Targeting the 3D genome by anthracyclines for chemotherapeutic effects

Minkang Tan[1,6], Shengnan Sun[1,6], Yuchen Liu[5,6], Andrea A. Perreault[2,3], Douglas H. Phanstiel[3,4], Liping Dou[5]*, Baoxu Pang[1,7]*

[1]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands

[2]Department of Biology, Elon University, Elon, NC, USA

[3]Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

[4]Thurston Arthritis Research Center, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

[5]Department of Hematology, Chinese PLA General Hospital, Beijing, China

[6]These authors contributed equally

[7]Lead Contact

*Correspondence: b.pang@lumc.nl (B.P.), lipingruirui@163.com (L.D.)

Abstract

The chromatins are folded into three-dimensional (3D) structures inside cells, which coordinates the regulation of gene transcription by the non-coding regulatory elements. Aberrant chromatin 3D folding has been shown in many diseases, such as acute myeloid leukemia (AML), and may contribute to tumorigenesis. The anthracycline topoisomerase II inhibitors can induce histone eviction and DNA damage. We performed genome-wide high-resolution mapping of the chemotherapeutic effects of various clinically used anthracycline drugs. ATAC-seq was used to profile the histone eviction effects of different anthracyclines. TOP2A ChIP-seq was used to profile the potential DNA damage regions. Integrated analyses show that different anthracyclines have distinct target selectivity on epigenomic regions, based on their respective ATAC-seq and ChIP-seq profiles. We identified the underlying molecular mechanism that unique anthracycline variants selectively target chromatin looping anchors via disrupting CTCF binding, suggesting an additional potential therapeutic effect on the 3D genome. We further performed Hi-C experiments, and data from K562 cells treated with the selective anthracycline drugs indicate that the 3D chromatin organization is disrupted. Furthermore, AML patients receiving anthracycline drugs showed altered chromatin structures around potential looping anchors, which linked to distinct clinical outcomes. Our data indicate that anthracyclines are potent and selective epigenomic targeting drugs and can target the 3D genome for anticancer therapy, which could be used for personalized medicine to treat tumors with aberrant 3D chromatin structures.

Introduction

The human genome is packaged into a highly organized structure inside the nucleus, which forms a hierarchical epigenetic regulation layer to guide biology processes like cell development and differentiation[1]. Emerging techniques such as ChIA-PET, 3C, 4C and Hi-C[2], which assay the genome-wide contact frequency of genomic loci, have enabled studying nuclear organization at fine scale. Analysis of high-resolution map of the spatial genome structure revealed the three-dimensional (3D) genome architecture can be divided into chromosome territories[3], compartments A/B representing transcriptionally active/inactive chromatin[2], topologically associating domains (TADs)[4], architectural stripes[5] and chromatin loops[6]. Architectural proteins like cohesin and CCCTC-binding factor (CTCF) play important roles in mediating the 3D genome foldings[7]. The different layers within the 3D genome structure facilitate the complex regulation of gene transcription within cells.

Accumulating evidence shows that cancer cells have aberrant 3D genome structure alterations via the accumulation of genetic mutations or other mechanism. The aberrant chromatin interactions in cancer cells hijack the non-coding regulatory elements such as silencers[8] and enhancers for gene deregulation to promote tumor development and progression[9,10]. This pathogenic mechanism of dysregulation of cellular function through chromatin interactions also implies the translational and clinical importance to understand the 3D structure in the cancer genome, and to explore the 3D genome as the potential new targets for clinical intervention. Recently, several studies reported that small-molecule drugs targeting epigenetic factors will change cancer genome organization with varying effects[9,11-13]. These epigenetic drugs function either by inhibiting targeted transcriptional activators to selectively block the activity of enhancers[9,12,13] or by dissolving the liquid-liquid phase separation condensates to hinder CTCF-independent loop formation[11]. However, no clinically used therapeutics have been shown to directly affect the 3D genome.

Anthracyclines are a class of conventional anti-cancer chemotherapeutic agents which have been widely used in oncology for many decades because of their broad spectrum and high efficacy. They are still the standard practice in cancer treatment, especially for aggressive tumors like acute myeloid leukemia (AML). Daunorubicin

(Daun) and doxorubicin (Doxo) are the first two anthracyclines used in the clinics. Due to their strong side effects, many anthracycline variants were developed with less toxicity such as aclarubicin (Acla), amrubicin (Amr), epirubicin (Epi) and idarubicin (Ida). These drugs intercalate into the DNA and most of them induce DNA double-strand breaks (DSBs) via interference with the catalytic cycle of topoisomerase IIα (TOP2A)[14]. Recently, a novel mechanism of action of anthracyclines was discovered to induce tumor cell death via histone eviction from chromatin[15]. Of note, the anthracycline variant Acla, which is used in the clinic to treat AML, was shown to only evict histone without damaging the DNA[15,16]. And different anthracycline variants may also have selectivity in targeting distinct genomic regions for their therapeutic effects[15,17].

In this study, we show that anthracycline drugs, as exemplified by Acla and Daun, redistributed their canonical target protein TOP2A around chromatin looping anchors, as shown by ChIP-seq experiments against endogenously tagged TOP2A in human erythroleukemia K562 cells. ATAC-seq analysis revealed the loose nucleosome compaction and decreased CTCF footprints signal at CTCF-bound loop anchor regions, which is associated with the formation of insulative chromatin domain boundaries[18]. Such loss of CTCF chromatin binding was further confirmed by western blotting and ChIP-seq against CTCF, indicating that the 3D chromatin structure may be affected by these drugs. Indeed, in situ Hi-C experiments revealed massive spatial genome alterations following Acla and Daun treatment in K562 cells, including weakened insulation of TAD boundaries and disruption of looping structures. The integrated analysis combining in situ Hi-C, ATAC-seq, and ChIP-seq suggested that the decreased CTCF abundance is the major factor driving the decoupling of chromatin loops. To test if anthracycline drugs can also induce 3D chromatin structure changes in patients, we analyzed chromatin accessibility changes in tumor cells isolated from AML patientss after Daun treatment using ATAC-seq, where CTCF-bound regulatory elements were also affected. We also identified potential 3D chromatin changes in AML patient tumor cells that were associated with the clinical outcomes of patients receiving anthracycline-based chemotherapy. These findings indicate that anthracycline has the potential to target the 3D chromatin structure for their therapeutic effects, and such property could be exploited to treat other cancers with aberrant 3D chromatin changes.

Results

Anthracyclines redistribute TOP2A at chromatin looping anchors

Type II topoisomerases α (TOP2A) enzyme tend to overexpress in tumor cells, which make it a good target for anti-cancer treatment. Most anthracyclines function by inhibiting TOP2A to exert anti-cancer activity[14]. TOP2A also plays important roles in many biological processes, e.g. DNA replication, DNA repair, gene transcription and genome organization[19]. To directly monitor the effects of anthracyclines on the interplay between TOP2 and chromatin, we performed genome-wide mapping of TOP2A binding sites using chromatin immunoprecipitation followed by sequencing (ChIP-seq) in K562 cells with TOP2A endogenously tagged with FLAG tag (**Methods**). It has been shown that promoter-associated TOP2A plays a vital role to resolve the DNA torsional stress arising by transcription processes, i.e. DNA supercoiling[19,20], and indeed, we observed that TOP2A accumulated at the promoter regions, and were redistributed upon Acla and Daun treatments (**Supplementary Fig. 1a,b**). TOP2A shares potential functional redundancy as TOP2B[21], which was shown to interact with CTCF and cohesin at TAD boundaries[19] to resolve the topological obstacles and DNA torsional stress induced by loop extrusion[22,23]. We then checked the enrichment of TOP2A around TADs and observed a notable absence of TOP2A binding at TAD borders bound by CTCF and cohesin after Acla and Daun treatments while not after Amr treatment, which only induce DNA damage without histone eviction (**Fig. 1a**). When all the chromatin looping anchors bound by CTCF and cohesin were further examined, a reduction of TOP2A ChIP-seq signal was observed (**Fig. 1b**). To further characterize the genomic redistribution pattern of TOP2A quantitively, we performed differential TOP2A binding analysis (**Methods**), and observed that Acla and Daun induced massive perturbation on TOP2A binding on chromatin genome-wide, whilst Amr did not show notable effect compared to the untreated cells (**Supplementary Fig. 1c**). Permutation-based association test showed that the stable (TOP2 peaks which were not affected by the respective drug treatments) and lost TOP2A peaks from Acla and Daun treatments significantly associated with CTCF binding sites while the gained TOP2A peaks did not show association with CTCF peaks (**Fig. 1c**). Motif enrichment analysis also showed that the lost TOP2A binding significantly enriched for the motifs such as CTCF, CTCFL

and GATA2 (**Fig. 1d**), which further suggests that the Acla- or Daun-induced redistribution of TOP2A may affect TAD borders and looping anchors marked by CTCF. Interestingly, such potential effects are shared by other clinically-administered anthracyclines that are capable to induce histone evictions (**Supplementary Fig. 1d**)[16]. These data suggest that most anthracycline drugs may redistribute TOP2A from TAD borders and looping anchor regions associated with CTCF.

Anthracyclines treatments affect CTCF foot-printing on chromatin

Recent study suggested anthracyclines present a selective property of evicting histone from chromatin which is dependent on the local chromatin structure[17]. It is possible that anthracycline-mediated TOP2A depletion from chromatin looping regions as marked by CTCF binding may coincide with the disruption of CTCF binding directly. ATAC-seq was then used to profile the chromatin structures to infer the changes of genome-wide transcription factor binding upon anthracycline treatment. We exposed K562 cells to Acla, Amr or Daun for 2 hours and then performed ATAC-seq. The distribution of ATAC-seq fragment size from cells treated with anthracyclines that can induce histone eviction (i.e. Acla and Daun) showed a distinct distribution pattern compared with cells treated with Amr (induces DNA breaks only) and untreated cells (**Fig. 2a**), confirming that the periodicity of compact nucleosome structures was affected by histone eviction[24]. In addition, the enrichments of Tn5 insertions in seven classes of functional chromatin states previously defined[25] were also substantially changed in the chromatin state-specific way (**Fig. 2b**). For instances, CTCF-bound regions which are preferably enriched for shorter fragments and depleted for longer fragments for the superiority of CTCF to organize nucleosomes in its vicinity[26,27]. Acla and Daun treatments change the local chromatin packed by nucleosomes around CTCF-bound regions (**Fig. 2b**). We also performed the nucleosome positioning analysis around TSSs from ATAC-seq data and showed the disruption pattern of the nucleosome-free regions (NFRs) at TSSs which are flanked by two well-positioned nucleosome arrays (**Supplementary Fig. 2e**) as expected[28], supporting the global chromatin changes induced by anthracyclines. The orthogonal analysis on the 15 chromatin states inferred by chromHMM[29] show similar conclusion in the insulator elements (**Supplementary Fig. 2a-d**). To test the chromatin structure changes around looping regions, we further performed nucleosome occupancy mapping from the ATAC-seq data using

NucleoATAC[30] in CTCF-bound chromatin looping anchor regions (**Fig. 2c**) and the data showed substantial decrease of nucleosome occupancy around the CTCF-bound chromatin looping anchor regions (**Fig. 2c**). Such structure changes might be accompanied with the decreasing insulation strength of chromatin domain boundaries[18]. To unbiasedly analyze the TF binding changes, we further used the ATAC-seq data to map Tn5 cleavage frequencies around the CTCF motif genome-wide through TF footprint analysis by HINT-ATAC[31]. Interestingly, the aggregated footprints in CTCF motif suggest Acla and Daun may weaken the CTCF occupancy on the chromatin by altering the local chromatin structure[27] (**Fig. 2d**). Together, these data indicate the chromatin structure around looping anchors and CTCF binding may be affected by Acla and Daun treatments.

Anthracyclines evict CTCF from chromatin

To directly test if CTCF binding is affected by anthracyclines, K562 cells were treated with Daun for 2 hours, and then the cytosolic and chromatin fractions of the cells were separated and probed for CTCF. The results showed that CTCF in the chromatin fraction decreased accompanied with the increase of CTCF in the cytoplasmic fraction, compared to the untreated cells (**Fig. 3a,b**). This data suggested anthracyclines as exemplified by Daun can partially evict the endogenous CTCF from the chromatin resulting in the elevated level of chromatin-unbound CTCF in the cytoplasm. To pinpoint the disruption of CTCF binding in a genome-wide fashion, we profiled the CTCF binding on the chromatin by ChIP-seq, and observed massive loss of CTCF from the chromatin upon Acla or Daun treatments (**Fig. 3c**). Furthermore, we also observed some new CTCF binding sites on the chromatin after drug treatments (**Fig. 3d**), which may also facilitate the chromatin loops reorganization. CTCF is one of the key factors for TAD boundaries and chromatin looping[32,33], which shapes the architecture of gene regulatory landscape. We observed that CTCF ChIP-seq signals were generally decreased in loop anchors and TAD boundaries upon Acla and Daun treatment (**Fig. 3e, upper panels**). Furthermore, promoter-proximal CTCF binding usually defines the engagement specificity of enhancer-promoter interactions[34,35], and a reduction of CTCF ChIP-seq signals at the transcription start sites (TSSs) was also observed upon Acla and Daun treatment (**Fig. 3e, lower panels**). The data indicates the anthracyclines may have the potential to interfere with the cis-regulatory interactions landscape by affecting

CTCF binding. Because TOP2A was shown to be redistributed from chromatin looping regions and TAD boundaries which are marked by CTCF (**Fig. 1**), we further zoomed in the regions where both TOP2A and CTCF colocalize. We also investigated the colocalization of TOP2A and CTCF[22,36] response to the two drugs by combining the TOP2A and CTCF ChIP-seq data. Our analysis showed that both TOP2A and CTCF were depleted from these regions (**Fig. 3f**). Thus, our data suggests the CTCF binding on chromatin was disrupted by the anthracyclines which may lead to 3D genome reorganization.

Anthracyclines target distinct genomic patterns of CTCF binding sites

To characterize the differential responses of chromatin-binding CTCF to anthracyclines therapies, we performed differential binding analysis of CTCF ChIP-seq data (**Methods**). We observed that the lost CTCF peaks upon drug treatment were more present in distal intergenic or intron regions, whilst the newly gained CTCF peaks and stable CTCF peaks were more enriched at promoter regions (**Fig. 4a**). These data indicate that distal looping regions may be preferably affected by these drugs. We scanned for conserved binding motifs within the peak center (200 bp window) of differential CTCF peaks using FIMO[37] and found that the stable CTCF peaks show significantly higher match scores (**Fig. 4b**) and generally more motif occurrences (**Supplementary Fig. 3a-d**). In addition, we found the second bit guanine base within the CTCF core motif might contribute to the sensitivity of CTCF to Daun perturbation (**Fig. 4c**) while not in Acla perturbation (**Supplementary Fig. 3e**). The CTCF binding affinity was shown to be associated with the methylation levels of two cytosines (C2: position 2 and C12: position 12) within the 15-bp consensus core motif as indicated (**Fig. 4c**) [38]. We observed the CpG methylation level of C2 is significantly higher within the stable CTCF peaks than the lost CTCF peaks after Acla treatment, while at C12 there was no difference (**Fig. 4d**). It has been shown before that methylation level at C2 is determent for the binding affinity of CTCF[38]. The data suggest that the stable binding of CTCF during Acla treatment may indicate a high binding affinity on chromatin and reduced sensitivity to DNA methylation at the C2 CpG site. Interestingly, the CpG methylation levels of C2 and C12 are both significantly higher within the gained CTCF peaks than the lost or stable peaks after both Acla and Daun treatments (**Fig. 4d, Supplementary Fig. 3f**), showing the potential of anthracycline drugs may redistribute CTCF to its unspecific

recognition sites. As we showed before, the anthracycline drugs with histone eviction property disrupt TOP2A occupancy and nucleosome package surrounding the CTCF binding sites (**Fig. 1,2**). Using the micrococcal nuclease digestion with deep sequencing (MNase-seq) and TOP2A ChIP-seq data generated in K562 cells, we profiled the nucleosome occupancy (**Fig. 4e, Supplementary Fig. 3g**) and TOP2A binding affinity (**Fig. 4f, Supplementary Fig. 3h**) around differentially affected CTCF binding sites upon Duan and Acla treatments. Our analysis indicates that the stable CTCF binding sites correlate with high-abundant nucleosome protection and stronger affinity of TOP2A binding, which suggests that the local chromatin environment also affect the ability of anthracyclines to disrupt CTCF binding.

