# Review

**Cell**
PRESS

# Chromatin Insulators:
# Linking Genome Organization to Cellular Function

Jennifer E. Phillips-Cremins[1,2] and Victor G. Corces[1,*]
[1]Department of Biology, Emory University, Atlanta, GA 30322, USA
[2]Program in Systems Biology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605-0103, USA
*Correspondence: vcorces@emory.edu
http://dx.doi.org/10.1016/j.molcel.2013.04.018

A growing body of evidence suggests that insulators have a primary role in orchestrating the topological arrangement of higher-order chromatin architecture. Insulator-mediated long-range interactions can influence the epigenetic status of the genome and, in certain contexts, may have important effects on gene expression. Here we discuss higher-order chromatin organization as a unifying mechanism for diverse insulator actions across the genome.

## Introduction

Genomes of metazoan organisms are packaged in a hierarchy of topological configurations that allow for intricate spatiotemporal regulation of complex nuclear functions such as transcription, replication, recombination, and DNA repair (Misteli, 2007). In G1, for example, the chromatin fiber must be arranged in a manner that is favorable for regulation by a complex cascade of transcription factors and chromatin-modifying enzymes. In principle, this architecture should also leave chromosomes poised for reorganization, condensation, decondensation, and reassembly during each subsequent S phase and mitosis. Reports of fractal organization of genomes provide a biophysical model for how this folding and unfolding can take place in a rapid and precise manner at a larger scale (Lieberman-Aiden et al., 2009), but the molecular mechanisms that regulate chromatin organization at a finer sub-Mb scale are poorly understood.

Understanding the organizing principles responsible for three-dimensional (3D) folding of chromatin remains an important and unachieved goal. Progress has been made possible in recent years by the development of a cadre of chromosome conformation capture (3C)-based molecular techniques that allow high-resolution mapping of inter- and intrachromosomal interactions (de Wit and de Laat, 2012; Dekker et al., 2002). This information, coupled with genome-wide maps of the distribution of chromatin binding proteins obtained by chromatin immunoprecipitation sequencing (ChIP-seq), has made it possible to correlate protein occupancy with large-scale structural features, as well as point-to-point looping interactions across the genome.

Here, we review evidence suggesting that insulator proteins have a conserved role across metazoans as architectural proteins that orchestrate chromatin organization. We use observations in mammals and in *Drosophila melanogaster* to discuss molecular mechanisms regulating the myriad of intra- and inter-chromosomal interactions between regulatory elements and insulator proteins across the genome. Mechanisms of chromatin folding are discussed in the context of transcription, but we note that similar principles could apply to other genome functions linked to chromatin organization (e.g., replication and recombination) (Gilbert et al., 2010; Jhunjhunwala et al., 2009).

## Modularity and Distribution of Insulators across the Genome

CTCF is considered the primary insulator in mammals. The protein is ubiquitously expressed across most mammalian tissues (Wendt et al., 2008) and is required for early mouse development (Fedoriw et al., 2004). Homozygous CTCF deletion results in early embryonic lethality (Heath et al., 2008; Splinter et al., 2006), and conditional knockdown in mouse oocytes leads to mitotic defects upon fertilization and delayed progression to the blastocyst stage (Fedoriw et al., 2004; Wan et al., 2008). In somatic cells, conditional knockout confirmed additional important roles for CTCF in cell-cycle progression, apoptosis, and differentiation (Heath et al., 2008; Soshnikova et al., 2010; Splinter et al., 2006). Thus, CTCF has a pervasive role across most cell types during mammalian development.

In addition to ubiquitous expression patterns, CTCF binding sites are also widely distributed across mammalian genomes. In the first ChIP-chip analysis of CTCF, Ren and colleagues reported ∼14,000 occupied sites in human IMR90 fibroblasts, with a genomic distribution of 46% intergenic, 22% intronic, 12% exonic, and 20% within 2.5 kb upstream of transcription start sites (TSSs) (Kim et al., 2007). Subsequent ChIP-seq analysis by Zhao and colleagues revealed ∼19,000–29,000 occupied sites in CD4[+], HeLa, and Jurkat cells, with genomic distributions of 49%–56% intergenic, 3%–4% exonic, 32%–33% intronic, and 7%–15% at TSSs (Barski et al., 2007; Cuddapah et al., 2009; Jothi et al., 2008). Most recently, Stamatoyannopoulos and colleagues found an average of 55,000 CTCF sites per cell type when comparing 19 different cell lines, whereas Chen et al. (2012) reported ∼66,800 occupied sites from each of 38 different cell types from the ENCODE project (Wang et al., 2012). Advances in molecular and computational ChIP-seq technologies are the most probable explanation for the marked increase in sensitivity for occupied CTCF sites across the genome.

In *Drosophila*, there are five subclasses of insulator sequences (reviewed in Gurudatta and Corces, 2009). Each subclass is defined by common accessory proteins and a unique DNA binding protein, including: suppressor of hairy wing (Su[Hw]), *Drosophila* CTCF (dCTCF), boundary-element-associated factor of 32 kDa (BEAF-32), GAGA binding factor, and Zeste-white 5

(Zw5). The biological significance of five different insulators is unclear. However, because some subclasses are only found in *Drosophila*, a leading hypothesis is that multiple insulators might be critical for precise regulation of compact genomes that have much less distance between genes than in vertebrates. For example, the BEAF-32 insulator has recently been linked to the independent regulation of adjacent genes transcribed in opposite orientations (Yang et al., 2012).

*Drosophila* insulator proteins are ubiquitously expressed, and null mutations (with the exception of Su[Hw]) result in lethality (Gurudatta and Corces, 2009). ChIP analyses indicated that insulators bind to 3,000–6,000 sites across the *Drosophila* genome and show distinct distributions with respect to genomic features (Bartkuhn et al., 2009; Bushey et al., 2009; Nègre et al., 2010; Wood et al., 2011). BEAF-32 is preferentially located in promoter regions, whereas Su(Hw) is biased toward intergenic regions. By contrast, dCTCF follows an intermediate distribution similar to that of CTCF in mammals. Distinct distributions of different subclasses of insulators may be important for predicting their function(s) in genome regulation.

Growing evidence also suggests a widespread role for TFIIIC as the most evolutionarily conserved insulator (Van Bortle and Corces, 2012). Protists, fungi, and plants lack CTCF (Heger et al., 2012) and may rely on TFIIIC to carry out aspects of insulator function. For example, transfer RNA (tRNA) genes containing TFIIIC binding sites show insulator activity in *S. cerevisiae*, *S. pombe*, and mammals (Raab et al., 2012). In addition, TFIIIC sites unrelated to tRNA genes have been mapped across the mouse genome and appear to correlate with CTCF, suggesting that these two insulators may cooperate at specific genomic sites (Moqtaderi et al., 2010). Interestingly, both cohesin and condensin interact with TFIIIC and are required for its function (D'Ambrosio et al., 2008), suggesting that this insulator might show mechanistic similarities to CTCF (discussed below).

## Controversy Surrounding Insulator Mechanisms of Action

Insulators have been linked to a vast range of genomic functions, including activation, repression, enhancer-blocking (EB) insulation, barrier insulation, promoter-proximal pausing, alternative splicing, and protection from DNA methylation. The molecular mechanisms by which insulators confer these pleiotropic effects across the genome remain poorly understood. It has been suggested that mammalian CTCF serves distinct functional purposes by binding to divergent consensus sequences and subsequently recruiting different binding partners and posttranslational modifications (Ohlsson et al., 2010). In the case of *Drosophila*, an influential idea is that the unique insulator subclasses may be responsible for performing distinct functions. Thus, an important unresolved question is whether insulators are true multivalent factors with the ability to perform many contrasting functions, or whether there is a single unifying mechanism that can explain these divergent roles.

## Mammalian CTCF and 3D Chromatin Architecture

A body of locus-specific and genome-wide evidence now points to a primary role for mammalian CTCF in genome organization.

Prior to the availability of deep-sequencing technologies, several clues had already emerged supporting this role. First, massspectroscopy analysis of Flag-tagged CTCF purified from HeLa cells revealed that CTCF can form both homodimers and multimers in vivo (Yusufzai et al., 2004). Second, yeast two-hybrid experiments demonstrated that CTCF has the capacity to bind other CTCF molecules in vitro (Yusufzai et al., 2004). Third, CTCF molecules bound to probes encoding divergent CTCF consensus sequences also dimerized in vitro (Pant et al., 2004). Finally, glutathione S-transferase (GST) pull-down assays revealed that the C terminus of CTCF binds to the 11-zinc-finger domain of CTCF in vitro (Pant et al., 2004). Together, these data provided the initial biochemical evidence in mammalian systems to support a role for CTCF in long-range looping interactions.

More recently, 3C-based methods have been leveraged to analyze higher-order chromatin architecture at kb resolution. Independent studies at the mouse β-globin, H19/insulin growth factor 2 (Igf2), and major histocompatibility complex class II genomic loci indicated that CTCF sites are important for long-range interactions between specific genomic elements (Kurukuti et al., 2006; Li et al., 2008; Majumder et al., 2008; Murrell et al., 2004; Splinter et al., 2006; Yoon et al., 2007). Subsequent studies have leveraged 3C to identify insulator-mediated contacts at many mammalian genomic loci, including, but not limited to, human β-globin (Hou et al., 2010), human apolipoprotein (Mishiro et al., 2009), human kcnq5 (Ren et al., 2012), human and mouse HoxA (Ferraiuolo et al., 2010; Kim et al., 2011), human insulin (Xu et al., 2012), and human Interferon-γ (Hadjur et al., 2009). In most of these studies, global RNAi-mediated knockdown of CTCF resulted in a marked reduction of the 3C signal, providing direct evidence that CTCF is required for at least some subset of long-range interactions.

