PNAS
www.pnas.org

Supplementary Information for

# Runx1 promotes murine erythroid progenitor proliferation and inhibits differentiation by preventing Pu.1 downregulation

**Authors:** Michael A Willcockson[1], Samuel J Taylor[1], Srikanta Ghosh[1], Sean E Healton[1], Justin C Wheat[1], Tommy J Wilson[2] and, Ulrich Steidl[1], Arthur I Skoultchi[1]

Dr. Arthur Skoultchi
Department of Cell Biology
Albert Einstein College of Medicine
1300 Morris Park Avenue, Bronx, NY 10461
Chanin 402
Telephone: (718) 430-2169
Fax: (718) 430-8574
arthur.skoultchi@einstein.yu.edu

Michael Willcockson
Department of Cell Biology
Albert Einstein College of Medicine
1300 Morris Park Avenue, Bronx, NY 10461
Chanin 402
Telephone: (718) 430-2169
Fax: (718) 430-8574
michael.willcockson@med.einstein.yu.edu

**This PDF file includes:**





www.pnas.org/cgi/doi/10.1073/pnas.1901122116

Supplementary Information Text
**Materials and Methods.**

### ATAC sequencing

ATAC seq was performed as in (1). Libraries were quantified using the Kappa Library Quantification kit and the average molecular sizes were determined by Bioanalyzer. ATAC-seq libraries were sequenced 2x50bp with an average coverage of 50 million reads per sample. Sequencing reads were aligned to the mouse genome (mm9) with Bowtie2(2). Picard was used to mark duplicates and Samtools was used to filter reads for low quality alignments, marked duplicates, and mitochondrial DNA(3). Homer was used to visualize chromatin accessibility, find peaks, calculate differential chromatin accessibility, calculate position-weighted matrices (*de novo* motifs) and match enriched DNA sequences to a database of known motifs (4). Motifs were discovered using Homer's default parameters that normalize motif frequency to a random set of GC-content normalized genomic sequences. Motifs were considered enriched if the background-subtracted frequency was at least 5% and the p-value was less than 1E-50. All of the scripts used are available on request.

### Cell culture

Primary fetal liver cells and ESEP were cultured at 37ºC, 5% $CO_2$ in StemPro SFM medium containing penicillin, streptomycin, glutamine (2mM), beta-mercaptoethanol (55uM), stem cell factor (100ng/ml), dexamethasone ($10^{-6}$M), erythropoietin (2U/ml), and insulin-like growth factor-1 (40ng/ml). To induce differentiation of the cells, mifepristone (3nM), insulin (10ug/mL) and erythropoietin (10U/mL) were substituted for stem cell factor, dexamethasone and insulin-like growth factor-1. For differentiation experiments, cells were washed twice with PBS before being transferred to the differentiation medium. Benzidine staining was performed by adding 2uL of 30% hydrogen peroxide to 100uL of 0.2% benzidine dihydrochloride in 3% glacial acetic acid and combining the mixture with 100uL of cells in medium. After 5 minutes the number of blue, benzidine-positive cells was determined by counting at least 100 cells with bright-field microscopy. NIH-3T3, RAW 264.7 macrophage and K562 were grown in Dulbecco's Modified Eagle's Medium containing penicillin, streptomycin and 10% fetal bovine serum. All cytokines were purchased from Peprotech. For proliferation assays, cells were counted by hemocytometer or by FACS analysis using the ratio of total cells to accuBeads (Fisher).

**Colony Assays**

Colony assays in primary cells were performed by plating 10,000 E14.5 fetal liver cells into M3234 methylcellulose. For CFUe colonies, methylcellulose was supplemented with 10U Epo and colonies were counted at day 3. All other colonies were grown in M3234 supplemented with 10U Epo, 20ng/mL IL-3, 20ng/mL IL-6, 100 nM Dex and 50 ng/mL SCF and counted after 9 days.

For colony assays performed in immortalized BFUe+Runx1 and BFUe+Pu.1 1,000 cells were plated in M3234 methylcellulose supplemented with 10U Epo, 20ng/mL IL-3, 20ng/mL IL-6, 100 nM Dex and 50 ng/mL SCF and counted after 9 days. Colonies were stained with benzidine and scored as large and small erythroid colonies. Representative images of small and large colonies are shown in **Figure S8F**.

**CRISPR/dCas9 TF blocking and cis-element CRISPR/Cas9-mediated deletion**

**CRISPR/dCas9 TF blocking ---** dCas9 was used to block Runx1 TF binding sites within the Pu.1 URE in K562 cells. We created a dCas9 vector (pHR-SFFV-dCas9-P2A-mCherry) by removing the Krab domain and glycine linker from pHR-SFFV-KRAB-dCas9-P2A-mCherry (Addgene #60954). gRNAs were cloned into pU6-sgRNA EF1Alpha-puro-T2A-BFP (Addgene #60955). See **Table S6** for gRNA sequences. 5 million K562 were washed with PBS and electroporated in buffer containing 5mM KCl, 15mM MgCl2, 15mM HEPES, 120 mM Na2HPO4/NaH2PO4 pH 7.4 and 50mM mannitol using the Amaxa protocol for K562 cells with 5μg pHR-SFFV-dCas9-P2A-mCherry and 5μg of sgRNA plasmid. After electroporation, cells were transferred to pre-warmed medium and after 48 hours of incubation at 37ºC cells with the highest 1% of expression of mCherry and BFP were isolated by FACS and assayed for Pu.1 mRNA by qRT-PCR.

**Cas9-mediated deletion ---** CRISPR/Cas9-mediated deletion was performed as in (5). Briefly, 5 million ESEP cells were washed with PBS and electroporated using the Amaxa kit and electroporation program for CD34+ cells with 5ug of plasmid DNA (FUCas9Cherry Addgene #70182) encoding CRISPR/Cas9 and 5ug of two plasmid DNAs encoding gRNAs that target regions just upstream and downstream of the desired deletion (**Table S4**). After electroporation, cells were placed in pre-warmed medium and after 48 hours of incubation at 37ºC single cells with the highest 1% of expression of mCherry and BFP were deposited by fluorescence-activated cell sorting (FACS) into 96-well tissue culture

plates. After 10-14 days 100 clones were picked and screened for successful deletion of the desired genomic locus using a set of PCR primers lying within, and external to, the targeted region (**Table S5**; representative screening results **Figure S4**).

**Ectopic expression**

Ectopic expression of proteins was carried out by infecting cells with recombinant lentiviruses produced using the 1501-Ef1a-IRES-GFP plasmid vector. Lentiviruses were produced in 293T cells by calcium phosphate transfection as in(6), concentrated and resuspended in StemPro SFM. Cells were infected by centrifugation for 90 minutes at 800g at 37ºC in proliferation medium supplemented with 5ug/mL polybrene and then further incubated overnight at 37ºC. Infected cells were isolated by FACS sorting for GFP-positive cells. Positive cells were defined by comparison with cells not infected with GFP virus.

**FISH-IF**

Cells were fixed in 3.2% (v/v) PFA (Electron Microscopy Sciences) in PBS with 1mM MgCl2 (PBSM) at room temperature (RT) for ten minutes. Cells were washed with 2 ml cold PBSM with 10mM glycine and permeabilized by incubation on ice for 20 minutes in PBSM with 0.1% (v/v) TritonX1000 and 2mM Vanadyl Ribonucleoside Complex (VRC). After washing with PBSM, cells were incubated at RT with prehybridization-30 buffer (prehyb-30; 30% formamide, 2X SSC, 2mM VRC) and then stained overnight at 37°C with hybridization buffer consisting of 10% Dextran Sulfate, 30% formamide, 2X SSC, 2mM VRC, 100ug/ml sheared ssDNA from salmon sperm, 100ug/ml *E.coli* tRNA, 10ug/ml molecular grade bovine serum albumin, and 200ng each of 70mer primary probe mixes against *PU.1*. Cells were washed twice for 20 minutes at 37°C with prehyb-30, and once with 2X SSC and then post fixed in 1% PFA in 2XSSC for 5 minutes, followed by washing in 2X SSC. Stained primary cells were washed with prehyb-10 (10% formamide, 2X SSC) for 10 minutes at 37°C and stained with 10% Dextran Sulfate, 10% formamide, 2X SSC, 2mM VRC, 100ug/ml sheared ssDNA from salmon sperm, 100ug/ml *E.coli* tRNA, 10ug/ml molecular grade bovine serum albumin, 20ng each of Cy3 20mer readout probes for *PU.1,* and anti-PU.1 (1:200) rabbit polyclonal antibody (T-21; Santa Cruz) for 6 hours at 37°C. Cells were washed with 1 wash prehyb-10 and 2 washes of 2X SSC (30 mins each) RT, followed by staining with 1 μg/ml of anti-rabbit Alexa Fluor 647. Cells were washed finally

in PBS 3 times and mounted with Prolong Diamond+DAPI, then imaged as previously described (7).

**Intracellular staining**

Intracellular staining for Pu.1 was performed on ESEP, RAW 264.7 macrophage, and HPC-7 cells with a conjugated Pu.1-Alexa 647 antibody at a dilution of 1:50 (Cell Signaling clone 9G7) using the BD Transcription Factor Buffer Set per the manufacturer's instructions. To prevent non-specific antibody binding, cells were incubated with IgG prior to incubation with the Pu.1 antibody. Cell fluorescence was measured using a BD FACS Aria II (APC channel).

