

# Pathogenic human variant that dislocates GATA2 zinc fingers disrupts hematopoietic gene expression and signaling networks

Mabel Minji Jung, ... , Sunduz Keles, Emery H. Bresnick

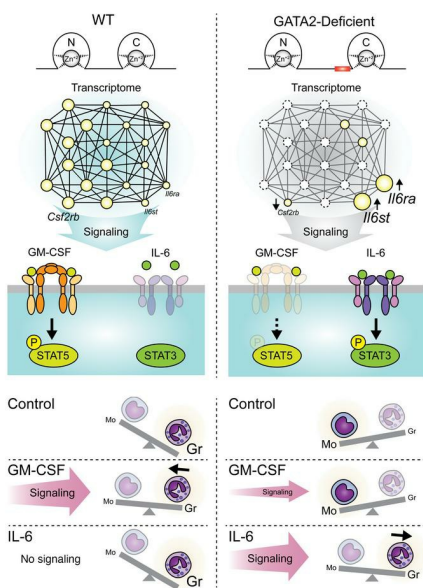
*J Clin Invest.* 2023;133(7):e162685. <https://doi.org/10.1172/JCI162685>.

Research Article

Genetics

Hematology

## Graphical abstract



Find the latest version:

<https://jci.me/162685/pdf>



# Pathogenic human variant that dislocates GATA2 zinc fingers disrupts hematopoietic gene expression and signaling networks

Mabel Minji Jung,<sup>1</sup> Siqi Shen,<sup>2</sup> Giovanni A. Botten,<sup>3</sup> Thomas Olender,<sup>4</sup> Koichi R. Katsumura,<sup>1</sup> Kirby D. Johnson,<sup>1</sup> Alexandra A. Soukup,<sup>1</sup> Peng Liu,<sup>2</sup> Qingzhou Zhang,<sup>4</sup> Zena D. Jensvold,<sup>5</sup> Peter W. Lewis,<sup>5</sup> Robert A. Beagrie,<sup>6</sup> Jason K.K. Low,<sup>7</sup> Lihua Yang,<sup>7</sup> Joel P. Mackay,<sup>7</sup> Lucy A. Godley,<sup>8</sup> Marjorie Brand,<sup>9</sup> Jian Xu,<sup>3</sup> Sunduz Keles,<sup>2</sup> and Emery H. Bresnick<sup>1</sup>

<sup>1</sup>Wisconsin Blood Cancer Research Institute, Department of Cell and Regenerative Biology, Carbone Cancer Center, and <sup>2</sup>Department of Biostatistics and Biomedical Informatics, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA. <sup>3</sup>Children's Medical Center Research Institute, Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, Texas, USA. <sup>4</sup>Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute—General Hospital, Ottawa, Ontario, Canada. <sup>5</sup>Department of Biomolecular Chemistry, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA. <sup>6</sup>MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom. <sup>7</sup>School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia. <sup>8</sup>Section of Hematology/Oncology, The University of Chicago, Chicago, Illinois, USA. <sup>9</sup>Department of Cell and Regenerative Biology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA.

Although certain human genetic variants are conspicuously loss of function, decoding the impact of many variants is challenging. Previously, we described a patient with leukemia predisposition syndrome (GATA2 deficiency) with a germline GATA2 variant that inserts 9 amino acids between the 2 zinc fingers (9aa-Ins). Here, we conducted mechanistic analyses using genomic technologies and a genetic rescue system with *Gata2* enhancer–mutant hematopoietic progenitor cells to compare how GATA2 and 9aa-Ins function genome-wide. Despite nuclear localization, 9aa-Ins was severely defective in occupying and remodeling chromatin and regulating transcription. Variation of the inter-zinc finger spacer length revealed that insertions were more deleterious to activation than repression. GATA2 deficiency generated a lineage-diverting gene expression program and a hematopoiesis-disrupting signaling network in progenitors with reduced granulocyte-macrophage colony-stimulating factor (GM-CSF) and elevated IL-6 signaling. As insufficient GM-CSF signaling caused pulmonary alveolar proteinosis and excessive IL-6 signaling promoted bone marrow failure and GATA2 deficiency patient phenotypes, these results provide insight into mechanisms underlying GATA2-linked pathologies.

## Introduction

In metal-binding proteins, the metal requirement to regulate protein structure/function links mechanisms governing metal homeostasis, proteome composition, and genome regulation. Hundreds of proteins harbor zinc “fingers” that coordinate zinc ions (1). These proteins include transcription factors, in which the finger mediates DNA and/or protein binding. Sequences within one or more fingers confer sequence-specific DNA binding (1). Artificial zinc finger proteins can be engineered with designer DNA binding specificities for experimental and therapeutic applications (2). Finger-altering genetic variation can change DNA binding specificity (3) and cause diseases (4–8). Sequences extrinsic to fingers can enhance DNA-binding affinity (9). Insertion of KTS residues between Wilms tumor 1 zinc fingers via alternative splicing alters DNA binding specificity (10). Unlike sophisticated zinc finger structure/function knowl-

edge, the impact of spacer length and sequence and alterations thereof cannot be predicted.

GATA transcription factors (11) harbor N- and C-zinc fingers, based on proximity to N- and C-termini (12, 13). In GATA1, which promotes erythroid, megakaryocytic, and mast cell development, the C-finger binds a small subset of WGATAR motifs (14, 15) in chromatin (16, 17). The N-finger binds the 9–zinc finger coregulator Friend of GATA1 (FOG1) (18, 19). Although FOG1 fingers do not appear to bind DNA, four are implicated in binding GATA1 (20). GATA2 promotes hematopoietic stem cell emergence in the embryo (21, 22) and hematopoiesis in adult mice (23, 24). Heterozygous GATA2 coding or enhancer germline mutations cause GATA2 deficiency syndrome involving immunodeficiency, myelodysplastic syndrome, and acute myeloid leukemia (25–29). Most mutations alter the GATA2 C-finger (30) and inhibit DNA binding (27, 31) and chromatin occupancy (32, 33). Although GATA2 mutational analyses revealed loss-of-function phenotypes that may underlie pathogenesis (30), GATA2 variants can partially retain or have supraphysiological activity (33, 34). GATA2 dysregulation is also implicated in myeloproliferative neoplasms (35).

We described a human variant that alters the distance between GATA2 zinc fingers. A family harbored a germline in-frame 9–amino acid insertion (9aa-Ins) between the N- and C-fingers

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Copyright:** © 2023, Jung et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

**Submitted:** June 14, 2022; **Accepted:** February 16, 2023; **Published:** April 3, 2023.

**Reference information:** *J Clin Invest.* 2023;133(7):e162685.

<https://doi.org/10.1172/JCI162685>.

(36). The 8-year-old proband had GATA2 deficiency with warts, lymphedema, and cytopenias, and the variant was detected in the asymptomatic mother. 9aa-Ins was defective in activating three GATA2 target genes tested. As GATA factor target genes are regulated via distinct mechanisms with context-dependent coregulator requirements (37), it is unpredictable whether 9aa-Ins would be defective in activating and repressing all target genes, whether defects reflect an inability to occupy and/or remodel chromatin, and whether it has activities distinct from GATA2 (ectopic). It is unclear whether 9aa-Ins defects would be recapitulated with any spacing alteration, whether select changes are tolerable, and whether a threshold exists beyond which alterations are inhibitory.

We analyzed how GATA2 and 9aa-Ins regulate chromatin and transcription genome-wide in a rescue system with GATA2-deficient progenitors. 9aa-Ins was defective in occupying and remodeling chromatin and regulating most target genes. GATA2 repressed expression of *Ebfl*, encoding a B lineage-regulatory transcription factor (38) and EBF1 target genes important in B cell biology, indicating a lineage-diverting expression program in GATA2-deficient fetal progenitors. *Csf2rb*, encoding the common  $\beta$  chain of granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin-3 (IL-3), and IL-5 receptors (39–41), exemplified a GATA2-activated gene compromised by the insertion, and this defect diminished GM-CSF signaling. In humans, defective GM-CSF signaling impairs the capacity of lung macrophages to consume surfactant, causing pulmonary alveolar proteinosis (42, 43), a severe GATA2 deficiency syndrome phenotype (44, 45). Genes encoding IL-6 receptor subunits (*Il6ra* and *Il6st*) (46, 47) were upregulated in GATA2-deficient cells, instigating IL-6 signaling, and 9aa-Ins did not repress expression. Ectopic IL-6 signaling drives immune dysregulation and bone marrow failure (48–50). Our analysis of a disease variant unveiled principles of GATA factor function and pathogenic mechanisms and will enable clinical genetic variant curation.

## Results

*GATA2 pathogenic variant with dislocated zinc fingers is severely defective genome-wide, yet retains activity at select loci.* The GATA2 pathogenic 9-aa insertion (p.A345delinsALLVAALLAA) lies within the spacer between zinc fingers proximal to the DNA-binding C-finger (Figure 1A). The GATA2 zinc finger and spacer sequences are conserved (Supplemental Figure 1A; supplemental material available online with this article; <https://doi.org/10.1172/JCI162685DS1>), and spacer sequence and length are conserved among the mammalian GATA factors. To test whether the 9aa insertion is uniquely deleterious, we generated retroviruses to express variants with reductions (8, 6, 4, and 2 aa) from the middle of 9aa-Ins (X = deleted aa; LLVAXLLAA, LLVXXXLAA, LLXXXXXAA, LXXXXXXXXA).

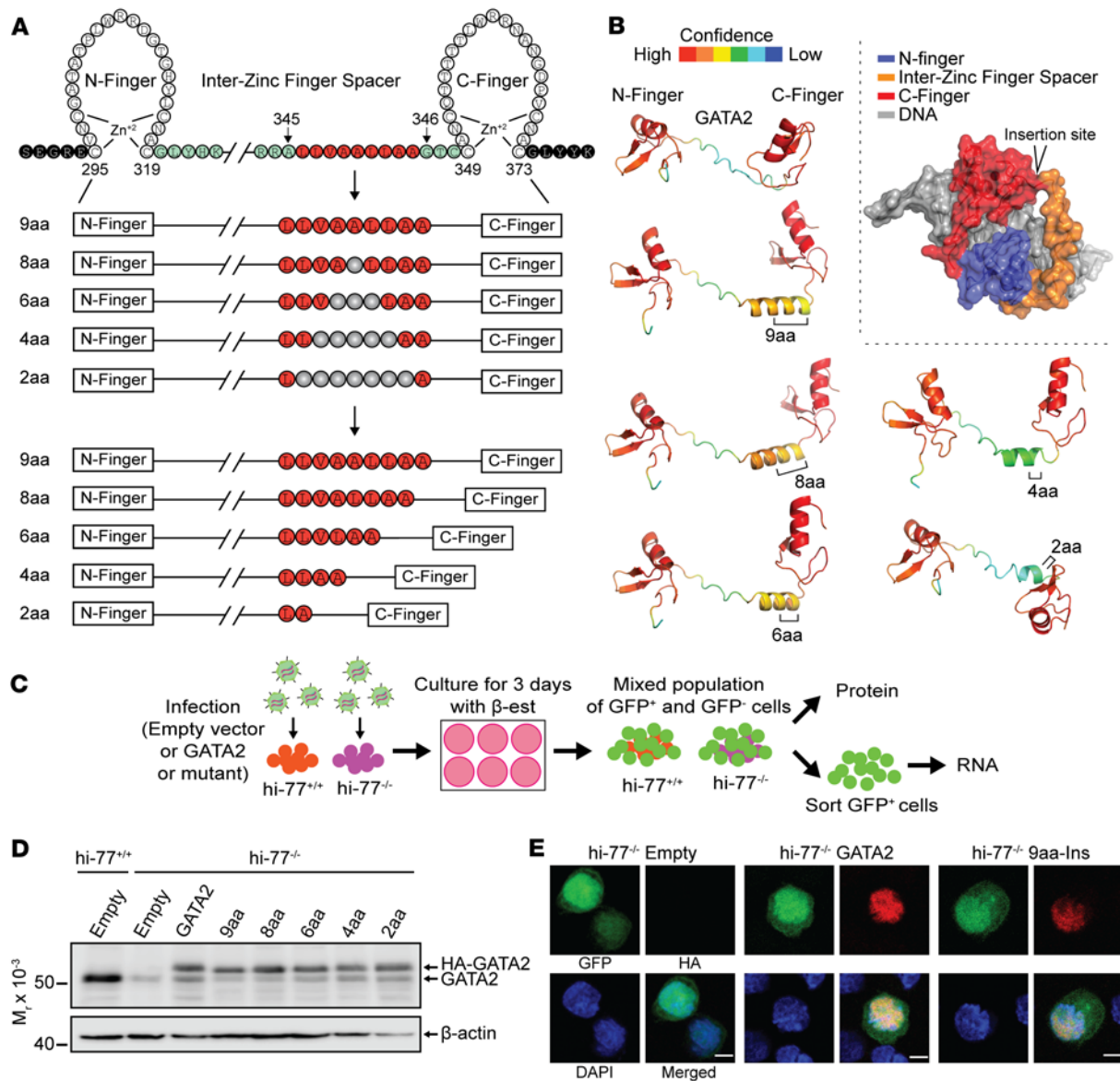
To assess whether the insertion alters the zinc finger domain conformation, we used the AlphaFold Protein Structure Database to predict structures of WT and 9aa-Ins variant domains (Figure 1B). The WT fingers display conformations similar to experimentally derived structures of other GATA zinc fingers (51–54). The spacer conformation is extended and predicted with low confidence, consistent with being disordered in solution and not discernible in crystal structures of GATA zinc finger proteins bound to DNA (51, 53). The 9aa insertion is predicted with intermediate confidence to form

an  $\alpha$  helix and induce 3 residues (RRA) from the WT spacer to adopt helicity (Figure 1B). This predicted helix was also observed with smaller insertion variants, with 8aa predicted to yield the longest helix and 2aa the smallest. Such a conformation might alter or generate GATA2-dependent macromolecular interactions (Figure 1B).

Given the invariant spacer, spacer length or sequence deviations might destabilize GATA2. We tested whether variants can be expressed in progenitors. We adapted our rescue assay with *Gata2*-77<sup>-/-</sup> primary fetal liver progenitors (33, 36) to HoxB8-immortalized mutant progenitors (hi-77<sup>-/-</sup>) (55) (Figure 1C). WT and mutant hi-77<sup>-/-</sup> cells were infected with control (empty) GFP-expressing retrovirus or retroviruses expressing GFP and HA-GATA2 or a variant. After culturing for 3 days, the population containing GFP<sup>+</sup> cells was analyzed by Western blotting (Figure 1D and Supplemental Figure 1B). The mean infection efficiency quantified by flow cytometry was 27% GFP<sup>+</sup> (Supplemental Figure 1C). Endogenous GATA2 mRNA and protein expression was approximately 75% lower in hi-77<sup>-/-</sup> versus hi-77<sup>+/+</sup> progenitors, consistent with -77 function to increase *Gata2* transcription in progenitors (55–57). As the HA tag reduced GATA2 mobility, anti-GATA2 antibody detected endogenous and exogenous GATA2. HA-tagged 9, 8, 6, 4, and 2aa variant levels resembled HA-GATA2 (Figure 1D) and were not destabilizing. Endogenous GATA2 levels in hi-77<sup>+/+</sup> cells resembled HA-GATA2 or variant levels in hi-77<sup>-/-</sup> (Supplemental Figure 1D). Using anti-HA to detect GATA2 or 9aa-Ins after expression in hi-77<sup>-/-</sup> cells and GFP to identify infected hi-77<sup>-/-</sup> cells, anti-HA staining in confocal microscopy was nuclear for GATA2 and 9aa-Ins (Figure 1E), with no immunoreactivity detected in control cells.

