ANNUAL REVIEWS

Regulation and Dysregulation of Chromosome Structure in Cancer

Denes Hnisz,[1] Jurian Schuijers,[1] Charles H. Li,[1,2] and Richard A. Young[1,2]

[1]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA; email: hnisz@wi.mit.edu, young@wi.mit.edu

[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Keywords

chromosome structure, insulated neighborhood, oncogene regulation

Abstract

Cancer arises from genetic alterations that produce dysregulated gene expression programs. Normal gene regulation occurs in the context of chromosome loop structures called insulated neighborhoods, and recent studies have shown that these structures are altered and can contribute to oncogene dysregulation in various cancer cells. We review the types of genetic and epigenetic alterations that influence neighborhood structures and contribute to gene dysregulation in cancer, present models for insulated neighborhoods associated with the most prominent human oncogenes, and discuss how such models may lead to further advances in cancer diagnosis and therapy.

INTRODUCTION

The idea that structural alterations of chromosomes may cause disease is nearly as old as the chromosome theory of inheritance (Boveri 1914). The first discovery of a chromosomal translocation, the Philadelphia chromosome, in the blood cells of a leukemia patient (Nowell & Hungerford 1960) stimulated further study of the potential roles of chromosome structural alterations in the neoplastic state of cancer cells. Such studies revealed that structural alterations of chromosomes often contribute to the dysregulation of cellular gene expression programs in cancer cells (Rabbitts 1994, Vogelstein & Kinzler 2004). More recently, chromosome conformation capture technologies, which detect DNA interactions genome-wide, have led to important new insights into the roles that chromosome structures play in normal gene control and have revealed how various alterations in chromosome structure contribute to gene dysregulation in disease (Bickmore & van Steensel 2013, Bonev & Cavalli 2016, Corces & Corces 2016, de Laat & Duboule 2013, Dixon et al. 2016, Gibcus & Dekker 2013, Gorkin et al. 2014, Groschel et al. 2014, Hnisz et al. 2016a, Krijger & de Laat 2016, Lupianez et al. 2015, Merkenschlager & Nora 2016, Ong & Corces 2014, Phillips-Cremins & Corces 2013, Valton & Dekker 2016).

Recent studies have revealed that interphase chromosomes are organized into thousands of DNA loops, which are anchored, in part, through the interactions of CTCF proteins that bind to two separate sites in DNA and to one another; these CTCF-CTCF interactions are reinforced by the cohesin complex. These loops generally contain one or more genes together with the regulatory elements that operate on them. The loop anchors constrain the regulatory elements to act predominantly on genes within the loop. In this manner, the anchors insulate genes and their regulatory elements from other regulatory elements located outside the neighborhood, and thus the CTCF-CTCF anchored loop structures have been called insulated neighborhoods.

We review recent evidence that genetic and epigenetic alterations can disrupt insulated neighborhoods in cancer cells and thereby contribute to the transformed phenotype. We present models for insulated neighborhoods associated with prominent human oncogenes and identify neighborhoods that are altered based on cancer genome sequence data. Finally, we discuss how knowledge of insulated neighborhoods may lead to further advances in cancer diagnosis and therapy.

CHROMOSOME STRUCTURES

Interphase chromosomes are organized in a hierarchy of structures, and these can play important roles in transcriptional regulation (**Figure 1**). Detailed descriptions of the various layers of genome organization and the history of this field can be found in other excellent reviews (Bickmore & van Steensel 2013, Cavalli & Misteli 2013, de Laat & Duboule 2013, Dekker & Heard 2015, Dixon et al. 2016, Gibcus & Dekker 2013, Gorkin et al. 2014, Merkenschlager & Nora 2016). We first provide a brief description of the layers of chromosome structural organization as background to the recent concept that chromosome loops are a structural and functional unit of gene control in mammalian cells.

Chromosomes in interphase nuclei tend not to intermingle but occupy distinct regions within the nuclear space (**Figure 1**). In situ hybridization and microscopy techniques revealed that these chromosome territories are a general feature in mammalian nuclei and that the territorial organization of chromosomes is maintained through cell division, although the positions of chromosome territories can be reshuffled (Cremer & Cremer 2010). At present, the mechanisms that maintain chromosome territories are unknown.

Chromosome conformation capture technologies initially revealed that interphase chromosomes are partitioned into megabase-sized folding entities termed topologically associating domains (TADs) (**Figure 1**) (Dixon et al. 2012, Nora et al. 2012). TADs are regions of DNA that
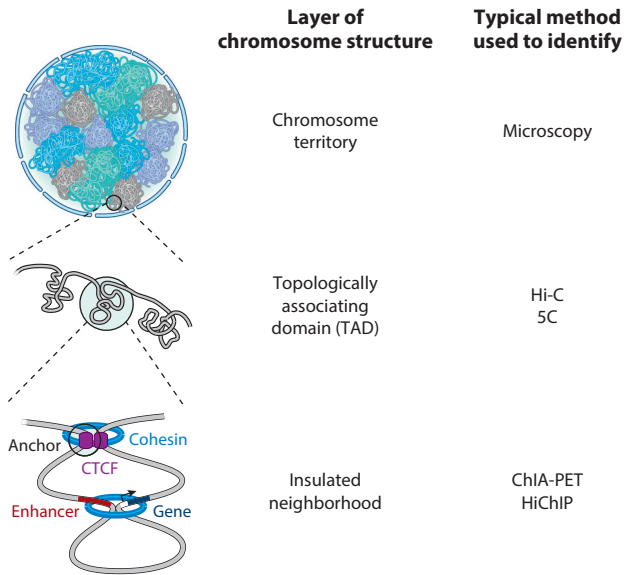
Layer of chromosome structure	Typical method used to identify
Chromosome territory	Microscopy
Topologically associating domain (TAD)	Hi-C 5C
Insulated neighborhood	ChIA-PET HiChIP

Figure 1

Chromosome structures. (*Left*) Hierarchy of chromosome structures: chromosome territories, TADs, and insulated neighborhoods. (*Right*) The experimental methods typically used to identify these structures. Abbreviations: ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; TAD, topologically associating domain.

show a high frequency of interactions relative to regions outside the TAD boundaries. Early studies reported about 2,000 TADs, which tend to have similar boundaries in all human cell types and contain on average eight genes whose expression is weakly correlated (Dixon et al. 2012, 2015). TADs were postulated to help constrain interactions between genes and their regulatory sequences (Dixon et al. 2012). The initial studies produced data at approximately 40-kb resolution, which was not sufficient to determine the mechanistic basis of TAD formation and maintenance, although an abundance of CTCF-bound sites was noted at TAD boundaries (Dixon et al. 2012).

Insights into the relationship between chromosome structure and gene regulation have emerged from studies that focused on the roles of chromosome-structuring proteins in DNA interactions and that used chromatin contact mapping technologies that provided a high-resolution view of DNA contacts associated with those proteins (**Figure 1**) (DeMare et al. 2013, Dowen et al. 2014, Handoko et al. 2011, Ji et al. 2016, Phillips-Cremins et al. 2013, Splinter et al. 2006, Tang et al. 2015, Tolhuis et al. 2002). These studies showed that chromosomes are organized into thousands of DNA loops, formed by the interaction of DNA sites bound by the CTCF protein and occupied by the cohesin complex. The anchors of these CTCF-CTCF loops function to insulate enhancers and genes within the loop from enhancers and genes outside the loop. These CTCF-CTCF loops have thus been called insulated neighborhoods, but they have also been called sub-TADs, loop domains, and CTCF-contact domains (Dowen et al. 2014, Phillips-Cremins et al. 2013, Rao et al. 2014, Tang et al. 2015). For the purposes of this review, we use the term "insulated neighborhoods" to describe CTCF-CTCF loops that encompass at least one gene.

