Supplemental Methods

Cell lines and *in vitro* human erythroid differentiation

K562 cells were cultured in RPMI 1640 medium supplemented with 10% FBS. The *in vitro* erythroid differentiation protocol was described in Kang et al[1] and Madzo et al[2]. Briefly, human primary erythroblasts were generated by culturing $CD34^+$ early hematopoietic stem and progenitor cells initially isolated from growth factor-mobilized peripheral blood (purchased from ALL Cells, Inc.) using a CliniMacs device. The culture contained 15% fetal bovine serum, 15% human serum, Isocove's modified Dulbecco's medium (IMDM), 10ng/mL interleukin-3, 2units/mL EPO, and 50ng/mL SCF. During the initial 8 days of culture, cells were fed on days 3 and 6 by adding equal volumes of fresh culture media supplemented with growth factors. However, no new interleukin-3 was added after the initial addition on day 0, and the amount of SCF added to the fresh media was gradually decreased at each feeding (day 3, 25ng/mL; day 6, 10ng/mL; day 8, 2ng/mL). The amount of EPO added was 2units/mL during each feeding until day 8 of culture. Cells were fed one more time on day 10 of culture by adding equal volumes of fresh media with only EPO (2units/mL) during this final feeding. Cells were collected at various time points during the culture and stained with benzidine and hematoxylin. Flow cytometry was done with CD71 and GlyA to monitor the differentiation program.

Hypoxia treatment of cell lines

Hypoxia samples were cultured with 1% O_2 and 5% CO_2 in a hypoxia chamber, maintained at 37°C, whereas normoxia samples were cultured with 21% O_2 and 5% CO_2 at 37°C.

hMe-SEAL and enrichment analysis

5-hmC pull-down and sequencing were done as previously described.[3] Briefly, genomic DNA was sheared to ~300bp, and T4 β-gluclosyltransferase was used to label 5-hmC with azide-modified glucose. Biotin-S-S-DBCO was covalently attached to the modified glucose via the Click reaction. Labeled DNA

fragments were then enriched using streptavidin beads and purified for next-generation sequencing along with input DNA controls. Raw sequences in fastq format were aligned to hg19 reference genome using Burrows-Wheeler Aligner (BWA).[4] 5-hmC peaks were called using MACS2[5] with input sequences as controls. In order to compare 5-hmC peaks in hypoxia and normoxia, a unified peak set was made by merging both normoxia and hypoxia peaks with bedtools merge.[6] Normoxic and hypoxic 5-hmC densities in peak regions were then quantified by counting aligned reads in each peak region with HTseq-count[7] and converting to FPKM.

The genomic distribution of 5-hmC peaks was examined by calculating enrichment of 5-hmC peaks in particular genomic elements. Enrichment was defined as *(Observed overlap)/(Expected overlap)*, where expected overlap assumed 5-hmC peaks are distributed randomly across the genome. Bedtools shuffle and bedtools intersect were used to estimate expected overlaps with Monte Carlo simulations.

RNA-seq and data processing

Sample RNA libraries were prepared using the KAPA mRNA HyperPrep Kit (Roche, KK8580) with 1ug input RNA and submitted for sequencing. Raw reads in fastq format were aligned to the hg19 reference genome using Tophat2[8] with Gencode gene and transcript annotation (GRCh37.p13)[9]. Gene expression was quantified and compared by tools in the Cufflinks package [10].

HIF-1α ChIP-seq and data processing

K562 cells cultured in normoxia or hypoxia (1% O_2) for 72 hours were fixed by 1% formaldehyde for 10 minutes. Nuclear extraction and sonication of the cross-linked chromatin were performed using the Covaris truChIP kit (Covaris, 520154) and the Covaris S220 sonicator. HIF-1α bound DNA was enriched with rabbit anti HIF-1α antibody (Abcam, ab2185) and purified for sequencing. Raw sequences in fastq format were

aligned to the hg19 reference genome using BWA and HIF-1α peaks were called using MACS2 with input sequences as control.

Accession Numbers

GSE142870: hMe-SEAL, RNA-seq, and ChIP-seq data from this study.

HPLC-MS/MS quantification of global 5-mC and 5-hmC

Genomic DNA was first hydrolyzed to single nucleosides with nuclease P1 (Sigma-Aldrich, N8630), phosphodiesterase I (Sigma-Aldrich, P3243), and alkaline phosphatase (Thermo Fisher, 18009027). Hydrolyzed nucleosides were separated by Acquility UPLC Oligonucleotide BEH C18 Column (Waters, 186003950) in the Agilent 1290 LC system before measured in a Agilent 6460 Triple-Quadrupole MS/MS. 5-mC and 5-hmC contents were quantified as percentage of all cytosine species.

qPCR quantification of gene expression

cDNA libraries were made from 1ug input RNA using High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher, 4368814). Gene expressions were quantified using Power SYBR Green PCR Master Mix (Thermo Fisher, 4367659) with 18S rRNA as the housekeeping gene.

CRISPR-Cas9 targeted deletion

CRISPR guide sequences were inserted to the lentiCRISPR v2 plasmid (Addgene #52961) according to the manufacturer's protocol.[11] Active lentivirus was produced in HEK293T cells to transduce K562 cells. Successful transduction was selected via puromycin resistance. Single-cell clones were isolated from the transduced population, and the targeted site was sequenced from each clone. Clones that lost the core "ACGT" motif on

both alleles of *TET3* were used in hypoxia experiments. S1/S2 double deletion clones were made from validated single deletion clones by targeting the intact binding site using CRISPR-Cas9. Primers used for sequencing are as follows: TET3_S1_F: TTTTGAGGGATTGGGGGCTT

TET3_S1_R:CAGCACACAAGAACCAGGTC

TET3_S2_F: TGCTGTTCATGCTTGAGGGA

TET3_S2_R: GCAAATCTGTCAGTGGCTGG

Sodium butyrate induced differentiation of K562 cells

Sodium butyrate (Sigma-Aldrich, 303410) was dissolved in water and filtered to prepare 1M stock solution. For the differentiation assay, the stock solution was added to cells in fresh media at a 1:1000 ratio to a final concentration of 1mM.[12] Water was used as negative control.

