

# Multiomic analysis of cohesin reveals that ZBTB transcription factors contribute to chromatin interactions

Rui Wang<sup>†</sup>, Qiqin Xu<sup>†</sup>, Chenlu Wang, Kai Tian, Hui Wang and Xiong Ji<sup>✉\*</sup>

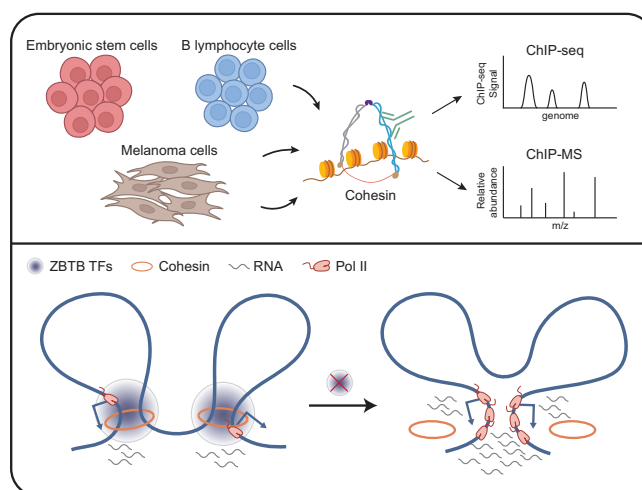
Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, School of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

Received December 15, 2022; Editorial Decision April 22, 2023; Accepted May 23, 2023

## ABSTRACT

One bottleneck in understanding the principles of 3D chromatin structures is caused by the paucity of known regulators. Cohesin is essential for 3D chromatin organization, and its interacting partners are candidate regulators. Here, we performed proteomic profiling of the cohesin in chromatin and identified transcription factors, RNA-binding proteins and chromatin regulators associated with cohesin. Acute protein degradation followed by time-series genomic binding quantitation and BAT Hi-C analysis were conducted, and the results showed that the transcription factor ZBTB21 contributes to cohesin chromatin binding, 3D chromatin interactions and transcriptional repression. Strikingly, multiomic analyses revealed that the other four ZBTB factors interacted with cohesin, and double degradation of ZBTB21 and ZBTB7B led to a further decrease in cohesin chromatin occupancy. We propose that multiple ZBTB transcription factors orchestrate the chromatin binding of cohesin to regulate chromatin interactions, and we provide a catalog of many additional proteins associated with cohesin that warrant further investigation.

## GRAPHICAL ABSTRACT



## INTRODUCTION

Chromatin structures are partitioned into hierarchical units, including chromosome territories, A/B compartments, topologically associating domains (TADs), enhancer–promoter loops, nucleosome clutches, globular chromatin domains and nucleosomes (1–12). Although enhancer- and promoter-mediated local chromatin architectures have been studied for several decades (13–18), the molecular basis of their establishment, maintenance and disruption remains unclear. To describe these mechanisms, the looping model, tracking model and phase separation-based compartmentalization model have been introduced (19–22). Many different chromatin structural regulators have been depleted in different studies, resulting in different effects on 3D genome organization (23–27). The dissection of these molecular mechanisms is crucial to understanding the regulation of the 3D genome during development and disease.

\*To whom correspondence should be addressed. Tel: +86 010 62756191; Email: [xiongji@pku.edu.cn](mailto:xiongji@pku.edu.cn)

<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Structural maintenance of the chromosome (SMC) complexes (cohesin, condensin and SMC5/SMC6) are crucial for the organization of interphase chromatin structures (28–30). Previous genomic distribution mapping efforts suggested that cohesin occupies promoters, enhancers and insulators (31–33). Cohesin is loaded by NIPBL and MAU2, stabilized by CTCF at insulators, and removed by WAPL (34–36). These direct regulators of cohesin have been shown to regulate global 3D genome organization (37–40). Previous studies have identified multiple transcription factors involved in 3D genome organization that interact directly or indirectly with cohesin. These factors, including CTCF, YY1, ZNF143, ADNP, MAZ, WIZ and BHLHE40 (41–51), usually occupy insulators or gene promoters, and their perturbation modulates the global 3D chromatin organization or affects a subset of chromatin interactions. We hypothesized that there are additional transcription factors involved in cohesin-mediated chromatin interactions. Therefore, the comprehensive identification of cohesin-interacting partners will provide a valuable resource for the identification of candidate regulators of 3D genome organization and will support the idea that the redundancy between cohesin and various transcription factors modulates local chromatin interactions or specific interactions under different conditions.

The ZBTB transcription factors Mod (mdg4) and Cp190 are known to play critical roles in the regulation of insulator activities in *Drosophila* (52–56). The zinc finger domain interacts with specific DNA sequences, and the BTB domain mediates homo- and heteromeric interactions (57,58). These properties might be responsible for the ability of ZBTB transcription factors to mediate 3D chromatin structures. There are >40 ZBTB transcription factors in mammals, and many of them have been reported to function in T-cell and B-cell lineage specification during development and diseases (59–62). The major focus was previously on their roles in transcription regulation. However, whether and how they regulate 3D chromatin structures remain unknown.

Here, we report that ZBTB21 facilitates the chromatin occupancy of cohesin and modulates 3D chromatin interactions and transcriptional repression. A number of ZBTB transcription factors (ZBTB7A, ZBTB7B, ZBTB11 and ZBTB35) have been shown to interact with cohesin. The double depletion of ZBTB21 and ZBTB7B further decreases the chromatin occupancy of cohesin. These results imply that various ZBTB transcription factors facilitate cohesin chromatin binding in addition to performing transcriptional regulatory activities that modulate chromatin interactions in mammalian cells.

## MATERIALS AND METHODS

### Cell culture

Mouse embryonic stem cells (mESCs) (obtained from Richard Young's lab, MIT) were cultured in knockout Dulbecco's modified Eagle's medium (DMEM; Gibco; 10829018-018), supplemented with 15% fetal bovine serum (FBS), PenStrep (Gibco, 15140-122), nucleoside (Sigma, ES-008), L-glutamine (Gibco, 25030-149), NEAAs (non-essential amino acids; Milipore, TMS-001-C),

CHIR99021 (Selleck, S1036), PD0325901 (Selleck, S1263),  $\beta$ -mercaptoethanol (Sigma, M3148) and mLIF (mouse leukemia inhibitory factor; Milipore, ESG1107) at 37°C with 5% CO<sub>2</sub>. To degrade the target proteins, we first treated the degon-tagged mESCs with 1  $\mu$ g/ml doxycycline (Sigma Aldrich, D9891) for 24 h, then added 500  $\mu$ M indole acetic acid (IAA; Sigma Aldrich, 115148) for the indicated times for different analyses. HEK293T (ATCC, CRL-3216) cells were cultured in DMEM (Gibco, 11995-065) supplemented with 10% FBS (Gibco, 10099-141) and PenStrep (Gibco, 15140-122) at 37°C with 5% CO<sub>2</sub>.

### Plasmid construction

The donors and single guide RNAs (sgRNAs) for the degon-tagged cells were generated as in the previous report (63). The homology arms were amplified from mouse genomic DNA with a length of 400 bp upstream or downstream of the targeted site. The homology arms were then ligated with mouse auxin-inducible degon (mAID)-green fluorescent protein (GFP) donor cassettes by Gibson assembly (NEB, E2611S). The ligated constructs were used to target the sequence near the TAG stop codon of the protein of interest by CRISPR/Cas9 gene editing to generate degon cell lines. The plasmid pMK243 (Addgene, 72835) was digested by restriction enzymes BglII and MluI (NEB, R0144S and R0198S). The coding sequences of the target protein and the GFP open reading frame (ORF) were ligated with digested pMK243 by Gibson assembly. The correct construct was used to generate Tet-on inducible expression of ZBTB stable HEK293T cell lines. For primers, see Supplementary Table S9.

### Stable cell line generation

The degon donors and corresponding sgRNAs were transfected into the mES-OsTIR1 cell line generated in a previous study (64). The transfected cells were selected with neomycin (Gibco, 21810-031), and homogenous clones were obtained after the validation by genotyping and Sanger sequencing. The HEK293T cell lines expressing a doxycycline-inducible GFP-tagged ZBTB proteins were generated by inserting the expression ORF into the AAVS1 locus by CRISPR/Cas9 gene editing. The clones were validated by genotyping and Sanger sequencing. In order to not express the ZBTB protein to very high levels, ZBTB-GFP heterozygous clones were selected for downstream analyses.

### Chromatin immunoprecipitation (ChIP)

The ChIP experiment was performed based on a previous publication (65). Briefly, the 1% formaldehyde-fixed cells were lysed by Nonidet P-40 (NP-40) lysis buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% NP-40). The cell lysates were added on top of a 1.25 ml sucrose cushion [24% sucrose (w/v) in NP-40 lysis buffer]. The nuclei were collected by centrifugation. The nuclear pellet was gently resuspended with 0.5 ml of glycerol buffer [20 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM EDTA, 0.85 mM dithiothreitol (DTT), 50% (v/v) glycerol] and 0.5 ml of nuclei lysis buffer (10 mM HEPES pH 7.6, 1 mM DTT, 7.5 mM MgCl<sub>2</sub>,

0.2 mM EDTA, 0.3 M NaCl, 1 M urea, 1% NP-40) to isolate chromatin. The chromatin pellets were resuspended in 1 ml of sonication buffer [20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA pH 8.0, 0.1% sodium dodecylsulfate (SDS), 1% Triton X-100]. Micrococcal nuclease (MNase; 40 U) was added to the chromatin extracts to digest chromatin. The chromatin was further fragmented by sonication with biorupter with the setting at the position high 30 s on/60 s off 15 times. The 5% lysate was saved as input: half for the input of protein analysis and half for the input of DNA analysis. Different antibodies were added to the soluble chromatin extracts and incubated at 4°C with rotation overnight. The next day, Protein G Dynabeads (Thermo Fisher, 10004D, #11205D) were added to the chromatin-antibody reaction and incubated for ~2 h at 4°C. Protein G beads were washed with sonification buffer, high-salt buffer, LiCl buffer and TE buffer, respectively. The ChIP products were eluted from beads by heating at 65°C for 15 min. De-cross-linked DNA was used for ChIP-seq library preparation, and eluted protein products were used for mass spectrometry (MS) analysis. High-salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA pH 8.0, 0.1% SDS, 1% Triton X-100), LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% NP-40) and TE buffer (1 mM EDTA, 10 mM Tris-HCl pH 8.0).

### BAT-Hi-C

The BAT Hi-C was performed by following the protocol published previously (66). The major difference between BAT Hi-C and *in situ* Hi-C is that BAT Hi-C employs a biotin-labeled bridge linker to mediate proximity ligation, while *in situ* Hi-C uses blunt end ligation. Briefly, the 1% formaldehyde-fixed cells were digested by AluI (NEB, #R0137S) for at least 12 h. Then, the restriction enzyme was inactivated by heat. The DNA overhangs were added with a base dA by Klenow (3'-5' exo-) enzyme [40 µl of NEB buffer 2, 8 µl of 10 mM dATP, 40 µl of 10% Triton X-100, 304 µl of H<sub>2</sub>O and 8 µl of Klenow (3'-5' exo-)]. The dA overhang DNA products were ligated with a bridge linker by T4 DNA ligase (NEB, #M0202L) at room temperature for at least 6 h. Then the chromatin fractions were isolated, and the unligated DNA was digested by lambda exonuclease. The ligated DNA was de-cross-linked at 65°C overnight. The ligated DNAs were purified, and then fragmented by sonication. The fragmented DNA was examined by 1% DNA agarose gel electrophoresis. The size of the DNA product was ~0.2–2 kb. The ligated DNA was purified by biotin-streptavidin pulldown. The DNA library was constructed with the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®. The 300–500 bp DNA products were purified by the Magen gel purification kit. The DNA was sequenced at Hiseq 4000, and 150 bp paired-end reads were obtained for downstream analyses.

### 4C-seq

The 4C-seq was described in a previous publication (67). Briefly, the cells were first treated appropriately and then cross-linked with 1% formaldehyde for 10 min at room temperature. The nuclei were isolated and treated with 25 µl of

10× cutsmart buffer and 100 U of NlaIII (NEB, #R0125S), and incubated overnight at 37°C, 900 rpm to digest the chromatin. The next day, the restriction enzyme was inactivated by heat. The proximal, digested chromatin fragments were ligated by T4 DNA ligase (NEB, M0202) by incubating the reaction for at least 6 h at room temperature. After ligation, the DNA was reverse cross-linked and extracted by an equal volume of phenol:chloroform:isoamyl alcohol (Solarbio, P1013). The purified DNA was fragmented by sonication with Biorupter at high energy 30 s on, 60 s off, for two cycles. The fragmented DNA was examined by 1% DNA agarose gel electrophoresis. The size of the DNA product was ~0.2–2 kb. To construct a 4C-seq library, a biotin-labeled primer targeting the region of interest was used for primer extension. The amplified DNA was purified by biotin-streptavidin pulldown. The biotin-labeled single-stranded DNA was ligated to an adapter by bridge linker-mediated ligation. The ligated DNA products were amplified by nested polymerase chain reaction (PCR), and 500–900 bp DNA products were purified by the Magen gel purification kit. The purified DNA was sequenced by Hiseq 4000 with 150 bp paired reads.

### Immunoprecipitation and western blotting

A total of 20 million native cells were used for immunoprecipitation experiments. A 1 ml aliquot of IP lysis buffer [20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, sodium pyrophosphate, β-glycerophosphate, EDTA, leupeptin, phenylmethylsulfonyl fluoride (PMSF) and protein inhibitor cocktail] was used to resuspend cell pellets by rotating the cell lysates at 4°C for 30 min. The supernatant (5% cell lysis supernatant was saved as an input sample) was collected. Then, different antibodies were added to the pre-cleared whole-cell lysates and incubated overnight with rotation at 4°C. The protein-antibody complexes were incubated with Protein G for ~2 h at 4°C. The beads were washed by sonification buffer four times and by high-salt buffer once. The supernatant was discarded, and a suitable volume of 1× SDS loading buffer was added. The eluted IP samples were boiled at 100°C for 5 min. The protein elution and input samples were used for western blotting analyses. Then the samples were run on different concentration SDS-polyacrylamide gel electrophoresis (PAGE) gels. They were transformed onto nitrocellulose membranes and the membranes were blocked with skim milk and then incubated with diluted primary antibody (1:1000–1:5000) overnight at 4°C. On the next day, the membranes were washed and incubated with secondary antibodies (1:5000). The membranes were washed and analyzed on a G.E AI 600 RGB imaging system.

### Live cell imaging

The cell lines that express proteins of interest with GFP tags were imaged on 35 mm glass-bottom dishes (MatTek, USA) by Nikon A1R-si. Before imaging, the cells were digested by 0.25% trypsin EDTA (Gibco, 25200-056) and transferred to the 35 mm glass-bottom dishes for ~6 h. Then the cells were imaged by a confocal microscope (Nikon A1R-si). Hoechst 33342 (Beyotime, C1029) was used to stain the nucleus. Laser: 405 nm, 488 nm.



### ChIP-qPCR and ChIP-mass spectrometry

ChIP-quantitative PCR (qPCR) primers were designed and validated by NCBI primer-blast. The Ct value of input DNA was used for normalization. Calculation formula:  $\%Input = 5\% \times 2^{-(Ct^{Input} - Ct^{Sample})}$ . The results of ChIP-qPCR were displayed by prism 8. At least two replicates were run for each qPCR. The statistical significance was evaluated by Student's *t*-test. The primers for the ChIP-qPCR can be found in Supplementary Table S9.

The suitable concentrations of SDS-PAGE protein gels to run ChIP protein preparations were chosen according to the molecular weight of target proteins. Cut gels of the full lane were used for MS analyses which were performed at the National Center for Protein Sciences at Peking University in Beijing, School of Life Science, Peking University) (Note: the part of the gels containing antibody was separated in to a single tube for MS analyses).

