Nucleic Acids Research, 2025, **53**, gkaf093
https://doi.org/10.1093/nar/gkaf093
Gene regulation, Chromatin and Epigenetics

The lncRNA DUBR is regulated by CTCF and coordinates chromatin landscape and gene expression in hematopoietic cells

Hober Nelson Núñez-Martínez [1], Gustavo Tapia-Urzúa [1], Ángel Josué Cerecedo-Castillo [1], Carlos Alberto Peralta-Alvarez [1,3], Georgina Guerrero [1], Maite Huarte [2], Félix Recillas-Targa [1,*]

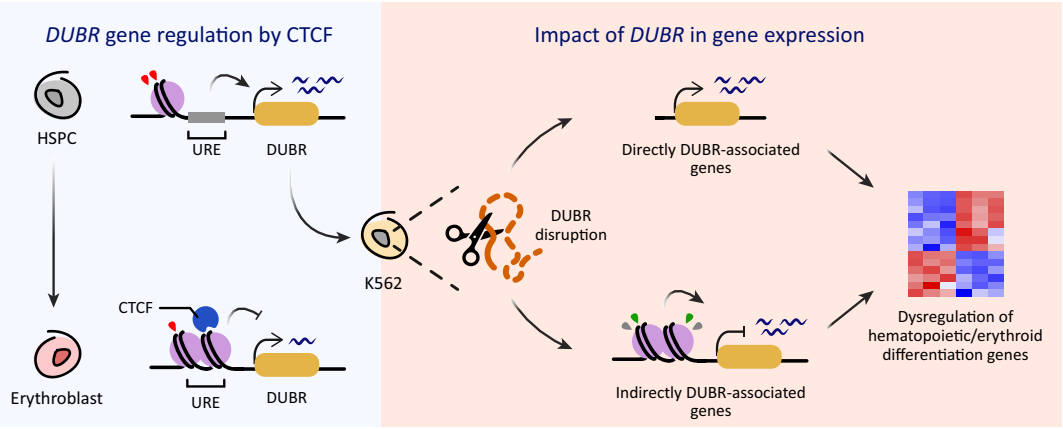
[1]Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, 04510, Ciudad de México, México
[2]Center for Applied Medical Research, Department of Gene Therapy and Regulation of Gene Expression, University of Navarra, Pamplona, 31008, Spain
[3]Unidad de Bioinformática y Manejo de la Información, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, 04510, Ciudad de México, México

[*]To whom correspondence should be addressed. Email: frecilla@ifc.unam.mx

Abstract

Master hematopoietic transcription factors (TFs) and long noncoding RNAs (lncRNAs) coordinate shaping lineage-specific gene expression programs during hematopoietic differentiation. The architectural protein CCCTC-binding factor (CTCF) has emerged as a pivotal regulator of gene expression in cell differentiation. However, the relationship and its regulatory effect of CTCF on lncRNA genes in hematopoiesis remain elusive. We demonstrated that CTCF constrains the lncRNA *DUBR* transcription throughout erythroid differentiation. *DUBR* is highly expressed in human hematopoietic stem and progenitor cells (HSPCs) but depleted in erythroblasts. *DUBR* perturbation dysregulates hematopoietic-erythroid cell differentiation genes and facilitates genome-wide activation of regulatory elements. A genomic map of RNA occupancy revealed that *DUBR* associates with a set of genes involved in regulating hematopoietic differentiation, including the erythroid repressor *HES1*, which targets a subset of regulatory elements of DUBR-dysregulated genes. Our results support the role of *DUBR* as a regulator of a hematopoietic differentiation gene program by coordinating the expression of genes and influencing their chromatin regulatory landscape.

Graphical abstract



Introduction

Hematopoiesis has been adopted as a paradigm to understand and unravel the regulatory mechanisms that establish gene expression programs driving lineage commitment and cell differentiation [1]. To generate all blood cell types, hematopoietic stem and progenitor cells (HSPCs) undergo extensive chromatin re-organization, as lineage-restricted gene expression programs are established [2–4]. The erythroid lineage

Received: December 18, 2023. **Revised:** January 24, 2025. **Editorial Decision:** January 28, 2025. **Accepted:** February 3, 2025
© The Author(s) 2025. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

produces red blood cells, or erythrocytes, which synthesize hemoglobin that transports oxygen to tissues. Erythroid cell formation, also known as erythropoiesis, initiates with the differentiation of HSPC to erythroid progenitors (EProg: BFU-E and CFU-E) that later generate erythroid precursors (EPrec) or erythroblasts, and subsequently erythrocytes [5]. Transcriptional and epigenetic regulators of erythropoiesis include lineage-specific transcription factors (TFs), chromatin remodelers, regulatory elements, as well as long noncoding RNAs (lncRNAs) [6, 7].

In mammals, lncRNAs fine-tune gene expression in a broad range of biological processes [8]. Multiple lncRNAs function as hematopoietic regulators [9–11]. For instance, *Spehd* is abundant in HSPC and required for multilineage cell differentiation, by regulating oxidative phosphorylation [12]. Similarly, the HSPC-specific lncRNA *lncHSC-1* inhibits myeloid differentiation, whereas *lncHSC-2* inhibits lymphoid T-cell commitment [13]. Diverse lncRNAs are also required at different stages of erythropoiesis [14–17]. For example, *Bloodlinc* regulates red blood cell development by promoting erythroblast proliferation and enucleation [18]; *lncEry* coordinates terminal erythroid differentiation and maturation by activating *Klf1* and globin genes [19]; *UCA1* controls heme biosynthesis during erythroblast development [20]; and *shlnc-EC6* drives erythroid enucleation via the Rac1/PIP5K signaling pathway [21]; whereas *GATA2AS* opposes erythroid differentiation by affecting erythroid TF binding at regulatory elements [22]. Despite the growing evidence of lncRNAs as regulators of erythroid differentiation, the functional relevance of most annotated lncRNAs in hematopoiesis and erythropoiesis remains largely unknown.

Dynamic occupancy of regulatory elements (enhancer and promoters) by key TFs including GATA-1, GATA-2, TAL1, KLF1, PU.1, and NF-E2 [23–27] drives erythroid commitment, differentiation, and maturation. GATA-1, the most studied erythroid-specific TF, is essential to establishing the erythroid transcriptional program [28]. GATA-1-binding sites are frequently co-occupied by the CCCTC-binding factor (CTCF) at regulatory elements enriched for histone H3 acetylated in lysine 27 (H3K27ac). GATA-1 and CTCF co-occupancy is essential to coordinate gene expression programs across erythroid cell differentiation [29–31]. CTCF is an architectural protein that contributes to the three-dimensional organization of the genome and controls gene expression, by facilitating enhancer–promoter communication, among other functions [32]. CTCF dynamically binds erythroid-specific regulatory elements to ensure proper erythroid lineage establishment [29, 33–35].

While the contributions of CTCF and lncRNAs in erythropoiesis have been studied independently, how they interplay to regulate erythroid cell commitment and differentiation remains largely unknown. In this work, we utilized human erythroleukemia K562 cells to study the functional role of CTCF in regulating the lncRNA *DUBR* and its impact on the hematopoietic transcriptional gene program. CTCF favors *DUBR* repression by reducing chromatin accessibility and histone H3K27ac deposition at an upstream regulatory element (URE) during erythroid differentiation. *DUBR* disruption affects a hematopoietic-erythroid gene expression program and cell proliferation. The genomic map of RNA occupancy reveals that DUBR associates with a subset of dysregulated genes identified upon *DUBR* disruption. Between the DUBR-associated genes, we identified the TF *HES1* involved in opposing erythroid differentiation. Chromatin immunoprecipitation (ChIP)-seq and transcriptomic (RNA-seq) data integration revealed an interplay between *DUBR* and *HES1* in coordinating a hematopoietic transcriptional gene program.

Materials and methods

Cell culture, lentiviral production, and cell infection

K562 cells were grown in ISCOVE (Gibco) supplemented with 10% fetal bovine serum (Biowest) and 1× penicillin–streptomycin (Biowest). HEK-293T cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Biowest) containing 10% of fetal bovine serum (FBS) (Gibco) and 1× penicillin–streptomycin (Biowest). Cells were maintained at 37°C and 5% CO_2. For lentiviral production in HEK-293T cells, 1×10^7 cells were transfected in 10 ml of final volume of complete media containing 0.3 M $MgCl_2$, 2× HEBS (HEPES Buffered Saline) [280 mM NaCl, 10 mM KCl, 1.5 mM Na_2HPO_4, 12 mM D-glucose and 50 mM HEPES (pH 7.05)], and a mix of plasmids as follows: 10 μg of the vector of interest, 3 μg of pMD2.G, and 6 μg of psPAX2. Transfection media was replaced by fresh complete media after overnight transfection. The supernatant containing the virus was harvested 48 h post-transfection and centrifuged for 90 min at 27000 rpm at 4°C. The pellet containing the viral particles was eluted with 1× phosphate-buffered saline (PBS) overnight at 4°C. This lentiviral supernatant was aliquoted and stored at −80°C. A total of 1×10^5 K562 cells per well in 6-well plates were infected with 1 ml of lentiviral supernatant and 2 ml of media supplemented with 8 μg/ml polybrene (Sigma). Twenty-four hours after infection, lentiviral media was replaced by fresh complete media supplemented with a selection agent (puromycin or hygromycin). Cells were maintained with the selection agent for 5 days before the experiment analysis.

Plasmid constructions

For Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) experiments, all single-guide RNAs (sgRNAs) employed in this study were designed with the CRISPOR web tool (http://crispor.tefor.net/) [36] using the hg19 genome annotation. SgRNAs with high specificity scores (>70) as recommended by the CRISPOR tool were purchased from Sigma. For genetic deletions with CRISPR–Cas9, sgRNAs used to generate C1, C2, ΔEX1, and ΔEX2 mutants were cloned into lentiCRISPRv2 (a gift from Feng Zhang, Addgene plasmid #52961). SgRNAs used for CRISPRa and CRISPRi were cloned into lentiSAM v2 (a gift from Adam Karpf, Addgene plasmid #92062) and pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro (a gift from Charles Gersbach, Addgene plasmid #71236), respectively. Cloning was carried out as previously described [37]. Briefly, sgRNAs synthesized as oligonucleotides (Sigma) were flanked by BbsI- or BsmBI-compatible overhangs and phosphorylated using T7 polynucleotide kinase (NEB) and annealed. Annealed sgRNAs were ligated into BbsI- or BsmBI-digested plasmid. Ligated plasmids were transformed into *Escherichia coli* TOP10 competent cells. Bacteria clones were picked and used for colony polymerase chain reaction (PCR) with U6 forward and sgRNA reverse oligonucleotides. Positive colonies were grown and used for plasmid purification. DUBR overexpression plasmid (pcDUBR) was constructed using PCR-amplified DUBR isoform 2 (DUBR-2). PCR fragments were cloned into pcDNA3.1(+) plasmid

(Invitrogen). All the primers and sgRNAs used in this study are listed in Supplementary Table S6.

RNA isolation from subcellular fractionation

RNA isolation from chromatin, nucleoplasm, and cytoplasm was carried out as previously described with some modifications [38]. Briefly, 2×10^7 K562 cells were washed twice with 1× PBS and centrifuged for 2 min at 5000 × *g* at 4°C. The cell pellet was resuspended in cytoplasmic lysis buffer (0.15% NP-40, 150 mM NaCl, 25 μM α-amanitin, 10 U SUPERase.IN, and 1× Complete protease inhibitor mix) for 5 min in ice. Sucrose buffer [10 mM Tris–HCl (pH 7.0), 150 mM NaCl, 25% sucrose, 25 μM α-amanitin, 10 U SUPERase.IN, and 1× Complete protease inhibitor mix] was added to the cell lysate and centrifuged for 10 min at 16000 × *g* at 4°C. The supernatant (cytoplasmic fraction) was stored at −70°C until RNA isolation. The nuclear pellet was washed with nuclei wash buffer (1× PBS, 0.1% Triton X-100, 1 mM EDTA, 25 μM α-amanitin, 10 U SUPERase.IN, and 1× Complete protease inhibitor mix) and centrifuged for 1 min at 1150 × *g* at 4°C. The cell pellet was resuspended in glycerol buffer [20 mM Tris–HCl (pH 8.0), 75 mM NaCl, 0.5 mM EDTA, 50% glycerol, 0.85 mM DTT, 25 μM α-amanitin, 10 U SUPERase.IN, and 1× Complete protease inhibitor mix]. Nuclei lysis buffer [1% NP-40, 20 mM HEPES (pH 7.5), 300 mM NaCl, 1 M urea, 0.2 mM EDTA, 1 mM DTT, 25 μM α-amanitin, 10 U SUPERase.IN, and 1× Complete protease inhibitor mix] was added and incubated on ice for 2 min and centrifuged for 2 min at 18500 × *g* at 4°C. The supernatant (nucleoplasmic fraction) and pellet (chromatin fraction) were separated and stored at −70°C until RNA isolation. For RNA isolation, each fraction was resuspended with TRIzol reagent (Invitrogen) as described below.

RNA isolation and RT-qPCR

Total RNA was isolated using TRIzol Reagent (Invitrogen) according to the manufacturer's protocol with minor modifications. Briefly, the cell pellet was resuspended in TRIzol Reagent and incubated at room temperature for 10 min. Chloroform (Invitrogen) was added, incubated at room temperature for 10 min, and centrifugated for 10 min 12000 × *g* at 4°C. The aqueous phase was resuspended in 2-propanol (Invitrogen) and centrifuged for 10 min at 12000 × *g* at 4°C. The RNA pellet was washed twice with 75% ethanol and resuspended in nuclease-free water. RNA was used directly to determine gene abundance by KAPA SYBR FAST One Step Kit (KAPA Byosystems) using the StepOne Real-Time PCR System. The endogenous RNA from the HPRT gene was utilized for internal normalization. Reverse transcription quantitative real-time PCR (RT-qPCR) data were analyzed by the ΔΔCt method [39]. Significance in gene expression was determined by Student's *t*-test by GraphPad Prism 9.0. All the primers used in this study are listed in Supplementary Table S6.

CRISPR–Cas9, CRISPRa, and CRISPRi assays

For CRISPR–Cas9 mediated deletions, sgRNAs cloned into lentiCRISPRv2 were used to generate lentiviral particles as described above. A total of 1×10^5 K562 cells per well were infected in 6-well plates with 1 ml of lentiviral supernatant in 2 ml of media supplemented with 8 μg polybrene (Sigma). Twenty-four hours after infection, lentiviral media was replaced by fresh complete media supplemented with a selection agent (puromycin or hygromycin). Cells were maintained with the selection agent. After 5 days of selection, a cell pellet was used for DNA extraction by phenol–chloroform. Desired deletions were analyzed by PCR genotypification (see Supplementary Table S6). For isolation of mutant cell clones, 1×10^4 pooled cells were serially diluted in a 96-well plate containing 100 μl of complete ISCOVE medium per well. After 2 weeks, single clones were identified by microscopy and expanded for subsequent genotypification by PCR. Deletion of each mutant cell clone was further characterized by cloning PCR fragments obtained from genotypification into pGEM-T Easy (Promega) and confirmed by Sanger sequencing. For CRISPRa and CRISPRi assays, sgRNAs overlapping the URE and the promoter region of DUBR were cloned and used for lentivirus production as described before. For CRISPRa, 1×10^5 K562 cells were transduced with lentivirus carrying lentiMPH v2 (a gift from Adam Karpf, Addgene plasmid #92065). Twenty-four hours post-transduction, cells were selected using Hygromycin B (Thermo Fisher Scientific) for 4 days. Next, 1×10^5 of selected cells were infected with lentivirus expressing sgRNAs cloned into lentiSAM v2. Twenty-four hours post-transduction, cells were selected using Puromycin (Sigma) for 4 days. For CRISPRi, 1×10^5 K562 cells were transduced with lentivirus-carrying sgRNAs cloned into pLV-hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro. Twenty-four hours post-transduction, cells were selected using Puromycin (Sigma) for 4 days.