**Mice and primary cell isolation**

Murine fetal livers were isolated from E14.5 embryos dissected from pregnant females produced from timed mating of mice between the ages of 8 and 25 weeks. Wild-type mice were C57/Bl6 (Jax), mice with GFP integrated at the Pu.1 locus were obtained from the Chan lab(8)and mice lacking the URE were obtained from the Tenen lab(9). All protocols were approved by the Institutional Animal Care and Use Committee at Einstein College of Medicine (#20180304).

Primary mouse hematopoietic cells were isolated from E14.5 murine fetal livers by FACS using either a MoFlo Astrios EQ or BD FACSAria II. Lineage positive cells were removed either by FACs sorting alone or a combination of magnetic depletion (Dynabeads Biotin Binder) followed by FACs sorting. Mature cells were stained with CD41-biotin, CD45R/B220-biotin, CD3e-biotin, CD11b/Mac-1-biotin, Gr-1-biotin, and Ter119-biotin and Avidin-APC-cy7. KSL, CMP, and GMP were isolated as in (10) and BFUe and late CFUe were isolated as in(11). Fetal liver progenitors were separated by FACS using antibodies for the following markers: Kit-PE-cy7, CD71 perCP-efluro710, Sca-1-APC, CD16/32-PE, CD34-FITC. DAPI was used as a live/dead marker. Hematopoietic populations were defined as: KSL- $Kit^{+}Sca^{+}Lin^{-}$, CMP- $Kit^{+}Sca^{-}Lin^{-}CD16/32^{-}CD34^{+}$, GMP - $Kit^{+}Sca^{-}Lin^{-}CD16/32^{+}CD34^{+}$, BFUe - $Kit^{+}Sca^{-}Lin^{-}CD16/32^{-}CD34^{-}CD71^{10\%low}$, late CFUe - $Kit^{+}Sca^{-}Lin^{-}CD16/32^{-}CD34^{-}CD71^{20\%high}$ (**Table S1 & S2**). FACS data was prepared in FloJo or FACS diva. FACS analysis was performed using a BD LSR II.

**RNA and qPCR analysis**

**mRNA RT-qPCR:** Total RNA was isolated from $10^4$-$10^6$ cells using Direct-zol isolation (Zymo) or the RNeasy Plus Micro Kit (Qiagen). cDNA was synthesized using oligo dT primers (Promega). PCR amplification was performed using Power Sybr (Fisher). For measurements with differentiating BFUe or ESEP cells, *Drosophila melanogaster* S2 cells were mixed with cells before RNA isolation in order to normalize the data, since the amounts of total RNA per cell and "housekeeping" transcripts per cell change during terminal differentiation, preventing utilization of an internal normalization control. RT-qPCR normalization was performed with Drosophila actin-specific primers. qPCR primers are indicated in **Table S3**

**Primary transcript RT-qPCR:** Pu.1 primary transcripts were assayed in ESEP before and after terminal erythroid differentiation by preparing ESEP as described above followed by total RNA isolation using Trizol and synthesizing cDNA with random hexamers (Promega). Cells were treated with 100mM DRB (1000x, 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) in culture medium for 3 hours to verify the expected instability of the Pu.1 primary transcript. PCR amplification was performed using Power Sybr (Fisher). qPCR primers are indicated in **Table S3**.

**RNAseq and ChIPseq analysis**

***Pu.1 TF correlation analysis in scRNAseq*:** scRNAseq data was obtained from Socolovsky, Klein and colleagues(12). Cell fate was measured using a clustering-based approach from the authors original analysis (Spring plots)(12, 13). Erythropoiesis was divided into 100 evenly sized bins based on Spring X and Spring Y excluding cells biased towards non-erythroid cells. Pu.1 correlation analysis was performed for all cells from multipotent progenitors (MPP) to BFUe (**Figure S6**). These fractions contained nearly all of the Pu.1 expressing cells capable of erythroid differentiation. Enrichment or depletion of TF expression in Pu.1 expressing cells was determined by calculating the residual of the linear model for the frequency of TF expression in $Pu.1^+$ cells verse the frequency of TF expression in $Pu.1^-$ cells. Correlation was determined using the Pearson Correlation. 10 factors were selected for additional experiments based on a combination of their Pearson correlation with Pu.1 and the magnitude of their enrichment or depletion in Pu.1 expressing cells (see **Figures 3 and S7**).

**ChIP-seq analysis**: Publicly available datasets were used for Pu.1(14) and Runx1(15) ChIPseq in MEL cells. Reads were aligned with Bowtie2(2), deduplicated with Picard, filtered with Samtools(3) and peaks were called with HOMER(4). Overlapping binding sites within a 1kb window were determined using custom scripts and Bedtools(16). Pu.1 and Runx1 gene targets were defined as genes with a TF binding site within a -20 to +10 kb window of each RefSeq(17) TSS. Gene Ontology(18) and KEGG pathway(19) enrichment for Runx1 and Pu.1 shared target genes was performed using David(20). A list of genes which are targeted by both Pu.1 and Runx1 is available as a supplemental file.

All scripts used in this study are available upon request.

**Statistical Analysis**

Data was analyzed and plotted in GraphPad Prism. Unless explicitly stated samples were compared using the unpaired Student's T-test and a $p<.05$ was considered significant. Statistical significance of the overlap between Pu.1 and Runx1 target genes was determined using a Fisher exact test. Statistical significance between the overlap of Pu.1 and Runx1 ChIPseq binding sites was determined using chi-squared analysis with Yates' correction.

**Western blot analysis**

Total protein was isolated from $2x10^6$ ESEP, RAW 264.7, and HPC7 cells by incubating whole cells on ice in 500uL of RIPA buffer [150 mM NaCl, 20 μM Tris pH8, 2 μM EDTA, 1% Triton X, .1% SDS] for 90 minutes, followed by centrifugation for 15 minutes at 20,000g. The supernatants were collected, and total protein concentrations were determined by the Bradford assay. 25ug of each sample were separated by SDS-PAGE and transferred to nitrocellulose by wet transfer at 100V for 1 hour. Membranes were blocked for 1 hour in Odyssey blocking buffer (LI-COR) and incubated with antibodies against Pu.1(1:1000, Santa Cruz T-21, rabbit) and Tubulin (1:10,000, mouse) followed by incubation with anti-mouse and anti-rabbit fluorescent secondary antibodies. Immunoblots were imaged with a LI-COR Odyssey instrument.

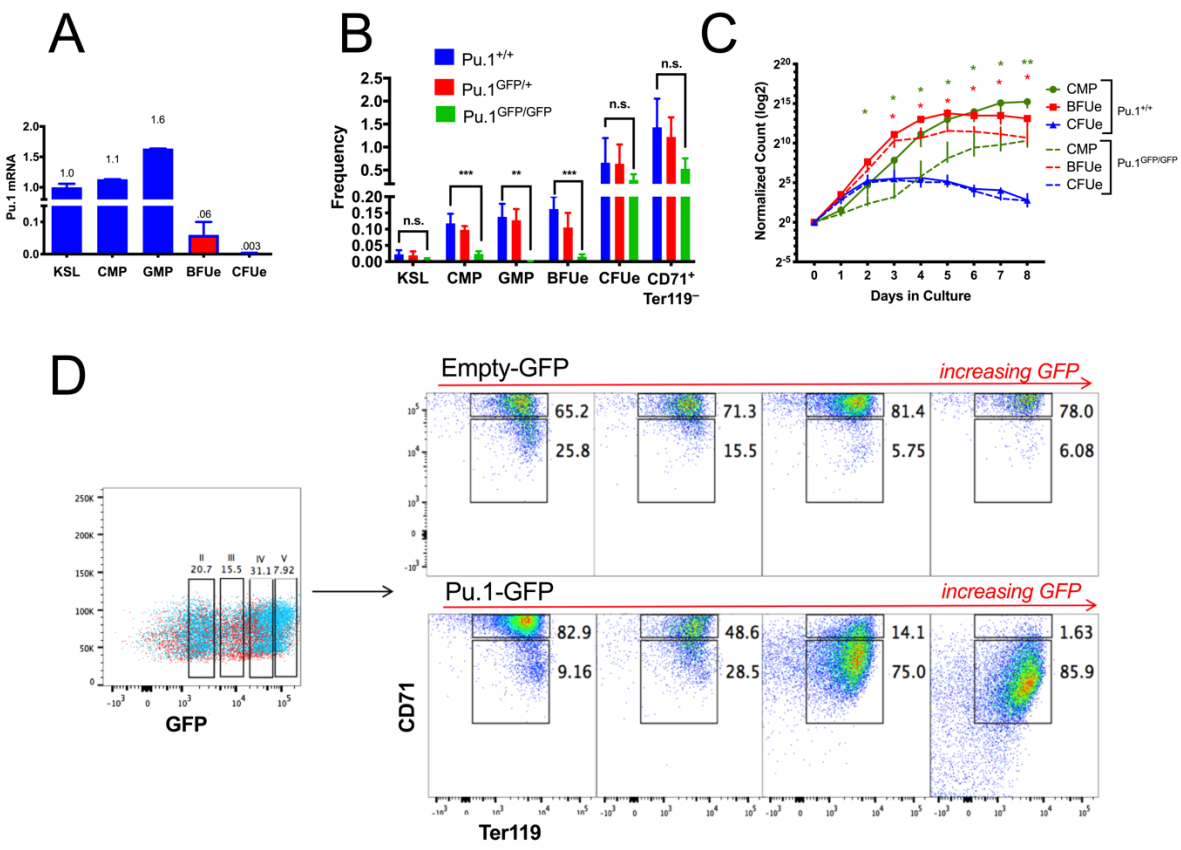


**Fig. S1. Pu.1 is required for proliferation of BFUe but sustained Pu.1 expression blocks terminal erythroid differentiation.**

**A.** KSL, CMP, GMP, BFUe and late CFUe were isolated by FACS from E14.5 fetal liver cells. Pu.1 mRNA levels were measured by qRT-PCR and normalized using the ΔΔCT approach to 3 housekeeping genes: Gapdh, Actin, and Tubulin (*n=3*).