### DNA library construction

ChIP DNA libraries were constructed according to the previously reported TELP library construction protocol (68) or the NEBNext® Ultra™ II DNA Library Prep Kit (NEB, E765S). Briefly, DNA was end-repaired, and a poly(C) tail was added to the purified end-repaired DNA products. Poly(C) tails were extended with biotin-labeled adapters. Biotin-labeled products were captured by streptavidin C1 beads (Invitrogen, 11206D). Quick ligase (NEB, M2200S) was used to ligate another adapter to the biotin-labeled products. The first-round PCR amplification was performed on beads. The second-round PCR amplification was carried out with the first-round PCR products. The final PCR products were run on a 2% DNA agarose gel. DNA products of 200–400 bp were purified by a Magen gel extraction kit. The NEB library construction was according to the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® by following the manufacturer's instructions. The purified DNA was sequenced by Hiseq 4000 with 150 paired-end reads.

### RNA-sequencing

For RNA-seq, the ZBTB21-degron mESCs were induced by doxycycline for 24 h, and treated with IAA for 24 h. The untreated cells and IAA-treated degron cells were collected. RNA extraction was performed with Trizol by following the manufacturer's instructions. Our RNA-seq library was prepared by using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA), which is regularly used in our lab (64,69,70). The libraries were prepared by following the manufacturer's instructions and sent to Novogene and sequenced on Illumina Hiseq 4000; 150 bp paired-end reads were obtained for downstream analyses.

### Chromatin fraction preparation

The chromatin was prepared by following the protocol published previously (65). A total of 10 million cells were incubated with NP-40 cell lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.01% NP-40 and protease inhibitors) to lyse cell membranes. The cell suspension

was transferred to a 1.25 ml sucrose cushion [24% sucrose (w/v) in NP-40 cell lysis buffer] to extract the nuclei. The supernatant (cytoplasmic fraction) was collected. The nuclei were gently resuspended with 0.5 ml of glycerol buffer [20 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT, 50% (v/v) glycerol]. A 0.5 ml aliquot of nuclei lysis buffer (10 mM HEPES pH 7.6, 1 mM DTT, 7.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 0.3 M NaCl, 1 M urea, 1% NP-40) was added to the resuspended nuclei. After centrifugation, the supernatant was saved as a nucleoplasmic fraction. The pellets were chromatin fractions used in this study; 1× SDS loading buffer was added to the chromatin pellets and boiled at 100°C for 5 min. Western blot was used to examine the target proteins.

### Chromatin RNA-seq (ChAR-seq)

For ChAR-seq, the ZBTB21-degron mESCs were induced by doxycycline for 24 h, and treated with IAA for 6 h. The chromatin was prepared by following the chromatin fraction preparation protocol as described above. The chromatin RNA was extracted from chromatin fractions with Trizol by following the manufacturer's instructions. The total RNA samples were firstly depleted of rRNAs and were constructed into next-generation sequencing libraries by following the experimental procedure for RNA-seq library preparation. The ChAR-seq libraries were sequenced in the same way as for RNA-seq libraries.

### ChIP-MS analysis

MS raw data were mapped to mouse or human proteins in the Uniprot database by the MS analysis platform in the National Center for Protein Sciences at Peking University, Beijing. Then, the peptide counts of all isoforms of a protein were combined to generate a single peptide-spectrum match (PSM) value for each protein. For mESCs, the candidate proteins were selected if the total peptide counts of the proteins in ChIP samples were at least twice more than those in input samples. For B16 and CH12F3 cells, the candidate proteins were identified by the R package DEP [*P*-value < 0.05, log<sub>2</sub> fold change (FC) > 1] (71). The missing data were imputed using the 'MinProb' method. The candidate proteins were annotated according to previous publications (72,73), Uniport, Genecards and g:Profiler databases. In addition, the Gene Ontology (GO) determination of candidate proteins was carried out by using R package Cluster Profiler (74).

### ChIP-seq data analysis

Raw ChIP-seq reads were firstly trimmed using Cutadapt software (75) and then mapped to the genome using Bowtie2 software (76) in 'very-sensitive-local' mode. ChIP-seqs that were performed in mESCs were mapped to the mm10 genome. ChIP-seqs that were performed in HEK293T cells were mapped to the hg19 genome. ChIP-seqs with spike-in control were mapped to both the mm10 and hg19 genome. Only uniquely and concordantly mapped reads were kept for further analysis. The mapped reads were converted to bigwig format using bamCoverage from

deepTools2 software (77) in 10 bp bins with parameters ‘–normalizeUsing’ set to RPGC. The mapped reads from ChIP-seq with spike-in control in mESCs were converted to bigwig format normalized by the spike-in reads mapped to the hg19 genome. Peaks were called by MACS2 software (78) using a *q* threshold of 0.05. The two replicates were filtered by IDR software (<https://github.com/nboley/idr>) with an *idr* threshold of 0.05 to obtain repeatable peaks. Peaks that overlapped with the mm10 or hg19 blacklist were removed from the peak lists.

The genomic distribution of peaks was annotated according to the priority of promoters, enhancers, CTCF sites, gene bodies and others. The *cis*-regulated elements were defined as below:

Promoters:  $\pm 1$  kb from transcription starting site (TSS) of Refseq genes.

Enhancers: H3K27ac peaks that did not overlap with promoters. Superenhancers were downloaded from a previous publication (79).

CTCF sites: CTCF peaks that did not overlap with enhancers and promoters.

Heatmaps and meta-gene profiles were calculated using computeMatrix of deepTools2 software centered around  $\pm 3$  or  $\pm 5$  kb of the peaks, SMC1 ChIA-PET anchors ( $n = 24\,036$ , in Figure 2H) and genes. ChIP-seq signals of cohesin, NIPBL and ZBTB21 were counted at active gene promoters. The correlation curves between cohesin, NIPBL and ZBTB21 were fitted and smoothed by LOESS regression. Motifs surrounding  $\pm 250$  bp of peak summits were predicted using Homer (<http://homer.ucsd.edu/homer/ngs/index.html>) software. The genes regulated by each peak were assigned by the R package ChIPseeker (80). GO analyses were performed using the R package ClusterProfiler (74).

To define ZBTB21-bound and unbound genes, genes were first filtered by requiring the existence of ZBTB21 peaks for the bound ones and no peaks for the unbound ones. Then the top 1000 genes were selected as bound genes based on the ChIP-seq signal of ZBTB21 at gene promoters. The bottom 1000 no peak genes were selected as the unbound genes. The mm10 gene set was obtained from the GENCODE database.

The differential cohesin peaks were first identified by the R package Diffbind using a false discovery rate (FDR) threshold of 0.05 to obtain more robust changes in 0–1, 0–6 and 1–6 h. These differential peaks were then merged and clustered using the R package hclust. For clustering analysis, a heatmap was used to visualize the changes of cohesin binding. Due to the high variability of cohesin binding signals at different regions, the *z*-score, which was calculated by  $(x - \text{mean})/\text{standard deviation (SD)}$ , was used to normalize the data in order to make the change patterns clearer. To determine the optimal cluster numbers, we used the ‘Elbow method’ to test a range of cluster numbers and finally selected three clusters to reflect the biologically meaningful changes.

Peaks of ZBTB21, cohesin and CTCF ChIP-seq datasets in Figure 2C were first merged and then clustered by their ChIP-seq signal using the R package k-means. The control peaks were randomly selected from the mm10 genome (excluding the merged peak regions) to the same number

of ZBTB21 peaks. ChIP-seq densities of cohesin, RAD21, NIPBL, WAPL and CTCF were then plotted according to the clustered and control categories. The heatmap in Figure 6C was generated using the same method utilizing the merged peaks of the five ZBTB transcription factors, cohesin and CTCF. The 1/2/3/4+ ZBTB peaks in Figure 6F were defined as the regions co-occupied by one, two, three, four and more than four ZBTB transcription factors.

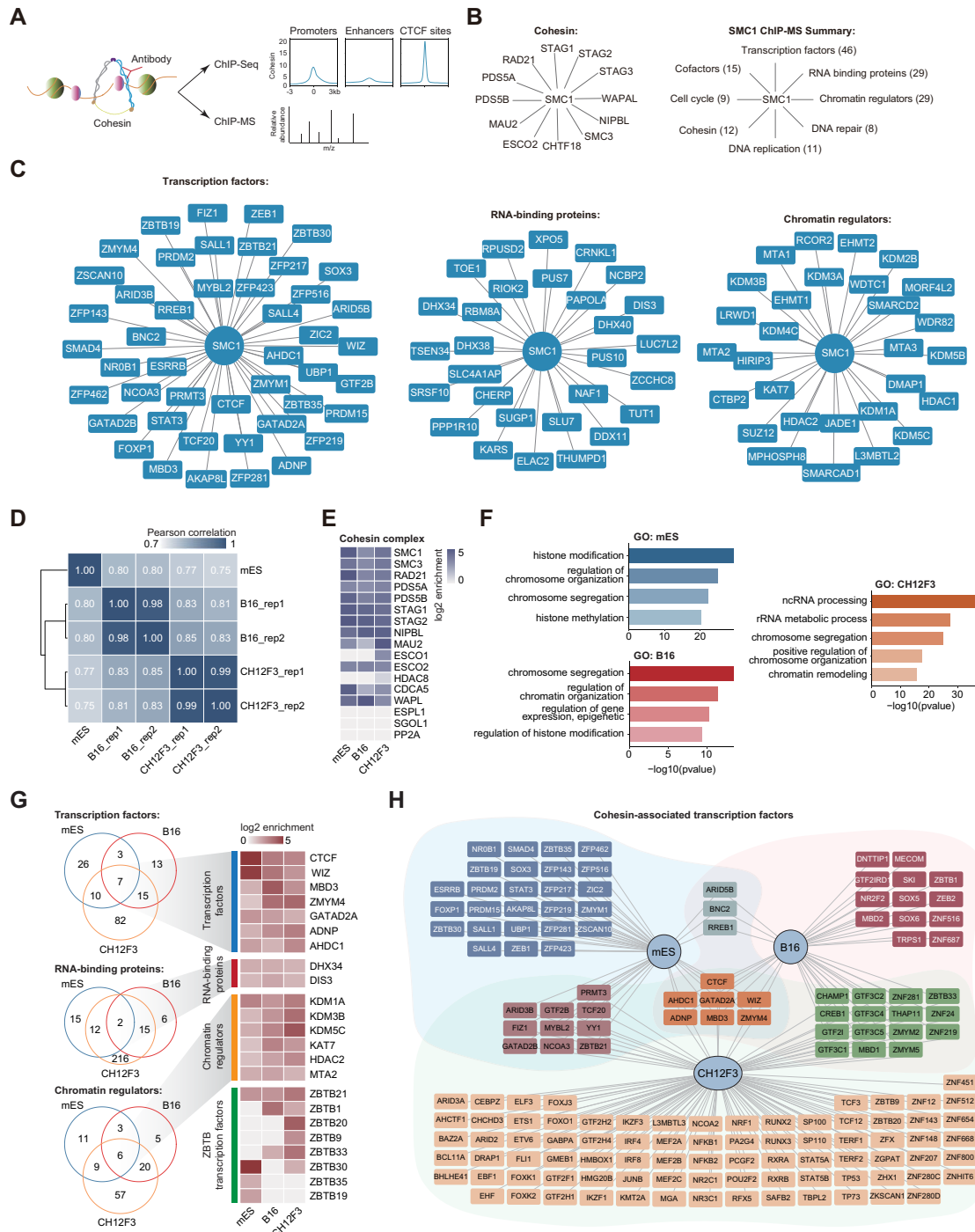
Statistical tests used in this study were all determined by two-side Wilcoxon test or Student’s test as mentioned in each figure legend.

## Hi-C data analysis

The bridge linkers of raw Hi-C reads were first trimmed using trimLinker of the ChIA-PET2 software (81). The processed reads were then mapped to the mm10 genome and filtered to obtain valid contact pairs using the HiC-Pro pipeline (82). RCP duplicated, dangling ended, re-ligated and self-circled read pairs were discarded for further analysis. Long-range ( $>20$  kb) intrachromosomal valid pairs of all the samples were randomly sampled to the same depth ( $n = 34$  M) and merged together between the two replicates ( $n = 68$  M) to obtain higher resolution. The resolution of our Hi-C library is  $\sim 25$  kb, so the interactions  $<20$  kb are below our Hi-C resolution, which are potential false-positive interactions. Therefore, we removed the interactions  $<20$  kb in our analyses. Many previous publications also removed the short-distance interactions for subsequent analysis (83–85). The merged read pairs were converted to ‘.hic’ format using hicpro2juicebox in HiC-Pro software and then converted to ‘.cool’ format using hic2cool software (<https://github.com/4dn-dcic/hic2cool>). The ‘.cool’ format Hi-C maps were balanced by cooler software (86). The snapshots of contact maps were visualized using Juicebox software (87) after Knight–Ruiz (KR, also known as balanced) normalization. The Hi-C maps were iteratively corrected using the ‘balance’ command from cooler software.

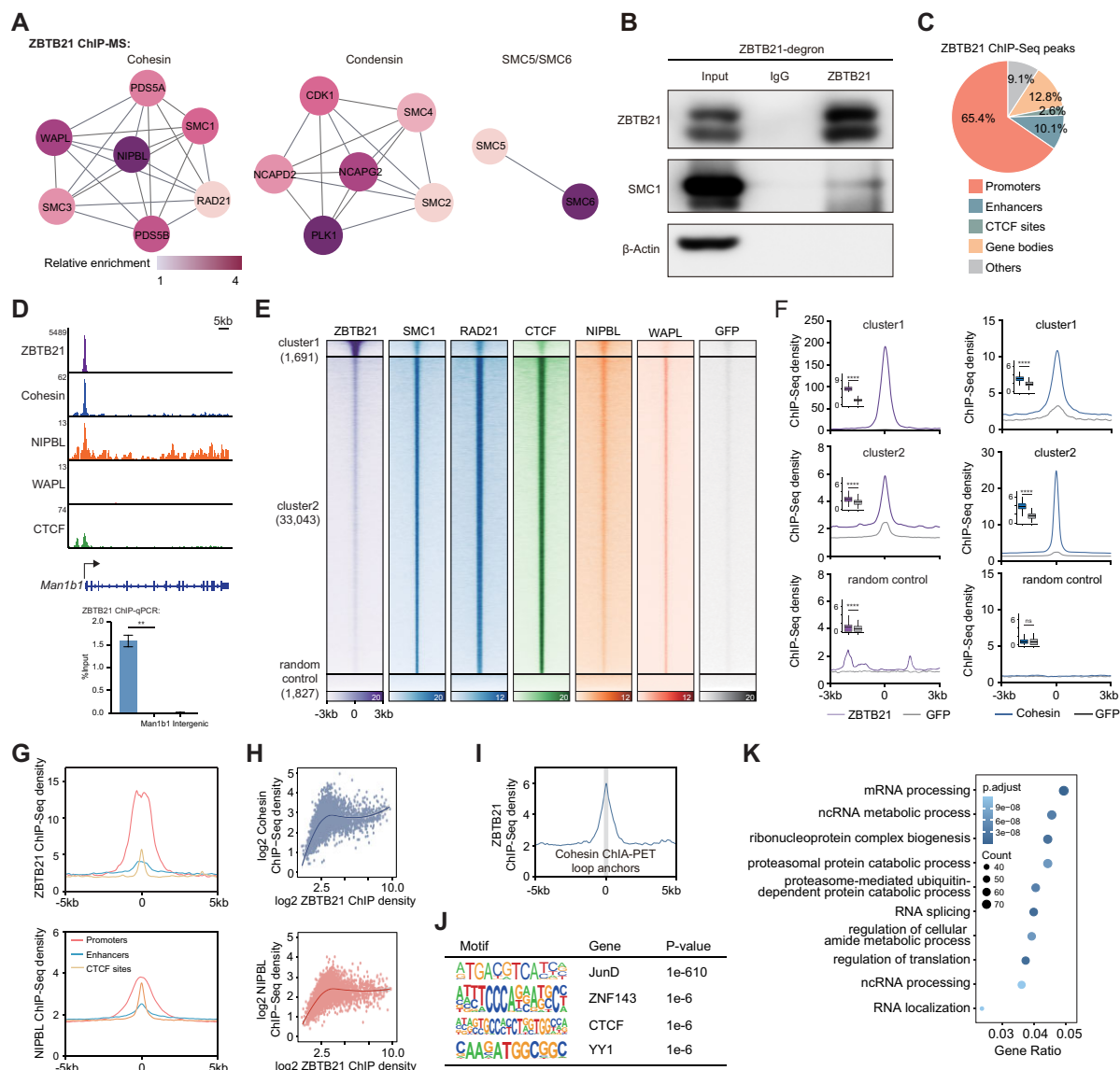
A/B compartments were called using call-compartments of Cooltools software (<https://github.com/open2c/cooltools>) and adjusted by H3K27ac ChIP-seq signals. The saddle plots were computed by compute-saddle of Cooltools software at 50 kb resolution and normalized by the expected matrix in ‘–quantiles’ mode. The enrichment score was calculated as  $(AA + BB)/2AB$ , where AA refers to the mean normalized density of bins in the top 20% eigenvector scores, BB refers to the mean normalized density of bins in the bottom 20% eigenvector scores, and AB refers to the mean normalized density of bins between AA and BB.

