+$Mac-1$^+$ (LSK SLAM) cells in fetal liver (FL) in WT (n=6) and *Gata2*$^{+/-}$ (n=9) at embryonic day (E)14 (mean± SEM). B) PCA of E14 samples show separation by genotype on principal component (PC) 1. BM: Bone marrow. C) Cytoscape Enrichment map representing significant networks (False discovery rate cutoff =0,25, P value cutoff=0,01, Edge cutoff=0,4, Node cutoff Q value =0,05), based on curated gene sets. Only clusters with more than 4 nodes were presented. D) Representative significantly different curated gene sets. NES: Normalized enrichment score FDR: False discovery rate. ***p<0.001.

Gata2$^{+/-}$ HSPCs remain less quiescent during adulthood

Next, adult HSPCs were investigated similarly. Consistent with prior studies (Guo et al., 2013; Rodrigues et al., 2005), adult *Gata2*$^{+/-}$ BM have reduced numbers of phenotypic HSCs (Fig. 2a), suggesting a comparable reduction in functional HSCs. However, blood values remain unchanged compared to WT (Suppl. Fig. 2a). To elucidate how GATA2 haploinsufficiency affects HSC number and behavior, we analyzed the transcriptome of LSK SLAM using RNA sequencing. Just like FL LSK SLAM, the PC analysis in LSK SLAM from adult mice highlights a separation of the two genotypes based on PC1 (Fig. 2b). Cytoscape analysis reveals the formation of networks between gene sets related to proliferation like “Cell cycle”, “proliferation” and “respiratory electron transport” and aberrant proliferation gene sets like “Protein degradation” and “DNA repair” (Fig. 2c). Gene sets related to “cell cycle mitotic”, ”Myc targets”, “respiratory electron transport” and “proteasome pathway” were greatly enriched in *Gata2*$^{+/-}$ compared to WT (Fig. 2d and Suppl. Fig 2b), indicating that proliferation is upregulated in *Gata2*$^{+/-}$ LSK SLAM cells compared to WT. To test if the hyperproliferative transcriptome of BM HSCs results in increased proliferation of *Gata2*$^{+/-}$ HSCs, we conducted

a cell cycle analysis using Ki-67 and DAPI (K. H. Kim & Sederstrom, 2015; Szade et al., 2016; Whitfield et al., 2006). *Gata2*$^{+/-}$ LSK SLAM show a significant loss of quiescent G0 cells (WT 73.8±2.2 vs *Gata2*$^{+/-}$ 52.2±7.6) and an acquisition of cells in the cell cycle as G1 (WT 13.8±1.6 vs *Gata2*$^{+/-}$ 26.5±5.0) (Fig. 3a-b). In the less primitive *Gata2*$^{+/-}$ LSK population, we detected a significant increase in S cell cycle phase cells (WT 8.7 ± 1.0 vs *Gata2*$^{+/-}$ 15.8 ± 1.6) and an overall increase in proliferative cells (Suppl. Fig. 3a-b). The increase in cycling cells was specific for the more immature HSPC compartments (LSK and LSK SLAM) and not broadly present in the hematopoietic compartment like the LK cell population (Suppl. Fig. 3c). Taken together, these data indicate that *Gata2*$^{+/-}$ HSPCs gradually leave their quiescent state starting from embryonic stages.

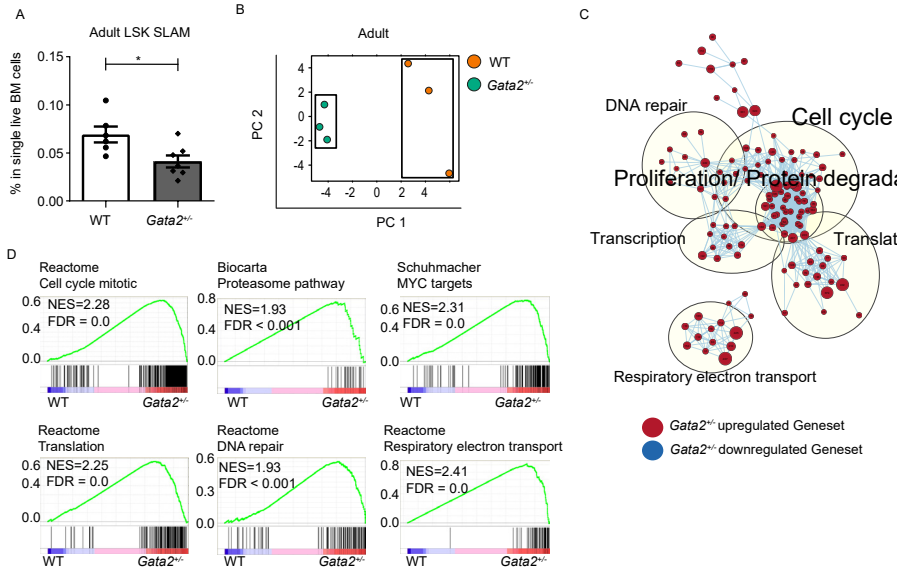


Figure 2

Figure 2. Adult *Gata2*$^{+/-}$ HSCs are transcriptionally hyperproliferative
A) Percentage of LSK SLAM cells in bone marrow (BM) in WT (n=6) and *Gata2*$^{+/-}$ (n=7) during adulthood (3-6 months) (mean± SEM). B) PCA of adult show separation by genotype on principal component (PC) 1. C) Cytoscape Enrichment map representing significant networks (False discovery rate cutoff =0,25, P value cutoff=0,01, Edge cutoff=0,4, Node cutoff Q value =0,05), based on curated gene sets. Only clusters with more than 10 nodes were presented. D) Representative significantly different curated gene sets. NES: Normalized enrichment score, FDR: False discovery rate. *p<0.05.

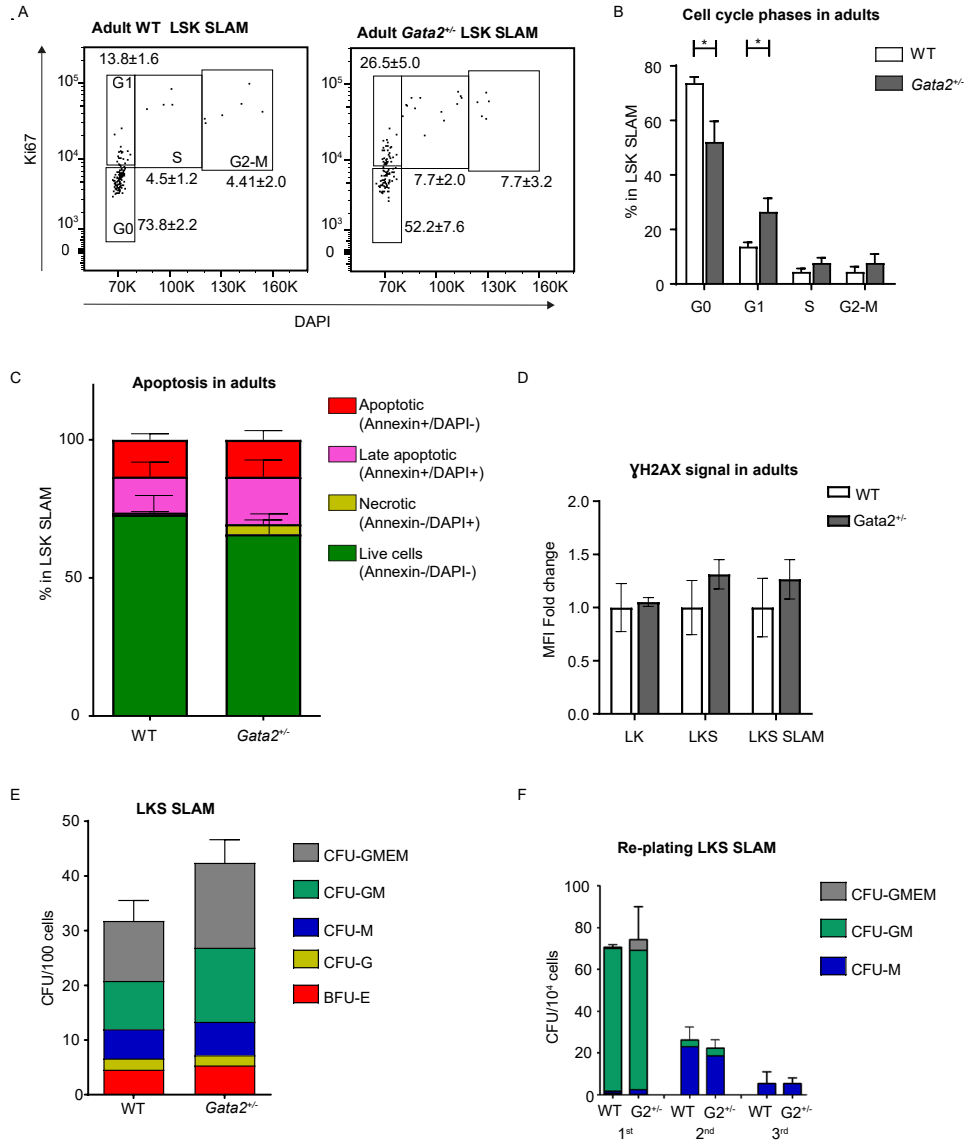


Figure 3. Adult *Gata2*$^{+/-}$ phenotypic HSCs have a reduced quiescence
A) Representative cell cycle flow cytometry plots of equal numbers of LSK SLAM in WT and *Gata2*$^{+/-}$ with gating strategy used. mean ±SEM. B) Column graph representing the difference found in the cell cycle phases. mean ±SEM, *p<0.05. C) Apoptosis analysis via flow cytometry of adult BM LSK SLAM in WT and *Gata2*$^{+/-}$ using AnnexinV and Dapi. mean±SEM. D) γH2AX MFI fold change by flow cytometry analysis of fixed adult BM in WT and *Gata2*$^{+/-}$ primitive hematopoietic compartments. E) Stack bar chart representing the CFU types obtained 10 days after plating 100 LSK SLAM in MethoCult. F) CFU types obtained after re-plating 10^4 cells isolate from LSK SLAM colonies. The same condition were used in 3 re-platings. BM; bone marrow MFI; mean fluorescent intensity. mean±SEM.

Gata2$^{+/-}$ HSPCs reduction does not affect myeloid progenitor potential

The loss of quiescence concomitant with the number reduction of *Gata2*$^{+/-}$ HSCs can have multiple explanations. Even without significant gene sets specifically related to apoptosis in *Gata2*$^{+/-}$ HSPCs transcriptome, we tested apoptosis as a possible mechanism reducing HSCs since it was previously suggested to affect *Gata2*$^{+/-}$ HSPCs (Rodrigues et al., 2005). We took advantage of the loss of phospholipid asymmetry of the plasma membrane in apoptotic cells and used the Annexin V affinity assay to detect differences in *Gata2*$^{+/-}$ apoptosis. However, In our hands, we could not detect a significant difference in LSK (Supp. Fig. 3d) and a LSK SLAM (Fig. 3c). The loss of quiescence can impair genomic integrity, inducing a DNA damage response (Lindahl & Barnes, 2000; Rossi et al., 2007). Considering that a gene set indicative of DNA repair is upregulated in *Gata2*$^{+/-}$ LSK SLAM cells (Fig. 2d), we decided to assess its involvement in the *Gata2*$^{+/-}$ phenotype. Ser139-phosphorylated H2AX histone (or γH2AX) accumulates in response to DNA damage and marks double strand breaks, which are hallmarks of proliferative stress (Mah et al., 2010; Rogakou et al., 1999). In a flow cytometry “DNA damage” assay, the γH2AX mean fluorescent intensity (MFI) tended to be higher in *Gata2*$^{+/-}$ HSPCs relative to controls, although no significant difference was detected (Fig. 3d). The lack of significant defects related to apoptosis or DNA damage, led us to speculate that the reduced HSPCs in *Gata2*$^{+/-}$ are an integral component of the haploinsufficient phenotype. Consequently, *Gata2*$^{+/-}$ HSPCs require higher proliferation in order to maintain adequate numbers of differentiated hematopoietic cells. Consequently, we next sought to investigate the role of Gata2 in the differentiation of phenotypic HSCs (LSK SLAM) and multipotent progenitors (MPPs or LSK CD48^{-}CD150^{-}) using an established methylcellulose assay. We analyzed differences via a Colony-Forming Unit (CFU) assay, using a semisolid medium implemented with cytokines, driving differentiation into myeloid and erythroid linages (MethoCult GF M3434). As input, we used an equal amount of LSK SLAM or MPPs and tested them in serial replating. No significant difference in the number or type of colonies was detected in the first plating for LSK SLAM (Fig. 3e) or MPPs (Supp. Fig. 3f) and following re-platings (Suppl. Fig. 3g), indicating an unaffected myeloid and erythroid differentiation potential in phenotypic HSCs and MPPs.

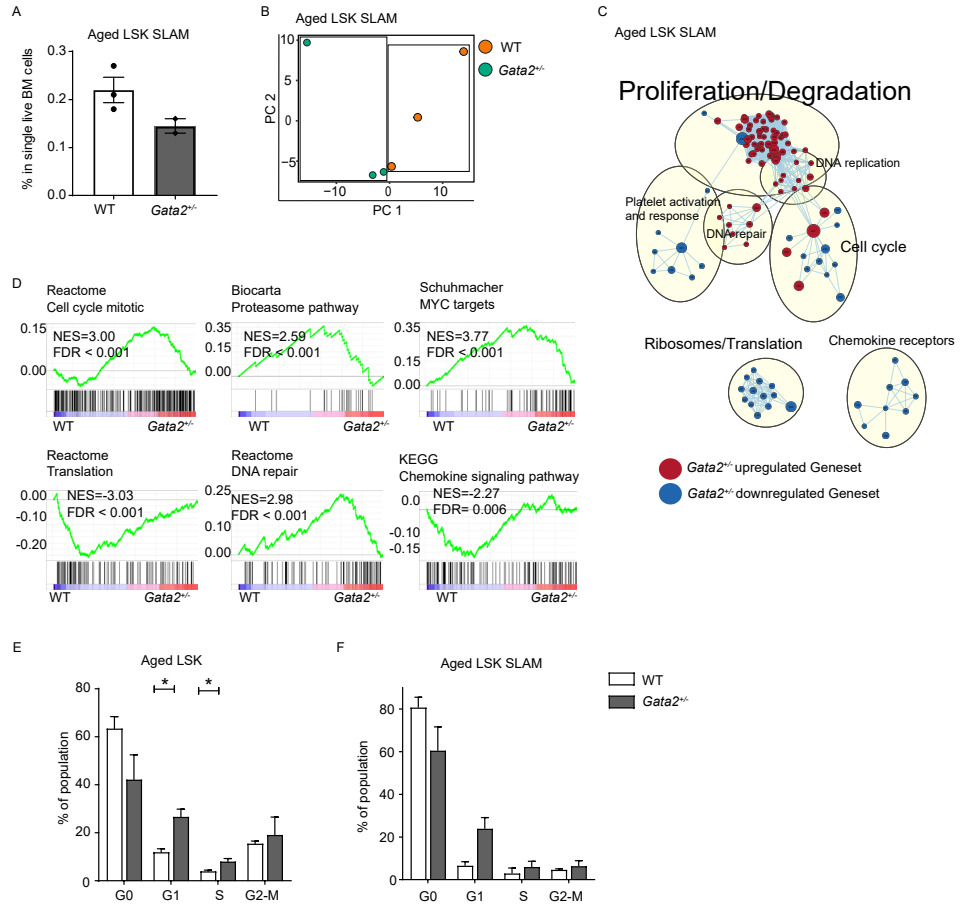


Figure 4. Aged *Gata2*$^{+/-}$ phenotypic HSCs maintain a reduced quiescence
A) Percentage of LSK SLAM cells in bone marrow (BM) in WT (n=3) and *Gata2*$^{+/-}$ (n=2) during aging (14-15 months) (mean± SEM). B) PCA of adult show separation by genotype on principal component (PC) 1. C) Cytoscape Enrichment map representing significant networks (False discovery rate cutoff =0,25, P value cutoff=0,01, Edge cutoff=0,4, Node cutoff Q value =0,05), based on curated gene sets. Only clusters with more than 8 nodes were presented. D) Representative significantly different curated gene sets. E) Column graph representing the difference found in the cell cycle phases of LSK and (F) LSK SLAM. mean ±SEM, *p<0.05. NES: Normalized enrichment score, FDR: False discovery rate.

