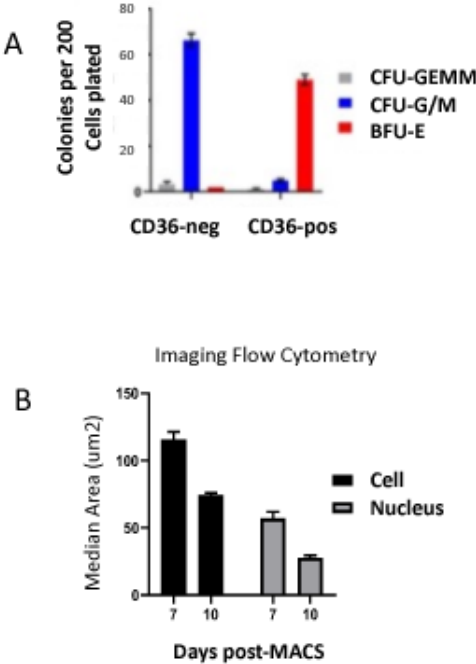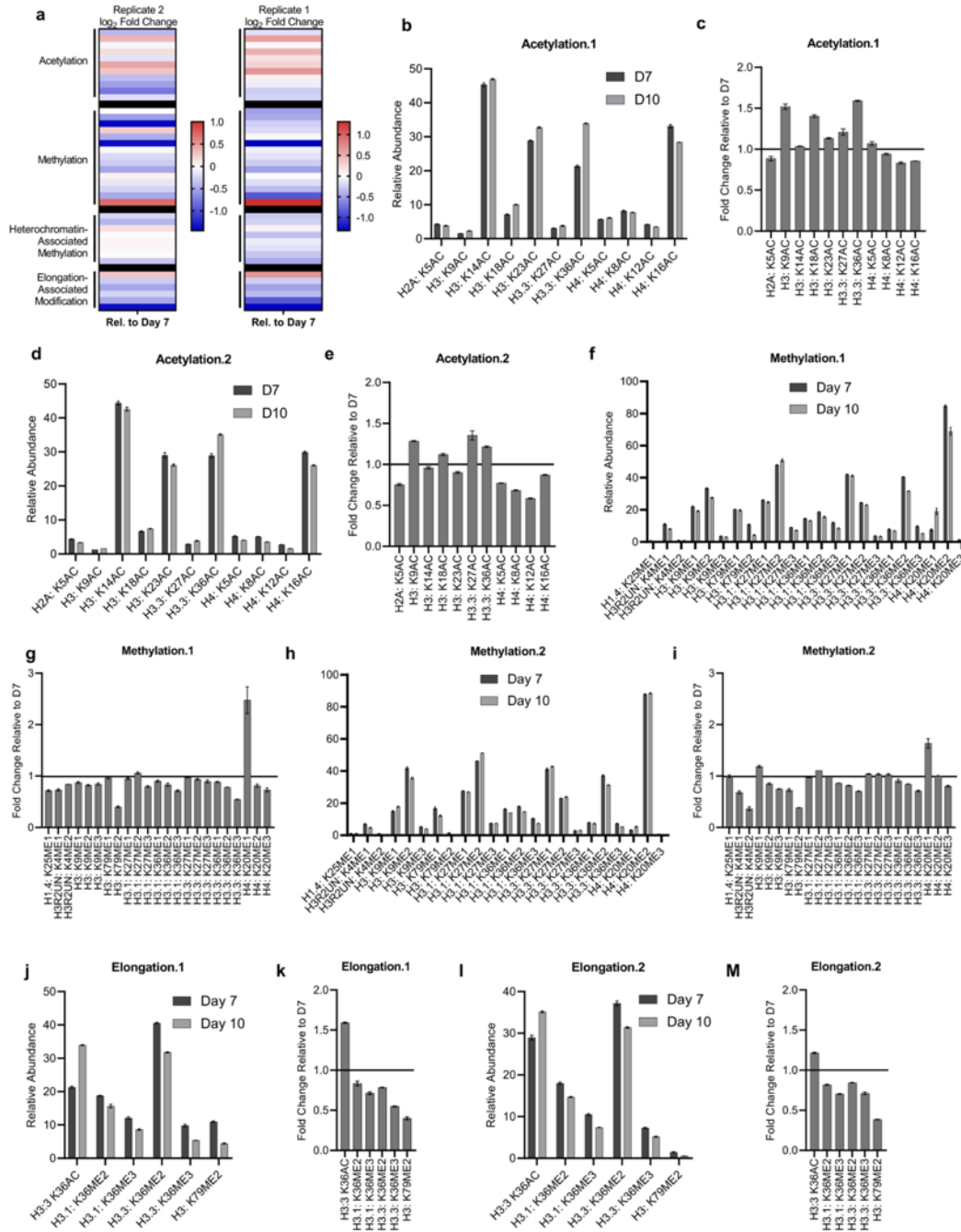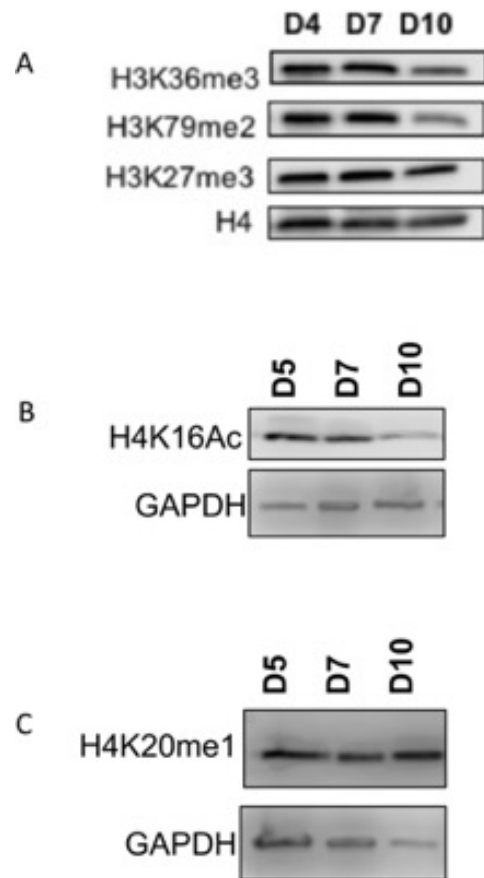**Supplemental Data**

Figure S1. Validation of CD36+ selected erythroid culture. (A) Colony forming ability of CD36+ selected cells. (B) Imaging flow cytometric analysis of cell and nuclear size on day 7 and day 10 of culture.

Figure S2. Replicates of ModSpec experiments. (A) Heat maps corresponding to the first and second biological replicate of ModSpec from CD36 synchronized cells on Day 7 and Day 10 of culture. Values represent $\log_2$ fold change of Day 10 compared to Day 7. Total abundance of acetylation marks, greater than 1%, and fold change at Day 10 relative to Day 7 are shown for the first (B and C) and second (D and E) replicates. Total abundance of methylation marks, greater than 1%, and fold change at Day 10 relative to Day 7 are shown for the first (F and G) and second (H and I) replicates. Total abundance of selected marks associated with transcription elongation, greater than 1%, and fold change at Day 10 relative to Day 7 are shown for the first (J and K) and second (L and M) replicate.

