Regular Article

MYELOID NEOPLASIA

AML displays increased CTCF occupancy associated with aberrant gene expression and transcription factor binding

Huthayfa Mujahed,[1] Sophia Miliara,[2] Anne Neddermeyer,[3] Sofia Bengtzén,[1] Christer Nilsson,[1] Stefan Deneberg,[1] Lina Cordeddu,[2] Karl Ekwall,[2] Andreas Lennartsson,[2,*] and Sören Lehmann[1,3,*]

[1]Center for Hematology and Regenerative Medicine, Department of Medicine, Huddinge, and [2]Department of Biosciences and Nutrition, Karolinska Institute, Stockholm, Sweden; and [3]Department of Medical Sciences, Hematology, Uppsala University, Uppsala, Sweden

---

KEY POINTS

- **AML cells display aberrant CTCF occupancy, and patients with *TET2* mutations show a particularly strong gain of CTCF binding in promoters.**

- **Aberrant CTCF binding is enriched at CEBPA, PU.1, and RUNX1 motifs, and azacitidine exerts profound effects on CTCF occupancy.**

CCTC-binding factor (CTCF) is a key regulator of gene expression through organization of the chromatin structure. Still, it is unclear how CTCF binding is perturbed in leukemia or in cancer in general. We studied CTCF binding by chromatin immunoprecipitation sequencing in cells from patients with acute myeloid leukemia (AML) and in normal bone marrow (NBM) in the context of gene expression, DNA methylation, and azacitidine exposure. CTCF binding was increased in AML compared with NBM. Aberrant CTCF binding was enriched for motifs for key myeloid transcription factors such as CEBPA, PU.1, and RUNX1. AML with *TET2* mutations was characterized by a particularly strong gain of CTCF binding, highly enriched for gain in promoter regions, while AML in general was enriched for changes at enhancers. There was a strong anticorrelation between CTCF binding and DNA methylation. Gain of CTCF occupancy was associated with increased gene expression; however, the genomic location (promoter vs distal regions) and enrichment of motifs (for repressing vs activating cofactors) were decisive for the gene expression pattern. Knockdown of CTCF in K562 cells caused loss of CTCF binding and transcriptional repression of genes with changed CTCF binding in AML, as well as loss of RUNX1 binding at RUNX1/CTCF-binding sites. In addition, CTCF knockdown caused increased differentiation. Azacitidine exposure caused major changes in CTCF occupancy in AML patient cells, partly by restoring a CTCF-binding pattern similar to NBM. We conclude that AML displays an aberrant increase in CTCF occupancy that targets key genes for AML development and impacts gene expression. (*Blood*. 2020;136(3):339-352)

Introduction

Acute myeloid leukemia (AML) is a hematological malignancy characterized by accumulation of immature blast cells in the blood and bone marrow (BM). AML development is driven by aberrant gene expression patterns caused by genetic and epigenetic changes of the genome.[1] The chromatin organizing protein CCTC-binding factor (CTCF) is a transcription factor (TF) and a key gene in the regulation of gene expression by forming 3-dimensional chromatin interactions.[2,3] CTCF serves as an anchor at boundaries for topological associated domains (TADs), which defines compartments of genomic regions that are transcriptionally coregulated.[4] Importantly, CTCF plays a key role in bringing promoters and enhancers together by either facilitating or inhibiting the contacts, as well as acting as an insulator through the isolation of genomic regions.[5] Together with cohesin, CTCF organizes the looping of DNA at a 40-kb level, and the orientation of the binding motif sequence is crucial for the loop formation.[6,7] However, CTCF does not seem to affect higher order genomic compartmentalization.[8] CTCF binding is at least in part sensitive to DNA methylation,[9,10] which

suggests that DNA methylation can affect the CTCF-binding landscape and thereby the shaping of chromatin architecture. CTCF can also protect DNA from being methylated, and loss of CTCF binding results in hypermethylation of CTCF-binding sites (CBSs).[11] CTCF has been implicated in carcinogenesis in general[11] and has shown to be critical for normal development and hematopoiesis.[12-14] Studies of aberrant CTCF binding in primary cancers have been scarce, and studies in AML are lacking.

AML is a heterogeneous disease characterized by chromosomal aberrations and/or genetic mutations in individual genes. Especially in cytogenetically normal AML, TF genes (ie, *CEBPA* and *RUNX1*), epigenetic genes (ie, *DNMT3A*, *TET2*, *AXLH1*, *IDH1*, and *IDH2*), and the nucleophosmin 1 gene (*NPM1*) are frequently mutated.[15] A recent study found that CTCF directly interacts with NPM1 and that this interaction is critical for insulating functions of CTCF.[16] Mutated NPM1 protein gets trapped in the cytosol and loses its ability to shuttle back to the nucleus.[17] AML with mutated *NPM1* has also been found to be associated with hypermethylation in AML.[18]

The TET2 (ten-eleven-translocation-2) enzyme converts 5-methylcytosine (5mC) DNA to 5-hydroxymethylcytosine (5hmC), which constitutes a step in the demethylation process.[19] Consequently, loss-of-function mutations in *TET2* cause global DNA hypermethylation.[20] The TET2 enzyme also regulates alternative splicing through oxidation of 5mC at CTCF-binding sites. Indeed, a decreasing TET2 activity level results in increased 5mC, which causes CTCF eviction and promotes exon exclusion.[21] Myelodysplastic syndrome and AML patients with *TET2* mutations have been suggested to respond better to treatment with hypomethylating agents.[22,23]

Taken together, we hypothesized that aberrant CTCF occupancy could be involved in regulating leukemogenic gene expression patterns in AML and that AML with mutant *NPM1* and *TET2* would be of special interest to study with regard to CTCF binding. Our study reveals that AML is associated with an aberrant CTCF occupancy pattern characterized mainly by gain of CTCF binding enriched for motifs for key myeloid TFs. *TET2*-mutated AML stood out as a type of AML with an especially strong CTCF binding, while AML with *NPM1* mutations did not display a distinct CTCF-binding pattern. In general, there was a strong anticorrelation between DNA methylation and CTCF binding, while the effect on gene expression was dependent on the location as well as binding on the motifs for the aberrant occupancy. In addition, exposure to azacitidine resulted in major changes in CTCF binding that were associated with the antiproliferative effects induced by the drug.

Methods

Patient material and normal CD34+ cells

BM mononuclear cells (MNCs) were obtained from 14 normal karyotype AML patients as well as 16 healthy donors after informed consent, following institutional review board approval and the Declaration of Helsinki recommendations. AML samples were vitally frozen and thawed before fluorescence-activated cell sorting (FACS) sorting, while normal BM (NBM) cells were sorted and used fresh.

FACS

BM-MNCs from AML patients were thawed and marked with anti-human antibodies against CD3, CD19, CD45 (all from BD Biosciences), CD33 (eBioscience), and NKp46 (Beckman Coulter) to purify AML cells and exclude normal T, B, and natural killer cells. After staining and sorting, AML cells were divided for RNA and DNA extraction and fixing for chromatin immunoprecipitation sequencing (ChIP-seq) analysis. Anti-CD61 (BD Biosciences) was used to analyze K562 differentiation.

DNA, RNA extraction, RNA-seq, and immunoblotting

Following sorting, cells were lysed with RLT buffer, and DNA and RNA were extracted using an AllPrep Kit (Qiagen) or TRIzol (Invitrogen) according to manufacturer protocol. RNA sequencing (RNA-seq) was performed as previously described.[24] Protein lysate was loaded on sodium dodecyl sulfate gel, and antibodies against CTCF (Abcam) and actin (Abcam) were used. For details, see supplemental Methods (available at the *Blood* Web site).

ChIP-seq

For more information about the fixing protocol and ChIP-seq, see supplemental Methods.

DNA methylation

Methylation assay was performed using Infinium MethylationEPIC BeadChip (Illumina). For more details see supplemental Methods.

Cell lines, CTCF knockdown, and azacitidine exposure

K562, HAP1^{TET2-}, and HAP1^{TET2+} cells were purchased from Horizon Discovery and cultured in Iscove Modified Dulbecco Medium (Gibco). CTCF was knocked down by small interfering RNA (siRNA) in K562 cells as described in supplemental Methods. Mononuclear cells from 4 patients were thawed and cultured with and without 1μM azacitidine every 12 hours for 5 days. After incubation with azacitidine, cells were FACS sorted and analyzed as described above (for details, see supplemental Methods).

RNA-seq data from the ClinSeq cohort

Expression data for the comparison of CTCF expression between AML^{TET2mut} and AML^{TET2wt} was retrieved from RNA-seq analysis from our previously published ClinSeq cohort.[24,25]

Data analysis

For a detailed description, see supplemental Methods. In short, DNA methylation data were processed using ChAMP[26] and Limma packages. Data were normalized using BMIQ[27] or FunctionalNormliazation[28] for treated samples. ChIP-seq reads were aligned against reference genome GRCh38 with bowtie.[29] Peaks were called with model-based analysis of ChIP-seq,[30] and the DiffBind package was used to run differential binding analysis. ChIPpeakAnno was used for downstream analysis of peaks. Raw sequencing read quality was assessed by FastQC.[31] For RNA-seq data, Star was used to align the reads to the reference genome GRCh38. FeatureCounts from subread was used to count the fragments in the exon regions as defined in the corresponding file. The edgeR package was used for RNA normalization using trimmed mean of M values to define a design matrix based on the experimental design, fit gene-wise generalized linear models, and conduct likelihood ratio tests for the selected group comparisons.

