haematologica
Journal of the Ferrata Storti Foundation

Enhancement of PRMT6 binding to a novel germline *GATA1* mutation associated with congenital anemia

by Yingsi Lu, Qingqing Zhu, Yun Wang, Meiling Luo, Junbin Huang, Qian Liang, Lifen Huang, Jing Ouyang, Chenxin Li, Nannan Tang, Yan Li, Tingting Kang, Yujia Song, Xiaoyu Xu, Liping Ye, Guoxing Zheng, Chun Chen, and Chengming Zhu

Received: September 4, 2023.
Accepted: February 14, 2024.

Citation: Yingsi Lu, Qingqing Zhu, Yun Wang, Meiling Luo, Junbin Huang, Qian Liang, Lifen Huang, Jing Ouyang, Chenxin Li, Nannan Tang, Yan Li, Tingting Kang, Yujia Song, Xiaoyu Xu, Liping Ye, Guoxing Zheng, Chun Chen, and Chengming Zhu. Enhancement of PRMT6 binding to a novel germline GATA1 mutation associated with congenital anemia.
Haematologica. 2024 Feb 22. doi: 10.3324/haematol.2023.284183 [Epub ahead of print]

Title page

Enhancement of PRMT6 binding to a novel germline *GATA1* mutation associated with congenital anemia

Running Title: Enhanced binding of PRMT6 to a novel mutant *GATA1*

Authors:

Yingsi Lu[1], Qingqing Zhu[1], Yun Wang[1], Meiling Luo[2], Junbin Huang[2], Qian Liang[3], Lifen Huang[2], Jing Ouyang[1], Chenxin Li[1], Nannan Tang[2], Yan Li[1], Tingting Kang[1], Yujia Song[1], Xiaoyu Xu[1,4], Liping Ye[1], #Guoxing Zheng[1], #Chun Chen[2] and #Chengming Zhu[1,*]

Affiliations:

[1]Scientific Research Center, The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen, Guangdong, China; [2]Pediatric Hematology Laboratory, Division of Hematology/Oncology, Department of Pediatrics, The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen, Guangdong, China; [3]Department of Spine Surgery, the First Affiliated Hospital, Shenzhen University, Shenzhen Second People's Hospital, Shenzhen, Guangdong, China; [4]Department of Obstetrics, The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen, Guangdong, China.

#Corresponding authors

Chengming Zhu, Scientific Research Center, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, China; E-mail address: zhuchm3@mail.sysu.edu.cn; and Chun Chen, Pediatric Hematology Laboratory, Division of Hematology/Oncology, Department of Pediatric, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, China; E-mail address: chenchun69@126.com; and Guoxing Zheng, Scientific Research Center, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, China; E-mail address: opqsky@126.com

Word Count Text: 3916

Word Count Abstract: 242

Figure count: 7 main manuscript, 6 supplemental figures

Table count: 0 main manuscript, 5 supplemental tables

Reference count: 48

Data-sharing statement

Processing data of quantitative mass spectrometry are available in Supplemental Table S5. All other data are available from the corresponding author upon reasonable request.

Acknowledgments

We thank members of the Zhu Lab for their feedback on this manuscript and Dr. Qiurong Yuan, laboratory physician at the Seventh Affiliated Hospital of Sun Yat-sen University, for her interpretation of the blood and bone marrow test results.

Author contributions

Y.Lu, Q.Z., Y.W., M.L., L.H., J.O., C.L., Y.Li, T.K., Y.S. performed experiments; Y.L., Y.W., J.H., C.Z. performed bioinformatics analyses; L.Y., N.T., X.X. managed and maintained research data; Y.Lu., C.Z. wrote the manuscripts; Y.Lu., L.Y., C.Z. revised the manuscripts; Y.Lu, G.Z., C.Z. designed the study; C.C., C.Z. directed the study. All authors contributed to the research and approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (82203807 and 82072905) and the Science, Technology and Innovation Commission of Shenzhen Municipality (JCYJ20180307150419435, JCYJ20210324123004011, JCYJ20220530144817040).

Conflict of Interest Disclosure

The authors declare no competing financial interests.

ABSTRACT

Mutations in the master hematopoietic transcription factor GATA1 are often associated with functional defects in erythropoiesis and megakaryopoiesis. In this study, we identified a novel GATA1 germline mutation (c.1162delGG, p.Leu387Leufs*62) in a patient with congenital anemia and occasional thrombocytopenia. The C-terminal GATA1, a rarely studied mutational region, undergoes frameshifting translation as a consequence of this double-base deletion mutation. To investigate the specific function and pathogenic mechanism of this mutant, in vitro mutant models of stable re-expression cells were generated. The mutation was subsequently validated to cause diminished transcriptional activity of GATA1 and defective differentiation of erythroid and megakaryocytes. Using proximity labeling and mass spectrometry, we identified selective alterations in the proximal protein networks of the mutant, revealing decreased binding to a set of normal GATA1-interaction proteins, including the essential co-factor FOG1. Notably, our findings further demonstrated enhanced recruitment of the protein arginine methyltransferase PRMT6, which mediates histone modification at H3R2me2a and represses transcription activity. We also found an enhanced binding of this mutant GATA1/PRMT6 complex to the transcriptional regulatory elements of GATA1's target genes. Moreover, treatment of the PRMT6 inhibitor MS023 could partially rescue the inhibited transcriptional and impaired erythroid differentiation caused by the GATA1 mutation. Taken together, our results provide molecular insights into erythropoiesis in which mutation leads to partial loss of GATA1 function and the broader role of PRMT6 and its inhibitor MS023 in congenital anemia, highlighting PRMT6 binding as a negative factor of GATA1 transcriptional activity in aberrant hematopoiesis.

INTRODUCTION

The differentiation process of erythrocytes and megakaryocytes shares the regulation of the hematopoietic transcription factor GATA1.[1] It activates globin genes and hemoglobin synthase to support the maturation of hemoglobin in erythroid differentiation during both embryonic development and adult stages.[2,3] Moreover, GATA1 plays a crucial role in regulating cell proliferation and apoptosis in erythroid cells by activating anti-proliferative and anti-apoptotic genes, ensuring stable erythroid development.[4,5] In the case of megakaryocytes, GATA1 is involved in their differentiation and plays a significant role in their proliferation, terminal maturation, and platelet production.[6-8]

Since Nichols[9] first reported a family affected with germline GATA1 V205M mutation, a series of congenital hematological disorders caused by GATA1 mutations have been observed in multiple families.[10-12] It is the discovery of these genetic cases and the exploration of the associated mechanisms that have propelled GATA1 to become one of the most extensively researched hematopoietic transcription factors to date. In earlier studies, the reported mutations were mainly focused on the N-terminal transactivation domain (TAD) and the two central zinc fingers (ZF). Mutations in the TAD are mainly related to Diamond-Blackfan anemia (DBA) with a short isoform of GATA1 (GATA1s),[13,14] while those in the zinc fingers are commonly reported in anemia or thrombocytopenia.[9,15] However, very few mutation cases and studies have been reported in the C-terminus of the GATA1 protein, and the specific function remains not fully defined.

Normal expression of erythrocyte- and megakaryocyte-associated genes depends on GATA1's interaction with co-factors and binding to GATA motifs in downstream genes' transcriptional regulatory regions, both of which are indispensable. A variety of molecules or complexes have been found to interact with GATA1 in transcriptional actions, including FOG1, known as the "Friend of GATA1",[16] the core subunit of the mediator complex MED1,[17] the hematopoietic transcription factor TAL1,[18] CBP/P300 acetyltransferase,[19-22] the SWI/SNF complex and the NuRD complex.[23]

PRMT6, a type I arginine methyltransferase, mediates the production of asymmetric dimethylarginines on histones, including H3R2me2a, which is recognized to be implicated in transcriptional repression.[24] Previous studies have found that PRMT6 binds to the transcription factor RUNX1 before megakaryocyte differentiation but disassociates during megakaryocyte differentiation and is recruited to the promoter of GYPA (CD235a).[25-27] During CD34+ hemopoiesis, PRMT6 inhibits erythroid gene expression and mediates H3R2me2a on GYPA and KLF1.[28] These results suggest that PRMT6 may act as a co-repressor by binding to hematopoietic transcription factors and participating in the regulation of lineage differentiation.

In this study, we identified a novel germline mutation of GATA1 in a patient with long-term transfusion-dependent anemia and occasional thrombocytopenia. Subsequently, we demonstrated in vitro that the mutation impeded the process of normal erythropoiesis and megakaryopoiesis. By employing the Turbo-ID proximity labeling approach to study protein interaction networks in the mutated state, we identified selective alterations in a set of established GATA1-interacting proteins and complexes with increased PRMT6 binding. These alterations are believed to be responsible for the compromised hematopoietic differentiation that results from the mutation.

METHODS

Ethical considerations

The study adhered to ethical considerations by obtaining informed consent from the parents of the proband and receiving approval from the Ethics Committee of the Seventh Affiliated Hospital, Sun Yat-sen University. The study followed the declaration of Helsinki and obtained the necessary permissions from the family to use clinical information. Bone marrow aspirates were obtained during diagnosis.

Cell culture

Cell lines of HEK-293T, HEL, and K562 cells were cultured as previously described.[29] All cell lines had been authenticated by STR analysis and were cultured with routine tests conducted for mycoplasma contamination by PCR-based assays (Yeason, Shanghai).

Constructs, lentivirus production, and infection

Construction methods of plasmids for GATA1 stable knockout, wild-type and mutant GATA1 stable overexpression, GATA1-Turbo-ID assay, PRMT6 overexpression, and knockdown are detailed in the supplemental file. For details of the sgRNAs and shRNAs sequence, see Supplemental Table S4. Lentiviruses preparation and modified spin-infection for suspension cells were performed as previously described.[30,31]

Deletion of GATA1 and the generation of re-expression stable cell models

CRISPR/Cas9-mediated GATA1 knockout in K562 and HEL cells (represented by K562-sgG1 and HEL-sgG1, respectively) was performed by lentiviral transduction with lentiCRISPR-v2-sgGATA1 plasmids. Then, re-expression was performed by lentivirus of wild-type and mutant GATA1, separately, in GATA1-depleted K562 and HEL cells.

Induction of erythroid and megakaryocytic differentiation

The erythroid induction of K562 and HEL cells was facilitated through hemin (MCE, Shanghai) treatment. Megakaryocytic induction of K562 and HEL cells was achieved by PMA (MCE, Shanghai). Benzidine cytochemical staining was performed to assess the level of hemoglobin in erythroid differentiation as previously described.[32]

Turbo-ID proximity labeling

Turbo-ID proximity labeling-based western blotting and mass spectrometry were

performed with the plasmid carrying a Turbo-ID fused wild-type or mutant GATA1 protein according to the previous description.[33] The EASY-nLCTM 1200 (Thermo Fisher, Germany) coupled with Orbitrap Exploris 480 (Thermo Fisher, Germany) platform was utilized for label-free mass spectrometry detection at Novogene Co., Ltd. (Tianjin, China).

Complete and detailed materials and methods of whole exome sequencing (WES) and Sanger sequencing, FACS analysis, RT-qPCR, Western blotting, Chromatin immunoprecipitation, and Co-immunoprecipitation analysis were described in the supplemental file.

Statistical analysis

Quantification of western blots was performed by ImageJ Software. All statistical tests were performed using SPSS 20.0 (IBM, USA) and GraphPad Prism 9 (GraphPad Software, CA) with a significance level of 0.05. After confirming the Shapiro–Wilk normality test, a t-test was performed. Mann–Whitney U test was used to compare sample groups for which normality was rejected.

RESULTS

1. Case description: a novel germline GATA1 mutation (c.1162delGG, p.Leu387Leufs*62) identified in a family associated with congenital anemia.

The patient investigated in the study was a 7-month-old male child with anemia for more than 6 months. The first bone marrow evaluation at 45 days found erythroid hyperplasia and thrombocytopenia. The second bone marrow stain at 6 months of age showed hypoplasia with a lower ratio of granulocytes to erythrocytes (17% vs. 1.5%) (Supplemental Figure S1A). The morphology of all cell types was generally normal in the bone marrow analysis (Supplemental Figure S1A). Relevant hematological

findings comprised normal hemoglobin electrophoresis results (HbA 78.0%, HbF 19.9%, HbA2 2.1%, at 45 days of age), normal G6PD level (4723 U/L), and negative direct Coombs test. Before transfusion, the patient exhibited consistently low RBC count (1.84-3.87×1012/L), hematocrit (18-35.3%), hemoglobin (56-110 g/L), and reticulocyte hemoglobin equivalent. The reticulocyte count and mean circulating volume (MCV) showed normal (80.8–97.9 fL) (Supplemental Table S1). Additionally, the patient had an occasionally low platelet count and MPV, persistently high lymphocyte count, and low eosinophil count (Supplemental Table S2). After haploidentical stem cell transplantation (haplo-HSCT) with post-transplant cyclophosphamide (PTCy) protocol at 1 year old, the patient became transfusion independent.

The patient's mother was found with anemia in the sixth month of gestation. His nonconsanguineous parents had three children, and their first male child, who also had anemia and short-term thrombocytopenia, died of severe anemia two months after birth (Supplemental Table S3). Whole exome sequencing (WES) was performed with peripheral blood from the patient and his parents to confirm the genetic cause of the anemia. A novel GATA1 mutation (c.1162delGG, p.Leu387Leufs*62, represented by Leu387fs below), which both the patient and his mother carried, was identified and confirmed by Sanger sequencing (Figure 1A). The variant has not been recorded in the ClinVar, HGMD mutation database, or the available literature. Its frequency data is absent in the 1,000 Genomes, ExAC, and gnomAD population database, indicating that the variant might be rare. For this mutation, the anemic proband was hemizygous, mother heterozygous, and father wild-type (Figure 1A and 1D). This mutation manifested as a double base deletion in the GATA1 coding region, caused a frameshift and extended translation (Figure 1B), and resulted in altered multiple species-conservative amino acids in the C-terminus (Figure 1C).

Given the crucial roles of GATA1 in the erythroid and megakaryocytic differentiation, its germline mutations are often associated with hematological disorders like anemia and thrombocytopenia,[10,34] consistent with the patient's anemic

phenotype and history of platelet reduction. Since its location on the X-chromosome, the germline mutation of the GATA1 gene has an X-linked inheritance pattern, which is compatible with the family (Figure 1D). Furthermore, a reduction of GATA1 expression level was found in the patient's bone marrow and peripheral blood, compared with healthy control (Figure 1E). Overall, these results suggest that the proband's anemia may be attributable to a decreased GATA1 dose and not just the mutation itself.

2.GATA1 Leu387fs mutation leads to impaired erythroid differentiation and increased apoptosis during erythropoiesis.

To investigate and clarify the effect of the identified GATA1 mutation, we employed the K562 and HEL cells, which were derived from myeloid leukemia patients,[35,36] express full-length wild-type GATA1,[37] and are capable of targeted differentiation into erythroid or megakaryocytes in vitro, as research tools.[38,39] We first stably depleted GATA1 in K562 and HEL cells via lentiCRISPR/cas9–v2 system (Supplemental Figure S2A). Subsequently, we rescued these cells with wild-type or Leu387fs mutant GATA1 to create stable cellular models to assess the mutational effects on erythropoiesis and megakaryopoiesis (Figure 2A, Supplemental Figure S2B-C).

