

## ARTICLE OPEN



## STEM CELL BIOLOGY

## Modeling GATA2 deficiency in mice: the R396Q mutation disrupts normal hematopoiesis

Trent Hall<sup>1</sup>, Rashid Mehmood<sup>1</sup>, Diana Sá da Bandeira<sup>1</sup>, Anitria Cotton<sup>1</sup>, Jonathon Klein<sup>2,3</sup>, Shondra M. Pruett-Miller<sup>2,3</sup>, Shai Izraeli<sup>4</sup>, Wilson K. Clements<sup>1</sup> and John D. Crispino<sup>1</sup>✉

© The Author(s) 2024

GATA2 deficiency is an autosomal dominant germline disorder of immune dysfunction and bone marrow failure with a high propensity for leukemic transformation. While sequencing studies have identified several secondary mutations thought to contribute to malignancy, the mechanisms of disease progression have been difficult to identify due to a lack of disease-specific experimental models. Here, we describe a murine model of one of the most common GATA2 mutations associated with leukemic progression in GATA2 deficiency, *Gata2*<sup>R396Q/+</sup>. While mutant mice exhibit mild defects in peripheral blood, they display significant hematopoietic abnormalities in the bone marrow, including a reduction in hematopoietic stem cell (HSC) function and intrinsic biases toward specific stem cell subsets that differ from previous models of GATA2 loss. Supporting this observation, single-cell RNA sequencing of hematopoietic progenitors revealed a loss of stemness, myeloid-bias, and indications of accelerated aging. Importantly, we show that *Gata2*<sup>R396Q/+</sup> exerts effects early in hematopoietic development, as mutant mice generate fewer HSCs in the aorta gonad mesonephros, and fetal liver HSCs have reduced function. This reduced and altered pool of HSCs could be potential contributors to leukemic transformation in patients, and our model provides a useful tool to study the mechanisms of malignant transformation in GATA2 deficiency.

*Leukemia* (2025) 39:734–747; <https://doi.org/10.1038/s41375-024-02508-z>

## INTRODUCTION

GATA2 is a hematopoietic transcription factor that is indispensable for hematopoietic stem cell (HSC) generation and maintenance [1], and alterations in its expression can lead to severe hematologic abnormalities. Over a decade ago, heterozygous GATA2 mutations leading to a loss of DNA binding ability were identified as drivers of familial myelodysplastic syndrome/acute myeloid leukemia (MDS/AML) in four families [2]. Germline GATA2 mutations have since been implicated in several autosomal dominant disorders with a propensity for familial MDS/AML [3–14], leading to the categorization of these disorders as “GATA2 deficiency” [15].

GATA2 deficiency is a pediatric/adolescent pre-leukemic disorder that is highly susceptible to transformation, with 75% of carriers developing a myeloid neoplasm by 20 years of age [16]. Studies of large European MDS cohorts revealed that 15% of MDS cases with excess blasts and 7% of all primary MDS cases contained germline GATA2 mutations [17], and that morbidity significantly increased in patients receiving HSC transplant post-adolescence [18], highlighting the need for early intervention. While several GATA2 variants are associated with GATA2 deficiency [16, 19–21], the most common mutations associated

with leukemia development are missense, hotspot mutations in the second zinc finger (ZF2) of GATA2 [18, 22, 23]. Molecular contributors to leukemic progression include cytogenetic abnormalities such as monosomy 7 or trisomy 8 followed by the acquisition of somatic mutations, including *ASXL1*, *STAG2*, and *SETBP1* [17, 18, 23–28]. However, the precise mechanisms underlying transformation are not fully understood.

The function of GATA2 as a potent regulator of HSC development and maintenance has been explored in several model systems [1, 29–34]. The first mammalian model of GATA2 deficiency disrupted the GATA2 intronic +9.5 enhancer site, which is altered in rare cases, and showed that baseline levels of GATA2 are necessary for HSC generation during development [35, 36]. Recently, two mouse models of missense ZF2 mutations were reported: L359V, which is an activating mutation commonly associated with myelomonocytic transformation of chronic myelogenous leukemia (CML) [37, 38], and R398W, a mutation commonly associated with MDS/AML formation in GATA2 deficiency [39]. Of note, *Gata2*<sup>R398W/+</sup> mice phenocopy the dendritic cell, monocyte, B and NK lymphoid (DCML) deficiency seen in some patients, and the mutant protein was shown to exert a dominant-negative effect on wild-type (WT) GATA2 by

<sup>1</sup>Department of Hematology, St. Jude Children’s Research Hospital, Memphis, TN, USA. <sup>2</sup>Center for Advanced Genome Engineering, St. Jude Children’s Research Hospital, Memphis, TN, USA. <sup>3</sup>Department of Cell and Molecular Biology, St. Jude Children’s Research Hospital, Memphis, TN, USA. <sup>4</sup>Department of Pediatric Hematology/Oncology, Schneider Children’s Medical Center of Israel, Tel Aviv University, Petah Tikva, Israel. ✉email: john.crispino@stjude.org

Received: 21 August 2024 Revised: 20 November 2024 Accepted: 17 December 2024

Published online: 7 January 2025

enhancing or silencing GATA2 target genes [39]. iPSC lines from GATA2 deficiency patients have been developed to model disease progression but show subtle hematological differences [40]. More recently, two heterozygous GATA2 mutant iPSC lines were developed with CRISPR/Cas9: R398W and R396Q [41], but the consequences of the mutations have yet to be reported. Here, we report the phenotype of a new GATA2 deficiency mouse model which mimics a common germline GATA2 missense mutation in ZF2 associated with MDS/AML progression, R396Q [22]. Using functional studies and single-cell RNA sequencing (scRNA-Seq), we show that *Gata2*<sup>R396Q/+</sup> mice display aberrant hematopoiesis throughout development due to abnormalities in lineage fate decisions and indications of accelerated aging within the HSC compartment, resulting in a loss of HSC stemness. However, the mice do not develop marked immunodeficiency or MDS/AML, indicating that additional drivers are necessary for disease.

## METHODS

### Mice

*Gata2*<sup>R396Q/+</sup> mice were created using CRISPR-Cas9 technology and direct zygote injection by the St. Jude Transgenic Core Facility. Briefly, a mixture of the 40 ng/ul 3X NLS SpCas9 protein (St. Jude Protein Production Core), 20 ng/ul chemically modified sgRNA (CAGE1313.Gata2.g2 – 5' AAGGGAUC-CAGACCCGGAU-3'; Synthego), and 10 ng/ul donor ssODN (CAGE1313.g2-mod.sense.ssODN-5'TAACAGGCCACTGACCATGAAGAAGGAGGATCCAGACCTGGAATCGCAAGATGTCCAGCAATCCAAGAAGAGCAAGAAAGGGCTG 3'; IDT AltR modification) was injected into fertilized zygotes as previously described [42]. Animals were genotyped by targeted deep sequencing using CAGE1313.Gata2.F – 5' 'ctacacgagctcttccgactctGGGGCTTTGGT GAGTCACATTGGGT 3' and CAGE1313.Gata2.R – 5' 'cagacgtgtgctcttccgactct-TAGGCGTGGGCGTGGGTAGGATGTG 3' primers at the Center for Advanced Genome Engineering (St. Jude) and resulting sequences analyzed by CRIS.py as previously described [43]. Animals positive for the desired R396Q modification were backcrossed to C57BL/6J mice and then bred to heterozygosity in a pathogen-free facility. B6.SJL (CD45.1<sup>+</sup>) mice were purchased from The Jackson Laboratory. For timed pregnancies, 1–3 female CD45.2<sup>+</sup> mice were placed with a *Gata2*<sup>R396Q/+</sup> male overnight. Females were assessed for a vaginal plug the following morning and prospective embryos were considered embryonic day 0.5 (E0.5). For all in vivo and ex vivo experiments, male and female mice were used and ranged from birth through one year. Animal studies were not blinded to the researcher. Animals were selected randomly from their litter for analysis. All animal experiments were approved by the St. Jude Children's Research Hospital Institutional Animal Care and Use Committee and comply with all relevant ethical regulations regarding animal research.

### Tissue collection and flow cytometry

Mice were bled from the retro-orbital plexus using heparinized capillary tubes and analyzed for complete blood counts (CBCs) using a Forcyte hematology analyzer. For flow cytometry, peripheral blood (PB) was lysed (eBioscience) then resuspended in FACS buffer (PBS/0.5% BSA/2 mM EDTA) and analyzed on a BD LSR Fortessa. All antibodies were purchased from BioLegend and catalog numbers are provided below. For transplant recipients, PB was stained with CD45.1-APC (110714), CD45.2-BV786 (109839), [B220 (103236), CD11b (101228), Gr-1 (108428)]-PerCP-Cy5.5, and [B220 (103222), CD4 (100528), CD8 (100722)]-PECy7. To isolate bone marrow (BM) hematopoietic cells in adult mice, tibias, femurs, and pelvic bones were removed and crushed in ice cold PBS, followed by passage through 70µm filters. Cells were then lysed, washed, and resuspended in FACS buffer. To isolate hematopoietic cells from fetal liver (FL), livers were gently pipetted in FACS buffer, lysed, passed through a 70µm filter, washed, and resuspended in FACS buffer. For flow cytometry of BM and FL hematopoietic progenitors, cells were stained with (1) Lineage (B220 (103244)), CD4 (100548), CD8 (100744), Gr-1 (108440), Ter119 (116239)-BV605, Sca1-PerCP-Cy5.5 (108124), c-Kit-APC780 (135108), CD135-APC (135310), CD48-A700 (103426), CD150-PE-Cy7 (115914) to identify LSK (Lineage<sup>+</sup>Sca1<sup>+</sup>c-Kit<sup>+</sup>), LT-HSC (LSKCD135<sup>+</sup>CD48<sup>+</sup>CD150<sup>+</sup>), ST-HSC (LSKCD135<sup>+</sup>CD48<sup>+</sup>CD150<sup>+</sup>), MPP2 (LSKCD135<sup>+</sup>CD48<sup>+</sup>CD150<sup>+</sup>), MPP3 (LSKCD135<sup>+</sup>CD48<sup>+</sup>CD150<sup>+</sup>), and MPP4 (LSKCD135<sup>+</sup>CD48<sup>+</sup>CD150<sup>+</sup>) cells or (2) Lineage (B220 (103244)), CD4 (100548), CD8 (100744), Gr-1 (108440), Ter119 (116239)-BV605, Sca1-PerCP-Cy5.5 (108124), c-Kit-APC780 (135108),

CD32/16-A700 (156620), CD34-A647 (343508), CD127-PE-Cy7 (135014) to identify LK (Lineage<sup>+</sup>Sca1<sup>+</sup>c-Kit<sup>+</sup>), CMP (LKCD34<sup>+</sup>CD32/16<sup>mid</sup>), GMP (LKCD34<sup>+</sup>CD32/16<sup>+</sup>), MEP (LKCD34<sup>+</sup>CD32/16<sup>+</sup>), and CLP (Lineage<sup>+</sup>c-Kit<sup>mid</sup>Sca1<sup>mid</sup>CD127<sup>+</sup>). For mature lineages, cells were stained with CD11b-FITC (101206), Gr-1-PerCP-Cy5.5 (108428), B220-BV786 (103246), Ter119-APC (116212), and CD71-BV786 (334102) to identify myeloid (CD11b<sup>+</sup>Gr-1<sup>+</sup>), B (B220<sup>+</sup>), or erythroid (R1, CD71<sup>+</sup>Ter119<sup>+</sup>; R2, CD71<sup>+</sup>Ter119<sup>+</sup>; R3, CD71<sup>+</sup>Ter119<sup>mid</sup>; R4, CD71<sup>+</sup>Ter119<sup>+</sup>; R5, CD71<sup>mid/low</sup>Ter119<sup>+</sup>). For analysis of transplanted bone marrow, CD45.1-BV786 (103149)/APC (110714)/FITC (110706) and CD45.2-PE (109808) were added. All flow data were analyzed using FlowJo version 10 (Treestar).