MTT cell proliferation assay

A total of 5×10^3 K562 cells per well were seeded in a 96-well plate. Cell proliferation was measured using the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay for 4 days using the Cell Proliferation Kit (Roche) according to manufacturer's protocol.

RNA sequencing and data analysis

Total RNA was isolated from nonedited (wild-type, WT) and *DUBR* mutants (ΔEX1 and ΔEX2) by triplicate using TRIzol reagent (Thermo Fisher) as described above. RNA sequencing (RNA-seq) libraries for all samples were prepared and sequenced by Novogene. Three independent libraries per condition were sequenced in an Illumina HiSeq 4000 platform as paired-end (25×10^6 reads per sample) 150 bp reads. Sequencing reads were aligned to the human genome assembly hg19 (GRCh37) using STAR with default parameters [40]. Mapped reads (bam files) were used for the read count at the gene level using featureCounts with human Gencode v36 annotation [41]. Differential expression analysis was conducted by edgeR using read count files [42]. Significant genes were identified by filtering in R as follows: false discovery rate (FDR) < 0.05, log2FC ± 2. RNA-seq signal tracks were created by merging bam files from replicates and converted to bigwig files using deepTools2 and normalized as counts per million (CPM) [43]. Heatmap was generated by transforming log2 CPM to *z*-score of differentially expressed genes using R. Volcano plot was generated using ggplot2 in R. Quality of RNA-seq reads generated in this study are included in Supplementary Table S1.

Chromatin immunoprecipitation

ChIP for CTCF and histone marks (H3K4me1, H3K4me3, H3K27ac, and H3K27me3) was performed as described with

some modifications [44]. Briefly, 4 × 10^7 K562 cells were crosslinked in 1% formaldehyde for 10 min at RT. Crosslink quenching was carried out with 0.125 M glycine for 5 min at 4°C. Cells were immediately centrifuged and washed twice with 1× PBS. The cell pellet was then resuspended in cell lysis buffer [10 mM Tris–HCl (pH 7.5), 10 mM NaCl, 0.3% NP-40, supplemented with protease inhibitors] and incubated for 30 min at 4°C. Nuclear fractions were isolated by centrifugation and dissolved in nuclear lysis buffer [50 mM Tris–HCl (pH 7.5), 10 mM EDTA, 1% sodium dodecyl sulfate (SDS), supplemented with protease inhibitors]. Chromatin was fragmented by sonication in a bioruptor for 12 cycles as follows: 30-seg on and 30-seg off. Chromatin fragmentation was evaluated by agarose gel electrophoresis. Insoluble chromatin was discarded by centrifugation at maximum speed; the supernatant was kept on ice. Fifty micrograms of chromatin per each IP was diluted 1:5 with dilution buffer (1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl, 150 mM NaCl, supplemented with protease inhibitors). Chromatin was precleared by adding 50 µl of blocked protein G/A beads for 2 h. The volume corresponding to 1% of chromatin used in each IP was saved as input. Chromatin was incubated overnight at 4°C with the corresponding primary antibodies. Next, 30 µl of blocked protein G/A beads were added to diluted chromatin and incubated for 2 h at 4°C. Beads were washed as follows: four washes with wash buffer I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl, 150 mM NaCl, supplemented with protease inhibitors) and a final wash with wash buffer II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl, 500 mM NaCl, supplemented with protease inhibitors). Beads along with the input chromatin were eluted in elution buffer (1% SDS and 100 mM $NaHCO_3$) and decrosslink buffer (200 mM Tris–HCl, 400 mM NaCl, 0.4% SDS, and 10 mM EDTA), incubated in a rotating wheel for 10 min at RT, and treated with RNase A (Ambion) for 1 h at 37°C and Proteinase K (NEB) for 4 h at 65°C. DNA was retrieved by adding 1:1 phenol:chloroform:isoamyl alcohol (Invitrogen) and mixed by rotation for 10 min at RT, and then centrifuged for 10 min at 12000 rpm at RT. The aqueous layer was retrieved. DNA was precipitated with 1 M ammonium acetate, glycogen (Roche), and 100% ethanol for 2 h at −70°C. The DNA pellet was obtained by centrifugation for 30 min at 12800 rpm and washed twice with 70% ethanol. The pellet was resuspended in nuclease-free water. For CTCF ChIP-qPCR, purified DNA from ChIP was used for real-time PCR (qPCR) using iTaq Universal SYBR Green Supermix (Bio-Rad) and oligonucleotides flanking the CTCF-binding motif. Fold enrichment of CTCF over the analyzed regions (URE and exon-2) on all the conditions (EV, C1, and C2) was compared to WT-CTCF. Oligonucleotides used for ChIP-qPCR are listed in Supplementary Table S1. The antibodies used in this study were as follows: anti-CTCF (5 µg, Millipore #07-729), anti-H3K4me3 (5 µg, Abcam #8580), anti-H3K4me1 (5 µg, Abcam #8895), anti-H3K27ac (5 µg, Abcam #4729), anti-H3K27me3 (5 µg, Abcam #6002), and anti-IgG (2 µg, Millipore #12-371).

ChIP-sequencing (ChIP-seq) and data analysis

Two ChIP-seq libraries were prepared per condition using the TruSeq Library Preparation Kit (Illumina), according to the manufacturer's protocol. Libraries were multiplexed and sequenced on an Illumina HiSeq 4000 platform as paired-end (25 × 10^6 reads per sample) 150 bp reads.

Sequencing reads were aligned to the human genome assembly hg19 (GRCh37) using bowtie2 with default parameters [45]. Mapped reads (bam files) were filtered using NGSUtils [46] as follows: keep mapping and properly paired read, discard secondary alignments, PCR duplicates, and low-quality mapped reads. ChIP-seq peaks were determined by MACS2 using an FDR < 0.05 [47]. Peaks with fold change >2 were used for downstream analysis. Peaks overlapped with ENCODE blacklisted regions were excluded. Bam files from each replicate were merged and used to generate read-density signal tracks using deepTools2 and normalized as CPM [43]. For regions with differential H3K27ac enrichment between K562 WT cells and ΔEX1, DiffBind was used from R/Bioconductor. The resulting peaks were filtered as follows: FDR < 0.01 and fold change ±2. Plot profiles and heatmaps of ChIP-seq read-density were generated by deepTools2 [43]. Significance was determined by the Mann–Whitney *U* test. CTCF peaks were determined as narrow peaks and histone marks as broad peaks. Motif analysis was conducted by *findMotifsGenome.pl* from HOMER using ChIP-seq peaks [48]. Chromatin segments were assigned by the Chromatin Hidden Markov Model (chromHMM) software as previously described [49]. The quality of ChIP-seq reads generated in this study is included in Supplementary Table S1.

Chromatin isolation by RNA purification

Chromatin isolation by RNA purification (ChIRP) was performed as previously described with some modifications [50]. Briefly, 1 × 10^8 K562 cells were crosslinked in 1% glutaraldehyde for 10 min at RT. Quenching was performed with 0.125 M glycine for 5 min, and cells were centrifuged at 2500 *g*. The cell pellet was washed twice with 1× PBS. The cell pellet was then resuspended in cell lysis buffer [10 mM Tris–HCl (pH 7.5), 10 mM NaCl, 0.3% NP-40, supplemented with protease inhibitors and Superase-in (Ambion)] and incubated for 30 min at 4°C. Nuclear cell fractions were isolated by centrifugation for 5 min at 2500 × *g*, 4°C. Nuclear pellet was resuspended in nuclear lysis buffer [50 mM Tris–HCl (pH 7.5), 10 mM EDTA, 1% SDS, supplemented with protease inhibitors and Superase-in (Ambion)]. Chromatin was fragmented by sonication in a bioruptor for 15 (30-seg on and 30-seg off). Chromatin fragmentation was evaluated by agarose gel electrophoresis. Insoluble chromatin was discarded by centrifugation at maximum speed, and the supernatant was kept on ice. Chromatin extract was diluted with hybridization buffer [750 mM NaCl, 1% SDS, 50 mM Tris–HCl (pH 7.0), 1 mM EDTA, 15% formamide, and supplemented with phenylmethylsulfonyl fluoride (PMSF), proteinase inhibitors, and Superase-in] and incubated with 100 pmol of each set of probes (EVEN, ODD, and LacZ) at 37°C for 4 h in a rotating wheel. Streptavidin magnetic beads (Invitrogen) were added to each hybridization reaction and incubated for 30 min at 37°C under rotation. Beads were pulled down and washed five times with wash buffer [2× SSC, 0.5% SDS, dithiothreitol (DTT), and PMSF]. For RNA elution, 10% of each bead sample was resuspended in RNA PK buffer [100 mM NaCl, 1 mM EDTA, 0.5% SDS, 10 mM Tris–HCl (pH 7.0)] and incubated for 45 min at 55°C. RNA was isolated by adding TRIzol Reagent (Invitrogen) and following the previous protocol. Retrieved RNAs were detected by RT-qPCR. For DNA elution,

the remaining bead samples were resuspended in DNA elution buffer (50 mM $NaHCO_3$, 200 mM NaCl, 1 mM EDTA, and 1% SDS), RNase A (Sigma–Aldrich), and RNase H (Sigma–Aldrich), and incubated for 30 min at 37°C. Reverse crosslink was carried out by adding Proteinase K (NEB) and incubating overnight at 65°C. DNA was isolated by phenol–chloroform as outlined before. All biotinylated probes were purchased from Sigma and listed in Supplementary Table S1.

ChIRP-seq and data analysis

ChIRP-seq libraries were prepared as described for ChIP-seq. EVEN, ODD, and Input libraries were prepared in duplicate. Libraries were multiplexed and sequenced on an Illumina HiSeq 500 platform as paired-end (25×10^6 reads per sample) 150 bp reads. Sequencing reads were aligned to the human genome assembly hg19 (GRCh37) using bowtie2 with default parameters [93]. Mapped reads (bam files) were filtered using NGSUtils as follows: keep mapping and properly paired read, discard secondary alignments, PCR duplicates, and low-quality mapped reads. Bam files for each replicate corresponding to EVEN, ODD, and INPUT were merged and used for peak calling [94]. ChIR-seq peaks for EVEN and ODD were determined by MACS2 using an FDR < 0.01 [47]. Peaks with fold change >5 over the input were used for downstream analysis. Peaks over ENCODE blacklisted regions were excluded. Bam files from each replicate were merged and used to generate read-density signal tracks using deepTools2 and normalized as CPM [43]. To identify high-confidence ChIRP peaks, we only analyzed overlapped regions and normalized signals between EVEN and ODD peaks. First, we identified overlapped EVEN and ODD peaks using deepTools2 (2031 peaks). Normalized EVEN and ODD ChIRP-seq signal over the resulting peaks was determined and used to calculate a fold change between both signals. Peaks with fold change between 0.5 and 2 were used for Pearson correlation. These peaks were used to generate heatmaps of read-density using deepTools2. The quality of ChIRP-seq reads generated in this study is included in Supplementary Table S1.

Gene ontology and transcription factor enrichment

Biological processes of differentially expressed genes were determined by Metascape (https://metascape.org) [51]. The top significant categories were selected with FDR < 0.05. TF enrichment analysis of gene sets was carried out by ChEA3 (https://maayanlab.cloud/chea3) [52].

Statistical analysis

Data represent the mean ± standard error of three biological replicates. Significance was determined by a two-tailed unpaired Student's *t*-test with Welch's correction using GraphPad Prism. For plot profiles of read-density over specific regions and box plots (or otherwise specified in figure legends), significance was determined by a two-sided Mann–Whitney test using R and GraphPad Prism 9. Hypergeometric and Pearson correlation test were carried out using R.

Additional datasets used in this study

Previously published data were reanalyzed and used in this study as follows: ChIP-seq for CTCF (GSE67893 and GSE51338); HES1 (GSE91470); GATA-1, NF-E2, and KLF1 (GSE43625); GATA-1 and H3K27ac (GSE211293); RUNX1 (GSE96253); NCOR1 (GSE92062); TAF1 (GSM803431); H3K27me3 and EZH2 in erythroblasts (GSE218231); H3K27ac and GATA-1 in HSPC, EProg, and EPrec (GSE214811). ATAC-seq for K562 (GSE170378); donors (GSE74912); HSPC-Ortho (GSE128266). RNA-seq for donors (GSE74246); K562 (GSE211316); MEL (GSE148421); HSPC (GSE183266); HSC-Erythroblasts (GSE60101); HSPC-EPrec (GSE124164); HSPC-BFUE-CFUE (GSE128268); proery-ortho (GSE53983); HUDEP2 (GSE213779): GATA-1 knockdown (GSE211319). STARR-seq in K562 (ENCSR858MPS). For K562 MeRIP-seq and MINT-seq (GSE137752). PCHiC in primary human blood cells (https://osf.io/u8tzp/); H3K27ac HiChIP (GSE101498); ChIA-PET for CTCF (GSE39495); GATA-1 HiChIP (GSE214807).

Results

CTCF is associated with changes in lncRNA gene expression during erythroid differentiation

Since CTCF binding to chromatin and lncRNA expression are dynamic during hematopoiesis, we hypothesized that differentially expressed lncRNAs along erythroid cell differentiation are regulated by CTCF. To identify genome-wide CTCF targets, we analyzed CTCF chromatin immunoprecipitation sequencing (ChIP-seq) data from human HSPC and *in vitro* differentiated erythroblasts [34]. Considering the differentially bound sites, CTCF increased its binding in 4245 sites, while decreasing only in 110 sites (FDR < 0.05 and Fig. 1A). Interestingly, chromatin accessibility was also dynamic in sites with differential CTCF binding (FDR < 0.05 and Supplementary Fig. S1A). Furthermore, analysis of public ChIP-seq data showed that these differentially bound sites are occupied by master erythroid TFs, such as GATA-1, NF-E2, and KLF1 in erythroblasts (Supplementary Fig. S1B) [25, 27, 30, 31, 53]. These data suggest an interplay between CTCF and master erythroid regulators to modulate chromatin accessibility of regulatory sites relevant to erythroid differentiation.

