Chromatin insulator elements: establishing barriers to set heterochromatin boundaries

Epigenomic profiling has revealed that substantial portions of genomes in higher eukaryotes are organized into extensive domains of transcriptionally repressive chromatin. The boundaries of repressive chromatin domains can be fixed by DNA elements known as barrier insulators, to both shield neighboring gene expression and to maintain the integrity of chromosomal silencing. Here, we examine the current progress in identifying vertebrate barrier elements and their binding factors. We overview the design of the reporter assays used to define enhancer-blocking and barrier insulators. We look at the mechanisms vertebrate barrier proteins, such as USF1 and VEZF1, employ to counteract Polycomb- and heterochromatin-associated repression. We also undertake a critical analysis of whether CTCF could also act as a barrier protein. There is good evidence that barrier elements in vertebrates can form repressive chromatin domain boundaries. Future studies will determine whether barriers are frequently used to define repressive domain boundaries in vertebrates.

KEYWORDS: barrier ▪ boundary ▪ chromatin ▪ CTCF ▪ domain ▪ insulator ▪ USF1 ▪ VEZF1

Gráinne Barkess & Adam G West*

Institute of Cancer Sciences, College of Medical, Veterinary & Life Sciences, McGregor Building, University of Glasgow, Glasgow, G11 6NT, Scotland, UK
**Author for correspondence:*
Tel.: +44 141 211 2974
Fax: +44 141 337 2494
adam.west@glasgow.ac.uk

Heterochromatin domains in higher eukaryotes

The requirement to package large eukaryotic chromosomes into a cell's nucleus whilst allowing access to DNA for processes such as transcription is thought to involve several layers of organization. Electron microscopy studies in the 1970s and 1980s led to the view that a higher order chromosome structure consisting of topologically independent looped domains exists [1–5]. This chromosomal domain model gained support from the first studies of chromatin accessibility at the model chicken *β-globin* gene locus [6]. Since that time, it has become clear that the regulation of a gene's transcription in higher eukaryotes is influenced by epigenetic modification of the chromatin template. The covalent modification of DNA or the histone proteins that package DNA can dominantly affect the recruitment and assembly of the transcription machinery and its regulatory components. In recent years, the sequencing of genomes and the development of powerful genomic technologies have opened up a new era of studying these epigenetic marks on a genome-wide level [7]. Epigenomic approaches draw upon the insight gained at model gene loci to analyze events on a genome-wide level, thus allowing researchers to understand the context in which gene regulation occurs.

Early epigenomic profiling studies have not only begun to reveal the complexity of combinatorial chromatin signatures that correlate with specific gene regulatory processes, but are also revealing a substantial degree of higher order chromosomal organization. The methylation of histone H3 at either lysine 9 or 27 is linked to HP1-associated heterochromatin and Polycomb-associated repressive chromatin, respectively [8]. Consistent with earlier studies on specific chromosomal regions [9,10], whole-genome epigenomic profiling has revealed that extensive domains of continuous H3K9me2, H3K9me3 or H3K27me3 enrichment are prevalent in the genomes of flies [11–13], plants [14–16], worms [17,18] and mammals [19–21]. Repressive chromatin domains spanning many hundreds of kilobases are frequently observed (Table 1). These domains relate to the gross 3D organization of the nucleus as there is considerable overlap between these domains and the chromosomal regions that associate with the nuclear lamina associated domains (LADs) [21,22].

Repressive chromatin domains across mammalian genomes undergo substantial redistribution upon cellular differentiation. Domains of H3K9me2 were reported to increase considerably in size and number in differentiated murine cells compared with pluripotent ES cells [21]. An independent study using a different bioinformatic tool to define chromatin domains from the same data indicated that the extent of H3K9me2 domains in murine ES cells may have been underestimated [23]. Nevertheless, support for

Future Medicine part of fsg

Table 1. Heterochromatin domains are prevalent in higher eukaryotes.