Results

CTCF binding is perturbed in AML with major gain of CTCF occupancy, especially in *TET2*-mutated AML

ChIP-seq for CTCF was performed on leukemic cells from 14 AML patients and CD34+ cells from the BM (NBM) of 5 healthy donors (Figure 1A). Samples were selected to represent subgroups of AML with *TET2*, and *NPM1* mutations (supplemental Table 1), and patients with *DNMT3A*, *IDH1*, and *IDH2* mutations were excluded. A CBS was defined as a ChIP-seq peak present in ≥2 samples among all analyzed samples. In total, 78 165 CBSs were identified in AML samples (AML^{all}, representing all AMLs regardless of mutations) and 61 376 in NBM samples (Figure 1B). Among CBSs, 17 507 were unique to AML^{all} and 718 were unique to NBM, while 60 658 were represented in both AML and NBM (Figure 1B). CBSs in AML^{all} vs NBM were significantly
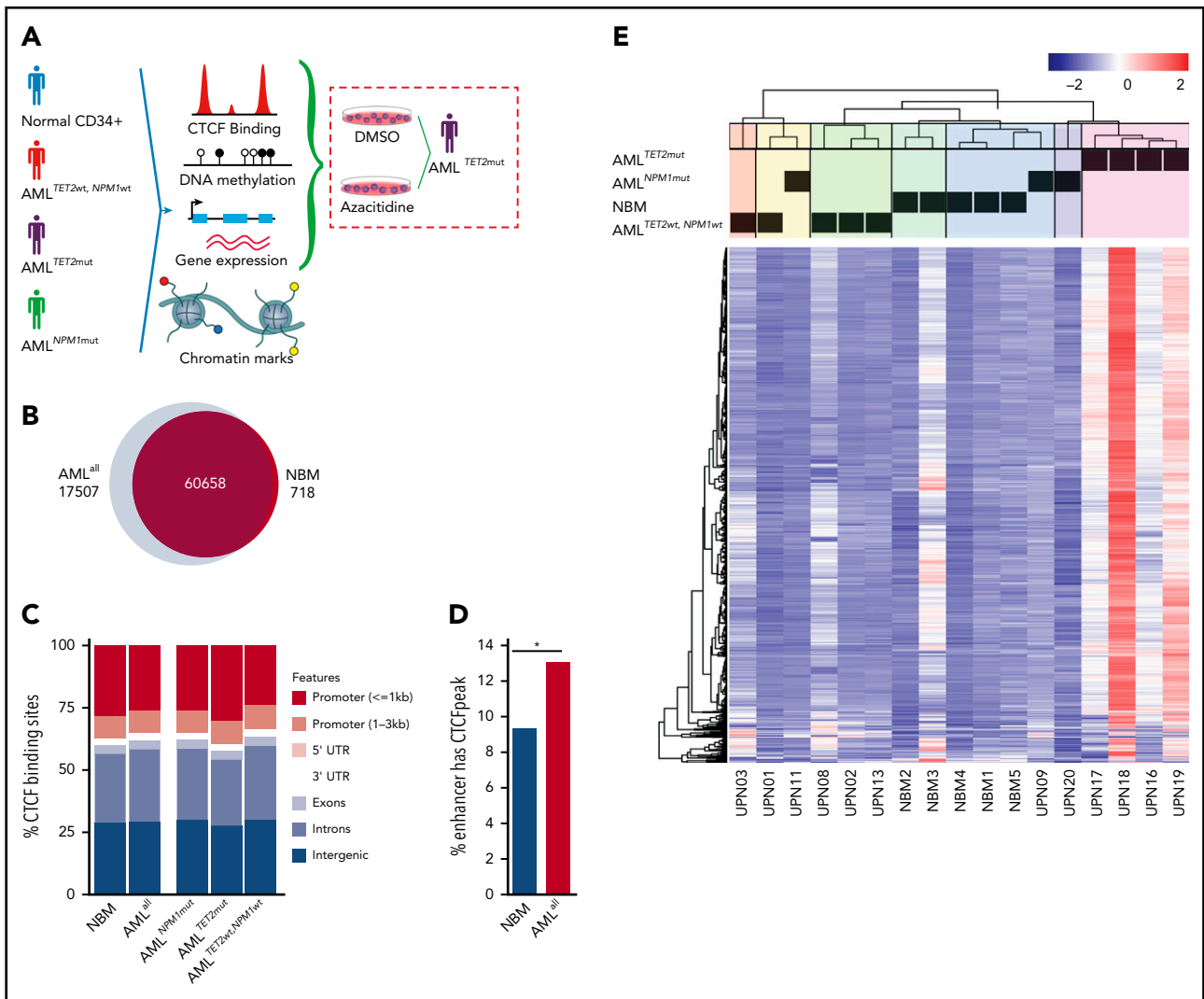
Figure 1. CTCF occupancy and its distribution within the genome. (A) Experimental outline illustrating the analyses of patient samples. (B) Venn diagram displaying unique and common CTCF peaks in AML[all] and NBM cells. (C) Distribution of CTCF-binding sites over genomic features. (D) Percentage of enhancers that contain a CTCF-binding site (Z-test, *P < .01). (E) Heatmap of unsupervised hierarchical clustering for the 1000 most variably bound CTCF peaks based on interquartile range measure. DMSO, dimethyl sulfoxide; UTR, untranslated region.

skewed toward increased binding in introns and intergenic sites (P < .05), while CBSs in *TET2*-mutated samples (AML[TET2mut]) showed enrichment in promoter regions compared with *TET2* wild-type samples (AML[TET2wt]) (Figure 1C; supplemental Table 2). Significantly more enhancers were bound to CTCF in AML[all] compared with NBM (Figure 1D). In an unsupervised clustering of the most variable CBSs, AML clustered separately from NBM, while AML[TET2mut] clustered separately from other AMLs and displayed a strong increase in CTCF occupancy. In contrast, AML with *NPM1* mutations (AML[NPM1mut]) did not cluster separately from *NPM1* wild-type samples (AML[NPM1wt]) (Figure 1E). A principal-component analysis (PCA) plot further confirmed these separations (supplemental Figure 1A). There was no difference in CTCF expression between AML[all] and NBM or between AML with or without *TET2* mutations, suggesting that differences in CTCF expression do not explain the pattern of gained CTCF occupancy (supplemental Figure 1B-C).

We then characterized differential CTCF occupancy and found 2064 differentially bound CTCF-binding sites (DBCs) between

AML[all] and NBM (false discovery rate [FDR] <0.05), where 1803 (87.4%) were gained and 261 (12.5%) lost in AML[all] (Figure 2A; supplemental Table 3A,D). We also compared AML[TET2mut] and AML[TET2wt] vs NBM, which revealed 15 021 DBCs (13 311 (88.6%) gained) for AML[TET2mut] vs NBM (Figure 2B; supplemental Table 3A,E) and 891 DBCs (539 (60.5%) gained) for AML[TET2wt] vs NBM (Figure 2C; supplemental Table 3A,F). In AML[TET2mut] vs AML[TET2wt], we identified 3217 DBCs, of which 2682 (83.6%) were gained and 526 (16.4%) lost in AML[TET2mut] (Figure 2D; supplemental Table 3A,G). In contrast, AML[NPM1mut] only showed 101 DBCs compared with AML[NPM1wt], confirming low impact on CTCF binding in AML[NPM1mut] (Figure 2E; supplemental Table 3A,H). The overlap between DBCs from the AML[all] vs NBM and the 2 mutation-specific comparisons was small (Figure 2F). Interestingly, the AML[TET2mut] vs AML[TET2wt] comparison revealed more unique DBCs than the AML[all] vs NBM comparison (3070 vs 1934), showing a greater difference in mutation related CTCF occupancy within AML rather than between AML and normal cells. Clustering of all DBCs is shown in supplemental Figure 2A.
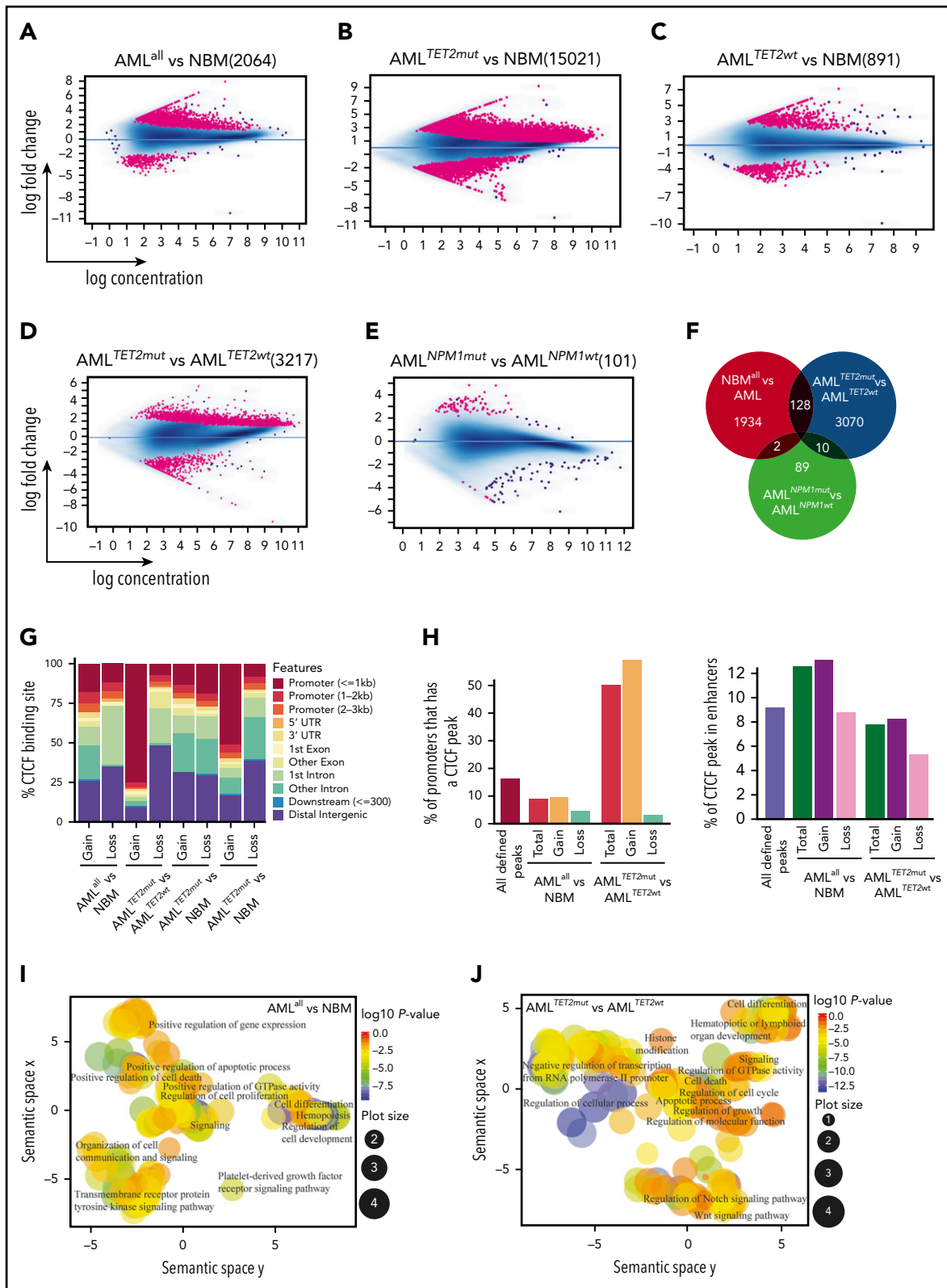
Figure 2. CTCF differential occupancy. (A-E) MA plots of differentially bound CTCF sites between AML[all] and NBM (A), AML[TET2mut] and NBM (B), AML[TET2wt] and NBM (C), AML[TET2mut] and AML[TET2wt] (D), and AML[NPM1mut] and AML[NPM1wt] (E) (FDR <0.05). (F) Venn diagram showing DBC overlap for comparisons between AML[all] vs NBM (pink circle), AML[TET2mut] vs AML[TET2wt] (blue circle), and AML[NPM1mut] and AML[NPM1wt] (green circle). (G) Distribution of gained and lost CTCF binding over genomic features for AML[all] and AML[TET2mut] specific DBCs. (H) Enrichment of differentially bound CTCF peaks in promoters (left) and enhancers (right) compared with all detected CTCF peaks. (I-J) Scatterplot of GO terms of gained DBCs in AML[all] vs NBM (I) and AML[TET2mut] vs AML[TET2wt] (J). Circle color indicates P value (−log10 P value) and circle size the frequency of the GO terms. (K) Motif analysis of all DBCs in AML[all] vs NBM. (L) ChIP quantitative PCR using anti-CTCF in K562-CTCF knockdown cells (*P < .05; Student t test). (M) ChIP quantitative PCR using anti-RUNX1 in K562-CTCF knockdown cells; scramble control was set to 1 (dashed line), and values were normalized against scramble control (*P < .05; Student t test).
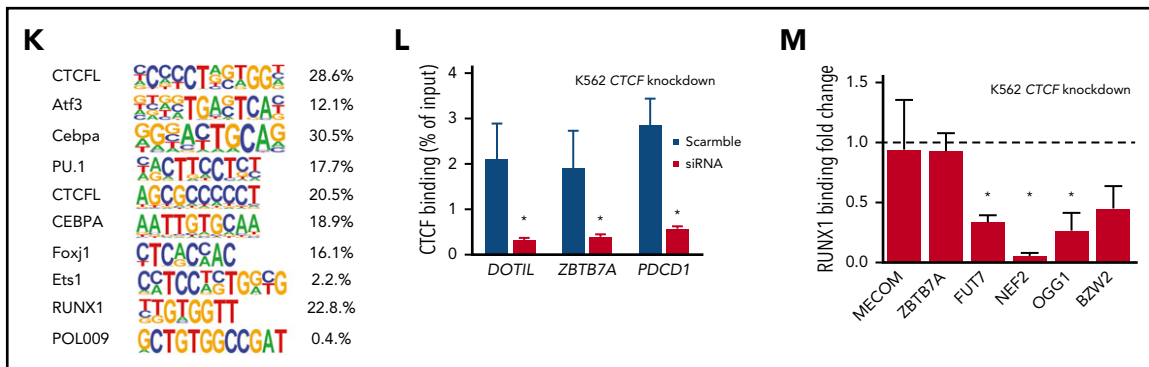
Figure 2. (Continued).