Handoko et al. (2011) used an independent technique termed chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) to find a subpopulation of CTCF-mediated chromatin interactions throughout the genome. The authors identified 1,816 high-confidence 3D interactions (1,480 intrachromosomal and 336 interchromosomal) connected by CTCF in mouse embryonic stem cells (ESCs). CTCF small interfering RNA in ESCs showed reduced interaction of specific inter- and intrachromosomal contacts selected for validation, suggesting that CTCF is essential for the formation of specific long-range interactions. Most recently, Dekker and colleagues reported that CTCF is highly enriched in long-range interactions between TSSs and distal regulatory elements throughout ENCODE pilot regions spanning 1% of the human genome (Sanyal et al., 2012).

Insights into the mechanisms governing insulator-mediated genome organization came with the discovery that CTCF colocalizes with cohesin at thousands of sites across mammalian genomes (Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wendt et al., 2008). Indeed, 50%–80% of CTCF-occupied sites overlap with cohesin, depending on the cell line. Because cohesin proteins are traditionally thought to function in sister-chromatid cohesion, an influential model suggests that CTCF recruits cohesins to DNA, and then, in turn, cohesin subunits form a ring-like structure that stabilizes higher-order organization of chromatin during interphase (Gause et al., 2008). Direct evidence from knockdown experiments is consistent

with this model thus far, in that RNAi for cohesin subunits results in the disruption of long-range looping interactions at several distinct loci (Hadjur et al., 2009; Hou et al., 2010; Mishiro et al., 2009; Nativio et al., 2009). Although it is clear that both cohesin-dependent and cohesin-independent CTCF sites can form long-range interactions, a critical unresolved question would be why some sites require cohesin and some sites do not. Altogether, these data provide unequivocal evidence that mammalian CTCF is involved in and essential for higher-order chromatin organization throughout the genome.

### Drosophila Insulators and 3D Chromatin Architecture

Direct and indirect evidence is also consistent with a role for *Drosophila* insulators in mediating 3D chromatin interactions. The scs and scs′ insulators flank *hsp70* genes and bind Zw5 and BEAF proteins, respectively. These insulators are separated by 15 kb on the linear genome but show high interaction frequency by 3C in *Drosophila* embryos (Blanton et al., 2003). Coimmunoprecipitation (coIP) and GST pull-down assays indicated that Zw5 and BEAF proteins can bind directly to each other in vitro and in vivo. These results support the idea that direct heterodimerization by insulator DNA-binding proteins could be one mechanism driving scs-scs′ looping interactions.

Another leading idea is that insulators form long-range interactions by recruiting cofactors such as centrosomal protein of 190 kDa (CP190) and/or modifier of mdg4 (Mod[mdg4]) to the DNA-binding proteins. Yeast two-hybrid experiments indicate that Mod(mdg4) proteins can form direct heterodimers with Su(Hw) and also homodimerize with each other (Gause et al., 2001; Ghosh et al., 2001). Furthermore, coIP, yeast two-hybrid, and affinity chromatography experiments suggested that CP190 can bind directly to CTCF, Su(Hw), and Mod(mdg4) insulator proteins (Gerasimova et al., 2007; Mohan et al., 2007; Pai et al., 2004). Different subclasses of insulators group together in *Drosophila* nuclei in subnuclear structures termed "insulator bodies." CP190 is essential for the formation of these bodies, suggesting that CP190 homodimerization may be a key mechanism in the formation of a 3D interaction network in *Drosophila* nuclei.

Contrary to vertebrates, *Drosophila* insulators do not appear to rely on cohesin to establish or maintain interactions with other sequences in the genome (Dorsett, 2009). We favor a model in which CP190 or Mod(mdg4) plays a similar functional role to cohesin, given that both proteins contain BTB domains that might be involved in mediating interinsulator interactions between independent genomic loci. In addition to CP190 and Mod(mdg4), several other proteins such as Chromator and L(3)mbt have recently been shown to interact or colocalize with *Drosophila* insulator proteins, but their possible role in chromatin organization has not been studied in detail (Gan et al., 2011). Importantly, a subset of so-called "aligned insulator elements" contain clustered occupied sites for CP190, Su(Hw), BEAF-32, and/or dCTCF in close proximity to each other within 100–300 bp-sized elements across the genome (Van Bortle et al., 2012). The presence of multiple insulator proteins clustered together within a small genomic element might give these sequences a unique role in 3D chromatin organization compared to sites that only bind single insulators (discussed in the following section).

### Mechanisms of Insulator-Mediated Chromatin Organization

#### Insulators as Boundaries of Higher-Order Topological Domains

Four independent studies have reported the discovery of highly self-interacting genomic units termed topologically associating domains (TADs) (Dixon et al., 2012; Hou et al., 2012; Nora et al., 2012; Sexton et al., 2012). Genomic sequences within TADs have a high frequency of 3D interactions with each other compared to sequences in adjacent TADs. Between TADs are distinct boundaries where chromatin interactions switch their directionality from an upstream bias (interactions within the current TAD) to a downstream bias (interactions within the adjacent TAD).

What are the mechanisms that causally demarcate the boundaries of TADs? In mammals, Ren and colleagues found that >75% of all TAD boundaries contain CTCF-occupied sites. Specifically, ~28% of all boundaries contain CTCF plus active housekeeping genes, ~20% contain CTCF and other genes, and ~28% contain CTCF alone without genes, whereas ~9% have only genes and ~15% do not display a particular mark (Dixon et al., 2012). In *Drosophila*, boundaries are enriched for BEAF-32, CTCF, and CP190, but not Su(Hw) (Sexton et al., 2012). There also appears to be an enrichment for aligned insulator elements containing binding sites for two or more insulator DNA-binding proteins plus CP190, whereas single-insulator sites are enriched inside TADs (Hou et al., 2012). On the basis of this finding, we hypothesize that clusters of mammalian CTCF and its numerous binding partners (reviewed in detail in Herold et al., 2012; Ohlsson et al., 2010) might also create tandemly aligned sites that contribute to the formation of boundaries between TADs in mammals.

It is also important to note that factors other than insulators also contribute to the formation of TAD boundaries. Indeed, in *Drosophila* and mammals, boundaries are also enriched in active genes, suggesting that high levels of transcription may be essential for the establishment and/or maintenance of topological domains. TAD boundaries are also enriched in tRNA genes and Alu/B1 and B2 SINE elements (Dixon et al., 2012), which contain binding sites for TFIIIC (Lunyak and Atallah, 2011). Because both CTCF and TFIIIC interact with cohesin and condensins (D'Ambrosio et al., 2008), we speculate that TFIIIC alone or in combination with CTCF might have a causal role in genome organization at TAD boundaries.

Together, these data suggest that insulators (with or without active transcription) might be important causal factors in the organization of topological transitions from one TAD to another. Several important questions related to the mechanisms of TAD boundary function remain to be answered: (1) Is CTCF necessary and/or sufficient for functional boundary formation? (2) What additional mechanisms (e.g., transcription) causally affect the specificity of CTCF action at TAD borders throughout the genome? (3) Are CTCF and active chromatin modifications a cause or consequence of boundary formation? and (4) How are insulators at TAD boundaries versus internal to TADs mechanistically distinguished?
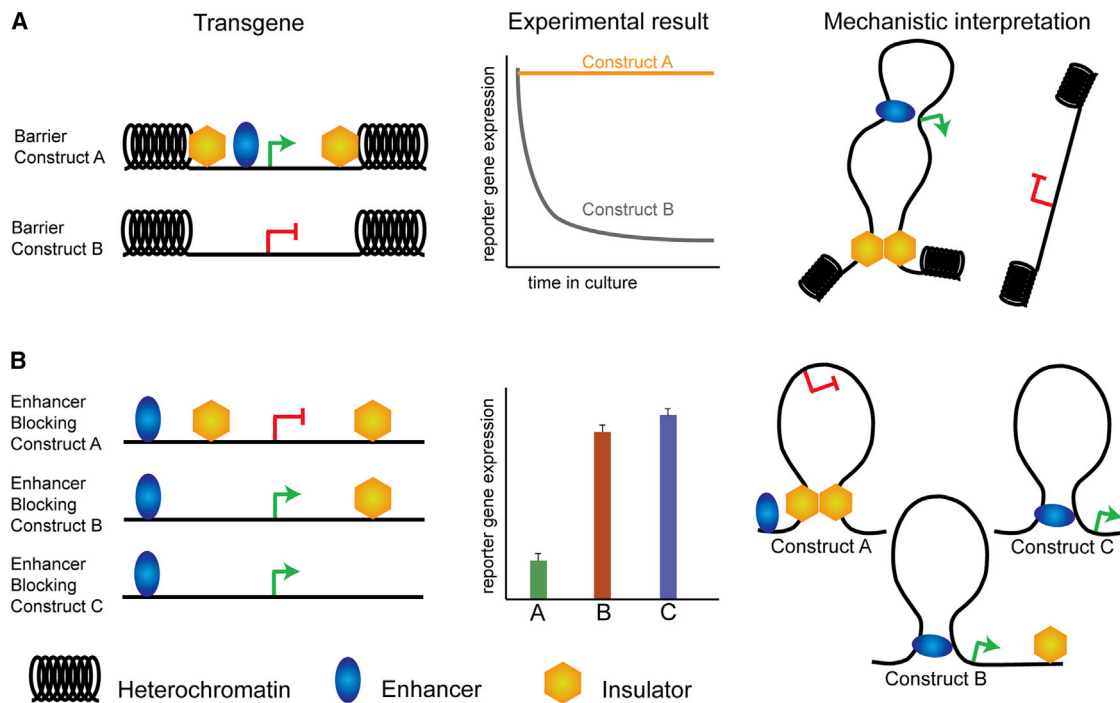
**Figure 1. Experimental Paradigms for Enhancer Blocking and Barrier Insulation**
(A) To test barrier insulator activity, a transgene flanked by insulators is randomly integrated into the genome. Multiple integration sites are considered to control for position-effect variegation. If sequences are true barrier insulators, reporter expression over time in culture should remain constant, whereas a control transgene that does not contain insulators should eventually be silenced by encroachment of heterochromatin.
(B) To test EB insulator activity, transgene constructs are designed by placing a putative insulator sequence in various positions with respect to an enhancer driving a reporter gene. The degree of "insulation" (or ability to abrogate the enhancer) is assayed as the level of reporter gene expression after transient transfection, or integration of the vector, into target cells. For ruling out the effects of position-independent silencing, results are compared to control constructs in which insulators are placed adjacent to, but not in between, linked enhancer-promoter sequences. Limitations of these assays for consideration during data interpretation include the spacing between elements, which does not mimic the endogenous locus; the integration of the reporter into multiple ectopic genomic locations; and the frequent use of heterologous enhancer-promoter sequences that also do not represent the genomic context of insulator sequences.