Anthracyclines impact 3D genome organization

As shown so far, anthracyclines exemplified by Acla and Duan may affect the chromatin structure around chromatin looping regions and TAD boundaries by redistributing TOP2A and depleting CTCF, it is possible that the 3D genome may be affected by these drugs. To test this hypothesis, K562 cells treated with Acla or Daun were processed for *in situ* Hi-C to assay the effects on chromatin folding structures (**Fig. 5a**). We observed a notable increase of *trans to cis* contacts ratio after Acla and Daun treatments (**Supplementary Fig. 4a**) based on high quality Hi-C datasets (**Supplementary Table 2**). When the intrachromosomal interactions were considered, we observed that the anthracycline treatments resulted in steep decays of interactions between two loci separated by less than 10 Kbp, and more interactions between loci separated by more than 1 Mbp (**Fig. 5b**). As the 3D folding of the chromatin can form distinct functional compartments such as A and B compartments, which are related to transcriptional activities[2], to assess whether the compartmentalization status of chromatin will change, we applied eigenvector decomposition analysis to determine the positions of A or B compartments. We observed 24.8% genomic bins switched the compartment status during Acla treatment (i.e. 12.3% A converted to B, and 12.5% B converted to A). While in Daun-treated cells, there were over 54.5% compartment status switched (i.e. 27.2% A to B, and 27.3% B to A) (**Fig. 5c**). However, it is unclear if the compartment switching was reflecting the transcriptional changes or just the results of the disrupted 3D folding. Such compartmentalization changes also aligned with the reduced level of

compartment separation by saddle plot analysis (**Fig. 5d**), and weaker TAD boundaries as measured by the diamond insulation approach[39] to detect boundaries in individual samples (**Fig. 5e**). To detect the differential insulated TAD boundaries, we first assigned the insulation score to each 5 Kbp bin and categorized the differential boundaries based on their deviations from untreated cells[40]. The analysis showed that most of TAD boundaries were lost by anthracycline treatments (i.e. 83.6% and 88.2% boundaries were lost in Acla and Daun treatment respectively) (**Fig. 5f**), which is correlated with the observation of CTCF depletion from the chromatin boundaries (**Fig. 3**). To investigate the effects on the residual TAD boundaries, we performed the correlation plot analysis of boundary strength scores (**Fig. 5g, Supplementary Fig. 4b**) and insulation scores (**Supplementary Fig. 4c,d**), and again the results indicate dampened insulation strength at residual TAD boundaries. Interestingly, we also observed a few TAD boundaries with stronger insulation scores, but no newly gained TAD boundaries during drug treatment (**Supplementary Fig. 4e,f**), indicating the redistributed CTCF (**Fig. 3**) did not establish new TAD boundaries. In general, our data aligned with the observation from targeted degradation of endogenous CTCF in mouse embryonic stem cells[41] and highlighted that the anthracyclines may perturb the 3D genome structure by evicting CTCF from its native chromatin context.

Anthracyclines affect CTCF-mediated *cis*-regulatory loops

To investigate whether chromatin looping structures within TADs were altered by anthracyclines, we further analyzed Hi-C data from anthracyclines-treated K562 cells using a random forest-based classifier Peakachu[42] which considered the sample-wise sequencing depth variations. A Gaussian mixture model was adapted to determine the significantly differential loops upon drug treatments, and visualized by aggregate peak analysis (APA) (**Fig. 6a, Supplementary Fig. 5a**)[10]. The analysis shows that these two drugs disrupted many native chromatin loops (24,460 and 28,025 lost loops in Acla- and Daun-treated samples, FDR < 0.05), whilst some new loops were also made (6,328 and 4,800 gained loops in Acla- and Daun-treated samples, FDR < 0.05). Again, within these lost loops, CTCF was also identified as one of the most associated factors, when integrative analysis was performed using ATAC-seq and ChIP-seq data (**Fig. 6b, Supplementary Fig. 5b**). The massive loss of CTCF peaks within the lost looping anchors was accompanied by loss of ATAC-

seq peaks with CTCF motifs. Taken together, our data suggest that the loss of looping structures was due to the depletion of CTCF from chromatin induced by anthracyclines (**Fig. 3**) as CTCF facilitates and stabilizes the looping structures[41]. On the contrary, the distances among newly formed loops upon Daun and Acla treatments were significantly larger than those of lost or stable loops (**Supplementary Fig. 5c,d**). To understand the potential perturbation effects on gene transcription by the anthracycline-disrupted chromatin looping structures, we performed KEGG pathway analysis on genes with their TSSs in contact with the lost looping anchors upon drug treatments (**Supplementary Fig. 5e,f**). Biological processes such as cell cycle, Wnt signaling and Hippo signaling pathways were affected by anthracycline-induced loop disruptions. Among the lost loops, we also found that the well-established promoter-enhancer[43] and enhancer-enhancer[44] contacts mediated by CTCF to regulate the *Myc* gene over-expression in K562 cells were also affected by Daun and Acla treatments (**Fig. 6c**). The differential loops and virtual 4C analyses anchored at *Myc* promoter showed that the interactions between *Myc* promoter and *Myc* enhancers were disrupted (**Fig. 6d**), which was also supported by the depletion of CTCF binding based on the ChIP-seq data. Interestingly, we also observed the disruption of interactions among *Myc* enhancers (**Fig. 6c**), which confer the enhancer network synergy and robustness regulation on *Myc* gene expression[44]. These data indicate that overexpression of oncogene *Myc* driven by super-enhancers that is common in cancer could be targeted by anthracycline-mediated disruption of long-range interactions and 3D chromatin folding. To extend this observation further, we further examined the effects induced by anthracyclines such as Daun and Acla on most of the potential *cis*-regulatory looping interactions, i.e. CTCF-CTCF, promoter-promoter, promoter-enhancer, enhancer-enhancer and putative silencer-promoter loops[45] by APA analysis. The results showed that the anthracyclines such as Daun and Acla can disrupt most of the interactions among cis-regulatory elements in a genome-wide fashion (**Fig. 6e, Supplementary Fig. 5g**). Together, our data suggests anthracyclines have the potential to be an effective therapy by disrupting the aberrant *cis*-regulatory chromatin looping structures mediated by CTCF to drive oncogene activation or tumor suppressor gene repression, which are key pathologic events in cancer[46].

Daun-induced 3D genome alterations correlate with therapeutic response

Anthracycline-based regimen remains the first-line treatment option for AML for its efficiency[16]. However, therapeutic response to anthracycline-based chemotherapies varies in cancer patients due to multiple factors[47,48]. To assess whether the anthracycline-induced chromatin looping and 3D genome structure changes contribute to the therapeutic response, we performed ATAC-seq on blasts isolated from AML patients, and compared the response to Daun. The CD33+ or CD34+ AML blast cells were isolated from peripheral blood or bone marrows of 6 complete remission (CR) and 6 refractory/relapsed (R/R) AML patients treated with the DDAG (i.e. Daunorubicin+Decitabine+Ara-C+G-CSF) or other anthracycline-based regimen using fluorescence-activated cell sorting (FACS) (**Supplementary Table 3**). The samples were first grouped based on the patients' response to induction therapy. This grouping allows for a comparative analysis of chromatin accessibility changes induced by Daun from the responders (CR) and non-responders (R/R). The distal binarization approach was performed[49], and 14,260 group-specific peaks were identified (**Supplementary Table 4**). Among the group-specific peaks, 3,357 regions were closed upon Daun treatment either in the R/R patients or in the CR patients (**Fig. 7a, Supplementary Fig. 6a**). Some of these closed regions may be because of the loss of TF binding and disruption of the 3D genome, similar to what we observed in the cell lines. To extract such information from CR or R/R patients, we performed the motif enrichment analysis in group-specific ATAC-seq peaks. The analysis indeed supported that CTCF is one of the major significantly affected TFs by Daun treatment (**Fig. 7b**). In addition, comparing the differences between the CR patients and the R/R patients, the CTCF motif is significantly enriched in the uniquely closed ATAC-seq regions from CR patient when treated with Daun (**Fig. 7b**), indicating additional chromatin loopings and 3D genome structures were targeted by Daun in CR patients. To reveal the transcriptome alterations by Daun treatment, we also performed RNA-seq analysis on the same AML samples. First, the change of expression of CTCF did not differ between CR and R/R groups (**Supplementary Fig. 7b**). When the gene pathway enrichment analysis was performed, as expected, the p53 signaling pathway was activated after Daun treatment in both CR and R/R patient groups in response to the DNA damage induced (**Supplementary Fig. 7c,d**). Patient-group-specific enriched pathways were also identified (**Supplementary Fig. 7e,f**), some of which may be the result of the 3D genome deregulation. To infer the potential transcriptional regulation effects of the Daun-induced 3D genome structure

changes in the patients, we performed the peak-gene correlation analysis between the ATAC-seq accessibility of peaks and expression of their potential gene targets, which could partially recapitulate the physical interactions identified from 3D chromosome conformation capture assay (i.e. in situ Hi-C or HiChIP) based on the bulk or single-cell ATAC-seq data[49,50]. Based on the peak-to-gene linkage analysis[49] using 20 paired RNA-seq and ATAC-seq samples, we identified 32,069 potential interactions between ATAC peaks and their target genes, where 4,298 peak-to-gene links were mediated by CTCF (i.e. either containing the CTCF motif or overlapping with CTCF ChIA-PET loops) (**Supplementary Table 5**). In addition to identify peak-to-gene links across all the samples, we can also further stratify the peak-to-gene links that uniquely been affected upon Daun treatment in either the R/R patients or in the CR patients. By overlapping the peak-to-gene links with the uniquely closed peak regions after Daun treatment in either CR or R/R patients (**Fig. 7a**), 139 and 211 peak-to-gene links composed by CTCF-bound regions were identified to be associated with the either CR or R/R patients response to Daun (**Supplementary Table 6**). To verify the regulatory interaction functions for the predicted peak-to-gene links, we selected six peaks that are potential CTCF-bound looping regions (**Supplementary Fig. 7a-f**) and predicted to regulate genes *DNM1*, *DNMT1*, *BAG2*, *RPS6KA1, METTL4* and *TRIM32* respectively (**Fig. 7c-e, Supplementary Fig. 7g-k**), and performed the CRISPR interference (CRISPRi) to target the peak-to-gene links by displacing CTCF from the targeted peak regions with dCas9-KRAB[51]. The two peaks linked to genes *DNM1* and *DNMT1* were specifically targeted in the R/R patients group (**Supplementary Fig. 7g,h**) and the four peaks linked to genes *TRIM32*, *BAG2*, *METTL4 and RPS6KA1* were specifically targeted in the CR patients group (**Fig. 7c-e, Supplementary Fig. 7i-k**) based on the previous analysis (**Fig. 7a**). For the peak-to-gene linkages specifically affected by Daun treatment in the R/R patient group (**Supplementary Fig. 7g,h**), we found that targeting the predicted CTCF-mediated peak-to-gene linkage via two different CRISPRi guide RNAs led to a significant reduction of *DNM1* gene expression in K562 cells (**Supplementary Fig. 7l**). The gene *DNM1* encodes the dynamin-1 from the dynamin subfamily of GTP-binding proteins and is involved in the endocytosis process. Downregulation of *DNM1* is associated with the inhibition of niche signals transduction mediated by the endocytosis and will potentially reshape the microenvironment of relapsed leukemia stem cells[52]. For the peak-to-gene linkages

specifically affected by Daun treatment in the CR patient group (**Fig. 7c-e, Supplementary Fig. 7i-k**), CRISPRi targeting the CTCF-mediated peak-to-gene linkage led to a significant reduction of *TRIM32* gene expression in K562 cells (**Fig. 7f**). The gene *TRIM32* encodes the E3 ubiquitin ligase and promotes p53 degradation through ubiquitination. Downregulation of *TRIM32* may enhance the p53-mediated cellular stress and thus promote p53-mediated apoptosis, cell cycle arrest and senescence[53]. The effects by Daun via disrupting the CTCF-mediated transcriptional regulatory loop of gene *TRIM32* may promote the sensitivity of CR patients to Daun treatment. Similar results were also shown for the gene *RPS6KA1* (**Supplementary Fig. 7f**). The gene *RPS6KA1* encodes a member of the ribosomal S6 kinase family of serine/threonine kinases and genome-wide dropout experiments[54,55] suggested that *RPS6KA1* is an AML-specific essential gene and inhibition of *RPS6KA1* prevents resistance and increases sensitivity of AML to chemotherapies[56]. The additional targeting of *RPS6KA1* by Daun via interrupting the 3D genome may be one of the reasons RR patients were more sensitive to the anthracycline-containing regimen (**Fig. 7d,e**). Similarly, targeting the predicted CTCF-mediated long-range interactions via one CRISPRi guide RNA also reduced the transcription of genes *BAG2* and *METTL4* which may exert additional effects on the CR patients to Daun-based chemotherapy[57,58](**Supplementary Fig. 7l**). These results indicate that the identified potential peak-to-gene may mediate long rang interactions in patients, and Daun-induced 3D genome structure changes may target these interactions to affect gene transcription and exert therapeutic effects. Together, we identified a set of distal regulatory regions in the noncoding genome which were targeted by Daun, such as CTCF-binding looping regions, using ATAC-seq in the primary blasts from AML patients. Our analysis also indicated the chromatin accessibility changes which were associated with CTCF binding induced by Daun treatment can be linked to the potential perturbation effects on the 3D regulatory interactions between susceptible genes and noncoding DNA elements. Overall, our data suggested anthracycline drugs as exemplified by Daun can target the 3D genome, which may contribute to the therapeutic effects of these chemotherapeutics for individual patient's response.

Discussion

Emerging evidence indicates that aberrant chromatin 3D structures in the human genome contribute to tumor development and malignancy[9,10]. Targeting the 3D genome of cancer cells is a promising strategy for anti-cancer drug discovery and development. Currently, there are not many options to directly target the 3D genome to treat cancer. We identified that anthracycline drugs can redistribute their canonical target TOP2A away from the chromatin looping regions. Our analysis of TOP2A ChIP-seq and ATAC-seq in K562 cells treated with Daun and Acla indicated that CTCF binding may be affected by these drugs. Indeed, both biochemical and ChIP-seq experiments against CTCF confirmed that the chromatin binding of CTCF was interrupted by Daun and Acla. As CTCF is one of the major factors involved in chromatin looping and long-range chromatin interactions, it is possible that the 3D genome is affected by these drugs. Indeed, direct profiling the 3D genome using Hi-C after cells were treated Daun and Acla showed that these anthracycline drugs induced massive global long-range interaction changes. Furthermore, there were also differences in targeting the 3D genome between Daun and Acla, suggesting the possibility of personalized application of anthracycline drugs in targeting patients specific 3D genome alterations in future.

To show that the 3D genome is possibly targeted in patients when treated with anthracycline drugs, blast cells isolated from patients were analyzed using ATAC-seq when treated with Daun. Similar to what was observed in K562 cells, many differentially affected the ATAC-seq regions bear the enrichment for CTCF motif, suggesting that distal interactions mediated by CTCF may also be targeted by anthracyclines exemplified by Daun in patients. Targeting these potential distal regulation regions marked by CTCF via CRISPRi affected key distal genes which may contribute to the therapeutic effects. Importantly, CTCF motifs were enriched in uniquely deregulated ATAC-seq regions in blast cells from CR patients treated with Daun (**Fig. 7a**), suggesting that affecting the 3D genome in the patients may contribute to the better therapeutic response in the CR patients.

The mechanism of action of anthracycline drugs is complex, including induction of DNA damage and histone eviction[15]. Here we also show that most anthracycline drugs can also disrupt 3D genome via targeting CTCF marked looping regions and TAD boundaries. Therefore, it is hard to distinguish the contribution of DNA damage,

histone eviction and 3D genome targeting to the final therapeutic effects. Nonetheless, it is possible that tumor cells relying on defined aberrant pathways for their transformation may show unique sensitivity to one of the mechanism of action of anthracycline drugs, as shown for AML patients where the 3D genome is hijacked for tumorigenesis[10]. Furthermore, different drug variants from the anthracycline family also had some distinct 3D genome targeting properties. Such targeting selectivity could be further explored in future for personalized medicine. ATAC-seq is a quick and reliable way to probe the anthracycline drug response for their 3D genome targeting effects, which may be used in the clinic to profile the patients and rational selection of the treatment options.

Our discovery on anthracycline drugs with 3D genome targeting property not only provides a new therapeutic option to target the aberrant 3D genome in the clinic, but also may stimulate the interests to identify more disease-driven aberrant 3D structures in cancer and possibly other diseases.

Methods

Primary sample collection

This clinical study was approved by the Human Subject Ethics Committee of Chinese PLA General Hospital. Leukemia blasts were obtained from the peripheral blood or bone marrow of 12 de novo or relapse AML patients with consent (**Supplementary Table 3**). The patients received an intravenous infusion of daunorubicin (60 mg/m^2) treatment within 3 hours. The AML blast cells were collected before treatment and 3 hours post-daunorubicin administration. The mononuclear cells were isolated from the peripheral blood or bone marrow of donors by density gradient centrifugation using Ficoll-Paque media (GE Healthcare). The CD33+ or CD34+ cells were enriched by magnetic bead (Miltenyi Biotec, Germany) selection.

Cell culture

K562 cells were cultured in RPMI 1640 + L-Glutamine (Gibco), 10% fetal bovine serum (Biowest), and 1% Penicillin-Streptomycin (Gibco). Cell density and culture conditions were maintained according to the ENCODE Cell Culture Guidelines. K562-TOP2A-3xFLAG tag cell line was generated as previously described[59].

Reagents

Aclarubicin was obtained from Santa Cruz Biotechnology (sc-200160). Daunorubicin was obtained from Sanofi-Aventis (the Netherlands). Amrubicin was obtained from Santa Cruz Biotechnology (sc-207289). All of these drugs were dissolved according to the manufacturer's formulation, aliquoted and stored at −20 °C for further use.