Figure S3. Abundance of selected histone PTMs as determined by western blot, during erythroid culture on indicated days.
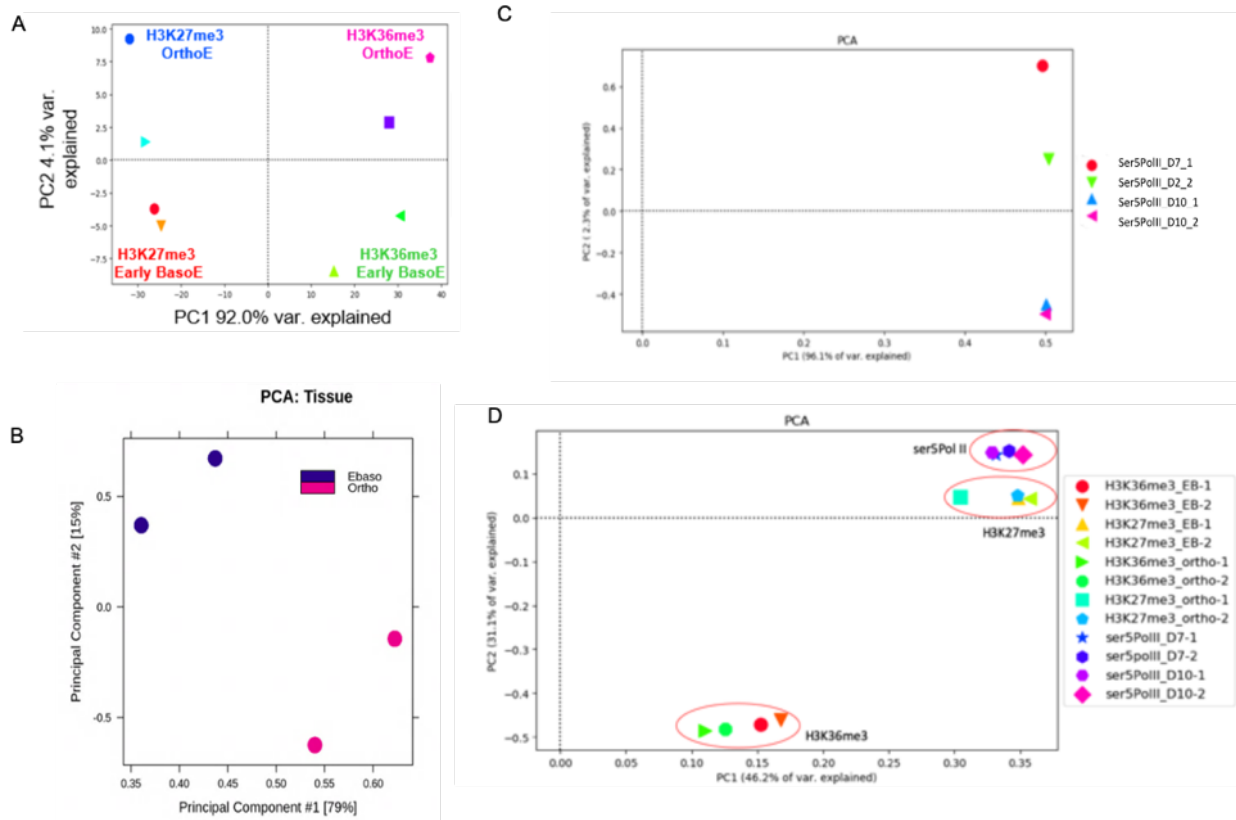
Figure S4. PCA analyses of H3K36me3-ChIP-seq, H3K27me3 ChIP-seq, ATAC-seq, and ser5 Pol II ChIP-seq studies. (A) PCA of H3K27me3 and H3K36me3 ChIP-seq experiments studies done in early basophilic and orthochromatic erythroblasts, sorted from CD34+ derived erythroid cultures. (B) PCA of ATAC-seq experiments, in early basophilic and orthochromatic erythroblasts, sorted from CD34+ derived erythroid cultures. (C) Ser5 Pol II ChIP-seq done using cells from the CD34+ culture system shown in 1A, on day 7 and day 10 of culture. (C) PCA analyses of all ChIP-seq experiments.
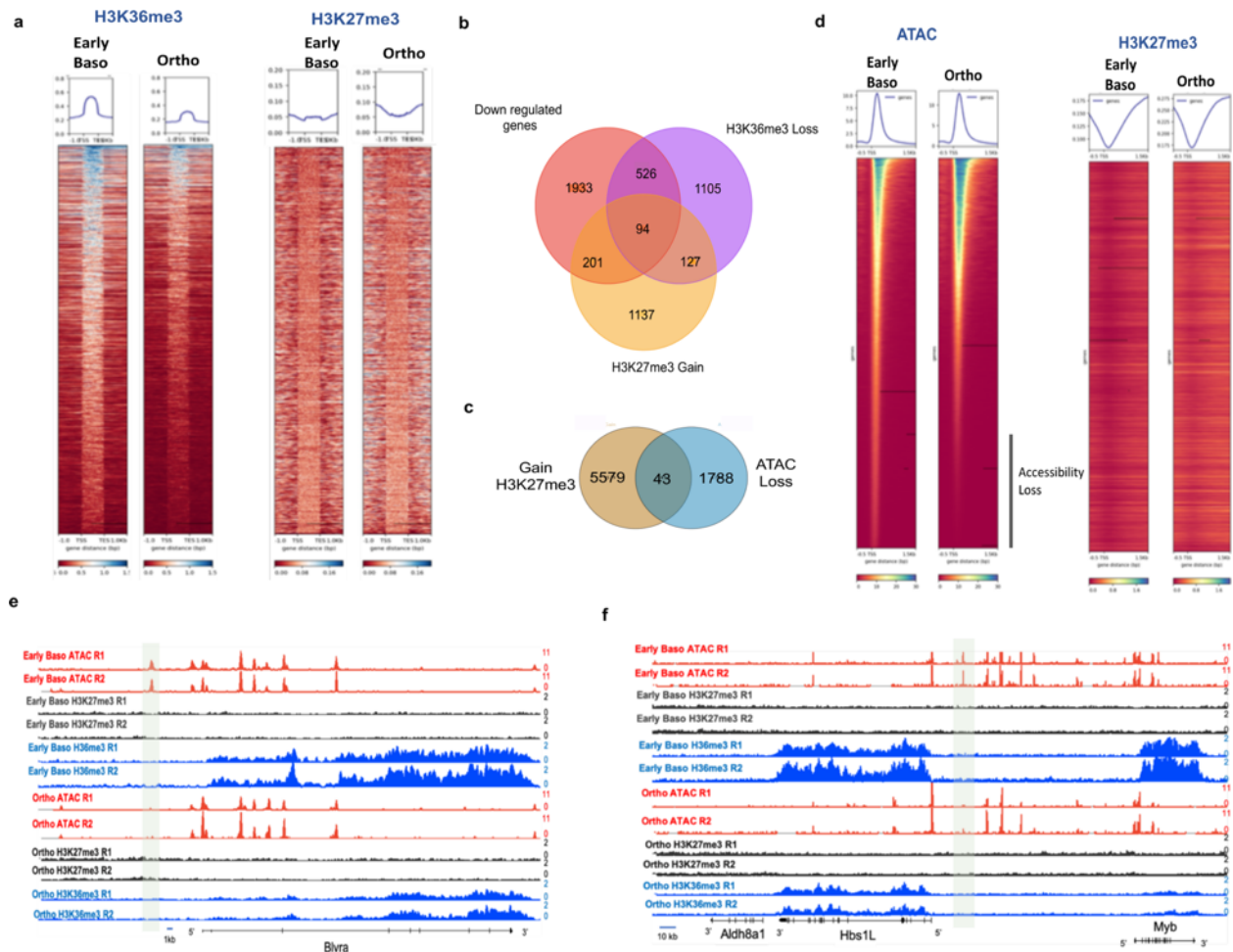
Figure S5. Decreased gene expression is correlated with loss of H3K36me3 without gain of H3K27me3. (A) Heat maps of H3K36me3 and H3K27me3 enrichment over 3,058 regions associated with genes whose expression decreases > 2-fold from intermediate (Early Basophilic erythroblast) to late (Orthochromatic) erythroblast. (B) Venn Diagram showing genes down regulated from basophilic erythroblast to orthochromatic erythroblast (red), genes that have loss of H3K36me3 (purple), and genes that have gains in H3K27me3(yellow). (C) Venn diagram showing intersection of H3K27me3 peaks and loss of ATAC peaks, as determined by MACS2. (D) Heat maps depicting chromatin accessibility in intermediate (early basophilic) and late (orthochromatic) erythroblasts (*left*), and corresponding H3K27me3 enrichment at those regions (*right*). (E-F) Changes in chromatin accessibility over the Biliverdin Reductase A (Blvra; E) and HSBl1-MYB (F) Loci. Green box highlights change in accessibility.
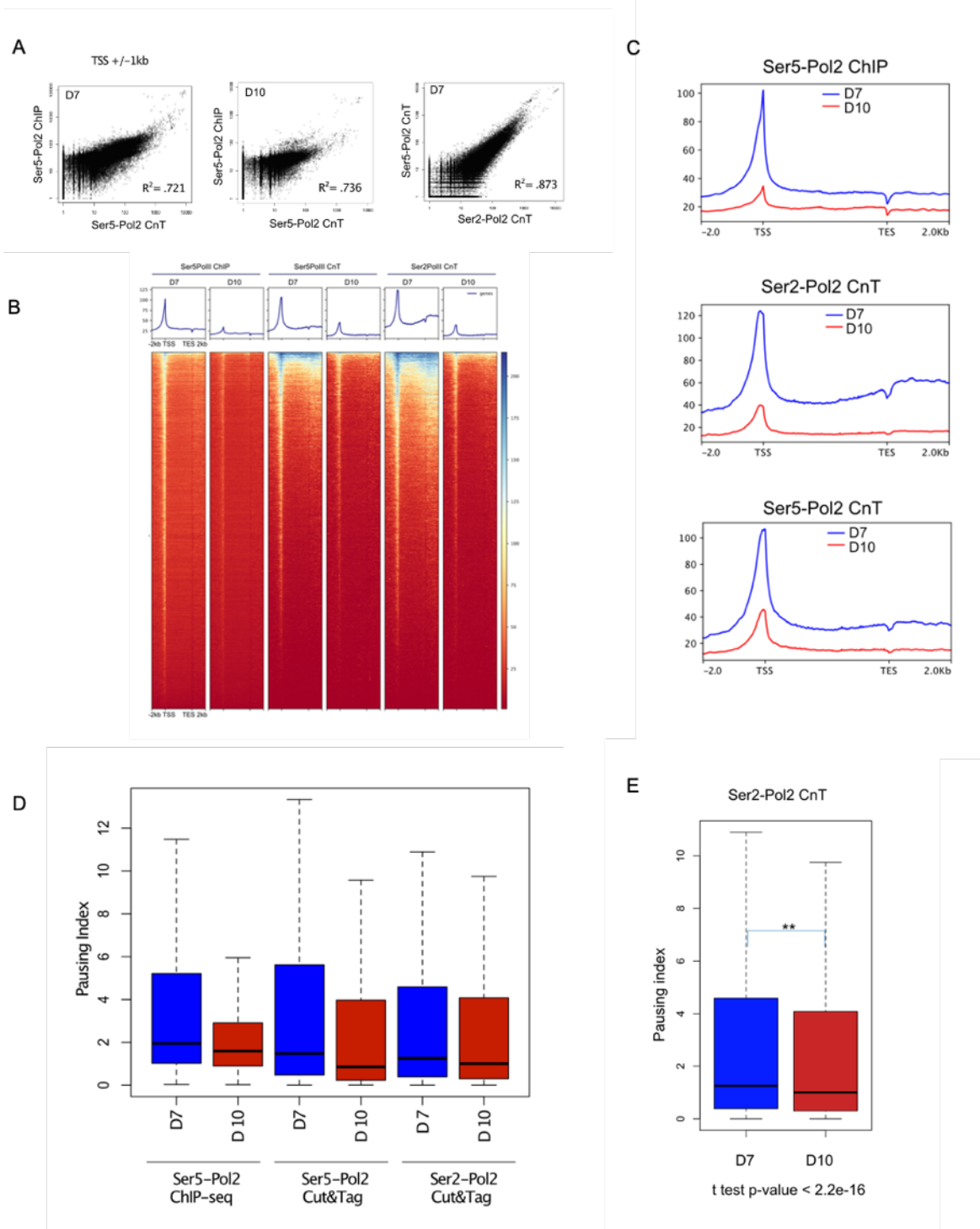
Figure S6. Genome wide profiling of RNA polymerase II (pol II) using ChIP-seq and cut and Tag. (A) Correlation plots of indicated ChIP-seq and cut and tag studies. (B) Heat maps of indicated ChIP seq and cut and tag studies. (C) Metagene plots of indicated ChIP-seq and cut and tag studies. (D-E) Pausing Index calculated in indicated populations. D7= Day 7 of erythroid culture system shown in Figure 1A. D10= Day 10 of erythroid culture system shown in Figure1A.
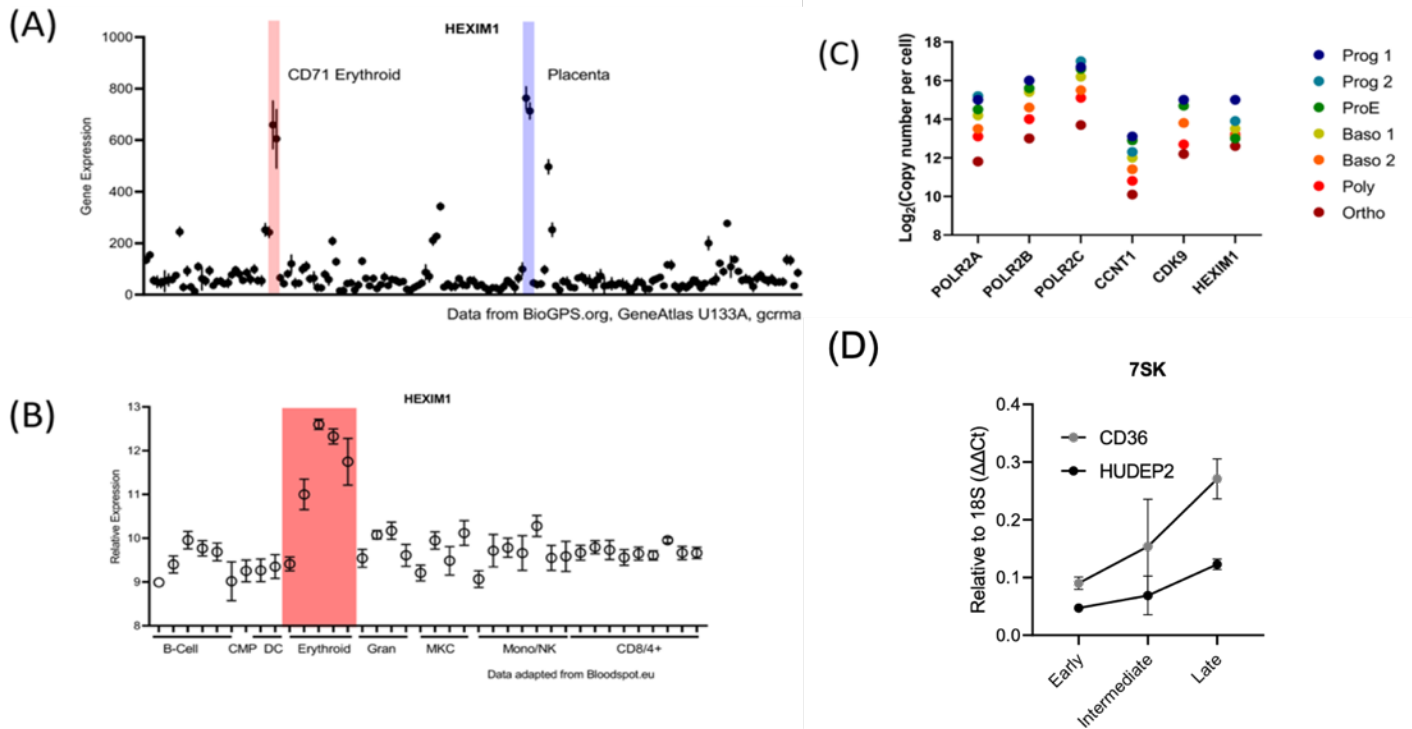
Figure S7. HEXIM1 expression in various tissues. (A) HEXIM1 mRNA expression in various tissues and cell lines from GeneAtlas (U133A; biogps.org[1]). (B) HEXIM1 mRNA expression in various hematopoietic cell types from GSE24759. (C) Protein expression of HEXIM1 and other important components of RNA polymerase II pausing regulation during terminal erythroid maturation, from [2] (D) 7SK levels during the erythroid maturation of CD36+ selected HSPCs and HUDEP2 cells.
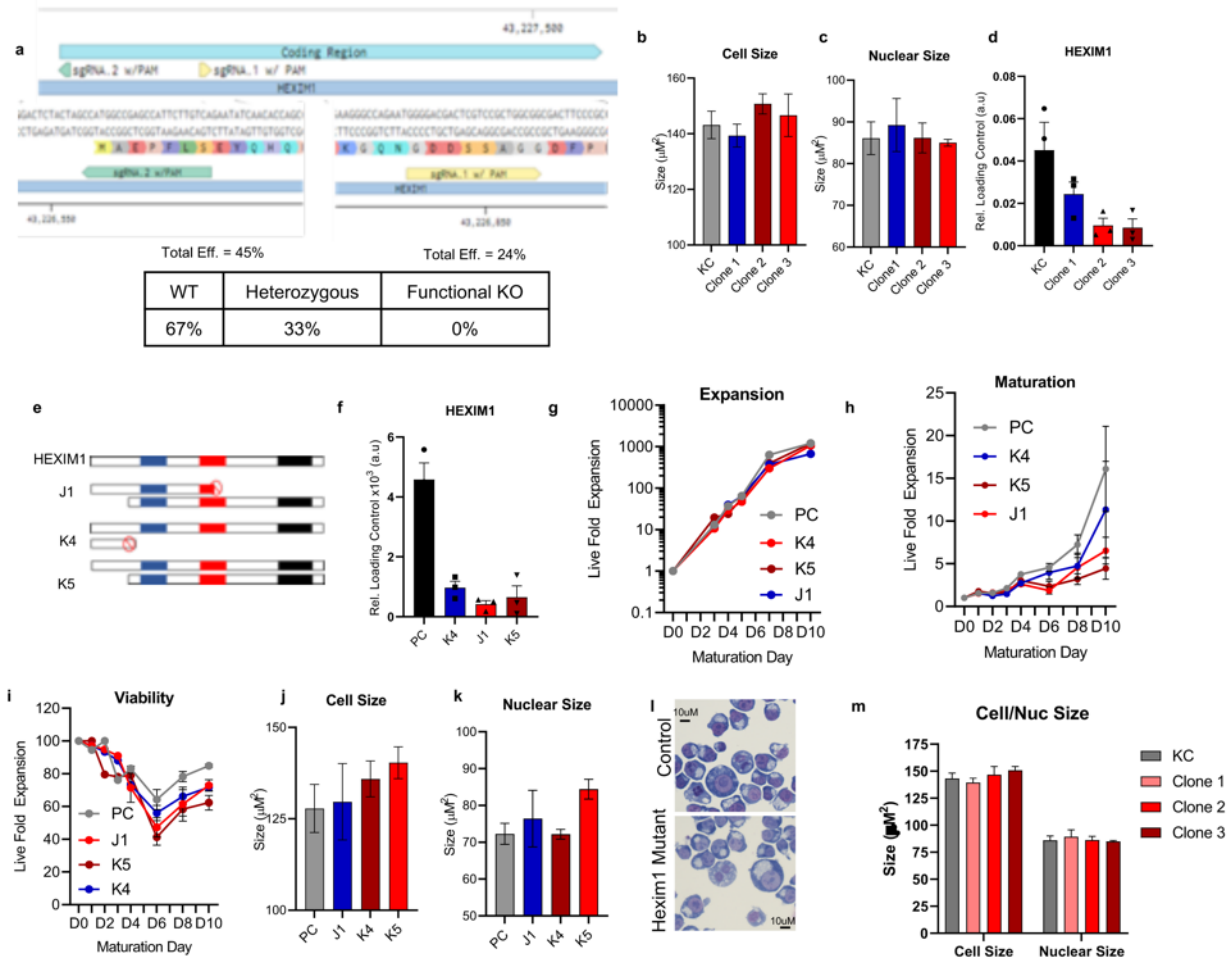