## Insulators as Borders of Functional Chromatin Domains

Evidence that CTCF organizes 3D chromatin topology across the genome raises important questions related to the putative role for CTCF as a barrier insulator. Barrier insulators were originally defined by their ability to protect transgenes from position-effect silencing (defined in Figure 1A). Initial studies in vertebrates reported >1 kb-sized genomic sequences containing CTCF binding sites that exhibited barrier activity in transgene assays (Cho et al., 2005; Filippova et al., 2005; Pikaart et al., 1998). Interestingly, specific CTCF binding sites could be deleted without affecting the barrier activity of a ~275 bp-sized chicken HS4 insulator sequence (Recillas-Targa et al., 2002). On the basis of this observation, it was first proposed that CTCF is nonessential for barrier function and works primarily through EB mechanisms (discussed below).

More recently, the role for insulators in barrier function was brought back into question with a genome-wide query (Cuddapah et al., 2009). Cuddupah et al. identified 30,000+ domains of the repressive chromatin modification H3K27me3 that ranged in size from 5 to 25 kb across the human genome. A search for proteins binding at the borders on either side of these H3K27me3 domains revealed 1,606 and 793 CTCF-occupied sites in CD4[+] and HeLa cells, respectively. Although this result is widely cited as supporting a genome-wide role for CTCF in

barrier function, a closer look suggests a different interpretation. Only a very small fraction of H3K27me3 domain borders contain CTCF: 1,578/39,900 (~4%) in CD4[+] cells and 771/32,704 (~2.4%) in HeLa cells. Furthermore, only a very small number of total CTCF binding sites (4%–6%) are associated with these chromatin borders. We interpret this result to support the original hypothesis (Recillas-Targa et al., 2002) that there must be other mechanisms, in addition to or independent from CTCF, that are essential for demarcating most of the H3K27me3-marked repressive domains.

What is known about the mechanism(s) by which the small fraction of CTCF sites at borders contribute to the establishment of functional domains of heterochromatin? Cuddapah et al. (2009) observed that the majority of their identified repressive domains were cell-type specific. When considering only borders bound by CTCF, only 5%–11% were constant between cells types and also had constant CTCF occupancy, whereas 23%–40% were cell-type specific and mirrored by cell-type-specific binding of CTCF. By contrast, 55%–66% of borders bound by CTCF displayed constant occupancy of CTCF but cell-type-specific boundary function. From a mechanistic perspective, these observations suggest that CTCF occupancy is probably not the critical factor for barrier function.

Two recent studies in *Drosophila* explored the mechanistic role for insulators in maintaining borders for domains of silencing histone modifications (Schwartz et al., 2012; Van Bortle et al., 2012). Pirrotta and colleagues analyzed H3K27me3 domains across the genome and reported that half of these domains showed a gradual decline, whereas half (~100) had edges that represented sharp transitions that could be characterized as borders. Computational analysis of the ~100 sharp borders demonstrated that ~33% display insulator binding and actively transcribed genes, whereas 44% have only actively transcribed genes and ~20% have only insulators. For borders marked only by insulator proteins, 75% showed slight spreading of the H3K27me3 into the surrounding regions after insulator knockdown. For the remaining borders that contain actively transcribed genes (with or without insulators), there was negligible effect on H3K27me3 spreading after insulator knockdown. Van Bortle et al. (2012) also explored H3K27me3 domains in *Drosophila* Kc cells with a focus on the role for insulators in demarcating domains. The authors reported that only 2% of independent CTCF sites without any other insulators were present within 5 kb of H3K27me3 borders, whereas 8% of aligned insulator sites (e.g., CTCF in combination with BEAF-32 and/or Su [Hw]) were located within 5 kb of H3K27me3 borders. In this study, downregulation of *Drosophila* insulators by RNAi, individually or in combination, resulted in a decrease of H3K27me3 levels within the domain but no clear spreading of the modification. These results suggest that the role for insulators may not be to maintain borders of repressive chromatin domains but rather to maintain the level of silencing by alternative mechanisms. One possibility is that insulators are involved in the recruitment of Polycomb (Pc)/H3K27me3 domains to Pc bodies, where clustering of these domains is necessary for the maintenance of silencing (Pirrotta and Li, 2012).

If insulators are not the primary contributor to barrier activity, then what other factors play a role? Seminal studies with the chicken HS4 insulator provided the first mechanistic clues. Whereas deletion of the CTCF site had no effect, deletion of either VEZF1 or USF1/USF2 binding sites markedly disrupted barrier activity (Dickson et al., 2010; Recillas-Targa et al., 2002; West et al., 2004). Intriguingly, the HS4 insulator contains both CTCF and high levels of active chromatin modifications H3K4me2, H3K9ac, and H3K14ac. USF proteins recruit H3K4-specific methyltransferase SET7/9 and the H3-specific acetyltransferase PCAF to HS4. Furthermore, deletion of the USF binding site resulted in loss of active histone modifications in addition to loss of barrier activity (Litt et al., 2001; West et al., 2004). Importantly, although positive histone modifications were necessary, they were not sufficient for insulator-barrier activity, because deletion of VEGZF1 sites also resulted in loss of barrier insulation without disrupting chromatin modifications (Dickson et al., 2010). These data are consistent with the idea that functional borders may be modular and require several different epigenetic mechanisms in addition to CTCF.

Together, these results suggest that barriers are complex genomic elements that require the combinatorial action of multiple proteins and chromatin modifications with distinct roles. A leading model from these studies is that multiple proteins (in addition to CTCF) bind to barrier insulator sequences to recruit chromatin-modifying enzymes that lay down persistent positive histone modifications to prevent the processive spread of heterochromatin. The primary role for insulators in these modular elements might be to mediate intra- or interchromosomal interactions that form the topological basis of barrier function, whereas the other proteins and epigenetic modifications have the causal responsibility of preventing the spread of the repressive chromatin mark.

### Blurring the Boundaries between Borders, Barriers, and Loops

Given the diversity of insulator actions across the genome, a critical question is whether the mechanisms by which insulators confer barrier function are the same as the mechanisms by which insulators demarcate TAD boundaries. In *Drosophila*, many TADs contain a heterogeneous mixture of chromatin states, suggesting that TADs do not always correspond to functional chromatin domains defined by specific histone modifications (Hou et al., 2012). In mice and humans, a small fraction of TADs correspond to domains of H3K9me3, whereas a large proportion contain a mixture of chromatin modifications (Dixon et al., 2012). These results suggest that only a subset of TADs might play a causal role in demarcating blocks of repressive chromatin modifications, whereas a larger subset of TADs have purposes broader than delimiting chromatin domains that have not yet been defined.

In the subset of cases where TADs align with domains of repressive chromatin modifications, do the modifications themselves dictate the formation of TADs, or do TAD boundaries causally mark heterochromatin borders? To directly test these questions, Heard, Dekker, and colleagues first knocked down enzymes that catalyze repressive histone modifications in ESCs (Nora et al., 2012). They showed that TADs were not disrupted on the X chromosome after reducing levels of H3K27me3 or H3K9me2, suggesting that repressive chromatin modifications are not causal for the creation of TAD boundaries. By contrast, the authors also reported that a 58 kb deletion at a boundary between two TADs on the X chromosome (depicted in Figure 2A) disrupts topological organization and results in numerous ectopic point-to-point looping interactions between adjacent TADs (depicted in Figure 2B) (Nora et al., 2012). Importantly, a CTCF-occupied site that has been shown by Lee and colleagues to have insulator activity in transgene assays binds within the deleted region (Spencer et al., 2011). This provides a powerful clue suggesting that the actual read out of an insulator transgene system can be functionally observed in the genome as a boundary between TADs. Together, these data suggest that information contained at the boundaries is essential for topological organization of TADs. However, it is not yet known whether deletion of TAD boundaries also results in the spread of heterochromatin marks into adjacent topological domains, and future genetic experiments will be important for assessing whether TADs are causal for demarcating blocks of repressive H3K9me3 or H3K27me3 across the genome.