**B.** The frequency of each cell type was determined by FACS analysis in E14.5 fetal liver from Pu.1$^{+/+}$, Pu.1$^{GFP/+}$and Pu.1$^{GFP/GFP}$ mice (*n=3-9*). The cell type marked "CD71$^{+}$Ter119$^{-}$" are Lin$^{-}$ Kit$^{+}$CD71$^{+}$Ter119$^{-}$.

**C.** CMP, BFUe, and late CFUe were isolated by FACS from E14.5 fetal livers obtained from Pu.1$^{+/+}$ and Pu.1$^{GFP/GFP}$ mice and the cells were cultured for 8 days in growth medium containing SCF, Dex, Epo, and Igf-1. Daily counts were performed using a hemocytometer (*n=3*).

**D.** BFUe were isolated from E14.5 fetal livers and infected with recombinant lentiviruses expressing either GFP (empty) or GFP and Pu.1. During the infection, cells were cultured in medium containing SCF and IL-3 for 24 hours before being transferred to medium containing EPO and cultured for an additional 24 hours to promote differentiation. FACS was used to determine the extent of differentiation as a function of GFP expression level (left) by measuring the ability of cultured BFUe to upregulate CD71 (right; Representative data from one of more than 3 experiments is shown. ns, not statistically significant; * $p \leq 0.05$; ** $p < 0.01$; *** $p < 0.001$.

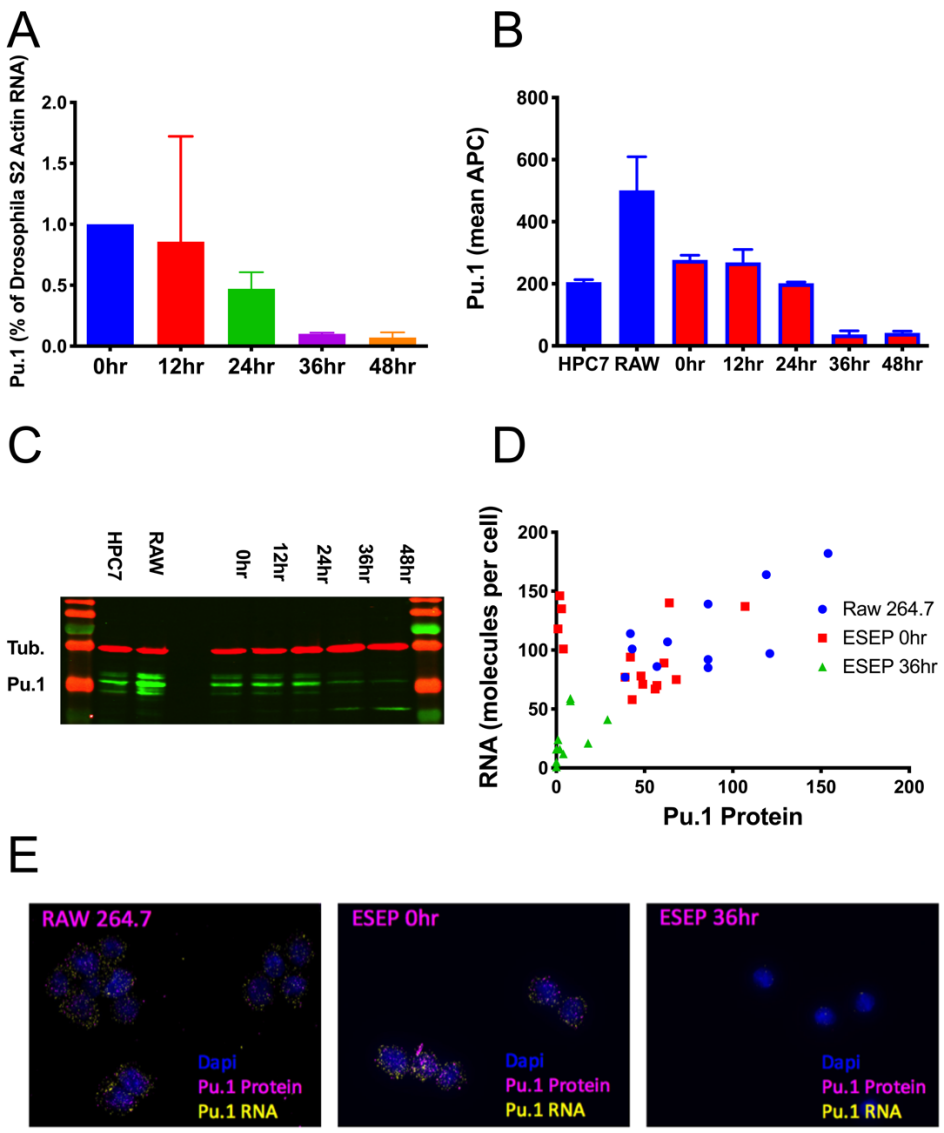


**Fig. S2. Pu.1 is downregulated during terminal erythroid differentiation of ESEP.**

ESEP were grown in medium containing SCF, Dex, Epo, and Igf-1. ESEP differentiation was triggered by transferring cells to medium containing high Epo, Mifepristone, and Insulin.

**A.** Pu.1 mRNA during ESEP differentiation. 10 million ESEP were collected and 1 million Drosophila S2 cells were added to each sample. Total RNA was isolated and qRT-PCR was carried out with primers for the Pu.1 gene and primers for Drosophila Actin. Pu.1 levels were determined by ΔΔCT approach with Drosophila Actin.

**B.** Pu.1 protein was quantified in ESEP during a differentiation time course as well as HPC-7 and RAW 264.7 macrophage by intracellular staining (BD TF intracellular staining kit) with Pu.1-APC followed by FACS analysis for mean APC fluorescence.

**C.** Pu.1 protein in ESEP, HPC-7, and RAW 264.7 macrophage quantified by western blot analysis.

**D. & E.** Pu.1 mRNA and protein in ESEP and RAW264.7 was simultaneously measured by FISH-IF as outlined in the Materials and methods. Epifluorescence images of Pu.1 mRNA and protein were quantified by Fish-Quant.

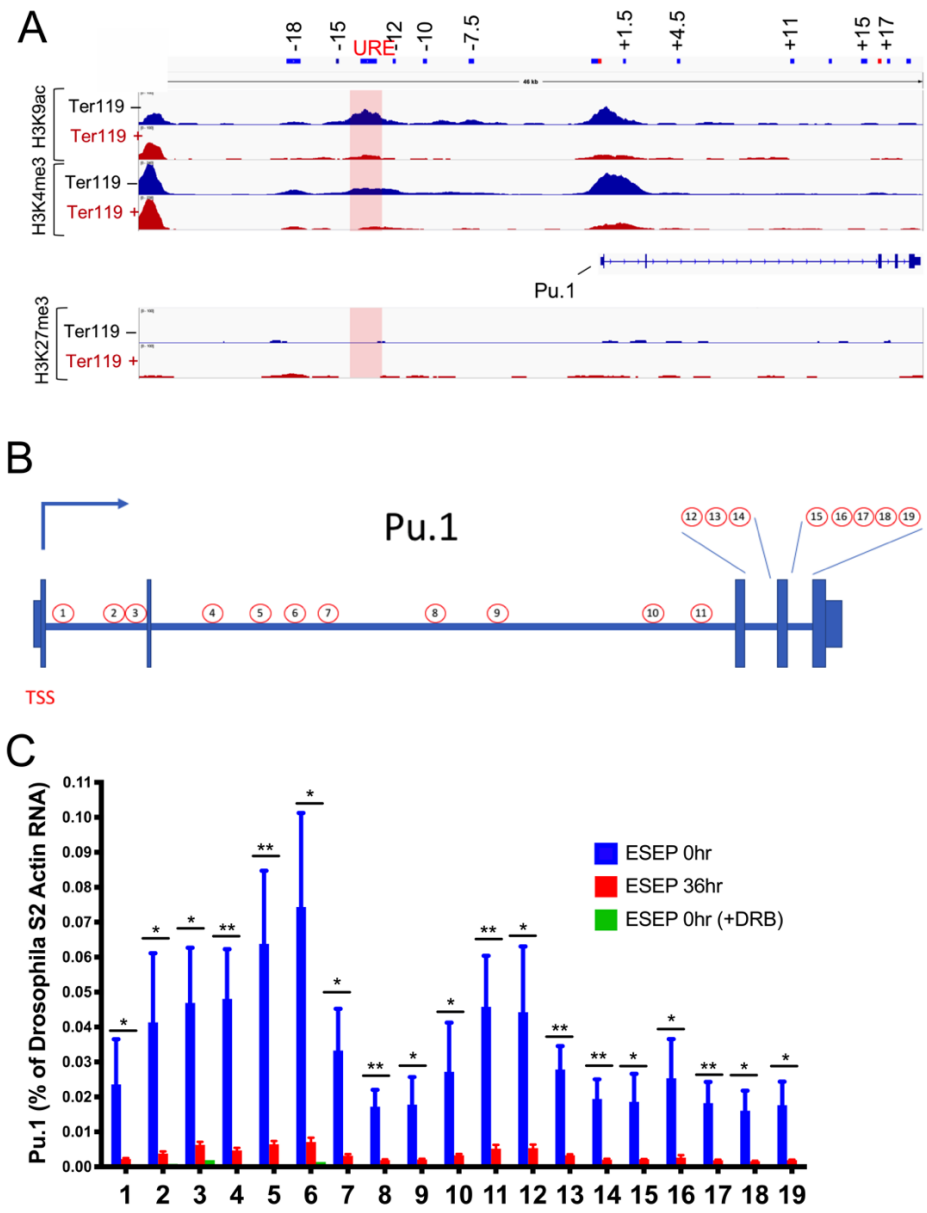


**Fig. S3. Pu.1 is downregulated during erythropoiesis at the level of transcription.**

**A.** H3K4me3, K3K9Ac and H3K27me3 ChIPseq in Ter119– and Ter119+ E14.5 fetal liver from Lodish and colleagues (21).

**B.** Schematic of the Pu.1 locus showing the locations of the 19 intronic primers used to detect Pu.1 primary transcripts in ESEP before and after terminal erythroid differentiation.

