

RESEARCH ARTICLE

CTCF and Cohesin^{SA-1} Mark Active Promoters and Boundaries of Repressive Chromatin Domains in Primary Human Erythroid Cells

Laurie A. Steiner^{1*}, Vincent Schulz², Yelena Makismova², Kimberly Lezon-Geyda², Patrick G. Gallagher^{2,3}

1 Department of Pediatrics, University of Rochester, Rochester, New York, United States of America, **2** Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut, United States of America, **3** Departments of Pathology and Genetics, Yale University School of Medicine, New Haven, Connecticut, United States of America

* Laurie_Steiner@urmc.rochester.edu



OPEN ACCESS

Citation: Steiner LA, Schulz V, Makismova Y, Lezon-Geyda K, Gallagher PG (2016) CTCF and Cohesin^{SA-1} Mark Active Promoters and Boundaries of Repressive Chromatin Domains in Primary Human Erythroid Cells. PLoS ONE 11(5): e0155378. doi:10.1371/journal.pone.0155378

Editor: Hodaka Fujii, Osaka University, JAPAN

Received: February 9, 2016

Accepted: April 27, 2016

Published: May 24, 2016

Copyright: © 2016 Steiner et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available via GEO (accession number GSE67893).

Funding: This work was supported in part by National Institutes of Health Grants K12HD000850 (LAS), HL65448 (PGG), and HL106184 (PGG). There was no additional external support received for this funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Background

CTCF and cohesin^{SA-1} are regulatory proteins involved in a number of critical cellular processes including transcription, maintenance of chromatin domain architecture, and insulator function. To assess changes in the CTCF and cohesin^{SA-1} interactomes during erythropoiesis, chromatin immunoprecipitation coupled with high throughput sequencing and mRNA transcriptome analyses via RNA-seq were performed in primary human hematopoietic stem and progenitor cells (HSPC) and primary human erythroid cells from single donors.

Results

Sites of CTCF and cohesin^{SA-1} co-occupancy were enriched in gene promoters in HSPC and erythroid cells compared to single CTCF or cohesin sites. Cell type-specific CTCF sites in erythroid cells were linked to highly expressed genes, with the opposite pattern observed in HSPCs. Chromatin domains were identified by ChIP-seq with antibodies against trimethylated lysine 27 histone H3, a modification associated with repressive chromatin. Repressive chromatin domains increased in both number and size during hematopoiesis, with many more repressive domains in erythroid cells than HSPCs. CTCF and cohesin^{SA-1} marked the boundaries of these repressive chromatin domains in a cell-type specific manner.

Conclusion

These genome wide data, changes in sites of protein occupancy, chromatin architecture, and related gene expression, support the hypothesis that CTCF and cohesin^{SA-1} have multiple roles in the regulation of gene expression during erythropoiesis including transcriptional regulation at gene promoters and maintenance of chromatin architecture. These

data from primary human erythroid cells provide a resource for studies of normal and perturbed erythropoiesis.

Introduction

The dynamic interplay between DNA methylation, histone modification, and chromatin structure are critical for establishing and maintaining appropriate patterns of mammalian gene expression. In vertebrates, the highly conserved, multifunctional CTCF-binding factor CTCF binds throughout the genome in a sequence-[\[1\]](#) and DNA methylation-specific manner. [\[2–4\]](#) CTCF has multiple functions including acting directly at gene promoters to regulate transcription, mediating long-range chromatin interactions, and it is the best characterized chromatin domain insulator-associated protein in vertebrates.

The cohesin complex plays numerous roles in mammalian gene regulation including promoting transcription factor binding at enhancers [\[5, 6\]](#) and promoting cell-type specific gene activation by facilitating DNA-promoter interactions through cell-type specific DNA-looping. [\[7, 8\]](#) CTCF may co-localize with cohesin [\[9–13\]](#) which then targets both proteins to specific sites in the genome. Interactions between the cohesin complex and CTCF mediate cell-type specific long-range chromatin contacts and modulate the enhancer-blocker activity of CTCF. [\[14–16\]](#) The cohesin complex is composed of four proteins Smc1, Smc3, Scc1, and either SA-1 or SA-2. [\[17\]](#) SA-1 and SA-2 are closely related homologs of Scc3, whose presence in cohesin complexes is mutually exclusive, leading to two highly related, but distinct complexes, cohesin^{SA-1} and cohesin^{SA-2} [\[18, 19\]](#) The SA-1 component of the cohesin complex has been shown to directly interact with CTCF, mediating many of the above functions. [\[9\]](#)

The goal of these studies was to gain insight into the roles of CTCF, cohesin^{SA-1}, and their association with gene expression and chromatin domain organization in erythroid development. Chromatin immunoprecipitation coupled with high throughput sequencing and mRNA transcriptome analyses via RNA-seq were performed in primary human hematopoietic stem and progenitor cells (HSPC) and primary human erythroid cells from single donors. Changes in sites of CTCF and cohesin^{SA-1} occupancy and their association with gene expression were observed. Cell type-specific CTCF sites in erythroid cells were linked to highly expressed genes. Repressive chromatin domains increased in both number and size during hematopoiesis, with many more repressive domains in erythroid cells than HSPCs. CTCF and cohesin^{SA-1} marked the boundaries of these repressive chromatin domains in a cell-type specific manner. These genomic data support the hypothesis that CTCF and cohesin^{SA-1} have multiple roles in the regulation of gene expression during erythropoiesis including transcriptional regulation at gene promoters and maintenance of chromatin architecture.

Methods

Cell selection and RNA analyses

Human CD34⁺-selected hematopoietic stem and progenitor cells (hereafter called HSPCs) isolated at >95% purity were obtained from the Yale Cooperative Center for Excellence in Molecular Hematology from unused clinical specimens. Erythroid progenitor cells were cultured and isolated as described. [\[20\]](#) Immunomagnetic bead selection was used to select a population of cells based on expression of CD71 (transferrin receptor) and CD235a (glycophorin A), representing the R3/R4 cell population of nucleated erythroid cells defined by Zhang *et al.* [\[21\]](#) at >95% purity as assessed by analytic FACS (Figure A in [S1 File](#)).

To avoid donor-to-donor variability observed in hematopoietic cells, including differences in age, gender, genetic background, etc., [22–24] studies, i.e. RNA-seq and ChIP-seq of CTCF and cohesin^{SA1}, were performed using CD34+ and erythroid cells derived from the same donor.

RNA was isolated and prepared for RNA-seq analyses as described.[20] Samples were sequenced on an Illumina HiSeq 2000 using 76bp-single end reads. FASTQ format sequencing reads were aligned to the hg19 genome, NCBI Build 37, using TopHat Version 2.0.4 software with default parameters except minimum anchor length of 12. The EdgeR program was used to identify differences in expression of RefSeq transcripts. Filtering included transcripts with >1 tag/million reads in 3 or more samples.

Chromatin immunoprecipitation and high throughput sequencing

ChIP assays were performed as previously described.[20, 25, 26] Samples were immunoprecipitated with antibody against CTCF (Creative Diagnostics, DMABT-H19813), the SA-1 subunit of cohesin (Abcam ab4457), trimethyl histone H3 lysine 27 (Abcam ab6002) or nonspecific rabbit IgG (sc-2091 Santa Cruz). DNA processing and high throughput sequencing were performed as described.[20] Because of the age, gender, and genetic background differences noted above, and the growing realization genetic variability influences epigenetic findings, [27] parallel RNA-seq and ChIP-seq of CTCF and cohesin^{SA1} data sets from individual donors were analyzed together.