Species	Genome composition	H3K9me2	H3K9me3	H3K27me3	LADs
Arabidopsis thaliana	157 Mb ~27,400 genes 5 pericentric chromosomes	Pericentric domains of up to 350 kb. Smaller domains (<5 kb) are found at genes and selected repeats [14]	Small domains (<5 kb) found at genes and selected repeats [15]	Small domains (<5 kb) at genes [15,16]	TBD
Caenorhabditis elegans	100 Mb ~20,000 genes 6 holocentric chromosomes	Up to 4-Mb domains on chromosome arms [17,18]	Up to 4-Mb domains on chromosome arms [17,18]	Up to >1-Mb domains on chromosome arms [17,18] Small domains at repressed centric gene loci	Up to 4-Mb domains on chromosome arms [17,18]
Drosophila melanogaster	165 Mb ~13,600 genes 4 pericentric chromosomes	Continuous pericentric domains up to 3.3 Mb [11–13] Cell type-specific gene domains up to 300 kb (m 14–78 kb)	Continuous pericentric domains up to 3.3 Mb [11–13] Cell type-specific gene domains up to 300 kb (m 14–78 kb)	Continuous domains typically <50 kb encompass repressed genes [12]	7–700-kb LADs (m 90 kb) cover 40% of genome [117]
Mus musculus	~3.4 Gb 36,817 genes 20 chromosome pairs	Continuous LOCK domains of 20 kb up to ~4 Mb Increase in number and size upon differentiation (m 43–235 kb) [21] Span ~4–46% of genome	TBD [20]	TBD [20]	Mostly continuous LADs of 40 kb–16 Mb (m ~400 kb) Span ~40% of genome [118]
Homo sapiens	~3.4 Gb 27,478 genes 23 chromosome pairs	Continuous LOCK domains of up to ~2.9 Mb, spanning up to ~20% of genome (m 54 kb) [21]	Up to 1.5-Mb domains that increase in number and size upon differentiation (m 7–12 kb) Span ~4–16% of genome [19]	Up to 370-kb domains that increase in number and size upon differentiation (m 9–16 kb) Span ~4–12% of genome [19]	Mostly continuous LADs of 100 kb–10 Mb (m 553 kb) [22] Large degree of overlap with H3K9me2 LOCK domains [21]

A summary of continuous domains of repressive chromatin marks or nuclear LADs observed from whole-genome studies in higher eukaryotes.
LAD: Lamina associated domain; LOCK: Large organized chromatin K9 modification; m: Median size of the domain; TBD: To be defined.

the model of heterochromatin-associated spreading during cellular differentiation was provided by a recent study of human cells [19]. Multiple analytical tools were used to show that both H3K9me3 and H3K27me3-enriched domains considerably increase in size and number in primary fibroblasts compared with ES cells. Perhaps unsurprisingly, the genes enveloped by heterochromatin-associated spreading in the differentiated murine and human cell types become silenced and include pluripotency-related genes. These studies indicate that the establishment of heterochromatin domains marked by H3K9me3 and H3K27me3 is critical for cell fate determination. This view is supported by the fact that H3K9 and H3K27 histone methyltransferases are essential for cellular differentiation and embryonic development in mammals [24–29].

Understanding how these domains are established and maintained is a key step in understanding the epigenetic basis of cellular identity. It will be important to understand how the boundaries of heterochromatin domains are established. Strong boundaries may be required to both maintain the integrity of heterochromatin domains that facilitate heritable gene silencing and to protect neighboring gene loci from pervasive heterochromatin spreading. In this article, we examine the current progress in identifying and understanding the mechanisms of elements that form chromatin boundaries in vertebrates. A great deal is also being learned from a number of elegant studies in *Drosophila melanogaster*, but this is beyond the scope of this article [30–33].

Defining chromatin insulator elements

The boundaries of heterochromatin domains can be formed in distinct ways. The simplest is where the balance between heterochromatin-promoting and chromatin opening processes at neighboring loci determine the boundary. Such boundaries vary in position and would appear as a broad transition in chromatin states when studying a population of cells. The lack of a fixed boundary is the basis of position-effect variegation (PEV), where the stochastic spread of heterochromatin formation results in the heritable silencing of a neighboring gene. Alternatively, the boundaries of heterochromatin domains can be fixed by the action of *cis*-regulatory elements. In some cases, these could be promoters or enhancers with specific characteristics that allow them to resist heterochromatin formation. However, it is anticipated that in most cases, heterochromatin boundaries would be formed by autonomous elements known as insulators.