To elucidate the mechanism by which CTCF and lncRNAs intersect as regulators of erythroid cell differentiation, we analyzed the distribution of CTCF differential peaks near lncRNAs annotated in the GENCODE database [54]. CTCF preferentially occupied regions distal to lncRNAs, which agrees with its role in maintaining chromatin interactions between promoters and distal regulatory elements [29, 32]. A smaller fraction (10.15%) of sites differentially bound by CTCF located within 5 kb of the transcriptional start site (TSS) of annotated lncRNAs (Fig. 1B, upper panel). We focused on these sites because being proximal to promoters, they likely regulate the expression of 454 lncRNAs. To assess the function of CTCF as a regulator of lncRNAs expression, we first analyzed public RNA-seq data from HSPC, *in vitro* differentiated EProg, and EPrec [24]. We found that 2108 lncRNAs were differentially expressed (1120 down-regulated and 988 up-regulated) in the HSPC-to-EPrec transition. A small fraction of those lncRNAs (85 lncRNAs) had proximal sites that were differentially bound by CTCF (Fig. 1B and Supplementary Fig. S1C). To uncover lncRNAs regulated by CTCF with important functions during erythropoiesis, we identified those whose genomic location and expression dynamics are conserved in mouse and human [55]. Among the initial 85 differentially expressed lncRNAs that display dynamic binding of CTCF at a proximal site, 31 were conserved

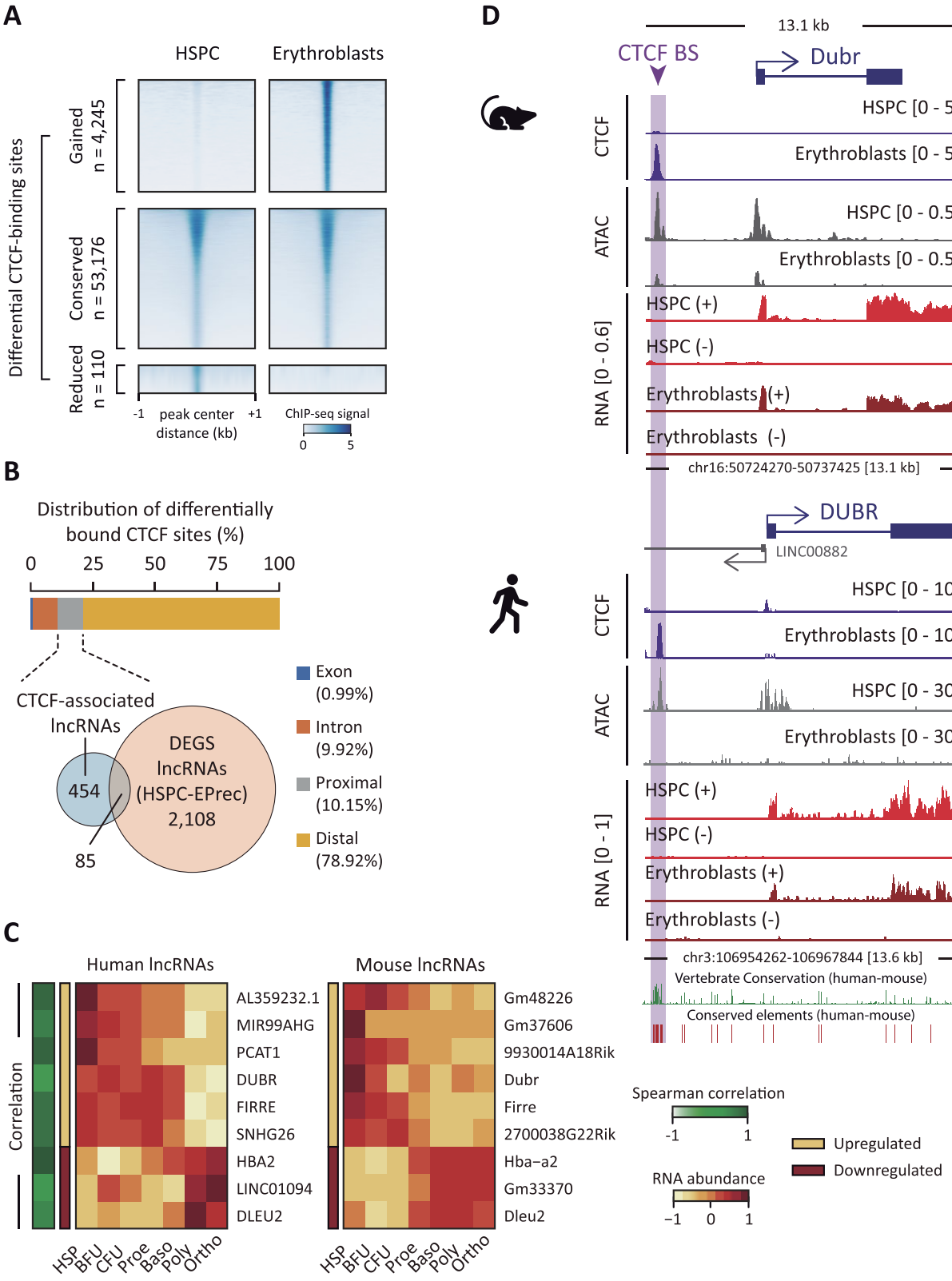


Figure 1. CTCF redistribution is associated with a set of lncRNAs during erythropoiesis. (**A**) Heatmaps of normalized CTCF ChIP-seq signals around differentially bound CTCF sites in erythroblasts compared with HSPC. Heatmaps are ranked by ChIP-seq signal. (**B**) Genomic distribution of dynamic CTCF-binding sites (upper panel) according to gene overlapping (exon or intron) or distance within 5 kb from TSS (proximal). Venn diagram (lower panel) of the overlap between lncRNAs associated with differentially bound CTCF sites ($n = 454$) and differentially expressed lncRNAs in EPrec compared with HSPC ($n = 2108$). (**C**) Heatmaps of RNA abundance (RNA-seq) of conserved human and mouse lncRNAs through erythropoiesis. The color key indicates the Spearman correlation coefficient of the gene abundance of each lncRNA and its mouse counterpart. The RNA abundance is shown as *z*-score of the log2 CPM. (**D**) Genomic snapshot of mouse (upper) and human (lower) *DUBR* loci showing the conserved CTCF-binding site by ChIP-seq signal as well as chromatin accessibility (ATAC-seq) and RNA signal (RNA-seq) obtained from HSPC and erythroblasts. The arrowhead points out the conserved CTCF site. RNA signal is shown in sense (+) and antisense (-) strands.

in mouse, but only 8 were differentially expressed during erythroid differentiation with a high correlation with their mouse syntenic counterparts (Fig. 1C). These included the previously described hematopoietic regulatory lncRNAs such as *DLEU2* and *FIRRE* (Fig. 1C and Supplementary Fig. S1D) [14,56]. Therefore, these results suggest that CTCF may regulate the expression of a set of lncRNAs with a conserved role during erythroid differentiation.

DUBR is down-regulated during erythroid cell differentiation

DUBR (annotated in human as *DPPA2* upstream binding RNA) or *Dum* (*Dppa2* upstream binding muscle lncRNA) in mice regulates *Dppa2incis* during myoblast differentiation [57]. *Dubr* is required for neuronal development [58]. However, its function in erythroid differentiation is unknown. *DUBR* expression was high in HSPC, decreased in EProg (BFU-E and CFU-E), and further downregulated in EPrec (proerythroblasts, basophilic erythroblasts, polychromatophilic erythroblasts, and orthochromatic erythroblasts) (Fig. 1C). Such a reduction in *DUBR* abundance inversely correlates with increased expression of globin genes, *HBA2* and *Hba-a2*, in human and mouse, respectively (Supplementary Fig. S1E). Nine isoforms of human *DUBR* have been annotated in GENCODE; however, only the isoform *DUBR-217*, which contains two exons, was expressed in human HSPC and rapidly decreased during *ex vivo* and *in vitro* erythroid differentiation (Supplementary Fig. S2). From the five isoforms of mouse *DUBR*, *Dubr-002* is highly expressed in HSPC (Supplementary Fig. S2). We observed that the sequence and exon structure are conserved between these human and mouse *DUBR* isoforms (*DUBR-217* and *Dubr-002*). Additionally, the genomic position of *DUBR* flanked by the protein-coding genes *CBLB* in the 5′ end and *BBX* in the 3′ end is also conserved (Supplementary Fig. S2). The human *DUBR* is divergently transcribed from another lncRNA, *LINC00882* (Fig. 1D). However, *LINC00882* is expressed at low levels in HSPC and erythroblasts (<1 CPM) and its genomic position is not conserved in mouse (Fig. 1D). Thus, *DUBR* is a conserved lncRNA that is strongly downregulated during erythroid differentiation. Nevertheless, the mechanisms behind *DUBR* downregulation and its role in erythroid differentiation have not been explored.

CTCF represses *DUBR* transcription in human K562 cells

CTCF is an essential TF that demarcates promoters and facilitates enhancer–promoter communication in erythroid genes throughout differentiation [29, 34, 35]. We observed that CTCF binds to an upstream regulatory DNA element (URE) from the TSS of *DUBR*, in both human and mouse erythroblasts, which coincides with a reduction of chromatin accessibility in these cells in comparison to human and mouse HSPC, respectively (Fig. 1D). Chromatin segmentation (ChromHMM) from the Epigenome Roadmap Project indicates that such regulatory element is active across different cell types and tissues (Fig. 2A, upper panel) [59]. Thus, CTCF could bind the URE to regulate DUBR expression during erythroid differentiation. To address this, we introduced a deletion to ablate the CTCF-binding motif in the URE in human erythroleukemic K562 cells using CRISPR–Cas9. We obtained two independent cell clones (C1 and C2) containing a 16 bp homozygous deletion overlapping two CTCF motifs predicted by JASPAR (Fig. 2A, bottom panel). Accordingly, CTCF binding to the URE was decreased in C1 and C2 cell clones (Fig. 2B). Interestingly, DUBR transcript levels were increased, whereas those of *LINC0082* decreased in C1 and C2 cells (Fig. 2C). This result was confirmed in two independent pools of mutant cells generated with two additional sgRNAs (sg1 and sg2) overlapping the CTCF motif (Supplementary Fig. S3A). The previous suggests that CTCF binds URE to counteract *DUBR* expression along erythroid differentiation, probably by modulating the regulatory activity of the URE. Therefore, we assessed ChIP-qPCR for histone marks in clone C1. Indeed, disruption of CTCF binding in clone C1 increased H3K27ac deposition at the CTCF-BS and *DUBR* promoter (P2) but not at *LINC00882* promoter (P1) (Supplementary Fig. S3B and C).

ATAC-seq showed an accessible chromatin structure of the *DUBR* URE region in HSPC. However, accessibility gradually decreased during differentiation until the URE was almost completely inaccessible in terminally differentiated erythroid cells (Fig. 2D, left panel). Loss of chromatin accessibility corresponded with decreased *DUBR* expression as shown by RNA-seq data (Fig. 2D, right panel). Loss of accessibility of the URE and decreased *DUBR* expression were also observed in erythroblasts relative to HSPC cells in donors (Supplementary Fig. S3B and C) [60]. This suggests that CTCF binding at URE reduces local accessibility and represses *DUBR* expression. To test the regulatory influence of the URE on *DUBR*, a transcriptional activator or repressor (CRISPRa and CRISPRi, respectively [61, 62]) was targeted to the URE via CRISPR–dCas9 (Fig. 2E). Indeed, targeting CRISPRi and CRISPRa to the URE reduced or increased *DUBR* expression, respectively (Fig. 2F and G). A similar effect, although to a lesser extent, was observed on *LINC00882*.

To further explore the contribution of the CTCF-BS on the regulatory activity of the URE, we analyzed the effects of CTCF-binding disruption on the deposition of histone modifications related to active regulatory elements (H3K4me3, H3K4me1, and H3K27ac) in ΔCTCF-BS clone C1 employing ChIP-qPCR. Disruption of CTCF binding results in increased accumulation of H3K27ac at both the URE and the *DUBR* promoter (P2), suggesting that URE activity is regulated by the dynamic binding of CTCF (Supplementary Figs. S3D–F). Additionally, analysis of Hi-C data revealed that the CTCF-BS was located close to a topologically associating domain (TAD) boundary displaying high contact frequency with downstream genes into the same TAD (Supplementary Fig. S3G). Moreover, CTCF-BS anchoring chromatin loops with convergent CTCF-binding sites spanning the TAD as observed by ChIA-PET data [63]. These CTCF-BS-mediated loops were supported by H3K27ac-dependent interactions detected by HiChIP data indicating that CTCF-BS demarcates a chromatin loop domain containing the URE, *DUBR*, and downstream genes within the TAD. Regarding *DUBR* and *LINC00882*, we observed that perturbation of CTCF-BS in C1 and C2 clones resulted in reduced changes in neighboring gene expression, both inside (*BBX* and *CD47*) and outside (*CBLB*) of the TAD (Supplementary Fig. S3H).

In addition to CTCF, analysis of ChIP-seq data from the ReMap database indicated that several hematopoietic and erythroid regulators bind at the URE in K562 cells [64]. For example, GATA-1 binds downstream of the CTCF site (GATA1-BS) in erythroid and K562 cells (Supplementary Fig. S3D).

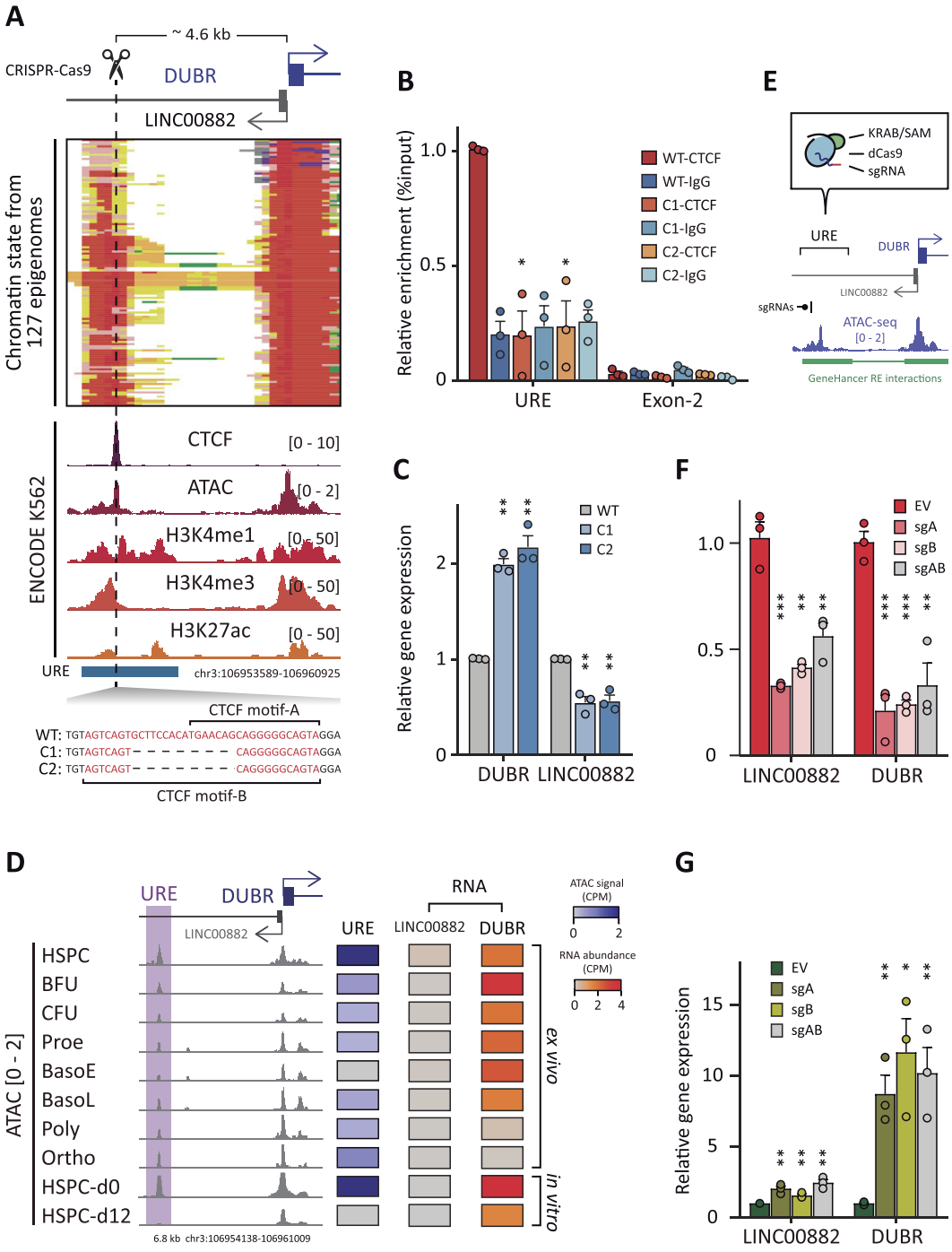


Figure 2. CTCF binding at the URE region leads to DUBR repression. (**A**) Genomic snapshot (upper panel) of chromatin-state annotation tracks of 127 epigenomes obtained from The Epigenome Roadmap Project. The segments point out the promoter and enhancer regions. ATAC-seq and ChIP-seq signal of CTCF and histone marks (H3K4me1, H3K4me3, and H3K27ac) occupancy in K562 cells from ENCODE project (middle panel). Sequences of CTCF motifs assigned by JASPAR (lower panel) overlapping the CTCF peak in the URE region as well as the sequences of mutant cell clones (C1 and C2) lacking 16 bp of the CTCF core motifs. (**B**) Fold enrichment of CTCF (ΔCTCF clone/WT) determined by ChIP-qPCR assay. Exon 2 of DUBR was used as a negative control region (Exon-2). Enrichment was normalized against the percentage of input for the indicated regions. The bar chart shows the mean ± SD of three independent experiments. Significance was calculated by unpaired Student's *t*-test. (**C**) RNA abundance determined by RT-qPCR analysis of K562 cells (WT) and ΔCTCF clones (C1 and C2) for the indicated transcripts. The bar chart shows the mean ± SD of three independent experiments. RNA abundance was normalized to the WT condition. Significance was calculated by unpaired Student's *t*-test. (**D**) Genomic snapshot (left) showing the ATAC-seq signal at the *DUBR* loci from *ex vivo* (HSPC-to-orthochromatic erythroblasts) and *in vitro* differentiated HSPC during 12 days. The highlighted region demarcates the URE region. Heatmaps of URE accessibility signal and *DUBR* and *LINC00882* (right) gene abundance (RNA-seq) from *in vivo* and *in vitro* differentiated cells. (**E**) Schematic representation of the CRISPRi and CRISPRa systems targeting the URE. K562 cells were infected with lentiviral vectors carrying out dCas9–KRAB (CRISPRi) or dCas9-SAM (CRISPRa) alongside two sgRNAs targeting the URE region demarcated by the ATAC-seq signal in K562 cells. The lower panel shows the promoter–enhancer interaction track, as predicted by the GeneHancer prediction model of the UCSC Genome Browser. (**F** and **G**) RNA abundance of *DUBR* and *LINC00882* determined by RT-qPCR analysis of K562 cells stably expressed the CRISPRi (F) and CRISPRa (G) systems. The bar chart shows the mean ± SEM of three independent experiments. RNA abundance was normalized to the empty vector (EV) condition. Significance was determined by unpaired Student's *t*-test as follows; nonsignificant (ns); $*P < .05$; $**P < .01$; $***P < .001$.