### Single-cell RNA sequencing (scRNA-seq)

LSK cells were sorted on a BD Aria cell sorter, and mRNA expression libraries were constructed according to the Chromium 10X single cell 3' reagent v3.1 protocol. Libraries were sequenced using Illumina NovaSeq paired-end sequencing per the 10X protocol to a minimum depth of 50,000 reads per cell. Raw reads from four WT and three *Gata2*<sup>R396Q/+</sup> libraries consisting of 2–3 mice (all mice were 8–10 week old females) were mapped to GRCh38 and gene expression matrices were obtained using count command in Cell Ranger (v.7.0.0). The filtered\_feature\_bc\_matrix was used in Seurat (V.4.0.3) with R (V.4.3.1) for downstream analysis. Low-quality cells were regressed out if the detected genes per cell were ≤ 200 or ≥ 6000, and the percentage of mitochondrial reads was greater than or equal to 5%. Harmony was used to integrate different replicates and remove the batch effects [44]. Default Seurat pipelines were followed to generate UMAP and clustering the cells into different groups. Differential genes between *Gata2*<sup>WT</sup> and *Gata2*<sup>R396Q/+</sup> were obtained using FindMarkers function in Seurat for each cluster. Gene set enrichment analyses were performed using GSEAPy (V.1.1.0) using python (V.3.11.6).

### Statistical analysis

Unless otherwise specified, data are presented as mean ± standard deviation with sample sizes and replicates indicated in figure legends. No statistical sample size calculation was performed to predetermine sample sizes. Variances were similar for the groups being compared. No samples or animals were excluded from analysis. Significance was determined via Mann-Whitney or Wilcoxon Signed Rank Test. We calculated the hscScore of each cell in our scRNA data using the published hscScore method [45]. hscScore is visualized using VlnPlot in Seurat with integration of the ggpubr package in R. A customized script is provided: ([https://github.com/gul Khan007/VlnPlot\\_pvalue\\_example](https://github.com/gul Khan007/VlnPlot_pvalue_example)).

Additional methods are provided in an online supplement.

## RESULTS

### *Gata2*<sup>R396Q/+</sup> mice display significant alterations in bone marrow hematopoiesis during adulthood

To model one of the most recurrent GATA2 ZF2 missense mutations associated with MDS/AML development in GATA2 deficiency, we utilized CRISPR/Cas9 genome editing and homology directed repair to induce a G to A point mutation at c.1087 in heterozygosity on the C57BL/6J background, resulting in the *Gata2*<sup>R396Q/+</sup> mutant mouse (Fig. 1A, Supplemental Fig. 1A). Digital droplet PCR confirmed that *Gata2*<sup>R396Q</sup> is expressed at 50% in *Gata2*<sup>R396Q/+</sup> mouse whole bone marrow cells (Fig. 1B). To assess the effects of this mutation on normal hematopoiesis, we first aged a cohort of mice and tracked changes in peripheral blood (PB) output over time through complete blood counts (CBC). Young (10–12 weeks) *Gata2*<sup>R396Q/+</sup> mice displayed a modest but significant reduction in white blood cell (WBC), neutrophil, and monocyte counts compared to WT littermates, with no differences in red blood cell (RBC) or platelet counts (Fig. 1C, top). The reduction in monocytes observed in young mice is intriguing, as many GATA2 deficiency patients experience severe monocytopenia [21]. Of note, the previously published *Gata2*<sup>R398W/+</sup> mouse model began showing reduced WBC counts at 6 months of age and severe monocytopenia at 12 months, while 3-month-old mutants had relatively normal CBCs [39], and *Gata2*<sup>+/-</sup> mice do not display significant differences in PB output [30]. While older mice (53–55 weeks) in our cohort continued to

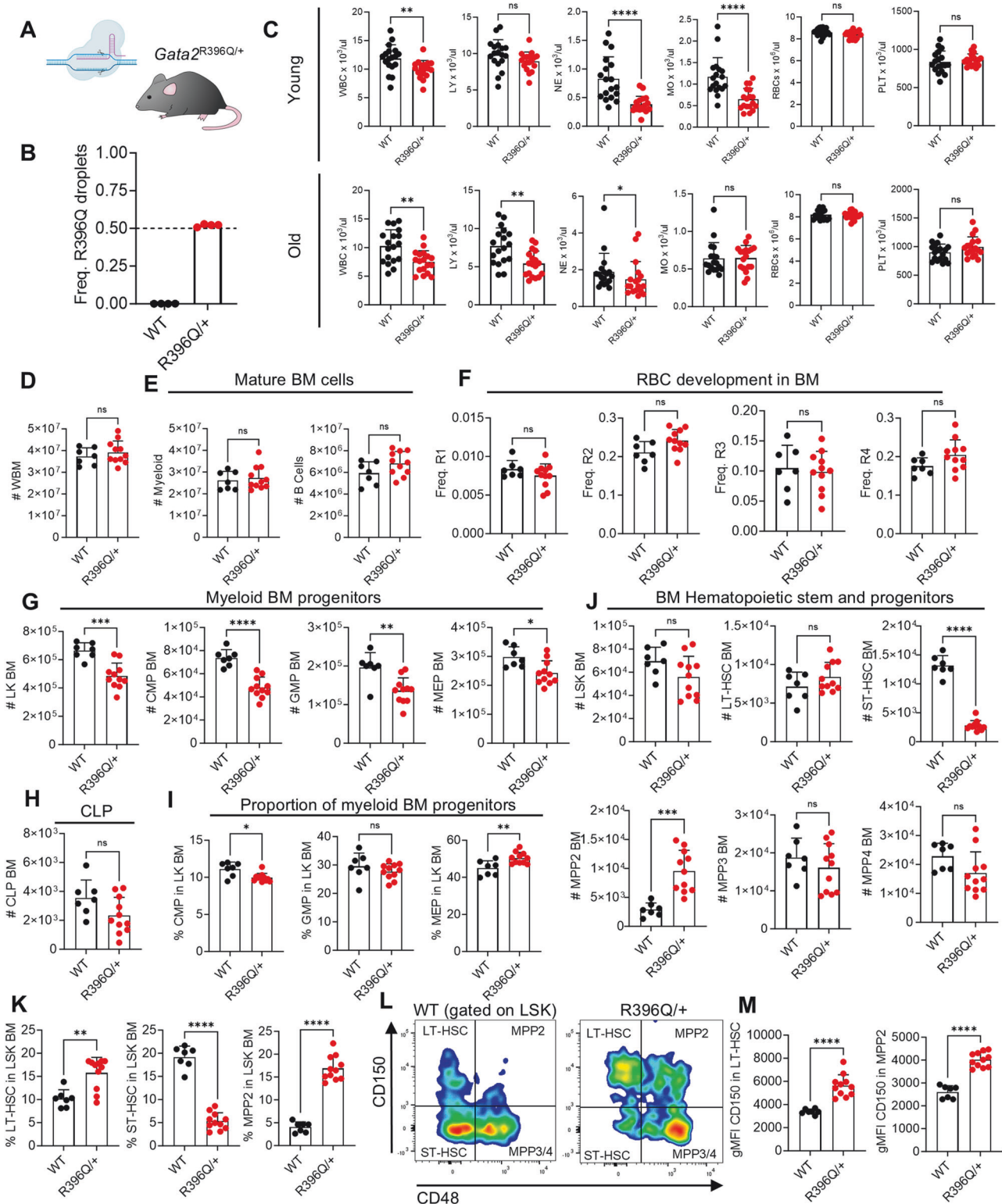


exhibit significantly reduced WBC, neutrophil, and lymphocyte counts, monocytopenia was not present by 12 months (Fig. 1C, **bottom**), suggesting that the R398W and R396Q mutations can elicit disparate phenotypes.

We next assessed the bone marrow (BM) hematopoietic compartment of 8–10 week old WT and *Gata2*<sup>R396Q/+</sup> littermates (Supplemental Fig. 1B–E). *Gata2*<sup>R396Q/+</sup> mice showed no significant differences in total BM cellularity, mature B or myeloid cells,

or RBC maturation (Fig. 1D–F), although there was a trend for increased B cells in the BM (Fig. 1E). However, mutant mice had significantly reduced numbers of committed myeloid progenitors, a trend for a reduction in common lymphoid progenitors (CLP), and a bias toward megakaryocytic/erythroid progenitors (MEP) at the expense of common myeloid progenitors (CMP) (Fig. 1G–I). More strikingly, the *Gata2*<sup>R396Q/+</sup> hematopoietic stem and progenitor cell (HSPC) compartment exhibited a significant



**Fig. 1** *Gata2*<sup>R396Q/+</sup> mice display significant alterations in bone marrow hematopoiesis during adulthood. **A** Schematic representing CRISPR/Cas9 editing of *Gata2* to produce *Gata2*<sup>R396Q/+</sup> heterozygous mice. **B** Frequency of R396Q positive droplets in WT and *Gata2*<sup>R396Q/+</sup> whole bone marrow (WBM) via digital droplet PCR ( $n = 4$ ). **C** Complete blood count (CBC) data from young (10–12 weeks) and old (53–55 weeks) mice ( $n = 18$ ). Absolute number of WBM cells **D**, mature cells **E**, maturing red blood cells (RBC) **F**, myeloid progenitors **G**, and common lymphoid progenitors (CLP) **H** in 8–10 week old mice ( $n = 7–11$ ). **I** Proportion of CMPs, GMPs, or MEPs among myeloid progenitors in the BM of 8–10 week old mice ( $n = 7–11$ ). **J** Absolute number of BM hematopoietic stem and progenitors (HSPCs) in 8–10 week old mice ( $n = 7–11$ ). **K** Proportion of LT-HSC, ST-HSC, and MPP2 cells among BM LSK cells in 8–10 week old mice ( $n = 7–11$ ). Representative flow plot **L** and quantification **M** of gMFI CD150 expression among BM HSPCs in 8–10 week old mice. ns, not significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$  WBC white blood cells, LY lymphocytes, NE neutrophils, MO monocytes, RBC red blood cells, PLT platelets, WBM whole bone marrow, LK lineage<sup>c</sup>-Kit<sup>+</sup>; CMP common myeloid progenitor, GMP granulocyte-monocyte progenitor, MEP megakaryocyte-erythroid progenitor, CLP common lymphoid progenitor, LSK lineage Sca-1<sup>+</sup>c-Kit<sup>+</sup>, LT-HSC long-term hematopoietic stem cell, ST-HSC short-term HSC, MPP2 multipotent progenitor 2, gMFI geometric mean fluorescence intensity.

reduction in the short-term HSC (ST-HSC) population and a significant increase in multipotent progenitor 2 (MPP2) cells, with a trend for increased long-term HSC (LT-HSC) and reduced total Lineage<sup>Sca1<sup>+</sup>c-Kit<sup>+</sup></sup> (LSK) cells (Fig. 1J). Gating on LSK cells revealed a bias in HSPC production toward LT-HSC and MPP2 cells at the expense of ST-HSCs (Fig. 1K), which appeared to be driven by an increase in expression of the SLAM family marker CD150 (Fig. 1L–M).