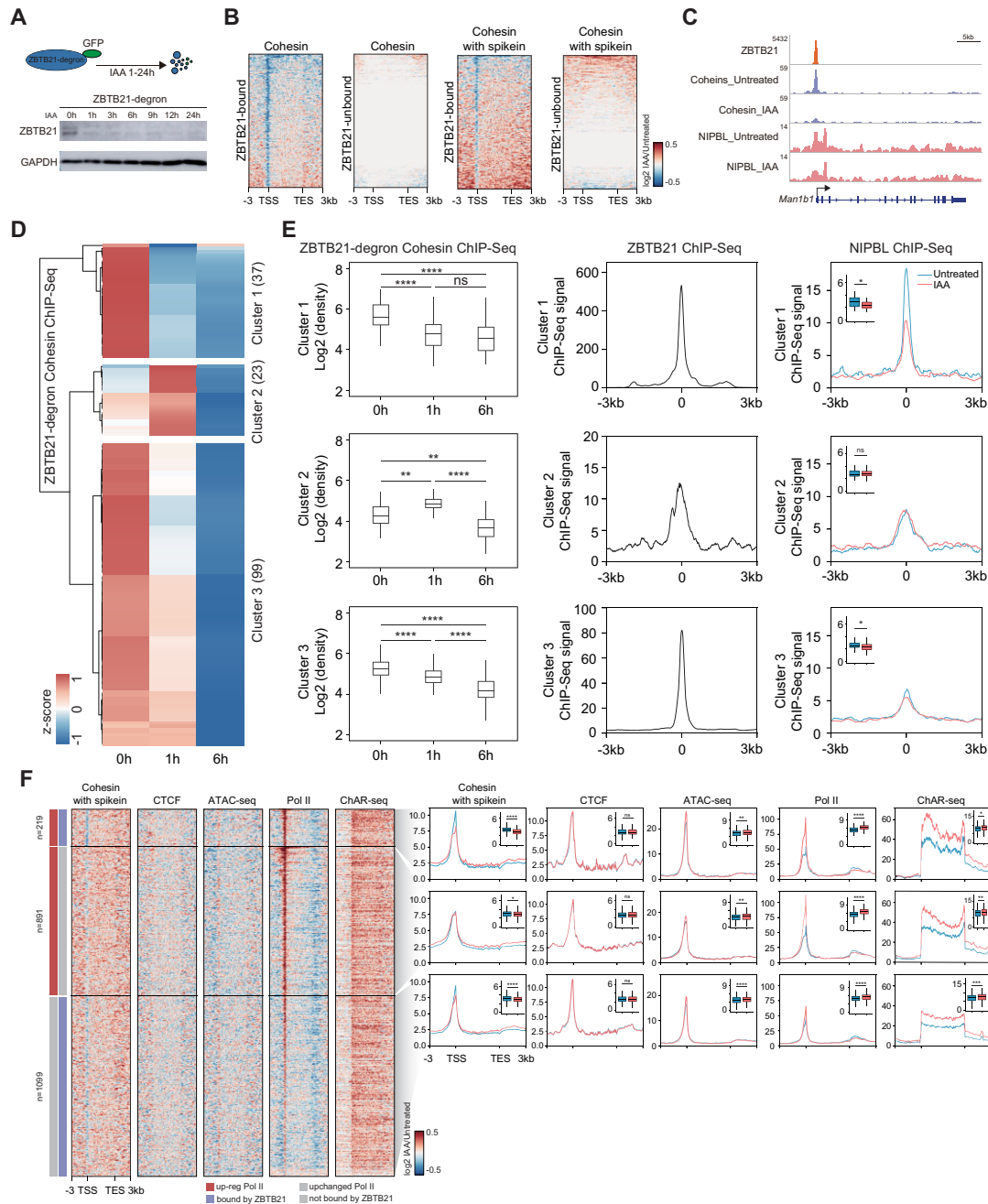
To plot meta-TAD, we firstly calculated the average contact probability for all loci at a certain distance as the expected Hi-C contact value for each chromosome. Then we transformed the KR-normalized Hi-C matrix into an observed/expected matrix by dividing each normalized observed value by its corresponding expected value at that distance. Second, we carried out a 2D meta feature analysis by piling up individual submatrices into an average matrix using Coolpup software (v0.9.5) with ‘–rescale’ and ‘–local’ parameters. In general, this is similar to meta-gene analysis, commonly performed for ChIP-seq data. The positions of



**Figure 1.** Cohesin ChIP-MS identifies candidate regulators of 3D chromatin organization. **(A)** Schematic representation of the cohesin chromatin proteomic profiling (ChIP-MS) procedure. Formaldehyde-cross-linked chromatin was sonicated, and ChIP was performed with specific antibodies (SMC1). The ChIP-purified DNA was subjected to DNA sequencing, and the ChIP-purified proteins were subjected to MS analyses. Meta-gene results showing the occupancy of cohesin (SMC1) in mESCs are presented in the right panel. **(B)** Cytoscape analysis of the cohesin ChIP-MS-identified cohesin components (left panel) and a summary of the categories of the identified proteins (right panel). The numbers within parentheses represent the numbers of identified proteins in each category identified by cohesin ChIP-MS. **(C)** Cytoscape analysis of cohesin ChIP-MS-identified transcription factors, RNA-binding proteins and chromatin regulators in mESCs. **(D)** The heatmap shows the Pearson correlation coefficients between each of the samples of mES, B16 and CH12F3 cells. The coefficients are clustered by hierarchical clustering. **(E)** The heatmap shows the log2 enrichment of the PSM values of cohesin components in ChIP samples over the input sample in mES, B16 and CH12F3 cells. **(F)** GO enrichment terms of the enriched proteins identified by cohesin ChIP-MS in mES, B16 and CH12F3 cells. **(G)** Venn diagram showing the shared transcription factors, RNA-binding proteins and chromatin regulators identified by cohesin ChIP-MS in mES, B16 and CH12F3 cell lines. The heatmap (right) shows the log2 enrichment of the PSM values of shared proteins and ZBTB transcription factors in ChIP samples over the input sample. **(H)** Network showing the transcription factors preferentially enriched in mES, B16 and CH12F3 cell lines. The transcription factors that were statistically significantly enriched are shown in the plot ( $P$ -value < 0.05,  $\log_2\text{FC}$  > 1). Transcription factors within different color backgrounds indicate that the proteins were detected in a specific cell line or different cell lines.

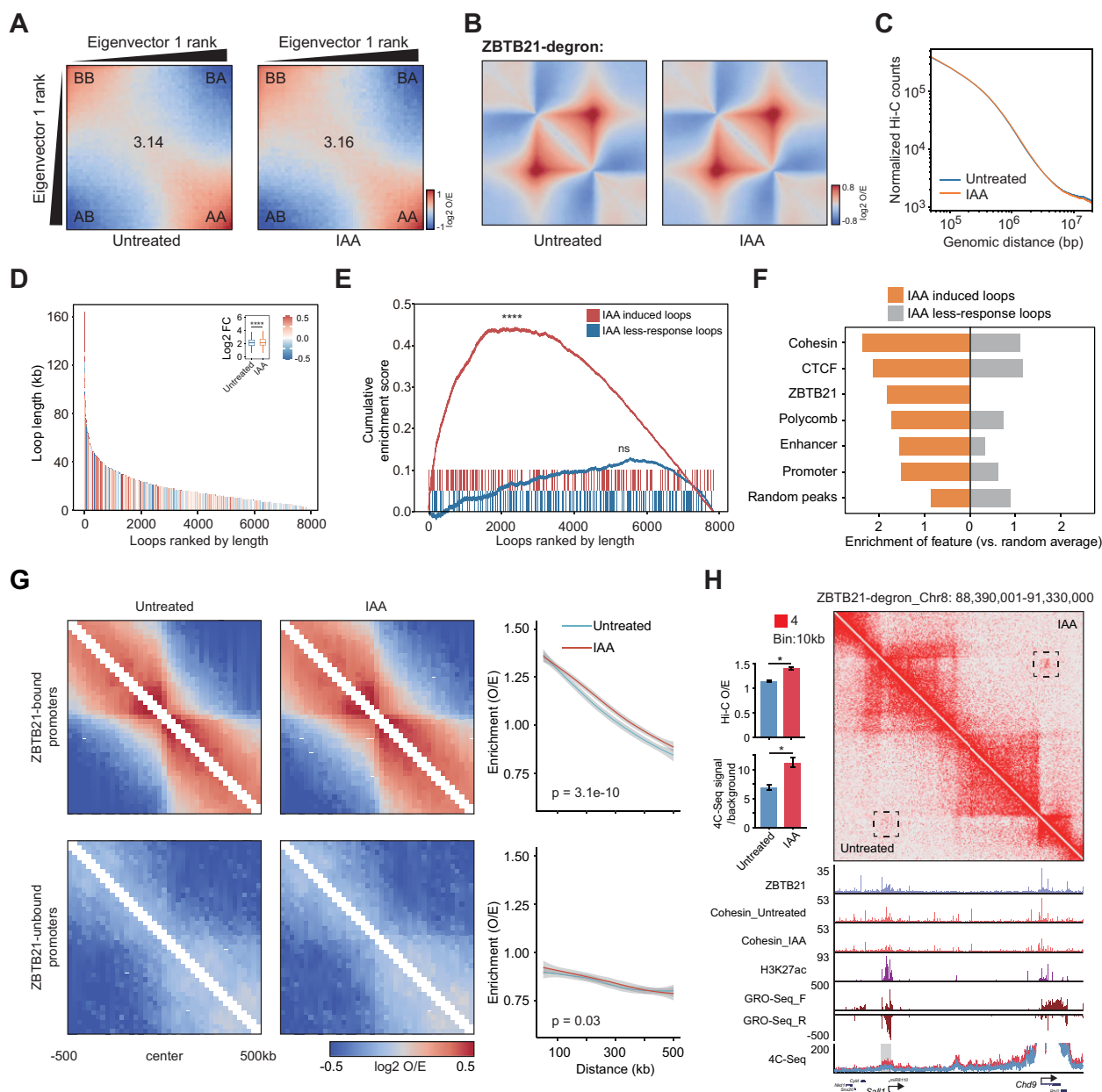




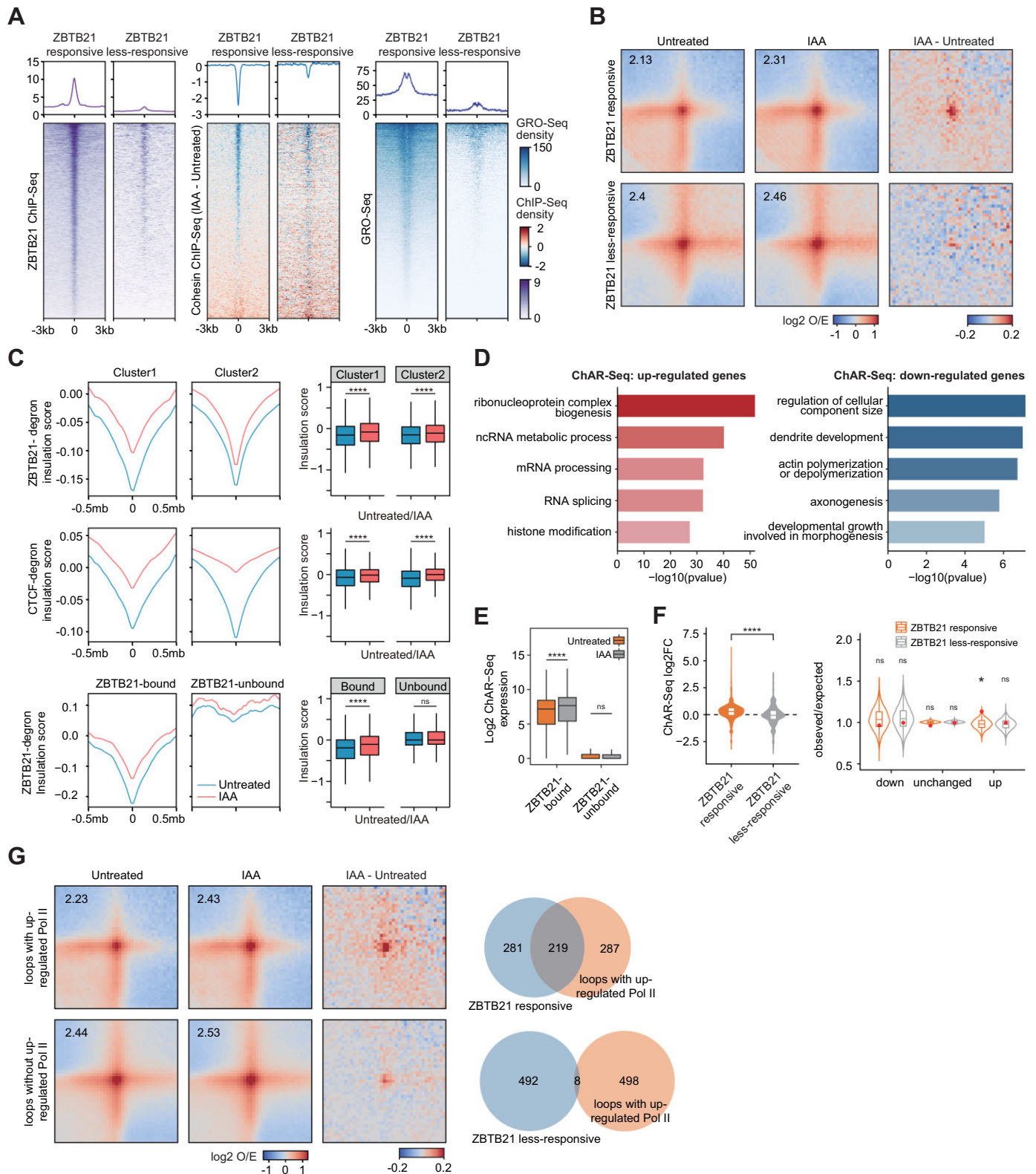


**Figure 3.** ZBTB21 depletion leads to decreased chromatin binding of cohesin. (A) Western blotting analysis of ZBTB21 degradation in ZBTB21-degron-GFP mESCs after IAA treatment at different time points. ZBTB21 was identified with a GFP antibody because it was endogenously GFP tagged, and GAPDH served as a loading control. (B) Heatmap showing the occupancy changes of cohesin with or without spike-in controls at ZBTB21-bound and ZBTB21-unbound genes. Each line of the heatmap indicates the log<sub>2</sub>FC in the ChIP-seq signal at a gene and its  $\pm 3$  kb flanking region. The color bar indicates the log<sub>2</sub>FC in ChIP-seq signals. For the definition of ZBTB21-bound and unbound genes, see the Materials and Methods. (C) ChIP-seq snapshots of the ZBTB21 signal and down-regulated cohesin and NIPBL signals upon ZBTB21 degradation at the *Man1b1* locus. (D) Hierarchical clustering of cohesin ChIP-seq density in peaks that were differentially regulated at 0–1, 0–6 or 1–6 h. The differentially regulated peaks (FDR  $< 0.05$ ) were identified by DiffBind, merged together and classified into three clusters (see the Materials and Methods). (E) Left: boxplot showing the log<sub>2</sub> cohesin ChIP-seq density of differentially regulated peaks in each cluster. Middle: meta-analyses displaying the mean ZBTB21 ChIP-seq densities at differentially regulated cohesin peaks in each cluster. Right: meta-analyses displaying the mean NIPBL ChIP-seq signals at differential cohesin peaks in each cluster after ZBTB21 degradation for 1 h. The inserts indicate the log<sub>2</sub> densities of ChIP-seq signals. Significance was determined by the Wilcoxon test (\*\*\*\* $P < 0.0001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ ; ns, not significant). (F) Heatmap showing RNA polymerase II (Pol II), cohesin, CTCF ChIP-seq, ATAC-seq and ChAR-seq signals among the genes with up-regulated Pol II (red bar) and ZBTB21 binding (purple bar) ( $n = 219$ ), genes with up-regulated Pol II and without ZBTB21 binding ( $n = 891$ ) and genes with less changed Pol II (below our cut-off) and ZBTB21 binding ( $n = 1099$ ) (left). The color bar indicates the log<sub>2</sub>FC in ChIP-seq, ATAC-seq or ChAR-seq signals. Genes with up-regulated Pol II binding were identified by the R package DiffBind (FDR  $< 0.05$ ). The meta-plots of the three clusters of genes are plotted on the right. The boxplot (insert) shows the changes in the ChIP-seq, ATAC-seq signals at  $\pm 100$  bp around the TSS, or ChAR-seq expression changes upon ZBTB21 degradation. Significance was determined by the Wilcoxon test (\*\*\*\* $P < 0.0001$ , \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ ).