ATAC-seq

The Omni-ATAC[60] was performed as described previously. Briefly, the 50,000 K562 (wild type) cells, Amr-, Acla-, and Daun-treated cells were harvested and washed using 1 ml cold resuspension buffer (RSB, 10 mM pH 7.4 TrisCl +10 mM NaCl + 3 mM MgCl2 + ddH_2O). Nuclei were purified using a 50 µl cold ATAC lysis buffer (RSB + 0.1% v/v Tween + 0.1% v/v NP40 + 0.01% v/v Digitonin) incubated on ice for 3 mins and washed in RSB with inverting 3 times. Nuclei were centrifuged for 10 mins at 500 rcf, 4 °C, and then discard supernatant carefully to avoid the pellet. The

pelleted nuclei were transposed as described in the protocol, making the transposition reaction master mix as 25 µl 2x TD buffer, 2.5 µl transposase (100 nM final), 16.5 µl PBS, 0.5 µl 1% digitonin, 0.5 µl 10% Tween-20, 5 µl water. Then nuclei pellets were resuspended in a 50 µl transposition reaction mix by triturating 6 times and incubated at 37°C for 30 mins on a thermomixer with 1000 rpm. Transposed fragments were purified using Zymo DNA Clean and Concentrator-5 kit (D4014) and amplified using PCR reaction as 20 µl transposed DNA, 2.5 µl 25 µM PCR primer 1 (AD 1), 2.5 µl 25 µM Barcoded PCR primer 2 (AD 2.1 - AD 2.5), and 25 µl NEB Next High-Fidelity 2x PCR Master. The thermal cycle is as follows, 1 cycle: 5 min 72°C, 30 sec 98°C; 5 cycles: 10 sec 98°C, 30 sec 63°C, 1 min 72°C. The concentration of pre-amplified fragments was determined by quantitative PCR (qPCR) and this concentration was used to estimate the total number of cycles required to obtain fragments. QPCR reaction was set up as 5 µl DNA from the last PCR step, 1.6 µl nuclease-free H_2O,1.25 µl 5 µM PCR primer 1 (AD 1), 1.25 µl 5 µM Barcoded PCR primer 2 (AD2.1 - AD 2.5,), 0.9 µl 10x SYBR Green I, and 5 µl NEB Next High-Fidelity 2x PCR Master Mix. Using a qPCR instrument, cycle as follows: 1 cycle: 30 sec 98°C, 20 cycles: 10 sec 98°C, 30 sec 63°C, 1 min 72°C. To calculate the additional number of cycles needed, plot linear Rn vs. Cycle, calculate the number of cycles, and determine the number corresponding to one-third of the maximum fluorescent intensity (should be less than 7). A second PCR was performed to amplify the pre-amplified fragments for the desired number of cycles. Second PCR fragments were purified using Zymo DNA Clean and Concentrator-5 kit (D4014) and were run on the E-gel, and sizes were selected between 150 bp to 800 bp. The cut of gel was purified using a MinElute Gel Extraction Kit (QIAGEN, 28606), and eluting in 15 µl of water. The DNA concentration (ng/µl) was measured using Qbit, and the quality of purified libraries was accessed using a Bioanalyzer High-Sensitivity DNA Analysis kit (Agilent).

ChIP-seq

The ChIP-seq was performed following the previous protocol[61], all lysis buffers should be supplemented with protease inhibitors (Complete, EDTA-free, Roche, #11873580001). The 108 WT K562 cells, Acla- and Daun-treated cells, and K562-TOP2A-3xFLAG tag cells were crosslinked with 1% formaldehyde and incubated for 10 mins at room temperature with constant gentle stirring to cross-link. The reactions

were quenched by adding 2 M glycine to a final concentration of 125 mM. Cells were then washed with PBS and pelleted. Cell pellet containing 10^8 cells was lysed with 10 ml of LB1 (50 mM Hepes–KOH, pH 7.5; 140 mM NaCl; 1 mM EDTA; 10% Glycerol; 0.5% NP-40 or Igepal CA-630; 0.25% Triton X-100) at 4°C for 10 mins. Nuclei were pelleted at 2000 rcf for 4 mins at 4°C and resuspended in 10 ml of LB2 (10 mM Tris–HCL, pH8.0; 200 mM NaCl; 1 mM EDTA; 0.5 mM EGTA) and rocked gently at 4 °C for 5 mins. Chromatin was pelleted at 2000 rcf for 4 mins at 4°C, resuspended in 2 ml of 1XRIPA buffer (Merck, 20-188) with washed sonification beads, and sheared with a Bioruptor® Pico sonication device. The Chromatin pellet should be kept in an ice-water bath during sonication at 4°C, and sonication using 16 cycles of 30s ON and 30s OFF (every 4 cycles stop the sonification and vortex briefly). The supernatant was collected by spinning down at 3000 rpm for 5 mins at 4°C and mixed with Triton X-100 to make the final concentration of Triton X-100 to 1% in 1.5 mL tubes. Insoluble chromatin and debris were removed by spinning down at 14,000 rpm for 15 mins at 4°C.

The supernatant was incubated with 100 µl antibody/magnetic bead mix (The magnetic beads should be pre-blocked with antibody against CTCF (Merck. 07-729) or Flag (Sigma, F1804) overnight or a minimum of 4 h on a rotating platform at 4°C. Washed antibody/magnetic beads 3 times in 1 ml block solution (0.5% BSA (w/v) in PBS), and resuspended in 100 µl block solution) in a 1.5 ml tube overnight on the rotator or rocker at 4°C. Bound fragments- antibody/magnetic beads were washed with 1 ml RIPA Buffer for 4-6 more times, then washed once with 1 ml TBS (20 mM Tris–HCl, pH 7.6; 150 mM NaCl). Followed by spinning at 960 rcf for 3 mins at 4°C and using the magnetic stand to remove any residual TBS buffer. Bound fragments-antibody/magnetic beads were resuspended 200 µl of elution buffer (50 mM Tris–HCl, pH 8; 10 mM EDTA; 1% SDS) for reverse crosslinking at 65 °C for 6-18 h with vortexing 900 rpm on a thermomixer with 1000 rpm (50 µl of the WCE was reversed the formaldehyde crosslinking simultaneously with the ChIP samples). The supernatant was transferred to a new tube and incubated with 200 µl of TE to dilute SDS in the elution buffer, 8 µl of 1 mg/ml RNaseA (Ambion Cat # 2271) at 37°C for 30 mins. Then incubated with 4 µl of 20 mg/ml proteinase K (Invitrogen, 25530-049) at 55°C for 1-2 h. The ChIP DNA was cleaned up using QIAquick PCR Purification Kit (QIAGEN, 28106) and eluted in 35 µl EB.

34 µl ChIP DNA was end-repaired in a 50 µl reaction, with 5 µl 10x Beads were reclaimed and washed once with 70% ethanol without mixing followed by aspirating ethanol. End-Repair Buffer, 5 µl 2.5 mM dNTP Mix, 5 µl,10 mM ATP, and 1 µl End-Repair Enzyme Mix (Lucigen,14422-5). Incubated at room temperature for 45 mins and purified using the QIAquick PCR Purification Kit (QIAGEN, 28106), eluting in 35 µl of EB. Then the 34 µl end-repaired ChIP DNA was added of 'A' base to 3' Ends in a 50 µl reaction, with 5 µl Klenow buffer (NEB2), 10 µl 1 mM dATP, 1 µl Klenow (3' to 5' exo minus). Incubated for 30 mins at 37°C and purified using the QIAquick PCR Purification Kit (QIAGEN, 28106) eluting in 12 µl of EB. The 12 µl purified DNA was ligated with an adapter in a 30 µl reaction, with 15 µl 2x DNA ligase buffer,1 µl Adapter oligo mix (40 ng/µl), and 2 µl T4 DNA ligase. Incubated at room temperature for 15 minutes and purified using the QIAquick PCR Purification Kit (QIAGEN, 28106), eluting in 25 µl of EB. The 23 µl purified DNA was combined and mixed with 25 µl Phusion High-Fidelity PCR Master Mix (Thermofisher, F531L), 1 µl 10 µM PCR primer 1.1, 1 µl 10 µM PCR primer 2.1 to a total 50 µl reaction volume. PCR procedures were 98 °C for 30 s, 15 cycles of 98 °C for 10 s, 65 °C for 30 s, 72 °C for 30 s, and 72 °C for 5 min (15 cycles could be used for high abundant ChIPed DNA, 17 cycles could be used for low abundant ChIPed DNA). The PCR product was purified using the QIAquick PCR Purification Kit (QIAGEN, 28106), and eluted in 19 µl of EB. The adapter-ligated DNA was run on a 2% Agarose E-gel (ThermoFisher, G402022) and excised DNA of the most abundant band (should be around 300 bp, within the range of 450 bp-600 bp) from the gel with a clean scalpel (be sure to take photos of the gel before and after the gel slices are excised). The cut of gel was purified using MinElute Gel Extraction Kit (QIAGEN, 28606), and eluting in 15 µl of EB. Measure the DNA concentration (ng/µl) using Qbit, and assess the quality of purified libraries using a Bioanalyzer High-Sensitivity DNA Analysis kit (Agilent).

In situ Hi-C

In situ Hi-C was performed as previously described[6]. Pellets were lysed in ice-cold Hi-C lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% IGEPAL CA630) with 50 µl of protease inhibitors for 15 mins on ice. Cells were pelleted and washed once more using the ice-cold Hi-C lysis buffer, then resuspended in 50 µl of 0.5% SDS and incubated at 62°C for 7 mins. Reactions were quenched with 145 µl water and

25 µl 10% Triton X-100 at 37°C for 15 mins. Chromatin was digested overnight with 25 µl of 10X NEBuffer 2 and 100 U of MboI at 37°C with rotation. The MboI of the reaction was inactivated at 62°C for 20 mins followed by cooling to RT. Fragment overhangs were repaired by adding 37.5 µl 0.4 mM biotin-14-dATP, 1.5 µl each 10 mM dCTP, dGTP, dTTP, 8 µl 5 U/µl DNA Polymerase I, Large (Klenow) fragment and incubating at 37°C for 1.5 h with rotation. Ligation was performed by adding 673 µl water, 120 µl 10x NEB T4 DNA ligase buffer, 100 µl 10% Triton X-100, 6 µl 20 mg/mL BSA, and 1 µl 2000 U/µl T4 DNA ligase and incubating at RT for 4 h with slow rotation. The ligated samples were pelleted at 2500 g and resuspended in 432 µl water with 18 µl 20 mg/mL proteinase K, 50 µl 10% SDS, and 46 µl 5M NaCl, and incubated at 55°C for 30 mins and then transferred to 68°C overnight. Cooling the samples to RT and added with 1.6x volumes of pure ethanol and 0.1x volumes of 3 M sodium acetate pH 5.2, and subsequently incubated at -80°C for over 4-6 h. Then samples were spun at max speed at 2°C for 15 mins and washed twice with 70% ethanol. The resulting pellet was dissolved in 130 µl of 10 mM Tris-HCl pH 8.0 incubated at 37°C for 1-2 h and stored at 4°C overnight. The resulting DNA was sheared using the Covaris LE220 (Covaris, Woburn, MA) to a fragment size of 300-500 bp in a Covaris microTUBE. The sheared DNA was transferred to a fresh tube and the Covaris microTUBE was rinsed with 70 µl of water and pooled with the sample. A 1:5 dilution of DNA was run on a 2% agarose gel to verify shearing and the sheared DNA was size selected using AMPure XP beads. 0.55x volumes of 2X concentrated AMPure XP beads were added to each reaction and incubated at RT for 5 mins. Beads were reclaimed on a magnet and the supernatant was transferred to a fresh tube. 30 µl of 2x concentrated AMPure XP beads were added and incubated for 5 mins at RT. Beads were reclaimed on a magnet and washed with fresh 70% ethanol. Then beads were dried for 5 mins at RT before DNA elution in 300 µl of 10 mM Tris-HCl pH 8. Undiluted DNA was run on a 2% agarose gel to verify size selection between 300-500 bp.150 µl of 10 mg/mL Dynabeads MyOne Streptavidin T1 beads were washed with 400 µl of 1x Tween washing buffer (TWB, 250 µl Tris-HCl pH 7.5, 50 µl 0.5 M EDTA, 10 ml 5 M NaCl, 25 µl Tween 20, 39.675 µl water). Beads were then resuspended in 300 µl of 2x Binding Buffer (500 µl Tris-HCl (pH 7.5), 100 µl 0.5 M EDTA, 20 m 5 M NaCl, 29.4 ml water), added to the DNA sample, and incubated at RT for 15 mins with rotation. DNA-bound beads were then washed twice with 600 µl of 1x TWB at 55°C for 2 mins with shaking and

resuspended in 100 µl 1x NEBuffer T4 DNA ligase buffer, transferred to a new tube, and reclaimed. Sheared ends were repaired by resuspending the beads in 88 µl of 1x NEB T4 DNA Ligase Buffer with 1mM ATP, 2 µl of 25 mM dNTP mix, 5 µl of 10 U/µl NEB T4 PNK, 4 µl of 3 U/µl NEB T4 DNA polymerase I, and 1µl of 5U/µl NEB DNA polymerase 1, large (Klenow) fragment and incubating at RT for 30 mins. Beads were washed two more times with 1x TWB for 2 mins at 55°C with shaking, following washed once with 100 µl of 1x NEBuffer 2, transferred to a new tube, and resuspended in 90 µl of 1x NEBuffer 2, 5 µl of 10 mM dATP, and 5 µl of NEB Klenow exo minus, and incubated at 37°C for 30 mins. Beads were washed two more times with 1x TWB for 2 mins at 55°C with shaking and washed in 100 µl 1x Quick Ligation Reaction Buffer, transferred to a new tube and reclaimed. Resuspended beads in 50 µl of 1x NEB Quick Ligation Reaction Buffer with 2 µl of NEB DNA Quick Ligase and 3 µl of an appropriate Illumina indexed adapter (TruSeq nano) incubating at RT for 15 mins. Beads were reclaimed and washed twice with 1x TWB for 2 mins at 55°C, following washed in 100 µl 10 mM Tris-HCl pH 8, transferred to a new tube, reclaimed, and resuspended in 50 µl of 10 mM Tris-HCl pH 8.

Hi-C libraries were amplified directly off T1 beads with 8 cycles in 5 µl of PCR primer cocktail, 20 µl of Enhanced PCR mix, and 25 µl of DNA on beads. The PCR settings were as follows, 3 mins at 95°C followed by 4-12 cycles of 20 s 98°C, 15 s at 60°C, and 30 s at 72°C. Samples were held at 72°C for 5 mins before lowering for holding at 4°C. Amplified samples were transferred to a new tube and brought to 250 µl in 10 mM Tris-HCl pH 8. The supernatant containing the amplified library was transferred to a new tube, and the beads were resuspended in 25 µl of 10 mM Tris-HCl pH 8 stored at -20°C. 0.7x volumes of warmed AMPure XP beads were added to the supernatant sample and incubated at RT for 5 mins. Beads were reclaimed and washed once with 70% ethanol without mixing followed by aspirating ethanol. Beads were resuspended in 100 µl of 10 mM Tris-HCl pH 8, 70 µl of fresh AMPure XP beads were added, and the solution was incubated for 5 mins at RT. Beads were reclaimed and washed twice with 70% ethanol without mixing. Beads were left to dry and DNA was eluted in 25 µl of 10 mM Tris-HCl pH 8. Using Qubit and Tapestation to quantify the resµlting libraries, the resulting libraries underwent paired-end 2x150 bp sequencing on an Illumina NovaSeq sequencer.

Western blot

Cells were washed with cold PBS and lysed in LB1 (50 mM Hepes–KOH, pH 7.5; 140 mM NaCl; 1 mM EDTA; 10% Glycerol; 0.5% NP-40 or Igepal CA-630; 0.25% Triton X-100) at 4°C for 10 mins. Nuclei were pelleted at 2000 rcf for 4 mins at 4°C and resuspended in LB2 (10 mM Tris–HCL, pH8.0; 200 mM NaCl; 1 mM EDTA; 0.5 mM EGTA) and rocked gently at 4 °C for 5 mins. Chromatin was pelted at 2000 rcf for 4 mins at 4°C and resuspended in the RIPA buffer (Merck,20-188). The concentration of protein was determined using the Pierce BCA protein assay kit (Thermofisher). Equal amounts of proteins were run on the 8% SDS-PAGE and were transferred to PVDF membranes. The membranes were blocked with 5% (wt/vol) nonfat milk in PBST for 1 h at room temperature and probed with primary antibodies against Flag (Sigma, F1804), CST 12364S) or β-actin (Sigma, A5441), 1:1000 dilution overnight at 4°C. The membranes were washed with PBST three times, and incubated with the appropriate secondary antibodies (1:10000 dilution) for 1 h at room temperature, and then washed with PBST three times. The resulting signal was visualized using the Pierce enhanced chemiluminescence (ECL) Plus Western Blotting Substrate (Thermofisher).

CRISPRi targeting daunorubicin-responsive peak-to-gene links

The CTCF motif residing in the targeted peaks were identified by FIMO with the cutoff P-value < 10^{-5}. The 150 bp motif flanking genomic sequence were used to identify candidate guide RNA sequences. Two non-overlapping guide RNA sequences of high specificity and targeting either strands were designed per targeted peak-to-gene link using CRISPick (https://portals.broadinstitute.org/gppx/crispick/public). Each guide RNA sequences were individually cloned into the lentiviral vector expressing sgRNA and dCas9-KRAB[62] (Addgene, #71236) after digesting with BsmBI restriction enzyme. To produce the lentivirus, the 293T cells were grown in 6 well plates. When cells reached 50% confluency, they were transfected with the mix of 1 µg of Lenti dCas9-KRAB-Puro plasmid, 0.6 µg of psPAX2, 0.4 µg of pCMV-VSV-G, and 6 µg PEI (Polyscience, 23966) in 50 µl of serum-free medium. Media supernatant containing the virus particles was collected on the third day after transfection. The K562 cells were transduced with the virus and selected under puromycin (2 µg/mL) treatment. After 2-3 total days of puromycin selection, puromycin-containing media was

removed and replaced with regular media lacking puromycin. Cells were recovered in the absence of puromycin. The cell transduced with lentiviral expressing guide RNA not targeting the human genome was used as the control cell line for qPCR analysis.

Quantitative RT-PCR

Total RNA was isolated from 10^6 cells using ISOLATE II RNA Mini Kit (Bioline, BIO-52073) according to the manufacturer's instructions. The cDNAs were synthesized from 600 ng of total RNA using a SuperScript™ IV VILO™ Master Mix (Invitrogen, 11756050) in 5 µl reaction, and were analyzed using SensiFAST SYBR No-ROX Kit (Bioline, BIO-98020) with gene expression primers for DNM1, DNMT1, BAG2, RPS6KA1, METTL4, TRIM32, and GAPDH (a housekeeping gene) on Biorad CFX Opus 384 Real-Time PCR Systems system. The sequences of the primers are in the Supplementary Table 7.