Previously, we demonstrated that 9aa-Ins is defective in activating 3 GATA2 target genes and promoting myeloerythroid progenitor activity when expressed in -77<sup>-/-</sup> progenitors (36). As context-dependent GATA factor mechanisms preclude predictions at a given locus (11, 37), and genetic variation can create ectopic functions (58), the relationship between defective 9aa-Ins activity and genome-wide function was unclear. We tested whether 9aa-Ins was defective in activation and/or repression and/or acquires ectopic activity. With the rescue assay, we expressed GATA2 or 9aa-Ins, isolated GFP<sup>+</sup> cells, and used RNA-Seq to test models (Figure 2A). We compared differentially expressed genes (DEGs) between hi-77<sup>-/-</sup> empty and hi-77<sup>+/+</sup> empty, between hi-77<sup>-/-</sup> empty and hi-77<sup>-/-</sup> GATA2, and between hi-77<sup>-/-</sup> empty and hi-77<sup>-/-</sup> 9aa-Ins, and amalgamated data from 4 biological replicates quantified by RSEM (59). The DEGs ( $|\log_2(\text{fold change})| \geq 1$  and adjusted *P* value < 0.05) were parsed into: -77-regulated, 2,084 DEGs; HA-GATA2-regulated, 2,138 DEGs; 9aa-Ins-regulated, 939 DEGs (Figure 2B). Analysis of gene expression changes ( $|\log_2(\text{fold change})| > 0$  and no adjusted *P* value cutoff) resulting from -77 deletion and HA-GATA2 expression in the rescue assay demonstrated that HA-GATA2 rescued 66% (4,768 of 7,258) and 68% (4,541 of 6,724) of genes activated and repressed, respectively, by endogenous GATA2 (Supplemental Figure 2A). There was greater overlap between -77- and HA-GATA2-regulated genes versus -77- and 9aa-Ins-regulated genes (916 vs. 377).

-77-, HA-GATA2-, and 9aa-Ins-regulated genes were parsed into activated and repressed (Figure 2B). The analysis revealed 2,138 GATA2-regulated genes, 525 GATA2- and 9aa-Ins-regulated genes, and 414 genes solely regulated by 9aa-Ins (ectopic) (Figure 2B). As GATA2 and 9aa-Ins regulated 525 genes, 9aa-Ins retained



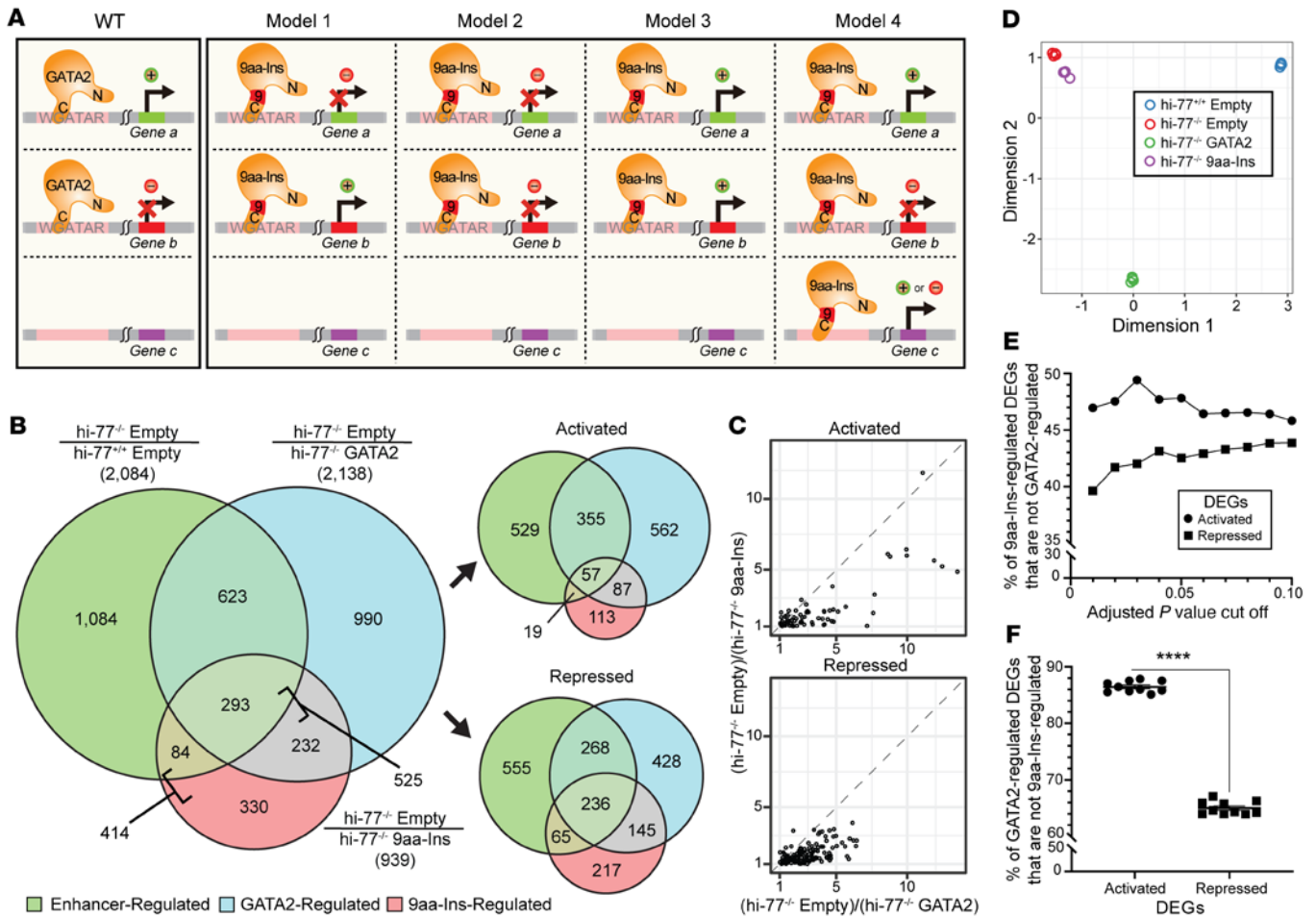
**Figure 1. Artificial GATA2 transcription factors with variable inter-zinc finger spacers.** (A) Mouse GATA2 9aa, 8aa, 6aa, 4aa, and 2aa insertion variants. The 9aa-Ins spacer variant models a human disease mutation (36). (B) AlphaFold prediction of structures. Model of GATA2 insertion site relative to the C-finger and N-finger. (C) Rescue assay with HoxB8-immortalized *Gata2* hi-77<sup>+/+</sup> and mutant (hi-77<sup>-/-</sup>) cells. β-est, β-estradiol. (D) Western blot with anti-GATA2 antibody of hi-77 cells expressing endogenous GATA2 with or without HA-tagged GATA2 or variants (*n* = 9). (E) Immunofluorescence analysis of HA-GATA2 localization in hi-77 cells. Scale bars: 5 μm.

activity to regulate a minority of GATA2-regulated genes (Supplemental Table 1). To identify 9aa-Ins-regulated genes, we excluded genes with low expression (transcripts per million [TPM] <1 in any replicate of hi-77<sup>-/-</sup> GATA2 for activation; hi-77<sup>-/-</sup> empty for repression) and compared fold change of hi-77<sup>-/-</sup> empty versus hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> empty versus hi-77<sup>-/-</sup> 9aa-Ins (Figure 2C). Of 215 genes, 79 were activated and 136 were repressed at 65% and 56%, respectively, of the level conferred by GATA2 (Figure 2C and Supplemental Figure 2B). Principal component analysis (PCA) demonstrated high data reproducibility (Figure 2D).

Since ectopic transcription factor activities may be significant (Figure 2A, model 4; and Supplemental Table 2), we tested whether the ectopic activity at 414 genes versus normal GATA2

activity was detected with *P* values from 0.01 to 0.1. As 47% of ectopically activated genes were retained at *P* = 0.1, and 40% of ectopically repressed genes at *P* = 0.01 (Figure 2E), ectopically 9aa-Ins-regulated genes emerged irrespective of stringency. Although GATA2-dependent repression is not understood, a similar number of genes were GATA2-activated (1,061 genes) and -repressed (1,077 genes), resembling GATA1 in erythroid cells (60, 61). GATA1-regulated genes, e.g., heme biosynthetic enzymes, hemoglobin subunits, and cytoskeletal components (16, 61–66), were not GATA2-regulated. One hundred forty-four of the 1,061 GATA2-activated genes (14%) were 9aa-Ins-activated, and 381 of 1,077 GATA2-repressed genes (35%) were 9aa-Ins-repressed. Analysis with *P* values from 0.01 to 0.1 revealed that the insertion



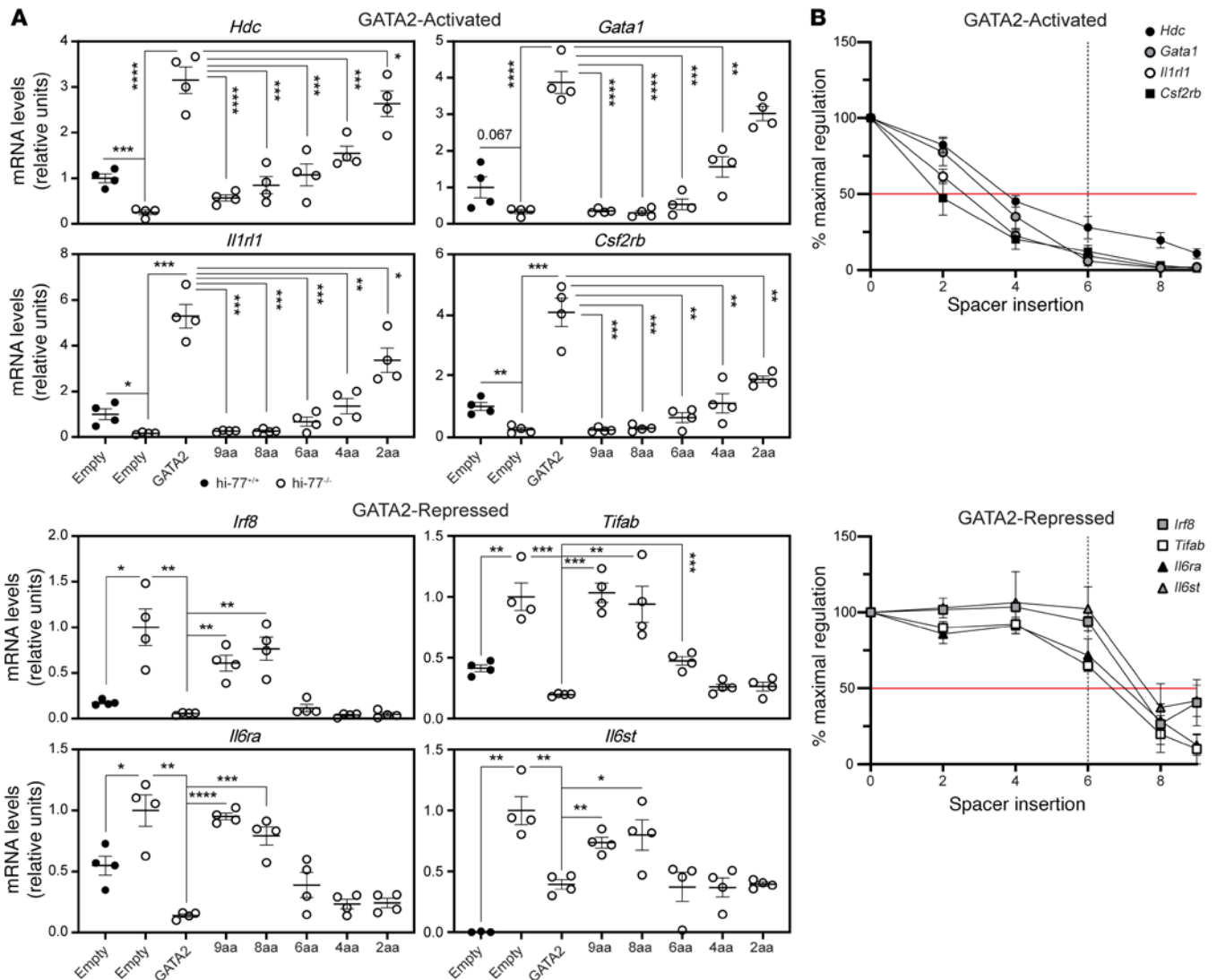


**Figure 2. GATA2 9aa-Ins disease variant is severely defective, but not entirely inactive, in genome regulation.** (A) Model 1, 9aa-Ins fails to regulate all GATA2-regulated genes; model 2, 9aa-Ins fails to repress GATA2-regulated genes; model 3, 9aa-Ins fails to activate GATA2-regulated genes; model 4, 9aa-Ins ectopically regulates genes that are not regulated by GATA2. (B) Overlap of DEGs that are  $-77^-$ , GATA2 $^-$ , and 9aa-Ins-regulated. RNA-Seq (4 biological replicates) of hi-77<sup>-/-</sup> with control vector (hi-77<sup>-/-</sup> empty), hi-77<sup>-/-</sup> with control vector (hi-77<sup>-/-</sup> empty), hi-77<sup>-/-</sup> with GATA2 (hi-77<sup>-/-</sup> GATA2), and hi-77<sup>-/-</sup> with 9aa-Ins (hi-77<sup>-/-</sup> 9aa-Ins). DEGs in hi-77<sup>-/-</sup> empty, hi-77<sup>-/-</sup> GATA2, and hi-77<sup>-/-</sup> 9aa-Ins were defined as  $|\log_2(\text{fold change})| \geq 1$  and adjusted *P* value < 0.05 relative to hi-77<sup>-/-</sup> empty. Each circle represents DEGs in the 3 categories: green circle, enhancer-regulated, (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> empty); blue circle, GATA2-regulated, (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> GATA2); pink circle, 9aa-Ins-regulated, (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> 9aa-Ins). DEGs were parsed into activated or repressed. DEG numbers are shown in parentheses. (C) Correlation plots depicting retention of 9aa-Ins-mediated activation and repression relative to GATA2. Comparison was calculated using  $|\log_2(\text{fold change})| \geq 1$  of (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> GATA2) and (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> 9aa-Ins). Genes were required to have TPM  $\geq 1$  in all replicates in hi-77<sup>-/-</sup> GATA2 for activation and hi-77<sup>-/-</sup> empty for repression. (D) PCA quantifying multidimensional scaling distances between transcriptomes (4 biological replicates). (E) Percentage of 9aa-Ins-regulated genes that are not GATA2-regulated analyzed by subtraction of overlap of (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> GATA2) from (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> 9aa-Ins). *P* cutoffs ranged from 0.01 to 0.1. (F) Percentage of GATA2-regulated genes that are not 9aa-Ins-regulated analyzed by subtraction of (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> 9aa-Ins) from (hi-77<sup>-/-</sup> empty)/(hi-77<sup>-/-</sup> GATA2). The same 10 *P* cutoffs as in E were used. The percentages of DEGs for each cutoff were parsed into activated or repressed. Statistical calculations used Mann-Whitney *U* test; \*\*\*\**P* < 0.0001.

impaired a greater percentage of activated versus repressed genes (86.4% of GATA2-activated genes were not 9aa-Ins-activated; 64.6% of GATA2-repressed genes were not 9aa-Ins-repressed; *P* =  $9.1 \times 10^{-5}$ ) (Figure 2F). Excluding genes with less than 1 TPM yielded identical conclusions (Supplemental Figure 2, C and D). Thus, without altering protein level and nuclear localization, zinc finger dislocation generated overt defects genome-wide.