Cytospin and benzidine-hematoxylin staining

K562 cells were counted, and 200,000 cells were collected by spinning at 400g. Culture media was discarded, and cells were re-suspended in 100 µL of PBS. The cells were spun onto glass slides at 600 rpm, which were allowed to dry for ~20 minutes before staining. The slides were stained in 0.2% benzidine solution in 0.5 M acetic acid for 4 minutes, followed by 2 minutes in 2% H_2O_2 in 70% ethanol, and 3 minutes washing in running tap water. After the wash, the slides were further stained in 1x hematoxylin solution for 1 minute, followed by another 3 minutes wash in running tap water. The slides were then air-dried and mounted.

Microscopy and quantification of benzidine staining

At least three randomly selected fields of stained cells were imaged at 400x magnification. Images were processed using the Color Deconvolution plugin in ImageJ software to separate hematoxylin and benzidine

staining to different channels.[13,14] The numbers of benzidine-positive cells and the intensity of benzidine staining were quantified in the benzidine channel, whereas the total cell number and abnormal nuclear morphology were quantified in the hematoxylin channel. Results of all fields from each slide were summed for the final quantification.

References

1. Kang J-A, Zhou Y, Weis TL, et al. Osteopontin regulates actin cytoskeleton and contributes to cell proliferation in primary erythroblasts. *J Biol Chem*. 2008;283(11):6997-7006.

2. Madzo J, Liu H, Rodriguez A, et al. Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell Rep*. 2014;6(1):231-244.

3. Song C-X, Szulwach KE, Fu Y, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*. 2011;29(1):68-72.

4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.

5. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.

6. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842.

7. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169.

8. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.

9. Frankish A, Diekhans M, Ferreira A-M, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res*. 2019;47(D1):D766-D773.

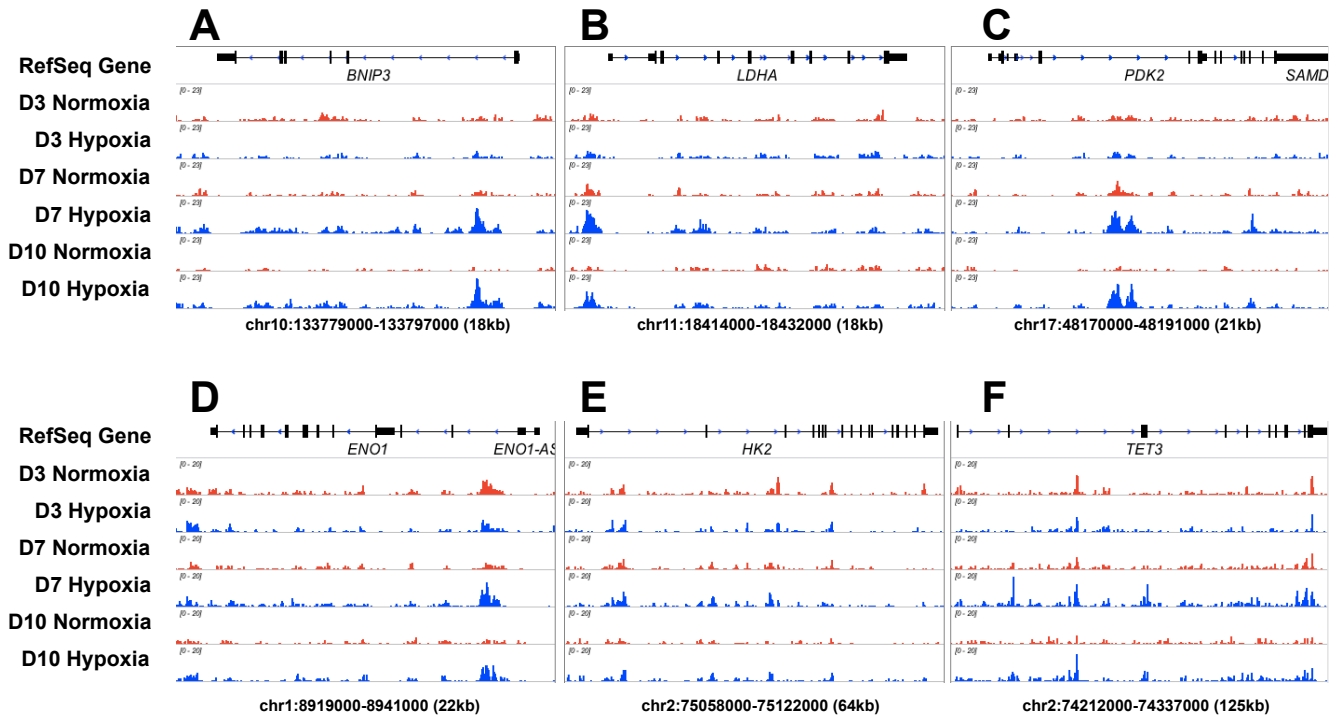
10. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7(3):562-578.

11. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014;11(8):783-784.

12. Shariati L, Modaress M, Khanahmad H, et al. Comparison of different methods for erythroid differentiation in the K562 cell line. *Biotechnol Lett*. 2016;38(8):1243-1250.

13. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol*. 2001;23(4):291-299.

14. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9(7):671-675.



Supplemental Figure 1: Representative 5-hmC distribution in hypoxia vs normoxia.

(A) 5-hmC distribution in and around *BNIP3* gene, showing normoxic and hypoxic conditions at days 3,7 and 10.