We hypothesize that CTCF demarcates topological units by simply mediating specific long-range interactions between boundaries on each side of a TAD. Although CTCF is enriched at TAD boundaries, putative boundary elements represent only a small fraction of total CTCF sites. Of the 32,000
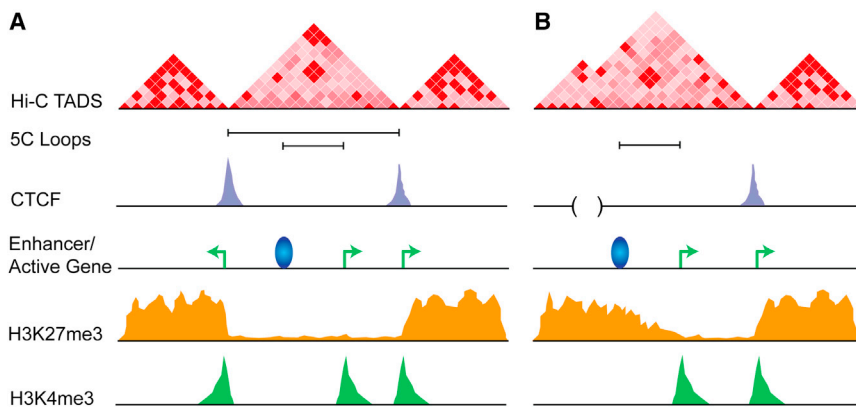
**Figure 2. Higher-Order Genome Organization as a Unifying Mechanism for Insulator Function**
(A) Model for CTCF in mediating TAD boundaries and intra-TAD genomic organization via long-range interactions in wild-type cells.
(B) Blurring of TAD boundaries after deletion of a 50–80 kb boundary between TADs that contains CTCF and active genes (adapted from Nora et al., 2012). In principle, the TAD boundary would also be responsible for demarcating the border of the repressive H3K27me3 mark. Thus, we hypothesize that deletion of the boundary would lead to aberrant heterochromatin spreading. TADs are defined by the genome-wide 3D mapping technology Hi-C. Counts are directly proportional to the frequency by which genomic fragments interact in the 3D nucleus. Deep red and light pink pixels depict high- and low-frequency interactions, respectively. Fragment-to-fragment looping interactions at kb resolution can be mapped by chromosome conformation capture carbon copy (5C) and are depicted by black bars connecting two genomic segments. The CTCF ChIP-seq track is in purple. The active enhancer is depicted as a blue ball. Active genes are depicted as green arrows. The ChIP-seq track for H3K27me3 repressive chromatin is in orange.

CTCF-occupied sites in the mouse genome identified by Ren and colleagues, only 4,800 (15%) are at boundaries, suggesting that the large majority of CTCF sites do not serve a boundary function (Dixon et al., 2012). Thus, we speculate that CTCF sites that are not at boundaries could participate in the formation of additional long-range looping interactions within TADs, between TADs, and possibly between chromosomes. Indirect evidence supporting this idea is underscored by a recent ChIA-PET analysis suggesting that a large proportion of CTCF-mediated loops are smaller than the Mb-length scale of TADs (Handoko et al., 2011). Furthermore, an analysis of the epigenetic states of chromatin around and within loops connected by CTCF in mouse ESCs revealed that >70% of CTCF-mediated looping interactions separate genomic regions enriched for divergent histone modifications. Together, these data provide preliminary evidence indicating that some proportion of CTCF-mediated long-range interactions serve to demarcate different chromatin domains. Correlating these interactions with TADs would shed further light into the link between TADs and looping of the 10 nm chromatin fiber.

Altogether, we favor a model in which traditional barrier and EB insulators primarily function to mediate higher-order genome folding and this 3D organization can result in a wealth of functional outcomes depending on the genomic context (reviewed in Phillips and Corces, 2009). We hypothesize a model in which CTCF is necessary but not sufficient for the formation of TADs through its primary role in long-range looping interactions (depicted in Figures 2A and 2B). Within this model, an alternative explanation for the enrichment of CTCF at boundaries windowed at 40 kb or greater is that CTCF is actually placed at the edges of each domain, and these occupied sites causally shape each TAD through a hierarchy of long-range looping interactions with genomic elements on the other edge of the TAD.

## Insulator-Mediated Chromatin Organization and Gene Regulation

Although significant evidence now points to a role for CTCF in genome organization, a much more controversial topic is related to whether and how insulator-mediated chromatin organization is linked to gene expression. This question remains unanswered in part due to significant technical challenges in analyzing insulator function at sub-kb resolution in the context of the 3D topology of an endogenous genomic locus. Emerging rules governing the epigenetic systems linking insulators, chromatin modifications, higher-order architecture, and gene expression are discussed in the sections that follow.

## Enhancer Blocking Is Rare and Context Dependent In Vivo

Insulators are traditionally thought to regulate gene expression through EB mechanisms. The classic definition of an enhancer blocker is the ability to block inappropriate communication between enhancers and promoters in a position-dependent manner (West et al., 2002). Most of our knowledge of EB insulators comes from experiments that rely on transgene constructs (described in Figure 1B). Use of this experimental paradigm has resulted in the identification of hundreds of EB insulator sequences to date (reviewed comprehensively in Herold et al., 2012; Ohlsson et al., 2010). However, finding sequences that confer EB function in a physiologically relevant system at an endogenous locus in vivo has been rare.

Two recent genome-wide analyses have provided new insights that now prompt us to re-evaluate the role for EB mechanisms in mammalian systems (Sanyal et al., 2012; Shen et al., 2012). Shen et al. used a computational approach to evaluate EB models. Under the assumption that all CTCF sites serve EB mechanisms, the correlation between the intensity of H3K4me1 signal at enhancers and polymerase II (Pol II) signal at promoters was calculated only for matched enhancer-promoter pairs within each CTCF-demarcated domain. The resulting Spearman's correlation coefficient was only slightly higher than that calculated for the random enhancer-promoter pairing control. Furthermore, >35% of enhancer-promoter pairs in CTCF-marked domains were anticorrelated, suggesting that an EB assumption results in a very large number of incorrectly paired regulatory sequences. Consistent with this finding, Sanyal et al. (2012) mapped long-range interactions between promoters and distal regulatory elements throughout the 30 Mb-sized ENCODE regions and found that >75% of identified looping

interactions pass over one or more CTCF-occupied sites. Together, these studies suggest that enhancer blocking is not a pervasive mechanism for most CTCF sites in mammalian systems.

A critical aspect of assessing the validity of the notion that insulators function through EB mechanisms is determining whether this mechanism exists in vivo at endogenous genomic loci in a developmentally relevant chromatin environment. The classical example supporting a role for EB insulators is the parent-of-origin-specific expression of *Igf2* and *H19* genes in mice. However, more recent evidence suggests that long-range epigenetic mechanisms could also account for the observed imprinted expression of these genes (reviewed in Phillips and Corces, 2009). Similarly, at the mouse β-globin locus, the *3′HS1* and *HS5* regulatory elements are gold-standard sequences with demonstrated EB activity in transgene assays (Bulger et al., 2003; Farrell et al., 2002). However, more recent evidence now demonstrates that *3′HS1* and *HS5* regulatory elements are connected in 3D space by long-range looping interactions mediated by CTCF (Splinter et al., 2006).

Are data at the β-globin locus more consistent with a model in which mammalian CTCF does not function through EB mechanisms, or a model in which EB insulation occurs as a consequence of looping interactions at endogenous loci? Interestingly, deletion of the 3′HS1 in vivo did not disrupt expression of β-globin or olfactory genes surrounding the β-globin locus during erythroid differentiation (Splinter et al., 2006). One caveat to this observation is that the undifferentiated erythroid progenitors used in these experiments do not normally express high levels of β-globin (Bulger and Groudine, 2010). Indeed, in a later study by Dean and colleagues, depletion of CTCF in K562 cells resulted in a reduction of β-globin gene expression (Hou et al., 2010). Nevertheless, neither deletion of a CTCF binding site nor global knockdown of CTCF in independent studies led to misregulation of the olfactory genes surrounding the β-globin locus. Together, these data support the idea that CTCF binding at 3′HS1 is necessary for long-range contacts between regulatory elements around developmentally regulated globin genes, but the function in this specific case is probably not EB insulation.

In contrast to mammalian systems, data consistent with a role for insulators in the proper regulation of gene expression through EB mechanisms in the endogenous context comes from now-classical studies of the homeotic genes of the bithorax complex in *Drosophila.* Spatiotemporal expression of these three genes involves an intricate collection of enhancers, Polycomb response elements, and insulators (Barges et al., 2000). dCTCF binds to the Fab-8 insulator between the *abd-A* and *Abd-B* genes, and mutations in CTCF result in abdominal homeotic phenotypes due to misexpression of *Abd-B* (Gerasimova et al., 2007; Mohan et al., 2007). More recent evidence in support of a role for insulators in EB was obtained from analyses of insulator protein binding throughout the genome after treatment of *Drosophila* Kc cells with the steroid hormone ecdysone (Wood et al., 2011). The ecdysone-inducible *Eip75B* gene encodes at least four different transcripts, named Eip75B-RA through Eip75B-RD. A poised but inactive CTCF site separates the Eip75B-RB promoter from regulatory sequences located within the locus. During the first 3 hr after ecdysone induction, the

Eip75B-RB RNA is transcribed. After 48 hr, the CTCF insulator becomes activated by recruitment of CP190, resulting in down-regulation of the Eip75B-RB promoter. CP190 recruitment occurs in parallel with increased interaction frequency between this site and another CP190 site <100 kb downstream as measured by 3C. RNAi for CP190 reduces the frequency of this 3D interaction down to wild-type levels. We hypothesize that, in this case, the inducible insulator present in the *Eip75B* locus could possibly inhibit the interaction between enhancers and promoters by creating loops through interactions with adjacent insulators (Wood et al., 2011).

