G&I Genomics & Informatics

REVIEW ARTICLE

CTCF, Cohesin, and Chromatin in Human Cancer

Sang-Hyun Song[1], Tae-You Kim[1,2,3]*

[1]Cancer Genomics Research Laboratory, Cancer Research Institute, Seoul National University, Seoul 03080, Korea, [2]Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 03080, Korea, [3]Department of Internal Medicine, Seoul National University Hospital, Seoul 03080, Korea

It is becoming increasingly clear that eukaryotic genomes are subjected to higher-order chromatin organization by the CCCTC-binding factor/cohesin complex. Their dynamic interactions in three dimensions within the nucleus regulate gene transcription by changing the chromatin architecture. Such spatial genomic organization is functionally important for the spatial disposition of chromosomes to control cell fate during development and differentiation. Thus, the dysregulation of proper long-range chromatin interactions may influence the development of tumorigenesis and cancer progression.

Keywords: cohesins, CTCF, gene expression regulation, higher-order chromatin structure, topologically associated domain

Introduction

Genomes form higher-order chromatin structures during cellular development and differentiation [1-4]. The spatial folding of chromosomes and their organization in the nucleus have profound effects on gene expression [5, 6]. It is now well established that chromosomal architecture is largely mediated by the CCCTC-binding factor (CTCF)/cohesin complex [7, 8]. CTCF, a zinc finger DNA-binding protein that functions in transcriptional repression, activation, and as an insulator that interferes with enhancer–promoter interactions [9], is needed for the recruitment of cohesin to chromatin [10]. In addition to its major influence on sister chromatid cohesion [11], cohesin affects gene transcription by facilitating long-range interactions between members of many developmentally regulated gene families [12-14].

Interphase chromosomes of higher eukaryotes are subdivided into evolutionarily conserved topologically associated domains (TADs) in the three-dimensional space of the nucleus (Fig. 1) [15, 16]. TADs are defined as self-associating chromosome segments enclosed by a chromatin loop in the megabase range detected by Hi-C methods [17]. TADs show a high frequency of interactions within domains and a low frequency of interactions among different domains [15, 17]. They are partitioned into several subcompartments related to their gene expression patterns and maintained across cell types, suggesting that TADs shape the regulatory landscape of the genome during development [6, 15, 16]. Interestingly, in the boundaries of TADs, there is an abundance of the architectural protein CTCF and cohesin [6, 15], suggesting that these proteins have a role in establishing the topological boundaries [17].
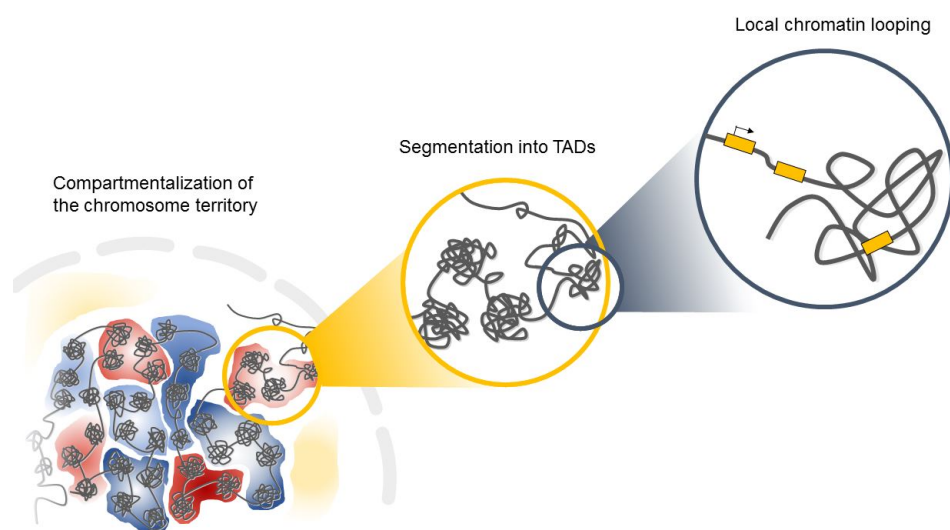
Increasing evidence from recent studies has indicated that genomic instability is spatially related to higher-order chromatin organization in cancer cells [18]. Previous studies of *c-Myc, BCL*, and immunoglobulin heavy-chain (*Igh*) (the most common translocation partners in various B-cell lymphomas) showed that these loci are preferentially localized in close spatial juxtaposition to each other in normal B cells [19, 20]. Moreover, the boundaries of copy-number alterations are preferentially localized in close spatial juxtaposition within the nucleus [21-23]. Thus, the continuous DNA damage and subsequent defective repair near double-strand break (DSB) sites might generate copy-number alterations of oncogenes during tumorigenesis [24, 25]. In addition, the recent discovery that DSBs and subsequent amplification of estrogen response elements can be generated by estrogen-induced long-range chromatin interactions in breast cancer [26] further indicates the importance of higher-order genome architecture for chromosomal rearrangement. Therefore, the close spatial proximity of genomic regions

may provide an opportunity for the formation of specific, cancer-related chromosomal translocation events during tumor development [18, 27, 28].