To assess the functional output of specific HSPCs, we performed single cell colony-forming unit (CFU) assays in methylcellulose with cytokines to induce differentiation (Supplemental Fig. 2A). Surprisingly, *Gata2*<sup>R396Q/+</sup> LT-HSCs and other HSPCs showed little to no difference in CFU potential, while MPP2s displayed a slight bias for GEMM and G/M/GM colonies at the expense of Erythroid/Meg colonies (Supplemental Fig. 2B, C). As MPP2s are known to have an intrinsic bias toward myeloid and erythroid/megakaryocyte differentiation [46], this result suggests that the immunophenotypic MPP2 population in *Gata2*<sup>R396Q/+</sup> BM—which has high expression of CD150—may contain a heterogeneous mixture of cells with disparate lineage potential in this in vitro assay. Together, these data show that the R396Q mutation induces significant changes in blood production during homeostasis, especially at the level of stem and progenitor cells.

#### *Gata2*<sup>R396Q/+</sup> bone marrow progenitors are functionally impaired

We next assessed the functional potential of the *Gata2*<sup>R396Q/+</sup> hematopoietic compartment by competitive transplantation assays with whole bone marrow (WBM) from WT and *Gata2*<sup>R396Q/+</sup> (CD45.2<sup>+</sup>) into lethally irradiated recipients (CD45.1<sup>+</sup>) (Fig. 2A). *Gata2*<sup>R396Q/+</sup> recipients showed a significant reduction in PB chimerism by 4 weeks post-transplant in two independent experiments (Fig. 2B, Supplemental Fig. 2D). While *Gata2*<sup>R396Q/+</sup> cells were outcompeted by competitor BM in all mature lineages as expected (Fig. 2C, left), they were still able to generate all mature lineages at the expected ratios (Fig. 2C, right), suggesting a defect in repopulation but not lineage potential after primary transplant. Assessment of the bone marrow at 20 weeks post-transplant revealed a reduction in chimerism for all BM compartments (Supplemental Fig. 2E–H). However, many of the intrinsic properties of *Gata2*<sup>R396Q/+</sup> seen during homeostasis were recapitulated post-transplant, including a preference for MPP2 production at the expense of ST-HSC production (Fig. 2D). This was again driven by an increase in CD150 expression in MPP2s (Fig. 2E, F), indicating the presence of LT-HSCs in *Gata2*<sup>R396Q/+</sup> BM with a propensity for differentiation into CD150<sup>hi</sup> progenitors.

We next isolated LSKs from primary recipients and performed secondary transplants to determine bona fide HSC potential (Fig. 2G). While there was a significant reduction in repopulating potential in the peripheral blood and bone marrow of *Gata2*<sup>R396Q/+</sup> secondary recipients compared to WT, long-term potential was still present by 20 weeks post-transplant (Fig. 2H, Supplemental

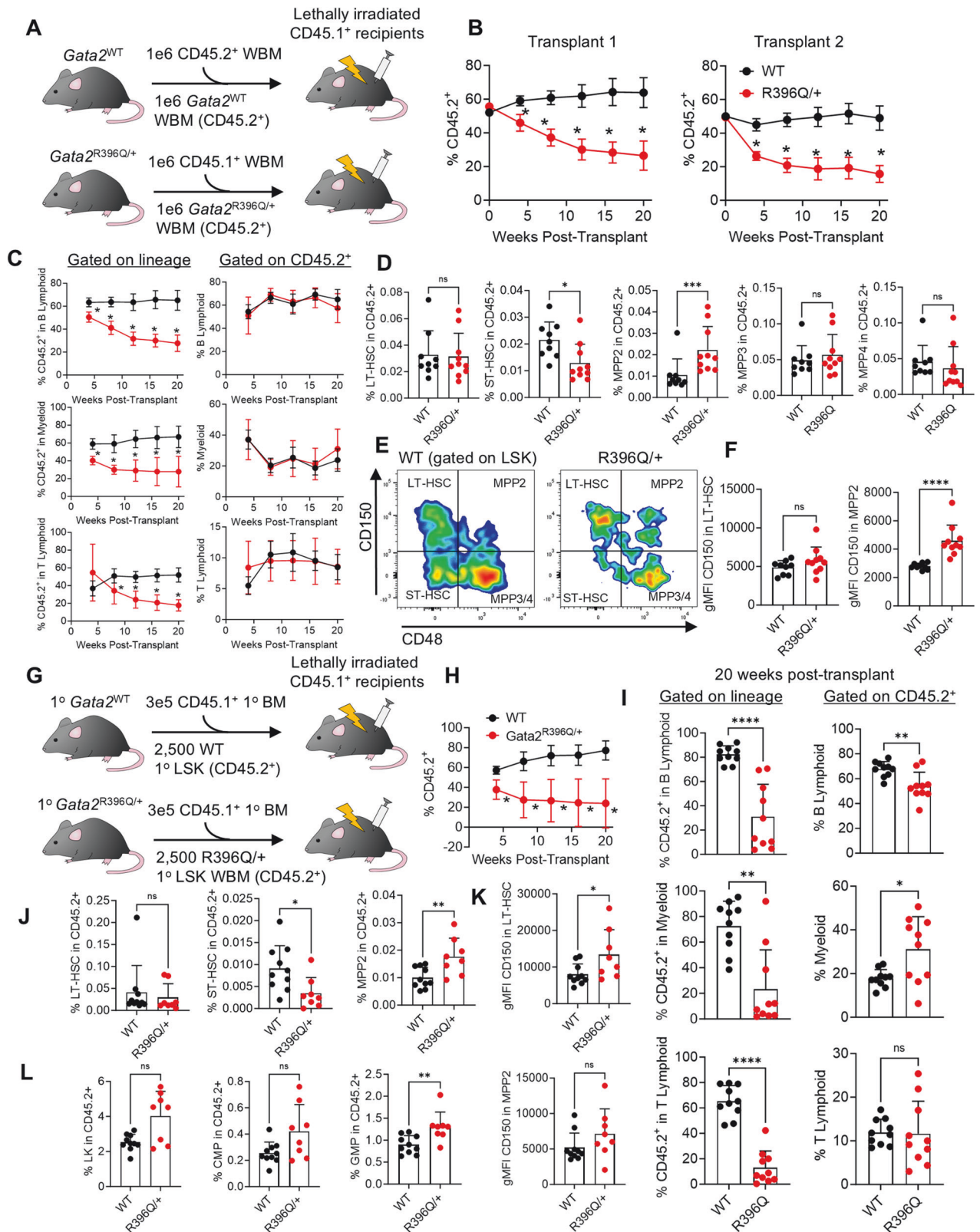
Fig. 3A–D). However, a slight but significant bias toward myeloid output at the expense of lymphoid differentiation was seen in the PB of *Gata2*<sup>R396Q/+</sup> secondary recipients (Fig. 2I). Consistent with primary recipients, *Gata2*<sup>R396Q/+</sup> recipients showed a bias toward MPP2 production at the expense of ST-HSCs, as well as a trend for increased CD150 expression among LT-HSCs and MPP2s (Fig. 2J–K). While not significant for all populations, secondary recipients also displayed an increase in the production of myeloid progenitors and a trend for a decrease in B cell output in the BM (Fig. 2L, Supplemental Fig. 3E). Coupled with the increase in CD150 expression among LT-HSC/MPP2 populations, this could be an indication of an accelerated aging phenotype within the *Gata2*<sup>R396Q/+</sup> HSC compartment, as aged HSCs are known to have increased CD150 expression and a myeloid bias [47].

Together, these data show that the stem cell compartment of *Gata2*<sup>R396Q/+</sup> BM is functionally impaired and contains intrinsic biases for progenitor output in vivo.

#### The transcriptional signature of *Gata2*<sup>R396Q/+</sup> hematopoietic progenitors indicate a myeloid bias and loss of HSC stemness

We next performed 10X scRNA-seq on sorted LSK cells from 8–10 week old WT and mutant bone marrow, revealing 22 distinct clusters (Fig. 3A). To annotate the clusters, we combined our dataset with recently published murine BM hematopoietic progenitor scRNA-seq data from the Göttgens group [48], and validated our results with known lineage markers (Fig. 3B, C, Supplemental Figs. 4–6) [48–53]. UMAP projections of the datasets revealed a myeloid progenitor bias in *Gata2*<sup>R396Q/+</sup> LSKs at the expense of erythroid/meg progenitors (Fig. 3B, D). *Gata2*<sup>R396Q/+</sup> LSKs also showed a decrease in the proportion of the earliest hematopoietic progenitors (HSC 1, HSC 2, Intermediate Progenitor (Int Prog) 1) and an increase in more committed progenitors (Int Prog 2–4, Lymphoid (Lymph) Prog) (Fig. 3D, Supplemental Fig. 6A). We also attempted to define immunophenotypic HSPC subsets (LT-HSC, MPP2, etc.) in our scRNA-seq data using publicly available datasets on Haemosphere (Supplemental Fig. 6B) [54]. Similar to our flow cytometry data (Figs. 1, 2), *Gata2*<sup>R396Q/+</sup> cells putatively annotated as LT-HSC, ST-HSC, and MPP2 expressed higher levels of *Slamf1* (CD150) (Supplemental Fig. 6C, D).

Given the reduced function of *Gata2*<sup>R396Q/+</sup> progenitors in vivo, we hypothesized that the transcriptional signature of HSC and Int Prog clusters would exhibit reduced stemness. To test this, we applied hscScore—a computational tool that scores cells in scRNA-seq data based on their similarity to the transcriptomes of validated HSCs—to our dataset [45]. hscScore validated our annotations, as the clusters with the highest hscScore were the most primitive (HSC 1, HSC 2, Int Prog 1, and Lymph Prog) (Fig. 3E). Meanwhile, hscScore was significantly lower in *Gata2*<sup>R396Q/+</sup> LSKs for all cells, Int Prog 1, Int Prog 2, and Lymph Prog clusters (Fig. 3F). Interestingly, while hscScore was significantly lower for the HSC 2 cluster in *Gata2*<sup>R396Q/+</sup> LSKs, the HSC 1 cluster showed no significant differences in hscScore between WT and mutant LSKs (Fig. 3F). This coupled with the reduction in the proportion of the



most primitive clusters in *Gata2*<sup>R396Q/+</sup> LSKs suggests that the reduced repopulating potential seen in our competitive transplants may be due to multiple factors: a reduction in the number of bona fide HSCs and a loss of stem cell potential, as more differentiated HSCs (represented by HSC 2) display lower HSC score.

### *Gata2*<sup>R396Q/+</sup> hematopoietic progenitors exhibit hallmarks of aging, stem cell exhaustion, and upregulation of *Gata2*

We next performed GSEA analysis on the HSC, Int Prog, and Lymph Prog clusters in our scRNA-seq dataset to determine significantly up- and down-regulated biological processes in *Gata2*<sup>R396Q/+</sup> cells.