*Inter-zinc finger spacing constraints for activation versus repression.* To establish whether the 9aa insertion is uniquely deleterious or whether smaller spacing alterations are inhibitory, we quantified expression of GATA2-activated (*Hdc*, *Gata1*, *Il1rl1*, *Csf2rb*) or -repressed (*Irf8*, *Tifab*, *Il6ra*, *Il6st*) genes detected by RNA-Seq (Fig-

ure 3A). Using the rescue assay, we compared activities of GATA2 and 2, 4, 6, 8, and 9aa variants to activate and repress transcription. Quantitative reverse transcriptase PCR (qRT-PCR) analysis with GATA2-expressing hi-77<sup>-/-</sup> cells confirmed *Hdc*, *Gata1*, *Il1rl1*, and *Csf2rb* activation and *Irf8*, *Tifab*, *Il6ra*, and *Il6st* repression. 9aa-Ins was largely defective. Whereas the 8, 6, and 4aa variants were defective in activating *Hdc*, *Gata1*, *Il1rl1*, and *Csf2rb* expression, the 2aa variant was active. The 6, 4, and 2aa variants shared GATA2 activity to repress *Irf8*, *Tifab*, *Il6ra*, and *Il6st*, differing from 8aa, which resembled the defective 9aa-Ins. Although 6aa reduced GATA2-mediated activation by more than 50%, it repressed GATA2 target genes by at least 50% (Figure 3B). GATA2-mediated



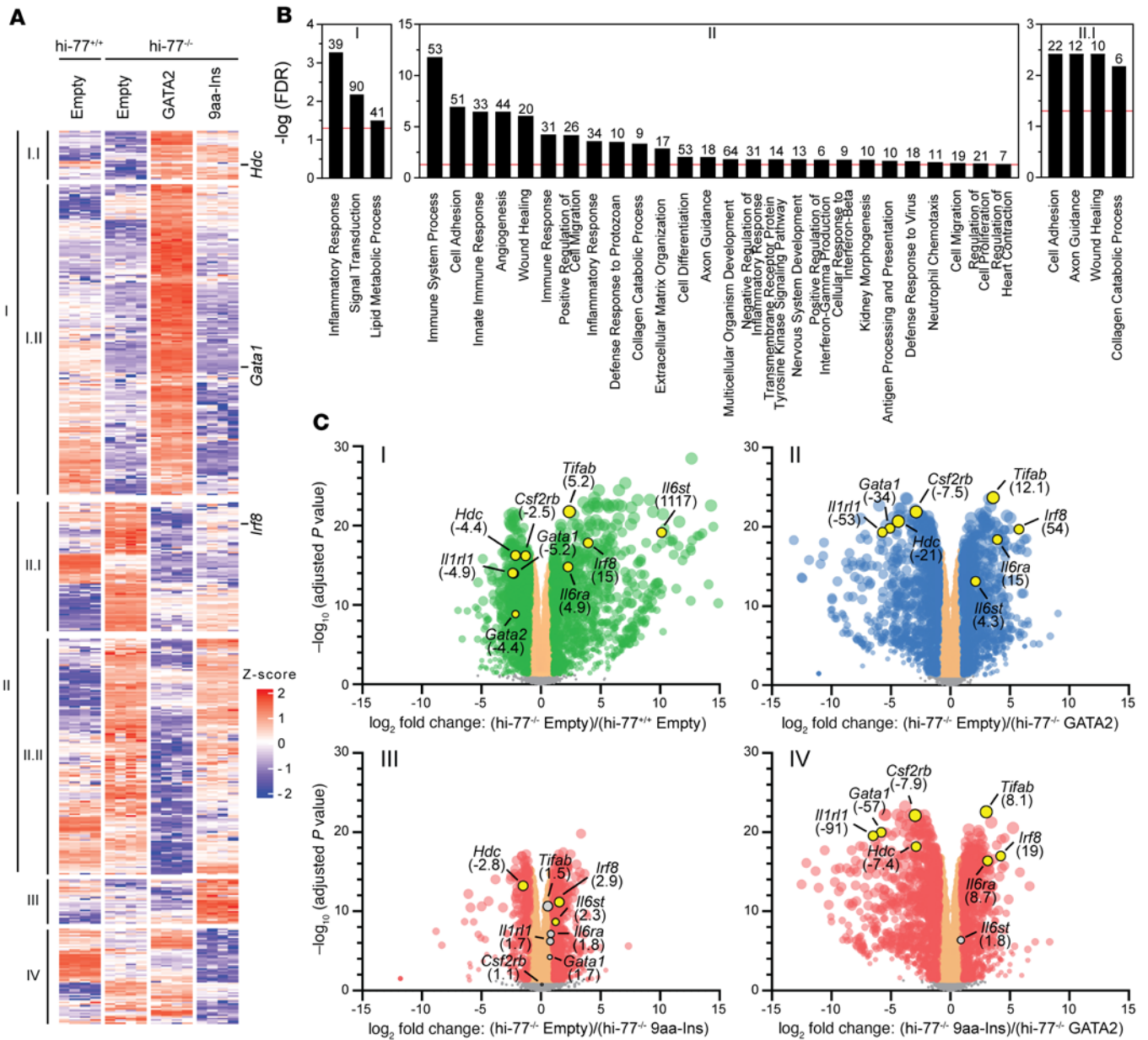
**Figure 3. Inter-zinc finger spacer constraints for GATA2-mediated activation versus repression.** (A) qRT-PCR analysis of mRNA expression in hi-77<sup>-/-</sup> cells rescued with GATA2 or variants (n = 4). (B) Comparison of the percentage maximal activation and repression calculated with data from A. 100% or 0% activation and repression determined by analysis of hi-77<sup>-/-</sup> cells with or without HA-GATA2. Red line, 50% regulation. Dotted line, 6aa insertion that impaired activation, but not repression, by more than 50%. Error bars represent mean ± SEM. Statistical calculations used unpaired 2-tailed Student's t test with Benjamini-Hochberg correction; \*P < 0.05; \*\*P < 0.01; \*\*\*P < 0.001; \*\*\*\*P < 0.0001.

activation was reduced by insertions ≥2aa, whereas repression tolerated ≤6aa insertions. These results illustrate how variants can disrupt certain molecular processes, while sparing others.

*Leveraging a GATA2 pathogenic variant to elucidate how GATA2 controls genome function.* 9aa-Ins could not establish the GATA2-dependent transcriptome, yet it retained some capacity to regulate select GATA2 target genes (Figure 4A). To unveil pathways/networks important for hematopoiesis, we stratified DEGs from hi-77<sup>-/-</sup> empty versus hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> empty versus hi-77<sup>-/-</sup> 9aa based on regulatory attributes and Gene Ontology to yield: I, GATA2-activated; I.I, GATA2- and 9aa-activated; I.II, only GATA2-activated; II, GATA2-repressed; II.I, GATA2- and 9aa-repressed; II.II, only GATA2-repressed; III, ectopically activated (9aa-Ins-activated but not GATA2-activated); IV, ectopically repressed (9aa-Ins-repressed but not GATA2-repressed) (Figure 4A). The 9aa insertion

impaired GATA2-mediated activation of genes related to G protein-coupled receptor signaling (e.g., *Cnr2*, *Slpr1*, *Slpr4*) and repression of innate immune genes (e.g., *Irf8*, *Tlr1*, *Tlr6*, *Tlr7*, *Il7*) (Figure 4B and Supplemental Figure 3). 9aa-Ins- but not GATA2-regulated genes (ectopic) did not reveal mechanistic insights. 9aa-Ins retained some activity to repress GATA2-repressed genes, including *Tgfa*, *Vwf*, *Fn1*, and *Tgfb3*, which mediate cell adhesion, wound healing, axon guidance, and collagen catabolic process (Figure 4B and Supplemental Figure 3), as well as the B-lineage genes *Pax5*, *Rag1*, and *Rag2* (Supplemental Table 1).

-77 deletion reduced *Gata2* expression 4.4-fold. We identified -77-regulated genes by comparing transcriptomes of hi-77<sup>-/-</sup> empty versus hi-77<sup>+/+</sup> empty cells. -77 deletion revealed 960 -77-activated and 1,124 -77-repressed genes (Figure 2B and Figure 4C). Many -77-activated (*Hdc*, *Gata1*, *Il1rl1*, *Csf2rb*) and -repressed



**Figure 4. Biological/mechanistic insights revealed from GATA2 disease variant transcriptomics. (A)** DEGs that are GATA2-regulated, GATA2- and 9aa-Ins-regulated, and 9aa-Ins-regulated from comparison of hi-77<sup>-/-</sup> empty (n = 4), hi-77<sup>-/-</sup> GATA2 (n = 4), and hi-77<sup>-/-</sup> 9aa-Ins (n = 4). Z score was calculated from each gene's log<sub>10</sub>(FPKM+10<sup>-3</sup>) from all RNA-Seq replicates. GATA2 or 9aa-Ins expression in hi-77<sup>-/-</sup> cells parsed the DEGs into: I, GATA2-activated; I.I, GATA2- and 9aa-Ins-activated; I.II, only GATA2-activated; II, GATA2-repressed; II.I, GATA2- and 9aa-Ins-repressed; II.II, only GATA2-repressed; III, ectopically activated; IV, ectopically repressed. **(B)** Gene Ontology (GO) analysis on 1,061 GATA2-activated, 1,077 GATA2-repressed, and 381 GATA2- and 9aa-Ins-repressed genes. Bar graphs represent -log(FDR) with a red line at FDR < 0.05 to determine statistical significance. Significant GO terms are presented. The number of genes comprised by each term is shown above the graphs. **(C)** Plots I, II, and III, expression changes from -77, GATA2, and 9aa-Ins regulation. 2,084 enhancer-regulated, 2,138 GATA2-regulated, and 939 GATA2-regulated genes from **B** are color-coded in green, blue, and pink, respectively. Plot IV, magnitude of expression between GATA2 and 9aa-Ins. DEGs are depicted in pink. GATA2-activated and -repressed DEGs are highlighted. Fold change relative to hi-77<sup>-/-</sup> empty or hi-77<sup>-/-</sup> GATA2 is shown in parentheses.

(*Irf8*, *Tifab*, *Il6ra*, *Il6st*) genes (Figure 4C, plot I) were GATA2-activated and -repressed upon rescue (Figure 4C, plot II). 9aa-Ins was impaired in activating and repressing these genes (Figure 4C, plot III), and at most genes, it was less active than GATA2 (Figure 4C, plot IV). These analyses established the GATA2 contribution to the progenitor transcriptome and extreme differences between GATA2 and 9aa-Ins.

Since 9aa-Ins was compromised in activating and repressing many loci, ATAC-Seq (assay for transposase-accessible chromatin using sequencing) was used to ask whether 9aa-Ins is defective in remodeling of chromatin or in steps after chromatin regulation. Integrating ATAC-Seq and RNA-Seq data linked differentially accessible peaks to DEGs. Peaks were called in the 4 replicates by MACS2 (Model-Based Analysis of ChIP-Seq), and Irreproducible



Discovery Rate (IDR) was used to identify 44,733 reproducible (master) peaks. PCA of peaks across the samples confirmed reproducibility and definitive separation among the samples (Figure 5A). Using DESeq2, differentially accessible ATAC-Seq peaks were determined by amalgamation of overlapping peaks from the replicates of hi-77<sup>-/-</sup> empty, hi-77<sup>-/-</sup> GATA2, and hi-77<sup>-/-</sup> 9aa-Ins. Differentially accessible peaks in hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins were defined as  $|\log_2(\text{fold change})| > 1$  and adjusted *P* value  $< 0.05$  relative to hi-77<sup>-/-</sup> empty.

To test whether GATA2-mediated activation and repression similarly involve chromatin remodeling, we analyzed ATAC-Seq peaks at GATA2-regulated genes by accessibility and magnitude of changes. At GATA2-activated and -repressed loci, we parsed differential peaks by accessibility and assigned them to the nearest DEG (Supplemental Methods). Analyzing differentially regulated peaks in hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty revealed peaks linked to 559 GATA2-activated and 436 GATA2-repressed DEGs (Figure 5B). GATA2-mediated activation (559 genes) more frequently involved chromatin opening versus closing (35% versus 8.6%), while GATA2-mediated repression (436 genes) more frequently involved closing versus opening (40% versus 5.0%). The proportion of GATA2-activated genes with opening was comparable to that of GATA2-repressed genes with closing (35% versus 40%).

To compare GATA2 and 9aa-Ins activities to remodel chromatin, we quantified chromatin changes at 1,613 DEGs that were exclusively GATA2-regulated (917 activated and 696 repressed) (Figure 2B). Normalized ATAC-Seq peak signals from 439 genes activated by GATA2, but not 9aa-Ins, increased significantly (Figure 5C; Wilcoxon's rank sum test *P* value  $< 2.1 \times 10^{-41}$ ). Normalized peak signals from 274 genes repressed by GATA2, but not 9aa-Ins, decreased significantly (Wilcoxon's rank sum test *P*  $< 2.0 \times 10^{-23}$ ). The analysis at promoters (-2 kb to +100 bp) yielded similar conclusions (Supplemental Figure 4A).

As 9aa-Ins was defective in activating and repressing most GATA2-regulated genes, we asked whether 9aa-Ins failed to remodel chromatin. Analyzing differentially accessible peaks in hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty and hi-77<sup>-/-</sup> 9aa-Ins/hi-77<sup>-/-</sup> empty revealed 62 GATA2- and 9aa-Ins-activated and 137 GATA2- and 9aa-Ins-repressed genes (Figure 5B). The proportion of 21 mutually activated genes with chromatin opening was only 3.8% of 559 GATA2-activated genes. Thirty-two mutually repressed genes with chromatin closing constituted 7.3% of 436 GATA2-repressed genes. The 21 mutually activated genes with chromatin opening and 32 mutually repressed genes with chromatin closing represented 34% and 23% within their cohort of mutually regulated genes. Comparison of GATA2-activated with mutually activated genes revealed a similar proportion of genes (35% vs. 34%) with chromatin opening (proportion test *P*  $< 0.08193$ ). Comparison of GATA2-repressed with mutually repressed genes revealed a decreased proportion of genes (40% vs. 23%) with chromatin closing (proportion test *P*  $< 1.495 \times 10^{-6}$ ). 9aa-Ins retained activity at a minority of loci.

To quantify chromatin changes, we analyzed 525 DEGs (144 activated and 381 repressed) that were GATA2- and 9aa-Ins-regulated (Figure 2B). Sixty activated and 128 repressed genes harbored differential peaks. ATAC-Seq data from 60 GATA2- and 9aa-Ins-activated genes revealed chromatin opening by GATA2 (*P*  $< 8.737 \times 10^{-9}$ ) and 9aa-Ins (*P*  $< 7.502 \times 10^{-8}$ ), while ATAC-Seq data from 128

GATA2- and 9aa-Ins-repressed genes revealed chromatin closing by GATA2 (*P*  $< 8.875 \times 10^{-15}$ ) and 9aa-Ins (*P*  $< 1.337 \times 10^{-7}$ ). Peak signals from promoters yielded similar conclusions (Supplemental Figure 4A). Since GATA2 and 9aa-Ins remodeled chromatin similarly at mutually regulated genes, we assessed whether 9aa-Ins defect in transcriptional regulation is linked to its inability to open or close chromatin at genes regulated by GATA2, but not 9aa-Ins. We compared hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty versus hi-77<sup>-/-</sup> 9aa-Ins/hi-77<sup>-/-</sup> empty ATAC-Seq data from 439 exclusively GATA2-activated and 274 exclusively GATA2-repressed genes (Figure 5C). In 439 exclusively GATA2-activated genes, 9aa-Ins-mediated chromatin opening was attenuated relative to GATA2 (*P*  $< 3.031 \times 10^{-5}$ ). In 274 exclusively GATA2-repressed genes, 9aa-Ins-mediated chromatin closing was attenuated (*P*  $< 4.081 \times 10^{-10}$ ) (Figure 5C). Similar conclusions emerged from promoter analyses (Supplemental Figure 4A). 9aa-Ins was defective in regulating chromatin at genes exclusively regulated by GATA2.

To test whether 9aa-Ins regulates chromatin at ectopic loci, ATAC-Seq peaks were linked to 414 ectopically regulated genes (132 activated, 282 repressed) (Figure 2B). ATAC-Seq analysis with hi-77<sup>-/-</sup> 9aa-Ins/hi-77<sup>-/-</sup> empty, but not hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty, revealed 78 ectopically activated and 83 ectopically repressed genes (Figure 5B). Fifty-five of 78 genes (70%) that were only activated by 9aa-Ins exhibited chromatin opening, and 13 of the 83 genes (16%) only repressed by 9aa-Ins exhibited chromatin closing. Analysis of 74 ectopically activated genes revealed increased accessibility (*P*  $< 7.502 \times 10^{-8}$ ), while peaks at 77 ectopically repressed genes revealed reduced accessibility (*P*  $< 1.337 \times 10^{-7}$ ) (Figure 5C). At select loci in which 9aa-Ins uniquely controlled transcription, it ectopically regulated chromatin.