Figure S8. HEXIM1 knockdown in HUDEP2 and KITCAT clones alters erythroid expansion and maturation. (A) Schematic representation of the two sgRNAs used and the targeting location within the coding region of HEXIM1. The table represents the percent of clonal population cells screened that resulted in WT clonal lines, homozygous edits, or function knockout or null lines. (B) Quantification of the cell size of KITCAT clones in expansion media by imaging flow cytometric analyses. (C) Quantification of nuclear size of KITCAT clones in expansion media. (D) Protein expression of HEXIM1 in three clonal KITCAT lines compared to control. Values represent normalization to a loading control. (E) Summary of genome edits in HUDEP2 clonal lines with HEXIM1 mutations. (F) Protein abundance in HUDEP2 clonal lines, quantified as in (D). Of note, genome edits may hinder the ability of the HEXIM1 antibody to bind the HEXIM1 protein. (G) Proliferation of clonal lines with HEXIM1 disruption in expansion conditions. (H) Proliferation of clonal lines with HEXIM1 disruption during terminal maturation (I) Viability of clonal lines with HEXIM1 disruption in maturation conditions. (J) Quantification of cell size of HEXIM1 disrupted HUDEP2 clonal lines via imaging flow cytometric analyses. (K) Quantification of nuclear size of HEXIM1 disrupted HUDEP2 clonal lines via imaging flow cytometric analyses. (L) Representative control and HEXIM1 mutant cells at Day 4 of maturation shown. Day 4 was selected as it is prior to significant loss of cell viability in culture. (M) Cell size and nuclear size were quantified by imaging flow cytometry and showed no significant change in morphologic maturation in HEXIM1 disruption.