Insulators can be defined as DNA elements that have a common ability to protect a gene from undesirable regulatory influences present in their chromosomal environment. Insulators are functionally defined in reporter gene assays that test for their ability to act as an enhancer blocker or a barrier to chromosomal silencing. In principle, these assays report two very different activities, but some insulators can function in both assays. This terminology can lead to confusion, so it is always important to determine the functions ascribed to any given insulator (Box 1). Enhancer-blocking assays score whether a putative insulator interferes with an enhancer's action on a reporter gene specifically when positioned between the two (Figure 1A). It is crucial that control experiments, where putative insulators are positioned outside of the linked enhancer and gene promoter, are performed to rule out any position-independent silencer activities [34,35].

While enhancer-blocking assays are relatively easy to perform, they may not accurately report the activity of insulators in their natural genomic context. A principal criticism of these assays is that putative insulators are rarely scored against the actual promoter/enhancer combinations found at their endogenous genomic locations, as these are often unknown. It was established

Box 1. Glossary of terms.

Chromatin domain
- A chromosomal region containing multiple genes or gene regulatory elements with a shared chromatin modification or structural state

Heterochromatin
- Transcriptionally repressive chromatin enriched in H3K9me2 or H3K9me3 that may be bound by HP1 and may have condensed conformation

Polycomb domain
- Transcriptionally repressive chromatin enriched in H3K27me3 that may be bound by Polycomb group proteins

Lamina associated domains
- Chromosomal domains that associate with the nuclear lamina, which tend to incorporate heterochromatin or Polycomb domains

DNaseI hypersensitive site
- A short region of genomic DNA (hundereds of base pairs) that is especially accessible, or structurally favorable, to nuclease digestion. DNaseI hypersensitive sites typically mark gene regulatory elements

Enhancer
- A gene regulatory element that upregulates the transcription of a nearby gene regardless of orientation relative to the gene. Can operate over tens to hundreds of kilobases

Locus control region
- A chromosomal region capable of providing position-independent, copy number-dependent expression of a transgene. Locus control regions are typically composed of multiple enhancers and can contain insulators

Silencer
- A regulatory element that downregulates the expression of a neighboring gene, often by establishing heterochromatin

Chromatin boundary
- A transition between two distinct chromatin states defined by histone modifications, DNA methylation or chromatin accessibility. Typically refers to a transition between a heterochromatin domain and an active gene locus. Boundary elements or factors are DNA sequences or proteins that locate at chromatin boundaries, which may have the properties of insulators

Enhancer-blocking insulator
- A regulatory element that is demonstrated in functional assays to interfere with the communication between a linked enhancer and promoter when positioned between them. The enhancer and promoter are not prevented from forming other regulatory interactions

Barrier insulator
- A regulatory element that is demonstrated in functional assays to block the spread of chromosomal silencing into a neighboring gene locus. Barriers can define heterochromatin boundaries