Given that a low level of spontaneous erythroid differentiation could occur in general K562 cells,[40] during cell culture, we observed that Leu387fs mutant cell pellets displayed a paler red color in comparison with the wild-type GATA1 cells and a darker red compared to GATA1-depleted cells (Figure 2B), implying defects of erythropoiesis with cells in the mutant state. To further explore the effect of GATA1 Leu387fs on erythropoiesis, we introduced hemin to stimulate erythroid maturation in both cell groups. As compared to the wild-type GATA1 controls, the mutant cells showed reduced hemoglobin production before and after induction (Figure 2C and Supplemental Figure S3A-B). The CD71+CD235a+ erythroid populations decreased remarkably during hemin-induced differentiation in mutant cells (Figure 2D and

Supplemental Figure S3C), consistent with the performance of K562 mutant cells without induction (Supplemental Figure S3D), indicating the Leu387fs mutation contributes to blocked erythroid differentiation.

Moreover, CD235a and the apoptotic marker Annexin V were used to label apoptotic erythrocytes in the flow assay to assess erythroid apoptosis. Both the mutant cells and the GATA1-depleted cells showed a distinct increase of apoptotic cells (AC) and the ratio of erythroid apoptotic cells (%EAC) during differentiation (Figure 2E and Supplemental Figure S3E-F).

Further evaluation of the transcriptional impact by GATA1 Leu387fs revealed that a subset group of GATA1 target genes, including a series of globin, heme synthase ALAS2, and ferrochelatase FECH, were detected with reduced expression in the mutant state in comparison to wild-type GATA1 (Figure 2F and Supplemental Figure S3G), indicating attenuated erythroid transcription activity of the GATA1 Leu387fs mutation. Taken together, these results demonstrated that the Leu387fs mutation caused a diminished transcriptional activity of GATA1 and a reduction in hemoglobin synthesis, which led to an impeded erythroid differentiation process and was associated with increased cell apoptosis.

3. The GATA1 Leu387fs mutation affects megakaryocytic differentiation.

To investigate the effect of the Leu387fs mutation on megakaryocyte differentiation, we used PMA as an inducer of megakaryocyte differentiation and treated both groups of K562 and HEL cells for 96h. Both groups of cells showed a significant reduction of $CD41^{high}$ megakaryocytic population in mutant cells (Figure 3A). To analyze ploidy changes during megakaryocytic differentiation in both groups under Leu387fs mutation, measurement of cellular DNA content was performed by flow cytometry analysis of PI dye showed a significant decrease of high ploidy populations (≥8N) by the GATA1 Leu387fs mutation, suggesting a defect in polyploidization during megakaryocyte differentiation (Figure 3B).

Moreover, the transcriptional expression of megakaryocyte- and platelet-related

genes, including platelet agglutinin GPIIb/GPIIIa (ITGA2B and ITGB3), thrombopoietin (TPOR, c-MPL) and megakaryocyte-related transcription factor FLI1, were further analyzed in both groups of model cells. The mutation suppressed the expression of ITGA2B and TPOR compared to wild-type GATA1, while the expression of megakaryocyte transcription factors FLI1 was diminished only in K562 cells by the mutation (Figure 3C). Altogether, these results indicated that the Leu387fs mutation could block megakaryocytic differentiation and impact some platelet-functional gene expression.

4. Alterations of the proximity protein networks by GATA1 Leu387fs mutation were found via Turbo-ID proximity labeling.

During specific transcriptional activities, GATA1 is required to assemble into transcriptional complexes by recruiting multiple co-factors/repressors. To investigate the exact mechanism of the hematopoietic defects and how the original protein networks would be changed in the GATA1 Leu387fs mutant state, the biotin ligase Turbo-ID proximity labeling method was applied. The wild-type and Leu387fs mutant GATA1-turboID fusion proteins were overexpressed separately in the GATA1-depleted K562 cells. Then the biotin-labeled proximal proteins were collected for mass spectrometry to identify the differential proteins of proximal networks between the wild-type and the mutant GATA1, seeking to discover the critical factors in the process of cell differentiation under hematopoietic abnormality (Figure 4A).

Preceding the mass spectrometry assays, the labeling system was tested to evaluate its efficiency, optimize the labeling time, and confirm that the fused Turbo-ID proteins remained in nuclear after transfection (Supplemental Figure S4A-B). Upon acquisition of mass spectrometry results, we found that the proximal proteins of the wild-type and mutant GATA1 were similar in properties, with only seven proteins unique to either group out of over 3000 identified in each sample (Supplemental Figure S4C). When comparing the quantitative results of proximal proteins in two groups, significantly increased and decreased proteins were found in the GATA1

Leu387fs group (Supplemental Figure S4D).

Subsequent analysis revealed that the reduced proximal protein in the Leu387fs mutation comprised a subset of established GATA1co-factors, including FOG1, an essential co-factor of GATA1, ATF2, a member of the CBP/P300 histone acetyltransferase complex that activates the promoter of GATA1 target genes, MED1, a core subunit of the intermediate complex that mediates the linkage between transcription factors and POLR1A and POLR2B, two subunits encoding RNA polymerases (Figure 4B). Moreover, multiple subunits of the SWI/SNF chromatin remodeling complex and two subunits of the NuRD complex, MTA2 and CHD4, also remarkably reduced in the Leu387fs mutation (Figure 4C). These data on altered binding of the known GATA1-interacting proteins/complexes appear to explain part of the findings in the above study, whereby changes in the binding of the GATA1 Leu387fs mutation to these known co-factors, lead to impaired GATA1-dominated transcriptional activity, ultimately causing dysfunctional differentiation.

Further investigation was conducted to analyze the enriched pathway and interaction network in the differential proximity proteins. These results indicated that the GATA1 Leu387fs mutation interfered with various mRNA-related pathways, which may further explain the transcriptional abnormalities caused by the mutation (Supplemental Figure S4E-F). The interacting networks significantly reduced by Leu387fs mutation included regulation of transcriptional activity, erythroid/megakaryocyte differentiation, mRNA splicing, and the electron transport chain, whereas, in the increased networks, we found alterations in hemoglobin, regulation of mRNA stability and methylation-related signals (Figure 4D). Within the mutation-increased methylation network of proximal proteins, we identified PRMT6, a member of the protein arginine methyltransferase family (PRMTs), which mediates the repressive histone H3R2 asymmetric dimethylation (H3R2me2a), previously reported to be closely related with erythroid and megakaryocytic differentiation, implying its enhanced binding to mutant GATA1 may play a role in the aberrant differentiation caused by the mutation (Figure 4D and Figure 5A).

5. Recruitment to PRMT6 was enhanced by the GATA1 Leu387fs mutation.

According to the search results of the Human Protein Atlas and Bloodspot databases, PRMT6 had low tissue specificity (Supplemental Figure S5A). Still, in bone marrow, PRMT6 showed high expression in erythroid cells (Supplemental Figure S5B-C), particularly in megakaryocyte-erythroid progenitors (Supplemental Figure S5D-F), implying that PRMT6 plays a role at this stage of differentiation. We then examined the binding of wild-type GATA1 to PRMT6 in K562 cells that underwent induced erythroid or megakaryocytic differentiation. As a result, we observed attenuated binding of GATA1 to PRMT6 under both differential conditions, which however could be due to the reduced PRMT6 expression after differentiation (Figure 5B).

To validate the mass spectrometry results, we performed the pull-down WB analysis with GATA1-depleted K562 cells based on the Turbo-ID method. Consistent with the mass spectrometry findings, the GATA1 Leu387fs mutation had an increased proximal PRMT6 compared to the wild type (Figure 5C). We also found the labeled biotinylated PRMT6 in the negative control, suggesting that PRMT6 may be a naturally biotinylated protein (Figure 5C).

Subsequently, Co-immunoprecipitation analysis revealed that Leu387fsGATA1 interacts with PRMT6 and is enhanced compared to wild-type GATA1. At the same time, the binding of mutant GATA1 to FOG1 was also weakened in the immunoprecipitation test (Figure 5D). Interaction studies using *E.coli*-produced GST-PRMT6 with in vitro-translated HA-tagged GATA1 indicated a weak direct interaction between the mutated GATA1 and GST-PRMT6 but significantly enhanced compared to wild-type GATA1 (Figure 5E). Preliminary exploration of the interaction domain between PRMT6 and mutant GATA1 by immunoprecipitation assays revealed that the mutant GATA1 interacts with amino acids 86-375 of PRMT6 (Figure 5F), while PRMT6 could interact with the N-terminal zinc finger of GATA1 (amino acids 84-258), the C-terminus of wild-type GATA1 (amino acids 292-413) and the C-

terminus of mutant GATA1 (292-449) (Figure 5G).

6. GATA1 Leu387fs leads to enhancement of PRMT6 and its driven H3R2me2a modification binding at the promoter/enhancer of the erythroid target genes.

In our study, the mutation was found to suppress CD235a (GYPA) and multiple globin gene expression. Upon the β-globin gene cluster, DNase I hypersensitive sites (HSs) in the locus control region (LCR) serve as distal enhancers that provide binding sites for GATA1, mediating the formation of loop chromatin structures (LCR-promoter loops) and activate downstream globin gene expression.[41,42] To elucidate the involvement of increased recruitment of PRMT6 by GATA1 Leu387fs in regulating the transcriptional activity of erythroid target genes, chromatin immunoprecipitation assays were performed in the re-expression K562 cells. CD235a, HBB, HBG1, and distal enhancer of β-globin genes LCR-HS2 were selected as target genes.

Within the results, we found that the Leu387fs mutation had no discernible impact on the binding of GATA1 itself to the promoters/distal enhancers of the target genes (Figure 6A). Significant binding was detected on these regulatory elements in wild-type and mutant GATA1 with comparable levels (Figure 6A). However, the binding of PRMT6 and its mediated repressive modification H3R2me2a to the regulatory elements of these target genes was significantly enhanced in mutant GATA1 cells compared to wild-type cells (Figure 6B-C). Similar results were observed in cells stimulated with hemin (Supplemental Figure 6A-B). In addition, we found that the binding of PRMT6 and H3R2me2a modifications was almost always reduced upon hemin induction in both wild-type and mutant cells (Data not shown). Together, these results suggest that the binding of PRMT6 and H3R2me2a to erythroid gene activation elements impedes the process of erythroid differentiation, and the mutation-induced transcriptional repression may be attributed to the increase of PRMT6 interaction with GATA1 Leu387fs.

7. Aberrant phenotypes in erythroid differentiation caused by GATA1 Leu387fs could

be partially rescued with PRMT6 inhibitor MS023

To clarify the impact of increased PRMT6 binding and its mediated H3R2me2a on the GATA1 Leu387fs, we introduced the PRMT6 inhibitor MS023 in the mutational complemented cells for investigation. Knockdown of PRMT6 in K562 cells resulted in reduced H3R2me2a and increased levels of γ-globin (Figure 7A). In previous reports, H3R2me2a modification was antagonistic to the activating histone modification H3K4me3[24], yet we found a significant reduction of H3K4me3 levels, but not of H3R2me2a levels, in both GATA1-depleted and mutant cells (Figure 7B). Decreased expression of H3R2me2a and increased expression of γ-globin were observed in K562 complemented cells treated with different concentration gradients of MS023 (Figure 7C). The expression of γ-globin was reduced in GATA1 Leu387fs cells compared to wild-type cells, and this difference was noticeably reduced following MS023 treatment (Figure 7C).

While MS023 did not entirely reverse the decline in erythroid cells induced by the mutation, a more significant increase in the CD71+CD235a+ population was observed in two dosage groups compared to wild-type cells following MS023 treatment (Figure 7D). A comparable pattern was identified in the analysis of hemoglobin synthesis (Figure 7E). Finally, the transcription levels of GATA1 erythroid target genes before and after MS023 treatment were also analyzed, in which the expression of CD235a and β-globin gene in mutant cells was reversed after MS023 treatment (Figure 7F). Though the expression of the γ-globin gene in the mutant cells remained lower following MS023 treatment than in the wild-type cells, its post-treatment gain in fold change of expression was significantly greater than that of the control cells (Figure 7F). Taking these results together, we concluded that PRMT6 inhibitor MS023 can partially rescue the erythroid defects caused by GATA1 Leu378fs mutation.

DISCUSSION

In this study, we identified a novel GATA1 germline mutation (c.1162delGG, p.Leu387Leufs*62) in a child with persistent anemia and occasional thrombocytopenia. Despite inducing a frameshift and the extending translation, this mutation maintains the complete integrity of the zinc fingers and the N-terminal transactivation domain of GATA1. Very few C-terminal GATA1 mutation reports and pathogenic investigations have been done. Out of the documented cases, T296P was found with thrombocytopenia and insignificant erythroid impact,[43] Stop414Arg caused an erythroid Lu(a-b-) phenotype and mild thrombocytopenia,[44] R307C/H caused mild thrombocytopenia and hemolytic anemia with elevated adenosine deaminase (ADA) levels.[45] Although these cases varied in appearance, they seemed to have a shared deficiency in platelet production. However, the mutational case we found primarily had anemia, but the thrombocytopenia did not persist, although we did find an impaired megakaryopoiesis by the mutation in vitro. The precise mechanism underlying the recovered platelet count of the patient remains unknown at this time and it could be the result of frequent transfusion or drug use.

Currently, the role of mutations in the C-terminal GATA1 has only been investigated to a limited extent. An early study involving transgenic mice with deletion of the C-Terminus showed impacts on embryonic hematopoiesis and regulation of megakaryocyte proliferation while suggesting that this region was trans-activated and had no effect on DNA binding or self-dimerization;[46] An earlier work on the “bloodless” zebrafish vltm651with partial deletion of the C-terminal gata1 showed complete inhibition of DNA binding and transcriptional activation of GATA1[47] and, more recently, the R307C/H mutation found in congenital hemolytic anemia showed reduced DNA binding and aberrant nuclear localization.[25,45] In our mutational investigation, we did not observe abnormalities in the cellular localization or the binding of GATA1 to several target genes. Undoubtedly, we still need additional support from global data regarding the DNA-binding of the mutant GATA1, and the specific function of the human C-terminal GATA1 remains for further clarification.

Changes in the protein-protein interaction network had been the focus of our

exploration into the pathogenesis of the novel identified mutation. With Turbo-ID proximity labeling, we found a subset of proteins in the GATA1 transcriptional complex showed reduced interaction with mutant GATA1, including FOG-1, whose binding site is located at the N-terminal zinc finger with GATA1, as previously demonstrated.[48]

Meanwhile, for the first time, we reported the protein-protein interaction of PRMT6/GATA1 and identified an increased aberrant binding of PRMT6 to mutant GATA1. We showed that PRMT6 could function as a co-repressor of mutant GATA1, achieving transcriptional repression of target genes through GATA1 binding with increased H3R2me2a modification to the regulatory elements of the GATA1 target genes. A similar mechanism was observed in the transcription factor RUNX1, which controls hematopoietic progenitors and megakaryocytes. When RUNX1 binds to PRMT6, RUNX1-mediated gene expression is blocked in megakaryocytes; however, upon dissociation, the active expression of these genes resumes.[25] These results suggest that PRMT6 acts as a negative transcription regulator during hematopoiesis and is responsible for stabilizing gene expression.

As stated above, the GATA1 Leu387fs mutation recruits more PRMT6, accompanied by decreased binding of known GATA1 co-factors (Figure 7G). It may explain why we used the PRMT6 inhibitor but could only partially restore erythroid gene expression. We have not yet learned whether PRMT6 competes with these co-factors for binding, which requires further experimental evidence. In addition, the process of how PRMT6 was recruited to specific transcription factors remains to be investigated.