Here, we briefly review recent works on the roles of CTCF and cohesin in genome folding, their global impact on gene expression, and their association with human disease.

Results

CTCF

Chromatin organizer roles of CTCF

CTCF is a sequence-specific DNA-binding protein that functions by utilizing an 11-zinc-finger domain [29]. Because CTCF was first identified at the 5´ and 3´ ends of the chicken β-globin locus [30] and the imprinted *Igf2/H19* locus [31, 32], it is known as an insulator protein that can block enhancer activity in eukaryotes [32]. Although insulators would be expected to be located in intergenic regions where they could act as barriers to block enhancer activity [30], genome-wide analysis indicated that CTCF-binding sites are present in genes and/or promoter regions as well as intergenic regions [10, 33]. More recent evidence revealed that almost 15% of CTCF-recognition sites are located near promoters and ~40% are within exons and introns [17], suggesting that CTCF has dynamic roles other than enhancer blocking activity.

While earlier studies implied that the distribution patterns of CTCF are similar to those of transcription activators or repressors, recently determined global distribution patterns suggested that CTCF-binding sites are not strongly correlated with general transcription factor occupancy [10]. Moreover, depletion of CTCF altered its histone acetylation and methylation profiles in the β-globin locus, but did not significantly affect β-globin expression [34, 35], suggesting
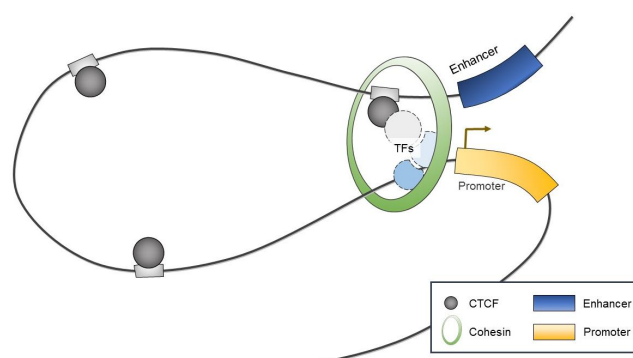
that CTCF has a role distinct from that of traditional regulatory proteins.

Interestingly, CTCF has been shown to serve as a chromatin organizer complex by linking chromosomal domains in the mouse/human β-globin cluster (Fig. 2) [36, 37]. During erythroid differentiation, CTCF is recruited and enables enhancers to physically access promoters of β-globin, which both influences transcription and contributes to cell-type-specific chromatin organization and function [36, 37]. Similarly, long-range interactions associated with CTCF have been observed within mammalian gene loci including the *Igf2/H19* imprinted control region [38, 39], the α-globin gene cluster in erythroid cells [40], and the *Igh* locus in B cells [41].

DNA methylation and CTCF binding

It has been known for many years that CTCF binding is abolished by the DNA methylation of CpG sites within the CTCF motif [32]. At the imprinted *Igf2/H19* locus, CTCF binds specifically to the unmethylated differentially methy-

lated region (DMR), which is required for the expression of *H19* on the maternal chromosome (Fig. 3A) [32, 42]. However, on the paternal allele, the methylated DMR prohibits CTCF enrichment and leads to IGF2 expression [30, 42], suggesting methylation-sensitive binding of CTCF at the target region. Interestingly, genome-wide association studies have identified that only a small subset of CTCF-binding sites are sensitive to the methylation status of DNA [8, 43].

Abnormal DNA methylation patterns of CTCF-binding sites are associated with transcriptional regulation of tumor suppressor or oncogenic genes in several human cancers [44]. CTCF plays an essential role in maintaining *INK/ARF* gene expression and disruption of its binding by DNA methylation contributes to the epigenetic silencing of *INK/ARF* genes in human breast cancer cells [45, 46]. Epigenetic inactivation of *RASSF1A* and *CDH1* also correlates with the epigenetic alteration of CTCF-recognition sites in human breast cancer [46]. Conversely, in one study, aberrant DNA methylation led to the prevention of CTCF-mediated silencing of the *BCL6* gene, thus increasing oncogenic *BCL6* expression in lymphoma [47].

