

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

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

<sup>1</sup>Department of Hematology, Erasmus MC, Rotterdam, The Netherlands; <sup>2</sup>Cancer Genome Editing Center, Erasmus MC, Rotterdam, The Netherlands; <sup>3</sup>Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, Oxford, United Kingdom; <sup>4</sup>Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, United Kingdom; and <sup>5</sup>Department of Pathology, Cancer Treatment Screening Facility, Erasmus MC Optical Imaging Centre, and <sup>6</sup>Department of Neurology, Cancer Treatment Screening Facility, Erasmus MC, Rotterdam, The Netherlands

## Key Points

- Gata2b is required for embryonic HSPC expansion, but not HSPC generation in zebrafish.
- Gata2b plays an instructive role in the lineage output of HSPCs in zebrafish.

The differentiation of hematopoietic stem cells (HSCs) 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 HSCs. GATA2 is expressed throughout development, but because of early embryonic lethality in mice, its role during adult hematopoiesis is incompletely understood. Zebrafish contains 2 orthologs of GATA2: Gata2a and Gata2b, which are expressed in different cell types. We show that the mammalian functions of GATA2 are split between these orthologs. 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 in HSPCs. Gata2b-deficient HSPCs showed impaired progression of the myeloid transcriptional program, concomitant with increased coexpression of lymphoid genes. This resulted 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. Our study helps to deconstruct the functions of GATA2 during hematopoiesis and shows that lineage differentiation flows toward a lymphoid lineage in the absence of Gata2b.

## Introduction

Hematopoietic stem cells (HSCs) have the capacity to self-renew and generate all lineages of the hematopoietic system.<sup>1</sup> 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.<sup>2-5</sup> It is currently under debate whether HSCs can flow between myeloid and lymphoid lineage commitment or whether the HSC pool consists of separate lymphoid- and myeloid-biased HSCs.<sup>6</sup>

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

Submitted 20 July 2020; accepted 22 February 2021; published online 25 June 2021.  
DOI 10.1182/bloodadvances.2020002993.

\*E.G. and C.K. contributed equally to this work.

Sequencing data have been deposited to ArrayExpress under accession #E-MTAB-10370.

The full-text version of this article contains a data supplement.

© 2021 by The American Society of Hematology

maintenance of HSCs.<sup>7-11</sup> A role for this transcription factor in myeloid/lymphoid commitment is supported by findings of reduced and impaired granulocyte-macrophage progenitors in *Gata2*<sup>+/-</sup> mice.<sup>12-14</sup> Conversely, retroviral-mediated overexpression of *Gata2* results in enhanced self-renewal of myeloid progenitors and a block in lymphoid differentiation.<sup>15</sup> Homozygous germ line deletion of *Gata2* in mice results in embryonic lethality at E10, just before the generation of the first HSCs.<sup>16</sup>

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

Zebrafish have 2 orthologs of GATA2: *Gata2a* and *Gata2b*. Previous studies have shown that *gata2b* is prominently expressed in hematopoietic stem and progenitor cells (HSPCs), whereas *gata2a* is mainly expressed in the vasculature, including the HE regulated by the conserved +9.5 enhancer previously identified in mice.<sup>24,25</sup> Knockdown of *gata2b* severely reduces definitive hematopoiesis during embryonic stages. Lineage tracing revealed that all definitive hematopoietic cells are derived from *gata2b*-expressing cells,<sup>24</sup> indicating that *Gata2b* is the predominant GATA2 ortholog required for the maintenance of HSCs.

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 instead diverges into a lymphoid program, indicated by coexpression 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. Single guide RNAs were designed using CHOP-CHOP software and prepared according to Gagnon et al,<sup>26</sup> with minor adjustments. All animal experiments were performed according to Dutch law.

### Quantitative reverse transcription polymerase chain reaction analysis

Total RNA was isolated from 6 pooled zebrafish embryos per genotype ( $n = 6$ ) using TRIzol Reagent (Life Technologies), and complementary DNA (cDNA) was synthesized using SuperScript III Reverse Transcriptase kit (Invitrogen). *gata2a* (forward primer, 5'-CAAACCTCCACAACGTCAACAG-3'; reverse primer, 5'-CCCTCACAGATCGTTTACTC-3') and *gata2b* (forward primer, 5'-TACA-CAATGTGAATCGCCCA-3'; reverse primer, 5'-GAAGGAG-GATGGTTTGTCTG-3') expression levels were normalized to *elfa*

(forward primer, 5'-CCGCTAGCATTACCCTCC-3'; reverse primer, 5'-CTTCTCAGGCTGACTGTG-3') expression.

### ISH and analysis

At appropriate stages, 0.003% 1-phenyl-2-thiourea-treated embryos were fixed overnight with 4% paraformaldehyde in phosphate-buffered saline (PBS) containing 3% sucrose and subsequently transferred to methanol. KM smears were fixed in methanol. *In situ* hybridization (ISH) of embryos was performed as previously described.<sup>27</sup> The *cmyb* and *runx1* probes were a gift from Roger Patient and were quantified as described previously.<sup>28</sup> ISH of KM smears was performed as follows: DIG-11-UTP-labeled *s100a10b* probe was incubated overnight at 68°C; slides were blocked at room temperature in MABT (sodium chloride, maleic acid, and 1% Tween 20), 2% bovine serum albumin, and sheep serum for a minimum of 3 hours, and  $\alpha$ DIG antibody was incubated overnight at 4°C; staining was developed in tris(hydroxymethyl)aminomethane (pH 9.5), magnesium chloride, sodium chloride, and Tween 20 with 5% polyvinyl alcohol, nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate-toluidine salt at room temperature for 2 days; and cells were counterstained with Nuclear Fast Red (Sigma Aldrich) and imaged using a Leica microscope (63 $\times$  magnification).

*s100a10b* probe synthesis. *s100a10b* was amplified from cDNA of adult KM (forward primer, 5'-GAG AGC AAT GGA GAC CCT GA-3'; reverse 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 labeling kit (Sigma-Aldrich). Sense probe was used as negative control.

### Transgenic lines, confocal imaging, and adult KM fluorescence-activated cell sorting analysis

Embryos were anesthetized using 160 mg/L of tricaine (3-amino benzoic acid ethylester) and selected for reporter expression. *Tg(fli:enhanced green fluorescent protein [eGFP])*<sup>29</sup> and *Tg(CD41:GFP)*<sup>30</sup>; *Tg(flt1:RFP)*<sup>31</sup> embryos were imaged in 0.25% agarose with tricaine and imaged using a Leica SP5 confocal microscope prewarmed at 28°C. *Tg(mpeg1.1:GFP)*,<sup>32</sup> *Tg(mpx:GFP)*,<sup>33</sup> and *Tg(lck:GFP)*<sup>34</sup> embryos were placed in a 96-well plate (Zfplate; Hashimoto Electronic Industry Co. Ltd, Matsusaka, Japan) and imaged using a spinning disk confocal high-throughput microscope system (Opera Phenix; Perkin Elmer) equipped with a dry 10 $\times$  objective (NA 0.3). B-cell populations were analyzed using *Tg(IgM:GFP)*<sup>35</sup> zebrafish. Adult zebrafish were euthanized, and KM was isolated and dissociated by pipetting in PBS/10% fetal calf serum. For live/dead discrimination, 0.5 mg/L of 7-amino-actinomycin D (BD Biosciences) or 1 mg/L of 4',6-diamidino-2-phenylindole was used. For the embryonic proliferation assay, 25 embryos per genotype were pooled in prewarmed PBS/10% fetal calf serum, and single-cell suspension was prepared by adding 1% from each collagenase (I, II, and IV; Sigma-Aldrich) and incubating for 45 minutes at 37°C. Proliferation was assessed after 4% paraformaldehyde fixation and  $\alpha$ -Ki67 staining for both embryonic and adult stages. The analysis was performed using FACSAriaIII (BD Biosciences).