Analyses of ChIP-seq results

The MACS program version 1.4.0rc2 was used to identify peaks with a p-value < 10e-5 and a fold enrichment > 6 for erythroid SA1 and > 8 for the other samples.[28] Quality control analyses of ChIP-seq data were performed using Picard MarkDuplicates (<http://broadinstitute.github.io/picard>), Phantompeakqualtools and the DiffBind package.[29, 30] The DiffBind analysis used fold-change filtered peaks with default parameters (minOverlap = 2). The best replicate for each condition was chosen for further analysis. Localization of CTCF and cohesin^{SA1} binding sites relative to known genes was done using the BEDTools software package.[31] Comparison of CTCF genome-wide binding data sets generated through the Broad Institute as part of the ENCODE consortium were acquired through the UCSC Genome Browser (<http://genome.ucsc.edu/>). Motif finding was done using the Homer software package.[32] Motifs discovered by Homer were compared against the Homer database of known motifs from TRANSFAC, JASPAR and public ChIP-seq data.[33] The Genomic Regions Enrichment Annotations Tool (GREAT) was used to analyze functional significance of *cis*-regulatory regions identified by ChIP-seq.[34] Broad regions of H3K27me3 binding were identified using SICER.[35] Regions with > 3 fold enrichment were merged with neighboring regions within 2000 bases, and the resulting regions larger than 2000 bases were used for H3K27me3 domain analysis. Co-localization *p*-values were obtained by randomization of genomic intervals within the human genome excluding gap regions for 1000 iterations.

Validation of ChIP-seq results

Primers were designed for representative binding regions for both CTCF and cohesin^{SA1} in the target genes identified by the MACS program (Table A in [S2 File](#)). Immunoprecipitated DNA was analyzed by quantitative real-time PCR as described.[25] All quantitative ChIP validation experiments were performed at least in triplicate.

Data access

The raw data files generated by RNA-seq and ChIP-seq analyses have been submitted to Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/> Reference series number GSE67893).

Results

CTCF and cohesin^{SA-1} ChIP-seq and mRNA expression analyses in human hematopoietic stem and progenitor (HSPC) and primary erythroid cells

ChIP-seq was performed utilizing antibodies specific for CTCF and the SA-1 component of the cohesin complex (cohesin^{SA-1}) to generate genome-wide maps of CTCF and cohesin^{SA-1} binding in primary human HSPC and erythroid cell chromatin. Quality control analyses of ChIP-seq data for read duplication, strand cross correlation, and principal components clustering demonstrated the data were of high quality (Table B in [S2 File](#) and Figure B-D in [S1 File](#)). Validation of CTCF and cohesin^{SA-1} enrichment at selected peaks was performed by quantitative ChIP PCR (Figure C in [S1 File](#) and Tables C and D in [S2 File](#)). In the replicate chosen for analyses, the MACS program identified 50,798 sites of CTCF and 42,072 sites of cohesin^{SA-1} occupancy in HSPC cell chromatin and 49,417 sites of CTCF and 40,511 sites of cohesin^{SA-1} occupancy in erythroid cell chromatin ($p < 10e-5$) (Table E in [S2 File](#)).

Transcriptome analyses were performed using mRNA isolated from human HSPC and erythroid cells using RNA-seq. In HSPC cells, 13,106 transcripts were detected (median count per million reads > 1), while in erythroid cells 12,790 transcripts were detected. Five thousand two hundred thirty two transcripts were differentially expressed by more than 2 fold between HSPC and erythroid cells, with 2289 genes up regulated in erythroid cells and 2943 down regulated in erythroid cells.

Sites of CTCF and cohesin^{SA-1} co-occupancy are enriched in gene promoters

Overlap of sites of CTCF and cohesin^{SA-1} occupancy were analyzed in HSPC and erythroid cells (Figure E in [S1 File](#)). In erythroid cells, more CTCF sites were co-occupied with cohesin^{SA-1} than CTCF sites lacking cohesin^{SA-1} (co-occupied: 26,658 vs. CTCF alone 22,869). In contrast, in HSPCs, the majority of CTCF sites lacked cohesin^{SA-1} co-occupancy (co-occupied: 18,179 vs CTCF alone: 29,000).

In both HSPC and erythroid cell chromatin, CTCF and cohesin^{SA-1} binding sites were enriched in 5' flanking regions and promoter regions, and, intergenic regions were underrepresented relative to genome composition ([Fig 1](#)). In both cell types, sites of CTCF and cohesin^{SA-1} co-occupancy were increased at gene promoters compared to singly occupied sites at gene promoters, 20% in HSPC cells and 31% in erythroid cells.

The Homer algorithm was utilized to identify over represented DNA motifs at sites of CTCF and cohesin^{SA-1} binding. In HSPC cell chromatin, the most common motif identified at co-occupied peaks and CTCF peaks without cohesin^{SA-1} was nearly identical to the CTCF consensus motif identified by Kim *et al.* in primary human fibroblasts (Figure F in [S1 File](#)). [1] The most common motif identified at cohesin^{SA-1} binding sites in HSPC cell chromatin was a BRCA1-binding motif. In erythroid cell chromatin, the most common motif identified at co-occupied peaks and CTCF peaks without cohesin^{SA-1} was CTCF, while the most common motif identified at cohesin^{SA-1} binding sites without CTCF was Sp1. Other over represented motifs are shown in Figure G in [S1 File](#).

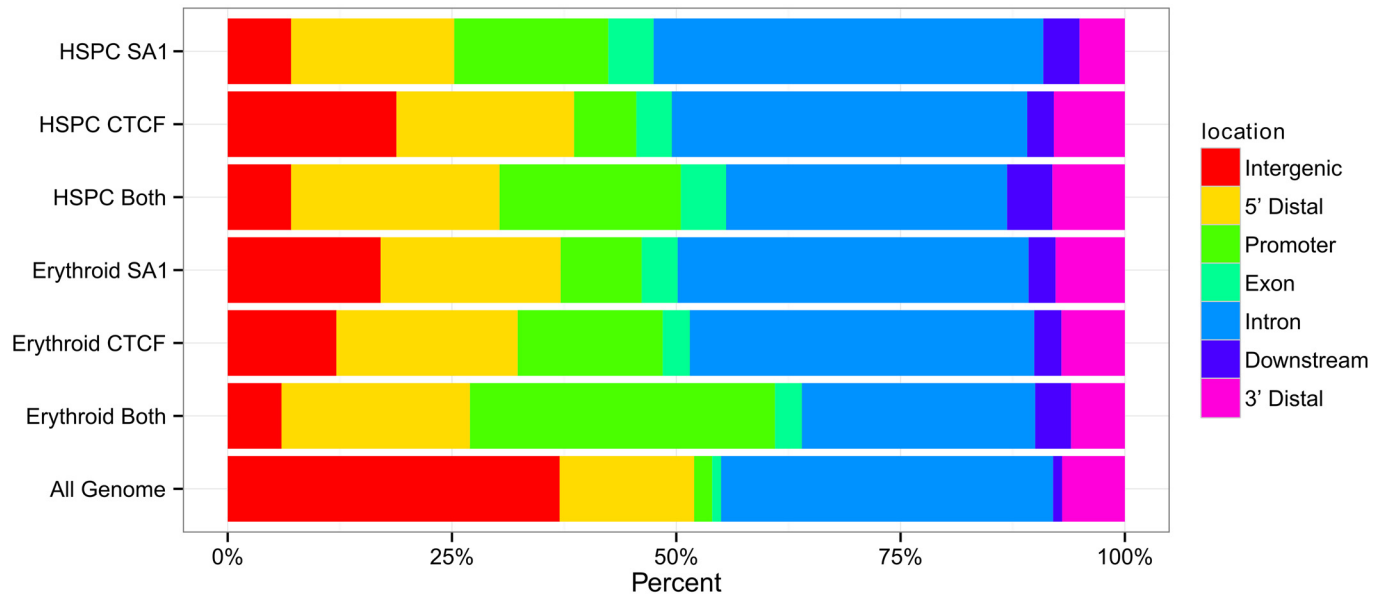


Fig 1. Distribution of CTCF and cohesin^{SA-1} occupancy and co-occupancy in human primary hematopoietic stem and progenitor cell (HSPC) and primary erythroid cell chromatin. ChIP-seq was performed with antibodies against CTCF and cohesin^{SA-1} in human primary hematopoietic stem and progenitor cell (HSPC) and primary erythroid cell chromatin. Sites of protein occupancy were determined by MACS. The human genome was portioned into seven bins relative to known genomic features associated with RefSeq genes. The percentage of the human genome represented by each bin was color coded, and the distribution of peaks of CTCF and cohesin^{SA-1} in each bin graphed on the color coded bar. Abbreviations: TSS: transcriptional start site; TES: transcriptional end site. Intergenic: >50Kb from a gene. 5' distal 1-50Kb upstream of TSS. Promoter: within 1Kb of TSS. Downstream: within 1 Kb of TES. 3' Distal 1-50Kb downstream of TES.