THE INSULATED NEIGHBORHOOD MODEL

Insulated neighborhoods are formed by an interaction between two DNA sites bound by the transcription factor CTCF and the cohesin complex (**Figure 1**) (Hnisz et al. 2016a). The term

insulated neighborhood is used to describe a CTCF-CTCF loop that encompasses at least one gene, but in some cases, these loops are nested within a larger CTCF-CTCF loop that can also be considered an insulated neighborhood. Insulated neighborhoods form the mechanistic basis of TADs: TADs typically consist of a large TAD-spanning CTCF-CTCF loop that can encompass additional smaller CTCF-CTCF loops (Hnisz et al. 2016a, Ji et al. 2016, Tang et al. 2015, Zuin et al. 2014). In human embryonic stem cells, there are at least 13,000 insulated neighborhoods, which range from 25 kb to 940 kb in size and contain from one to ten genes (Ji et al. 2016). The median insulated neighborhood is ∼190 kb and contains three genes. These numbers can vary depending on assumptions made when filtering genomic data but provide an initial description of genomic loops that is useful for further analysis.

Several lines of evidence argue that the CTCF-bound anchor sites of insulated neighborhoods insulate genes and regulatory elements within a neighborhood from those outside the neighborhood. Genome-wide analysis indicates that the majority (>90%) of enhancer-gene interactions occur within insulated neighborhoods (Dowen et al. 2014, Hnisz et al. 2016b, Ji et al. 2016). Perturbation of insulated neighborhood anchor sequences leads to changes in gene expression in the vicinity of the altered neighborhood boundary (Dowen et al. 2014, Ji et al. 2016, Narendra et al. 2015, Sanborn et al. 2015). Insulated neighborhood boundary elements coincide with the end points of chromatin marks that spread over regions of transcriptional activity or repression (Dowen et al. 2014). These lines of evidence indicate that the insulating function of the neighborhood loop anchors contributes to normal gene regulation.

Insulated neighborhoods and the CTCF-CTCF loops that form them are largely maintained during development, and the subset of CTCF sites that form neighborhood loop anchors shows little genetic variation in the germ line (Hnisz et al. 2016a). However, allele-specific CTCF binding contributes to the formation of allele-specific insulated neighborhoods at imprinted genes (Hnisz et al. 2016a), and cell type–specific CTCF binding and neighborhoods appear to make some contribution to cell-specific transcriptional programs (Bunting et al. 2016, Narendra et al. 2015, Splinter et al. 2006, Tolhuis et al. 2002, Wang et al. 2012).

Although the foregoing descriptions of CTCF-CTCF loops and TADs may suggest that these are static structures, several lines of evidence suggest that they are dynamic. Both CTCF and cohesin dynamically interact with DNA (Hansen et al. 2017), and as described below, their binding is influenced by several different factors and post-translational modifications. Modeling studies suggest that chromatin contact mapping data represent an assembly of configurations that can differ between individual cells in the cell population, between time points within the same cell, and between alleles of a locus within the same cell (**Figure 2**) (Fudenberg et al. 2016, Giorgetti et al. 2014, Imakaev et al. 2012, Naumova et al. 2013). Consequently, the loop models displayed in this review and in other reports represent the predominant configurations deduced from cell population data or, in some cases, a combination of configurations that are inferred from the data.

Insulated neighborhoods cover most of the genome, and thus genes that play prominent roles in cancer biology are typically found within them. These genes include, but are not limited to, *KRAS*, *NRAS*, and *BRAF*, which are members of the RAS and RAF pathway (**Figure 3a–c**) (Bos 1989, Downward 2003); *MYC*, the most frequently overexpressed and amplified human oncogene (**Figure 3d**) (Beroukhim et al. 2010); *TP53*, which encodes the p53 protein and is the most frequently mutated gene in all cancers (**Figure 3e**) (Lawrence et al. 2014); *EGFR*, which encodes the epidermal growth factor receptor, a major drug target (**Figure 3f**) (Lynch et al. 2004); and *CD274*, or programmed death-ligand 1 (PD-L1), and the gene encoding its receptor *PDCD1*, which are immune checkpoint targets for cancer immunotherapy (**Figure 3g,h**) (Hamid et al. 2013, Pardoll 2012). More detailed information on the structures of these loci is provided in **Supplemental Figure 1**. These models rely on data from a cell line but provide the reader with
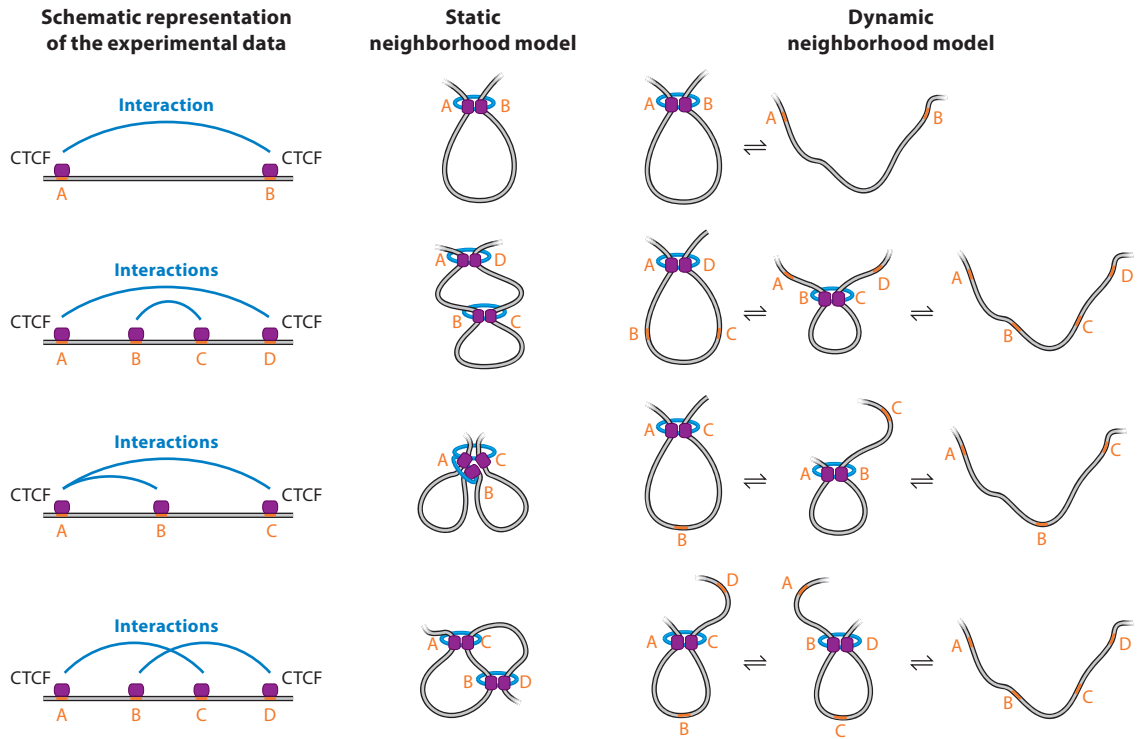
Figure 2

Insulated neighborhood models. (*Left*) Schematic representations of the experimental data and (*right*) models of their interpretation. Insulated neighborhoods are thought to be dynamic, and alternative neighborhood configurations indicated by the experimental data may occur in different cells of the population or within the same cell at different times or on different alleles within the same cell.

one view of the structural features of these loci and a potential foundation for further exploration of these structures in primary cells of various cancer types.

REGULATORS OF INSULATED NEIGHBORHOOD STRUCTURE

The proteins that are best understood to contribute to insulated neighborhood anchor structures are CTCF and cohesin, as discussed in more detail below. There are additional factors that have been implicated in establishing, maintaining, or modifying insulated neighborhood anchor structures (**Figure 4**). These include structural maintenance of chromosomes (SMC) proteins such as condensin II, the CTCF-like protein BORIS, poly (ADP-ribose) polymerase, DNA methylation, noncoding RNA species, and the process of transcription by RNA polymerase II.