### Single-cell RNA sequencing

A total of 70 000 single viable cells were sorted from 2 pooled KMs of female *Tg(CD41:GFP)* zebrafish and supplemented with 114 to

1607 CD41:GFP<sup>low</sup>-expressing cells. cDNA was prepared using the manufacturer's protocol (Chromium [version 2]; 10× Genomics) and sequenced on a Novaseq 6000 instrument (Illumina). Two wild-type (WT) replicates and 2 *gata2b*<sup>-/-</sup> replicates were sequenced with the following read depth: WT1, 52 384 reads per cell; WT2, 43 876 reads per cell; *gata2b* knockout 1, 53 836 reads per cell; and *gata2b* knockout 2, 46 761 reads per cell. Data were analyzed using the Seurat R package,<sup>36</sup> and a detailed description is provided in the data supplement.

## Statistics

All statistical analyses were carried out in GraphPad Prism 5 (GraphPad Software). Normally distributed data were analyzed using 1-way analysis of variance with Tukey multiple comparison test when comparing 3 sample sets or Student *t* test when comparing 2 sample sets. Data with nonnormal distribution were analyzed using a nonparametric Kruskal-Wallis with Dunn correction test.

## Results

### Generation of 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). Quantitative reverse transcription polymerase chain reaction analysis of *gata2b* on pooled WT and *gata2b*<sup>-/-</sup> embryos at 30 hpf indicated that *gata2b* expression level, a known transcriptional target of Gata2, was significantly reduced in mutant embryos (supplemental Figure 1A-B).<sup>25,37</sup> Hereafter, we refer to this mutant as *gata2b*<sup>-/-</sup>.

### Gata2b is dispensable in the generation of HSCs from HE

The first HSCs transdifferentiate from specialized HE cells in the aorta-gonad-mesonephros region, through a highly conserved process known as EHT.<sup>18,19,38-40</sup> In mice, *Gata2* is expressed in the endothelium, including the HE of the dorsal aorta,<sup>8</sup> and deletion of *Gata2* results in a reduction in HSC generation.<sup>7,8</sup> *cmyb* and *runx1* are 2 bona fide marker genes for HE at 26 hpf in zebrafish.<sup>18,41</sup> We quantified *cmyb* and *runx1* expression by measuring pixel intensity of the ISH staining compared with background.<sup>28</sup> 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 HE 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.<sup>29</sup> Consistent with our initial results, EHT events were not significantly reduced in *gata2b*<sup>-/-</sup> embryos compared with WT (*P* = .077; WT embryos, *n* = 18; *gata2b*<sup>-/-</sup> embryos, *n* = 18; Figure 1I-J; supplemental Table 1). We conclude that neither HE specification nor HSPC generation through EHT is impaired in *gata2b*<sup>-/-</sup> embryos.

### Gata2a is required for HE specification

GATA2 is required for the generation of HSCs in mice,<sup>7,16</sup> but *Gata2b*-deficient zebrafish have intact HE and EHT (Figure 1E-J). High maternal expression of *gata2b* has been reported previously,<sup>24</sup> and therefore, residual *Gata2b* protein levels could possibly rescue

EHT. However, maternal zygotic *gata2b*<sup>-/-</sup> zebrafish that do not contain functional maternally provided *gata2b* messenger RNA are viable and survive to Mendelian ratios (supplemental Figure 1C-D), indicating that maternal expression of *gata2b* does not contribute to embryonic hematopoiesis. By contrast, *gata2a* is expressed in HE and regulates *runx1* expression in HE.<sup>25</sup> Therefore, we analyzed *runx1* expression in *gata2a* mutants (*gata2a*<sup>i4/i4</sup>, lacking a conserved endothelial enhancer).<sup>25</sup> *runx1* expression at 26 hpf was reduced in *gata2a*<sup>i4/i4</sup> embryos compared with WT embryos (Figure 1K-L). This confirmed that endothelial expression of *gata2a*, but not *gata2b*, is required for specification of HE and that the different functions of mammalian GATA2 are separated between *Gata2a* and *Gata2b* in zebrafish.

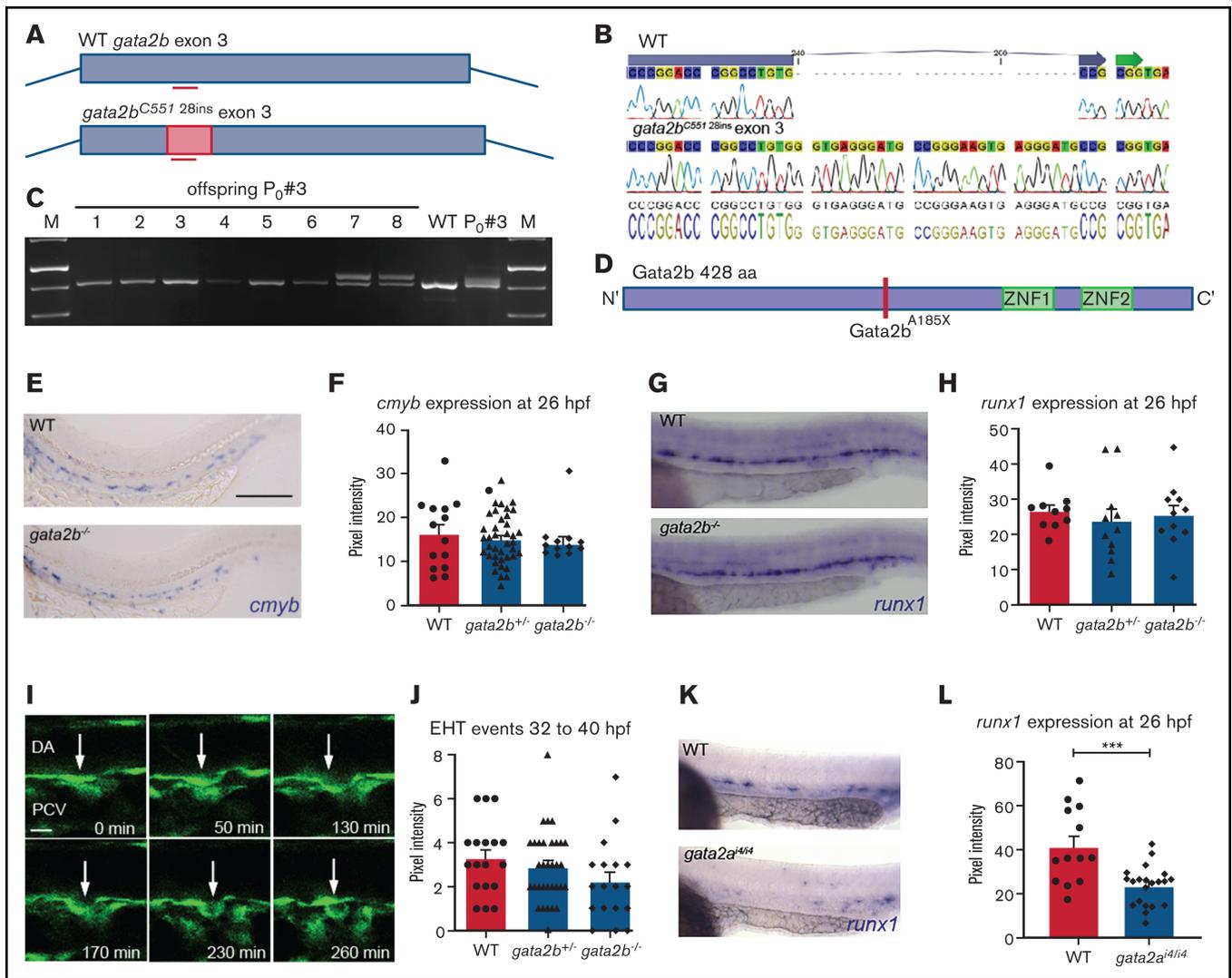
### Gata2b is required for the expansion of definitive HSPCs during CHT amplification phase

CHT is temporally and spatially analogous to mouse fetal liver where HSPCs undergo amplification.<sup>22</sup> We investigated whether the loss of *Gata2b* affects the number of definitive HSPCs in CHT between 2 and 3 days postfertilization (dpf). From 44 hpf onward, definitive HSPCs are marked by coexpression of the *Tg(CD41:GFP)*<sup>30</sup> marker and the arterial *Tg(fli:RFP)*<sup>31</sup> marker, because definitive HSPCs are derived from arteries.<sup>42</sup> CD41:GFP<sup>+</sup>Fli:RFP<sup>+</sup> cell numbers were similar in WT and *gata2b*<sup>-/-</sup> embryos at 52 to 54 hpf (Figure 2A-B; supplemental Table 1) and 56 to 58 hpf (Figure 2C; supplemental Table 1). However, at 76 hpf, CD41:GFP<sup>+</sup>Fli:RFP<sup>+</sup> cells were significantly reduced in CHT in *gata2b*<sup>-/-</sup> embryos compared with WT (Figure 2D; supplemental Table 1).