doi:10.1371/journal.pone.0155378.g001

A subset of CTCF sites are cell-type specific in HSPC and erythroid cell chromatin

CTCF has been reported to have sites of both cell-type specific and cell-type invariant binding, with ~40–60% of sites demonstrating cell-type specificity. Patterns of CTCF occupancy in HSPC and erythroid cell chromatin were compared to each other and to CTCF occupancy in several human ENCODE ChIP-seq data sets, including monocyte (CD14+), lymphoblastoid (G17828), embryonic stem cell (H1ES), human cardiac myocytes (HCM), human mammary fibroblasts (HMF), human umbilical vein endothelial (HUVEC), and normal human epidermal keratinocytes (NHEK) (Table 1). Cell type-specific CTCF sites were more common in HSPC cell chromatin, with 51% (25,912) of CTCF sites specific to HSPC cells, *i.e.* not present in any of the 7 ENCODE data sets. Twenty six percent (13,307) of the CTCF sites in HSPC cells were invariant, *i.e.* present in all 7 data sets compared to 39% (19,396) in erythroid cells. Typical cell-type specific and invariant CTCF binding sites are shown at several gene loci in erythroid cells (Fig 2).

Cell-type specific CTCF sites are near highly expressed genes in erythroid cells but not HSPCs

Levels of mRNA expression were assessed in genes within 1kb of cell type-specific or invariant CTCF sites. Genes linked to erythroid-specific CTCF sites were expressed at significantly higher levels than those with invariant CTCF sites (p -value < 2.2e-16). In contrast, in HSPCs, genes linked to cell type-specific CTCF sites were expressed at significantly lower levels than genes linked to invariant CTCF sites (p -value < 2.2e-16)(Fig 3). A series of network and pathway analyses were performed on genes with cell-type specific CTCF binding sites.[36]

Table 1. Comparison of CTCF binding sites in hematopoietic stem and progenitor cell (HSPC) and erythroid cell chromatin with CTCF binding in multiple cell types.

Cell Type	Data Source*	Number of CTCF Binding Sites Identified	Number of Common CTCF Sites in HSPC Cells	Number of Common CTCF Sites in Erythroid Cells
HSPC	This report	50,798	23,105 (47%)	
Erythroid	This report	49,417		23,183 (45%)
CD14	Broad	40,187	16,497 (32%)	24,867 (50%)
GM12878	Broad	37,570	19,064 (38%)	29,316 (59%)
H1ES	Broad	54,489	19,747 (39%)	29,898 (61%)
HCM	UWashington	44,259	19,799 (39%)	31,253 (63%)
HMF	UWashington	51,723	19,392 (38%)	31,512 (64%)
HUVEC	UWashington	66,858	22,240 (44%)	36,453 (74%)
NHEK	Broad	53,766	19,292 (38%)	31,723 (64%)
K562	UWashington	54,076	22,021 (43%)	34,091 (69%)

* ENCODE ChIP-seq data from multiple cell types obtained from the UCSC database were compared to binding sites in primary human HSPC progenitor cells and primary nucleated human erythroid cells. Percents are the fraction of common sites divided by the total number of HSPC or erythroid CTCF sites.

doi:10.1371/journal.pone.0155378.t001

Interestingly, genes within 1kb of erythroid cell-specific CTCF sites were highly significantly enriched for Gene Ontogeny Biological Process terms associated with hematopoiesis including “regulation of erythrocyte differentiation” and were enriched for Mouse Phenotype terms including “microcytic anemia” and “decreased mean corpuscular volume.”

There is poor correlation of CTCF occupancy between primary erythroid cells and K562 cells

These studies were performed in primary human hematopoietic cells rather than in cells from transformed lines. K562 erythroleukemia cells have been utilized as a model of erythroid cell genetics and epigenetics by ENCODE. When comparing CTCF occupancy in human primary erythroid cells to K562 cells, only 69% of sites were shared (Table 1).

Repressive chromatin domains increase in number during erythropoiesis

Cellular differentiation has been associated with reorganization and expansion of repressive chromatin domains in mammalian genome with silencing of the genes in the domain.[37–39] To examine repressive chromatin domains and their boundaries during hematopoiesis, ChIP-seq with an antibody against H3K27me3 as a marker of repressive chromatin was performed with HSPC and erythroid cells. Chromatin domains were identified using the Sicer program.[35] More H3K27me3 chromatin domains were identified in erythroid vs. HSPC cell chromatin (17,165 vs. 11,649, Table 2). In addition, average domain lengths were longer in erythroid compared to HSPC chromatin (12.2 vs. 8.3kb, Table 2), with the erythroid domains encompassing 6.7% of the genome compared to 3.1% in HSPC cells.

Of the 17,165 H3K27me3 domains identified in erythroid cell chromatin, 59% (10,146) were specific to erythroid cells (*i.e.* not in CD34 cells). Thus a large number of tissue-specific repressive chromatin domains are found in differentiated erythroid cells. There was a strong anti-correlation of H3K27me3 domains with gene expression. This difference was much greater in erythroid cells than HSPCs (Figure H in S1 File).