**C.** 10 million ESEP were isolated before and after culturing cells in medium containing mifepristone (3nM), insulin (10ug/mL) and erythropoietin (10U/mL) for 36 hours. ESEP were also cultured for 3 hours in medium containing stem cell factor (100ng/ml), dexamethasone ($10^{-6}$M), erythropoietin (2U/ml), insulin-like growth factor-1 (40ng/ml) and 100mM DRB (1000x, 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) which potently inhibits transcription. 1 million Drosophila S2 cells were added to each sample and total RNA was isolated with Trizol. Reverse transcription reactions were performed with random hexamers and qPCR was carried out with the indicated primers from the Pu.1 gene and primers for Drosophila Actin. Pu.1 levels were determined by ΔΔCT analysis with Drosophila Actin. ns, not statistically significant; * $p \leq 0.05$; ** $p < 0.01$; *** $p < 0.001$.

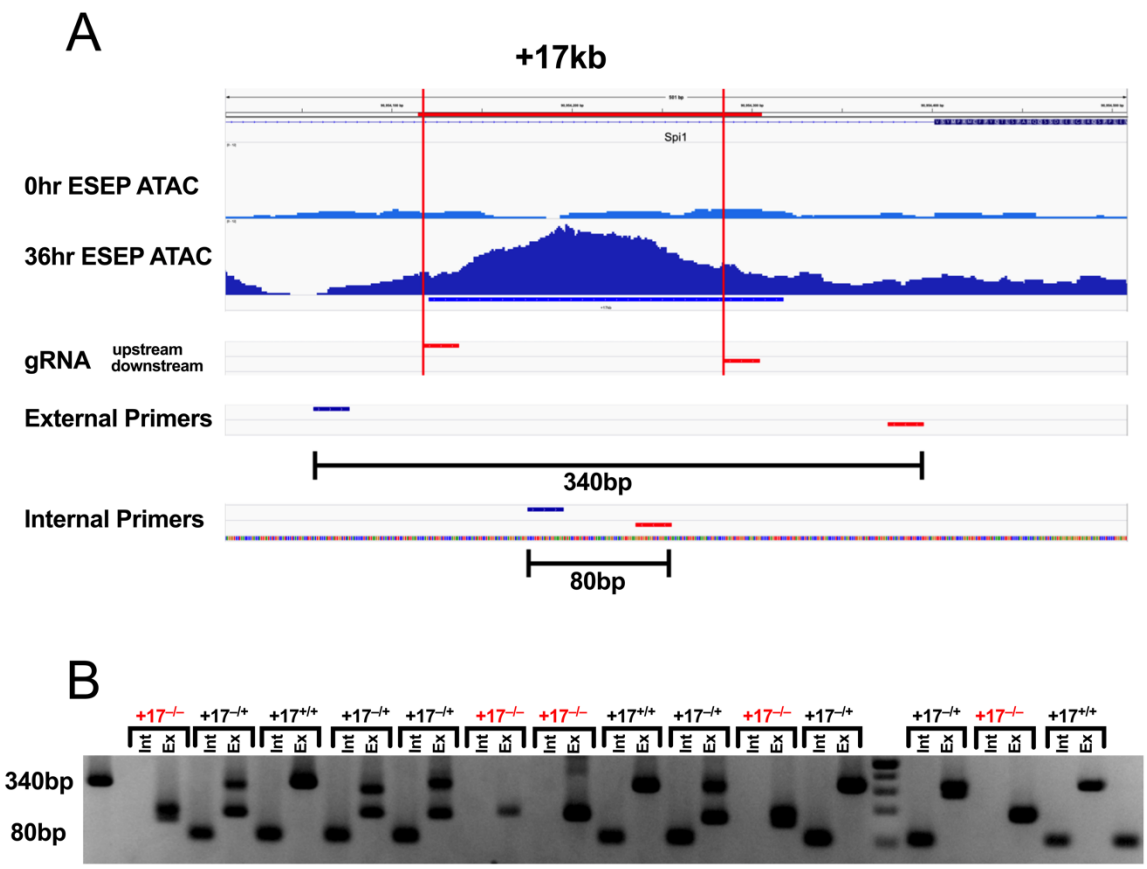


**Fig. S4. Efficient CRISPR/Cas9 mediated deletion of the chromatin accessible region 17kb downstream of the Pu.1 TSS**

**A.** Schematic showing the sites targeted by CRISPR/Cas9 and the internal and external PCR primers used to detect ESEP clones that successfully deleted the targeted loci.

**B.** Representative data of wildtype, heterozygous, and homozygous +17kb deletions. ESEP were electroporated with a gRNA upstream and downstream of the target loci as well as Cas9 as indicated in the Materials and methods. Successfully transfected clones were plated by FACS and allowed to undergo clonal amplification for 14 days. After clonal expansion genomic DNA was isolated from 100 clones as reported previously (5) and the DNA was amplified with the external and internal primers indicated in A. Clones with successful homozygous deletion of the target loci showed no amplification with primers inside the deleted region and an amplicon of decreased size for the primer set flanking the deleted region.

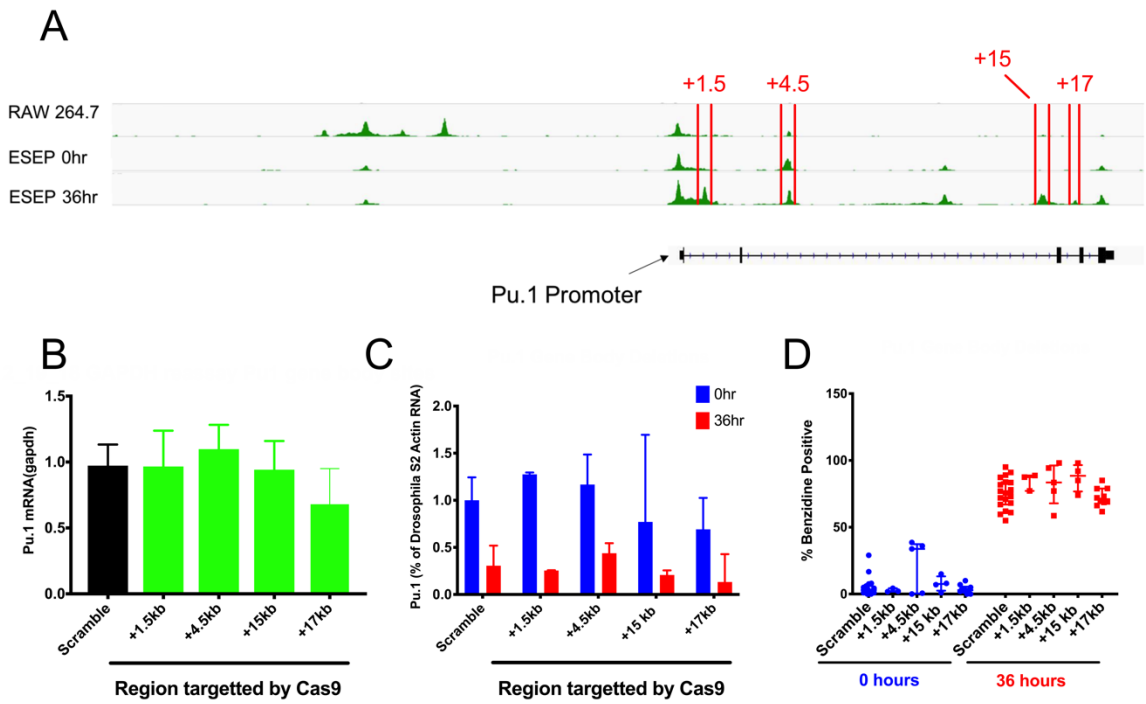


**Fig. S5. Deletion of Pu.1 gene body sites does not affect Pu.1 mRNA expression or terminal erythroid differentiation in ESEP.**

**A.** Schematic showing chromatin accessibility in ESEP before and after terminal erythroid differentiation as well as RAW 264.7. The chromatin accessible regions which were deleted in ESEP within the Pu.1 gene body at +1.5kb, +4.5kb, +15kb, and +17kb are indicated.

**B.** Pu.1 mRNA expression was measured by qRT-PCR in ESEP without each of the specified regions under proliferating conditions. Pu.1 mRNA was normalized using the ΔΔCT approach.

**C.** 10 million ESEP were collected and 1 million Drosophila S2 cells were added to each sample. Total RNA was isolated and qRT-PCR was carried out with primers for the Pu.1 gene and primers for Drosophila Actin. Pu.1 levels were determined by ΔΔCT approach with Drosophila Actin.