The number of definitive HSPCs expanded rapidly from 52 to 76 hpf in WT embryos (6.5-fold); in *gata2b*<sup>-/-</sup> 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.<sup>43,44</sup> A significant reduction in *cmyb* expression was detectable from 33 hpf onward in the aorta-gonad-mesonephros and CHT regions of *gata2b*<sup>-/-</sup> embryos compared with WT (Figure 2E-H; supplemental Table 1). This analysis detected a reduction in *cmyb* expression levels rather than quantifying HSPC numbers. However, because the number of CD41:GFP<sup>+</sup>Fli:RFP<sup>+</sup> 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<sup>+</sup> cells at 75 hpf in WT and *gata2b*<sup>-/-</sup> embryos. This analysis showed that *Gata2b*-deficient CD41:GFP<sup>+</sup> cells had an increased proportion of cells in the G<sub>0</sub> phase of cell cycle, explaining the reduction of HSPCs at 3 dpf (Figure 2I-K). At 5 dpf, the difference in proliferation was no longer detectable, although HSPC numbers were still reduced (Figure 2I; data not shown).

### Single-cell RNA sequencing identifies lymphoid bias at the expense of myeloid lineage output in *gata2b*<sup>-/-</sup> KM

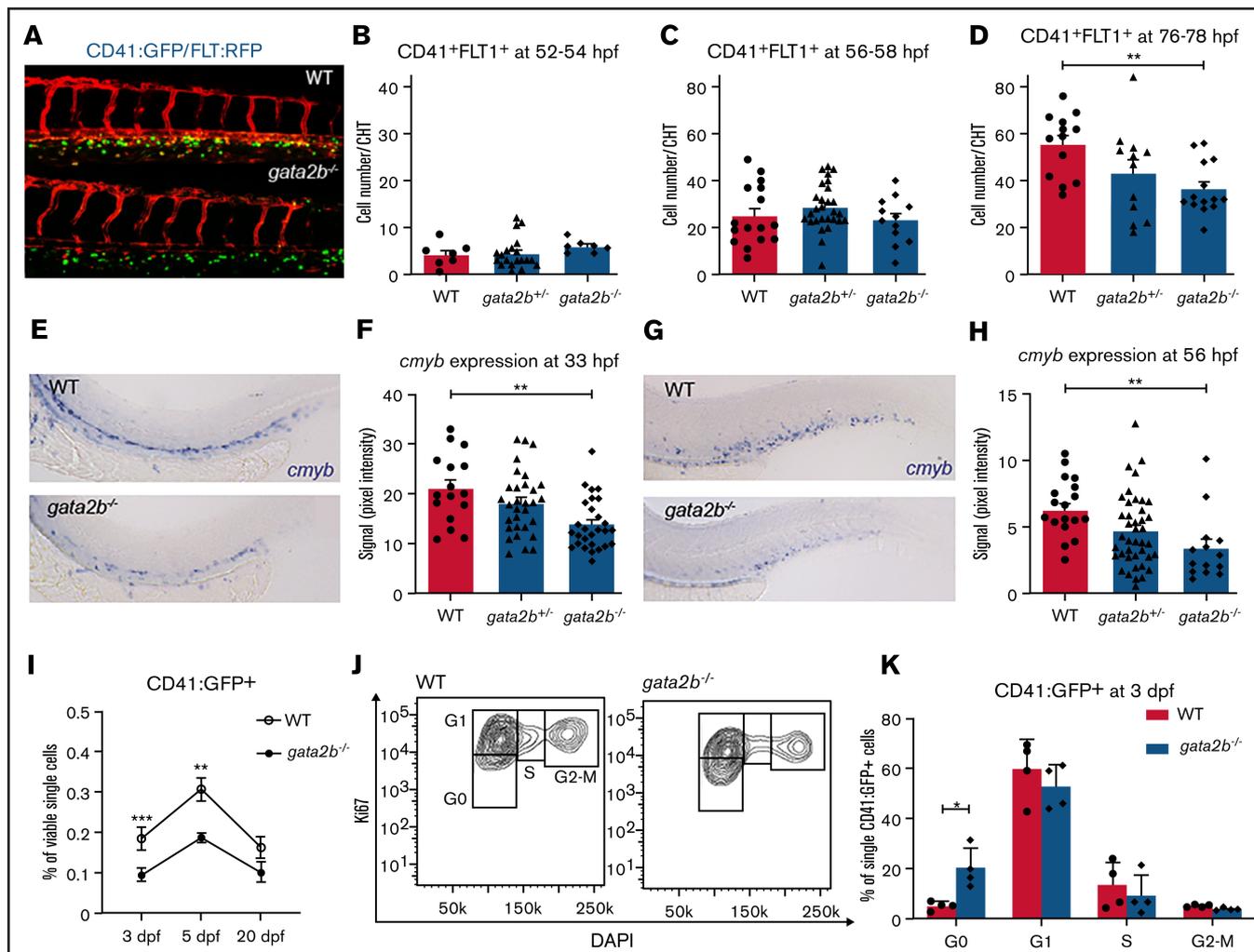
Because the functions of GATA2 are separated between *Gata2a* and *Gata2b* in zebrafish, and *Gata2b*-deficient zebrafish are viable, we can uniquely assess the function of *Gata2b* in adult hematopoiesis. To investigate the hematopoietic lineages in an unbiased manner and assess the impact of *Gata2b* deficiency on the transcriptional profile of hematopoietic progenitors and differentiated cells, the progenitor population, including lymphocytes from KM,



**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 by the blue arrow at the top of the sequence, and sequencing data from *gata2b*<sup>-/-</sup> DNA showing a 28-nucleotide integration. (C) Gel picture showing genotyping polymerase chain reaction of founder 3 and the F1 with a 28-bp integration in embryos 7 and 8. (D) *Gata2b* mutation leading to a STOP codon abrogating the protein before the DNA and protein binding zinc fingers. (E) Representative image of ISH for *cmyb* expression in WT and *gata2b*<sup>-/-</sup> embryos at 26 hours postfertilization (hpf). Bar indicates 100  $\mu$ m (F) Quantitation of *cmyb* signal intensity relative to background in WT and *gata2b*<sup>-/-</sup> embryos at 26 hpf. (G) Representative image of *runx1* expression in WT and *gata2b*<sup>-/-</sup> embryos at 26 hpf. (H) Quantitation of *runx1* signal intensity relative to background in WT and *gata2b*<sup>-/-</sup> embryos at 26 hpf. (I) Example of endothelial-to-hematopoietic transition (EHT) event from WT *Tg(fli1a:eGFP)* transgenic zebrafish. Time indicated at the bottom right corner in minutes. Scale bar represents 10  $\mu$ m. Arrow indicates endothelial cell undergoing hematopoietic transition. (J) Quantitation of EHT events between 32 and 40 hpf in WT, *gata2b*<sup>+/-</sup>, and *gata2b*<sup>-/-</sup> embryos. (K) Representative example of *runx1* expression in WT and *gata2a*<sup>4/4</sup> embryos at 26 hpf in the aorta-gonad-mesonephros region. (L) Quantitation of signal intensity relative to background cells in WT and *gata2a*<sup>4/4</sup> embryos at 26 hpf, where each dot represents 1 embryo ( $41.4 \pm 4.8$  and  $23.5 \pm 2.0$ ; WT, n = 13; *gata2a*<sup>4/4</sup>, n = 21). Error bars represent standard error of the mean. \*\*\**P* < .001. DA, dorsal aorta; PVC, posterior cardinal vein; sp, somite pair.