The concept that the methylation-sensitive binding of CTCF controls gene expression by changing the chromatin architecture has been supported by the finding that CTCF alters the chromatin architecture [8]. For instance, in the *Igf2/H19* locus, *Igf2* imprinting on the maternal allele is performed by perturbing the proper long-range chromatin interactions between the *Igf2* gene and a distal enhancer through the formation of chromatin loops mediated by CTCF (Fig. 3A) [38, 39]. However, on the paternal chromosome, CTCF enrichment at the DMR and insulator looping are prevented by DNA methylation, thus ensuring physical interaction between the *Igf2* gene and the distal enhancer and inducing the exclusive expression of the paternal allele. Similarly, nucleotide excision repair factor-mediated DNA demethylation at the promoter region induces the enrichment of CTCF and consequently the formation of a looping structure and controls gene expression at the *RAR* β 2 locus [48]. We also found that epigenetic silencing of *PTGS2* correlates with the loss of CTCF binding by DNA methylation at the promoter region, thereby producing an inappropriate higher-order chromatin structure in human gastric cancer cells (Fig. 3B) [49].

Somatic mutations at CTCF-binding sites

In several studies, somatic mutations at the coding region of the *CTCF* gene were detected in acute leukemia and individuals with intellectual disability [50-52]. However, a high frequency of recurrent mutations in the CTCF-binding site has been more profoundly found in human cancer [53]. Unsurprisingly, single-nucleotide polymorphisms also confer disease susceptibility in humans by decreasing the methylation level at differentially methylated CTCF-binding sites such as rs2334499 in the 11p15 region [54]. Since genetic and/or epigenetic alterations frequently occur in the CTCF anchor region in various human cancers [7, 8, 55], these mutations can influence gene expression and tumor progression by abrogating the CTCF-mediated spatial folding of chromosomes [56].

Architectural role of CTCF

There is direct evidence that CTCF can physically interact with other transcriptional regulators, such as the zinc finger protein Yin Yang 1 (YY1), as an X chromosome binary switch [57]. CTCF also forms a complex with the SNF2-like chromodomain helicase protein (CHD8) [58] and the methyl-CpG–binding protein Kaiso [59] through the zinc-finger domain [10]. Thus, CHD8 enhances insulator activity, whereas Kaiso has a negative effect on the CTCF-mediated enhancer blocking activity [58, 59].
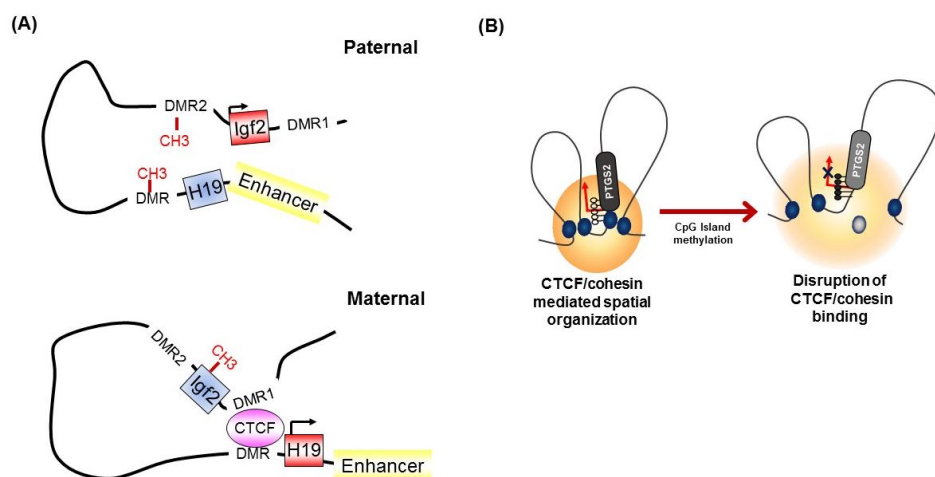
Fig. 3. (A, B) Methylation-sensitive binding of CCCTC-binding factor (CTCF).

Interestingly, the C-terminus of CTCF preferentially interacts with the STAG1 or STAG2 subunit of cohesin [60]. Furthermore, recent genome-wide studies mapping the binding sites of CTCF revealed that CTCF often colocalizes with the cohesin complex throughout the genome [61-63]. Thus, CTCF is generally thought to be required for the localization of cohesin at its binding sites [64].