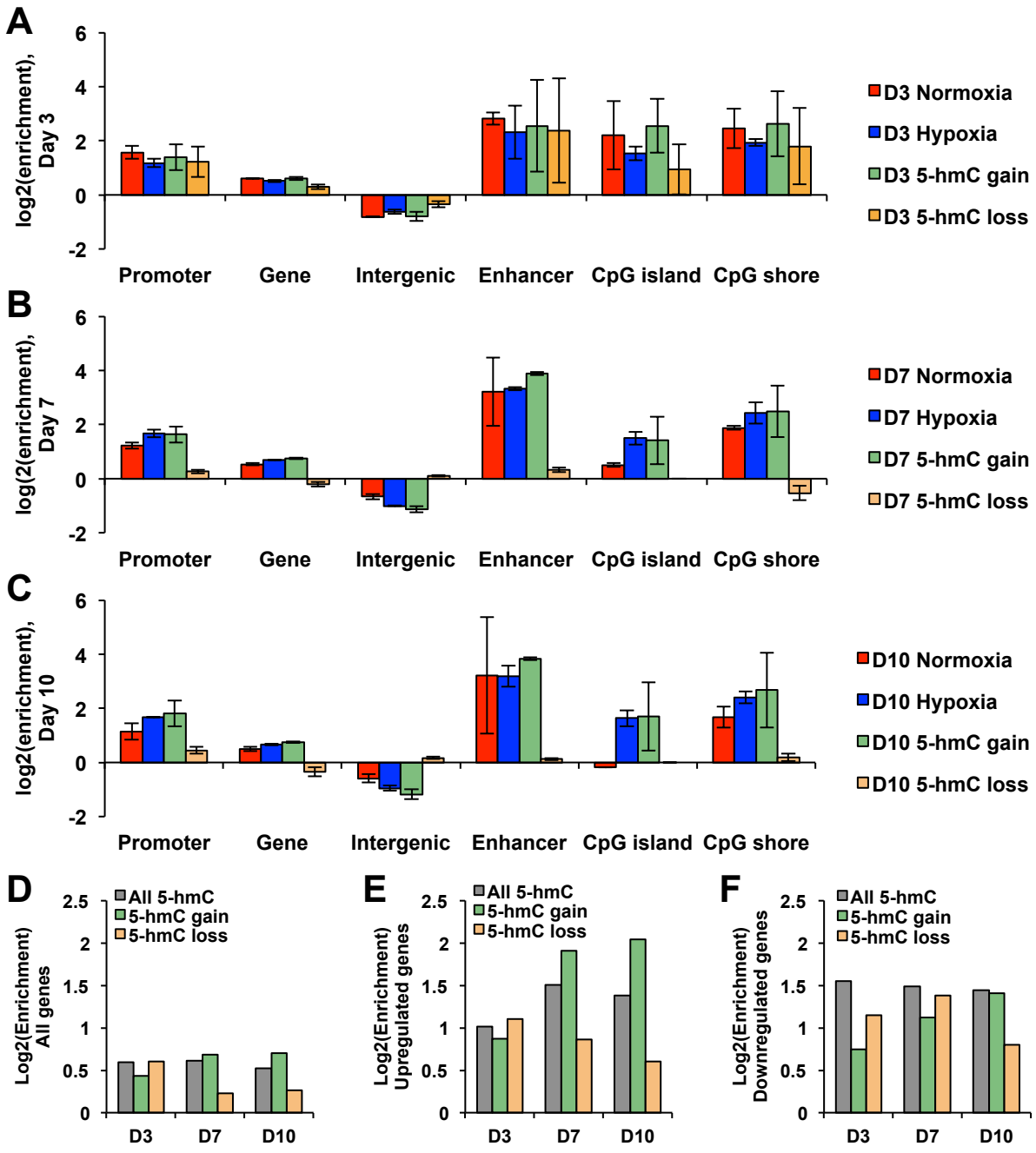
(B) 5-hmC distribution in and around *LDHA* gene, showing normoxic and hypoxic conditions at days 3,7 and 10.

(C) 5-hmC distribution in and around *PDK2* gene, showing normoxic and hypoxic conditions at days 3,7 and 10.

(D) 5-hmC distribution in and around *ENO1* gene, showing normoxic and hypoxic conditions at days 3,7 and 10.

(E) 5-hmC distribution in and around *HK2* gene, showing normoxic and hypoxic conditions at days 3,7 and 10.

(F) 5-hmC distribution in and around *TET3* gene, showing normoxic and hypoxic conditions at days 3,7 and 10.



Supplemental Figure 2: Genomic distribution of 5-hmC peaks.

(A) Day 3 5-hmC enrichment in genomic regions, normoxia vs. hypoxia, as well as 5-hmC peaks that gained or lost 5-hmC in hypoxia.

