Supplemental materials for Khandros *et al,* Understanding Heterogeneity of Fetal Hemoglobin Induction through Comparative Analysis of F- and A-erythroblasts.

SUPPLEMENTAL METHODS

HUDEP cell culture:

Cells were maintained in StemSpan Serum Free Medium (SFEM, StemCell Technologies, Cat. #09650) supplemented with 50 ng/ml human Stem Cell Factor (hSCF, Peprotech, Cat #300-07), 1 μM dexamethasone (Sigma, Cat #D4902), 1μg/ml doxycycline (Sigma, Cat #D9891), 3 IU/ml erythropoietin (Amgen, Cat #55513-144-10), 1% penicillin/streptomycin (ThermoFisher, Cat #15140122). Cells were kept at a density of less than 1.0 million cells/ml. Differentiation was achieved by growing cells for 5-7 days in IMDM (Mediatech, Cat #MT10016CV) supplemented with 50 ng/ml hSCF for the first 3 days, 3 IU/ml erythropoietin, 5% fetal bovine serum, 320 μg/ml holo-transferrin (Sigma, Cat # T4132), 1% penicillin/streptomycin, 10 μg/ml (2 IU/mL) heparin (Sigma, Cat # H3149), 10 μg/ml insulin (Sigma, Cat # I9278), 1 μg/ml doxycycline. Media was changed once at day 3 of differentiation to remove hSCF.

HUDEP2C single cell clones

Single-cell derived clonal lines were generated from parental HUDEP2C cells by limiting dilution. 11 clonal lines were initially selected, expanded, and frozen down. Subclones of selected clones were similarly generated. Clones were expanded and differentiated as described above. For HbF inducer experiments, cells were treated with either vehicle, hydroxyurea, or pomalidomide for 2 days of expansion followed by 7 days of differentiation. Hydroxyurea (Sigma, Cat # H8627) was used at a final concentration of 50 μM and pomalidomide (Sigma, Cat # P0018) was used at a final concentration of 10 μM. Each clone was tested for HbF expression on day 7 of differentiation by western blotting and flow cytometry as described below, and by ion exchange HPLC. All clones were analyzed in 3 independent experiments from separately thawed frozen stock. Western blots from differentiated and undifferentiated cells were done as noted below.

CD34+ HSPC-derived erythroid cell culture

Frozen purified CD34+ cells purified from peripheral blood mononuclear cells of single donors were obtained from the Fred Hutchinson Cooperative Center for Excellence in Hematology (Seattle, WA). Cells were thawed at day 0 according to standard protocols and cultured in three phases in IMDM supplemented with 3 IU/ml erythropoietin, 5% human AB serum (Sigma, Cat # H4522), 10 μg/ml heparin, 10 μg/ml insulin. Phase-I medium was supplemented with 100 ng/ml hSCF, 1 ng/ml IL-3 (Peprotech, Cat. # 200-03), 250 μg/ml holo-transferrin. Phase- II medium was supplemented with 100 ng/ml hSCF and 250μg/ml holo-transferrin. Phase- III medium was supplemented with 1.25 mg/ml holo-transferrin. Cells were maintained in Phase-I from the day of thawing (day 0) until day 8. Cells were then transitioned to Phase-II media, and finally transitioned to Phase-III media on day 13 of culture. For experiments involving pharmacologic treatment, drugs were added at day 6 and continued until the day of collection, with media changes containing fresh drug every 2 days. Hydroxyurea (Sigma, Cat # H8627) was used at a final concentration of 50 micromolar. Pomalidomide (Sigma, Cat # P0018) was used at a final concentration of 1 micromolar. For RNP electroporation, $CD34^+$ cells were expanded for a total of 5 days in media consisting of SFEM with CC100 supplement (StemCell, Cat # 02690), with electroporation done on day 3 of expansion. Cells were subsequently transitioned to Phase I media.

Cell fixation/permeabilization, staining, and sorting

50-200 million HUDEP2C or primary cultured cells were used for each experiment. Cells were washed twice with PBS and incubated in PBS with Live/Dead fixable Near-IR or Aqua dead cell stain (ThermoFisher, Cat # L34976, L34966) at 4ºC for 30 minutes. Cells were then washed once with nuclease-free 1x PBS (prepared from ThermoFisher Cat #AM9625) and fixed in 2-4% methanol-free formaldehyde (ThermoFisher, Cat # 28906) in PBS at 50-100 million cells/mL with agitation at 4ºC for 30 minutes. We used 4% for initial experiments in HUDEP2 cells but subsequently found no difference with 2% formaldehyde fixation and used 2% formaldehyde fixation for primary cells. Following fixation, cells were washed three times with PBS/0.2% BSA and

permeabilized with PBS/0.2% BSA containing 0.1% Triton-X100 (Thermo, Cat # HFH10) for 5 minutes on ice and washed twice more with PBS/0.2% BSA. Antibody staining was done in PBS/0.2% BSA with a cell concentration of 50 million/mL for 45 minutes on ice with antibodies to HbF, CD71, and CD235 (see table below). Cells were then washed twice with PBS/BSA and resuspended in PBS/BSA for sorting. All sorting was done on a BD FACSAria Fusion sorter (BD, CHOP flow cytometry core facility). Cells were sorted into tubes coated with PBS/BSA to help reduce cell loss, washed twice, and processed immediately or stored at -80ºC until downstream analysis. For RNA purification, all wash buffers beginning with fixation were prepared with nuclease-free water and contained recombinant RNasin RNAse inhibitor (ThermoFisher, Cat # N2515) at 1:100 concentration. Permeabilization buffer, antibody staining solution, sorting buffer, and collection buffer all contained RNAsin Plus (ThermoFisher, Cat # N2615) at 1:25 concentration. For protein purification, all buffers contained Complete EDTA-free protease inhibitor cocktail (Roche, Cat #11873580001).

Antibodies for flow cytometry and sorting

Target	Fluorophore	Supplier and catalog #	Dilution
CD71, human	PE	Biolegend, 334106	1:100
CD235, human	PE-Cy7	Biolegend, 306620	1:100
HbF	APC	Thermo, MHFH05	1:10
HbF	AF647 (custom conjugated)	Novus Biological, NB110-41084	1:250

RNA purification

RNA from sorted cells was purified using the RecoverAll Total Nucleic Acid Isolation Kit for FFPE (ThermoFisher, Cat #AM1975) following the supplied protocol, except the crosslinking reversal was done at 50ºC for 45 minutes without a step at a higher temperature (we found that addition of 85 ºC decrosslink step or incubation at 50ºC for longer than 1 hour gave lower RINs). ERCC spike-in controls (LifeTechnologies) were diluted 1:10 and added to the lysates prior to column purification at 1 microliter per 1 million cells. We found that using only a single DNAse step included with the kit led to genomic DNA contamination if more than 1 million cells were being processed, so an additional DNase treatment step was done using the DNA-free DNA Removal Kit (ThermoFisher, Cat #AM1906) after RNA elution. Resulting RNA quality was analyzed using the Agilent 2100 Bioanalyzer with the RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA). Sorting and processing steps were optimized to maintain RINs around 9.

cDNA synthesis, RT-PCR, ddPCR

cDNAs were prepared by reverse-transcription using iScript Supermix (Bio-Rad, Cat. #1708841) using 100-200 ng RNA. We found that using standard SYBR Green chemistry for qPCR led to low amplification efficiencies and that better results were obtained using Taqman primer and probe combinations. Quantitative real-time PCR was done using 1 ng RNA-equivalent cDNA with Taqman Gene Expression Master Mix (ThermoFisher, Cat #4369016) and the Taqman gene expression assays listed in the table below. Reactions were run on a ViiA7 Real Time qPCR machine (ThermoFisher Scientific). Quantification was performed with the standard curve method using serial dilutions of pooled cDNA from A- and F-cells at different maturation stages. Data for each assay of interest were normalized to average expression of three housekeeping genes (ACTB, HPRT, and RPS18), which were determined to be equally expressed between F- and A-cells at different maturation stages by RNA-seq. For miRNA analysis, cDNA was prepared using the Taqman Advanced miRNA cDNA synthesis kit (ThermoFisher, Cat #A28007) and quantitative PCR reactions done using Taqman Fast Advanced Master Mix (ThermoFisher, Cat # 4444557). Quantification was performed using standard curves as above, and data for each assay of interest was normalized to average expression of 4 miRNAs that are stable either between fetal and adult erythroid cells or across erythroid differentiation (miR-16a, miR-103a, miR-144, miR-191). For ddPCR analysis, reactions were assembled with ddPCR Supermix for Probes (Bio-Rad), with multiplexed FAM-labeled hemoglobin transcript assay and VIC-labeled AHSP assay. 0.01-0.1 ng RNA-equivalent cDNA was used for duplicate reactions. Droplets were generated and analyzed using the QX200 AutoDG Droplet Digital PCR system (Biorad).

Taqman gene expression assays

Target	Assay ID

ACTB	Hs01060665_g1
AHSP	Hs00372339_g1
BCL11A	Hs01093197_m1
BGLT3	Hs04406765_s1
CA1	Hs01100176_m1
CHD3	Hs01050212_m1
CHD4	Hs00172349_m1
GATAD2A	Hs00214293_m1
GATAD2B	Hs00372672_m1
HBA1/2	Hs00361191_g1
HBB	Hs00747223_g1
HBD	Hs00426283_m1
HBG1/2	Hs00361131_g1
HDAC1	Hs00606262_g1
HDAC2	Hs00231032_m1
HPRT	Hs02800695_m1
KLF1	Hs00610592_m1
LIN28B	Hs01013729_m1
MBD2	Hs00969366_m1
MBD3	Hs00922219_m1
MTA1	Hs00950776_m1
MTA2	Hs00191018_m1
MTA3	Hs00383033_m1
RBBP7	Hs00171476_m1
RPS18	Hs01375212_g1
ZBTB7A	Hs00792219_m1

Taqman miRNA expression assays

Target	**Assay ID**
hsa-miR-16-5p	477860_mir
hsa-miR-144-5p	477914_mir
hsa-miR-103a-3p	478253_mir
hsa-miR-191-5p	477952_mir
hsa-let-7a-5p	478575_mir
hsa-let-7b-5p	478576_mir
hsa-let-7c-5p	478577_mir
hsa-let-7d-5p	478439_mir
hsa-let-7e-5p	478579_mir
hsa-let-7g-5p	478580_mir

RNA-seq

Approximately 100 ng of treated RNA was depleted for rRNA and globin RNA using the Ribo-Zero removal reagents and protocol from the ScriptSeq Complete kit (Illumina Cat # BHMR1224), followed by purification using the RNeasy MinElute Clean Up kit (Qiagen Cat # 74202). Sequencing libraries were then constructed from depleted RNA using the TruSeq Stranded Total RNA kit (Illumina Cat # 20020598) and PCR-amplified for 13 cycles. The size and quality of each library were then evaluated by Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA), and quantified using qPCR. Libraries were sequenced in paired-end mode on the NextSeq 500 to generate 2 x 76 bp reads using Illumina-supplied kits as appropriate. The sequence reads were processed using the ENCODE3 long RNA-seq pipeline (https://www.encodeproject.org/pipelines/ENCPL002LPE/). Gene set Gencode v29 was used for mapping the reads to the human genome (hg38 assembly) using STAR, followed by RSEM for gene quantifications.

Protein purification and mass spectrometry

Cells sorted for protein purification were washed twice with PBS containing protease inhibitors. Protein extraction and crosslink reversal were done using the Qproteome FFPE Tissue Kit (Qiagen, Cat #37623) following the manufacturer instructions. Protein concentrations were quantified using a modified Lowry protein assay with the RC DC Protein Assay Kit II (Biorad, Cat # 5000122). For HUDEP2 cells, F- and A-cells from three subclones were sorted on day 6 of differentiation. 20 micrograms of protein were used per sample and size-fractionated on a 10% Bis-Tris NuPage gel (Thermo, Cat # NP0301) into fractions of 2 millimeters each. Mass spectrometry sample preparation and analysis were done by the Proteomics and Metabolomics Facility at the Wistar Institute (Philadelphia, PA). Each lane was cut into a total of 8 fractions, reduced with TCEP, alkylated with iodoacetamide, and digested with trypsin. Tryptic digests were analyzed by LC-MS/MS using a 163 min LC run on a Thermo Q Exactive Plus mass spectrometer. MS/MS data were searched against the UniProt human database (10/01/2017) using MaxQuant 1.6.2.3. Protein quantification was performed using unique+razor peptides. Protein and peptide false discovery rate was set at 1%. Filtering was done to require a protein to be identified in all 3 replicates in at least one group (A or F). 6120 total proteins were identified. Normalization was done using the MaxLFQ algorithm. LFQ intensity was subsequently $\log_2$ transformed and missing values were imputed. Statistical comparisons were done using Student's t-test, and proteins with t-test p-value <0.05 and at least 2-fold change were identified. In addition, a more stringent Student's T-test Significant (using permutation-based FDR of 0.05, s0 of 0.1, and number of randomizations of 250) was also performed.