Primary AML samples ATAC-seq library construction

The Hyperactive ATAC-Seq Library Prep Kit for Illumina (Nanjing Vazyme Biotech Co, TD711-01) was used for AML sample ATAC-seq library construction. Briefly, cells were gently thawed at RT to maximize the cell viability recovery and counted. For each sample library preparation, 10^5 live cells were centrifuged down and washed twice with 50 µl precooled Tagment DNA & Wash Buffer. For samples with treatment *in vitro*, the 10^5 live cells were treated with 1 ug/ml Daun for 3h before harvesting for library preparation. The cell pellet was resuspended in 50 µl precooled Lysis Buffer (48.5 µl RS Buffer, 0.5 µl 10% NP40, 0.5 µl 10% Tween-20 and 0.5 µl 1% Digitonin) on ice for 5 min. The cells were centrifuged at 2,300 rpm (500 × g) for 10 min at 4°C and gently removed the supernatant. Then, the cells were resuspended in 50 µl of precooled fragmentation mix (16.5 µl TW Buffer, 0.5 µl 10% Tween-20, 0.5 µl 1% Digitonin, 18.5 µl Nuclease-free ddH_2O, 10 µl 5 × TTBL and 4 µl TTE Mix V50). Gently mix the pipette mix up and down 10 times followed by incubation in waterbath at 37°C for 30 min. After the tagmentation reaction, mix well by gently pipetting up and down and transfer to a 200 µl PCR tube. After the transfer, add 5 µl of Stop Buffer to each PCR tube and resuspend the cells then incubate at RT for 5 min. To extract the fragmented DNA, 100 µl of ATAC DNA Extract Beads mixed by vortexing was pipetted to 55 µl of the terminated sample and then

incubated at RT for 5 min. Wash the magnetic beads twice with 200 μl of freshly prepared 80% ethanol, incubate at RT for 30 sec, and carefully remove the supernatant. After air-dry for 5 min, remove the reaction tube from the magnetic rack and elute with 26 μl of Nuclease-free ddH_2O followed by incubation at RT for 5 min. Centrifuge the reaction tube briefly and place it on the magnetic rack to separate the magnetic beads from the solution. Wait 5 min until the solution becomes clear, then carefully pipette 24 μl of the supernatant into a new PCR tube. Library was amplified using 60 ul PCR reaction as 20 μl Purified Fragmented DNA, 30 μl 2×CAM, 5 μl N5XX PCR primer, and 5 μl N7XX PCR primer. The PCR program is as follows: 3 min 72°C, 3 min 95°C; 12 cycles of 10 sec 98°C and 5 sec 60°C, 1 min 72°C. The PCR product was size selected using ATAC DNA Clean Beads by mixing 33 μl (0.55 ×) magnetic beads into 60 μl PCR products gently followed by incubation at RT for 5 min. Wash the magnetic beads twice with 200 μl of freshly prepared 80% ethanol, incubate at RT for 30 sec, and carefully remove the supernatant. After air-dry for 5 min, remove the reaction tube from the magnetic rack and elute with 22 μl of Nuclease-free ddH_2O followed by incubation at RT for 5 min. Centrifuge the reaction tube briefly and place it on the magnetic rack to separate the magnetic beads from the solution. Wait 5 min until the solution becomes clear, then carefully pipette 20 μl of the supernatant into a new sterile tube. The quality control of the DNA library was performed using Bioanalyzer.

Primary AML samples RNA library construction and sequencing

Total RNA was extracted using TRIzol Reagent (Invitrogen, USA) according to the manufacturer's instructions. After total RNA was extracted, mRNA was purified from total RNA (5ug) using Dynabeads Oligo (dT) (Thermo Fisher, CA, USA) with two rounds of purification. Fragmentation of mRNA was performed using divalent cations under elevated temperature (Magnesium RNA Fragmentation Module (NEB, cat.e6150, USA) under 94℃ 5-7 min). Then the cleaved RNA fragments were reverse-transcribed to cDNA by SuperScript™ II Reverse Transcriptase (Invitrogen, cat. 1896649, USA), which were next used to synthesize U-labeled second-stranded DNAs with E. coli DNA polymerase I (NEB, cat.m0209, USA), RNase H (NEB, cat.m0297, USA) and dUTP Solution (Thermo Fisher, cat.R0133, USA). An A-base was then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters. Each adapter contained a T-base overhang for ligating the

adapter to the A-tailed fragmented DNA. Dual-index adapters were ligated to the fragments, and size selection was performed with AMPureXP beads. After the heat-labile UDG enzyme (NEB, cat.m0280, USA) treatment of the U-labeled second-stranded DNAs, the ligated products were amplified with PCR by the following conditions: initial denaturation at 95℃ for 3 min; 8 cycles of denaturation at 98℃ for 15 sec, annealing at 60℃ for 15 sec, and extension at 72℃ for 30 sec; and then final extension at 72℃ for 5 min. The average insert size for the final cDNA libraries was 300 ± 50 bp. At last, we performed the 2×150bp paired-end sequencing (PE150) on an Illumina Novaseq™ 6000 (LC-Bio Technology CO., Ltd., Hangzhou, China) following the vendor's recommended protocol.

Informatics analysis

ATAC-seq data processing

The ATAC-seq samples from K562 cells were processed using the ENCODE Data Coordination Center (DCC) ATAC-seq pipeline (v1.10.0) (https://github.com/ENCODE-DCC/atac-seq-pipeline). Briefly, the ATAC-seq reads were aligned to the human reference genome (GRCh37/hg19) using Bowtie2[63]. Duplicated and mitochondrial reads were filtered. The ATAC-seq fragment sizes for all the samples were corrected for the Tn5 offset ("+" stranded +4 bp, "-" stranded -5 bp) and plotted using the geom_density() function in the R package ggplot2.

Enrichment analysis of Tn5 insertion sizes within chromatin states

The ATAC-seq insertion size enrichment analysis in different chromatin states were performed as described[24]. First, the Tn5 insertion size ranging from 50 bp to 750 bp were divided into 100 bins and the percent of insertions in each bin were calculated followed by normalizing to the maximal percent within each state. Finally, the enrichment of different Tn5 insertion sizes within each state was calculated by normalizing to the genome-wide different Tn5 fragment sizes.

Nucleosome positioning analysis

To define the regions of nucleosome analysis, broad peaks were first called from all the ATAC-seq samples identically using MACS2[64] with the settings of '-p 0.05 --broad --broad-cutoff 0.05 --keep-dup all' and then extended by 200 bp on either

sides. The python package NucleoATAC[30] (v0.3.4) was used to call nucleosome positions and occupancy within the defined regions from low-quality reads filtered and non-mitochondrial BAM files. The predicted nucleosome occupancy scores signal tracks from different samples by NucleoATAC around interesting regions were compared and plotted by deepTools[65].

TF footprinting analysis

The TF footprints were identified using HINT-ATAC[31] within the broad peaks called by MACS2 with the settings of '-p 0.05 --broad --broad-cutoff 0.05 --keep-dup all'. The TF motifs from JASPAR database[66] were used to match within the predicted footprints to infer the bound (active) TFs. The average Tn5 cleavage profiles around binding sites of particular TFs were plotted and the TF activity scores were calculated to compare TF binding strength under different drug treatment conditions.

Differential chromatin accessibility analysis

The R package csaw[67] were used for differential chromatin accessibility analysis as suggested[68]. Briefly, filtered BAM files of two biological replicates from two conditions are specified as the inputs for the csaw pipeline. The genome were segmented into 300bp windows and reads count within each window were computed for each sample. The windows with low reads enrichment (< 3-fold threshold relative to 3 Kbp neighborhood) were filtered. TMM normalization was implemented for read counts within each window and de novo locally enriched windows were used as the query regions for differential accessibility comparison. Differentially accessible nearby windows within 100 bp apart were merged. Significant differential regions were filtered using the threshold of FDR < 0.05.

ChIP-seq data processing and visualization

The ChIP-seq data were processed using the ENCODE Data Coordination Center (DCC) ChIP-seq pipeline (https://github.com/ENCODE-DCC/chip-seq-pipeline2) (v1.9.0). Briefly, the ChIP-seq reads were aligned to the human reference genome (GRCh37/hg19) using Bowtie2[63]. Duplicate reads were removed using Picard MarkDuplicates (https://broadinstitute.github.io/picard/). Peaks of each biological replicate were called against the whole-cell lysates (WCL) replicates using SPP[69] with the parameters '-npeak 300000 -speak 155 -fdr 0.01'. Reproducible peaks were

intersected from two biological replicates. The blacklisted regions described by ENCODE were discarded.

The ChIP-seq signal tracks were generated from BAM files using the command 'bedtools genomecov -scale' with read count per million (CPM) normalization and converted to bigwig files using bedGraphToBigWig[70]. The genome snapshots were created with Integrative Genomics Viewer (IGV) browser[71]. The heatmap of ChIP-seq signal in the regions of interest and similarity comparisons of genome-wide ChIP-seq signal distribution were computed by deepTools[65].

Differential ChIP-seq peaks calling

The R package csaw[67] were used for differential peak calling for TOP2A and CTCF ChIP-seq experiments. Briefly, filtered BAM files of two biological replicates from two conditions are specified as the inputs for the csaw pipeline. The genome was segmented into 20bp windows and reads count within each window were computed for each sample. The windows with low reads enrichment (< 3-fold threshold relative to 2Kb neighborhood) were filtered. TMM normalization was implemented for read counts within each window and de novo locally enriched windows were used as the query regions for differential binding affinity comparison. Differential nearby windows within 50 bp apart were merged. Significant differential ChIP-seq peaks were filtered using the threshold of FDR < 0.05. Motif enrichment analysis within the differential ChIP-seq peaks was performed by HOMER[72].

Permutation association test for ChIP-seq peak sets

The associations between two ChIP-seq peak sets were determined by calculating the significance of overlapping peaks based on the genome coordinations. The function enrichPeakOverlap() in R package ChIPseeker[73] were used to shuffle the genome coordinations of the target peak set and calculate the overlaps of the query peak set with the shuffled target peak set. The permutation step was repeated for 50,000 times to estimate the null distribution. Then the P-values were calculated and adjusted by the Benjamini-Hochberg (BH) method.

Hi-C data processing and visualization

In situ Hi-C datasets were processed using a modified version of the Juicer Hi-C pipeline (https://github.com/EricSDavis/dietJuicer) with default parameters as

previously described[74]. Reads were aligned to the GRCh37/hg19 human reference genome with bwa (v0.7.17) and MboI was used as the restriction enzyme. Four biological replicates were aligned and merged. Chromatin loops were called using the supervised machine learning framework called Peakachu[42] which has been pre-trained at different sequencing depth to overcome bias from different samples. The replicates-merged, MAPQ > 30 filtered and Knight & Ruiz methods (KR) balanced .hic matrices were used as inputs for calling loops at 10Kbp resolution with respective sequencing depth models (control: 1.2 billion model, Acla: 900 million model, Daun: 850 million model) and the cut-off of loop probability above 0.9. This resulted in 37,479 loops for control, 15,347 loops for Acla-treated and 9,438 loops for Daun-treated samples after filtering.

For visualizing the Hi-C matrices, the R package plotgardener[75] was used with the KR-normalized Hi-C matrices as input. To get the idea of global chromatin folding pattern change after anthracycline treatment, we plotted the *P(s)* curves of samples either with or without anthracycline treatment using the cooltools workflow[76] (https://github.com/open2c/open2c_examples/blob/master/contacts_vs_distance.ipynb), which compared the rate of decay of contacts with the genomic separation distance across different samples.

A/B compartments calling and visualization

The A/B compartment profiles were calculated as previously described[77]. Briefly, the genome was segmented into 100 Kbp bins using bedtools[78] and the genomic bins were overlapped with gene annotation to count gene density within each bin Juicertools[74] was used to calculate the first eigenvector of each chromosome from KR-balanced .hic matrix input. The correlation coefficient between the first eigenvector and gene density in each chromosome was calculated to correct the sign assignment of the first eigenvector. The eigenvector from all the chromosomes was combined into a single bedgraph file and then converted into a bigwig file using bedGraphToBigWig[70]. The differential compartment analysis were performed by correlating the interaction profile of a locus in control Hi-C matrix to the interaction of same locus in the drug-treated Hi-C matrix using HiC1Dmetrics[79] command “h1d two CD” with the resolution size of 100 Kbp. The cutoff of pearson correlation coefficient below 0.4 was used to identify the substantial compartment status switch.

To generate the saddle plot, the 100 Kbp genomic bins were sorted by the eigenvector values and divided into 100 quantiles with similar eigenvector values. The average observed/expected contact frequency between bins grouped by eigenvector values were shown after filtering the 2.5% of genomic bins with the lowest or highest eigenvector values. Inactive chromatin regions (B compartment) with low eigenvector values are on the top and left and active chromatin regions (A compartment) with high eigenvector values are on the bottom and right.

TAD boundaries calling and visualization

To identify the TAD boundaries, we use the diamond insulation algorithm[39] implemented in cooltools pipeline[76] with the input of KR-balanced Hi-C matrix and the parameters setting of 5 Kbp bin size and 100 Kbp window size. The genome-wide boundary scores (*BS*) and insulation scores (*IS*) were calculated as output. The 5Kbp bins satisfying the cutoff of $0.2 \leq IS < 0.5$ and $IS \geq 0.5$ were called as weak and strong boundaries respectively as suggested (https://data.4dnucleome.org/resources/data-analysis/insulation_compartment_scores). The differential boundary analysis was performed as previously described[40]. Briefly, lost boundaries by anthracycline treatments were defined by the diamond insulation algorithm, i.e. same locus bins not called as boundaries in drug-treated samples but called as boundaries in the control sample. Residual boundaries by anthracycline treatments were defined if the locus bins called as boundaries in drug-treated samples. No gained boundaries were reported in Acla- and Daun-treated samples by the diamond insulation pipeline. The bins with $BS^{control} - BS^{drug} \geq 0.4$ and $IS^{control} - IS^{drug} \leq -0.1$ were defined as weakened or less insulated boundaries by drug treatment. The bins with $BS^{control} - BS^{drug} \leq -0.4$ and $IS^{control} - IS^{drug} \geq 0.1$ were defined as weakened or less insulated boundaries by drug treatment. The other residual boundaries were regarded as stable.

The correlation of boundary scores and insulation scores of residual TAD boundaries after anthracycline treatment were plotted using the R package ggplot2. The local pileups of lost or differential boundaries were calculated using coolpup.py[80] with the parameters '--local --ignore_diags 0 --view --flank 500000' and visualized with plotpup.py.

Differential chromatin loops calling

To identify the drug-induced lost and gained loops, we relied on the genome-wide loop probability prediction by Peakachu[42], which is positively associated with loop intensity and the fold-change of Peakachu probability at the same locus in two individual samples can be used to measure the looping dynamics[10] (i.e. gained or lost). A Gaussian mixture model was used to fit the reciprocal fold-changes of the Peakachu probabilities in the merged loop regions which was called either from control or anthracycline-treated samples. The threshold of FDR below 0.05 was used as the threshold to extract the significant differential loops. This procedure resulted in 6,328 gained and 24,460 lost loops by Acla treatment and 4,800 gained and 28,025 lost loops by Daun treatment. The non-overlapped loops were regarded as stable loops.

Aggregate peak analysis

To assess the quality of differential loops calling, we generated the aggregate peak analysis (APA) plots and scores by using the *apa-analysis* script in HiCPeaks (https://github.com/XiaoTaoWang/HiCPeaks/) with the default parameters. Briefly, the pixels within 5 bins in both directions were extracted (11x11) and the normalized contact signals at each position were calculated by dividing the observed contacts by the expected contacts at the distance accordingly. The mean of observed over expected ratios at each position of the matrix was assigned to the pixel and plotted out. The APA score was calculated by dividing the center pixel value by the mean value of the nine pixels (3x3) in the lower right section. The z-score was calculated by dividing the difference between center pixel value and mean of the lower right section by the standard deviation of the lower right section. The p-value was converted from z-score using one-side t-test.

The APA analysis around *cis*-regulatory loops were performed as described above. To determine different types of *cis*-regulatory loops, the CTCF-CTCF loop was called by overlapping the both anchors of Hi-C loop with CTCF ChIP-seq peaks, the promoter-promoter loop was called by overlapping both anchors of Hi-C loop with the promoter regions (within -1500 to +500 bp from TSS) presenting active promoter mark (H3K4me3 peaks), the enhancer-promoter loops by overlapping one anchor of Hi-C loop with distal regions (excluding promoters) presenting active promoter mark (H3K27ac peaks), the other anchor with active H3K4me3 promoters, the enhancer-

enhancer loops by overlapping both anchors of Hi-C loops with the active H3K27ac distal enhancers.

Virtual 4C analysis at *Myc* promoter

The track for the virtual 4C analysis anchored at *Myc* promoter based on Hi-C data was generated using CoolBox[81] with default parameters. The 5 Kbp resolution Hi-C matrix was used as input and the validated 7 *Myc* enhancers[43] were highlighted.

KEGG pathway enrichment in differential chromatin loops

KEGG pathway enrichment was performed on the genes whose promoters (within -1500 to +500 bp from the TSSs) overlapped either a gained or lost loop anchor after the drug treatments using clusterProfiler[82] (v4.12). KEGG pathway terms were filtered using the cutoff of adjusted p-value below 0.05.

AML RNA-seq data processing

The adaptors were trimmed from paired-end RNA-seq reads using Cutadapt (v1.9). The trimmed reads were aligned to the human reference genome (GRCh37/hg19) using HISAT2[83] (v2.2.1). The mapped reads were assembled using StringTie[84] (v2.1.6) with default parameters. Then, all transcriptomes from all samples were collapsed to reconstruct a comprehensive transcriptome using gffcompare[85] (v0.9.8). After the final transcriptome was generated, StringTie and ballgown[86] were used to estimate the expression levels of all transcripts and perform expression abundance for mRNAs by calculating FPKM (fragment per kilobase of transcript per million mapped reads) value.