Indeed, similar to the ablation of the CTCF-binding site in C1 and C2 clones, GATA-1 knockdown increased *DUBR* expression in K562 cells (Supplementary Fig. S3I) [23]. Interestingly, GATA-1 HiChIP from EProg revealed chromatin interactions from the GATA1-BS with the *BBX* gene similar to the observed with CTCF-BS (Supplementary Fig. S3G) [28]. Overall, these results suggest that CTCF and GATA-1 cooperate to limit the expression of *DUBR* throughout erythroid differentiation. This raises a potential involvement of *DUBR* in this process.

DUBR perturbation affects hematopoietic gene expression and proliferation

LncRNAs are emerging as important regulators of erythroid differentiation [14, 65]. *DUBR* is highly expressed in HSPC and rapidly downregulated during erythroid differentiation (Supplementary Fig. S1E). This suggests that *DUBR* might function in the regulation of gene expression during hematopoiesis. To test this, we assessed the effect of *DUBR* inactivation in erythroid and hematopoietic gene expression. We targeted *DUBR* using two genetic approaches. In the first approach, we used CRISPR–Cas9-mediated gene editing to delete ~600 bp including exon 1 of the *DUBR* locus (ΔEX1) in K562 cells. The second approach deleted ~600 bp overlapping the 3′ end corresponding to the second exon of the isoform 2 (ΔEX2) (Supplementary Fig. S4A). RNA-seq analysis showed that while ΔEX1 mutants displayed undetectable *DUBR* levels (<1 CPM), transcript levels were unaffected in the ΔEX2 mutant compared to WT cells (Fig. 3A). ΔEX1 cells contain a heterozygous deletion of exon 1 and the promoter sequence of *DUBR*, enriched for H3K4me3, which led to decreased *DUBR* transcription (Supplementary Fig. S4A–C). In contrast, ΔEX2 cells harbor a homozygous deletion of a nonconserved 3′ region of exon 2, an m6A modification site [58, 66], and a predicted Polycomb Repressive Complex 2 (PRC2) RNA-interacting region (Supplementary Fig. S4A–C) [67]. While ΔEX2 cells preserve the promoter sequence and transcribe *DUBR* at levels comparable to WT cells, they lack 319 bp that could exert regulatory functions.

RNA-seq analysis showed that a significant proportion of genes were dysregulated in both ΔEX1 and ΔEX2 in comparison to WT cells (673 upregulated and 543 downregulated) (Fig. 3B and Supplementary Tables S2 and S3). Gene ontology analyses revealed that those genes are involved in the cell cycle, DNA repair, hematopoiesis, and hemostasis processes (Fig. 3C). Interestingly, up- and downregulated genes were enriched for direct targets of well-known hematopoietic regulators such as SPI1, TAL1, RUNX1, GATA-1, and KLF1 [26, 68] (Fig. 3D). Importantly, cell proliferation was impaired in ΔEX1 and ΔEX2, compared to nonedited (EV) and WT cells (Fig. 3E), which was consistent with the downregulation of cell cycle-related genes and previously reported roles of *DUBR* in regulating cell division (Fig. 3C) [69–71]. To validate these findings, we generated additional ΔEX1 cell clones (ΔEX1-B and ΔEX1-C) containing heterozygous deletions around the exon 1 of *DUBR* (Supplementary Fig. S4D). ΔEX1-B and ΔEX1-C clones displayed reduced *DUBR* transcript levels as well as for genes involved in the cell cycle progression, including *CDK6*, an essential cell-cycle kinase cyclin-dependent kinase with reported roles in quiescence, proliferation, self-renewal, and differentiation of hematopoietic cells [72, 73]; *ASNS*, an enzyme critical for protein synthesis during cell division and hematopoietic regeneration [74]; *CDC7*, a protein kinase necessary for DNA replication during S-phase and cell division in hematopoietic cells [75]; *E2F7*, a TF required for cell cycle progression and regulation of genes involved in DNA replication, metabolism, and DNA repair [76]; *HES1*, a TF required for cell cycle and hematopoietic differentiation [77–79]. We also validated the upregulation of the hematopoietic regulators *GATA-1*, *RUNX1*, and *SPI1*, which were enriched at promoters of dysregulated genes (Fig. 3D and F) [26, 68]. Additionally, we observed upregulation of the adult globin genes *HBA2*, *HBB,* and *HBG1*, a hallmark of erythroid differentiation (Fig. 3F) [5]. Altogether, these findings indicate that maintaining *DUBR* levels and integrity is required to properly express genes controlling hematopoietic cell proliferation and differentiation.

Characterization of genome-wide changes in the regulatory landscape following *DUBR* perturbation

To further dissect the effect of *DUBR* perturbation in gene regulation, we defined genomic regulatory regions by their activity based on ChromHMM segmentation [49]. In addition, we performed ChIP-seq for the histone marks H3K4me3 and H3K4me1 and analyzed their enrichment ratio to define promoters and enhancers in WT, ΔEX1, and ΔEX2 cells (Supplementary Fig. S5A). Moreover, we generated ChIP-seq data for H3K27ac and H3K27me3 to distinguish active from inactive chromatin regions, respectively [80]. As expected, we observed a high correlation between these histone marks related to active regulatory elements (H3K27ac, H3K4me3, and H3K4me1) (Supplementary Fig. S5B). In line with this observation, we found that H3K27ac was enriched in active promoters, in which H3K4me3 was higher than H3K4me1; and in enhancers, in which H3K4me1 was higher than H3K4me3; in contrast, H3K27ac was depleted in repressed regions, in which H3K27me3 was enriched (Supplementary Fig. S5A). We performed differential H3K27ac enrichment analysis in the ΔEX1 clone due to undetectable *DUBR* transcript levels in contrast to the ΔEX2 clone (Fig. 3A). Interestingly, the H3K27ac profile was altered in ΔEX1 in comparison to WT cells, with a major number of sites gaining H3K27ac (n = 4000 peaks) and a lower number showing a reduction in the mark (n = 1589 peaks) (Supplementary Fig. S5C and Supplementary Table S4). We then assessed the effects associated with altered H3K27ac in promoters marked by H3K4me3, enhancers enriched by H3K4me1, and repressed regions marked by H3K27me3 in ΔEX1 and ΔEX2 compared to WT cells. ChIP-seq revealed that sites with higher H3K27ac also showed increased H3K4me3 and H3K4me1, while decreased H3K27me3 upon DUBR perturbation (Fig. 4A and Supplementary Fig. S5D). Conversely, sites with lower H3K27ac tracked with reduced H3K4me3 and H3K4me1 but increased H3K27me3 in ΔEX1 and ΔEX2 compared to WT. Additionally, ATAC-seq and STARR-seq data showed that these sites with differential H3K27ac deposition overlap accessible chromatin regions with transcriptional activity, suggesting that these represent putative regulatory elements (Supplementary Fig. S5E and F). Notably, some regulatory elements initially inactive may gain H3K27ac in ΔEX1 and ΔEX2 cells, since most of the increased H3K27ac accumulation occurred at regions identified as low signal corresponding to chromatin segments with low ChIP-seq signal for the tested histone marks in WT cells (Fig. 4B). Therefore, *DUBR*

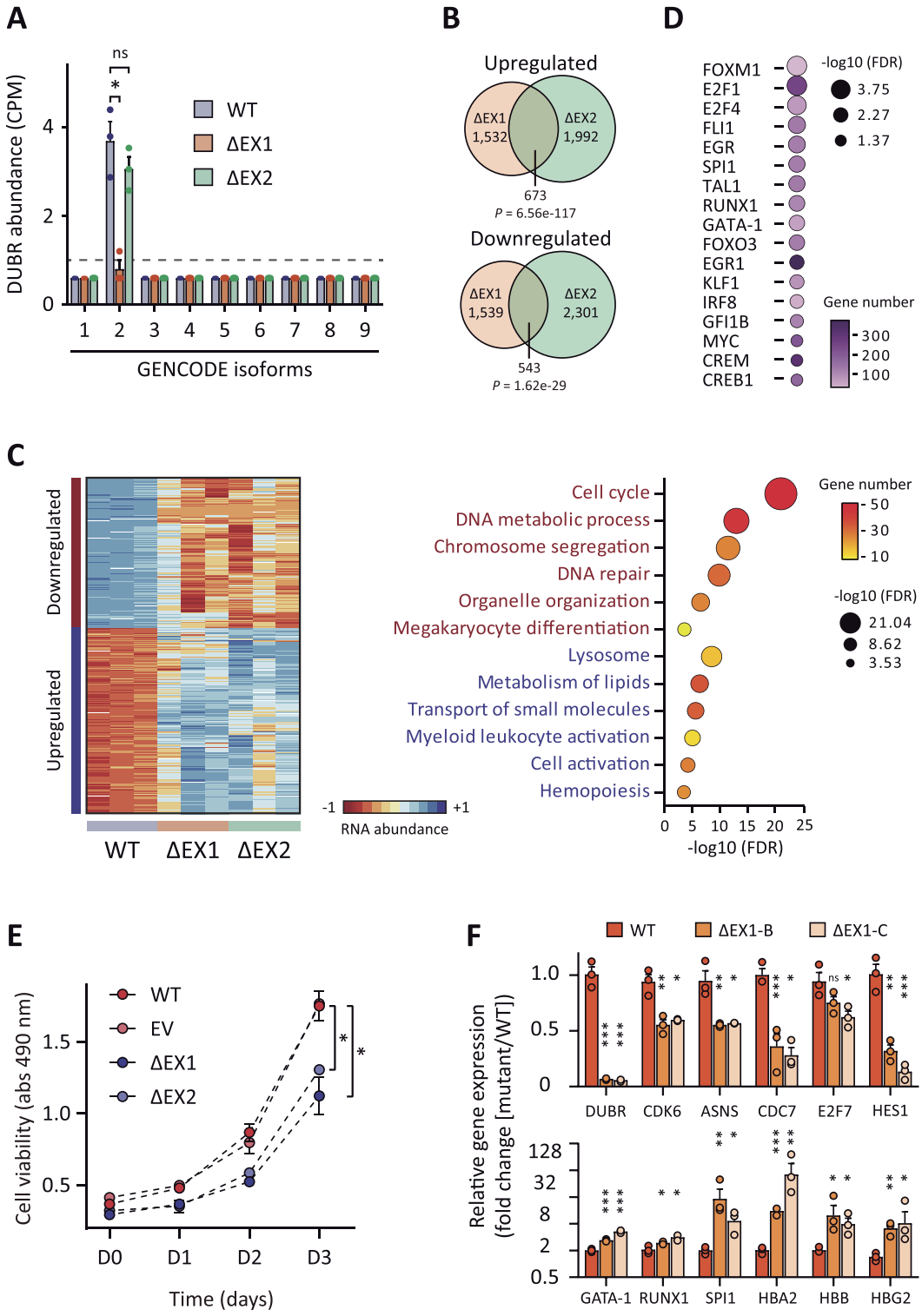


Figure 3. Genome-wide gene expression changes upon *DUBR* ablation. (**A**) RNA abundance of *DUBR* isoforms annotated by GENCODE v36 in K562 cells (WT) and DUBR mutant clones (ΔEX1 and ΔEX2) determined by RNA-seq. The bar chart shows the mean ± SEM of three independent sequenced RNA-seq libraries. Significance was calculated by unpaired Student's *t*-test. (**B**) Venn diagram showing overlapping differentially expressed genes between ΔEX1 and ΔEX2 cell clones as analyzed by RNA-seq. Significance was calculated by a hypergeometric test. (**C**) Heatmap (left) of significant (FDR < 0.05) up- (fold change > 2) and downregulated (fold change < 0.5) genes in ΔEX1 and ΔEX2 clones by RNA-seq. The color key indicates RNA abundance as *z*-score of log2 CPM. Bubble plot (right) of the top significant biological processes (GO terms) identified by gene ontology analysis for the up- (673) and downregulated (543) genes showed in panel (B). Significance (FDR) was determined by the Benjamini–Hochberg test in Metascape [51]. (**D**) Bubble plot of TF enrichment analysis by ENCODE and ChEA3 databases using the up- (673) and downregulated (543) genes in panel (B). Significance (FDR) was calculated by a hypergeometric test. (**E**) MTT proliferation assay of K562 WT, control (EV), and cell clones (ΔEX1 and ΔEX2 clones). The graph shows the mean ± SEM of three independent measurements. Significance was calculated at data points corresponding to day 3 (D3) by unpaired Student's *t*-test. (**F**) RNA abundance determined by RT-qPCR from differentially expressed genes identified in the RNA-seq analysis in two independent ΔEX1 clones (B and C). The bar chart shows the mean ± SEM of three independent experiments. RNA abundance was normalized to the nonedited condition (EV). Significance was determined by unpaired Student's *t*-test as follows; nonsignificant (ns); $*P < .05$; $**P < .01$; $***P < .001$.