Western Blotting

We used 10-20 micrograms of protein lysate for regular western blots, and 2 micrograms for hemoglobin western blotting. Protein lysates were prepared in protein sample loading buffer (LI-COR, Cat # 928-40004) and resolved on NuPAGE Bis-Tris 4-12% polyacrylamide gels (ThermoFisher, Cat # NP0322) using MES SDS running buffer (ThermoFisher, Cat # NP0002). Proteins were transferred to Immobilon-FL PVDF membrane (Millipore Sigma, Cat #IFPL00010) and total protein staining was done using Revert 700 total protein stain (LI-COR Cat # 926-11011). After stain removal, membranes were blocked with Odyssey blocking buffer (LI-COR Cat #927-50000) and incubated in primary antibody (see list) in blocking buffer overnight at 4°C. After washing with Tris-buffered saline with Tween-20, secondary antibody staining was done using species-appropriate fluorescent secondary antibody at a 1:10,000 dilution in blocking buffer for one hour at room temperature followed by three washes. Fluorescent imaging was done using the Odyssey imaging system (LI-COR). Samples from fixed and sorted cells did give bands that were less sharp than those from fixed cells. Detection of some substrates was also limited by sample contamination with mouse and sheep IgG that was used for sorting. For semi-quantitative western blot analysis of globin chains, standard curves were prepared using γ-globin purified from cord blood hemolysates (gift from Dr. Kazuhiko Adachi) and β-globin purified from HbA (Gift from Dr. Mitchell J. Weiss).

Primary antibodies for western blotting

Target	**Species/clonality**	**Supplier and catalog #**	**Dilution**
Alpha hemoglobin	Goat polyclonal	Novus Biologicals, NB119-41083	1:2000
Beta hemoglobin	Mouse monoclonal (37-8)	Santa Cruz Biotechnology, SC-21757	1:1000
Beta hemoglobin	Mouse monoclonal	Novus Biologicals, H00003043-M12	1:2000
Gamma hemoglobin	Sheep polyclonal	Novus Biologicals, NB110-41084	1:5000
Gamma hemoglobin	Rabbit monoclonal (D4K7X)	Cell Signaling, 39386	1:2000
BCL11A	Rabbit monoclonal (D4E3P)	Cell Signaling, 75432	1:1000
BCL11A	Mouse monoclonal (14B5)	Abcam, ab19487	1:1000
LIN28B	Rabbit Monoclonal	Cell Signaling, 4196	1:1000
LRF	Hamster monoclonal (13E9)	Thermo, 14-3309-82	1:1000
Tubulin	Rabbit polyclonal	Cell Signaling, 2146S	
Actin	Mouse monoclonal (C4)	Santa Cruz Biotechnology, SC-47778	1:1000
RPL19	Rabbit polyclonal	Proteintech, 14701-1-AP	1:1000

Secondary antibodies for western blotting:

Target Species	Host Species	Fluorescent label	Supplier and catalog #
Mouse	Donkey	IRDye 800CW	LI-COR, 926-32212
Mouse	Donkey	IRDye 680RD	LI-COR, 926-68072
Rabbit	Donkey	IRDye 800CW	LI-COR, 926-32213
Rabbit	Donkey	IRDye 680RD	LI-COR, 926-68074
Sheep/goat	Donkey	IRDye 800CW	LI-COR, 926-32214
Sheep/goat	Donkey	IRDye 680RD	LI-COR, 926-68074
Hamster	Donkey	DyLight 680	Rockland, 620-144-440

HPLC Analysis

Hemolysates were prepared from 1-2 million cells by osmotic lysis with 0.05x PBS on ice for 30 minutes, and clarified by centrifugation at 20,000g. Clarified hemolysates from experimental samples were analyzed by cation-exchange high-performance liquid chromatography (HPLC). We utilized a Hitachi D-7000 Series (Hitachi Instruments, Inc., San Jose, CA), and a weak cation-exchange column (Poly CAT A: 35 mm x 4.6 mm, Poly LC, Inc., Columbia, MD). Hemoglobin isotype peaks were eluted with a linear gradient of phase B from 0% to 80% at A_{410nm} (Mobile Phase A: 20 mM Bis-Tris, 2 mM KCN, pH 6.95; Phase B:20 mM Bis-Tris, 2 mM KCN, 0.2 M sodium chloride, pH 6.55). Clarified lysates from normal human cord blood samples (high Hb F content), as well as a commercial standard containing approximately equal amounts of Hb F, A, S and C (Helena Laboratories, Beaumont, TX), were utilized as reference isotypes.

LIN28B Targeting in HUDEP and primary cells

LIN28B editing in HUDEP1 and HUDEP2 cells stably expressing spCas9 was done as previously described using lentiviral expression of gRNA listed below.[1,2] Two non-targeting controls as well as two controls that cut in gene introns and are not predicted to disrupt gene function were used. LIN28B gRNAs were based on published sequences.[3,4] Cells were infected with lentivirus containing gRNA, sorted for GFP expression, and expanded for 7 days prior to differentiation or generation of single-cell derived clones by limiting dilution. We did note that HUDEP2 cells tended to increase LIN28B expression over time when expressing any gRNA (including non-targeting controls), especially with single cell cloning. These cells also expressed high HbF levels on differentiation. Genomic DNA extracted from single cell clones was analyzed for editing by PCR amplification and sequencing of 400-500 nt regions containing the gRNA recognition site. Sequences were analyzed using Synthego ICE CRISPR analysis tool (https://www.synthego.com/products/bioinformatics/crispr-analysis). LIN28B is triplicated in HUDEP2 cells, so clones with a 2-3 successfully edited alleles with predicted frameshift mutations were selected. For LIN28B overexpression experiments, pMSCV-PIG-LIN28B plasmid was obtained from Addgene (Plasmid # 64795) and virus prepared using standard protocols.

For primary cell experiments, cells were electroporated with RNP complexes generated using 300 pmol of Synthego modified gRNA (Synthego, Menlo Park, CA) and 50 pmol of IDT Alt-R S.p. HiFi Cas9 Nuclease V3 (IDT, Coralville, IA, Cat # 1081061). 50,000 cells were electroporated with the RNP complex using the P3 Primary Cell 4D-Nucleofector X Kit S, (Lonza, Cat # V4XP-3032) with the Amaxa 4D Nucleofector (Lonza) using program DZ-100. Cells were expanded and differentiated as per standard protocol, and editing efficiency assessed after 6 days as above. Cells were analyzed for HbF expression on day 12 of culture.

gRNA sequences for HUDEP experiments

gRNA name	Target	Guide sequence
NT1	Non-targeting control	GTATTACTGATATTGGTGGG
NT2	Non-targeting control	GTAGCGAACGTGTCCGGCGT
INT1	AHSP intron 2	AGACACGGGATACAATGCAG
INT2	LIN28B intron	TTATGTTAGGGGATTTGCAG
LIN28B gRNA1	LIN28B, exon 2	CATCGACTGGAATATCCAAG
LIN28B gRNA2	LIN28B, exon 3	CAGAGCAAACTATTCATGGA
LIN28B gRNA3	LIN28B, start codon	TGAGGGCCCGTGGGGCAACA
LIN28B gRNA4	LIN28B, exon 3	AATGATTACCTATCTCCCTT

gRNA sequences for CD34 erythroid culture experiments

gRNA name	Target	Guide sequence
NT1	Non-targeting control	GCACUACCAGAGCUAACUCA
AHSPint2	AHSP intron 2	AGACACGGGATACAATGCAG
hBCL11A+58	BCL11A +58 enhancer	CTAACAGTTGCTTTTATCAC
LIN28B gRNA1	LIN28B, exon 2	CATCGACTGGAATATCCAAG
LIN28B gRNA4	LIN28B, exon 3	AATGATTACCTATCTCCCTT

SUPPLEMENTAL REFERENCES

1. Lan X, Khandros E, Huang P, et al. The E3 ligase adaptor molecule SPOP regulates fetal hemoglobin levels in adult erythroid cells. *Blood Adv*. 2019;3(10):1586–1597.

2. Grevet JD, Lan X, Hamagami N, et al. Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin regulator in human erythroid cells. *Science*. 2018;361(6399):285–290.

3. Tan FE, Sathe S, Wheeler EC, et al. A Transcriptome-wide Translational Program Defined by LIN28B Expression Level. *Mol Cell*. 2018;73(2):304-313.e3.

4. Zhang J, Ratanasirintrawoot S, ram Chandrasekaran, et al. LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. *Cell Stem Cell*. 2016;19(1):66–80.

5. Huang P, Keller CA, Giardine B, et al. Comparative analysis of three-dimensional chromosomal architecture identifies a novel fetal hemoglobin regulatory element. *Gene Dev*. 2017;31(16):1704–1713.

6. Lessard S, Beaudoin M, Orkin SH, Bauer DE, Lettre G. 14q32 and let-7 microRNAs regulate transcriptional networks in fetal and adult human erythroblasts. *Hum Mol Genet*. 2018;27(8):1411–1420.

7. Xu J, Shao Z, Glass K, et al. Combinatorial Assembly of Developmental Stage-Specific Enhancers Controls Gene Expression Programs during Human Erythropoiesis. *Dev Cell*. 2012;23(4):796–811.

SUPPLEMENTAL TABLES

Supplemental Table 1: RNA-seq expression of selected transcripts in sorted day 7 differentiated HUDEP2C cells. Data are shown as average normalized A-cell and F-cell reads for 3 replicates, log_2 fold change and DESeq adjusted p-values.

HUDEP2C DAY 7	ID	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
HBF REGULATORS	BCL11A	2721.2	2230.1	-0.2885	0.00018	0.017
	ZBTB7A	502.8	537.9	0.0994	0.47	1
	MYB	1427.9	1764.4	0.3237	0.00053	0.039
	EIF2AK1	18143.0	19296.3	0.0831	0.28	1
	ATF4	15486.7	15302.3	-0.0177	0.81	1
	SPOP	2014.1	1912.0	-0.0758	0.36	1
	KLF1	15590.3	14669.8	-0.0756	0.4	1
	SOX6	12845.4	11197.4	-0.2033	0.002	0.11
	LIN28B	4.9	62.7	3.6798	1.3E-10	0.00000014
	IGF2BP1	909.3	947.0	0.0548	0.6	1
	HMGA2	21.5	18.9	-0.7323	0.35	1
NURD COMPLEX	CHD3	2443.5	2615.4	0.0961	0.22	0.98
	CHD4	6771.6	7869.5	0.2168	0.00054	0.039
	GATAD2A	646.5	668.3	0.0297	0.82	1
	GATAD2B	647.0	687.0	0.0793	0.5	1
	HDAC1	3407.8	3229.8	-0.078	0.32	1
	HDAC2	3694.7	3889.2	0.0687	0.39	1
	MBD2	1780.4	1661.2	-0.0958	0.28	1
	MBD3	642.8	653.7	0.0364	0.79	1
	MTA1	222.1	199.8	-0.1421	0.46	1
	MTA2	2901.7	2975.9	0.0409	0.62	1
	MTA3	102.5	91.7	-0.1454	0.6	1
	RBBP4	7058.8	7104.2	0.0051	0.94	1
	RBBP7	4737.9	5128.5	0.1139	0.085	0.79
TRANSCRIPTION FACTORS	ZFPM1	105.5	104.2	0.0132	0.96	1
	RUNX1	155.2	178.6	0.2096	0.32	1
	IKZF1	3855.1	4017.7	0.0691	0.41	1
LSD1/COREST	KDM1A	2719.8	2618.2	-0.0591	0.44	1
	RCOR1	2625.8	2357.4	-0.1716	0.044	0.62
NCOR/SMRT COMPLEX	NCOR1	11599.5	11423.0	-0.0304	0.66	1
	NCOR2	115.6	136.8	0.2919	0.25	1
	TBL1XR1	13724.6	13369.5	-0.0307	0.7	1
	TBL1X	652.9	678.2	0.0659	0.61	1
	CORO2A	13.2	22.5	0.7517	0.22	0.98
	ZBTB33	2334.1	2597.8	0.159	0.044	0.62
SIN3 COMPLEX	SIN3A	2142.1	2209.9	0.0296	0.75	1
	SIN3B	84.0	100.6	0.2481	0.38	1
OTHER COREPRESSORS	BCOR	2119.6	2151.1	0.0177	0.84	1
	TRIM28	1159.1	1162.2	0.0179	0.87	1
SWI/SNF COMPLEX	SMARCB1	471.8	480.9	0.0273	0.86	1
	SMARCA4	2724.6	2881.9	0.0881	0.27	1
	SMARCA2	1684.1	1636.2	-0.0415	0.64	1
	SMARCE1	4799.2	4772.0	-0.0059	0.94	1
	SMARCD1	225.3	224.1	0.0269	0.89	1
	SMARCD2	1689.0	1589.9	-0.0854	0.36	1
	SMARCC1	1885.4	2254.9	0.2596	0.0031	0.14
	SMARCC2	2633.7	2780.1	0.0758	0.31	1
	PBRM1	13599.2	11894.8	-0.1945	0.0014	0.082
	ASH2L	4480.0	4948.3	0.1482	0.036	0.57
OTHER NUCLEAR FACTORS	DNMT1	9986.3	10715.4	0.1031	0.1	0.83
	KDM3A	7030.2	7073.5	0.01	0.88	1
	KDM3B	3670.3	3676.7	0.0012	0.99	1
	KDM5A	8707.2	8152.7	-0.0947	0.13	0.89
	CARM1	471.3	445.9	-0.0726	0.61	1
	YLPM1	3000.5	3143.7	0.0606	0.45	1
	MSH2	3055.9	3178.8	0.0591	0.42	1

Supplemental Table 2: Mass spectrometry expression of selected proteins in sorted day 6 differentiated HUDEP2C cells. Data are shown as average LFQ for F-cells and A-cells (3 replicates), absolute fold change and t-test p-values. ND: not detected.