TET2-mutated AML is characterized by an enrichment of CTCF occupancy in promoters

We further analyzed the distribution of DBCs over genomic features and again, AML^{TET2mut} presented with a distinct pattern. Three-quarters (74%) of gained DBCs in AML^{TET2mut} vs AML^{TET2wt} were located within promoters and only 12% in distal intergenic regions, while almost half (48%) of the lost CBSs were located in distal intergenic regions and only 7% in promoter regions (Figure 2G; supplemental Table 3B). Gained and lost CBSs in AML^{all} vs NBM were distributed more similarly but with slightly more gain in promoters and "other" introns and more losses in distal intergenic regions and first introns ($P < .05$) (Figure 2G; supplemental Table 3B). Enhancers, as defined by the FANTOM project,[32] gained CTCF binding in AML^{all} as well as AML^{TET2mut} (supplemental Table 3C). DBCs were enriched for enhancers in AML^{all} vs NBM, while DBCs for AML^{TET2mut} vs AML^{TET2wt} were highly enriched for promoters ($P<.001$; Figure 2H). Gene Ontology (GO) analysis of gained AML^{all} specific as well as gained AML^{TET2mut}-specific DBCs related to cell proliferation, cell death, and cell differentiation, while AML^{TET2mut}-specific DBCs also showed changes in Notch and Wnt signaling pathways[33,34] (Figure 2I-J; supplemental Table 4A-B). Scatterplots of GO terms of lost sites as well as list of all GO terms for lost and gained are shown in supplemental Figure 2B-C and supplemental Table 4A-D.

To better understand the differences in CTCF binding seen in AML^{TET2mut}, we studied CTCF binding by ChIP polymerase chain reaction (PCR) in TET2-knockout HAP1 cells. CTCF binding was found to be increased at some CBSs, such as DOT1L, but also unchanged as well as decreased at other loci (supplemental Figure 3F). Thus, TET2 activity seems to affect CTCF binding, while the pattern of gain and loss appear to be context dependent. The results also suggest that other factors than the direct enzymatic activity of TET2 could be responsible for the change in CTCF occupancy in AML^{TET2mut}.

Differentially bound CTCF sites are enriched at key myeloid TFs, and loss of CTCF binding leads to partial loss of RUNX1 binding and altered cellular functions

The top motif for DBCs in AML^{all} vs NBM included CEBPA, PU.1, ETS1, and RUNX1, which are TFs known to be involved in AML development[35,36] (Figure 2K; supplemental Table 5). Other top

motifs such as ATF3, DDIT3 (C/EBPzeta, CHOP), and FOSL2[37] were significantly upregulated in the AML^{all} samples (supplemental Figure 3A), which may contribute to their AML-specific motif enrichment. CCCTC-binding factor like (CTCFL) that share the same binding motif as CTCF,[38] was the most significantly enriched motif (Figure 2K). CEBPA motifs were found in the top gained, as well as top lost, DBCs, whereas BCL11A, Klf4, and Plagl1 were enriched for gained and AP-1 and PU.1 for lost DBCs (supplemental Figure 3C-D). DBCs for AML^{TET2mut} vs AML^{TET2wt} were also enriched for several important TFs suggested to play roles in hematopoiesis or leukemia, such as the KLF7,[39] SOX4,[40] FOXH1,[41] HIC1,[42] and BCL11A[43] (supplemental Figure 3E).

To investigate whether a loss of CTCF binding also affected binding of neighboring TFs, we knocked down CTCF by siRNA in K562 cells and analyzed binding of CTCF (Figure 2L) and RUNX1 (Figure 2M) at cobinding regions by ChIP-PCR. CTCF knockdown caused a loss of CTCF binding as well as loss of RUNX1 binding at most, but not all, investigated sites (Figure 2L; supplemental Figure 3B), which shows RUNX1 binding to be dependent on CTCF occupancy in a context-dependent manner.

CTCF binding is known to impact the surrounding chromatin. In order to study histone changes at gained and lost DBCs, ChIP-seq was performed for H3K18ac, H3K27ac, and H3K4me1 in 5 AML samples (UPN01, UPN03, UPN15, UPN17, and UPN20) and H2A.Z in 4 AML samples (UPN01, UPN15, UPN17, and UPN20) and 4 NBM samples (Figure 3A). Differential binding to histone marks was calculated for DBCs between AML^{all} and NBM samples. Histone marks associated with active transcription and accessible chromatin, such as H3K18acetylation, H3K27acetylation, and H3K4me1, increased at gained CTCF-binding sites, while they decreased at lost DBCs in AML. The same pattern was thus observed for the promoter associated mark (H3K18acetylation), as well as the active enhancer mark (H3K27acetylation). Histone variant H2A.Z has been shown to increase chromatin accessibility for transcriptionally activating as well as repressive complexes,[44] and we found lost CTCF binding in AML^{all} to be associated with decreased H2A.Z occupancy (Figure 3A).