AML ATAC-seq data processing

The ATAC-seq samples from primary AML patients were identically processed using the Snakemake pipeline (https://github.com/porchard/ATACseq-Snakemake). Briefly, the ATAC-seq adapters were trimmed from the paired-end ATAC-seq reads using cta (v0.1.2) (https://github.com/ParkerLab/cta). The trimmed reads were aligned to the human reference genome (GRCh37/hg19) using BWA-MEM. The duplicated reads were removed using Picard MarkDuplicates. The low-quality mapped and mitochondrial reads were filtered.

Reproducible ATAC-seq peaks calling from AML samples

Peaks were called using MACS2 with the parameters of '-p 0.01 --shift -75 --extsize 150 --nomodel --call-summits --nolambda --keep-dup all', then followed by blacklisted regions filtering recommended by ENCODE[87] and peaks mapped to chromosome Y removal. To create a single merged ATAC-seq peak set for all samples, we first categorize all the samples based on a) the patient's response assessment of daunorubicin-based chemotherapy: complete response (CR) as sensitive group, partial response (PR) and no response (NR) as refractory or relapsed (RR) group, b) the in vitro treatment with daunorubicin or not into 4 groups, i.e. CR, CR_D, RR, RR_D. All input peak summits from MACS2 calling were identically extended by 250 bp on either side to a final width of 501 bp and the non-overlapped, merged peak set was identified using the iterative overlap approach[88] (https://github.com/corceslab/ATAC_IterativeOverlapPeakMerging) which results in the union set of 60,881 reproducible peaks.

Reads-in-peaks normalization of AML ATAC-seq signal tracks

The ATAC-seq signal tracks were generated by the reads-in-peaks normalization approach[49] as the patient's primary sample of varying quality was not comparable for the varying signal-to-noise ratio, i.e. fraction of mapped reads in called peaks (FRiP). Briefly, the genome was segmented into 100 bp bins and the Tn5 offset-corrected ATAC-seq reads falling in bins overlapping peak regions and background regions were quantified separately. The ratio between peak reads sum and background reads sum was calculated as the scale factor to normalize each sample read depth in peak regions or background into the constant total 30 million reads per sample. The bins of read counts for each sample were converted into bigwig files using bedGraphToBigWig[70].

Construction of ATAC-seq peak accessibility matrix

To quantify the Tn5 insertions within the merged peak set for each sample, first the BAM files for 30 samples were identically converted to BEDPE format and adjusted for the Tn5 offset ("+" stranded +4 bp, "-" stranded -5 bp). Then the number of Tn5 cleavage falling in each peak was counted for each sample which was collapsed into a 60,881×30 peaks-by-samples count matrix. The raw count matrix was normalized first using $\log_2$(CPM) with the prior count of 5 and then quantile normalized using the

normalize.quantiles() function in R package preprocessCore. The PCA analysis was performed using the prcomp() function in R package ggfortify.

Distal binarization identifies daunorubicin-responsive peaks

To identify differential peaks that are unique to multiple groups, we applied the classification method termed "distal binarization" as previously described[49]. Briefly, the intra-group mean $\mu_{i,j}$ and intra-group standard deviation $\sigma_{i,j}$ of normalized ATAC-seq signal in each peak *i* across all the samples in each group *j* were calculated. For each peak *i*, the groups were ranked by the intra-group mean ($\mu_{i,j}^{1st}$, $\mu_{i,j}^{2nd}$,..., $\mu_{i,j}^{kth}$,...,$\mu_{i,j}^{Nth}$) and iterated from the group j^{1st} with *1st* lowest mean ATAC-seq signal. The criteria to define the break point is if the next-lowest intra-group mean $\mu_{i,j}^{(k+1)th}$ of the next group *(k+1)th* is greater than the maximum of other *k* lower intra-group means plus their intra-group standard deviations $max(\mu_{i,j}^{1st} + \sigma_{i,j}^{1st}, \ldots, \mu_{i,j}^{kth} + \sigma_{i,j}^{kth})$. A sign matrix was assigned with "0" before break point and "1" after break point. To assess the statistical significance of the binary assigned group-unique peaks, the limma's eBayes differential test was performed on the normalized ATAC-seq signal matrix with the contrast matrix input which is defined as the binary assigned differential groups combination. The binary assigned group-unique peaks below the FDR cutoff of 0.05 were kept and sorted based on the binary assigned differential groups combinations. The distal binarization method results in 14,260 significant group-unique peaks. The sorted ATAC-seq signal matrix of binary assigned group-unique peaks was plotted using the R package ComplexHeatmap[89]. Daunorubicin-responsive ATAC-seq peaks were identified based on the binary assigned group combinations, for examples, peaks which were assigned to the combination of "CR:1, CR_D:1, RR:1, RR_D:0" were extracted as the refractory/relapsed AML patients-unique daunorubicin-deactivated ATAC-seq peaks.

Motif enrichment analysis in daunorubicin-responsive peaks

The motifs enrichment analysis within group-unique peak sets was performed first using the R package motifmatchr to compute the occurrence of known motifs from CIS-BP database. To identify the motif specifically enriched within the group-unique peak sets, the background motif occurrences within the total merged ATAC-seq peak set were computed as well. The hypergeometric test of the motif representation

within the group-unique peak sets compared to the total merged peak set was calculated using the R package phyper with the cutoff of P-value < 10^{-5}. The motif enriched in at least one but not all groups were sorted based on $\log_{10}$(P-value) and plotted using the R package ComplexHeatmap[89].

Correlation analysis to predict daunorubicin-responsive peak-to-gene links

To infer the causal regulatory links between daunorubicin-responsive ATAC-seq peaks and target genes, we performed the correlation analysis between chromatin accessibility and gene expression using 20 cases of paired RNA-seq and ATAC-seq data from primary AML patients as previously described[49]. First, the ATAC-seq peak accessibility count matrix was normalized using $\log_2$(CPM) with the prior count of 5 and the RNA-seq gene expression count matrix was normalized using $\log_2$(RPKM+1). Second, the less variable (below 25th quantile) ATAC-seq peaks and genes were filtered from both matrices as they will arise false-positive predictions based on the correlation approach. Third, all the potential interactions between ATAC-seq peaks and gene promoters (within -1500 to +500 bp from the TSSs) within 0.5 Mbp were extracted and the Pearson correlation was calculated for each potential interaction. To assess the statistical significance of identified correlations between ATAC-seq peak accessibility and gene expression, the null distribution was constructed on each single chromosome by correlating the chromatin accessibility of random 10000 peaks from the other different chromosomes and the gene expression of each gene in this chromosome. The correlation model of non-specific interactions between ATAC-seq peak chromatin accessibility and gene expression on a single chromosome were estimated by the mean and standard deviation calculated from the “trans” peak-to-gene interactions. The cutoff of p-value below 0.05 was applied to filter the non-significant correlations. The peak-to-gene links falling within the gene promoter regions were filtered to identify potential long-range interactions between distal regulatory elements (e.g. CTCF loops) and target genes. This results in 32,060 significant peak-to-gene links in total. To identify the daunorubicin-responsive peak-to-gene links, the group-unique differential peaks from distal binarization were overlapped with the predicted peak-to-gene links. The potential interactions driven by CTCF from daunorubicin-responsive peak-to-gene links were further extracted by overlapping the CTCF motif scanned by FIMO with the cutoff of p-value below 10^{-5}.

The predicted peak-to-gene links were visualized using the R package plotgardener[75].

Statistics and reproducibility

The western blots for CTCF, and β-actin were repeated three times, and similar results were obtained (**Fig. 3a**). Two biological replications were applied to each ATAC-seq and ChIP-seq experiment in K562 cells. Four biological replicates were applied to each in situ Hi-C experiment in K562 cells. Three biological replications were applied to each qPCR experiment in CRISPRi-editing K562 cells. All other statistics and replications are specified in each method and figure legend. External data used to reproduce the figures is listed in the Supplementary Table 8.

Data availability

The raw sequencing data including ATAC-seq and TOP2A ChIP-seq of cell lines generated in the previous study[59] were deposited at GEO under accession code GSE240443. The raw sequencing data including CTCF ChIP-seq and Hi-C of cell lines and processed ATAC-seq and RNA-seq data of primary samples generated in this study are available at GEO under the accession code GSE276066.

Code availability

The scripts for reproducing the genomic data analysis are available on GitHub at (https://github.com/MorganTan95/ANTH_3D_genome_paper).

Acknowledgements

This work was supported by the Gisela Thier Fellowship from Leiden University Medical Center and the ERC Starting Grant 950655-Silencer from the European Research Council (all awarded to B.P.), and by the National Institutes of Health (R35-GM128645 to D.H.P. and 5T32 CA217824 to A.A.P.).

Author contribution

B.P. conceptualized the project. M.T., S.S., Y.L., A.A.P. and B.P conducted investigations. M.T., S.S., A.A.P., and D.H.P conducted formal analysis. M.T., S.S., and B.P. wrote the original draft. B.P., M.T., S.S., D.H.P. and reviewed and edited the manuscript. B.P., D.L. and D.H.P. supervised the project. B.P. acquired funding.

Declarations of Interests

The authors declare no competing interests.

References

1 Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat Rev Mol Cell Biol* **20**, 535-550, doi:10.1038/s41580-019-0132-4 (2019).

2 Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293, doi:10.1126/science.1181369 (2009).

3 Meaburn, K. J. & Misteli, T. Cell biology: chromosome territories. *Nature* **445**, 379-781, doi:10.1038/445379a (2007).

4 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

5 Vian, L. *et al.* The Energetics and Physiological Impact of Cohesin Extrusion. *Cell* **173**, 1165-1178 e1120, doi:10.1016/j.cell.2018.03.072 (2018).

6 Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).

7 Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep* **15**, 2038-2049, doi:10.1016/j.celrep.2016.04.085 (2016).

8 Pang, B. & Snyder, M. P. Systematic identification of silencers in human cells. *Nature Genetics* **52**, 254-263, doi:10.1038/s41588-020-0578-5 (2020).

9 Kloetgen, A. *et al.* Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat Genet* **52**, 388-400, doi:10.1038/s41588-020-0602-9 (2020).

10 Xu, J. *et al.* Subtype-specific 3D genome alteration in acute myeloid leukaemia. *Nature* **611**, 387-398, doi:10.1038/s41586-022-05365-x (2022).

11 Ahn, J. H. *et al.* Phase separation drives aberrant chromatin looping and cancer development. *Nature* **595**, 591-595, doi:10.1038/s41586-021-03662-5 (2021).

12 Sungalee, S. *et al.* Histone acetylation dynamics modulates chromatin conformation and allele-specific interactions at oncogenic loci. *Nat Genet* **53**, 650-662, doi:10.1038/s41588-021-00842-x (2021).

13 Wang, J. *et al.* Epigenomic landscape and 3D genome structure in pediatric high-grade glioma. *Sci Adv* **7**, doi:10.1126/sciadv.abg4126 (2021).

14 Nitiss, J. L. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer* **9**, 338-350, doi:10.1038/nrc2607 (2009).

15 Pang, B. *et al.* Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. *Nat Commun* **4**, 1908, doi:10.1038/ncomms2921 (2013).

16 Qiao, X. *et al.* Uncoupling DNA damage from chromatin damage to detoxify doxorubicin. *Proc Natl Acad Sci U S A* **117**, 15182-15192, doi:10.1073/pnas.1922072117 (2020).

17 Pang, B., de Jong, J., Qiao, X., Wessels, L. F. & Neefjes, J. Chemical profiling of the genome with anti-cancer drugs defines target specificities. *Nat Chem Biol* **11**, 472-480, doi:10.1038/nchembio.1811 (2015).

18 Oberbeckmann, E., Quililan, K., Cramer, P. & Oudelaar, A. M. In vitro reconstitution of chromatin domains shows a role for nucleosome positioning in 3D genome organization. *Nat Genet* **56**, 483-492, doi:10.1038/s41588-023-01649-8 (2024).

19 Pommier, Y., Nussenzweig, A., Takeda, S. & Austin, C. Human topoisomerases and their roles in genome stability and organization. *Nat Rev Mol Cell Biol* **23**, 407-427, doi:10.1038/s41580-022-00452-3 (2022).

20 Teves, S. S. & Henikoff, S. Transcription-generated torsional stress destabilizes nucleosomes. *Nat Struct Mol Biol* **21**, 88-94, doi:10.1038/nsmb.2723 (2014).

21 Uuskula-Reimand, L. & Wilson, M. D. Untangling the roles of TOP2A and TOP2B in transcription and cancer. *Sci Adv* **8**, eadd4920, doi:10.1126/sciadv.add4920 (2022).

22 Canela, A. *et al.* Genome Organization Drives Chromosome Fragility. *Cell* **170**, 507-521 e518, doi:10.1016/j.cell.2017.06.034 (2017).

23 Gothe, H. J. *et al.* Spatial Chromosome Folding and Active Transcription Drive DNA Fragility and Formation of Oncogenic MLL Translocations. *Mol Cell* **75**, 267-283 e212, doi:10.1016/j.molcel.2019.05.015 (2019).

24 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218, doi:10.1038/nmeth.2688 (2013).

25 Hoffman, M. M. *et al.* Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res* **41**, 827-841, doi:10.1093/nar/gks1284 (2013).

26 Clarkson, C. T. *et al.* CTCF-dependent chromatin boundaries formed by asymmetric nucleosome arrays with decreased linker length. *Nucleic Acids Res* **47**, 11181-11196, doi:10.1093/nar/gkz908 (2019).

27 Fu, Y., Sinha, M., Peterson, C. L. & Weng, Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet* **4**, e1000138, doi:10.1371/journal.pgen.1000138 (2008).

28 Lai, W. K. M. & Pugh, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* **18**, 548-562, doi:10.1038/nrm.2017.47 (2017).

29 Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* **28**, 817-825, doi:10.1038/nbt.1662 (2010).

30 Schep, A. N. *et al.* Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res* **25**, 1757-1770, doi:10.1101/gr.192294.115 (2015).

31 Li, Z. *et al.* Identification of transcription factor binding sites using ATAC-seq. *Genome Biol* **20**, 45, doi:10.1186/s13059-019-1642-2 (2019).

32 Dubois, F., Sidiropoulos, N., Weischenfeldt, J. & Beroukhim, R. Structural variations in cancer and the 3D genome. *Nat Rev Cancer* **22**, 533-546, doi:10.1038/s41568-022-00488-9 (2022).

33 Huang, H. *et al.* CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. *Nat Genet* **53**, 1064-1074, doi:10.1038/s41588-021-00863-6 (2021).
34 Kubo, N. *et al.* Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat Struct Mol Biol* **28**, 152-161, doi:10.1038/s41594-020-00539-5 (2021).
35 Oh, S. *et al.* Enhancer release and retargeting activates disease-susceptibility genes. *Nature* **595**, 735-740, doi:10.1038/s41586-021-03577-1 (2021).
36 Uuskula-Reimand, L. *et al.* Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biol* **17**, 182, doi:10.1186/s13059-016-1043-8 (2016).
37 Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018, doi:10.1093/bioinformatics/btr064 (2011).
38 Hashimoto, H. *et al.* Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Mol Cell* **66**, 711-720 e713, doi:10.1016/j.molcel.2017.05.004 (2017).
39 Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240-244, doi:10.1038/nature14450 (2015).
40 Zhou, Y. *et al.* EBF1 nuclear repositioning instructs chromatin refolding to promote therapy resistance in T leukemic cells. *Mol Cell* **82**, 1003-1020 e1015, doi:10.1016/j.molcel.2022.01.015 (2022).
41 Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944 e922, doi:10.1016/j.cell.2017.05.004 (2017).
42 Salameh, T. J. *et al.* A supervised learning framework for chromatin loop detection in genome-wide contact maps. *Nat Commun* **11**, 3428, doi:10.1038/s41467-020-17239-9 (2020).
43 Fulco, C. P. *et al.* Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**, 769-773, doi:10.1126/science.aag2445 (2016).
44 Lin, X. *et al.* Nested epistasis enhancer networks for robust genome regulation. *Science* **377**, 1077-1085, doi:10.1126/science.abk3512 (2022).
45 Pang, B. & Snyder, M. P. Systematic identification of silencers in human cells. *Nat Genet* **52**, 254-263, doi:10.1038/s41588-020-0578-5 (2020).
46 Corces, M. R. & Corces, V. G. The three-dimensional cancer genome. *Curr Opin Genet Dev* **36**, 1-7, doi:10.1016/j.gde.2016.01.002 (2016).
47 Desmedt, C. *et al.* Multifactorial approach to predicting resistance to anthracyclines. *J Clin Oncol* **29**, 1578-1586, doi:10.1200/JCO.2010.31.2231 (2011).
48 Nielsen, D., Maare, C. & Skovsgaard, T. Cellular resistance to anthracyclines. *Gen Pharmacol* **27**, 251-255, doi:10.1016/0306-3623(95)02013-6 (1996).
49 Corces, M. R. *et al.* The chromatin accessibility landscape of primary human cancers. *Science* **362**, doi:10.1126/science.aav1898 (2018).
50 Wu, H. *et al.* Highly sensitive single-cell chromatin accessibility assay and transcriptome coassay with METATAC. *Proc Natl Acad Sci U S A* **119**, e2206450119, doi:10.1073/pnas.2206450119 (2022).

51 Tarjan, D. R., Flavahan, W. A. & Bernstein, B. E. Epigenome editing strategies for the functional annotation of CTCF insulators. *Nat Commun* **10**, 4258, doi:10.1038/s41467-019-12166-w (2019).

52 Tremblay, C. S. *et al.* Small molecule inhibition of Dynamin-dependent endocytosis targets multiple niche signals and impairs leukemia stem cells. *Nat Commun* **11**, 6211, doi:10.1038/s41467-020-20091-6 (2020).

53 Liu, J. *et al.* E3 ubiquitin ligase TRIM32 negatively regulates tumor suppressor p53 to promote tumorigenesis. *Cell Death Differ* **21**, 1792-1804, doi:10.1038/cdd.2014.121 (2014).