HUDEP2C DAY 6	ID	A-cell log_2 LFQ	F-cell log_2 LFQ	Fold change	P-value
HBF REGULATORS	BCL11A	21.84	22.58	1.677	0.5144
	ZBTB7A	25.59	25.67	1.058	0.6494
	MYB	ND	ND		
	EIF2AK1	26.74	27.01	1.201	0.1035
	ATF4	ND	ND		
	SPOP	ND	ND		
	KLF1	28.42	28.51	1.066	0.5427
	SOX6	25.06	25.55	1.395	0.1400
	LIN28B	ND	ND		
	IGF2BP1	31.42	31.41	-1.004	0.9764
	HMGA2	ND	ND		
NURD COMPLEX	CHD3	32.17	32.02	-1.107	0.6508
	CHD4	31.75	31.66	-1.066	0.4993
	GATAD2A	29.39	29.39	-1.003	0.9729
	GATAD2B	27.88	27.76	-1.082	0.7273
	HDAC1	27.39	27.78	1.308	0.2285
	HDAC2	30.80	30.84	1.024	0.8886
	MBD2	26.35	26.64	1.218	0.2632
	MBD3	28.69	28.83	1.102	0.6094
	MTA1	27.70	27.63	-1.053	0.3626
	MTA2	31.09	31.04	-1.040	0.7303
	MTA3	22.95	23.15	1.146	0.5365
	RBBP4	31.87	31.84	-1.021	0.8619
	RBBP7	32.20	32.09	-1.077	0.5300
TRANSCRIPTION FACTORS	ZFPM1	27.35	27.23	-1.092	0.5420
	RUNX1	ND	ND		
	IKZF1	28.19	28.10	-1.065	0.4647
LSD1/COREST	KDM1A	30.20	30.10	-1.070	0.4419
	RCOR1	28.63	28.66	1.024	0.7810
NCOR/SMRT COMPLEX	NCOR1	29.28	29.07	-1.156	0.2536
	NCOR2	25.25	25.10	-1.114	0.1910
	TBL1XR1	28.59	28.59	1.002	0.9649
	TBL1X	ND	ND		
	CORO2A	ND	ND		
	ZBTB33	20.86	20.87	1.009	0.9722
SIN3 COMPLEX	SIN3A	28.61	28.64	1.025	0.7216
	SIN3B	ND	ND		
OTHER COREPRESSORS	BCOR	25.02	25.23	1.156	0.2800
	TRIM28	33.40	33.51	1.084	0.4471
SWI/SNF COMPLEX	SMARCB1	28.93	28.84	-1.060	0.7515
	SMARCA4	30.84	30.96	1.087	0.2760
	SMARCA2	28.81	28.87	1.040	0.6311
	SMARCE1	29.20	28.96	-1.184	0.1656
	SMARCD1	26.24	26.23	-1.007	0.9556
	SMARCD2	29.74	29.74	1.002	0.9917
	SMARCC1	29.34	29.31	-1.021	0.8121
	SMARCC2	31.06	31.29	1.175	0.0838
	PBRM1	29.63	29.75	1.084	0.1620
	ASH2L	28.20	28.26	1.041	0.6871
OTHER NUCLEAR FACTORS	DNMT1	30.53	30.75	1.162	0.0490
	KDM3A	24.33	24.31	-1.010	0.8692
	KDM3B	29.06	29.20	1.097	0.1391
	KDM5A	22.21	23.44	2.336	0.2924
	CARM1	28.68	28.68	-1.002	0.9866
	YLPM1	29.45	29.47	1.012	0.9044
	MSH2	29.60	29.82	1.169	0.2531

Supplemental Table 3: RNA-seq expression of selected markers of erythroid maturation in sorted cells from CD34 erythroid cultures at days 8 (A), 11 (B), and 14 (C). Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values.

(A) CD34 Day 8	ID	Day 8 A-cell reads	Day 8 F-cell reads	Log_2 fold change	P value	P_{adj} value
Erythroid up	AHSP	5551.4	6963.2	0.323	0.00000038	0.00025
	SLC4A1	3151.6	4434.2	0.475	6.7E-10	0.0000011
	ANK1	16912.6	17062.8	0.046	0.78	1
	ALAS2	3775.6	4932.5	0.416	0.00000007	0.000057
	EPB42	2425.4	2824.5	0.227	0.011	0.43
	CA1	537.1	627.8	0.317	0.079	0.89
	SLC25A37	20232.6	22457.6	0.155	0.0044	0.26
	GYPA	3953.6	4202.4	0.087	0.23	1
	GATA1	3643.1	3906.4	0.091	0.26	1
	KLF1	9384.0	10803.7	0.198	0.0029	0.21
	FOXO3	3876.3	3957.6	0.030	0.69	1
	KEL	2991.5	3202.6	0.099	0.22	1
Erythroid down	KIT	7040.8	7518.7	0.106	0.45	1
	JUN	100.0	43.8	-1.158	0.0059	0.29
	MYB	8891.4	8012.5	-0.149	0.011	0.42
	MYC	5540.5	5630.3	0.043	0.59	1
	PCNA	11323.0	11837.6	0.063	0.27	1
	GATA2	694.4	606.8	-0.192	0.21	1
	CASP3	5298.6	5152.0	-0.041	0.54	1
	CD44	2314.3	2106.1	-0.073	0.49	1
	TFRC	241245.9	253735.5	0.077	0.31	1

(B) CD34 Day 11	ID	Day 11 A-cell reads	Day 11 F-cell reads	Log_2 fold change	P value	P_{adj} value
Erythroid up	AHSP	15442.3	14346.1	-0.102	0.03	0.91
	SLC4A1	95359.2	84721.1	-0.170	0.00054	0.075
	ANK1	20909.4	19616.5	-0.095	0.073	1
	ALAS2	37489.8	34445.2	-0.117	0.014	0.62
	EPB42	16077.9	14772.8	-0.121	0.0078	0.44
	CA1	8569.3	7708.5	-0.156	0.0019	0.17
	SLC25A37	120295.5	112451.7	-0.096	0.058	1
	GYPA	24479.0	22307.6	-0.134	0.0053	0.34
	GATA1	2978.7	2831.2	-0.075	0.31	1
	KLF1	13029.4	12744.8	-0.034	0.5	1
	FOXO3	8668.8	7439.8	-0.221	0.0000087	0.0028
	KEL	6452.8	6291.5	-0.035	0.54	1
Erythroid down	KIT	274.3	399.1	0.541	0.0065	0.39
	JUN	39.0	18.5	-1.181	0.083	1
	MYB	463.9	542.4	0.225	0.17	1
	MYC	328.7	398.5	0.276	0.15	1
	PCNA	6929.2	7216.2	0.059	0.25	1
	GATA2	60.1	45.1	-0.391	0.43	1
	CASP3	1140.6	1267.1	0.151	0.17	1
	CD44	753.8	716.8	-0.080	0.56	1
	TFRC	171292.0	173399.6	0.017	0.74	1

(C) CD34 Day 14	ID	Day 14 A-cell reads	Day 14 F-cell reads	Log_2 fold change	P value	P_{adj} value
Erythroid up	AHSP	17278.3	18326.6	0.0818	0.17	0.67
	SLC4A1	128133.1	104930.2	-0.2885	0.000017	0.0005
	ANK1	28482.6	21782.1	-0.385	4E-09	0.00000027
	ALAS2	55349.3	46224.1	-0.2585	0.000032	0.00084
	EPB42	23031.3	18919.2	-0.2831	0.0000036	0.00012
	CA1	7705.9	7347.6	-0.126	0.078	0.46
	SLC25A37	154942.5	130521.8	-0.2473	0.00028	0.0059
	GYPA	28897.4	27346.6	-0.0824	0.24	0.75
	GATA1	2819.7	2646.9	-0.0911	0.26	0.78
	KLF1	14657.0	13483.0	-0.1201	0.036	0.28
	FOXO3	20435.8	12794.5	-0.6736	1.6E-30	7.1E-28
	KEL	5806.5	5566.3	-0.0608	0.34	0.85
Erythroid down	KIT	129.8	232.1	0.7148	0.19	0.69
	JUN	58.9	34.5	-0.734	0.15	0.62
	MYB	197.9	257.5	0.368	0.11	0.55
	MYC	190.9	271.6	0.516	0.029	0.25
	PCNA	6576.5	7764.9	0.2393	0.000077	0.0018
	GATA2	59.0	54.8	-0.1036	0.82	0.99
	CASP3	1228.8	1328.1	0.1109	0.31	0.82
	CD44	1262.8	834.8	-0.6054	0.00000021	0.0000098
	TFRC	213988.0	212634.7	-0.0088	0.9	1

Supplemental Table 4: RNA-seq expression of selected transcripts in day 8 sorted CD34$^+$ erythroid cultures. Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values. ND: not detected.

CD34$^+$ DAY 8	ID	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
HBF REGULATORS	BCL11A	4997.9	4077.6	-0.258	0.17	1
	ZBTB7A	696.0	578.7	-0.2192	0.17	1
	MYB	8891.4	8012.5	-0.1493	0.011	0.42
	EIF2AK1	5522.5	5566.7	0.0118	0.87	1
	ATF4	6967.4	7481.4	0.0993	0.11	0.97
	SPOP	1893.2	1921.9	0.0223	0.82	1
	KLF1	9384.0	10803.7	0.198	0.0029	0.21
	SOX6	1244.4	1088.4	-0.1905	0.12	0.97
	LIN28B	ND	ND			
	IGF2BP1	ND	ND			
	HMGA2	836.7	648.7	-0.3513	0.015	0.48
NURD COMPLEX	CHD3	4560.1	4723.0	0.0515	0.47	1
	CHD4	14752.8	15364.5	0.0513	0.42	1
	GATAD2A	1968.5	1762.0	-0.0891	0.44	1
	GATAD2B	1136.0	948.6	-0.1956	0.15	1
	HDAC1	3165.0	3576.8	0.1791	0.039	0.69
	HDAC2	3972.4	3439.9	-0.1744	0.042	0.72
	MBD2	2013.4	1824.3	-0.1403	0.15	1
	MBD3	2758.3	3138.9	0.1751	0.043	0.72
	MTA1	834.2	766.0	-0.0716	0.63	1
	MTA2	5790.1	6001.5	0.0467	0.47	1
	MTA3	242.3	237.2	-0.0401	0.87	1
	RBBP4	9750.8	8895.9	-0.1256	0.04	0.69
	RBBP7	8625.4	9088.6	0.0766	0.2	1
TRANSCRIPTION FACTORS	ZFPM1	777.2	771.8	0.0788	0.63	1
	RUNX1	2053.3	1740.6	-0.2095	0.044	0.73
	IKZF1	6207.7	5952.1	-0.0469	0.52	1
LSD1/COREST	KDM1A	8095.9	8350.4	0.0447	0.45	1
	RCOR1	2318.3	2109.5	-0.1227	0.2	1
NCOR/SMRT COMPLEX	NCOR1	8605.6	7997.1	-0.1065	0.07	0.85
	NCOR2	638.5	524.5	-0.1521	0.41	1
	TBL1XR1	10101.4	8767.7	-0.2085	0.0013	0.13
	TBL1X	1323.4	1205.9	-0.1269	0.27	1
	CORO2A	174.2	184.1	0.0747	0.78	1
	ZBTB33	2232.8	2061.9	-0.1149	0.21	1
SIN3 COMPLEX	SIN3A	3130.8	2916.1	-0.0968	0.23	1
	SIN3B	291.4	265.1	-0.1147	0.61	1
OTHER COREPRESSORS	BCOR	1658.3	1471.9	-0.1653	0.13	1
	TRIM28	6617.4	6750.5	0.0667	0.34	1
SWI/SNF COMPLEX	SMARCB1	1290.7	1257.6	-0.0441	0.7	1
	SMARCA4	5362.7	5029.4	-0.0874	0.19	1
	SMARCA2	6414.2	5525.9	-0.2145	0.00079	0.097
	SMARCE1	4940.4	4679.7	-0.0788	0.26	1
	SMARCD1	685.4	619.3	-0.064	0.7	1
	SMARCD2	3140.0	3408.5	0.116	0.16	1
	SMARCC1	6534.5	6247.1	-0.0593	0.36	1
	SMARCC2	2709.4	2313.6	-0.2216	0.01	0.41
	PBRM1	6102.9	6008.0	-0.0226	0.72	1
	ASH2L	2852.4	3081.7	0.111	0.19	1
OTHER NUCLEAR FACTORS	DNMT1	10869.7	10651.8	-0.0375	0.55	1
	KDM3A	4057.7	4089.9	0.0215	0.77	1
	KDM3B	4755.0	4857.0	0.0359	0.61	1
	KDM5A	6349.3	6387.1	0.0119	0.85	1
	CARM1	939.4	896.6	-0.0848	0.52	1
	YLPM1	4548.7	4475.8	-0.0139	0.85	1
	MSH2	6743.1	7122.6	0.0793	0.2	1

Supplemental Table 5: RNA-seq expression of selected transcripts in day 11 sorted $CD34^+$ erythroid cultures. Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values. ND: not detected.