Aged Gata2$^{+/-}$ maintain reduced phenotypic HSCs and loss of quiescence without developing overt hematological anomalies

In mice, the consequences of GATA2 haploinsufficiency on HSCs are less noticeable compared to humans due to the shorter longevity of the hematopoietic system (Collin et al., 2015). Our results pointed out higher proliferation in *Gata2*$^{+/-}$ HSCs. Consequently, we hypothesized necessary longer periods of time to exacerbate the mutational burden and observe further anomalies in *Gata2*$^{+/-}$. Aging mice up to 15 months did not result in any blood abnormality (Suppl. Fig. 4a). Instead, the BM from aged mice maintained fewer

phenotypic HSCs (Fig. 4a), consistent with a persistently decreased HSC compartment throughout *Gata2*$^{+/-}$ life. Transcriptionally, *Gata2*$^{+/-}$ aged LSK SLAM are distinct from WT based on PC 1 (Fig. 4b). Overall, aged *Gata2*$^{+/-}$ gene sets related to cell cycle, DNA replication and proliferation remain elevated compared to WT (Fig. 4c). This is evident with "cell cycle mitotic", "synthesis of DNA", "mitotic G1_G1/S phases" and "MYC targets" gene sets upregulation (Fig. 4d and Suppl. Fig. 4b). Although, rare cell cycle related gene sets, like "Chang cycling genes", appeared unperturbed or downregulated in *Gata2*$^{+/-}$ (Suppl. Fig 4b). Similarly to adult *Gata2*$^{+/-}$, "proteasome pathway", and "DNA repair" were upregulated in aged *Gata2*$^{+/-}$ (Fig. 4d). For example "proteasome" and "metabolism of amino acids and derivates" were upregulated, while "Translation" and "peptide chain elongation" were downregulated (Suppl. Fig 4b). Indicating an overall alteration of protein synthesis and degradation. Likewise, the upregulation of "base excision repair" and "regulation of apoptosis" (Suppl. Fig 4b) suggests higher genome instability. To confirm the loss of HSCs quiescence in aged *Gata2*$^{+/-}$, we performed a cell cycle analysis. It confirmed that aged *Gata2*$^{+/-}$ mice possess significantly more LSK cells in S phase and G1 phase (Fig. 4e), with *Gata2*$^{+/-}$ LSK SLAM maintain a similar trend (Fig 4f). To summarize, *Gata2*$^{+/-}$ aged HSCs are less quiescent than WT and transcriptionally acquire characteristics of genome instability and protein degradation.

Heterozygous loss of GATA2 recapitulates human haploinsufficiency phenotype after transplantation of aged HSCs

Inferring that HSCs steady state is suboptimal to reproduce and promote the progression of the human Gata2 deficiency, we performed sequential BM transplantation. 3x10^6 nucleated BM cells of 14.5 to 15 months old mice of each genotype were transplanted into lethally irradiated WT mice (Fig. 5a). Chimerism, determined by CD45.2 expression (compared to CD45.1 expression from recipient mice cells) was over 90% from 2 months after transplantation (Fig. 5b). Concomitant to the high level of chimerism of donor BM cells, a specific loss of white blood cells (WBC) in *Gata2*$^{+/-}$ BM recipients' blood was observed without any other apparent cellular anomaly. No difference in red blood cells (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular volume (MCV) or platelets (Fig. 5c) was measured. To confirm that the WBC are reduced, blood smears were performed. Blood smears from *Gata2*$^{+/-}$ aged transplanted mice showed clear reductions in white blood cellularity (Fig. 5d), confirming the leukopenia. This result shows that, after BM transplantation, *Gata2*$^{+/-}$ HSCs fitness is impaired, causing bone marrow failure.

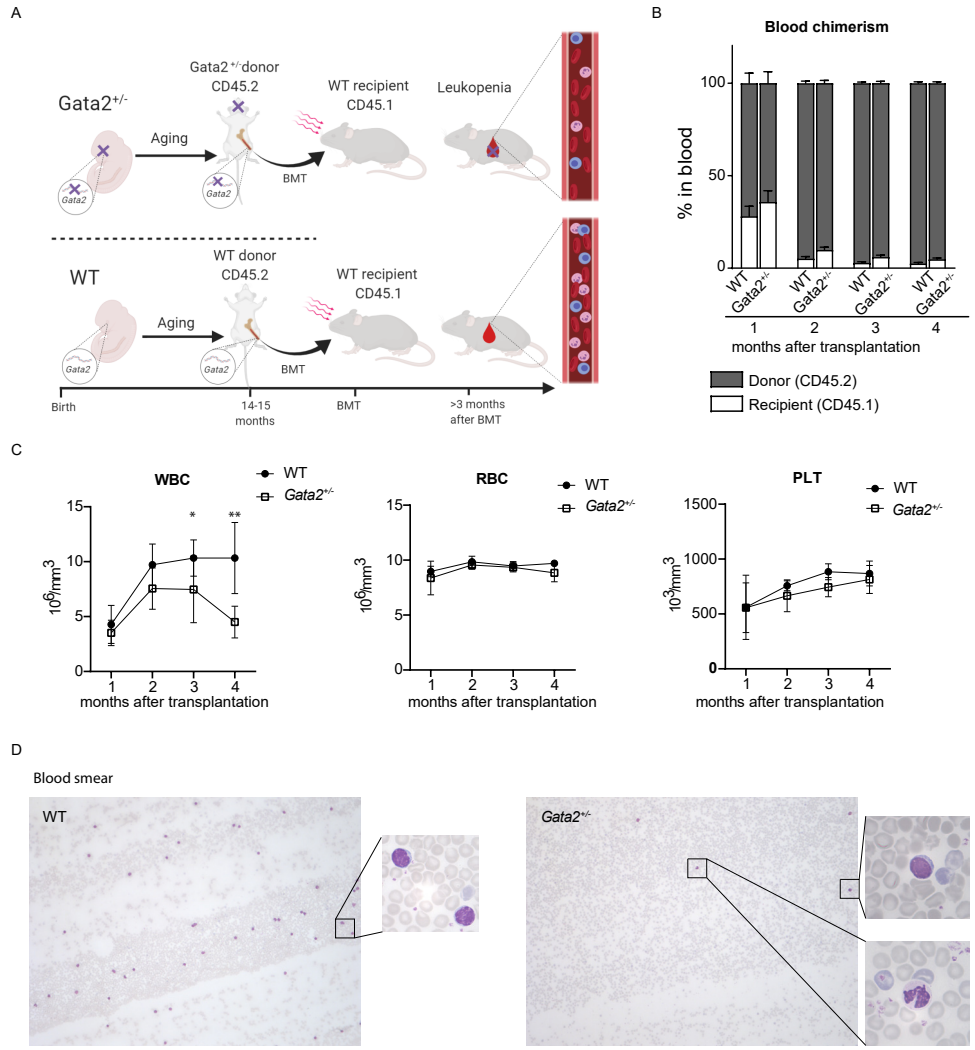


Figure 5. *Gata2*$^{+/-}$ bone marrow cells cause leukopenia in transplantation recipients

A) Graphic representation of WT and *Gata2*$^{+/-}$ mouse aging and following CD45.2 bone marrow transplantation (BMT) in CD45.1 lethally irradiated recipients. Only mice transplanted with *Gata2*$^{+/-}$ cells develop leukopenia. B) Blood chimerism in WT recipients of *Gata2*$^{+/-}$ and WT BM transplantation 1 to 4 months after transplantation. C) Blood values of BM transplantation recipients. Mice were transplanted using either WT or *Gata2*$^{+/-}$ aged donors bone marrow. WBC:white blood cells, RBC: red blood cells, HGB: hemoglobin, PLT:platelets. ** = $p < 0.01$, * = $p < 0.05$. D) Representative blood smears of mice transplanted with WT or *Gata2*$^{+/-}$ BM cells 3 months after transplantation.

DISCUSSION

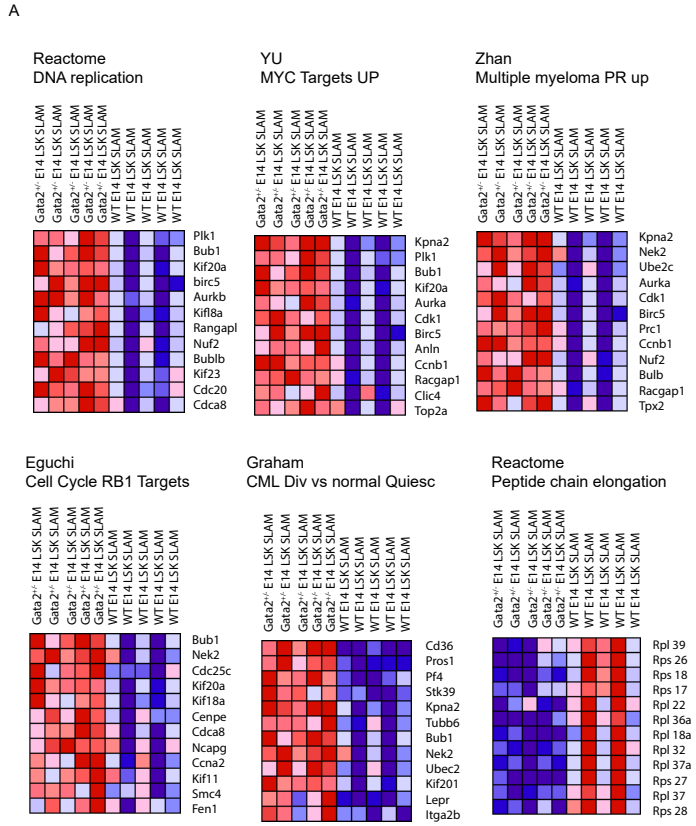
Since the generation of the first *Gata2* deficient mouse model in 1994 (Tsai et al., 1994), GATA2 and related pathologies have been extensively studied. Despite characterizing major *Gata2*$^{+/-}$ HSPCs traits (Ling et al., 2004; Rodrigues et al., 2005), the *Gata2*$^{+/-}$ mouse model thus far did not recapitulate the hematologic deficiencies of GATA2 deficiency syndrome. Here, we examined the effect of aging on Gata2 haploinsufficient HSCs and found that the fitness of aged *Gata2*$^{+/-}$ HSCs is impaired, resulting in a failure to produce sufficient numbers of WBC. We assessed the underlying mechanism affecting Gata2 haploinsufficient HSPCs via RNA sequencing (RNA seq) and revealed that, starting from embryonic development, *Gata2*$^{+/-}$ phenotypic HSPCs are quantitatively scarce and hyperproliferative compared to WT.

The onset of leukopenia after transplantation of aged *Gata2*$^{+/-}$ HSCs reflects a major aspect of human GATA2 deficiency (Collin et al., 2015; Rachel E. Dickinson et al., 2014; Nováková et al., 2016). Mononuclear cytopenia frequently precedes MDS/AML in GATA2 deficiency (Nováková et al., 2016; Wlodarski et al., 2017), and the decline in absolute B, CD4 T, monocyte and NK cells is highly correlated with disease presence in GATA2 deficient patients (Donadieu et al., 2018; Spinner et al., 2014). Our finding represents the first observable blood abnormality derived from GATA2 haploinsufficient mice. The fact that we observe for the first time a blood condition derived by *Gata2*$^{+/-}$ cells is a key component for the reproducibility of the GATA2 deficiency in a mouse model.

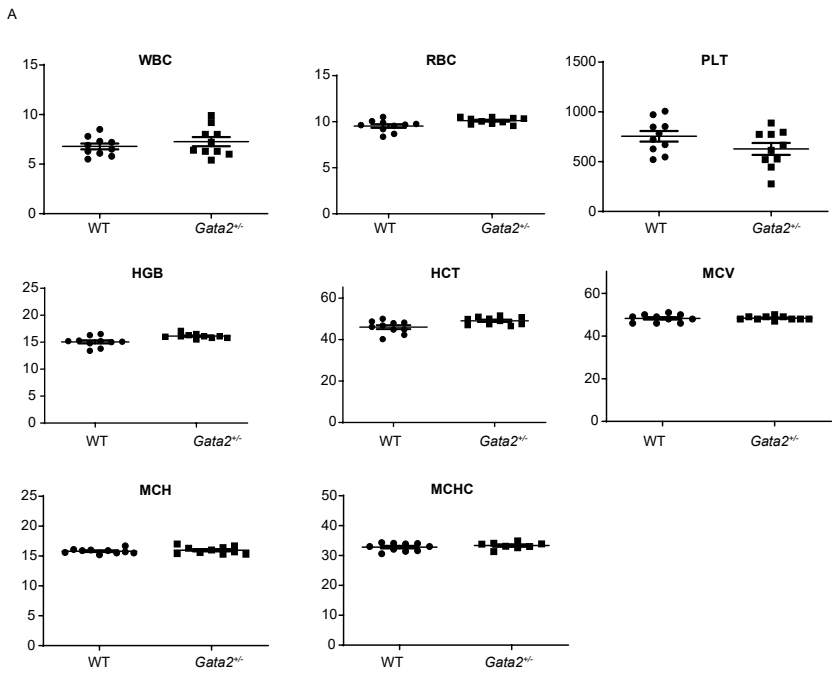
Our results indicate Gata2 as a critical factor for the maintenance of HSCs fitness upon aging. While aged *Gata2*$^{+/-}$ mice did not show an apparent hematological phenotype, the transplantation forced proliferation of HSCs, resulting in an additional source of stress and bone marrow failure. This results revealed the necessity of supplementary stress factors, such as bone marrow transplantation, to recapitulate human GATA2 deficiency symptoms in mouse. Therefore, we do not exclude that a different combination of stress factors could trigger a similar outcome.

Although BM transplantation is evidently an artificial source of stress, collectively our studies suggest that the accumulation of proliferative stress supports the leukopenia in *Gata2*$^{+/-}$ recipients. If this is true, infections or diseases causing a call for stem cell activity might promote the development of a phenotype in GATA2 patients, and possibly explain the variability in penetrance and expressivity. Considering the lack of insults the mouse immune system suffers in the aseptic conditions of the animal facility, it is foreseeable that HSCs are not challenged in this conditions and do not cause development of leukopenia.

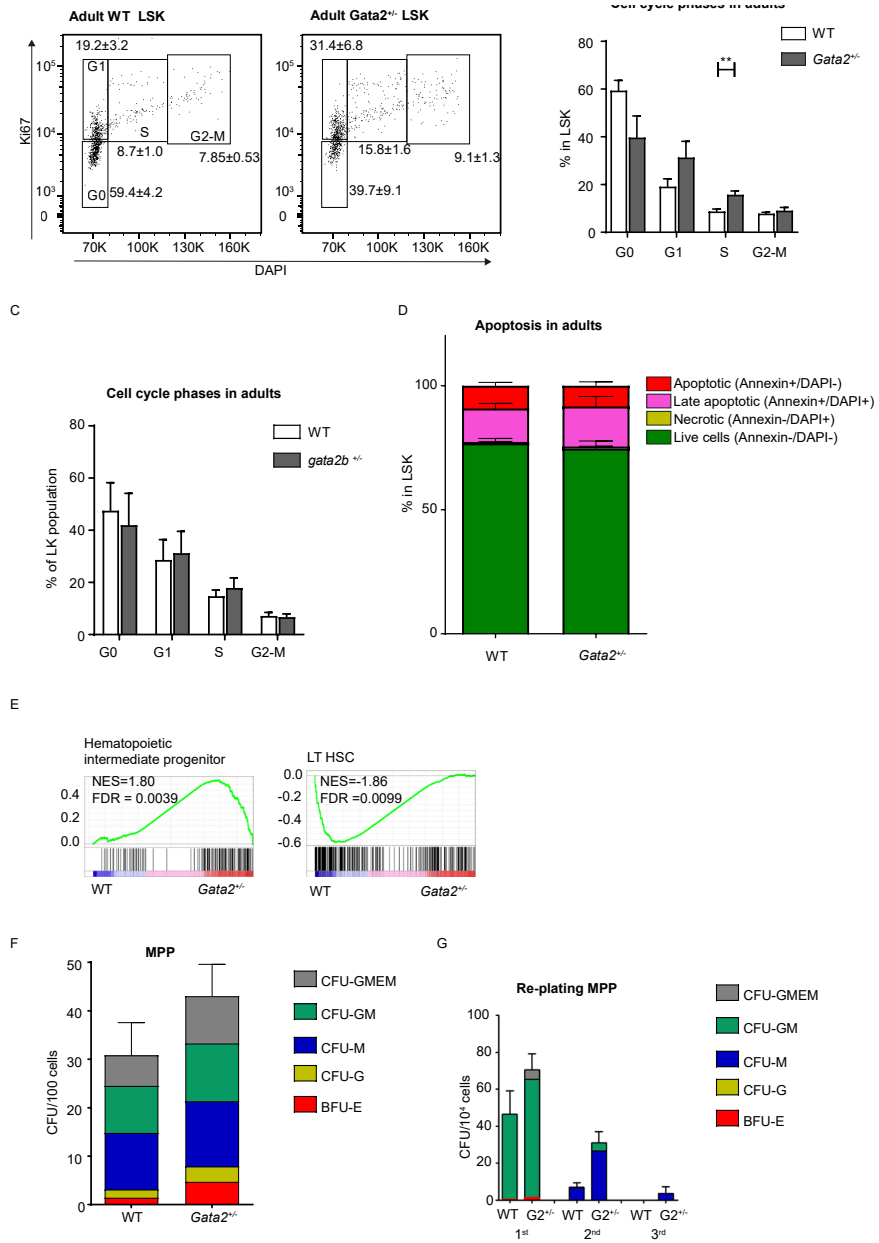
In summary, we recapitulated a manifestation of the human GATA2 disorder in mice using bone marrow transplantation of aged *Gata2*$^{+/-}$ HSCs. Moreover, we identified the loss of quiescence as an attribute of the reduced *Gata2*$^{+/-}$ HSPCs compartment. Taken together, we provide a consistent methodology for reproducing a GATA2 deficiency associated disorder in *Gata2*$^{+/-}$ mice and identify HSPCs aberrancies as a possible cause of the disease outcome.