The cohesin complex

Cohesin as a sister chromatid cohesion molecule

Cohesin, a large ring-shaped molecule that can bind DNA strands, is a multi-subunit protein complex composed of two structural maintenance of chromosomes (SMC) molecules, SMC1 and SMC3, either stromal antigen (STAG) STAG1 or STAG2, and the kleisin subunit RAD21 (Fig. 4) [9]. Cohesin is required to mediate sister chromatid cohesion for proper chromosome segregation in the S phase until cell division [65]. The cohesin complex was also found to be involved in efficient DNA DSB repair [64]. Since cohesin is important for holding sister chromatids together following DNA replication, mutational inactivation of the cohesin complex causes genomic instability and aneuploidy during cell cycle progression in human diseases [11]. For example, Cornelia de Lange syndrome, a rare autosomal-dominant developmental disorder, is caused by mutation of *SMC1*, *SMC3*, *Rad21*, *NIPBL*, or *HDAC8*, which encode core components of the cohesin complex or proteins that interact with this complex [66, 67]. Somatic mutations in the cohesin subunits have also been frequently found in several different human tumor types [68-71].

Cohesin in transcriptional regulation

Although its role in chromatid cohesion during mitosis is well established [65], cohesin was also found to bind thousands of sites on interphase chromosomes [4]. Indeed,

Fig. 4. A schematic representation of the cohesin complex. Smc, structural maintenance of chromosomes; Stag, stromal antigen.

cohesin interacts with the Mediator complex, a transcriptional coactivator [72], and co-occupies enhancer and promoter regions with it to regulate tissue-specific gene expression [4]. High-throughput chromatin immunoprecipitation–sequence analyses also revealed that cohesin remains bound at the transcription factor-binding sites through replication to facilitate the re-establishment of transcription factor clusters after DNA replication and cell division [12]. This suggests that the cohesin complex acts as a transcriptional regulator in cellular proliferation, differentiation, and development [8].

Colocalization of cohesin with CTCF

CTCF was originally known as a cohesin loading factor [10, 64] because genome-wide studies revealed that cohesin globally colocalizes extensively with CTCF throughout the genome [61, 62, 73]. However, CTCF depletion did not completely impair the entire association of cohesin with chromatin [61, 73-75]. Instead, depletion of CTCF was shown to reduce the enrichment of cohesin at only a small proportion of cohesin-binding sites [61, 74, 75], indicating that CTCF facilitates the distribution of cohesin to specific sites on chromosomes [8]. In contrast, the depletion of RAD21, a core subunit of the cohesin complex [9], does not disrupt the enrichment of CTCF, suggesting that CTCF binding is independent of the presence of cohesin on chromatin [8]. In this context, the following question arises: What is the essential role of cohesin at CTCF-binding sites?

Role of CTCF/cohesin in genome folding

Apart from its major function in sister chromatid cohesion, it has recently been shown that cohesin acts in concert with CTCF to affect higher-order chromosome architecture by forming long-range chromosomal interactions in many developmentally regulated gene families [76]. For example, cohesin has been shown to play a critical role in maintaining CTCF-mediated higher-order chromatin conformation at the β-globin and *Igf2/H19* loci [37, 39, 41, 74]. CTCF and cohesin also stabilize the rearrangement of *Igh* and T-cell receptor loci via long-range chromatin interactions [41, 77, 78].

Although it is not yet clear how CTCF/cohesin mediate chromatin looping, CTCF may first bind between two CTCF sites and form a complex with cohesin through its C-terminal region [60, 79, 80]. Considering the evidence that cohesin can tether DNA molecules together [81], a study appeared to show that cohesin stabilized long-range chromatin interactions by anchoring DNA strands together within a closed ring structure among CTCF/cohesin localization sites [11, 64, 76]. Interestingly, CTCF/cohesin-mediated chromatin looping preferentially occurs between CTCF sites with
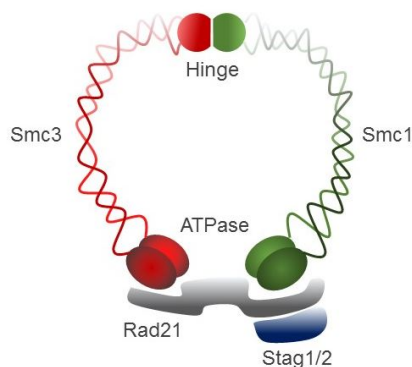
convergent CTCF DNA motifs [82]. Thus, inverting one site of convergent CTCF-binding sites changes the chromatin architecture and can alter gene expression [5, 83].