REFERENCES

1. Doré LC, Crispino JD. Transcription factor networks in erythroid cell and megakaryocyte development. Blood. 2011;118(2):231-239.
2. Pevny L, Simon MC, Robertson E, et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature. 1991;349(6306):257-260.
3. Gutierrez L, Caballero N, Fernandez-Calleja L, Karkoulia E, Strouboulis J. Regulation of GATA1 levels in erythropoiesis. IUBMB Life. 2020;72(1):89-105.
4. Papetti M, Wontakal SN, Stopka T, Skoultchi AI. GATA-1 directly regulates p21 gene expression during erythroid differentiation. Cell Cycle. 2010;9(10):1972-1980.
5. Gregory T, Yu C, Ma A, et al. GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. Blood. 1999;94(1):87-96.
6. Vyas P, Ault K, Jackson CW, Orkin SH, Shivdasani RA. Consequences of GATA-1 deficiency in megakaryocytes and platelets. Blood. 1999;93(9):2867-2875.
7. Meinders M, Hoogenboezem M, Scheenstra MR, et al. Repercussion of Megakaryocyte-Specific Gata1 Loss on Megakaryopoiesis and the Hematopoietic Precursor Compartment. PloS one. 2016;11(5):e0154342-e0154342.
8. Muntean AG, Crispino JD. Differential requirements for the activation domain and FOG-interaction surface of GATA-1 in megakaryocyte gene expression and development. Blood. 2005;106(4):1223-1231.
9. Nichols KE, Crispino JD, Poncz M, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat Genet. 2000;24(3):266-270.
10. Ling T, Crispino JD. GATA1 mutations in red cell disorders. IUBMB Life. 2020;72(1):106-118.
11. Crispino JD, Horwitz MS. GATA factor mutations in hematologic disease. Blood. 2017;129(15):2103-2110.
12. Songdej N, Rao AK. Inherited platelet dysfunction and hematopoietic transcription factor mutations. Platelets. 2017;28(1):20-26.
13. Ludwig LS, Gazda HT, Eng JC, et al. Altered translation of GATA1 in Diamond-Blackfan anemia. Nat Med. 2014;20(7):748-753.
14. Sankaran VG, Ghazvinian R, Do R, et al. Exome sequencing identifies GATA1 mutations resulting in Diamond-Blackfan anemia. J Clin Invest. 2012;122(7):2439-2443.
15. Del Vecchio GC, Giordani L, De Santis A, De Mattia D. Dyserythropoietic anemia and thrombocytopenia due to a novel mutation in GATA-1. Acta Haematol. 2005;114(2):113-116.
16. Tsang AP, Fujiwara Y, Hom DB, Orkin SH. Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. Genes Dev. 1998;12(8):1176-1188.
17. Stumpf M, Yue X, Schmitz S, et al. Specific erythroid-lineage defect in mice conditionally deficient for Mediator subunit Med1. Proc Natl Acad Sci U S A. 2010;107(50):21541-21546.
18. Green AR, Lints T, Visvader J, Harvey R, Begley CG. SCL is coexpressed with

GATA-1 in hemopoietic cells but is also expressed in developing brain. Oncogene. 1992;7(4):653-660.
19. Ahringer J. NuRD and SIN3 histone deacetylase complexes in development. Trends Genet. 2000;16(8):351-356.
20. Hong W, Nakazawa M, Chen YY, et al. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. EMBO J. 2005;24(13):2367-2378.
21. Rodriguez P, Bonte E, Krijgsveld J, et al. GATA-1 forms distinct activating and repressive complexes in erythroid cells. EMBO J. 2005;24(13):2354-2366.
22. Snow JW, Kim J, Currie CR, Xu J, Orkin SH. Sumoylation regulates interaction of FOG1 with C-terminal-binding protein (CTBP). J Biol Chem. 2010;285(36):28064-28075.
23. Yan B, Yang J, Kim MY, et al. HDAC1 is required for GATA-1 transcription activity, global chromatin occupancy and hematopoiesis. Nucleic Acids Res. 2021;49(17):9783-9798.
24. Guccione E, Bassi C, Casadio F, et al. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. Nature. 2007;449(7164):933-937.
25. Herglotz J, Kuvardina ON, Kolodziej S, et al. Histone arginine methylation keeps RUNX1 target genes in an intermediate state. Oncogene. 2013;32(20):2565-2575.
26. Lausen J. Contributions of the histone arginine methyltransferase PRMT6 to the epigenetic function of RUNX1. Crit Rev Eukaryot Gene Expr. 2013;23(3):265-274.
27. Pencovich N, Jaschek R, Tanay A, Groner Y. Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models. Blood. 2011;117(1):e1-14.
28. Herkt SC, Kuvardina ON, Herglotz J, et al. Protein arginine methyltransferase 6 controls erythroid gene expression and differentiation of human CD34(+) progenitor cells. Haematologica. 2018;103(1):18-29.
29. Cui JW, Vecchiarelli-Federico LM, Li YJ, Wang GJ, Ben-David Y. Continuous Fli-1 expression plays an essential role in the proliferation and survival of F-MuLV-induced erythroleukemia and human erythroleukemia. Leukemia. 2009;23(7):1311-1319.
30. Moffat J, Grueneberg DA, Yang X, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell. 2006;124(6):1283-1298.
31. Liu L, Lin B, Yin S, et al. Arginine methylation of BRD4 by PRMT2/4 governs transcription and DNA repair. Sci Adv. 2022;8(49):eadd8928.
32. Orkin SH, Harosi FI, Leder P. Differentiation in erythroleukemic cells and their somatic hybrids. Proc Natl Acad Sci U S A. 1975;72(1):98-102.
33. Doerr A. Proximity labeling with TurboID. Nat Methods. 2018;15(10):764.
34. Freson K, Wijgaerts A, Van Geet C. GATA1 gene variants associated with thrombocytopenia and anemia. Platelets. 2017;28(7):731-734.
35. Martin P, Papayannopoulou T. HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression. Science. 1982;216(4551):1233-1235.

36. Lozzio CB, Lozzio BB. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. Blood. 1975;45(3):321-334.
37. Calligaris R, Bottardi S, Cogoi S, Apezteguia I, Santoro C. Alternative translation initiation site usage results in two functionally distinct forms of the GATA-1 transcription factor. Proc Natl Acad Sci U S A. 1995;92(25):11598-11602.
38. Zhu Y, Wang D, Wang F, et al. A comprehensive analysis of GATA-1-regulated miRNAs reveals miR-23a to be a positive modulator of erythropoiesis. Nucleic Acids Res. 2013;41(7):4129-4143.
39. Vilaboa N, Bermejo R, Martinez P, Bornstein R, Calés C. A novel E2 box-GATA element modulates Cdc6 transcription during human cells polyploidization. Nucleic Acids Res. 2004;32(21):6454-6467.
40. Green AR, DeLuca E, Begley CG. Antisense SCL suppresses self-renewal and enhances spontaneous erythroid differentiation of the human leukaemic cell line K562. EMBO J. 1991;10(13):4153-4158.
41. Liu X, Zhang Y, Chen Y, et al. In Situ Capture of Chromatin Interactions by Biotinylated dCas9. Cell. 2017;170(5):1028-1043.e19.
42. Vakoc CR, Letting DL, Gheldof N, et al. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol Cell. 2005;17(3):453-462.
43. Jurk K, Adenaeuer A, Sollfrank S, et al. Novel GATA1 Variant Causing a Bleeding Phenotype Associated with Combined Platelet α-/δ-Storage Pool Deficiency and Mild Dyserythropoiesis Modified by a SLC4A1 Variant. Cells. 2022;11(19):3071.
44. Singleton BK, Roxby DJ, Stirling JW, et al. A novel GATA1 mutation (Stop414Arg) in a family with the rare X-linked blood group Lu(a-b-) phenotype and mild macrothrombocytic thrombocytopenia. Br J Haematol. 2013;161(1):139-142.
45. Ludwig LS, Lareau CA, Bao EL, et al. Congenital anemia reveals distinct targeting mechanisms for master transcription factor GATA1. Blood. 2022;139(16):2534-2546.
46. Kaneko H, Kobayashi E, Yamamoto M, Shimizu R. N- and C-terminal transactivation domains of GATA1 protein coordinate hematopoietic program. J Biol Chem. 2012;287(25):21439-21449.
47. Lyons SE, Lawson ND, Lei L, et al. A nonsense mutation in zebrafish gata1 causes the bloodless phenotype in vlad tepes. Proc Natl Acad Sci U S A. 2002;99(8):5454-5459.
48. Muntean AG, Crispino JD. Differential requirements for the activation domain and FOG-interaction surface of GATA-1 in megakaryocyte gene expression and development. Blood. 2005;106(4):1223-1231.

FIGURE LEGENDS

Figure 1. A novel germline mutation of GATA1 was identified in a patient with congenital anemia.

A. Sanger sequencing confirmed the existence of the novel GATA1 mutation (c. 1162delGG). The proband and his mother were found to have hemizygous and heterozygous mutations in GATA1, respectively, with the double guanine base deletion identified at the codon position highlighted in the red box.

B. Schematic presentation of the wild-type and mutated (p. Leu387Leufs*62) forms of GATA1 protein was provided, with the red region indicating the area of translation frameshift due to the mutation. AD, activation domain; N-ZF, N-terminal zinc finger; C-ZF, C-terminal zinc finger.

C. As indicated by the black triangle, the mutation occurred at the carboxyl terminus of the GATA1 protein, resulting in a frameshift in translation (amino acids in red) with changes of multiple conserved amino acids (blue).

D. A two-generation pedigree of the family shows the anemic proband carrying the hemizygous mutation indicated by the black arrow, the proband's mother carrying the X-linked heterozygous mutation represented by the dotted circle, and the proband's older brother who died of severe anemia with an unknown genotype indicated by the light gray box with a slash.

E. Immunoblot analysis of GATA1 expression levels in bone marrow (BM) and peripheral blood (PB) of the proband and healthy controls. The bands were quantified by Image J software, and the intensities were normalized to the proband.

Figure 2. The Leu387fs mutation in GATA1 resulted in decreased differentiation, increased cell death, and transcriptional abnormalities during erythropoiesis when compared to wild-type cells.

A. Schematic workflow of constructing re-expression mutant GATA1 model in K562 and HEL cells. The obtained GATA1 depleted cells, wild-type, and mutant GATA1

complement cells were denoted as sgG1, sgG1/WT, and sgG1/MU, respectively. The process of erythropoiesis and megakaryocytic differentiation were simulated using hemin (40μM) and PMA (5 nM), respectively.

B. The cell appearance of the constructed K562 re-expression model for the GATA1 mutation was observed during spontaneous erythroid differentiation.

C. The positive rate of benzidine staining for hemoglobin of K562 cells was measured through a violin diagram, which indicated significant differences between the hemin-treated and untreated groups.

D. Flow cytometric analysis of erythroid differentiation was carried out to quantify the percentage of surface CD71+CD235a+ population in K562 cells treated with hemin.

E. Flow cytometric analysis of erythroid apoptosis in mutant K562 cells. Representative FACS plots displaying the percentages of AnnexinV and CD235a expressed in mutant and wild-type cells after hemin treatment for 48 hours: AnnexinV+ as apoptotic cells (%AC), and measured AnnexinV+CD235a+/CD235a+ as erythroid apoptotic cells (%EAC). The bar graphs (right) show the statistical analysis of apoptotic cells (%AC) and erythroid apoptotic cells (%EAC).

F. RT-qPCR was performed on HEL mutant group cells to assess the expression of target genes known to play a role in erythroid differentiation after 96h hemin induction. Results were normalized with each value of the housekeeping gene GAPDH.

All the results were presented as mean ± SD of 3 independent replicates. The statistical significance of the findings was indicated. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns, no significance.

Figure 3. The GATA1 Leu387fs mutation impedes megakaryocytic differentiation.

A. Flow cytometry plots display the megakaryocytic CD41a^{high} proportion in K562 and HEL cells treated with PMA for 96h. The statistical results are depicted in the bar

graph.

B. The ploidy analysis of mutant cells during PMA-induced megakaryocytic differentiation is represented by flow cytometry analysis of DNA content staining with PI dye. The left peak arose from 2N (G1) cells, the middle peak from 4N (G2) cells, and the right peak from ≥8N (polyploid cells). The bar graph quantified the percentage of each peak in flow cytometry.

C. RT-qPCR analysis shows expression levels of megakaryocyte and platelet-related genes in mutant cells after PMA induction for 4 days, normalized to the housekeeping gene GAPDH.

All the data are shown as mean ± SD of 3 replicates. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns, no significance.

Figure 4. Turbo-ID proximity labeling was utilized to identify the altered interaction co-regulators of the GATA1 Leu387fs mutation.

A. Schematic of Turbo-ID proximity labeling workflow. The wild-type and Leu387fs GATA1-Turbo fusion protein were overexpressed in GATA1-depleted K562 cells separately. After transfection, the cells were pulsed with biotin for 30 minutes, and the biotinylated proteins were subsequently collected and isolated by streptavidin beads. These proteins were then subjected to mass spectrometry and verified for interest regulators by western blot analysis.

B. The mutant GATA1 altered the proximity to a set of known co-factors/interacting proteins and RNA polymerase. Bar graphs showing the relative amounts of biotinylated proximal proteins. *, $p < 0.05$; ***, $p < 0.001$.

C. The relative amounts of subunits of the Switch/Sucrose Non-Fermentable (SWI/SNF) complex and the Nucleosome Remodeling Complex (NuRD) complex proximal to the wild-type and mutant GATA1 were also compared, as illustrated in the upper and lower heat maps, respectively. Z-Score shows fold differences in relative quantitative protein values and color intensity (high: red; low: blue). The asterisks indicate proteins with a significant difference.

D. The significantly decreased/increased proximal proteins were uploaded with GATA1 to STRING for network analysis. A medium level of confidence (0.400) was established for the database searches. Connected lines indicate protein-protein interactions and unconnected nodes in the network were hidden. The significantly reduced proximal proteins were involved in the networks of transcriptional regulatory activity, erythroid/megakaryocyte differentiation, mRNA splicing, and electron transport chain, while the significantly increased proximal proteins were involved in networks such as methylation, mRNA stability regulation, and hemoglobins.

Figure 5. The Leu387fs mutation in GATA1 resulted in enhanced interaction with arginine methyltransferase PRMT6, compared to wild-type GATA1.

A. The TuboID-mass spectrometry result showed a significant increase in the relative amount of biotinylated PRMT6 protein with the mutant GATA1. ***$p < 0.001$.

B. Co-immunoprecipitation analysis of GATA1 binding to PRMT6 during erythropoiesis (E) or megakaryopoiesis (M) in K562 cells. The intensities of bands were corrected by the input H3 and normalized to untreated K562 cells.

C. The proximity between GATA1 Leu387fs mutation and PRMT6 was confirmed in K562 GATA1-depleted cells, using pull-down immunoblot analysis based on the Turbo-ID method. The intensities of bands were calibrated by the input H3 and normalized to the untransfected K562-sgG1 cells.

D. Co-immunoprecipitation analysis of mutant GATA1 with PRMT6 in HEK-293T cells. The experiments involved co-expression of Flag-tagged PRMT6 alongside HA-tagged wild-type or mutant GATA1, with transfected alone as a control. Pull-down assays were performed with Flag and HA tags, respectively, to ascertain PRMT6's enhanced interaction with mutant GATA1. The results of the HA-tagged GATA1 precipitation experiment also revealed a decrease in mutant GATA1's binding ability to FOG1. The intensities of bands were calibrated by the input H3 and normalized to the sample transfected with both mutant GATA1 and PRMT6.

E. The GST-pulldown assay was utilized to compare the in vitro binding ability of

GST-PRMT6 with wild-type or mutant GATA1. Labeled values are quantified intensities of bands normalized to the sample GST-wild-type GATA1.

F. Different Flag-tagged truncated PRMT6 was co-expressed with HA-tagged mutant GATA1 in IP assay to explore the binding site of mutant GATA1 at PRMT6. Lable values are band intensities calibrated by the internal reference protein and followed normalization.