In addition, we studied whether cellular functions can be altered by CTCF knockdown in K562 cells. Transient knockdown by siRNA was effective at the mRNA and protein levels (Figure 3B-C) and caused a significant increase of the megakaryocytic
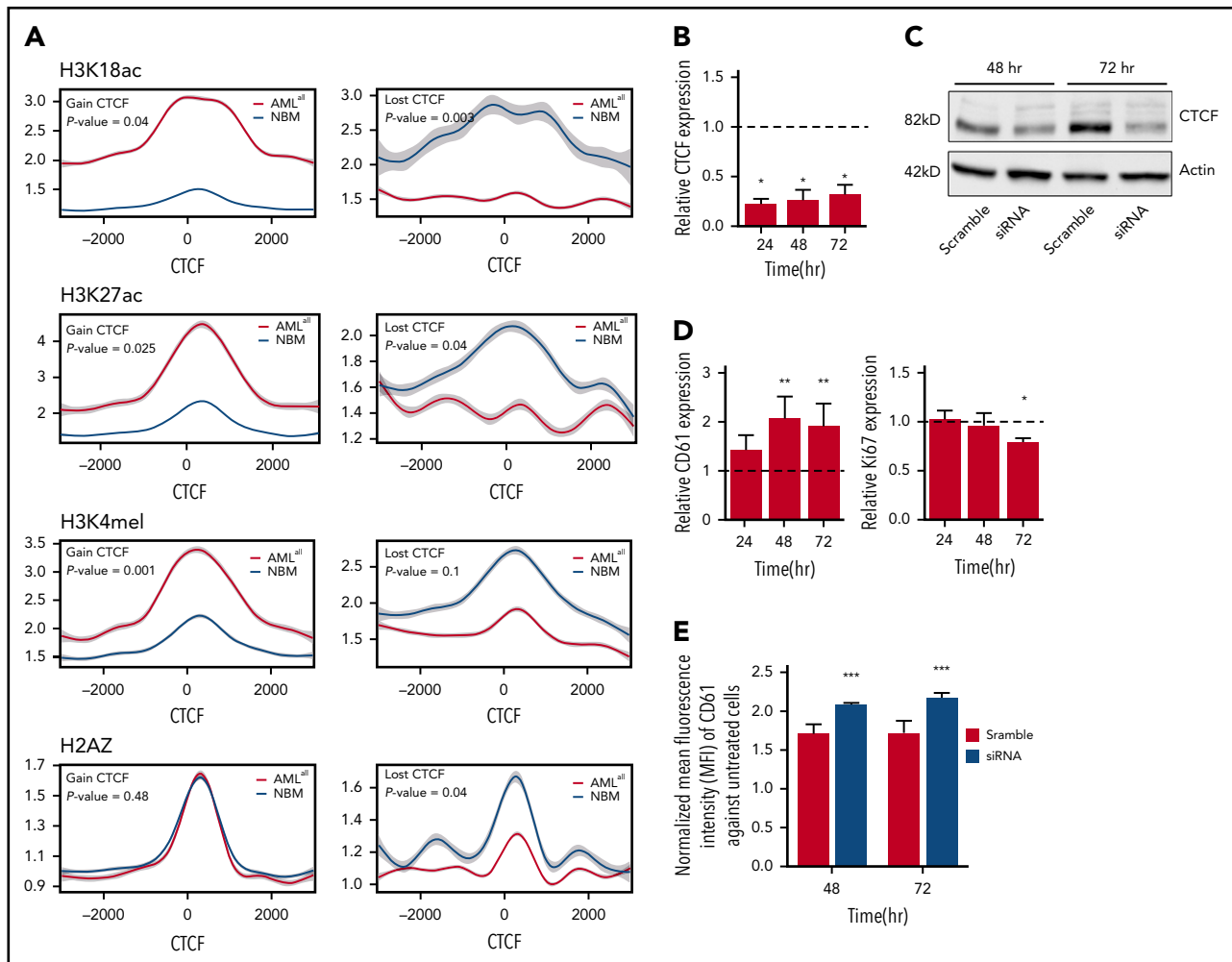
Figure 3. CTCF changes and their correlation with histone marks and effects on cell differentiation. (A) ChIP-seq on histone marks (H3K18ac, H3K27ac, H3K4me1, and H2AZ) displays changes in histone modifications around gained (left) and lost (right) CTCF-binding sites between AML^{all} and NBM. (B) CTCF expression by real-time quantitative PCR following its knockdown in K562 cells. Relative gene expression was calculated using $2^{-\Delta\Delta Ct}$ against GAPDH, while scramble control was set to 1 (dashed line) and knockdown was normalized against the scramble ($P < .05$; Student t test). (C) Western plot showing CTCF protein levels following CTCF knockdown. (D) CD61 and Ki67 expression by real-time quantitative PCR in K562 after CTCF knockdown; normalization was done as described in panel B. (E) Detection of surface marker CD61 on K562-CTCF knockdown by FACS. Mean florescence intensity was used to calculate the expression level of CD61 after normalization against untransfected K562 cells ($P < .001$; Student t test). *$P < .05$, **$P < .01$, ***$P < .001$.

differentiation marker CD61, a key differentiation marker in K562 cells (Figure 3D-E). In addition, the proliferation marker Ki67 was decreased after 72 hours (Figure 3D). Together, this suggests that knockdown of CTCF causes altered cellular function, particularly differentiation, in a leukemic cell line. It should be noted that altered cellular functions after CTCF knockdown have been shown to be dependent on cell type, type and efficacy of knockdown, and the time course.[8]
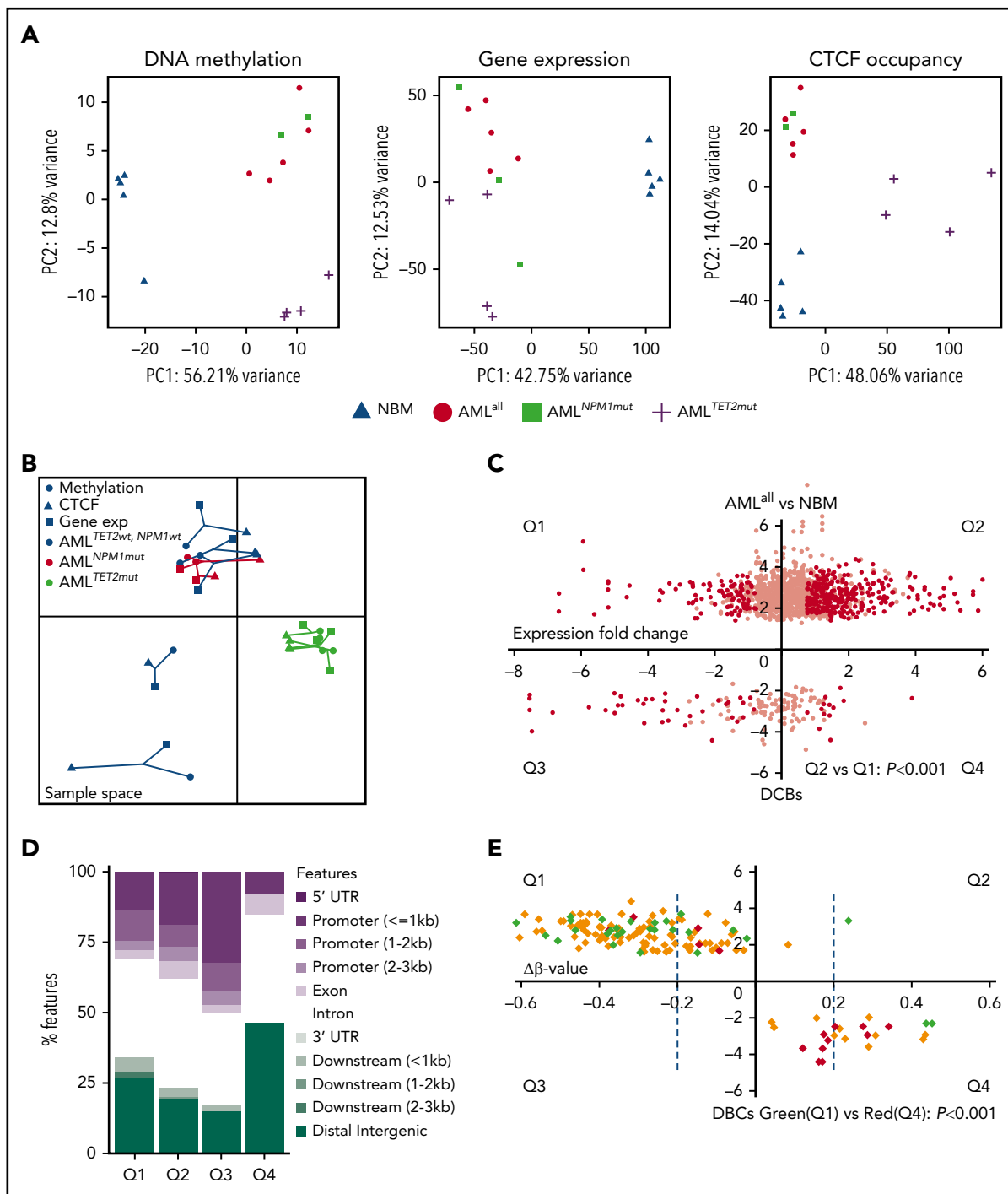
DNA methylation and gene expression analyses in AML and NBM cells

To integrate CTCF binding with DNA methylation and gene expression, EPIC methylation arrays and RNA-seq were performed on the same AML samples that had undergone ChIP-seq. AML^{all} samples displayed hypomethylation compared with NBM, with 70% of differentially methylated cytosine guanine dinucleotide (CpGs) (DMCs) being hypomethylated (supplemental Figure 4A). Similarly, 84.8% of DMCs in AML^{NPM1mut} were hypomethylated compared with AML^{NPM1wt}, while 93.2% of DMCs in AML^{TET2mut} were hypermethylated compared with

AML^{TET2wt} (supplemental Figure 4A). Hypomethylation in AML^{TET2mut} occurred preferentially in the promoters, whereas hypermethylation was overrepresented in distal regions (supplemental Figure 4B). Differentially expressed genes (DEGs) were also defined between groups by RNA-seq data (supplemental Figure 4C and supplemental Table 6). PCA analyses using data from each omics analysis (data on DBCs, DMCs, and DEG) show the degree of separation between groups and underline the separation of NBM and AML^{TET2mut} with regard to CTCF binding and DNA methylation (Figure 4A).

CTCF binding and DNA methylation anticorrelates in AML and NBM cells

To identify corelationships between multiple omics datasets, we used the integrative multiple coinertia analysis (Figure 4B). This analysis integrated differential loci for all samples that were covered by all 3 omics analyses (ChIP-seq, EPIC methylation array, and RNA-seq), which included all AML samples. Most strikingly, AML^{TET2mut} samples were highly correlated and distinguished from other AML samples.

Figure 4. Correlations among CTCF occupancy, DNA methylation, and gene expression. (A) PCA analysis of DMCs (left), DEGs (middle), and DBCs (right) based on analyses from EPIC DNA methylation array, RNA-seq, and ChIP-seq for CTCF, respectively. (B) Multiple coinertia analysis of DMCs, DBCs, and DEGs based on EPIC DNA methylation, RNA-seq, and ChIP-seq data, where data from all 3 analyses performed in the same samples were integrated. Each patient is represented by filled circles for DNA methylation data, triangles for ChIP-seq data, and filled squares for gene expression data, while the length of the line that connects the 3 omics types is proportional to the divergence of the data in the same sample. (C) Dot plot of DBCs (AML[all] vs NBM) in relation to the change in expression of the related genes comparing AML[all] and NBM. The x-axis represents fold change of gene expression, and the y-axis fold change in CTCF occupancy (FDR <0.05). Black dots indicate genes with statistically significant change expression (FDR <0.05, fold change >1.5). χ^2 test was used to calculate statistical significance between Q1 and Q2 (P < .001). (D) Bar plot showing the genomic distribution of DBCs with differentially expressed gene in panel C. (E) Dot plot of DBCs in relation to fold change of β-values (Δβ-value) for AML[all] vs NBM, where the x-axis represents fold change in β-value and the y-axis fold change in CTCF occupancy. Statistical significance is indicated by dashed blue lines for Δβ-value (FDR <0.05, fold change >1.5). Green and red diamonds represent up- or downregulated genes, respectively, while yellow diamonds represent genes with no change in gene expression. χ^2 test used to calculate statistical significance in gene expression between Q1 and Q4 (P < .001). PC, principal component.