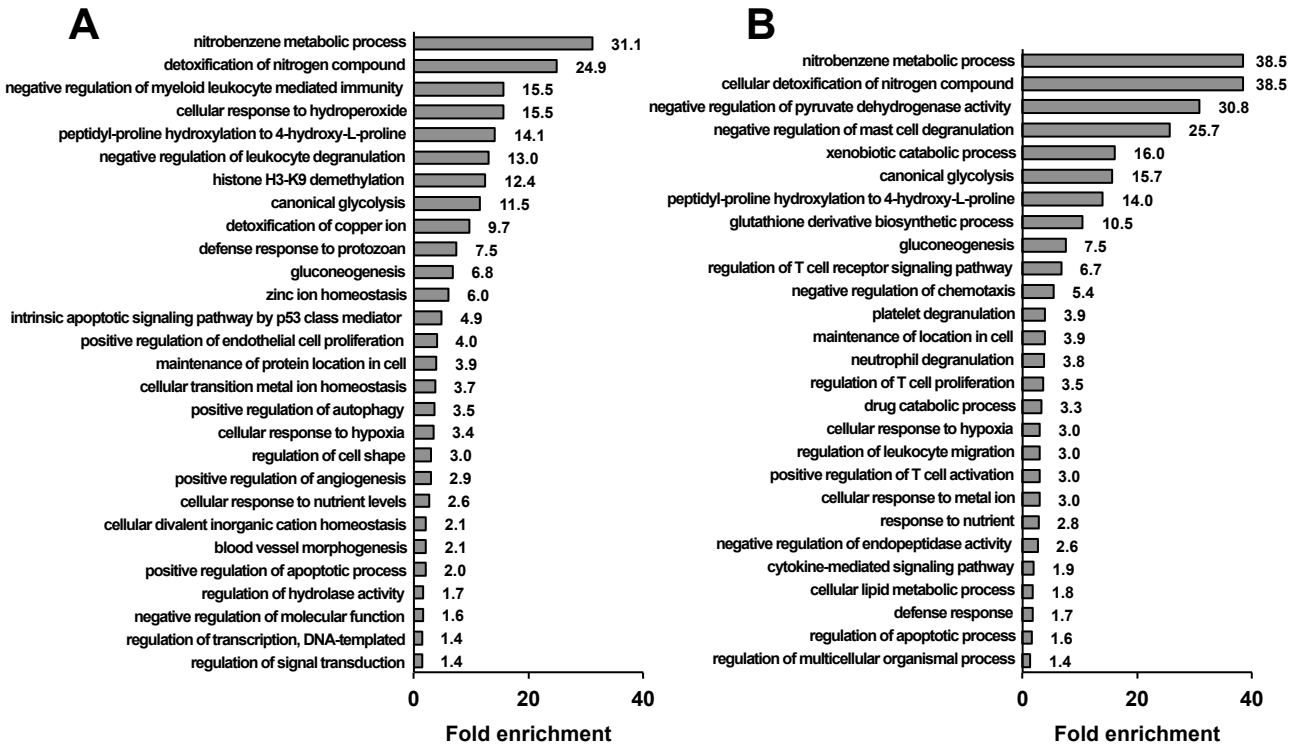
(B) Day 7 5-hmC enrichment in genomic regions, normoxia vs. hypoxia, as well as 5-hmC peaks that gained or lost 5-hmC in hypoxia.

(C) Day 10 5-hmC enrichment in genomic regions, normoxia vs. hypoxia, as well as 5-hmC peaks that gained or lost 5-hmC in hypoxia.

(D) Enrichment of 5-hmC peaks in all genes.

(E) Enrichment of 5-hmC peaks in genes upregulated by hypoxia.

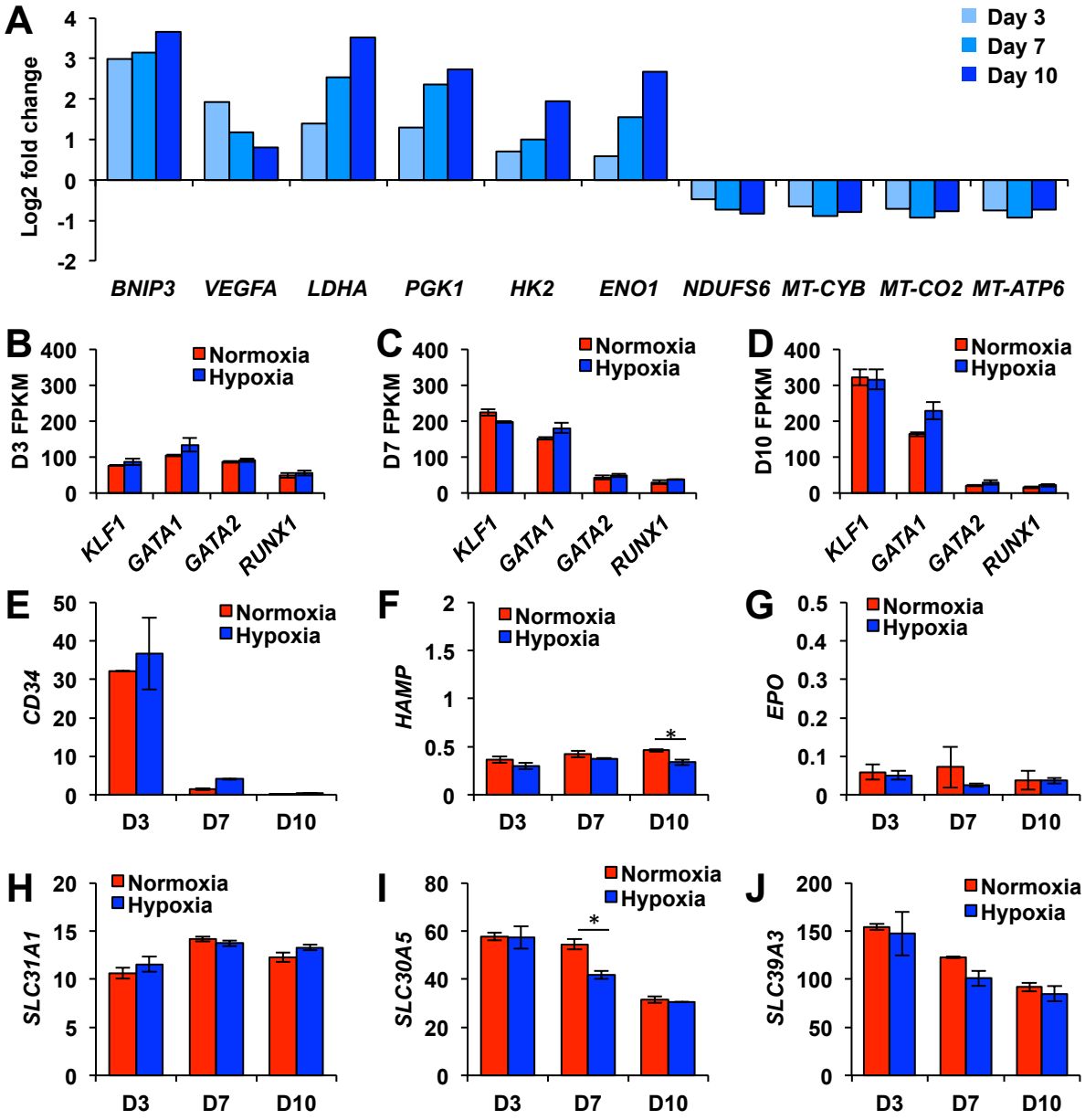
(F) Enrichment of 5-hmC peaks in genes downregulated by hypoxia.