(E) Downregulated genes Hexim1 OE

Pathway	Adj P value	Database
Phosphatidylinositol signaling	0.00008	BioPlanet 2019
Sulfide Oxidation	0.001	BioPlanet 2019
Fc gamma receptor mediated phagocytosis	0.001	BioPlanet 2019
Platelet activation	0.001	BioPlanet 2019
Mitochondrial role in apoptotic signaling	0.001	BioPlanet 2019
Apoptosis	0.005	MSigDB Hallmark 2020
Estrogen Response	0.005	MSigDB Hallmark 2020
P53 pathway	0.001	MSigDB Hallmark 2020

(F) Upregulated genes Hexim1 +/-

Pathway	Adj P value	Database
Estrogen Signaling	0.00004	KEGG 2019
GnRH Signaling	0.0009	KEGG 2019
Inflammatory mediator of TRP channels	0.001	KEGG2019
Calcium channel inhibitor activity	0.04	GO Molecular Function 2018
Actin cross link formation	0.001	GO Biologic Process 2018
Microtubule Cytoskeleton	0.001	GO Cellular Component 2018
Spindle pole	0.01	GO Cellular Component

Figure S9. Perturbation of HEXIM1 Expression and impacts on gene expression. (A) Protein expression of HEXIM1 in three clonal KITCAT lines compared to control, as shown in Figure 5A, with the inclusion of HEXIM1 OE. (B) Protein expression in HEXIM1 OE and YYFF lines, with Ser2 Pol II as shown in figure 7.(C) Quantification of HEXIM1 levels in HEXIM1 OE and YYFF lines relative to GAPDH control. (D) qPCR of indicated genes in HEXIM1 OE cells. (E) Pathways enriched in genes downregulated in HEXIM1 OE compared to EV control. (F) Pathways enriched in upregulated genes in HEXIM1 +/- compared to EV control.
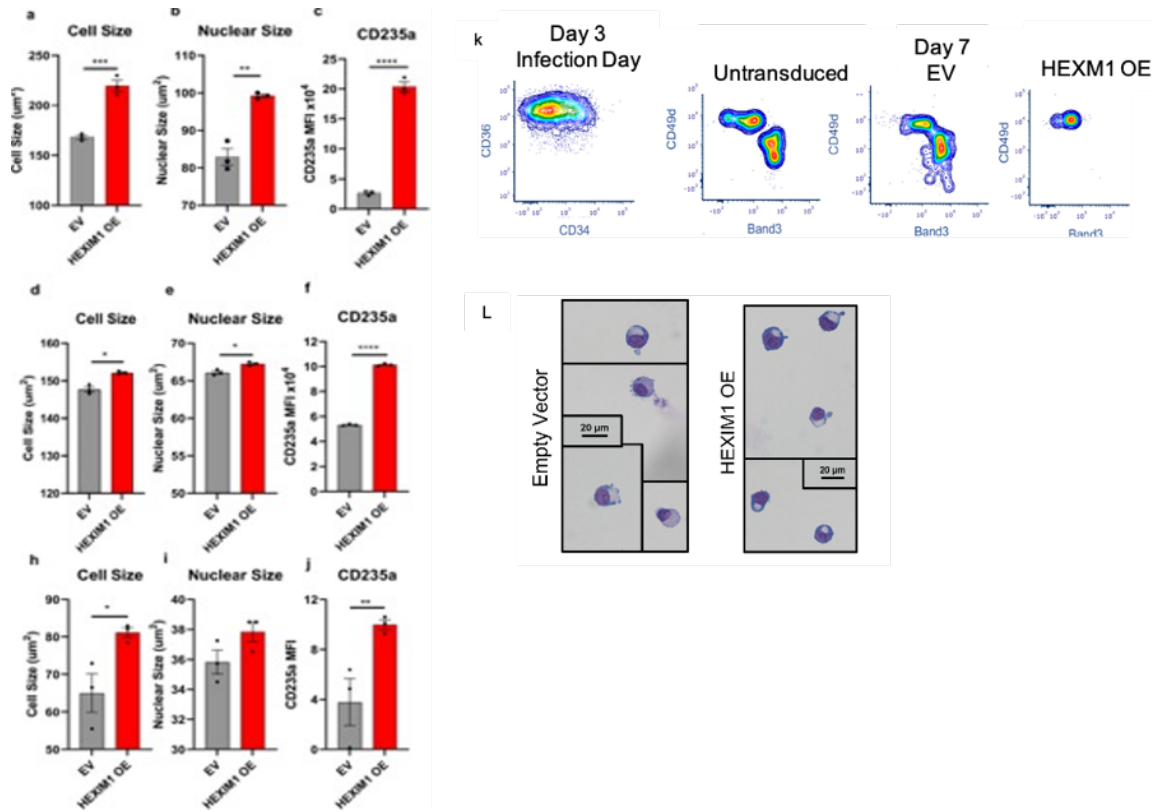
Figure S10. Hexim1 overexpression disrupts terminal erythroid maturation. (A-J) Cell size, nuclear size and CD235a cell surface intensity are quantified using imaging flow cytometric analyses at each stage of maturation, (Day 4 (A-C), Day 7 (D-F), and Day 10 (H-J)) following HEXIM1 OE in HUDEP2 cells. (K) Flow cytometric analyses of Band 3 and CD44d expression following HEXIM1 OE in the CD34+ erythroid culture system shown in 1A. Cells were transduced on Day 3 following CD36 selection, and flow cytometric analyses was conducted on day 7. (L) Cytospins of Hexim1 overexpression and control cells on Day 7 after CD36 selection.

.