**Fig. 2 The hematopoietic compartment of *Gata2*<sup>R396Q/+</sup> bone marrow displays functional defects.** **A** Schematic of primary competitive transplant of WT or *Gata2*<sup>R396Q/+</sup> WBM. **B** Peripheral blood chimerism in recipients of WT or *Gata2*<sup>R396Q/+</sup> WBM over 20 weeks of transplant. Two independent transplants shown ( $n = 9$ –10 recipients/condition in each transplant). **C** Peripheral blood output of recipients, showing CD45.2<sup>+</sup> chimerism within myeloid and lymphoid lineages (left) or lineage output from CD45.2<sup>+</sup> cells (right). Representative transplant is shown ( $n = 9$ –10 recipients). **D** Percent output of hematopoietic stem and progenitor cells (HSPCs) from CD45.2<sup>+</sup> cells in recipient bone marrow 20 weeks post-transplant. Representative transplant is shown ( $n = 9$ –10 recipients). **E** and quantification **F** of gMFI CD150 expression among BM HSPCs in recipients of WT or *Gata2*<sup>R396Q/+</sup> bone marrow 20 weeks post-transplant. Representative transplant is shown ( $n = 9$ –10 recipients). **G** Schematic for secondary transplantation of WT or *Gata2*<sup>R396Q/+</sup> LSK cells. **H** Peripheral blood chimerism in secondary transplant ( $n = 10$  recipients per condition). **I** Peripheral blood output of recipients, showing CD45.2<sup>+</sup> chimerism within myeloid and lymphoid lineages (left) or lineage output from CD45.2<sup>+</sup> cells (right) ( $n = 10$  recipients per condition). **J** Percent output of HSPCs from CD45.2<sup>+</sup> cells in secondary recipient BM 20 weeks post-transplant ( $n = 8$ –10 recipients). **K** Quantification of gMFI CD150 expression among BM HSPCs in secondary recipients 20 weeks post-transplant ( $n = 8$ –10 recipients). **L** Percent output of myeloid progenitors from CD45.2<sup>+</sup> cells in secondary recipient BM 20 weeks post-transplant ( $n = 8$ –10 recipients). ns not significant, \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . WBM whole bone marrow, LK lineage c-Kit<sup>+</sup>, LSK lineage Sca-1<sup>+</sup>c-Kit<sup>+</sup>, LT-HSC long-term hematopoietic stem cell, ST-HSC short-term HSC, MPP2 multipotent progenitor 2, gMFI geometric mean fluorescence intensity, CMP common myeloid progenitor, GMP granulocyte-macrophage progenitors.

*Gata2*<sup>R396Q/+</sup> HSC 1 and HSC 2 clusters showed an upregulation of GO terms involved in mitochondrial activity, and downregulation of terms involved in HSC signature, response to cytokines, and differentiation (Fig. 4A). Similar terms were found for Int Prog 1/ Lymph Prog clusters as well as Int Prog 2/3/4 clusters (Fig. 4B, C). Combined with the myeloid bias and reduced hscScore seen in our scRNA-seq data, these terms suggest a phenotype of aging and stem cell exhaustion within the *Gata2*<sup>R396Q/+</sup> HSC compartment. For instance, increased mitochondrial activity has been shown to be associated with HSC aging [55], and aged HSCs lose self-renewal capacity and are myeloid-biased [47, 56, 57]. A blunted response to cytokines may also be indicative of a stem cell compartment that is on the pathway to senescence. Further GSEA analysis revealed a significant enrichment of BM HSC aging pathway in *Gata2*<sup>R396Q/+</sup> HSC, Int Prog, and Lymph Prog clusters (Fig. 4D). Indeed, all *Gata2*<sup>R396Q/+</sup> HSPCs displayed significantly increased oxidative stress (Supplemental Fig. 7A), which may be indicative of premature exhaustion of the stem cell compartment [58]. Also, acute treatment of *Gata2*<sup>R396Q/+</sup> mice with lipopolysaccharide (LPS)—which drives quiescent HSCs into cycling—resulted in a significant reduction in LT-HSC recovery compared to WT controls, suggesting that the mutant HSC compartment contains fewer quiescent HSCs (Fig. 4E, F).

To further identify transcriptional differences between WT and mutant LSKs, we looked for genes specifically upregulated in *Gata2*<sup>R396Q/+</sup> clusters. As expected, the myeloid progenitor marker *Ctsg* was upregulated in progenitor cells surrounding the Myeloid Prog cluster in *Gata2*<sup>R396Q/+</sup> cells (Fig. 5A), validating an enhanced myeloid bias. Interestingly, *Ly6c1*—a cell-surface marker for granulocyte and monocyte/macrophage populations—was exclusively expressed in *Gata2*<sup>R396Q/+</sup> progenitor cells surrounding the Myeloid Prog cluster (Fig. 5B). Another highly upregulated gene was *Tgfb3* (Fig. 5C), also known as betaglycan, an important co-receptor for TGF $\beta$  receptors [59]. TGF $\beta$  plays several roles in HSPC function and has been shown to inhibit the lymphoid/megakaryocyte/erythroid differentiation pathways while stimulating the monocyte/DC pathway [59], which is reflected in our scRNA-seq dataset. Interestingly, *Gata2* expression was significantly increased in several *Gata2*<sup>R396Q/+</sup> clusters, including HSC 1, Int Prog 1, and Lymph Prog (Fig. 5D, E, Supplemental Fig. 7B). Increased *Gata2* expression within hematopoietic progenitors was also seen in the *Gata2*<sup>R398W/+</sup> mouse [39]. To determine if this was due to enhanced WT or mutant *Gata2* expression, we sorted HSPCs and performed droplet digital PCR (ddPCR) for mutant allele expression. *Gata2*<sup>R396Q/+</sup> mice displayed varying mutant expression ratios depending on the cell type, with a lower mutant ratio in ST-HSCs and a significantly higher ratio in MPPs (Supplemental Fig. 7C). Given that our scRNA-Seq data showed increased *Gata2* expression in all putatively annotated HSPC populations (Supplemental Fig. 7D), this suggests that the degree of WT *Gata2*

expression may be more critical in some populations (ST-HSC), while mutant expression may be permissible or even advantageous in others (MPP2).

Interestingly, the Lymph Prog cluster appeared to upregulate *Gata2* exclusively in *Gata2*<sup>R396Q/+</sup> cells (Fig. 5D). We therefore subclustered on the Lymph Prog cluster and identified two unique clusters, defined by *Gata2* expression (Fig. 5F, left). These cells expressed lymphoid (*Dnrtt*) and myeloid (*Ctsg*) markers and appeared to arise from a small population of HSCs based on *Procr* expression (Fig. 5F), indicating the presence of a unique *Gata2*-expressing progenitor population confined to the *Gata2*<sup>R396Q/+</sup> LSK compartment that may contribute to the enhanced myeloid potential seen in our scRNA-seq dataset.

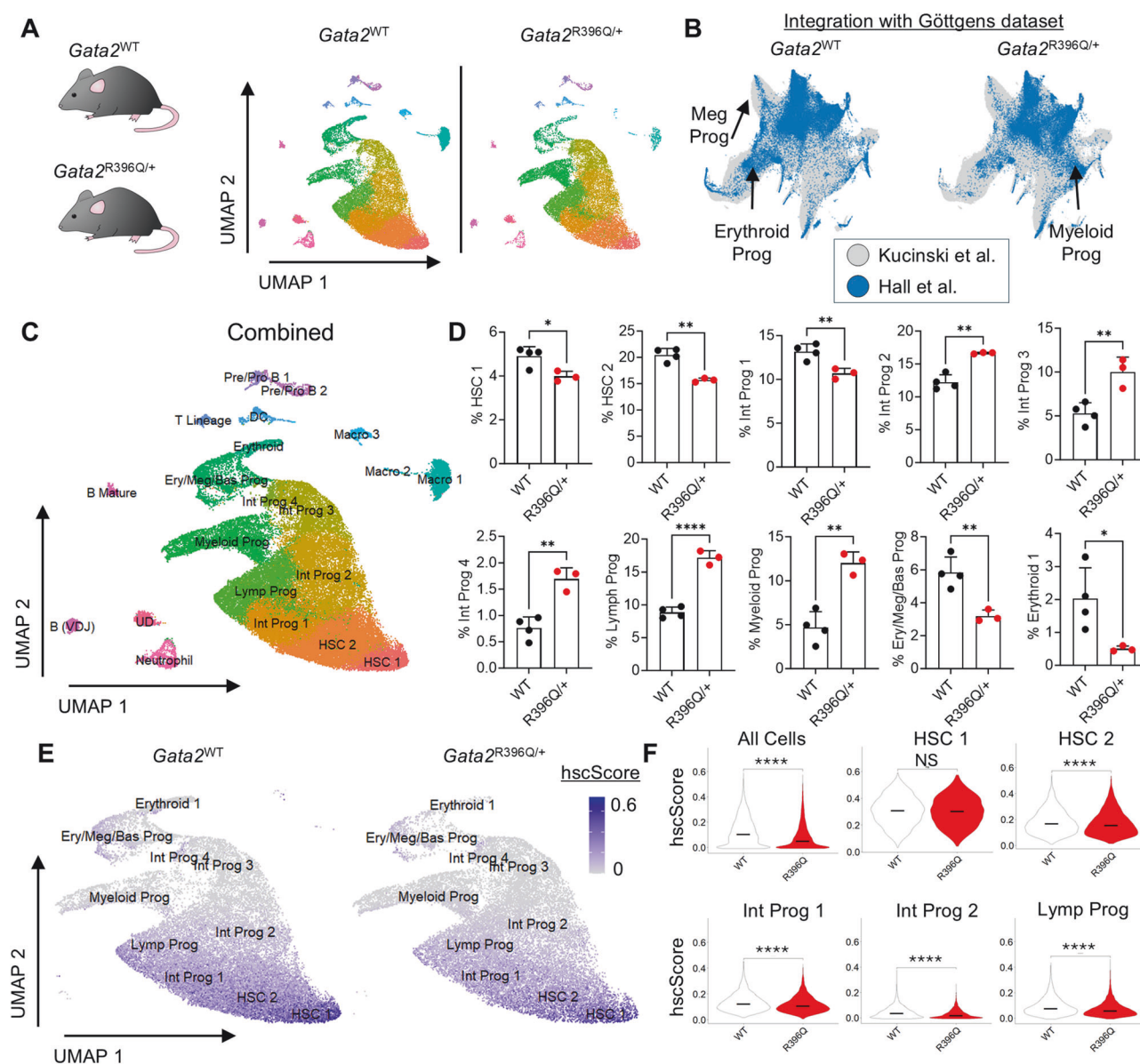
Together, these data provide evidence for an HSPC compartment exhibiting stem cell exhaustion and aging-related signatures in *Gata2*<sup>R396Q/+</sup> BM that may contribute to the phenotypic and functional defects seen in vivo.

### ***Gata2*<sup>R396Q/+</sup> affects fetal liver hematopoietic development and function**

As GATA2 deficiency is primarily a pediatric/adolescent disease, we investigated whether the effects of *Gata2*<sup>R396Q/+</sup> on hematopoiesis could be detected prenatally. To do this, we first performed flow cytometry and CFU assays on embryonic day 14.5 (E14.5) fetal liver (FL) cells (Fig. 6A), as at E14.5 the definitive hematopoietic compartment is largely confined to the FL [60]. While there were no significant differences in FL cellularity, *Gata2*<sup>R396Q/+</sup> cells showed a significant increase in all HSPC subtypes except MPP3/4 cells (Fig. 6B). Similarly to BM HSPCs, E14.5 FL HSPCs showed a bias for LT-HSC and MPP2 differentiation, with significantly increased CD150 expression in MPP2s (Fig. 6C–E), indicating that this intrinsic property of *Gata2*<sup>R396Q/+</sup> HSPCs begins during embryonic hematopoiesis. While there was an increase in immunophenotypic HSPCs, CFU assays revealed a functional deficiency in *Gata2*<sup>R396Q/+</sup> E14.5 FL progenitors in terms of GEMM and BFU-E production (Fig. 6F). Similar to the BM compartment, these data suggest the expansion of a functionally compromised pool of HSPCs.