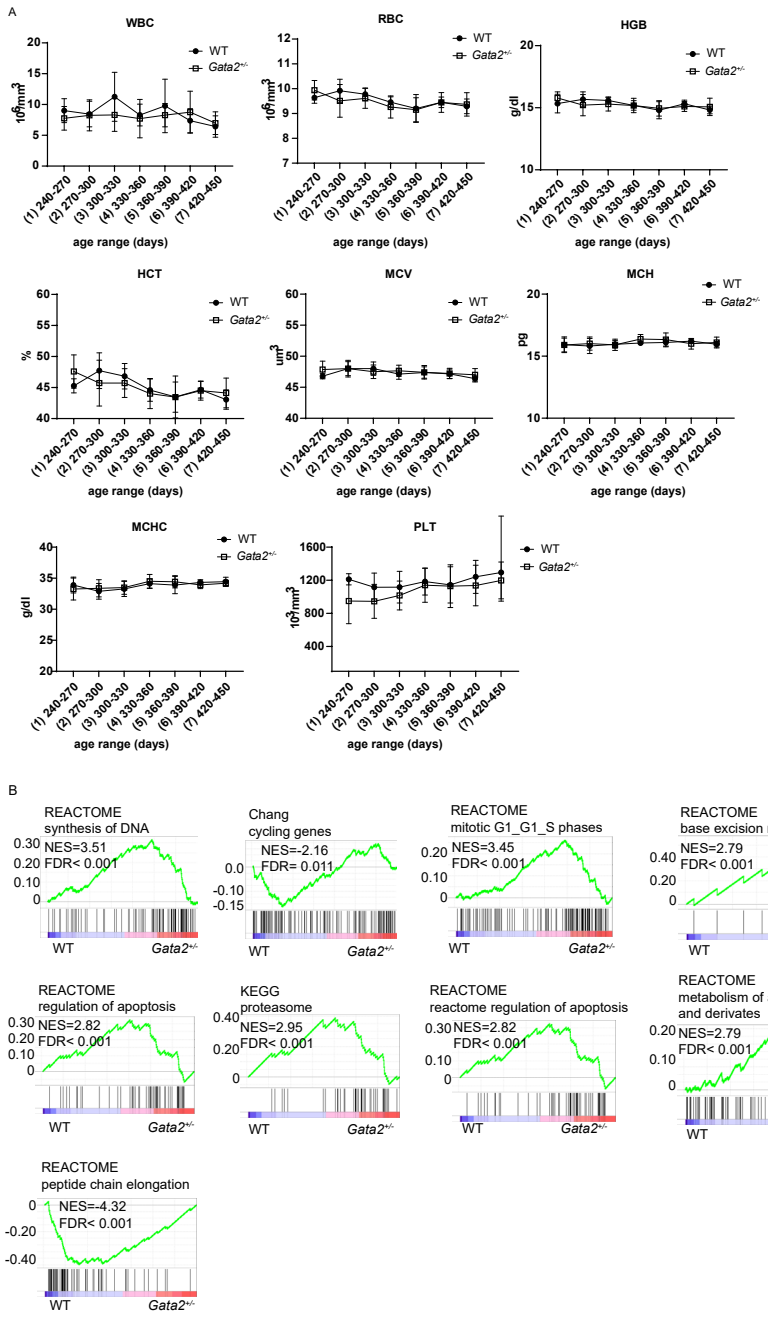
Supplementary Figure 1. A) Top 12 *Gata2*$^{+/-}$ upregulated genes in significantly upregulated gene sets (or Top 12 downregulated genes in significantly downregulated gene sets). Red and blue square intensity represents respectively gene upregulation or downregulation.



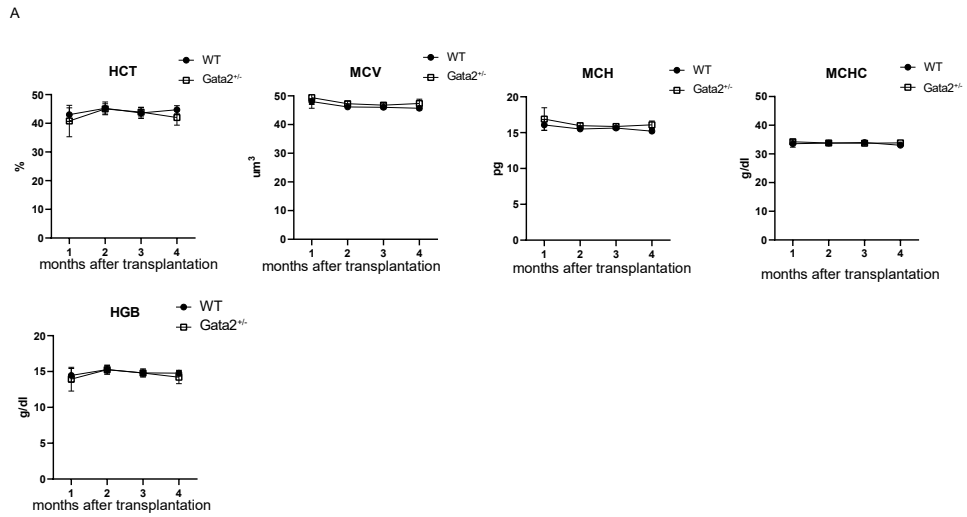
Supplementary Figure 2. A) Blood values of adult WT and *Gata2*$^{+/-}$ mice. HCT: Hematocrit, MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration, HGB: Hemoglobin, WBC: white blood cells, RBC: red blood cells, HGB: hemoglobin, PLT:platelets. B) Top 12 *Gata2*$^{+/-}$ upregulated genes in significantly upregulated gene sets. Red and blue square intensity represents respectively gene upregulation or downregulation.



Supplementary figure 3. A) Representative cell cycle flow cytometry plots of equal numbers of LSK in WT and *Gata2*$^{+/-}$ with gating strategy used. mean ±SEM. B) Column graph representing the difference found in the LSK and C) LK cell cycle phases. mean ±SEM. D) Apoptosis analysis via flow cytometry of fixed adult BM in WT and *Gata2*$^{+/-}$ using AnnexinV and Dapi. mean±SEM. E) Representative significantly different curated gene sets. NES: Normalized enrichment score, FDR: False discovery rate. F) Stack bar chart representing the CFU types obtained 10 days after plating 100 MPPs in MethoCult. G) CFU types obtained after re-plating 10^4 cells isolate from MPPs colonies. The same condition were used in 3 re-platings. ** = $p < 0.01$.



Supplementary figure 4. A) Blood values of aged WT and *Gata2*$^{+/-}$ mice. HCT: Hematocrit, MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration, HGB: Hemoglobin, WBC: white blood cells, RBC: red blood cells, HGB: hemoglobin, PLT:platelets. B) Representative significantly different curated gene sets of aged LSK SLAM. NES: Normalized enrichment score, FDR: False discovery rate.



Supplementary figure 5.
A) Blood values of BM transplantation recipients. Mice were transplanted using aged donors either WT or *Gata2*$^{+/-}$ BM. HCT: Hematocrit, MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration, HGB: Hemoglobin.

REFERENCES

Bhagwat, A. S., & Vakoc, C. R. (2015). Targeting Transcription Factors in Cancer. In *Trends in Cancer* (Vol. 1, Issue 1, pp. 53–65). Cell Press. https://doi.org/10.1016/j.trecan.2015.07.001

Collin, M., Dickinson, R., & Bigley, V. (2015). Haematopoietic and immune defects associated with *GATA2* mutation. *British Journal of Haematology*, *169*(2), 173–187. https://doi.org/10.1111/bjh.13317

Cortés-Lavaud, X., Landecho, M. F., Maicas, M., Urquiza, L., Merino, J., Moreno-Miralles, I., & Odero, M. D. (2015). GATA2 Germline Mutations Impair GATA2 Transcription, Causing Haploinsufficiency: Functional Analysis of the p.Arg396Gln Mutation . *The Journal of Immunology*, *194*(5), 2190–2198. https://doi.org/10.4049/jimmunol.1401868

de Pater, E., Kaimakis, P., Vink, C. S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P. S., Camper, S. A., Speck, N., & Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. *Journal of Experimental Medicine*, *210*(13), 2843–2850. https://doi.org/10.1084/jem.20130751

Dickinson, Rachel E., Milne, P., Jardine, L., Zandi, S., Swierczek, S. I., McGovern, N., Cookson, S., Ferozepurwalla, Z., Langridge, A., Pagan, S., Gennery, A., Heiskanen-Kosma, T., Ḧaṁaläinen, S., Sepp̈anen, M., Helbert, M., Tholouli, E., Gambineri, E., Reykdal, S., Gottfreósson, M., … Collin, M. (2014). The evolution of cellular deficiency in GATA2 mutation. *Blood*, *123*(6), 863–874. https://doi.org/10.1182/blood-2013-07-517151

Dickinson, Rachel Emma, Griffin, H., Bigley, V., Reynard, L. N., Hussain, R., Haniffa, M., Lakey, J. H., Rahman, T., Wang, X. N., McGovern, N., Pagan, S., Cookson, S., McDonald, D., Chua, I., Wallis, J., Cant, A., Wright, M., Keavney, B., Chinnery, P. F., … Collin, M. (2011). Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood*, *118*(10), 2656–2658. https://doi.org/10.1182/blood-2011-06-360313

Donadieu, J., Lamant, M., Fieschi, C., De Fontbrune, F. S., Caye, A., Ouachee, M., Beaupain, B., Bustamante, J., Poirel, H. A., Isidor, B., Neste, E. Van Den, Neel, A., Nimubona, S., Toutain, F., Barlogis, V., Schleinitz, N., Leblanc, T., Rohrlich, P., Suarez, F., … Pasquet, M. (2018). Natural history of GATA2 deficiency in a survey of 79 French and Belgian patients. *Haematologica*, *103*(8), 1278–1287. https://doi.org/10.3324/haematol.2017.181909

Dzierzak, E., & Speck, N. A. (2008). Of lineage and legacy: The development of mammalian hematopoietic stem cells. In *Nature Immunology* (Vol. 9, Issue 2, pp. 129–136). NIH Public Access. https://doi.org/10.1038/ni1560

Ema, H., & Nakauchi, H. (2000). Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood*, *95*(7), 2284–2288. https://doi.org/10.1182/blood.v95.7.2284.007k14_2284_2288

Guo, G., Luc, S., Marco, E., Lin, T. W., Peng, C., Kerenyi, M. A., Beyaz, S., Kim, W., Xu, J., Das, P. P., Neff, T., Zou, K., Yuan, G. C., & Orkin, S. H. (2013). Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. *Cell Stem Cell*, *13*(4), 492–505. https://doi.org/10.1016/j.stem.2013.07.017

Hahn, C. N., Chong, C. E., Carmichael, C. L., Wilkins, E. J., Brautigan, P. J., Li, X. C., Babic, M., Lin, M., Carmagnac, A., Lee, Y. K., Kok, C. H., Gagliardi, L., Friend, K. L., Ekert, P. G., Butcher, C. M., Brown, A. L., Lewis, I. D., To, L. B., Timms, A. E., … Scott, H. S. (2011). Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nature Genetics*, *43*(10), 1012–1019. https://doi.org/10.1038/ng.913

Hsu, A. P., Sampaio, E. P., Khan, J., Calvo, K. R., Lemieux, J. E., Patel, S. Y., Frucht, D. M., Vinh, D. C., Auth, R. D., Freeman, A. F., Olivier, K. N., Uzel, G., Zerbe, C. S., Spalding, C., Pittaluga, S., Raffeld, M., Kuhns, D. B., Ding, L., Paulson, M. L., … Holland, S. M. (2011). Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*, *118*(10), 2653–2655. https://doi.org/10.1182/blood-2011-05-356352

Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R. A., Turner, M. L., & Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. *Journal of Experimental Medicine*, *208*(12), 2417–2427. https://doi.org/10.1084/jem.20111688

Kim, I., He, S., Yilmaz, Ö. H., Kiel, M. J., & Morrison, S. J. (2006). Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. *Blood*, *108*(2), 737–744. https://doi.org/10.1182/blood-2005-10-4135

Kim, K. H., & Sederstrom, J. M. (2015). Assaying cell cycle status using flow cytometry. *Current Protocols in Molecular Biology*, *2015*, 28.6.1-28.6.11. https://doi.org/10.1002/0471142727.mb2806s111

Kumaravelu, P., Hook, L., Morrison, A. M., Ure, J., Zhao, S., Zuyev, S., Ansell, J., & Medvinsky, A. (2002). Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): Role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver (Development) vol. 129 (4891-. *Development*, *130*(2), 425.