$CD34^+$ DAY 11	ID	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
HBF REGULATORS	BCL11A	5385.4	4993.0	-0.1065	0.063	1
	ZBTB7A	471.3	428.2	-0.1497	0.4	1
	MYB	463.9	542.4	0.2245	0.17	1
	EIF2AK1	9583.5	8792.6	-0.123	0.018	0.7
	ATF4	5544.9	5306.3	-0.0634	0.25	1
	SPOP	1201.0	1171.2	-0.0364	0.75	1
	KLF1	13029.4	12744.8	-0.0336	0.5	1
	SOX6	13673.9	12282.6	-0.1548	0.00074	0.094
	LIN28B	ND	ND			
	IGF2BP1	ND	ND			
	HMGA2	561.2	519.3	-0.1069	0.51	1
NURD COMPLEX	CHD3	1075.2	1062.3	-0.0145	0.9	1
	CHD4	3559.8	3828.1	0.1037	0.12	1
	GATAD2A	352.3	323.7	-0.1215	0.54	1
	GATAD2B	298.9	338.4	0.1697	0.42	1
	HDAC1	2269.7	2262.7	-0.0032	0.97	1
	HDAC2	1080.3	1017.5	-0.078	0.53	1
	MBD2	840.9	866.7	0.0405	0.76	1
	MBD3	842.3	912.1	0.115	0.36	1
	MTA1	168.0	184.5	0.1387	0.61	1
	MTA2	2198.1	2396.3	0.1249	0.13	1
	MTA3	79.2	78.2	-0.0152	0.97	1
	RBBP4	3755.3	3659.6	-0.0368	0.59	1
	RBBP7	4024.7	3880.0	-0.0526	0.41	1
TRANSCRIPTION FACTORS	ZFPM1	317.4	323.1	0.018	0.93	1
	RUNX1	185.3	179.6	-0.0463	0.86	1
	IKZF1	4603.4	4523.3	-0.0264	0.67	1
LSD1/COREST	KDM1A	2504.2	2622.0	0.0663	0.39	1
	RCOR1	916.9	924.4	0.0014	0.99	1
NCOR/SMRT COMPLEX	NCOR1	5013.5	4922.5	-0.026	0.67	1
	NCOR2	176.7	155.8	-0.1918	0.49	1
	TBL1XR1	3046.2	2825.4	-0.1082	0.14	1
	TBL1X	270.8	293.3	0.1049	0.63	1
	CORO2A	14.2	29.8	1.1055	0.14	1
	ZBTB33	1084.7	1150.0	0.0907	0.43	1
SIN3 COMPLEX	SIN3A	1349.0	1286.7	-0.0706	0.5	1
	SIN3B	48.9	72.0	0.4994	0.3	1
OTHER COREPRESSORS	BCOR	1197.6	1186.6	-0.013	0.91	1
	TRIM28	1129.4	1305.3	0.2085	0.056	1
SWI/SNF COMPLEX	SMARCB1	400.6	448.8	0.1646	0.36	1
	SMARCA4	1360.0	1404.5	0.0488	0.64	1
	SMARCA2	1699.8	1742.3	0.0335	0.72	1
	SMARCE1	2293.9	2256.3	-0.0247	0.76	1
	SMARCD1	169.4	144.6	-0.2213	0.44	1
	SMARCD2	1308.2	1332.2	0.0268	0.8	1
	SMARCC1	400.6	448.8	0.1646	0.36	1
	SMARCC2	1027.4	1059.9	0.0448	0.7	1
	PBRM1	6342.4	6126.3	-0.0505	0.35	1
	ASH2L	2652.3	2573.9	-0.0433	0.58	1
OTHER NUCLEAR FACTORS	DNMT1	5302.6	5527.6	0.0605	0.3	1
	KDM3A	6534.1	6022.3	-0.1173	0.027	0.86
	KDM3B	2651.0	2631.0	-0.0109	0.89	1
	KDM5A	5105.5	4949.0	-0.0442	0.44	1
	CARM1	416.5	385.9	-0.128	0.49	1
	YLPM1	1289.3	1395.7	0.1136	0.3	1
	MSH2	2471.4	2572.3	0.0577	0.46	1

Supplemental Table 6: RNA-seq expression of selected transcripts in day 14 sorted CD34$^+$ erythroid cultures. Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values. ND: not detected.

CD34$^+$ DAY 14	ID	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
HBF REGULATORS	BCL11A	6409.5	5349.4	-0.2664	0.00011	0.0026
	ZBTB7A	652.7	504.5	-0.3733	0.024	0.22
	MYB	197.9	257.5	0.368	0.11	0.55
	EIF2AK1	13172.1	10367.2	-0.3451	2.9E-09	0.0000002
	ATF4	5074.7	5379.5	0.0841	0.21	0.72
	SPOP	1333.9	1257.1	-0.0856	0.42	0.89
	KLF1	14657.0	13483.0	-0.1201	0.036	0.28
	SOX6	20408.0	15903.2	-0.36	3.2E-09	0.00000022
	LIN28B	ND	ND			
	IGF2BP1	ND	ND			
	HMGA2	514.0	474.6	-0.1149	0.49	0.92
NURD COMPLEX	CHD3	2081.1	2142.3	0.0429	0.63	0.96
	CHD4	49.9	56.5	0.1661	0.72	0.97
	GATAD2A	434.4	393.7	-0.142	0.41	0.89
	GATAD2B	418.5	312.5	-0.4201	0.023	0.21
	HDAC1	2270.1	2373.3	0.0645	0.45	0.91
	HDAC2	1211.0	1057.9	-0.1941	0.083	0.47
	MBD2	820.5	806.8	-0.0263	0.84	0.99
	MBD3	941.3	989.1	0.0723	0.55	0.94
	MTA1	137.3	155.6	0.1821	0.53	0.93
	MTA2	2396.2	2722.8	0.1835	0.027	0.23
	MTA3	41.2	56.0	0.4079	0.4	0.89
	RBBP4	3818.7	3758.7	-0.0242	0.75	0.98
	RBBP7	4141.0	4313.4	0.0576	0.42	0.89
TRANSCRIPTION FACTORS	ZFPM1	329.0	297.5	-0.1486	0.74	0.98
	RUNX1	134.3	123.3	-0.1231	0.68	0.97
	IKZF1	5811.1	5323.9	-0.128	0.062	0.4
LSD1/COREST	KDM1A	134.3	123.3	-0.1231	0.68	0.97
	RCOR1	1006.8	934.8	-0.1069	0.37	0.87
NCOR/SMRT COMPLEX	NCOR1	5526.6	5001.4	-0.1446	0.026	0.23
	NCOR2	191.1	146.9	-0.3654	0.18	0.67
	TBL1XR1	4601.3	3444.1	-0.4143	0.00000014	0.0000067
	TBL1X	299.1	259.7	-0.2036	0.33	0.84
	CORO2A	8.2	13.2	0.7557	0.46	0.91
	ZBTB33	1099.8	1238.3	0.1719	0.15	0.62
SIN3 COMPLEX	SIN3A	1450.5	1282.7	-0.1774	0.087	0.48
	SIN3B	38.4	44.2	0.2012	0.7	0.97
OTHER COREPRESSORS	BCOR	1133.8	996.4	-0.1898	0.1	0.52
	TRIM28	1106.7	1186.6	0.1013	0.37	0.87
SWI/SNF COMPLEX	SMARCB1	389.1	409.3	0.0737	0.68	0.97
	SMARCA4	1240.7	1280.7	0.0482	0.65	0.96
	SMARCA2	1596.3	1361.0	-0.2314	0.022	0.21
	SMARCE1	2441.3	2422.3	-0.0113	0.89	1
	SMARCD1	188.0	162.4	-0.211	0.42	0.89
	SMARCD2	1266.1	1445.5	0.1881	0.11	0.54
	SMARCC1	1101.2	1070.3	-0.0411	0.73	0.97
	SMARCC2	1543.1	1271.8	-0.2793	0.0065	0.083
	PBRM1	7194.9	6794.7	-0.083	0.19	0.69
	ASH2L	3054.9	3101.8	0.0241	0.76	0.98
OTHER NUCLEAR FACTORS	DNMT1	5071.7	5431.0	0.099	0.13	0.58
	KDM3A	2744.5	2365.1	-0.2147	0.0088	0.1
	KDM3B	2981.2	2447.6	-0.2847	0.00052	0.01
	KDM5A	5537.2	5114.6	-0.109	0.11	0.55
	CARM1	518.8	411.8	-0.3316	0.045	0.33
	YLPM1	1306.4	1304.7	0.0061	0.96	1
	MSH2	2381.3	2619.9	0.1401	0.1	0.52

Supplemental Table 7: RNA-seq expression of differentially expressed transcripts in day 11 sorted $CD34^+$ erythroid cultures treated with hydroxyurea. Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values. Significance is defined by greater than 1.5-fold change and DESeq P_{adj} <0.05.