CTCF

CTCF is a zinc-finger transcription factor that was originally identified as a repressor of the c-*MYC* oncogene (Baniahmad et al. 1990, Lobanenkov et al. 1990). CTCF is conserved in eukaryotes from *Drosophila* to *Homo sapiens*, is essential for embryonic development in mammals, and is ubiquitously expressed in all cells (Ghirlando & Felsenfeld 2016). CTCF has long been described as a component of insulators, which are DNA elements that can block the ability of enhancers to activate genes when placed between them (Bell et al. 1999). Several recent reviews provide
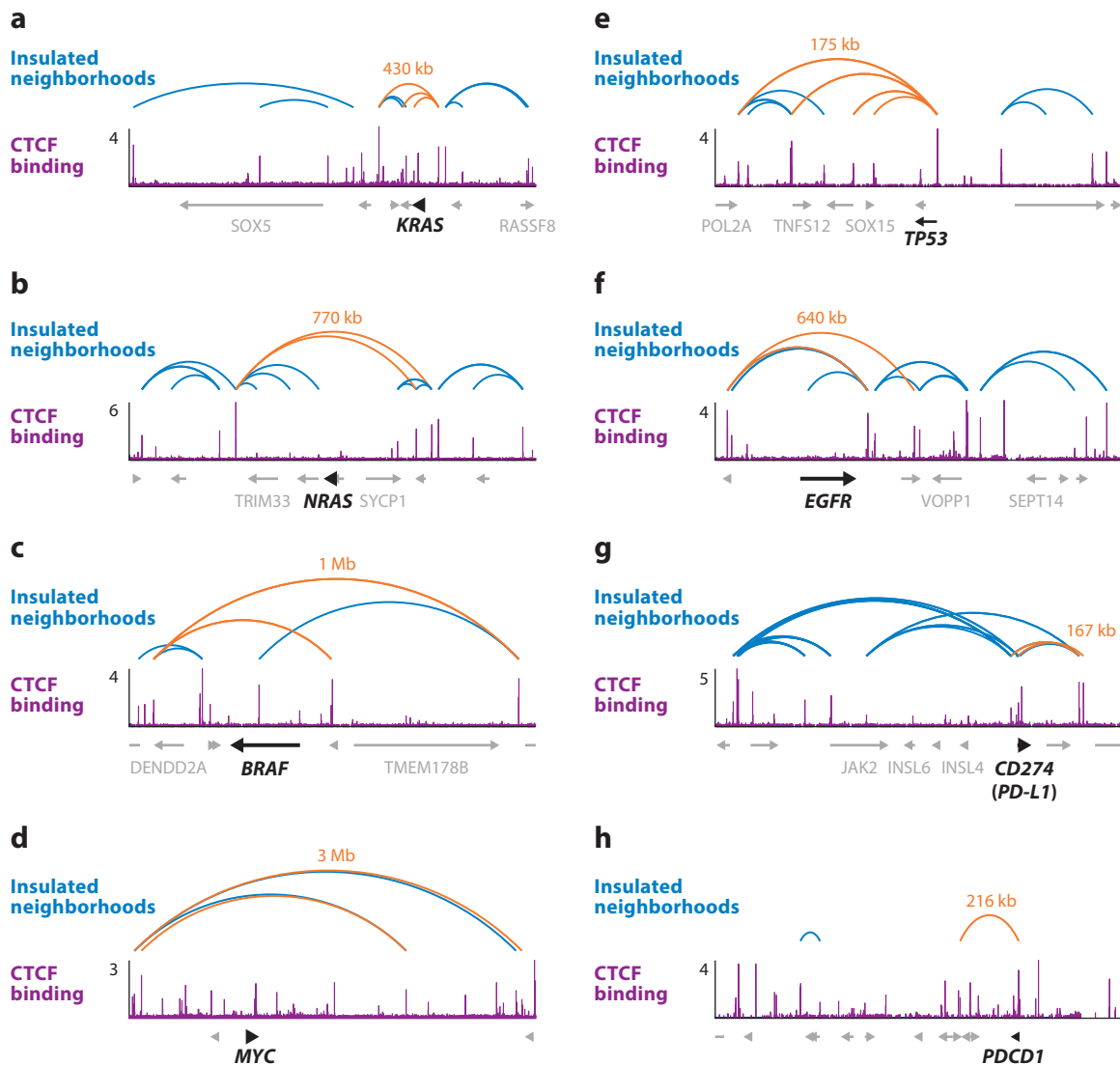
Figure 3

Insulated neighborhoods containing genes with prominent roles in cancer. Insulated neighborhoods were identified from high-confidence interactions detected in CTCF ChIA-PET data in GM12878 cells (Tang et al. 2015). Insulated neighborhoods are depicted as arcs, with those containing the gene of interest in orange. The length of the largest such neighborhood is noted. CTCF binding profiles (ChIP-seq) are displayed in gene tracks below the insulated neighborhood arcs (ENCODE Project Consortium 2012), and read density is measured as reads per million mapped reads. The genes with prominent roles in cancer are depicted with black arrows and identified in black font. Only a subset of neighborhoods at each locus is shown for simplicity; more detailed information can be found in **Supplemental Figure 1**. The genomic coordinates (hg19 genome assembly) of the displayed loci are (*a*) *KRAS*, chromosome (chr) 12:23,328,472–26,234,964; (*b*) *NRAS*, chr1:114,471,740–116,063,184; (*c*) *BRAF*, chr7:140,149,898–141,280,727; (*d*) *MYC*, chr8:127,797,231–130,842,492; (*e*) *TP53*, chr17:7,398,136–7,751,726; (*f*) *EGFR*, chr7:54,800,278–56,193,912; (*g*) *CD274*, chr9:4,706,510–5,693,885; and (*h*) *PDCD1*, chr2:241,704,545–243,199,373. Abbreviations: ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; ChIP-seq, chromatin immunoprecipitation and sequencing.

Figure 4

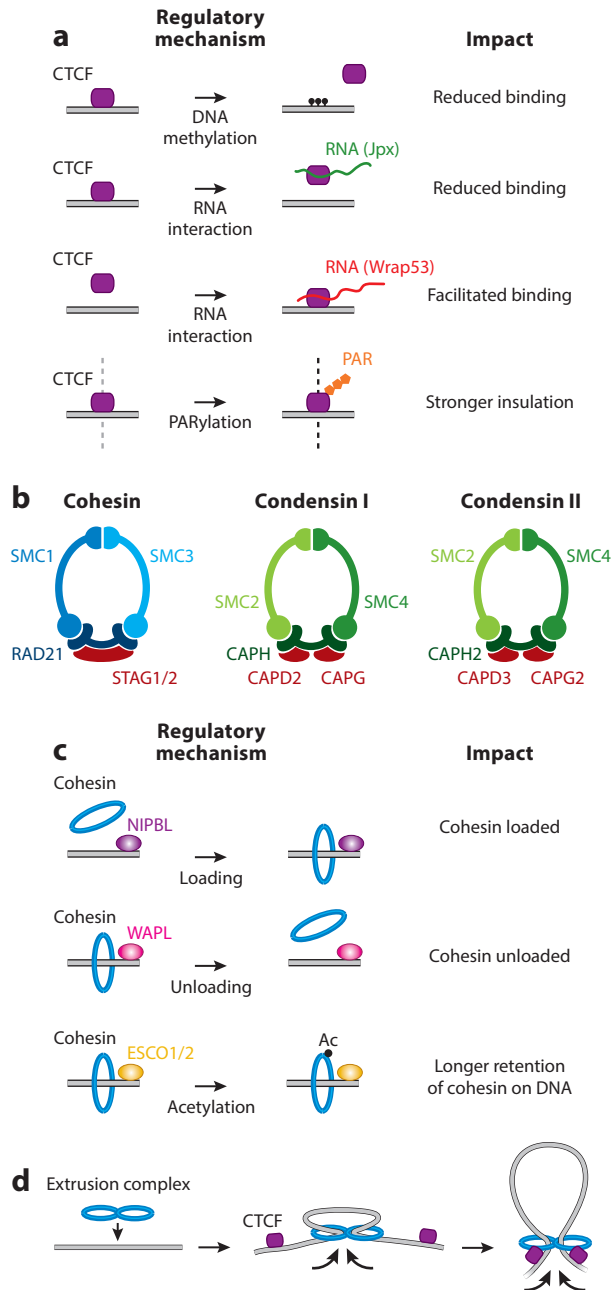Chromosome structure regulators. (*a*) Regulatory mechanisms and their impact on the chromosome structure regulator CTCF. (*b*) Schematic models of the composition of the structural maintenance of chromosomes family members cohesin, condensin I, and condensin II. (*c*) Regulatory mechanisms and their impact on the chromosome structure regulator cohesin. (*d*) Model of DNA loop formation by loop extrusion.

more detailed information and historical perspectives on CTCF (Ghirlando & Felsenfeld 2016, Merkenschlager & Odom 2013, Ong & Corces 2014, Phillips & Corces 2009).