We next looked at E12.5 FL: the murine timepoint at which definitive HSCs have migrated into the FL niche (Fig. 6G) [60]. Similar to E14.5 FL, there was an expansion of LSK and LT-HSC populations in *Gata2*<sup>R396Q/+</sup> embryos (Fig. 6H, I). However, the geometric mean fluorescence intensity (gMFI) of CD150 within LT-HSC and MPP2 cells was not significantly enhanced, suggesting that this phenomenon does not occur until later in FL hematopoiesis (Fig. 6J). E12.5 FL cells showed a significant reduction in GEMM and BFU-E CFUs (Fig. 6K), and transplantation of E12.5 FL cells into irradiated recipients resulted in a significant reduction in repopulating potential (Fig. 6L–N). Similar to primary BM transplants, *Gata2*<sup>R396Q/+</sup> E12.5 FL cells displayed reduced





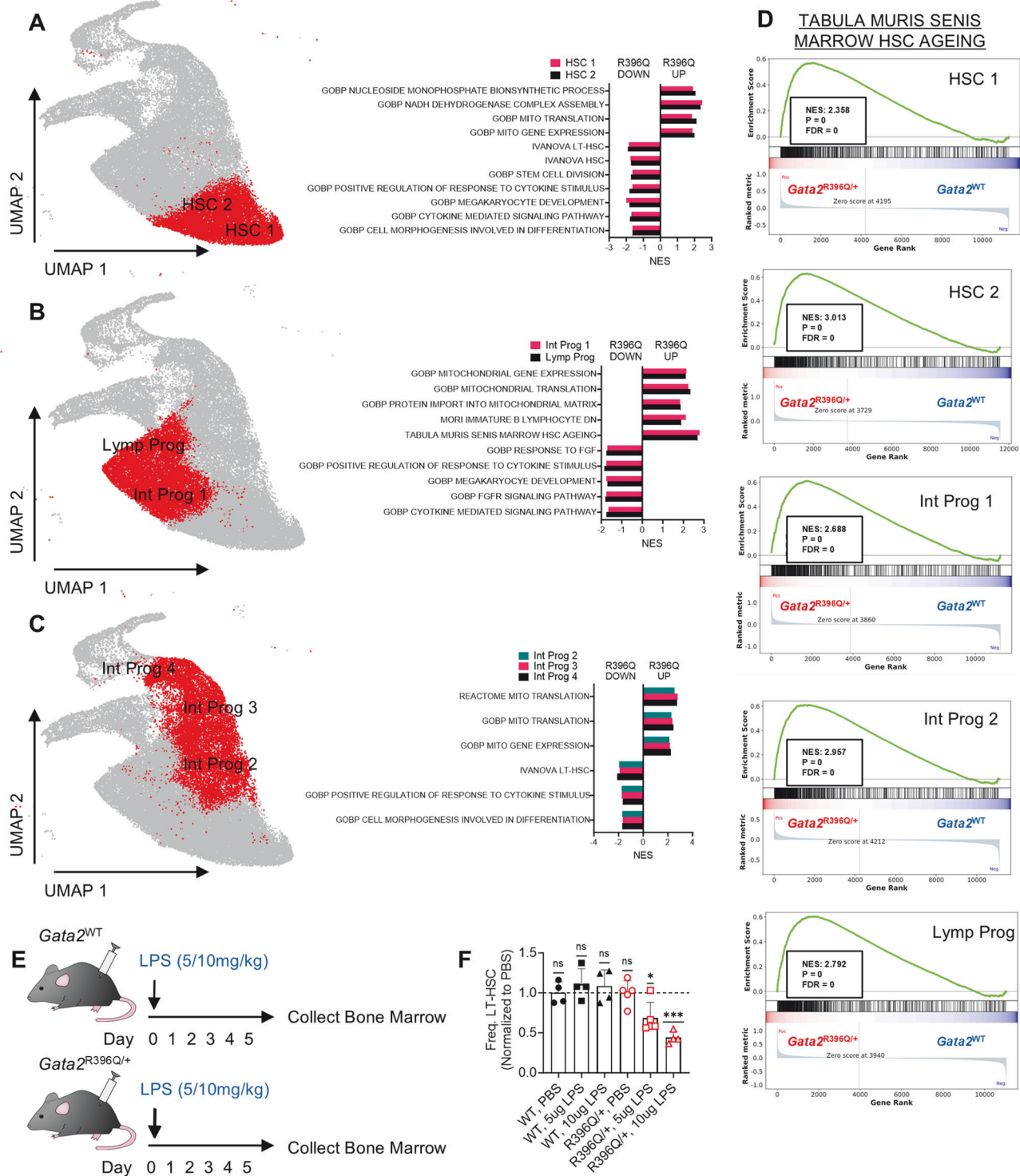
**Fig. 3** The transcriptional signature of *Gata2*<sup>R396Q/+</sup> hematopoietic progenitors reveals myeloid bias and a loss of HSC stemness. **A** UMAP projections of scRNA-seq from sorted WT or *Gata2*<sup>R396Q/+</sup> BM LSK cells from 8–10 week old mice ( $n = 3$ –4 samples per condition). **B** UMAP projections of WT or *Gata2*<sup>R396Q/+</sup> LSKs (blue) overlaid onto previously published scRNA-seq data from Kucinski et al. (gray), with indications of specific lineages. **C** Cluster annotations for LSK cells from WT and *Gata2*<sup>R396Q/+</sup>. **D** Proportions of specific cell types among WT or *Gata2*<sup>R396Q/+</sup> LSKs ( $n = 3$ –4). **E** FeaturePlot displaying hscScore values among specific clusters in WT or *Gata2*<sup>R396Q/+</sup> LSK cells. **F** Quantification and comparison of hscScores between specific clusters of WT or *Gata2*<sup>R396Q/+</sup> LSKs. ns not significant, \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . Prog progenitor, Mono monocyte, DC dendritic cell, pDC plasmacytoid dendritic cell, Ery/Meg/Bas erythroid, megakaryocyte, basophil, Int intermediate, Lymph lymphoid, HSC hematopoietic stem cell.

competitive potential but a retained ability to produce all major lineages at expected ratios, suggesting that the pool of E12.5 FL HSCs is reduced but not fully dysfunctional.

#### *Gata2*<sup>R396Q/+</sup> abrogates definitive HSC production

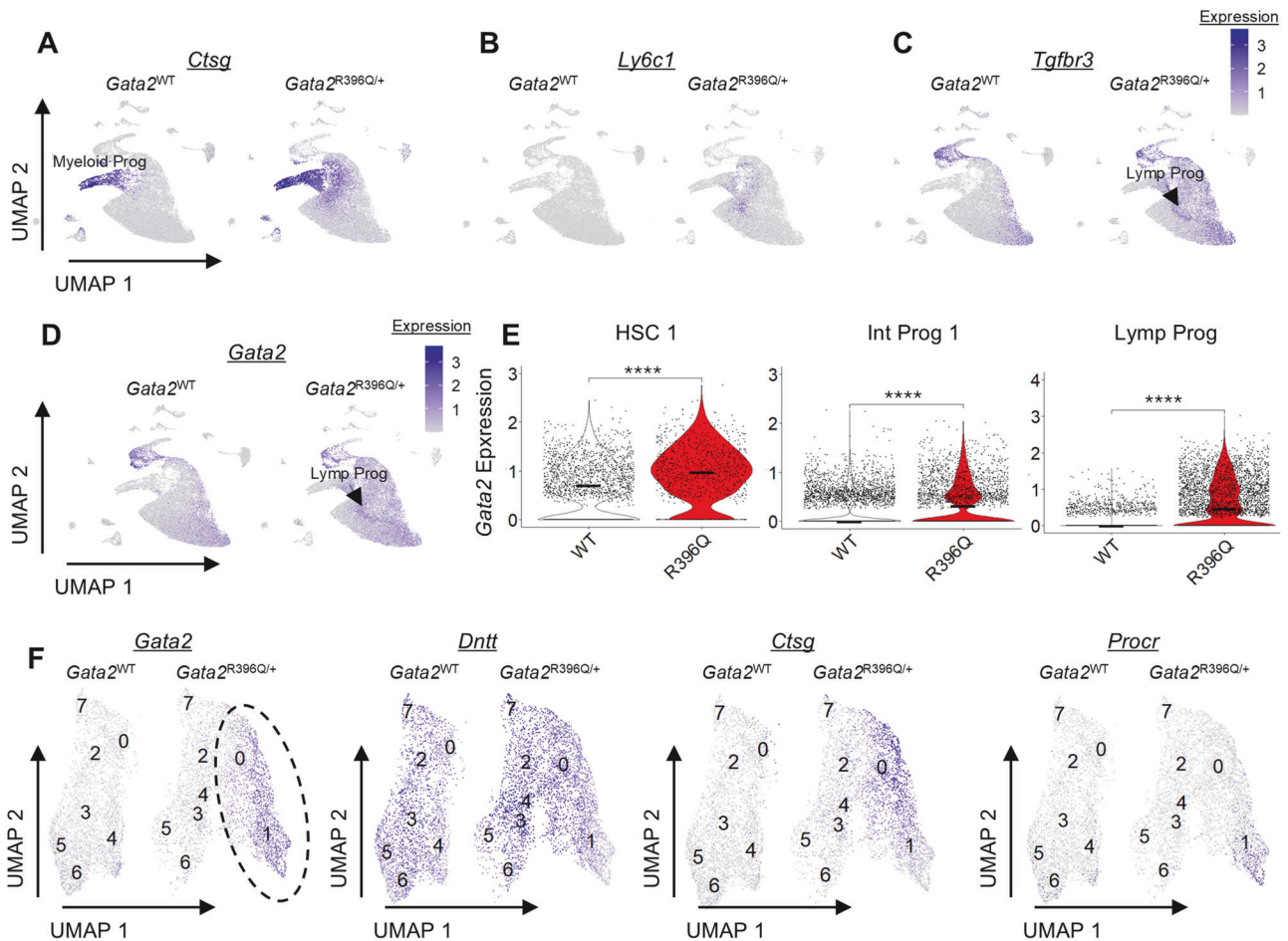
Given the reduced function of *Gata2*<sup>R396Q/+</sup> E12.5 FL cells and the known role of GATA2 in hematopoietic development [1], we interrogated the generation of definitive HSCs in the aorta gonad mesonephros (AGM) of WT and mutant mice. First, we performed transplants of E11 AGM into lethally irradiated recipients with support BM cells (Fig. 7A). While 6/23 recipients of WT AGM showed PB chimerism  $> 1\%$ , 0/14 recipients of *Gata2*<sup>R396Q/+</sup> AGM had sufficient levels of donor chimerism (Fig. 7B), suggesting either a functional deficiency or lack of definitive HSCs in the AGM.