F. Different HA-tagged truncated GATA1 was co-expressed with Flag-tagged PRMT6 in IP assay to explore the binding domain of PRMT6 at mutant GATA1. Lable values are band intensities calibrated by the internal reference protein and followed normalization.

All the band intensities were quantified using Image J software.

Figure 6. The occupancy changes of PRMT6 and H3R2me2a modification on the promotor/enhancer of GATA1 erythroid target genes by GATA1 Leu387fs mutation.

ChIP-qPCR assays were performed on the re-expression wild-type and mutant K562 cells during spontaneous differentiation to determine the binding of HA-tagged GATA1, PRMT6, and H3R2me2a modification on the target genes HBG1, HBB, CD235a (GYPA) promoter, and LCR-HS2 distal enhancer.

A. HA-tagged GATA1 binding results without erythroid induction.

B. PRMT6 binding results without erythroid induction.

C. H3R2me2a binding results with histone H3 serving as the control in spontaneous differentiation.

The data are presented as mean ± SD of 3 independent replicates by percent enrichment relative to input, and statistical significance is denoted by asterisks. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$; ns, differences not statistically significant.

Figure 7. Abnormal erythroid differentiation caused by GATA1 Leu387fs mutation can be partially rescued with MS023.

A. In K562 cells, stable knockdown of PRMT6 led to an increase in γ-globin (HBG1) expression and a decrease in the H3R2me2a modification, as observed in western blot analysis.

B. Western blot analysis was performed to determine the expression levels of histone modifications H3K4me3 and H3R2me2a in mutant GATA1 re-expression cells. The resulting values were measured using Image J software and normalized with the NC group.

C. Western blot analysis was conducted to detect the expression levels of γ-globin (HBG1), PRMT6, and H3R2me2a following treatment with the PRMT6 inhibitor MS023 at indicated concentrations (0.01μM, 0.1μM, and 1μM, respectively) in both the re-expression wild-type (WT) and mutant (MU) K562 cells for 4 days. The empty and DMSO treatment groups were used as controls. The indicated values were quantified using Image J software and normalized with the WT cells without any treatment.

D. Erythroid differentiation was analyzed via flow cytometry following treatment with the indicated concentrations of MS023 in the wild-type- and mutant-complementary cells for 4 days. The percentages of surface marker CD71+CD235a+ cells were quantified, and the fold changes of CD71+CD235a+ proportion by MS023 treatment compared to DMSO control were also calculated. The presented data are the mean ± SD of 3 independent replicates.

E. The positive rate of benzidine staining between re-expressing wild-type and mutant K562 cells treated with 1μM MS023 for 4 days was compared. DMSO was used as the negative control. The fold changes of positive rate by MS023 treatment compared with DMSO control were also calculated. The data presented are the mean ± SD of 3 independent replicates.

F. RT-qPCR analysis was performed to measure the relative expression levels of CD235a (GYPA), β-globin (HBB), and γ-globin (HBG1) at the mRNA level following treatment of cells with 1μM MS023 for 4 days. DMSO was used as the negative control. The data presented are the mean ± SD of 3 independent replicates.

G. Schematic representation of the novel GATA1 mutation's mechanism. In its wild-type state, GATA1 engages in transcriptional regulatory complex formations with co-factors, including FOG1, to ensure proper transcriptional activities for the expression of downstream target genes. However, in its mutated state, GATA1 decreases its binding capabilities to certain original co-factors. Still, increased binding to the arginine methyltransferase PRMT6 facilitates a higher level of histone H3R2me2a modification located on the promoter of the GATA1 target genes, which effectively inhibits the regular transcription process.

Statistical significance is denoted by asterisks. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$; ns, no significance.

Figure 1

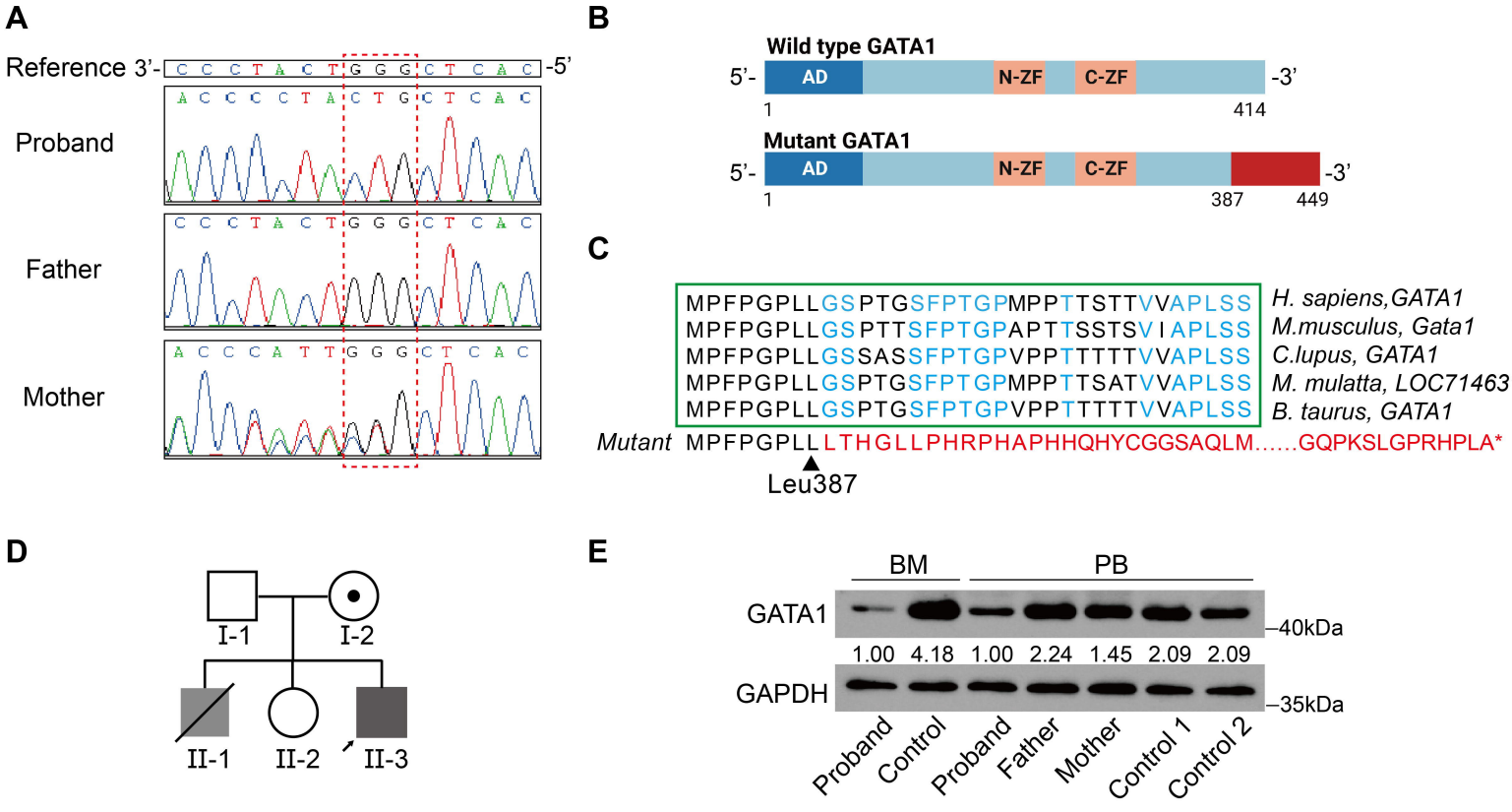

A
Reference 3'-
C C C T A C T G G C T C A C
-5'
Proband
A C C C C T A C T G C T C A C
Father
C C C T A C T G G C T C A C
Mother
A C C C A T T G G C T C A C
B
Wild type GATA1
5'-
AD
N-ZF
C-ZF
-3'
1
414
Mutant GATA1
5'-
AD
N-ZF
C-ZF
-3'
1
387
449
C
MPFPGPLLGSPTGSFPTGPMPPTTSTTVVAPLSS
MPFPGPLLGSPTTSFPTGPAPTTSSTSVIAPLSS
MPFPGPLLGSSASSFPTGPVPPTTTTTVVAPLSS
MPFPGPLLGSPTGSFPTGPMPPTTSATVVAPLSS
MPFPGPLLGSPTGSFPTGPVPPTTTTTVVAPLSS
H. sapiens,GATA1
M.musculus, Gata1
C.lupus, GATA1
M. mulatta, LOC714631
B. taurus, GATA1
Mutant
MPFPGPLLLTHGLLPHRPHAPHHQHYCGGSAQLM.....GQPKSLGPRHPLA*
Leu387
D
I-1
I-2
II-1
II-2
II-3
E
BM
PB
GATA1
40kDa
1.00 4.18 1.00 2.24 1.45 2.09 2.09
GAPDH
35kDa
Proband
Control
Proband
Father
Mother
Control 1
Control 2

Figure 2

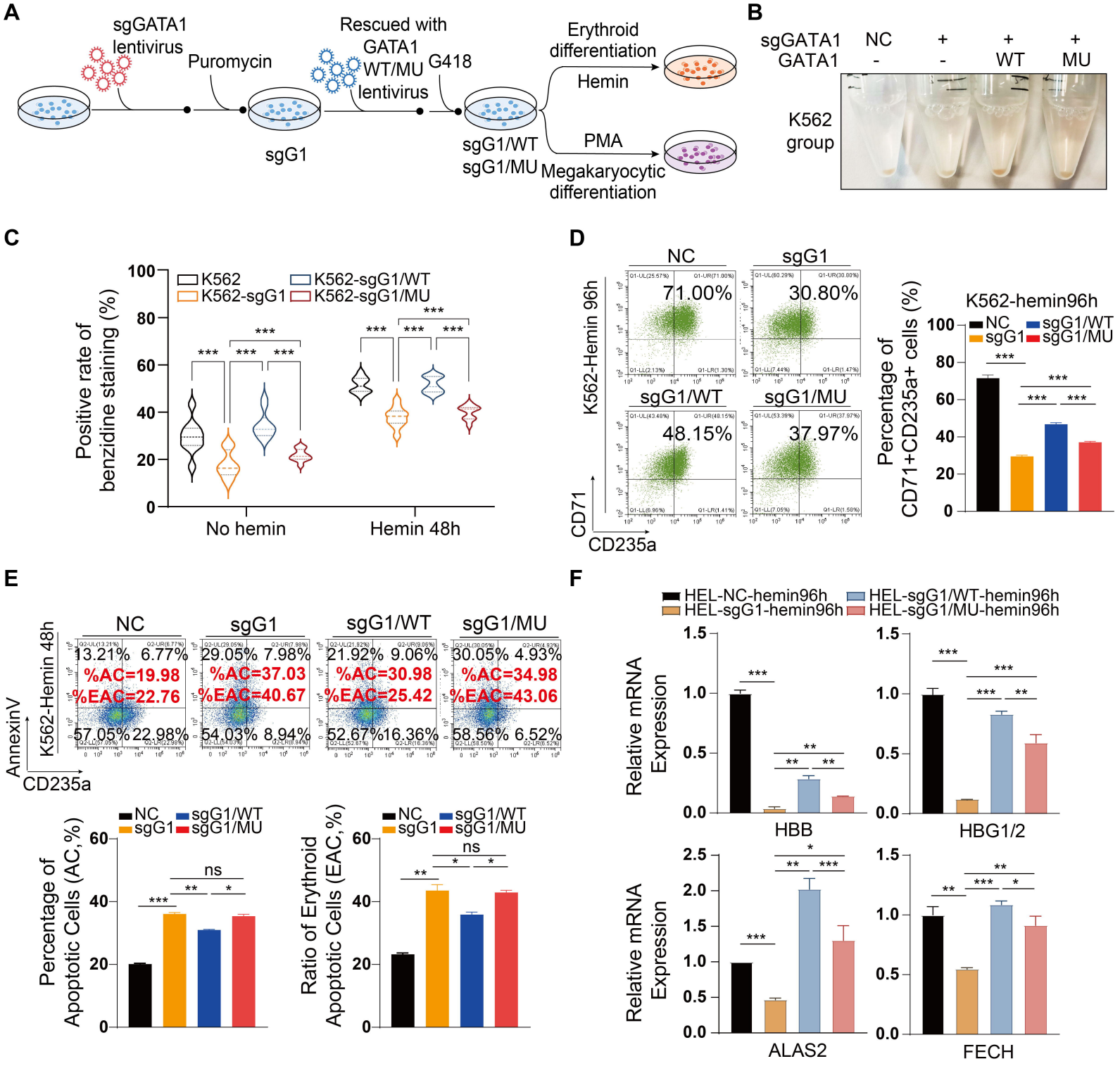

A
sgGATA1 lentivirus
Puromycin
sgG1
Rescued with GATA1 WT/MU lentivirus
G418
sgG1/WT
sgG1/MU
Erythroid differentiation
Hemin
PMA
Megakaryocytic differentiation
B
sgGATA1
GATA1
NC
-
+
-
+
WT
+
MU
K562 group
C
K562
K562-sgG1
K562-sgG1/WT
K562-sgG1/MU
Positive rate of benzidine staining (%)
No hemin
Hemin 48h
D
NC
sgG1
sgG1/WT
sgG1/MU
71.00%
30.80%
48.15%
37.97%
K562-Hemin 96h
CD71
CD235a
K562-hemin96h
Percentage of CD71+CD235a+ cells (%)
E
K562-Hemin 48h
AnnexinV
CD235a
%AC=19.98
%EAC=22.76
%AC=37.03
%EAC=40.67
%AC=30.98
%EAC=25.42
%AC=34.98
%EAC=43.06
Percentage of Apoptotic Cells (AC,%)
Ratio of Erythroid Apoptotic Cells (EAC,%)
F
HEL-NC-hemin96h
HEL-sgG1-hemin96h
HEL-sgG1/WT-hemin96h
HEL-sgG1/MU-hemin96h
Relative mRNA Expression
HBB
HBG1/2
ALAS2
FECH