$CD34^+$ DAY 11 HU	ID	ENSEMBL	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
F-CELL ENRICHED	AC104389.5	ENSG00000284931	18.4	2717.7	7.223	3.3E-121	3.7E-117
	HBG2	ENSG00000196565	30.2	4318.4	6.912	6E-35	3.3E-31
	HBG1	ENSG00000213934	23.5	2703.2	6.844	4.7E-149	1E-144
	BGLT3	ENSG00000260629	12.3	1099.8	6.509	3.2E-69	2.4E-65
	AC022400.3	ENSG00000242288	5.6	74.0	4.167	0.0000017	0.00081
	AC135983.3	ENSG00000254912	9.7	62.2	2.656	0.00015	0.035
	COL6A5	ENSG00000172752	125.8	282.4	1.249	0.000059	0.016
	NPIPB12	ENSG00000169203	138.8	299.5	1.102	0.00019	0.041
	MT-ND1	ENSG00000198888	2488.3	4645.1	0.853	5.7E-12	7.4E-09
	CHD7	ENSG00000171316	1473.0	2527.4	0.779	2.8E-09	0.000002
	MT-ND5	ENSG00000198786	6940.9	12101.1	0.762	2.2E-12	3E-09
	PRKCB	ENSG00000166501	905.2	1505.6	0.747	0.0000018	0.00085
	KIT	ENSG00000157404	963.2	1618.4	0.747	0.0000011	0.00054
	AHCY	ENSG00000101444	589.6	970.9	0.719	0.000054	0.015
	MT-ND4	ENSG00000198886	12847.1	20479.6	0.655	7.5E-15	1.3E-11
	MT-ND2	ENSG00000198763	4720.2	7448.6	0.633	9.8E-10	0.0000008
	LAPTM5	ENSG00000162511	1015.5	1562.8	0.622	0.000035	0.01
A-CELL ENRICHED	SNORA28	ENSG00000272533	96.3	20.96	-2.254	0.000055	0.015
	MIAT	ENSG00000225783	2247.1	859.76	-1.380	8.4E-22	2.3E-18
	PTPRF	ENSG00000142949	358.1	140.73	-1.349	0.0000011	0.00054
	TRIB3	ENSG00000101255	635.2	249.21	-1.349	1.4E-09	0.0000011
	C5	ENSG00000106804	738.9	304.06	-1.283	2.3E-09	0.0000017
	TMCC2	ENSG00000133069	1040.9	461.27	-1.172	1.8E-10	0.00000017
	RGPD6	ENSG00000183054	622.4	297.54	-1.146	0.000009	0.0034
	TCP11L2	ENSG00000166046	2037.6	945.88	-1.106	1.6E-14	2.6E-11
	RNU12	ENSG00000276027	6950.03	3372.91	-1.043	5.5E-26	2E-22
	IFIT1B	ENSG00000204010	1798.00	905.42	-0.995	3.6E-11	3.9E-08
	RNU5B-1	ENSG00000200156	16545.6	8590.36	-0.952	2.6E-31	1.1E-27
	SNORA27	ENSG00000207051	1603.9	889.74	-0.840	6.5E-08	0.00004
	RNU5A-1	ENSG00000199568	14896.0	8555.48	-0.820	2.6E-17	5.3E-14
	SNORD15B	ENSG00000207445	19697.4	11332.50	-0.814	1.7E-19	4.3E-16
	SCARNA1	ENSG00000252947	1291.2	736.58	-0.806	0.0000008	0.00041
	RNU5E-1	ENSG00000199347	13557.4	7773.62	-0.803	2.8E-22	8.8E-19
	RNU5E-6P	ENSG00000202444	6716.3	3869.43	-0.787	4.5E-10	0.0000004
	CDKN2D	ENSG00000129355	1227.8	723.70	-0.785	0.0000094	0.0035
	SNORA74B	ENSG00000212402	2302	1345.08	-0.781	7.4E-09	0.0000051
	CEBPB	ENSG00000172216	864.2	510.84	-0.779	0.000049	0.013
	PLCL2	ENSG00000154822	959.5	567.36	-0.761	0.000024	0.0075
	AL360012.1	ENSG00000270103	20410.4	12011.85	-0.756	1.4E-17	3.2E-14
	FAM129A	ENSG00000135842	4289.8	2560.48	-0.742	2.4E-11	2.8E-08
	GABARAPL1	ENSG00000139112	1217.1	747.16	-0.704	0.000019	0.0062
	SNORA81	ENSG00000221420	12176.2	7561.87	-0.698	3.9E-11	4.1E-08
	SNORA23	ENSG00000201998	1627.2	992.69	-0.698	0.0000058	0.0023
	AC006064.5	ENSG00000276232	12655.2	8070.7	-0.695	7.6E-11	7.7E-08
	SNORA73B	ENSG00000200087	57808.3	36358.8	-0.673	8E-17	1.5E-13
	SCARNA16	ENSG00000275143	3837.4	2521.93	-0.667	0.0000037	0.0016
	SNORD3A	ENSG00000263934	100308	63760.75	-0.651	7.6E-12	9.3E-09
	SNORA74A	ENSG00000200959	3717.6	2319.29	-0.629	0.0000006	0.00033

	FBXO30	ENSG00000118496	3936.2	2612.03	-0.606	0.0000081	0.0031
	SNORA22	ENSG00000206634	7746.5	5074.82	-0.604	3.5E-08	0.000023
	SNORD17	ENSG00000212232	27732.2	18770.6	-0.594	5.7E-10	0.00000049
	PHLPP2	ENSG00000040199	3811.6	2529.13	-0.587	0.0000003	0.00015
	UBE2H	ENSG00000186591	4475.9	2981.36	-0.585	6.5E-08	0.00004

Supplemental Table 8: RNA-seq expression of selected transcripts in day 11 sorted $CD34^+$ erythroid cultures treated with hydroxyurea. Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values. ND: not detected.

$CD34^+$ DAY 11 HYDROXYUREA	ID	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
HBF REGULATORS	BCL11A	8227.1	8792.2	0.098	0.26	1
	ZBTB7A	1426.1	1571.5	0.1216	0.43	1
	MYB	1015.9	1323.9	0.3718	0.021	0.79
	EIF2AK1	18181.4	16107.6	-0.1759	0.028	0.88
	ATF4	9975.2	9339.8	-0.0964	0.26	1
	SPOP	2210.0	2103.6	-0.0727	0.58	1
	KLF1	25003.8	24132.3	-0.0532	0.55	1
	SOX6	24383.8	20290.2	-0.2646	0.00059	0.095
	LIN28B	ND	ND			
	IGF2BP1	ND	ND			
	HMGA2	1169.4	978.3	-0.2404	0.15	1
NURD COMPLEX	CHD3	3555.3	3785.1	0.0908	0.4	1
	CHD4	10843.1	11084.9	0.0322	0.7	1
	GATAD2A	1151.3	1315.5	0.1897	0.22	1
	GATAD2B	772.7	946.1	0.2803	0.12	1
	HDAC1	4046.6	3877.2	-0.0633	0.56	1
	HDAC2	3389.0	3402.1	0.006	0.96	1
	MBD2	1641.0	1716.4	0.0712	0.63	1
	MBD3	2414.0	2405.7	-0.0017	0.99	1
	MTA1	417.2	466.8	0.1466	0.51	1
	MTA2	7096.2	7299.6	0.0415	0.65	1
	MTA3	160.3	237.6	0.5703	0.057	1
	RBBP4	8790.4	8741.8	-0.0072	0.94	1
	RBBP7	9172.9	8644.8	-0.0865	0.32	1
TRANSCRIPTION FACTORS	ZFPM1	869.0	1078.0	0.3025	0.078	1
	RUNX1	377.6	434.1	0.1964	0.39	1
	IKZF1	10107.1	9995.2	-0.0179	0.83	1
LSD1/COREST	KDM1A	5120.0	5284.9	0.0457	0.64	1
	RCOR1	1411.2	1384.1	-0.0269	0.86	1
NCOR/SMRT COMPLEX	NCOR1	11293.8	11290.1	-0.0085	0.92	1
	NCOR2	380.2	475.3	0.3168	0.15	1
	TBL1XR1	5982.6	5502.2	-0.1367	0.72	1
	TBL1X	751.1	877.9	0.2115	0.26	1
	CORO2A	34.3	62.6	0.8601	0.13	1
	ZBTB33	2213.7	2236.0	0.0144	0.91	1
SIN3 COMPLEX	SIN3A	3197.6	3318.5	0.0527	0.64	1
	SIN3B	109.7	114.3	0.0605	0.87	1
OTHER COREPRESSORS	BCOR	2570.6	2437.5	-0.0997	0.46	1
	TRIM28	2992.8	3419.8	0.1966	0.091	1
SWI/SNF COMPLEX	SMARCB1	1044.4	1094.2	0.0715	0.66	1
	SMARCA4	4007.0	4472.9	0.1548	0.19	1
	SMARCA2	3808.8	3826.0	0.0066	0.95	1
	SMARCE1	6129.0	5853.8	-0.0636	0.51	1
	SMARCD1	381.7	377.2	-0.0193	0.93	1
	SMARCD2	3274.5	3256.4	-0.007	0.95	1
	SMARCC1	2477.3	2831.2	0.1906	0.12	1
	SMARCC2	3668.8	3831.0	0.0596	0.58	1
	PBRM1	11533.5	10217.7	-0.1708	0.05	1
	ASH2L	5424.7	5043.1	-0.1053	0.28	1
OTHER NUCLEAR FACTORS	DNMT1	17162.4	17319.5	0.0026	0.98	1
	KDM3A	5238.2	4674.4	-0.1647	0.098	1
	KDM3B	6296.1	6379.7	0.019	0.84	1
	KDM5A	8610.3	8114.9	-0.0854	0.34	1
	CARM1	848.2	766.9	-0.1601	0.37	1
	YLPM1	3712.0	4024.4	0.1126	0.3	1
	MSH2	4436.7	4285.5	-0.051	0.63	1

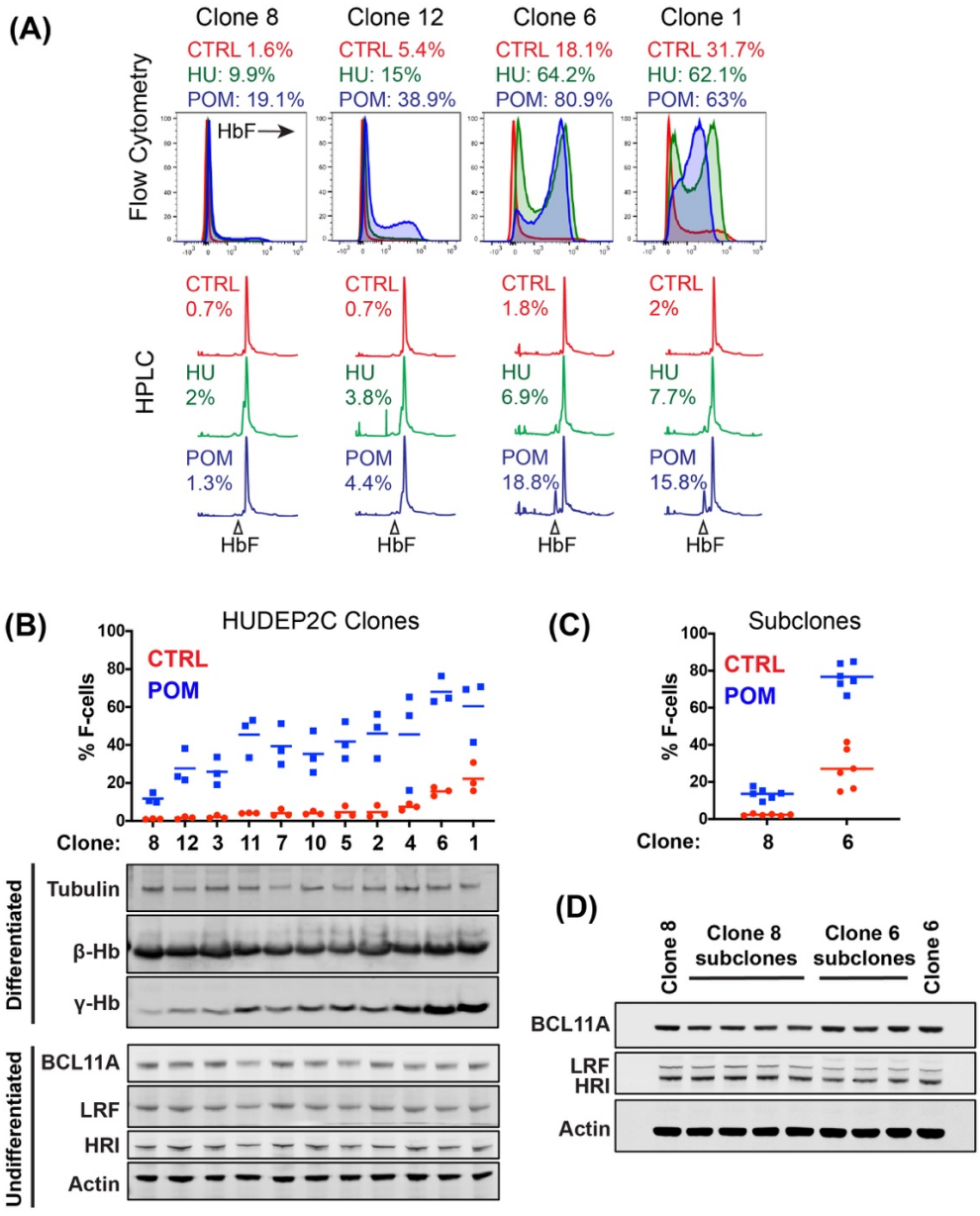
Supplemental Table 9: RNA-seq expression of selected transcripts in day 11 sorted $CD34^+$ erythroid cultures treated with pomalidomide. Data are shown as average normalized A-cell and F-cell reads for 2 replicates, log_2 fold change and DESeq adjusted p-values. ND: not detected.