Supplemental Figure 3: GO enrichment analysis of genes upregulated in hypoxia.

(A) GO term enrichment for all genes upregulated in hypoxia at day 7. All terms shown have FDR<0.05.

(B) GO term enrichment for all genes upregulated in hypoxia at day 10. All terms shown have FDR<0.05.



Supplemental Figure 4: Hypoxia does not change expressions of erythroid transcription factors.

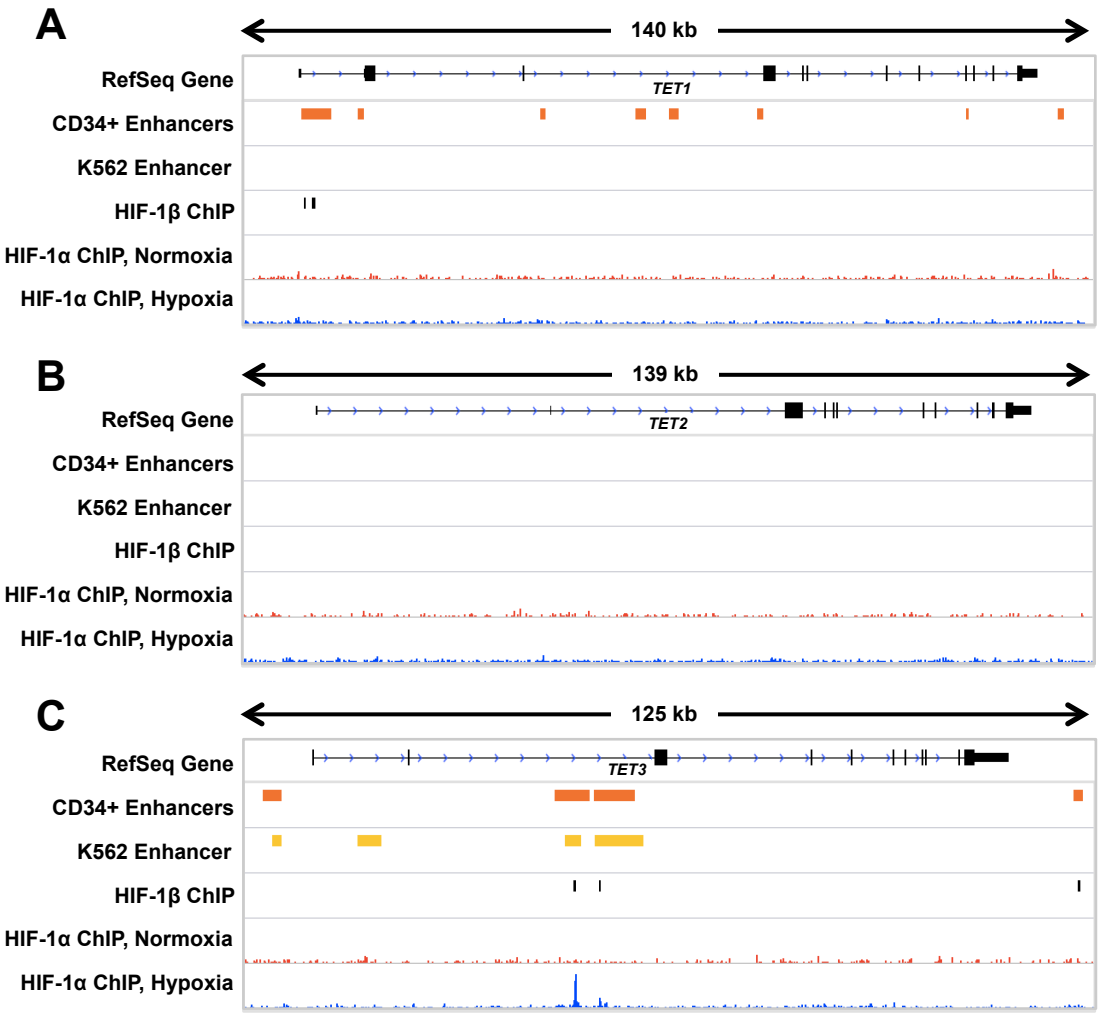
(A) Hypoxia treatment upregulates HIF-1α target genes and genes encoding glycolysis enzymes, while downregulates genes encoding proteins in the electron transport chain.

(B-D) Expression of erythroid transcription factors *KLF1*, *GATA1*, *GATA2*, and *RUNX1* was not significantly changed by hypoxia at day 3 (B), day 7 (C), or day 10 (D).

(E) HSPC marker CD34 expression rapidly decreased after day 3 of erythroid differentiation.

