Supplemental Figure Legends

Supplemental Figure 1. An unbiased CRISPR/Cas9 screen for epigenetic regulators identifies KMT2D as a critical regulator of erythropoiesis.

(A). Gene ontology analysis of the biological process of all significant positive regulators identified in the screen. The relative P value was calculated from the enrichR online software.

(B). KMT2D mRNA levels detected by RT-qPCR at different stages after induced erythroid maturation of $CD34^{+}$ HPSCs.

(C). CCK8 assay displaying the proliferation status of HUDEP-2 cells treated with KMT2D-sg1, KMT2D-sg2 and Ctrl. sgRNAs and cultured in maintenance medium over a time course of 5 days. n.s., not significant, unpaired Student's t-test.

(D). Bar plot displaying the indel frequency obtained from two different KMT2D sgRNAs at day 0 and day 6 of differentiation. n.s. not significant, unpaired Student's t-test.

(E). Quantitative flow cytometry analysis of HUDEP-2 cells treated with Ctrl, KMT2D-sg1, and KMT2D-sg2 sgRNAs and stained for the erythroid differentiation marker CD235a.**$P < 0.01$, unpaired Student's t-test.

Supplemental Figure 2. KMT2D is required for activation of erythroid signature genes and proper chromatin accessibility of GATA1.

(A). Principal component analysis of RNA-seq data from non-targeting sgRNA control (Ctrl) and KMT2D-sgRNA2 mediated bulk knockout (KMT2D KO) HUDEP-2 cells before (day 0) and after three days of maturation (day 3).

(B). Gene set enrichment analysis (GSEA) of RNA-seq data from KMT2D KO versus control HUDEP-2 cells before (day 0) and after three days of maturation (day 3).

(C). Gene ontology analysis of the biological process from the downregulated genes of **Figure 2B** before (day 0) and after three days of maturation (day 3). The relative P value was calculated from the enrichR online software.

(D). TF enrichment analysis of the downregulated genes of **Figure 2B** before (day 0) and after three days of maturation (day 3). The relative P value was calculated from the enrichR online software.

(E). Heatmap centered at ATAC-seq nucleosome-free peak summits for 119 increased regions and 659 decreased regions from non-targeting sgRNA control (Ctrl) and KMT2D-sgRNA2 mediated bulk knockout (KMT2D KO) in undifferentiated HUDEP-2 cells.

(F). Volcano plot of motif enrichment analysis of ATAC-seq comparisons: unchanged control regions versus decreased regions in the HUDEP-2 cells. The *P* values and odds ratios were calculated via Fisher's exact tests to compare the frequency of regions containing a motif with those that do not. Each dot represents a motif in the database.

(G). Homer motif enrichment analysis of the top 10 TF motifs at ATAC down regions. P values, hypergeometric test.

Supplemental Figure 3. Acute depletion of KMT2D selectively suppresses the expression of GATA1 targets.

(A). Schematic diagram of the knock-in design and genotyping PCR primer design for the KMT2D^{AID2} reporter allele.

(B). Genotyping PCR confirming the successful integration of the EGFP-AID cassette into the endogenous KMT2D locus for two K562 cell clones and one HUDEP-2 cell clone. PCR products from the 5' and 3' knock-in boundaries were sequenced to verify the seamless knock-in of the EGFP-AID cassette.

(C). Flow cytometry analysis of AID-KMT2D K562 cells following 1 μM 5-Ph-IAA treatment. The fluorescence of EGFP-AID-KMT2D was detected before 5-Ph-IAA treatment (black histogram), after 8 hours of 5-Ph-IAA treatment (red histogram) and following washout of 5-Ph-IAA after 48 hours (blue histogram).

(D). Western blot showing KMT2D, KDM6A (UTX) and H3 protein expression in nuclear extracts of AID-KMT2D K562 cells before treatment with 5-Ph-IAA, after treatment with 5-Ph-IAA for 8 hours and following washout of 5-Ph-IAA after 48 hours.

(E). Immunoprecipitation-mass spectrometry (IP-MASS) data of KMT2D immunoprecipitations (IPs) from nuclear extracts of AID-KMT2D K562 cells identifies all members of the KMT2D complex. KMT2D IPs were normalized to control IgG IPs.

(F). Interaction Network Analysis of IP-MS data performed on KMT2D relative to IgG IPs from nuclear extracts of AID-KMT2D K562 cells.

(G). Schematic diagram of the experimental process for inducing erythroid differentiation of AID-KMT2D HUDEP2 cells.