$CD34^+$ DAY 11 POMALIDOMIDE	ID	A-cell reads	F-cell reads	Log_2 fold change	P value	P_{adj} value
HBF REGULATORS	BCL11A	7705.4	6987.2	-0.1717	0.047	0.33
	ZBTB7A	1390.6	1299.4	-0.1341	0.31	0.87
	MYB	1145.3	1767.6	0.6272	9.4E-08	0.0000066
	EIF2AK1	16935.7	13904.5	-0.2889	0.000043	0.0015
	ATF4	11488.6	10690.4	-0.105	0.12	0.57
	SPOP	2063.2	1816.1	-0.177	0.11	0.54
	KLF1	20747.6	16598.0	-0.3641	0.058	0.37
	SOX6	22869.1	14768.9	-0.633	9.4E-20	3.9E-17
	LIN28B	ND	ND			
	IGF2BP1	ND	ND			
	HMGA2	833.2	747.3	-0.1892	0.24	0.79
NURD COMPLEX	CHD3	1098.5	1128.6	0.04	0.77	1
	CHD4	9919.5	11862.1	0.2588	0.00013	0.0038
	GATAD2A	1131.4	1256.6	0.145	0.26	0.81
	GATAD2B	828.3	900.9	0.1258	0.41	0.95
	HDAC1	3980.2	3624.5	-0.1374	0.11	0.54
	HDAC2	3832.3	3738.7	-0.0361	0.68	1
	MBD2	2335.9	1986.0	-0.2311	0.021	0.19
	MBD3	1893.6	1970.0	0.0272	0.81	1
	MTA1	445.5	536.5	0.3049	0.14	0.6
	MTA2	6086.6	6379.2	0.0568	0.46	0.98
	MTA3	233.0	315.3	0.437	0.084	0.46
	RBBP4	9487.7	9272.2	-0.0354	0.61	1
	RBBP7	7660.8	7482.8	-0.0333	0.64	1
TRANSCRIPTION FACTORS	ZFPM1	648.9	837.1	0.4051	0.017	0.16
	RUNX1	491.3	754.5	0.5949	0.0043	0.061
	IKZF1	17281.2	17942.2	0.0536	0.51	0.99
LSD1/COREST	KDM1A	5206.1	5306.8	0.0244	0.76	1
	RCOR1	1768.9	1702.9	-0.0623	0.58	1
NCOR/SMRT COMPLEX	NCOR1	13322.6	12553.3	-0.0841	0.22	0.75
	NCOR2	325.4	447.4	0.467	0.032	0.25
	TBL1XR1	8088.8	6964.9	-0.2216	0.0017	0.029
	TBL1X	904.9	1140.9	0.326	0.018	0.17
	CORO2A	29.1	58.2	0.9132	0.14	0.6
	ZBTB33	2682.8	2460.7	-0.1222	0.19	0.71
SIN3 COMPLEX	SIN3A	3586.4	3367.6	-0.0906	0.28	0.84
	SIN3B	117.7	166.0	0.4221	0.25	0.8
OTHER COREPRESSORS	BCOR	2915.2	2540.9	-0.1966	0.033	0.26
	TRIM28	3116.0	3739.8	0.2632	0.0019	0.032
SWI/SNF COMPLEX	SMARCB1	979.9	1034.9	0.0837	0.55	1
	SMARCA4	4112.4	4668.6	0.1974	0.021	0.19
	SMARCA2	3774.2	3924.7	0.0564	0.53	1
	SMARCE1	5890.7	5678.4	-0.0556	0.46	0.98
	SMARCD1	447.8	542.8	0.2719	0.16	0.65
	SMARCD2	2838.1	2820.0	-0.0231	0.81	1
	SMARCC1	2688.7	3341.3	0.3102	0.00093	0.018
	SMARCC2	4193.5	3962.3	-0.0792	0.33	0.88
	PBRM1	12006.3	9512.7	-0.3328	0.0000012	0.000063
	ASH2L	5135.1	4799.1	-0.0963	0.21	0.74
OTHER NUCLEAR FACTORS	DNMT1	14704.1	14357.4	-0.0355	0.6	1
	KDM3A	5560.9	4490.9	-0.3014	0.00011	0.0032
	KDM3B	5804.4	5956.7	0.0375	0.61	1
	KDM5A	9246.1	7586.6	-0.2833	0.00016	0.0044
	CARM1	793.5	675.1	-0.2919	0.11	0.54
	YLPM1	3955.6	4343.2	0.1349	0.091	0.49
	MSH2	5032.3	4692.6	-0.1272	0.14	0.6

SUPPLEMENTAL FIGURES

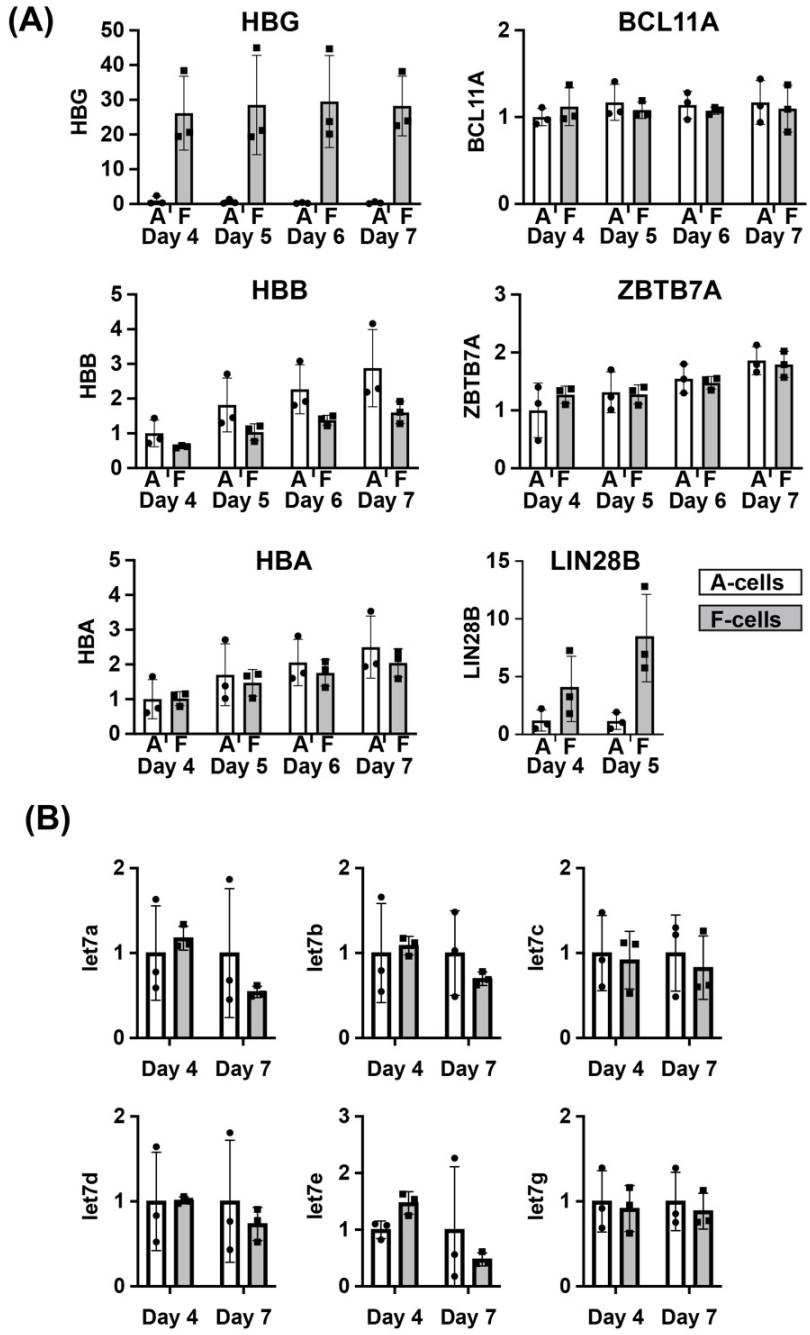
Supplemental Figure 1: Generation and analysis of HUDEP2C single-cell derived clonal lines. We generated a total of 11 single-cell derived clones from HUDEP2 cells stably expressing spCas9. When differentiated, the clonal lines had different HbF expression and differential responses to HbF inducers hydroxyurea and pomalidomide. **(A)** HbF flow cytometry and ion-exchange HPLC of the lowest and highest HbF-expressing clones after differentiation in the presence of vehicle (CTRL), hydroxyurea (HU), or pomalidomide (POM). Percentage of HbF-positive cells is indicated for flow cytometry and percentage of HbF is indicated for HPLC plots. **(B)** Analysis of HbF regulator expression in HUDEP2C clones. Top: percentage of F-cells following differentiation in the presence of vehicle (CTRL) or pomalidomide (POM) for each of 11 tested clones. Data points are shown for 3 separate experiments and mean for each clone. Bottom: western blot analysis of clones after and before differentiation. **(C)** The highest and lowest HbF clones were subsequently subcloned into 6 single-cell derived subclone lines, which were then analyzed for percentage of F-cells following differentiation in the presence of vehicle (CTRL) or pomalidomide (POM). Data points are individual subclones. **(D)** Western blot analysis of key HbF regulators in undifferentiated subclone cell lines, with original clones included for comparison. Subclones of clone 6 were used for HUDEP2 experiments.

Figure S1

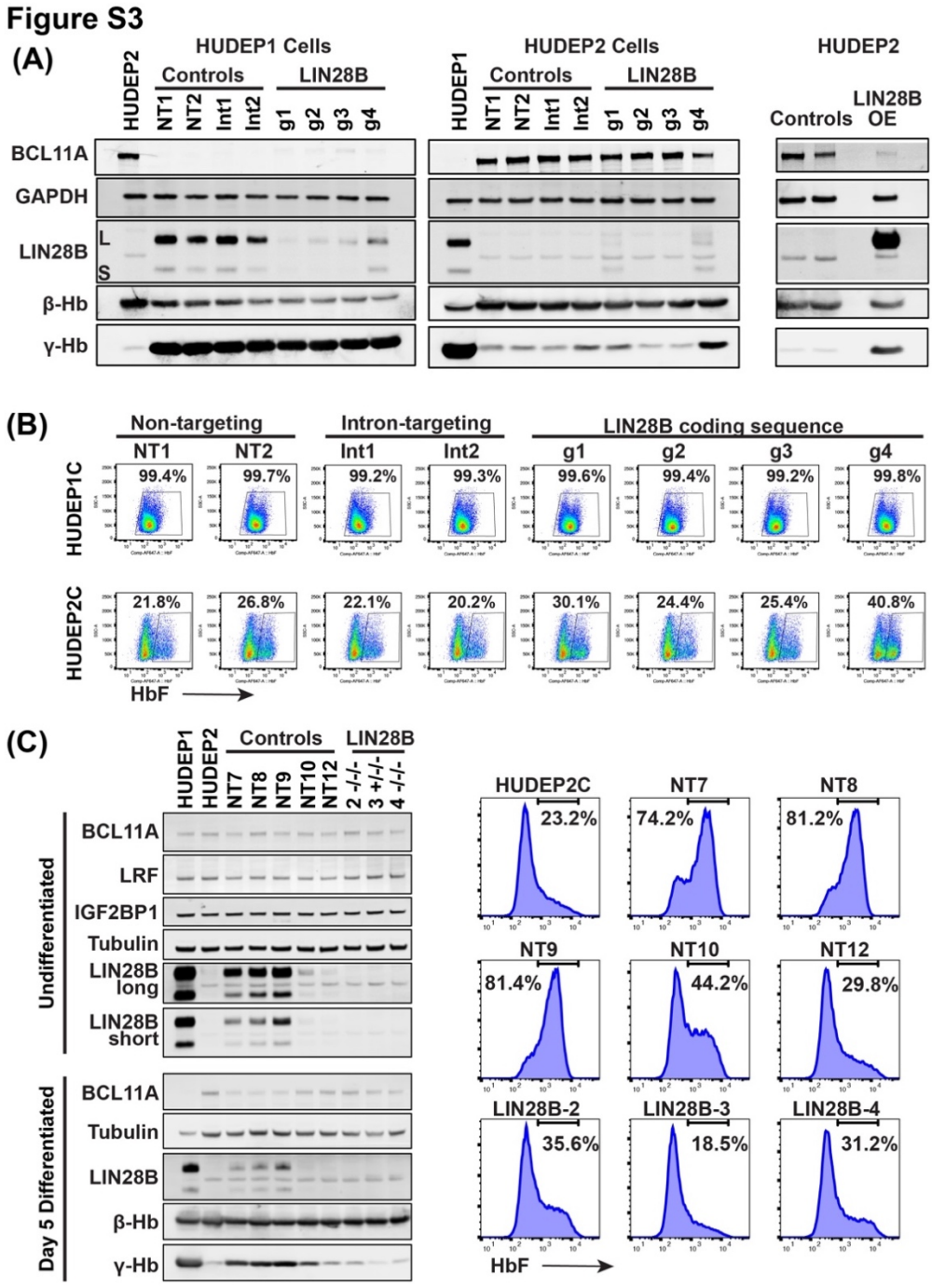


Supplemental Figure 2: (A) HUDEP2 F-cell sort validation. Three independent HUDEP2 single-cell derived clones (#1, 4, and 6, see supplemental figure 1) were used for a time-course F-cell sorting experiment, with matched F- and A-cells sorted daily on days 4-7 of differentiation. RT-PCR for globin transcript and HbF regulators was done for each time point. Data are normalized to average of 3 housekeeping gene transcripts and shown relative to day 4 A-cells. Plots show individual clone data points, mean, and SD. **(B)** Analysis of let7 miRNA family members in sorted HUDEP2 F- and A-cells at days 4 and 7 of differentiation. Data are normalized to 4 stable miRNA transcripts and shown relative to A-cells.