Much more experimental work is necessary for understanding EB mechanisms in light of the newly identified TADs. Data so far suggest that canonical EB insulation is not a widespread regulatory mechanism in mammals. However, we cannot yet rule out the possibility that in certain specific cases, insulator-mediated looping interactions can prevent inappropriate enhancer-promoter interactions. In the case of *Drosophila*, studies so far underscore the importance of insulators in the proper regulation of developmental genes. One possibility is that EB mechanisms are much more widespread in *Drosophila* due to the compact genome and greater need for regulating enhancer specificity. To date, the enhancers that are blocked by fly EB insulators have not yet been identified, so we cannot rule out the possibility that in flies, as in mammals, insulators are simply regulating the higher-order architecture that is critical for proper expression of a genomic locus. Another interesting possibility is that enhancers are not necessarily specific to a single gene and, instead, may activate all promoters within a genomic locus defined by a TAD. In this model, enhancers would have limited ability to sample the genomic space outside of the current TAD, suggesting that the CTCF sites causal for TAD boundary formation might also be the sites involved in EB insulation.

### Insulators Directly Tether Promoters to Distal Regulatory Elements

A leading idea is that constitutive insulator sites across cell types might be involved in constitutive long-range interactions that connect the larger architectural framework of the genome, whereas the cell-type-specific insulator sites have a more important role in regulating gene expression. Initial genome-wide studies observed ∼40%–70% overlap in CTCF-occupied sites between cell types and used this information to support the claim that CTCF binding is largely invariant (Kim et al., 2007). However, because 30%–60% can be cell-type specific, we hypothesize that variable CTCF sites play a functionally important role in genome regulation.

In a recent study by Stamatoyannopoulos and colleagues (Wang et al., 2012), CTCF occupancy was mapped by ChIP-seq in 19 different cell types and showed much more widespread differential occupancy than previously suggested. This result was confirmed in parallel by a study in which CTCF binding was compared across 38 different cell types from the ENCODE project. CTCF overlap between cell types was generally >50%, but an overlap as high as 80% was observed between similar cell lines (Chen et al., 2012). For example, two lymphocyte cell lines (GM12875 versus GM12873) showed ∼80% overlap, whereas overlap was as low as 25% in unrelated lineages (GM12801 versus HepG2). Together, these data indicate that

cell-type-specific CTCF sites are more widespread than previously thought and as important as constitutive CTCF sites.

A clue to the functional significance of cell-type-specific CTCF sites comes from the mouse ENCODE project, where ~34,000 and ~41,000 occupied sites were reported in ESCs and mouse embryonic fibroblasts, respectively (Shen et al., 2012). Interestingly, only 50%–60% of CTCF sites were found to be common between these cells, and these constitutive sites were biased toward promoters (22% enhancers, 36% promoters, 42% other). By contrast, cell-type-specific CTCF sites significantly overlapped with enhancers (50% enhancers, 24% promoters, 26% other). Thus, one possible reason for cell-type-specific CTCF sites would be to organize proper chromatin configurations that enable enhancer-promoter interactions instead of blocking them.

Consistent with the idea that CTCF helps tether promoters to distal genomic elements, Handoko et al. (2011) reported a small subset of CTCF-mediated loops that function to bring p300-bound enhancers in close spatial proximity to target genes. Similarly, Tijan and colleagues reported that TAF3, a component of the basal TFIID transcriptional machinery, binds directly to CTCF in coIP assays (Liu et al., 2011). Genome-wide analysis indicated that TAF3 is enriched at core promoters marked by TFIID subunits and H3K4me3, as well as at CTCF sites distal to promoters in ESCs. 3C analysis at *Mapk3* and *Psmd1* genes indicated that a distal element co-occupied by both CTCF and TAF3 can form a long-range interaction with a TAF3-bound promoter. This looping interaction might be functional, as a combination of CTCF and TAF3 knockdown reduced expression of *Mapk3* and *Psmd1* genes, presumably through disruption of loop formation. Although it is not yet clear whether distal sites are functional enhancers or some other regulatory element, these data provide evidence that CTCF sites can tether distal regulatory elements to promoters to regulate gene expression.

Finally, evidence also exists for CTCF-mediated loops that tether distal promoters to another promoter. This is perhaps best exemplified at the insulin (INS) locus in human pancreatic β-cells (Xu et al., 2011). The *SYT8* gene is important for insulin secretion in response to glucose, and it is located 300 kb from the *INS* gene. Felsenfeld and colleagues used 3C and 4C to show that *INS* and *SYT8* genes physically interact (Xu et al., 2011). This interaction is mediated by CTCF and increases in response to glucose. Depletion of CTCF or inactivation of the *INS* promoter results in a decrease of *SYT8* transcription. The results suggest that in addition to enhancer-promoter interactions, CTCF can also help in the coordination of gene expression by mediating long-distance interactions between the promoters of distally separated genes.

### A Role for CTCF in Segregating Enhancers during Limb Patterning

The mouse HoxD locus has 13 genes arranged in descending order (*HoxD13* to *HoxD1*) from 5′ (centromeric) to 3′ (telomeric) on the linear DNA (described in Figure 3A). Precise spatiotemporal expression of different *HoxD* isoforms within this locus is critical for proper development of proximal and distal segments of the limb bud (Spitz et al., 2005; Tarchini and Duboule, 2006).

Duboule and colleagues hypothesized that CTCF binding sites operate through EB mechanisms to insulate the early *HoxD1–9*

genes (expressed in the proximal limb bud and regulated by a 3′ gene desert containing early limb telomeric enhancers) from *Evx2* and the late *HoxD12–10* genes (expressed in the distal limb bud and regulated by a 5′ gene desert containing late limb centromeric enhancers) (Montavon et al., 2011). Consistent with this idea, ChIP-chip analysis (Soshnikova et al., 2010) in distal embryonic limb buds at embryonic day 10.75 (E10.75) revealed that seven of the nine *HoxD* genes were flanked by CTCF binding sites, whereas there were four CTCF sites in the centromeric gene desert and seven CTCF sites in the telomeric gene desert (Figure 3A).

To analyze the causal role for CTCF at the specific time when an EB mechanism would be needed in development, CTCF alleles were conditionally deleted in distal mouse limb buds at E10.75 with Cre-recombinase under control of the Prx1 promoter. Almost complete knockdown of CTCF messenger RNA (mRNA) and protein expression was achieved, and ChIP-chip confirmed a 96% loss of the wild-type CTCF-occupied sites. At E10.75, mutant and wild-type limb buds showed similar sizes and few differences in apoptosis. However, by E11.5, there was extensive apoptosis that ultimately resulted in severe shortening of fore- and hindlimbs. Thus, CTCF knockdown during development results in severe effects on the phenotype at organ, tissue, and cellular levels.

Analysis of gene expression at E10.75 after CTCF knockdown revealed 220 downregulated and 177 upregulated genes with an at least 1.5-fold change and an enrichment for genes involved in apoptotic pathways, oxidative-stress pathways, and mitochondrial functions. Several genes with CTCF binding in their TSS, such as *Evx2* and *HoxD13*, showed reduced expression by 4.5- and 2.4-fold, respectively, after CTCF knockdown. By contrast, the more telomeric *HoxD* genes that were not active in the limb bud, such as *HoxD8* and *HoxD9*, showed a 2.4- and 3-fold increase in expression, respectively. Importantly, although transcription was markedly deregulated upon CTCF knockdown, minimal changes in the spatial patterns of *HoxD* expression were observed. Similair to wild-type animals, *Evx2* and *HoxD13* expression remained localized to the distal limb bud, and *HoxD9* expression remained localized to the proximal limb bud. Before the discovery of the organization of the 3D genome into TADs, these results were originally interpreted to suggest that CTCF did not function as a canonical enhancer blocker but might function though more general transcription mechanisms.

To consider the above in the context of 3D genome organization, we compared the HoxD locus with TADs mapped by Hi-C in mouse ESCs and mouse cortex (Dixon et al., 2012). Intriguingly, the *HoxD* genes and multiple CTCF sites fall directly at the boundary between two TADs (Figure 3B). Gene deserts containing telomeric and centromeric enhancers coincide remarkably well with the interior of each adjacent TAD. This striking colocalization suggests that CTCF also has a role in higher-order organization at the *HoxD* locus, but further studies will be needed for exploring the causality in this relationship. Because recent reports suggest that insulators often work in combination with active genes to create boundaries between TADs, it is very probable that simple loss of CTCF sites would not be sufficient to disrupt the topological organization leading to aberrant
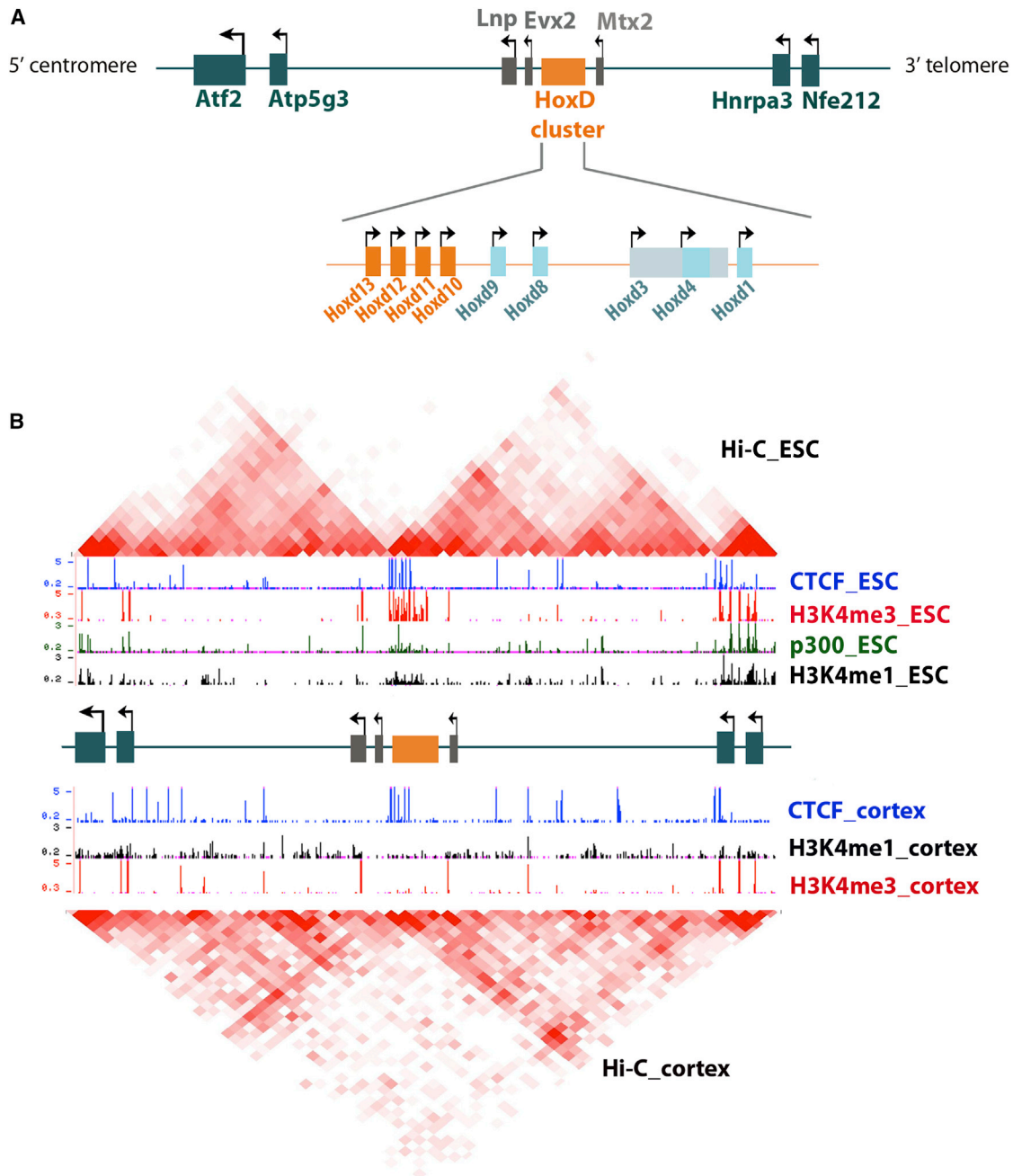
**Figure 3. Role of CTCF in Higher-Order Chromatin Architecture at the Mouse HoxD Locus**