CTCF/cohesin-mediated abnormal higher-order chromosome structure during tumor development

Dysregulation of the components of the cohesin complex might promote genomic instability by perturbing proper long-range chromatin interactions, which can confer the spatial proximity required for the rejoining of DSBs during chromosomal rearrangement [18]. More recent studies support a similar role for the CTCF/cohesin-mediated chromatin loop as a regulator of genome integrity [84]. They found that chromosome loop anchors bound by CTCF and cohesin are vulnerable to continuous DNA breaks and translocation breakpoint regions in various cancers are enriched at these loop anchors [84]. Similarly, we found that cohesin-mediated chromatin organization and DNA replication are important for stabilizing gene amplification in cancer cells with chromosomal instability (Fig. 5) [85]. Although a high frequency of recurrent mutations and deletions of the components of the cohesin complex was identified in human diseases [70, 71, 86], aberrant overexpression of the cohesin complex [9] was also frequently detected in various human malignancies [85]. Interestingly, we found that overexpression of the cohesin complex in mesenchymal cancer cells induces mesenchymal to epithelial transition–specific expression patterns and dynamic cohe-

Fig. 5. Cohesin-mediated higher-order chromatin structures are important for the expression and presence of high-level gene amplification in cancer cells with chromosomal instability. HSRs, homogeneously staining regions; DMs, extrachromosomal double minutes.
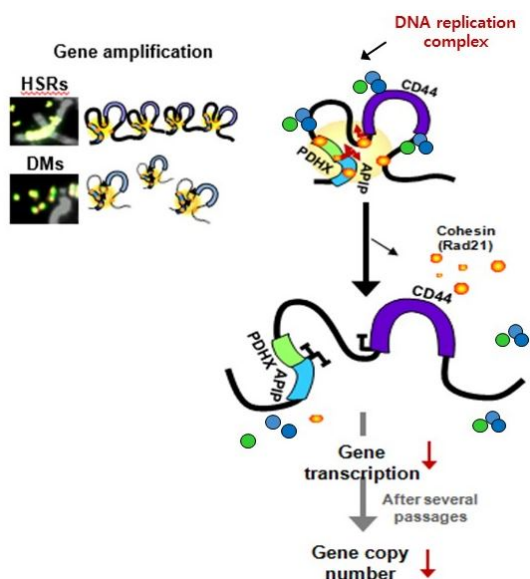
sin-mediated chromatin structures are responsible for the initiation and regulation of essential epithelial to mesenchymal transition–related cell fate changes in human cancer (Fig. 6) [87].

An increasing amount of recent evidence has indicated that CTCF and cohesin are enriched at TADs [6, 15]. Accordingly, cohesin-associated CTCF loops occur within TADs and enhancers generally interact with genes within these loops [8]. Interestingly, recurrent mutations occur frequently within CTCF anchor sites adjacent to oncogenes or cancer-associated genes [55]. Mutation in the isocitrate dehydrogenase (IDH) gene promotes susceptibility of the CTCF-binding sites to DNA methylation and the loss of CTCF binding resulting in the disruption of TAD organization in human gliomas [88]. Somatic mutations change oncogene-containing insulated neighborhoods, thereby allowing improper activation of proto-oncogenes by enhancers located within different TADs [89]. Furthermore, disruption of CTCF-mediated TAD formation by human noncoding disease variants elicits pathogenic phenotypes, providing the mechanistic linkage between spatial genomic organization and genetic alterations that influence gene expression [90].

Conclusion

Over the past few years, substantial progress has been made in understanding three-dimensional genome architecture. However, its role in cancer remains incompletely understood. An intriguing issue in this context is that, although most topological boundaries are enriched for the binding of CTCF, only 15% of CTCF-binding sites are located within TAD boundaries [15], suggesting that additional factors other than the CTCF/cohesin complex might be required to establish the topological domain structure of the genome [17]. Further work is needed to clarify the mechanisms underlying this level of chromosomal organization, and to what extent it generally contributes to the transcriptional regulation of genes during tumorigenesis.