**D.** Each of the ESEP cell lines indicated was grown under proliferating and differentiation conditions as indicated in the Materials and methods. The percent of hemoglobin producing cells for each cell line was determined by benzidine assay.

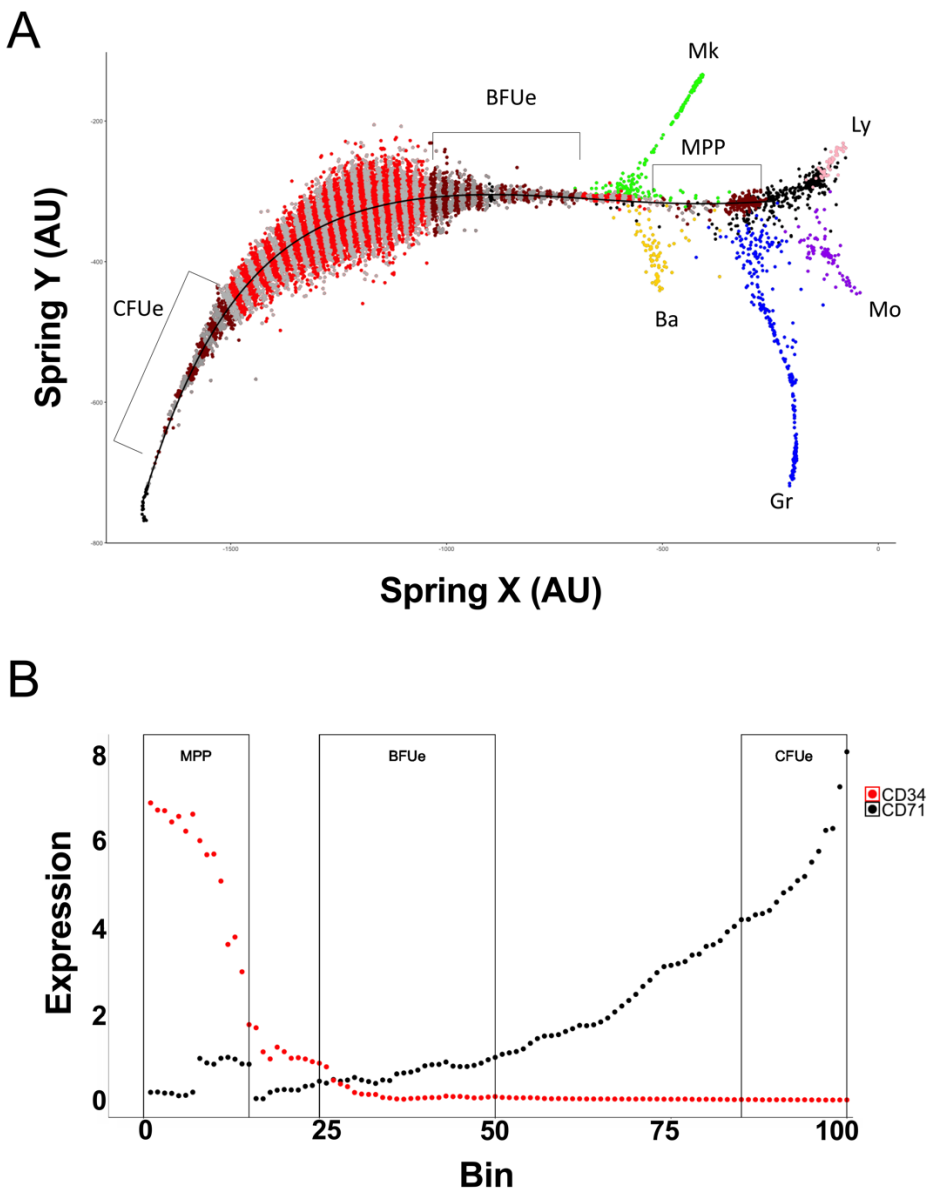


**Fig. S6. MPP, BFUe, and late CFUe defined by scRNAseq can be distinguished by variable expression of CD34 and CD71, markers used to separate them by FACS.**

**A.** scRNAseq mRNA counts for each gene in Kit+ fetal liver cells were obtained from Drs. Socolovsky, Klein and colleagues (12, 13). Cells were grouped based on their mRNA profile using the "SPRING plot" defined by the original authors (12). All cells biased towards non-erythroid lineages were removed and the full progression of erythropoiesis was divided in 100 bins. MPP, BFUe and late CFUe were defined as the earliest population of cells, the first erythroid committed population and the last, most differentiated, population of $Kit^+$ cells, respectively.

**B.** scRNAseq mRNA expression for CD34 and CD71 is indicated for each bin defined in (A). Bins containing MPP, BFUe, and late CFUe are highlighted. mRNA expression was smoothed using a rolling average over 8 adjacent bins.

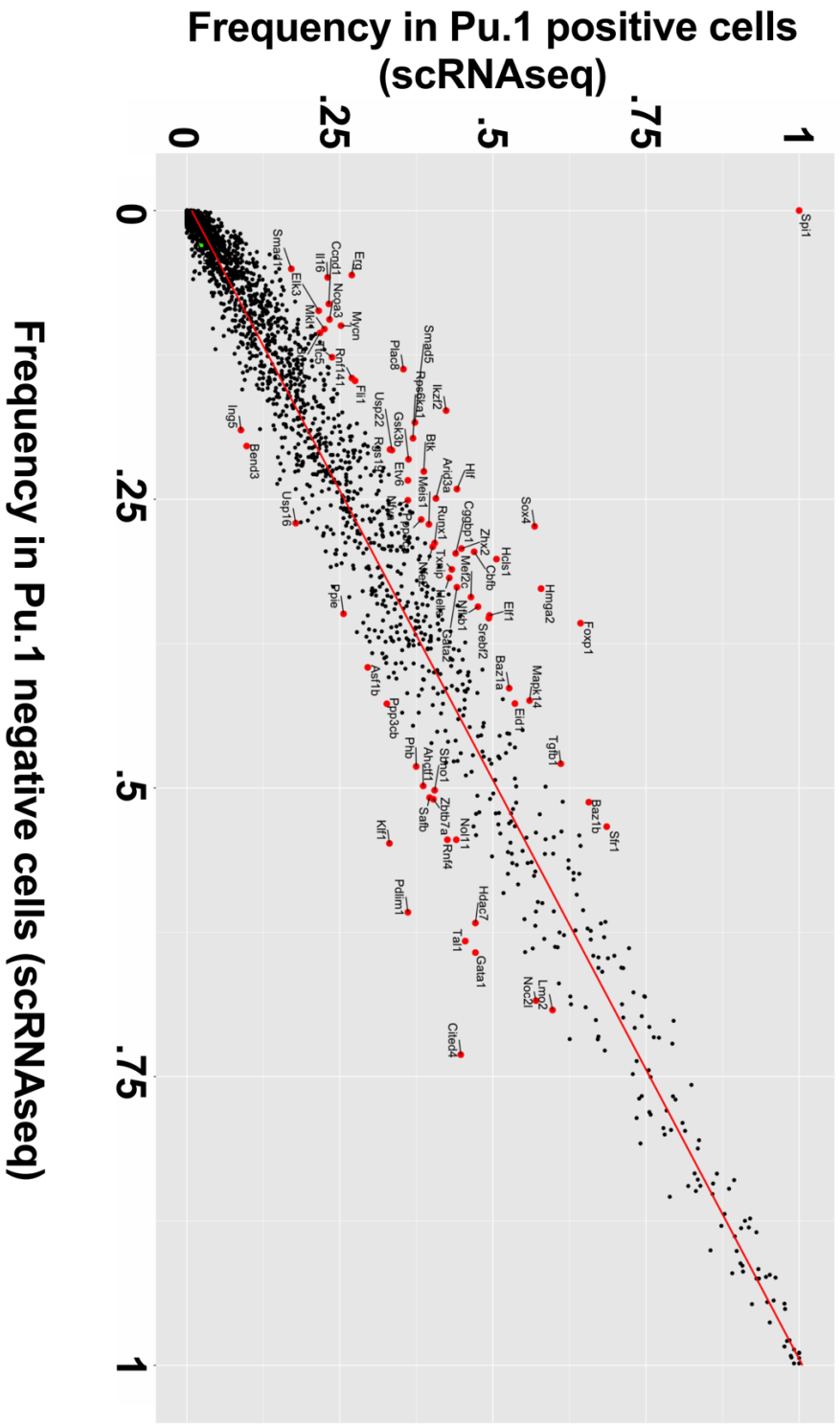


**Fig. S7. mRNA expression frequency for TFs in Pu.1 expressing and non-expressing cells**
scRNAseq mRNA counts for each gene in Kit$^{+}$ fetal liver cells were obtained from Drs. Socolovsky, Klein and colleagues (12, 13). TFs were subset from all other genes using GO annotations for "DNA-binding transcription factor" (GO:0003700). Cells from MPP to BFUe – containing essentially all Pu.1 expressing cells and many Pu.1 non-expressing cells – were divided into two sets: Pu.1 expressing cells and Pu.1 non-expressing cells. For each TF the frequency of its own expression in Pu.1 expressing and Pu.1 non-expressing cells is plotted. Factors which were over or under represented in Pu.1 expressing cells were determine by calculating the residual value from the linear model (indicated in red).