Lindahl, T., & Barnes, D. E. (2000). Repair of endogenous DNA damage. *Cold Spring Harbor Symposia on Quantitative Biology*, *65*, 127–133. https://doi.org/10.1101/sqb.2000.65.127

Ling, K. W., Ottersbach, K., Van Hamburg, J. P., Oziemlak, A., Tsai, F. Y., Orkin, S. H., Ploemacher, R., Hendriks, R. W., & Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *Journal of Experimental Medicine*, *200*(7), 871–882. https://doi.org/10.1084/jem.20031556

Mah, L. J., El-Osta, A., & Karagiannis, T. C. (2010). γh2AX: A sensitive molecular marker of DNA damage and repair. In *Leukemia* (Vol. 24, Issue 4, pp. 679–686). Nature Publishing Group. https://doi.org/10.1038/leu.2010.6

Mazumder, T. H., Uddin, A., & Chakraborty, S. (2016). Transcription factor gene GATA2: Association of leukemia and nonsynonymous to the synonymous substitution rate across five mammals. *Genomics*, *107*(4), 155–161. https://doi.org/10.1016/j.ygeno.2016.02.001

Medvinsky, A., & Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*, *86*(6), 897–906. https://doi.org/10.1016/S0092-8674(00)80165-8

Nakamura-Ishizu, A., Takizawa, H., & Suda, T. (2014). The analysis, roles and regulation of quiescence in hematopoietic stem cells. In *Development (Cambridge)* (Vol. 141, Issue 24, pp. 4656–4666). Company of Biologists Ltd. https://doi.org/10.1242/dev.106575

Nováková, M., Žaliová, M., Suková, M., Wlodarski, M., Janda, A., Froňková, E., Campr, V., Lejhancová, K., Zapletal, O., Pospíšilová, D., Černá, Z., Kuhn, T., Švec, P., Pelková, V., Zemanová, Z., Kerndrup, G., van den Heuvel-Eibrink, M., van der Velden, V., Niemeyer, C., … Mejstříková, E. (2016). Loss of B cells and their precursors is the most constant feature of GATA-2 deficiency in childhood myelodysplastic syndrome. *Haematologica*, *101*(6), 707–716. https://doi.org/10.3324/haematol.2015.137711

Ogawa, M. (1993). Differentiation and proliferation of hematopoietic stem cells. In *Blood* (Vol. 81, Issue 11, pp. 2844–2853). https://doi.org/10.1182/blood.v81.11.2844.2844

Orford, K. W., & Scadden, D. T. (2008). Deconstructing stem cell self-renewal: Genetic insights into cell-cycle regulation. In *Nature Reviews Genetics* (Vol. 9, Issue 2, pp. 115–128). https://doi.org/10.1038/nrg2269

Ostergaard, P., Simpson, M. A., Connell, F. C., Steward, C. G., Brice, G., Woollard, W. J., Dafou, D., Kilo, T., Smithson, S., Lunt, P., Murday, V. A., Hodgson, S., Keenan, R., Pilz, D. T., Martinez-Corral, I., Makinen, T., Mortimer, P. S., Jeffery, S., Trembath, R. C., & Mansour, S. (2011). Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nature Genetics*, *43*(10), 929–931. https://doi.org/10.1038/ng.923

Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., & Demchak, B. (2019). Cytoscape Automation: Empowering workflow-based network analysis. *Genome Biology*, *20*(1). https://doi.org/10.1186/s13059-019-1758-4

Rodrigues, N. P., Janzen, V., Forkert, R., Dombkowski, D. M., Boyd, A. S., Orkin, S. H., Enver, T., Vyas, P., & Scadden, D. T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*, *106*(2), 477–484. https://doi.org/10.1182/blood-2004-08-2989

Rogakou, E. P., Boon, C., Redon, C., & Bonner, W. M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. *Journal of Cell Biology*, *146*(5), 905–915. https://doi.org/10.1083/jcb.146.5.905

Rossi, D. J., Bryder, D., Seita, J., Nussenzweig, A., Hoeijmakers, J., & Weissman, I. L. (2007). Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature*, *447*(7145), 725–729. https://doi.org/10.1038/nature05862

Seita, J., & Weissman, I. L. (2010). Hematopoietic stem cell: Self-renewal versus differentiation. In *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* (Vol. 2, Issue 6, pp. 640–653). NIH Public Access. https://doi.org/10.1002/wsbm.86

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Spinner, M. A., Sanchez, L. A., Hsu, A. P., Shaw, P. A., Zerbe, C. S., Calvo, K. R., Arthur, D. C., Gu, W., Gould, C. M., Brewer, C. C., Cowen, E. W., Freeman, A. F., Olivier, K. N., Uzel, G., Zelazny, A. M., Daub, J. R., Spalding, C. D., Claypool, R. J., Giri, N. K., … Holland, S. M. (2014). GATA2 deficiency: A protean disorder of hematopoiesis, lymphatics, and immunity. *Blood*, *123*(6), 809–821. https://doi.org/10.1182/blood-2013-07-515528

Szade, K., Bukowska-Strakova, K., Zukowska, M., Jozkowicz, A., & Dulak, J. (2016). Analysis of cell cycle status of murine hematopoietic stem cells. In *Methods in Molecular Biology* (Vol. 1516, pp. 91–99). Humana Press Inc. https://doi.org/10.1007/7651_2016_361

Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., & Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, *371*(6494), 221–226. https://doi.org/10.1038/371221a0

Whitfield, M. L., George, L. K., Grant, G. D., & Perou, C. M. (2006). Common markers of proliferation. In *Nature Reviews Cancer* (Vol. 6, Issue 2, pp. 99–106). Nat Rev Cancer. https://doi.org/10.1038/nrc1802

Wlodarski, M. W., Collin, M., & Horwitz, M. S. (2017). GATA2 deficiency and related myeloid neoplasms. *Seminars in Hematology*, *54*(2), 81–86. https://doi.org/10.1053/j.seminhematol.2017.05.002

Wlodarski, M. W., Hirabayashi, S., Pastor, V., Starý, J., Hasle, H., Masetti, R., Dworzak, M., Schmugge, M., Van Den Heuvel-Eibrink, M., Ussowicz, M., De Moerloose, B., Catala, A., Smith, O. P., Sedlacek, P., Lankester, A. C., Zecca, M., Bordon, V., Matthes-Martin, S., Abrahamsson, J., … Niemeyer, C. M. (2016). Prevalence, clinical characteristics, and prognosis of GATA2-related myelodysplastic syndromes in children and adolescents. *Blood*, *127*(11), 1387–1397. https://doi.org/10.1182/blood-2015-09-669937

5

General discussion

SUMMARY

Hematopoietic stem cells (HSCs) provide the millions of differentiated blood cells we need in our everyday life. In blood we distinguish the lymphoid and myeloid lineage. The lymphoid lineage generates adaptive immune cells like T cells and B cells. The myeloid lineage generates erythrocytes for oxygen transport, megakaryocytes for coagulation and innate immune cells like neutrophils and monocytes. These lineages have to be produced at the right proportion to maintain these processes. Transcriptional regulation of HSCs is an essential mechanism to maintain a balanced production of blood cells, and the transcription factor GATA2 plays a key function in this process. GATA2 directs embryonic and adult hematopoiesis in vertebrates by occupying GATA-DNA motifs in numerous genes. Homozygous GATA2 mutations in mice are embryonically lethal, making it difficult to assess its function in the different aspects of hematopoiesis. Heterozygous germline *GATA2* mutations in humans are the cause of a series of disorders called GATA2 deficiency syndromes. They lead to a spectrum of phenotypes including lymphedema, cytopenias and, in more than 80% of patients, MDS or leukemia. To date, there is no clear correlation between the patient phenotype and the type of GATA2 mutation. The exact role of GATA2 in HSCs generation, expansion, self-renewal and differentiation is still unknown.

In this thesis, we focus on the hematopoietic consequences of GATA2 deficiency in zebrafish (**chapter 2-3**) and mouse (**chapter 4**).

Zebrafish have two orthologues of GATA2; i.e., Gata2a and Gata2b. Previous studies have shown that *gata2b* is prominently expressed in HSPCs, whereas *gata2a* is mainly expressed in the vasculature. Being the predominant GATA2 orthologue required for the maintenance of hematopoietic stem cells, we targeted *gata2b* to generate a knockout zebrafish line adopted in **chapters 2-3**. Interestingly, the homozygous Gata2b mutant is viable, allowing us to assess the functionality of Gata2b in adult hematopoiesis.

In **chapter 2** we characterize the effects of Gata2b homozygous knockout (*gata2b*$^{-/-}$) in the hematopoietic system. Gata2b is required for the proliferation of hematopoietic stem and progenitor cells (HSPCs) during embryonic HSPCs expansion in the caudal hematopoietic tissue. Adult *gata2b*$^{-/-}$ HSPCs are more proliferative and their kidney marrow shows a lymphoid bias and a block in neutrophil differentiation. Single cell transcriptome analysis showed that the HSPCs were the origin of the increased lymphoid lineage output in *gata2b*$^{-/-}$ kidney marrow cells, due to a failure to initiate myeloid gene expression to sufficient levels and a subsequent co-expression of both myeloid and lymphoid genes in *gata2b*$^{-/-}$ HSPCs.

To evaluate the transcriptional and phenotypic consequences of half Gata2b dose in zebrafish hematopoiesis, in **chapter 3** we assessed *gata2b* heterozygous zebrafish (*gata2b*$^{+/-}$). We hypothesized that, with half Gata2b dose, zebrafish would still present reduced myeloid differentiation, although to a lesser extent than when knocking out *gata2b* homozygously. Instead, we observed an unprecedented erythroid dysplasia in the kidney marrow, but a conservation of the major differentiation lineages. The HSPC compartment also showed transcriptional changes fitting with the transcriptional changes as found in MDS.

The clear dichotomy in phenotypes between the heterozygous Gata2b mutant and the homozygous Gata2b mutant show that, depending on its dose, Gata2 has distinct functions in the maintenance of HSPCs and the subsequent differentiation into the different hematopoietic lineages. This could possibly explain the variable phenotypes found in GATA2 patients.

Mice, like humans, possess one copy of the *Gata2* gene. *Gata2* knock out (*Gata2*$^{-/-}$) mouse embryos succumb before the generation of HSCs, while *Gata2* heterozygous knock-out mice (*Gata2*$^{+/-}$) do not develop noticeable hematological phenotypes. We hypothesized that a transcriptional analysis of HSPCs would reveal unnoticeable peculiarities caused by *Gata2* heterozygosity. In **chapter 4** we found that *Gata2*$^{+/-}$ HSPCs are hyperproliferative and to determine the long term consequences of this condition, we let *Gata2*$^{+/-}$ mice age up to 14-15 months. *Gata2*$^{+/-}$ aged HSPCs remain hyperproliferative and caused a failure to produce sufficient numbers of white blood cells after bone marrow transplantation.

The onset of leukopenia after transplantation of aged *Gata2*$^{+/-}$ HSCs reflects a major aspect of human GATA2 deficiency and provides a consistent methodology for reproducing a GATA2 associated disorder in mice. Further experiments will be necessary to better characterize the leukopenia identified and to test if other stress factors can contribute to a similar outcome in a reduced time-span. Since aging is necessary for the leukopenia to occur, it is would be interesting addressing its role in the phenotype development such as bone marrow niche-mediated inflammation.

Collectively our findings provide insights into the biology of GATA2 and the hematopoietic consequences following its deficiency. They establish the human disease relevance of the *Gata2*$^{+/-}$ mouse model and spotlight an essential role for Gata2 in HSPCs lineage output/proliferative stress in human GATA2 deficiency. These findings pave the way for further research on the synergism between HSPCs reduction, loss of quiescence and altered lineage output caused by Gata2 deficiency.

5.2 GENERAL DISCUSSION

One gene, multiple roles in hematopoietic cell development and functions

Transcription factors like GATA2 have key roles in directing the expression of lineage specific target-genes in the hematopoietic system. Gata2 has several different functions in the hematopoietic system. Gata2 is required for HSCs generation (Tsai et al., 1994), survival (de Pater et al., 2013), proper proliferation and differentiation (this thesis). In humans, around 90% of germline *GATA2* pathogenic mutations are either missense or truncating, and considered to cause GATA2 haploinsufficiency (Cortés-Lavaud et al., 2015; Leubolt et al., 2020; Wlodarski et al., 2017). Treatment of the diverse clinical presentations remains limited and the only potential cure is allogeneic HSC transplantation (Hirabayashi et al., 2017). Both hereditary and sporadic forms of GATA2 deficiency exist. Their clinical features include immunodeficiency, pulmonary, vascular/lymphatic dysfunctions and a strong propensity to develop MDS and AML (Spinner et al., 2014)(Fig.1).

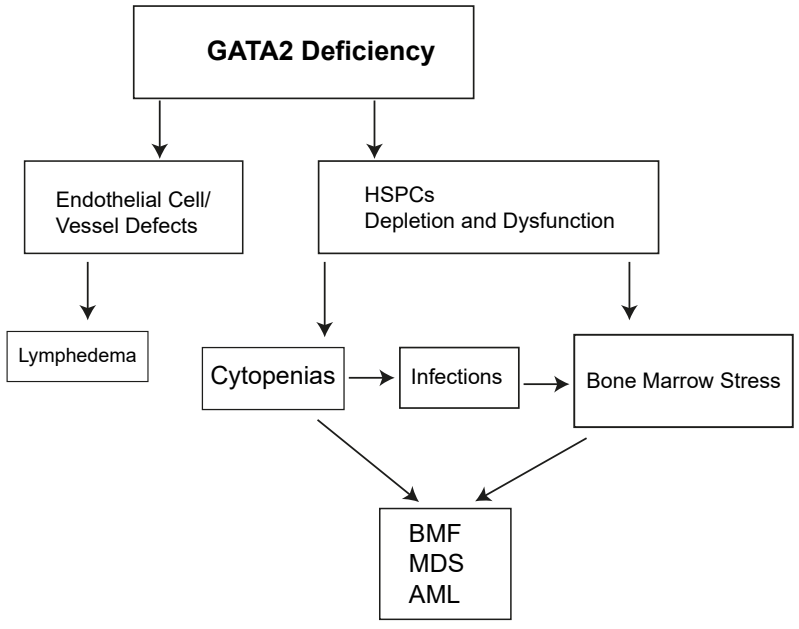


Figure 1. Effects of GATA2 deficiency. Adapted from McReynolds et al (McReynolds et al., 2018). GATA2 deficiency patients have lymphedema secondary to effects on the vasculature, likely due to GATA2 activity at the endothelial cell level. Loss of GATA2 protein leads to hematopoietic stem and progenitor cell (HSPC) loss and dysfunction. This depletion can support cytopenias—B cells, dendritic cells, natural killer cells, and monocytes. The resultant immunodeficiency drives infections. Overall, these factors lead to bone marrow stress, hypocellularity and the ultimate developmentof bone marrow failure (BMF), myelodysplastic syndrome (MDS), and acute myeloid leukemia (AML). The combination of continued HSC depletion, infections, and cytopenias stress the bone marrow creating an environment supportive for the development of MDS, which may lead to AML or CMML.

Since first described, GATA2 haploinsufficiency provides management difficulties due to the challanges in delineating the genotype to phenotype relationship. Given the heterogeneity in the mutational spectrum, age of onset and clinical manifestations, a standardized analysis of the effects of GATA2 mutations remains difficult. For instance, even members of the same family (identical GATA2 mutation) can present with different clinical outcomes, with disease latencies varying from 14 to 74 years of age (Mutsaers et al., 2013). The reasons for this wide array of phenotypes remain unexplained at present. Possibly, chronic pathogen challenges of the immune system (highly frequent in GATA2 patients) could provide a pathophysiologic basis for myeloid neoplasms (Kristinsson et al., 2011) in association with the GATA2 haploinsufficiency. Indeed, pancytopenia and bone marrow suppression are known complications of chronic infectious diseases (Achi et al., 2013), and inflammatory stimulation and dysregulation of interferon signaling contribute to depletion of hematopoietic stem and progenitors (Essers et al., 2009; Lin et al., 2014; Matatall et al., 2016; Sato et al., 2009). On the other hand, mechanisms controlling the rate of allelic expression might have a role in disease progression. Our data suggests that different dosages of Gata2b have different effects on the hematopoietic system. In **chapter 2** we show that the lack of wild type *gata2b* expression in *gata2b* homozygous knockout zebrafish causes a shift in differentiation, while the heterozygous expression of the wild type *gata2b* causes erythrodysplasia (**chapter 3**)**.** The distinct phenotypes demonstrate a gata2b dose-dependent effect on zebrafish hematopoiesis. Indeed, a recent study showed that leukemic blasts of AML patients display GATA2 allele specific expression more often than any other myeloid or cancer-related gene, and GATA2 mutated patients preferably express the mutant allele (Mulet-Lazaro 2021). These results suggest that the ratio of mutated vs wildtype allele expression could cause dose dependent effects.

Interplay of different GATA family members in controlling hematopoiesis

All members of the GATA family of transcription factors (GATA1 to GATA6) possess zinc-fingers domains responsible for binding the consensus DNA sequence (A/T)GATA(A/G). Given the simplicity of this sequence, GATA binding sites can be found scattered through the entire genome. Transcription factors can replace each other on these DNA binding sites but do not bind with the same affinity to all GATA sites. Distinct activities of the GATA members are predominantly determined by their temporal expression at distinct stages of hematopoietic development (Katsumura & Bresnick, 2017). GATA1, GATA2 and GATA3 are the hematopoietic regulators in this family. GATA1 is mostly involved in the development of erythroid (Fujiwara et al., 1996) and megakaryocyte progenitors (Martin et al., 1990), and required for their differentiation (Pevny et al., 1991; Shivdasani et al., 1997). While GATA3 has an essential function in T cell differentiation (Ho et al., 2009), together with Gata2, it can partially rescue the erythroid phenotype in Gata1 deficient mice (Ferreira et al., 2007; Takahashi et al., 2000; Tsai et al., 1998). A strict regulation has been attributed to Gata1 and

Gata2 during erythroid differentiation since these transcription factors regulate the same binding sites in different steps of differentiation, in a process called GATA switch (Grass et al., 2006). These findings suggest that Gata2 is positively regulated by its protein product via binding its promoter (Grass et al., 2003; Vicente et al., 2012), while it gets repressed by GATA1, in partly by disrupting its autoregulation (Bresnick & Johnson, 2019). The erythroid dysplasia observed in *gata2b*$^{+/-}$ mice (**chapter 3**) might be suggestive of an aberrant Gata1 compensation in zebrafish. The finding that Gata1 is not differentially expressed in bulk or scRNA sequencing of zebrafish *gata2b*$^{+/-}$ HSPCs may argue against this hypothesis. On the other hand, it is possible to have Gata1 compensation even if the level of its expression remains unaltered. An increased association rate of Gata1 to gene promoters in response to *gata2b* heterozygosity (Hasegawa et al., 2016) can make transcriptional upregulation not necessary for compensation. Consequently, we do not exclude that altered Gata factors binding to DNA can be causative of the phenotype.