**Figure 4.** ZBTB21 insulates small-scale 3D chromatin interactions. (A) Saddle plot showing the change in compartmental strength upon ZBTB21 degradation. The enrichment score was calculated as  $(AA + BB)/2AB$ . (B) Meta-TAD analysis showing that the TAD strength does not change upon ZBTB21 degradation. The genomic positions of TADs were downloaded from Bonev *et al.* (83). (C) Hi-C enrichment counts at different genomic distances before and after ZBTB21 degradation generated by HiCExplorer. The x-axis starts from 50 kb. Hi-C maps with 50 kb resolution were used for the plot. (D) Changes in loop strength upon ZBTB21 degradation. The loops from untreated and IAA samples were first merged and then ranked by their loop strength. The x-axis indicates the rank of the loops, the y-axis indicates the loop length and the color indicates the  $\log_2FC$  in loop strength. The boxplot (inset) shows the changes in loop strength. Significance was determined by the Wilcoxon test ( $****P < 0.0001$ ). (E) GSEA indicated that the gained loops in IAA samples were distributed at the top of all the merged loops ranked by length (see the Materials and Methods). The less responsive loops are shown as a negative control. Statistics were computed by permuting the loop labels and recomputing the enrichment score of the IAA-induced or less-responsive loop sets to generate a null hypothesis (89). (F) Enrichment of ChIP-seq peaks at anchors of induced loops that arise after ZBTB21 degradation and loops less responsive to ZBTB21. The enrichment of the peaks was calculated using the number of ChIP-seq peaks of the *cis*-regulatory elements inside the loops normalized by random averaging. Random averaging was defined by permuting the loops across the genome 1000 times and calculating the average number of ChIP-seq peaks at these random loop anchors. (G) Aggregate target-centered Hi-C interaction maps showing the changes in contact strength upon ZBTB21 degradation around ZBTB21-bound and ZBTB21-unbound promoters. Pile-up maps were plotted at a 25 kb resolution and normalized by the expected matrix. The color bar indicates the  $\log_2$  ratio of observed/expected density. The right panel shows the decay curve of the Hi-C interaction frequencies extracted from the observed/expected matrix in the corresponding regions. Significance was determined by the Wilcoxon test in paired mode. (H) Hi-C contact maps for the regions of chromosome 8 (88–91 Mb) at a 10 kb resolution in untreated and IAA-treated (6 h) ZBTB21-degron cell lines. Genome browser snapshots of ZBTB21, cohesin (SMC1) and H3K27ac ChIP-seq signals, GRO-seq (GRO-seq\_F for signals of the sense strand, GRO-seq\_R for signals of the antisense strand) and 4C-seq signals. A subset of active genes is shown. The quantification of increased regions (indicated by a dashed rectangle or gray bar) of Hi-C (top) and 4C-seq signals (bottom) before and after ZBTB21 degradation are displayed in the left panel. Significance was determined by Student's *t*-test ( $*P < 0.05$ ). The error bars indicate the SD. The red rectangle indicates the normalized Hi-C contact frequency.



**Figure 5.** ZBTB21 depletion affects 3D chromatin interactions of ZBTB21-bound promoters. (A) Heatmap and aggregate signal analysis showing ZBTB21 ChIP-seq (left) and differential cohesin over 0–6 h (middle) and GRO-seq (right) centered around ATAC-seq peaks in the top 500 loops that contain both ZBTB21-bound and down-regulated cohesin peaks (ZBTB21-responsive loops) and the bottom 500 loops that do not contain ZBTB21 and down-regulated cohesin peaks (less ZBTB21-responsive loops). The color bar indicates ChIP-seq density. For the definition of responsive and less-responsive loops, see the Materials and Methods and the loop information in Supplementary Table S7. GRO-seq data were downloaded from a previous publication (128). (B) APA plot indicating the changes in loop strength upon ZBTB21 degradation in ZBTB21-responsive (upper panel) and less ZBTB21-responsive (bottom

TADs were downloaded from Bonev *et al.* (83). Each TAD was padded by the neighboring regions with the same length as the TAD. Hi-C O/E maps at 25 kb resolution were used for the analysis.

Loops were called in 5, 10 and 25 kb resolution using Mustache software (88) with parameters ‘-pt’ and ‘-st’ set to 0.1 and 0.88. Loops of different resolutions were extended to the lowest resolution (25 kb), and merged together for further analysis. Loops from untreated and IAA-treated samples were first merged together and then ranked and plotted by their loop length in Figure 4D. The color of each loop indicates the log2FC of loop strength before and after ZBTB21 degradation. The gene set enrichment analysis (GSEA) in Figure 4E indicates that IAA-induced loops (the loops that occurred after ZBTB21 degradation) distribute at the top longest loops among all the merged loops. Loops were ranked by their loop length and marked in red if they were hits in IAA-gained loop sets. The less responsive loops (see below) serve as negative controls. Statistics were computed by permuting the loop labels and recomputing the enrichment score of the IAA-induced or less responsive loop sets to generate a null hypothesis (89). Local Hi-C interactions of promoters, enhancers, CTCF sites and superenhancers were calculated by Coolpup software at 25 kb resolution and normalized by the expected matrix. The interaction frequency enrichment curves were fitted using all the bins that interacted with the bin containing the peaks.

ZBTB21-responsive loops were defined as the top 500 loops that contained both ZBTB21 peaks and down-regulated cohesin peaks; less ZBTB21-responsive loops were defined as the bottom 500 loops that did not contain ZBTB21 peaks and down-regulated cohesin peaks. Loops were ranked by the sum of the ZBTB21 ChIP-seq signal and the absolute value of log2FC of cohesin ChIP-seq signal upon ZBTB21 degradation at the ATAC-Seq peaks inside each loop. Because the loop length distribution analyses indicated that loops >500 kb showed more changes, loops >500 kb were classified into ZBTB21-responsive and less-responsive loops. Down-regulated cohesin peaks were identified using cohesin ChIP-seqs before and after ZBTB21 degradation for 6 h by Diffbind software using a *P*-threshold of 0.05. ZBTB21, differential cohesin and

GRO-seq signals were plotted by heatmaps centered around the ATAC-seq peaks in ZBTB21-responsive ( $n = 7973$ ) and less-responsive ( $n = 1112$ ) loops. GRO-seq signals were not classified into + or – strands due to the inability to distinguish ATAC-seq peaks between the + or – strands. APA plots of loops were calculated by Coolpup software at 25 kb resolution and normalized by the expected matrix. The enrichment score was computed as the mean normalized density of the central 3\*3 bins.

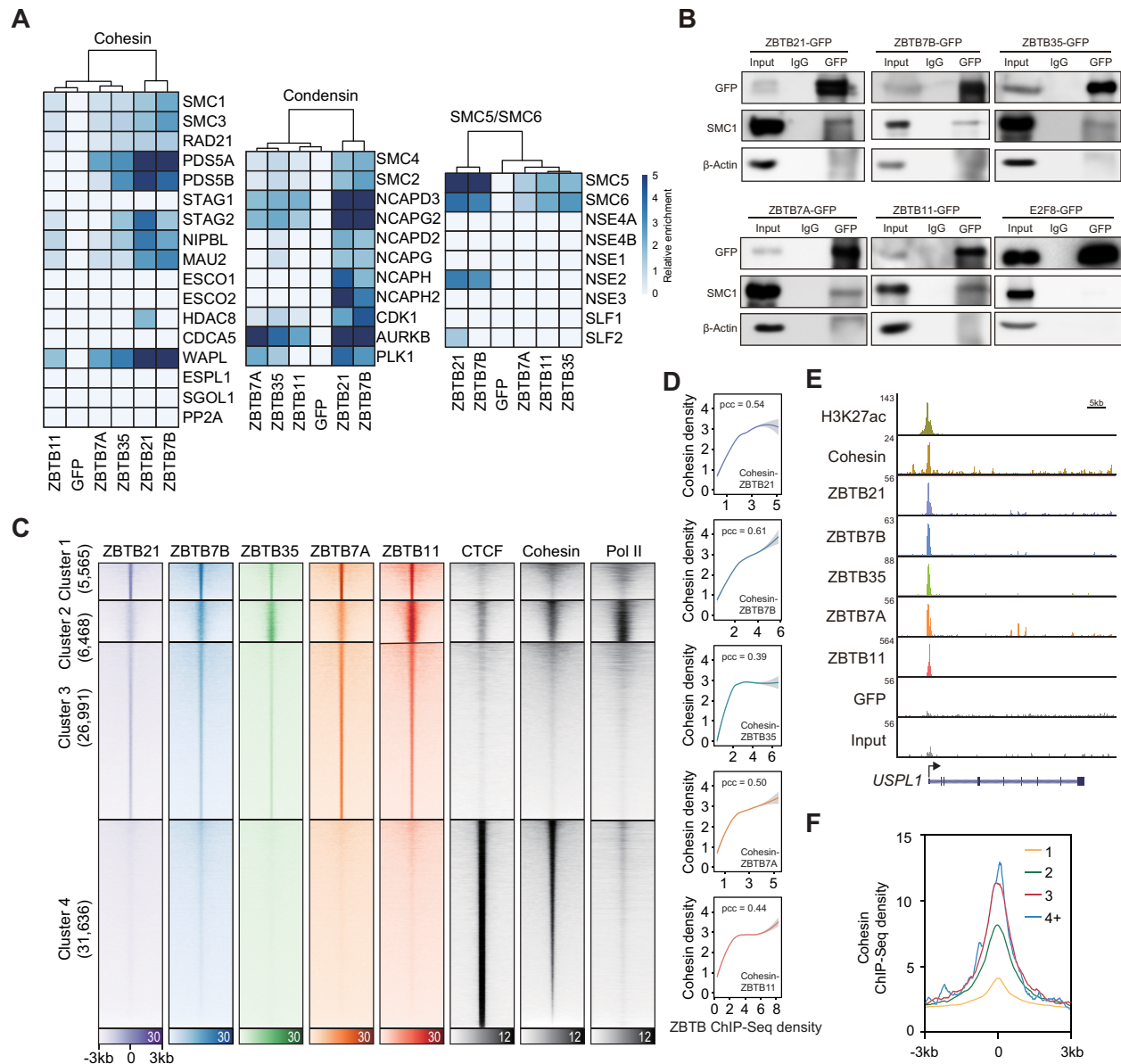
The genes were assigned to ZBTB21-responsive and less-responsive loops and plotted using the absolute value of log2 ChAR-seq fold change upon ZBTB21 degradation for 6 h (Figure 5D). The enrichment of down-regulated, unchanged and up-regulated genes was shown as an observed/expected ratio in the same way as in a previous publication (90). The percentages of genes in each category of loops (observed) were normalized by the genome-wide percentage of the genes in each category (expected). The red dots indicate the observed/expected ratio of ZBTB21-responsive and less-responsive loops. ZBTB21-responsive and less-responsive loops were randomly permuted 10 000 times across the genome to generate distributions of the ratio (the violin plot), and significance was determined by the distributions. The insulation scores of the ice-normalized Hi-C matrix of ZBTB21 were computed by cworld software (<https://github.com/dekkerlab/cworld-dekker>) at 25 kb resolution. The insulation scores of the CTCF Hi-C matrix were computed by cworld software at 20 kb resolution. CTCF Hi-C data were downloaded from a previous publication (25).

#### 4C data analysis

Raw 4C reads were first filtered and trimmed by the enzyme sites using trimLinker from ChIA-PET2 (81) software. The paired-end reads were selected for downstream analysis by Cutadapt (75) software if the 5' end of reads1 contains bait sites. The processed reads2 were then mapped to the mm10 genome by bowtie2 (76) software using ‘-very-sensitive-local’ parameters. PCR duplicates, self-religations and reads longer than the library size were discarded. The bigwig files were generated by bamCoverage of deepTools

panel) loops. Pile-up maps were plotted at 25 kb resolution and normalized by the expected matrix. The color bar indicates the log2 ratio of the observed density/expected density. The third column shows the subtracted value of loop strength extracted from the observed/expected matrix in IAA-treated and untreated samples. (C) Aggregated insulation score changes around peaks in each cluster identified in Figure 2C upon ZBTB21 and CTCF degradation. Aggregated insulation score changes around ZBTB21-bound and ZBTB21-unbound promoters are shown at the bottom. The insulation scores of the ZBTB21 Hi-C maps were generated at 25 kb resolution. The insulation scores of the CTCF Hi-C maps were generated at 20 kb resolution. CTCF Hi-C data were downloaded from previous publications (25,41). Quantification results of the insulation scores are shown on the right. Significance was determined by the Wilcoxon test (\*\*\*\* $P < 0.0001$ ). (D) GO analysis of up- and down-regulated genes identified by ChAR-seq upon ZBTB21 depletion. The DEGs were identified by DEseq2. (E) Boxplot showing the expression changes revealed by ChAR-seq for ZBTB21-bound and ZBTB21-unbound genes before and after ZBTB21 degradation. Significance was determined by the Wilcoxon test (\*\*\*\* $P < 0.0001$ , ns,  $P > 0.05$ ). (F) The violin plots (left) show the expression changes of the genes in ZBTB21-responsive and less-responsive loops. Significance was determined by the Wilcoxon test (\*\*\*\* $P < 0.0001$ ). The violin plots (right) show that genes in ZBTB21-responsive loops are more enriched in up-regulated genes. The observed/expected ratio was calculated by using the percentages of down-regulated, unchanged and up-regulated genes in loops (observed) normalized by the genome-wide percentage of the genes in each category (expected). The red dots indicate the observed/expected ratios of ZBTB21-responsive and less-responsive loops. ZBTB21-responsive and less-responsive loops were subjected to random permutation 10 000 times across the genome to illustrate the distributions of the ratios (violin plot), and significance was determined by the distributions. Up- or down-regulated genes were identified by DEseq2 (FDR < 0.05). (G) APA plot indicating the changes in loop strength upon ZBTB21 degradation in loops with (upper panel) and loops without (bottom panel) up-regulated Pol II. Up-regulated Pol II was identified by Diffbind software at gene promoters. Genes longer than 3 kb were selected for the analysis. Pile-up maps were plotted at 25 kb resolution and normalized by the expected matrix. The color bar indicates the log2 ratio of the observed density/expected density. The third column shows the subtracted value of loop strength in IAA-treated and untreated samples. Venn diagrams showing the overlap of ZBTB21-responsive loops and loops with up-regulated Pol II (upper panel) and the overlap of less ZBTB21-responsive loops and loops with up-regulated Pol II (bottom panel).