were isolated and processed for single-cell RNA sequencing (supplemental Figure 2A-B). To enrich the scarce HSC population, we used pooled KM from 2 WT and *gata2b*<sup>-/-</sup> *Tg(CD41:GFP)* zebrafish per sample and included all CD41:GFP<sup>low</sup>-expressing cells present in the KM pool, because these cells have been shown to contain transplantable HSCs<sup>30</sup> (supplemental Figure 2A,C). This resulted in a mild enrichment of phenotypic HSCs, from 0.21% to 0.5% to 0.46% to 2.73% of CD41:GFP<sup>low</sup> cells within the total

progenitor population. We identified 20 different cell clusters using the nearest neighbor algorithm in the R Seurat package<sup>36</sup> (Figure 3A; supplemental Figure 2G; supplemental data for 3D interactive graph). Most progenitors that were sequenced expressed previously characterized differentiation markers<sup>45-49</sup> (Figure 3B-E; supplemental Figure 2H). We identified 2 HSPC populations. These clusters were characterized by the robust expression of HSC genes, like *fli1a* and *meis1b*<sup>48-50</sup> (Figure 3F-G) and *gata2b* (Figure 3H),

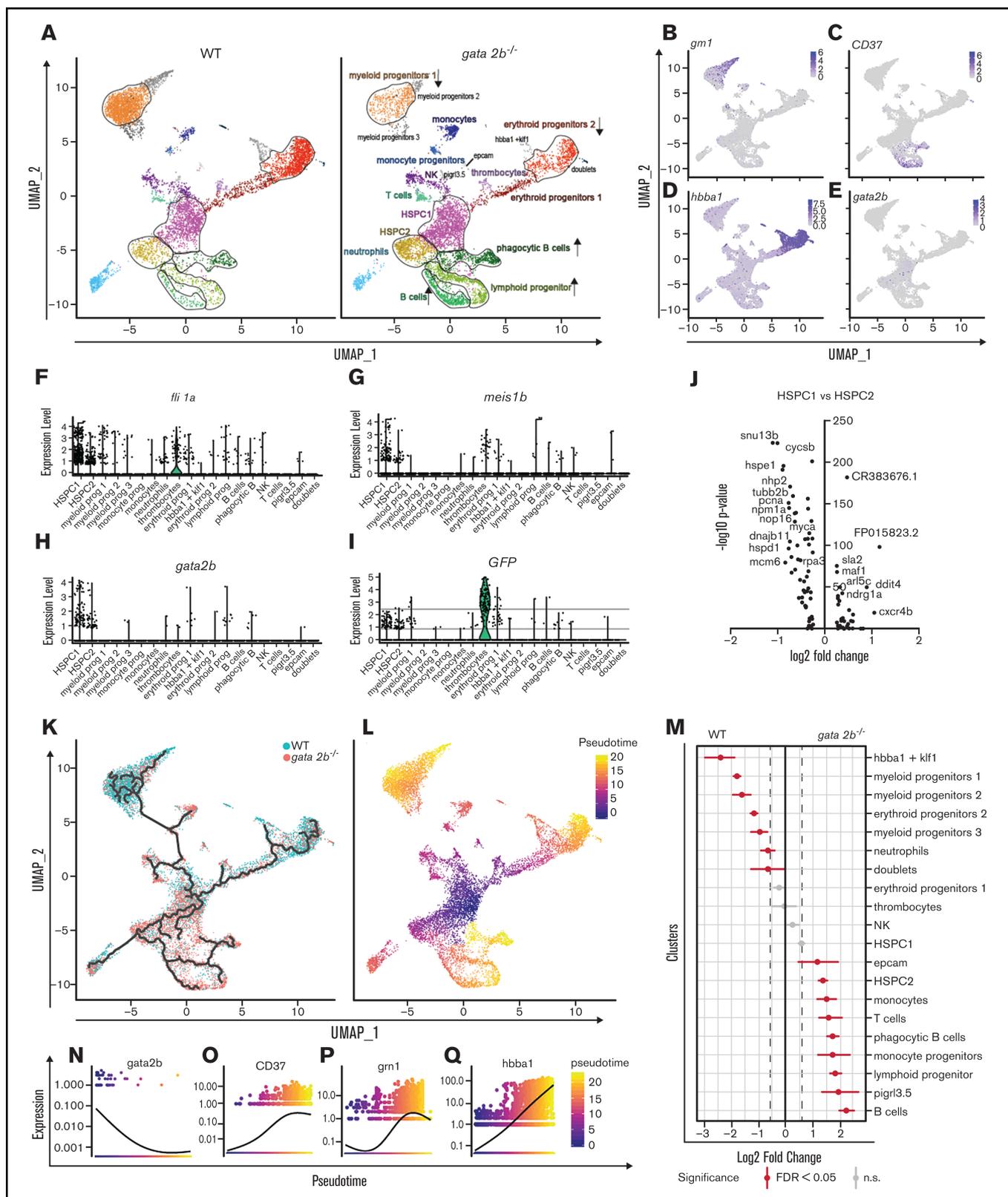


**Figure 2. HSPC numbers are reduced in CHT after 3 dpf in *gata2b*<sup>-/-</sup> embryos.** (A) Example of *Tg*(CD41:GFP); *Tg*(*Flt*:RFP) expression in CHT of WT and *gata2b*<sup>-/-</sup> embryos at 76 hpf. (B-D) Quantitation of GFP<sup>+</sup>RFP<sup>+</sup> cells in WT, *gata2b*<sup>+/-</sup>, and *gata2b*<sup>-/-</sup> embryos at different time points, where each dot represents 1 embryo: 52 to 54 hpf (8.50 ± 1.7 vs 11.86 ± 1.0) (B), 56 to 58 hpf (25.19 ± 3.2 vs 23.25 ± 2.9) (C), and 76 to 78 hpf (55.46 ± 3.8 vs 36.50 ± 3.0; *P* < .001) (D). Representative example of *cmyb* expression in WT and *gata2b*<sup>-/-</sup> embryos at 33 hpf (E), and quantitation of signal intensity relative to background at 33 hpf (F), where each dot represents 1 embryo. Representative example of *cmyb* expression in WT and *gata2b*<sup>-/-</sup> embryos at 56 hpf (G), and quantitation of signal intensity relative to background at 56 hpf (H), where each dot represents 1 embryo. (I) Quantitation of CD41:GFP<sup>+</sup> cell percentages by flow cytometry at 3, 5, and 20 dpf. (J) Cell-cycle analysis by flow cytometry of Ki67 and 4',6'-diamidino-2-phenylindole (DAPI) staining of CD41:GFP<sup>+</sup> cells in WT and *gata2b*<sup>-/-</sup> embryos at 75 hpf. (K) Bar graph representing the quantitation of cell cycle of CD41:GFP<sup>+</sup> cells in WT and *gata2b*<sup>-/-</sup> embryos at 75 hpf. Bars represent mean ± standard error of the mean, with each dot indicating 1 pooled sample. \**P* < .05, \*\**P* < .01.

concomitant with intermediate levels of GFP derived from the CD41:GFP transgene (Figure 3I) and low expression of differentiation markers (Figure 3B-E; supplemental Figure 2H). Compared with HSPC2s, HSPC1s exhibited a lower expression of metabolic and proliferation markers like *pcna* and *myca*, suggesting that HSPC1s are more quiescent than HSPC2s (Figure 3J). Therefore, lineage trajectory analysis was started from this cluster, identifying separate lineage differentiation trajectories for the erythroid, myeloid, and lymphoid lineages (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 toward the lymphoid lineage in *gata2b*<sup>-/-</sup> cells at the expense of the myeloid lineage compared with WT (Figure 3A,K-M;

supplemental Figure 2I; supplemental Table 2). The largest differences were observed in 3 clusters expressing high levels of *gm1* (Figure 3B), indicating that these clusters contained myeloid progenitors and were overrepresented by WT cells (Figure 3A,K,M; supplemental Figure 2I). We defined these 3 clusters expressing myeloid-specific genes with slight differences in their expression patterns as myeloid progenitors 1, 2, and 3 (Figure 3A; supplemental Figure 2H). The *gm1*-expressing cluster also showed high expression of *s100a10b*, a potential new marker for these cells (supplemental Figure 3A). Expression analysis of KM smears showed that *s100a10b* was expressed in the neutrophil lineage (supplemental Figure 3B). The B-cell clusters were overrepresented by *gata2b*<sup>-/-</sup> cells and showed very high expression of immunoglobulin heavy variable 1-4 (*ighv1-4*; supplemental Figure 2H),