When one gene copy is not enough

For most genes, a single copy is enough for normal development and most heterozygous knockout mice result in no detectable phenotype (Seidman & Seidman, 2002). When a single copy of the wild type allele in diploid organisms is not sufficient to produce the standard phenotype (Morrill & Amon, 2019), haploinsufficiency occurs. About 300 human genes are known to be haploinsufficient (Dang et al., 2008) and, when mutated, can become the source of multiple health issues. Why only some genes cause haploinsufficiency and the mechanism behind their regulation, is not clear. The most recent hypothesis called “dose-stabilizing” assumes that haploinsufficient genes have a narrow gene dosage range where gene products are limiting for fitness when under-expressed, while higher gene products are toxic, possibly because they interfere with cellular functions or because of a detrimental production and degradation of the excess product (Morrill & Amon, 2019).

In humans, GATA2 heterozygosity mainly causes haploinsufficiency (Cortés-Lavaud et al., 2015). Like other haploinsufficient genes, GATA2 balanced expression is necessary for proper homeostasis and consequently more than 90% of GATA2 mutant patients develop a life-threatening hematologic disease by 40 years of age (Donadieu et al., 2018). However, the age of onset and clinical manifestations widely vary and a small percentage of carriers remains asymptomatic. Using reverse genetics in inbred animal models we standardized the effects of Gata2 knockout. Whereas in zebrafish Gata2b heterozygous fish show erythrodysplasia (**chapter 3**), in mice, *Gata2* heterozygosity (*Gata2*$^{+/-}$) causes the development of a hematologic phenotype only under stress conditions. However, some genes can result in haploinsufficiency only under specific conditions. For instance, in yeast there is little overlap between haploinsufficient genes based on growth medium differences (Deutschbauer et al., 2005). Possibly, *Gata2*$^{+/-}$ mice did not meet the conditions to cause a phenotype like human GATA2 heterozygous patients. Based on the observation that

GATA2 patients’ phenotypes are strongly correlated with age, in **chapter 4** we promoted Gata2 haploinsufficiency by ageing *Gata2*$^{+/-}$ mice and after bone marrow transplantation, observing a hematologic phenotype. These results support the notion that a Gata2 knockout background might require further challenges for haploinsufficiency to occur.

Proliferative and differentiation aspects of Gata2 dosage in hematopoiesis

HSCs are multipotent cells which rarely enter cell cycle (Orford & Scadden, 2008). However, a combination of different factors shapes their potency and dormancy. Both of these aspects are affected by Gata2 mutations in this thesis. In accordance to the dose-stabilizing hypothesis of haploinsufficiency, GATA2 overexpression and downregulation can both cause adverse consequences to the hematopoietic system. GATA2 can be found overexpressed in the bone marrow of acute myeloid leukemia (AML) patients and in leukemic blasts (Bullinger et al., 2004; Luesink et al., 2012; Vicente et al., 2012), while GATA2 gain of function mutations can be observed in chronic myeloid leukemia (Zhang et al., 2008). GATA2 overexpression *in vitro* can cause cell cycle arrest and reduced colony forming capacity (Ezoe et al., 2002; Heyworth et al., 1999) and *in vivo* enforced GATA2 expression causes inhibition of cell cycle and differentiation (Persons et al., 1999; Tipping et al., 2009). If increasing Gata2 expression promotes quiescence, predictably a decrease in Gata2 expression would promote proliferation. Whereas a finding reported decreased cell cycle activity in Gata2$^{+/-}$ HSPCs (Rodrigues et al., 2005), in **chapter 4**, we identify that, starting from embryonic development, *Gata2*$^{+/-}$ phenotypic HSPCs are quantitatively scarce but lose quiescence compared to WT. These contrasting results have not been clarified yet and might be caused by multiple factors like the HSPC population taken in analysis, the growth conditions of the animals, or the different assays used. Consistent with the loss of quiescence, in **chapter 2 and 3,** by knocking out Gata2b in zebrafish we observe a consistent transcriptional increase in proliferation. In **chapter 2,** Gata2b deficient cells almost entirely lose a quiescent HSPC subcluster and gain a proliferative subcluster while i**n chapter 3** we confirm a pivotal role for gata2b in cell cycle, observing DNA replication and ribosome biogenesis as the main upregulated transcriptionally characteristics representing *gata2b*$^{+/-}$ HSPCs.

When considering the mechanisms governing Gata2 deficient proliferation, we hypothesized that the reduced HSPC population observed in Gata2$^{+/-}$ mouse embryos (de Pater et al., 2013; Gao et al., 2013, this thesis) and adults (Rodrigues et al., 2005, this thesis) could trigger a loss in quiescence as a compensatory mechanism. Gata2b deficient zebrafish instead, after HSCs amplification, re-acquire WT HSPCs numbers throughout adulthood, excluding that the reduction in HSPCs is a direct signal inducing their hyperproliferation in zebrafish. While the Gata2$^{+/-}$ proliferative changes observed *in vitro* suggest intrinsic signalling as an important factor in this regulation, extrinsic signalling might as well play an important role. Stress factors such as infections, cytopenias, ageing, chemotherapy, and bone marrow transplantation can force them out of dormancy and contribute actively to

hematopoiesis (Essers et al., 2009; Singh et al., 2020; Wilson et al., 2008). For example, cytokine signalling is a well-studied regulator of HSCs dormancy and could be involved in the proliferative phenotype observed in Gata2 deficiency. As cytokine signalling can direct lineage choice (Mossadegh-Keller et al., 2013; Rieger et al., 2009), it would be interesting to explore the niche signalling in this context.

Like proliferation, differentiation has been observed to depend on Gata2 dosage. For example, GATA2 knockdown accelerates adipocyte differentiation from mesenchymal stromal cells (Kamata et al., 2014; Tong et al., 2000). While increasing Gata2 expression promotes quiescence, a lower level of overexpression promotes myeloid self-renewal and a block in lymphoid differentiation (Nandakumar et al., 2015). In humans, a lymphoid/myeloid lineages loss has been observed previously in GATA2 deficient patients (Bigley & Collin, 2011), and recently appreciated transcriptionally in GATA2 deficient CD34+ cells (Wu et al., 2020). In *gata2b* knockout zebrafish (**chapter 2**), we observe repercussions involving analogous processes, with a lymphoid biased differentiation and a block in myeloid differentiation. These results are consistent with an essential function for Gata2b in maintaining myeloid differentiation and the lympho-myeloid commitment balance. In **chapter 3,** *gata2b* heterozygosity does not causes major shifts in differentiation, but supports the generation of erythroid dysplastic and (occasionally) myeloid dysplastic cells**.** In **chapter 4** we show that mouse *Gata2* heterozygous HSCs fail to produce sufficient white blood cells. In conclusion, previous research and the data presented in this thesis support the notion that Gata2 contributes to the regulation of HSC proliferation and coordinates their commitment toward lymphoid and myeloid lineages.

From embryonic to chronic Gata2 deficiency: Consequences for the development of malignancy

While acute stress can cause a burst of cycling HSCs, a chronic stress exposes HSCs to sustained proliferative signals which can result in loss of fitness and exhaustion (Baldridge et al., 2010; Ruzankina & Brown, 2007). The chronic hyperproliferation that we observe in Gata2 deficient HSPCs is possibly the cause of the reduced self-renewal capacity of the Gata2$^{+/-}$ HSCs. Considering the germline nature of Gata2 pathogenic mutations and of the mutants used in this thesis, we tested whether embryonic defects might affect adult hematopoiesis. Besides, to assess whether longer time periods were required for haploinsufficiency to emerge, we aged heterozygous gata2b zebrafish and Gata2 mice.

One of the main advantages of using zebrafish as an animal model is its transparency and the possibility to observe its embryogenesis in vivo. In **chapter 2** we show that the generation of HSCs through endothelial to hematopoietic transition (EHT) tends to be reduced both in gata2b knockout heterozygous and homozygous zebrafish. The number of EHT events is directly proportional to the expression of the WT *gata2b* allele, indicating a supportive function for this transcription factor in the generation of hematopoietic stem cells. Instead,

we show that the expression of the other Gata2 orthologue, Gata2a, is necessary for the specification of hemogenic endothelium. In **chapter 4** we confirm that Gata2$^{+/-}$ mice have an obvious loss of HSPCs during embryogenesis (Ling et al., 2004) and conclude that Gata2 causes a reduction in HSPC generation during embryogenesis concomitant with higher proliferative signatures, already starting at the fetal liver stage.

Supported by the fact that GATA2 haploinsufficient patients' phenotypes worsen with age, we aged *Gata2*$^{+/-}$ mice to assess the consequences of the accumulation of defects. The combination of transcriptional and functional analysis pointed to a loss of quiescence and proliferative stress. However, nor previous research (Ling et al., 2004; Rodrigues et al., 2005) nor our results highlighted anomalies in the blood of adult *Gata2*$^{+/-}$. Repetitive BM transplantation causes a continuous call for stem cell function and, just like a loss in quiescence, can cause additive cell-autonomous effects, such as epigenetic dysregulation, mutations, and telomere erosion (Ermolaeva et al., 2018; Mejia-Ramirez & Florian, 2020). However, *Gata2*$^{+/-}$ repetitive bone marrow transplantation (BMT) of *Gata2*$^{+/-}$ young adults did not support the development of a blood disorder in recipients (Rodrigues et al., 2005). These results indicate that proliferative stress, on its own, is not enough to cause a disorder in *Gata2*$^{+/-}$. Instead, Gata2 heterozygous HSCs are exhausted only after ageing and transplantation. Ageing can cause an additional HSPC extrinsic mis-regulation like a chronically inflamed, oxidative stressed microenvironment (Kovtonyuk et al., 2016; Tower, 2012), which can induce maladaptive transformations in HSPCs (Raaijmakers et al., 2010; Walkley et al., 2007; Zambetti et al., 2016). GATA2 is expressed in both hematopoietic and non-hematopoietic cells, including mesenchymal, endothelial, adipocyte, eosinophil, erythroid progenitor and megakaryocytic cells (Minami et al., 2004; Tong et al., 2000; Tsai & Orkin, 1997). These cells comprise a relevant part of the HSCs microenvironment but play a still unexplored role in the pathogenesis of germline GATA2 deficiency. It would be interesting to explore whether the aging microenvironment alone or the fact that the environment is also Gata2 haploinsufficient, contribute to the HSC exhaustion phenotype.

Conclusion and perspectives:

The work described in this dissertation was aimed at better understanding the consequences of GATA2 deficiency. For this purpose, we adopted Gata2 deficient zebrafish and mouse models, with a particular focus on the transcriptional consequences in Gata2 deficient HSPCs.

In **chapter 2,** Gata2b homozygous knockout (*gata2b*$^{-/-}$) zebrafish revealed a role for this transcription factor in proliferation and lineage commitment of early hematopoietic progenitor cells. Adult *gata2b*$^{-/-}$ kidney marrow shows a lymphoid bias and a block in neutrophil differentiation. In **chapter 3**, Gata2b heterozygous knockout (*gata2b*$^{+/-}$) zebrafish exposed the necessity of unaltered levels of Gata2b to avoid the development of dysplasia. Gata2b heterozygosity causes dysplasia in erythroid progenitors of zebrafish kidney marrow.

Taken together, the findings in chapter 2 and 3 highlight the importance of adequate levels of Gata2b in directing HSCs commitment and differentiation of myeloid, lymphoid and erythroid lineage cells. Moreover, we reveal that Gata2a is required during embryogenesis for the specification of hemogenic endothelium, upstream of Gata2b, while Gata2b acquires significant functions during HSPCs expansion and adulthood.

Different dosages of Gata2b result in distinct phenotypes, including cytopenia, loss of differentiation and dysplasia, suggesting a possible relation of GATA2 levels to the clinical outcome of GATA2 deficient patients. Variation of WT GATA2 expression levels occur frequently in GATA2 patients via allele-specific expression (ASE), reaching a 95% chance when associated with *CEBPA* double mutations (Mulet-Lazaro et. al 2021). *GATA2* is the most recurrent gene with ASE in AML and the variations in its expression levels are crucial aspects for understanding the mechanism responsible for the manifestations of GATA2 deficiency. We addressed this aspect analysing the consequences of different gata2b expression levels in zebrafish.

The lack of gross morphologic abnormalities or MDS\AML in our fish does not undervalue our research. Instead it highlights how GATA2 de-regulation represents a predisposition which requires further stress factors, like an infection challenge or secondary mutations, to cause the development of severe cytopenias and MDS/AML. Future experiments in this direction are likely necessary to determine the relation of GATA2 predisposition with environmental factors for the phenotype occurrence, particularly considering the variable penetrance observed in patients.

In **chapter 4** we examined the effects of Gata2 haploinsufficiency at different time points in mice and found that the fitness of aged *Gata2*$^{+/-}$ HSCs is impaired, resulting in a failure to produce sufficient numbers of WBC after bone marrow transplantation. The onset of leukopenia is the first observable mouse blood phenotype derived from GATA2 haploinsufficient HSCs and reflects a major aspect of human GATA2 deficiency. This cytopenic model represents a valuable method for recapitulating characteristics of human GATA2 deficiency in rodents. In patients, possibly the combination of low HSCs numbers and proliferative stress supports the cytopenias and marrow dysplasia, leading to clonal hematopoiesis and malignant transformation. In our model, the increased proliferative stress and aging likely promote leukopenia. Aging is associated with the accumulation of somatic mutations and we are curious to test for acquired driver mutations in our model. Another hallmark related to aging and malignancy is clonal hematopoiesis. If the leukopenia we observed in our model represents a pre-malignant state, evidence for an increased clonal hematopoiesis would be strong evidence in its support.

REFERENCES:

Achi, H. V., Ahui, B. J. M., Anon, J. C., Kouassi, B. A., Dje-Bi, H., & Kininlman, H. (2013). La pancytopénie: Une complication grave de la miliaire tuberculeuse. *Revue Des Maladies Respiratoires*, *30*(1), 33–37. https://doi.org/10.1016/j.rmr.2012.08.008

Baldridge, M. T., King, K. Y., Boles, N. C., Weksberg, D. C., & Goodell, M. A. (2010). Quiescent haematopoietic stem cells are activated by IFN-γ in response to chronic infection. *Nature*, *465*(7299), 793–797. https://doi.org/10.1038/nature09135

Bigley, V., & Collin, M. (2011). Dendritic cell, monocyte, B and NK lymphoid deficiency defines the lost lineages of a new GATA-2 dependent myelodysplastic syndrome. In *Haematologica* (Vol. 96, Issue 8, pp. 1081–1083). https://doi.org/10.3324/haematol.2011.048355

Bresnick, E. H., & Johnson, K. D. (2019). Blood disease–causing and –suppressing transcriptional enhancers: General principles and GATA2 mechanisms. In *Blood Advances* (Vol. 3, Issue 13, pp. 2045–2056). American Society of Hematology. https://doi.org/10.1182/bloodadvances.2019000378

Bullinger, L., Döhner, K., Bair, E., Fröhling, S., Schlenk, R. F., Tibshirani, R., Döhner, H., & Pollack, J. R. (2004). Use of Gene-Expression Profiling to Identify Prognostic Subclasses in Adult Acute Myeloid Leukemia. *New England Journal of Medicine*, *350*(16), 1605–1616. https://doi.org/10.1056/nejmoa031046

Cortés-Lavaud, X., Landecho, M. F., Maicas, M., Urquiza, L., Merino, J., Moreno-Miralles, I., & Odero, M. D. (2015). GATA2 Germline Mutations Impair GATA2 Transcription, Causing Haploinsufficiency: Functional Analysis of the p.Arg396Gln Mutation. *The Journal of Immunology*, *194*(5), 2190–2198. https://doi.org/10.4049/jimmunol.1401868

Dang, V. T., Kassahn, K. S., Marcos, A. E., & Ragan, M. A. (2008). Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *European Journal of Human Genetics*, *16*(11), 1350–1357. https://doi.org/10.1038/ejhg.2008.111

de Pater, E., Kaimakis, P., Vink, C. S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P. S., Camper, S. A., Speck, N., & Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. *Journal of Experimental Medicine*, *210*(13), 2843–2850. https://doi.org/10.1084/jem.20130751