Several lines of evidence suggest that CTCF contributes to the formation and maintenance of chromosome structures such as TADs and the insulated neighborhoods that comprise them. The majority of the boundary regions of TADs are bound by CTCF (Dixon et al. 2012, 2015; Nora et al. 2012), and global depletion of CTCF perturbs the insulating properties of TADs (Zuin et al. 2014). The CTCF protein is able to form homodimers, and thus, physical interactions between two CTCF molecules bound at two genomic locations can participate in the formation of DNA loops (Hou et al. 2008, Palstra et al. 2003, Splinter et al. 2006, Yusufzai et al. 2004).

Chromatin immunoprecipitation and sequencing experiments indicate that approximately 50,000–80,000 sites are bound by CTCF in mammalian genomes (Kim et al. 2007). However, functional assays of insulator function found that only a minority of these sites act as insulators (Liu et al. 2015) or participate in the formation of insulated neighborhood boundaries (Ji et al. 2016). It is possible that two CTCF sites need to be in a specific orientation in order for the CTCF proteins to interact and have an insulating function (Dekker & Mirny 2016, Fudenberg et al. 2016, Sanborn et al. 2015).

The ability of CTCF to bind its DNA sequence motif and participate in insulator function is influenced by DNA methylation and protein modification (**Figure 4a**). CTCF binds to hypomethylated regions of the genome (Mukhopadhyay et al. 2004), and mechanistic studies of the H19/IGF2A imprinted locus have revealed that methylation of DNA is sufficient to prevent CTCF from binding to the methylated allele (Bell & Felsenfeld 2000, Hark et al. 2000, Kanduri et al. 2000, Szabo et al. 2000). CTCF can be poly(ADP-ribosyl)ated (PARylated), and at the imprinted *H19/IGF2A* locus, PARylation of CTCF regulates its insulator function (**Figure 4a**), which is associated with its ability to form DNA loops at the locus (Yu et al. 2004). Studies in *Drosophila* have identified additional proteins that interact with CTCF, including DNA helicases, nucleophosmin, and topoisomerase (Phillips-Cremins & Corces 2013), but whether such proteins associate with CTCF in human cells and modulate its function remains to be investigated.

Transcription by RNA polymerase II has been reported to evict CTCF from specific sites (Lefevre et al. 2008), and various RNA species can enhance or reduce CTCF binding at specific loci. The Tsix, Xite, and Xist RNAs produced during X chromosome inactivation can recruit CTCF to the X-inactivation center (Kung et al. 2015), whereas the Jpx RNA evicts CTCF from the Xist promoter (Sun et al. 2013) (**Figure 4a**). An antisense transcript (*Wrap53*) produced at the *TP53* locus was found to contribute to CTCF binding (Saldaña-Meyer et al. 2014) (**Figure 4a**).

The CTCF gene has an ortholog in mammals called CTCFL or BORIS, which may also participate in DNA loops. Whereas CTCF is ubiquitously expressed in all cell types, the expression of BORIS is thought to be restricted to male germ cells (Loukinov et al. 2002). BORIS appears to bind the same DNA sequence as CTCF, and its expression is mutually exclusive with CTCF during germ-cell development (Loukinov et al. 2002).

Cohesin

Cohesin is a multiprotein complex that belongs to the SMC family of proteins (**Figure 4b**), whose members are conserved in both prokaryotes and eukaryotes (Nasmyth & Haering 2009). Cohesin consists of a tripartite ring of three subunits—SMC1, SMC3, and RAD21—which in human cells is bound by accessory factors that include STAG1 or STAG2. Cohesin was initially studied for its role in sister chromatid cohesion and was later found to play important roles in gene regulation (Dorsett & Merkenschlager 2013, Hirano 2006, Merkenschlager & Odom 2013, Nasmyth & Haering 2009, Uhlmann 2016).

Cohesin forms a ring whose internal dimensions are sufficient to entrap two DNA molecules, providing a model to explain how it contributes to DNA loops, but it is also possible that two connected cohesin rings function in DNA loop formation (Nasmyth & Haering 2009). Cohesin is loaded onto DNA by the SMC loading factor NIPBL (**Figure 4c**), which is associated with the Mediator cofactor, which mediates interactions between enhancers and promoters at active genes (Kagey et al. 2010). Disruption of cohesin perturbs enhancer-promoter interactions and gene expression (Kagey et al. 2010, Seitan et al. 2011, Zuin et al. 2014).

Cohesin is also associated with CTCF-bound sites and contributes to insulation when two CTCF-bound sites interact to form the anchors of a DNA loop (Parelho et al. 2008, Rubio et al. 2008, Wendt et al. 2008). The SMC-loading factor NIPBL is not found at CTCF sites, so it is possible that cohesin is loaded at transcriptionally active sites and then migrates to CTCF bound sites, where further movement is inhibited. The STAG1/2 subunits of cohesin can engage in direct physical interaction with CTCF (Xiao et al. 2011), which may contribute to stable CTCF-cohesin association. DNA loop extrusion models have been proposed to account for the formation of DNA loops; these models posit that where cohesin is initially loaded, extrusion of DNA through the cohesin ring (or multiple connected cohesin rings) would drive cohesin migration to two CTCF-bound sites where, if the sites were properly oriented for CTCF-CTCF interaction, the DNA loop would be anchored (Dekker & Mirny 2016, Fudenberg et al. 2016, Sanborn et al. 2015).

The regulation of cohesin has been studied primarily in the context of its role in sister chromatid cohesion, but these regulatory features may also contribute to cohesin regulation in enhancer-promoter and CTCF-CTCF interactions. For example, the SMC3 subunit of cohesin can be acetylated by the ESCO family of acetyltransferases and deacetylated by HDAC8, and the acetylation is important for the normal retention of cohesin on DNA and sister chromatid cohesion (**Figure 4c**) (Deardorff et al. 2012). Furthermore, cohesin is removed from chromatin in the mitotic prophase by the unloading factor Wapl, and depletion of Wapl leads to gross chromosome organization defects in interphase nuclei (Tedeschi et al. 2013).

Additional SMC complexes have been implicated in the control of chromosome organization. Vertebrate cells have two condensin complexes (**Figure 4b**). Although condensin I is excluded from interphase nuclei, condensin II is loaded, like cohesin, onto interphase chromatin by Nipbl at active enhancer-promoter interactions (Dowen et al. 2013). The contributions of condensin II to gene regulation, DNA looping, and larger chromosome structures are not yet understood.

MUTATIONS IN STRUCTURING COMPONENTS AND NEIGHBORHOOD BOUNDARIES IN CANCER

Translocations of large portions of chromosome arms have been described for decades in tumor cells, but only recently were mutations in chromosome structure regulators and their binding sites described and appreciated for their potential impact on specific chromosome structures. In this section, we describe the spectrum of mutations that impact neighborhood regulators and neighborhood boundary sites in tumor genomes, and we review evidence suggesting that these mutations contribute to tumor development (**Table 1**; see also **Supplemental Table 1**).