(A) Two-dimensional organization of the HoxD locus. The *HoxD* gene cluster is a developmentally regulated locus that must be partitioned into discrete regulatory landscapes. 3′ Hox genes (*HoxD9*–*HoxD1*) are activated during early limb bud development via enhancers in a gene desert region on the 3′ side of the cluster toward the telomeres. 5′ Hox genes (*HoxD13*–*HoxD10*), as well as adjacent *Lnp* and *Evx2* genes, are activated later in development during patterning of digits, and this wave of transcription is controlled by different enhancers in a gene desert region on the 5′ side of the cluster toward the centromeres. Centromeric enhancers have been well characterized: there is a distal GCR (global control region) 180 kb upstream of *HoxD13* that contains multiple enhancers, as well as a proximal enhancer 50 kb upstream from *HoxD13*.

(B) Three-dimensional organization of the HoxD locus. TADs identified with Hi-C analyses by Dixon et al. (2012) are shown in ESCs (top) and a heterogeneous population of cells from the mouse cortex (bottom) with counts ranging from low (white) to high (deep red). Genome-browser tracks from Dixon et al. are also displayed for CTCF, H3K4me3, H3K4me1, and p300 in ESCs.

enhancer-promoter contacts. We speculate that true deregulation of the spatial patterns of *HoxD* gene expression would occur if the TAD organization was disrupted by dual knockdown

of transcription and CTCF. We also note that CTCF knockdown still results in marked misregulation of several developmental genes responsible for limb patterning, as well as many genes
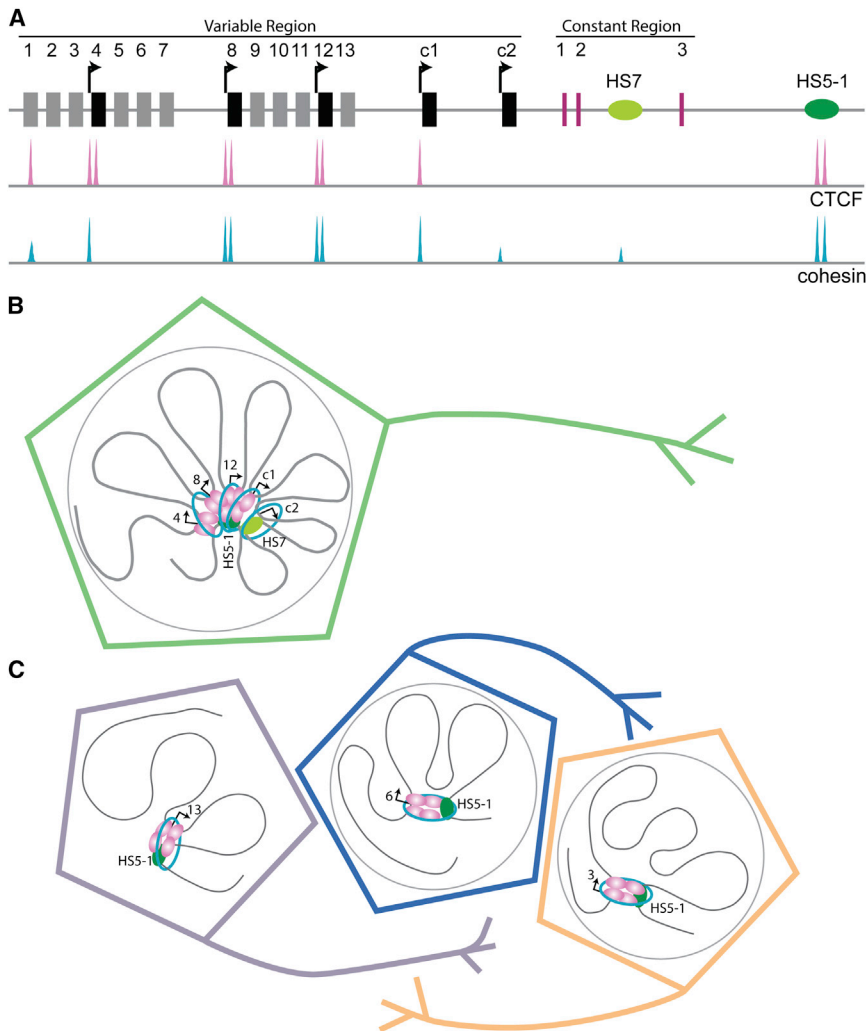
**Figure 4. Role for CTCF in Higher-Order Chromatin Architecture at the Human Pcdh Locus**

*Pcdh* genes encode a large number of calcium-dependent cell-cell adhesion molecules important in establishing neural diversity.

(A) Two-dimensional organization of the human Pcdh locus. Two enhancer elements downstream of the 13 alternative isoforms (1–13) and two c-type ubiquitous isoforms (c1 and c2) have been identified as necessary for appropriate tissue-specific expression: HS7 in the intron between constant exons 2 and 3 (light green) and HS5-1 downstream of constant exon 3 (dark green) (Ribich et al., 2006). In the diploid human neuroblastoma cell line SK-N-SH, CTCF- and cohesin-occupied sites have been mapped with ChIP-seq (adapted from Guo et al., 2012).

(B) Model for the role of enhancer-promoter looping interactions in variable exon expression in the diploid human neuroblastoma cell line SK-H-SH (adapted from Guo et al., 2012).

(C) Model for neural diversity created by alternative Pcdh isoform expression through looping interactions between alternative promoters and the downstream HS5-1 enhancer.

dividual neurons. The authors addressed this issue by using the human diploid neuroblastoma cell line SK-N-SH with stable expression of a select number of specific alternative and ubiquitous *Pcdh* isoforms (Figure 4A) (Guo et al., 2012). The power of this model system is that it represents an example of a simple expression pattern in a single neuron, thus enabling the study of the mechanisms of neural diversity for one specific clonal scenario.

Using ChIP-seq, the authors discovered that dual CTCF/cohesin sites are bound to the TSS and first exon of α4, α8, and α12 isoforms in SK-N-SH cells. CTCF/cohesin binding correlated with transcription, as these three isoforms were highly expressed compared to the other variable exons. Additionally, a single CTCF/cohesin site was mapped at the c1 isoform, and a CTCF-independent cohesin site was mapped at the c2 isoform. Both of these c-type ubiquitous isoforms were also expressed at low levels.

In addition to alternative promoters, it was also discovered that the HS7 enhancer contains a CTCF-independent cohesin binding site, whereas the HS5-1 enhancer contains dual CTCF/cohesion-occupied sites. (Figure 4A) (Guo et al., 2012). A previous study by the authors in mice had demonstrated that deletion of the CTCF-bound HS5-1 enhancer deregulates expression of specific *Pcdhα* isoforms and markedly disrupts CTCF binding at the promoters of these genes even though they are separated by up to 250 kb from the enhancer (Kehayova et al., 2011). This result suggested that CTCF-mediated long-range mechanisms may play a role in alternative promoter selection. To test this idea, the authors leveraged 3C to demonstrate that the HS5-1 enhancer forms strong 3D contacts with α8 and α12 and weaker interactions with α4, αc1, and αc2, whereas

encoding many basic cellular processes, ultimately leading to massive apoptosis. Because deregulation is often seen at genes that have CTCF bound to their promoter, we speculate that CTCF knockdown might disrupt long-range interactions that directly connect enhancers to promoters within TADs.