DYSREGULATED CTCF OCCUPANCY IN AML

🩸 blood® 16 JULY 2020 | VOLUME 136, NUMBER 3 **345**
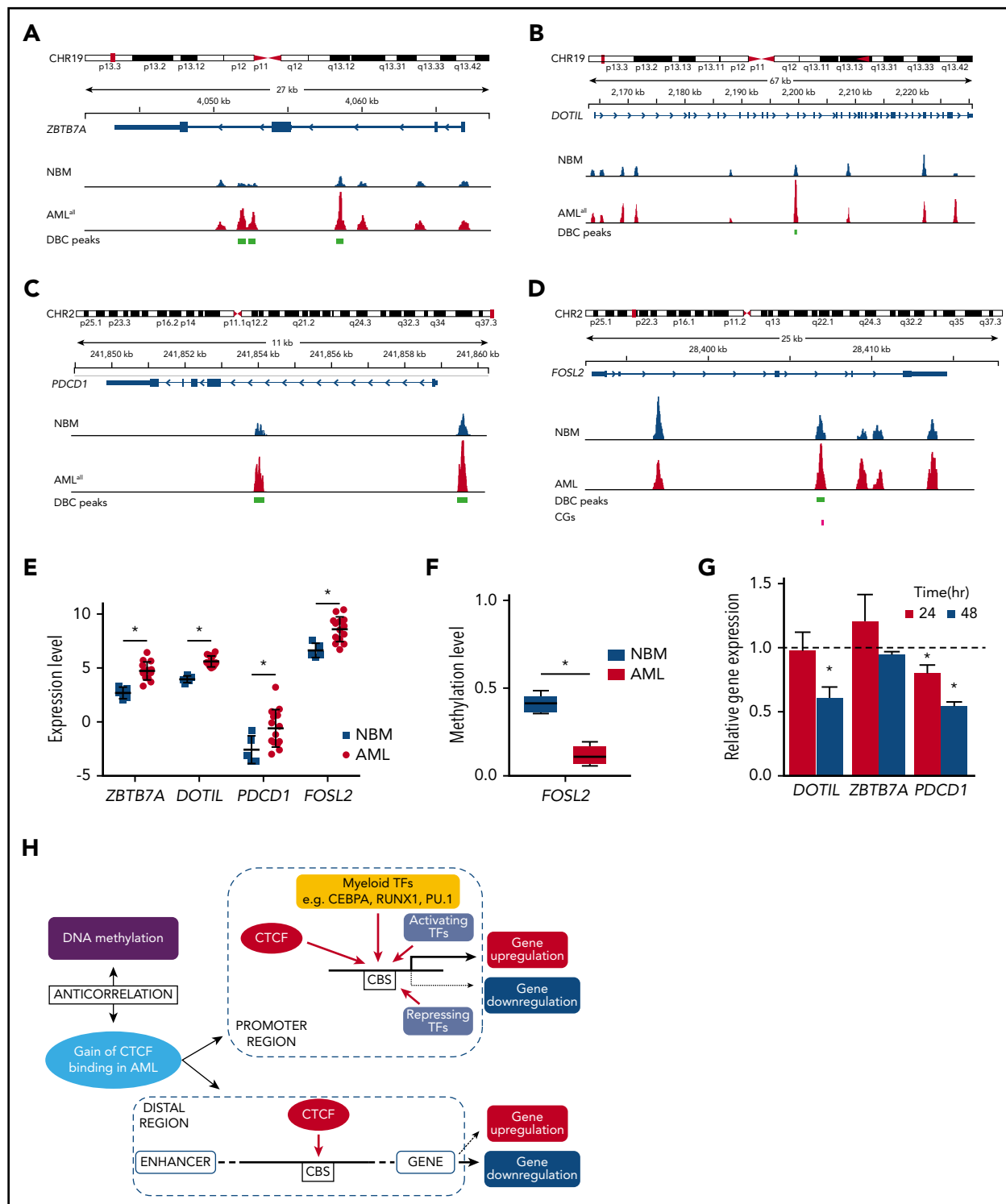
Figure 5. Genes with aberrant CTCF occupancy and changed gene expression in AML. (A-D) Integrative Genomics Viewer tracks visualize CTCF peaks from ChIP-seq results of AML^{all} and NBM for 4 genes: *ZBTB7A* (A), *DOT1L* (B), *PDCD1* (C), and *FOSL2* (D). Green boxes indicate differentially bound peaks, and pink boxes indicate a CpG detected by the EPIC array that overlaps with a DBC peak. (E) Expression levels of the 4 genes in panels A-D from RNA-seq data (*FDR <0.05, fold change >1.5). (F) Box plot of β-values for the CpG in *FOSL2* (*FDR <0.05). (G) Gene expression by quantitative PCR following CTCF knockdown in K562 cells. Relative gene expression was calculated using $2^{-\Delta\Delta Ct}$ against GAPDH, while scramble control was set to 1 (dashed line) and knockdown was normalized against the scramble control ($P < .05$; Student *t* test). (H) Hypothetical model of aberrant CTCF binding and how it relates to DNA methylation and gene expression in AML. CHR, chromosome.

In order to investigate how CTCF correlates with DNA methylation and gene expression, we integrated the data sets by overlapping CTCF peaks with CpGs in the EPIC array and related genes from RNA-seq. Regardless of sample type (AML^{all} or NBM), there was a strong anticorrelation between CTCF binding and DNA methylation (within 30 bp from the center of the CTCF
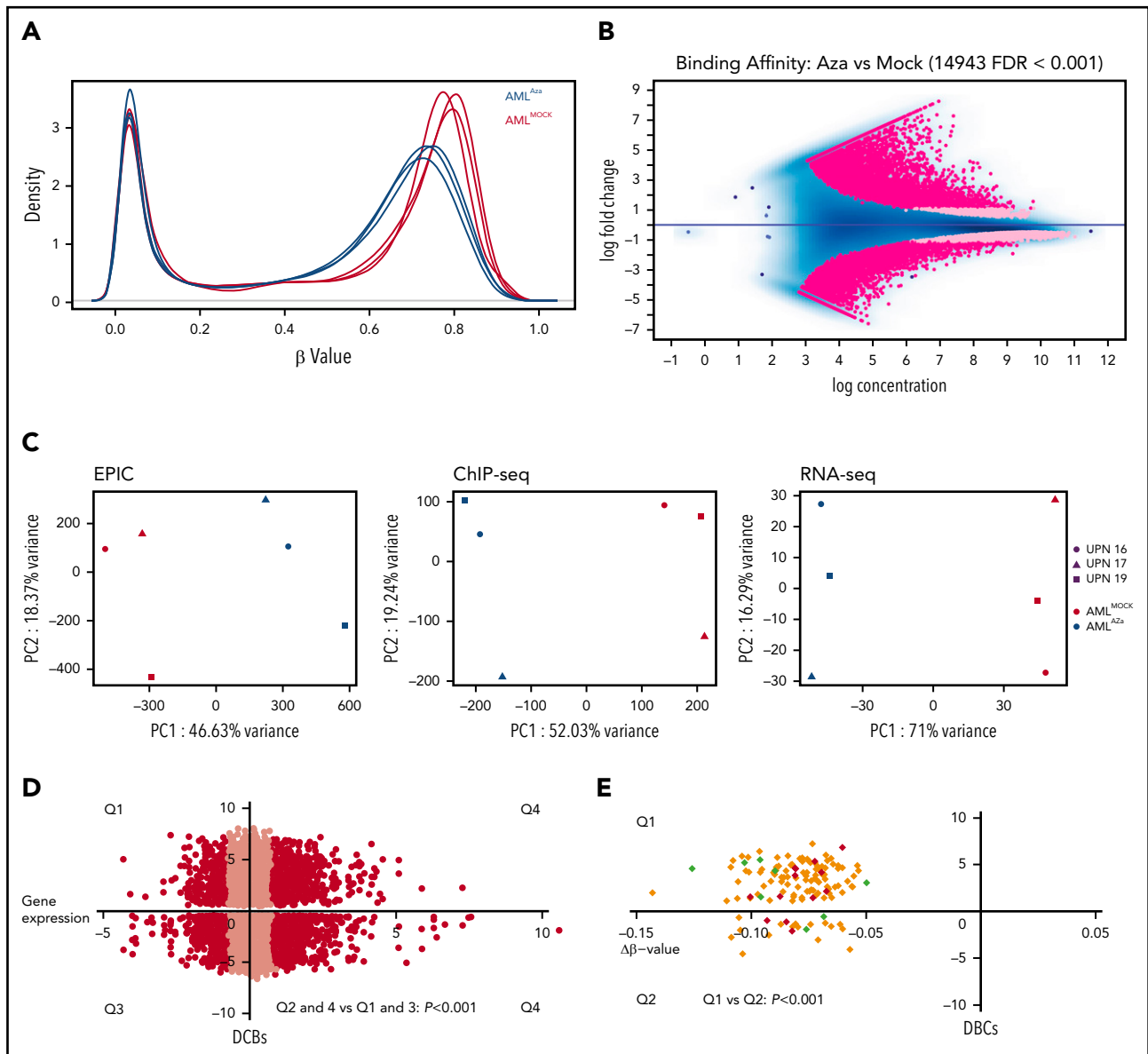
Figure 6. Azacitidine induces hypomethylation and changes in CTCF occupancy. (A) Density plots for β-values of all normalized CpGs (n = 723 768) in azacitidine and mock-treated AML samples. (B) MA plot of differentially bound CTCF sites between azacitidine- and mock-treated samples (FDR <0.001). Dark pink dots show peaks with fold change >1.5. (C) PCA analysis of differentially methylated sites (left), differentially bound sites (middle), and differentially expressed genes (right) based on EPIC array data, ChIP-seq for CTCF, and RNA-seq, respectively, in azacitidine- vs mock-treated AML samples. (D) Dot plot of DBCs in relation to changes in gene expression between azacitidine- vs mock-treated samples. The x-axis represents fold change in gene expression and the y-axis fold change in CTCF occupancy (FDR <0.001). Black dots represent genes with statistically significant changes in gene expression (FDR <0.05, fold change >1.5); χ^2 test was used to calculate statistical significance between up- and downregulated genes ($P < .001$). (E) Dot plot of DBCs in relation to fold change in DNA methylation by β-values (Δβ-value) for azacitidine- vs mock-treated samples. The x-axis represents Δβ-value (FDR<0.1) and the y-axis fold change in CTCF occupancy (FDR <0.001). Green and red dots represent up- or downregulated genes, respectively (FDR <0.05, fold change >1.5); χ^2 test was used to calculate statistical significance between Q1 and Q2 ($P < .001$).

peak; $P = 2.2\text{-}4.2 \times 10^{-16}$) (supplemental Figure 5). CTCF binding was dominantly found at unmethylated sites with β-values <0.10. Anticorrelation between CTCF binding and DNA methylation was also seen for DBCs in AML[all] in enhancer locations (supplemental Figure 6B).