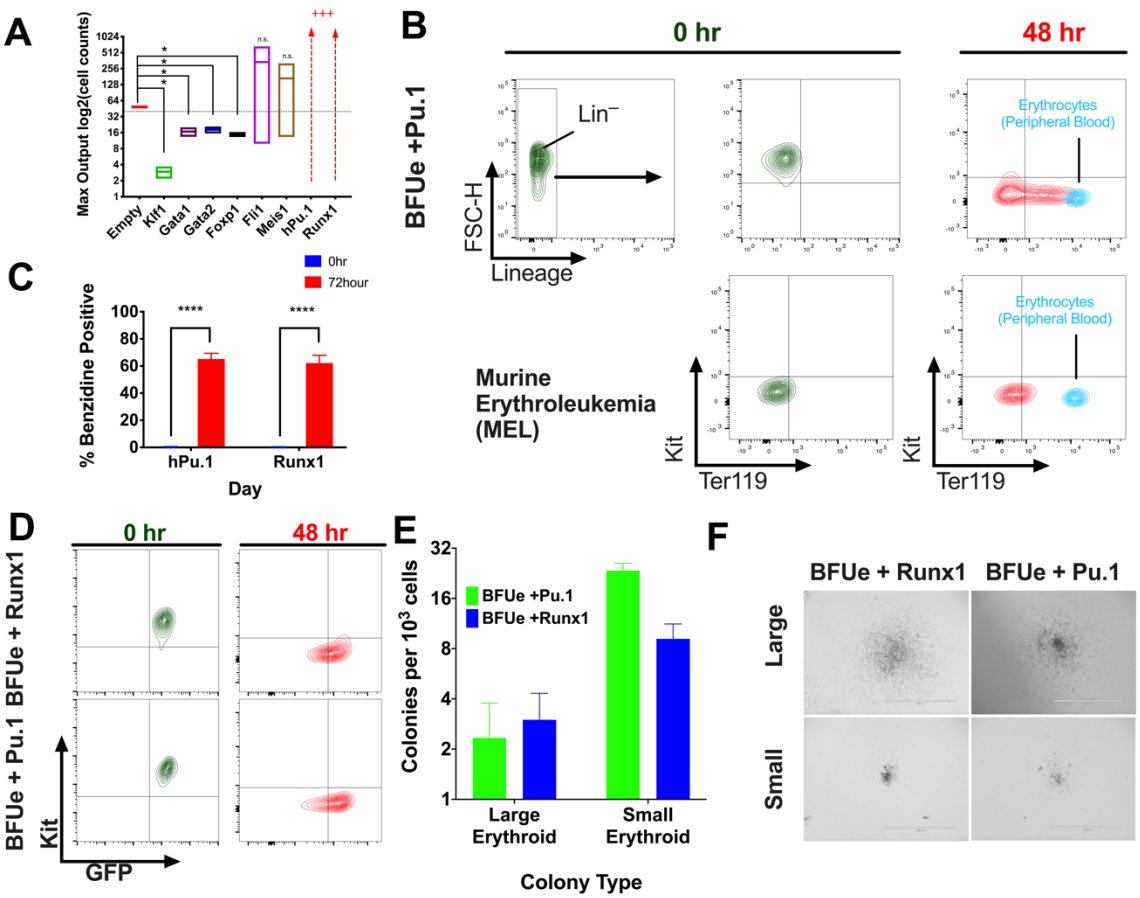


**Fig. S8. Immortalized cell lines initiated in BFUe with ectopic expression of Pu.1 or Runx1 produce a mixture of large and small erythroid colonies in colony assays.**

**A.** BFUe were isolated from E14.5 fetal livers and infected with recombinant lentiviruses encoding the indicated TF. After 48 hours of incubation in "proliferation medium", GFP+ cells were collected by FACS and further incubated in the same medium. (Empty, Pu.1, Runx1 *n=3,* Others *n=2*). The maximum number of progeny from a single cell during 18 days of culture is indicated for each recombinant lentivirus infected culture. **B.** Immortalized cells initiated in BFUe with ectopic expression of Pu.1 were proliferated *ex vivo* for four months in "proliferation medium". Kit, Ter119, and expression of lineage markers (Gr1, CD11b, CD3e, B220) were measured before and after 48 hours of differentiation in media containing Epo, Mifepristone, and Insulin. Representative FACS plots are shown (*n=3*). As a control, murine peripheral blood was stained for Kit and Ter119 and is shown in blue. Murine erythroleukemia cells were similarly stained and are shown before and after 48 hours of differentiation in medium containing 2% DMSO. **C.** BFUe cell lines were prepared using ectopic expression of Pu.1 and Runx1 as in B. Hemoglobin production was measured using benzidine staining before or after 72 hours of differentiation in media containing Epo, Mifepristone, and Insulin. **D.** Immortalized cells were prepared as in C. Kit and GFP were measured before and after 48 hours of differentiation in media containing Epo, Mifepristone, and Insulin. Representative FACS plots are shown (*n=3*). **E. & F.** Immortalized cells initiated in BFUe with ectopic expression of Pu.1 or Runx1 were proliferated *ex vivo* for four months in "proliferation medium". The following experiments were performed for both of the immortalized cell types. Cells were transferred to M3234 methylcellulose supplemented with 10U Epo, 20ng/mL IL-3, 20ng/mL IL-6, 100 nM Dex and 50 ng/mL SCF. Colonies were counted after 9 days. All colonies were benzidine positive and quantification of large and small erythroid colonies is shown (*n=2*). Representative images of large and small day 9 colonies are shown for both cell types.

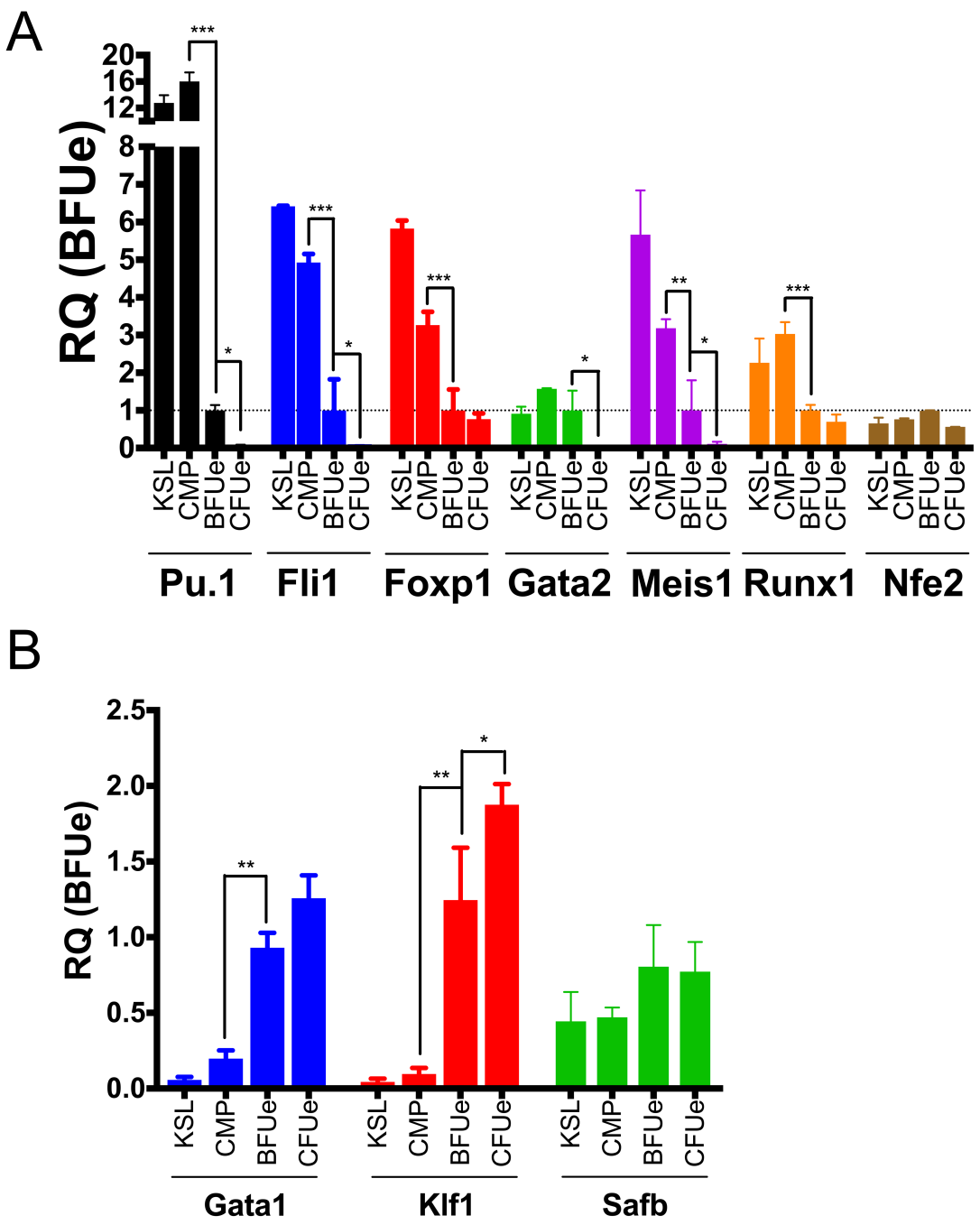


**Fig. S9. RNA expression of candidate Pu.1 regulators in KSL, CMP, BFUe and late CFUe**

**A. and B.** qRT-PCR was used to measure mRNA levels for each of the candidate TFs that were predicted to decrease (left) or increase (right) during erythropoiesis in KSL, CMP, BFUe and late CFUe from E14.5 fetal livers (see Fig. 3). RQ = relative quantification.