▶ Supplemental Material

CTCF Mutations

Mutations in the CTCF gene have been reported in breast cancer, endometrial cancer (Lawrence et al. 2014, Walker et al. 2015), prostate cancer (Filippova et al. 1998), Wilms' tumor (Filippova et al. 2002), and head and neck carcinomas (Lawrence et al. 2014). These mutations are predominantly missense or nonsense and thus predicted to impair CTCF function (Lawrence et al. 2014). Some tumor cell mutations occur within the zinc fingers of CTCF and may selectively perturb

Table 1   Mutations in structuring components and neighborhood boundaries in cancer

Factor	Cancer types	Types of mutation	Effects on gene	Hits	Sample size	Frequency (%)	Reference(s)
CTCF	Endometrial cancer	Unreported	Unreported	46	248	18.5	Lawrence et al. 2014
	Endometrial cancer, endometrioid	Unreported	Frameshift, missense, splice, read through	136	538	25.3	Walker et al. 2015
	Head and neck cancer	Unreported	Unreported	12	384	3.1	Lawrence et al. 2014
	Uterine corpus endometrial carcinoma	Substitution, indel	Unreported	43	230	18.7	Kandoth et al. 2013
ESCO1	Endometrial cancer	Substitution	Nonsense, missense	4	107	3.7	Price et al. 2014
NIPBL	CRC	Substitution, indel	Missense, frameshift	4	132	3	Barber et al. 2008
UCC		Substitution	Nonsense, missense, exon junction	5	131	3.8	Cancer Genome Atlas Res. Netw. 2014
RAD21	AML	Substitution, indel	Frameshift, nonsense, splice	7	157	4.5	Kon et al. 2013
		Unreported	Unreported	6	196	3.1	Lawrence et al. 2014
	UCC	Substitution	Missense, nonsense	5	131	3.8	Cancer Genome Atlas Res. Netw. 2014
SMC1A	AML	Substitution	Missense, nonsense	7	200	3.5	Cancer Genome Atlas Res. Netw. 2013
		Substitution	Unreported	7	200	3.5	Kandoth et al. 2013
		Unreported	Unreported	7	196	3.6	Lawrence et al. 2014
	CRC	Substitution	Missense	4	132	3	Barber et al. 2008
SMC3	AML	Substitution	Missense, nonsense	7	200	3.5	Cancer Genome Atlas Res. Netw. 2013
		Substitution, indel	Missense, nonsense, splice, frameshift	14	450	3.1	Thota et al. 2014
		Substitution	Unreported	7	200	3.5	Kandoth et al. 2013
		Unreported	Unreported	7	196	3.6	Lawrence et al. 2014
STAG1	UCC	Substitution	Missense, exon junction	4	131	3.1	Cancer Genome Atlas Res. Netw. 2014
STAG2	AML	Substitution, indel	Frameshift, nonsense, splice	10	157	6.4	Kon et al. 2013
		Substitution	Nonsense	6	200	3	Cancer Genome Atlas Res. Netw. 2013
		Substitution, indel	Nonsense, frameshift, splice, deletion	23	450	5.1	Thota et al. 2014

(Continued)

Table 1 (*Continued*)

Factor	Cancer types	Types of mutation	Effects on gene	Hits	Sample size	Frequency (%)	Reference(s)
		Substitution, indel	Nonsense, frameshift, splice, missense, deletion	24	299	8	Lindsley et al. 2015
		Substitution	Unreported	6	200	3	Kandoth et al. 2013
		Unreported	Unreported	6	196	3.1	Lawrence et al. 2014
	AML, de novo	Unreported	Frameshift, nonsense, missense	10	197	5.1	Kihara et al. 2014
	Bladder cancer	Substitution, indel	Missense, nonsense, exon junction, frameshift	25	111	22.5	Solomon et al. 2013
		Substitution, indel	Missense, nonsense, frameshift, deletion, splice	67	307	21.8	Taylor et al. 2014
	Ewing sarcoma	Substitution, indel, duplication	Nonsense, missense, exon junction, frameshift, exon duplication	19	112	17	Tirode et al. 2014
		Substitution, indel, duplication	Nonsense, missense, exon junction, frameshift, exon duplication, in-frame deletion	41	199	20.6	Tirode et al. 2014
	GBM	Substitution, indel	Unreported	12	290	4.1	Kandoth et al. 2013
		Unreported	Unreported	12	291	4.1	Lawrence et al. 2014
		Unreported	Unreported	12	291	4.1	Brennan et al. 2013
	MDS	Substitution, indel	Frameshift, nonsense, splice	13	224	5.8	Kon et al. 2013
		Substitution, indel	Missense, nonsense, frameshift, splice	30	386	7.8	Thota et al. 2014
		Substitution, indel	Splice, nonsense, frameshift	9	150	6	Walter et al. 2013
		Unreported	Unreported	71	944	7.5	Haferlach et al. 2014
	MDS/MPNs	Substitution	Missense, nonsense, splice	6	169	3.6	Thota et al. 2014
	Papillary RCC	Unreported	Unreported	8	157	5.1	Cancer Genome Atlas Res. Net. 2016
	UCC	Substitution, indel	Nonsense, missense, exon junction, frameshift	14	131	10.7	Cancer Genome Atlas Res. Netw. 2014

Listed are cancer types in which mutations in CTCF, subunits of cohesin, and cohesin regulators have been reported. Only studies that included at least 100 samples and mutations that reached at least 3% frequency are included. The complete list of mutations is provided in **Supplemental Table 1**. Abbreviations: AML, acute myeloid leukemia; CRC, colorectal cancer; GBM, glioblastoma multiforme; MDS, myelodysplastic syndrome; MPNs, myeloproliferative neoplasms; RCC, renal cell carcinoma; UCC, urothelial carcinoma.

certain neighborhoods because they affect CTCF binding at only a subset of sites (Filippova et al. 2002). The loss of a CTCF allele can occur in some tumor types, suggesting that CTCF may act as a haploinsufficient tumor suppressor (Filippova et al. 1998). Consistent with this notion, mice heterozygous for the CTCF gene display an increased susceptibility to develop tumors in various radiation and chemical-based cancer induction models (Kemp et al. 2014). Dysregulated expression of the germ line–specific CTCF ortholog BORIS has been reported in several cancer types (Simpson et al. 2005), but it is not yet clear that this contributes to tumorigenesis.

Cohesin Mutations

Mutations in the cohesin complex occur in acute myeloid leukemia (AML) (Cancer Genome Atlas Res. Netw. 2013), myeloid dysplastic syndrome (Kon et al. 2013), bladder cancer (Guo et al. 2013), breast cancer (Stephens et al. 2012), colorectal cancer (Barber et al. 2008), and Ewing sarcoma (Crompton et al. 2014), among others (**Table 1**; see also **Supplemental Table 1**). In AML, mutations in all four cohesin subunits (SMC1A, SMC3, RAD21, and STAG2) have been reported, whereas in solid tumors mutations of the STAG2 subunit occur most frequently. Most mutations in the cohesin subunits are missense, nonsense, or truncating (Lawrence et al. 2014), suggesting a loss-of-function effect, which is consistent with the reduced level of DNA-bound cohesin reported in cohesin-mutant AML cells (Kon et al. 2013).

Recent studies indicate that the tumor-promoting effect of at least a subset of cohesin alterations is linked to its roles in gene regulation and chromosome structure rather than its roles in proper chromosome segregation. The classic role of cohesin in sister chromatid cohesion would predict that cohesin mutations in cancer contribute to the neoplastic state through defects in chromosome segregation and consequent aneuploidy. However, modeling *SMC3* mutations that occur in AML has revealed no association with chromosome segregation defects and aneuploidy but has revealed an association with the alteration of the gene expression program of the leukemia cells (Mazumdar et al. 2015, Mullenders et al. 2015, Viny et al. 2015). Furthermore, analysis of *STAG2*-mutant bladder cancer did not reveal any association of the *STAG2* mutation with chromosome segregation defects and aneuploidy (Balbás-Martínez et al. 2013).

CTCF Binding Site Mutations

Nucleotide substitutions in the DNA binding site of CTCF occur in the genomes of several cancer types (**Supplemental Table 1**), and such substitutions appear to be especially enriched in CTCF binding sites that form insulated neighborhood boundaries. Nucleotide substitutions in CTCF binding sites have been reported in colorectal cancer (Katainen et al. 2015), gastrointestinal cancer (Umer et al. 2016), esophageal cancer (Hnisz et al. 2016b), liver cancer (Hnisz et al. 2016b, Katainen et al. 2015), and melanoma (Poulos et al. 2016). Although the functional impact of these mutations has not been investigated in depth, the observation that insulated neighborhood boundary CTCF sites are conserved and show limited germ-line variation (Ji et al. 2016), together with evidence that some of these mutations are recurrent, suggests that many of the somatic CTCF site mutations in tumor cells contribute to the neoplastic state by perturbing insulated neighborhoods.