### A Role for CTCF in Alternative Promoter Selection to Generate Neural Diversity

Enormous diversity in neuronal phenotypes is generated during development through combinatorial expression of ~50 protocadherin (Pcdh) isoforms from three primary gene clusters (*Pcdhα*, *Pcdhβ*, and *Pcdhγ*). For example, diverse combinations of *Pcdhα* isoforms are created in humans through stochastic alternative promoter choice from the 15-variable first exons, followed by alternative splicing of the chosen alternative exons to three downstream constant exons (described in Figure 4A).

Maniatis and colleagues hypothesized that CTCF plays a critical role in alternative promoter selection during the generation of mammalian neural diversity (Guo et al., 2012). This question has been notoriously difficult to study due to the technical challenges of mapping protein binding and 3D chromatin interactions in in-

interactions with the inactive isoforms were undetected (Figure 4B) (Guo et al., 2012). Similarly, 3C analysis also detected 3D interactions between the HS7 enhancer and α8, α12, αc1, and αc2, but not with αc4. Importantly, lentiviral small hairpin RNA (shRNA) knockdown of CTCF led to reduced expression of several alternative *Pcdhα* isoforms and disruption of the 3D interactions at this locus.

Taken together, these data indicate that CTCF binding mirrors alternative promoter expression and that the binding of this protein is required for proper expression of these genes. Because this region was also shown by Ruan and colleagues to have numerous CTCF-mediated looping interactions (Handoko et al., 2011), this supports the idea that CTCF has an important role in creating the unique 3D configurations favorable for alternative promoter choice and expression of *Pcdhα* during the generation of neural diversity (Figure 4C).

### Insulators Influence Epigenetic States to Regulate Alternative Splicing

One way that CTCF influences gene expression is through the organization of long-range interactions that influence the epigenetic state of specific genomic loci (reviewed in Phillips and Corces, 2009). In a more recent exciting example of the interplay between insulators and epigenetic modifications, CTCF and DNA methylation were recently functionally linked to alternative splicing during lymphocyte development (Shukla et al., 2011). The CD45 gene was used as a model system because exclusion of exons 4, 5, and 6 is correlated with differentiation of peripheral lymphocytes. The long form of CD45 (containing exon 4) is expressed early in development, and the short form of CD45 (lacking exons 4, 5, and 6) is expressed in terminally differentiated lymphocytes. Prior to this study, proteins involved in exclusion of exons 4 and 6 during terminal differentiation were known, but exon 5 exclusion appeared to be controlled through different unknown mechanisms. Interestingly, Shukla et al. found that CTCF binds specficially to exon 5 and might be casually linked to splicing. CTCF binding correlated with inclusion of exon 5 in the CD45 transcript, whereas disruption of CTCF binding resulted in exclusion of exon 5 and a shortened CD45 transcript.

What is the mechanism by which CTCF affects splicing? Importantly, exon 5 showed high levels of DNA methylation and was not bound by CTCF at late stages in development when the exon was excluded from the CD45 transcript. Depletion of the DNA maintenace methyltransferase DNMT1 led to reduced DNA methylation, reacquisition of CTCF binding, and a subsequent increase in inclusion of exon 5 in the CD45 transcripts. Mechanistically, in vitro biochemical studies supported the idea that CTCF facilitates inclusion by promoting transient pausing of Pol II, while also allowing subsequent Pol II elongation after pausing. Altogether, this study provides evidence linking the epigenetic system of CTCF occupancy and DNA methylation to polymerase pausing and splicing. Independent studies indicate that CTCF may facilitate Pol II pausing in other genomic contexts (Kang and Lieberman, 2011; Paredes et al., 2012; Wada et al., 2009), and it will be very interesting to see results from more detailed genome-wide analyses of splicing, insulators, epigenetic modifications, and higher-order chromatin organization.

### Conclusions

A growing body of evidence now supports the idea that insulators are multifaceted regulatory sequences that modulate a variety of nuclear processes by mediating long-range interactions between distant sites in the genome. We favor a unifying mechanism for insulators in the formation of inter- and intrachromosomal interactions, with the underlying consensus and the local chromatin environment at a particular position in the genome providing specificity for protein conformations, binding partners, and posttranslational modifications that yield context-dependent effects on gene expression.

Given the global role for insulators in orchestrating genome organization, it is surprising that insulator knockdown appears to have only modest gobal effects on gene expression. One probable explanation is that the small effects on gene expression observed so far are a consequence of only partial knockdown of the proteins. Another probable explanation is that insulators are modular elements that work in conjunction with other mechanisms to perform diverse functions. Thus, insulator-protein knockdown might have only a partial effect accounted for by redundant mechanisms. Finally, because CTCF primarily functions to facilitate genome organization, we envision that only a fraction of CTCF-mediated interactions are important for gene expression, whereas another fraction might be important for genome topology, but not specifically required for cellular function.

A critical issue in dissecting the role of insulators in nuclear biology is to understand whether genome function is an effector that determines its three-dimensional organization, or whether insulator proteins play a structural role to instruct patterns of organization that then allow specific functional outcomes. As is often the case, the answer is probably a combination of both options. The role for insulators in defining TADs is often redundant with high gene density and transcription and, therefore, function appears to be an important contributor to the establishment and/or maintenance of topological chromosome domains. At a local level, regulation of enhancer-promoter interactions may first require the establishment of contacts between regulatory sequences by insulators, cohesin, and/or mediator, suggesting that function in some cases might be a consequence of architecture. Future studies leveraging deep sequencing in combination with genetic and biochemical perturbation studies should yield valuable insights into the causes and consequences of genome organization.

### REFERENCES

Barges, S., Mihaly, J., Galloni, M., Hagstrom, K., Müller, M., Shanower, G., Schedl, P., Gyurkovics, H., and Karch, F. (2000). The Fab-8 boundary defines

the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. Development 127, 779–790.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837.

Bartkuhn, M., Straub, T., Herold, M., Herrmann, M., Rathke, C., Saumweber, H., Gilfillan, G.D., Becker, P.B., and Renkawitz, R. (2009). Active promoters and insulators are marked by the centrosomal protein 190. EMBO J. 28, 877–888.

Blanton, J., Gaszner, M., and Schedl, P. (2003). Protein:protein interactions and the pairing of boundary elements in vivo. Genes Dev. 17, 664–675.

Bulger, M., and Groudine, M. (2010). Enhancers: the abundance and function of regulatory sequences beyond promoters. Dev. Biol. 339, 250–257.

Bulger, M., Schübeler, D., Bender, M.A., Hamilton, J., Farrell, C.M., Hardison, R.C., and Groudine, M. (2003). A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. Mol. Cell. Biol. 23, 5234–5244.

Bushey, A.M., Ramos, E., and Corces, V.G. (2009). Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions. Genes Dev. 23, 1338–1350.

Chen, H., Tian, Y., Shu, W., Bo, X., and Wang, S. (2012). Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. PLoS ONE 7, e41374.

Cho, D.H., Thienes, C.P., Mahoney, S.E., Analau, E., Filippova, G.N., and Tapscott, S.J. (2005). Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol. Cell 20, 483–489.

Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K., and Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. 19, 24–32.

D'Ambrosio, C., Schmidt, C.K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., and Uhlmann, F. (2008). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. Genes Dev. 22, 2215–2227.

de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev. 26, 11–24.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science 295, 1306–1311.

Dickson, J., Gowher, H., Strogantsev, R., Gaszner, M., Hair, A., Felsenfeld, G., and West, A.G. (2010). VEZF1 elements mediate protection from DNA methylation. PLoS Genet. 6, e1000804.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380.

Dorsett, D. (2009). Cohesin, gene expression and development: lessons from Drosophila. Chromosome Res. 17, 185–200.

Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. Mol. Cell. Biol. 22, 3820–3831.

Fedoriw, A.M., Stein, P., Svoboda, P., Schultz, R.M., and Bartolomei, M.S. (2004). Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science 303, 238–240.

Ferraiuolo, M.A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X.Q., Nadler, M., Blanchette, M., and Dostie, J. (2010). The three-dimensional architecture of Hox cluster silencing. Nucleic Acids Res. 38, 7472–7484.

Filippova, G.N., Cheng, M.K., Moore, J.M., Truong, J.P., Hu, Y.J., Nguyen, D.K., Tsuchiya, K.D., and Disteche, C.M. (2005). Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. Dev. Cell 8, 31–42.

Gan, M., Moebus, S., Eggert, H., and Saumweber, H. (2011). The Chriz-Z4 complex recruits JIL-1 to polytene chromosomes, a requirement for interband-specific phosphorylation of H3S10. J. Biosci. 36, 425–438.

Gause, M., Morcillo, P., and Dorsett, D. (2001). Insulation of enhancer-promoter communication by a gypsy transposon insert in the Drosophila cut gene: cooperation between suppressor of hairy-wing and modifier of mdg4 proteins. Mol. Cell. Biol. 21, 4807–4817.

Gause, M., Schaaf, C.A., and Dorsett, D. (2008). Cohesin and CTCF: cooperating to control chromosome conformation? Bioessays 30, 715–718.

Gerasimova, T.I., Lei, E.P., Bushey, A.M., and Corces, V.G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. Mol. Cell 28, 761–772.

Ghosh, D., Gerasimova, T.I., and Corces, V.G. (2001). Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. EMBO J. 20, 2518–2527.

Gilbert, D.M., Takebayashi, S.I., Ryba, T., Lu, J., Pope, B.D., Wilson, K.A., and Hiratani, I. (2010). Space and time in the nucleus: developmental control of replication timing and chromosome architecture. Cold Spring Harb. Symp. Quant. Biol. 75, 143–153.

Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K.E., Li, W., Myers, R.M., Maniatis, T., and Wu, Q. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. Proc. Natl. Acad. Sci. USA 109, 21081–21086.

Gurudatta, B.V., and Corces, V.G. (2009). Chromatin insulators: lessons from the fly. Brief. Funct. Genomics Proteomics 8, 276–282.