Figure 3

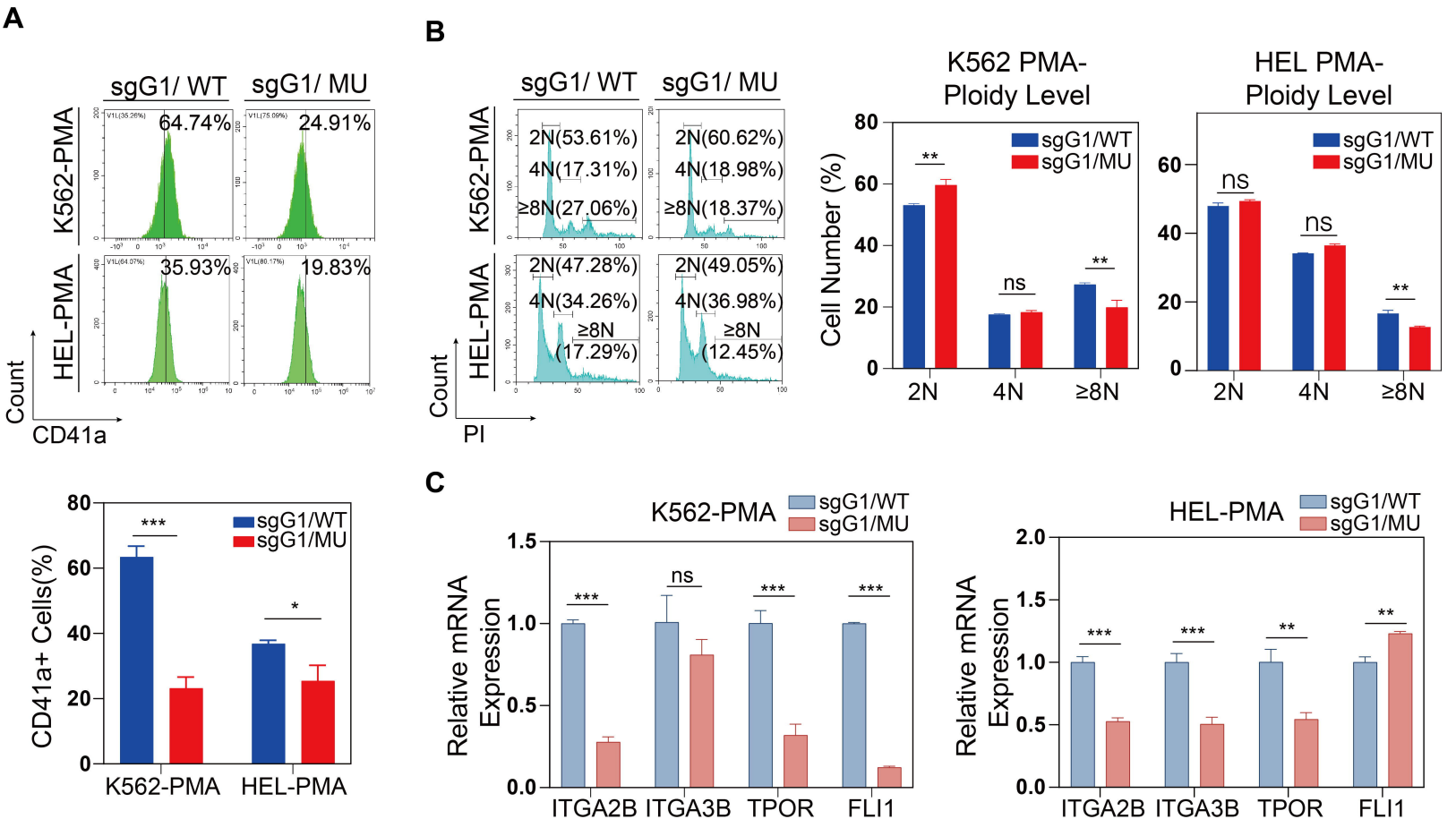

A
sgG1/ WT
sgG1/ MU
K562-PMA
HEL-PMA
64.74%
24.91%
35.93%
19.83%
Count
CD41a
B
2N(53.61%)
4N(17.31%)
≥8N(27.06%)
2N(60.62%)
4N(18.98%)
≥8N(18.37%)
2N(47.28%)
4N(34.26%)
≥8N (17.29%)
2N(49.05%)
4N(36.98%)
≥8N (12.45%)
PI
K562 PMA- Ploidy Level
HEL PMA- Ploidy Level
Cell Number (%)
sgG1/WT
sgG1/MU
2N
4N
≥8N
ns
CD41a+ Cells(%)
C
Relative mRNA Expression
ITGA2B
ITGA3B
TPOR
FLI1

Figure 4

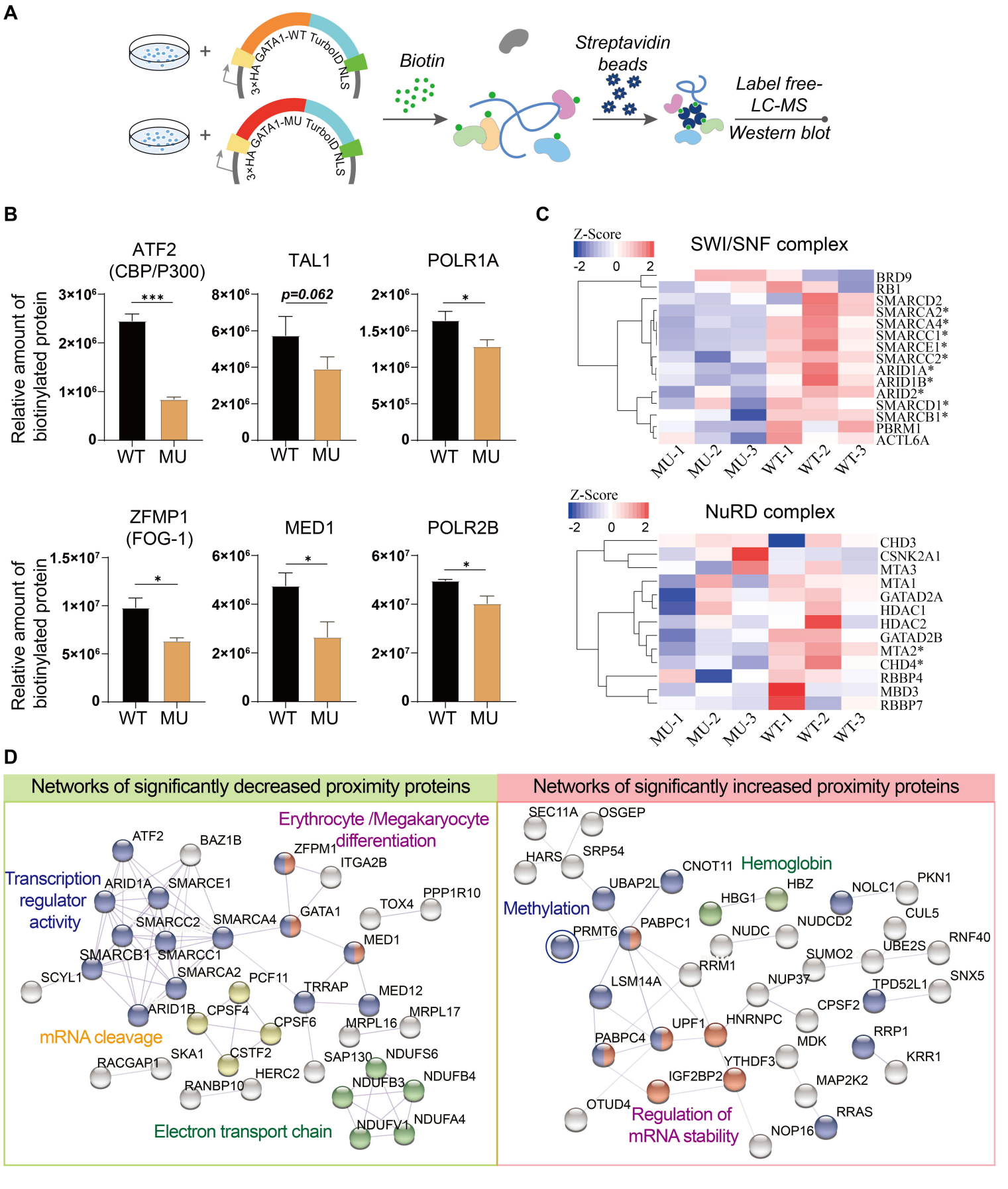

A
3×HA GATA1-WT TurboID NLS
3×HA GATA1-MU TurboID NLS
Biotin
Streptavidin beads
Label free-LC-MS
Western blot
B
Relative amount of biotinylated protein
ATF2 (CBP/P300)
TAL1
POLR1A
ZFMP1 (FOG-1)
MED1
POLR2B
***
p=0.062
*
WT
MU
C
Z-Score
SWI/SNF complex
BRD9
RB1
SMARCD2
SMARCA2*
SMARCA4*
SMARCC1*
SMARCE1*
SMARCC2*
ARID1A*
ARID1B*
ARID2*
SMARCD1*
SMARCB1*
PBRM1
ACTL6A
NuRD complex
CHD3
CSNK2A1
MTA3
MTA1
GATAD2A
HDAC1
HDAC2
GATAD2B
MTA2*
CHD4*
RBBP4
MBD3
RBBP7
MU-1
MU-2
MU-3
WT-1
WT-2
WT-3
D
Networks of significantly decreased proximity proteins
Networks of significantly increased proximity proteins
Erythrocyte /Megakaryocyte differentiation
Transcription regulator activity
mRNA cleavage
Electron transport chain
Hemoglobin
Methylation
Regulation of mRNA stability
ATF2
BAZ1B
ZFPM1
ITGA2B
ARID1A
SMARCE1
TOX4
PPP1R10
GATA1
SMARCC2
SMARCA4
MED1
SMARCB1
SMARCC1
SCYL1
SMARCA2
PCF11
TRRAP
MED12
ARID1B
CPSF4
CPSF6
MRPL16
MRPL17
SKA1
CSTF2
RACGAP1
SAP130
NDUFS6
HERC2
NDUFB3
NDUFB4
RANBP10
NDUFV1
NDUFA4
SEC11A
OSGEP
HARS
SRP54
CNOT11
UBAP2L
HBG1
HBZ
NOLC1
PKN1
CUL5
PRMT6
PABPC1
NUDC
NUDCD2
UBE2S
RNF40
SUMO2
RRM1
LSM14A
NUP37
TPD52L1
SNX5
CPSF2
UPF1
HNRNPC
PABPC4
MDK
RRP1
KRR1
YTHDF3
IGF2BP2
MAP2K2
OTUD4
RRAS
NOP16

Figure 5

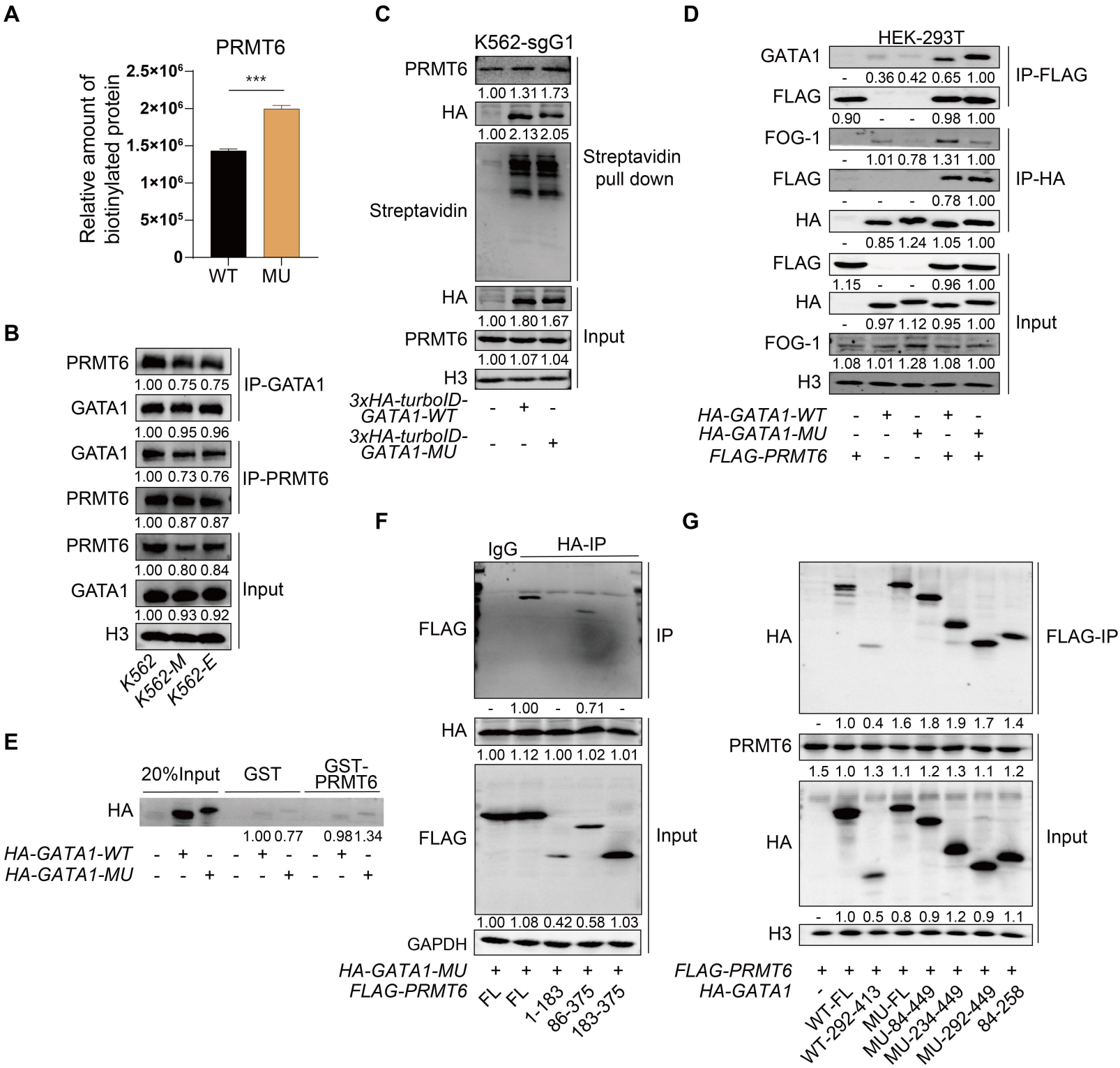

A
PRMT6
Relative amount of biotinylated protein
2.5×10⁶
2×10⁶
1.5×10⁶
1×10⁶
5×10⁵
0
***
WT
MU
B
PRMT6
1.00 0.75 0.75
IP-GATA1
GATA1
1.00 0.95 0.96
GATA1
1.00 0.73 0.76
IP-PRMT6
PRMT6
1.00 0.87 0.87
PRMT6
1.00 0.80 0.84
GATA1
1.00 0.93 0.92
Input
H3
K562
K562-M
K562-E
C
K562-sgG1
PRMT6
1.00 1.31 1.73
HA
1.00 2.13 2.05
Streptavidin
Streptavidin pull down
HA
1.00 1.80 1.67
PRMT6
1.00 1.07 1.04
Input
H3
3xHA-turboID-GATA1-WT - + -
3xHA-turboID-GATA1-MU - - +
D
HEK-293T
GATA1
- 0.36 0.42 0.65 1.00
IP-FLAG
FLAG
0.90 - - 0.98 1.00
FOG-1
- 1.01 0.78 1.31 1.00
FLAG
IP-HA
- - - 0.78 1.00
HA
- 0.85 1.24 1.05 1.00
FLAG
1.15 - - 0.96 1.00
HA
- 0.97 1.12 0.95 1.00
Input
FOG-1
1.08 1.01 1.28 1.08 1.00
H3
HA-GATA1-WT - + - + -
HA-GATA1-MU - - + - +
FLAG-PRMT6 + - - + +
E
20%Input
GST
GST-PRMT6
HA
1.00 0.77 0.98 1.34
HA-GATA1-WT - + - - + - - + -
HA-GATA1-MU - - + - - + - - +
F
IgG
HA-IP
FLAG
- 1.00 - 0.71 -
IP
HA
1.00 1.12 1.00 1.02 1.01
FLAG
1.00 1.08 0.42 0.58 1.03
Input
GAPDH
HA-GATA1-MU + + + + +
FLAG-PRMT6
FL
FL
1-183
86-375
183-375
G
HA
FLAG-IP
- 1.0 0.4 1.6 1.8 1.9 1.7 1.4
PRMT6
1.5 1.0 1.3 1.1 1.2 1.3 1.1 1.2
HA
Input
- 1.0 0.5 0.8 0.9 1.2 0.9 1.1
H3
FLAG-PRMT6 + + + + + + + +
HA-GATA1 -
WT-FL
WT-292-413
MU-FL
MU-84-449
MU-234-449
MU-292-449
84-258