(H). Proliferation assay of AID-KMT2D HUDEP2 cells after six days of induced erythroid expansion. AID-KMT2D HUDEP2 cells were treated with either DMSO or 5-Ph-IAA. n.s., not significant, unpaired Student's t-test.

(I). Proliferation assay of AID-KMT2D HUDEP2 cells after six days of induced erythroid maturation. AID-KMT2D HUDEP2 cells were treated with either DMSO or 5-Ph-IAA (upper panel); cell pellets of AID-KMT2D HUDEP-2 cells at day 6 of induced erythroid maturation treated with either DMSO or 5-Ph-IAA (lower panel). n.s., not significant, unpaired Student's t-test.

(J). RT-qPCR showing mRNA levels of *SLC4A1, GYPA, HBA, and HBB* normalized to β-Actin in AID-KMT2D HUDEP2 cells at day 3 and day 6 of induced erythroid maturation treated either with DMSO or 5-Ph-IAA. ****$P < 0.0001$, unpaired Student's t-test.

(K). Principal component analysis of RNA-seq data from AID-KMT2D HUDEP-2 and K562 cells treated with 5-Ph-IAA for 0, 6, 12, or 24 hours.

Supplemental Figure 4. KMT2D and GATA1 co-occupy erythroid-expressed genes.

(A). Principal component analysis from KMT2D ChIP-seq of AID-KMT2D HUDEP-2 cells treated with or without 1 μM 5-Ph-IAA for 24 hours.

(B). Genomic heatmaps of KMT2D peaks from three replicates centered on the summit of KMT2D bound regions in AID-KMT2D HUDEP-2 cells untreated or treated with 5-Ph-IAA for 24 hours.

(C). Homer motif enrichment analysis of the top 10 TF motifs at KMT2D bound enhancers and promoters. P values, hypergeometric test.

(D). Co-immunoprecipitation using MYC^{Tag} in K562-MYC^{Tag}-GATA1 overexpression cells, followed by Western blot detection of proteins interacting with GATA1, including KDM6A (UTX) and KMT2D.

(E). Co-immunoprecipitation using GATA1 antibody in K562 cells, followed by Western blot detection of GATA1, KDM6A (UTX), and KMT2D protein levels.

(F). Box plots show the basal expression levels of genes linked to KMT2D unbound Down DEGs and KMT2D bound Down DEGs from **Figure 4H** in HUDEP-2 cells and in $CD34^{+}$ HSPCs undergoing erythroid induction. **** $P < 0.0001$ according to the Wilcoxon test.

Supplemental Figure 5. KMT2D is required for enhancer activation of select GATA1 target genes.

(A). Gene ontology analysis of the enriched biological processes from KMT2D-dependent active enhancers and KMT2D-independent active enhancers in HUDEP-2 cells. The relative *FDR* value was calculated via the GREAT database.

(B). Homer motif enrichment analysis of the top 10 TF motifs at KMT2D-dependent active enhancers and KMT2D-independent active enhancers. P values, hypergeometric test.

(C). Genome browser tracks of the *SLC4A1* loci. Displayed are the ChIP-seq profiles of H3K4me1, H3K27ac, and GATA1 in the non-targeting sgRNA control (Ctrl) and KMT2D-sg2 mediated knockout (KO) HUDEP-2 cells and KMT2D signal in AID-KMT2D HUDEP-2 cells untreated with 5-Ph-IAA. The grey rectangles highlight the three active enhancer regions (S1, S2, S3), respectively.

(D). RT-qPCR verifying the relative fold enrichment of H3K4me1 and H3K27ac signal on three active enhancer regions (S1, S2, S3) of *SLC4A1* loci by CUT&RUN assay in in AID-KMT2D HUDEP-2 cells before and after treated with 5-Ph-IAA for 24 hours. ****P < 0.0001, unpaired Student's t-test.

(E). RT-qPCR verifying the relative fold enrichment of GATA1 signal on three example genes *SLC4A1, ZFPM1, and EPOR* loci by CUT&RUN assay in in AID-KMT2D HUDEP-2 cells before and after treated with 5-Ph-IAA for 24 hours. n.s. not significant, unpaired Student's t-test.

Supplemental Figure 6. KMT2D is required for the survival and maturation of primary human erythroblasts.

(A). Representative flow cytometry plots of $CD34^+$ HSPCs electroporated with RNPs consisting of Cas9 and either non-targeting control (Ctrl) or KMT2D (KMT2D-sg2) sgRNAs and stained for CD34 over seven days expansion in maintenance medium.