We then performed imaging of E10 AGMs to determine the number of intra-aortic hematopoietic clusters in WT and *Gata2*<sup>R396Q/+</sup> embryos. *Gata2*<sup>R396Q/+</sup> embryos contained significantly fewer intra-aortic hematopoietic clusters than WT littermates (Fig. 7C, D), suggesting that the R396Q mutation results in a loss or stalling of definitive HSC production. Finally, given the reduced production of intra-aortic hematopoietic clusters and reduced functional potential of FL and BM HSCs in *Gata2*<sup>R396Q/+</sup>, we hypothesized that the absolute number of HSCs was reduced into adulthood. To test this, we performed a limiting dilution transplant from adult BM (Fig. 7E). Indeed, the number of predicted HSCs was significantly reduced from ~1700 HSCs in WT BM to ~570 HSCs in *Gata2*<sup>R396Q/+</sup> BM (Fig. 7F). These data show that the *Gata2*<sup>R396Q/+</sup> mutation results in a reduction in definitive HSC



**Fig. 4** *Gata2*<sup>R396Q/+</sup> hematopoietic progenitors exhibit an accelerated aging phenotype. UMAP projections and GSEA analysis of HSC 1/HSC 2 **A**, Int Prog 1/ Lymph Prog **B**, and Int Prog 2/3/4 **C** clusters, showing terms negatively and positively enriched in *Gata2*<sup>R396Q/+</sup> populations. **D** GSEA plots showing positive enrichment for Marrow HSC Aging in *Gata2*<sup>R396Q/+</sup> cells from the indicated clusters. **E** Schematic for LPS in vivo treatment. **F** Frequency of LT-HSC cells normalized to PBS controls 5 days post in vivo LPS treatment of WT and *Gata2*<sup>R396Q/+</sup> 8–10 week old mice ( $n = 4-5$  per condition). Terms have an FDR < 0.05 and Gene % > 15%. ns not significant, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . HSC hematopoietic stem cell, Lymph lymphoid, Prog progenitor, Int intermediate, LT-HSC long-term HSC, PBS phosphate buffered saline, LPS lipopolysaccharide.





**Fig. 5** *Gata2*<sup>R396Q/+</sup> hematopoietic progenitors upregulate *Gata2*. **A** FeaturePlots displaying the expression of *Ctsf* in WT and *Gata2*<sup>R396Q/+</sup> LSK cells. **B** FeaturePlots displaying the expression of *Ly6c1* in WT and *Gata2*<sup>R396Q/+</sup> LSK cells. **C** FeaturePlots displaying the expression of *Tgfb3* in WT and *Gata2*<sup>R396Q/+</sup> LSK cells. **D** FeaturePlot displaying the expression of *Gata2* in WT and *Gata2*<sup>R396Q/+</sup> LSK cells. **E** Quantification and comparison of *Gata2* expression for the indicated clusters in WT and *Gata2*<sup>R396Q/+</sup> cells. **F** FeaturePlots displaying the expression of *Gata2*, *Dntt*, *Ctsf*, and *Procr* among Lymph Progenitor cells in WT and *Gata2*<sup>R396Q/+</sup> samples. Dotted circle indicates a unique population found among the *Gata2*<sup>R396Q/+</sup> Lymph Progenitor cluster. ns not significant, \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001; \*\*\*\**p* < 0.0001. HSC hematopoietic stem cell, Lymph lymphoid, Progenitor, Int intermediate.

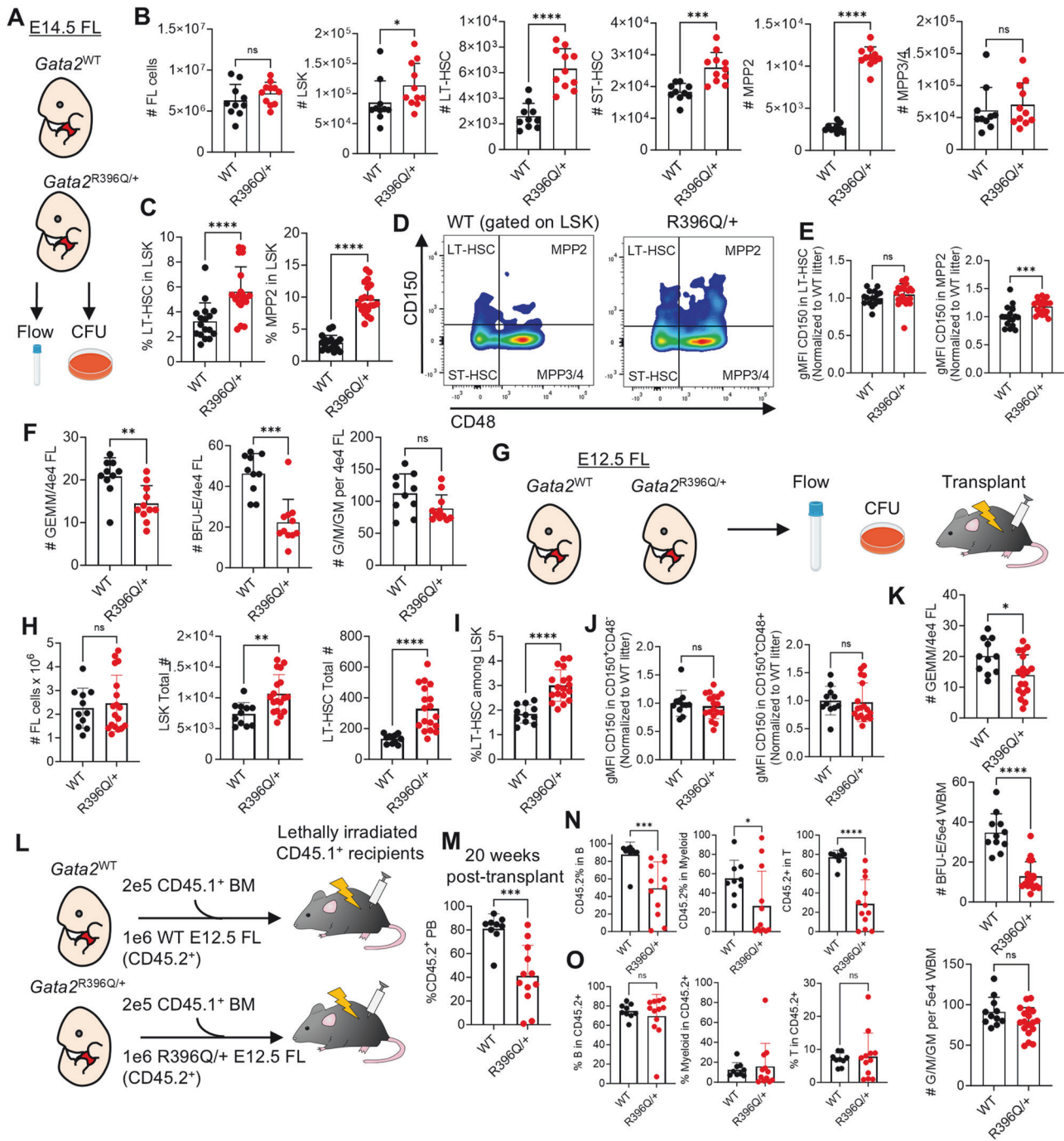
production at the earliest stage of hematopoietic development, resulting in downstream effects on the hematopoietic system.

## DISCUSSION

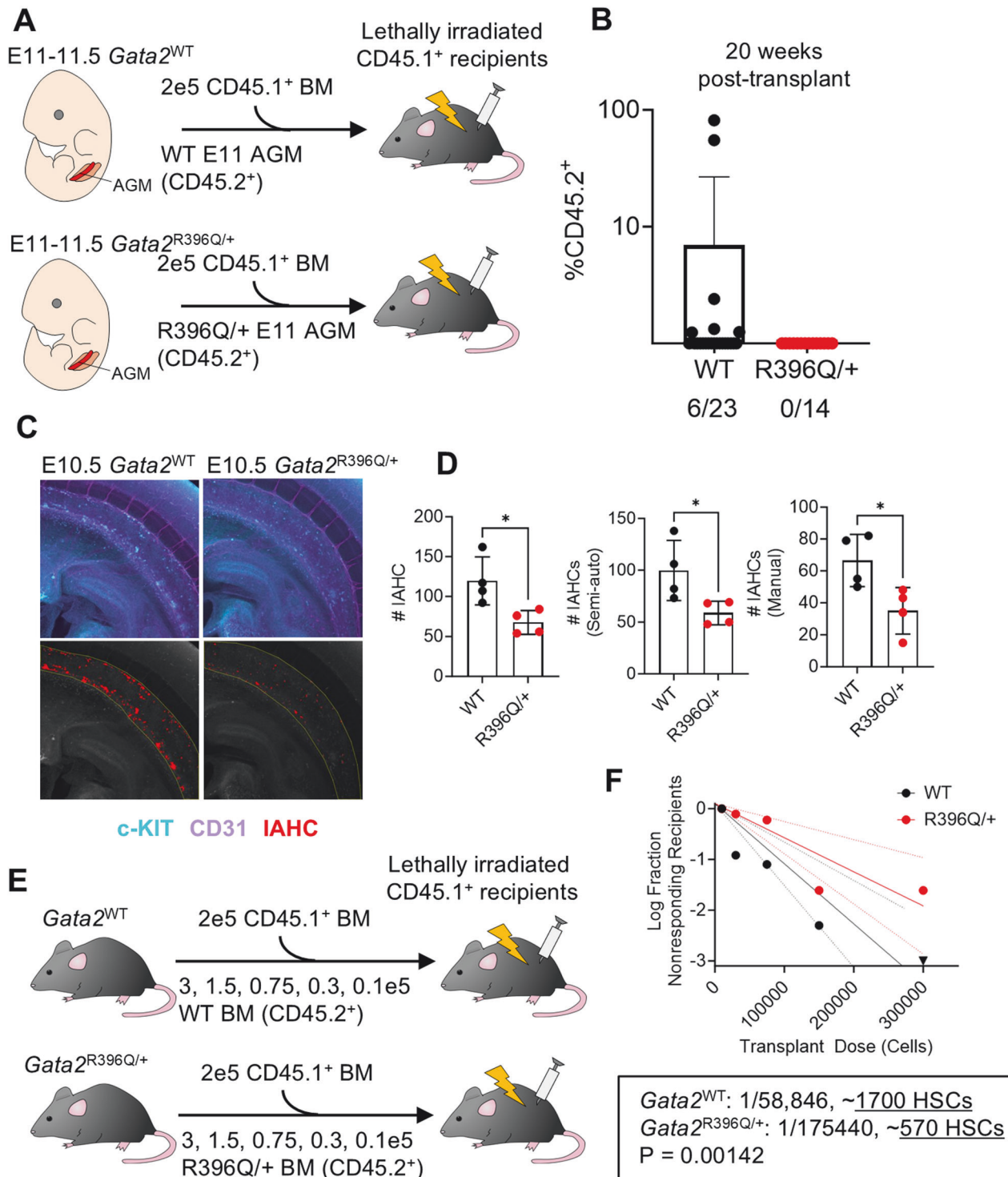
Here, we report the generation of a mouse model of the recurrent GATA2 deficiency mutation, *Gata2*<sup>R396Q/+</sup>. Through phenotypic, functional, and transcriptional analysis, we show that mice with the R396Q mutation generate fewer definitive HSCs that have reduced function at embryonic and adult timepoints, possibly due to an accelerated aging phenotype resulting in myeloid-bias and stem cell exhaustion (Fig. 8). These observations make the *Gata2*<sup>R396Q/+</sup> mouse an intriguing model for leukemic development in GATA2 deficiency. For instance, aberrant hematopoiesis begins *in utero* in our model, which will allow for modeling of leukemic progression during developmental stages more characteristic of human disease, such as early postnatal and adolescence. The decrease in definitive HSC production and lifelong numbers of HSCs as shown by our limiting dilution assay suggests that the clonal complexity of the hematopoietic system may be reduced. While this will need to be further validated experimentally through clonal tracing studies, reduced clonal complexity may provide an impetus for accelerated progression after the appropriate secondary events are applied to *Gata2*<sup>R396Q/+</sup> cells. Likewise, an aged hematopoietic compartment

would be more susceptible to transformation. Finally, the enrichment of mitochondrial activity terms in *Gata2*<sup>R396Q/+</sup> progenitors is similar to observations seen in transgenic *Erg/Gata2*<sup>+/-</sup> and *Erg/Gata2*<sup>L395V/+</sup> mouse models [61], which exhibited accelerated AML. Our future studies will focus on modeling leukemic development by adding putative drivers of malignancy in this disease, such as loss of orthologous chromosome 7 genes or mutations in *Asx1*, *Setbp1*, and *Stag2* [23], followed by identifying therapeutic strategies to ameliorate disease progression.