Figure 6

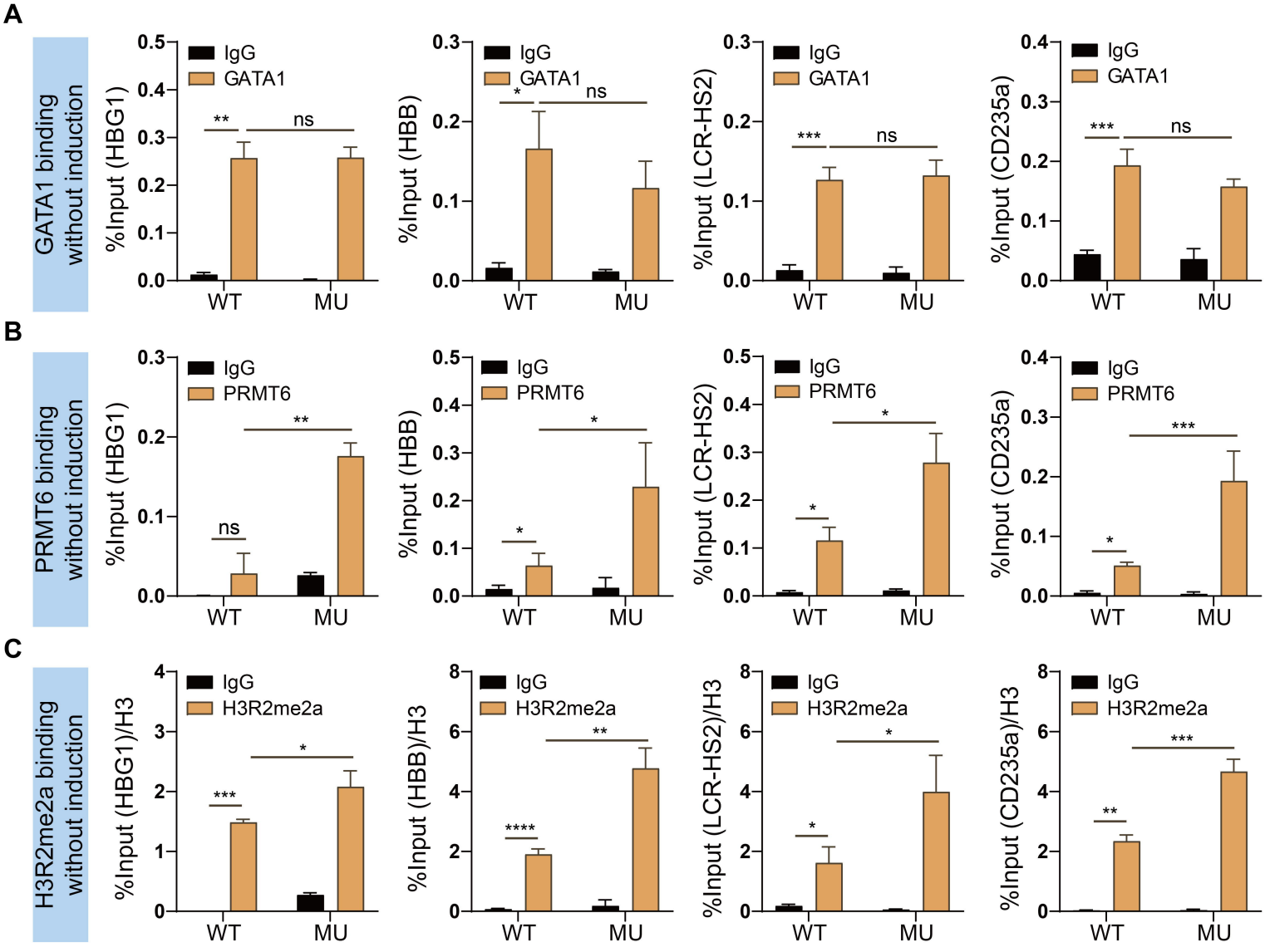
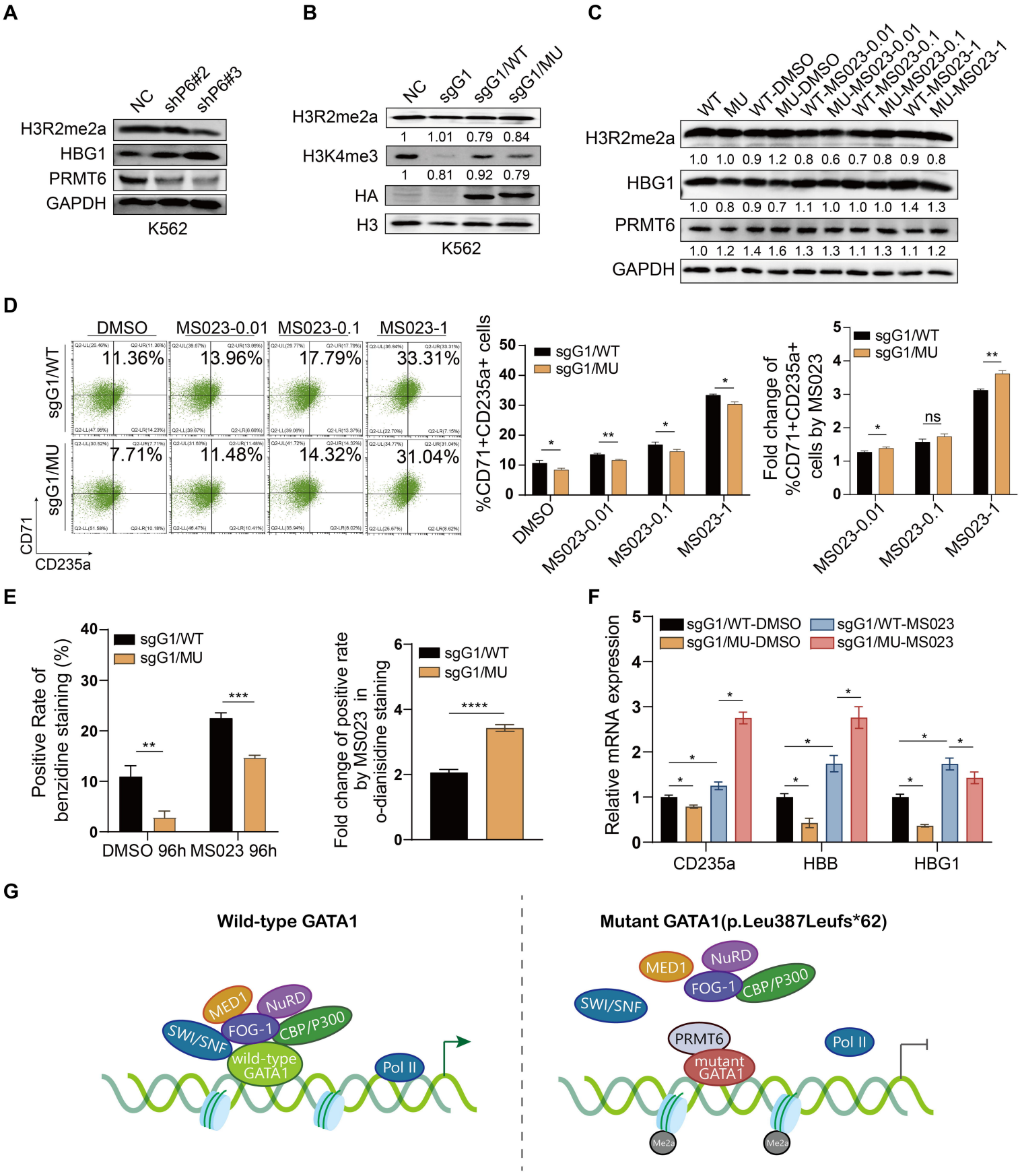

A
GATA1 binding without induction
%Input (HBG1)
%Input (HBB)
%Input (LCR-HS2)
%Input (CD235a)
IgG
GATA1
WT
MU
B
PRMT6 binding without induction
PRMT6
C
H3R2me2a binding without induction
%Input (HBG1)/H3
%Input (HBB)/H3
%Input (LCR-HS2)/H3
%Input (CD235a)/H3
H3R2me2a

Figure 7


A
NC
shP6#2
shP6#3
H3R2me2a
HBG1
PRMT6
GAPDH
K562
B
NC
sgG1
sgG1/WT
sgG1/MU
H3R2me2a
1 1.01 0.79 0.84
H3K4me3
1 0.81 0.92 0.79
HA
H3
K562
C
WT
MU
WT-DMSO
MU-DMSO
WT-MS023-0.01
MU-MS023-0.01
WT-MS023-0.1
MU-MS023-0.1
WT-MS023-1
MU-MS023-1
H3R2me2a
1.0 1.0 0.9 1.2 0.8 0.6 0.7 0.8 0.9 0.8
HBG1
1.0 0.8 0.9 0.7 1.1 1.0 1.0 1.0 1.4 1.3
PRMT6
1.0 1.2 1.4 1.6 1.3 1.3 1.1 1.3 1.1 1.2
GAPDH
D
DMSO
MS023-0.01
MS023-0.1
MS023-1
sgG1/WT
sgG1/MU
11.36%
13.96%
17.79%
33.31%
7.71%
11.48%
14.32%
31.04%
CD71
CD235a
%CD71+CD235a+ cells
sgG1/WT
sgG1/MU
DMSO
MS023-0.01
MS023-0.1
MS023-1
Fold change of %CD71+CD235a+ cells by MS023
ns
E
Positive Rate of benzidine staining (%)
sgG1/WT
sgG1/MU
DMSO 96h
MS023 96h
Fold change of positive rate by MS023 in o-dianisidine staining
F
Relative mRNA expression
sgG1/WT-DMSO
sgG1/MU-DMSO
sgG1/WT-MS023
sgG1/MU-MS023
CD235a
HBB
HBG1
G
Wild-type GATA1
MED1
NuRD
FOG-1
CBP/P300
SWI/SNF
wild-type GATA1
Pol II
Mutant GATA1(p.Leu387Leufs*62)
MED1
NuRD
FOG-1
CBP/P300
SWI/SNF
PRMT6
mutant GATA1
Pol II
Me2a
Me2a

SUPPLEMENTAL DATA

Detailed materials and methods

Whole exome sequencing (WES) and Sanger sequencing

Genomic DNA extracted from whole blood was ultrasonically fragmented, then selected for 400-600 bp with AMPure XP beads (Agencourt), followed by end repair, sequencing adaptor ligation, and PCR amplification according to standard library construction protocols to obtain a preliminary sequencing library. IDT probes (xGen Exome Research Panel v1) were used to capture exon-targeting regions. The exome libraries of the patient and her parents were then sequenced with the NovaSeq 6000 (Illumina) platform. The sequencing raw reads were filtered by the fastp software for quality control, the Burrows-Wheeler-Alignment (BWA) tool to align with the reference genome GRCh37/hg19, and the Genome Analysis Toolkit (GATK) tool to annotate SNPs and Indels, finally screened for confident variants. The variant effect predictor (VEP) tool was applied in further variant annotation to find the variant associated with the anemia phenotype.

Constructs

Human GATA1 gRNAs were cloned into the lentiCRISPR-v2 puro vector (Addgene plasmid, #52961) as previously described[1]. Full-length human wild-type GATA1 and mutant GATA1 were amplified with the cDNA derived from K562 cells and peripheral blood of the proband, respectively, together with an HA tag, then cloned into either the XbaI and NotI sites of pCDH-MSCV lentiviral vector (applied for invitro model construction) or the NotI and XbaI sites of pCDNA3.1 transient transfection vector (used for protein interaction study). 3xHA-G1WT/G1MU-Turbo-NLS was generated by subcloning the cDNA of the G1WT/G1MU gene from the pCDH vector into the NheI site of pCDNA3-3xHA-TurboID-NLS (Addgene plasmid, # 107171). PRMT6-Flag was generated by the cDNA for PRMT6 subcloned in the pCDNA 3.1-Flag vector. shRNAs targeting PRMT6 were cloned into pLKO.1-hygro vector. All

plasmids had been verified by sequencing before use.

Lentivirus production and infection

Low-passage HEK-293T cells were seeded in 10 ml of DMEM complete media and allowed to reach a 70% confluency in a 10cm culture dish. Lentiviral plasmids were co-transfected with pMD2.G and psPAX2 following polyethyleneimine (PEI) transfection protocol. After 48-hour transfection, the viral supernatant was collected and filtered through a 0.45-μm filter (Merck Millipore). The viral supernatants were then added directly to cells in 6-well plates with polybrene (Sigma-Aldrich) supplemented to a concentration of 2 μg/ml, and the suspension cells were spin-infected at 1500 rpm for 90 min at room temperature. Cells were replaced with a complete medium after 24 hours and selected by corresponding antibiotics the next day.

CRISPR/Cas9-mediated deletion of GATA1 in K562 and HEL cells

GATA1 knock-out in K562 and HEL cells (K562-sgG1 and HEL-sgG1) was performed by lentiviral transduction as described above with lentiCRISPR-v2-sgGATA1 plasmids. 72 hours after infection, cells underwent selection using 2 μg/ml puromycin (Beyotime) for 7 days and were subsequently amplified, then evaluated protein levels of GATA1 by western blot analysis. For details of the sgRNAs sequence see Supplemental Table S4.

Generation of re-expression stable cells

For reconstituted with wild-type GATA1 and GATA1-Leu387fs in GATA1-deficient K562 and HEL cells generated as above mentioned, packaging plasmids were co-transfected with either pCDH-MSCV-HA-G1WT or pCDH-MSCV-HA-G1MU into HEK-293T cells. 48 h post-transfection, GATA1-depleted cells were infected with harvested viral particles for 24 h in the presence of polybrene. Positive stable cell lines were subsequently selected by G418 (70 μg/ml) for 7 days. The GATA1 protein levels of the stable re-expression cells were then evaluated by WB analysis.

Benzidine staining

Benzidine cytochemical staining was performed to assess the level of hemoglobin[2]. Briefly, cells were washed once with PBS and resuspended in a ratio of 1:1 with benzidine solution (1.46 ml acetic acid, 48.54 ml ddH_2O, 100 mg benzidine dihydrochloride (Sigma-Aldrich)) and a final concentration of 0.1% H_2O_2. Stained cells were incubated in the dark at room temperature for 2 min. Positive staining cells (dark blue or brown) were photographed and quantified via an upright microscope (Leica, DM4B) at magnification x100, with positively stained cells manifesting blue or dark brown. Over 100 total cells were counted for each image.

FACS analysis

The cultured cells were washed with PBS and were blocked with 0.5% FBS for 15 minutes on ice before staining. Erythroid differentiation examinations utilized APC-CD235a (eBioscience) and either PE-Cyanine7-CD71 (eBioscience) or FITC-CD71 (Biolegend). Megakaryocytic induction analysis was investigated using APC-CD41a (eBioscience). To assess erythroid apoptosis, Annexin V analysis was performed wherein cells were washed with 1x Annexin V binding buffer and then stained with FITC-Annexin V (BD, 556420) and APC-CD235a (eBioscience) following the manufacturer's protocols. For Megakaryocytic ploidy analysis (DNA content analysis), cells were permeabilized using cold 70% ethanol and stained with propidium iodide (50 μg/mL) along with RNase A. Flow cytometry was conducted utilizing the CytoFLEX (Beckman Coulter) platform, and subsequent data analysis was performed using CytExpert software (Beckman Coulter) or Flowjo software.

Real-time quantitative PCR

Real-time quantitative PCR was performed by isolating total RNA from cells, performing reverse transcription, and the quantitative PCR assay according to the manufacturer's instructions (AG, China). All primers are listed in Supplementary Table

S3. Normalization was done to the housekeeping gene GAPDH for gene expression levels. The quantitative PCR was analyzed on Bio-Rad CXF96, and the supporting software (Bio-Rad CXF Manager) was used for analysis.