(B). Representative flow cytometry plots and analysis of $CD34^+$ HSPCs electroporated with RNPs consisting of Cas9 and either non-targeting control (Ctrl) or KMT2D (KMT2D-sg2) sgRNAs and stained for apoptosis. n.s not significant, unpaired Student's t-test.

(C). Cell counting assay showing the proliferation ability of $CD34^+$ HSPCs treated with RNPs consisting of Cas9 and either non-targeting control (Ctrl) or KMT2D (KMT2D-sg2) sgRNAs in myeloid differentiation medium (left panel). Indel efficiency of $CD34^+$ HSPCs edited at various time points of induced myeloid maturation (right right). $^{**}P < 0.01$, $^{***}P < 0.001$, unpaired Student's t-test for the growth curve while paired Student's t-test for the indels.

(D). Representative flow cytometry plots and analysis of $CD34^+$ HSPCs and stained for CD11b after fourteen days of myeloid differentiation. $^{*}P < 0.05$, unpaired Student's t-test.

(E). RT-qPCR from $CD34^+$ HSPCs treated with luciferase control (Ctrl) and KMT2D (KMT2D-sh1 and KMT2D-sh2) shRNAs showing *KMT2D* mRNA expression levels normalized to β-Actin after seven days of induced maturation. Data are shown as the mean ± SEM of three replicates. $^{**}P < 0.01$, $^{***}P < 0.001$, unpaired Student's t-test.

(F). Colony formation ability of $CD34^+$ HSPCs after lentiviral transduction with shRNAs against luciferase (Ctrl) or KMT2D (KMT2D-sh1 and KMT2D-sh2). $^{*}P < 0.05$, $^{**}P < 0.01$, unpaired Student's t-test.

(G). Representative flow cytometry plots showing CD235a, CD49d and Band3 expression at the indicated time points after induced erythroid maturation of $CD34^+$ HSPCs transduced with lentiviral shRNAs against luciferase (Ctrl) and KMT2D (KMT2D-sh1 and KMT2D-sh2). $^{**}P < 0.01$, $^{***}P < 0.001$, unpaired Student's t-test.

(H). May-Grünwald-Giemsa–stained erythroblasts at day 14 after induced erythroid maturation of $CD34^+$ HSPCs treated with shRNAs against luciferase (Ctrl) or KMT2D (KMT2D-sh1 and KMT2D-sh2). Red arrows denote immature erythroblasts. Scale bar, 10 μM.

(I). Cell pellets at day 10 of induced erythroid maturation after lentiviral transduction with sRNAs against luciferase (Ctrl) or KMT2D (KMT2D-sh1 and KMT2D-sh2).

Supplemental Table Legends (separate files)

Supplemental Table 1: Analysis of CRISPR screen using single-guide (sg) RNAs targeting epigenetic modifiers.

Supplemental Table 2: RNA transcriptome and ATAC-seq analysis of wild type and KMT2D-sgRNA2 mediated knockout (KMT2D KO) HUDEP-2 cells.

Supplemental Table 3: Gene Ontology analysis and Transcription factors occupancy analysis for the de-regulated genes between wild type and KMT2D-sgRNA2 mediated knockout (KMT2D KO) HUDEP-2 cells.

Supplemental Table 4: KMT2D interacting proteins list by the Immunoprecipitation mass spectrometry (IP-Mass) in AID-KMT2D K562 cells.

Supplemental Table 5: The RNA transcriptome analysis of AID-KMT2D HUDEP-2 cells with 5-Ph-IAA treatment for 6 hours, 12 hours and 24 hours vs. without 5-Ph-IAA treatment grown in the expansion culture.

Supplemental Table 6: RNA transcriptome analysis of AID-KMT2D K562 cells with 5-Ph-IAA treatment for 6 hours, 12 hours and 24 hours vs. without 5-Ph-IAA treatment grown in the expansion culture.

Supplemental Table 7: Transcription factors occupancy analysis for the de-regulated genes in AID-KMT2D K562 and HUDEP-2 cells with time serial analysis.

Supplemental Table 8: Integration analysis of ChIP-seq data for GATA1, H3K4me1, H3K27ac and KMT2D.

Supplemental Table 9: Table S9, Integration analysis of all ChIP-seq data for KMT2D, GATA1, H3K4me1, H3K27ac.

Supplemental Table 10: Oligonucleotides, DNA primers, and antibodies used in this study. Oligonucleotides were used for sgRNA vector construction; DNA primers were

used for the analysis of Cas9-mediated indels by Sanger sequencing; antibodies were used for Western blotting, immunoprecipitation.