Histone Mark	Relative to Day 7
H2A: K5AC	0.81887488
H3: K9AC	1.41134533
H3: K14AC	0.99809536
H3: K18AC	1.26769577
H3: K23AC	1.01766154
H3.3: K27AC	1.27803015
H3.3: K36AC	1.37472482
H4: K5AC	0.92864487
H4: K8AC	0.84285474
H4: K12AC	0.73516993
H4: K16AC	0.86350677
H1.4: K25ME1	0.87396678
H3R2UN: K4ME1	0.71199262
H3R2UN: K4ME2	0.6152212
H3: K9ME1	1.00186305
H3: K79ME1	0.85976045
H3: K79ME2	0.39860199
H3.1: K27ME1	0.96364685
H3.1: K36ME1	0.88312415
H3.1: K36ME2	0.8258482
H3.1: K36ME3	0.70954113
H3.3: K27ME1	1.00999653
H3.3: K36ME1	0.89770853
H3.3: K36ME2	0.81277679
H3.3: K36ME3	0.61849163
H4: K20ME1	2.22479993
H3: K9ME2	0.84014352
H3: K9ME3	0.78972069
H3.1: K27ME2	1.0835637
H3.1: K27ME3	0.88377148
H3.3: K27ME2	0.98702548
H3.3: K27ME3	0.95323781
H4: K20ME2	0.91161782
H4: K20ME3	0.75959821
H3:3 K36AC	1.37472482
H3.1: K36ME2	0.8258482
H3.1: K36ME3	0.70954113
H3.3: K36ME2	0.81277679
H3.3: K36ME3	0.61849163
H3: K79ME2	0.39860199

Table S1. Fold change of histone marks during terminal erythroid maturation.

ChIP Primers		
	Forward	Reverse
GYPA Promoter	TCACTGCAAGGAACAGGTTG	GGCTCCACAACAGCTACCTC
GYPA Gene Body	CTGCATATGTGTCCCGTTTG	GGAGGGATGTGGGAGAGTTT
RPS19 Promoter	ACACTCCGGGAGAAGGAAAC	GGTGTCTAGTGAGGGGTGGA
RPS19 Gene Body	CTATGGGGGACGTCAGAGAA	ACCCATCTTGGTCCTTTTCC

Table S2. Primers used in ChIP studies

Supplemental Methods

Cell Culture

HUDEP2 and KITCAT Culture

HUDEP2 cells[3] were obtained from RIKEN BioResource Center. KITCAT lines were derived from HUDEP2 cells as previously described[4] . Expansion and maturation conditions were previously described. [4,5]

Primary Erythroid Culture

CD34 HSPCs were provided by the Yale Cooperative Center of Excellence in Hematology, and subjected to erythroid culture following CD36 selection as outlined. [2] CD34 cells for analyses of HEXIM1 OE and KD were partially expanded for in H3000 supplemented with cc100 (Stem Cell Technologies) for 3 days to increase yield, prior to entering the CD36 synchronization protocol. Cell morphology and immunophenotype were determined to be consistent with CD34 cells without initial expansion (data not shown). For the H3K36me3 and H3K27me3 ChIP-seq, early basophilic erythroblasts and orthochromatic erythroblasts were sorted from CD34+-derived erythroid cultures as described. [6]

Mass Spectrometry Studies (Mod Spec, Performed by Active Motif)

Cell Preparation for Mass Spectrometry Experiments

Mass spectrometry was performed on cells from Day 7 and Day 10 of the erythroid maturation of the CD36+ selected cells, as shown in figure 1A and described in Gautier Cell Reports 2016.[2] 2-3 million cells from Day 7 and Day 10 from two biologically distinct cultures were submitted for assessment of the abundance of histone modifications via mass spectrometry using the Mod Spec Service (Active Motif, https://www.activemotif.com/catalog/1235/mod-spec).

Histone Extraction and Mass Spectrometry

Bulk histones were extracted using acid extraction, propionylated and subjected to trypsin digestion, as described.[7] Briefly, histones were extracted by incubating samples in 0.2M sulfuric

acid with intermittent vortexing, at room temperature for 1 hour. Histones were then precipitated on ice by the addition of trichloroacetic acid (TCA), and recovered by centrifugation at 10,000 x g for 5 minutes at 4°C. The pellet was then washed once with 1mL cold acetone/0.1% HCl, followed by two washed with 100% acetone, and then air dried in a clean hood. The histones were propionylated by adding 1:3 v/v propionic anhydride/2-propanol. The pH was maintained at ~8 by incrementally adding ammonium hydroxide, and subsequently dried in a SpeedVac concentrator. The sample was resuspended in 100mM ammonium bicarbonate, with the pH adjusted to 7-8 with ammonium hydroxide, prior to trypsin digestion and drying in a SpeedVac concentrator. The pellet was resuspended in 100mM ammonium bicarbonate and propionylated a second time by adding 1:3 v/v propionic anhydride/2-propanol with the pH maintained at ~8 with the addition of ammonium hydroxide. Following drying in a SpeedVac concentrator, the histones were resuspended in 0.1% TFA in H_2O for mass spectrometry analysis. Samples were analyzed on a triple quadrupole (QqQ) mass spectrometer (Thermo Fisher Scientific TSQ Quantiva) directly coupled with an UltiMate 3000 Dionex nano-liquid chromatography system. Three technical replicates were analyzed by mass spectrometry for each sample.

Mass Spectrometry Data Analyses

Raw mas spec files were analyzed with Skyline,[8] using Savitzky-Golay smoothing. Skyline peak area assignments for all monitored peptide transitions were manually confirmed before proceeding with analysis, with multiple peptide transitions quantified for each amino acid. The sum of peak areas of corresponding peptide transitions were first quantified for all monitored amino acid residues, including both modified and unmodified forms. The sum of all modified forms was then calculated for each amino acid to represent the total pool of modifications for that residue. Finally, each modification is then represented as a percentage of the total pool of modifications. This process was carried out on three separate mass spec runs for each biologic replicate. The data was returned as relative abundance of histone modifications with the standard deviation of the technical replicates. All modifications with relative abundance above 1% are reported here.

Cytopsins

Cytospins of 2 x 10^3 cells were performed following a wash in PBS followed by a 10-minute incubation in ice in PBS supplemented with 0.01% glucose and 0.03% BSA. Cells placed on slide using Cytospin (Shandon Cytospin 2) at 300 RPM for 2 minutes. Cells air dried for 15 minutes and then fixed in 100% methanol for 5 minutes. Cells stained in 1:20 Giemsa (Millipore-Sigma, GS1L) for 30 minutes and then washed in water. Images taken at 200x magnification on a Nikon DS-Fi1 camera using NIS elements software (Nikon).

Functional Studies (Hexim1 genome editing, knockdown, and overexpression)

HEXIM1 variant clonal lines were created by selecting two sgRNAs templates from CRISPRscan.org (PMID:26322839) and sgRNA was synthesized using protocol derived from[9] . Ribonucleoproteins and sgRNA were introduced into KITCAT and HUDEP2 cells as previously described.[4] Cells were recovered in expansion medium for 3 days prior to generation of clonal lines through single cell dilution. Clonal lines were screened for deletions in HEXIM1 following allelic separation via TOPO cloning (TA cloning Kit, Invitrogen). DNA isolated from individual bacteria colonies following cloning (NEB, Miniprep Kit) was Sanger sequenced to determine specific genome edits in clonal population (Genewiz). 30 total clonal populations were screened with no total HEXIM1 knockouts found.

cDNA for WT HEXIM1 or HEXIM1 harboring the YYFF substitution was cloned into pReceiver-Lv165 overexpression constructs (GeneCopoeia, provided by Palis Laboratory, University of Rochester). The virus was derived and transfected into polyclonal HUDEP2 cells as previously described [10]. Clonal populations of HUDEP2 cells were obtained using single cell dilution.