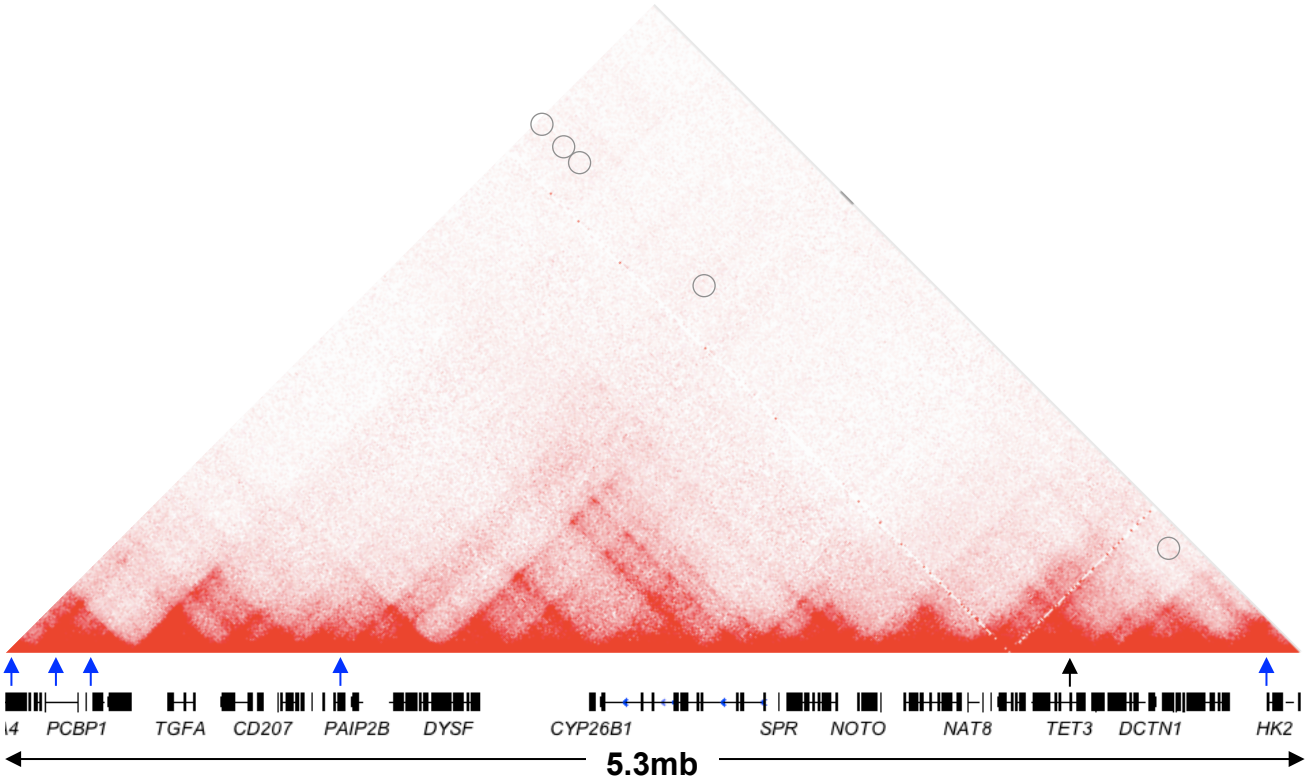
(F-G) *Hepcidin* (*HAMP*) and *erythropoietin* (*EPO*) expression were low in erythrocytes and were not induced by hypoxia.

(H-I) Expression of the copper transporter *SLC31A1* and zinc transporters *SLC30A5* and *SLC39A3* were not induced by hypoxia.



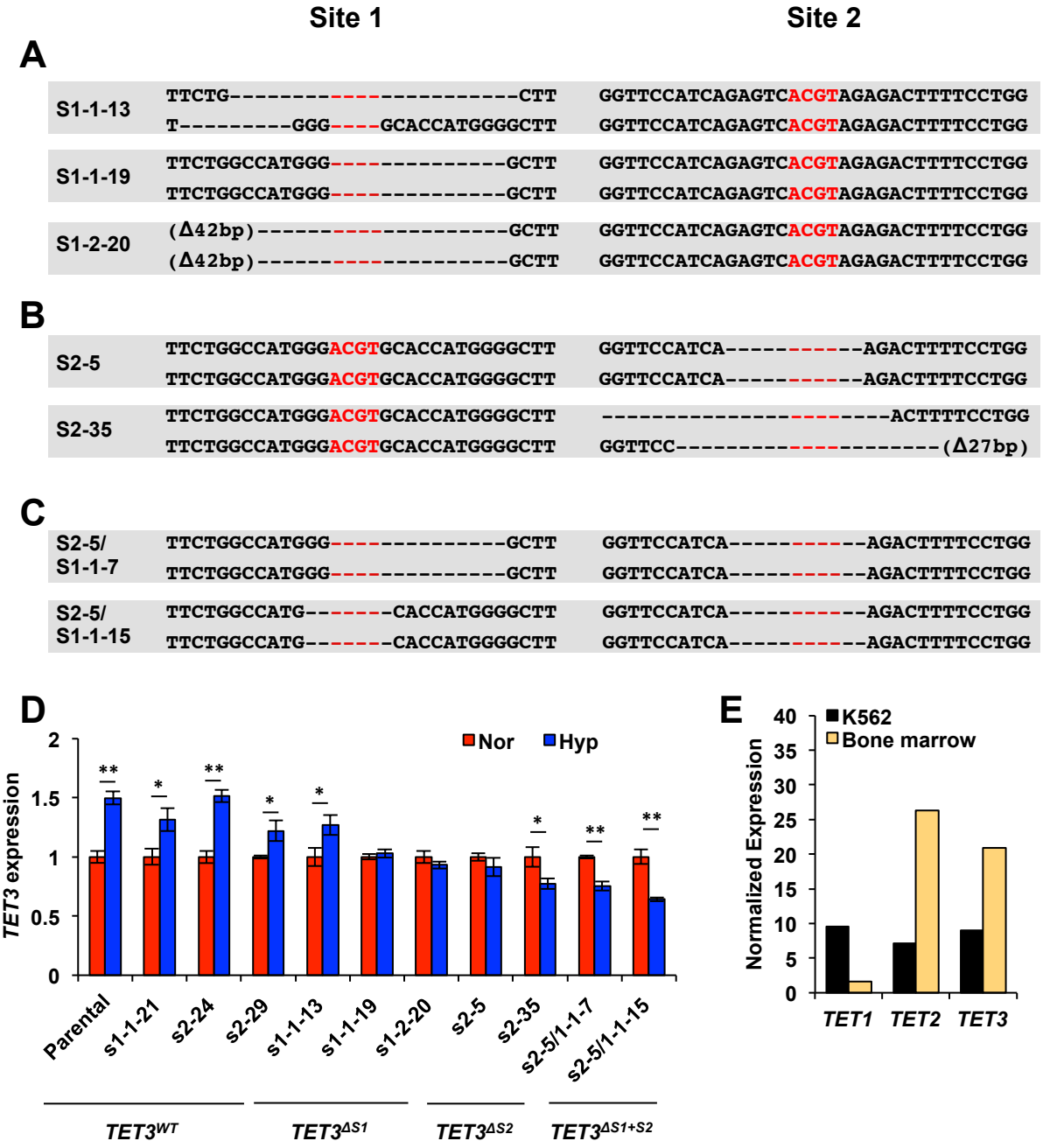
Supplemental Figure 5: HIF-1 only binds *TET3*, but not *TET1* and *TET2*, in K562 cells.