To understand why AML[TET2mut] exhibited strong gain of CTCF occupancy and still showed genome-wide DNA hyper-methylation, we investigated the degree of overlap between DMCs and CBSs. Importantly, of the hypermethylated DMCs in AML[TET2mut], only 7% (62 of 809) overlapped with a CBS. In

contrast, 31% of hypomethylated DMCs in AML[all] vs NBM were located in a CBS (supplemental Figure 6A). Also, gains of CTCF in AML[TET2mut] were primarily seen at promoters where hypo-methylation was overrepresented in AML[TET2mut] (Figure 2G; supplemental Figure 4B). In addition, DBCs in promoters showed little overlap with DMCs in AML[TET2mut] vs AML[TET2wt] (supplemental Figure 6E). Furthermore, gained DBC in AML[TET2mut] occurs at sites that are hypomethylated in both AML[TET2mut] and AML[TET2wt] (supplemental Figure 6F). Thus, hypermethylation in AML[TET2mut] occurs dominantly outside of CTCF-binding sites. The same degree of anticorrelation

between CTCF binding and DNA methylation is also seen in both AML[TET2mut] and AML[TET2wt] (supplemental Figure 6F-G).

Increased CTCF binding in promoter regions is associated with DNA hypomethylation and increased gene expression, while CTCF changes in distal regions anticorrelate with gene expression

We next investigated whether aberrant CTCF binding was related to proleukemic gene expression patterns. Therefore, all DBCs were investigated for changes in expression of related genes. In AML[all], gain of CTCF binding correlated with transcriptional upregulation (quadrant 2 [Q2] vs Q1; $P < .001$), while loss correlated with transcriptional downregulation (Q4 vs Q3; $P < .05$) (Figure 4C). Still, some genes displayed anticorrelation between CTCF binding and gene expression. DBCs in promoters were more likely to show a positive correlation to gene expression, while changes in distal intergenic regions were more likely to show an anticorrelation (Figure 4D). This finding could reflect the different roles of CTCF as a TF at promoters and as an insulator at gene distal locations where the location in relation to TAD boundaries likely is of importance.[8,45] In addition, DBCs that correlate with active transcription were associated with binding motifs of transcriptional activators, such as RUNX2, FOSL2, OSR1, and FOXA2,[37,46-48] while DBCs that anticorrelate with gene expression are associated with motifs of transcriptional repressors, such as BCL6, PRDM1, and YY2[49-54] (supplemental Table 8A-B).

DNA methylation data were also integrated with both CTCF binding and gene expression for DBCs in AML[all] samples. Almost all sites that gained CTCF binding were hypomethylated (Figure 4E), while all sites that lost CTCF binding were hypermethylated. Gene expression decreased at hypermethylated sites that lost CTCF binding, while it was increased at hypomethylated sites that gained CTCF binding (Figure 4E). For AML[TET2mut] vs AML[TETwt] same trends were found (supplemental Figure 6C-D), but due to low overlap between AML[TET2mut] specific DBCs and DMCs, changes were not statistically significant.

Among genes that increased regarding both CTCF binding and gene expression in AML were DOT1L, a histone H3K79 methyltransferase related to leukemia development and a well-known therapeutic target[55]; and ZBTB7A, known to be a proto-oncogene, overexpressed in leukemia[56,57] and PDCD-1 (PD-1), which is a key check point in immunomodulation (Figure 5A-C,E). In addition, FOSL2, known to promote cell growth and constitutively expressed in leukemia,[58] was hypomethylated and gained both CTCF binding and gene expression (Figure 5D-F).

In order to investigate the causal relationship between differential CTCF binding and changed gene expression, we knocked down CTCF in K562 cells and studied the expression of genes at DBCs that had also displayed altered gene expression in AML[all] vs NBM. Knockdown led to loss of CTCF at DBCs (Figure 2L) as well as downregulation of related genes (Figure 5G). However, the impact on expression is likely promoter regulation dependent, since not all genes that lose CTCF binding become transcriptionally repressed. A hypothetical model, based on ours and previously published results on CTCF, is shown in Figure 5H and outlines the relationships among aberrant CTCF binding, DNA methylation, and gene expression in AML.

## Azacitidine treatment induces profound changes in CTCF occupancy in AML patient cells

The demethylating agent azacitidine is widely used for treatment of myeloid malignancies, including AML.[22] Therefore, we further investigated how azacitidine exposure affects CTCF occupancy and the resulting gene expression. We incubated samples from 4 AML[TET2mut] patients with and without azacitidine in vitro (at concentrations inducing antileukemic effects; see supplemental Figure 7A) and analyzed samples by ChIP-seq for CTCF, RNA-seq, and EPIC methylation array. Firstly, we confirmed that azacitidine exposure caused demethylation (Figure 6A). Azacitidine also caused a significant change in CTCF binding with 14 943 DBCs. There was an equal number of gained and lost CBSs (7474 vs 7469 lost; FDR <0.001; Figure 6B), but DBCs with more than twofold change were enriched for gained sites (7073 gained vs 5251 lost, dark pink). Top DBCs are listed in supplemental Table 7. A PCA analysis showed that azacitidine had a significant impact on CTCF occupancy, DNA methylation, and transcriptomic level (Figure 6C). Both gain and loss of CTCF binding was significantly associated with increased expression ($P < .001$; Figure 6D). In addition to enrichment for the CTCF motif ($P < 1 \times 10^{-16}$), repressed genes were significantly enriched for the transcriptional repressor BCL11A ($P < 1 \times 10^{-11}$) (supplemental Table 8C) while transcriptionally activated genes were enriched for motifs for transcriptional activators such as PLAGL1 ($P < 1 \times 10^{-19}$) and POU2F2 ($P < 1 \times 10^{-12}$) (supplemental Table 8D). GO term analysis on DBCs that also changed expression upon azacitidine exposure showed terms related to apoptosis, cell death, and cell differentiation among upregulated genes and GO terms related to DNA biosynthesis and DNA replication among downregulated genes (supplemental Figure 7B-C and supplemental Table 9A-B).

We then specifically searched for genes that were demethylated by azacitidine and concomitantly changed CTCF binding and gene expression. Among top genes were KLF6 (Figure 7A-C), SP140, and SLC2A6. SP140 has shown to be specifically methylated in TET2 mutant hematological malignancies.[59,60] SLC2A6 has been implicated in solid tumors and is a marker for chronic lymphocytic leukemia with trisomy 12.[61,62] Most interestingly, KLF6 is epigenetically dysregulated and methylated in AML.[63,64] Furthermore, among CTCF sites gained with azacitidine, 56% (4162 of 7473) were sites normally bound in NBM (Figure 7D), which suggests that azacitidine restores some of the aberrant CTCF-binding pattern in AML.

Discussion

CTCF plays a critical role in gene regulation by shaping the structure of chromatin through DNA looping.[2,3] For the first time, we here investigated CTCF occupancy in AML patient cells and normal CD34[+] BM cells. We investigated the differential CTCF binding in normal vs AML cells and in subgroups of AML cells and integrated CTCF occupancy with genome-wide data on gene expression and DNA methylation. As CTCF binding has been related to DNA methylation,[9,10] we also explored the effects of the demethylating drug azacitidine on CTCF binding and related gene expression patterns.

AML cells displayed an aberrant CTCF-binding pattern compared with normal CD34[+] BM cells with a significant gain of
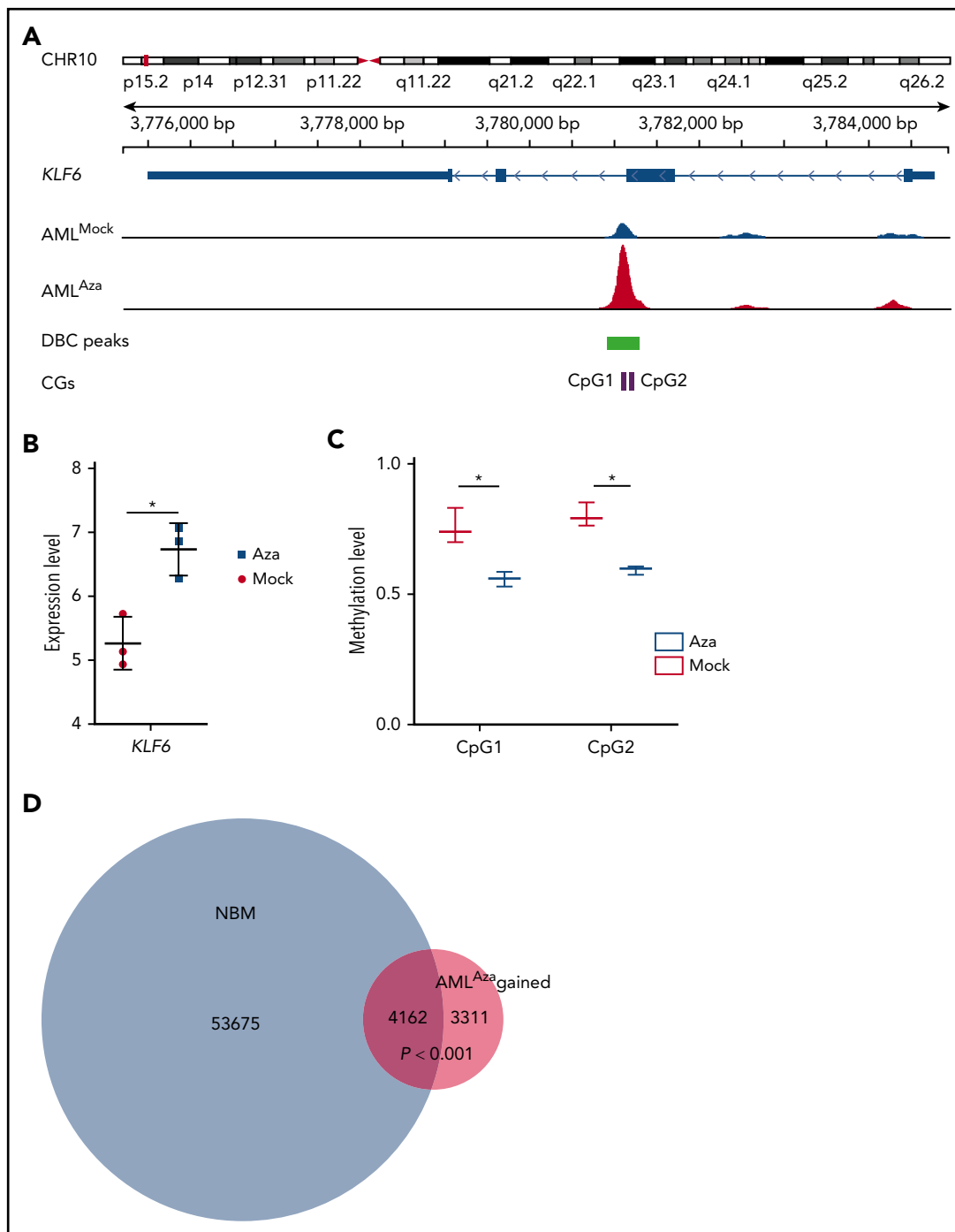
Figure 7. Azacitidine demethylates, causes gain in CTCF occupancy, and increases *KLF6* gene expression. (A) Integrative Genomics Viewer tracks visualize a CTCF peak, the green box represents a differentially bound peak, and the pink boxes a detected CpG by the EPIC array, overlapping with the DBC peak. (B) Expression levels of KLF6 gene (*FDR <0.05, fold change >1.5). (C) Methylation of the 2 CpGs shown in panel A as measured by β-values (*FDR <0.1). (D) Venn diagram showing the overlap CTCF-binding sites gained with azacitidine treatment (red circle) and CTCF-binding sites bound in NBM. There are significant number of gained CTCFs in AML^{aza} that overlap with NBM (Z-test, $P < .001$).