**Figure 6.** Additional ZBTB factors interact with cohesin. (A) Heatmap depicting the proteomic detection of the SMC complex for ZBTB protein ChIP-MS in HEK293T cells. The blue color intensity indicates the relative enrichment (IP/input). The scale bar is shown on the top right. Scale bar: relative enrichment (IP/input). The subunits of cohesin, condensin and the SMC5/SMC6 complex were based on a previous publication (129). (B) Western blotting was used to examine the interactions among ZBTB-GFPs and cohesin (SMC1) in native ZBTB-GFP HEK293T cells. The input cell lysate was obtained from HEK293T cells with stable Tet-on-induced expression of the protein of interest. IP samples were obtained via GFP antibody-IP. IgG and  $\beta$ -actin were used as negative controls. A 5% input was loaded. (C) Heatmap showing the occupancy of the ZBTB transcription factors cohesin, CTCF and Pol II among merged peaks of the ZBTB transcription factors cohesin and CTCF. The peaks of the ZBTB transcription factors cohesin and CTCF were first merged and then clustered by their ChIP-seq signal using the k-means method. Pol II ChIP-seq signals were plotted according to each cluster. (D) The curves show the correlations between ZBTB ChIP-seq density and cohesin ChIP-seq density at active gene promoters. The curves were fitted and smoothed by LOESS regression. The Pearson correlation coefficient is shown. (E) ChIP-seq snapshots of the co-localization among ZBTB21, ZBTB7B, ZBTB35, ZBTB7A, ZBTB11, the GFP control and the input at the USPL1 locus. (F) Metagenesis analysis indicating the cohesin ChIP-seq signals at 1/2/3/4+ ZBTB ChIP-seq peaks and cohesin peaks. 1/2/3/4+ ZBTB peaks represent the regions co-occupied by one, two, three, four or more than four ZBTB transcription factors.

(91) software and normalized to the 1× coverage of the mm10 genome using ‘-binSize 10 -normalizeUsing RPGC’ parameters.

For displaying 4C track, we smoothed the signals in the same way as in a previous publication (41), i.e. the genome to be displayed was binned using a 50 bp sliding window. At each bin, the smoothed signal was calculated as the mean signal of a 5 kb region around the bin. For 4C quantification, the signal at a specific region (chr8:88940000–89060000) was normalized by the signal of a 4 Mb downstream background region (chr8:84940000–85060000) using the non-smoothed normalized signals.

### RNA-seq data analysis

Raw RNA-seq reads were first trimmed using Cutadapt software and then mapped to the mm10 genome using STAR software (92). Primarily mapped reads without PCR duplication were kept for further analysis. The signals of genes downloaded from GENCODE vM25 were obtained using featureCounts software. Differentially expressed genes (DEGs) were identified by the R package DEseq2 (93) with an FDR threshold of 0.05.

### ChAR-seq data analysis

Raw ChAR-seq reads were first trimmed using Cutadapt software and then mapped to the mm10 and dm6 genome using bowtie2 software (v2.3.5.1). Uniquely mapped reads without PCR duplication were kept for further analysis. The expression of the mm10 genes was counted by featureCounts software at gene body regions [+300 from the TSS to the transcription end site (TES)] and normalized by the spike-in reads mapped to the dm6 genome in each sample. DEGs were identified by the R package DEseq2 (93) with an FDR threshold of 0.05.

### Distribution percentage, CRISPR score and expression analysis

ChIP-seq peaks of transcription factors in Supplementary Figure S1 were downloaded from the ENCODE and Cistrome databases (Supplementary Table S8). ChIP-seqs of H3K27ac and CTCF in different cell lines, as well as genes in the mm9/mm10 and hg19/hg38 genome, were downloaded to generate the genomic regions of promoters, enhancers and CTCF sites in corresponding cell lines. Peaks were annotated as these *cis*-elements and clustered based on the percentage of occupancy using the R package hclust. Gene expression data of the transcription factors in 53 different tissues were obtained from the GTEx database. All the samples were grouped and averaged by tissues. The expression percentage of each gene was defined as the percentage of tissues whose averaged transcripts per million (TPM) value was >10. CRISPR scores of human genes were obtained from a previous publication (94), which were computed by log2FC in abundance of each sgRNA between the initial and final populations. Genes were ranked by their CRISPR scores and defined as essential if their CRISPR scores were <0 and the *P*-value was <0.05.

## RESULTS

### Cohesin ChIP-MS identifies known and potential regulators of 3D chromatin organization

To comprehensively identify protein factors that contribute to 3D genome organization, we performed cohesin ChIP-MS in mESCs (Figure 1A) because cohesin plays a broad role in chromatin architecture (84,95). Cohesin ChIP-MS identified almost all known components and direct regulators of cohesin complexes (Figure 1B). The genomic distribution revealed by cohesin ChIP-seq among enhancers, promoters and CTCF sites was consistent with that obtained in previous studies (31–33) (Figure 1A; Supplementary Table S1). These results indicated the success of our cohesin ChIP experiment. The identified proteins were then grouped into different categories according to the UniProt, GeneCards and PubMed databases (Figure 1B; Supplementary Table S2; see also the Materials and Methods). Many known chromatin organization regulators were identified in our cohesin ChIP-MS data; for example, the known chromatin architecture proteins CTCF, YY1, ZFP143, ADNP and WIZ were among the transcription factors identified by cohesin ChIP-MS (Figure 1C) (25,41–44,47). The exosome complex was predicted to regulate superenhancer chromosomal interactions (96,97). The exosome component DIS3 was among the RNA-binding proteins identified by cohesin ChIP-MS (Figure 1C). Together, these analyses suggest that our cohesin ChIP-MS data in mESCs are reliable.

To provide further information on cohesin-interacting proteins, we performed genomic localization, expression specificity and essential score analyses. Genomic localization is expected to reveal the types of chromatin interactions regulated by these proteins, but not all these proteins have reliable ChIP-seq datasets in mESCs. We therefore searched the ChIP-seq datasets for cohesin-associated proteins from different cell lines included in the ENCODE and Cistrome databases, and calculated the occupancy percentages of promoter-, enhancer- and CTCF-binding regions in the corresponding cells (Supplementary Figure S1A). We performed gene expression analyses across different tissues by using gene expression datasets from the GTEx database (Supplementary Figure S1B) and conducted CRISPR score analyses to evaluate the essentiality of the cohesin-associated proteins (Supplementary Figure S1C) (94). We hypothesized that constitutively expressed, essential chromatin structural proteins, such as CTCF and YY1, play a general role in 3D genome organization (25,41,42), while less essential specifically expressed proteins, such as ADNP and AP-1, play a cell type-specific role or regulate a subset of chromatin interactions (44,48). Although these results were generated from different cells, they integrate useful information for investigating these cohesin interactors in the future.

It would be informative to compare our cohesin ChIP-MS data with the cohesin interactome in different cell types. Previous studies have identified the cohesin interactome in human HCT116 and MOLT-4 cells (98,99). However, these MS data are not comparable with our data, as they were generated by different experimental protocols and in different species. We therefore performed cohesin ChIP-MS

analysis in mouse B16 cells from skin and CH12F3 B lymphocytes by using a similar approach to the one we used for mESCs. We used the R package DEP (Differential Enrichment analysis of Proteomics data) to perform statistical analyses of these MS data ( $P$ -value  $< 0.05$ ,  $\log_2\text{FC} > 1$ ) and identified 260 and 1332 cohesin-interacting proteins in B16 and CH12F3 cells, respectively (the number of interacting proteins was quite different, possibly due to the different efficacy/accuracy of cohesin ChIP-MS in different cells). Our cohesin ChIP-MS data were quite reproducible and were strongly enriched in known cohesin components (Figures 1D, E), indicating that the quality of these MS data was good. We further compared the cohesin ChIP-MS data of B16 cells and CH12F3 cells with the cohesin ChIP-MS data of mESCs. GO analysis showed that the factors identified by cohesin ChIP-MS were mainly enriched in chromosome organization, histone modification and non-coding RNA processing terms in different cell types (Figure 1F). The factors that have previously been demonstrated to regulate cohesin, namely CTCF, WIZ and ADNP, were identified in three different cohesin ChIP-MS preparations (25,41–44,47). In addition to those known cohesin-interacting proteins, additional transcription factors, RNA-binding proteins and chromatin regulators were also detected in the cohesin ChIP-MS preparations (Figure 1G). Moreover, many transcription factors were preferentially detected in different cell lines (Figure 1H), which may be candidate regulators of cell type-specific 3D chromatin structures. These analyses provide a resource listing constitutive and cell type-specific cohesin-interacting proteins in different cell types.

Many of the proteins identified by cohesin ChIP-MS have not previously been reported to play roles in chromatin organization or to function together with cohesin. We detected four ZBTB transcription factors, ZBTB19, ZBTB21, ZBTB30 and ZBTB35, in mESCs (Figure 1C, G), which were of considerable interest to us because ZBTB transcription factors are known to regulate insulator activities, as suggested by the functions of ZBTB factors in *Drosophila* (52–56). Insulator function has been proposed to involve determining 3D chromatin organization (4,100,101). Since ZBTB21 was detected in cohesin ChIP-MS preparations in three different cell types (Figure 1G), we primarily investigated the molecular functions of ZBTB21 in this study.

### ZBTB21 interacts with cohesin at active promoters

ZBTB21 ChIP-MS and ChIP-seq were performed to validate the associations between ZBTB21 and cohesin. We first introduced a degron–GFP tag into the C-terminus of endogenous ZBTB21 in mESCs (Supplementary Figure S2A). Genotyping analyses confirmed the success of homozygous knockin (Supplementary Figure S2B). Then, widely used GFP antibodies were chosen to perform ChIP-MS and ChIP-seq in ZBTB21-degron–GFP cells. Consistent with the cohesin ChIP-MS results, cohesin components were preferentially identified in ZBTB21 ChIP-MS preparations (Figure 2A; Supplementary Table S2). Interestingly, condensin and SMC5/SMC6 components were also detected in ZBTB21 ChIP-MS preparations (Figure 2A). The protein–protein interactions between ZBTB21 and cohesin were further validated through IP–western blotting analy-

ses of native cell lysates but not IgG IP preparations (Figure 2B), which served as a negative control. To gain deeper insight into ZBTB21-mediated cohesin regulation, we performed ZBTB21 immunoprecipitation after ZBTB21 depletion or treatment with different nucleases. The results showed that ZBTB21 forms specific protein–protein interactions with cohesin (Figures S2C, D). The genomic localization of ZBTB21 was carefully characterized to obtain functional insights into ZBTB21-mediated regulation. Meta-gene analysis and a pie chart displaying the genomic distribution of the ZBTB21 peaks indicated that ZBTB21 preferentially bound gene promoters (Figure 2C). We identified ZBTB21 chromatin-binding sites in mESCs (Supplementary Tables S1 and S3). These binding sites could be independently validated by ChIP-qPCR (Figure 2D; Supplementary Figure S2E). Furthermore, genomic localization analyses indicated that ZBTB21-occupied regions showed weak binding of cohesin and the cohesin loading factor NIPBL, which was significant compared with the signals in randomly selected regions and GFP ChIP-seq signals (Figure 2E). ZBTB21 peaks belonged to Cluster 1 and were associated with weak cohesin ChIP-seq signals. Cluster 2 did not have ZBTB21 peaks and was associated with strong cohesin binding signals (Figure 2F).

Furthermore, ZBTB21 and NIPBL preferentially occupy the gene promoters (Figure 2G). As the ZBTB21 ChIP-seq density increased, the chromatin binding strength of cohesin and NIPBL at gene promoters first increased and was then maintained at a fairly consistent level (Figure 2H). ZBTB21-binding sites were enriched at the cohesin loop anchors identified by cohesin ChIA-PET (Figure 2I). Motif analyses of the ZBTB21 ChIP-seq peaks revealed the enrichment of transcription factor [e.g. JUND (a component of the AP-1 family), ZNF143, CTCF and YY1] functions in 3D chromatin organization (Figure 2J). GO analyses of the ZBTB21-bound gene indicated that these genes were involved in RNA metabolism pathways (Figure 2K). These results suggested that ZBTB21 was associated with cohesin chromatin binding at promoters, indicating potential involvement in cohesin promoter occupancy.

### ZBTB21 depletion leads to decreased chromatin binding of cohesin

To obtain functional insights into the interactions between ZBTB21 and cohesin, we performed cohesin (SMC1) and NIPBL ChIP-seq after ZBTB21 depletion (Supplementary Tables S1 and S4). The western blotting results confirmed the acute depletion of ZBTB21 proteins in our ZBTB21-degron mESCs (Figure 3A). We also performed cohesin ChIP-seq with equal amounts of *Drosophila* cells spiked-in. The tornado plot displaying the spike-in and non-spike-in cohesin ChIP-seq signal changes at the ZBTB21-bound and unbound promoters consistently showed that ZBTB21 depletion decreased cohesin binding at ZBTB21-bound promoters but not for the unbound promoters (Figure 3B). The chromatin binding of NIPBL at ZBTB21-bound gene promoters also displayed a weak decrease (Supplementary Figure S2F). Screenshots of specific genes showed that ZBTB21 depletion led to decreased cohesin and NIPBL binding at the *Man1b1*, *Mmab* and *Znr1* gene promoters



(Supplementary Figure S2G). We also performed cohesin ChIP after ZBTB21 depletion, and ZBTB21 resumed after withdrawing the IAA. The results showed that cohesin chromatin binding decreased after ZBTB21 depletion and then resumed after the withdrawal of IAA (Supplementary Figure S2H), suggesting that the roles of ZBTB21 in cohesin chromatin binding are specific. Collectively, these results support that ZBTB21 specifically regulates cohesin chromatin occupancy, even though direct protein–protein interactions with recombinant proteins are currently lacking.

We next carried out time-course cohesin ChIP-seq after ZBTB21 depletion to determine the dynamic effects on cohesin. The clustering analyses of the differential peaks indicated the existence of three clusters (Figure 3D, E): in Cluster 1, the cohesin signals rapidly decreased, and the peaks displayed high-density ZBTB21 signals; in Cluster 2, the cohesin signals first increased and then decreased, and the peaks displayed less ZBTB21 binding; in Cluster 3, the cohesin signals gradually decreased, and the peaks displayed an intermediate level of ZBTB21 binding. Moreover, there was an obvious decrease in NIPBL binding in Cluster 1, a small but significant decrease in Cluster 3 and no significant change in Cluster 2 after the rapid depletion of ZBTB21 (Figure 3E). We then performed Pol II and CTCF ChIP-seq and ATAC-seq to obtain functional insights into ZBTB21-mediated cohesin regulation (Figure 3F). The results revealed that ZBTB21 depletion leads to global transcriptional activation, as shown by the increased signals of Pol II ChIP-seq and ChAR-seq, suggesting that ZBTB21 functions in transcription repression, which is consistent with the findings of a previous study (102), but could not simply explain its effects on cohesin chromatin binding. The ATAC-seq signals and CTCF chromatin binding did not show obvious changes after ZBTB21 depletion, indicating that CTCF and chromatin accessibility also could not explain ZBTB21-mediated cohesin binding.

To gain functional insights into ZBTB21 chromatin binding, we performed total RNA-seq and ChAR-seq to enrich nascent RNAs after ZBTB21 depletion (Supplementary Table S5). The differential analyses showed that 473 DEGs in ChAR-seq have ZBTB21 binding, but those are a subset of ZBTB21-bound genes. The number of those genes was higher than the number of genes with similar RNA-seq results (Supplementary Figure S2I). These results are consistent with a recent finding that there is a poor correlation between transcription factor binding and acutely affected gene expression (103). We also examined the expression changes of subunits of cohesin, condensin and SMC5/6, and found that the expression of most of the subunits did not change. Some of the subunits showed up-regulation, but less than our cut-off (Supplementary Figure S2J; Supplementary Table S5). However, we could not completely rule out other secondary effects of ZBTB21-mediated transcriptional repression that might contribute to the reduction in cohesin binding.

### ZBTB21 contributes to 3D chromatin interactions

To explore the direct roles of ZBTB21 in 3D chromatin organization, we performed BAT Hi-C after ZBTB21 depletion for 6 h. The BAT Hi-C method, which was pre-

viously developed by our laboratory (64,66), can generate high-resolution chromatin interaction maps in a more economical way than the classical *in situ* Hi-C method. The Hi-C data were processed via previously established pipelines (64,66). The mapping rate was high and was quite reproducible between the two replicates (Supplementary Figure S3A; Supplementary Tables S1 and S6). The saddle plot and meta-TAD analyses indicated that ZBTB21 depletion did not obviously affect large-scale chromatin interactions in mESCs, such as A/B compartments and TADs (Figures 4A, B), and the decay curve of Hi-C interaction frequencies across different genomic distances likewise indicated no change (Figure 4C), suggesting that ZBTB21 was not required for global, large-scale chromatin interactions in mESCs.