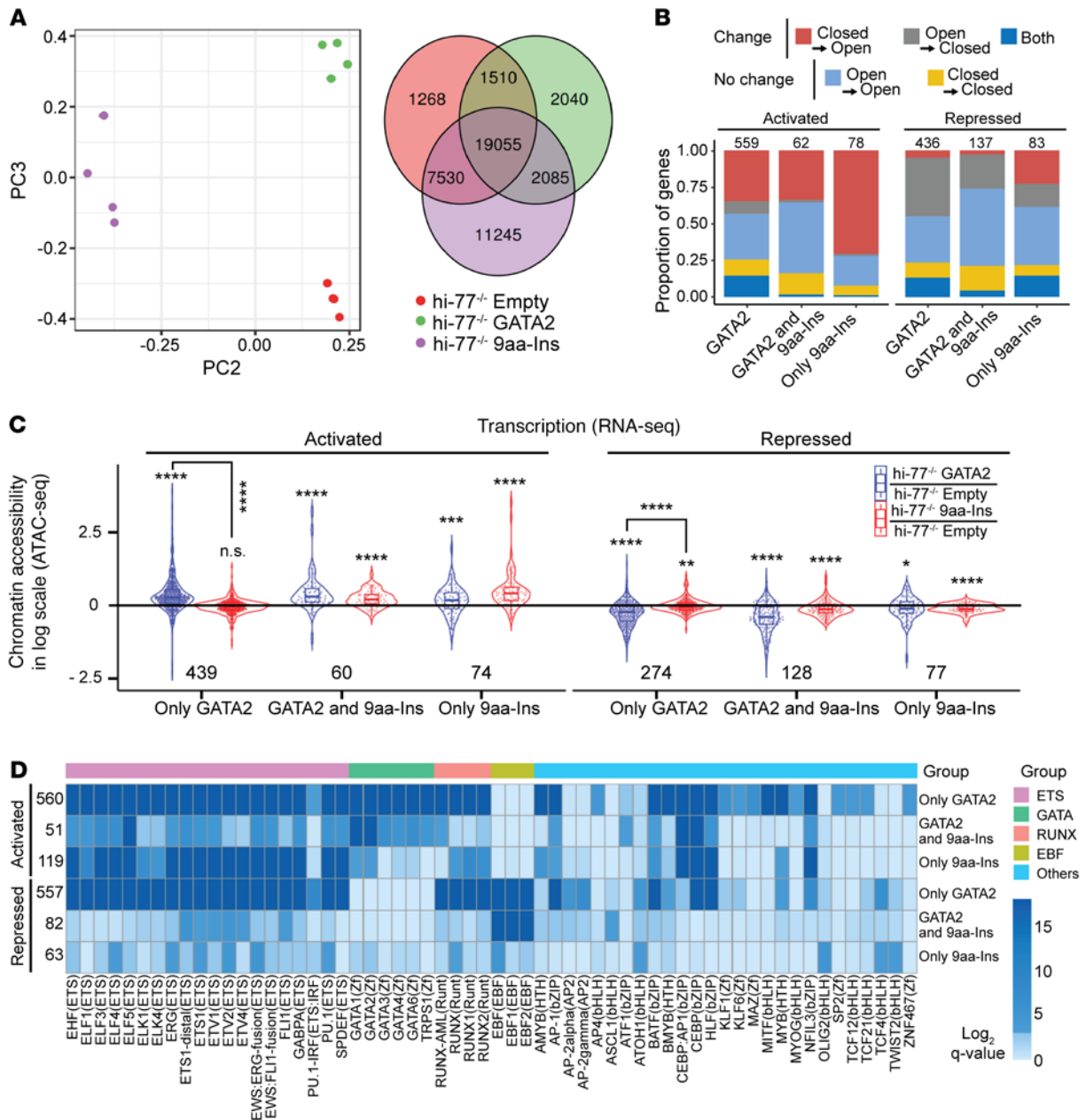
Motif analysis was conducted with 560 and 557 differentially accessible peaks linked to GATA2-activated and -repressed genes, respectively (not 9aa-Ins-regulated). The GATA2-binding WGATAR motif was enriched in ATAC-Seq peaks at activated loci (Figure 5D) and underrepresented in peaks from repressed loci (Figure 5D), which was unpredictable since GATA2-dependent repression mechanisms are not established. De novo motif finding from enrichment (Supplemental Figure 4B) and discriminative (Supplemental Figure 4C) analyses yielded similar results. ETS motifs were enriched at sites within GATA2-repressed loci (Figure 5D) to a greater extent than GATA2-activated loci (Supplemental Figure 4C).

As WGATAR was enriched at GATA2-activated loci, we asked whether E-box-spacer-WGATAR composite elements and double WGATAR motifs were also enriched. We identified sequences with 6- to 14-bp spacers between E-box and WGATAR and up to 5 bp spacer between 2 WGATARS. Composite elements resided at 11.3% and 6.46% and double WGATAR motifs at 1.25% and 0.72% of ATAC-Seq peaks at GATA2-activated and -repressed loci, respectively.

Given the ETS motif enrichment at GATA2-repressed loci, and ETS protein activation functions (67), we tested whether GATA2 downregulates genes encoding ETS factors. While 15 of 26 ETS factors (*Erg*, *Etv3*, *Etv4*, *Etv5*, *Etv6*, *Elf1*, *Elf2*, *Elf4*, *Elk1*, *Elk3*, *Elk4*, *Ets2*, *Gabpa*, *Spi1*, *Fli1*) were expressed (TPM  $\geq 1$  in all conditions) in progenitors, GATA2 and 9aa-Ins did not alter their expression (Supplemental Table 3).

9aa-Ins regulated chromatin accessibility and activated or repressed a minority of GATA2-regulated genes (Figure 5B and





**Figure 5. Multiomic analysis with GATA2 disease variant reveals principles of GATA2 function through chromatin.** (A) PCA quantifying multidimensional scaling distances between differential chromatin accessibility (ATAC-Seq, GEO GSE201968) with  $n = 4$  biological replicates of hi-77<sup>-/-</sup> empty, hi-77<sup>-/-</sup> GATA2, hi-77<sup>-/-</sup> 9aa-Ins. Venn diagram depicts overlap. (B) Chromatin transitions of genes activated or repressed by GATA2, both GATA2 and 9aa-Ins, or only 9aa-Ins. Number of genes comprised by each group is shown above the graphs. (C) GATA2 and 9aa-Ins impact on chromatin accessibility. hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty signal or hi-77<sup>-/-</sup> 9aa-Ins/hi-77<sup>-/-</sup> empty signal was determined at genes activated/repressed by only GATA2, GATA2 and 9aa-Ins, and only 9aa-Ins conditions by amalgamation of ATAC-Seq peaks linked to DEGs. Statistical calculations to measure chromatin accessibility (>0 or <0 for differential accessibility) used Wilcoxon's rank sum test. Comparisons between 2 groups used Wilcoxon's signed rank test. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . (D) Motif enrichment analysis at differentially accessible loci activated or repressed by only GATA2, GATA2 and 9aa-Ins, or only 9aa-Ins. Number of peaks comprised by each group is on the left of the heatmap. Peaks less than 100 kb from the start site were analyzed.

Supplemental Table 1). Fifty-one differentially accessible ATAC-Seq peaks at GATA2- and 9aa-Ins-activated genes harbored fewer WGATAR motifs (Figure 5D) versus only GATA2-activated genes (Supplemental Figure 4C). Eighty-two differentially accessible peaks at GATA2- and 9aa-Ins-repressed genes harbored significantly fewer ETS motifs versus only GATA2-repressed loci (Figure 5D and Supplemental Figure 4C). Comparing motifs at loci reg-

ulated by exclusively GATA2, GATA2 and 9aa-Ins, or exclusively 9aa-Ins, enriched motifs (Figure 5D) were not detected from discriminative analyses (Supplemental Figure 4C) of 119 differentially accessible peaks at ectopically 9aa-Ins-activated genes and 63 differentially accessible peaks at ectopically 9aa-Ins-repressed loci.

To test whether defective transcriptional and chromatin regulation by 9aa-Ins involves impaired DNA binding, we generated

recombinant GATA2 and 9aa-Ins double-zinc finger proteins in *E. coli*, purified to about 80% purity (Supplemental Figure 5A), and conducted electrophoretic mobility shift assay (EMSA). One-dimensional <sup>1</sup>H NMR spectra detected comparable folding between WT and 9aa-Ins, with additional signals representing the inserted 9 aa residues; the zinc fingers were not destabilized (Supplemental Figure 5B). EMSA with increasing protein concentrations and probes harboring single WGATAR sequences (AGATAA or TGATAA) of the GATA2-activated *Kit* locus (-114 kb) or double GATA motif (GGATAAAGATC) (68) revealed that 9aa-Ins zinc fingers had reduced DNA binding capacity with TGATAA and double GATA probes. WT GATA2 fingers did not stably bind the AGATAA probe; although the oligonucleotide was bound, stable complexes were unresolved (Supplemental Figure 5C).

To establish whether the 9aa insertion affects chromatin occupancy, we conducted CUT&Tag with anti-HA antibody in hi-77<sup>-/-</sup> cells expressing GATA2 or 9aa-Ins. Peaks were called with MACS3 and merged with HOMER mergePeaks (69) and Diffbind (70) (6,014 GATA2 and 1,051 9aa-Ins peaks; Supplemental Figure 6A). Linking CUT&Tag peaks to the nearest GATA2-regulated genes from RNA-Seq yielded 561 GATA2 and 35 9aa-Ins peaks at activated loci and 190 GATA2 and 27 9aa-Ins peaks at repressed loci (Supplemental Figure 6B). As only 33 GATA2 and 9aa-Ins peaks at activated loci and 20 peaks at repressed loci were detected, the 9aa insertion reduced occupancy at 94% (528 of 561) and 90% (170 of 190) of activated and repressed loci, respectively. ChIP-qPCR with hi-77<sup>-/-</sup> cells expressing comparable HA-GATA2 and 9aa-Ins levels (Supplemental Figure 6C) confirmed the loss of 9aa-Ins occupancy (*Gata1* and *Hdc* promoters, *Kit* -114) (Supplemental Figure 6D). Comparison of activated versus repressed loci revealed a 3.0-fold increase in GATA2 occupancy sites at activated loci (561 vs. 190) (Supplemental Figure 6B), despite a similar WGATAR distribution at activated and repressed loci (363 of 561 peaks [69%] at activated loci and 104 of 190 peaks [61%] at repressed loci; 2-sample test for equality of proportions with continuity correction,  $P = 0.08$ ).

At loci in which GATA2 remodeled chromatin, GATA2 occupied the GATA2-activated loci *Far2*, *Gata1*, *Hdc*, and *Kit* with differentially accessible ATAC-Seq peaks in hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty (Figure 6). GATA2 did not occupy the GATA2-repressed loci *Ifi209*, *Tifab*, *Nlrp1a*, and *Trem1* with ATAC-Seq peaks that were less accessible in hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty (Figure 6). 9aa-Ins lacked occupancy at these loci (Figure 6). The lack of chromatin occupancy at select repressed loci suggests that GATA2-regulated genome repression also occurs indirectly.

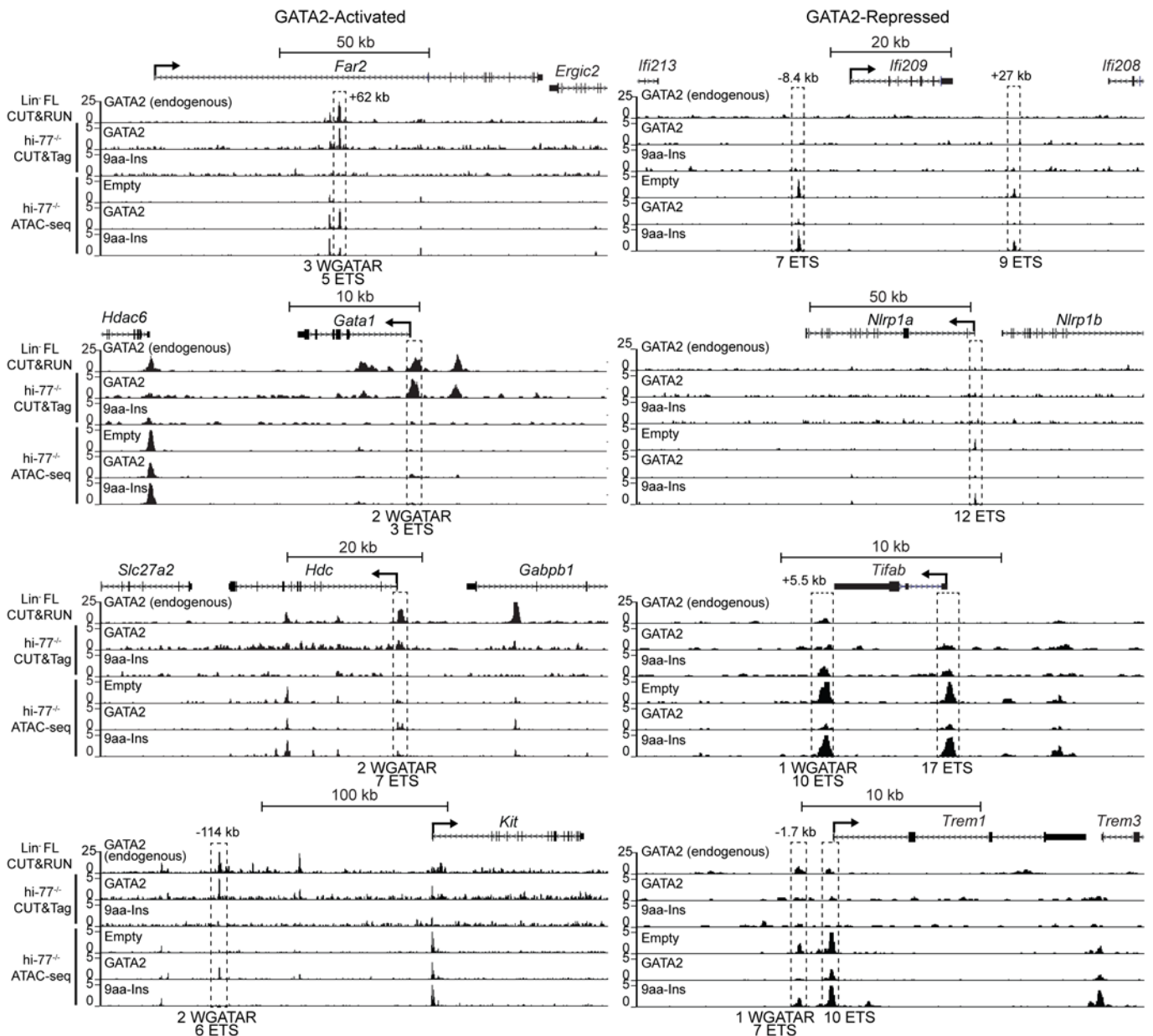
*GATA2 suppresses a lineage-diverting gene expression program.* To elucidate mechanisms underlying GATA2-mediated repression, we evaluated motifs enriched at GATA2-repressed loci. Motifs for the B-lineage developmental regulator EBF1 (38, 71, 72) were enriched at repressed loci (Figure 5D and Supplemental Figure 4C). *Ebf1* expression was low (7.5 TPM) in hi-77<sup>+/+</sup> cells, as expected for a lymphopoiesis-driving gene. *Ebf1* expression was upregulated 2.5-fold in hi-77<sup>-/-</sup> cells ( $P = 0.0002$ ) (Figure 7A) and 4.0-fold in -77<sup>-/-</sup> primary fetal liver progenitors ( $P = 0.028$ ) (55). GATA2 expression in the rescue assay reduced *Ebf1* expression by 84% ( $P < 0.0001$ ) (Figure 7A). The genes upregulated in hi-77<sup>-/-</sup> cells and repressed by GATA2 included a cohort with expression enriched in B-lineage cells (haemosphere.org) (*Ebf1*, *Igll1*, *Vpreb3*,

*Pax5*, *Myl4*, *Rag1*, and *Rag2*), which are vital for B-lineage biology and B-lineage/myeloid genes (*Cd79a*, *Cd79b*, *Mef2c*, and *Irf8*) (Figure 7A). GATA2 occupied the *Ebf1* promoter and remodeled chromatin (Supplemental Figure 7). In pre-B cells, EBF1 occupied GATA2-regulated ATAC-Seq peaks at *Igll1*, *Vpreb3*, *Pax5*, *Cd79a*, and *Cd79b* (Supplemental Figure 7). To test whether elevated EBF1 suffices to activate B-lineage genes in progenitors, we expressed HA-EBF1 in hi-77<sup>+/+</sup> cells. *Ebf1* transcripts were 6.5-fold higher in HA-EBF1-expressing versus control (empty vector) cells. *Gata2* transcripts were unchanged. HA-EBF1 increased expression of B-lineage (*Igll1*, *Vpreb3*, *Pax5*, *Myl4*) and B-lineage/myeloid genes (*Cd79a* and *Cd79b*) (Figure 7B).