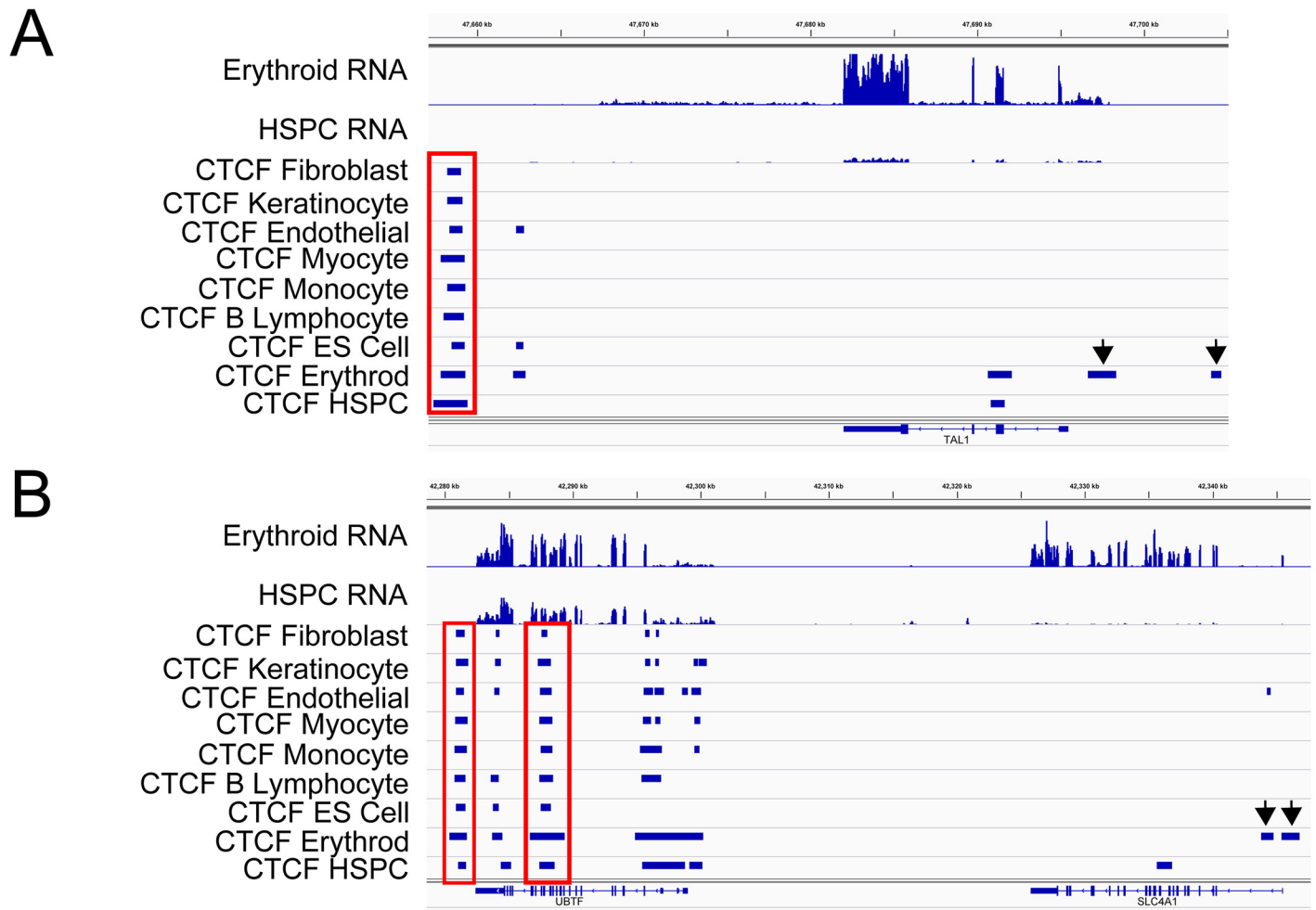


Fig 2. Invariant and cell type-specific CTCF sites. Patterns of CTCF occupancy in HSPC and erythroid cell chromatin were compared to CTCF occupancy in several human ENCODE ChIP-seq data sets, including fibroblast, keratinocyte, endothelial, myocyte, monocyte, lymphocyte, embryonic stem (ES) cell, erythroid and HSPC cells. **A.** At the *TAL1* locus, a 3' site of invariant CTCF binding marked by the rectangle is present in all cell types. Two sites of erythroid-specific CTCF binding, denoted by the arrows, are present 5' of the gene. Corresponding RNA-seq tracks in HSPC and erythroid cells are shown at the top. Genomic coordinates: Chr1:47,600–47,700. **B.** At the *UBTF* and *SLC4A1* loci, there are 2 sets of invariant CTCF binding, marked by rectangles, 3' of the *UBTF* locus, present in all cell types. One site of erythroid-specific CTCF binding, denoted by the arrows, are present 5' of the *SLC4A1* locus. Corresponding RNA-seq tracks in HSPC and erythroid cells are shown at the top. Genomic coordinates: Chr17:42,280–42,340.

doi:10.1371/journal.pone.0155378.g002

CTCF and cohesin^{SA-1} mark the boundaries of chromatin domains in a cell-type specific manner

In some cell types, CTCF has been observed to mark the boundaries of repressive chromatin domains in a cell-type specific manner.[40] To determine whether CTCF and cohesin^{SA-1} are present at domain boundaries in HSPC and erythroid cell chromatin, CTCF and cohesin^{SA-1} binding sites were mapped onto chromatin domains defined by H3K27me3 modification. Binding sites within 1 kb of a domain boundary were considered to mark the boundary of the domain.

There were 4,832 and 3,888 CTCF sites that marked domain boundaries in HSPC and erythroid cells, respectively (Table 2 and Fig 4). These CTCF sites were cell-type specific, as only 711 sites were shared between HSPCs and erythroid cells. Cohesin^{SA-1} was also found at

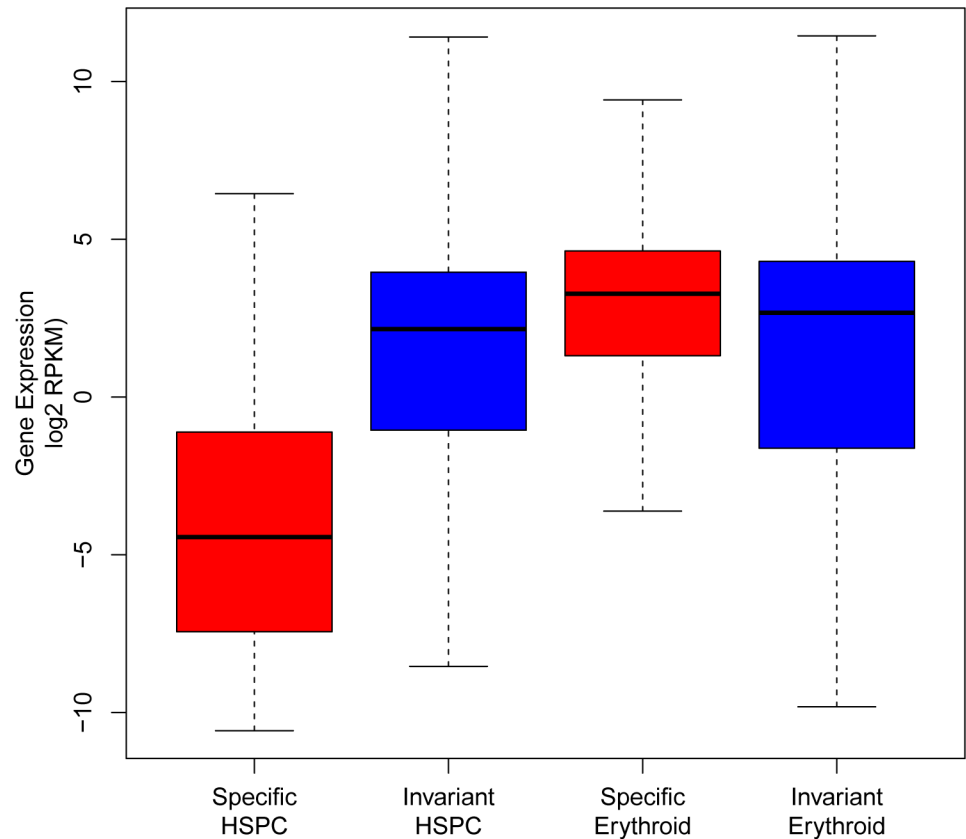


Fig 3. Gene expression and CTCF occupancy. Gene expression levels in primary human HSPC and erythroid cell mRNA were correlated with sites of CTCF occupancy by class within 1kb of the transcription start site.

doi:10.1371/journal.pone.0155378.g003

domain boundaries, present at 5093 boundaries in HSPC cells and 3854 boundaries in erythroid cells.

CTCF frequently co-localized with cohesin^{SA-1} at domains, with 54% of CTCF sites at boundaries (p -value <0.001) and 56% of CTCF sites at boundaries (p -value <0.001) binding both proteins in HSPC and erythroid cell chromatin, respectively (Table 2). An example of CTCF and cohesin^{SA-1} at an erythroid-specific boundary is shown at the ankyrin-1 (*ANK1*) locus in Fig 5. Multiple tissue-specific “exon 1s” are found at the 5’ end of the *ANK1* gene which all join in frame to exon 2, creating cDNA transcripts with unique 5’ ends. In erythroid cells, the sequence surrounding and including a neural-specific *ANK1* exon 1, located 5’ of the erythroid exon 1, is in a region of repressive chromatin, heavily modified by H3K27me3 (Fig 5,

Table 2. Repressive chromatin domains and CTCF and cohesin^{SA-1} occupancy at domain boundaries.