K562 HIF-1α ChIP-seq revealed no binding site at *TET1* (A) or *TET2* (B). *TET3* (C) is the only *TET* genes that was bound by HIF-1α in K562 cells



Supplemental Figure 6: Hi-C heatmap of the genomic region near *TET3*.

K562 Hi-C heatmap showing 5.3 megabases (chr2:70,000,000-75,300,000) containing *TET3* and the closest HIF1 binding sites (blue arrows) . Grey circles indicate parts of the heatmap that quantify *TET3*:HIF1 binding site interactions. Hi-C data from Rao et al., 2014; Visualization by Juicebox (Robinson et al., 2018)



Supplemental Figure 7: CRISPR deletion of HIF-1α binding site 1 and/or site 2.

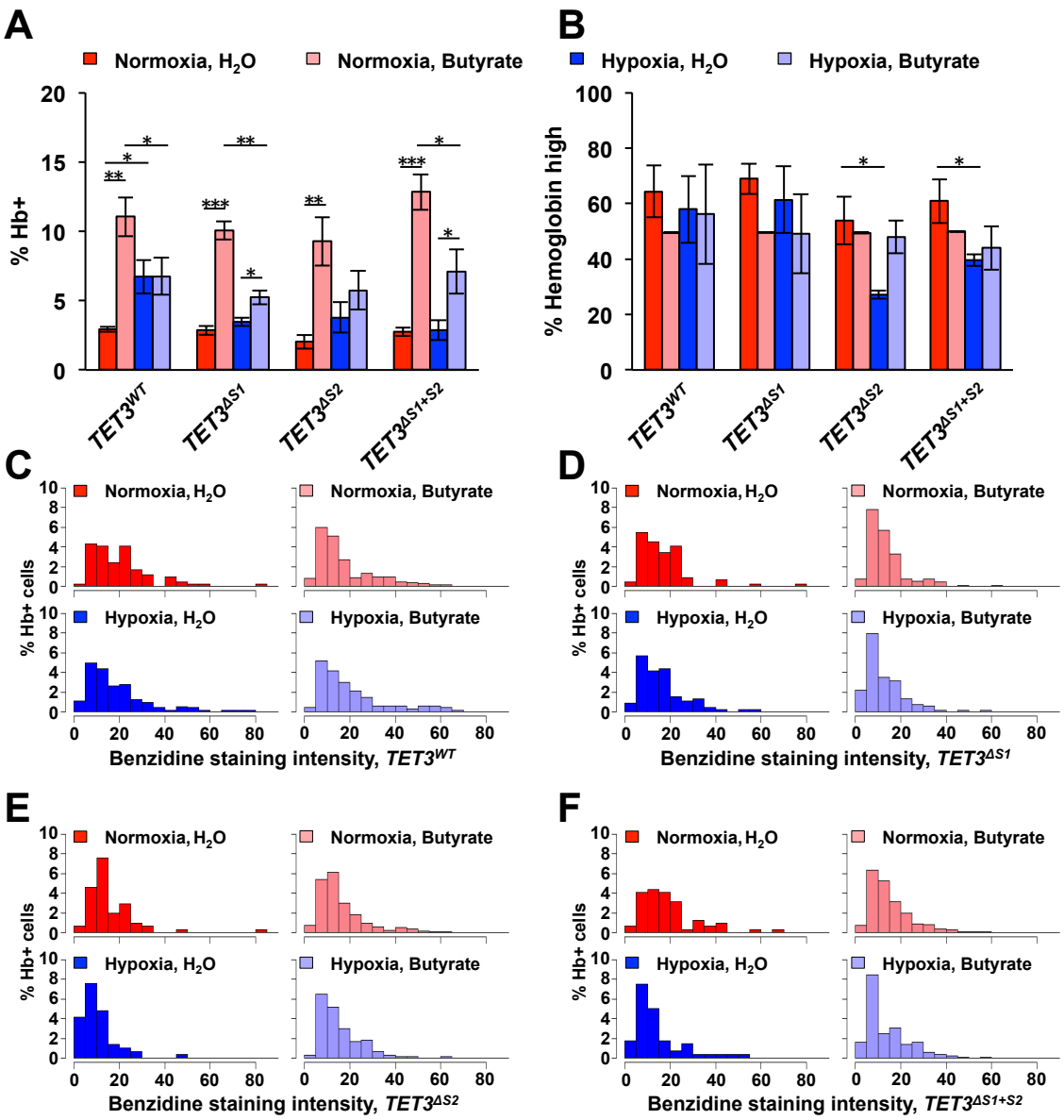
(A) *ΔS1* clones deletions.

(B) *ΔS2* clones deletions.

(C) *ΔS1+S2* clones (made from S2-5 clone) deletions. Core ACGT motif is shown in red letters. Total numbers of bases deleted are indicated if deletion extends beyond sequences shown.

(D) *TET3* expression of each individual clones in normoxia vs. hypoxia.