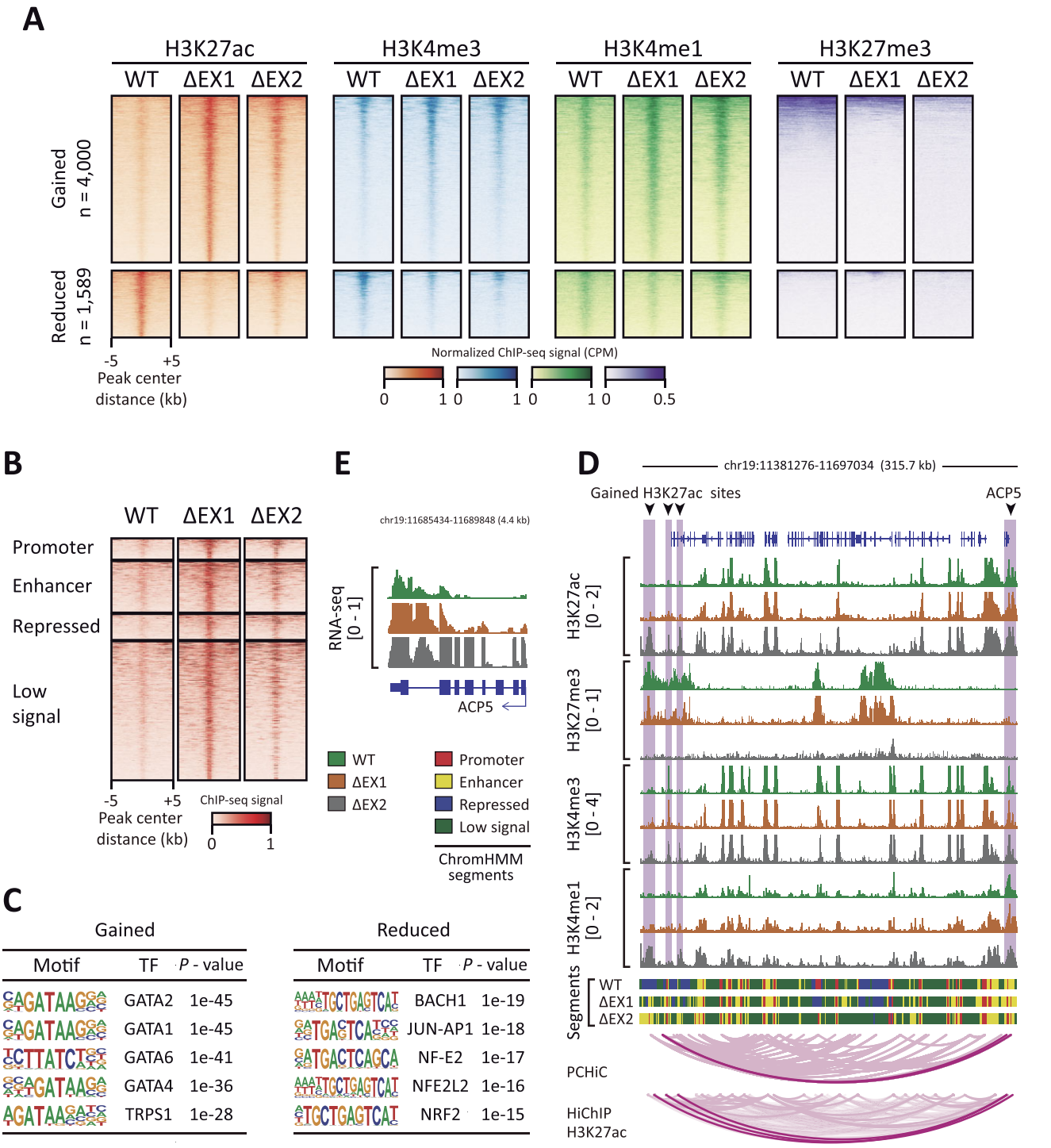


Figure 4. Genome-wide changes in the chromatin landscape upon *DUBR* perturbation. (**A**) Heatmap and profile plots of histone H3K27ac, H3K4me3, H3K4me1, and H3K27me3 ChIP-seq signals at gained and reduced histone H3K27ac peaks in WT, ΔEX1, and ΔEX2 conditions. Heatmaps are ranked by histone H3K27ac signal. (**B**) Heatmaps of histone H3K27ac ChIP-seq signal over given segments (promoter, enhancer, repressed, and low signal) overlapping gained and reduced histone H3K27ac peaks identified upon *DUBR* depletion (WT versus ΔEX1). Chromatin segments were detected by the ChromHMM algorithm. (**C**) Top TF-binding motifs enriched at gained and reduced histone H3K27ac peaks upon DUBR perturbation detected by Homer. Significance was obtained by a hypergeometric distribution test. (**D**) Genomic snapshot of chromatin landscape of *ACP5* loci and its distal-associated regulatory element. Arrowheads and highlighted lines point out the main differences in H3K27ac according to the ChIP-seq signals and ChromHMM segments. *ACP5* promoter-centered interactions identified by PCHi-C and histone H3K27ac HiChIP in K562 cells are shown as arcs. (**E**) Genomic snapshot of *ACP5* gene showing the RNA-seq signal in WT, ΔEX1, and ΔEX2 conditions.

perturbation may induce widespread changes in the activity of regulatory DNA elements affecting hematopoietic differentiation. To assess the functional relevance of these elements, we analyzed their dynamic H3K27ac deposition during erythropoiesis in HSPC, EProg, and EPrec. Interestingly, heatmaps showed increased and decreased H3K27ac signal at gained and reduced H3K27ac sites, respectively, in EProg and EPrec compared with HSPC (Supplementary Fig. S6A).

Analysis of TF motifs revealed a significant enrichment of regulators involved in erythroid differentiation, particularly GATA-1 and GATA-2, at sites with increased H3K27ac marks [81–83]. In contrast, sites with reduced H3K27ac were enriched in motifs for HSPC regulators such as NF-E2 and NRF2 [84, 85] (Fig. 4C). During erythropoiesis, GATA-2-binding sites are replaced by GATA-1 to facilitate erythroid maturation, a process known as GATA factor switching, which occurs between the early (EProg) and late (EPrec) stages [86]. Consistent with this notion, public ChIP-seq data analysis showed GATA-2 binding signal at differential H3K27ac sites in HSPC and EProg but not in EPrec (Supplementary Fig. S6B). In contrast, the GATA-1 signal was enriched mainly in EPrec, with a higher signal at sites with increased H3K27ac upon *DUBR* perturbation rather than at decreased (Supplementary Fig. S6C). Moreover, GATA-1 knockdown in K562 decreased H3K27ac deposition at gained H3K27ac sites (Supplementary Fig. S6D) [23]. This suggests that the sites where H3K27ac changed in response to *DUBR* perturbation may correspond to regulatory elements controlling erythroid differentiation.

DUBR perturbation influences the H3K27ac landscape

Enhancers are key determinants of cell identity during hematopoiesis through interacting with their cognate gene promoters [26]. A small subset of ∼5% of differential H3K27ac sites was at gene promoters, whereas most were distal (predominantly 5–500 kb upstream and downstream from annotated genes) indicating a potential distal-mediated gene regulation (Supplementary Fig. S7A). To assess the function of enhancers activated upon DUBR ablation in hematopoietic gene expression, we analyzed Promoter Capture Hi-C data (PCHi-C) in human blood cells to map loops anchored at regions with gained and reduced H3K27ac upon *DUBR* disruption and link gene promoters to putative enhancers [87]. Notably, a subset of loop anchors was enriched for H3K27ac, depleted for H3K27me3, and structured in accessible chromatin (Supplementary Fig. S7B). This suggests that these loop anchors harbor transcriptionally active regions associated with putative regulatory elements. Moreover, ∼50% of increased H3K27ac sites interacting with anchors scored significantly higher for CHiCAGO interactions more common in erythroblasts (Ery) and megakaryocytes (Mk), indicating that the frequency of these enhancer-promoter interactions tends to be cell-type specific. In contrast, reduced H3K27ac sites interacting with anchors scored higher for interactions in monocytes (Mon) and neutrophils (Neu) (Supplementary Fig. S7C). To assess the regulatory potential of the regions with differential H3K27ac enrichment in controlling long-range gene transcription, we cross-referenced the regions interacting with anchors against genes dysregulated globally upon *DUBR* ablation in ΔEX1 and ΔEX2 cells compared with WT. We observed that some genes whose promoters interact with differential H3K27ac sites were deregulated (182 upregulated and 107 downregulated genes). Examples of upregulated genes include the hematopoietic growth factor (*ACP5*) involved in erythroid commitment, whose promoter interacts with a distal region displaying enhancer activity in ΔEX1 and EX2 (Fig. 5E and D) [88]; the promoter of the lipoprotein receptor-related protein 1 (*LRP1*), a regulator of autophagy during erythroid maturation, was found interacting with downstream enhancer and promoter elements (Supplementary Fig. S7E) [89]; the *NOX5*, an oxidase required for endothelial cell proliferation, was surrounded by several proximal enhancers in ΔEX1 and ΔEX2 (Supplementary Fig. S7F) [90]. In contrast, the reduced gene expression was observed in several genes associated with reduced H3K27ac sites including *HAND2*, a TF required for cell fate specification during embryonic development, and *MSN*, a membrane-associated cytoskeleton protein involved in vascular endothelial development (Supplementary Fig. S7D, G, and H) [91, 92]. Overall, the disruption of *DUBR* had a large effect on the regulatory landscape of K562 cells linked to hematopoietic differentiation-related genes. We reasoned that, due to the widespread effects on the transcriptome and the chromatin landscape upon *DUBR* perturbation, a fraction of these represents direct and indirect effects as a consequence of dysregulation of transcriptional regulators.

The genomic map of RNA occupancy reveals DUBR-associated genes

DUBR was enriched in nuclear fractions, specifically in the chromatin fraction, of K562 cells, suggesting a function in transcriptional gene regulation (Fig. 5A) [93, 94]. To identify the DUBR-associated genes, we mapped genome-wide DUBR–chromatin interactions in K562 cells through chromatin isolation followed by RNA purification (ChIRP-seq) [50]. First, we confirmed that *DUBR*, but not *GAPDH* or *MALAT1*, was specifically retrieved from K562 cells using DUBR EVEN and ODD probes (Fig. 5B). In addition, none of these RNAs were retrieved with a negative control probe targeting LacZ (Fig. 5B). To identify DUBR-occupied regions with high confidence, we exclusively analyzed the ChIRP peaks identified by both probes. There were 2031 overlapped DUBR peaks between EVEN and ODD probes (Fig. 5C and Supplementary Table S5). Using the nearest-neighbor gene approach, we assigned DUBR-bound regions with their potential target genes (DUBR-associated genes) [95]. This analysis showed that DUBR locates distal, predominantly 5–500 kb from TSS genes, whereas only ∼10% of DUBR peaks were at gene promoters (Fig. 5D). Gene ontology analysis revealed that these genes were significantly enriched in processes related to nervous and skeletal system development and cell cycle, which is consistent with previously reported roles of *DUBR* in driving neuronal development, muscle regeneration, and cell division (Supplementary Fig. S8A) [57, 58, 69–71]. Interestingly, a set of genes was enriched in the circulatory system process and cell cycle. By intersecting these DUBR-associated genes with dysregulated genes in ΔEX1 and ΔEX2, we identified only 129 directly DUBR-associated genes (Fig. 5E). Thus, genes whose expression was affected during *DUBR* perturbation but were not assigned to a DUBR peak are indirectly associated genes. We hypothesize that a fraction of the detected DUBR peaks may represent nonfunctional sites in hematopoietic but could be in other tissues, where

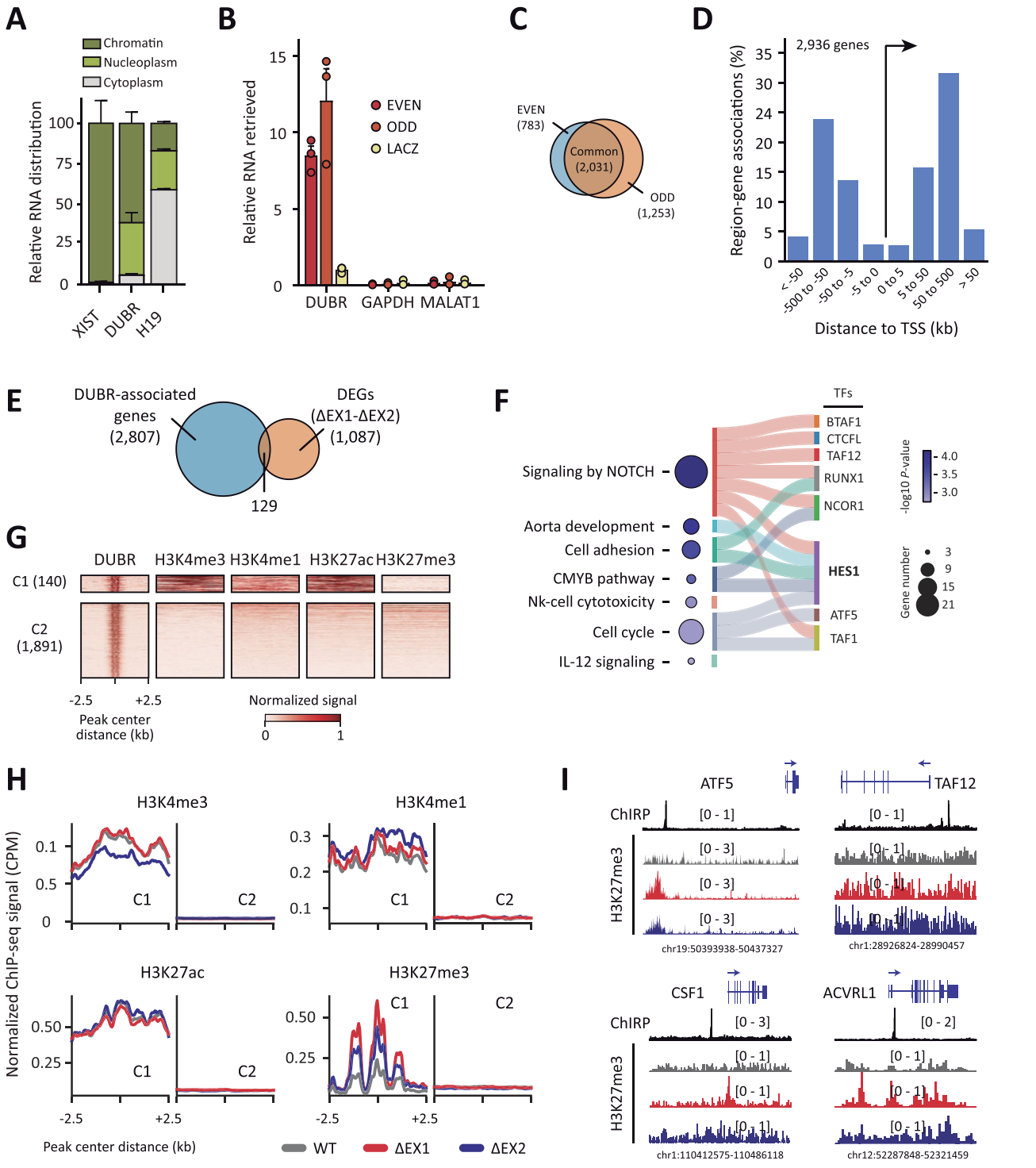


Figure 5. Characterization of *DUBR* genomic occupancy. (**A**) Subcellular distribution of *DUBR* and control lncRNAs (*XIST* and *H19*) determined by RT-qPCR of RNA isolated from subcellular fractionation of K562 cells. The bar chart shows the mean ± SEM of three independent experiments. (**B**) Relative levels of RNA retrieved from *DUBR* and control RNAs (*GAPDH* and *MALAT1*) by ChIRP and determined by RT-qPCR. Two sets of independent (EVEN and ODD) and negative control (LACZ) probes were used. The bar chart shows the mean ± SEM of three independent experiments. (**C**) Venn diagram of common DUBR ChIRP peaks detected in EVEN and ODD probes. (**D**) Relative distances of DUBR ChIRP peaks to TSS of annotated genes by GREAT algorithm. (**E**) Venn diagram showing the overlapped genes between DUBR-associated genes identified by ChIRP-seq and differentially expressed genes identified in ΔEX1 and ΔEX2 compared to WT. Gene–peak associations were determined by the GREAT algorithm [95]. (**F**) Sankey and dot plots showing the top of gene ontology terms (GO) for direct DUBR-associated genes identified in Fig. 5E. Examples of TFs are highlighted in a Sankey diagram. Significance (*P*-value) was determined by the Benjamini–Hochberg test in Metascape [51]. (**G**) Heatmaps of ChIRP-seq and ChIP-seq signals at *DUBR* peaks grouped into two clusters (C1 and C2) according to their signals by the *k*-means algorithm. (**H**) Profile plots of ChIP-seq signal at DUBR peaks grouped in C1 and C2 clusters. (**I**) Genomic snapshots of representative direct DUBR-associated genes showing the signals for ChIRP-seq and ChIP-seq (H3K27me3).

additional or tissue-specific factors are necessary to coordinate the DUBR-mediated gene regulation.