Apart from modeling transformation, many questions remain about the role of the R396Q mutation in preparing the hematopoietic system for transformation. Our scRNA-seq data suggests that *Gata2* is upregulated in several BM HSPC populations of *Gata2*<sup>R396Q/+</sup> mice. Therefore, the hematologic abnormalities seen in our model are not due to a lack of *Gata2* expression as in heterozygous KO models. It will be important to determine if there are any changes in DNA binding ability or targets of the R396Q variant *in vivo* that can contribute to a pre-leukemic state. Indeed, it is known that upregulation of GATA2 in humans is found in hematologic malignancies [62]. Isolation of specific populations of *Gata2*<sup>hi</sup> cells from *Gata2*<sup>R396Q/+</sup> HSPCs, such as the Lymph Progenitor population, will also be useful in answering questions about the cell-of-origin in this disease. Finally, we have only assessed the hematopoietic compartment in this study, but *Gata2* is also

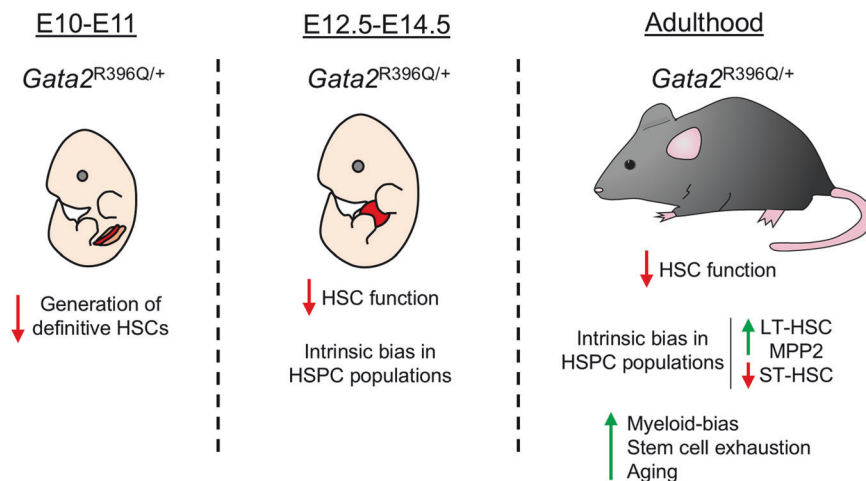


**Fig. 6** *Gata2*<sup>R396Q/+</sup> impairs fetal liver hematopoietic development and function. **A** Schematic of experiments utilizing WT and *Gata2*<sup>R396Q/+</sup> E14.5 fetal liver (FL). **B** Absolute numbers of total FL, LSK, and hematopoietic stem and progenitor (HSPC) cells from WT and *Gata2*<sup>R396Q/+</sup> littermates at E14.5 (*n* = 10–11 embryos). **C** Percentage of LT-HSC or MPP2 among LSK cells in E14.5 FL (*n* = 17–20 embryos). **D** Representative flow plot and normalized quantification of gMFI CD150 among WT and *Gata2*<sup>R396Q/+</sup> FL HSPCs at E14.5 (*n* = 17–20 embryos). **E** Number of CFU subtypes from E14.5 FL in WT or *Gata2*<sup>R396Q/+</sup> embryos (*n* = 10–12 embryos). **F** Schematic of experiments utilizing WT and *Gata2*<sup>R396Q/+</sup> E12.5 FL. **G** Absolute numbers of total FL, LSK, and LT-HSC in E12.5 FL (*n* = 11–18 embryos). **H** Percentage of LT-HSC among LSK cells in E12.5 FL (*n* = 11–18 embryos). **I** Normalized quantification of gMFI CD150 among WT and *Gata2*<sup>R396Q/+</sup> FL HSPCs at E12.5 (*n* = 11–18 embryos). **J** Number of CFU subtypes from E12.5 FL in WT or *Gata2*<sup>R396Q/+</sup> embryos (*n* = 11–18 embryos). **K** Schematic of transplantation of E12.5 FL cells. **L** Peripheral blood CD45.2<sup>+</sup> chimerism in recipients of E12.5 FL (*n* = 9–12 recipients). **M** Lineage output from CD45.2<sup>+</sup> cells (*n* = 9–12 recipients). **N** Lineage output from CD45.2<sup>+</sup> cells (*n* = 9–12 recipients). **O** Lineage output from CD45.2<sup>+</sup> cells (*n* = 9–12 recipients). **P** Lineage output from CD45.2<sup>+</sup> cells (*n* = 9–12 recipients). ns not significant, \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001; \*\*\*\**p* < 0.0001. E14.5, embryonic day 14.5; FL fetal liver, CFU colony-forming unit, LSK lineage Sca-1<sup>+</sup>c-Kit<sup>+</sup>, LT-HSC long-term hematopoietic stem cell, ST-HSC short-term HSC, MPP2 multipotent progenitor 2, gMFI geometric mean fluorescence intensity, GEMM granulocyte/erythrocyte/monocyte/megakaryocyte, BFU-E burst-forming unit-erythroid, G/M/GM granulocyte/monocyte.



**Fig. 7** *Gata2*<sup>R396Q/+</sup> negatively affects definitive HSC production. **A** Schematic for transplantation of E11 AGM from WT or *Gata2*<sup>R396Q/+</sup> into irradiated recipients. **B** Peripheral blood CD45.2<sup>+</sup> chimerism in recipients of AGM from E11-E11.5 (42-48 somite pairs (sp)) WT or *Gata2*<sup>R396Q/+</sup> embryos. Adequate reconstitution defined as  $\geq 1\%$  CD45.2<sup>+</sup> ( $n = 14-23$  recipients). **C** Representative imaging of intra-aortic hematopoietic clusters (IAHC) in E10.5 AGM (35-37 sp) of WT or *Gata2*<sup>R396Q/+</sup> embryos (blue, c-KIT; purple, CD31; red, IAHC). **D** Quantification of IAHCs in "C" using automatic, semi-automatic, or manual methods ( $n = 4$  embryos per condition). **E** Schematic of limiting dilution transplant of whole BM from WT or *Gata2*<sup>R396Q/+</sup> mice, 8-10 weeks old. **F** Limiting dilution analysis plot ( $n = 6-10$  recipients per dilution). Analysis performed using Extreme Limiting Dilution Analysis (ELDA) [65]. \* $p < 0.05$ . E11 embryonic day 11, IAHC intra-aortic hematopoietic cluster.





**Fig. 8** *Gata2*<sup>R396Q/+</sup> disrupts normal hematopoiesis. During HSC generation at around embryonic day 10 (E10), *Gata2*<sup>R396Q/+</sup> mice produce fewer definitive HSCs in the aorta gonad mesonephros (AGM). HSCs show reduced functional potential in the fetal liver at E12.5-E14.5, as well as intrinsic biases toward particular HSPC populations. During adulthood, the continued effects of dysregulated HSC generation manifest as reduced HSC function, biases toward LT-HSC and MPP2 production at the expense of ST-HSC production (driven by enhanced CD150 expression), myeloid bias, and suggestions of stem cell exhaustion/aging.

expressed in the BM niche—especially endothelial cells [63]. Given the role of inflammation in MDS/AML progression [64], future studies will interrogate the niche components of *Gata2*<sup>R396Q/+</sup> mice for potential contributions to a pre-leukemic state.

#### DATA AVAILABILITY

RNA-Seq data have been deposited to Gene Expression Omnibus (GSE275213). The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### REFERENCES

1. Tsai F-Y, Keller G, Kuo FC, Weiss M, Chen J, Rosenblatt M, et al. An early hematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371:221–6.
2. Hahn CN, Chong CE, Carmichael CL, Wilkins EJ, Brautigan PJ, Li XC, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet*. 2011;43:1012–7.
3. Bigley V, Haniffa M, Doulatov S, Wang XN, Dickinson R, McGovern N, et al. The human syndrome of dendritic cell, monocyte, B and NK lymphoid deficiency. *J Exp Med*. 2011;208:227–34.
4. Mansour S, Connell F, Steward C, Ostergaard P, Brice G, Smithson S, et al. Emberger syndrome—primary lymphedema with myelodysplasia: report of seven new cases. *Am J Med Genet A*. 2010;152A:2287–96.
5. Vinh DC, Patel SY, Uzel G, Anderson VL, Freeman AF, Olivier KN, et al. Autosomal dominant and sporadic monocytopenia with susceptibility to mycobacteria, fungi, papillomaviruses, and myelodysplasia. *Blood*. 2010;115:1519–29.
6. Dickinson RE, Griffin H, Bigley V, Reynard LN, Hussain R, Haniffa M, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood*. 2011;118:2656–8.
7. Hsu AP, Sampaio EP, Khan J, Calvo KR, Lemieux JE, Patel SY, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*. 2011;118:2653–5.
8. Ostergaard P, Simpson MA, Connell FC, Steward CG, Brice G, Woollard WJ, et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat Genet*. 2011;43:929–31.
9. Mace EM, Hsu AP, Monaco-Shawver L, Makedonas G, Rosen JB, Dropulic L, et al. Mutations in GATA2 cause human NK cell deficiency with specific loss of the CD56(bright) subset. *Blood*. 2013;121:2669–77.
10. Pasquet M, Bellanne-Chantelot C, Tavittian S, Prade N, Beaupain B, Larochelle O, et al. High frequency of GATA2 mutations in patients with mild chronic neutropenia evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. *Blood*. 2013;121:822–9.
11. Ganapathi KA, Townsley DM, Hsu AP, Arthur DC, Zerbe CS, Cuellar-Rodriguez J, et al. GATA2 deficiency-associated bone marrow disorder differs from idiopathic aplastic anemia. *Blood*. 2015;125:56–70.
12. Koegel AK, Hofmann I, Moffitt K, Degar B, Duncan C, Tubman VN. Acute lymphoblastic leukemia in a patient with MonoMAC syndrome/GATA2 haploinsufficiency. *Pediatr Blood Cancer*. 2016;63:1844–7.
13. Novakova M, Zaliouva M, Sukova M, Wlodarski M, Janda A, Fronkova E, et al. Loss of B cells and their precursors is the most constant feature of GATA-2 deficiency in childhood myelodysplastic syndrome. *Haematologica*. 2016;101:707–16.
14. Esparza O, Xavier AC, Atkinson TP, Hill BC, Whelan K. A unique phenotype of T-cell acute lymphoblastic leukemia in a patient with GATA2 haploinsufficiency. *Pediatr Blood Cancer*. 2019;66:e27649.
15. Khoury JD, Solary E, Abela O, Akkari Y, Alaggio R, Apperley JF, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. *Leukemia*. 2022;36:1703–19.
16. Sahoo SS, Kozyra EJ, Wlodarski MW. Germline predisposition in myeloid neoplasms: Unique genetic and clinical features of GATA2 deficiency and SAMD9/SAMD9L syndromes. *Best Pract Res Clin Haematol*. 2020;33:101197.
17. Wlodarski MW, Hirabayashi S, Pastor V, Stary J, Hasle H, Masetti R, et al. Prevalence, clinical characteristics, and prognosis of GATA2-related myelodysplastic syndromes in children and adolescents. *Blood*. 2016;127:1387–97.
18. Donadieu J, Lamant M, Fieschi C, de Fontbrune FS, Caye A, Ouachee M, et al. Natural history of GATA2 deficiency in a survey of 79 French and Belgian patients. *Haematologica*. 2018;103:1278–87.
19. Brown AL, Hahn CN, Scott HS. Secondary leukemia in patients with germline transcription factor mutations (RUNX1, GATA2, CEBPA). *Blood*. 2020;136:24–35.
20. Kotmayer L, Romero-Moya D, Marin-Bejar O, Kozyra E, Catala A, Bigas A, et al. GATA2 deficiency and MDS/AML: Experimental strategies for disease modelling and future therapeutic prospects. *Br J Haematol*. 2022;199:482–95.
21. Calvo KR, Hickstein DD. The spectrum of GATA2 deficiency syndrome. *Blood*. 2023;141:1524–32.
22. Chong CE, Venugopal P, Stokes PH, Lee YK, Brautigan PJ, Yeung DTO, et al. Differential effects on gene transcription and hematopoietic differentiation correlate with GATA2 mutant disease phenotypes. *Leukemia*. 2018;32:194–202.
23. Largeaud L, Collin M, Monselet N, Vergez F, Fregona V, Larcher L, et al. Somatic genetic alterations predict hematological progression in GATA2 deficiency. *Haematologica*. 2023;108:1515–29.
24. Bodor C, Renneville A, Smith M, Charazac A, Iqbal S, Etancelin P, et al. Germ-line GATA2 p.THR354MET mutation in familial myelodysplastic syndrome with acquired monosomy 7 and ASXL1 mutation demonstrating rapid onset and poor survival. *Haematologica*. 2012;97:890–4.
25. West RR, Hsu AP, Holland SM, Cuellar-Rodriguez J, Hickstein DD. Acquired ASXL1 mutations are common in patients with inherited GATA2 mutations and correlate with myeloid transformation. *Haematologica*. 2014;99:276–81.