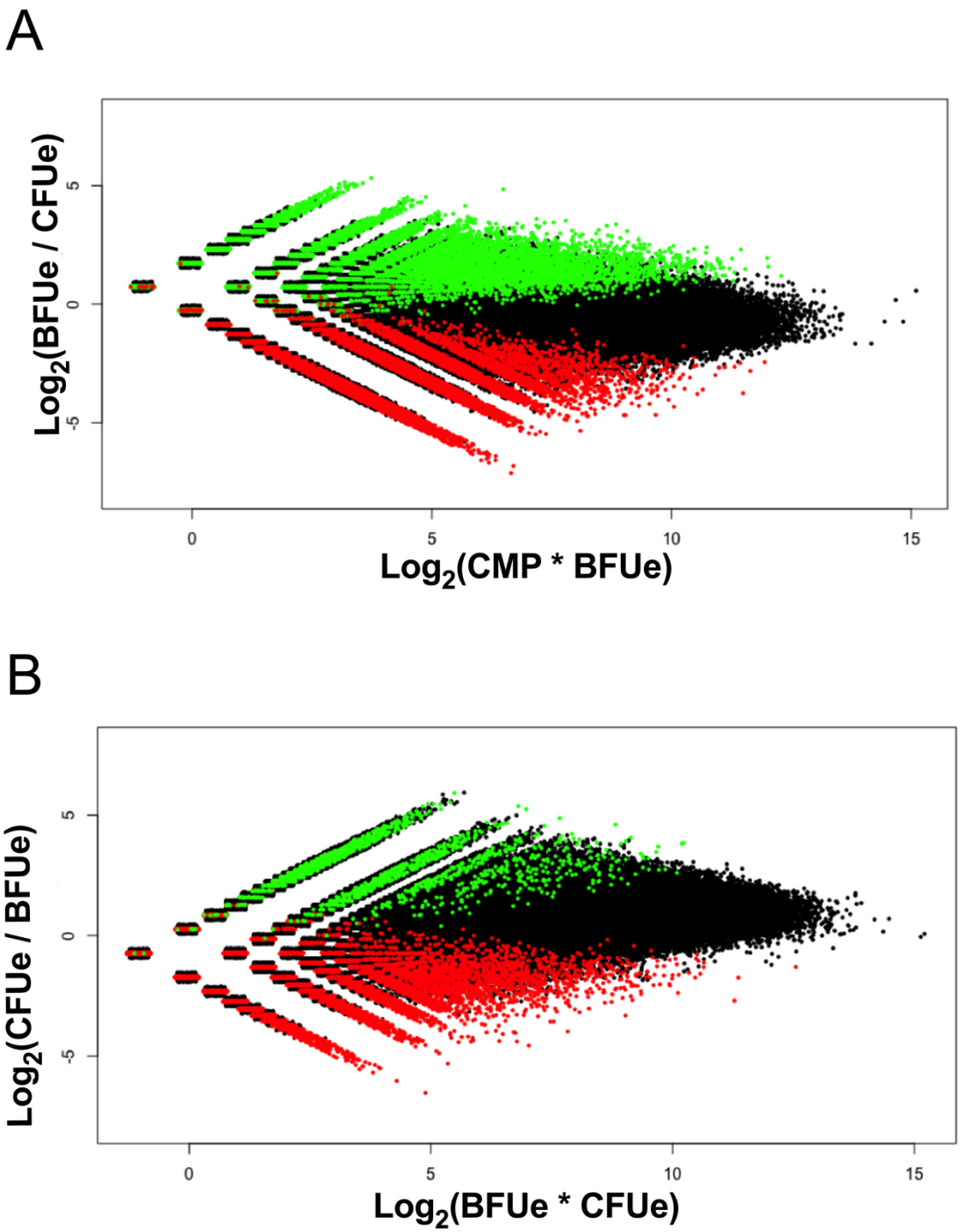


**Fig. S10. ATACseq analysis of CMP, BFUe, and late CFUe shows large changes in chromatin accessibility between cell types.**

**A. and B.** Chromatin accessibility was measured using ATACseq in E14.5 fetal liver from CMP, BFUe and late CFUe. Differential peaks, indicated in red and green, were determined by HOMER using default parameters. The scatter plot contains jitter, or noise, to minimize the overlap caused by integer numbers of reads.

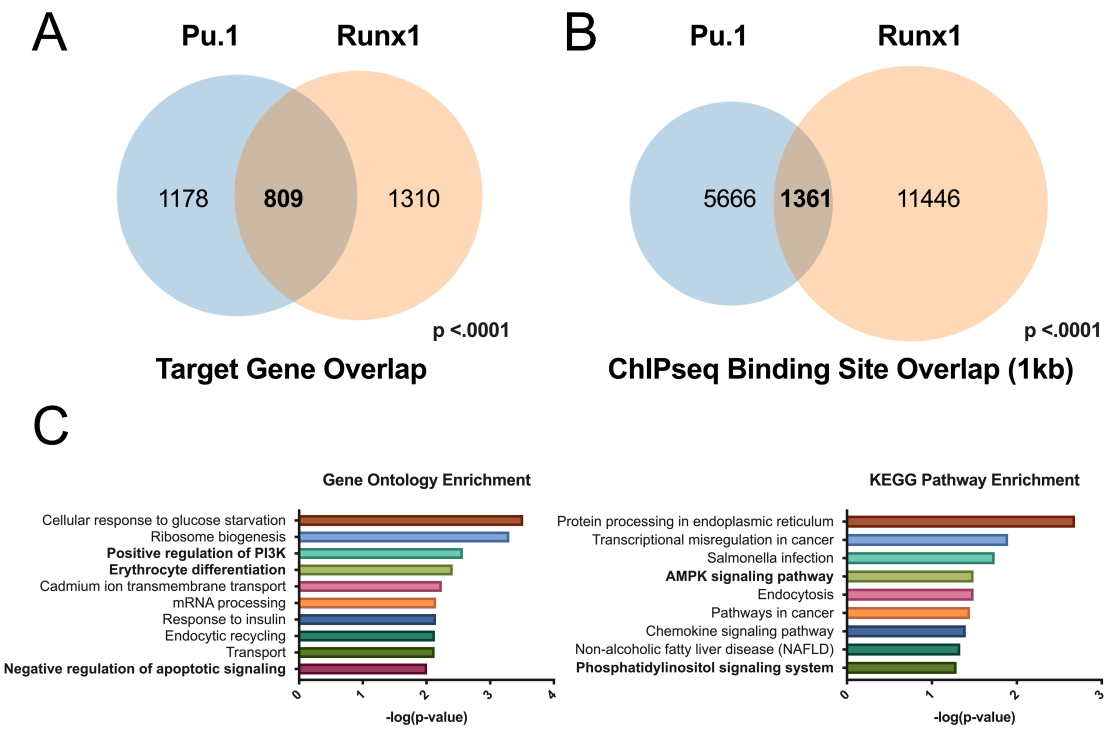


**Fig. S11. Loss of chromatin accessible regions containing the Runx and Pu.1 ETS motif is the dominant chromatin change during erythropoiesis.**

**A. and B.** Venn diagram showing the overlap of Pu.1 and Runx1 target genes and TF binding sites. Statistical significance of the overlap between Pu.1 and Runx1 target genes was determined using a Fisher exact test. Statistical significance between the overlap of Pu.1 and Runx1 ChIPseq binding sites was determined using chi-squared analysis with Yates' correction.

**C.** Gene Ontology and KEGG pathway enrichment of Pu.1 and Runx1 target genes was performed using DAVID.

**Table S1.** Hematopoietic populations were identified using FACS markers.

| **Population** | **Markers** |
|---|---|
| KSL | $Kit^{+}Sca^{+}Lin^{-}$ |
| CMP | $Kit^{+}Sca^{-}Lin^{-}CD16/32^{-}CD34^{+}$ |
| GMP | $Kit^{+}Sca^{-}Lin^{-}CD16/32^{+}CD34^{+}$ |
| BFUe | $Kit^{+}Sca^{-}Lin^{-}CD16/32^{-}CD34^{-}CD71^{10\%low}$ |
| Late CFUe | $Kit^{+}Sca^{-}Lin^{-}CD16/32^{-}CD34^{-}CD71^{20\%high}$ |
| Lineage | CD41-biotin, CD45R/B220-biotin, CD3e-biotin, CD11b/Mac-1-biotin, Gr-1-biotin, and Ter119-biotin |

**Table S2.** Antibodies used for FACS analysis of murine blood

| **Target** | **Clone** | **Manufacturer** |
|---|---|---|
| CD41-biotin | eBioMWReg30 | eBioscience |
| CD45R/B220-biotin | RA3-6B2 | eBioscience |
| CD3e-biotin | 145-2C11 | eBioscience |
| CD11b/Mac-1-biotin | M1/70 | eBioscience |
| Gr-1-biotin | RB6-8C5 | eBioscience |
| Ter119-biotin | TER-119 | eBioscience |
| Avidin-APC-cy7 | - | eBioscience |
| Kit-PE-cy7 | 2B8 | BD Biosciences |
| CD71 perCP-efluro710 | R17217 | eBioscience |
| Sca-1-APC | D7 | eBioscience |
| CD16/32-PE | 93 | eBioscience |
| CD34-FITC | RAM34 | eBioscience |