PU.1 regulates B-lineage and myeloid genes and binds ETS motifs that exist at GATA2-repressed loci (Figure 5D and Supplemental Figure 4C). GATA2-PU.1 antagonism can determine myeloid fate (73–75), and GATA2 loss does not alter PU.1 levels (55). At loci in which GATA2 remodeled chromatin, PU.1 occupied *Ebf1*, *Vpreb3*, *Pax5*, *Myl4*, *Cd79a*, *Cd79b*, *Mef2c*, *Irf8*, *Tifab*, and *Csf2rb* (Supplemental Figure 7). To determine whether downregulating PU.1 in hi-77<sup>-/-</sup> cells reduces expression of B-lineage genes upregulated in GATA2-deficient progenitors, we ablated the -14 kb upstream regulatory element (URE) from *Spi1* (encoding PU.1) in hi-77<sup>-/-</sup> cells (76–78) (Figure 7C). *Spi1* expression decreased 2.2-fold in  $\Delta$ URE cells versus hi-77<sup>-/-</sup> without affecting *Gata2* expression. GATA2-repressed B-lineage genes (*Vpreb3*, *Pax5*, *Rag2*, *Cd79a*, *Cd79b*, *Mef2c*, *Irf8*, and *Tifab*) were downregulated in the  $\Delta$ URE cells (Figure 7D). *Csf2rb*, a GATA2-activated gene (Figure 4C), was unaffected (Figure 7D). These results support a model in which GATA2 represses *Ebf1* expression and antagonizes PU.1 in fetal progenitors to suppress a B-lineage expression program that is discordant with myeloid differentiation (Figure 7E).

*GATA2 deficiency establishes a hematopoiesis-disrupting cytokine signaling network.* GATA2 regulated genes encoding signaling proteins that were not 9aa-Ins-regulated (Supplemental Table 4). These genes included GATA2-activated *Csf2rb* (Figure 4C), encoding the shared common  $\beta$  chain of GM-CSF, IL-3, and IL-5 receptors (39, 40, 79). Since reduced GM-CSF signaling causes pulmonary alveolar proteinosis, a GATA2 deficiency syndrome phenotype (44, 45), we analyzed the underlying mechanism. To test whether GATA2 occupies *Csf2rb*, we assessed GATA2 CUT&RUN peaks in fetal liver Lin<sup>-</sup> progenitors (Lin<sup>-</sup> FL), GATA2 and HA CUT&Tag peaks in hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins, ATAC-Seq peaks in hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty, and ChIP-Seq from human CD34<sup>+</sup> and peripheral blood-derived erythroblast cells (80–82). GATA1 and GATA2 occupied sites upstream of *Csf2rb* (Figure 8A) and regulated chromatin accessibility. Comparing WT versus GATA2-deficient conditions (55, 83) or G1E-ER-GATA1 erythroblasts (60) revealed GATA2 and GATA1 induction of *Csf2rb* expression (Figure 8B).

By contrast to reduced *Csf2rb* expression in GATA2-deficient progenitors, genes encoding IL-6 receptor  $\alpha$  and  $\beta$  subunits (*Il6ra* and *Il6st*) were upregulated due to impaired GATA2-mediated repression (Figure 4C). *Il6st* encodes GPI30, a receptor subunit for IL-6 family members (IL-6, IL-11, and IL-27 among others) (49). Elevated IL-6 in acute myeloid leukemia (AML) induces bone marrow failure (49, 84) and promotes myelodysplastic syndrome (MDS) to AML progression (50). CUT&RUN peaks in Lin<sup>-</sup> progenitors, GATA2



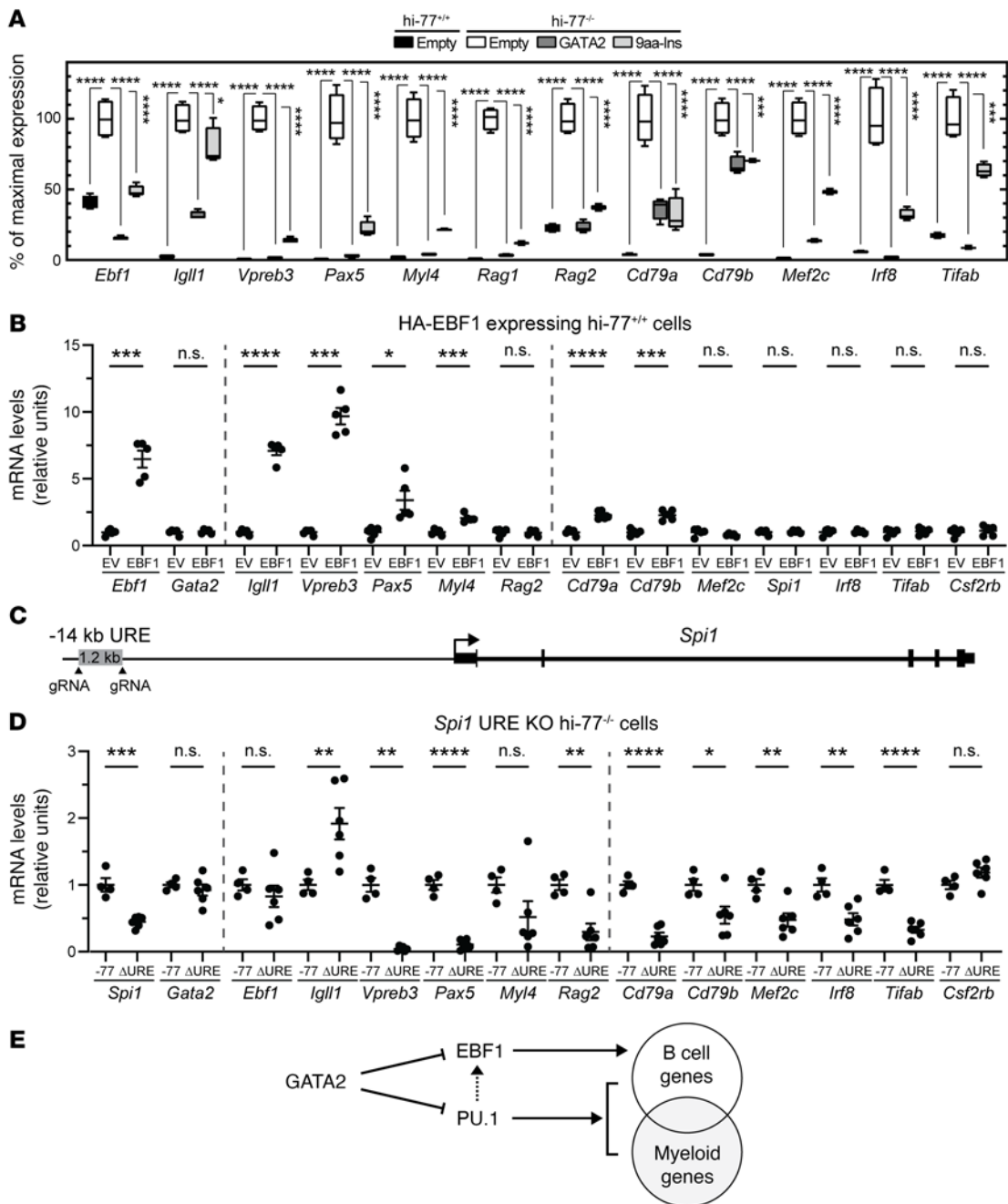
**Figure 6. GATA2 occupancy and GATA2-regulated chromatin remodeling at GATA2-activated and -repressed loci.** GATA2 CUT&RUN with fetal liver Lin<sup>-</sup> erythroid progenitors (Lin<sup>-</sup> FL) and HA CUT&Tag with hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins revealed GATA2 occupancy at GATA2-activated but not -repressed loci corresponding to ATAC-Seq profiles of GATA2-activated and GATA2-repressed genes. Peaks near the genes are boxed by dashed lines. WGATAR and ETS motifs located at peak sites are tabulated below.

and HA CUT&Tag peaks in hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins, and differentially accessible ATAC-Seq peaks in hi-77<sup>-/-</sup> GATA2/hi-77<sup>-/-</sup> empty revealed GATA2 occupancy and regulated chromatin at *Il6ra* intronic sites and 3'-UTR, and *Il6st* upstream sites (Figure 9, A and B). GATA2 ChIP-Seq with CD34<sup>+</sup> cells (79) revealed intronic and upstream peaks for *Il6ra* and *Il6st*, respectively (Figure 9, A and B). *Il6ra* and *Il6st* mRNA levels were upregulated in GATA2-deficient immortalized and primary progenitors (Figure 9C).

We analyzed GM-CSF signaling with hi-77<sup>+/+</sup> empty, hi-77<sup>-/-</sup> empty, and hi-77<sup>-/-</sup> GATA2 cells treated with vehicle (veh) or GM-CSF for 15 minutes and quantified STAT5 phosphorylation (p-STAT5) (Figure 10, A and B). In GM-CSF-treated cells, the

p-STAT5/total STAT5 ratio in hi-77<sup>+/+</sup> empty cells increased versus hi-77<sup>-/-</sup> empty cells. Signaling was indistinguishable between hi-77<sup>+/+</sup> empty and hi-77<sup>-/-</sup> GATA2 cells, demonstrating that GATA2 elevates signaling. As GATA2 increased *Csf2rb* expression and GM-CSF signaling, and GATA2-deficient progenitors had reduced signaling, these results inform the molecular underpinnings of a pathogenic phenotype.

Commensurate with upregulated *Il6ra* and *Il6st* mRNA levels (Figure 9C), IL-6/STAT3 signaling was upregulated. A 15-minute treatment with IL-6 increased STAT3 phosphorylation (p-STAT3) in hi-77<sup>-/-</sup> empty, but not hi-77<sup>-/-</sup> GATA2, cells (Figure 10, C and D). IL-6 induced p-STAT3 in hi-77<sup>-/-</sup> 9aa-Ins, but not hi-77<sup>-/-</sup> GATA2,

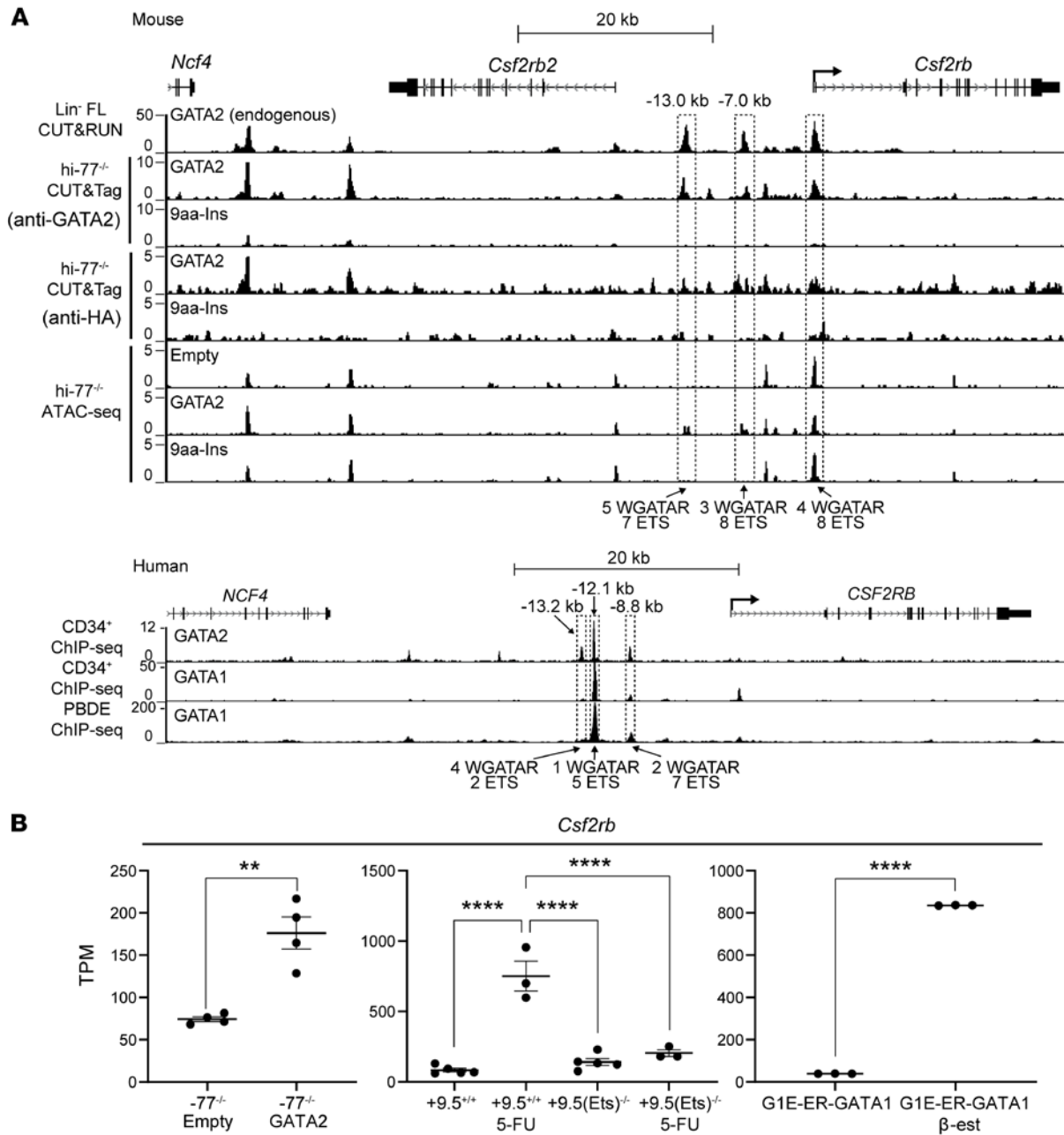


**Figure 7. GATA2 opposes a B-lineage gene expression program.** (A) Expression of GATA2-repressed CLP and B-lineage or myeloid genes in hi-77<sup>+/+</sup> empty, hi-77<sup>-/-</sup> empty, hi-77<sup>-/-</sup> GATA2, and hi-77<sup>-/-</sup> 9aa-Ins from the RNA-Seq of Figure 2. Average TPM of hi-77<sup>-/-</sup> empty is presented as 100% of maximal expression. For multiple comparisons with hi-77<sup>-/-</sup> empty control, statistics were calculated using 1-way ANOVA followed by Dunnett's test; \**P* < 0.05; \*\*\**P* < 0.001; \*\*\*\**P* < 0.0001. (B) mRNA levels of GATA2-repressed genes in hi-77<sup>+/+</sup> cells infected with retrovirus to express EBF1 or empty vector (EV). Error bars represent mean ± SEM. Statistical calculations used unpaired 2-tailed Student's *t* test. Welch's correction was applied when variances were unequal (*Ebf1*, *Igll1*, *Vpreb3*, *Pax5*); \**P* < 0.05; \*\*\**P* < 0.001; \*\*\*\**P* < 0.0001. (C) Targeted ablation of a *Spi1* (encoding PU.1) enhancer -14 kb upstream regulatory element (ΔURE) (77, 78). (D) GATA2-repressed target gene mRNA levels in hi-77<sup>-/-</sup> cells lacking PU.1 enhancer. Error bars represent mean ± SEM. Statistical calculations used unpaired 2-tailed Student's *t* test. Welch's correction was applied when variances were unequal (*Vpreb3*); \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001; \*\*\*\**P* < 0.0001. (E) GATA2-mediated repression with GATA2 occupying and repressing *Ebf1* expression and opposing PU.1-mediated activation. PU.1 induction of *Ebf1* expression (75, 76, 123) is indicated by a dashed line.

cells (Supplemental Figure 8, A and B). As elevated IL-6 signaling promotes bone marrow failure and MDS to AML progression (50), high IL-6 signaling in GATA2-deficient fetal progenitors may have pathogenic implications.