Western blot analysis

The cell pellets were lysed with RIPA buffer (Epizyme, Shanghai) supplemented with a protease inhibitor cocktail (Beyotime, Shanghai) for 30 minutes on ice. Proteins were subsequently collected through 12000 rpm centrifugation for 10 minutes at 4 ℃. Quantitation was carried out by making use of the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) and followed by 10% or 12.5% SDS-PAGE electrophoresis separation. The antibodies and the dilution multiple utilized in the study were as follows: anti-GATA1 (Abcam, ab181544, 1:10000), anti-FOG1 (Abcam, 1:1000), anti-GAPDH (Proteintech, 60004-1, 1:10000), anti-HA (CST, C29F4,1:1000), anti-histone-H3 (Proteintech, 17168-1-AP, 1:1000), anti-HBG1 (Proteintech,25728-1-AP,1:500), anti-PRMT6 (Proteintech, 15395-1-AP, 1:1000), anti-H3R2me2a (ABclonal, A3155, 1:1000), anti-FLAG (CST, 9A3, 1:1000), HRP Streptavidin (BioLegend, 405210, 1:1000). For secondary antibodies, Dylight 800-goat anti-rabbit IgG (Abbkine, A23920, 1:10000) or Dylight 800-goat anti-mouse IgG (Abbkine, A23910, 1:10000) were employed based on the species.

Turbo-ID proximity labeling

To verify the cellular location of biotinylated proteins by immunofluorescent staining, HEK-293T cells were seeded on glass coverslips in a 24-well plate for transfection the following day. 1 μg Turbo-ID fusion constructs were introduced through transfection via PEI for each well. The medium was replaced with 50 μM biotin (MCE)-supplemented complete DMEM media for 30min further culture post-36h-transfection. After this, the labeling reaction was ceased by removing the media and rinsing with 500μL cold PBS five times. Immunostaining was performed post-fixation and cell permeabilization using paraformaldehyde and cold methanol using rabbit mAb

anti-HA-tag primary antibodies (CST, #3724). The following secondary antibodies were used: Alexa Fluor 488 Streptavidin (Yeason, #35103ES60), goat anti-rabbit IgG (H+L) cross Alexa Fluor Plus 594 (Thermo Fisher, #A32740). Lastly, 1 μg/ml DAPI was utilized to stain the nuclei.

To pull down the biotinylated proteins for western blot analysis, 5×10^6 post-transfected cells were incubated with 50 μM biotin for 30 minutes to label proximal proteins, subsequently washed with cold PBS, and lysis with 250 μL RIPA lysis buffer. After sonication and centrifugation at 12,000 rpm for 10 minutes, the supernatant was collected, and the protein concentration was then measured using the BCA assay kit (Thermo Scientific, #23225). The input sample was acquired from the lysate. Subsequently, 25 μL of streptavidin magnetic beads (MCE, #HY-K0208) were incubated with 500 μg protein from lysates overnight at 4℃ to enrich biotinylated proteins. Beads were then washed in order with RIPA lysis buffer (1 ml, twice, 2 minutes each), 0.1 M Na2CO3 solution (1ml, once for 10s), 2M urea in PH 8.0 Tris-HCl (1ml, once for 10s), and RIPA lysis buffer (1 ml, twice, 2 minutes each) at last. The beads were then resuspended in 3 x protein loading buffer supplemented with 2 mM biotin and 20 mM DTT, boiled at 100 ℃ for 10 minutes. Finally, the eluate was collected by a magnetic rack and analyzed by western blotting to confirm biotinylated proteins.

To perform the Turbo-ID proximity labeling-based mass spectrometry assays, the plasmid of Turbo-ID fused wild-type or mutant GATA1 protein was transfected in stable GATA1-depleted K562 cells using liposomal transfection reagents (Yeason, #40802ES02) according to the product instructions, respectively. Following individual biotin labeling according to the abovementioned conditions, 1×10^7 cells were collected for protein collection and quantification. 5 mg total protein was incubated overnight at 4°C with 100 μL streptavidin magnetic beads. After the enrichment of biotinylated proteins, the beads were washed with RIPA lysis buffer and ready for mass spectrometry.

Label-free mass spectrometry for protein identification and quantification.

The EASY-nLCTM 1200 (Thermo Fisher, Germany) coupled with Orbitrap Exploris 480 (Thermo Fisher, Germany) platform was utilized for label-free mass spectrometry detection at Novogene Co., Ltd. (Beijing, China). Employing Proteome Discoverer, the mass tolerance for precursor ion was ten ppm, and the mass tolerance for production was 0.02 Da. Fixed modifications for carbamidomethyl, dynamic modification for methionine (M) oxidation, and loss of methionine at the N-terminal were specified. Up to 2 missed cleavage sites were permitted. To ensure high-quality analysis results, PD software was used to filter retrieval results to identify Peptide Spectrum Matches (PSMs) with more than 99% credibility. Identified proteins were required to contain at least one unique peptide, and both the identified PSMs and proteins were retained and subjected to FDR no more than 1.0%. T-tests were employed for statistical analysis of protein quantitation results, with differentially expressed proteins (DEP) defined as those whose quantitation was significantly different between experimental and control groups ($p < 0.05$ and $FC > 1.25$ or $FC < 0.75$ [foldchange, FC]).

Co-immunoprecipitation and immunoprecipitation analysis

HEK-293T cells were transfected with appropriate plasmids and lysed by co-immunoprecipitation lysis buffer NETN (20 mM Tris-HCl (pH 8.0), 0.5% NP-40, 100 mM NaCl, 0.5 mM EDTA) complemented with the protease inhibitor cocktail (Beyotime, Shanghai). The cell lysates were then incubated with a primary antibody and protein A/G agarose (ABmart, A10001M). The agarose beads were washed three times with the NETN buffer and subjected to SDS-PAGE.

Chromatin immunoprecipitation

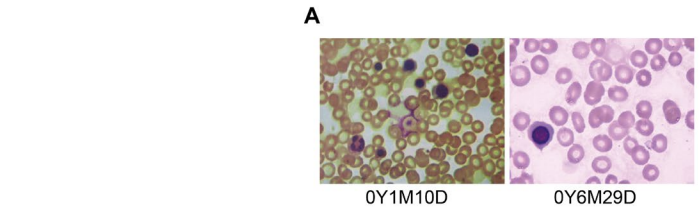
ChIP was performed with the BeyoChIP™ ChIP Assay kit (Beyotime, P2080S) and the following antibodies: anti-H3R2me2a (ABclonal, A3155), anti-PRMT6 (Proteintech, 15395-1-AP), anti-HA (CST, C29F4), anti-H3 (Proteintech, 17168-1-AP) and normal

rabbit IgG (CST, #2729). All primers are listed in Supplemental Table S3.

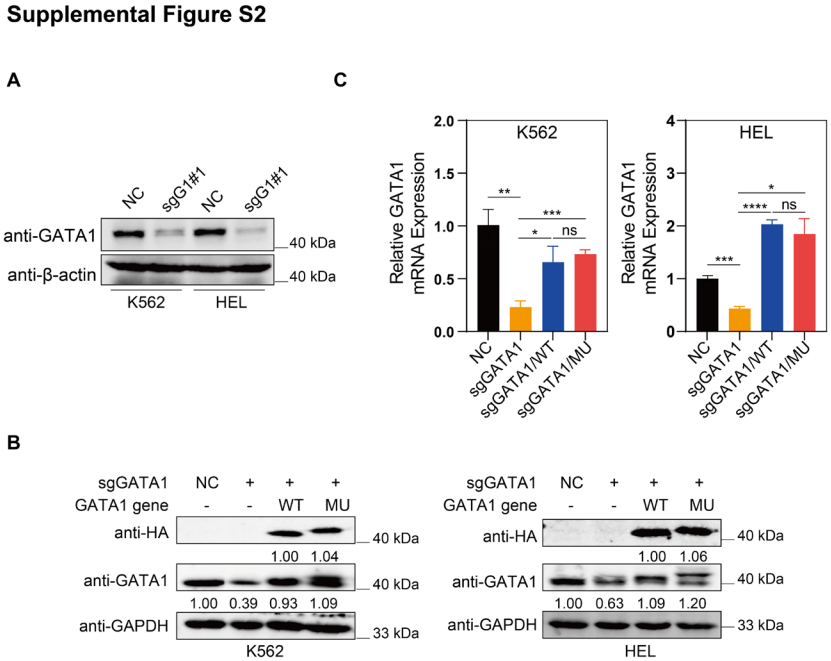
Supplemental References

1. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nature methods*. 2014/08/01 2014;11(8):783-784. doi:10.1038/nmeth.3047
2. Orkin SH, Harosi FI, Leder P. Differentiation in erythroleukemic cells and their somatic hybrids. *Proc Natl Acad Sci U S A*. Jan 1975;72(1):98-102. doi:10.1073/pnas.72.1.98

Supplemental Figure S1

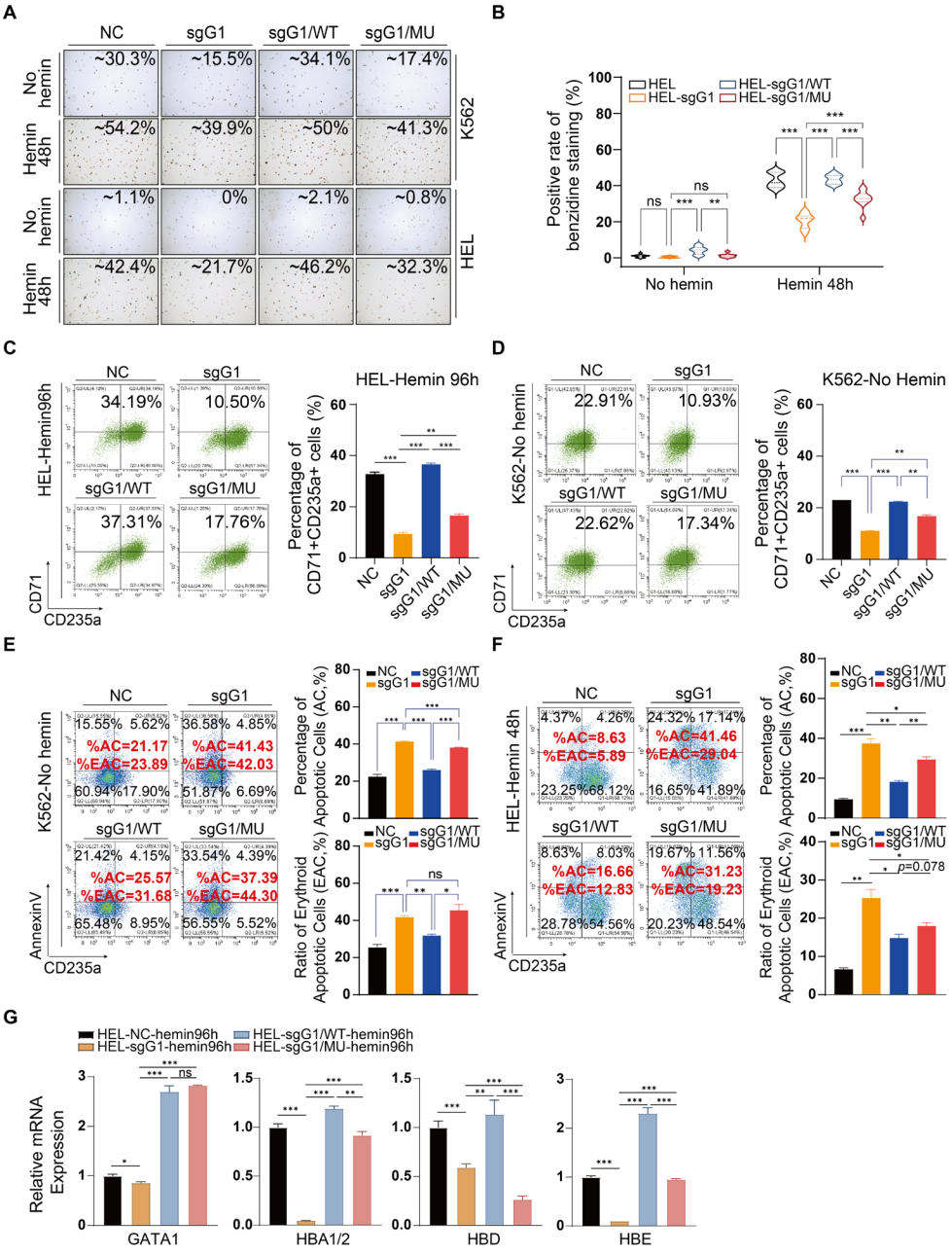


Supplemental Figure S2



Supplemental Figure S3

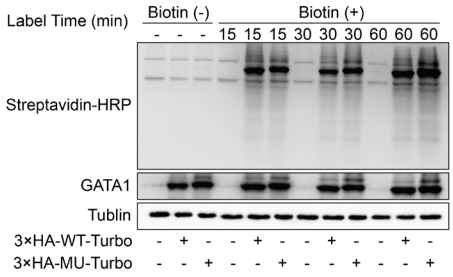
Supplemental Figure S3



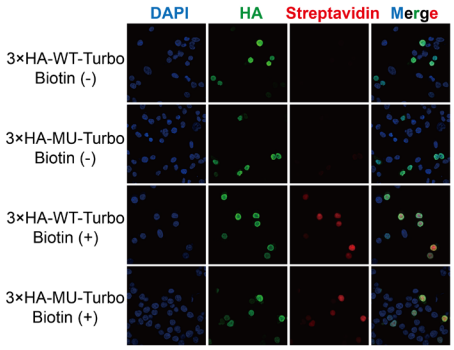
Supplemental Figure S4

Supplemental Figure S4

A



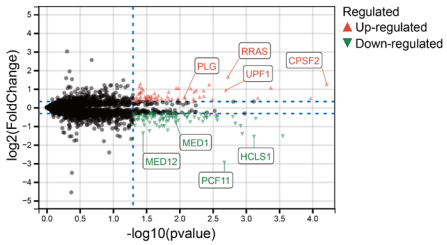
B



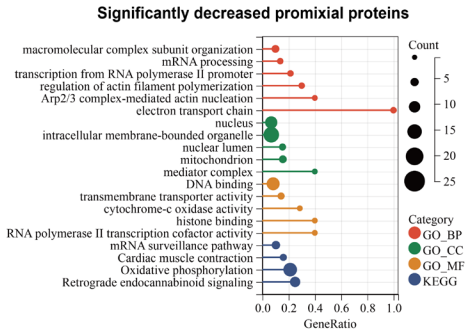
C

Streptavidin pulldown proximity labeling proteins	Turbo-3xHA-GATA1-WT			Turbo-3xHA-GATA1-MU		
Replicate number	#1	#2	#3	#1	#2	#3
Total number of identified proteins	3473	3414	3400	3417	3443	3410
HBG1	0	0	0	2691464.794	1889042.761	2016545.286
MPC1	0	0	0	2295566.46	1950745.601	2592849.807
NDUFS6	2421278.25	1622373.123	1704854.652	0	0	0
FBXL12	738894.3125	704674.6836	803811.6507	0	0	0
PIGK	516467.1875	1043007.591	1004770.237	0	0	0
SKA1	2203040	1959978.64	1608287.985	0	0	0
ITGA2B	672053.5625	741038.8296	784033.81	0	0	0
.....						