**Table S3.** qPCR primers for qRT-PCR analysis of mRNAs and the Pu.1 primary transcript.

| Target | Forward | Reverse | Comments |
|---|---|---|---|
| Pu.1 (mouse) | TAAGGGAAGCACATCCGGGG | CACGTCCTCGATACTCCCAT | Main |
| Gapdh (mouse) | GCCTTGACTGTGCCGTTG | CCAGCCTCGTCCCGTAGAC | |
| ActB (mouse) | GCCTCCTAGCACCATGAAGAT | GGGTGTAAAACGCAGCTCAGTAAC | |
| | | | |
| Act (Fly) | GCCACACGCAGCTCATTGTAG | AAGTTGCTGCTCTGGTTGTCG | S2 spike |
| Pu.1 Intron #1 Primer#1 | ACTGGCCTAGGAGTCTCTGG | GCCAGGCCACAAACTGAATG | Pu.1 unspliced transcript |
| Pu.1 Intron #1 Primer#2 | ATCCCCAAAATCTGCCTGGG | GCCGCTGTCCATCTAGTCTC | Pu.1 unspliced transcript |
| Pu.1 Intron #1 Primer#3 | CCACTGTGTCCCTGAATGCT | CGAGCCTGAGCTGTTGAAGA | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#4 | CCTCCGAGTGTCCCTTGTTC | GTGTCTCTTCCCTCTCCCCT | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#5 | GCCTTTCCCATCCTTCTGCT | GAGGTAGGGGCACTCAGAGA | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#6 | ATGGGTGTGTGGAAGCTCTG | GTGTCTGGCTTCCTGCTCTT | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#7 | ACCTCAGGCAAAGCCTTCTC | CAGGGCTCTGAAAGGACAGG | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#8 | GTCATGCTTGGTGGCAAGTG | CCTGGATCTTTGGTGTCCCC | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#9 | ACCACCTAGCCTCCTTAGCA | ACCCCACAGCATCTATCCCT | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#10 | TCCTGGACACTCTCTCTGGG | TTGGGTGCCCAGGCAATTAT | Pu.1 unspliced transcript |
| Pu.1 Intron #2 Primer#11 | TCCGGGTTTGTCTGAAGCTC | ACCAAGGCTCTGTCCATGTG | Pu.1 unspliced transcript |

| | | | |
|---|---|---|---|
| Pu.1 Intron #2 Primer#12 | AGTCCCTCAGATAGG CAGCA | CCCAACTCTAGGAAGC AGCC | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#13 | ACCTTAGGGACCCAC ACCTT | AGAGAAGATGGCCGA GGACT | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#14 | AGTCCTCGGCCATCT TCTCT | ATAGGGAGCCAGGAG GTCTG | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#15 | CAGCTCATTGCTCCC CAGAA | CTGTGGCCTCTGGTTT AGGG | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#16 | TAGGGTCCAGCATCA CCGAT | CCAGTTGACCTTGGCT AGGA | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#17 | AGGCCTGGGTGTATA CAGGA | ATCCAAAGGCCTTCAC ACGC | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#18 | AGGCCTTTGGATAGG CATGT | CAAGTGCACACCCCTC CATA | Pu.1 unspliced transcript |
| Pu.1 Intron #3 Primer#19 | GGTGTGCACTTGAGT GGGAG | GGTGTGTGGACAGGA TGCAA | Pu.1 unspliced transcript |
| | | | |
| Foxp1 | ACGGATCAGCCATCC AGAAC | CTGGTGCCTCTCCGTT GG | Pu.1 candidate regulators |
| Pu.1 (human) | GCAGATGCACGTCCT CGATA | GGACAGGGATGGGTA CTGGA | Pu.1 candidate regulators |
| Gata-2 | GGCAACCCTTACTAC GCCAA | CTGTGCAACAAGTGTG GTCG | Pu.1 candidate regulators |
| Nfe2 | GTTGTTGGCACAGTA TCCGC | CTTGCGACAGTTTTGG GCTG | Pu.1 candidate regulators |
| Klf1 | CTGTGACTGGAGGTT CGCTC | TCAGAGCGTGAAAAAG CACG | Pu.1 candidate regulators |
| Gata-1 | CCCACCTCTATCAGG GCCTA | GAGGTTGTAGGCGAT CCCAG | Pu.1 candidate regulators |
| Fli1 | GGCAGCCCCACAAAA TCAAC | TGTGGTCATACTCCCG CTTG | Pu.1 candidate regulators |
| Runx1 | CAAATCCGCCACAAG TTGCC | CGGGCTTGGTCTGAT CATCTA | Pu.1 candidate regulators |

| | | | |
|---|---|---|---|
| Meis | TCTGCACTCGCATCAGTACC | AAGCGTCATTGACCGAGGAA | Pu.1 candidate regulators |
| Safb (human) | ACAGGACACAAGTAGCGTGG | TTGACTCGCTGGCCAAATCT | Pu.1 candidate regulators |

**Table S4.** CRISPR-Cas9 gRNAs used for cis-element deletion within the Pu.1 gene body of ESEP

| **Name** | **upstream gRNA** | **downstream gRNA** |
|---|---|---|
| +1.5 upstream gRNA | CAGCCCAGAATCCCTTAGAG | GTGTGCTGCCTGGAAACCCA |
| +4.5 upstream gRNA | GGGGAAAATAAGTATAACGC | ACTGCATAAAGGTAAAAGTG |
| +15 upstream gRNA | TGTGGAAGGTGATGAACCAG | GAGTCTGTGGAAGGTCACAG |
| +17 upstream gRNA | TTCTGGGGAGCAATGAGCTG | CAGTCTGGACCAAGACTCAG |
| *A G was added to each sequence 5' for optimal transcription from the U6 promoter. | | |

**Table S5.** PCR screening primers external (flanking) and internal to the DNA-region being deleted by CRISPR-Cas9.

| | External Primers | | Internal Primers | |
|---|---|---|---|---|
| **Name** | **Forward** | **Reverse** | **Forward** | **Reverse** |
| +1.5 | ACACTAGCAGG TGGGAGGAT | GTGTGGGAACT GGAGGTCTG | GTGTTGGGGC AGTGAGGAAA | AGGGCTGGAA GGGATAGAGA |
| +4.5 | GGAAAGCTTCC AAATGCCTTCTT | AGGAATTCTGG GCTTTCGCC | GCAGAGTTGG AGGACTGACC | GCAACCGGGA CTGAGAAAGA |
| +15 | ACAGCATATTT GGCCCCTCC | TAAGCTCCCAC TCTAGACCCA | AGTCCCTCAG ATAGGCAGCA | CCCAACTCTA GGAAGCAGCC |
| +17 | ACTGCCCATTC ATTGGCTCA | AGGGAAGAACA GGCGTGAAG | CCCCTTTGCAT GCGTAAAGA | AGTTACCGAG CCTACCACCC |

**Table S6.** CRISPR-dCas9 gRNA sequences used to block Runx1 binding sites within the URE of human K562.

| **human URE** | **Sequence** | Notes |
|---|---|---|
| gRNA #1 | TTACCACAGGCACCGCACA G | Blocks Runx BS #2 |
| gRNA #2 | ACATGCTTCCTGTGGTGACT | Blocks Runx BS #1 |
| gRNA #3 | GCACTACCAGAGCTAACTCA | Scramble sequence |
| *A G was added to each sequence 5' for optimal transcription from the U6 promoter. | | |

**References**


1. Buenrostro JD, Wu B, Chang HY, & Greenleaf WJ (2015) ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 109:21 29 21-29.
2. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* 9(4):357-359.
3. Li H*, et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078-2079.
4. Heinz S*, et al.* (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* 38(4):576-589.
5. Bauer DE, Canver MC, & Orkin SH (2015) Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. *Journal of visualized experiments : JoVE* (95):e52118.
6. Wang X & McManus M (2009) Lentivirus production. *Journal of visualized experiments : JoVE* (32).
7. Eliscovich C, Shenoy SM, & Singer RH (2017) Imaging mRNA and protein interactions within neurons. *Proc Natl Acad Sci U S A* 114(10):E1875-E1884.
8. Back J, Allman D, Chan S, & Kastner P (2005) Visualizing PU.1 activity during hematopoiesis. *Experimental hematology* 33(4):395-402.
9. Rosenbauer F*, et al.* (2004) Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nature genetics* 36(6):624-630.
10. David Lara-Astiaso AW, Erika Lorenzo-Vivas, Irina Zaretsky, Diego Adhemar Jaitin, Eyal David, Hadas Keren- Shaul, Alexander Mildner, Deborah Winter, Steffen Jung, Nir Friedman, Ido Amit (2014) Chromatin state dynamics during blood formation. *Science*.
11. Flygare J, Rayon Estrada V, Shin C, Gupta S, & Lodish HF (2011) HIF1alpha synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. *Blood* 117(12):3435-3444.
12. Tusi BK*, et al.* (2018) Population snapshots predict early haematopoietic and erythroid hierarchies. *Nature* 555(7694):54-60.
13. Weinreb C, Wolock S, Tusi BK, Socolovsky M, & Klein AM (2018) Fundamental limits on dynamic inference from single-cell snapshots. *Proc Natl Acad Sci U S A* 115(10):E2467-E2476.
14. Wontakal SN*, et al.* (2012) A core erythroid transcriptional network is repressed by a master regulator of myelo-lymphoid differentiation. *Proc Natl Acad Sci U S A* 109(10):3832-3837.
15. van Riel B*, et al.* (2012) A novel complex, RUNX1-MYEF2, represses hematopoietic genes in erythroid cells. *Molecular and cellular biology* 32(19):3814-3822.
16. Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841-842.
17. O'Leary NA*, et al.* (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research* 44(D1):D733-745.

18. Ashburner M*, et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25(1):25-29.
19. Kanehisa M, Furumichi M, Tanabe M, Sato Y, & Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research* 45(D1):D353-D361.
20. Dennis G, Jr.*, et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology* 4(5):P3.
21. Wong P*, et al.* (2011) Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. *Blood* 118(16):e128-138.