We tested whether defective GM-CSF and IL-6 signaling impacts differentiation. hi-77<sup>+/+</sup> versus hi-77<sup>-/-</sup> cells were treated with GM-CSF or IL-6, and granulocytic and monocytic differentiation was analyzed by flow cytometry with CD11b, CD115, Ly6C,

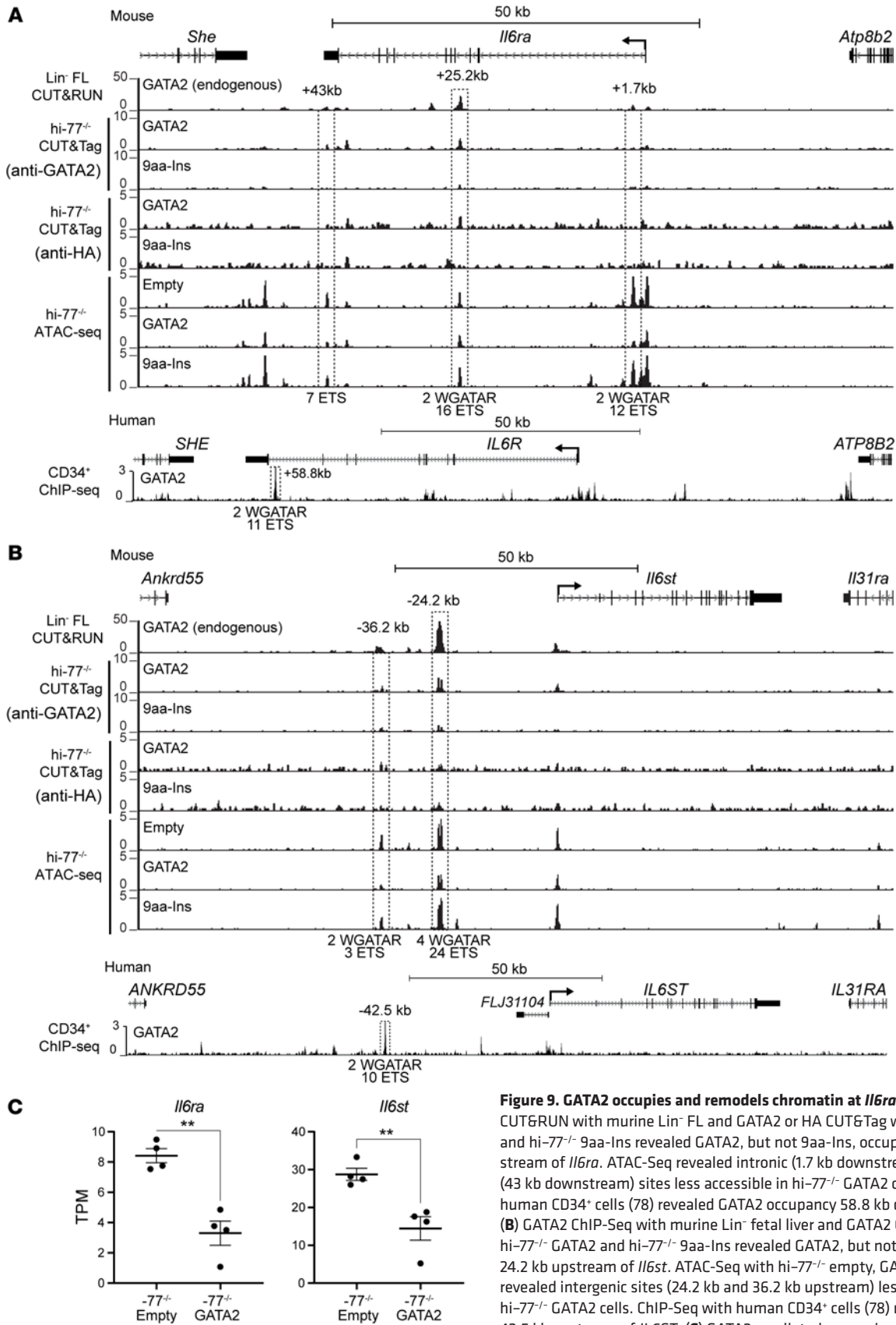




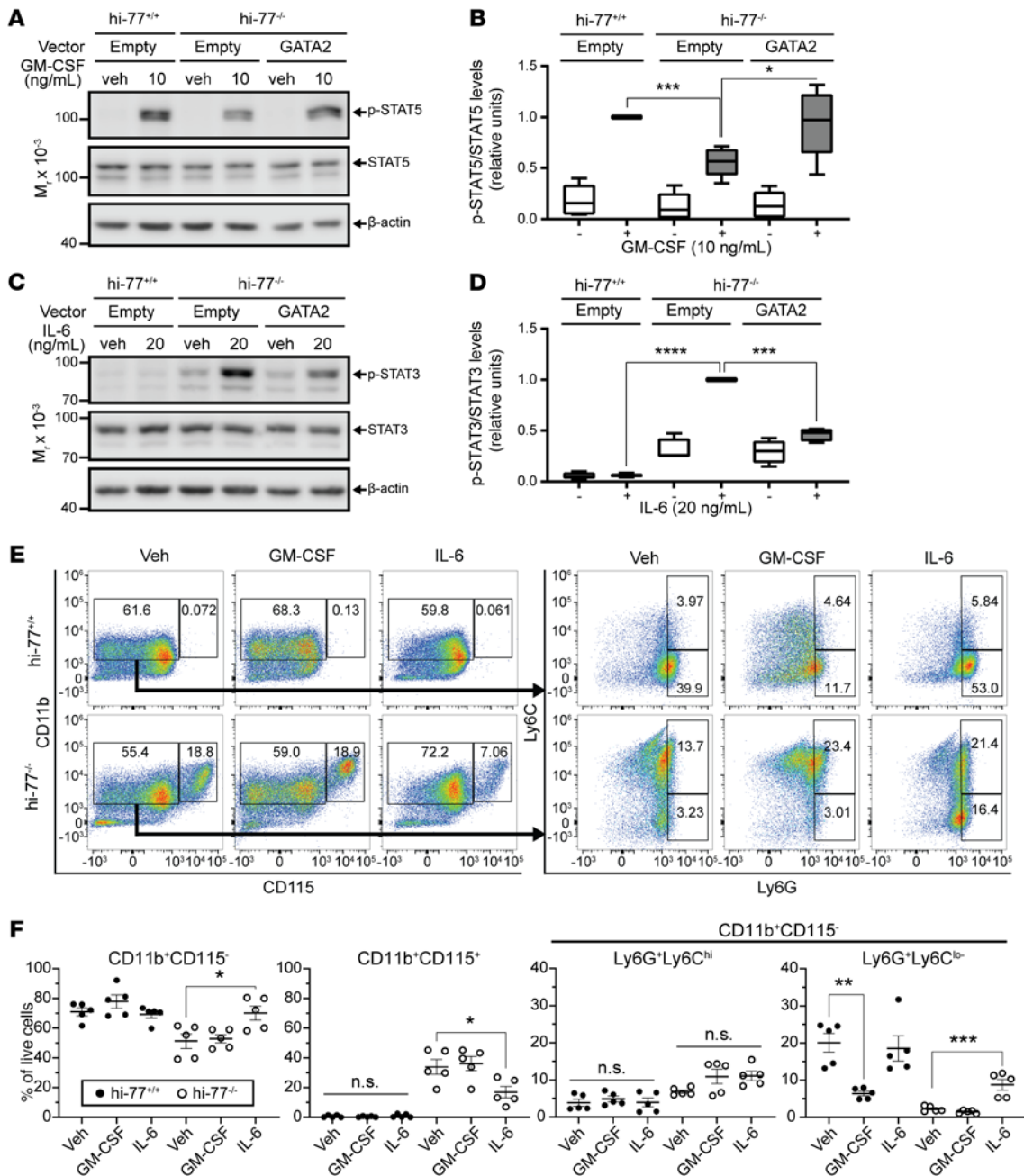
**Figure 8. GATA1 and GATA2 occupy and remodel *Csf2rb* chromatin. (A)** In mice, GATA2 CUT&RUN with Lin<sup>-</sup> FL and GATA2 or HA CUT&Tag with hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins revealed GATA2, but not 9aa-Ins, occupancy at the start site and 7 kb and 13 kb upstream of *Csf2rb*. ATAC-Seq revealed sites 7 kb and 13 kb upstream that were accessible only in hi-77<sup>-/-</sup> GATA2 cells. ChIP-Seq with human CD34<sup>+</sup> cells (78) revealed GATA2 occupancy 8.8, 12.1, and 13.2 kb upstream of *CSF2RB*. GATA1 ChIP-Seq with CD34<sup>+</sup> cells (79) and peripheral blood-derived erythroblast (PBDE) cells (79) revealed GATA1 occupancy 8.8 and 12.1 kb upstream. **(B)** GATA2 regulation of *Csf2rb* (RNA-Seq). GATA2-mediated activation of *Csf2rb* in primary Lin<sup>-</sup> -77<sup>-/-</sup> progenitors with or without GATA2 (-77<sup>-/-</sup> empty vs. -77<sup>-/-</sup> GATA2) (61), WT or 9.5(Ets) motif-mutant bone marrow LSK cells with 5-FU [9.5<sup>+/+</sup> 5-FU vs. 9.5(Ets)<sup>-/-</sup> 5-FU] (82), and G1E-ER-GATA1 erythroblasts with or without β-estradiol (G1E-ER-GATA1 vs. G1E-ER-GATA1 β-est) (68). To compare differences between 2 groups, statistical calculations used unpaired 2-tailed Student's *t* test. For multiple comparisons, unpaired 1-way ANOVA was used, followed by Tukey's test; \*\**P* < 0.01; \*\*\*\**P* < 0.0001.

and Ly6G markers (Supplemental Figure 9, A and B). Cells were cultured for 3 days with GM-CSF or IL-6. WT and hi-77<sup>-/-</sup> cells exhibited greater granulocytic (CD11b<sup>+</sup>CD115<sup>-</sup>) and monocytic (CD11b<sup>+</sup>CD115<sup>+</sup>) differentiation, respectively (Figure 10, E and F). In WT cells, GM-CSF reduced CD11b<sup>+</sup>CD115<sup>-</sup>Ly6G<sup>+</sup>Ly6C<sup>lo</sup>- granulocytic cells (Figure 10, E and F). GM-CSF treatment of hi-77<sup>-/-</sup> cells

had no effect. IL-6 suppressed the generation of CD11b<sup>+</sup>CD115<sup>+</sup> monocytic cells to a greater extent with hi-77<sup>-/-</sup> (Figure 10, E and F) versus WT cells and increased CD11b<sup>+</sup>CD115<sup>-</sup>Ly6G<sup>+</sup>Ly6C<sup>lo</sup>- granulocytic cells (Figure 10, E and F). Thus, GATA2 deficiency in fetal progenitors generated an aberrant transcriptome that dysregulated progenitor responsiveness to cytokines.



**Figure 9. GATA2 occupies and remodels chromatin at *Il6ra* and *Il6st*.** (A) GATA2 CUT&RUN with murine Lin<sup>-</sup> FL and GATA2 or HA CUT&Tag with hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins revealed GATA2, but not 9aa-Ins, occupancy 25.2 kb downstream of *Il6ra*. ATAC-Seq revealed intronic (1.7 kb downstream) and 3'-UTR (43 kb downstream) sites less accessible in hi-77<sup>-/-</sup> GATA2 cells. ChIP-Seq with human CD34<sup>+</sup> cells (78) revealed GATA2 occupancy 58.8 kb downstream of *IL6R*. (B) GATA2 ChIP-Seq with murine Lin<sup>-</sup> fetal liver and GATA2 CUT&Tag with hi-77<sup>-/-</sup> GATA2 and hi-77<sup>-/-</sup> 9aa-Ins revealed GATA2, but not 9aa-Ins, occupancy 24.2 kb upstream of *Il6st*. ATAC-Seq with hi-77<sup>-/-</sup> empty, GATA2, and 9aa-Ins revealed intergenic sites (24.2 kb and 36.2 kb upstream) less accessible in hi-77<sup>-/-</sup> GATA2 cells. ChIP-Seq with human CD34<sup>+</sup> cells (78) revealed occupancy 42.5 kb upstream of *IL6ST*. (C) GATA2-mediated repression of *Il6ra* and *Il6st* in primary Lin<sup>-</sup> -77<sup>-/-</sup> hematopoietic progenitors with and without GATA2 (-77<sup>-/-</sup> empty vs. -77<sup>-/-</sup> GATA2) (55). Statistical calculations used unpaired 2-tailed Student's *t* test; \*\**P* < 0.01.

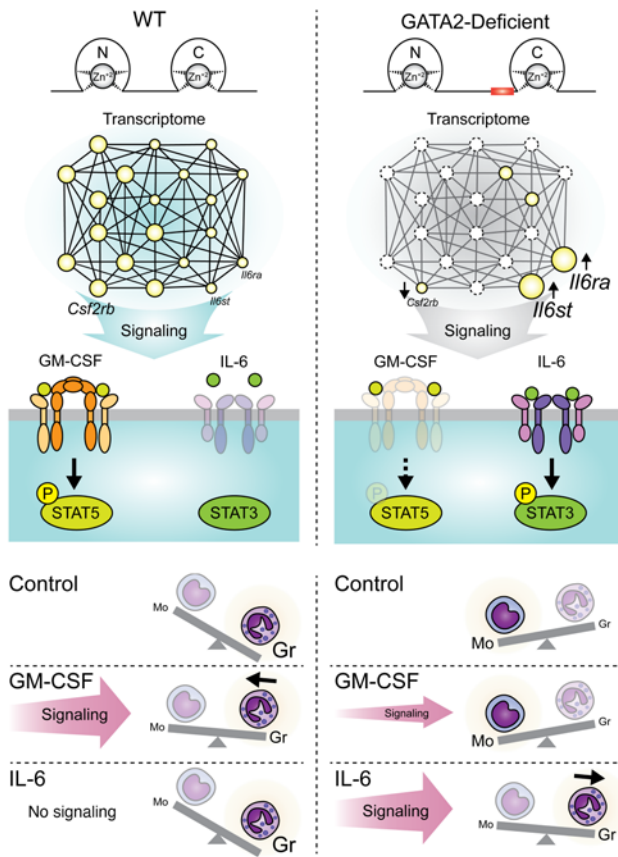


**Figure 10. GATA2-mediated regulation of cellular signaling and differentiation.** (A) Western blot to detect GM-CSF–induced STAT5 phosphorylation ( $n = 6$ ). (B) p-STAT5 quantification. Results were normalized to GM-CSF–treated hi-77<sup>+/+</sup> empty (box-and-whisker plots with bounds from the 25th to the 75th percentiles, the median line, and whiskers ranging from minimum to maximum values) ( $n = 6$ ). (C) Western blot to detect IL-6–induced STAT3 phosphorylation (p-STAT3) ( $n = 4$ ). (D) p-STAT3 quantification. Results were normalized to IL-6–treated hi-77<sup>-/-</sup> empty (box-and-whisker plots with bounds from the 25th to the 75th percentiles, the median line, and whiskers ranging from minimum to maximum values) ( $n = 4$ ). Statistical comparisons in B and D used paired 2-tailed Student’s *t* tests with Benjamini-Hochberg correction; \* $P < 0.05$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . (E) Flow cytometric plots of CD11b<sup>+</sup>CD115<sup>-</sup> (granulocytic) and CD11b<sup>+</sup>CD115<sup>+</sup> (monocytic) differentiated progenitors cultured for 3 days in control, GM-CSF–containing, or IL-6–containing media ( $n = 5$ ). Plots (gated at CD11b<sup>+</sup>CD115<sup>-</sup>) of Ly6G<sup>+</sup>Ly6C<sup>hi</sup> and Ly6G<sup>+</sup>Ly6C<sup>lo</sup> differentiated progenitors cultured for 3 days in control, GM-CSF–containing, or IL-6–containing media ( $n = 5$ ). (F) Quantification of CD11b<sup>+</sup>CD115<sup>-</sup>, CD11b<sup>+</sup>CD115<sup>+</sup>, CD11b<sup>+</sup>CD115<sup>-</sup>Ly6G<sup>+</sup>Ly6C<sup>hi</sup>, and CD11b<sup>+</sup>CD115<sup>-</sup>Ly6G<sup>+</sup>Ly6C<sup>lo</sup> populations. Error bars represent mean  $\pm$  SEM. For multiple comparisons with vehicle-treated control, statistics were calculated using 1-way ANOVA followed by Dunnett’s test; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

Discussion

Coding and noncoding variants elicit loss-of-function, gain-of-function, or composite phenotypes that initiate or promote pathogenesis or create a disease predisposition. Our analysis of how

a blood disease–causing variant affects GATA factor activity indicates that GATA2 zinc finger dislocation conforms to a model involving hypomorphic and neomorphic attributes. 9aa-Ins was inactive or had reduced activity at most targets, yet it retained



**Figure 11. Model of GATA2-regulated genome function, cytokine signaling, and progenitor differentiation.** GATA2 deficiency disrupts progenitor cell genome regulation. GATA2 loss decreases *Csf2rb* expression and GM-CSF signaling. GATA2, but not 9aa-Ins, elevates *Csf2rb* expression. GATA2 loss elevates *Il6ra* and *Il6st* expression and IL-6 signaling. WT GATA2, but not 9aa-Ins, reduces *Il6ra* and *Il6st* expression. These alterations impact differentiation and may impact function of progenitor-derived progeny. Normal progenitors exhibit predominantly granulocytic potential, and GM-CSF promotes granulopoiesis. IL-6 does not induce signaling in WT progenitors nor impact differentiation. GATA2-deficient progenitors exhibit predominantly monocytic (Mo) potential, and IL-6 promotes granulocytic (Gr), at the expense of monocytic, differentiation. GM-CSF does not induce signaling in GATA2-deficient progenitors nor impact differentiation. As CSF2RB and IL6ST are shared by additional receptors, their dysregulation will impact a broader ensemble of signaling systems to yield an aberrant network that may disrupt fetal hematopoiesis.

activity at a minority of targets and acquired ectopic activity. 8aa-Ins was defective, 2aa-Ins functioned normally, and 4 and 6aa-Ins preferentially disrupted activation.