26. Pastor V, Hirabayashi S, Karow A, Wehrle J, Kozyra EJ, Nienhold R, et al. Mutational landscape in children with myelodysplastic syndromes is distinct from adults: specific somatic drivers and novel germline variants. *Leukemia*. 2017;31:759–62.
27. McReynolds LJ, Yang Y, Yuen Wong H, Tang J, Zhang Y, Mule MP, et al. MDS-associated mutations in germline GATA2 mutated patients with hematologic manifestations. *Leuk Res*. 2019;76:70–5.
28. West RR, Calvo KR, Embree LJ, Wang W, Tuschong LM, Bauer TR, et al. ASXL1 and STAG2 are common mutations in GATA2 deficiency patients with bone marrow disease and myelodysplastic syndrome. *Blood Adv*. 2022;6:793–807.
29. Tsai F-Y, Orkin SH. 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*. 1997;89:3636–43.
30. Rodrigues NP, Janzen V, Forkert R, Dombkowski DM, Boyd AS, Orkin SH, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*. 2005;106:477–84.
31. Eich C, Arlt J, Vink CS, Solaimani Kartalaei P, Kaimakis P, Mariani SA, et al. In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. *J Exp Med*. 2018;215:233–48.
32. Gioacchino E, Koyunlar C, Zink J, de Looper H, de Jong M, Dobrzycki T, et al. Essential role for Gata2 in modulating lineage output from hematopoietic stem cells in zebrafish. *Blood Adv*. 2021;5:2687–700.
33. Takai J, Shimada T, Nakamura T, Engel JD, Moriguchi T. Gata2 heterozygous mutant mice exhibit reduced inflammatory responses and impaired bacterial clearance. *iScience*. 2021;24:102836.
34. Gioacchino E, Zhang W, Koyunlar C, Zink J, de Looper H, Gussinklo KJ, et al. GATA2 heterozygosity causes an epigenetic feedback mechanism resulting in myeloid and erythroid dysplasia. *Br J Haematol*. 2024;205:580–593.
35. Johnson KD, Hsu AP, Ryu MJ, Wang J, Gao X, Boyer ME, et al. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J Clin Invest*. 2012;122:3692–704.
36. Gao X, Johnson KD, Chang YI, Boyer ME, Dewey CN, Zhang J, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med*. 2013;210:2833–42.
37. Zhang S-J, Ma L-Y, Huang Q-H, Li G, Gu B-W, Gao X-D, et al. Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci USA*. 2008;105:2076–81.
38. Fu YK, Tan Y, Wu B, Dai YT, Xu XG, Pan MM, et al. Gata2-L359V impairs primitive and definitive hematopoiesis and blocks cell differentiation in murine chronic myelogenous leukemia model. *Cell Death Dis*. 2021;12:568.
39. Hasegawa A, Hayasaka Y, Morita M, Takenaka Y, Hosaka Y, Hirano I, et al. Heterozygous variants in GATA2 contribute to DCML deficiency in mice by disrupting tandem protein binding. *Commun Biol*. 2022;5:376.
40. Jung M, Cordes S, Zou J, Yu SJ, Guitart X, Hong SG, et al. GATA2 deficiency and human hematopoietic development modeled using induced pluripotent stem cells. *Blood Adv*. 2018;2:3553–65.
41. Castano J, Romero-Moya D, Richaud-Patin Y, Giorgetti A. Generation of two heterozygous GATA2 CRISPR/Cas9-edited iPSC lines, R398W and R396Q, for modeling GATA2 deficiency. *Stem Cell Res*. 2021;55:102445.
42. Zheng M, Karki R, Kancharana B, Berns H, Pruett-Miller SM, Kanneganti TD. Caspase-6 promotes activation of the caspase-11-NLRP3 inflammasome during gram-negative bacterial infections. *J Biol Chem*. 2021;297:101379.
43. Narina S, Connelly JP, Pruett-Miller SM. High-throughput analysis of CRISPR-Cas9 editing outcomes in cell and animal models using CRISpy. *Methods Mol Biol*. 2023;2631:155–82.
44. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16:1289–96.
45. Hamey FK, Gottgens B. Machine learning predicts putative hematopoietic stem cells within large single-cell transcriptomics data sets. *Exp Hematol*. 2019;78:11–20.
46. Pietras EM, Reynaud D, Kang YA, Carlin D, Calero-Nieto FJ, Leavitt AD, et al. Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell*. 2015;17:35–46.
47. Beerman I, Bhattacharya D, Zandi S, Sigvardsson M, Weissman IL, Bryder D, et al. Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc Natl Acad Sci USA*. 2010;107:5465–70.
48. Kucinski I, Campos J, Barile M, Severi F, Bohin N, Moreira PN, et al. A time- and single-cell-resolved model of murine bone marrow hematopoiesis. *Cell Stem Cell*. 2024;31:244–59.e10.
49. Cabezas-Wallscheid N, Klimmeck D, Hansson J, Lipka DB, Reyes A, Wang Q, et al. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell*. 2014;15:507–22.
50. Wilson NK, Kent DG, Buettner F, Shehata M, Macaulay IC, Calero-Nieto FJ, et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell*. 2015;16:712–24.
51. Nestorowa S, Hamey FK, Pijuan-Sala B, Diamanti E, Shepherd M, Laurenti E, et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*. 2016;128:e20–31.
52. Dahlin JS, Hamey FK, Pijuan-Sala B, Shepherd M, Lau WWY, Nestorowa S, et al. A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood*. 2018;131:e1–e11.
53. Baccin C, Al-Sabah J, Velten L, Helbling PM, Grunschlag F, Hernandez-Malmierca P, et al. Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat Cell Biol*. 2020;22:38–48.
54. Choi J, Baldwin TM, Wong M, Bolden JE, Fairfax KA, Lucas EC, et al. Haemopedia RNA-seq: a database of gene expression during haematopoiesis in mice and humans. *Nucleic Acids Res*. 2019;47:D780–D5.
55. Morganti C, Ito K. Mitochondrial contributions to hematopoietic stem cell aging. *Int J Mol Sci*. 2021;22:1117.
56. Dykstra B, Olthof S, Schreuder J, Ritsema M, de Haan G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J Exp Med*. 2011;208:2691–703.
57. Ganuza M, Hall T, Finkelstein D, Wang YD, Chabot A, Kang G, et al. The global clonal complexity of the murine blood system declines throughout life and after serial transplantation. *Blood*. 2019;133:1927–42.
58. Ito K, Hirao A, Arai F, Takubo K, Matsuoka S, Miyamoto K, et al. Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nat Med*. 2006;12:446–51.
59. Bataller A, Montalban-Bravo G, Soltysiak KA, Garcia-Manero G. The role of TGFbeta in hematopoiesis and myeloid disorders. *Leukemia*. 2019;33:1076–89.
60. Medvinsky A, Rybtsov S, Taoudi S. Embryonic origin of the adult hematopoietic system: advances and questions. *Development*. 2011;138:1017–31.
61. Rein A, Geron I, Kugler E, Fishman H, Gottlieb E, Abramovich I, et al. Cellular and metabolic characteristics of pre-leukemic hematopoietic progenitors with GATA2 haploinsufficiency. *Haematologica*. 2023;108:2316–30.
62. Vicente C, Vazquez I, Conchillo A, Garcia-Sanchez MA, Marcotegui N, Fuster O, et al. Overexpression of GATA2 predicts an adverse prognosis for patients with acute myeloid leukemia and it is associated with distinct molecular abnormalities. *Leukemia*. 2012;26:550–4.
63. Kanki Y, Kohro T, Jiang S, Tsutsumi S, Mimura I, Suehiro J, et al. Epigenetically coordinated GATA2 binding is necessary for endothelium-specific endomucin expression. *EMBO J*. 2011;30:2582–95.
64. Craver BM, El Alaoui K, Scherber RM, Fleischman AG. The critical role of inflammation in the pathogenesis and progression of myeloid malignancies. *Cancers*. 2018;10:104.
65. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods*. 2009;347:70–8.

## ACKNOWLEDGEMENTS

We would like to thank the Flow Cytometry team in the St. Jude Department of Hematology for their assistance in cell sorting for several of the experiments described in this manuscript. This research was supported by the National Institutes of Health (R35CA253096), the Samuel Waxman Cancer Research Foundation, and the St. Jude/American Lebanese Syrian Associated Charities (ALSAC) to J.D.C. The Center for Advanced Genome Engineering, Hartwell Center, Protein Production Shared Resource, and the Flow Cytometry and Cell Sorting Shared Resource at St. Jude are supported in part by P30 CA021765. T.H. is supported by a career development award from the Leukemia and Lymphoma Society. D.SdB is supported by a fellowship from the American Heart Association (23POST1011028). W.K.C. is supported by the National Institutes of Health (R01DK113973). S.I. is an Israel Cancer Research Foundation Professor. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

## AUTHOR CONTRIBUTIONS

T.H. performed the experiments, analyzed the data, and contributed to writing the manuscript. R.M. provided bioinformatic analysis and contributed to writing the manuscript. D.SdB. and A.C. performed experiments and contributed to writing the manuscript. J.K. and S.M.P.-M. created the mouse model and contributed to writing the manuscript. S.I., W.K.C., and J.D.C. conceived the experiments, analyzed the data, and contributed to writing the manuscript.

## COMPETING INTERESTS

J.D.C. receives research support from Syndax, consulting fees from Cellarity, and is a scientific advisor for Alethiomics.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41375-024-02508-z>.

**Correspondence** and requests for materials should be addressed to John D. Crispino.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024