	HSPC cells	Erythroid Cells
Number of Domains	11649	17165
Average Domain Length	8.3 kb	12.2 kb
CTCF at Domain Boundaries	4832	3888
Cohesin ^{SA-1} at Domain Boundaries	5093	3854
CTCF- Cohesin ^{SA-1} co-localization at Domain Boundaries	2602	2180

doi:10.1371/journal.pone.0155378.t002

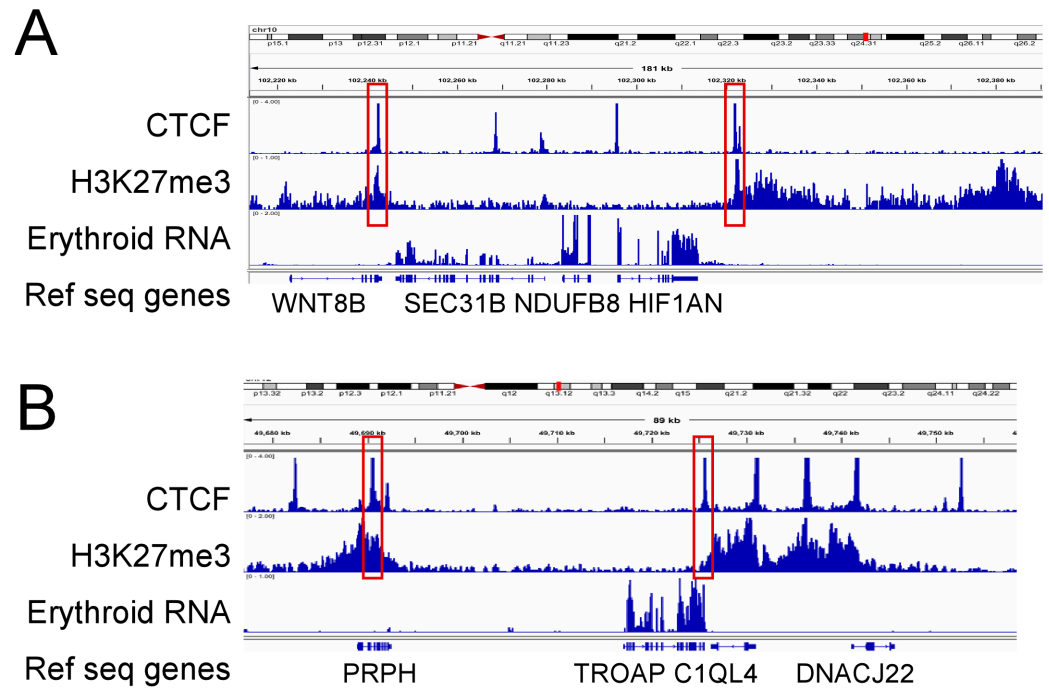


Fig 4. Repressive chromatin domains and CTCF occupancy. Representative integrated genome viewer (IGV) views of CTCF occupancy, repressive chromatin domains marked by H3K27me3 enrichment, and gene expression determined by RNA-seq in erythroid cells. **A.** Repressive chromatin domains marked by CTCF occupancy at their boundaries flank the *SEC31B*, *NDUFB8*, and *HIF1AN* genes. These 3 genes are expressed in erythroid cells, while the *WNT8B* gene, located in a repressive chromatin domain, is not. Genomic coordinates: Chr10:102,220–102,380. **B.** Repressive chromatin domains marked by CTCF occupancy at their boundaries flank the *TROAP* and *C1QL4* genes. These 2 genes are expressed in erythroid cells, while the flanking *PRPH* and *DNACJ22* genes, located in flanking repressive chromatin domains, are not. Genomic coordinates: Chr12:49,680–49,760.

doi:10.1371/journal.pone.0155378.g004

top). At the boundary of this repressive chromatin domain are a pair of CTCF/cohesin^{SA-1} sites, present in erythroid but not HSPC chromatin, followed by the transcribed exons of the *ANK1* gene. *ANK1* is not expressed in HSPCs and this entire region is modified by H3K27 trimethylation (Fig 5, bottom). This region has been shown to functionally act as a barrier insulator *in vitro* and *in vivo*. [41] Together, these data indicate CTCF and cohesin^{SA-1} mark the boundaries of some repressive chromatin domains in a cell-type specific manner.

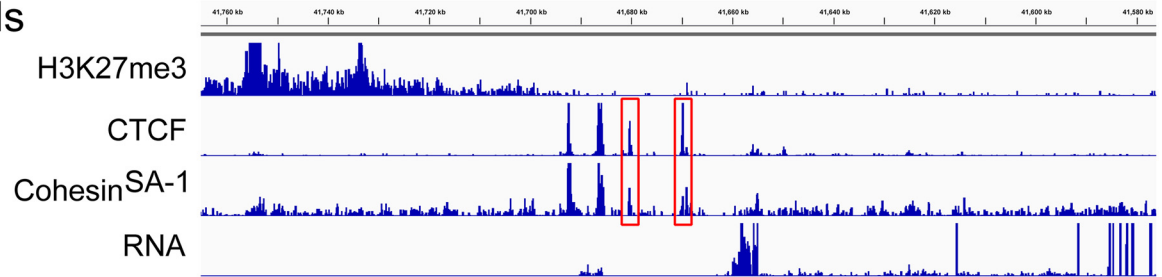
Discussion

CTCF and cohesin^{SA-1} are distributed widely throughout the genomes of human HSPC hematopoietic stem and progenitor cells and differentiating erythroid cells. The finding of large numbers of co-occupied sites present at gene promoters in erythroid cells, not a common finding in all cell types studied to date, [42] is consistent with the recent observation that the cohesin complex is present at enhancers and active gene promoters. [8]

Although there were many shared sites of CTCF and cohesin^{SA-1} co-occupancy in both cell types, the majority of CTCF and cohesin^{SA-1} sites lacked the other protein. Similar to other highly differentiated cell types, cell-type specific CTCF sites were far more common in erythroid cell chromatin than HSPCs.

Detailed genome wide epigenetic studies have revealed a complex, higher order of chromosomal organization, with numerous, extensive chromatin domains. Repressive heterochromatin

Erythroid Cells



HSPCs

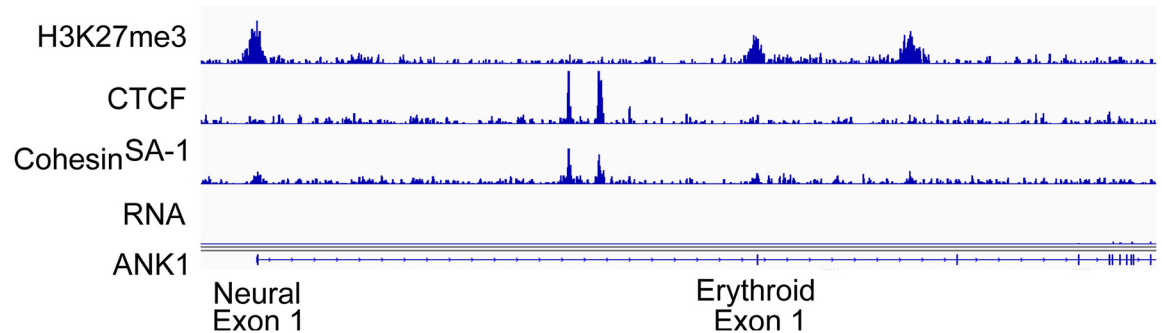


Fig 5. Repressive chromatin domains and CTCF-cohesin^{SA-1} co-occupancy. Representative integrated genome viewer (IGV) views of CTCF and cohesin^{SA-1} occupancy, repressive chromatin domains marked by H3K27me3 enrichment, and gene expression determined by RNA-seq in HSPC and erythroid cells. Multiple tissue-specific “exon 1s” are found at the 5’ end of the *ANK1* gene which all join in frame to exon 2, creating cDNA transcripts with unique 5’ ends. In erythroid cells (top), the sequence surrounding and including a neural-specific *ANK1* exon 1, located 5’ of the erythroid exon 1, is in a region of repressive chromatin, heavily modified by H3K27me3. At the boundary of this repressive chromatin domain are a pair of CTCF/cohesin^{SA-1} sites, present in erythroid but not HSPC chromatin, followed by the transcribed exons of the *ANK1* gene. *ANK1* is not expressed in HSPCs and this entire region is modified by H3K27 trimethylation (bottom). Genomic coordinates: Chr8:41,760–41,580.

doi:10.1371/journal.pone.0155378.g005

domains, defined by posttranslational histone modifications such as dimethylation of histone H3 lysine 9, trimethylation of histone H3 lysine 9, and trimethylation of histone H3 lysine 27, may extend over megabases in human cells.[37, 38] Studies comparing human embryonic stem cells to differentiated cell types have suggested that repressive chromatin domains increase in number and size with cellular differentiation, [37, 39] with silencing of the genes contained in the heterochromatin. In our studies, the number of H3K27me3 repressive chromatin domains doubled during erythroid development indicating that acquisition of repressive chromatin domains during erythropoiesis parallels embryonic stem cell development.