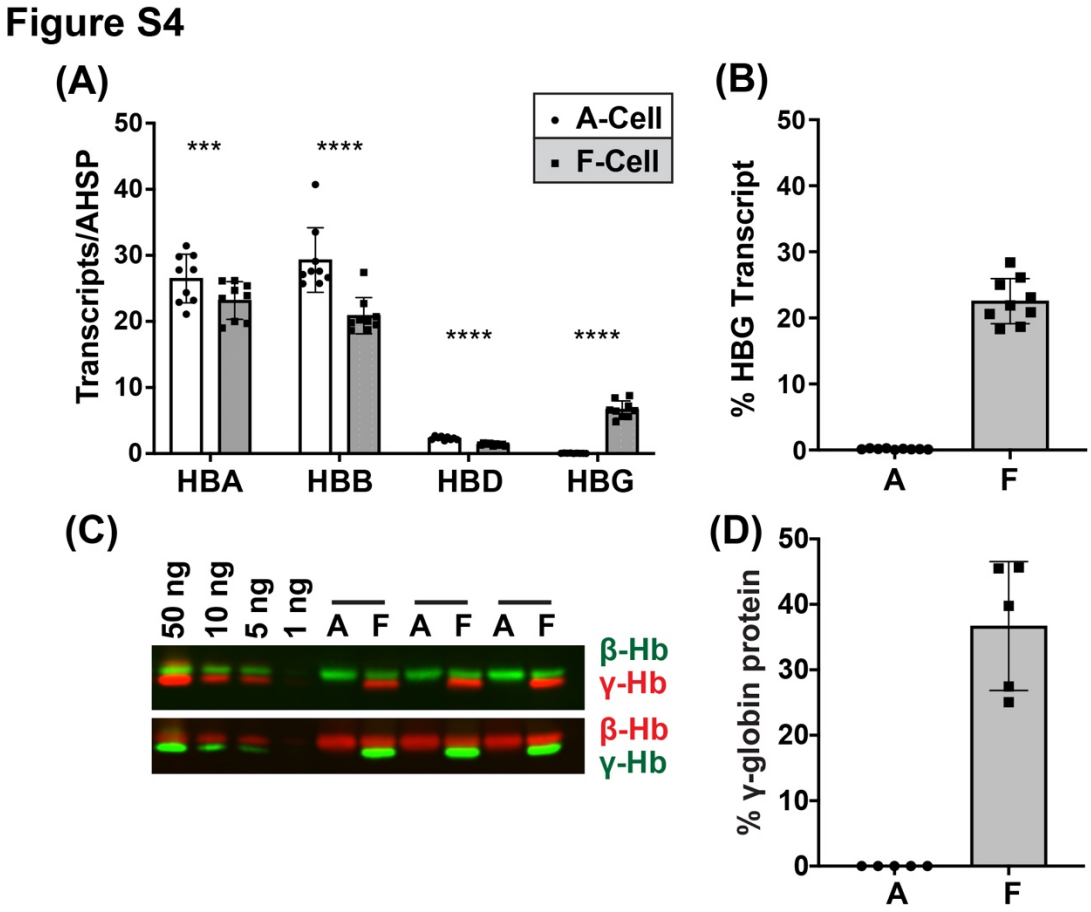
Figure S2



Supplemental Figure 3: LIN28B manipulation in HUDEP cells. HUDEP1 and HUDEP2 cells stably expressing spCas9 were infected with LRG lentivirus expressing non-targeting gRNAs (NT1 and NT2), gRNAs targeting the introns of AHSP and LIN28B (Int1 and Int2), and four gRNAs targeting exons 2 and 3 of LIN28B. In addition, we infected HUDEP2 cells with a retrovirus driving LIN28B expression. After sorting and one week of expansion, cells were differentiated for 7 days and analyzed by western blot (**A**) and flow cytometry (**B**). Of note, gRNA1 and 4 tended to lead to higher LIN28B expression in HUDEP2 cells despite good knockdown efficiency in HUDEP1 cells. HUDEP2C cells expressing either non-targeting gRNA or LIN28B gRNA were single-cell cloned and expanded prior to analysis (**C**). Single-cell clones tended to drift to high LIN28B and HbF expression. LIN28B is triplicated in HUDEP2 cells. Several clones with mutations predicted to completely disrupt LIN28B expression in all three alleles continued to express γ-globin and form F-cells.

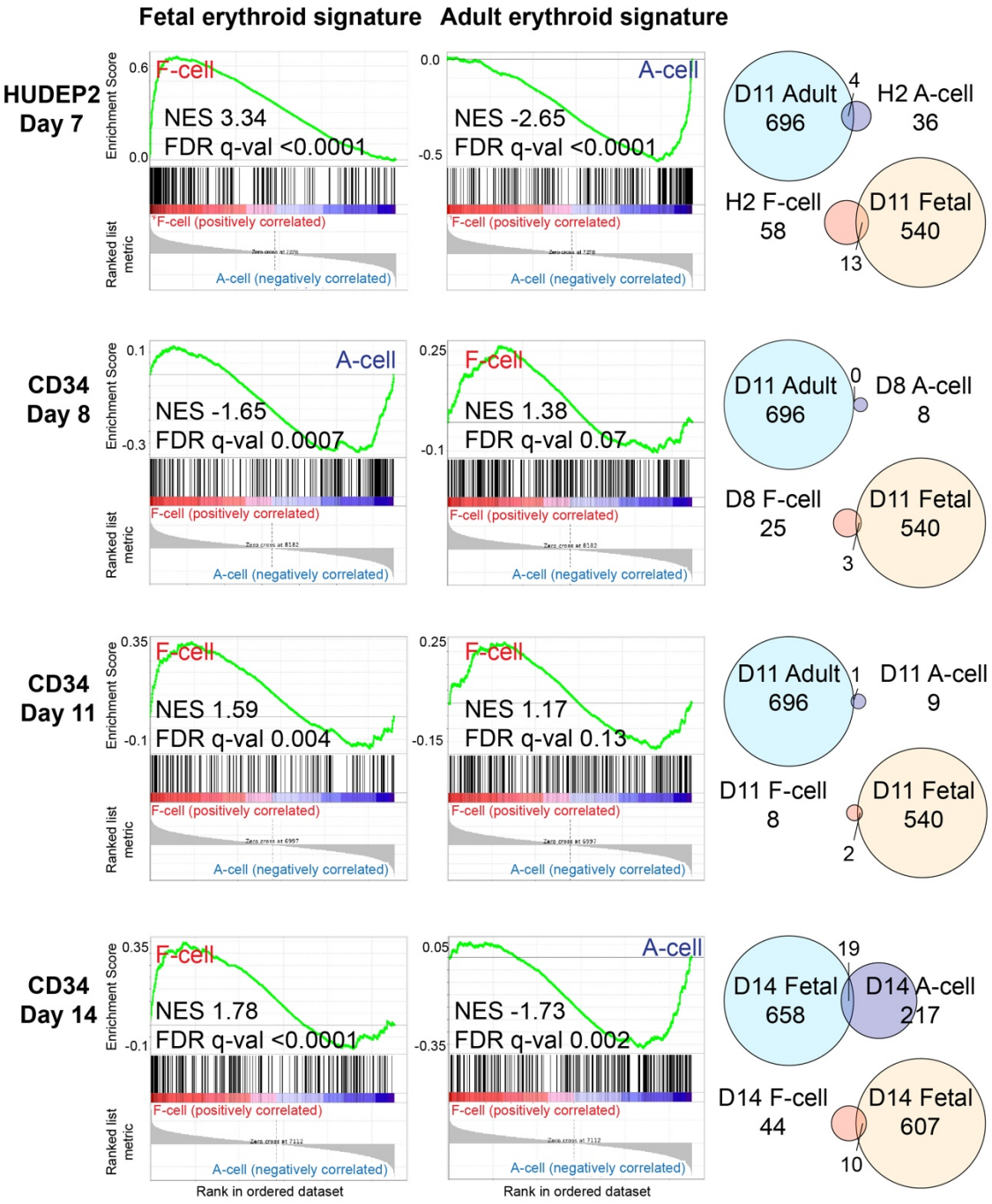


Supplemental Figure 4: Analysis of F-cell globin content. (A) Droplet digital RT-PCR analysis of transcripts for HBA1/2, HBB, HBD, and HBG1/2 from day 11 sorted A- and F-cells in 9 separate CD34 donors across four experiments. Data are shown relative to multiplexed AHSP transcripts as mean ± SD. *** p<0.001 versus A-cells, **** p<0.001 versus A-cells by Student's t-test. (B) Percentage of HBG transcripts (relative to HBB+HBD+HBG) in day 11 erythroblasts. (C) Semi-quantitative western blotting for γ-globin and β-globin in sorted F- and A-cells from triplicate day 11 CD34 cultures. Two different antibodies were used for each globin chain in duplicate blots. Standard curves were generated using purified human γ- and β-globin. (D) Estimated percentage of γ-globin protein (relative to β-globin + γ-globin) in day 11 erythroblasts from five donors.



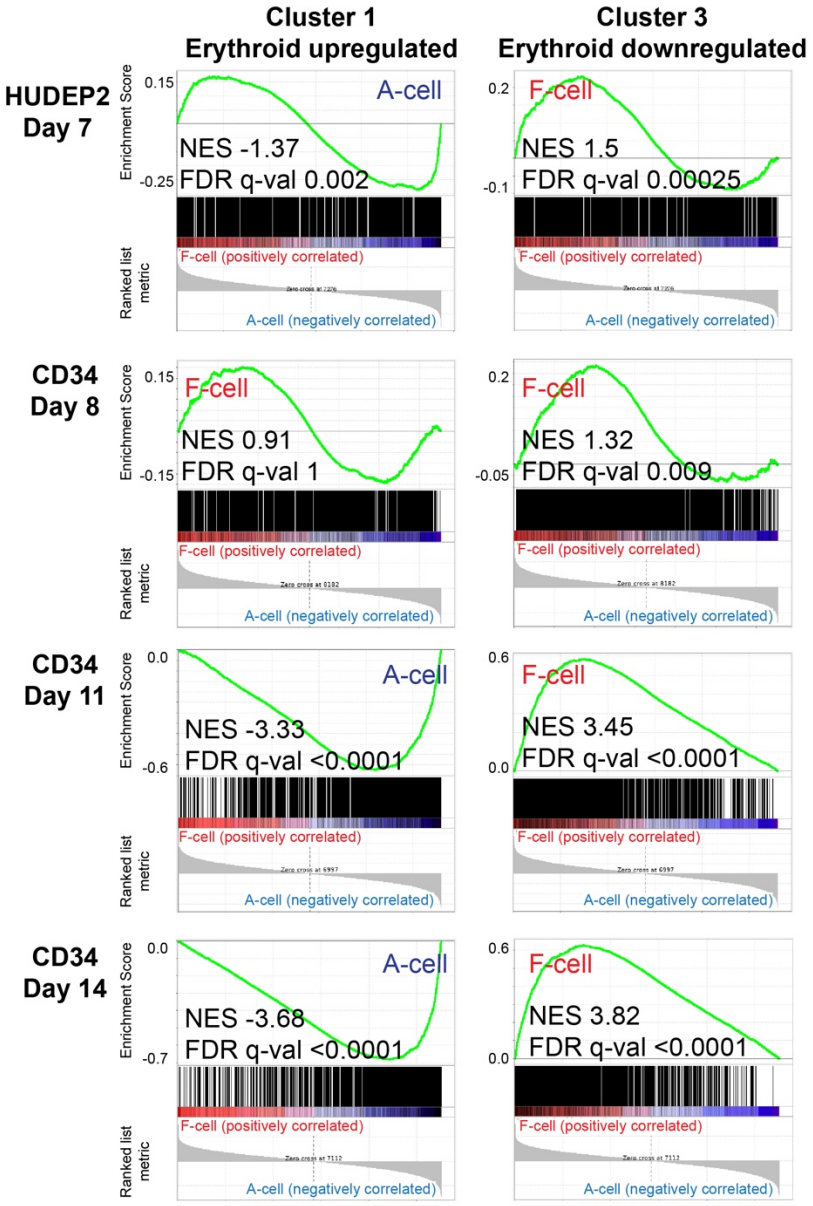
Supplemental Figure 5: Comparison of F- and A-cell transcriptomes to fetal and adult erythroblasts. Gene set enrichment analysis of CD34-derived F-cells using fetal and adult erythroid expression signatures was done using previously generated transcriptome data from fetal liver and adult peripheral blood derived CD34 cultures at 11 and 14 days of differentiation.[5] Top 250 fetal enriched or adult enriched transcripts were used as fetal and adult signatures, respectively. Venn diagrams show overlap of differentially expressed transcripts from fetal or adult erythroblasts with those from F- and A-cells. DEGs were defined as having FDR < 0.05 and >1.5-fold change for all data sets.