CTCF occupancy that was enriched in enhancers. Relevant to this, we have previously shown the importance of enhancer usage in both normal myelopoiesis[65] and AML.[66] AML samples with *TET2* mutations showed a particularly strong gain of CTCF occupancy, and this gain was highly enriched at promoter locations. As NPM1 binds to CTCF,[16] we hypothesized that mutant *NPM1* could impact on CTCF binding through dysfunctional CTCF nuclear transport. However, *NPM1* mutations did not

significantly impact CTCF binding in our analysis; thus, AML^{NPM1mut} vs AML^{NPM1wt} samples were not studied further in detail.

The top 10 motifs for aberrant CTCF occupancy in AML^{all} included well-known myeloid TFs known to play a role in AML such as CEBPA, PU.1, and RUNX1, which suggest CTCF binding to be involved in the leukemogenic process. The top motifs specific for AML^{TET2mut} were also TFs known to be involved in leukemia

transformation such as SOX4,[40] FOXH1,[41] and HIC1.[42] Furthermore, GO analysis of genes with aberrant CTCF binding and gene expression in AML[TET2mut] were enriched for Notch and Wnt signaling. Moreover, GO analysis of differential CTCF binding in AML[all] as well as AML[TET2mut] showed effects on GO terms related to cell proliferation, cell death, and apoptosis. In addition, sites that changed CTCF binding in AML also displayed epigenetic and chromatin alterations by changes in histone marks. Marks for active chromatin were increased at gained CTCF sites, whereas they were lost at lost CTCF sites. In total, the relation of differential CTCF binding to transformative myeloid TFs and GO terms relevant for leukemia transformation, as well as chromatin changes, suggests that CTCF binding could be involved in regulating leukemogenic gene expression patterns in AML. As a next step, we knocked down CTCF in K562 cells and showed that the knockdown caused decreased CTCF binding as well as RUNX1 binding at specific RUNX1 targets. This shows that occupancy of TFs that bind in the vicinity to CTCF can display CTCF dependency, albeit not at all sites. Knockdown also caused downregulation of genes that were differentially bound to CTCF and differentially expressed in AML. With regard to cellular functions, knockdown caused an increase in cell differentiation and decrease in Ki67 expression, a marker for proliferation. This suggests that changed CTCF binding can alter cellular functions in a leukemogenic direction. Previous studies have shown growth inhibitory effect upon CTCF knockdown[8,67,68]; however this is dependent on the efficacy of the knockdown, cell type, and time course.[8,67,68]

CTCF binding has previously been related to DNA methylation,[10,11,69] and indeed, a strong correlation between CTCF occupancy and DNA methylation was found. Previous studies aiming to decipher the hierarchy of the events have shown that CTCF binding can regulate DNA methylation[67] but also that DNA methylation can direct CTCF binding.[68] In line with previous reports,[70] we show global hypomethylation in AML, which could theoretically be consistent with an increase in CTCF binding. Conversely, AML[TET2mut] patients exhibited global hypermethylation but still an increase in CTCF occupancy. We hypothesized that this unexpected correlation was due to AML[TET2mut]-specific methylation occurring outside of CTCF-binding sites and that CTCF changes occur at sites lacking CpGs, making them insensitive to DNA methylation. Indeed, only 7% of the hypermethylated sites in AML[TET2mut] overlapped with CTCF binding. This suggests that hypomethylation in AML[TET2mut] occurs mainly outside of CTCF-binding sites and thus does not contribute to the changes in CTCF binding. Still, the role of hydroxymethylation on CTCF binding needs to be further investigated.

CTCF regulates gene expression through multiple mechanisms[2,3,5] and we show that a change in CTCF occupancy can be associated to both up- and downregulation of related genes. However, gain of CTCF binding was more commonly associated with gene upregulation and loss of CTCF binding with downregulation. Still, the outcome on gene expression of changes in CTCF binding was not random. Gains in promoter regions were more likely related to upregulation of genes, whereas downregulation was more common at genomic regions distal to promoters. One could speculate whether this relates to CTCF as an TF and an activator of gene expression at promoters and as an insulator located at more promoter-distant regions.[45] In line with this, depletion of CTCF has previously been shown to decrease gene expression near

transcription start sites and induce gene expression more distal to transcription start sites, especially when occurring at a TAD boundary separating an enhancer from a promoter.[8] Similar findings with respect to CTCF binding and gene expression were recently found in embryonic stem cells in relation also to DNA methylation.[68]

Changes in gene expression were also associated with binding motifs of DBCs, where gene upregulation was associated with motifs for activating cofactors and downregulation for motifs of repressing cofactors. Thus, the genomic location where aberrant CTCF binding occurs and what motifs that are targeted emerge as key factors of how it affects gene expression. When additionally integrating DNA methylation, the relation between CTCF binding and gene expression was even more unambiguous, where loss of methylation was more clearly associated with gain of CTCF binding and increased gene expression. All of these findings demonstrate the complex and multifaceted role of CTCF in gene regulation and leukemia.

Exposure of azacitidine in AML patient cells exerted profound effects on CTCF binding and significantly changed CTCF occupancy at almost 15 000 sites. Again, binding motifs for DBCs were enriched for activators when genes were upregulated and for repressors when genes were downregulated by azacitidine. Genes that changed both CTCF binding and gene expression with azacitidine were enriched for genes regulating cell proliferation and apoptosis, indicating that changes in CTCF are implicated in the antileukemic effects of the drug. All genes that were found to be demethylated, and concomitantly increased CTCF binding and gene expression by azacitidine treatment, are genes known to be implicated in hematological cancers. Of special interest is *KLF6*, which is epigenetically dysregulated in AML as well as during aging. Downregulation of KLF6 in CD34[+] cells is enough to change the phenotype and induce a AML like gene expression pattern.[63,64] Also, among CBSs gained with azacitidine, more than half were bound in NBM, suggesting that azacitidine restores a normal CTCF-binding pattern.

In summary, we show for the first time aberrant occupancy of the key transcriptional regulator CTCF in primary AML cells with a general gain of CTCF binding enriched for enhancers. AML with *TET2* mutations stands out as a subtype of AML with particularly high CTCF occupancy, which dominantly occurs in promoters. The underlying mechanism for this prominent gain of CTCF occupancy is unclear and needs further investigation. Clearly, the aberrant CTCF binding is associated with changed gene expression patterns, specifically with gene upregulation in sites that gain CTCF binding near promoters in AML. How this relates to detailed changes in chromatin structure, loop formation, and connections between promoters and enhancers needs to be further studied using methods deciphering DNA interactions, such as Hi-C. In conclusion, our results suggest that CTCF could be involved in the leukemogenesis in AML. However, the role of CTCF for initiating and/or driving leukemia needs further elucidation.

Acknowledgments

MUJAHED et al

Authorship

Contribution: H.M., A.L., and S.L. designed the research; H.M., S.M., A.N., S.B., C.N., S.D., L.C., K.E., A.L., and S.L. performed the research; S.B., C.B., S.D., L.C., and S.L. contributed vital material and collected data; H.M., S.M., A.N., A.L., and S.L. analyzed and interpreted data; and H.M., A.L., and S.L. wrote the paper

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: H.M., 0000-0001-9057-8604; A.N., 0000-0002-1109-1660; C.N., 0000-0003-0695-0050.

Correspondence: Sören Lehmann, Department of Medical Sciences, Uppsala University, Uppsala University Hospital, Entrance 101, 1st Floor, 751 85 Uppsala, Sweden; e-mail: soren.lehmann@ki.se.

Footnotes

REFERENCES

1. Sun Y, Chen BR, Deshpande A. Epigenetic regulators in the development, maintenance, and therapeutic targeting of acute myeloid leukemia. *Front Oncol.* 2018;8:41.

2. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of chromosomal domains by loop extrusion. *Cell Rep.* 2016;15(9):2038-2049.

3. Lu Y, Shan G, Xue J, Chen C, Zhang C. Defining the multivalent functions of CTCF from chromatin state and three-dimensional chromatin interactions. *Nucleic Acids Res.* 2016;44(13):6200-6212.

4. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012;485(7398):376-380.

5. Kim TH, Abdullaev ZK, Smith AD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell.* 2007;128(6):1231-1245.

6. Guo Y, Xu Q, Canzio D, et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell.* 2015;162(4):900-910.

7. Rao SS, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping [published correction appears in *Cell.* 2015;162(3):687-688]. *Cell.* 2014;159(7):1665-1680.

8. Nora EP, Goloborodko A, Valton AL, et al. Targeted degradation of CTCF decouples Local insulation of chromosome domains from genomic compartmentalization. *Cell.* 2017;169(5):930-944.e922.