To further explore the possible transcriptional regulatory function of DUBR, we performed an integrative analysis of the DUBR peaks and ChIP-seq for histone modifications in K562 cells. We identified a subset of DUBR peaks (C1, $n = 140$ peaks) clustering with regulatory elements-related histone modifications (H3K4me3, H3K4me1, and H3K27ac), whereas the majority appeared depleted for these histone modifications (C2, $n = 1891$ peaks) (Fig. 5G). This suggests that DUBR associates with chromatin regions independently of its deposited histone marks. Notably, the considerable number of DUBR peaks that were neither assigned to dysregulated genes upon *DUBR* perturbation nor bound to regulatory regions suggests that these DUBR peaks may represent off-target events due to multiple probes used in the ChIRP assay or other experimental artifacts or nonfunctional DUBR sites in K562 [96, 97].

We next compared histone modification signals around DUBR peak clusters (C1 and C2), revealing that only H3K27me3 showed increased signals in ΔEX1 and ΔEX2 compared with WT cells in the C1 DUBR peak cluster (Fig. 5H). Similarly, analysis of public ChIP-seq data in erythroblasts showed a similar H3K27me3 enrichment around DUBR peaks, which was co-occupied by the histone methyltransferase EZH2 (Supplementary Fig. S8B) [98]. These observations suggest that *DUBR* perturbation may facilitate the deposition of H3K27me3 at regulatory elements of directly DUBR-associated genes. This regulatory mechanism has been described for lncRNAs acting as regulators of histone methyltransferase (HMT) activity and recruitment of PRC2 to trimethylate H3K27 on chromatin [96, 99–101]. Thus, the absence of DUBR may enable the HMT of PRC2 on a subset of DUBR sites. However, further work will be necessary to unravel this apparent mechanism by which *DUBR* modulates H3K27me3 deposition at these regions. Among the dysregulated genes with increased H3K27me3 coverage upon DUBR perturbation, including the TFs *ATF5* and *TAF12* involved in the regulation of cell proliferation [102, 103]; the hematopoietic cytokine colony-stimulating factor 1 (*CSF1*) [104]; and the activin protein *ACVRL1* implicated in the hematopoietic endothelial transition (Fig. 5I) [105]. Notably, DUBR was associated with similar coverage at these sites regarding the effect in gene expression upon *DUBR* perturbation in ΔEX1 and ΔEX2 compared with WT cells (Supplementary Figs. S8C and D). This observation indicates that not all the changes in gene expression were explained by the increased deposition of H3K27me3 at their associated DUBR sites. Instead, additional mechanisms may result in modulating changes in gene expression that do not reflect their chromatin status.

HES1 indirectly contributes to defining a DUBR hematopoietic transcriptional program

Gene ontology analysis of directly DUBR-associated genes indicated no evidence of TFs essential for erythroid cell differentiation [106]. Instead, these genes were involved in key pathways critical for the early stages of blood cell development, including the NOTCH pathway (Fig. 5F) [107, 108]. We selected *HES1* as a representative direct DUBR-associated gene involved in NOTCH and CMYB signaling pathways, aorta development, cell adhesion, and cell cycle (Fig. 5F).

HES1 was positively regulated by *DUBR* as validated by CRISPRi and CRISPRa targeting the endogenous *DUBR* promoter using two pairs of sgRNAs (Fig. 6B and C). This effect was further validated by overexpressing *DUBR* through transfecting a plasmid containing the isoform 2 (*DUBR-2*) into K562 cells (Supplementary Fig. S8E). However, the *DUBR* rescue assay in ΔEX1 cells was unsuccessful in recovering the basal *HES1* gene expression compared to WT cells, indicating that the overexpression of *DUBR* alone was unable to revert the loss of *HES1* gene expression upon *DUBR* depletion. Alternatively, *HES1* transcription is repressed by GATA-1 during the transition of EProg to EPrec, suggesting that additional regulators such as GATA-1 are implicated in *HES1* transcription [109]. This observation supports an interplay between *DUBR* and erythroid regulators to coordinate a global transcriptional program. Importantly, *HES1* was positively regulated by *DUBR* through erythroid differentiation from HSPC to erythroblasts (EPrec and D14 days of differentiation) and EProg (HUDEP-2) to erythroblasts (D9) as determined by analysis of public RNA-seq data (Supplementary Fig. S8F) [22]. The decrease in *HES1* gene expression was accompanied by increased H3K27me3 deposition at the DUBR site and *HES1* gene body in ΔEX1 and ΔEX2 compared with WT (Fig. 6A). A similar effect in H3K27me3 deposition and EZH2 binding was observed in erythroblasts (Fig. 6A and Supplementary Fig. S8G). These findings suggested that *DUBR* regulates the expression of the *HES1* gene by associating with its promoter region in the chromatin. However, this potential regulatory mechanism warrants further experimental validation.

HES1 is an essential TF for HSPC development and erythroid differentiation [79, 109–111]. Given these functions in hematopoietic cells, we hypothesized that the widespread dysregulation of gene expression observed following *DUBR* perturbation might be a result of reduced *HES1* levels and its transcriptional activity in hematopoietic genes. To test this possibility, we analyzed the HES1 genome-wide binding in K562 cells revealing that HES1 predominantly bound at regulatory elements structured on accessible chromatin, marked by H3K27ac but not by H3K27me3 (Supplementary Fig. S8H) [112]. Moreover, HES1-binding sites were enriched for GATA-1 and GATA-2 motifs (Supplementary Fig. S8I). Published ChIP-seq data for GATA-1 in K562 demonstrate that a fraction of HES1 sites were cobound with GATA-1 (Supplementary Fig. S8H). Importantly, HES1 interacts with GATA-1 to inhibit the deposition of H3K27ac by dissociating GATA-1 from the P300/CBP complex [111]. As a result of inhibiting the GATA-1 activity, HES1 impairs the development and differentiation of erythroid cells [77, 111]. In line with this evidence, we analyzed the binding of HES1 at differential H3K27ac sites identified upon *DUBR* perturbation. Noticeably, HES1 was enriched at gained and reduced H3K27ac sites where GATA-1 was also bound (Fig. 6D). This analysis suggested that the loss of *HES1* abundance followed by *DUBR* perturbation facilitates the transcriptional activity of GATA-1 by influencing the deposition of H3K27ac at regulatory elements. Heatmaps of ChIRP-seq showed undetectable DUBR signal at differential H3K27ac sites, suggesting that the widespread changes in H3K27ac occupancy followed *DUBR* perturbation were a result of indirect effects such as the activity of transcriptional regulators such as HES1 and GATA-1 (Fig. 6D). Moreover, we observed that 25% of dysregulated genes in ΔEX1 and ΔEX2 compared to WT

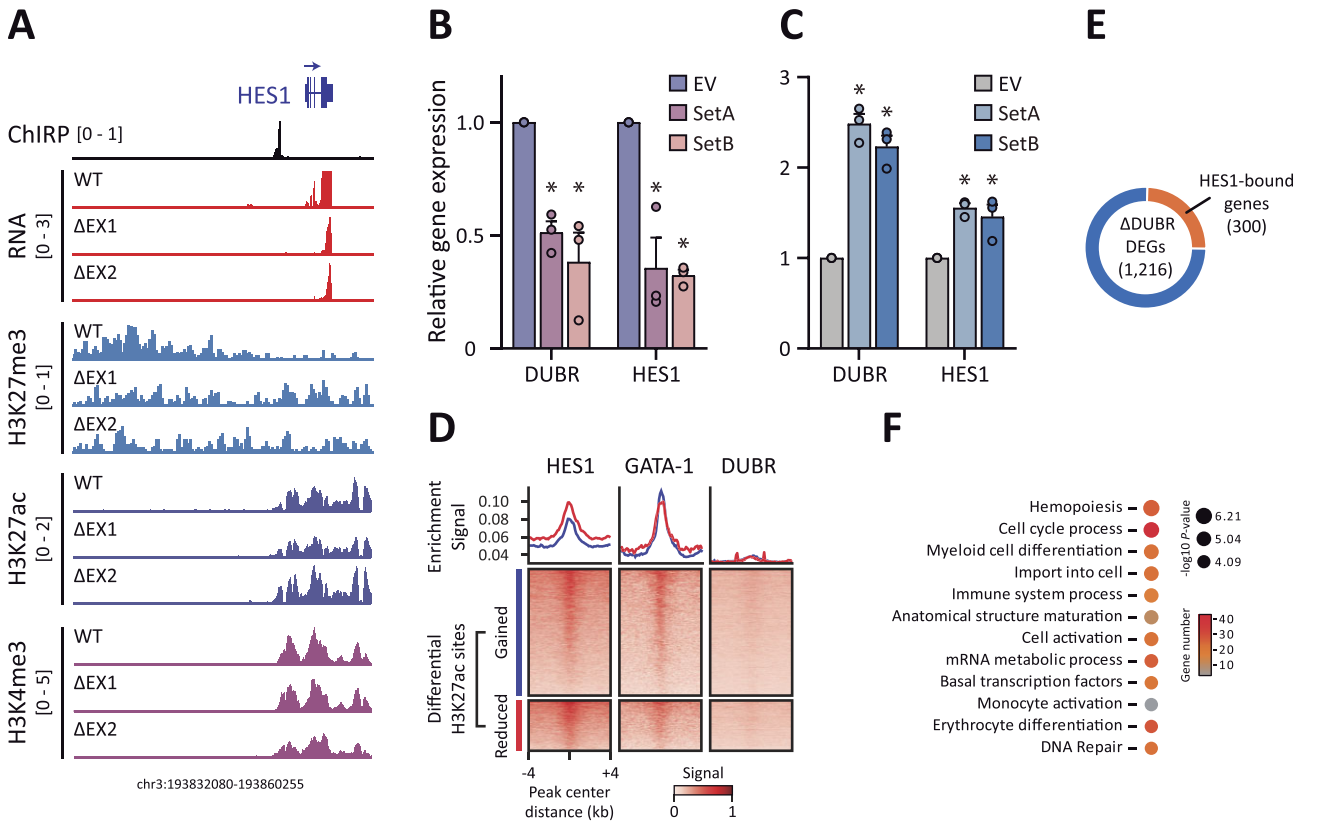


Figure 6. A gene transcriptional program mediated by the interplay between *HES1* and *DUBR*. (**A**) Genomic snapshot of *DUBR* gene landscape showing the signals for ChIRP-seq, RNA-seq, and ChIP-seq in WT, ΔEX1, and ΔEX2 conditions. (**B** and **C**) RT-qPCR analysis for *HES1* and *DUBR* transcripts in K562 cells stably expressing dCas9–KRAB (CRISPRi) and dCas9–SAM (CRISPRa) fusion protein and infected with lentiviral vectors carrying two sets of independent sgRNAs (setA and setB) targeting the *DUBR* promoter or an EV. The bar chart shows the mean ± SEM of three independent experiments. Significance was calculated by unpaired Student's *t*-test (**P* < .05). (**D**) Heatmap and profile plots of ChIP-seq and ChIRP-seq signals at gained and reduced H3K27ac sites upon *DUBR* disruption. Heatmaps are ranked based on the HES1 signal. (**E**) The proportion of dysregulated genes upon *DUBR* disruption (ΔEX1 and ΔEX2) that are targeted by HES1 at their promoter–proximal gene regions (within 5 kb from TSS) are analyzed. (**F**) Bubble plot of biological processes of dysregulated genes upon *DUBR* perturbation and targeted by HES1 at their promoter–proximal gene regions.

were bound by HES1 at their promoter–proximal regions (Fig. 6E). These genes were enriched for gene ontologies related to hematopoiesis, cell cycle, and myeloid–erythroid differentiation (Fig. 6F). Altogether, these data suggest that *HES1* may contribute to defining a DUBR-mediated gene program by indirectly cooperating with *DUBR*.

In addition to *HES1*, heatmaps of public ChIP-seq data in K562 showed that other direct DUBR-associated genes including *NCOR1*, *RUNX1*, and *TAF1* showed binding signals at differential H3K27ac sites identified upon *DUBR* perturbation (Supplementary Fig. S8J). This indicates that additional regulators may contribute to the indirect transcriptional program mediated by *DUBR*.

Discussion

Hematopoietic differentiation is controlled by complex regulatory circuits [113]. Important master hematopoietic regulators such as GATA-1, GATA-2, TAL1, PU.1, and KLF1 [7, 114–116] orchestrate gene expression through binding to regulatory elements (promoter and enhancers) of erythroid genes at different stages of hematopoiesis [1, 26]. The nuclear factor CTCF is dynamically expressed during erythroid differentiation and its downregulation inhibits erythropoiesis [33]. Mechanistically, CTCF mediates the promoter–enhancer interactions to fine-tune the erythroid gene expression program [29, 34]. Importantly, CTCF colocalizes with cell-specific TFs at regulatory elements to facilitate cell identity gene expression [117]. This cooperative function of CTCF is essential to orchestrate cell fate specification and differentiation. Motivated by these findings, we investigated the poorly explored function of CTCF in coordinating lncRNA gene expression during erythropoiesis and how such regulation impacts the erythroid transcriptional gene program.

We demonstrated crosstalk between CTCF and lncRNAs, particularly with *DUBR*. Functionally, CTCF directly binds to an URE to prevent *DUBR* transcription by reducing local chromatin accessibility and H3K27ac deposition. This regulatory effect mediated by CTCF occurs in a chromatin three-dimensional context, in which CTCF configures the *DUBR* loci to coordinate its expression during erythropoiesis (Fig. 7) [118]. The functional characterization of lncRNAs underpins the essential functions of these RNA molecules in shaping gene expression during erythroid differentiation [19, 20, 22, 65]. This suggests an interplay between CTCF and *DUBR* to control a network of genes, in which *DUBR* represents a further level of gene regulation.

DUBR was previously identified as a lncRNA in mouse, where it plays key roles in coordinating myoblast differentiation and neuronal development [57, 58]. These findings and

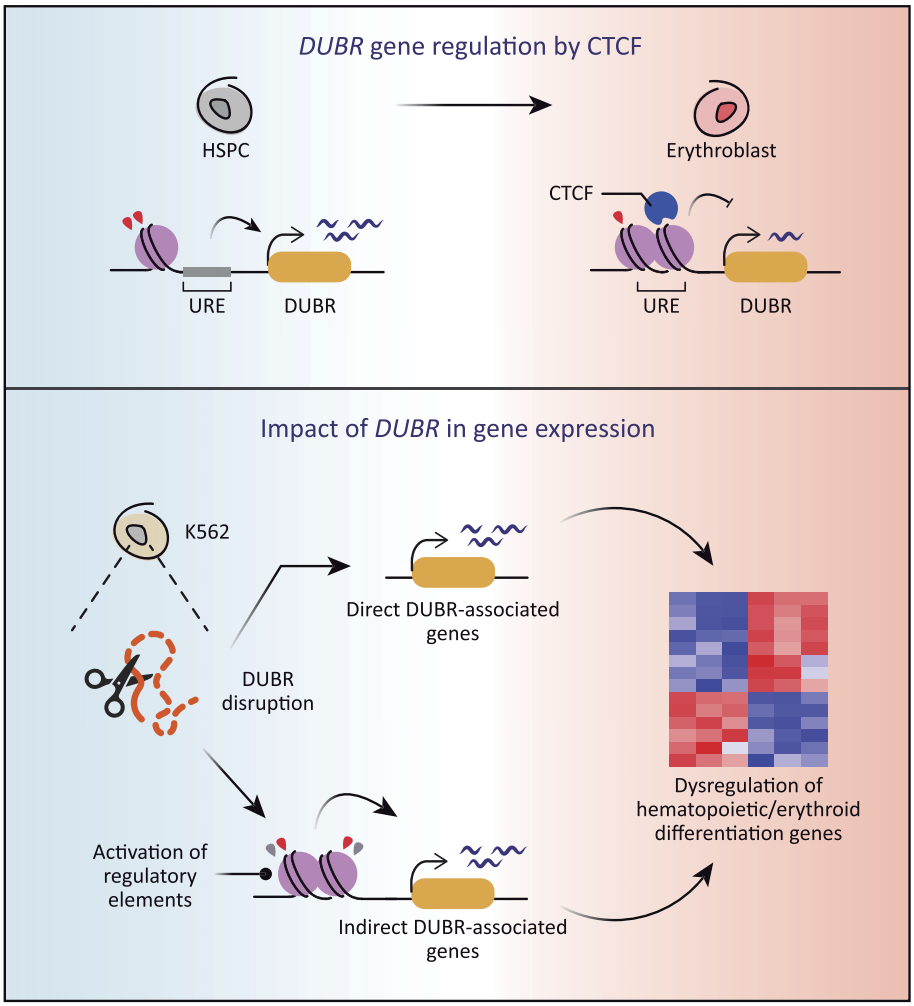


Figure 7. Graphical model of *DUBR* gene regulation by CTCF and its impact in hematopoietic transcriptome. During human erythroid differentiation, CTCF binds to an URE to limit *DUBR* transcription in erythroblasts. *DUBR* perturbation induces widespread changes in hematopoietic–erythroid gene expression as a result of dysregulation of DUBR-associated genes and activation of regulatory elements.

the dynamic expression of *DUBR* throughout the erythropoiesis suggested a functional relevance of this lncRNA in this process. The perturbation of *DUBR* abundance and its sequence result in widespread dysregulation of genes involved in hematopoietic differentiation and cell cycle. Among the induced genes upon DUBR perturbation, we identified essential erythroid regulators, such as GATA-1, *RUNX1*, and *SPI1* as well as the classical globin genes, supporting that *DUBR* perturbation contributes to modulating a hematopoietic–erythroid differentiation gene program.