(E) Comparison of the expression of *TET* genes in K562 cells vs human bone marrow.

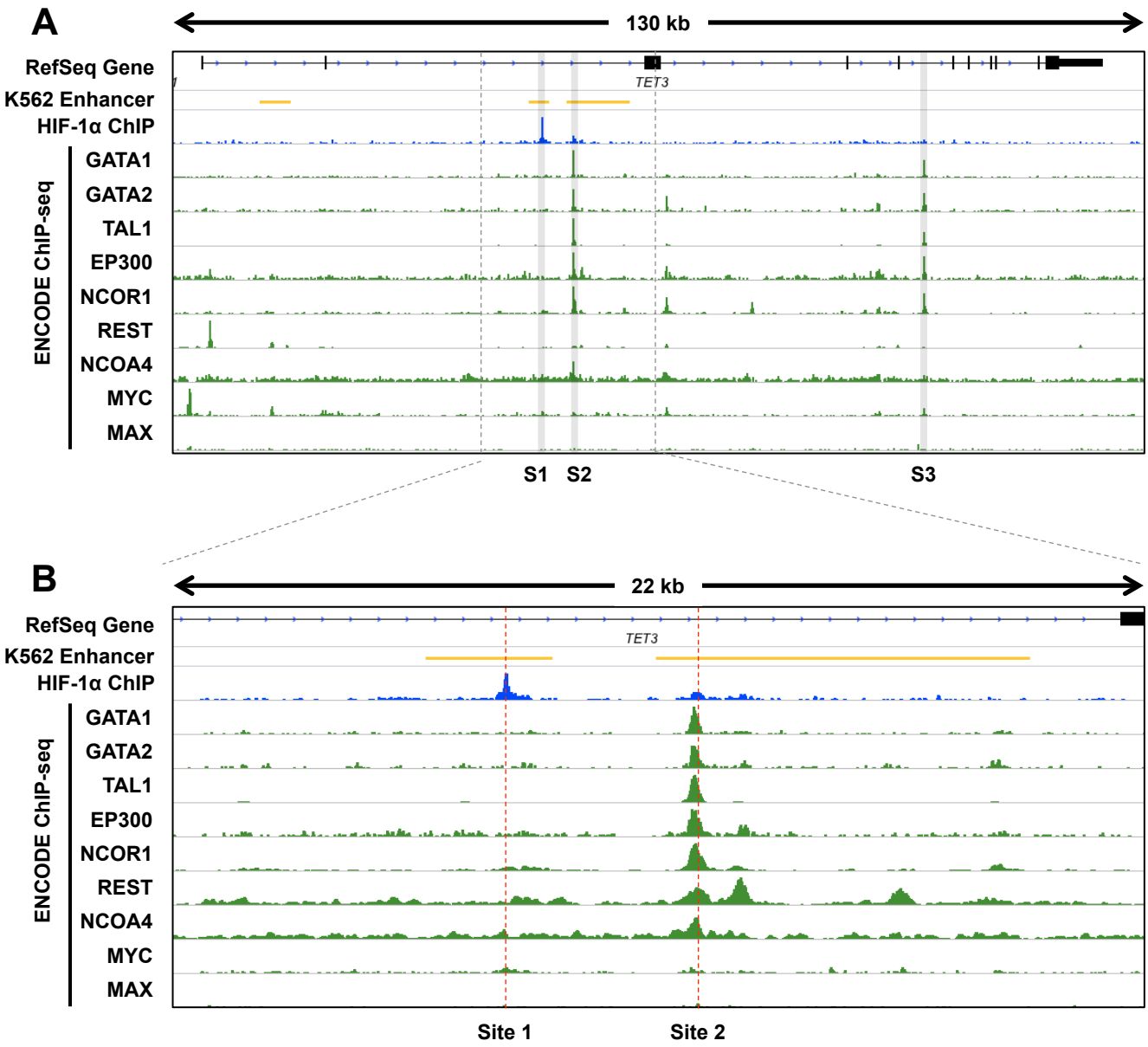


Supplemental Figure 8: Hemoglobin accumulation in Hb+ cells by benzidine staining intensity.

(A) Percent of cells producing hemoglobin in all cells.

(B) Percent of cells with high hemoglobin accumulation in all hemoglobin producing cells. Normoxic butyrate treated cells were used to establish the range of range of benziding staining in each replicate. High hemoglobin cells were defined as having higher benziding staining intensity than 50% of Hb+ cells in the normoxic butyrate treated cells.

(C-F) Benzidine staining intensity in *TET3*^{WT} (C), *TET3*$^{\Delta S1}$ (D), *TET3*$^{\Delta S2}$ (E), and *TET3*$^{\Delta S1+S2}$ (F) cells under normoxia or hypoxia, treated with water or 1 mM sodium butyrate.



Supplemental Figure 9: Binding of other transcription factors at Sites 1-3

(A) ENCODE ChIP-seq data overlaid on HIF-1α ChIP-seq results at *TET3*. Site 1 is largely unique to HIF-1, whereas Site 2 has multiple transcription factors binding sites very close to HIF-1 binding site. Site 3 lacks HIF-1 binding, but is occupied by multiple other transcription factors.

(B) Zoomed in view of Sites 1 and 2. Dotted red lines indicate the positions of the ACGT motifs that were deleted by CRISPR. Note that the summits of GATA1/2, TAL1, and EP300 are slightly upstream of the HIF-1 binding site, which is likely due to the two GATA-boxes (GATAAG) being upstream (-60 bp and -105 bp) of the HIF-1 binding site.

ENCODE alignment file ID for each track:

GATA1- ENCFF942NUX; GATA2- ENCFF990JRT; TAL1- ENCFF389WLJ; EP300- ENCFF703ULA; NCOR1- ENCFF246WTN; REST- ENCFF686NWE; NCOA4- ENCFF662JKA; MYC- ENCFF058VAU; MAX- ENCFF000YTL.