Chromatin loops were then compared before and after ZBTB21 depletion with IAA. Chromatin loops were identified with Mustache software (88), and Hiccups produced results with a similar trend. There were 5606 loops identified in the untreated condition and 6423 loops identified under IAA treatment. Overlap analyses indicated that 2867 loops were induced after ZBTB21 depletion. Density distribution analyses and loop strength changes across loops of different lengths indicated that loop strength increased slightly after ZBTB21 depletion (Figure 4D; Supplementary Figure S3B). GSEA showed that IAA-induced loops were more enriched in larger loops (Figure 4E). Accordingly, the average loop size increased from 417 kb to 455 kb after IAA treatment of ZBTB21-degron cells (Supplementary Figure S3C). We then investigated the features of the new Hi-C loops. The results showed that the new Hi-C loops were more enriched with regulatory elements (such as CTCF/cohesin peaks, enhancers, promoters or polycomb regions) than the less ZBTB21-responsive loops (Figure 4F).

We performed chromatin interaction analyses focusing on *cis*-regulatory elements (enhancer, promoter and CTCF-binding regions) using the ZBTB21 depletion Hi-C dataset. The results showed that ZBTB21 depletion increased the Hi-C interaction signals at promoters, enhancers, CTCF sites and superenhancers (Supplementary Figure S3D). The aggregated analysis of Hi-C interaction frequencies centered on ZBTB21-bound promoters compared with promoters without ZBTB21 binding showed that ZBTB21 depletion significantly increased 3D chromatin interactions at ZBTB21-bound promoters but not at unbound promoters (Figure 4G). For example, at highly transcribed gene loci (including *Lrrc75a*, *Ncor1* and *Pigl*), the Hi-C interaction frequencies increased, and these regions also showed ZBTB21 binding and decreased cohesin binding after the depletion of ZBTB21 (Supplementary Figure S3E). At the *Sall1* and *Chd9* loci, we observed active transcription, ZBTB21 binding and decreased binding of cohesin after ZBTB21 depletion. The interactions between *Sall1* and *Chd9* increased after IAA treatment of ZBTB21-degron cells, which could be independently validated by 4C-seq after ZBTB21 depletion (Figure 4H). Although the effects of ZBTB21 on chromatin interactions are modest, they are statistically significant, which is reminiscent of a previous report that subtle changes in 3D chromatin structures can have a large functional impact, for example, on transcription (104). It seems that all cohesin peaks in these regions,

independent of ZBTB binding, are smaller in IAA-treated samples than in untreated samples, suggesting that ZBTB21 may regulate cohesin loading or non-specifically modulate cohesin chromatin binding.

We next defined ZBTB21-responsive and less-responsive loops based on the existence of ZBTB21 ChIP-seq peaks and down-regulated cohesin peaks (Figure 5A; Supplementary Table S7). The ZBTB21-responsive loops showed higher transcriptional activity (Figure 5A). Then, the loop interaction frequencies displayed in APA plots indicated that ZBTB21 depletion caused greater increases in the loop strength of the ZBTB21-responsive loops than in that of the less-responsive loops (Figure 5B). These results suggested that ZBTB21 binding is associated with decreased cohesin occupancy and increased loop strength after the acute depletion of ZBTB21. Genomic regions with more ZBTB21 binding but less CTCF binding exhibited a greater increase in insulation scores after ZBTB21 depletion. Regions with higher densities of CTCF binding but lower densities of ZBTB21 binding displayed a greater increase in insulation scores after CTCF depletion (Figure 5C). The insulation scores at ZBTB21-bound promoters also showed a more significant increase than those at promoters not bound by ZBTB21 after ZBTB21 depletion (Figure 5C). The insulation scores of ZBTB21-unbound promoters also appeared to be sensitive to ZBTB21 depletion, indicating that the effects of ZBTB21 on 3D chromatin interactions were not specific. The GO enrichment analysis of the DEGs showed that the up-regulated genes were enriched in terms such as ribonucleoprotein complex biogenesis, mRNA processing and histone modification, and the down-regulated genes were enriched in regulation of cellular component size, dendrite development and actin polymerization or depolymerization (Figure 5D). We then explored the correlation between gene expression and the chromatin occupancy of ZBTB21. We found that the expression of ZBTB21-bound genes tended to increase more upon ZBTB21 depletion (Figure 5E). Genes associated with ZBTB21-responsive loops were also more enriched for up-regulated genes than those associated with less-responsive loops (Figure 5F). These results collectively suggested that ZBTB21 is associated with low levels of CTCF and active transcription, which involves the genomic distribution of cohesin and changes in local chromatin interactions between promoters and their regulatory elements within the range of several hundred kilobases. These changes in cohesin distribution and chromatin architecture have been extensively associated with cohesin effects on cell type-specific programs in many different cell systems, including mESCs and human cancer cells, as reported previously (105–107). The biological relevance of these regulatory mechanisms is worthy of further investigation in various cell systems in the future.

Further analyses indicated that 3D chromatin interactions increased more at genes with increased Pol II occupancy, and the ZBTB21-responsive loops also overlapped more with the loops of genes with increased Pol II occupancy. These results indicated that the increased long-range interactions were correlated with transcriptional activation after ZBTB21 activation (Figure 5G). However, it is still unclear whether the increased long-range interactions were

the cause or consequence of transcription activation after ZBTB21 depletion.

### More ZBTB transcription factors are associated with the chromatin binding of cohesin

There are >40 ZBTB transcription factors in mammalian cells (57,58). Inspired by the results obtained for ZBTB21, we hypothesized that ZBTB transcription factors might generally function together with cohesin. We therefore chose more essential ZBTB transcription factors based on their CRISPR scores (94) or cell type-specific functions (59,108), including ZBTB7A, ZBTB7B, ZBTB11 and ZBTB35, and generated GFP-tagged stable HEK293T cell lines with inducible expression of these four factors as well as ZBTB21 (Supplementary Figure S4A). For side-by-side comparison of the protein interacting partners among different ZBTB factors, clones with relatively similar, reasonable (not too high or too low) expression levels were used for downstream analyses (Supplementary Figure S4B). We then performed GFP ChIP-MS analysis of these cell lines. The results showed that these ZBTB transcription factor-interacting partners shared many transcription factors and chromatin regulators (Supplementary Figure S4C). Surprisingly, ZBTB ChIP-MS also identified many components of cohesin, condensin and SMC5/SMC6 complexes that were not found by GFP ChIP-MS performed in a cell line that expressed only GFP as a negative control (Figure 6A; Supplementary Table S2). IP-western blotting experiments further validated the interactions of these components with cohesin (Figure 6B). E2F8 IP did not detect SMC1 [E2F8 was shown not to connect to cohesin components in a previous study (109)]. Furthermore, we expressed these ZBTB factors with a smaller tag (HA tag) in HEK293T cells, and HA antibody immunoprecipitation followed by western blotting showed that these ZBTB factors pulled down cohesin. Cohesin IP in HEK293T cells reciprocally verified these interactions (Supplementary Figure S5A). Together, these results indicate that ZBTB transcription factors specifically interact with cohesin.

ChIP-seq experiments were then performed for ZBTB7A, ZBTB7B, ZBTB11, ZBTB21 and ZBTB35. Their genomic binding sites were distributed at promoters, enhancers and CTCF sites in the genome, and this binding could be validated by ChIP-qPCR with two different antibodies (Supplementary Figure S5B), suggesting that we captured specific binding sites of these ZBTB factors. The analyses of genomic distributions, motifs and GO functions indicated that different ZBTB factors recognized specific DNA sequences and were associated with diverse biological functions (Supplementary Figure S6A–E). Interestingly, the ZBTB21 genomic distribution and GO functions seemed to differ between 293T cells and mESCs. Furthermore, the chromatin binding strength of these ZBTB factors was positively correlated with the expression levels of the target genes, which was consistent with previously known roles of ZBTB in recruiting transcriptional regulators (Supplementary Figure S6A–E) (108,110,111).

Low levels of cohesin were observed at active promoters where ZBTB factors co-localized, but the highest cohesin binding was observed at CTCF sites where ZBTB

factor binding was very weak (Figure 6C–E), and the chromatin binding strength was positively associated with the occupancy of active promoters by cohesin (Figure 6D). Our k-means clustering analyses indicated that the five ZBTB transcription factors appeared to co-localize in the genome but also revealed preferentially occupied sites (Figure 6C). ZBTB11 and ZBTB35 showed preferential binding clusters and stronger co-localization with Pol II, consistent with the recent identification of their functions in transcriptional regulation (112). We predicted that if other ZBTB transcription factors affected cohesin similarly to ZBTB21, then sites showing higher levels of ZBTB factor binding would also show more cohesin binding. Indeed, cohesin binding gradually increased as more ZBTB factors were bound (Figure 6F).

#### Acute ZBTB21 and ZBTB7B depletion leads to a further decrease in cohesin occupancy

We chose to degrade ZBTB21 and ZBTB7B because they showed stronger ChIP-MS signals for the cohesin and condensin subunits than the other ZBTB proteins that we investigated (Figure 6A). We first degraded ZBTB7B in mESCs (Figure 7A) and then performed cohesin ChIP-seq after ZBTB7B depletion. The illustration of meta- and single-gene examples showed that ZBTB depletion led to a slight but significant decrease in cohesin binding at ZBTB7B-binding sites (Figure 7B, C). To directly investigate the functional relationships among ZBTB transcription factors, we generated a double degradation system to simultaneously deplete ZBTB21 and ZBTB7B by knocking degron tags into the C-termini of the two genes in mESCs (Figure 7D). Western blotting confirmed the simultaneous degradation of ZBTB21 and ZBTB7B (Figure 7E).

We first performed chromatin fractionation analyses after ZBTB21 and ZBTB7B depletion. Double degradation caused changes in the protein levels of chromatin-bound SMC1 and SMC4 that were not statistically significant (Figure 7F; Supplementary Figure S7A). Consistently, the total protein levels of the cohesin and condensin components remained similar to those in the whole-cell extract (Supplementary Figure S7B), suggesting that the protein levels of the cohesin and condensin components SMC1 and SMC4, at least, did not obviously change after ZBTB21 depletion. We then performed cohesin ChIP-seq after ZBTB21 and ZBTB7B depletion. The specific gene locus and meta-gene analyses indicated that double depletion caused a greater decrease in cohesin binding than the depletion of ZBTB21 alone for the ZBTB21-bound promoters (Figure 7G, H; Supplementary Figure S7C). However, cohesin binding did not show obvious changes among genes that were not bound by ZBTB21 (Supplementary Figure S7C), which is consistent with the local changes in cohesin binding after ZBTB21 depletion.

## DISCUSSION

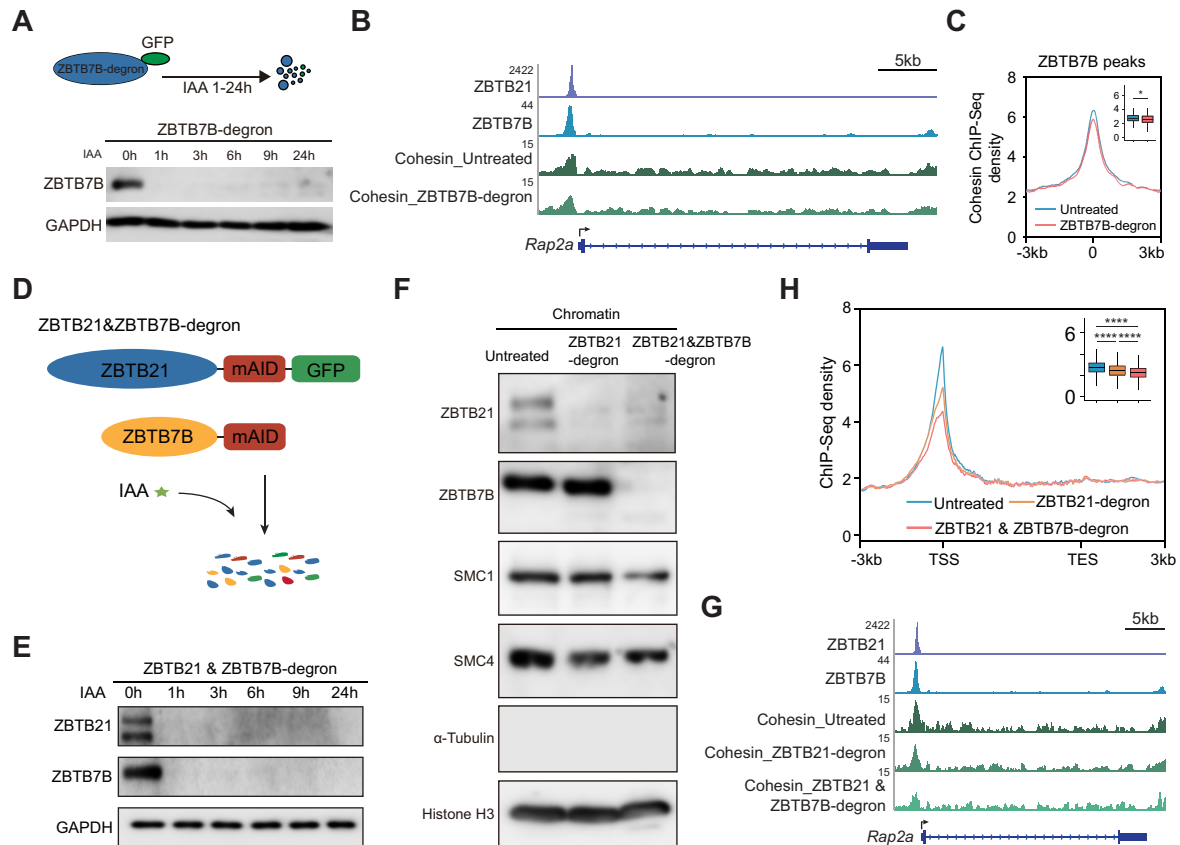
Cohesin plays critical roles in diseases and cell identity (113,114), and its direct regulators have been functionally linked to 3D chromatin structures (25,45,47). The proteomic profiling of cohesin-interacting partners can improve our understanding of the molecular mechanisms

underlying 3D chromatin organization. Here, we report dozens of cohesin-interacting proteins associated with chromatin, and we reveal that ZBTB transcription factors interact with cohesin. The depletion of ZBTB21 leads to a decrease in the occupancy of cohesin and to increased local 3D chromatin interactions for ZBTB21-bound promoters and transcriptional activation. Moreover, the double depletion of ZBTB21 and ZBTB7B causes a further decrease in cohesin chromatin binding. Our results reveal that cohesin interacts and cooperates with ZBTB transcription factors to contribute to 3D chromatin architecture and gene expression regulation.

The well-established CTCF-cohesin complex is responsible for the formation of large-scale chromatin structures, such as TADs. However, the molecular basis of small-scale chromatin interactions (such as chromatin loops associated with active genes and sub-TAD structures) has remained unclear. Many cofactors and chromatin regulators have been implicated in the formation of biomolecular condensates to mediate local chromatin structures (20,115,116). However, recent experiments involving the chemical inhibition and rapid protein degradation of these factors did not reveal obvious effects on small-scale chromatin structures (23,24,117,118). In this study, we provide evidence indicating that ZBTB21 is both physically and functionally linked to cohesin: (i) cohesin ChIP-MS analysis identified ZBTB21 (Figure 1C, G); (ii) ZBTB21 ChIP-MS and IP-western blotting confirmed the interaction of ZBTB21 with cohesin (Figure 2A, B); (iii) ZBTB21 ChIP-seq showed the co-occupancy of cohesin with ZBTB21 at promoters (Figure 2C, D); (iv) acute ZBTB21 depletion decreased cohesin and NIPBL binding at ZBTB21-bound genes (Figure 3B, E; Supplementary Figure S2E–G); and (v) rapid depletion of ZBTB21 enhanced the 3D chromatin interactions of ZBTB21-bound active regions (Figure 5A, B). The increases in chromatin interactions in certain regions might be due to the clustering of chromatin regions modified with similar histone marks after the removal of ZBTB21/cohesin-mediated insulation, which is consistent with a model of genome compartmentalization (119,120), and the induced loops were indeed enriched in H3K4me3-, H3K27me3- and H3K27ac-modified regions (Figure 4F). Our findings were also consistent with the previous observation of strengthened interactions between superenhancers after cohesin depletion (26).