**Figure 3. Single-cell analysis reveals that *gata2b*<sup>-/-</sup> KM cells are overrepresented in lymphoid lineage clusters and reduced in erythroid and myeloid lineage clusters compared with WT.** (A) Split uniform manifold approximation projection (UMAP) of WT and *gata2b*<sup>-/-</sup> cells with cluster indication of enriched (arrow up) or reduced (arrow down) cell clusters in *gata2b*<sup>-/-</sup> cells. (B-E) Pooled WT and *gata2b*<sup>-/-</sup> UMAP feature analysis with gradual gene expression in shades of blue. Expression pattern of granululin (*gm1*) (B), cluster of differentiation 37 (*CD37*) (C), hemoglobin  $\beta$  adult (*hbba1*) (D), and *gata2b* (E). (F-I) Violin plots representing the expression levels of genes within the different clusters; each dot represents expression in 1 cell. *Fli-1* proto-oncogene (*fli1a*) (F), meis homeobox 1b (*meis1b*) (G), *gata2b* (H), and *GFP* (I),

*CD37* (Figure 3C), and *pax5* (data not shown), indicating that these were bona fide B-cell populations. Interestingly, we found a population of phagocytic B cells previously identified in teleosts, which express both *mpeg1.1* and B-cell markers.<sup>51</sup> Lineage trajectory analysis indicated that these cells descended from lymphoid progenitors (Figure 3C,K,L). When pseudotime analyses was performed for WT and *gata2b*<sup>-/-</sup> cells separately, the phagocytic B cells showed a lineage differentiation trajectory not only from lymphoid progenitors, but also from the HSPC1 population, indicating a skewing in *gata2b*<sup>-/-</sup> HSPCs directly toward the lymphoid lineage (supplemental Figure 2J-K).

### Lack of Gata2b leads to reduced neutrophil numbers and increased lymphoid progenitors in adult KM

Because single-cell RNA sequencing analysis showed a major switch in lineage differentiation, we asked whether hematopoietic differentiation was affected in the adult *gata2b*<sup>-/-</sup> KM using scatter profile, transgenic marker, and morphological analyses.<sup>30,32,33,52</sup> Although *gata2b*<sup>-/-</sup> embryos did not show signs of altered lineage differentiation up to 5 dpf (supplemental Figure 4A-F), scatter profiles of adult *gata2b*<sup>-/-</sup> zebrafish KM showed a significant reduction in the myeloid population (Figure 4A-B; Table 1) and a relative increase in the scatter population containing HSPCs and lymphoid cells at 4 months postfertilization and onward (Figure 4A,C; 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,<sup>33,53</sup> and *Tg(mpeg1.1:GFP)* expression, marking monocytes and phagocytic B cells,<sup>32,54</sup> were assessed. No significant difference was observed in *mpeg:GFP*<sup>+</sup> cells between WT and *gata2b*<sup>-/-</sup> KM (supplemental Figure 4O-Q). *gata2b*<sup>-/-</sup> zebrafish showed a severe reduction in *mpx:GFP*<sup>+</sup> neutrophils in the KM at 4 months postfertilization (Figure 4D-F; Table 1). Sorted *mpx:GFP*<sup>+</sup> cells from these zebrafish showed that the remaining *gata2b*<sup>-/-</sup> *mpx:GFP*<sup>+</sup> cells did not reach WT levels of GFP and had a more immature neutrophil morphology and a block at the promyelocyte stage (Figure 4G-H; supplemental Figure 3E), 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,M).

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

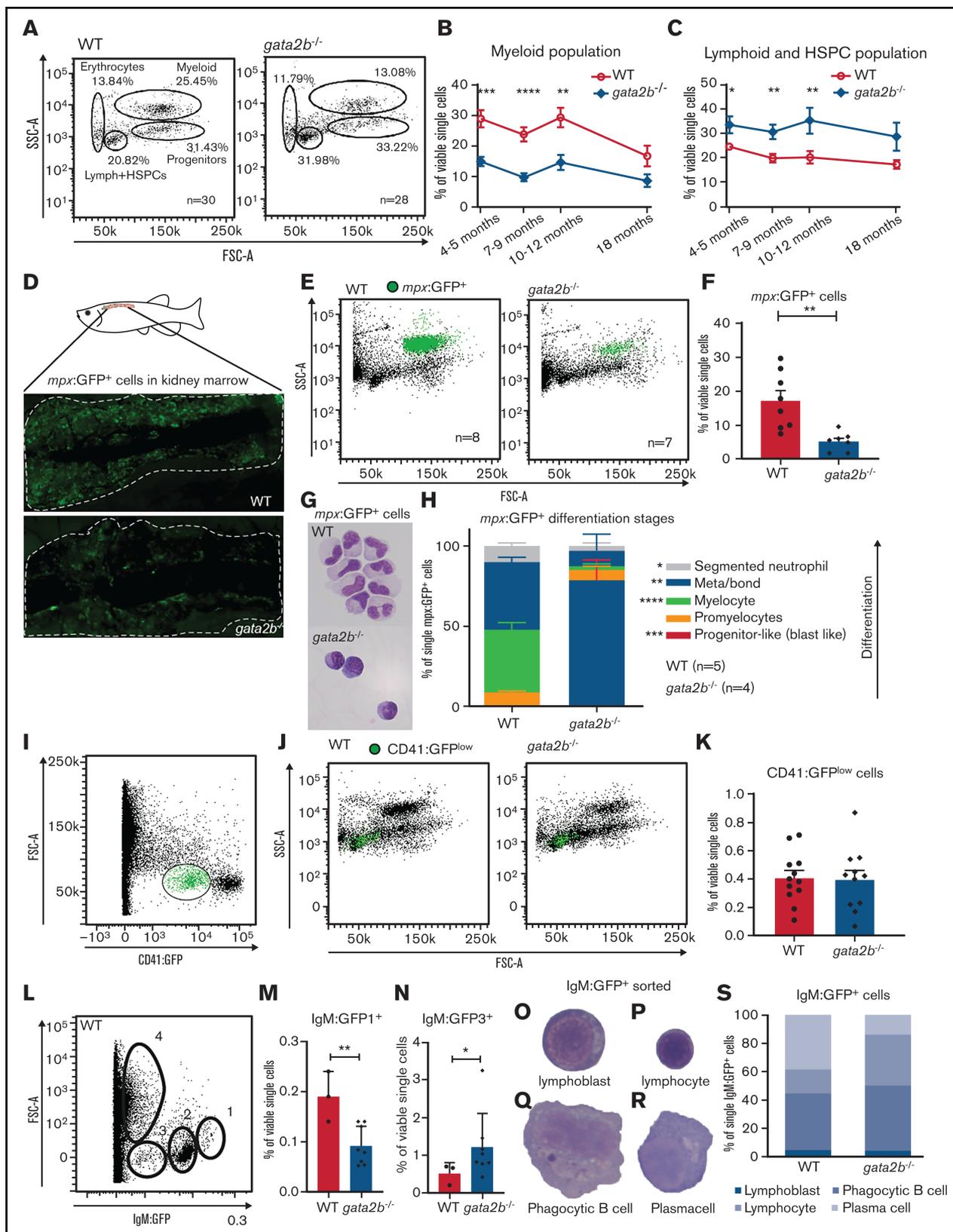
*Tg(IgM:GFP)*, marking B cells,<sup>35</sup> and *Tg(Ick:GFP)*, marking T cells,<sup>34</sup> were used to assess lymphoid differentiation. We did not find an increase in the *Ick:GFP*<sup>+</sup> population (supplemental Figure 4M-N; Table 1). However, in *Tg(IgM:GFP)* zebrafish, we identified several populations of *IgM:GFP*<sup>+</sup> cells, with a significant increase in immature *IgM:GFP*<sup>+</sup> cells (*IgM:GFP3* fraction; Figure 4L-N; supplemental Figure 4G-M). We could classify the different *IgM:GFP*<sup>+</sup> populations as lymphoblastic cells, lymphocytes, plasma cells, and phagocytic B cells<sup>51,55</sup> (Figure 4O-S). In particular, phagocytic B cells were increased in *gata2b*<sup>-/-</sup> KM compared with WT, but mature plasma cells were significantly reduced (Figure 4S; *P* < .01 and *P* < .001, respectively). Although a majority of cells in the lymphoid and HSPC populations were not marked by known lymphoid lineage markers *IgM:GFP* or *Ick:GFP*, we could still detect a significant increase in immature B cells, confirming the increase in lymphoid output in KM in *gata2b*<sup>-/-</sup> zebrafish compared with 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 have shown that blocked neutrophil differentiation results in a shift toward monocytic lineage differentiation.<sup>56</sup> Our data indeed show an increase in monocyte progenitors and monocytes (Figure 3M). However, we also detected a shift toward 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 had identified as HSPC1s and HSPC2s (Figure 3A).