54 Tzelepis, K. *et al.* A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Rep* **17**, 1193-1205, doi:10.1016/j.celrep.2016.09.079 (2016).

55 Wang, T. *et al.* Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. *Cell* **168**, 890-903 e815, doi:10.1016/j.cell.2017.01.013 (2017).

56 Weidenauer, K. *et al.* The ribosomal protein S6 kinase alpha-1 (RPS6KA1) induces resistance to venetoclax/azacitidine in acute myeloid leukemia. *Leukemia* **37**, 1611-1625, doi:10.1038/s41375-023-01951-8 (2023).

57 Malvi, P., Wang, B., Shah, S. & Gupta, R. Dissecting the role of RNA modification regulatory proteins in melanoma. *Oncotarget* **10**, 3745-3759 (2019).

58 Wang, F., Wang, F., Zhang, S. & Xu, X. MicroRNA-325 inhibits the proliferation and induces the apoptosis of T cell acute lymphoblastic leukemia cells in a BAG2-dependent manner. *Exp Ther Med* **21**, 631, doi:10.3892/etm.2021.10063 (2021).

59 Qiao, X. *et al.* Diversifying the anthracycline class of anti-cancer drugs identifies aclarubicin for superior survival of acute myeloid leukemia patients. *Mol Cancer* **23**, 120, doi:10.1186/s12943-024-02034-7 (2024).

60 Corces, M. R. *et al.* An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**, 959-962, doi:10.1038/nmeth.4396 (2017).

61 Schmidt, D. *et al.* ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods* **48**, 240-248, doi:10.1016/j.ymeth.2009.03.001 (2009).

62 Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods* **12**, 1143-1149, doi:10.1038/nmeth.3630 (2015).

63 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

64 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

65 Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-191, doi:10.1093/nar/gku365 (2014).

66 Castro-Mondragon, J. A. *et al.* JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **50**, D165-D173, doi:10.1093/nar/gkab1113 (2022).

67 Lun, A. T. & Smyth, G. K. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res* **44**, e45, doi:10.1093/nar/gkv1191 (2016).

68 Reske, J. J., Wilson, M. R. & Chandler, R. L. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. *Epigenetics Chromatin* **13**, 22, doi:10.1186/s13072-020-00342-y (2020).
69 Kharchenko, P. V., Tolstorukov, M. Y. & Park, P. J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol* **26**, 1351-1359, doi:10.1038/nbt.1508 (2008).
70 Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204-2207, doi:10.1093/bioinformatics/btq351 (2010).
71 Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26, doi:10.1038/nbt.1754 (2011).
72 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
73 Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383, doi:10.1093/bioinformatics/btv145 (2015).
74 Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* **3**, 95-98, doi:10.1016/j.cels.2016.07.002 (2016).
75 Kramer, N. E. *et al.* Plotgardener: cultivating precise multi-panel figures in R. *Bioinformatics* **38**, 2042-2045, doi:10.1093/bioinformatics/btac057 (2022).
76 Open2C *et al.* Cooltools: Enabling high-resolution Hi-C analysis in Python. *PLoS Comput Biol* **20**, e1012067, doi:10.1371/journal.pcbi.1012067 (2024).
77 Miura, H., Poonperm, R., Takahashi, S. & Hiratani, I. Practical Analysis of Hi-C Data: Generating A/B Compartment Profiles. *Methods Mol Biol* **1861**, 221-245, doi:10.1007/978-1-4939-8766-5_16 (2018).
78 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
79 Wang, J. & Nakato, R. HiC1Dmetrics: framework to extract various one-dimensional features from chromosome structure data. *Brief Bioinform* **23**, doi:10.1093/bib/bbab509 (2022).
80 Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* **36**, 2980-2985, doi:10.1093/bioinformatics/btaa073 (2020).
81 Xu, W. *et al.* CoolBox: a flexible toolkit for visual analysis of genomics data. *BMC Bioinformatics* **22**, 489, doi:10.1186/s12859-021-04408-w (2021).
82 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).
83 Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915, doi:10.1038/s41587-019-0201-4 (2019).
84 Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* **33**, 290-295, doi:10.1038/nbt.3122 (2015).
85 Pertea, G. & Pertea, M. GFF Utilities: GffRead and GffCompare. *F1000Res* **9**, doi:10.12688/f1000research.23297.2 (2020).

86 Frazee, A. C. *et al.* Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol* **33**, 243-246, doi:10.1038/nbt.3172 (2015).

87 Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep* **9**, 9354, doi:10.1038/s41598-019-45839-z (2019).

88 Grandi, F. C., Modi, H., Kampman, L. & Corces, M. R. Chromatin accessibility profiling by ATAC-seq. *Nat Protoc* **17**, 1518-1552, doi:10.1038/s41596-022-00692-9 (2022).

89 Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849, doi:10.1093/bioinformatics/btw313 (2016).

Figure legends

Figure 1. Anthracyclines redistribute TOP2A at chromatin looping anchors

(a) Snapshot of TOP2A ChIP-seq data from untreated (Ctr), Amr-, Acla- and Daun-treated K562 cells. ChIP-seq against CTCF, cohesin subunits RAD21 and SMC3, topologically associating domains (TADs) in K562 cells and gene annotation are shown.

(b) TOP2A ChIP-seq profile around chromatin loop anchors from untreated (Ctr), Amr-, Acla- and Daun-treated K562 cells. Chromatin looping anchors were defined as CTCF and cohesin subunit RAD21 co-binding sites from ChIP-seq experiments. CTCF motifs in chromatin loop anchors were identified by FIMO with the threshold of *P*-value < 10^{-5} and normalized TOP2A ChIP-seq read signals were shown within 1 kbp upstream (+) and downstream (-) from the CTCF motif center.

(c) Association of TOP2A peaks from untreated (Ctr) K562 cells and differential TOP2A peaks from Acla- and Daun-treated K562 cells with CTCF ChIP-seq peaks. The cohesin subunits RAD21 and SMC3 were used as positive control for CTCF association. The *y*-axis indicates the significance of the enrichment of TOP2A within CTCF peaks. The *P*-values were calculated by permutation test (*n*=50,000) and multiple testing correction by the Benjamini-Hochberg approach. The red horizontal line indicates the cutoff of *P*-value 0.05. Circle size indicates the ratio of CTCF peaks covered by the respective TOP2A peaks (scale, 0-1).

(d) TF motif enrichment in lost TOP2A peaks from Acla- and Daun-treated K562 cells calculated by HOMER. The ranking order of enriched TF motifs was shown in the top panel. The motif and *P*-value of the top 5 enriched motifs were shown in the bottom panel.

Figure 2. Anthracyclines affect CTCF foot-printing on chromatin

(a) ATAC-seq fragments size distribution from untreated (Ctr), Amr-, Acla- and Daun-treated K562 cells.

(b) Normalized ATAC-seq fragment enrichment from untreated (Ctr), Amr-, Acla- and Daun-treated K562 cells in seven classes of chromatin states defined in K562 cells. CTCF: CTCF enriched element, TSS: Predicted promoter region including transcription start site, Weak enhancer: Predicted weak enhancer or open chromatin *cis*-regulatory element, Enhancer: Predicted enhancer, Promoter flanking: Predicted promoter flanking region, Transcribed: Predicted transcribed region, Repressed: Predicted repressed or low activity region. The Tn5 insertion size ranging from 50 bp to 750 bp were divided into 100 bins and the percent of insertions in each bin were normalized to the maximal percent within each state. The enrichment of Tn5 insertions in different-size bins was calculated by dividing the normalized percent of insertions for each state by the genome-wide normalized percent of insertions. scale: -4 to +4.

(c) Nucleosome occupancy analysis around chromatin looping anchors from untreated (Ctr), Amr-, Acla- and Daun-treated K562 cells. Chromatin looping anchors were defined as CTCF and cohesin subunit RAD21 co-binding sites based on ChIP-seq data. CTCF motifs in chromatin looping anchors were identified by FIMO with the threshold of *P*-value<10^{-5} and nucleosome occupancy was shown within 1 kbp upstream (+) and downstream (-) from the CTCF motif center.

(d) TF footprinting analysis around CTCF motifs based on ATAC-seq data (n=16,814) from untreated (Ctr), Amr-, Acla- and Daun-treated K562 cells. The normalized Tn5 cleavage frequency was shown within 100 bp upstream (+) and downstream (-) from the CTCF motif center.

Figure 3. Anthracyclines evict CTCF from chromatin

(a) CTCF protein distribution in the cytoplasmic and chromatin fraction in untreated (Ctr) and Daun-treated K562 cells by western blot. Anti-β-actin was used as the western blot loading control.

(b) Quantification of the western blot results from 3 biological replicate experiments. **P*=0.0161, calculated by two-way ANOVA.

(c) Venn diagram of CTCF peaks called from untreated (Ctr), Acla- and Daun-treated K562 cells. Two biological replicates were performed for each condition.

(d) Snapshot of normalized CTCF ChIP-seq signals from untreated, Acla- and Daun-treated K562 cells at *RXRA* and *EPS15L1* loci. ENCODE ChIP-seq data against cohesin subunits RAD21 and SMC3, topologically associating domains (TADs) in K562 cells, and gene annotation are shown.

(e) Density plot of CTCF ChIP-seq signals at cohesin RAD21 binding sites, TAD boundaries and transcription start sites from untreated (Ctr), Acla- and Daun-treated K562 cells. The TAD boundaries were identified from published in situ Hi-C data in K562 using the diamond insulation approach at the resolution of 5 Kbp. The local minima of the prominence track for log2 (insulation score) above 0.5 threshold are defined as strong boundaries, and for log2 (insulation score) above 0.2 and below 0.5 threshold are defined as weak boundaries. The TSSs of genes were categorized into high-, middle- and low-expression classes based on the K562 RNA-seq data.

(f) TOP2A and CTCF redistribution at the TOP2A-CTCF interaction sites upon Acla and Daun treatment. The TOP2A-CTCF interaction sites were annotated as the TOP2A and CTCF co-binding regions based on the ChIP-seq data from untreated K562 cells. The normalized TOP2A (upper panel) and CTCF (lower panel) ChIP-seq signals were shown within ± 2 kbp surrounding the TOP2A-CTCF interaction peak center.

Figure 4. Anthracyclines target distinct genomic patterns of CTCF binding sites

(a) Distribution of differential CTCF peaks (gained, lost and stable) in Acla- and Daun-treated cells compared with untreated cells.

(b) FIMO scores at CTCF motifs detected in differential CTCF binding sites in K562 cells. Left: Acla-treatment. Right: Daun-treatment. The CTCF motifs were detected using FIMO with the threshold of *P*-value < 10^{-5}. *P*-values were calculated using the two-sided Wilcoxon rank-sum test.

(c) Core motif patterns in differential CTCF binding sites (gained, lost and stable) in Daun-treated K562 cells. The PWM matrix of core motifs within the different CTCF binding categories matched with the classic CTCF motif was calculated and visualized by FIMO. The two CpG sites (C2 and C12) within the core motif of CTCF were indicated by red dots.

(d) The original DNA methylation level of two key CpG sites (C2 and C12) within the core motif of Acla-induced differential CTCF binding sites (gained, lost and stable) in K562 cells. *P*-values were calculated using the two-sided Wilcoxon rank-sum test.

(e) The original nucleosome occupancy profiles of Acla-induced differential CTCF binding sites (gained, lost and stable) based on ENCODE MNase-seq data. The MNase-seq signal was shown within ± 1.5 kbp from differential CTCF peak centers.

(f) The original TOP2A occupancy profiles around Acla-induced differential CTCF binding sites (gained, lost and stable) based on the TOP2A ChIP-seq data in K562 cells. The TOP2A ChIP-seq signal was shown within ± 1.5 kbp from differential CTCF peak centers.

Figure 5. Anthracyclines impact 3D genome organization

(a) Snapshot of Hi-C maps. From top to bottom, Hi-C matrices of untreated (Ctr), Acla- and Daun-treated K562 cells.

(b) Decay curves of chromatin interaction frequency with distance in untreated (Ctr), Acla- and Daun-treated K562 cells. The *y*-axis represents the logarithm scale of contact frequency from respective Hi-C data and *x*-axis represents logarithm scale of the genomic distance between loop anchors (bp).

(c) A/B compartment switching in Acla- and Daun-treated K562 cells. Percentage of switched compartment status (A-to-B or B-to-A) and the stable A/B compartment states (A-to-A or B-to-B) were shown. KR normalized Hi-C matrices with the resolution of 100 Kbp were used to identify compartment states. The first eigenvector value was used to separate A (positive) or B (negative) compartment states.

(d) Saddle plot of separation strength between the A and B compartments. From left to right, saddle plot of untreated (Control), Acla- and Daun-treated K562 cells. The average contact frequencies between 100 Kbp bins are normalized by expected contact frequency based on genomic distance. Bins are sorted by their first eigenvector value and divided into 100 quantiles. The red area in the upper left and lower right corners indicate enriched B-B and A-A interactions respectively. The blue area in the upper right and lower left corners indicate depleted B-A or A-B interactions. The numbers in the corners represent the relative strength of A-A interactions compared to A–B interactions (lower right) and the relative strength of B-B interactions compared to B–A interactions (upper left).

(e) Snapshot of insulation score signal tracks called from KR balanced Hi-C matrices with 5 Kbp bin size and 100 Kbp window size using diamond-insulation algorithm. Upper panel: Hi-C contact map at the chr5:111,000,000-116,000,000 locus. Lower panel: insulation profiles of untreated (Control), Acla- and Daun-treated K562 cells in the same region as indicated by the respective colors. The insulation score reflects contact frequencies between upstream and downstream loci. The bins with the lowest insulation scores locally (local minima) are defined as boundaries.

(f) Changes of insulation strength at TAD boundaries from Acla- and Daun-treated K562 cells. Four classes of TAD boundaries were shown: lost (failed to identify in drug-treated Hi-C matrix), weakened (control-drug boundary score >= 0.4 and control-drug insulation score <= -0.1), stable and enhanced (control-drug boundary score <= -0.4 and control-drug insulation score >= 0.1). Genome-wide insulation scores and boundary scores were calculated by KR balanced Hi-C matrices with 5 Kbp bin size and 100 Kbp window size using diamond-insulation algorithm.

(g) Scatterplot of boundary strength scores of residual TAD boundaries identified from Hi-C matrices from untreated (*x* axis) and Acla-treated (*y* axis). The regression line was plotted in the solid line. The dashed diagonal line was plotted as the reference. The color indicates the spatial density of nearby points.

Figure 6. Anthracyclines affect CTCF-mediated *cis*-regulatory loops

(a) Aggregate peak analysis of rewired chromatin loops in Daun-treated K562 cells (gained, lost and stable) within ± 50 Kbp window. Differential loops were determined by diffPeakachu when the fold-change of probability score deviates from the Gaussian mixture model with a *P*-value < 0.05.

(b) Epigenomic signatures of lost loops in Daun-treated K562 cells. Upper panel: pie charts of Daun-induced differential (gained, lost and stable) ATAC-seq peaks (left) and CTCF ChIP-seq peaks (right) within the lost looping anchors. Lower panel: motif enrichment in Daun-induced differential ATAC-seq peaks (from top to bottom: lost, stable and gained). *P*-values were calculated using hypergeometric distribution compared with background peaks and significantly enriched TF motifs were determined with the cutoff *p*-value below 10^{-10}.

(c) From top to bottom, Hi-C contact profiles from untreated (Ctr), Acla-treated and Daun-treated K562 cells at *Myc* TAD locus. The *Myc* enhancer-promoter loops are indicated by the red dots and enhancer-enhancer loops are indicated by the red rectangles.

(d) Virtual 4C analysis from the Hi-C data with the viewpoint at the *Myc* gene locus. The interactions between the *Myc* promoter and 7 validated *Myc* enhancers are indicated by red arrows. Normalized CTCF ChIP-seq and virtual 4C signal tracks from untreated (Ctr), Acla- and Daun-treated K562 cells are shown.

(e) Aggregate peak analysis of CTCF-CTCF loops within ± 50 Kbp window are shown for untreated (Ctr), Acla- and Daun-treated cells. The CTCF-CTCF loops were defined as looping anchors called from Hi-C that overlap with the CTCF peaks from ChIP-seq.

Figure 7. Daun-induced 3D genome alterations in AML correlates with therapeutic response

(a) Heatmap of chromatin accessibility signals in AML-patient-specific Daun-suppressed ATAC peak regions. The distal binarization method was used to identify patient-specific Daun-responsive ATAC peaks. The limma eBayes test was performed to determine the statistical significance of differential ATAC peaks

between patient and treatment groups. The threshold of FDR below 0.05 was used to capture the group-specific significant peaks. The peaks that were significantly suppressed either in the R/R AML group or in the CR AML group are shown.