Deutschbauer, A. M., Jaramillo, D. F., Proctor, M., Kumm, J., Hillenmeyer, M. E., Davis, R. W., Nislow, C., & Giaever, G. (2005). Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics*, *169*(4), 1915–1925. https://doi.org/10.1534/genetics.104.036871

Donadieu, J., Lamant, M., Fieschi, C., De Fontbrune, F. S., Caye, A., Ouachee, M., Beaupain, B., Bustamante, J., Poirel, H. A., Isidor, B., Neste, E. Van Den, Neel, A., Nimubona, S., Toutain, F., Barlogis, V., Schleinitz, N., Leblanc, T., Rohrlich, P., Suarez, F., … Pasquet, M. (2018). Natural history of GATA2 deficiency in a survey of 79 French and Belgian patients. *Haematologica*, *103*(8), 1278–1287. https://doi.org/10.3324/haematol.2017.181909

Ermolaeva, M., Neri, F., Ori, A., & Rudolph, K. L. (2018). Cellular and epigenetic drivers of stem cell ageing. In *Nature Reviews Molecular Cell Biology* (Vol. 19, Issue 9, pp. 594–610). Nature Publishing Group. https://doi.org/10.1038/s41580-018-0020-3

Essers, M. A. G., Offner, S., Blanco-Bose, W. E., Waibler, Z., Kalinke, U., Duchosal, M. A., & Trumpp, A. (2009). IFNα activates dormant haematopoietic stem cells in vivo. *Nature*, *458*(7240), 904–908. https://doi.org/10.1038/nature07815

Ezoe, S., Matsumura, I., Nakata, S., Gale, K., Ishihara, K., Minegishi, N., Machii, T., Kitamura, T., Yamamoto, M., Enver, T., & Kanakura, Y. (2002). GATA-2/estrogen receptor chimera regulates cytokine-dependent growth of hematopoietic cells through accumulation of p21WAF1 and p27Kip1 proteins. *Blood*, *100*(10), 3512–3520. https://doi.org/10.1182/blood-2002-04-1177

Ferreira, R., Wai, A., Shimizu, R., Gillemans, N., Rottier, R., Von Lindern, M., Ohneda, K., Grosveld, F., Yamamoto, M.,

& Philipsen, S. (2007). *Dynamic regulation of Gata factor levels is more important than their identity*. https://doi.org/10.1182/blood-2006-11-060491

Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C., & Orkin, S. H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(22), 12355–12358. https://doi.org/10.1073/pnas.93.22.12355

Gao, X., Johnson, K. D., Chang, Y. I., Boyer, M. E., Dewey, C. N., Zhang, J., & Bresnick, E. H. (2013). Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *Journal of Experimental Medicine*, *210*(13), 2833–2842. https://doi.org/10.1084/jem.20130733

Grass, J. A., Boyer, M. E., Pal, S., Wu, J., Weiss, M. J., & Bresnick, E. H. (2003). GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 8811–8816. https://doi.org/10.1073/pnas.1432147100

Grass, J. A., Jing, H., Kim, S.-I., Martowicz, M. L., Pal, S., Blobel, G. A., & Bresnick, E. H. (2006). Distinct Functions of Dispersed GATA Factor Complexes at an Endogenous Gene Locus. *Molecular and Cellular Biology*, *26*(19), 7056. https://doi.org/10.1128/MCB.01033-06

Hasegawa, A., Kaneko, H., Ishihara, D., Nakamura, M., Watanabe, A., Yamamoto, M., Trainor, C. D., & Shimizu, R. (2016). GATA1 Binding Kinetics on Conformation-Specific Binding Sites Elicit Differential Transcriptional Regulation. *Molecular and Cellular Biology*, *36*(16), 2151–2167. https://doi.org/10.1128/mcb.00017-16

Heyworth, C., Gale, K., Dexter, M., May, G., & Enver, T. (1999). A GATA-2/estrogen receptor chimera functions as a ligand-dependent negative regulator of self-renewal. *Genes and Development*, *13*(14), 1847–1860. https://doi.org/10.1101/gad.13.14.1847

Hirabayashi, S., Wlodarski, M. W., Kozyra, E., & Niemeyer, C. M. (2017). Heterogeneity of GATA2-related myeloid neoplasms. In *International Journal of Hematology* (Vol. 106, Issue 2, pp. 175–182). Springer Tokyo. https://doi.org/10.1007/s12185-017-2285-2

Ho, I. C., Tai, T. S., & Pai, S. Y. (2009). GATA3 and the T-cell lineage: Essential functions before and after T-helper-2-cell differentiation. In *Nature Reviews Immunology* (Vol. 9, Issue 2, pp. 125–135). NIH Public Access. https://doi.org/10.1038/nri2476

Kamata, M., Okitsu, Y., Fujiwara, T., Kanehira, M., Nakajima, S., Takahashi, T., Inoue, A., Fukuhara, N., Onishi, Y., Ishizawa, K., Shimizu, R., Yamamoto, M., & Harigae, H. (2014). GATA2 regulates differentiation of bone marrow-derived mesenchymal stem cells. *Haematologica*, *99*(11), 1686–1696. https://doi.org/10.3324/haematol.2014.105692

Katsumura, K. R., & Bresnick, E. H. (2017). The GATA factor revolution in hematology. In *Blood* (Vol. 129, Issue 15, pp. 2092–2102). American Society of Hematology. https://doi.org/10.1182/blood-2016-09-687871

Kovtonyuk, L. V., Fritsch, K., Feng, X., Manz, M. G., & Takizawa, H. (2016). Inflamm-aging of hematopoiesis, hematopoietic stem cells, and the bone marrow microenvironment. In *Frontiers in Immunology* (Vol. 7, Issue NOV). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2016.00502

Kristinsson, S. Y., Björkholm, M., Hultcrantz, M., Derolf, Å. R., Landgren, O., & Goldin, L. R. (2011). Chronic immune stimulation might act as a trigger for the development of acute myeloid leukemia or myelodysplastic syndromes. *Journal of Clinical Oncology*, *29*(21), 2897–2903. https://doi.org/10.1200/JCO.2011.34.8540

Leubolt, G., Redondo Monte, E., & Greif, P. A. (2020). GATA2 mutations in myeloid malignancies: Two zinc fingers in many pies. In *IUBMB Life* (Vol. 72, Issue 1, pp. 151–158). Blackwell Publishing Ltd. https://doi.org/10.1002/iub.2204

Lin, F. C., Karwan, M., Saleh, B., Hodge, D. L., Chan, T., Boelte, K. C., Keller, J. R., & Young, H. A. (2014). IFN-γ causes aplastic anemia by altering hematopoietic stem/progenitor cell composition and disrupting lineage differentiation. *Blood*, *124*(25), 3699–3708. https://doi.org/10.1182/blood-2014-01-549527

Ling, K. W., Ottersbach, K., Van Hamburg, J. P., Oziemlak, A., Tsai, F. Y., Orkin, S. H., Ploemacher, R., Hendriks, R. W., & Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *Journal of Experimental Medicine*, *200*(7), 871–882. https://doi.org/10.1084/jem.20031556

Luesink, M., Hollink, I. H. I. M., Van Der Velden, V. H. J., Knops, R. H. J. N., Boezeman, J. B. M., De Haas, V., Trka, J., Baruchel, A., Reinhardt, D., Van Der Reijden, B. A., Van Den Heuvel-Eibrink, M. M., Zwaan, C. M., & Jansen, J. H. (2012). High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. *Blood*, *120*(10), 2064–2075. https://doi.org/10.1182/blood-2011-12-397083

Martin, D. I. K., Zon, L. I., Mutter, G., & Orkin, S. H. (1990). Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature*, *344*(6265), 444–447. https://doi.org/10.1038/344444a0

Matatall, K. A., Jeong, M., Chen, S., Sun, D., Chen, F., Mo, Q., Kimmel, M., & King, K. Y. (2016). Chronic Infection Depletes Hematopoietic Stem Cells through Stress-Induced Terminal Differentiation. *Cell Reports*, *17*(10), 2584–2595. https://doi.org/10.1016/j.celrep.2016.11.031

McReynolds, L. J., Calvo, K. R., & Holland, S. M. (2018). Germline GATA2 Mutation and Bone Marrow Failure. In *Hematology/Oncology Clinics of North America* (Vol. 32, Issue 4, pp. 713–728). W.B. Saunders. https://doi.org/10.1016/j.hoc.2018.04.004

Mejia-Ramirez, E., & Florian, M. C. (2020). Understanding intrinsic hematopoietic stem cell aging. In *Haematologica* (Vol. 105, Issue 1, pp. 22–37). Ferrata Storti Foundation. https://doi.org/10.3324/haematol.2018.211342

Minami, T., Murakami, T., Horiuchi, K., Miura, M., Noguchi, T., Miyazaki, J. I., Hamakubo, T., Aird, W. C., & Kodama, T. (2004). Interaction between Hex and GATA transcription factors in vascular endothelial cells inhibits flk-1/KDR-mediated vascular endothelial growth factor signaling. *Journal of Biological Chemistry*, *279*(20), 20626–20635. https://doi.org/10.1074/jbc.M308730200

Morrill, S. A., & Amon, A. (2019). Why haploinsufficiency persists. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(24), 11866–11871. https://doi.org/10.1073/pnas.1900437116

Mossadegh-Keller, N., Sarrazin, S., Kandalla, P. K., Espinosa, L., Richard Stanley, E., Nutt, S. L., Moore, J., & Sieweke, M. H. (2013). M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature*, *497*(7448), 239–243. https://doi.org/10.1038/nature12026

R, Mulet-Lazaro, van Herk S, Erpelinck C, Bindels E, Sanders MA, Vermeulen C, Renkens I, et al. 2021. "Allele-Specific Expression of GATA2 Due to Epigenetic Dysregulation in CEBPA Double-Mutant AML." *Blood* 138 (2). https://doi.org/10.1182/BLOOD.2020009244.

Mutsaers, P. G. N. J., Van De Loosdrecht, A. A., Tawana, K., Bödör, C., Fitzgibbon, J., & Menko, F. H. (2013). Highly variable clinical manifestations in a large family with a novel GATA2 mutation. *Leukemia*, *27*(11), 2247–2248. https://doi.org/10.1038/leu.2013.105

Nandakumar, S. K., Johnson, K., Throm, S. L., Pestina, T. I., Neale, G., & Persons, D. A. (2015). Low-level GATA2 overexpression promotes myeloid progenitor self-renewal and blocks lymphoid differentiation in mice. *Experimental Hematology*, *43*(7), 565-577.e10. https://doi.org/10.1016/j.exphem.2015.04.002

Orford, K. W., & Scadden, D. T. (2008). Deconstructing stem cell self-renewal: Genetic insights into cell-cycle regulation. In *Nature Reviews Genetics* (Vol. 9, Issue 2, pp. 115–128). Nat Rev Genet. https://doi.org/10.1038/nrg2269

Persons, D. A., Allay, J. A., Allay, E. R., Ashmun, R. A., Orlic, D., Jane, S. M., Cunningham, J. M., & Nienhuis, A. W. (1999). Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood*, *93*(2), 488–499. https://doi.org/10.1182/blood.v93.2.488.402k17_488_499

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., & Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, *349*(6306), 257–260. https://doi.org/10.1038/349257a0

Raaijmakers, M. H. G. P., Mukherjee, S., Guo, S., Zhang, S., Kobayashi, T., Schoonmaker, J. A., Ebert, B. L., Al-Shahrour, F., Hasserjian, R. P., Scadden, E. O., Aung, Z., Matza, M., Merkenschlager, M., Lin, C., Rommens, J. M., & Scadden,

D. T. (2010). Bone progenitor dysfunction induces myelodysplasia and secondary leukaemia. *Nature*, *464*(7290), 852–857. https://doi.org/10.1038/nature08851

Rieger, M. A., Hoppe, P. S., Smejkal, B. M., Eitelhuber, A. C., & Schroeder, T. (2009). Hematopoietic cytokines can instruct lineage choice. *Science*, *325*(5937), 217–218. https://doi.org/10.1126/science.1171461

Rodrigues, N. P., Janzen, V., Forkert, R., Dombkowski, D. M., Boyd, A. S., Orkin, S. H., Enver, T., Vyas, P., & Scadden, D. T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*, *106*(2), 477–484. https://doi.org/10.1182/blood-2004-08-2989

Ruzankina, Y., & Brown, E. J. (2007). Relationships between stem cell exhaustion, tumour suppression and ageing. In *British Journal of Cancer* (Vol. 97, Issue 9, pp. 1189–1193). Nature Publishing Group. https://doi.org/10.1038/sj.bjc.6604029

Sato, T., Onai, N., Yoshihara, H., Arai, F., Suda, T., & Ohteki, T. (2009). Interferon regulatory factor-2 protects quiescent hematopoietic stem cells from type i interferon-dependent exhaustion. *Nature Medicine*, *15*(6), 696–700. https://doi.org/10.1038/nm.1973

Seidman, J. G., & Seidman, C. (2002). Transcription factor haploinsufficiency: when half a loaf is not enough. *Journal of Clinical Investigation*, *109*(4), 451–455. https://doi.org/10.1172/jci15043

Shivdasani, R. A., Fujiwara, Y., McDevitt, M. A., & Orkin, S. H. (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO Journal*, *16*(13), 3965–3973. https://doi.org/10.1093/emboj/16.13.3965

Singh, S., Jakubison, B., & Keller, J. R. (2020). Protection of hematopoietic stem cells from stress-induced exhaustion and aging. *Current Opinion in Hematology*, *27*(4), 225–231. https://doi.org/10.1097/MOH.0000000000000586

Spinner, M. A., Sanchez, L. A., Hsu, A. P., Shaw, P. A., Zerbe, C. S., Calvo, K. R., Arthur, D. C., Gu, W., Gould, C. M., Brewer, C. C., Cowen, E. W., Freeman, A. F., Olivier, K. N., Uzel, G., Zelazny, A. M., Daub, J. R., Spalding, C. D., Claypool, R. J., Giri, N. K., … Holland, S. M. (2014). GATA2 deficiency: A protean disorder of hematopoiesis, lymphatics, and immunity. *Blood*, *123*(6), 809–821. https://doi.org/10.1182/blood-2013-07-515528

Takahashi, S., Shimizu, R., Suwabe, N., Kuroha, T., Yoh, K., Ohta, J., Nishimura, S., Lim, K.-C., Engel, J. D., & Yamamoto, M. (2000). GATA factor transgenes under GATA-1 locus control rescue germline GATA-1 mutant deficiencies. *Blood*, *96*(3), 910–916. https://doi.org/10.1182/blood.v96.3.910

Tipping, A. J., Pina, C., Castor, A., Hong, D., Rodrigues, N. P., Lazzari, L., May, G. E., Jacobsen, S. E. W., & Enver, T. (2009). High GATA-2 expression inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. *Blood*, *113*(12), 2661–2672. https://doi.org/10.1182/blood-2008-06-161117

Tong, Q., Dalgin, G., Xu, H., Ting, C. N., Leiden, J. M., & Hotamisligil, G. S. (2000). Function of GATA transcription factors in preadipocyte-adipocyte transition. *Science*, *290*(5489), 134–138. https://doi.org/10.1126/science.290.5489.134

Tower, J. (2012). Stress and stem cells. *Wiley Interdisciplinary Reviews: Developmental Biology*, *1*(6), 789–802. https://doi.org/10.1002/wdev.56

Tsai, F. Y., Browne, C. P., & Orkin, S. H. (1998). Knock-in mutation of transcription factor GATA-3 into the GATA-1 locus: Partial rescue of GATA-1 loss of function in erythroid cells. *Developmental Biology*, *196*(2), 218–227. https://doi.org/10.1006/dbio.1997.8842

Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., & Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, *371*(6494), 221–226. https://doi.org/10.1038/371221a0

Tsai, F. Y., & Orkin, S. H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood*, *89*(10), 3636–3643. https://doi.org/10.1182/blood.v89.10.3636.3636_3636_3643

Vicente, C., Vazquez, I., Conchillo, A., García-Sánchez, M. A., Marcotegui, N., Fuster, O., González, M., Calasanz, M. J., Lahortiga, I., & Odero, M. D. (2012). Overexpression of GATA2 predicts an adverse prognosis for patients with acute myeloid leukemia and it is associated with distinct molecular abnormalities. In *Leukemia* (Vol. 26, Issue 3, pp. 550–554). https://doi.org/10.1038/leu.2011.235