We obtained 25 kb resolution Hi-C maps in the current study, which is not high enough. This could be why we did not observe robust effects on 3D chromatin loop changes. Higher resolution chromatin interaction mapping (such as Micro-C) and elucidation of the detailed protein–protein interactions between cohesin and ZBTB21 would be valuable to clarify the molecular mechanism of ZBTB21 function in the future. A previous study showed that transcription factors are enriched at local cohesin-binding sites, and the pioneer transcription factors OCT4 and SOX2 create an open chromatin conformation for cohesin binding (90). We identified many transcription factors in cohesin ChIP-MS preparations and showed decreased cohesin binding at ZBTB21-binding sites after acute depletion of ZBTB21, which is consistent with the results described by Liu *et al.*, indicating that ZBTB21 may also function as a pioneer





**Figure 7.** ZBTB21 and ZBTB7B are redundantly responsible for the chromatin binding of cohesin. (A) Western blotting analysis of ZBTB7B degradation in ZBTB7B-deg GFP mESCs after IAA treatment at different time points. ZBTB7B was identified with a GFP antibody because it was endogenously GFP tagged, and GAPDH served as a loading control. (B) Snapshots of cohesin ChIP-seq signals at the *Rap2a* locus in untreated and ZBTB7B-depleted conditions. (C) The meta-gene displaying the cohesin ChIP-seq signals at  $\pm 3$  kb around the ZBTB7B binding peaks under untreated and ZBTB7B depletion (IAA-treated) conditions. The lines indicate the mean cohesin ChIP-seq signal under each condition. The boxplot (inset) shows the changes in the ChIP-seq signal at  $\pm 100$  bp around ZBTB7B peaks upon ZBTB7B degradation. Significance was determined by the Wilcoxon test ( $*P < 0.05$ ). (D) Diagrammatic illustration of the double degradation of ZBTB21 and ZBTB7B. (E) Western blotting analysis of ZBTB21 and ZBTB7B protein degradation in ZBTB21 and ZBTB7B double-deg mESCs. ZBTB21 was detected with a GFP antibody because it was endogenously GFP tagged, ZBTB7B was identified with endogenous antibodies, and GAPDH served as a loading control. (F) Western blotting analysis of the chromatin fractions of targeted proteins in the ZBTB21-deg and ZBTB21 and ZBTB7B-deg cell lines under untreated and (6 h) IAA-treated conditions. SMC1 is a cohesin component. SMC4 is a condensin complex subunit.  $\alpha$ -Tubulin is a cytoplasmic marker. Histone H3 is a chromatin marker. (G) Snapshots of cohesin ChIP-seq signals at the *Rap2a* locus in untreated, ZBTB21-depleted and ZBTB21- and ZBTB7B-double-depleted conditions. (H) Meta-gene analysis of cohesin ChIP-seq density at ZBTB21-bound genes upon ZBTB21 or ZBTB21 and ZBTB7B degradation. The lines indicate the mean cohesin ChIP-seq signal under each condition. The boxplot (inset) shows the changes in the ChIP-seq signal at  $\pm 100$  bp around TSS regions upon ZBTB21 or ZBTB21 and ZBTB7B degradation. Significance was determined by the Wilcoxon test ( $***P < 0.0001$ ).

factor (similar to OCT4 and SOX2) to facilitate cohesin binding.

ZBTB transcription factors have been reported to function in several critical developmental processes and diseases. For example, many ZBTB transcription factors, including ZBTB7A, ZBTB7B and ZBTB35, are required for the proliferation and differentiation of T cells (59,121); ZBTB7A is involved in tumorigenesis and pluripotency control through signaling pathways or chromatin remodelers (108,122,123); ZBTB7B drives brown fat development via long non-coding RNAs (111); ZBTB11 regulates neutrophil development, and its mutation causes intellectual disability (124,125); and ZBTB21 is associated with the pathogenesis of congenital heart defects in Down syndrome (126). Their effects on 3D chromatin interactions are probably cell type specific because the distributions of ZBTB ChIP-seq peaks in mESCs and HEK293T cells are different. It would be interesting to

deplete these ZBTB factors during development or disease progression in future research. The molecular mechanisms of ZBTB functions have mostly been investigated in the context of transcriptional regulation. Even though the chromatin interaction analyses were performed within 6 h, these secondary effects on mature RNAs could be limited. We could not eliminate the possibility that misregulated gene expression may also contribute to their effects on cohesin binding and 3D chromatin interactions. ZBTB21 has zinc finger domains that bind DNA, and its BTB domains mediate protein–protein interactions. Moreover, *Drosophila* insulator proteins such as Mod (mdg4) and Cp190 are also ZBTB transcription factors and have been shown to function as regulators of insulator activities (52–54). Previous studies have shown that ZBTB factors are critical for the cell fate determination of specific lineages. Although the effects of ZBTB21 depletion on cohesin chromatin binding

are minor, we envisaged that ZBTB21 proteins may function as insulator proteins, similar to previously reported ZBTB factors in *Drosophila*, which might prevent cohesin extrusion at critical chromatin regulatory elements. This potential mechanism, together with transcriptional repressor activity, would be important for the specific gene expression program during different developmental processes.

## DATA AVAILABILITY

All the data that support the findings of this study are available from the corresponding authors upon reasonable request. Raw sequencing data can be found in the GEO database: GSE184272. Access to the mass spectrometry dataset is available in ProteomeXchange: PXD028860, PXD036004, PXD035893. Raw blot, gel and microscopy image data can be found in Mendeley data: DOI: 10.17632/cwpb6kmxv7.2

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank the members of the Ji laboratory for engaging in helpful discussions. We thank Dr J.H. and B.Y.L. for their help with the analyses conducted in this study. We thank the National Center for Protein Sciences at Peking University in Beijing, China, and Ms L.Y.D. and H.Y. for their help with flow cytometry sorting, D.L. for help with mass spectrometry, and G.L.L. for help with biochemical assays.

*Author contributions:* X.J. conceived and supervised the project. R.W. performed most of the experiments. Q.Q.X. performed most of the bioinformatic analyses. C.L.W. helped with the mass spectrometry analyses. K.T. generated the CTCF ChIP-seq dataset and H.W. generated the ATAC-seq dataset. X.J. wrote the manuscript with input from R.W. and Q.Q.X., and with help from the other authors.

## FUNDING

National Natural Science Foundation of China [31871309 and 32170569]; the Ministry of Science and Technology of China [2017YFA0506600]; the Qidong-SLS Innovation Fund; the Peking-Tsinghua Center for Life Sciences; and the Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education at Peking University School of Life Sciences [to X.J.].

*Conflict of interest statement.* None declared.

## REFERENCES

- Rowley, M.J. and Corces, V.G. (2018) Organizational principles of 3D genome architecture. *Nat. Rev. Genet.*, **19**, 789–800.
- Hildebrand, E.M. and Dekker, J. (2020) Mechanisms and functions of chromosome compartmentalization. *Trends Biochem. Sci.*, **45**, 385–396.
- Yu, M. and Ren, B. (2017) The three-dimensional organization of mammalian genomes. *Annu. Rev. Cell Dev. Biol.*, **33**, 265–289.
- Ghirlando, R. and Felsenfeld, G. (2016) CTCF: making the right connections. *Genes Dev.*, **30**, 881–891.
- Hnisz, D., Day, D.S. and Young, R.A. (2016) Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell*, **167**, 1188–1200.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O. *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**, 289–293.
- Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H. *et al.* (2009) An oestrogen-receptor- $\alpha$ -bound human chromatin interactome. *Nature*, **462**, 58–64.
- Baù, D., Sanyal, A., Lajoie, B.R., Capriotti, E., Byron, M., Lawrence, J.B., Dekker, J. and Marti-Renom, M.A. (2011) The three-dimensional folding of the  $\alpha$ -globin gene domain reveals formation of chromatin globules. *Nat. Struct. Mol. Biol.*, **18**, 107–114.
- Castells-Garcia, A., Ed-Daoui, I., González-Almela, E., Vicario, C., Ottetstrom, J., Lakadamyali, M., Neguembor, M.V. and Cosma, M.P. (2022) Super resolution microscopy reveals how elongating RNA polymerase II and nascent RNA interact with nucleosome clutches. *Nucleic Acids Res.*, **50**, 175–190.
- Dimitrov, S.I., Russanova, V.R. and Pashev, I.G. (1987) The globular domain of histone H5 is internally located in the 30 nm chromatin fiber: an immunochemical study. *EMBO J.*, **6**, 2387–2392.
- Liu, S., Lin, X. and Zhang, B. (2022) Chromatin fiber breaks into clutches under tension and crowding. *Nucleic Acids Res.*, **50**, 9738–9747.
- Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M. and Cosma, M.P. (2015) Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell*, **160**, 1145–1158.
- Vernimmen, D. and Bickmore, W.A. (2015) The hierarchy of transcriptional activation: from enhancer to promoter. *Trends Genet.*, **31**, 696–708.
- Schoenfelder, S. and Fraser, P. (2019) Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.*, **20**, 437–455.
- Pombo, A. and Dillon, N. (2015) Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.*, **16**, 245–257.
- Ren, B. and Yue, F. (2015) Transcriptional enhancers: bridging the genome and phenome. *Cold Spring Harbor Symp. Quant. Biol.*, **80**, 17–26.
- Zabidi, M.A. and Stark, A. (2016) Regulatory enhancer–core–promoter communication via transcription factors and cofactors. *Trends Genet.*, **32**, 801–814.
- Levine, M., Cattoglio, C. and Tjian, R. (2014) Looping back to leap forward: transcription enters a new era. *Cell*, **157**, 13–25.
- Smallwood, A. and Ren, B. (2013) Genome organization and long-range regulation of gene expression by enhancers. *Curr. Opin. Cell Biol.*, **25**, 387–394.
- Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K. and Sharp, P.A. (2017) A phase separation model for transcriptional control. *Cell*, **169**, 13–23.
- Benabdallah, N.S. and Bickmore, W.A. (2015) Regulatory domains and their mechanisms. *Cold Spring Harbor Symp. Quant. Biol.*, **80**, 45–51.
- Mirny, L.A., Imakaev, M. and Abdennur, N. (2019) Two major mechanisms of chromosome organization. *Curr. Opin. Cell Biol.*, **58**, 142–152.
- El Khattabi, L., Zhao, H., Kalchschmidt, J., Young, N., Jung, S., Van Blerkom, P., Kieffer-Kwon, P., Kieffer-Kwon, K.R., Park, S., Wang, X. *et al.* (2019) A pliable mediator acts as a functional rather than an architectural bridge between promoters and enhancers. *Cell*, **178**, 1145–1158.
- Jaeger, M.G., Schwalb, B., Mackowiak, S.D., Velychko, T., Hanzl, A., Imrichova, H., Brand, M., Agerer, B., Chorn, S., Nabet, B. *et al.* (2020) Selective Mediator dependence of cell-type-specifying transcription. *Nat. Genet.*, **52**, 719–727.
- Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A. and Bruneau, B.G. (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, **169**, 930–944.

26. Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D. *et al.* (2017) Cohesin loss eliminates all loop domains. *Cell*, **171**, 305–320.
27. Hsieh, T.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Darzacq, X. and Tjian, R. (2022) Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nat. Genet.*, **54**, 1919–1932.
28. Merckenschlager, M. and Nora, E.P. (2016) CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.*, **17**, 17–43.
29. Dekker, J. and Mirny, L. (2016) The 3D genome as moderator of chromosomal communication. *Cell*, **164**, 1110–1121.
30. Wood, A.J., Severson, A.F. and Meyer, B.J. (2010) Condensin and cohesin complexity: the expanding repertoire of functions. *Nat. Rev. Genet.*, **11**, 391–404.
31. Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T. *et al.* (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, **132**, 422–433.
32. Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T. *et al.* (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**, 796–801.
33. Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S. *et al.* (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**, 430–435.
34. Davidson, I.F. and Peters, J.M. (2021) Genome folding through loop extrusion by SMC complexes. *Nat. Rev. Mol. Cell Biol.*, **22**, 445–464.
35. Seitan, V.C. and Merckenschlager, M. (2012) Cohesin and chromatin organisation. *Curr. Opin. Genet. Dev.*, **22**, 93–100.
36. Shi, Z., Gao, H., Bai, X.C. and Yu, H. (2020) Cryo-EM structure of the human cohesin–NIPBL–DNA complex. *Science*, **368**, 1454–1459.
37. Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yáñez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B. *et al.* (2017) The cohesin release factor WAPL restricts chromatin loop extension. *Cell*, **169**, 693–707.
38. Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L. *et al.* (2017) Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, **551**, 51–56.
39. Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J. *et al.* (2017) Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.*, **36**, 3573–3599.
40. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J. and Yu, H. (2019) Human cohesin compacts DNA by loop extrusion. *Science*, **366**, 1345–1349.
41. Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L. *et al.* (2017) YY1 is a structural regulator of enhancer–promoter loops. *Cell*, **171**, 1573–1588.
42. Beagan, J.A., Duong, M.T., Titus, K.R., Zhou, L., Cao, Z., Ma, J., Lachanski, C.V., Gillis, D.R. and Phillips-Cremins, J.E. (2017) YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res.*, **27**, 1139–1152.
43. Bailey, S.D., Zhang, X., Desai, K., Aid, M., Corradin, O., Cowper-Sal Lari, R., Akhtar-Zaidi, B., Scacheri, P.C., Haibe-Kains, B. and Lupien, M. (2015) ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. *Nat. Commun.*, **2**, 6186.
44. Kaaij, L.J.T., Mohn, F., van der Weide, R.H., de Wit, E. and Bühler, M. (2019) The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse. *Cell*, **178**, 1437–1451.
45. Xiao, T., Li, X. and Felsenfeld, G. (2021) The Myc-associated zinc finger protein (MAZ) works together with CTCF to control cohesin positioning and genome organization. *Proc. Natl Acad. Sci. USA*, **118**, e2023127118.
46. Hu, G., Dong, X., Gong, S., Song, Y., Hutchins, A.P. and Yao, H. (2020) Systematic screening of CTCF binding partners identifies that BHLHE40 regulates CTCF genome-wide distribution and long-range chromatin interactions. *Nucleic Acids Res.*, **48**, 9606–9620.
47. Justice, M., Carico, Z.M., Stefan, H.C. and Dowen, J.M. (2020) A WIZ/cohesin/CTCF complex anchors DNA loops to define gene expression and cell identity. *Cell Rep.*, **31**, 107503.
48. Phanstiel, D.H., Van Bortle, K., Spacek, D., Hess, G.T., Shamim, M.S., Machol, I., Love, M.I., Aiden, E.L., Bassik, M.C. and Snyder, M.P. (2017) Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell*, **67**, 1037–1048.
49. Xiang, J.F. and Corces, V.G. (2021) Regulation of 3D chromatin organization by CTCF. *Curr. Opin. Genet. Dev.*, **67**, 33–40.
50. Dong, X., Guo, R., Ji, T., Zhang, J., Xu, J., Li, Y., Sheng, Y., Wang, Y., Fang, K., Wen, Y. *et al.* (2022) YY1 safeguard multidimensional epigenetic landscape associated with extended pluripotency. *Nucleic Acids Res.*, **50**, 12019–12038.
51. Song, Y., Liang, Z., Zhang, J., Hu, G., Wang, J., Li, Y., Guo, R., Dong, X., Babarinde, I.A., Ping, W. *et al.* (2022) CTCF functions as an insulator for somatic genes and a chromatin remodeler for pluripotency genes during reprogramming. *Cell Rep.*, **39**, 110626.
52. Gerasimova, T.I., Gdula, D.A., Gerasimov, D.V., Simonova, O. and Corces, V.G. (1995) A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell*, **82**, 587–597.
53. Pai, C.Y., Lei, E.P., Ghosh, D. and Corces, V.G. (2004) The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol. Cell*, **16**, 737–748.
54. Van Bortle, K., Nichols, M.H., Li, L., Ong, C.T., Takenaka, N., Qin, Z.S. and Corces, V.G. (2014) Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.*, **15**, R82.
55. Bag, I., Chen, S., Rosin, L.F., Chen, Y., Liu, C.Y., Yu, G.Y. and Lei, E.P. (2021) M1BP cooperates with CP190 to activate transcription at TAD borders and promote chromatin insulator activity. *Nat. Commun.*, **12**, 4170.
56. Kaushal, A., Mohana, G., Dorier, J., Özdemir, I., Omer, A., Cousin, P., Semenova, A., Taschner, M., Dergai, O., Marzetta, F. *et al.* (2021) CTCF loss has limited effects on global genome architecture in Drosophila despite critical regulatory functions. *Nat. Commun.*, **12**, 1011.
57. Maeda, T. (2016) Regulation of hematopoietic development by ZBTB transcription factors. *Int. J. Hematol.*, **104**, 310–323.
58. Siggs, O.M. and Beutler, B. (2012) The BTB-ZF transcription factors. *Cell Cycle*, **11**, 3358–3369.
59. Cheng, Z.Y., He, T.T., Gao, X.M., Zhao, Y. and Wang, J. (2021) ZBTB transcription factors: key regulators of the development, differentiation and effector function of T cells. *Front. Immunol.*, **12**, 713294.
60. Zhu, C., Chen, G., Zhao, Y., Gao, X.M. and Wang, J. (2018) Regulation of the development and function of B cells by ZBTB transcription factors. *Front. Immunol.*, **9**, 580.
61. Lee, S.U. and Maeda, T. (2012) POK/ZBTB proteins: an emerging family of proteins that regulate lymphoid development and function. *Immunol. Rev.*, **247**, 107–119.
62. Kelly, K.F. and Daniel, J.M. (2006) POZ for effect—POZ-ZF transcription factors in cancer and development. *Trends Cell Biol.*, **16**, 578–587.
63. Natsume, T., Kiyomitsu, T., Saga, Y. and Kanemaki, M.T. (2016) Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep.*, **15**, 210–218.
64. Jiang, Y., Huang, J., Lun, K., Li, B., Zheng, H., Li, Y., Zhou, R., Duan, W., Wang, C., Feng, Y. *et al.* (2020) Genome-wide analyses of chromatin interactions after the loss of Pol I, Pol II, and Pol III. *Genome Biol.*, **21**, 158.
65. Ji, X., Dadon, D.B., Abraham, B.J., Lee, T.I., Jaenisch, R., Bradner, J.E. and Young, R.A. (2015) Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proc. Natl Acad. Sci. USA*, **112**, 3841–3846.
66. Huang, J., Jiang, Y., Zheng, H. and Ji, X. (2020) BAT Hi-C maps global chromatin interactions in an efficient and economical way. *Methods*, **170**, 38–47.
67. Jain, S., Ba, Z., Zhang, Y., Dai, H.Q. and Alt, F.W. (2018) CTCF-binding elements mediate accessibility of RAG substrates during chromatin scanning. *Cell*, **174**, 102–116.