9. Engel N, West AG, Felsenfeld G, Bartolomei MS. Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. *Nat Genet.* 2004;36(8):883-888.

10. Wang H, Maurano MT, Qu H, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* 2012;22(9):1680-1688.

11. Kemp CJ, Moore JM, Moser R, et al. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.* 2014;7(4):1020-1029.

12. Torrano V, Chernukhin I, Docquier F, et al. CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *J Biol Chem.* 2005;280(30):28152-28161.

13. Splinter E, Heath H, Kooren J, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* 2006;20(17):2349-2354.

14. Kim TG, Kim S, Jung S, et al. CCCTC-binding factor is essential to the maintenance and quiescence of hematopoietic stem cells in mice. *Exp Mol Med.* 2017;49(8):e371.

15. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med.* 2016;374(23):2209-2221.

16. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell.* 2004;13(2):291-298.

17. Falini B, Mecucci C, Tiacci E, et al; GIMEMA Acute Leukemia Working Party. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med.* 2005;352(3):254-266.

18. Marcucci G, Yan P, Maharry K, et al. Epigenetics meets genetics in acute myeloid leukemia: clinical impact of a novel seven-gene score. *J Clin Oncol.* 2014;32(6):548-556.

19. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature.* 2010;466(7310):1129-1133.

20. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell.* 2010;18(6):553-567.

21. Marina RJ, Sturgill D, Bailly MA, et al. TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing. *EMBO J.* 2016;35(3):335-355.

22. Bejar R, Lord A, Stevenson K, et al. TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. *Blood.* 2014;124(17):2705-2712.

23. Itzykson R, Kosmider O, Cluzeau T, et al; Groupe Francophone des Myelodysplasies (GFM). Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. *Leukemia.* 2011;25(7):1147-1152.

24. Wang M, Lindberg J, Klevebring D, et al. Validation of risk stratification models in acute myeloid leukemia using sequencing-based molecular profiling. *Leukemia.* 2017;31(10):2029-2036.

25. Wang M, Lindberg J, Klevebring D, et al. Development and validation of a novel RNA sequencing-based prognostic score for acute myeloid leukemia. *J Natl Cancer Inst.* 2018;110(10):1094-1101.

26. Tian Y, Morris TJ, Webster AP, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics.* 2017;33(24):3982-3984.

27. Abdel-Wahab O, Mullally A, Hedvat C, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood.* 2009;114(1):144-147.

28. Fortin JP, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* 2014;15(12):503.

29. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-359.

30. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc.* 2012;7(9):1728-1740.

31. Andrews S. FastQC. http://wwwbioinformaticsbabrahamacuk/projects/fastqc/. 2010.

32. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature.* 2014;507(7493):455-461.

33. Lobry C, Oh P, Mansour MR, Look AT, Aifantis I. Notch signaling: switching an oncogene to a tumor suppressor. *Blood.* 2014;123(16):2451-2459.

34. Li C, Lan Y, Schwartz-Orbach L, et al. Overlapping requirements for Tet2 and Tet3 in normal development and

hematopoietic stem cell emergence. *Cell Rep.* 2015;12(7):1133-1143.

35. Hughes JM, Legnini I, Salvatori B, et al. C/EBPα-p30 protein induces expression of the oncogenic long non-coding RNA UCA1 in acute myeloid leukemia. *Oncotarget.* 2015; 6(21):18534-18544.

36. Sood R, Kamikubo Y, Liu P. Role of RUNX1 in hematological malignancies. *Blood.* 2017; 129(15):2070-2082.

37. Wang J, Sun D, Wang Y, et al. FOSL2 positively regulates TGF-β1 signalling in non-small cell lung cancer. *PLoS One.* 2014;9(11):e112150.

38. Bergmaier P, Weth O, Dienstbach S, et al. Choice of binding sites for CTCFL compared to CTCF is driven by chromatin and by sequence preference. *Nucleic Acids Res.* 2018; 46(14):7097-7107.

39. Schuettpelz LG, Gopalan PK, Giuste FO, Romine MP, van Os R, Link DC. Kruppel-like factor 7 overexpression suppresses hematopoietic stem and progenitor cell function. *Blood.* 2012;120(15): 2981-2989.

40. Lu JW, Hsieh MS, Hou HA, Chen CY, Tien HF, Lin LI. Overexpression of SOX4 correlates with poor prognosis of acute myeloid leukemia and is leukemogenic in zebrafish. *Blood Cancer J.* 2017;7(8):e593.

41. Loizou E, Banito A, Livshits G, et al. A gain-of-function p53 mutant oncogene promotes cell fate plasticity and myeloid leukemia through the pluripotency factor Foxh1. *Cancer Discov.* 2019;9(7):962-979.

42. Britschgi C, Jenal M, Rizzi M, et al. HIC1 tumour suppressor gene is suppressed in acute myeloid leukaemia and induced during granulocytic differentiation. *Br J Haematol.* 2008;141(2):179-187.

43. Tao H, Ma X, Su G, et al. BCL11A expression in acute myeloid leukemia. *Leuk Res.* 2016;41: 71-75.

44. Hu G, Cui K, Northrup D, et al. H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. *Cell Stem Cell.* 2013;12(2):180-192.

45. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet.* 2014;15(4): 234-246.

46. Vorvis C, Hatziapostolou M, Mahurkar-Joshi S, et al. Transcriptomic and CRISPR/Cas9 technologies reveal FOXA2 as a tumor suppressor gene in pancreatic cancer. *Am J Physiol Gastrointest Liver Physiol.* 2016;310(11): G1124-G1137.

47. Kuo YH, Zaidi SK, Gornostaeva S, Komori T, Stein GS, Castilla LH. Runx2 induces acute myeloid leukemia in cooperation with Cbfbeta-SMMHC in mice. *Blood.* 2009; 113(14):3323-3332.

48. Zhang Y, Yuan Y, Liang P, et al. OSR1 is a novel epigenetic silenced tumor suppressor regulating invasion and proliferation in renal cell carcinoma. *Oncotarget.* 2017;8(18): 30008-30018.

49. Di Zazzo E, De Rosa C, Abbondanza C, Moncharmont B. PRDM proteins: molecular mechanisms in signal transduction and transcriptional regulation. *Biology (Basel).* 2013; 2(1):107-141.

50. Green MR, Vicente-Dueñas C, Romero-Camarero I, et al. Transient expression of Bcl6 is sufficient for oncogenic function and induction of mature B-cell lymphoma. *Nat Commun.* 2014;5(1):3904.

51. Hurtz C, Hatzi K, Cerchietti L, et al. BCL6-mediated repression of p53 is critical for leukemia stem cell survival in chronic myeloid leukemia. *J Exp Med.* 2011;208(11): 2163-2174.

52. Chen L, Shioda T, Coser KR, Lynch MC, Yang C, Schmidt EV. Genome-wide analysis of YY2 versus YY1 target genes. *Nucleic Acids Res.* 2010;38(12):4011-4026.

53. Dhordain P, Lin RJ, Quief S, et al. The LAZ3(BCL-6) oncoprotein recruits a SMRT/mSIN3A/histone deacetylase containing complex to mediate transcriptional repression. *Nucleic Acids Res.* 1998;26(20): 4645-4651.

54. Lee SH, Cho YC, Jeong HM, et al. Yin-Yang 1 and Yin-Yang 2 exert opposing effects on the promoter activity of interleukin 4. *Arch Pharm Res.* 2016;39(4):547-554.

55. Wong M, Polly P, Liu T. The histone methyltransferase DOT1L: regulatory functions and a cancer therapy target. *Am J Cancer Res.* 2015; 5(9):2823-2837.

56. Maeda T, Hobbs RM, Merghoub T, et al. Role of the proto-oncogene Pokemon in cellular transformation and ARF repression. *Nature.* 2005;433(7023):278-285.

57. Hojo N, Tatsumi N, Moriguchi N, et al. A Zbtb7a proto-oncogene as a novel target for miR-125a. *Mol Carcinog.* 2016;55(12): 2001-2009.

58. Higuchi T, Nakayama T, Arao T, Nishio K, Yoshie O. SOX4 is a direct target gene of FRA-2 and induces expression of HDAC8 in adult T-cell leukemia/lymphoma. *Blood.* 2013; 121(18):3640-3649.

59. Yamazaki J, Taby R, Vasanthakumar A, et al. Effects of TET2 mutations on DNA methylation in chronic myelomonocytic leukemia. *Epigenetics.* 2012;7(2):201-207.

60. Yamazaki J, Taby R, Jelinek J, et al. Hypomethylation of TET2 target genes identifies a curable subset of acute myeloid leukemia. *J Natl Cancer Inst.* 2015;108(2): djv323.

61. Porpaczy E, Bilban M, Heinze G, et al. Gene expression signature of chronic lymphocytic leukaemia with trisomy 12. *Eur J Clin Invest.* 2009;39(7):568-575.

62. Yang W, Liu H, Duan B, et al. Three novel genetic variants in NRF2 signaling pathway genes are associated with pancreatic cancer risk. *Cancer Sci.* 2019;110(6): 2022-2032.

63. Matsumoto N, Kubo A, Liu H, et al. Developmental regulation of yolk sac hematopoiesis by Kruppel-like factor 6. *Blood.* 2006;107(4):1357-1365.

64. Adelman ER, Huang HT, Roisman A, et al. Aging human hematopoietic stem cells manifest profound epigenetic reprogramming of enhancers that may predispose to leukemia. *Cancer Discov.* 2019;9(8): 1080-1101.

65. Rönnerblad M, Andersson R, Olofsson T, et al; FANTOM consortium. Analysis of the DNA methylome and transcriptome in granulopoiesis reveals timed changes and dynamic enhancer methylation. *Blood.* 2014;123(17): e79-e89.

66. Qu Y, Siggens L, Cordeddu L, et al. Cancer-specific changes in DNA methylation reveal aberrant silencing and activation of enhancers in leukemia. *Blood.* 2017;129(7): e13-e25.

67. Stadler MB, Murr R, Burger L, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature.* 2011; 480(7378):490-495.

68. Wiehle L, Thorn GJ, Raddatz G, et al. DNA (de) methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. *Genome Res.* 2019;29(5):750-761.

69. Rasmussen KD, Jia G, Johansen JV, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev.* 2015;29(9):910-922.

70. Qu Y, Lennartsson A, Gaidzik VI, et al. Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics.* 2014;9(8):1108-1119.