Our data indicate that many repressive chromatin domains in HSPC and erythroid cells have cell-type specific CTCF and cohesin^{SA-1} occupancy at their boundaries, suggesting that these proteins play a role in either domain establishment or maintenance. A subset of CTCF sites has been mapped to domain boundaries in T lymphocytes, HeLa cells, and Jurkat cells, leading to speculation that CTCF plays an important role in chromatin insulator function.[40] CTCF is not required for the barrier activity of the chicken HS4 insulator.[43] However, other reports have implicated a role for CTCF in barrier function, [40, 44] although this has not been supported by direct evidence.[45] Finally, it has been suggested that cohesin proteins may act as transcriptional insulators, [11] but again, studies providing direct evidence to support this hypothesis are lacking. Unraveling the numerous role(s) of CTCF and cohesin^{SA-1} at domain boundaries will provide considerable insight into our understanding of higher order chromatin structure and function. These genome wide datasets in human primary hematopoietic cells are excellent resources for future studies.

Much of the currently available data on chromatin architecture and transcription factor occupancy have been generated by ENCODE, which primarily utilized transformed cell lines for their studies. These studies were performed in primary human hematopoietic cells rather than in cells from transformed lines. K562 erythroleukemia cells, derived from a patient with chronic myelogenous leukemia in blast crisis and often used as surrogates for studies of erythroid gene function and regulation, have been utilized as a model of erythroid cell genetics and epigenetics by ENCODE. When comparing CTCF occupancy in human primary erythroid cells to K562 cells, only 69% of sites were shared. The lack of more extensive overlap may reflect developmental differences, as K562 cells are at significantly earlier stage of differentiation than R3/R4 erythroid cells, differences between primary cells and an immortalized cell line due to acquired aneuploidy, and/or other related changes acquired over time.[46]

Alterations in higher-order genome organization leading to perturbation in gene expression are being recognized as important mechanisms of inherited and acquired disease.[47] Because of their critical roles in organizing and maintaining higher order chromatin structure and regulating appropriate patterns of gene expression, perturbation of the structure or function of CTCF or cohesin^{SA-1} have been associated with disease phenotypes. Disruption or deletion of CTCF-associated insulators have been described in human disease such as loss of function of the DM1 insulator in myotonic dystrophy, and chromosomal deletions or translocations of regions containing CTCF binding sites in Beckwith-Wiedemann syndrome, Wilms' tumor, and other various cancers.[44, 48] Perturbation of associated *cis*-sequences regulating their binding are another predicted mechanism of disease, [49, 50] as shown in a subset of cases of hereditary spherocytosis.[51]

Defects of the cohesin complex, collectively termed the “cohesinopathies” have been associated with several disorders with prominent developmental defects.[52] Roberts syndrome/SC-phocomelia and Cornelia de Lange syndrome patients suffer from mutations in cohesin complex-associated pathway proteins. Detailed analyses of these disorders indicate that distinct from its role in chromosome segregation, abnormalities of the cohesin network that alter gene expression and genome organization may underlie cohesinopathies.[53] Synthesis of data from detailed patient genetic studies and from functional genomics studies, such as these hematopoietic cell data sets, which identify regions of DNA with regulatory potential throughout the genome, will provide critical insight into our understanding of the complex mechanisms of genetic variation in inherited and acquired disease.

Conclusions

Sites of CTCF and cohesin^{SA-1} occupancy, associated chromatin architecture, and related gene expression changed during erythropoiesis. Repressive chromatin domains increased in both number and size during hematopoiesis, with many more repressive domains in erythroid cells than HSPCs, with CTCF and cohesin^{SA-1} marking the boundaries of these repressive chromatin domains in a cell-type specific manner. These genomic data support the hypothesis that CTCF and cohesin^{SA-1} have multiple roles in the regulation of gene expression during erythropoiesis. Obtained from primary human erythroid cells, these datasets provide an important resource for studies of normal and perturbed erythropoiesis.

Supporting Information

S1 File. Supporting Figures. (Figure A) Analytic flow activated cell sorting analyses of cultured human primary erythroid cells. (Figure B) Correlation heat map using affinity (read count) data. (Figure C) Principal components analysis of affinity (read count) data. (Figure D) Heat map of affinity (read count) data for individual sites in individual ChIP-seq samples.

(Figure E) Sites of CTCF and cohesin^{SA-1} occupancy in HSPCs and erythroid cells. (Figure F) Motif analysis using the Homer algorithm in HSPC. (Figure G). Motif analysis using the Homer algorithm in erythroid cells. (Figure H) Correlation of repressive domains and gene expression in HSPCs and erythroid cells. (PDF)

S2 File. Supporting Tables. (Table A) PCR primers for CTCF and cohesin^{SA-1} validation. (Table B) Read Count, Duplication and Strand Cross Correlation Analyses. (Table C) Quantitative ChIP Validation of CTCF binding Sites. (Table D) Quantitative ChIP Validation of cohesin^{SA-1} binding sites. (Table E) Summary of ChIP seq results. (DOCX)

Acknowledgments

This work was supported in part by National Institutes of Health Grants K12HD000850, HL65448, and HL106184.

Author Contributions

Conceived and designed the experiments: LAS KLG YM PGG. Performed the experiments: LAS KLG YM. Analyzed the data: LAS VPS PGG.

References

1. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*. 2007; 128(6):1231–45. Epub 2007/03/27. doi: [10.1016/j.cell.2006.12.048](https://doi.org/10.1016/j.cell.2006.12.048) PMID: [17382889](https://pubmed.ncbi.nlm.nih.gov/17382889/); PubMed Central PMCID: [PMC2572726](https://pubmed.ncbi.nlm.nih.gov/PMC2572726/).
2. Mukhopadhyay R, Yu W, Whitehead J, Xu J, Lezcano M, Pack S, et al. The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res*. 2004; 14(8):1594–602. doi: [10.1101/gr.2408304](https://doi.org/10.1101/gr.2408304) PMID: [15256511](https://pubmed.ncbi.nlm.nih.gov/15256511/); PubMed Central PMCID: [PMC509268](https://pubmed.ncbi.nlm.nih.gov/PMC509268/).
3. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*. 2000; 405(6785):482–5. doi: [10.1038/35013100](https://doi.org/10.1038/35013100) PMID: [10839546](https://pubmed.ncbi.nlm.nih.gov/10839546/).
4. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*. 2000; 405(6785):486–9. doi: [10.1038/35013106](https://doi.org/10.1038/35013106) PMID: [10839547](https://pubmed.ncbi.nlm.nih.gov/10839547/).
5. Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, et al. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res*. 2013; 23(12):2066–77. doi: [10.1101/gr.161620.113](https://doi.org/10.1101/gr.161620.113) PMID: [24002784](https://pubmed.ncbi.nlm.nih.gov/24002784/); PubMed Central PMCID: [PMC3847776](https://pubmed.ncbi.nlm.nih.gov/PMC3847776/).
6. Faure AJ, Schmidt D, Watt S, Schwalie PC, Wilson MD, Xu H, et al. Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Res*. 2012; 22(11):2163–75. doi: [10.1101/gr.136507.111](https://doi.org/10.1101/gr.136507.111) PMID: [22780989](https://pubmed.ncbi.nlm.nih.gov/22780989/); PubMed Central PMCID: [PMC3483546](https://pubmed.ncbi.nlm.nih.gov/PMC3483546/).
7. Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K. Effects of sister chromatid cohesion proteins on cut gene expression during wing development in *Drosophila*. *Development*. 2005; 132(21):4743–53. Epub 2005/10/07. doi: [10.1242/dev.02064](https://doi.org/10.1242/dev.02064) PMID: [16207752](https://pubmed.ncbi.nlm.nih.gov/16207752/); PubMed Central PMCID: [PMC1635493](https://pubmed.ncbi.nlm.nih.gov/PMC1635493/).
8. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 467(7314):430–5. PMID: [20720539](https://pubmed.ncbi.nlm.nih.gov/20720539/). doi: [10.1038/nature09380](https://doi.org/10.1038/nature09380)
9. Rubio ED, Reiss DJ, Welcsh PL, Distechi CM, Filippova GN, Baliga NS, et al. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A*. 2008; 105(24):8309–14. doi: [10.1073/pnas.0801273105](https://doi.org/10.1073/pnas.0801273105) PMID: [18550811](https://pubmed.ncbi.nlm.nih.gov/18550811/); PubMed Central PMCID: [PMC2448833](https://pubmed.ncbi.nlm.nih.gov/PMC2448833/).
10. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, et al. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res*. 20(5):578–88. PMID: [20219941](https://pubmed.ncbi.nlm.nih.gov/20219941/). doi: [10.1101/gr.100479.109](https://doi.org/10.1101/gr.100479.109)

11. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008; 451(7180):796–801. doi: [10.1038/nature06634](https://doi.org/10.1038/nature06634) PMID: [18235444](https://pubmed.ncbi.nlm.nih.gov/18235444/).
12. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*. 2008; 132(3):422–33. Epub 2008/02/02. doi: [10.1016/j.cell.2008.01.011](https://doi.org/10.1016/j.cell.2008.01.011) PMID: [18237772](https://pubmed.ncbi.nlm.nih.gov/18237772/).
13. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J*. 2008; 27(4):654–66. Epub 2008/01/26. doi: [10.1038/emboj.2008.1](https://doi.org/10.1038/emboj.2008.1) PMID: [18219272](https://pubmed.ncbi.nlm.nih.gov/18219272/); PubMed Central PMCID: PMC2262040.
14. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, et al. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature*. 2009; 460(7253):410–3. Epub 2009/05/22. doi: [10.1038/nature08079](https://doi.org/10.1038/nature08079) PMID: [19458616](https://pubmed.ncbi.nlm.nih.gov/19458616/); PubMed Central PMCID: PMC2869028.
15. Hou C, Dale R, Dean A. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci U S A*. 2010; 107(8):3651–6. Epub 2010/02/06. doi: [10.1073/pnas.0912087107](https://doi.org/10.1073/pnas.0912087107) PMID: [20133600](https://pubmed.ncbi.nlm.nih.gov/20133600/); PubMed Central PMCID: PMC2840441.
16. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet*. 2014; 15(4):234–46. doi: [10.1038/nrg3663](https://doi.org/10.1038/nrg3663) PMID: [24614316](https://pubmed.ncbi.nlm.nih.gov/24614316/); PubMed Central PMCID: PMC24610363.
17. Peters JM, Tedeschi A, Schmitz J. The cohesin complex and its roles in chromosome biology. *Genes Dev*. 2008; 22(22):3089–114. Epub 2008/12/06. doi: [10.1101/gad.1724308](https://doi.org/10.1101/gad.1724308)
18. Losada A, Yokochi T, Kobayashi R, Hirano T. Identification and characterization of SA/Scs3p subunits in the *Xenopus* and human cohesin complexes. *J Cell Biol*. 2000; 150(3):405–16. Epub 2000/08/10. PMID: [10931856](https://pubmed.ncbi.nlm.nih.gov/10931856/); PubMed Central PMCID: PMC2175199.
19. Sumara I, Vorlaufer E, Gieffers C, Peters BH, Peters JM. Characterization of vertebrate cohesin complexes and their regulation in prophase. *J Cell Biol*. 2000; 151(4):749–62. Epub 2000/11/15. PMID: [11076961](https://pubmed.ncbi.nlm.nih.gov/11076961/); PubMed Central PMCID: PMC2169443.
20. Su MY, Steiner LA, Bogardus H, Mishra T, Schulz VP, Hardison RC, et al. Identification of biologically relevant enhancers in human erythroid cells. *The Journal of biological chemistry*. 2013; 288(12):8433–44. doi: [10.1074/jbc.M112.413260](https://doi.org/10.1074/jbc.M112.413260) PMID: [23341446](https://pubmed.ncbi.nlm.nih.gov/23341446/); PubMed Central PMCID: PMC3605659.
21. Zhang J, Socolovsky M, Gross AW, Lodish HF. Role of Ras signaling in erythroid differentiation of mouse fetal liver cells: functional analysis by a flow cytometry-based novel culture system. *Blood*. 2003; 102(12):3938–46. doi: [10.1182/blood-2003-05-1479](https://doi.org/10.1182/blood-2003-05-1479) PMID: [12907435](https://pubmed.ncbi.nlm.nih.gov/12907435/).
22. Koller MR, Manchel I, Brott DA, Palsson B. Donor-to-donor variability in the expansion potential of human bone marrow cells is reduced by accessory cells but not by soluble growth factors. *Exp Hematol*. 1996; 24(13):1484–93. PMID: [8950231](https://pubmed.ncbi.nlm.nih.gov/8950231/).
23. Yao YG, Kajigaya S, Feng X, Samsel L, McCoy JP Jr., Torelli G, et al. Accumulation of mtDNA variations in human single CD34+ cells from maternally related individuals: effects of aging and family genetic background. *Stem cell research*. 2013; 10(3):361–70. doi: [10.1016/j.scr.2013.01.006](https://doi.org/10.1016/j.scr.2013.01.006) PMID: [23455392](https://pubmed.ncbi.nlm.nih.gov/23455392/); PubMed Central PMCID: PMC2154056.
24. Sudo K, Yasuda J, Nakamura Y. Gene expression profiles of cryopreserved CD34(+) human umbilical cord blood cells are related to their bone marrow reconstitution abilities in mouse xenografts. *Biochem Biophys Res Commun*. 2010; 397(4):697–705. doi: [10.1016/j.bbrc.2010.06.010](https://doi.org/10.1016/j.bbrc.2010.06.010) PMID: [20570655](https://pubmed.ncbi.nlm.nih.gov/20570655/).
25. Steiner LA, Maksimova Y, Schulz V, Wong C, Raha D, Mahajan MC, et al. Chromatin Architecture and Transcription Factor Binding Regulate Expression of Erythrocyte Membrane Protein Genes. *Mol Cell Biol*. 2009. PMID: [19687298](https://pubmed.ncbi.nlm.nih.gov/19687298/).
26. Steiner LA, Schulz VP, Maksimova Y, Wong C, Gallagher PG. Patterns of Histone H3 Lysine 27 Monomethylation and Erythroid Cell Type-specific Gene Expression. *The Journal of biological chemistry*. 2011; 286(45):39457–65. Epub 2011/09/23. doi: [10.1074/jbc.M111.243006](https://doi.org/10.1074/jbc.M111.243006) PMID: [21937433](https://pubmed.ncbi.nlm.nih.gov/21937433/).
27. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. 2013; 503(7475):290–4. doi: [10.1038/nature12644](https://doi.org/10.1038/nature12644) PMID: [24141950](https://pubmed.ncbi.nlm.nih.gov/24141950/); PubMed Central PMCID: PMC3838900.
28. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology*. 2008; 9(9):R137. Epub 2008/09/19. doi: [10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137) PMID: [18798982](https://pubmed.ncbi.nlm.nih.gov/18798982/); PubMed Central PMCID: PMC2592715.
29. Marinov GK, Kundaje A, Park PJ, Wold BJ. Large-scale quality analysis of published ChIP-seq data. *G3*. 2014; 4(2):209–23. doi: [10.1534/g3.113.008680](https://doi.org/10.1534/g3.113.008680) PMID: [24347632](https://pubmed.ncbi.nlm.nih.gov/24347632/); PubMed Central PMCID: PMC3931556.

30. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*. 2012; 481(7381):389–93. doi: [10.1038/nature10730](https://doi.org/10.1038/nature10730) PMID: [22217937](https://pubmed.ncbi.nlm.nih.gov/22217937/); PubMed Central PMCID: PMC3272464.
31. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010; 26(6):841–2. Epub 2010/01/30. doi: [10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033) PMID: [20110278](https://pubmed.ncbi.nlm.nih.gov/20110278/); PubMed Central PMCID: PMC2832824.
32. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings / International Conference on Intelligent Systems for Molecular Biology; ISMB International Conference on Intelligent Systems for Molecular Biology*. 1994; 2:28–36. Epub 1994/01/01. PMID: [7584402](https://pubmed.ncbi.nlm.nih.gov/7584402/).
33. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol*. 2007; 8(2):R24. Epub 2007/02/28. doi: [10.1186/gb-2007-8-2-r24](https://doi.org/10.1186/gb-2007-8-2-r24) PMID: [17324271](https://pubmed.ncbi.nlm.nih.gov/17324271/); PubMed Central PMCID: PMC281852410.
34. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010; 28(5):495–501. Epub 2010/05/04. doi: [10.1038/nbt.1630](https://doi.org/10.1038/nbt.1630) PMID: [20436461](https://pubmed.ncbi.nlm.nih.gov/20436461/).
35. Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics*. 2009; 25(15):1952–8. doi: [10.1093/bioinformatics/btp340](https://doi.org/10.1093/bioinformatics/btp340) PMID: [19505939](https://pubmed.ncbi.nlm.nih.gov/19505939/); PubMed Central PMCID: PMC2732366.
36. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009; 4(1):44–57. Epub 2009/01/10. doi: [10.1038/nprot.2008.211](https://doi.org/10.1038/nprot.2008.211) PMID: [19131956](https://pubmed.ncbi.nlm.nih.gov/19131956/).
37. Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, Pelizzola M, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell*. 2010; 6(5):479–91. Epub 2010/05/11. doi: [10.1016/j.stem.2010.03.018](https://doi.org/10.1016/j.stem.2010.03.018) PMID: [20452322](https://pubmed.ncbi.nlm.nih.gov/20452322/); PubMed Central PMCID: PMC2867844.
38. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. 2007; 448(7153):553–60. Epub 2007/07/03. doi: [10.1038/nature06008](https://doi.org/10.1038/nature06008) PMID: [17603471](https://pubmed.ncbi.nlm.nih.gov/17603471/); PubMed Central PMCID: PMC2921165.
39. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet*. 2009; 41(2):246–50. Epub 2009/01/20. doi: [10.1038/ng.297](https://doi.org/10.1038/ng.297) PMID: [19151716](https://pubmed.ncbi.nlm.nih.gov/19151716/); PubMed Central PMCID: PMC2632725.
40. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res*. 2009; 19(1):24–32. doi: [10.1101/gr.082800.108](https://doi.org/10.1101/gr.082800.108) PMID: [19056695](https://pubmed.ncbi.nlm.nih.gov/19056695/); PubMed Central PMCID: PMC2612964.
41. Gallagher PG, Steiner LA, Liem RI, Owen AN, Cline AP, Seidel NE, et al. Mutation of a barrier insulator in the human ankyrin-1 gene is associated with hereditary spherocytosis. *The Journal of clinical investigation*. 2010; 120(12):4453–65. Epub 2010/11/26. doi: [10.1172/JCI42240](https://doi.org/10.1172/JCI42240) PMID: [21099109](https://pubmed.ncbi.nlm.nih.gov/21099109/); PubMed Central PMCID: PMC2993586.
42. Dorsett D, Merkenschlager M. Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr Opin Cell Biol*. 2013; 25(3):327–33. doi: [10.1016/j.ceb.2013.02.003](https://doi.org/10.1016/j.ceb.2013.02.003) PMID: [23465542](https://pubmed.ncbi.nlm.nih.gov/23465542/); PubMed Central PMCID: PMC3691354.
43. Recillas-Targa F, Pikaart MJ, Burgess-Beusse B, Bell AC, Litt MD, West AG, et al. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A*. 2002; 99(10):6883–8. doi: [10.1073/pnas.102179399](https://doi.org/10.1073/pnas.102179399) PMID: [12011446](https://pubmed.ncbi.nlm.nih.gov/12011446/); PubMed Central PMCID: PMC2124498.
44. Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, et al. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet*. 2001; 28(4):335–43. Epub 2001/08/02. doi: [10.1038/ng570](https://doi.org/10.1038/ng570) PMID: [11479593](https://pubmed.ncbi.nlm.nih.gov/11479593/).
45. Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet*. 2006; 7(9):703–13. Epub 2006/08/16. doi: [10.1038/nrg1925](https://doi.org/10.1038/nrg1925) PMID: [16909129](https://pubmed.ncbi.nlm.nih.gov/16909129/).
46. Lozzio CB, Lozzio BB. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*. 1975; 45(3):321–34. Epub 1975/03/01. PMID: [163658](https://pubmed.ncbi.nlm.nih.gov/163658/).
47. Misteli T. Higher-order genome organization in human disease. *Cold Spring Harbor perspectives in biology*. 2010; 2(8):a000794. Epub 2010/07/02. doi: [10.1101/cshperspect.a000794](https://doi.org/10.1101/cshperspect.a000794) PMID: [20591991](https://pubmed.ncbi.nlm.nih.gov/20591991/); PubMed Central PMCID: PMC2908770.
48. Prawitt D, Enklaar T, Gartner-Rupprecht B, Spangenberg C, Oswald M, Lausch E, et al. Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. *Proc Natl Acad Sci U S A*. 2005; 102(11):4085–90. Epub 2005/03/04. doi: [10.1073/pnas.0500037102](https://doi.org/10.1073/pnas.0500037102) PMID: [15743916](https://pubmed.ncbi.nlm.nih.gov/15743916/); PubMed Central PMCID: PMC554791.

49. Epstein DJ. Cis-regulatory mutations in human disease. *Brief Funct Genomic Proteomic*. 2009; 8(4):310–6. doi: [10.1093/bfgp/elp021](https://doi.org/10.1093/bfgp/elp021) PMID: [19641089](https://pubmed.ncbi.nlm.nih.gov/19641089/); PubMed Central PMCID: PMCPMC2742803.
50. Wray GA. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet*. 2007; 8(3):206–16. Epub 2007/02/17. doi: [10.1038/nrg2063](https://doi.org/10.1038/nrg2063) PMID: [17304246](https://pubmed.ncbi.nlm.nih.gov/17304246/).
51. Gallagher PG, Steiner LA, Liem RI, Owen AN, Cline AP, Seidel NE, et al. Hereditary spherocytosis due to mutation in a barrier insulator in the human ankyrin-1 gene. *The Journal of clinical investigation*. 2010;In press.
52. Remeseiro S, Cuadrado A, Losada A. Cohesin in development and disease. *Development*. 2013; 140(18):3715–8. doi: [10.1242/dev.090605](https://doi.org/10.1242/dev.090605) PMID: [23981654](https://pubmed.ncbi.nlm.nih.gov/23981654/).
53. Skibbens RV, Colquhoun JM, Green MJ, Molnar CA, Sin DN, Sullivan BJ, et al. Cohesinopathies of a feather flock together. *PLoS Genet*. 2013; 9(12):e1004036. doi: [10.1371/journal.pgen.1004036](https://doi.org/10.1371/journal.pgen.1004036) PMID: [24367282](https://pubmed.ncbi.nlm.nih.gov/24367282/); PubMed Central PMCID: PMCPMC3868590.