Figure S5

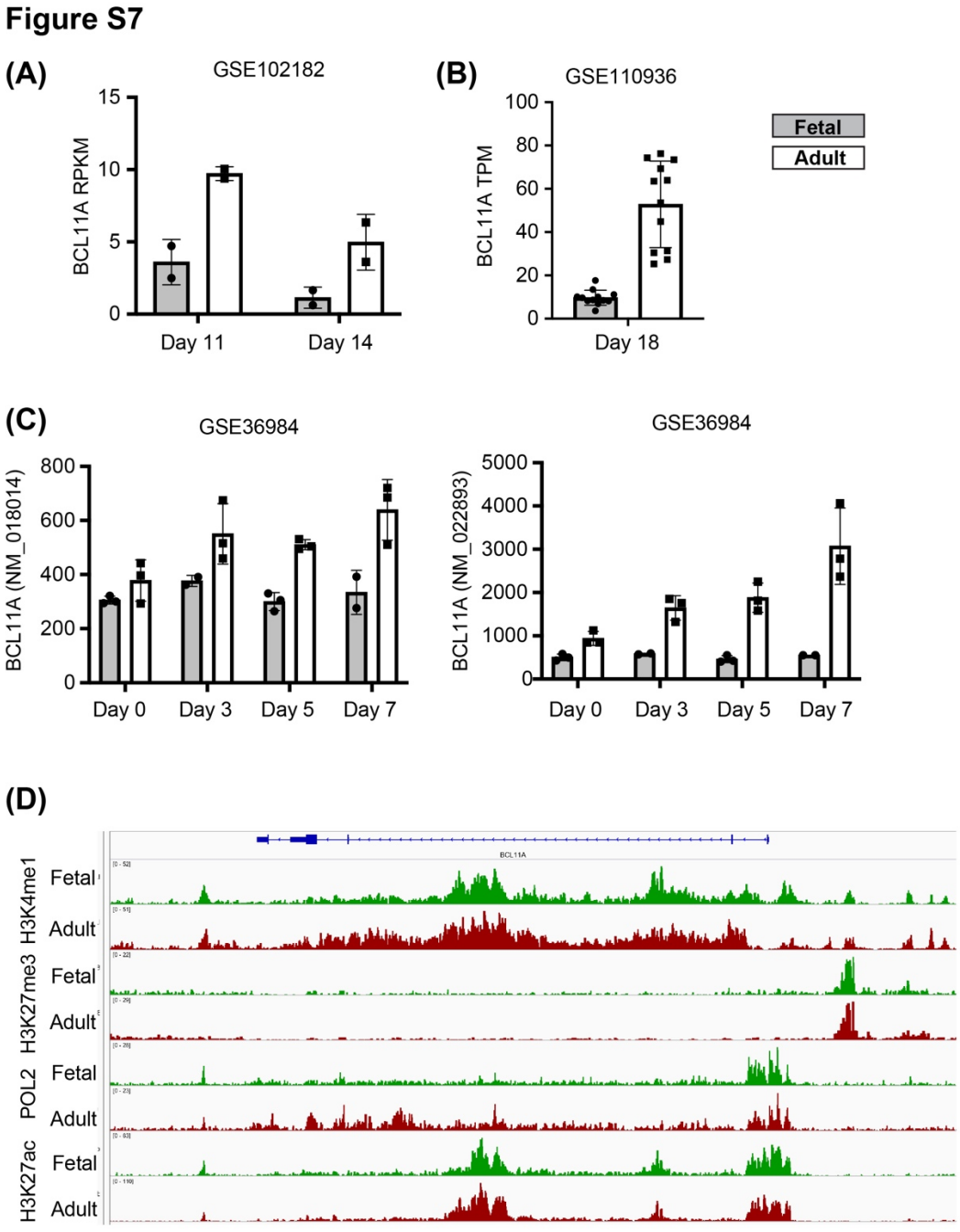


Supplemental Figure 6: Gene set enrichment analysis of CD34- and HUDEP2-derived F-cells using erythroid differentiation transcriptional signatures. RNA-seq analysis of human CD34 erythroid differentiation was used to generate expression clusters. Cluster 1 contains genes that are upregulated during erythroid differentiation and cluster 3 contains genes that are downregulated during erythroid differentiation.

Figure S6

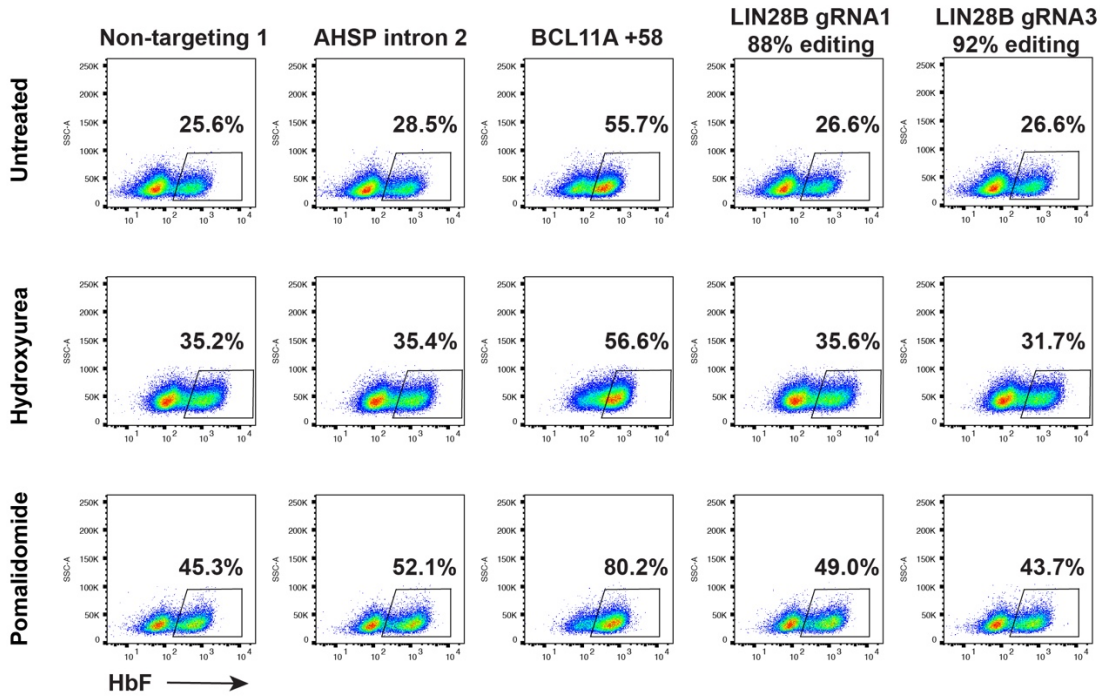


Supplemental Figure 7: Analysis of BCL11A expression in previously published studies of fetal and adult erythroblasts. **(A)** RNA-seq analysis of CD34 erythroid cultures from fetal liver and adult peripheral blood (GEO accession GSE102182).[5] Cells were analyzed at days 11 and 14 of differentiation. Day 11: 3.3 fold, FDR 3.8E-17; Day 14: 4.5 fold, FDR 2.6E-11. **(B)** RNAseq of CD34 erythroid cultures from fetal liver and adult bone marrow (GEO accession GSE110936).[6] Cells were studied at day 18 of culture (day 12 of differentiation following 6-day expansion). **(C)** Microarray data from a differentiation timecourse of CD34 erythroid cells derived from fetal liver and adult peripheral blood (GEO accession GSE36984).[7] Two different probes for BCL11A are shown. **(D)** Chromatin landscape of the BCL11A locus in fetal and adult erythroblasts (GEO accession GSE36984).[7]

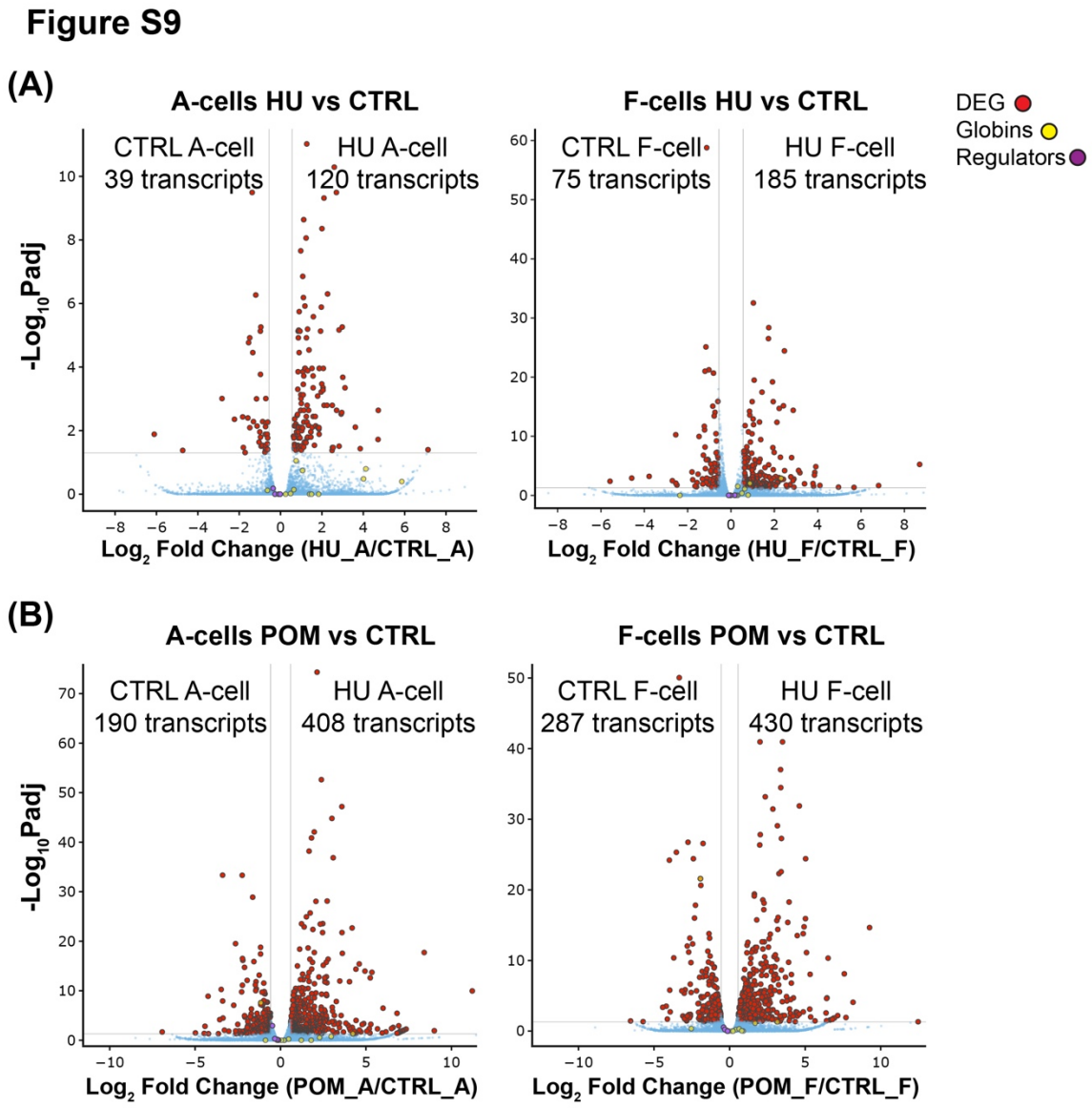


Supplemental Figure 8: LIN28B ablation in primary erythroid cultures. CD34+ HSPC-derived erythroid cells were electroporated with RNP complexes containing negative control gRNA (non-targeting, AHSP intron 2), positive control gRNA targeting the +58 enhancer of BCL11A, and two gRNAs against LIN28B. TIDE analysis after 7 days in culture showed high editing efficiency for LIN28B. Cells were subsequently cultured with vehicle, hydroxyurea, and pomalidomide from days 6-11 prior to flow cytometry analysis for HbF expression at day 12.

Figure S8



Supplemental Figure 9: Comparison of A-cells and F-cells following drug treatment. (A) RNA-seq data comparing A-cells between HU and CTRL conditions (left) and F-cells between HU and CTRL conditions (right). (B) RNA-seq data comparing A-cells between POM and CTRL conditions (left) and F-cells between POM and CTRL conditions (right). Average data from two replicates is shown. DEG: differentially expressed genes, FDR <0.05 and fold change >1.5. Globin transcripts and known HbF regulators (BCL11A, ZBTB7A, EIF2AK1, and SPOP) are highlighted.



Supplemental Figure 10: Gene set enrichment analysis of RNA-seq data from hydroxyurea and pomalidomide-treated F- and A-cells using erythroid differentiation transcript signatures (A) and fetal and adult erythropoiesis signatures (B). Venn diagrams show overlap of differentially expressed transcripts from fetal or adult erythroblasts with those from F- and A-cells. DEGs were defined as having FDR < 0.05 and >1.5 fold change.

Figure S10