HSPC1s made up 18% of the total analyzed KM population. Because HSCs are a rare population of cells based on transplantation studies, we hypothesized that HSPC1s also contain other progenitor cells. We subclustered HSPC1s to subdivide these progenitor cells. In this way, clustering was not based on gene expression differences found in comparison with more committed cells, but only based on gene expression differences within the HSPC1 population. Eight subclusters were identified based on differential gene expression analysis (supplemental Figure 5A). These subclusters were classified as a quiescent subcluster with very low gene expression, an HSC subcluster with expression of *meis1b* and *fli1a*, 3 myeloid subclusters, 1 lymphoid subcluster, a proliferative subcluster, and an undefined subcluster (Figure 5B,E,F; supplemental Figure 5A). Interestingly, proportion analysis of WT and *gata2b*<sup>-/-</sup> 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 *cepa* in *gata2b*<sup>-/-</sup> HSPC1s (Figure 5G), but not a clear upregulation of lymphoid genes. When the same comparison was performed in HSPC2s (supplemental Figure 6A-B), we detected a larger reduction in the myeloid gene expression program in the entire HSPC2 cluster

**Figure 3. (Continued).** indicating CD41:GFP<sup>low</sup> cells. (J) Volcano plot comparing HSPC1s vs HSPC2s. At the left of the y-axis are genes in HSPC1s with an average logarithmic fold change less than -0.25, and to the right are genes with a logarithmic fold change >0.25 compared with HSPC2s. (K) Lineage differentiation trajectory depicted on UMAP with WT cells in blue and *gata2b*<sup>-/-</sup> cells in pink. (L) Pseudotime analysis assuming HSPC1s as a starting point. (M) Quantitation of proportions of distribution between WT and *gata2b*<sup>-/-</sup> cells in the different clusters. Significant differences are indicated in red. (N-Q) Pseudotime analysis of gene expression in lineage trajectory analysis of *gata2b* (N), *CD37* (O), *grn1* (P), and hemoglobin  $\beta$  adult (*hbba1*) (Q). FDR, false discovery rate; n.s., not significant.



**Figure 4.** *Gata2b* deficiency results in decreased myeloid differentiation in adult zebrafish KM. (A) Gating strategy of fluorescence-activated cell sorting analysis of whole KM of WT and *gata2b*<sup>-/-</sup> zebrafish. Percentages represent the average of all zebrafish analyzed per genotype. (B-C) Quantitation as percentages of single viable cells over time of myeloid (B) and lymphoid (C) and HSPC populations. (D) Representative example of *Tg(mpx:GFP)* expression in WT and *gata2b*<sup>-/-</sup> zebrafish KM by fluorescence microscopy. (E) Forward (FSC) and side scatter (SSC) profiles of *Tg(mpx:GFP)* expression in WT and *gata2b*<sup>-/-</sup> zebrafish KM in green. (F) Quantitation of

**Table 1. Adult hematopoietic cell quantitations**

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

Data are given as mean ± standard error of the mean. n indicates number of zebrafish used in analysis.

\*\**P* < .01, \*\*\*\**P* < .0001. If data are normally distributed, we used 1-way analysis of variance with Tukey posttest. If data are not normally distributed, we used Kruskal-Wallis with Dunn posttest.

(supplemental Figure 6C-D) and found that *gata2b*<sup>-/-</sup> cells expressed higher levels of lymphoid genes like IKAROS family zinc finger 1 (*ikzf1*), *tcer1g1*, *ighv1-4*, *ccr9a*, and *xbp1* (supplemental Figure 6E-J). Pseudotime analysis showed the differentiation trajectories within the HSPC1 cluster when started from the quiescent subcluster, containing most CD41:GFP-expressing cells (Figure 3E-F,H-I). When inferring gene expression in pseudotime analysis of HSPC1s, we found that *gata2b* and *meis1* expression were highest in the quiescent population and decreased during differentiation (Figure 5J-L). Interestingly, *gata2a* expression did 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*<sup>-/-</sup> unique subcluster (Figure 5N-O). Assessing proliferation by flow cytometry of CD41:GFP<sup>low</sup> cells in WT and *gata2b*<sup>-/-</sup> KM, we found that *gata2b*<sup>-/-</sup> CD41:GFP<sup>low</sup> cells had increased numbers of cells in S phase, indicating that *gata2b*<sup>-/-</sup> 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 toward the lymphoid lineage.

### Differential gene expression analysis reveal decreased myeloid marker expression in *gata2b*<sup>-/-</sup> HSPCs and aberrant coexpression of myeloid and lymphoid genes

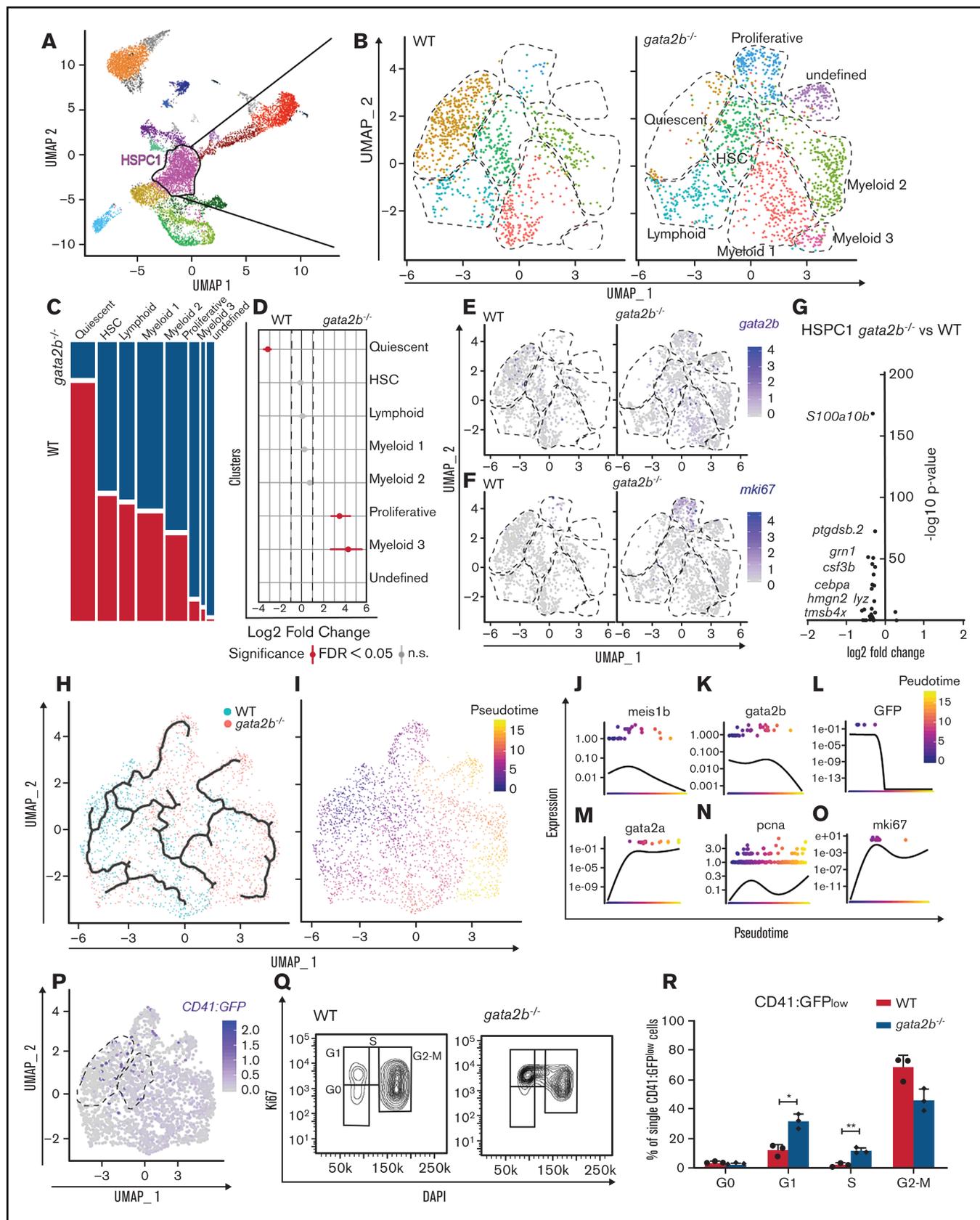
Overall, the expression of myeloid genes in *gata2b*<sup>-/-</sup> HSPC1s was reduced, but the percentage of *gata2b*<sup>-/-</sup> HSPC1s with detectable expression of myeloid genes such as *grn1* was increased