Hadjur, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A.G., and Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature 460, 410–413.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F., et al. (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. Nat. Genet. 43, 630–638.

Heath, H., Ribeiro de Almeida, C., Sleutels, F., Dingjan, G., van de Nobelen, S., Jonkers, I., Ling, K.W., Gribnau, J., Renkawitz, R., Grosveld, F., et al. (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. EMBO J. 27, 2839–2850.

Heger, P., Marin, B., Bartkuhn, M., Schierenberg, E., and Wiehe, T. (2012). The chromatin insulator CTCF and the emergence of metazoan diversity. Proc. Natl. Acad. Sci. USA 109, 17507–17512.

Herold, M., Bartkuhn, M., and Renkawitz, R. (2012). CTCF: insights into insulator function during development. Development 139, 1045–1057.

Hou, C., Dale, R., and Dean, A. (2010). Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proc. Natl. Acad. Sci. USA 107, 3651–3656.

Hou, C., Li, L., Qin, Z.S., and Corces, V.G. (2012). Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol. Cell 48, 471–484.

Jhunjhunwala, S., van Zelm, M.C., Peak, M.M., and Murre, C. (2009). Chromatin architecture and the generation of antigen receptor diversity. Cell 138, 435–448.

Jothi, R., Cuddapah, S., Barski, A., Cui, K., and Zhao, K. (2008). Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. Nucleic Acids Res. 36, 5221–5231.

Kang, H., and Lieberman, P.M. (2011). Mechanism of glycyrrhizic acid inhibition of Kaposi's sarcoma-associated herpesvirus: disruption of CTCF-cohesin-mediated RNA polymerase II pausing and sister chromatid cohesion. J. Virol. 85, 11159–11169.

Kehayova, P., Monahan, K., Chen, W., and Maniatis, T. (2011). Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. Proc. Natl. Acad. Sci. USA 108, 17195–17200.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231–1245.

Kim, Y.J., Cecchini, K.R., and Kim, T.H. (2011). Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. Proc. Natl. Acad. Sci. USA 108, 7391–7396.

Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkov, V., Reik, W., and Ohlsson, R. (2006). 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. USA 103, 10684–10689.

Li, T., Hu, J.F., Qiu, X., Ling, J., Chen, H., Wang, S., Hou, A., Vu, T.H., and Hoffman, A.R. (2008). CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. Mol. Cell. Biol. 28, 6473–6482.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science 293, 2453–2455.

Liu, Z., Scannell, D.R., Eisen, M.B., and Tjian, R. (2011). Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. Cell 146, 720–731.

Lunyak, V.V., and Atallah, M. (2011). Genomic relationship between SINE retrotransposons, Pol III-Pol II transcription, and chromatin organization: the journey from junk to jewel. Biochem. Cell Biol. 89, 495–504.

Majumder, P., Gomez, J.A., Chadwick, B.P., and Boss, J.M. (2008). The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. J. Exp. Med. 205, 785–798.

Mishiro, T., Ishihara, K., Hino, S., Tsutsumi, S., Aburatani, H., Shirahige, K., Kinoshita, Y., and Nakao, M. (2009). Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. EMBO J. 28, 1234–1245.

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell 128, 787–800.

Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heinl, N., Bardenhagen, I., Leers, J., White, R.A., Renkawitz-Pohl, R., Saumweber, H., and Renkawitz, R. (2007). The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. EMBO J. 26, 4203–4214.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F., and Duboule, D. (2011). A regulatory archipelago controls Hox genes transcription in digits. Cell 147, 1132–1145.

Moqtaderi, Z., Wang, J., Raha, D., White, R.J., Snyder, M., Weng, Z., and Struhl, K. (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. Nat. Struct. Mol. Biol. 17, 635–640.

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

Nativio, R., Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.M., and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS Genet. 5, e1000739.

Nègre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., et al. (2010). A comprehensive map of insulator elements for the Drosophila genome. PLoS Genet. 6, e1000814.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381–385.

Ohlsson, R., Lobanenkov, V., and Klenova, E. (2010). Does CTCF mediate between nuclear organization and gene expression? Bioessays 32, 37–50.

Pai, C.Y., Lei, E.P., Ghosh, D., and Corces, V.G. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. Mol. Cell 16, 737–748.

Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanenkov, V., and Ohlsson, R. (2004). 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. 24, 3497–3504.

Paredes, S.H., Melgar, M.F., and Sethupathy, P. (2012). Promoter proximal CTCF binding is associated with an increase in the transcriptional pausing index. Bioinformatics. Published online October 9, 2012. http://dx.doi.org/10.1093/bioinformatics/bts596.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 132, 422–433.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell 137, 1194–1211.

Pikaart, M.J., Recillas-Targa, F., and Felsenfeld, G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev. 12, 2852–2862.

Pirrotta, V., and Li, H.B. (2012). A view of nuclear Polycomb bodies. Curr. Opin. Genet. Dev. 22, 101–109.

Raab, J.R., Chiu, J., Zhu, J., Katzman, S., Kurukuti, S., Wade, P.A., Haussler, D., and Kamakaka, R.T. (2012). Human tRNA genes function as chromatin insulators. EMBO J. 31, 330–350.

Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. Proc. Natl. Acad. Sci. USA 99, 6883–6888.

Ren, L., Wang, Y., Shi, M., Wang, X., Yang, Z., and Zhao, Z. (2012). CTCF mediates the cell-type specific spatial organization of the Kcnq5 locus and the local gene regulation. PLoS ONE 7, e31416.

Ribich, S., Tasic, B., and Maniatis, T. (2006). Identification of long-range regulatory elements in the protocadherin-alpha gene cluster. Proc. Natl. Acad. Sci. USA 103, 19719–19724.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc. Natl. Acad. Sci. USA 105, 8309–8314.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature 489, 109–113.

Schwartz, Y.B., Linder-Basso, D., Kharchenko, P.V., Tolstorukov, M.Y., Kim, M., Li, H.B., Gorchakov, A.A., Minoda, A., Shanower, G., Alekseyenko, A.A., et al. (2012). Nature and function of insulator protein binding sites in the Drosophila genome. Genome Res. 22, 2188–2198.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 148, 458–472.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., and Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature 488, 116–120.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature 479, 74–79.

Soshnikova, N., Montavon, T., Leleu, M., Galjart, N., and Duboule, D. (2010). Functional analysis of CTCF during mammalian limb development. Dev. Cell 19, 819–830.

Spencer, R.J., del Rosario, B.C., Pinter, S.F., Lessing, D., Sadreyev, R.I., and Lee, J.T. (2011). A boundary element between Tsix and Xist binds the chromatin insulator Ctcf and contributes to initiation of X-chromosome inactivation. Genetics 189, 441–454.

Spitz, F., Herkenne, C., Morris, M.A., and Duboule, D. (2005). Inversion-induced disruption of the Hoxd cluster leads to the partition of regulatory landscapes. Nat. Genet. 37, 889–893.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev. 20, 2349–2354.

Stedman, W., Kang, H., Lin, S., Kissil, J.L., Bartolomei, M.S., and Lieberman, P.M. (2008). Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. EMBO J. 27, 654–666.

Tarchini, B., and Duboule, D. (2006). Control of Hoxd genes' collinearity during early limb development. Dev. Cell 10, 93–103.

Van Bortle, K., and Corces, V. (2012). tDNA insulators and the emerging role of TFIIIC in genome organization. Transcription 3, 277–284.

Van Bortle, K., Ramos, E., Takenaka, N., Yang, J., Wahi, J.E., and Corces, V.G. (2012). Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. Genome Res. 22, 2176–2187.

Wada, Y., Ohta, Y., Xu, M., Tsutsumi, S., Minami, T., Inoue, K., Komura, D., Kitakami, J., Oshida, N., Papantonis, A., et al. (2009). A wave of nascent transcription on activated human genes. Proc. Natl. Acad. Sci. USA 106, 18357–18361.

Wan, L.B., Pan, H., Hannenhalli, S., Cheng, Y., Ma, J., Fedoriw, A., Lobanenkov, V., Latham, K.E., Schultz, R.M., and Bartolomei, M.S. (2008). Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. Development 135, 2729–2738.

Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res. 22, 1680–1688.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 451, 796–801.

West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. Genes Dev. 16, 271–288.

West, A.G., Huang, S., Gaszner, M., Litt, M.D., and Felsenfeld, G. (2004). Recruitment of histone modifications by USF proteins at a vertebrate barrier element. Mol. Cell 16, 453–463.

Wood, A.M., Van Bortle, K., Ramos, E., Takenaka, N., Rohrbaugh, M., Jones, B.C., Jones, K.C., and Corces, V.G. (2011). Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. Mol. Cell 44, 29–38.

Xu, Z., Lefevre, G.M., and Felsenfeld, G. (2012). Chromatin structure, epigenetic mechanisms and long-range interactions in the human insulin locus. Diabetes Obes. Metab. 14(Suppl 3), 1–11.

Xu, Z., Wei, G., Chepelev, I., Zhao, K., and Felsenfeld, G. (2011). Mapping of INS promoter interactions reveals its role in long-range regulation of SYT8 transcription. Nat. Struct. Mol. Biol. 18, 372–378.

Yang, J., Ramos, E., and Corces, V.G. (2012). The BEAF-32 insulator coordinates genome organization and function during the evolution of Drosophila species. Genome Res. 22, 2199–2207.

Yoon, Y.S., Jeong, S., Rong, Q., Park, K.Y., Chung, J.H., and Pfeifer, K. (2007). Analysis of the H19ICR insulator. Mol. Cell. Biol. 27, 3499–3510.

Yusufzai, T.M., Tagami, H., Nakatani, Y., and Felsenfeld, G. (2004). CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Mol. Cell 13, 291–298.