68. Peng, X., Wu, J., Brunmeir, R., Kim, S.Y., Zhang, Q., Ding, C., Han, W., Xie, W. and Xu, F. (2015) TELP, a sensitive and versatile library construction method for next-generation sequencing. *Nucleic Acids Res.*, **43**, e35.
69. Wang, C., Xu, Q., Zhang, X., Day, D.S., Abraham, B.J., Lun, K., Chen, L., Huang, J. and Ji, X. (2022) BRD2 interconnects with BRD3 to facilitate Pol II transcription initiation and elongation to prime promoters for cell differentiation. *Cell. Mol. Life Sci.*, **79**, 338.
70. Yang, B., Li, B., Jia, L., Jiang, Y., Wang, X., Jiang, S., Du, S., Ji, X. and Yang, P. (2020) 3D landscape of Hepatitis B virus interactions with human chromatin. *Cell Discov.*, **6**, 95.
71. Zhang, X., Smits, A.H., van Tilburg, G.B., Ovaa, H., Huber, W. and Vermeulen, M. (2018) Proteome-wide identification of ubiquitin interactions using UbiA-MS. *Nat. Protoc.*, **13**, 530–550.
72. Deplancke, B., Alpern, D. and Gardeux, V. (2016) The genetics of transcription factor DNA binding variation. *Cell*, **166**, 538–554.
73. Du, X. and Xiao, R. (2020) An emerging role of chromatin-interacting RNA-binding proteins in transcription regulation. *Essays Biochem.*, **64**, 907–918.
74. Yu, G., Wang, L.G., Han, Y. and He, Q.Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**, 284–287.
75. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **2011**, 17.
76. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
77. Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F. and Manke, T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
78. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
79. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I. and Young, R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
80. Yu, G., Wang, L.G. and He, Q.Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.
81. Li, G., Chen, Y., Snyder, M.P. and Zhang, M.Q. (2017) ChIA-PET2: a versatile and flexible pipeline for ChIA-PET data analysis. *Nucleic Acids Res.*, **45**, e4.
82. Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J. and Barillot, E. (2015) HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.*, **16**, 259.
83. Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A. *et al.* (2017) Multiscale 3D genome rewiring during mouse neural development. *Cell*, **171**, 557–572.
84. Ji, X., Dadon, D.B., Powell, B.E., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., Hnisz, D., Pegoraro, G., Lee, T.I. *et al.* (2016) 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell*, **18**, 262–275.
85. Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y. *et al.* (2017) Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature*, **547**, 232–235.
86. Abdennur, N. and Mirny, L.A. (2020) Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics*, **36**, 311–316.
87. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S. and Aiden, E.L. (2016) Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.*, **3**, 99–101.
88. Roayaei Ardakany, A., Gezer, H.T., Lonardi, S. and Ay, F. (2020) Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation. *Genome Biol.*, **21**, 256.
89. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA*, **102**, 15545–15550.
90. Liu, N.Q., Maresca, M., van den Brand, T., Braccioli, L., Schijns, M., Teunissen, H., Bruneau, B.G., Nora, E.P. and de Wit, E. (2021) WAPL maintains a cohesin loading cycle to preserve cell-type-specific distal gene regulation. *Nat. Genet.*, **53**, 100–109.
91. Ramírez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F. and Manke, T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
92. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
93. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
94. Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S. and Sabatini, D.M. (2015) Identification and characterization of essential genes in the human genome. *Science*, **350**, 1096–1101.
95. Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K. *et al.* (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*, **159**, 374–387.
96. Davidson, L., Francis, L., Cordiner, R.A., Eaton, J.D., Estell, C., Macias, S., Cáceres, J.F. and West, S. (2019) Rapid depletion of DIS3, EXOSC10, or XRN2 reveals the immediate impact of exoribonucleolysis on nuclear RNA metabolism and transcriptional control. *Cell Rep.*, **26**, 2779–2791.
97. Szczepinska, T., Kalisiak, K., Tomecki, R., Labno, A., Borowski, L.S., Kulinski, T.M., Adamska, D., Kosinska, J. and Dziembowski, A. (2015) DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. *Genome Res.*, **25**, 1622–1633.
98. Kim, J.S., He, X., Liu, J., Duan, Z., Kim, T., Gerard, J., Kim, B., Pillai, M.M., Lane, W.S., Noble, W.S. *et al.* (2019) Systematic proteomics of endogenous human cohesin reveals an interaction with diverse splicing factors and RNA-binding proteins required for mitotic progression. *J. Biol. Chem.*, **294**, 8760–8772.
99. Panigrahi, A.K., Zhang, N., Otta, S.K. and Pati, D. (2012) A cohesin–RAD21 interactome. *Biochem. J.*, **442**, 661–670.
100. Chen, D. and Lei, E.P. (2019) Function and regulation of chromatin insulators in dynamic genome organization. *Curr. Opin. Cell Biol.*, **58**, 61–68.
101. Cubeñas-Potts, C. and Corces, V.G. (2015) Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett.*, **589**, 2923–2930.
102. Wang, J., Kudoh, J., Takayanagi, A. and Shimizu, N. (2005) Novel human BTB/POZ domain-containing zinc finger protein ZNF295 is directly associated with ZFP161. *Biochem. Biophys. Res. Commun.*, **327**, 615–627.
103. Maresca, M., van den Brand, T., Li, H., Teunissen, H., Davies, J. and de Wit, E. (2022) Pioneer activity distinguishes activating from non-activating pluripotency transcription factor binding sites. *bioRxiv* doi: <https://doi.org/10.1101/2022.07.27.501606>, 27 July 2022, preprint: not peer reviewed.
104. Xiao, J.Y., Hafner, A. and Boettiger, A.N. (2021) How subtle changes in 3D structure can create large changes in transcription. *Elife*, **10**, e64320.
105. Cuadrado, A., Giménez-Llorente, D., Kojic, A., Rodríguez-Corsino, M., Cuartero, Y., Martín-Serrano, G., Gómez-López, G., Martí-Renom, M.A. and Losada, A. (2019) Specific contributions of cohesin-SA1 and cohesin-SA2 to TADs and polycomb domains in embryonic stem cells. *Cell Rep.*, **27**, 3500–3510.
106. Viny, A.D., Bowman, R.L., Liu, Y., Lavallée, V.P., Eisman, S.E., Xiao, W., Durham, B.H., Navitski, A., Park, J., Braunstein, S. *et al.* (2019) Cohesin members Stag1 and Stag2 display distinct roles in chromatin accessibility and topological control of HSC self-renewal and differentiation. *Cell Stem Cell*, **25**, 682–696.
107. Richart, L., Lapi, E., Pancaldi, V., Cuenca-Ardura, M., Pau, E.C., Madrid-Mencia, M., Neyret-Kahn, H., Radvanyi, F., Rodríguez, J.A., Cuartero, Y. *et al.* (2021) STAG2 loss-of-function affects short-range genomic contacts and modulates the basal–luminal transcriptional

- program of bladder cancer cells. *Nucleic Acids Res.*, **49**, 11005–11021.
108. Yu, S., Zhou, C., Cao, S., He, J., Cai, B., Wu, K., Qin, Y., Huang, X., Xiao, L., Ye, J. *et al.* (2020) BMP4 resets mouse epiblast stem cells to naive pluripotency through ZBTB7A/B-mediated chromatin remodelling. *Nat. Cell Biol.*, **22**, 651–662.
  109. Yan, J., Enge, M., Whittington, T., Dave, K., Liu, J., Sur, I., Schmierer, B., Jolma, A., Kivioja, T., Taipale, M. *et al.* (2013) Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell*, **154**, 801–813.
  110. Varier, R.A., Carrillo de Santa Pau, E., van der Groep, P., Lindeboom, R.G., Matarese, F., Mensinga, A., Smits, A.H., Edupuganti, R.R., Baltissen, M.P., Jansen, P.W. *et al.* (2016) Recruitment of the mammalian histone-modifying EMSY complex to target genes is regulated by ZNF131. *J. Biol. Chem.*, **291**, 7313–7324.
  111. Li, S., Mi, L., Yu, L., Yu, Q., Liu, T., Wang, G.X., Zhao, X.Y., Wu, J. and Lin, J.D. (2017) Zbtb7b engages the long noncoding RNA Blnc1 to drive brown and beige fat development and thermogenesis. *Proc. Natl Acad. Sci. USA*, **114**, E7111–E7120.
  112. Garipler, G., Lu, C., Morrissey, A., Lopez-Zepeda, L.S., Vidal, S.E., Aydin, B., Stadtfeld, M., Ohler, U., Mahony, S., Sanjana, N.E. *et al.* (2020) The BTB transcription factors ZBTB11 and ZFP131 maintain pluripotency by pausing POL II at pro-differentiation genes. bioRxiv doi: <https://doi.org/10.1101/2020.11.23.391771>, 23 November 2020, preprint: not peer reviewed.
  113. Carico, Z.M., Stefan, H.C., Justice, M., Yimit, A. and Down, J.M. (2021) A cohesin cancer mutation reveals a role for the hinge domain in genome organization and gene expression. *PLoS Genet.*, **17**, e1009435.
  114. Rittenhouse, N.L., Carico, Z.M., Liu, Y.F., Stefan, H.C., Arruda, N.L., Zhou, J. and Down, J.M. (2021) Functional impact of cancer-associated cohesin variants on gene expression and cellular identity. *Genetics*, **217**, iyab025.
  115. Ahn, J.H., Davis, E.S., Daugird, T.A., Zhao, S., Quiroga, I.Y., Uryu, H., Li, J., Storey, A.J., Tsai, Y.H., Keeley, D.P. *et al.* (2021) Phase separation drives aberrant chromatin looping and cancer development. *Nature*, **595**, 591–595.
  116. Wang, J., Yu, H., Ma, Q., Zeng, P., Wu, D., Hou, Y., Liu, X., Jia, L., Sun, J., Chen, Y. *et al.* (2021) Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions. *Cell Stem Cell*, **28**, 1868–1868.
  117. Crump, N.T., Ballabio, E., Godfrey, L., Thorne, R., Repapi, E., Kerry, J., Tapia, M., Hua, P., Lagerholm, C., Filippakopoulos, P. *et al.* (2021) BET inhibition disrupts transcription but retains enhancer–promoter contact. *Nat. Commun.*, **12**, 223.
  118. Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Darzacq, X. and Tjian, R. (2021) Enhancer-promoter interactions and transcription are maintained upon acute loss of CTCF, cohesin, WAPL, and YY1. bioRxiv doi: <https://doi.org/10.1101/2021.07.14.452365>, 14 July 2021, preprint: not peer reviewed.
  119. Nichols, M.H. and Corces, V.G. (2021) Principles of 3D compartmentalization of the human genome. *Cell Rep.*, **35**, 109330.
  120. Rowley, M.J., Nichols, M.H., Lyu, X., Ando-Kuri, M., Rivera, I.S.M., Hermetz, K., Wang, P., Ruan, Y. and Corces, V.G. (2017) Evolutionarily conserved principles predict 3D chromatin organization. *Mol. Cell*, **67**, 837–852.
  121. Iguchi, T., Aoki, K., Ikawa, T., Taoka, M., Taya, C., Yoshitani, H., Toma-Hirano, M., Koiwai, O., Isobe, T., Kawamoto, H. *et al.* (2015) BTB-ZF protein Znf131 regulates cell growth of developing and mature T cells. *J. Immunol.*, **195**, 982–993.
  122. Singh, A.K., Verma, S., Kushwaha, P.P., Prajapati, K.S., Shuaib, M., Kumar, S. and Gupta, S. (2021) Role of ZBTB7A zinc finger in tumorigenesis and metastasis. *Mol. Biol. Rep.*, **48**, 4703–4719.
  123. Liu, X.S., Liu, Z., Gerarduzzi, C., Choi, D.E., Ganapathy, S., Pandolfi, P.P. and Yuan, Z.M. (2016) Somatic human ZBTB7A zinc finger mutations promote cancer progression. *Oncogene*, **35**, 3071–3078.
  124. Fattahi, Z., Sheikh, T.I., Musante, L., Rasheed, M., Taskiran, I.I., Harripaul, R., Hu, H., Kazeminasab, S., Alam, M.R., Hosseini, M. *et al.* (2018) Biallelic missense variants in ZBTB11 can cause intellectual disability in humans. *Hum. Mol. Genet.*, **27**, 3177–3188.
  125. Keightley, M.C., Carradice, D.P., Layton, J.E., Pase, L., Bertrand, J.Y., Wittig, J.G., Dakic, A., Badrock, A.P., Cole, N.J., Traver, D. *et al.* (2017) The Pu.1 target gene Zbtb11 regulates neutrophil development through its integrase-like HHCC zinc finger. *Nat. Commun.*, **8**, 14911.
  126. Sailani, M.R., Makrythanasis, P., Valsesia, A., Santoni, F.A., Deutsch, S., Popadin, K., Borel, C., Migliavacca, E., Sharp, A.J., Duriaux Sail, G. *et al.* (2013) The complex SNP and CNV genetic architecture of the increased risk of congenital heart defects in Down syndrome. *Genome Res.*, **23**, 1410–1421.
  127. Down, J.M., Bilodeau, S., Orlando, D.A., Hübner, M.R., Abraham, B.J., Spector, D.L. and Young, R.A. (2013) Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. *Stem Cell Rep.*, **1**, 371–378.
  128. Sigova, A.A., Abraham, B.J., Ji, X., Molin, B., Hannett, N.M., Guo, Y.E., Jangi, M., Giallourakis, C.C., Sharp, P.A. and Young, R.A. (2015) Transcription factor trapping by RNA in gene regulatory elements. *Science*, **350**, 978–981.
  129. Haering, C.H. and Gruber, S. (2016) SnapShot: SMC protein complexes part I. *Cell*, **164**, 326–326.