(Figures 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 coexpress myeloid and lymphoid genes before lineage decision.<sup>57,58</sup> Although WT cells had a clear dichotomy in expression of myeloid and lymphoid genes, *gata2b*<sup>-/-</sup> HSPCs had a higher fraction of cells coexpressing lymphoid and myeloid genes (Figure 6B-E). For example, increased coexpression of a phagocytic B-cell marker, *igic1s1*, could be detected in *gata2b*<sup>-/-</sup> HSPC1s together with the myeloid marker *cebpb* (Figure 6B). This result suggests that the loss of *Gata2b* does not halt HSPC differentiation, but instead redirects this toward another lineage. Interestingly, when we inferred pseudotime analysis of only WT and only *gata2b*<sup>-/-</sup> cells, this was exactly what we found. In *gata2b*<sup>-/-</sup> cells, phagocytic B cells could be formed from both lymphoid progenitors, as well as HSPC1s, as opposed to WT phagocytic B cells (supplemental Figure 2J-K). On the basis of these data, we conclude that the lymphoid bias in *gata2b*<sup>-/-</sup> zebrafish KM initiates 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, which redirects HSPCs toward a lymphoid fate.

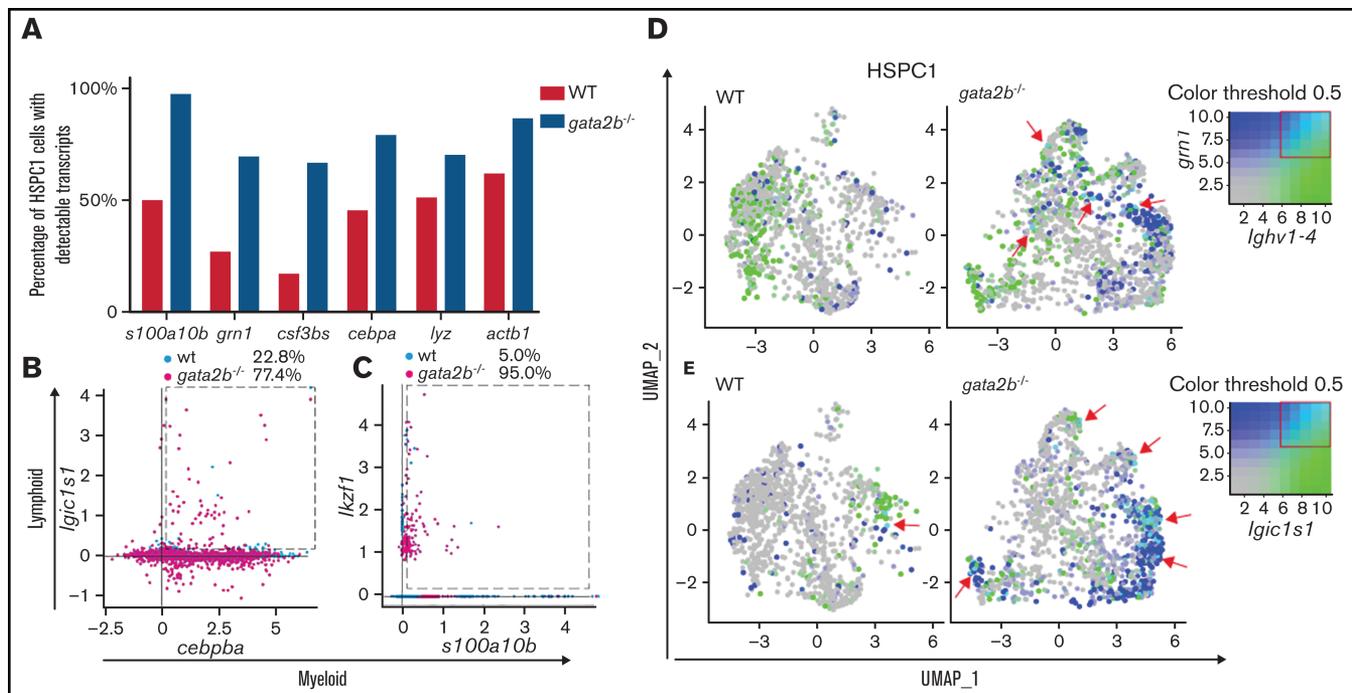
### Discussion

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 CHT. However, during adulthood, *Gata2b* is required for the quiescent HSPC population, and in its absence, HSPCs are more proliferative. In addition, *Gata2b*-deficient KM from adult zebrafish showed a

**Figure 4. (Continued).** *Tg(mpx:GFP)*<sup>+</sup> cells expressed as percentage in single viable cells. Each dot represents KM analysis of 1 zebrafish. (G) Representative figure of sorted *Tg(mpx:GFP)*<sup>+</sup> cells from WT and *gata2b*<sup>-/-</sup> zebrafish KM after May Grunwald-Giemsa (MGG) staining. (H) Quantification of *Tg(mpx:GFP)*<sup>+</sup> cells from WT and *gata2b*<sup>-/-</sup> zebrafish KM based on the differentiation phenotype using MGG staining. (I) Gating strategy for CD41:GFP<sup>low</sup>-expressing cells in total KM in green. (J) FSC and SSC plots of WT and *gata2b*<sup>-/-</sup> KM cells and CD41:GFP<sup>low</sup>-expressing cells in green. (K) Quantification of the frequency of CD41:GFP<sup>low</sup> cells in single live cells of total KM. Each dot represents KM analysis of 1 zebrafish. (L) FSC/GFP scatter profile of *Tg(IgM:GFP)* WT KM. (M) Quantitation of gating 1 of *Tg(IgM:GFP)* WT and *gata2b*<sup>-/-</sup> KM as percentage of single viable cells. Each dot represents KM analysis of 1 zebrafish. (N) Quantitation of gating 3 of *Tg(IgM:GFP)* WT and *gata2b*<sup>-/-</sup> KM as percentage of single viable cells. Each dot represents KM analysis of 1 zebrafish. (O-R) Representative image of sorted IgM:GFP<sup>+</sup> cells indicating lymphoplasmic cell (O), lymphocyte (P), phagocytic B cell (Q), and plasma cell (R). (S) Quantitation of sorted IgM:GFP<sup>+</sup> cells per genotype. Error bars represent standard error of the mean. \**P* < .05, \*\**P* < .01, \*\*\**P* < .001, \*\*\*\**P* < .0001.