Cell numbers during expansion and maturation recorded using daily counts and viability tracking, following trypan staining, and presented as total live amplification accounting for cell dilution. Doubling time determined over 10 days of consecutive counts using the formula below; where $N(t)$ is the number of cumulative cells at time t adjusting for dilution, T_d is the doubling period in hours, N_0 is the initial number of cells, and t is time.

$N(t) = N_0{}^{2t/Td}$

For functional studies in CD34 derived erythroid cultures, CD34+ cells were cultured as in [2], and transduced with lentivirus on day 3 following CD36 selection, with flow analyses of transduced cultures on day 7 following CD36 selection.

Colony assays were performed as previously described.[11]

Biochemical Analysis

Protein extracts from cultured cells collected in RIPA buffer (Cell Signaling Technology) and sonicated for 30 seconds on and 30 seconds off for a total of 10 cycles in a Biorupter (Diagenode). Lysates were then resolved, transferred and blotted. Antibodies were purchased from Cell Signaling Technologies (Rpb1 (S2) 13499, Rpb1 (S5) 13523, H3K79me2 5427, H3K36me3 4909, H3K27me3 9733, H4K16Ac 13534, Total H4 2935) and Millipore (H4K20me1 07-1570). Blots imaged using C-DiGit Blot Scanner (Li-Cor) after five-minute exposure to WesternSure PREMIUM Chemilunescent Substrate (Li-Cor). Digital quantification was done using D-DiGit Image Acquisition Software (Li-Cor) and all protein concentrations normalized to total H4 or GAPDH as a loading control.

Flow Cytometry

Cultured cells analyzed for immunophenotype following staining as previously described[2,4] with antibodies for CD36-FITC (clone CB38, BD Biosciences), CD235a-PECy7 (clone GA-R2, BD Biosciences), CD49d-PE (clone MZ18-24A9, Miltenyi Biotec), and BAND3 (Bric 200, IBGRL). BAND3 surface protein expression was analyzed using a secondary antibody (anti-IgG APC-Cy7, clone MOPC-21, BD Biosciences). Additionally, DAPI (ThermoFisher Scientific, D1306) and DRAQ5 (ThermoFisher Scientific, 65-0880-92) were added to identify live cells and nuclei, respectively. Cells were run either on an Image Stream X (Amnis/EMD Millipore) and analyzed with IDEAS 6.3 (Amnis/EMD Millipore); or an LSRII (BD Biosciences) and analyzed with FCS Express (DeNovo Software, v7). Cytoplasmic and nuclear area were determined as previously described.[4]

Chromatin Immunoprecipitation

For the ser2 and ser5 RNA polymerase II ChIP, standard ChIP protocol was followed with minor modifications.[10] Briefly, approximately 13 million cells from Day 7 and Day 10 of CD36 synchronization protocol were crosslinked with 1% formaldehyde and lysed to isolate chromatin. Sonication was performed (Diagenode Bioruptor) on high setting for 35 cycles of 30 seconds on and 30 seconds off. The samples were diluted, and pre-cleared overnight with unmodified beads (Invitrogen). The immunoprecipitation was performed with preconjugated AG Magnetic beads

(Invitrogen 10004D) to phospho RPB1 CTD ser5 (clone D9N5, Cell Signaling Technologies) overnight at 4°C. The beads-antibody-chromatin complex was subjected to a series of salt/TE washes and eluted in a fresh RT elution buffer. The samples were treated w/ RNase A, reverse crosslinked with proteinase K and cleaned up with a PCR purification kit (New England Biolabs). For ChIP qPCR, 50ng of ChIP DNA was amplified with the Whole Genome Amplification Kit (WGA2 Sigma). For the RNA polymerase II ChIP-seq, the ChIP DNA was library prepared using the Illumina TrueSeq Library Prep kit. The indexed ChIP DNA was then submitted to Genewiz. For ChIP qPCR, primer sets used are shown in table S2.

For the H3K36me3 and H3K27me3 ChIP-seq, early basophilic erythroblasts and orthochromatic erythroblasts were sorted from CD34+-derived erythroid cultures as previously described, [6] using antibodies against H3K36me3 (Active Motif, 61101) and H3K27me3 (Abcam ab6002) in early basophilic and orthochromatic erythroblasts. ChIP was performed following the ChIPmentation protocol as outlined [ChIPmentation CeMMv1.14 (September 2016)]. Following the ChIP protocol, libraries were prepared and sequenced at the Yale Center for Genome Analyses.

Cut&Tag

CUT&Tag was performed in duplicate for each antibody using approximately 100,000 cells following the Bench top V.1 protocol (Kaya-Okur et al, 2019[12] and https://www.protocols.io/view/bench-top-cut-amp-tag-bcuhiwt6?version_warning=no). Briefly, cells were incubated with Concanavalin A beads, and then primary antibodies overnight at 4 degrees, secondary antibodies for 1h at RT, and pA-Tn5 adapter complex for 1h at RT. Tagmentation was induced for 1h, and DNA was extracted with phenol-chloroform-isomyl alcohol. Libraries were amplified using NEBNext HiFi 2X Master mix, pooled, and cleaned up using 1.3X vol. SPRI beads. Antibodies for cut&tag included Ser2-Pol2: Cell Signaling 13499s, Ser5-Pol2: Cell Signaling 13523s.

Statistics and Reproducibility

Multiple HEXIM1 KITCAT clones were derived and analyzed separately to account for clonal differences between cells lines. All cell culture experiments were performed and analyzed at separate times to account for medium and temporal variation. This included Mod Spec analysis which was done from two separate CD34 donor sources.

Comparisons were performed using GraphPad Prism 8 with the specific number of replicates shown in figures and in figure legends. All data obtained is shown, excluding several western blots in which a representative image is shown. One-way ANOVA, two-way ANOVA, or student's t-test were used when appropriate. Significance thresholds are referenced in figure legends.

Bioinformatic Analyses

ChIP-Seq

Single end raw reads were processed for quality and residual adapter sequencing using fastp v0.19.6 (--trim_poly_g -x --cut_window_size 4 -3 -l 35). Quality reads were aligned to the hg19 reference genome using bowtie2[13] and filtered for quality based on MAPQ > 10. Significantly enriched histone modifications were identified using MACS2[14] (--broad --broad-cutoff 0.1 -f BAM -B --SPMR -g hs) with each biological replicate and the corresponding total input control. In addition, replicated regions were also identified by providing MACS2 with all biological replicates for each mark to generate continuous coverage and enrichments for each condition (used to generate figures). For the RNA polymerase II ChIP-seq, alignment and peak calling was done using fastp and MACS 2 on the basepairtech.com platform. Genome wide distribution of

significantly enriched regions was evaluated using CEAS[15] (v1.0.2). Pathway analyses were completed using GREAT [16](version 4.0.4). Heat maps were generated using DeepTools.[17]

Cut&Tag

Fastq files were aligned to hg19 using Bowtie2, and PCR duplicates were removed using Picard. Read count normalization was performed on alignment files. Replicate merged bigwig files were generated using deepTools bigWigMerge with adjust=1.0. To correlate CUT&Tag and ChIP-seq datasets, occupancy for Ser5-Pol2 ChIP-seq, Ser5-Pol2 CUT&Tag, and Ser2-Pol2 Cut&Tag was calculated at promoters for all hg19 Refseq gene regions. Bed files with TSS +/- 1kb and CUT&Tag or ChIP-seq merged bigwig files were used with the multiBigwigSummary BED-file function of deepTools and parameter outRawCounts. Dot plots and pearsons correlation values were generalted in R.