Epigenetic Alteration of CTCF Binding

Because DNA hypermethylation is a feature of many cancer types and DNA methylation reduces CTCF binding, insulated neighborhood structures may be compromised in cells with

hypermethylated DNA (**Supplemental Table 1**). Indeed, in a subset of gliomas that harbor mutations in the *IDH1* gene, tumor-specific hypermethylation is associated with the disruption of CTCF binding, alteration of chromosome structure, and dysregulation of oncogene expression (Flavahan et al. 2016).

Mutations in Regulators of CTCF and Cohesin

Several regulators of CTCF and cohesin have been recently implicated in cancer, and it is possible that mutations of some of these regulators contribute to the neoplastic state through the alteration of insulated neighborhoods. For example, nucleophosmin, a direct physical interaction partner of CTCF, is frequently mutated in AML (Cancer Genome Atlas Res. Netw. 2013). Many noncoding RNA species are implicated in cancer development (Lin & He 2017), and CTCF interacts with thousands of RNAs, some of which impact its binding to DNA (Kung et al. 2015), so it is plausible that the dysregulation of noncoding RNAs in tumor cells contributes to alterations of chromosome structures. ESCO1, the enzyme that acetylates cohesin, is mutated in a subset of endometrial cancers (Price et al. 2014). In summary, defects in a diverse set of mechanisms may contribute to alterations of chromosome structure in cancer cells.

IMPACTS OF NEIGHBORHOOD ALTERATIONS IN CANCER

The structural and functional impact of mutations in chromosome structuring components and in neighborhood boundaries is only beginning to be studied in cancer, but it is useful nonetheless to consider models that explain how these mutations can contribute to gene dysregulation in tumor cells (**Figure 5**). In some instances, these models are supported by experimental data; in others, they are predictive and await further study.

About half of T cell acute lymphoblastic leukemias contain mutations that activate the *TAL1* oncogene (Armstrong & Look 2005, Van Vlierberghe & Ferrando 2012). In a subset of these leukemias, the *TAL1* oncogene is activated by microdeletions that remove the boundary of an insulated neighborhood containing the *TAL1* gene. The disruption of the boundary leads to inappropriate contacts between the *TAL1* gene and regulatory elements normally located outside the *TAL1* neighborhood (**Figure 5a**) (Hnisz et al. 2016b). Similar microdeletions that disrupt an insulated neighborhood boundary also occur encompassing the *LMO2* oncogene in these T-cell leukemias (Hnisz et al. 2016b).

Epigenetic alteration of CTCF binding sites at insulated neighborhood anchors can also lead to oncogene activation. A subset of gliomas harbor a mutation in the *IDH1* gene, and this mutation is associated with DNA hypermethylation (Cancer Genome Atlas Res. Netw. 2015, Dang et al. 2009). A recent study found that in *IDH1*-mutant gliomas, the methylation of an insulated neighborhood boundary encompassing the *PDGFRA* oncogene leads to a loss of the insulating property of the neighborhood, inappropriate contacts between *PDGFRA* and an upstream enhancer normally located outside the *PDGFRA* neighborhood, and elevated expression of *PDGFRA* (**Figure 5a**) (Flavahan et al. 2016).

Our understanding of insulated neighborhoods in normal gene control suggest additional models for the impact of genetic or epigenetic alterations of neighborhood structures in gene dysregulation in neoplastic cells; these have yet to be reported in cancer cells and thus serve merely as predictions. For example, alterations of neighborhood boundaries or their components may lead to the activation of silent oncogenes by enabling enhancers within the neighborhood to activate genes that are normally located outside the neighborhood (**Figure 5b**). Genetic or epigenetic perturbation of insulated neighborhood boundaries can lead to a downregulation of genes found
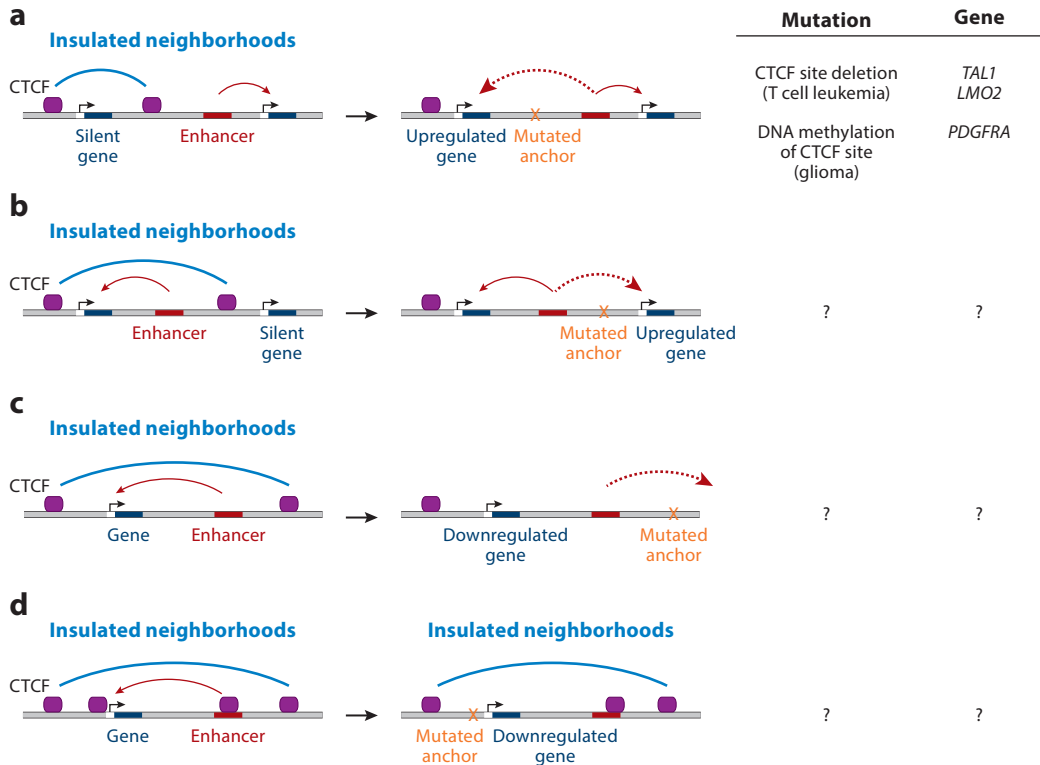
Figure 5

Insulated neighborhood models for gene regulation and dysregulation in cancer. (*a*) Disruption of an insulated neighborhood boundary leads to upregulation of a gene that was in the neighborhood due to inappropriate contact with an enhancer that was outside the neighborhood. (*b*) Disruption of an insulated neighborhood boundary leads to upregulation of a gene that was outside the neighborhood due to inappropriate contact with an enhancer that was inside the neighborhood. (*c*) Disruption of an insulated neighborhood boundary leads to downregulation of a gene that used to be inside the neighborhood. (*d*) Disruption of a CTCF anchor that mediates enhancer-promoter interactions within a neighborhood leads to downregulation of a gene within the neighborhood.

in neighborhoods (**Figure 5***c*) (Dowen et al. 2014) and, thus, potentially the loss of expression of a tumor suppressor. A minority of enhancers and promoters are bound by CTCF, and contacts between such enhancer-promoter pairs can be facilitated by CTCF-CTCF interactions (Guo et al. 2015), so disruption of these interactions can contribute to gene dysregulation (**Figure 5***d*). Because the process of transcription and the presence of RNA species can affect CTCF binding, altered transcription in the cancer state may also be responsible for changes in neighborhood structure and function.

FUTURE CHALLENGES: CHROMOSOME STRUCTURES IN CANCER DIAGNOSTICS AND THERAPY

Genetic and epigenetic alterations of insulated neighborhoods can lead to dysregulation of prominent oncogenes that drive tumorigenesis. These new insights suggest new approaches to identify mechanisms associated with gene dysregulation in cancer and new approaches to target dysregulated expression of oncogenes and tumor suppressors.

Neighborhood Alteration in Cancer: Identification of Oncogenes and Dependencies

Somatic mutations and epigenetic alterations that perturb insulated neighborhood boundaries may be useful to identify oncogenes and dependencies in cancers whose development and progression are not well understood. The finding that insulated neighborhood boundary alterations occur at oncogene loci in leukemia and gliomas (Flavahan et al. 2016, Hnisz et al. 2013) suggests that neighborhoods with recurrently altered boundaries can identify new oncogenic drivers.