**Figure 5. Subclustering of HSPC1s demonstrates the loss of *gata2b*-expressing quiescent subcluster and appearance of proliferative subcluster.** (A) Cluster selection for subclustering. (B) Reclustering of the HSPC1 population split between WT and *gata2b*<sup>-/-</sup> cells. (C) Genotype distribution of each of the clusters, with WT cells in white and *gata2b*<sup>-/-</sup> cells in black. (D) Quantitation of proportions of distribution between WT and *gata2b*<sup>-/-</sup> cells in the different clusters. Significant



**Figure 6. Gata2b-deficient HSPC1s show coexpression of myeloid and lymphoid gene expression programs.** (A) Bar graph representing the percentage of WT HSPC1s in white and *gata2b*<sup>-/-</sup> HSPC1s in gray, with at least 1 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. Coexpression analysis of the lymphoid gene *igic1s1* with the myeloid gene *cebpb* (B) and the lymphoid gene *ikzf1* (C) with the myeloid gene *s100a10b*. Values represent percentages of WT and *gata2b*<sup>-/-</sup> HSPC1s coexpressing myeloid and lymphoid genes (within the dashed box). WT and *gata2b*<sup>-/-</sup> feature analysis representing coexpression analysis of the lymphoid gene *ighv1-4*, with the myeloid gene *grn1* (D) and the phagocytic B-cell marker immunoglobulin light  $\kappa$  constant 1-s1 (*igic1s1*) with the myeloid gene *s100a10b* (E), with myeloid genes in blue and lymphoid genes in green. Coexpression of the myeloid and lymphoid genes is represented in turquoise and indicated by red arrows in the WT and *gata2b*<sup>-/-</sup> feature analysis. Coloring threshold set in quantiles, with minimum cutoff of q10 and maximum cutoff of q90.

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*<sup>-/-</sup> KM cells as a result of a failure to increase myeloid gene expression to sufficient levels and a subsequent coexpression of both myeloid and lymphoid genes in *gata2b*<sup>-/-</sup> 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.<sup>59</sup> However, later reports have suggested that some transcription factors have a reinforcing activity for terminal differentiation and have proposed that microenvironmental or upstream regulators are decisive in lineage commitment.<sup>60</sup> 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 CD41:GFP<sup>low</sup> population showed increased proliferation. Loss of quiescence in HSPCs then increased the expression of commitment genes, resulting in cells coexpressing lymphoid and myeloid lineage markers as detected in *gata2b*<sup>-/-</sup> HSPCs, and Gata2b is therefore an essential cell-intrinsic regulator of lineage output in HSPCs.

In mice, GATA2 is also required for the maintenance of HSCs after they are generated.<sup>7</sup> During embryonic hematopoiesis, the number and percentage of HSPCs are reduced because of reduced proliferation, but during adult stages, Gata2b-deficient HSCs as marked by CD41:GFP<sup>low</sup> expression are not reduced, and proliferation is

**Figure 5. (Continued).** differences are indicated in red. (E-F) WT and *gata2b*<sup>-/-</sup> feature analysis with gradual gene expression in shades of blue within HSPC1s of *gata2b* (E) and *mki67* (F). (G) Volcano plot comparing HSPC1 *gata2b*<sup>-/-</sup> vs WT. At the left of the y-axis is gene expression in *gata2b*<sup>-/-</sup> HSPC1s with an average logarithmic fold change less than  $-0.25$ , and to the right is gene expression with a logarithmic fold change  $>0.25$  compared with WT HSPC1s. Each dot represents a gene. (H) Lineage differentiation trajectory depicted on uniform manifold approximation projection, with WT cells in blue and *gata2b*<sup>-/-</sup> cells in pink. (I) Pseudotime analysis assuming the quiescent population as starting point. (J-O) Gene expression analysis on pseudotime analysis with *meis1b* (J), *gata2b* (K), *GFP* (L), *gata2a* (M), *pcna* (N), and *ki67* (O). (P) WT and *gata2b*<sup>-/-</sup> 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<sup>low</sup> cells in adult WT and *gata2b*<sup>-/-</sup> KM cells. (R) Bar graph representing the quantitation of cell cycle of CD41:GFP<sup>low</sup> cells in adult WT and *gata2b*<sup>-/-</sup> KM cells. Bars represent mean  $\pm$  standard error of the mean; each dot indicates analysis from 1 zebrafish. \* $P < .05$ , \*\* $P < .01$ . FDR, false discovery rate; n.s., not significant.

increased, probably responsible for the normalization in HSC numbers (Figures 4I-K and 5Q-R). We did not find upregulation of *gata2a* in these cells as a rescue mechanism (Figure 5M). Single-cell transcriptome analysis identified several HSPC populations with unique transcriptional signatures. Interestingly, the CD41:GFP<sup>low</sup>-expressing cells were scattered among different HSPC populations. Transplantation data suggest that only a minority of these cells are bona fide HSCs.<sup>30,61</sup> 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 zebrafish.

Myeloid lineage differentiation was not entirely 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 have found that if neutrophil development is blocked, myeloid differentiation progresses toward monocytic lineage.<sup>56</sup> In addition to an increase in monocytic progenitors, we also detected a redirection of lineage differentiation at a much earlier state, leading to increases in B-cell populations (Figure 3M; supplemental Figure 2J-K). This indicates that in *Gata2b*-deficient HSPCs, a reprogramming occurs both in immature cells to delineate lineage differentiation toward the lymphoid lineage and 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 plasma cells were reduced (Figure 4S). Along with the severe neutropenia, this is similar to conditions of patients with MonoMAC syndrome, which is characterized by neutropenia, monocytopenia, and dendritic cell- and B-cell lymphopenias.<sup>62-64</sup> This syndrome is caused by haploinsufficiency of the *GATA2* transcription factor. Despite the severe neutropenia, no infections were observed in *Gata2b*-deficient zebrafish, probably because of the SPF conditions of the animal facility.

In conclusion, we find that *Gata2b* is required for proliferation of the HSPC pool in CHT and vital for myeloid lineage differentiation in the adult, both in the HSPC compartment and for terminal differentiation.

## 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, Oeohara 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.

Loss of *Gata2b* consequently induces a differentiation diversion toward the lymphoid lineage.

## Acknowledgments

The authors thank members of the E.d.P., I.P.T., Raaijmakers, and Schneider-Kramann laboratories for helpful discussions; Tom Cupe-do (Erasmus MC) and Elaine Dzierzak (University of Edinburgh) for careful reading of the manuscript; and the Experimental Animal Facility of Erasmus MC for animal husbandry and the Erasmus Optical Imaging Center for confocal microscopy services. The visual abstract was created with BioRender.com.

This research was supported by the European Hematology Association (junior nonclinical research fellowship) (E.d.P.), the Dutch Cancer Foundation/Alpe d'HuZes (SK10321) (E.d.P.), the British Heart Foundation (BHF IBSR fellowship FS/13/50/30436) (R.M. and C.B.M.), the Wellcome Trust (PhD scholarship #WT102345/Z/13/Z) (T.D.), the Daniel den Hoed Foundation for support of the Cancer Genome Editing Center (I.P.T.), and the Josephine Nefkens Foundation for purchase of the Chromium (10× Genomics) (I.P.T.).

## Authorship

Contribution: E.d.P., E.G., and C.K. conceived the study; E.G., C.K., H.d.L., J.Z., D.B., T.D., C.B.M., P.M.H.v.S., M.E.v.R., and E.B. performed experiments; E.G., C.K., M.d.J., R.H., C.B.M., R.M., K.J.G., and E.d.P. analyzed results; R.M., P.J.F., and I.P.T. provided resources; E.G., C.K., and E.d.P. wrote the manuscript; and I.P.T. revised the manuscript.

ORCID profiles: C.K., 0000-0002-9006-1135; J.Z., 0000-0001-6829-244X; M.d.J., 0000-0002-6145-6154; T.D., 0000-0002-7708-3188; R.M., 0000-0002-4223-8506.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Emma de Pater, Erasmus MC, Wijtemaweg 80, Rotterdam, 3015 CN, The Netherlands; e-mail: e.depater@erasmusmc.nl.

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.e1-10.
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 [published correction appears in *PLoS One*. 2014;9(8):e106396]. *PLoS One*. 2014;9(5):e98186.
27. Chocron S, Verhoeven MC, Rentsch F, Hammerschmidt M, Bakkens 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):bio031096.
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-e56.
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 USA*. 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-e11.
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 USA*. 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. Intraortic 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 (plgR) and a family of plgR-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, Iyer 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 [published correction appears in *Nature*. 2019;569(7755):E3]. *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.