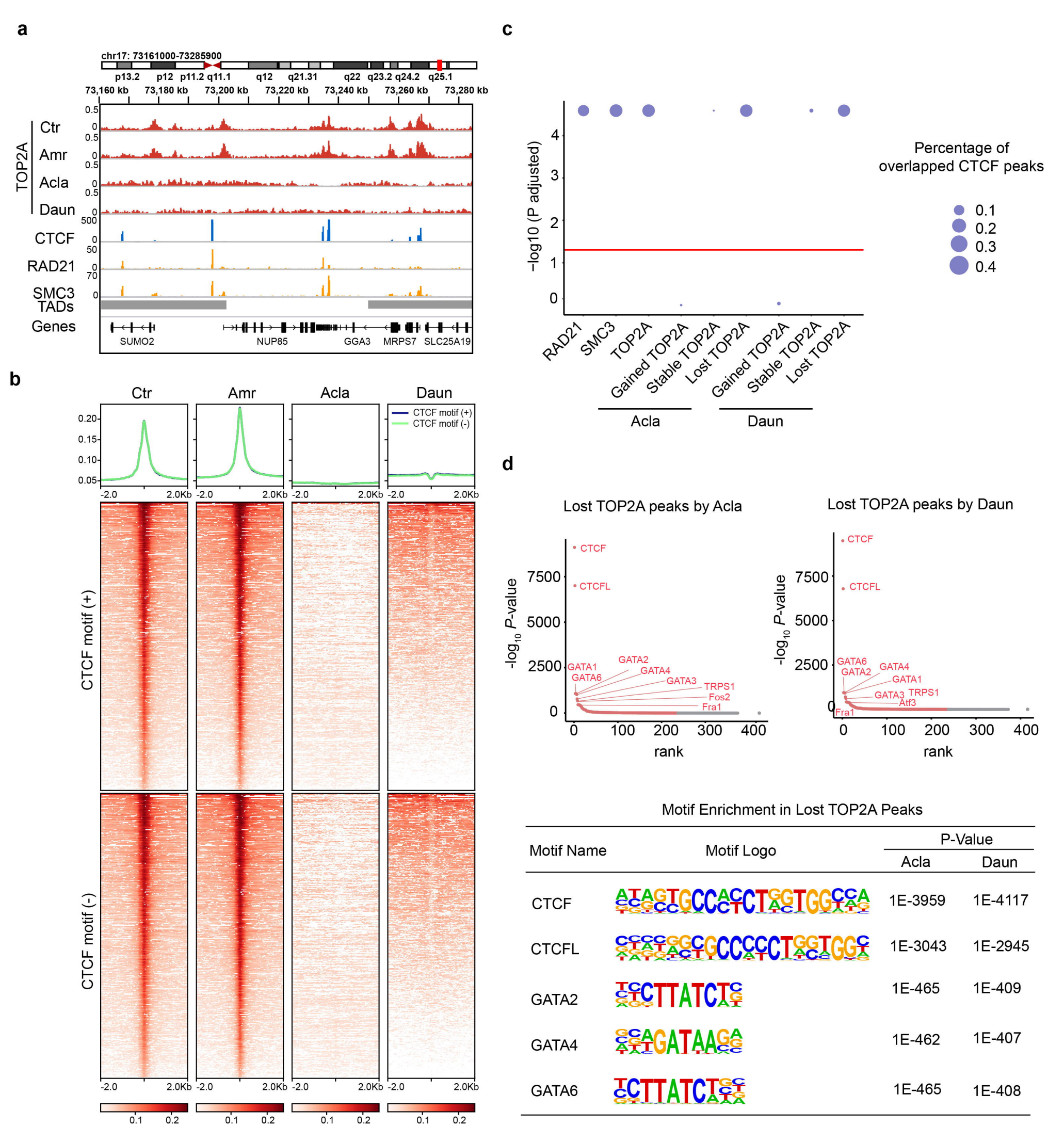
(b) TF motif enrichment in patient-subtype-specific Daun-suppressed ATAC peaks. From top to bottom: significantly enriched TF motifs in either shared, R/R AML- or CR AML-specific Daun-suppressed ATAC peak sets are shown. *P*-values were calculated using hypergeometric distribution compared with background peaks and significantly enriched TF motifs were determined with the cutoff *p*-value below 10^{-10}.

(c) Predicted peak-to-gene links regulating the *TRIM32* gene. All the potential interactions between *TRIM32* gene and ATAC peaks within the ± 0.5 Mbp window relative to the TSS of the *TRIM32* gene were tested. The Pearson correlation coefficients were calculated to correlate the ATAC-seq accessibility of each peak with the RNA-seq expression of *TRIM32* gene using 20 paired ATAC-seq and RNA-seq patient samples. The statistical significance was assessed by comparing with the null distribution of correlations between trans-peak-to-gene links. The significant peak-to-gene links were determined using the cutoff *p*-value below 0.05. Arch height of peak-to-gene links represents the Pearson correlation coefficients. The color indicates the $\log_{10}$ (*p*-value).

(d) Scatterplot of peak-to-gene links predictive for *TRIM32* gene expression based on the ATAC-seq accessibility and RNA-seq gene expression from the individual patients. The Pearson correlation coefficient (PCC) and *p*-value are included. The color indicates the different patient-treatment groups. Each dot represents an individual sample.

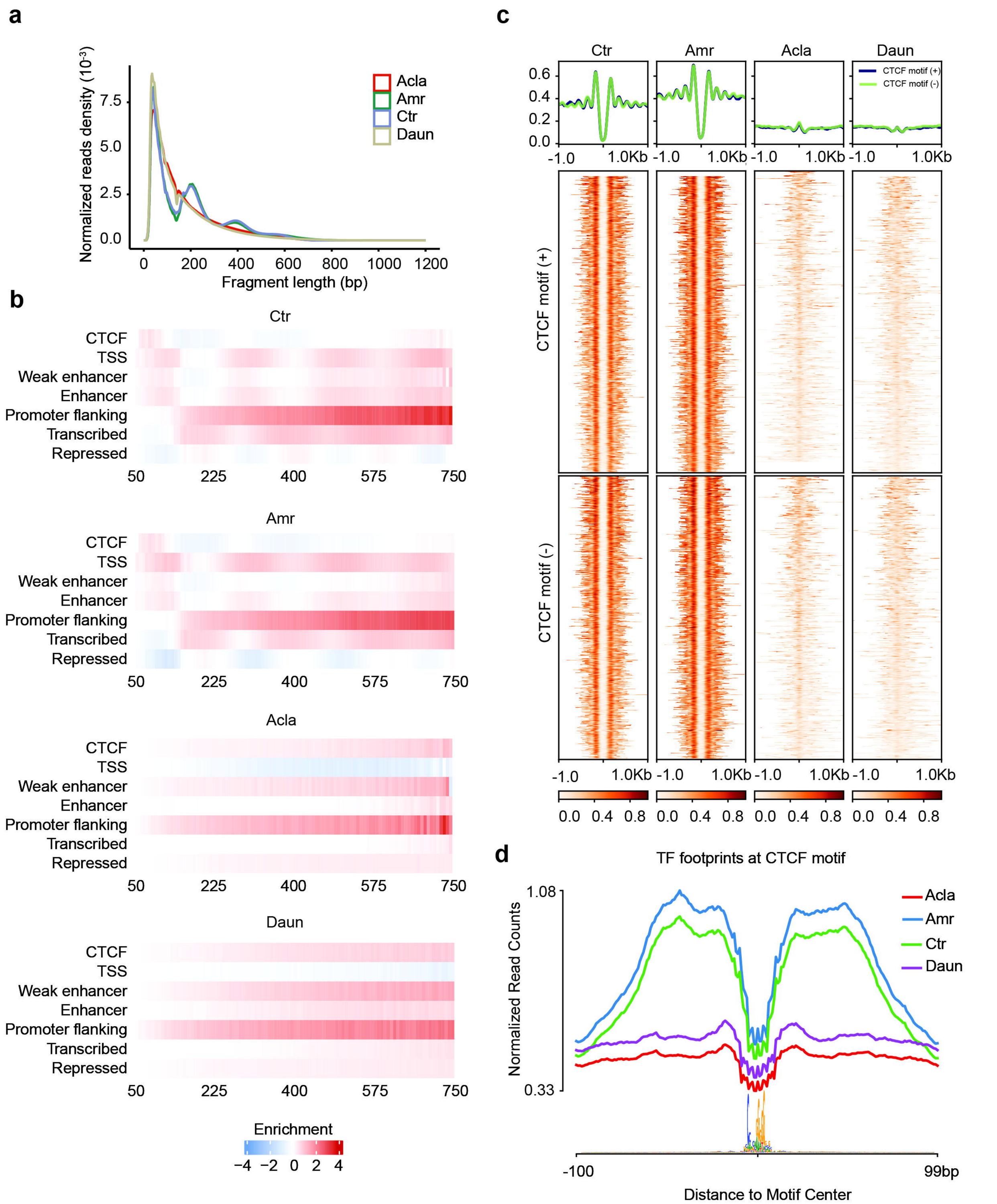
(e) ATAC-seq accessibility within the CTCF motif-containing peaks in different patient groups. The selected peak is predicted to regulate *TRIM32* gene expression by the peak-to-gene link analysis. The signals from each group was normalized to their respective basal mean of ATAC-seq accessibility. The p-values were computed by two-tailed Student's t-test. The color indicates the different patient-treatment groups. Each dot represents an individual sample.

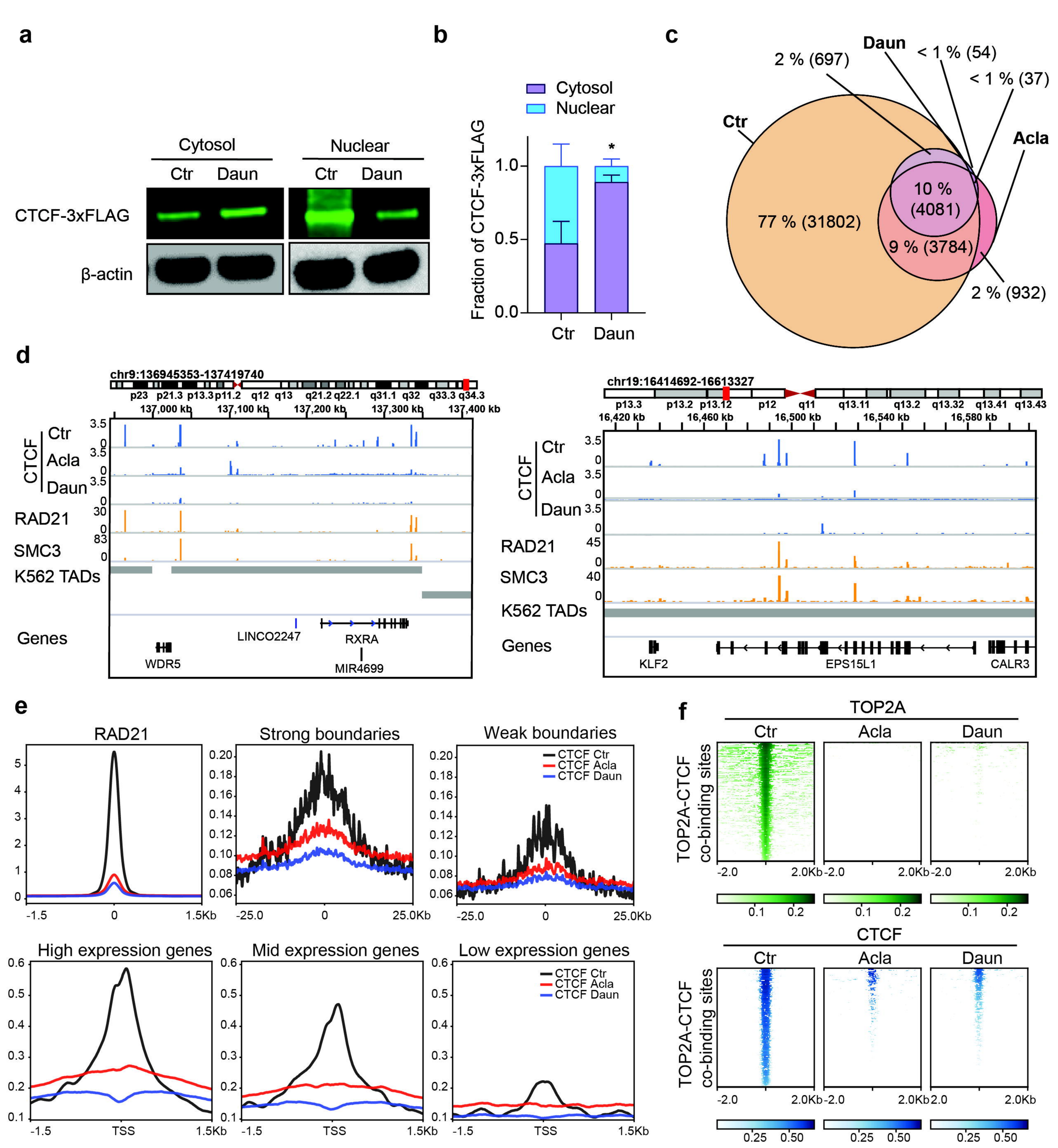
(f) Gene expression changes measured by qPCR. K562 cells with CRISPRi targeting CTCF-motif-containing ATAC-seq peak that are predicted to regulate the *TRIM32* gene in AML patients. NT: K562 cells with dCas9-KRAB and non-targeting RNA, L1: K562 cells with dCas9-KRAB and guide RNA targeting the CTCF-motif-containing peak. Error bars represent the standard deviation of 3 biological replicates. *$P<0.05$ calculated by two-tailed Student's t-test.