Cancer cells can become highly dependent on certain gene products during cancer progression. The disrupted neighborhood around *PDGFRA* in *IDH1*-mutant glioma cells is associated with the sensitivity of these cells to PDGFRA inhibitors (Flavahan et al. 2016). This suggests that cancer dependencies engendered by similar mechanisms will be revealed through investigation of neighborhood alterations in cancer cells. Progress here will require improved understanding of the mutational landscape of noncoding regions of the genome where most neighborhood boundaries are located and of the epigenetic mechanisms that impact neighborhood boundary function.

Cancer Susceptibility

Although cancer development occurs as a consequence of somatic alterations of the genome, DNA variants in the germ line contribute to the susceptibility of cancer development. Recent evidence that DNA polymorphisms in noncoding DNA are linked to cancer predisposition (Oldridge et al. 2015), that rare germ line variants occur in insulated neighborhood CTCF sites (Ji et al. 2016), and that such variants can impact neighborhoods (Tang et al. 2015) indicates that some of the genetic variation that contributes to cancer susceptibility may occur in insulated neighborhoods.

Epigenetic Editing of CTCF Anchors

Targeted disruption of CTCF binding and neighborhood integrity, with predictable effects on gene dysregulation, has been demonstrated through targeted methylation with a catalytically inactive Cas9 (dCas9)-DNA-metyltransferase-3 fusion protein (Liu et al. 2016). Targeted demethylation with a dCas9-TET fusion protein reversed this effect, allowing CTCF binding and insulated neighborhood formation (Liu et al. 2016). This suggests that targeted methylation and demethylation of CTCF binding sites could be used to alter CTCF-CTCF loops that form either insulated neighborhoods or enhancer-promoter interactions. Such epigenetic editing tools might evolve to be useful for therapeutic purposes in cancer and in other diseases where gene dysregulation is key to the disease state.

LITERATURE CITED

Armstrong SA, Look AT. 2005. Molecular genetics of acute lymphoblastic leukemia. *J. Clin. Oncol.* 23:6306–15

Balbás-Martínez C, Sagrera A, Carrillo-de-Santa-Pau E, Earl J, Márquez M, et al. 2013. Recurrent inactivation of *STAG2* in bladder cancer is not associated with aneuploidy. *Nat. Genet.* 45:1464–69

Baniahmad A, Steiner C, Kohne AC, Renkawitz R. 1990. Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. *Cell* 61:505–14

Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, et al. 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *PNAS* 105:3443–48

Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405:482–85

Bell AC, West AG, Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98:387–96

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

Bickmore WA, van Steensel B. 2013. Genome architecture: domain organization of interphase chromosomes. *Cell* 152:1270–84

Bonev B, Cavalli G. 2016. Organization and function of the 3D genome. *Nat. Rev. Genet.* 17:661–78

Bos JL. 1989. *ras* oncogenes in human cancer: a review. *Cancer Res.* 49:4682–89

Boveri T. 1914. *Zur Frage der Entstehung maligner Tumoren*. Jena, Ger.: Gustav Fischer

Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, et al. 2013. The somatic genomic landscape of glioblastoma. *Cell* 155:462–77

Bunting KL, Soong TD, Singh R, Jiang Y, Beguelin W, et al. 2016. Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. *Immunity* 45:497–512

Cancer Genome Atlas Res. Netw. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* 368:2059–74

Cancer Genome Atlas Res. Netw. 2014. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* 507:315–22

Cancer Genome Atlas Res. Netw. 2015. Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* 372:2481–98

Cancer Genome Atlas Res. Netw. 2016. Comprehensive molecular characterization of papillary renal-cell carcinoma. *N. Engl. J. Med.* 374:135–45

Cavalli G, Misteli T. 2013. Functional implications of genome topology. *Nat. Struct. Mol. Biol.* 20:290–99

Corces MR, Corces VG. 2016. The three-dimensional cancer genome. *Curr. Opin. Genet. Dev.* 36:1–7

Cremer T, Cremer M. 2010. Chromosome territories. *Cold Spring Harb. Perspect. Biol.* 2:a003889

Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, et al. 2014. The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov*. 4:1326–41

Dang L, White DW, Gross S, Bennett BD, Bittinger MA, et al. 2009. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* 462:739–44

de Laat W, Duboule D. 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502:499–506

Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, et al. 2012. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* 489:313–17

Dekker J, Heard E. 2015. Structural and functional diversity of Topologically Associating Domains. *FEBS Lett*. 589:2877–84

Dekker J, Mirny L. 2016. The 3D genome as moderator of chromosomal communication. *Cell* 164:1110–21

DeMare LE, Leng J, Cotney J, Reilly SK, Yin J, et al. 2013. The genomic landscape of cohesin-associated chromatin interactions. *Genome Res*. 23:1224–34

Dixon JR, Gorkin DU, Ren B. 2016. Chromatin domains: the unit of chromosome organization. *Mol. Cell* 62:668–80

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* 518:331–36

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, et al. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376–80

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

Dowen JM, Bilodeau S, Orlando DA, Hubner MR, Abraham BJ, et al. 2013. Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. *Stem Cell Rep.* 1:371–78

Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, et al. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159:374–87

Downward J. 2003. Targeting RAS signalling pathways in cancer therapy. *Nat. Rev. Cancer* 3:11–22

ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74

Filippova GN, Lindblom A, Meincke LJ, Klenova EM, Neiman PE, et al. 1998. A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. *Genes Chromosom. Cancer* 22:26–36

Filippova GN, Qi CF, Ulmer JE, Moore JM, Ward MD, et al. 2002. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer Res.* 62:48–52

Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, et al. 2016. Insulator dysfunction and oncogene activation in *IDH* mutant gliomas. *Nature* 529:110–14

Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2016. Formation of chromosomal domains by loop extrusion. *Cell Rep.* 15:2038–49

Ghirlando R, Felsenfeld G. 2016. CTCF: making the right connections. *Genes Dev.* 30:881–91

Gibcus JH, Dekker J. 2013. The hierarchy of the 3D genome. *Mol. Cell* 49:773–82

Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, et al. 2014. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* 157:950–63

Gorkin DU, Leung D, Ren B. 2014. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* 14:762–75

Groschel S, Sanders MA, Hoogenboezem R, de Wit E, Bouwman BA, et al. 2014. A single oncogenic enhancer rearrangement causes concomitant *EVI1* and *GATA2* deregulation in leukemia. *Cell* 157:369–81

Guo G, Sun X, Chen C, Wu S, Huang P, et al. 2013. Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. *Nat. Genet.* 45:1459–63

Guo Y, Xu Q, Canzio D, Shou J, Li J, et al. 2015. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162:900–10

Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, et al. 2014. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* 28:241–47

Hamid O, Robert C, Daud A, Hodi FS, Hwu WJ, et al. 2013. Safety and tumor responses with lambrolizumab (anti–PD-1) in melanoma. *N. Engl. J. Med.* 369:134–44

Handoko L, Xu H, Li G, Ngan CY, Chew E, et al. 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* 43:630–38

Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* 6:e25776

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* 405:486–89

Hirano T. 2006. At the heart of the chromosome: SMC proteins in action. *Nat. Rev. Mol. Cell Biol.* 7:311–22

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, et al. 2013. Super-enhancers in the control of cell identity and disease. *Cell* 155:934–47

Hnisz D, Day DS, Young RA. 2016a. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* 167:1188–200

Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, et al. 2016b. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351:1454–58

Hou C, Zhao H, Tanimoto K, Dean A. 2008. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *PNAS* 105:20398–403

Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, et al. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* 9:999–1003

Ji X, Dadon DB, Powell BE, Fan ZP, Borges-Rivera D, et al. 2016. 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* 18:262–75

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467:430–35

Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, et al. 2013. Mutational landscape and significance across 12 major cancer types. *Nature* 502:333–39

Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, et al. 2000. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.* 10:853–56

Katainen R, Dave K, Pitkanen E, Palin K, Kivioja T, et al. 2015. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat. Genet.* 47:818–21