Analysis of RNA occupancy revealed that DUBR associates with a set of genomic sites including regulatory elements of genes with functions in NOTCH and CMYB pathways and cell cycle, which are essential biological processes in hematopoiesis [108, 119, 120]. Data integration of ChIRP-seq and ChIP-seq suggested that a fraction of DUBR sites were associated with changes in H3K27me3 deposition upon *DUBR* perturbation. Recent evidence strongly suggested that lncRNAs can control the activity of PRC2 by interfering with its recruitment, loading, and eviction from chromatin, and H3K27me3 deposition [99, 121, 122]. In line with this evidence, we propose that the deleted region in ΔEX2, corresponding to an *in silico* predicted PRC2-binding site, and the presence of *DUBR* is required for counteracting H3K27me3 deposition at DUBR sites. It is important to note that only a subset of directly DUBR-associated genes was downregulated upon *DUBR* perturbation, whereas others were upregulated or unaffected indicating that other DUBR sites appear to function by additional mechanisms that warrant further investigation.

Data integration from ChIRP-seq and RNA-seq indicates that only 10% of dysregulated genes following DUBR disruption were directly associated with DUBR sites, while the majority resulted from indirect regulatory effects. We selected *HES1* as a representative direct DUBR-associated gene due to its essential role in hematopoiesis through its function as a transcriptional repressor of erythroid differentiation [110]. Almost 25% of dysregulated genes that followed *DUBR* perturbation were bound by HES1 at its promoter–proximal regions. This indicates that *HES1* is part of the transcriptional network coordinated by *DUBR*.

Erythropoiesis is driven by the activity of master erythroid TFs at their regulatory elements [1]. Dysregulation of gene expression upon *DUBR* perturbation was accompanied by the activation of regulatory elements demarcated by increased H3K27ac deposition at GATA-2/GATA-1 sites, indicating

that these regulatory elements are involved in erythroid commitment [86]. Analysis of ChIP-seq data revealed HES1 but not DUBR binding at these H3K27ac sites, where HES1 may interfere with the activity of GATA-1 [111]. This reinforces our model in which *HES1* indirectly cooperates with *DUBR* in coordinating a hematopoietic–erythroid gene expression program. Thus, *HES1* may represent a layer of gene regulation to coordinate the transcription of the other downstream genes.

Our study systematically characterizes the role of CTCF in regulating the lncRNA *DUBR* and its impact on hematopoietic–erythroid gene expression underlining their implications in the cell differentiation biology.

Acknowledgements

We dedicate this study to the memory of our mentor, Gary Felsenfeld, who passed away on May 1, 2024. We thank Paul Delgado Olguin and Sylvia Patricia Garza Manero for the critical reading of the manuscript. We are grateful to Amaya Abad for NGS library preparation and the Molecular Biology Unit at the Instituto de Fisiología Celular, UNAM for the services provided for the sequencing experiments. H.N.N-M. is a PhD student from Programa de Doctorado en Ciencias Biomédicas (515022416), Universidad Nacional Autónoma de México, and received fellowship 574041 (CVU 661245) from Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCyT).

Author contributions: H.N.N-M. designed, performed, and analyzed most of the experiments. G.T-U. assisted with genome editing assays. J.A.C-C. assisted with ChIP-seq and library preparation. C.A.P-A. assisted with RNA-seq analysis. G.G. performed lentiviral-based experiments. M.H. supervised H.N.N-M. in ChIRP-seq assay. F.R-T. conceived and designed the study, supervised the work, and obtained funding. H.N.N-M. and F.R.-T. wrote the manuscript with the feedback from all authors.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

Consejo Nacional de Humanidades, Ciencias y Tecnologías [Investigación en Ciencia Basica 2017-2018 (A1-S-11844) and Fronteras de la Ciencia 2015 (Proyecto 290)]; Fundación Miguel Alemán A.C. (2018); Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (DGAPA-PAPIIT IN203620 and IN206823); Fondo Europeo de Desarrollo Regional/Ministerio de Ciencia, Innovación y Universidades—Agencia Estatal de Investigación grant BFU2017–82773-P; European Research Council Consolidator grant 771425; and Project PID2023-147980NB-I00 financed by MICIU/AEI /10.13039/501100011033. Funding to pay the Open Access publication charges for this article was provided by the Universidad Nacional Autónoma de México (UNAM-NAM24005).

Data availability

The raw and processed data files generated in this study are available in NCBI GEO with the accession numbers: RNA-seq (GSE268502), ChIP-seq (GSE268500), and ChIRP-seq (GSE268501). Custom scripts used in the analysis are available at https://zenodo.org/records/14767259.

References

1. Georgolopoulos G, Psatha N, Iwata M *et al.* Discrete regulatory modules instruct hematopoietic lineage commitment and differentiation. *Nat Commun* 2021;**12**:6790. https://doi.org/10.1038/s41467-021-27159-x
2. Martin EW, Rodriguez YBA, Reggiardo RE *et al.* Dynamics of chromatin accessibility during hematopoietic stem cell differentiation into progressively lineage-committed progeny. *Stem Cells* 2023;**41**:520–39. https://doi.org/10.1093/stmcls/sxad022
3. Goode DK, Obier N, Vijayabaskar MS *et al.* Dynamic gene regulatory networks drive hematopoietic specification and differentiation. *Dev Cell* 2016;**36**:572–87. https://doi.org/10.1016/j.devcel.2016.01.024
4. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E *et al.* Immunogenetics. Chromatin state dynamics during blood formation. *Science* 2014;**345**:943–49. https://doi.org/10.1126/science.1256271
5. Dzierzak E, Philipsen S. Erythropoiesis: development and differentiation. *Cold Spring Harb Perspect Med* 2013;**3**:a011601–. https://doi.org/10.1101/cshperspect.a011601
6. Wells M, Steiner L. Epigenetic and transcriptional control of erythropoiesis. *Front Genet* 2022;**13**:805265. https://doi.org/10.3389/fgene.2022.805265
7. Hattangadi SM, Wong P, Zhang L *et al.* From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood* 2011;**118**:6258–68. https://doi.org/10.1182/blood-2011-07-356006
8. Statello L, Guo CJ, Chen LL *et al.* Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* 2021;**22**:96–118. https://doi.org/10.1038/s41580-020-00315-9
9. Li W, Ren Y, Si Y *et al.* Long non-coding RNAs in hematopoietic regulation. *Cell Regener* 2018;**7**:27–32. https://doi.org/10.1016/j.cr.2018.08.001
10. Schwarzer A, Emmrich S, Schmidt F *et al.* The non-coding RNA landscape of human hematopoiesis and leukemia. *Nat Commun* 2017;**8**:218. https://doi.org/10.1038/s41467-017-00212-4
11. Alvarez-Dominguez JR, Lodish HF. Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. *Blood* 2017;**130**:1965–75. https://doi.org/10.1182/blood-2017-06-788695
12. Delas MJ, Jackson BT, Kovacevic T *et al.* lncRNA Spehd regulates hematopoietic stem and progenitor cells and is required for multilineage differentiation. *Cell Rep* 2019;**27**:719–729. https://doi.org/10.1016/j.celrep.2019.03.080
13. Luo M, Jeong M, Sun D *et al.* Long non-coding RNAs control hematopoietic stem cell function. *Cell Stem Cell* 2015;**16**:426–38. https://doi.org/10.1016/j.stem.2015.02.002
14. Alvarez-Dominguez JR, Hu W, Yuan B *et al.* Global discovery of erythroid long noncoding RNAs reveals novel regulators of red cell maturation. *Blood* 2014;**123**:570–81. https://doi.org/10.1182/blood-2013-10-530683
15. Arriaga-Canon C, Fonseca-Guzman Y, Valdes-Quezada C *et al.* A long non-coding RNA promotes full activation of adult gene expression in the chicken alpha-globin domain. *Epigenetics* 2014;**9**:173–81. https://doi.org/10.4161/epi.27030
16. Paralkar VR, Mishra T, Luan J *et al.* Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic

development. *Blood* 2014;**123**:1927–37. https://doi.org/10.1182/blood-2013-12-544494
17. Xu C, Shi L. Long non-coding RNAs during normal erythropoiesis. *Blood Sci* 2019;**1**:137–40. https://doi.org/10.1097/BS9.0000000000000027
18. Alvarez-Dominguez JR, Knoll M, Gromatzky AA *et al.* The super-enhancer-derived alncRNA-EC7/bloodlinc potentiates red blood cell development in *trans*. *Cell Rep* 2017;**19**:2503–14. https://doi.org/10.1016/j.celrep.2017.05.082
19. Yang S, Sun G, Wu P *et al.* WDR82-binding long noncoding RNA lncEry controls mouse erythroid differentiation and maturation. *J Exp Med* 2022;**219**. https://doi.org/10.1084/jem.20211688
20. Liu J, Li Y, Tong J *et al.* Long non-coding RNA-dependent mechanism to regulate heme biosynthesis and erythrocyte development. *Nat Commun* 2018;**9**:4386. https://doi.org/10.1038/s41467-018-06883-x
21. Wang C, Wu X, Shen F *et al.* Shlnc-EC6 regulates murine erythroid enucleation by Rac1–PIP5K pathway. *Dev Growth Differ* 2015;**57**:466–73. https://doi.org/10.1111/dgd.12225
22. Liu G, Kim J, Nguyen N *et al.* Long noncoding RNA GATA2AS influences human erythropoiesis by transcription factor and chromatin landscape modulation. *Blood* 2024;**143**:2300–13. https://doi.org/10.1182/blood.2023021287
23. Kim YW, Kang J, Kim A. Hematopoietic/erythroid enhancers activate nearby target genes by extending histone H3K27ac and transcribing intergenic RNA. *FASEB J* 2023;**37**:e22870. https://doi.org/10.1096/fj.202201891R
24. Romano O, Petiti L, Felix T *et al.* GATA factor-mediated gene regulation in human erythropoiesis. *iScience* 2020;**23**:101018. https://doi.org/10.1016/j.isci.2020.101018
25. Kim YW, Yun WJ, Kim A. Erythroid activator NF-E2, TAL1 and KLF1 play roles in forming the LCR HSs in the human adult beta-globin locus. *Int J Biochem Cell Biol* 2016;**75**:45–52. https://doi.org/10.1016/j.biocel.2016.03.013
26. Huang J, Liu X, Li D *et al.* Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. *Dev Cell* 2016;**36**:9–23. https://doi.org/10.1016/j.devcel.2015.12.014
27. Kang Y, Kim YW, Yun J *et al.* KLF1 stabilizes GATA-1 and TAL1 occupancy in the human beta-globin locus. *Biochim Biophys Acta* 2015;**1849**:282–89. https://doi.org/10.1016/j.bbagrm.2014.12.010
28. Li D, Zhao XY, Zhou S *et al.* Multidimensional profiling reveals GATA1-modulated stage-specific chromatin states and functional associations during human erythropoiesis. *Nucleic Acids Res* 2023;**51**:6634–53. https://doi.org/10.1093/nar/gkad468
29. Qi Q, Cheng L, Tang X *et al.* Dynamic CTCF binding directly mediates interactions among *cis*-regulatory elements essential for hematopoiesis. *Blood* 2021;**137**:1327–39. https://doi.org/10.1182/blood.2020005780
30. Kim YW, Kang Y, Kang J *et al.* GATA-1-dependent histone H3K27 acetylation mediates erythroid cell-specific chromatin interaction between CTCF sites. *FASEB J* 2020;**34**:14736–49. https://doi.org/10.1096/fj.202001526R
31. Kang Y, Kim YW, Kang J *et al.* Erythroid specific activator GATA-1-dependent interactions between CTCF sites around the beta-globin locus. *Biochim Biophys Acta* 2017;**1860**:416–26. https://doi.org/10.1016/j.bbagrm.2017.01.013
32. Arzate-Mejia RG, Recillas-Targa F, Corces VG. Developing in 3D: the role of CTCF in cell differentiation. *Development* 2018;**145**:dev137729. https://doi.org/10.1242/dev.137729
33. Torrano V, Chernukhin I, Docquier F *et al.* CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *J Biol Chem* 2005;**280**:28152–61. https://doi.org/10.1074/jbc.M501481200
34. Steiner LA, Schulz V, Makismova Y *et al.* CTCF and CohesinSA-1 mark active promoters and boundaries of repressive chromatin domains in primary human erythroid cells. *PLoS One* 2016;**11**:e0155378. https://doi.org/10.1371/journal.pone.0155378
35. Lee J, Krivega I, Dale RK *et al.* The LDB1 complex co-opts CTCF for erythroid lineage-specific long-range enhancer interactions. *Cell Rep* 2017;**19**:2490–502. https://doi.org/10.1016/j.celrep.2017.05.072
36. Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* 2018;**46**:W242–45. https://doi.org/10.1093/nar/gky354
37. Shalem O, Sanjana NE, Hartenian E *et al.* Genome-scale CRISPR–Cas9 knockout screening in human cells. *Science* 2014;**343**:84–7. https://doi.org/10.1126/science.1247005
38. Mayer A, Churchman LS. A detailed protocol for subcellular RNA sequencing (subRNA-seq). *CP Mol Biol* 2017;**120**:4 29 21–24 29 18. https://doi.org/10.1002/cpmb.44
39. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 2008;**3**:1101–8. https://doi.org/10.1038/nprot.2008.73
40. Dobin A, Davis CA, Schlesinger F *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15–21. https://doi.org/10.1093/bioinformatics/bts635
41. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;**30**:923–30. https://doi.org/10.1093/bioinformatics/btt656
42. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–40. https://doi.org/10.1093/bioinformatics/btp616
43. Ramirez F, Ryan DP, Gruning B *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;**44**:W160–65. https://doi.org/10.1093/nar/gkw257
44. Blecher-Gonen R, Barnett-Itzhaki Z, Jaitin D *et al.* High-throughput chromatin immunoprecipitation for genome-wide mapping of *in vivo* protein–DNA interactions and epigenomic states. *Nat Protoc* 2013;**8**:539–54. https://doi.org/10.1038/nprot.2013.023
45. Langmead B, Trapnell C, Pop M *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;**10**:R25. https://doi.org/10.1186/gb-2009-10-3-r25
46. Breese MR, Liu Y. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics* 2013;**29**:494–96. https://doi.org/10.1093/bioinformatics/bts731
47. Zhang Y, Liu T, Meyer CA *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;**9**:R137. https://doi.org/10.1186/gb-2008-9-9-r137
48. Heinz S, Benner C, Spann N *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;**38**:576–89. https://doi.org/10.1016/j.molcel.2010.05.004
49. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 2012;**9**:215–16. https://doi.org/10.1038/nmeth.1906
50. Chu C, Qu K, Zhong FL *et al.* Genomic maps of long noncoding RNA occupancy reveal principles of RNA–chromatin interactions. *Mol Cell* 2011;**44**:667–78. https://doi.org/10.1016/j.molcel.2011.08.027
51. Zhou Y, Zhou B, Pache L *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019;**10**:1523. https://doi.org/10.1038/s41467-019-09234-6
52. Keenan AB, Torre D, Lachmann A *et al.* ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res* 2019;**47**:W212–24. https://doi.org/10.1093/nar/gkz446