Walkley, C. R., Olsen, G. H., Dworkin, S., Fabb, S. A., Swann, J., McArthur, G. A. A., Westmoreland, S. V., Chambon, P., Scadden, D. T., & Purton, L. E. (2007). A Microenvironment-Induced Myeloproliferative Syndrome Caused by Retinoic Acid Receptor γ Deficiency. *Cell*, *129*(6), 1097–1110. https://doi.org/10.1016/j.cell.2007.05.014

Wilson, A., Laurenti, E., Oser, G., van der Wath, R. C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C. F., Eshkind, L., Bockamp, E., Lió, P., MacDonald, H. R., & Trumpp, A. (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell*, *135*(6), 1118–1129. https://doi.org/10.1016/j.cell.2008.10.048

Wlodarski, M. W., Collin, M., & Horwitz, M. S. (2017). GATA2 deficiency and related myeloid neoplasms. *Seminars in Hematology*, *54*(2), 81–86. https://doi.org/10.1053/j.seminhematol.2017.05.002

Wu, Z., Gao, S., Diamond, C., Kajigaya, S., Chen, J., Shi, R., Palmer, C., Hsu, A. P., Calvo, K. R., Hickstein, D. D., Holland, S. M., & Young, N. S. (2020). Sequencing of RNA in single cells reveals a distinct transcriptome signature of hematopoiesis in GATA2 deficiency. *Blood Advances*, *4*(12), 2702–2716. https://doi.org/10.1182/bloodadvances.2019001352

Zambetti, N. A., Ping, Z., Chen, S., Kenswil, K. J. G., Mylona, M. A., Sanders, M. A., Hoogenboezem, R. M., Bindels, E. M. J., Adisty, M. N., Van Strien, P. M. H., van der Leije, C. S., Westers, T. M., Cremers, E. M. P., Milanese, C., Mastroberardino, P. G., van Leeuwen, J. P. T. M., van der Eerden, B. C. J., Touw, I. P., Kuijpers, T. W., … Raaijmakers, M. H. G. P. (2016). Mesenchymal Inflammation Drives Genotoxic Stress in Hematopoietic Stem Cells and Predicts Disease Evolution in Human Pre-leukemia. *Cell Stem Cell*, *19*(5), 613–627. https://doi.org/10.1016/j.stem.2016.08.021

Zhang, S. J., Ma, L. Y., Huang, Q. H., Li, G., Gu, B. W., Gao, X. D., Shi, J. Y., Wang, Y. Y., Gao, L., Cai, X., Ren, R. B., Zhu, J., Chen, Z., & Chen, S. J. (2008). Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(6), 2076–2081. https://doi.org/10.1073/pnas.0711824105

A

Addendum

ENGLISH SUMMARY

Hematopoietic stem cells (HSCs) provide the millions of differentiated blood cells we need in our everyday life. In blood we distinguish the lymphoid and myeloid lineage. The lymphoid lineage generates adaptive immune cells like T cells and B cells. The myeloid lineage generates erythrocytes for oxygen transport, megakaryocytes for coagulation and innate immune cells like neutrophils and monocytes. These lineages have to be produced at the right proportion to maintain these processes. Transcriptional regulation of HSCs is an essential mechanism to maintain a balanced production of blood cells, and the transcription factor GATA2 plays a key function in this process. GATA2 directs embryonic and adult hematopoiesis in vertebrates by occupying GATA-DNA motifs in numerous genes. Homozygous GATA2 mutations in mice are embryonically lethal, making it difficult to assess its function in the different aspects of hematopoiesis. Heterozygous germline *GATA2* mutations in humans are the cause of a series of disorders called GATA2 deficiency syndromes. They lead to a spectrum of phenotypes including lymphedema, cytopenias and, in more than 80% of patients, MDS or leukemia. To date, there is no clear correlation between the patient phenotype and the type of GATA2 mutation. The exact role of GATA2 in HSCs generation, expansion, self-renewal and differentiation is still unknown.

In this thesis, we focus on the hematopoietic consequences of GATA2 deficiency in zebrafish (**chapter 2-3**) and mouse (**chapter 4**).

Zebrafish have two orthologues of GATA2; i.e., Gata2a and Gata2b. Previous studies have shown that *gata2b* is prominently expressed in HSPCs, whereas *gata2a* is mainly expressed in the vasculature. Being the predominant GATA2 orthologue required for the maintenance of hematopoietic stem cells, we targeted *gata2b* to generate a knockout zebrafish line adopted in **chapters 2-3**. Interestingly, the homozygous Gatat2b mutant is viable, allowing us to assess the functionality of Gata2b in adult hematopoiesis.

In **chapter 2** we characterize the effects of Gata2b homozygous knockout (*gata2b*$^{-/-}$) in the hematopoietic system. Gata2b is required for the proliferation of hematopoietic stem and progenitor cells (HSPCs) during embryonic HSPCs expansion in the caudal hematopoietic tissue. Adult *gata2b*$^{-/-}$ HSPCs are more proliferative and their kidney marrow shows a lymphoid bias and a block in neutrophil differentiation. Single cell transcriptome analysis showed that the HSPCs were the origin of the increased lymphoid lineage output in *gata2b*$^{-/-}$ kidney marrow cells, due to a failure to initiate myeloid gene expression to sufficient levels and a subsequent co-expression of both myeloid and lymphoid genes in *gata2b*$^{-/-}$ HSPCs.

To evaluate the transcriptional and phenotypic consequences of half Gata2b dose in zebrafish hematopoiesis, in **chapter 3** we assessed *gata2b* heterozygous zebrafish (*gata2b*$^{+/-}$). We hypothesized that, with half Gata2b dose, zebrafish would still present reduced myeloid differentiation, although to a lesser extent than when knocking out *gata2b* homozygously. Instead, we observed an unprecedented erythroid dysplasia in the kidney marrow, but a

conservation of the major differentiation lineages. The HSPC compartment also showed transcriptional changes fitting with the transcriptional changes as found in MDS.

The clear dichotomy in phenotypes between the heterozygous Gata2b mutant and the homozygous Gata2b mutant show that, depending on its dose, Gata2 has distinct functions in the maintenance of HSPCs and the subsequent differentiation into the different hematopoietic lineages. This could possibly explain the variable phenotypes found in GATA2 patients.

Mice, like humans, possess one copy of the *Gata2* gene. *Gata2* double knock out (*Gata2*$^{-/-}$) mouse embryos succumb before the generation of HSCs, while *Gata2* heterozygous knock-out mice (*Gata2*$^{+/-}$) do not develop noticeable hematological phenotypes. We hypothesized that a transcriptional analysis of HSPCs would reveal unnoticeable peculiarities cause by *Gata2* heterozygosity. In **chapter 4** we found that *Gata2*$^{+/-}$ HSPCs are hyperproliferative and to determine the long term consequences of this condition, we let *Gata2*$^{+/-}$ age up to 14-15 months. *Gata2*$^{+/-}$ aged HSPCs remain hyperproliferative and cause a failure to produce sufficient numbers of white blood cells after bone marrow transplantation.

The onset of leukopenia after transplantation of aged *Gata2*$^{+/-}$ HSCs reflects a major aspect of human GATA2 deficiency and provides a consistent methodology for reproducing a GATA2 associated disorder in mice. Further experiments will be necessary to better characterize the leukopenia identified and to test if other stress factors can contribute to a similar outcome in a reduced time-span. Since aging is necessary for the leukopenia to occur, it is would be interesting addressing its role in the phenotype development such as bone marrow niche-mediated inflammation.

Collectively our findings provide insights into the biology of GATA2 and the hematopoietic consequences following its deficiency. They establish the human disease relevance of the *Gata2*$^{+/-}$ mouse model and spotlight an essential role for Gata2 in HSPCs lineage output/ proliferative stress in human GATA2 deficiency. These findings pave the way for further research on the synergism between HSPCs reduction, loss of quiescence and altered lineage output caused by Gata2 deficiency.

NEDERLANDSE SAMENVATTING

Hematopoëtische stamcellen (HSC's) leveren de miljoenen gedifferentieerde bloedcellen die we in ons dagelijks leven nodig hebben. In bloed kunnen we een onderscheid maken tussen bloedcellen van lymfoïde en myeloïde afkomst. De lymfoïde tak levert adaptieve immuuncellen zoals T-cellen en B-cellen. De myeloïde tak genereert erytrocyten voor zuurstoftransport, megakaryocyten voor coagulatie, en aangeboren immuuncellen zoals neutrofielen en monocyten. Deze verschillende bloedcellen moeten in de juiste verhouding worden geproduceerd om deze processen in stand te houden. Transcriptionele regulatie van HSC's is een essentieel mechanisme om een gebalanceerde productie van bloedcellen te behouden, en de transcriptiefactor GATA2 vervult een sleutelrol in dit proces.

GATA2 stuurt embryonale en volwassen hematopoëse in gewervelden door de GATA-DNA-motieven van talrijke genen te bezetten. Homozygote GATA2-mutaties zijn in muizen embryonaal dodelijk, waardoor het moeilijk is om de functie van GATA2 in de verschillende aspecten van hematopoëse vast te stellen. Heterozygote kiembaan GATA2-mutaties in mensen zijn de oorzaak van een reeks aandoeningen die als GATA2-deficiëntiesyndromen worden aangeduid. Ze leiden tot een spectrum van fenotypes, waaronder lymfoedeem, cytopenieën, en bij meer dan 80% van de patiënten, MDS of leukemie.

Tot op heden is er geen duidelijk verband tussen het fenotype van de patiënt en het type GATA2-mutatie. De exacte rol van GATA2 bij het genereren, vermenigvuldigen, zelf-vernieuwen en differentiëren van HSC's is nog onbekend. In dit proefschrift richten we ons op de hematopoëtische gevolgen van GATA2-deficiëntie bij zebravissen (**hoofdstuk 2-3**) en muizen (**hoofdstuk 4**).

Zebravissen hebben twee orthologen van GATA2: Gata2a en Gata2b. Eerdere studies hebben aangetoond dat *gata2b* prominent tot expressie komt in HSPC's, terwijl *gata2a* voornamelijk tot expressie komt in het bloedvatenstelsel. Aangezien dat de eerste dé GATA2 ortholoog is die nodig is voor het behoud van hematopoëtische stamcellen, hebben we ons gericht op *gata2b* om een knock-out zebravislijn te generen waar gebruik van is gemaakt in **hoofdstukken 2-3**. Opvallend is dat de homozygote Gatat2b-mutant levensvatbaar is, waardoor we de functie van Gata2b in volwassen-hematopoëse kunnen bepalen. In **hoofdstuk 2** karakteriseren we de effecten van Gata2b homozygote knock-out (*gata2b*$^{-/-}$) in het hematopoëtische systeem. Gata2b is nodig voor de proliferatie van hematopoëtische stam- en voorlopercellen (HSPC's) tijdens embryonale HSPC-expansie in het caudale hematopoëtische weefsel. Volwassen *gata2b*$^{-/-}$ HSPC's zijn meer proliferatief en hun niermerg vertoont een lymfoïde neiging en een blokkade in neutrofieldifferentiatie. Eencellige transcriptoomanalyse toonde aan dat de HSPC's de bron waren van de verhoogde lymfoïdecellen productie in *gata2b*$^{-/-}$ niermergcellen, als gevolg van het niet voldoende induceren van myeloïde genexpressie en een daaropvolgende co-expressie van zowel myeloïde als lymfoïde genen in *gata2b*$^{-/-}$ HSPC's.

Om de transcriptionele en fenotypische gevolgen van een halve dosis Gata2b voor de hematopoëse van de zebravis te evalueren, hebben we in **hoofdstuk 3** *gata2b* heterozygote zebravis (*gata2b* $^{+/-}$) onderzocht. Onze hypothese was dat zebravissen met een halve Gata2b-dosis ook een verminderde myeloïde differentiatie zouden vertonen, zij het in mindere mate dan bij het homozygote deficiëntie van gata2b. In plaats daarvan observeerden we een ongekende erytroïde dysplasie in het niermerg, maar een behoud van de belangrijkste differentiatielijnen.

Het HSPC-compartiment vertoonde ook transcriptionele veranderingen die overeenkomen met de transcriptionele veranderingen gevonden in humane MDS. Dit zou mogelijk de variabele fenotypes kunnen verklaren die worden aangetroffen bij GATA2-patiënten.

Muizen bezitten, net als mensen, één kopie van het Gata2-gen. Gata2 dubbele knock-out (Gata2$^{-/-}$) muizenembryo's bezwijken vóór de generatie van HSC's, terwijl Gata2 heterozygote knock-out muizen (Gata2$^{+/-}$) geen merkbare hematologische fenotypes ontwikkelen. We hypothetiseerden dat een transcriptionele analyse van HSPC's onopgemerkte eigenaardigheden, die door Gata2 heterozygotie werden veroorzaakt, zou onthullen.

In hoofdstuk 4 ontdekten we dat *Gata2*$^{+/-}$ HSPC's hyperproliferatief zijn en om de langetermijngevolgen van deze aandoening te bepalen, hebben we *Gata2*$^{+/-}$ muizen tot 14-15 maanden oud laten worden. *Gata2*$^{+/-}$ verouderde HSPC's blijven hyperproliferatief en veroorzaken een onvoldoende productie van wittebloedcel aantallen na beenmergtransplantatie. Het begin van leukopenie na transplantatie van verouderde Gata2$^{+/-}$ HSC's weerspiegelt een belangrijk aspect van menselijke GATA2-deficiëntie en biedt een consistente methodologie voor het reproduceren van een GATA2-geassocieerde aandoening in muizen. Verdere experimenten zullen nodig zijn om de geïdentificeerde leukopenie beter te karakteriseren en om uit te zoeken of andere stressfactoren kunnen bijdragen aan een vergelijkbaar fenotype in een kortere tijdspanne. Aangezien veroudering noodzakelijk is om leukopenie te laten optreden, zou het interessant zijn om de rol hiervan in de ontwikkeling van het fenotype, zoals bij niche-gemedieerde ontsteking van het beenmerg, aan te pakken.

Gezamenlijk bieden onze bevindingen inzicht in de biologie van GATA2 en de hematopoëtische gevolgen die na deficiëntie hiervan ontstaan. Ze stellen de relevantie van het Gata2$^{+/-}$ muismodel voor de menselijke ziekte vast en benadrukken een essentiële rol voor Gata2 in HSPC- myeloïde/lymfoïde productie en proliferatieve stress in humane GATA2-deficiëntie. Deze bevindingen maken de weg vrij voor verder onderzoek naar de synergie tussen HSPCs-reductie, verlies van rust, en gewijzigde myeloïde/lymfoïde verhouding veroorzaakt door Gata2-deficiëntie.

LIST OF ABBREVIATIONS

AGM	Aorta-gonad-mesonephrons
AML	Acute myeloid leukemia
BM	Bone marrow
CD	Cluster of differentiation
CFU	Colony-forming unit
CHT	Caudal hematopoietic tissue
CLP	Common lymphoid progenitor
CML	Chronic myelogenous leukemia
CMML	Chronic myelomonocytic leukemia
CMP	Common myeloid progenitor
DC	Dendritic cell
E	Embryonic day
EHT	Endothelial to hematopoietic transition
ES	Enrichment score
FACS	Fluorescence activated cell sorting
FC	Fold change
FDR	False discovery rate
FPKM	Fragments per kilobase of exon per million fragments mapped
FWD	Forward
G-CSF	Granulocyte colony-stimulating factor
GFP	Green fluorescent protein
GO	Gene ontology
GSEA	Gene set enrichment analysis
γH2AX	Ser139-phosphorylated H2AX histone
HSC	Hematopoietic stem cell
HPC	Hematopoietic progenitor cell
HPV	Human papillomavirus
HSPC	Hematopoietic stem and progenitor cell
LOF	Loss of function
LSK	Lineage negative C-kit positive Sca-1 positive
LT-HSC	Long-term HSC
MDS	Myelodysplastic syndromes
MFI	Mean fluorescence intensity
MGG	May Grünwald Giemsa
MonoMac	Monocytopenia and Mycobacterium Avium infection
MPP	Multipotent progenitor
NK	Natural killer cell

NTM	nontuberculous mycobacterial
PCA	Principal component analysis
RBC	Red blood cell
REV	Reverse
ROS	Reactive oxygen species
SPF	Specific-pathogen-free
ST-HSC	Short-term HSC
WBC	White blood cell
WT	Wild type
ZF	Zinc finger

CURRICULUM VITAE

Emanuele Gioacchino was born on the 25th of May 1989 in Urbino, Italy. After receiving his high school diploma from "Istituto tecnico Mazzocchi" (Ascoli Piceno, Italy) in 2008, he studied Biology at University of New Orleans, University of Manchester, and University of Tor Vergata (Rome), obtaining the bachelor degree from the latter in 2012. He pursued his master degree in Molecular Biology at the University of Tor Vergata, completing an internship in the research group of Prof. Castagnoli (Rome) and an internship in the research group of Louise Purton (St Vincent's Institute of medical research, Melbourne). In 2015 he defended his thesis titled "Impact of Hoxa1 expression on MDS" and graduating *cum laude*. In 2016 he moved to the Netherlands where he was appointed PhD candidate in the research group of Dr. De Pater at the department of Hematology of Erasmus Medical Center (Rotterdam). Here he focused on the role of Gata2 deficiency in mouse and zebrafish hematopoiesis.