GATA2 activated and repressed a similar number of genes in progenitors, and 9aa-Ins was incapable of remodeling chromatin and regulating transcription at most loci. 9aa-Ins retained activity at a small cohort of GATA2 targets and ectopic loci. An analysis of several target genes for GATA2 disease mutants T354M and R307W suggested a loss-of-function and gain-of-function phenotype (33). Given GATA2 deficiency syndrome variable penetrance and complex phenotypes (44, 45), hypomorphic and neomorphic attributes may contribute to this complexity. Frameshift mutations can ablate one allele (30), and epigenetic repression of the second allele can

exacerbate phenotypes (85). Considering the approximately 2,000 GATA2-regulated genes, lowering GATA2, ectopic activity, or both may corrupt GATA2 genetic networks.

As GATA2-mediated repression is not understood, our demonstration that regulated chromatin sites at GATA2-activated versus -repressed genes were differentially enriched in WGATAR motifs was surprising. Sites at activated genes harbored more WGATAR than sites at repressed genes. GATA1 and GATA2 colocalize on chromatin with other transcription factors and coregulators at sites that often, but not always, contain E-box-spacer-WGATAR composite elements (16, 61, 86–90). The ETS factor FLI1 occupies these sites and is implicated in GATA1- and GATA2-mediated activation (91, 92). Although the ETS factor PU.1 can antagonize GATA1 or GATA2 (74, 93, 94), it can cooperate with GATA2 in activation (95). We tested whether GATA2 regulates expression of ETS factors in this system, and it did not. EBF motifs were enriched at GATA2-repressed loci, and GATA2 occupied *Ebf1* and decreased expression of EBF1, which promotes B-lineage differentiation (38, 72), and EBF1 target genes, including the B-lineage developmental regulator *Pax5* (96) and V-D-J recombination regulators *Rag1* and *Rag2* (97). Ectopic EBF1 expression sufficed to induce the B-lineage expression program (Figure 7B). Although GATA2 deficiency did not alter PU.1 levels, lowering PU.1 levels by ablating a PU.1 enhancer, in the context of *Gata2* -77 enhancer loss, attenuated B-lineage gene upregulation (Figure 7D). PU.1 occupied select genes that were upregulated in GATA2-deficient progenitors. Besides GATA2 inducing hematopoietic stem cell (HSC) generation during embryogenesis, maintaining HSCs in adults and promoting myeloerythroid progenitor differentiation (11), our results support a paradigm in which GATA2 suppresses a B-lineage program that may oppose myeloerythroid differentiation. In B cells, Bach1 and Bach2 suppress a myeloid program (98). As not all GATA2-repressed genes exhibit B lineage-enriched expression, this mechanism constitutes one mode of GATA2-mediated repression, but other mechanisms likely operate at distinct target genes.

GATA2 controls hematopoiesis through cell-intrinsic activities in hematopoietic stem and progenitor cells (HSPCs) (99, 100), yet many questions remain regarding the physiological and pathological mechanisms. *Gata2* mutant transcriptomes (21, 29, 55–57, 83, 101–103), GATA2 mutant patient samples (104), and GATA2 chromatin occupancy (16, 105, 106) have revealed GATA2-activated and -repressed genes. Despite inferences from gene expression alterations, critical mediators of GATA2 activities are unknown. In GATA2-deficient fetal progenitors, with or without 9aa-Ins, inflammatory and innate immune components were upregulated. In a double-knockout rescue system to test whether upregulation of the innate immune gene activator IRF8 (107) underlies defective differentiation of -77<sup>-/-</sup> progenitors (108), ablating *Irf8* partially but inefficiently rescued granulopoiesis without restoring erythropoiesis (108). Thus, a GATA2-dependent, IRF8-independent mechanism is also important. The dysregulated cytokine receptors described herein (Supplemental Table 4) may alter GATA2-deficient progenitor responsiveness to extrinsic signals, and our analyses detected such defects.

GATA2, but not 9aa-Ins, normalized cytokine receptor gene expression in GATA2-deficient fetal progenitors. The upregulated genes included those encoding IL-6 receptor subunits. By contrast to *Il6ra*, encoding a dedicated IL-6 receptor subunit, *Il6st* encodes



IL6ST (GPI30), a subunit of multiple IL-6/IL-12 receptor family members (47, 109). Inflammatory stimuli increase IL-6 elaboration from bone marrow microenvironment cells (48). IL-6 acts on HSPCs to divert functions into emergency granulopoiesis, in which granulocyte-monocyte progenitors (GMPs) disproportionately generate neutrophils (110). In GATA2-deficient fetal progenitors, *Il6ra* and *Il6st* upregulation endowed progenitors with IL-6 signaling, which counteracted the monocytic fate, enabling granulopoiesis, commensurate with IL-6 promotion of granulopoiesis *in vivo*.

Besides granulopoiesis, high IL-6 induces anemia of inflammation (111, 112). IL-1 $\beta$  pro-proliferative actions on AML progenitors induce inflammatory cytokines including IL-6 (113). AML cell-derived IL-6 was linked to bone marrow failure in a xenograft model. Antibody-mediated neutralization of IL-6 reversed anemia, extending survival (84), and IL-6 promotes MDS progression to AML (50). As elevated IL-6 signaling in GATA2-deficient fetal progenitors suppresses hematopoiesis, in principle, IL-6 may mediate GATA2 pathologies during fetal development.

The mechanisms of how GATA2 deficiency and pathogenic GATA2 variants impact processes during human fetal development are not established. Considering proliferative fetal versus quiescent bone marrow HSCs (114, 115) and differential sensitivity of fetal and adult HSCs to FLT3-ITD-induced oncogenesis (116), fetal and adult HSC genomes may differ in sensitivity to dysregulated GATA2. With IL-6, granulopoiesis persisted with GATA2-deficient fetal progenitors (Figure 11), and neutrophils exist in pediatric and adult GATA2 deficiency patients, despite dendritic cell, monocyte, NK cell, and lymphoid cell reductions (44, 45). The neutrophil persistence may reflect sustained granulopoietic activity in GATA2-deficient myeloid progenitors when inflammatory cytokines, e.g., IL-6, drive the process. Otherwise, GATA2-deficient fetal progenitors exhibit a monocytic fate *ex vivo*, although whether the mutant progeny function normally is unknown (55).

By contrast to elevated IL-6 signaling in GATA2-deficient fetal progenitors, the GM-CSF receptor common  $\beta$  subunit (CSF2RB) decreased, reducing GM-CSF signaling. Whereas GM-CSF induced myeloid differentiation of WT progenitors, signaling-defective GATA2-deficient progenitors were less responsive (Figure 11). CSF2RB is an IL-3 and IL-6 receptor subunit (39–41), and its deficiency would impact signaling networks involving multiple cytokine receptors, analogous to multicomponent aberrations expected from elevated IL6ST, shared by IL-6/IL-12 receptor family members (49). During inflammation, GM-CSF signaling, requiring CSF2RB, elevates neutrophils, monocytes, eosinophils, dendritic cells, and myeloid-derived suppressor cells (117). CSF2RB mediates FLT3-ITD leukemogenic activity by promoting STAT5 phosphorylation (118).

Attenuated GM-CSF signaling and acquired IL-6 signaling, with each system linked to additional cytokine receptors, illustrate complex signaling perturbations caused by GATA2 deficiency in fetal progenitors. 9aa-Ins ectopically repressed expression of cytokine receptors (*Il7ra* and *Il27ra*) and increased *Pglrp1* expression, encoding a conserved peptidoglycan binding protein implicated in *Drosophila* Toll-like receptor (119) and mammalian (120) TNF receptor-1 regulation (Supplemental Table 2). 9aa-Ins increased *Mrgpra2b* expression, encoding an unstudied member of the Itch G protein-coupled receptor family (121) with high expression in myeloid progenitors (haemosphere.org). 9aa-Ins exacerbates the

scope of dysregulated signaling in GATA2-deficient fetal progenitors. It is attractive to propose that in GATA2-deficient patients, defective signaling networks attenuate progenitor responsiveness to stress-derived signals, e.g., from pathogens, and responsiveness to other genetic or epigenetic aberrations, thereby causing or contributing to a bone marrow failure and leukemia predisposition.

## Methods

Additional details can be found in Supplemental Methods.

**Immortalized cell culture.** ER-HoxB8-immortalized (hi) progenitors were generated from mouse fetal liver Lin<sup>-</sup> cells immortalized by retroviral expression of estrogen-regulated HoxB8 (122). Cells were cultured in Opti-MEM (Gibco, Thermo Fisher Scientific) with 10% FBS, 1% penicillin-streptomycin, 1% SCF-conditioned medium, 30 mM  $\beta$ -mercaptoethanol, 1  $\mu$ M  $\beta$ -estradiol, and 500  $\mu$ g/mL G418. Cells were cultured in a humidified 5% CO<sub>2</sub> incubator at 37°C.

**GATA2 rescue assay.** GATA2 or spacer mutants were expressed in -77 enhancer-deleted GATA2-depleted -77<sup>-/-</sup> hematopoietic progenitors by infection of hi-77<sup>-/-</sup> cells with retrovirus harboring murine *Gata2* or mutant cDNA in the murine stem cell virus plasmid with GFP (MSCV-PIG; ref. 33). Ecotropic virus was packaged in 293T cells, and retrovirus-containing supernatants were collected 48 hours after transfection. Cells were transferred to IMDM containing 2% FBS and incubated with supernatant by spinoculation for 90 minutes at 1,315g at 30°C. Cells were cultured for 3 days in media described above. GFP<sup>+</sup> and GFP<sup>-</sup> cells were subjected to protein analysis. GFP<sup>+</sup> cells were sorted by FACS with a FACSAria cell sorter (BD Biosciences), and RNA was isolated with TRIzol (Invitrogen, Thermo Fisher Scientific).

**Genomic analysis.** Sample preparation and data processing of RNA-Seq (Gene Expression Omnibus [GEO] GSE199464), ATAC-Seq (GEO GSE201968), CUT&RUN (GEO GSE17138), and CUT&Tag (GEO GSE224904) are described in Supplemental Methods.

**Statistics.** qRT-PCR and flow cytometric analysis results were presented as the mean  $\pm$  SEM. Statistical comparisons for RT-qPCR used unpaired 2-tailed Student's *t* tests with Benjamini-Hochberg correction. For quantitative flow cytometric analysis, Dunnett's multiple-comparison test was performed to vehicle-treated control. Western blot signals were presented as box-and-whisker plots with bounds from the 25th to the 75th percentiles, the median line, and whiskers ranging from minimum to maximum values. Dunnett's test (for GATA2 rescue) or paired 2-tailed Student's *t* tests with Benjamini-Hochberg correction (for cytokine signaling) were performed to control conditions. All statistics had significance cutoff of *P* less than 0.05 and were calculated using Prism software (GraphPad Software).

**Study approval.** Animal protocols were approved by the University of Wisconsin–Madison IACUC in accordance with the Association for Assessment and Accreditation of Laboratory Animal Care International regulations.

## Author contributions

MMJ and EHB designed the study and wrote the manuscript with input from all authors. MMJ performed experiments and the analyzed data. SS and SK designed analytical strategies and conducted statistical/computational analyses. RAB generated GATA2 CUT&RUN data. PL analyzed RNA-Seq data. GAB and JX conducted ATAC-Seq. TO and MB conducted CUT&Tag. TO and QZ analyzed CUT&Tag data. ZDJ and PWL facilitated generation of

ChIP-Seq data. JPM generated the GATA2 molecular model. JKLL, LY, and JPM conducted EMSA. KRK generated cell lines, HA-GATA2, and 9aa-Ins and conducted ChIP-qPCR. KDJ conducted EBF1 and PU.1 experiments. AAS and KDJ facilitated flow analyses. LAG identified the 9aa-Ins patient.

## Acknowledgments

This work is supported by NIH grants R01DK068634, R01DK050107, U01CA257666, and Carbone Cancer Center P30CA014520, an ASH Scholar Award (to AAS), the Leukemia and Lymphoma Society Career Development Program (to AAS), NIH shared instru-

mentation grant 1S1ORR025483-01, and Edward P. Evans Foundation grants. GAB was supported by an American Heart Association Predoctoral Fellowship (827324). JX is a Scholar of the Leukemia Lymphoma Society and is supported by NIH grants R01DK111430, R01CA230631, R01CA259581, and R21AI158240 and CPRIT grants RP190417, RP220337, and RP220375.

Address correspondence to: Emery H. Bresnick, University of Wisconsin School of Medicine and Public Health, 1111 Highland Avenue, 4009 WIMR, Madison, Wisconsin 53705, USA. Phone: 608.265.6446; Email: ehbresni@wisc.edu.