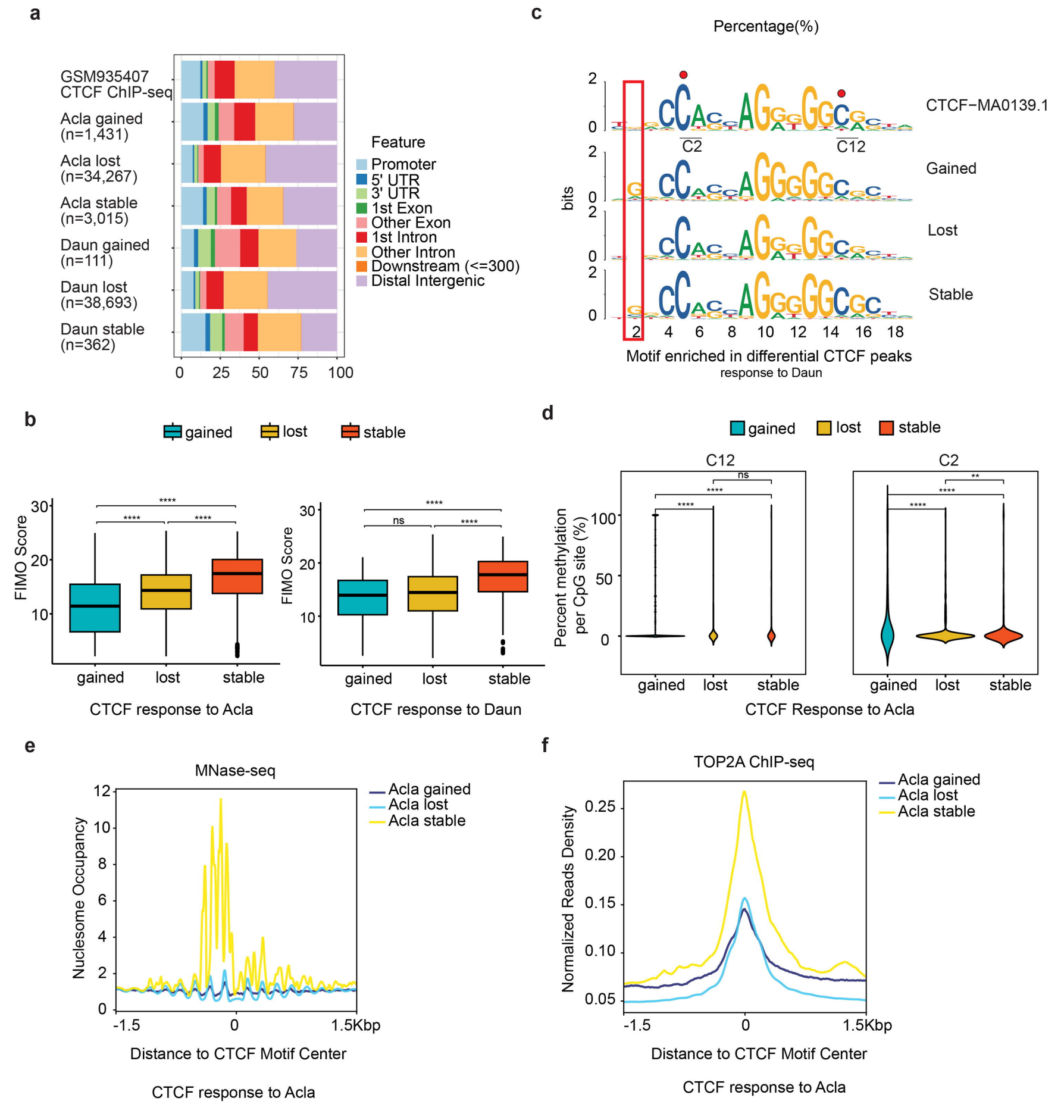


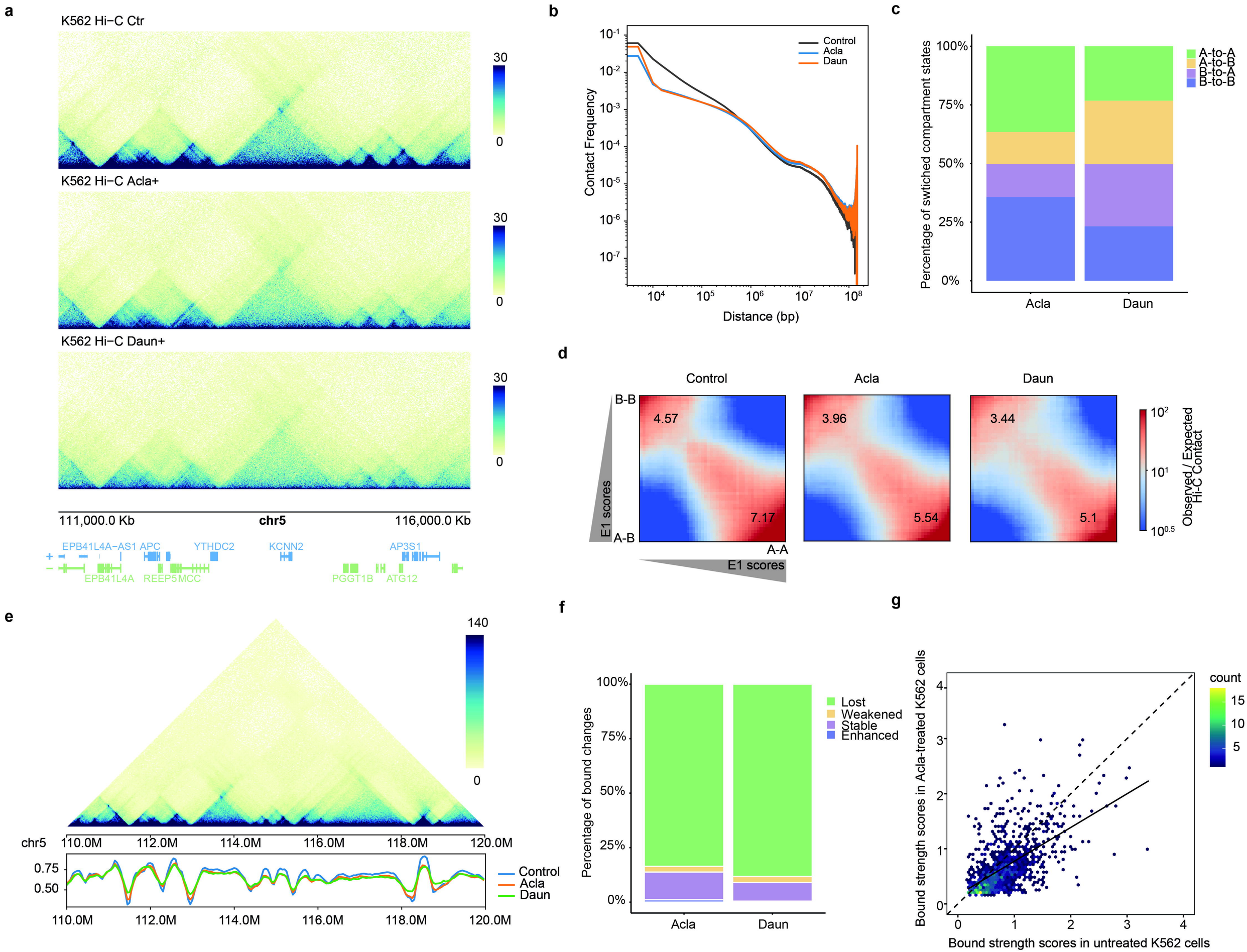
Motif Enrichment in Lost TOP2A Peaks

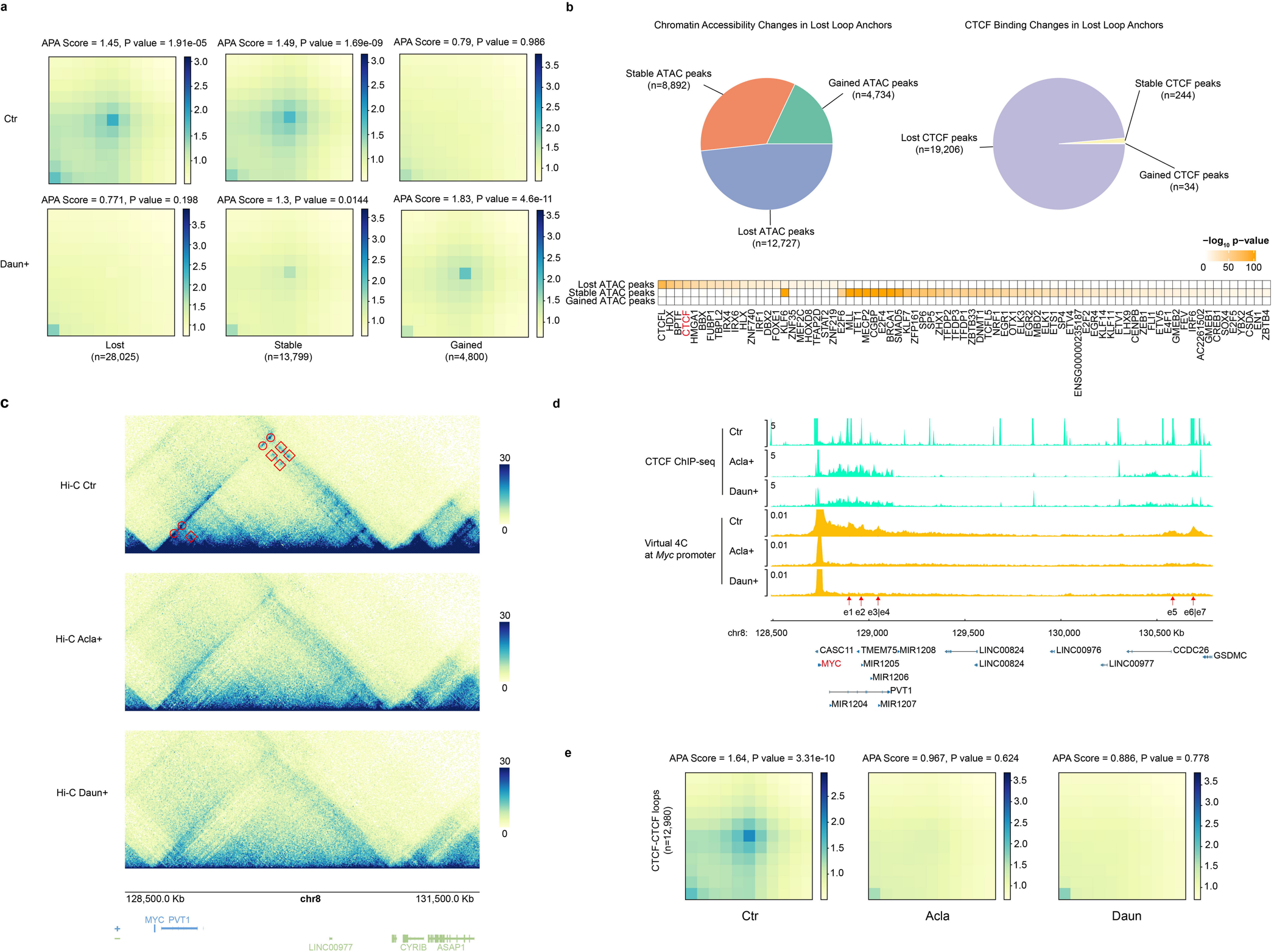
Motif Name	Motif Logo	P-Value	
		Acla	Daun
CTCF		1E-3959	1E-4117
CTCFL		1E-3043	1E-2945
GATA2		1E-465	1E-409
GATA4		1E-462	1E-407
GATA6		1E-465	1E-408

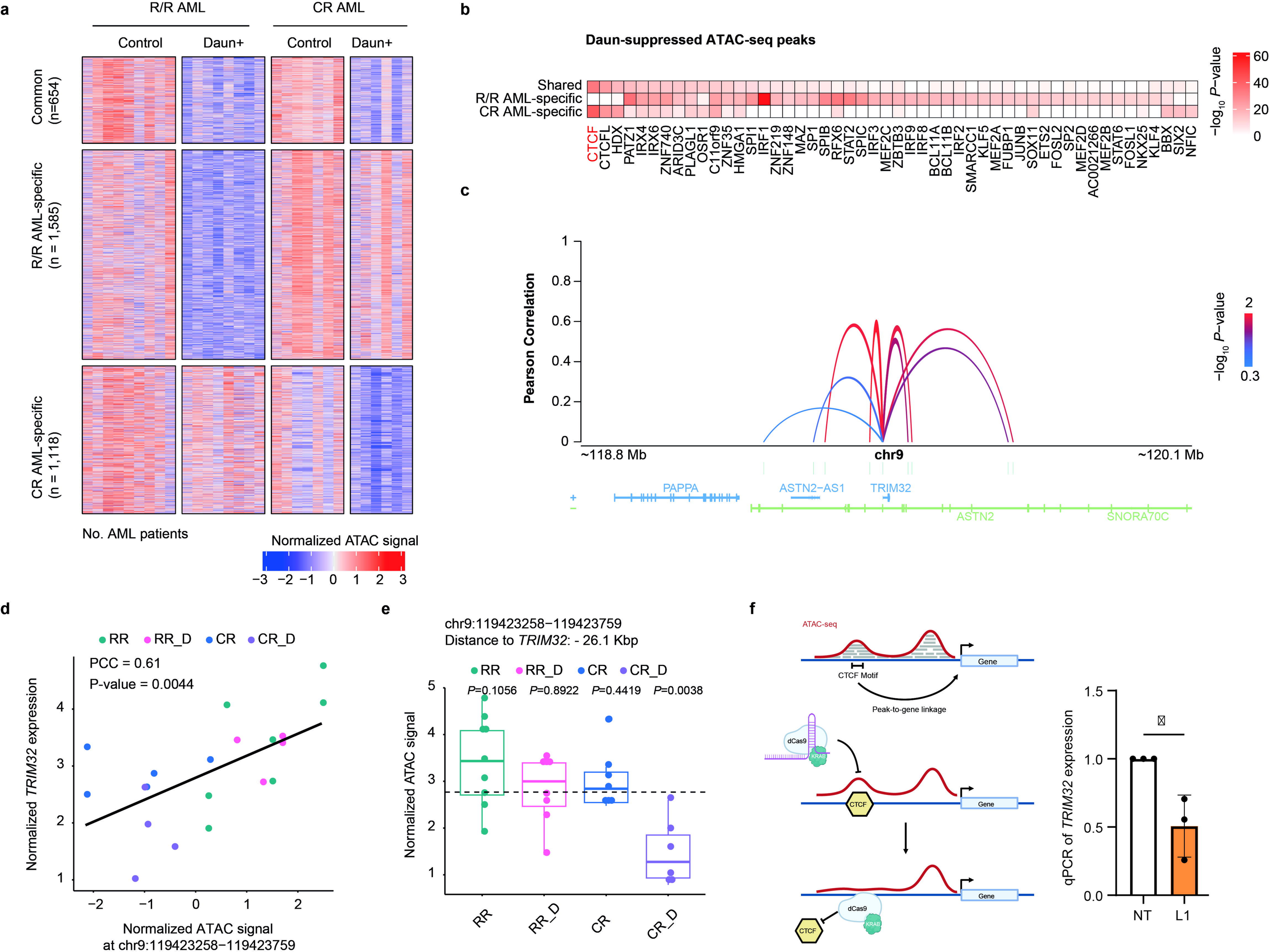

a
Normalized reads density (10⁻³)
Fragment length (bp)
Acla
Amr
Ctr
Daun
b
Ctr
Amr
Acla
Daun
CTCF
TSS
Weak enhancer
Enhancer
Promoter flanking
Transcribed
Repressed
Enrichment
c
CTCF motif (+)
CTCF motif (-)
1.0Kb
d
TF footprints at CTCF motif
Normalized Read Counts
Distance to Motif Center
-100
99bp


a
Cytosol
Nuclear
Ctr
Daun
Ctr
Daun
CTCF-3xFLAG
β-actin
b
Cytosol
Nuclear
Fraction of CTCF-3xFLAG
1.0
0.5
0.0
*
Ctr
Daun
c
Daun
< 1 % (54)
2 % (697)
< 1 % (37)
Ctr
Acla
10 %
(4081)
77 % (31802)
9 % (3784)
2 % (932)
d
chr9:136945353-137419740
p23 p21.3 p13.3 p11.2 q12 q13 q21.2 q22.1 q31.1 q32 q33.3 q34.3
137,000 kb 137,100 kb 137,200 kb 137,300 kb 137,400 kb
CTCF
Ctr
Acla
Daun
RAD21
SMC3
K562 TADs
Genes
LINCO2247
RXRA
WDR5
MIR4699
chr19:16414692-16613327
p13.3 p13.2 p13.12 p12 q11 q13.11 q13.2 q13.32 q13.41 q13.43
16,420 kb 16,460 kb 16,500 kb 16,540 kb 16,580 kb
CTCF
Ctr
Acla
Daun
RAD21
SMC3
K562 TADs
Genes
KLF2
EPS15L1
CALR3
e
RAD21
Strong boundaries
Weak boundaries
CTCF Ctr
CTCF Acla
CTCF Daun
-1.5
0
1.5Kb
-25.0
0
25.0Kb
High expression genes
Mid expression genes
Low expression genes
TSS
f
TOP2A
Ctr
Acla
Daun
TOP2A-CTCF co-binding sites
-2.0
2.0Kb
0.1
0.2
CTCF
0.25
0.50


a
GSM935407 CTCF ChIP-seq
Acla gained (n=1,431)
Acla lost (n=34,267)
Acla stable (n=3,015)
Daun gained (n=111)
Daun lost (n=38,693)
Daun stable (n=362)
0 25 50 75 100
Feature
Promoter
5' UTR
3' UTR
1st Exon
Other Exon
1st Intron
Other Intron
Downstream (<=300)
Distal Intergenic
c
Percentage(%)
bits
CTCF−MA0139.1
Gained
Lost
Stable
C2
C12
2 4 6 8 10 12 14 16 18
Motif enriched in differential CTCF peaks
response to Daun
b
gained lost stable
FIMO Score
CTCF response to Acla
CTCF response to Daun
d
gained lost stable
C12
C2
Percent methylation per CpG site (%)
CTCF Response to Acla
e
MNase-seq
Acla gained
Acla lost
Acla stable
Nucleosome Occupancy
-1.5 0 1.5Kbp
Distance to CTCF Motif Center
CTCF response to Acla
f
TOP2A ChIP-seq
Acla gained
Acla lost
Acla stable
Normalized Reads Density
-1.5 0 1.5Kbp
Distance to CTCF Motif Center
CTCF response to Acla


a
K562 Hi−C Ctr
K562 Hi−C Acla+
K562 Hi−C Daun+
30
0
111,000.0 Kb
chr5
116,000.0 Kb
EPB41L4A−AS1
APC
YTHDC2
KCNN2
AP3S1
EPB41L4A
REEP5
MCC
PGGT1B
ATG12
b
Control
Acla
Daun
Contact Frequency
Distance (bp)
c
Percentage of swtiched compartment states
100%
75%
50%
25%
0%
A-to-A
A-to-B
B-to-A
B-to-B
Acla
Daun
d
Control
Acla
Daun
B-B
A-B
A-A
E1 scores
4.57
7.17
3.96
5.54
3.44
5.1
Observed / Expected Hi-C Contact
e
140
0
chr5
110.0M
112.0M
114.0M
116.0M
118.0M
120.0M
0.75
0.50
Control
Acla
Daun
f
Percentage of bound changes
Lost
Weakened
Stable
Enhanced
g
Bound strength scores in Acla-treated K562 cells
Bound strength scores in untreated K562 cells
count
15
10
5


a
APA Score = 1.45, P value = 1.91e-05
APA Score = 1.49, P value = 1.69e-09
APA Score = 0.79, P value = 0.986
Ctr
APA Score = 0.771, P value = 0.198
APA Score = 1.3, P value = 0.0144
APA Score = 1.83, P value = 4.6e-11
Daun+
Lost (n=28,025)
Stable (n=13,799)
Gained (n=4,800)
b
Chromatin Accessibility Changes in Lost Loop Anchors
Stable ATAC peaks (n=8,892)
Gained ATAC peaks (n=4,734)
Lost ATAC peaks (n=12,727)
CTCF Binding Changes in Lost Loop Anchors
Stable CTCF peaks (n=244)
Lost CTCF peaks (n=19,206)
Gained CTCF peaks (n=34)
−log10 p−value
Lost ATAC peaks
Stable ATAC peaks
Gained ATAC peaks
c
Hi-C Ctr
Hi-C Acla+
Hi-C Daun+
128,500.0 Kb
chr8
131,500.0 Kb
MYC PVT1
LINC00977
CYRIB ASAP1
d
CTCF ChIP-seq
Ctr
Acla+
Daun+
Virtual 4C at Myc promoter
e1 e2 e3|e4
e5 e6|e7
chr8: 128,500
129,000
129,500
130,000
130,500 Kb
CASC11 TMEM75 MIR1208 LINC00824 LINC00976 CCDC26 GSDMC
MYC MIR1205 LINC00824 LINC00977
MIR1206
PVT1
MIR1204 MIR1207
e
APA Score = 1.64, P value = 3.31e-10
APA Score = 0.967, P value = 0.624
APA Score = 0.886, P value = 0.778
CTCF-CTCF loops (n=12,980)
Ctr
Acla
Daun


a
R/R AML
CR AML
Control
Daun+
Control
Daun+
Common (n=654)
R/R AML-specific (n = 1,585)
CR AML-specific (n = 1,118)
No. AML patients
Normalized ATAC signal
−3 −2 −1 0 1 2 3
b
Daun-suppressed ATAC-seq peaks
Shared
R/R AML-specific
CR AML-specific
CTCF
CTCFL
HDX
PATZ1
IRX4
IRX6
ZNF740
ARID3C
PLAGL1
OSR1
C11orf9
ZNF35
HMGA1
SPI1
IRF1
ZNF219
ZNF148
MAZ
SP1
SPIB
RFX6
STAT2
SPIC
IRF3
MEF2C
ZBTB3
IRF9
IRF8
BCL11A
BCL11B
IRF2
SMARCC1
KLF5
MEF2A
FUBP1
JUNB
SOX11
ETS2
FOSL2
SP2
MEF2D
AC0021266
MEF2B
STAT6
FOSL1
NKX25
KLF4
BBX
SIX2
NFIC
−log10 P−value
60
40
20
0
c
Pearson Correlation
1
0.8
0.6
0.4
0.2
0
~118.8 Mb
chr9
~120.1 Mb
PAPPA
ASTN2−AS1
TRIM32
ASTN2
SNORA70C
−log10 P−value
2
0.3
d
RR
RR_D
CR
CR_D
PCC = 0.61
P-value = 0.0044
Normalized TRIM32 expression
Normalized ATAC signal
at chr9:119423258−119423759
e
chr9:119423258−119423759
Distance to TRIM32: - 26.1 Kbp
RR
RR_D
CR
CR_D
P=0.1056
P=0.8922
P=0.4419
P=0.0038
Normalized ATAC signal
f
ATAC-seq
Gene
CTCF Motif
Peak-to-gene linkage
dCas9
KRAB
CTCF
qPCR of TRIM32 expression
1.5
1.0
0.5
0.0
NT
L1