Scholarships and Awards

Dutch Hematology Congress, Oral presentation in 2019 and 2020. **International society of experimental hematology,** poster presentation in 2018. **Laziodisu Scholarship,** Lazio region, Rome. Every year since 2009 until 2015. **Erasmus Scholarships,** University of Barcelona, February-July 2015. **Overseas Thesis Scholarship,** University of Tor Vergata, Rome. January 2015. **Overseas Scholarship,** University of Tor Vergata, Rome. September-December 2014. **Erasmus Scholarship,** University of Manchester: September 2010-June 2011. **St Charles Scholarship,** University of New Orleans, LA. September 2008- May 2009

PHD PORTFOLIO

Name: Emanuele Gioacchino
Erasmus MC department: Hematology
Research school: Molecular Medicine (MolMed)

PhD period: May 2016-May 2021
Promoter: Prof.dr. I.P. Touw
Supervisor: Dr. Emma de Pater

1. PhD Training	**Year**	**ECTS**
General courses		
Laboratory Animal Science (Art.9)	2016	4.2
Biomedical Research Techniques	2016	1.5
Biomedical English writing course	2019	2.5
Annual course on Molecular Medicine	2017	0.7
Biomedical Research Techniques	2016	1.5
Photoshop and Illustrator workshop	2019	0.3
In-depth Courses and Workshops		
Molecular aspects of hematological disorders (2x)	2016-2017	1.4
BD Flow Cytometry Course	2017	0.2
Species-Specific Laboratory fish and claw frogs course (Hubrecht institute)	2016	0.6
Galaxy for NGS	2017	1
Microscopic Image analysis: from Theory to Practice	2017	0.8
Basic course on R	2019	1.8
Scientific Meetings Department of Hematology		
Work discussion (Weekly)	2016-2020	8
Journal club/literature discussion (bi-monthly)	2016-2020	7
PhD lunch with invited speaker (Monthly)	2016-2020	2.5
Erasmus Hematology Lectures (Monthly)	2016-2020	2
Hematology floor meetings (Weekly)	2016-2020	8
National/International conferences		
Molecular Medicine Day (2x) (Rotterdam)	2017-2018	0.6
Dutch hematology congress (3x) (Arnhem)	2017-2019	0.9
International society of experimental hematology (Frankfurt)	2017	0.3
European Hematology Association (Amsterdam)	2019	0.3
Presentations/posters		
Departmental work discussion (Oral 7x) (Rotterdam)	2016-2020	3.5
Journal clubs (Oral 4x) (Rotterdam)	2016-2020	2

Dutch Hematology Congress (Oral) (Arnhem)	2018	1
European Hematology Association (Poster) (Amsterdam)	2019	1
International society of experimental hematology (Poster) (Frankfurt)	2017	1
Hematology floor meetings (Oral 4x)	2016-2020	2
2. Teaching, Supervision & Organization Activities		
Organization of PhD day	2017	0.2
Organization and supervision PhD lunch with invited speakers (3x)	2017-2019	0.6
Total		**57.4**

LIST OF PUBLICATIONS

Emanuele Gioacchino*, Cansu Koyunlar*, Joke Zink, Hans de Looper, Madelon de Jong, Tomasz Dobrzycki, Christopher B Mahony, Remco Hoogenboezem, Dennis Bosch, Paulina M H van Strien, Martin E van Royen, Pim J French, Eric Bindels, Kirsten J Gussinklo, Rui Monteiro, Ivo P Touw and Emma de Pater. Essential role for Gata2 in modulating lineage output from hematopoietic stem cells identified in zebrafish. *Blood Adv* (2021) 5 (13): 2687–2700.

Emanuele Gioacchino, Cansu Koyunlar, Joke Zink, Hans de Looper, Remco Hoogenboezem, Eric Bindels, Ivo Touw and Emma de Pater. Zebrafish gata2b haploinsufficiency results in erythroid-dysplasia. Manuscript submitted.

Emanuele Gioacchino, Cansu Koyunlar, Hans de Looper, Joke Peulen, Mariëtte Ter Borg, Remco Hoogenboezem, Dennis Bosch, Paulette van Strien, Kirsten J Gussinklo, Eric Bindels, Elaine Dzierzak, Ivo Touw and Emma de Pater. GATA2 haploinsufficiency reduces fitness of aged hematopoietic stem cells. Manuscript in submission.

ACKNOWLEDGMENTS

The PhD experience has transformed me during the years and drastically made me grow as a man and researcher. Schopare. I would like to briefly thank all of you for making these years memorable.

First of all a heartfelt thanks to my supervisor Dr. Emma de Pater. Dear **Emma**, you have guided my research under every aspect, from zebrafish handling to microscopy, writing, presenting and more. I still smile remembering your joy after the PCR confirming germline transmission of the *gata2b* mutation. You spent so much time training me and correcting my thesis that we *almost* understand each other without introducing the topic. As the first PhD student of your group it has been interesting observing the responsibility and developments associated with the creation of a new research group. I really appreciated your encouragement during my underwhelming reaction to results and I admired your capacity to deal with your family and your job in an apparent effortless manner. I wish you satisfying progresses at the transplantation lab and all the best for the continuation of your research group in the future.

My warmest thanks goes to my promotor Prof. dr. **Ivo Touw**. Your enthusiasm for science has been inspiring for me and probably for a generation of students. Thank you for all your support, key comments and suggestions for my dissertation. Your active presence during the Friday morning seminars was much better than a coffee. Thanks to you I had to (secretly) learn how to make the original carbonara. I still hope for you to try it once.

I would like to thank the members of the doctoral reading committee, Prof. dr. **Ruud Delwel**, Prof. dr. J.N.J. **Sjaak Philipsen**, Prof. dr. **Anna Bigas** for scrupulously and efficiently reading my dissertation. A big thanks to all the PIs in the Hematology Department, including **Mark**, **Rebecca**, **Tom**, **Mojca**, **Peter**, **Frank**, **Moniek**. Your questions and suggestions during our meetings helped me better understand how to approach science and how to reason my research.

This thesis would not be complete without expressing a huge gratitude to the dj, artist, paranymph and soon to be Dr. Koyunlar. Dear **Cansu**, the first thing I knew about you was that you like to play the piano☺ Soon after I discovered your passion for scientific research, art, music and one of my best amigos and lovers **Raz**. You have been my closest colleague but also one of my best friends and confidant. The joy you transmit every day is obvious to everybody and helped me to overcome the Dutch toneless days. Thank you for all the GATA2 discussions, the experiments and laughter we had together in the last 4 years. You are evolving fast as a scientist and will defend your thesis very soon. I'm sure you will make your daddy proud!

Doctor **Hans** you were the first person teaching me and showing me the lab tools. Since then you have been a solid reference point for our experiments and always with extreme calm and reliability. A huge thanks for doing the hard work in the generation of the

zebrafish *gata2b* mutant and a myriad of thanks for all your help and lighting the mood of the lab together with **Sensei. Dr Eric** you are possibly the most hardworking person in the department. I want to especially thank you for helping me and Cansu with the single cell RNA sequencing. Your offstage presence in the laboratory is of vital importance and allowed plenty of students (me included) to conclude their experiments.

Joke: Our dear fishwoman. Thank you for taking good care and respect of our little friends. You had to learn to handle many aspects in our group. I am particularly thankful for the analysis of the kidney marrow cells, it must have been tedious and not a joke. Good luck with your future career.

Thank you **Paulette** for helping me with the mouse stuff I couldn't do and suerte being a great double mama.

One, two, three and to the four Dr D is at the door. **Dennis** Keep It Real.

I owe a huge amount of gratitude to the people that helped me during these years and taught me how to use softwares and machines. Thank you **Roger** and **Remco** for your help in the bioinformatics support. You were always very busy but always found some time to help me in the bulk and single cell RNA sequencing. Another big thanks to **Michael** for your flow cytometry tricks and your patience with mine and Cansu's sample analysis. Thank you **Madelon** for the script you provided us and the support in the single cell analysis. I have always been a big fan of your clothing style. To **Claire** for helping with flow cytometry and antibodies. I still owe you a tiramisu. Good luck with your new family and your beautiful Nora.

Thank you Patty **Patricia** for giving me a very good impression of the Department after my first interview, for your tips and for your contagious laughter. Good luck in Utrecht!

Dear **Eline**, my trusty presence to my right side. We have been sitting next to each other for a looong time. I hope one day to be as productive as you at work. Thanks for all the suggestions you gave me, short but to the point!

A big thanks to all the remaining members of the Hematology Department who helped me one way or another like **Tessa**, **Paola, Andrea**, **Bella**, **Inge**, **Jaqueline**, **Helene and** the other students and technicians.

To **Egied**, thank you for your help in the dissertation design and the patience with my continuous modifications.

Dr. Mr. Keane, what would I be doing without you? It has been an honor to teach you the "care-free, bordering on the lackadaisical, attitude" and how to play basketball. I'm very grateful for the graphpad and flojo tricks you taught me in the asian office when Ping was too annoyed to repeat twice, and for the endless suggestions you always provided. Thanks to your gift of the gab I met most of my friends in Rotterdam. I think the 43 shared whatsapp groups summarize well how many extra lab activities we have been through. You will always be a honor member of my community. With your new job you became more serious and less of a party king. Materialism is a false prophet. Keep fit and eat your vitamins. I hope you and Nina will take care of each other like you did in the past and will settle in a more latin country.

Dr. Mista P.i.n.g. you are not just a friend but represent for me hard work, perfectionism and honor. You are a nice character and someway always give me joy. I'm grateful for your tips and vision of the scientific world and for the FACS and flojo tricks I was eager to learn from you. My honor for being your host in Italy and at Bababoom. Claudio and Jacopo keep on asking me about you! Even if we have opposite views on priorities we found an unexpected connection which I hope will last even if you won't invite me to your house anymore (I'm again sorry about the plant;)). Talking about plants… Sap Mr. **Onnoooo!** Your bush has been a point of reference both during festivals and during my lab days. You have been always very kind to the newbies like me and were available for helping with confocal, stress assays and a bunch of other technical issues. Some time has passed since our first out-the-lab activity in Rochussentraat. Our shared interests in music, travelling, nature and fun gave us plenty of opportunities to have some good memories outside the lab which I hope to repeat as soon as possible. What about when a guy started crying for a drum solo at the jazz festival in Tivoli? Thanks for being such a kind, solid and happy column in my years in Roffa. I am sure I will keep on seeing you and your bush clones in the future!

Multumesc dragostea mea **Raz**, together with Ivasito you made me enter into a new world in Rotterdam and I will always be grateful for it. I wish I could spend more time with that serious funny face of yours and get lost in your beard.

A big thanks to all the bisons **Alex and Tania, Alicja, the Frank, Ian my man, Enrico, Irene and family, la bella e bona Isabona, Kasia, Lena, Lennon, Pablo, Jana, Rodrigo** etc. and to the Italian crew **Enrico, Francesca, Agnese, Matteo, Livia, Simon, Giorgia, Marta, Roberta, Elena,Pazzo, Carlos and Joana, Geoffrey.** All of you made me feel so loved and exponentially boosted my well-being, thank you chicos y chicas!

Francesca grazie per storare tutti I miei beni materiali e …. per i caffe'. Sei sempre bella e carica e mi fai sentire a casa al dodicesimo piano. I know you expect me to write a long ringraziamento per you. Sorry to disappoint tu, but I tuoi rimangono longer than Francois. Statte tranquilla e daje forte. Te vulem ben.

A big hug to all my ex flatmates **Thomas, Matteo, Lina and Gizem.** Good times damn… good times! Thomas desaparecido fatti sentire cazzo che ci manchi. Matteo thank you in advance for hiring me if I can't find a job. E' stato un piacere spiegarti come fare le rovesciate. Thanks to Dr **Sebasito** for introducing me to roffa underground scene and taking care of our whatever disease we get. Without you and Lina Gina Pina I would have done 2 vaccine chip doses instead of one. Muchas gracias Pina for all the love you transmitted in these years and for ensuring that me and Thomas were going to sleep on time. Grazie to my short lasting landlord Gizem for hosting me and, just in a few days, for having recharged me of energies and self-trust.

To all the desaparecidos like my dear **Ivasito, Teresa, Roberto, Monica, Adrian, Mira.**

You are not forgotten! Your presence is still very vivid in Rotterdam for many of us.

Thanks to **Lorenzo** for not caring at all about my work and dissertation. You showed me how it is possible to be more careless than me. This made you a valuable paraparanymph. We were meant to be together from the start and for a variety of reasons we got to know each other pretty well. It's sometimes scary how many passions we share and how many things you manage to do better than me at the end. We are so similar and so different at the same time but we are both big fans of *lafiga.* Suerte with your last experiments y con la guapasita **elenita**.

Cristina you have been next to me in the most tough years of my research and we had a beautiful experience together. I hope you find the right connection with nature and the best compromise with commodity. Lemme know what you find.

Thanks to the **Almondestraat** crew who gave me a house when I was jobless and allowed me to finish my thesis with a roof over my head. You can be destroyed but not forgotten. Ciao to **Pietro, Gaia, Francisco, David, Izzy and the rest** it has been some good saunas and some good parties. 217 respect.

The odd case of discovering a passion after 4 years. **Francois,** I love you as much as a straight man can love another man. More words would only underrepresent the connection we surprisingly found my dear paraparanymph.

Grazie a tutti gli italiani d'Italia come il **Fiorellin**, **Tota and family**, lu scem di **Luzi and family**, il magico **Teo**, la stramamma **Feta e il Mauri**, la strasabrina **Sabrina**, la neomamma **Farciola**, i polish **Zu** e **Polska**, il **Cicchen**, il Garcia lorca **Paolo e Maria**. Avete provato un tutti I modi a farmi fare solo festa ma alla fine sono riuscito a districarmi pure con gli studi. Vi voglio bbe'!

Obviously grazie al faro **Anna**, la **Maiorca** e la mia cara ammmericana **Giulia**. Grazie per avermi fatto laureare, senza di voi starei ancora facendo l' esame di informatica a Tor Vergata.

Thanks to el mi amor **Josephine** who supported and encouraged me during the last years. Besides making the amazing gata2b animation and my thesis cover you assured to always make me feel loved and still teach me how to be kind. I hope you will soon appreciate how smart and interesting you are and how lucky you make me feel. We already had quite some adventures together and more will come! No matter what happens and where it will happen we can count on each other. T'amo tanto.

E chiaramente il ringraziamento **piú** grande va alla mia famiglia che ancora crede quando dico che faccio lo scienziato. Un grazie alla bella sorella **Gioel** per avermi sempre dato il buon esempio e un obiettivo da raggiungere. Lo sai solo tu quanto ti voglio bene. Mo non esprimo tutte le emozioni sennó mi richiedi la tesina per approfondire. Un abbraccio a **Kirsten** che e' tanto una brava uagliona e che trova la mia simpatia quando gli altri non la colgono.

A **michela, giuliano e andrea** che continuano ad ospitarmi a Roma facendomi sentire a casa e che sono rimasti sempre al mio fianco. Alla **nonna Adina** pe tutt lu magna'che c' ha fatto e tutto l' amore che ci hai dato e che continuiamo ad usare. M capisc sol tu.

Un grazie infinito ai miei genitori che facendomi sentire di essere fieri di me mi hanno trasmesso una enorme fiducia in me stesso e la capacitá di andare oltre molti ostacoli.

Grazie a "il padre" il **Papá** che ci ha sempre supportato negli studi e aiutato in mille altri modi. Hai sempre trovato il modo di dedicarci il tuo tempo e di giocare con noi. Il patre esemplare!

E ovviamente il grazie **piú** grande va a **MAMMA',** o core mio e la cosa piu' cara che ho. Sei sempre andata oltre I tuoi limiti per assicurarti che stessimo bene. Anche se alle volte non hai i riconoscimenti che meriti ricorda che i tuoi figli ti amano tanto.