- Klug A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem.* 2010;79:213–231.
- Wu H, et al. Building zinc fingers by selection: toward a therapeutic application. *Proc Natl Acad Sci U S A.* 1995;92(2):344–348.
- Gillinder KR, et al. Promiscuous DNA-binding of a mutant zinc finger protein corrupts the transcriptome and diminishes cell viability. *Nucleic Acids Res.* 2017;45(3):1130–1143.
- Lentaigne C, et al. Germline mutations in the transcription factor IKZF5 cause thrombocytopenia. *Blood.* 2019;134(23):2070–2081.
- Shaw-Smith C, et al. GATA4 mutations are a cause of neonatal and childhood-onset diabetes. *Diabetes.* 2014;63(8):2888–2894.
- Hughes MR, et al. Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. *Science.* 1988;242(4886):1702–1705.
- Borg J, et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet.* 2010;42(9):801–805.
- Nichols KE, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA-1. *Nat Genet.* 2000;24(3):266–270.
- Omichinski JG, et al. A small single-“finger” peptide from the erythroid transcription factor GATA-1 binds specifically to DNA as a zinc or iron complex. *Proc Natl Acad Sci U S A.* 1993;90(5):1676–1680.
- Drummond IA, et al. DNA recognition by splicing variants of the Wilms’ tumor suppressor, WT1. *Mol Cell Biol.* 1994;14(6):3800–3809.
- Katsumura KR, et al. The GATA factor revolution in hematology. *Blood.* 2017;129(15):2092–2102.
- Evans T, Felsenfeld G. The erythroid-specific transcription factor Eryfl: a new finger protein. *Cell.* 1989;58(5):877–885.
- Tsai SF, et al. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature.* 1989;339(6224):446–451.
- Ko LJ, Engel JD. DNA-binding specificities of the GATA transcription factor family. *Mol Cell Biol.* 1993;13(7):4011–4022.
- Merika M, Orkin SH. DNA-binding specificity of GATA family transcription factors. *Mol Cell Biol.* 1993;13(7):3999–4010.
- Fujiwara T, et al. Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell.* 2009;36(4):667–681.
- Hewitt KJ, et al. The hematopoietic stem and progenitor cell cisrome: GATA factor-dependent cis-regulatory mechanisms. *Curr Top Dev Biol.* 2016;118:45–76.
- Tsang AP, et al. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell.* 1997;90(1):109–119.
- Crispino JD, et al. Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex. *Mol Cell.* 1999;3(2):219–228.
- Cantor AB, et al. Distinct domains of the GATA-1 cofactor FOG-1 differentially influence erythroid versus megakaryocytic maturation. *Mol Cell Biol.* 2002;22(12):4268–4279.
- Gao X, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med.* 2013;210(13):2833–2842.
- de Pater E, et al. Gata2 is required for HSC generation and survival. *J Exp Med.* 2013;210(13):2843–2850.
- Rodrigues NP, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood.* 2005;106(2):477–484.
- Ling KW, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med.* 2004;200(7):871–882.
- Dickinson RE, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood.* 2011;118(10):2656–2658.
- Hsu AP, 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.
- Hahn CN, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet.* 2011;43(10):1012–1017.
- Ostergaard P, et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat Genet.* 2011;43(10):929–931.
- Johnson KD, et al. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J Clin Invest.* 2012;122(10):3692–3704.
- Bresnick EH, et al. Human GATA2 mutations and hematologic disease: how many paths to pathogenesis? *Blood Adv.* 2020;4(18):4584–4592.
- Chong CE, et al. Differential effects on gene transcription and hematopoietic differentiation cor-
- relate with GATA2 mutant disease phenotypes. *Leukemia.* 2018;32(1):194–202.
- Katsumura KR, et al. Molecular basis of cross-talk between oncogenic Ras and the master regulator of hematopoiesis GATA-2. *EMBO Rep.* 2014;15(9):938–947.
- Katsumura KR, et al. Human leukemia mutations corrupt but do not abrogate GATA-2 function. *Proc Natl Acad Sci U S A.* 2018;115(43):E10109–E10118.
- Zhang SJ, et al. Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci U S A.* 2008;105(6):2076–2081.
- Li L, et al. HMG1A1 chromatin regulators induce transcriptional networks involved in GATA2 and proliferation during MPN progression. *Blood.* 2022;139(18):2797–2815.
- Cavalcante de Andrade Silva M, et al. Breaking the spatial constraint between neighboring zinc fingers: a new germline mutation in GATA2 deficiency syndrome. *Leukemia.* 2021;35(1):264–268.
- DeVilbiss AW, et al. Navigating transcriptional coregulator ensembles to establish genetic networks: a GATA factor perspective. *Curr Top Dev Biol.* 2016;118:205–244.
- Treiber T, et al. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poisoning of chromatin. *Immunity.* 2010;32(5):714–725.
- Kitamura T, et al. Expression cloning of the human IL-3 receptor cDNA reveals a shared beta subunit for the human IL-3 and GM-CSF receptors. *Cell.* 1991;66(6):1165–1174.
- Hayashida K, et al. Molecular cloning of a second subunit of the receptor for human granulocyte-macrophage colony-stimulating factor (GM-CSF): reconstitution of a high-affinity GM-CSF receptor. *Proc Natl Acad Sci U S A.* 1990;87(24):9655–9659.
- Tavernier J, et al. A human high affinity interleukin-5 receptor (IL5R) is composed of an IL5-specific alpha chain and a beta chain shared with the receptor for GM-CSF. *Cell.* 1991;66(6):1175–1184.
- Dranoff G, et al. Involvement of granulocyte-macrophage colony-stimulating factor in pulmonary homeostasis. *Science.* 1994;264(5159):713–716.
- Stanley E, et al. Granulocyte/macrophage colony-stimulating factor-deficient mice show no major perturbation of hematopoiesis but develop a characteristic pulmonary pathology. *Proc Natl Acad Sci U S A.* 1994;91(12):5592–5596.
- Dickinson RE, et al. The evolution of cellular deficiency in GATA2 mutation. *Blood.*

- 2014;123(6):863–874.
45. Spinner MA, et al. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. *Blood*. 2014;123(6):809–821.
  46. Yamasaki K, et al. Cloning and expression of the human interleukin-6 (BSF-2/IFN beta 2) receptor. *Science*. 1988;241(4867):825–828.
  47. Taga T, et al. Interleukin-6 triggers the association of its receptor with a possible signal transducer, gp130. *Cell*. 1989;58(3):573–581.
  48. Caiado F, et al. Inflammation as a regulator of hematopoietic stem cell function in disease, aging, and clonal selection. *J Exp Med*. 2021;218(7):e20201541.
  49. Murakami M, et al. Pleiotropy and specificity: insights from the interleukin 6 family of cytokines. *Immunity*. 2019;50(4):812–831.
  50. Mei Y, et al. Bone marrow-confined IL-6 signaling mediates the progression of myelodysplastic syndromes to acute myeloid leukemia. *J Clin Invest*. 2022;132(17):e152673.
  51. Omichinski JG, et al. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science*. 1993;261(5120):438–446.
  52. Kowalski K, et al. The solution structure of the N-terminal zinc finger of GATA-1 reveals a specific binding face for the transcriptional co-factor FOG. *J Biomol NMR*. 1999;13(3):249–262.
  53. Wilkinson-White L, et al. GATA1 directly mediates interactions with closely spaced pseudopalindromic but not distantly spaced double GATA sites on DNA. *Protein Sci*. 2015;24(10):1649–1659.
  54. Chen Y, et al. DNA binding by GATA transcription factor suggests mechanisms of DNA looping and long-range gene regulation. *Cell Rep*. 2012;2(5):1197–1206.
  55. Johnson KD, et al. Constructing and deconstructing GATA2-regulated cell fate programs to establish developmental trajectories. *J Exp Med*. 2020;217(11):e20191526.
  56. Johnson KD, et al. Cis-regulatory mechanisms governing stem and progenitor cell transitions. *Sci Adv*. 2015;1(8):e1500503.
  57. Mehta C, et al. Integrating enhancer mechanisms to establish a hierarchical blood development program. *Cell Rep*. 2017;20(12):2966–2979.
  58. Freed-Pastor WA, Prives C. Mutant p53: one name, many proteins. *Genes Dev*. 2012;26(12):1268–1286.
  59. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
  60. Welch JJ, et al. Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood*. 2004;104(10):3136–3147.
  61. Tanimura N, et al. Mechanism governing heme synthesis reveals a GATA factor/heme circuit that controls differentiation. *EMBO Rep*. 2016;17(2):249–265.
  62. Yu M, et al. Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell*. 2009;36(4):682–695.
  63. Cheng Y, et al. Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res*. 2009;19(12):2172–2184.
  64. Tanimura N, et al. GATA/Heme multi-omics reveals a trace metal-dependent cellular differentiation mechanism. *Dev Cell*. 2018;46(5):581–594.
  65. Liao R, et al. Discovering how heme controls genome function through heme-omics. *Cell Rep*. 2020;31(13):107832.
  66. An X, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*. 2014;123(22):3466–3477.
  67. Hollenhorst PC, et al. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem*. 2011;80:437–471.
  68. Onodera K, et al. Conserved structure, regulatory elements, and transcriptional regulation from the GATA-1 gene testis promoter. *J Biochem*. 1997;121(2):251–263.
  69. Heinz S, et al. Simple combinations of lineage determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576–589.
  70. Gentleman RC, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 2004;5(10):R80.
  71. Hagman J, et al. Cloning and functional characterization of early B-cell factor, a regulator of lymphocyte-specific gene expression. *Genes Dev*. 1993;7(5):760–773.
  72. Lin H, Grosschedl R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature*. 1995;376(6537):263–267.
  73. Zhang P, et al. Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc Natl Acad Sci U S A*. 1999;96(15):8705–8710.
  74. Walsh JC, et al. Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates. *Immunity*. 2002;17(5):665–676.
  75. Roessler S, et al. Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Mol Cell Biol*. 2007;27(2):579–594.
  76. Wang H, et al. IRF8 regulates B-cell lineage specification, commitment, and differentiation. *Blood*. 2008;112(10):4028–4038.
  77. Li Y, et al. Regulation of the PU.1 gene by distal elements. *Blood*. 2001;98(10):2958–2965.
  78. Rosenbauer F, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet*. 2004;36(6):624–630.
  79. Robb L, et al. Hematopoietic and lung abnormalities in mice with a null mutation of the common beta subunit of the receptors for granulocyte-macrophage colony-stimulating factor and interleukins 3 and 5. *Proc Natl Acad Sci U S A*. 1995;92(21):9565–9569.
  80. Choudhuri A, et al. Common variants in signaling transcription-factor-binding sites drive phenotypic variability in red blood cell traits. *Nat Genet*. 2020;52(12):1333–1345.
  81. Canver MC, et al. Variant-aware saturating mutagenesis using multiple Cas9 nucleases identifies regulatory elements at trait-associated loci. *Nat Genet*. 2017;49(4):625–634.
  82. Kang YA, et al. Autophagy driven by a master regulator of hematopoiesis. *Mol Cell Biol*. 2012;32(1):226–239.
  83. Soukup AA, et al. Single-nucleotide human disease mutation inactivates a blood-regenerative GATA2 enhancer. *J Clin Invest*. 2019;129(3):1180–1192:122694.
  84. Zhang TY, et al. IL-6 blockade reverses bone marrow failure induced by human acute myeloid leukemia. *Sci Transl Med*. 2020;12(538):eaax5104.
  85. Al Seraihi AF, et al. GATA2 monoallelic expression underlies reduced penetrance in inherited GATA2-mutated MDS/AML. *Leukemia*. 2018;32(11):2502–2507.
  86. Wadman IA, et al. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J*. 1997;16(11):3145–3157.
  87. Wozniak RJ, et al. Context-dependent GATA factor function: combinatorial requirements for transcriptional control in hematopoietic and endothelial cells. *J Biol Chem*. 2007;282(19):14665–14674.
  88. Wozniak RJ, et al. Molecular hallmarks of endogenous chromatin complexes containing master regulators of hematopoiesis. *Mol Cell Biol*. 2008;28(21):6681–6694.
  89. Hewitt KJ, et al. Hematopoietic signaling mechanism revealed from a stem/progenitor cell cis-trome. *Mol Cell*. 2015;59(1):62–74.
  90. Tijssen MR, et al. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev Cell*. 2011;20(5):597–609.
  91. Marks-Bluth J, et al. SMAD1 and SMAD5 expression is coordinately regulated by FLI1 and GATA2 during endothelial development. *Mol Cell Biol*. 2015;35(12):2165–2172.
  92. Wang X, et al. Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors. *EMBO J*. 2002;21(19):5225–5234.
  93. Chou ST, et al. Graded repression of PU.1/Sfp1 gene transcription by GATA factors regulates hematopoietic cell fate. *Blood*. 2009;114(5):983–994.
  94. Wheat JC, et al. Single-molecule imaging of transcription dynamics in somatic stem cells. *Nature*. 2020;583(7816):431–436.
  95. Ohmori S, et al. GATA2 and PU.1 collaborate to activate the expression of the mouse *Ms4a2* gene, encoding FcεRIβ, through distinct mechanisms. *Mol Cell Biol*. 2019;39(22):e00314–19.
  96. Souabni A, et al. Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. *Immunity*. 2002;17(6):781–793.
  97. Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem*. 2002;71:101–132.
  98. Itoh-Nakadai A, et al. The transcription repressors Bach2 and Bach1 promote B cell development by repressing the myeloid program. *Nat Immunol*. 2014;15(12):1171–1180.
  99. Tsai FY, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371(6494):221–226.
  100. 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(10):3636–3643.
  101. Menendez-Gonzalez JB, et al. Gata2 as a crucial regulator of stem cells in adult hematopoiesis and acute myeloid leukemia. *Stem Cell Reports*.

- 2019;13(2):291–306.
102. Rodrigues NP, et al. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood*. 2008;112(13):4862–4873.
103. Soukup AA, et al. Conditionally pathogenic genetic variants of a hematopoietic disease-suppressing enhancer. *Sci Adv*. 2021;7(50):eabk3521.
104. Wu Z, et al. Sequencing of RNA in single cells reveals a distinct transcriptome signature of hematopoiesis in GATA2 deficiency. *Blood Adv*. 2020;4(12):2656–2670.
105. Linnemann AK, et al. Genetic framework for GATA factor function in vascular biology. *Proc Natl Acad Sci U S A*. 2011;108(33):13641–13646.
106. Jing H, et al. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol Cell*. 2008;29(2):232–242.
107. Kurotaki D, Tamura T. Transcriptional and epigenetic regulation of innate immune cell development by the transcription factor, interferon regulatory factor-8. *J Interferon Cytokine Res*. 2016;36(7):433–441.
108. Johnson KD, et al. GATA2 deficiency elevates interferon regulatory factor-8 to subvert a progenitor cell differentiation program. *Blood Adv*. 2022;6(5):1464–1473.
109. Hibi M, et al. Molecular cloning and expression of an IL-6 signal transducer, gp130. *Cell*. 1990;63(6):1149–1157.
110. Manz MG, Boettcher S. Emergency granulopoiesis. *Nat Rev Immunol*. 2014;14(5):302–314.
111. Gardenghi S, et al. Distinct roles for hepcidin and interleukin-6 in the recovery from anemia in mice injected with heat-killed *Brucella abortus*. *Blood*. 2014;123(8):1137–1145.
112. Nemeth E, et al. IL-6 mediates hypoferrremia of inflammation by inducing the synthesis of the iron regulatory hormone hepcidin. *J Clin Invest*. 2004;113(9):1271–1276.
113. Carey A, et al. Identification of interleukin-1 by functional screening as a key mediator of cellular expansion and disease progression in acute myeloid leukemia. *Cell Rep*. 2017;18(13):3204–3218.
114. Kim I, et al. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell*. 2007;130(3):470–483.
115. Bowie MB, et al. Identification of a new intrinsically timed developmental checkpoint that reprograms key hematopoietic stem cell properties. *Proc Natl Acad Sci U S A*. 2007;104(14):5878–5882.
116. Porter SN, et al. Fetal and neonatal hematopoietic progenitors are functionally and transcriptionally resistant to *Flt3*-ITD mutations. *Elife*. 2016;5:e18882.
117. Dougan M, et al. GM-CSF, IL-3, and IL-5 family of cytokines: regulators of inflammation. *Immunity*. 2019;50(4):796–811.
118. Charlet A, et al. The IL-3, IL-5, and GM-CSF common receptor beta chain mediates oncogenic activity of FLT3-ITD-positive AML. *Leukemia*. 2022;36(3):701–711.
119. Michel T, et al. Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature*. 2001;414(6865):756–759.
120. Yashin DV, et al. Tag7 (PGLYRP1) in complex with Hsp70 induces alternative cytotoxic processes in tumor cells via TNFR1 receptor. *J Biol Chem*. 2015;290(35):21724–21731.
121. Cao C, et al. Structure, function and pharmacology of human itch GPCRs. *Nature*. 2021;600(7887):170–175.
122. Wang GG, et al. Quantitative production of macrophages or neutrophils ex vivo using conditional Hoxb8. *Nat Methods*. 2006;3(4):287–293.
123. Medina KL, et al. Assembling a gene regulatory network for specification of the B cell fate. *Dev Cell*. 2004;7(4):607–617.