53. Su MY, Steiner LA, Bogardus H *et al.* Identification of biologically relevant enhancers in human erythroid cells. *J Biol Chem* 2013;**288**:8433–44. https://doi.org/10.1074/jbc.M112.413260
54. Frankish A, Carbonell-Sala S, Diekhans M *et al.* GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res* 2023;**51**:D942–49. https://doi.org/10.1093/nar/gkac1071
55. Hezroni H, Koppstein D, Schwartz MG *et al.* Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep* 2015;**11**:1110–22. https://doi.org/10.1016/j.celrep.2015.04.023
56. Lewandowski JP, Lee JC, Hwang T *et al.* The Firre locus produces a *trans*-acting RNA molecule that functions in hematopoiesis. *Nat Commun* 2019;**10**:5137. https://doi.org/10.1038/s41467-019-12970-4
57. Wang L, Zhao Y, Bao X *et al.* LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. *Cell Res* 2015;**25**:335–50. https://doi.org/10.1038/cr.2015.21
58. Huang J, Jiang B, Li GW *et al.* m(6)A-modified lincRNA Dubr is required for neuronal development by stabilizing YTHDF1/3 and facilitating mRNA translation. *Cell Rep* 2022;**41**:111693. https://doi.org/10.1016/j.celrep.2022.111693
59. Roadmap Epigenomics C, Kundaje A, Meuleman W *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* 2015;**518**:317–30.
60. Corces MR, Buenrostro JD, Wu B *et al.* Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet* 2016;**48**:1193–203. https://doi.org/10.1038/ng.3646
61. Konermann S, Brigham MD, Trevino AE *et al.* Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex. *Nature* 2015;**517**:583–88. https://doi.org/10.1038/nature14136
62. Larson MH, Gilbert LA, Wang X *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* 2013;**8**:2180–96. https://doi.org/10.1038/nprot.2013.132
63. Li G, Ruan X, Auerbach RK *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 2012;**148**:84–98. https://doi.org/10.1016/j.cell.2011.12.014
64. Hammal F, de Langen P, Bergon A *et al.* ReMap 2022: a database of human, mouse, drosophila and arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res* 2022;**50**:D316–25.
65. Ren Y, Zhu J, Han Y *et al.* Regulatory association of long noncoding RNAs and chromatin accessibility facilitates erythroid differentiation. *Blood Adv* 2021;**5**:5396–409. https://doi.org/10.1182/bloodadvances.2021005167
66. Xiong F, Wang R, Lee JH *et al.* RNA m(6)A modification orchestrates a LINE-1-host interaction that facilitates retrotransposition and contributes to long gene vulnerability. *Cell Res* 2021;**31**:861–85. https://doi.org/10.1038/s41422-021-00515-8
67. Tu S, Yuan GC, Shao Z. The PRC2-binding long non-coding RNAs in human and mouse genomes are associated with predictive sequence features. *Sci Rep* 2017;**7**:41669. https://doi.org/10.1038/srep41669
68. Wilson NK, Foster SD, Wang X *et al.* Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* 2010;**7**:532–44. https://doi.org/10.1016/j.stem.2010.07.016
69. Stojic L, Lun ATL, Mascalchi P *et al.* A high-content RNAi screen reveals multiple roles for long noncoding RNAs in cell division. *Nat Commun* 2020;**11**:1851. https://doi.org/10.1038/s41467-020-14978-7
70. Zhu S, Li W, Liu J *et al.* Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR–Cas9 library. *Nat Biotechnol* 2016;**34**:1279–86. https://doi.org/10.1038/nbt.3715
71. Li Y, Gao X. LINC00883 promotes drug resistance of glioma through a microRNA-136/NEK1-dependent mechanism. *Front Oncol* 2021;**11**:692265. https://doi.org/10.3389/fonc.2021.692265
72. Mayer IM, Doma E, Klampfl T *et al.* Kinase-inactivated CDK6 preserves the long-term functionality of adult hematopoietic stem cells. *Blood* 2024;**144**:156–70. https://doi.org/10.1182/blood.2023021985
73. Scheicher R, Hoelbl-Kovacic A, Bellutti F *et al.* CDK6 as a key regulator of hematopoietic and leukemic stem cell activation. *Blood* 2015;**125**:90–101. https://doi.org/10.1182/blood-2014-06-584417
74. Qi L, Martin-Sandoval MS, Merchant S *et al.* Aspartate availability limits hematopoietic stem cell function during hematopoietic regeneration. *Cell Stem Cell* 2021;**28**:1982–99. https://doi.org/10.1016/j.stem.2021.07.011
75. Martin JC, Sims JR, Gupta A *et al.* CDC7 kinase (DDK) inhibition disrupts DNA replication leading to mitotic catastrophe in Ewing sarcoma. *Cell Death Discov* 2022;**8**:85. https://doi.org/10.1038/s41420-022-00877-x
76. Westendorp B, Mokry M, Groot Koerkamp MJ *et al.* E2F7 represses a network of oscillating cell cycle genes to control S-phase progression. *Nucleic Acids Res* 2012;**40**:3511–23. https://doi.org/10.1093/nar/gkr1203
77. Zhu AZ, Ma Z, Wolff EV *et al.* HES1 is required for mouse fetal hematopoiesis. *Stem Cell Res Ther* 2024;**15**:235. https://doi.org/10.1186/s13287-024-03836-8
78. Ma Z, Xu J, Wu L *et al.* Hes1 deficiency causes hematopoietic stem cell exhaustion. *Stem Cells* 2020;**38**:756–68. https://doi.org/10.1002/stem.3169
79. Yu X, Alder JK, Chun JH *et al.* HES1 inhibits cycling of hematopoietic progenitor cells via DNA binding. *Stem Cells* 2006;**24**:876–88. https://doi.org/10.1634/stemcells.2005-0598
80. Lavarone E, Barbieri CM, Pasini D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat Commun* 2019;**10**:1679. https://doi.org/10.1038/s41467-019-09624-w
81. Menendez-Gonzalez JB, Vukovic M, Abdelfattah A *et al.* Gata2 as a crucial regulator of stem cells in adult hematopoiesis and acute myeloid leukemia. *Stem Cell Rep.* 2019;**13**:291–306. https://doi.org/10.1016/j.stemcr.2019.07.005
82. Tremblay M, Sanchez-Ferras O, Bouchard M. GATA transcription factors in development and disease. *Development* 2018;**145**:dev164384. https://doi.org/10.1242/dev.164384
83. Iwasaki H, Mizuno S, Wells RA *et al.* GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* 2003;**19**:451–62. https://doi.org/10.1016/S1074-7613(03)00242-5
84. Di Tullio A, Passaro D, Rouault-Pierre K *et al.* Nuclear factor erythroid 2 regulates human HSC self-renewal and T cell differentiation by preventing NOTCH1 activation. *Stem Cell Rep* 2017;**9**:5–11. https://doi.org/10.1016/j.stemcr.2017.05.027
85. Tsai JJ, Dudakov JA, Takahashi K *et al.* Nrf2 regulates haematopoietic stem cell function. *Nat Cell Biol* 2013;**15**:309–16.
86. Suzuki M, Kobayashi-Osaki M, Tsutsumi S *et al.* GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation. *Genes Cells* 2013;**18**:921–33. https://doi.org/10.1111/gtc.12086
87. Javierre BM, Burren OS, Wilder SP *et al.* Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* 2016;**167**:1369–84. https://doi.org/10.1016/j.cell.2016.09.037

88. Bazer FW, Worthington-White D, Fliss MF *et al.* Uteroferrin: a progesterone-induced hematopoietic growth factor of uterine origin. *Exp Hematol* 1991;**19**:910–15.
89. Grosso RA, Caldarone PVS, Sanchez MC *et al.* Hemin induces autophagy in a leukemic erythroblast cell line through the LRP1 receptor. *Biosci Rep* 2019;**39**:BSR20181156. https://doi.org/10.1042/BSR20181156
90. BelAiba RS, Djordjevic T, Petry A *et al.* NOX5 variants are functionally active in endothelial cells. *Free Radical Biol Med* 2007;**42**:446–59. https://doi.org/10.1016/j.freeradbiomed.2006.10.054
91. Liu N, Barbosa AC, Chapman SL *et al.* DNA binding-dependent and -independent functions of the Hand2 transcription factor during mouse embryogenesis. *Development* 2009;**136**:933–42. https://doi.org/10.1242/dev.034025
92. Chen Y, Wang J, Zhang L *et al.* Moesin is a novel biomarker of endothelial injury in sepsis. *J Immunol Res* 2021;**2021**:6695679. https://doi.org/10.1155/2021/6695679
93. Werner MS, Sullivan MA, Shah RN *et al.* Chromatin-enriched lncRNAs can act as cell-type specific activators of proximal gene transcription. *Nat Struct Mol Biol* 2017;**24**:596–603. https://doi.org/10.1038/nsmb.3424
94. Simon MD, Pinter SF, Fang R *et al.* High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* 2013;**504**:465–69. https://doi.org/10.1038/nature12719
95. McLean CY, Bristor D, Hiller M *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010;**28**:495–501. https://doi.org/10.1038/nbt.1630
96. Deforzh E, Kharel P, Zhang Y *et al.* HOXDeRNA activates a cancerous transcription program and super enhancers via genome-wide binding. *Mol Cell* 2024; **84**:3950–66. https://doi.org/10.1016/j.molcel.2024.09.018
97. Pan H, Wang H, Zhang X *et al.* Chromosomal instability-associated MAT1 lncRNA insulates MLL1-guided histone methylation and accelerates tumorigenesis. *Cell Rep* 2022;**41**:3950–66. https://doi.org/10.1016/j.celrep.2022.111829
98. Qin K, Lan X, Huang P *et al.* Molecular basis of polycomb group protein-mediated fetal hemoglobin repression. *Blood* 2023;**141**:2756–70. https://doi.org/10.1182/blood.2022019578
99. Lee YW, Weissbein U, Blum R *et al.* G-quadruplex folding in Xist RNA antagonizes PRC2 activity for stepwise regulation of X chromosome inactivation. *Mol Cell* 2024;**84**:1870–85. https://doi.org/10.1016/j.molcel.2024.04.015
100. Teo WW, Cao X, Wu CS *et al.* Non-coding RNA LEVER sequestration of PRC2 can mediate long range gene regulation. *Commun Biol* 2022;**5**:343. https://doi.org/10.1038/s42003-022-03250-x
101. Kang X, Zhao Y, Van Arsdell G *et al.* Ppp1r1b-lncRNA inhibits PRC2 at myogenic regulatory genes to promote cardiac and skeletal muscle development in mouse and human. *RNA* 2020;**26**:481–91. https://doi.org/10.1261/rna.073692.119
102. Zhang S, Chen JJ. Requirement of activating transcription factor 5 for murine fetal liver erythropoiesis. *Br J Haematol* 2020;**188**:582–85. https://doi.org/10.1111/bjh.16202
103. Xu Y, Milazzo JP, Somerville TDD *et al.* A TFIID-SAGA perturbation that targets MYB and suppresses acute myeloid leukemia. *Cancer Cell* 2018;**33**:13–28. https://doi.org/10.1016/j.ccell.2017.12.002
104. Rieger MA, Hoppe PS, Smejkal BM *et al.* Hematopoietic cytokines can instruct lineage choice. *Science* 2009;**325**:217–18. https://doi.org/10.1126/science.1171461
105. Lempereur A, Canto PY, Richard C *et al.* The TGFbeta pathway is a key player for the endothelial-to-hematopoietic transition in the embryonic aorta. *Dev Biol* 2018;**434**:292–303. https://doi.org/10.1016/j.ydbio.2017.12.006
106. Dore LC, Crispino JD. Transcription factor networks in erythroid cell and megakaryocyte development. *Blood* 2011;**118**:231–39. https://doi.org/10.1182/blood-2011-04-285981
107. Huang C, Yang D, Ye GW *et al.* Vascular Notch signaling in stress hematopoiesis. *Front Cell Dev Biol* 2020;**8**:606448. https://doi.org/10.3389/fcell.2020.606448
108. Lampreia FP, Carmelo JG, Anjos-Afonso F. Notch signaling in the regulation of hematopoietic stem cell. *Curr Stem Cell Rep* 2017;**3**:202–9.
109. Ross J, Mavoungou L, Bresnick EH *et al.* GATA-1 utilizes Ikaros and polycomb repressive complex 2 to suppress Hes1 and to promote erythropoiesis. *Mol Cell Biol* 2012;**32**:3624–38. https://doi.org/10.1128/MCB.00163-12
110. Lee JB, Werbowetski-Ogilvie TE, Lee JH *et al.* Notch-HES1 signaling axis controls hemato-endothelial fate decisions of human embryonic and induced pluripotent stem cells. *Blood* 2013;**122**:1162–73. https://doi.org/10.1182/blood-2012-12-471649
111. Ishiko E, Matsumura I, Ezoe S *et al.* Notch signals inhibit the development of erythroid/megakaryocytic cells by suppressing GATA-1 activity through the induction of HES1. *J Biol Chem* 2005;**280**:4929–39. https://doi.org/10.1074/jbc.M406788200
112. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**:57–74. https://doi.org/10.1038/nature11247
113. Laurenti E, Gottgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* 2018;**553**:418–26.
114. Vagapova ER, Spirin PV, Lebedev TD *et al.* The role of TAL1 in hematopoiesis and leukemogenesis. *Acta Naturae* 2018;**10**:15–23. https://doi.org/10.32607/20758251-2018-10-1-15-23
115. Burda P, Laslo P, Stopka T. The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. *Leukemia* 2010;**24**:1249–57. https://doi.org/10.1038/leu.2010.104
116. Zhu J, Emerson SG. Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene* 2002;**21**:3295–313. https://doi.org/10.1038/sj.onc.1205318
117. Liu Y, Wan X, Li H *et al.* CTCF coordinates cell fate specification via orchestrating regulatory hubs with pioneer transcription factors. *Cell Rep* 2023;**42**:113259. https://doi.org/10.1016/j.celrep.2023.113259
118. Hnisz D, Day DS, Young RA. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* 2016;**167**:1188–200. https://doi.org/10.1016/j.cell.2016.10.024
119. Treichel S, Filippi MD. Linking cell cycle to hematopoietic stem cell fate decisions. *Front Cell Dev Biol* 2023;**11**:1231735. https://doi.org/10.3389/fcell.2023.1231735
120. Wang X, Angelis N, Thein SL. MYB—a regulatory factor in hematopoiesis. *Gene* 2018;**665**:6–17. https://doi.org/10.1016/j.gene.2018.04.065
121. Lee Y, Lee JT. PRC2–RNA interactions: viewpoint from YongWoo Lee and Jeannie T. Lee. *Mol Cell* 2024;**84**:3586–92. https://doi.org/10.1016/j.molcel.2024.09.006
122. Deforzh E, Kharel P, Zhang Y *et al.* HOXDeRNA activates a cancerous transcription program and super enhancers via genome-wide binding. *Mol Cell* 2024;**84**:3950–66. https://doi.org/10.1016/j.molcel.2024.09.018

Received: December 18, 2023. **Revised:** January 24, 2025. **Editorial Decision:** January 28, 2025. **Accepted:** February 3, 2025
© The Author(s) 2025. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.