RNA-Seq

Raw reads were downloaded from the sequence read archive (SRA) using fastqDump and the following accessions (SRR1106084, SRR1106085, SRR1106086, SRR1106087, SRR1106088 SRR1106089, SRR1106090, SRR1106091, SRR1106092, SRR1106093, SRR1106094, SRR1106095, SRR1106096, SRR1106097, SRR1106098) and low quality/adparter sequences were removed using Trimmomatic (TRAILING:13, LEADING:13, ILLUMINACLIP:${SCRIPT_DIR}/adapters.fa:2:30:10, SLIDINGWINDOW:4:20, MINLEN:35). Quality reads were aligned to the hg19 reference genome using STAR (readFilesCommand zcat, runThreadN ${CPUS}, runMode alignReads , genomeDir hg19, outSAMtype BAM Unsorted, outSAMstrandField intronMotif, outFilterIntronMotifs RemoveNoncanonical, outReadsUnmapped Fastx). Reads aligned within gene regions were quantified using the featureCounts function of the subread package. Differential expression was conducted using DESeq2[18] and significantly differentially expressed genes were defined as any gene with ab absolute log2FoldChange > 0 and adjusted p-value < 0.05.

Differential Binding

Differentially bound histone marks throughout erythroid maturation (Basophil vs. Orthochromatic) were identified using a default methods outlined by DiffBind[19] (v2.12.0) using the R framework (v3.6.0). Significantly differentially (FDR < 0.05 and Fold > 0) bound regions were annotated with the closest hg19 TSS using homer (v4.9.1). Genes associated with differentially bound regions were compared to genes differentially expressed throughout maturation (abs(log2FoldChange) >0 and adj p-value < 0.05).

Pausing Index Calculations and Heat Maps

For pausing index, occupancy for Ser5-Pol2 ChIP-seq, Ser5-Pol2 CUT&Tag, and Ser2-Pol2 Cut&Tag was calculated from TSS–30bp to TES for all hg19 Refseq gene regions. Bed files with chromStart=TSS+300 and chromEnd=TES, merged bigwig files for CUT&Tag and ChIP-seq, and RNA-seq bigwig files were used with the computeMatrix function of deepTools and the following parameters: scale-regions, regionBodyLength=300, before=330, bin size=30, outFileNameMatrix, and outFileSortedRegions. Subsequent analyses took place using R. Average Ser5-Pol2 or Ser2-Pol2 occupancy was calculated for TSS regions (TSS-30bp to TSS+300bp) and gene body regions (TSS+300bp to TES), and average RNA-seq expression was calculated from TSS to TES. Pausing index values were calculated as a ratio of TSS region to gene body region occupancy (pseudocounts=0.1). Subsets of gene regions with PI>4 were determined and plotted as a percentage of total. To compare pausing index values with gene

expression, Log2FC RNA was calculated comparing Day 10 to Day 7 (pseudocounts=0.1). Genes were defined as upregulated (Log2FC>0) or downregulated (Log2FC<0). For boxplots, statistical significance was determined using 2-sided Students t.test. Dot plots were generated for TSS change (defined as TSS region occupancy Day 10/Day 7) and gene body change (defined as gene body region occupancy Day 10/Day 7). For pausing index GO analysis, genes were filtered for minimum average expression on Day 7 (RNA>0.001). For downregulated pausing index>4 GO analysis, the subset of downregulated genes (Log2FC<0) was determined and analyzed with GREAT[16] to identify Gene Ontology terms.

Heatmaps and metagene plots were generated using the computeMatrix scale-regions function of deepTools.[17] Bed files with chromStart=TSS and chromEnd=TES, and merged bigwig files for CUT&Tag and ChIP-seq, were used the following parameters: beforeRegionStartLength 2000, regionBodyLength 4000, and afterRegionStartLength 2000. Matrix files were used with plotheatmap and plotProfile of deepTools.

Supplemental References

1.   Wu C, Jin X, Tsueng G, Afrasiabi C, Su AI. BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res.* 2016;44(D1):D313-316.
2.   Gautier EF, Ducamp S, Leduc M, et al. Comprehensive Proteomic Analysis of Human Erythropoiesis. *Cell Rep.* 2016;16(5):1470-1484.
3.   Kurita R, Suda N, Sudo K, et al. Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One.* 2013;8(3):e59890.
4.   Couch T, Murphy Z, Getman M, Kurita R, Nakamura Y, Steiner LA. Human erythroblasts with c-Kit activating mutations have reduced cell culture costs and remain capable of terminal maturation. *Exp Hematol.* 2019.
5.   Couch T, Murphy Z, Getman M, Kurita R, Nakamura Y, Steiner LA. Human erythroblasts with c-Kit activating mutations have reduced cell culture costs and remain capable of terminal maturation. *Exp Hematol.* 2019;74:19-24 e14.
6.   Schulz VP, Yan H, Lezon-Geyda K, et al. A Unique Epigenomic Landscape Defines Human Erythropoiesis. *Cell Rep.* 2019;28(11):2996-3009 e2997.
7.   Zheng Y, Thomas PM, Kelleher NL. Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nat Commun.* 2013;4:2203.
8.   MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics.* 2010;26(7):966-968.
9.   Gundry MC, Brunetti L, Lin A, et al. Highly Efficient Genome Editing of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. *Cell Rep.* 2016;17(5):1453-1461.
10.   Malik J, Getman M, Steiner LA. Histone methyltransferase Setd8 represses Gata2 expression and regulates erythroid maturation. *Mol Cell Biol.* 2015;35(12):2059-2072.
11.   Myers J CT, Murphy Z, Malik J, Getman M, Steiner LA. The histone methyltransferase Setd8 alters the chromatin landscape and regulates the expression of key transcription factors during erythroid development. . *Epigenetics and Chromatin.* 2020;in press.
12.   Kaya-Okur HS, Wu SJ, Codomo CA, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun.* 2019;10(1):1930.
13.   Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-359.
14.   Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.

15. Ji X, Li W, Song J, Wei L, Liu XS. CEAS: cis-regulatory element annotation system. *Nucleic Acids Res.* 2006;34(Web Server issue):W551-554.
16. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol.* 2010;28(5):495-501.
17. Ramirez F, Ryan DP, Gruning B, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 2016;44(W1):W160-165.
18. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
19. Stark R BG. DiffBind: differential binding analysis of ChIP-Seq peak data. 2011.