Kemp CJ, Moore JM, Moser R, Bernard B, Teater M, et al. 2014. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.* 7:1020–29

Kihara R, Nagata Y, Kiyoi H, Kato T, Yamamoto E, et al. 2014. Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients. *Leukemia* 28:1586–95

Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, et al. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128:1231–45

Kon A, Shih LY, Minamino M, Sanada M, Shiraishi Y, et al. 2013. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat. Genet.* 45:1232–37

Krijger PH, de Laat W. 2016. Regulation of disease-associated gene expression in the 3D genome. *Nat. Rev. Mol. Cell Biol.* 17:771–82

Kung JT, Kesner B, An JY, Ahn JY, Cifuentes-Rojas C, et al. 2015. Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* 57:361–75

Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, et al. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505:495–501

Lefevre P, Witham J, Lacroix CE, Cockerill PN, Bonifer C. 2008. The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Mol. Cell* 32:129–39

Lin CP, He L. 2017. Noncoding RNAs in cancer development. *Annu. Rev. Cancer Biol.* 1:163–84

Lindsley RC, Mar BG, Mazzola E, Grauman PV, Shareef S, et al. 2015. Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. *Blood* 125:1267–76

Liu M, Maurano MT, Wang H, Qi H, Song CZ, et al. 2015. Genomic discovery of potent chromatin insulators for human gene therapy. *Nat. Biotechnol.* 33:198–203

Liu XS, Wu H, Ji X, Stelzer Y, Wu X, et al. 2016. Editing DNA methylation in the mammalian genome. *Cell* 167:233–47.e217

Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, et al. 1990. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5′-flanking sequence of the chicken c-*myc* gene. *Oncogene* 5:1743–53

Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, et al. 2002. BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *PNAS* 99:6806–11

Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161:1012–15

Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, et al. 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 350:2129–39

Mazumdar C, Shen Y, Xavy S, Zhao F, Reinisch A, et al. 2015. Leukemia-associated cohesin mutants dominantly enforce stem cell programs and impair human hematopoietic progenitor differentiation. *Cell Stem Cell* 17:675–88

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

Merkenschlager M, Odom DT. 2013. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* 152:1285–97

Mukhopadhyay R, Yu W, Whitehead J, Xu J, Lezcano M, et al. 2004. The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res.* 14:1594–602

Mullenders J, Aranda-Orgilles B, Lhoumaud P, Keller M, Pae J, et al. 2015. Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J. Exp. Med.* 212:1833–50

Narendra V, Rocha PP, An D, Raviram R, Skok JA, et al. 2015. CTCF establishes discrete functional chromatin domains at the *Hox* clusters during differentiation. *Science* 347:1017–21

Nasmyth K, Haering CH. 2009. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* 43:525–58

Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, et al. 2013. Organization of the mitotic chromosome. *Science* 342:948–53

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485:381–85

Nowell PC, Hungerford DA. 1960. Chromosome studies on normal and leukemic human leukocytes. *J. Natl. Cancer Inst.* 25:85–109

Oldridge DA, Wood AC, Weichert-Leahey N, Crimmins I, Sussman R, et al. 2015. Genetic predisposition to neuroblastoma mediated by a *LMO1* super-enhancer polymorphism. *Nature* 528:418–21

Ong CT, Corces VG. 2014. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* 15:234–46

Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. 2003. The β-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* 35:190–94

Pardoll DM. 2012. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* 12:252–64

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132:422–33

Phillips JE, Corces VG. 2009. CTCF: master weaver of the genome. *Cell* 137:1194–211

Phillips-Cremins JE, Corces VG. 2013. Chromatin insulators: linking genome organization to cellular function. *Mol. Cell* 50:461–74

Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, et al. 2013. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153:1281–95

Poulos RC, Thoms JA, Guan YF, Unnikrishnan A, Pimanda JE, Wong JW. 2016. Functional mutations form at CTCF-cohesin binding sites in melanoma due to uneven nucleotide excision repair across the motif. *Cell Rep.* 17:2865–72

Price JC, Pollock LM, Rudd ML, Fogoros SK, Mohamed H, et al. 2014. Sequencing of candidate chromosome instability genes in endometrial cancers reveals somatic mutations in *ESCO1*, *CHTF18*, and *MRE11A*. *PLOS ONE* 8:e63313

Rabbitts TH. 1994. Chromosomal translocations in human cancer. *Nature* 372:143–49

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665–80

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, et al. 2008. CTCF physically links cohesin to chromatin. *PNAS* 105:8309–14

Saldaña-Meyer R, González-Buendía E, Guerrero G, Narendra V, Bonasio R, et al. 2014. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* 28:723–34

Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *PNAS* 112:E6456–65

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, et al. 2011. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* 476:467–71

Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ. 2005. Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* 5:615–25

Solomon DA, Kim JS, Bondaruk J, Shariat SF, Wang ZF, et al. 2013. Frequent truncating mutations of *STAG2* in bladder cancer. *Nat. Genet.* 45:1428–30

Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, et al. 2006. CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. *Genes Dev.* 20:2349–54

Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, et al. 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486:400–4

Sun S, Del Rosario BC, Szanto A, Ogawa Y, Jeon Y, Lee JT. 2013. Jpx RNA activates *Xist* by evicting CTCF. *Cell* 153:1537–51

Szabo P, Tang SH, Rentsendorj A, Pfeifer GP, Mann JR. 2000. Maternal-specific footprints at putative CTCF sites in the *H19* imprinting control region give evidence for insulator function. *Curr. Biol.* 10:607–10

Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, et al. 2015. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* 163:1611–27

Taylor CF, Platt FM, Hurst CD, Thygesen HH, Knowles MA. 2014. Frequent inactivating mutations of *STAG2* in bladder cancer are associated with low tumour grade and stage and inversely related to chromosomal copy number changes. *Hum. Mol. Genet.* 15:1964–74

Tedeschi A, Wutz G, Huet S, Jaritz M, Wuensche A, et al. 2013. Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature* 501:564–68

Thota S, Viny AD, Makishima H, Spitzer B, Radivoyevitch T, et al. 2014. Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood* 124:1790–98

Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, et al. 2014. Genomic landscapes of Ewing sarcoma defines an aggressive subtype with co-association of *STAG2* and *TP53* mutations. *Cancer Discov.* 4:1342–53

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. 2002. Looping and interaction between hypersensitive sites in the active β-globin locus. *Mol. Cell* 10:1453–65

Uhlmann F. 2016. SMC complexes: from DNA to chromosomes. *Nat. Rev. Mol. Cell Biol.* 17:399–412

Umer HM, Cavalli M, Dabrowski MJ, Diamanti K, Kruczyk M, et al. 2016. A significant regulatory mutation burden at a high-affinity position of the CTCF motif in gastrointestinal cancers. *Hum. Mutat.* 37:904–13

Valton AL, Dekker J. 2016. TAD disruption as oncogenic driver. *Curr. Opin. Genet. Dev.* 36:34–40

Van Vlierberghe P, Ferrando A. 2012. The molecular basis of T cell acute lymphoblastic leukemia. *J. Clin. Investig.* 122:3398–406

Viny AD, Ott CJ, Spitzer B, Rivas M, Meydan C, et al. 2015. Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *J. Exp. Med.* 212:1819–32

Vogelstein B, Kinzler KW. 2004. Cancer genes and the pathways they control. *Nat. Med.* 10:789–99

Walker CJ, Miranda MA, O'Hern MJ, McElroy JP, Coombes KR, et al. 2015. Patterns of *CTCF* and *ZFHX3* mutation and associated outcomes in endometrial cancer. *J. Natl. Cancer Inst.* 107:djv249

Walter MJ, Shen D, Shao J, Ding L, White BS, et al. 2013. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia* 27:1275–82

Wang H, Maurano MT, Qu H, Varley KE, Gertz J, et al. 2012. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* 22:1680–88

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

Xiao T, Wallace J, Felsenfeld G. 2011. 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.* 31:2174–83

Yu W, Ginjala V, Pant V, Chernukhin I, Whitehead J, et al. 2004. Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat. Genet.* 36:1105–10

Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell* 13:291–98

Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, et al. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *PNAS* 111:996–1001