While increasing evidence has recently indicated that enhancers are located near oncogenic genes and exhibit a large number of variants associated with diseases [91], a more complete understanding of how epigenetic alteration of enhancers directly participates in the development and onset of genome reorganization during tumor progression remains to be obtained. Based on this point, recently developed CRISPR/Cas9-based epigenome-editing technology has attracted considerable interest because this approach can acutely modify the epigenetic landscape of specific regulatory elements [92]. For example, targeted editing of the DNA methylation status of CTCF-binding sites changes CTCT recruitment, thereby altering the expression of genes
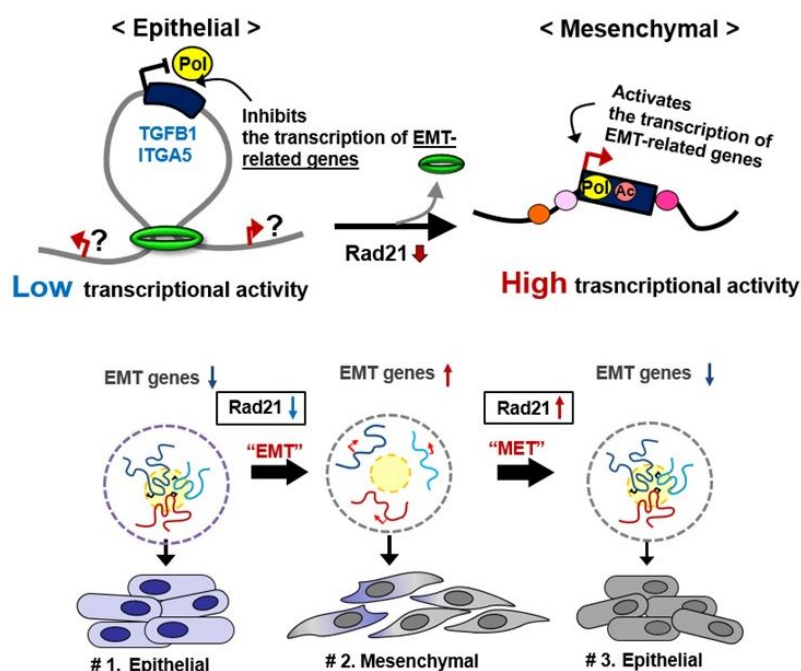
Fig. 6. Cohesin-mediated dynamic chromatin architecture of the *TGFB1* and *ITGA5* genes associated with epithelial to mesenchymal transition (EMT) plasticity. MET, mesenchymal to epithelial transition.

by influencing the organization of higher-order chromatin structures [93]. Furthermore, a high-throughput CRISPR activation system was also used to reveal how noncoding variation associated with human immune dysfunction alters stimulation-dependent enhancer function [94]. Thus, highly specific CRISPR/Cas9-based epigenome-editing technology may serve as an attractive treatment option for epigenetic-based cancer therapies in the coming years.

ORCID: Sang-Hyun Song: http://orcid.org/0000-0003-23 64-5592; Tae-You Kim: http://orcid.org/0000-0002-9986-2 684

Authors' contribution

Conceptualization: SHS, TYK
Data curation: SHS, TYK
Formal analysis: SHS, TYK
Funding acquisition: TYK
Methodology: SHS, TYK
Writing – original draft: SHS, TYK
Writing – review & editing: SHS, TYK

Acknowledgments

References

1. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, *et al*. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 2013;502:59-64.

2. de Wit E, Bouwman BA, Zhu Y, Klous P, Splinter E, Verstegen MJ, *et al*. The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* 2013;501:227-231.

3. Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, *et al*. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* 2010;42:53-61.

4. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, *et al*. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 2010;467:430-435.

5. de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, *et al*. CTCF binding polarity determines chromatin looping. *Mol Cell* 2015;60:676-684.

6. Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, *et al*. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 2013;153:1281-1295.

7. Hnisz D, Day DS, Young RA. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* 2016;167:1188-1200.

8. Merkenschlager M, Nora EP. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu Rev Genomics Hum Genet* 2016;17:17-43.

9. Dorsett D, Strom L. The ancient and evolving roles of cohesin in gene expression and DNA repair. *Curr Biol* 2012;22: R240-R250.

10. Phillips JE, Corces VG. CTCF: master weaver of the genome.

Cell 2009;137:1194-1211.

11. Mehta GD, Kumar R, Srivastava S, Ghosh SK. Cohesin: functions beyond sister chromatid cohesion. *FEBS Lett* 2013;587: 2299-2312.

12. Yan J, Enge M, Whitington T, Dave K, Liu J, Sur I, *et al*. Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* 2013;154: 801-813.

13. Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, *et al*. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res* 2013;23:2066-2077.

14. Schaaf CA, Kwak H, Koenig A, Misulovin Z, Gohara DW, Watson A, *et al*. Genome-wide control of RNA polymerase II activity by cohesin. *PLoS Genet* 2013;9:e1003382.

15. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, *et al*. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;485:376-380.

16. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, *et al*. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 2012;485:381-385.

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

18. Wijchers PJ, de Laat W. Genome organization influences partner selection for chromosomal rearrangements. *Trends Genet* 2011;27:63-71.

19. Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat Genet* 2003;34:287-291.

20. Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, *et al*. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol* 2007; 5:e192.

21. Fudenberg G, Getz G, Meyerson M, Mirny LA. High order chromatin architecture shapes the landscape of chromosomal alterations in cancer. *Nat Biotechnol* 2011;29:1109-1113.

22. De S, Michor F. DNA replication timing and long-range DNA interactions predict mutational landscapes of cancer genomes. *Nat Biotechnol* 2011;29:1103-1108.

23. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, *et al*. The landscape of somatic copy-number alteration across human cancers. *Nature* 2010;463:899-905.

24. Alabert C, Groth A. Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* 2012;13:153-167.

25. Albertson DG. Gene amplification in cancer. *Trends Genet* 2006;22:447-455.

26. Hsu PY, Hsu HK, Lan X, Juan L, Yan PS, Labanowska J, *et al*. Amplification of distant estrogen response elements deregulates target genes associated with tamoxifen resistance in breast cancer. *Cancer Cell* 2013;24:197-212.

27. Misteli T, Soutoglou E. The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nat Rev Mol Cell Biol* 2009;10:243-254.

28. Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, Simon AC, *et al*. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 2012;148: 908-921.

29. Awad TA, Bigler J, Ulmer JE, Hu YJ, Moore JM, Lutz M, *et al*. Negative transcriptional regulation mediated by thyroid hormone response element 144 requires binding of the multivalent factor CTCF to a novel target DNA sequence. *J Biol Chem* 1999;274:27092-27098.

30. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 1999;98:387-396.

31. Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, Wolffe A, *et al*. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr Biol* 2000;10:853-856.

32. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 2000;405:482-485.

33. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, *et al*. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 2008;133: 1106-1117.

34. Bulger M, Schübeler D, Bender MA, Hamilton J, Farrell CM, Hardison RC, *et al*. A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. *Mol Cell Biol* 2003;23: 5234-5244.

35. Farrell CM, West AG, Felsenfeld G. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol Cell Biol* 2002;22:3820-3831.

36. Hou C, Zhao H, Tanimoto K, Dean A. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A* 2008;105:20398-20403.

37. Hou C, Dale R, Dean A. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci U S A* 2010;107:3651-3656.

38. Murrell A, Heeson S, Reik W. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat Genet* 2004;36:889-893.

39. Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, *et al*. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A* 2006;103:10684-10689.

40. Valadez-Graham V, Razin SV, Recillas-Targa F. CTCF-dependent enhancer blockers at the upstream region of the chicken alpha-globin gene domain. *Nucleic Acids Res* 2004;32:1354-1362.

41. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, *et al*. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A* 2011;108:9566-9571.

42. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000; 405:486-489.

43. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, *et al*. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* 2012;22:1680-1688.

44. Wang XQ, Dostie J. Chromosome folding and its regulation in health and disease. *Curr Opin Genet Dev* 2017;43:23-30.

45. Rodriguez C, Borgel J, Court F, Cathala G, Forné T, Piette J. CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus. *Biochem Biophys Res Commun* 2010;392:129-134.

46. Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell* 2009;34:271-284.

47. Lai AY, Fatemi M, Dhasarathy A, Malone C, Sobol SE, Geigerman C, *et al*. DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. *J Exp Med* 2010;207:1939-1950.

48. Le May N, Fradin D, Iltis I, Bougnères P, Egly JM. XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Mol Cell* 2012;47:622-632.

49. Kang JY, Song SH, Yun J, Jeon MS, Kim HP, Han SW, *et al*. Disruption of CTCF/cohesin-mediated high-order chromatin structures by DNA methylation downregulates PTGS2 expression. *Oncogene* 2015;34:5677-5684.

50. Mullighan CG, Zhang J, Kasper LH, Lerach S, Payne-Turner D, Phillips LA, *et al*. *CREBBP* mutations in relapsed acute lymphoblastic leukaemia. *Nature* 2011;471:235-239.

51. Dolnik A, Engelmann JC, Scharfenberger-Schmeer M, Mauch J, Kelkenberg-Schade S, Haldemann B, *et al*. Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. *Blood* 2012;120:e83-e92.

52. Gregor A, Oti M, Kouwenhoven EN, Hoyer J, Sticht H, Ekici AB, *et al*. De novo mutations in the genome organizer CTCF cause intellectual disability. *Am J Hum Genet* 2013;93:124-131.

53. Katainen R, Dave K, Pitkänen E, Palin K, Kivioja T, Välimäki N, *et al*. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat Genet* 2015;47:818-821.

54. Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, *et al*. Parental origin of sequence variants associated with complex diseases. *Nature* 2009;462:868-874.

55. Ji X, Dadon DB, Powell BE, Fan ZP, Borges-Rivera D, Shachar S, *et al*. 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* 2016;18:262-275.

56. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, *et al*. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* 2015;163:1611-1627.

57. Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 2007;25:43-56.

58. Ishihara K, Oshimura M, Nakao M. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell* 2006;23:733-742.

59. Defossez PA, Kelly KF, Filion GJ, Pérez-Torrado R, Magdinier F, Menoni H, *et al*. The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *J Biol Chem* 2005;280:43017-43023.

60. Xiao T, Wallace J, Felsenfeld G. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol Cell Biol* 2011;31:2174-2183.

61. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, *et al*. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 2008;132:422-433.

62. Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, *et al*. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* 2008;105:8309-8314.

63. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J* 2008;27:654-666.

64. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. *Annu Rev Genet* 2009;43:525-558.

65. Merkenschlager M, Odom DT. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* 2013;152:1285-1297.

66. Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet* 2004;36:636-641.

67. Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, *et al*. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 2008;134:521-533.

68. Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, *et al*. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci U S A* 2008;105:3443-3448.

69. Cancer Genome Atlas Research Network, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, *et al*. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013;368:2059-2074.

70. Kon A, Shih LY, Minamino M, Sanada M, Shiraishi Y, Nagata Y, *et al*. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet* 2013;45:1232-1237.

71. Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, *et al*. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* 2011;333:1039-1043.

72. Taatjes DJ. The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends Biochem Sci* 2010;35:315-322.

73. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, *et al*. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 2008;451:796-801.

74. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, *et al*. Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet* 2009;5:e1000739.

75. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, *et al*. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 2009;460:410-413.

76. Dorsett D, Merkenschlager M. Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr Opin Cell Biol* 2013;25:327-333.

77. Ribeiro de Almeida C, Stadhouders R, de Bruijn MJ, Bergen IM, Thongjuea S, Lenhard B, *et al*. The DNA-binding protein CTCF limits proximal Vkappa recombination and restricts

kappa enhancer interactions to the immunoglobulin kappa light chain locus. *Immunity* 2011;35:501-513.

78. Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, *et al*. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* 2011;476:467-471.

79. Pant V, Kurukuti S, Pugacheva E, Shamsuddin S, Mariano P, Renkawitz R, *et al*. Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of *de novo* methylation upon maternal inheritance. *Mol Cell Biol* 2004;24:3497-3504.

80. Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, *et al*. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* 2006;312:269-272.

81. Haering CH, Farcas AM, Arumugam P, Metson J, Nasmyth K. The cohesin ring concatenates sister DNA molecules. *Nature* 2008;454:297-301.

82. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, *et al*. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;159:1665-1680.

83. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, *et al*. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 2015;162:900-910.

84. Canela A, Maman Y, Jung S, Wong N, Callen E, Day A, *et al*. Genome organization drives chromosome fragility. *Cell* 2017; 170:507-521.e518.

85. Yun J, Song SH, Kang JY, Park J, Kim HP, Han SW, *et al*. Reduced cohesin destabilizes high-level gene amplification by disrupting pre-replication complex bindings in human cancers with chromosomal instability. *Nucleic Acids Res* 2016;44:558-572.

86. Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, *et al*. RAD21 mutations cause a human cohesinopathy. *Am J Hum Genet* 2012;90:1014-1027.

87. Yun J, Song SH, Kim HP, Han SW, Yi EC, Kim TY. Dynamic cohesin-mediated chromatin architecture controls epithelial-mesenchymal plasticity in cancer. *EMBO Rep* 2016;17:1343-1359.

88. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, *et al*. Insulator dysfunction and oncogene activation in *IDH* mutant gliomas. *Nature* 2016;529: 110-114.

89. Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, Li CH, *et al*. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 2016;351:1454-1458.

90. Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, *et al*. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 2015;161:1012-1025.

91. Pott S, Lieb JD. What are super-enhancers? *Nat Genet* 2015; 47:8-12.

92. Amabile A, Migliara A, Capasso P, Biffi M, Cittaro D, Naldini L, *et al*. Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. *Cell* 2016;167:219-232.e214.

93. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, *et al*. Editing DNA methylation in the mammalian genome. *Cell* 2016;167: 233-247.e217.

94. Simeonov DR, Gowen BG, Boontanrart M, Roth TL, Gagnon JD, Mumbach MR, *et al*. Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature* 2017;549: 111-115.