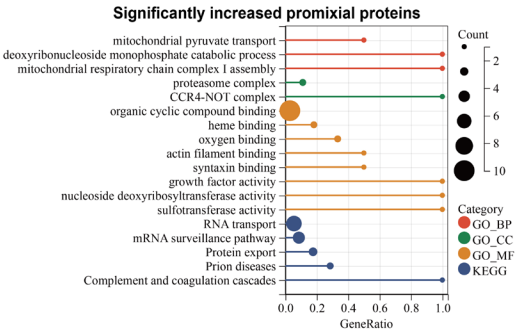
D



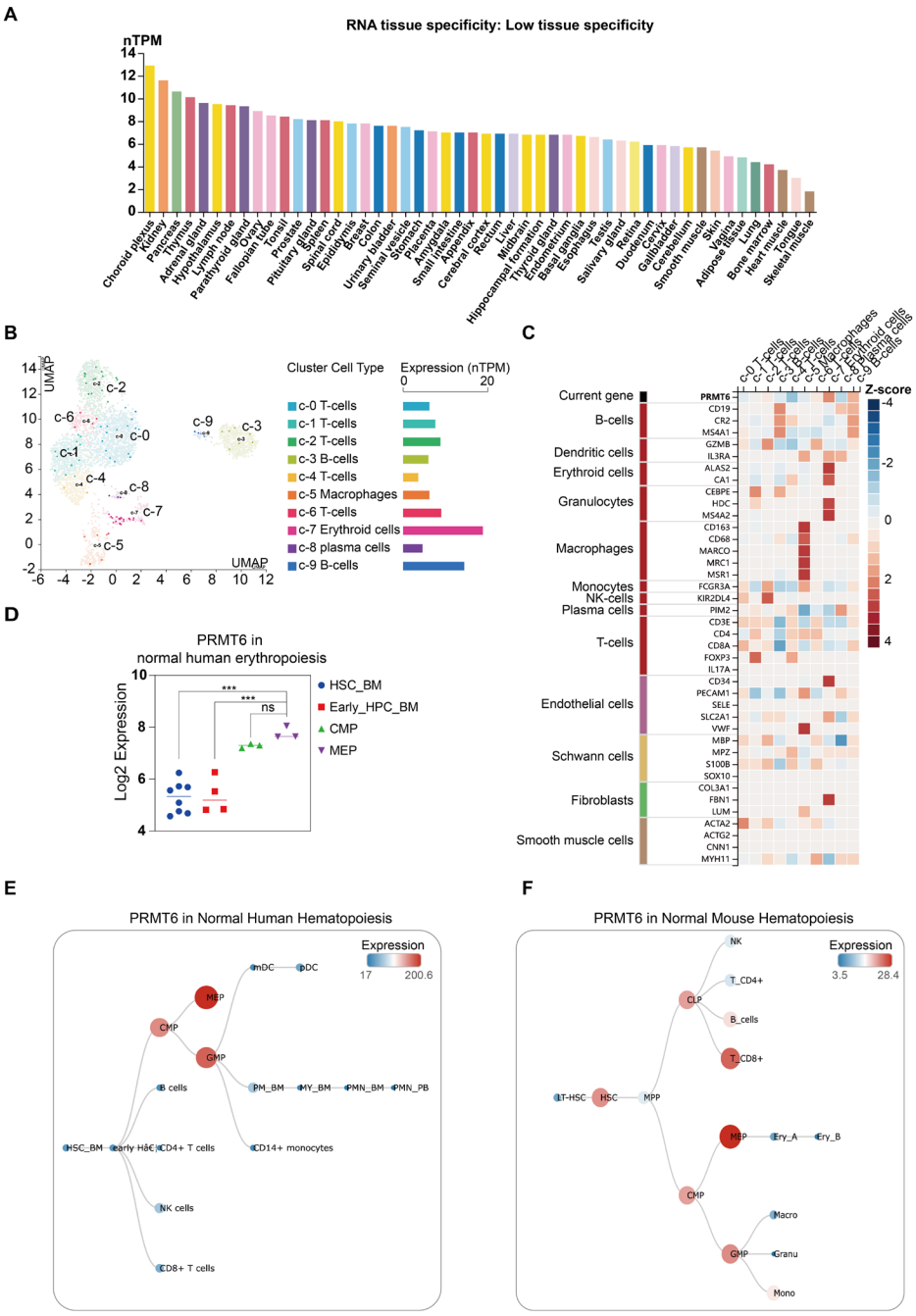
E



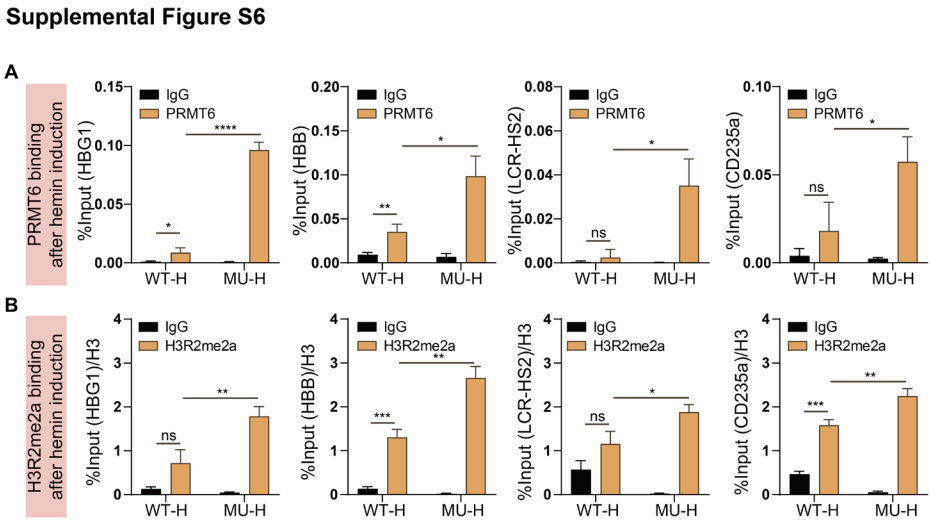
F



Supplemental Figure S5



Supplemental Figure S6



Supplemental Figure Legends

Supplemental Figure S1. The erythrocyte characteristics were observed in the bone marrow of the proband.

A. To examine erythrocyte characteristics, Wright-Giemsa staining was conducted on bone marrow smears at the first month (0Y1M10D) and the sixth month (0Y6M29D) after birth.

Supplemental Figure S2. Verification of re-expression GATA1 Leu387fs model in K562 and HEL cells by western blot and RT-qPCR.

A. Western blot showed the verification of the CRISPR/cas9-mediated GATA1-depleted effect in K562 and HEL cells, and the wild-type K562 and HEL cells were taken as control, respectively.

B. The re-expression effect of the wild-type and mutant GATA1 was detected through western blotting analysis, with GAPDH used as the loading control. The intensity of bands was quantified and normalized to NC control.

C. RT-qPCR assays were conducted to evaluate the relative GATA1 expression level in both GATA1-depleted and re-expression cells. The RT-qPCR results are presented as mean ± SD of 3 independent replicates. Statistical significance was determined as follows: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$; ns, no statistically significant.

Supplemental Figure S3. GATA1 Leu387fs mutant cells exhibited impaired differentiation during erythropoiesis compared to wild-type cells, as shown in Figure 2.

A. The representative benzidine staining images in both mutant model cell sets under no induction and 48h hemin induction conditions at a 100X magnification.

B. The positive rate of benzidine staining for hemoglobin of HEL cells was measured through a violin diagram, which indicated significant differences between the hemin-

treated and untreated groups.

C. Flow cytometric analysis of erythroid differentiation was carried out to quantify the percentage of surface CD71+CD235a+ population in HEL cells treated with hemin.

D. Flow cytometry analysis of spontaneous erythroid differentiation in K562 mutant cells. The bar graph presents the analyzed percentage of surface CD71+CD235a+ cells.

E. Flow cytometry analysis of erythroid apoptosis in K562 mutant cells without erythroid induction. The representative FACS plots show the percentages of Annexin V and CD235a expressed on the surface of K562 mutant and wild-type cells following 1% serum-induced apoptosis for 48 hours. The bar graphs describe the statistical analysis of apoptotic cells (AC%) and erythroid apoptotic cells (EAC%), as calculated in Figure 2E.

F. Flow cytometric analysis of erythroid apoptosis in erythroid-induced mutant HEL cells. Representative FACS plots displaying the percentages of AnnexinV and CD235a expressed in mutant and wild-type cells after hemin treatment for 48 hours. The bar graphs (right) show the statistical analysis of apoptotic cells (%AC) and erythroid apoptotic cells (%EAC), as calculated in Figure 2E.

G. RT-qPCR was performed on HEL mutant group cells to assess the expression of target genes after 96h hemin induction, related to Figure 2F. Results were normalized with each value of the housekeeping gene GAPDH.

The data are shown as mean ± SD of 3 independent replicates, and the statistical significance is represented by asterisks, *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns, no statistical significance.

Supplemental Figure S4. Proximity labeling experiments by Turbo-ID method, related to Figure 4.

A. Assessment of the Turbo-ID labeling system in HEK293T cells by transfecting either wild-type or Leu387fs mutant GATA1 protein fused with 3xHA-tagged Turbo-ID. Upon transfection, biotin was added for 0, 15, 30, and 60 minutes to enable labeling. We evaluated the transfection level and abundance of biotinylated protein by analyzing

the GATA1 and streptavidin bands, respectively.

B. Confocal fluorescence imaging was performed to visualize the cellular localization of the transfected GATA1-Turbo-ID fusion protein (green) and the biotinylated proteins (red) at a magnification of 200X. Biotin-labeled and unlabeled cells were fixed and stained with anti-HA-Alexa Fluor 488 and streptavidin–Alexa Fluor 594, respectively. The cellular nucleus was marked with DAPI (blue).

C. The label-free quantitative mass spectrometry results showed the most significant differential biotinylated proteins between the wild-type and mutant groups, as illustrated in the summary table. Both groups were the results of three independent replicates.

D. Volcano plot showing the differential proximity proteins in the mass spectrometry result. The y-axis shows the log2-fold change in protein abundance, and the x-axis shows the -log10 adjusted p-value (Student's two-sided t-test with Benjamini–Hochberg adjustment for multiple comparisons). Significantly decreased proximity proteins ($p < 0.05$, Foldchange < 0.75) in the mutant cells are shown by green inverted triangles, increased proximity proteins ($p < 0.05$, Foldchange > 1.25) are shown with red triangles, and proteins with insignificant differences are represented with black dots. Several proteins with marked changes in proximity are indicated.

E, F. GO and KEGG analysis of significantly decreased and increased proximal proteins by GATA1 Leu387fs mutation identified in the Turbo-ID-based mass spectrum assay. The top remarkably enriched processes or terms are shown in the plots.

Supplemental Figure S5. Evaluation of PRMT6 expression in tissue cells and during development.

A. mRNA expression profiles of PRMT6 in various human tissues from the Consensus transcriptomics datasets of Human Protein Atlas (available from https://www.proteinatlas.org/), ranked by normalized expression (nTPM) level.

B. Analysis of PRMT6 expression from the bone marrow scRNA-seq datasets on Human Protein Atlas. 9 single-cell clusters (c1-c9) were observed and visualized by a Uniform Manifold Approximation and Projection (UMAP) plot. The color intensity of

individual cells depends on the percentage of maximum expression (log2(read_count + 1)/log2(max(read_count + 1) * 100) in five different bins (i.e., <1%, <25%, <50%, <75%, and ⩾75%). Bar graph illustrating PRMT6 expression for each bone marrow cell cluster: pTPM, protein-transcripts per million. Image credit: Human Protein Atlas.

C. Heat map of Z-scores for PRMT6 expression in bone marrow cell clusters relative to other established cell type markers. Image credit: Human Protein Atlas.

D. Log2 expression of PRMT6 in different stages of human erythropoiesis (one-way ANOVA, Dunnett test), taken from the Bloodspot database (https://www.bloodspot.eu/). HSC_BM, hematopoietic stem cells from bone marrow; Early_HPC_BM, Hematopoietic progenitor cells from bone marrow; CMP, Common myeloid progenitor cell; MEP, Megakaryocyte-erythroid progenitor cell. The data are shown as mean ± SD, and the statistical significance is represented by asterisks, ***, $p < 0.001$; ns, no statistical significance.

E. The mRNA expression of PRMT6 in normal human and murine hematopoietic cells. The size and color of the dots correlate with the abundance of PMRT6 mRNAs. Data were obtained from the Bloodspot database. HSC_BM, Hematopoietic stem cells from bone marrow; early HPC_BM, Hematopoietic progenitor cells from bone marrow; CMP, Common myeloid progenitor cell; GMP, Granulocyte monocyte progenitors; MEP, Megakaryocyte-erythroid progenitor cell; PM_BM, Promyelocyte from bone marrow; MY_BM, Myelocyte from bone marrow; PMN_BM, Polymorphonuclear cells from bone marrow; PMN_PB, Polymorphonuclear cells from peripheral blood; NK cells, CD56+ natural killer cells; mDC, CD11c+ myeloid dendritic cells; pDC, CD123+ plasmacytoid dendritic cells.

F. The mRNA expression of PRMT6 in murine hematopoietic cells. The size and color of the dots correlate with the abundance of PMRT6 mRNAs. Data were obtained from the Bloodspot database. LT-HSC, Long term hematopoietic stem cell; HSC, Hematopoietic stem cell; MPP, multipotent progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte monocyte progenitor; Macro, bone marrow macrophages; Granu, granulocytes; Mono, Monocytes; B_cells,

B cells; T_CD4+, CD4+ T cells; T_CD8+, CD8+ T cells; NK, NK cells; MEP, Megakaryocyte-erythroid progenitor; Ery_A, erythrocytes A (CD71+); Ery_B, erythrocytes B (CD71-).

Supplemental Figure S6. The occupancy changes of PRMT6 and H3R2me2a modification on the promotor/enhancer of GATA1 erythroid target genes by GATA1 Leu387fs mutation, related to Figure 5.

ChIP-qPCR assays were performed for the binding of HA-tagged GATA1, PRMT6, and H3R2me2a modification on the target genes HBG1, HBB, CD235a (GYPA) promoter, and LCR-HS2 distal enhancer in the re-expression wild-type and mutant K562 cells post 40 μM hemin-induced erythroid differentiation for 48h.

A. PRMT6 binding results post hemin induction.

B. H3R2me2a binding results with histone H3 as control after hemin induction.

The data are presented as mean ± SD of 3 independent replicates by percent enrichment relative to input, and statistical significance is denoted by asterisks. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$; ns, differences not statistically significant.

Supplemental Tables provided as Excel files only

Supplemental Table S1. RBC parameters at different periods of the proband.

Abnormal parameters are indicated by bold font. Normal reference ranges are shown in the table header. RBC, red blood cell count; Hb, hemoglobin concentration; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-SD, red cell distribution width-standard deviation; RDW-CV, red cell distribution width-coefficient of variation; RET#, absolute reticulocyte count; RET%, reticulocyte percentage; RET-He, reticulocyte hemoglobin equivalent.

Supplemental Table S2. Mainly hematological indices besides red blood cells of the

proband.

Abnormal parameters are indicated by bold font. Normal reference ranges are shown in the table header. LYMPH#, absolute lymphocyte count; LYMPH%, lymphocytes percentage; NEUT#, absolute neutrophil count; NEUT%, neutrophil percentage; EO#, absolute eosinophils count; BASO#, absolute basophil granulocyte count; PLT, Platelet Count; MPV, mean platelet volume; PDW, Platelet Distribution Width.

Supplemental Table S3. Partial hematology results of the proband's brother.

Abnormal parameters are indicated by bold font. Normal reference ranges are shown in the table header.NEUT#, absolute neutrophil count; PLT, Platelet Count; Hb, hemoglobin concentration.

Supplemental Table S4. Sequences of Oligos applied in this study.

Sequences of Oligos were applied for sgRNA-knockout, shRNA-knockdown, RT-qPCR, and chIP-qPCR experiments in this study.

Supplemental Table S5. Relative protein quantification based on analysis of Turbo-ID mass spectrometry.

Protein: protein IDs identified to the sequence; Gene: gene name; MU_1, MU_2, MU_3: relative protein expression of three individual mutant GATA1 samples; WT_1, WT_2, WT_3: relative protein expression of three individual wild-type GATA1samples; MU.vs.WT FC, fold change of protein expression in the mutant group compared to the wild-type group; MU.vs.WT Pvalue is the p-value for the two-group test; MU.vs.WT log2FC, logarithms of fold change by the mutant group compared to the wild-type group.