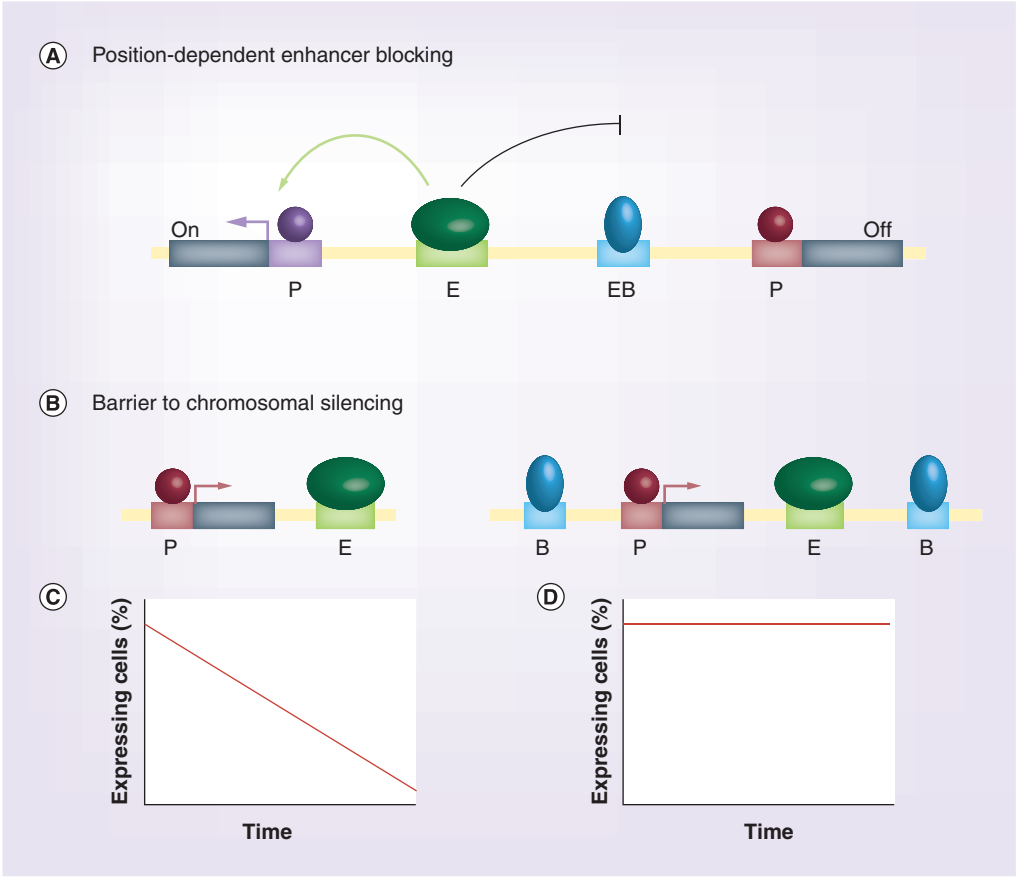

Figure 1. Reporter transgene assays used to define two properties of insulators.
(A) Enhancer-blocking assays score for the ability of a putative EB to interfere with the action of an E on a linked P (red) specifically when placed between the two. Unlike silencers, an enhancer-blocker would not interfere with enhanced transcription when positioned outside of the E, for example, allowing enhancement of the lilac gene shown. **(B)** Barrier assays score for the ability of a B to protect a reporter transgene from chromosomal position effect silencing. **(C)** Randomly integrated transgenes that lack insulation tend to silence over time in culture. **(D)** Transgenes shielded by barrier insulators retain their expression, regardless of chromosomal position. Barrier insulators must be placed on both flanks of the transgene as chromosomal silencing can encroach from either direction. B: Putative barrier insulator; E: Enhancer; EB: Enhancer-blocking insulator; P: Promoter.

some time ago that the strength and characteristics of the enhancer/promoter combination can influence the assay's outcome. The use of a weak transgenic enhancer can allow the scoring of weak insulator activity, but there may not be such activity at the endogenous chromosomal locus [36]. Furthermore, the compact spacing of reporter constructs compared with endogenous chromosomal loci makes it difficult to accurately score the functions of putative insulator elements, especially those that function via chromosomal loop formation (discussed below). While enhancer-blocking assays are a useful tool when identifying novel elements, care must be taken when translating the results of these assays to understanding the roles of elements in their natural chromosomal contexts.

Barrier assays score whether insulators can protect a reporter gene from chromosomal position-effect silencing. The first position effect assay was developed to test the ability of *Drosophila* insulators to protect a randomly integrated reporter gene from chromosomal silencing [37,38]. Similarly, barrier assays in vertebrates typically involve the random integration of a reporter construct in transfected cells. The lack of transgene targeting does mean that the nature of chromosomal silencing will vary depending on whether the transgene has integrated into a site of Polycomb silencing, constitutive or facultative heterochromatin [8]. Studies of *β-globin* transgenes in chicken cells have shown that the silencing of transgenes typically involves histone deacetylation followed by heterochromatin-associated H3K9 methylation and DNA methylation [39,40]. Putative barrier elements are placed at both flanks of the reporter gene as heterochromatin-associated

silencing can encroach from either direction (FIGURE 1B) [41]. Ideally, studies will include the analysis of several integration events for each construct to control for the variable silencing potential of different integration sites [41,42]. It is also good practice to test whether newly identified barrier elements harbor enhancer activity in separate reporter assays. Barrier assays are time consuming given the scale of the experiments and the fact that chromosomal position effect silencing often accumulates over a period of several weeks, depending on the strength of the transgene and the site of integration. It should be noted that the relative strength of the reporter gene/enhancer and the CpG content of the promoter will influence the sensitivity of barrier assays, which may be a potential source of discrepancy between different studies. Another factor that must be considered when designing barrier assay constructs is the spacing of the various elements [43].

Alternative assays have been devised to test whether putative insulators can act as barriers to a known source of chromosomal silencing. One episomal assay involves the positioning of a putative barrier between a transgene and lexA operator sequences, which are bound by known repressor proteins that are fused to a lexA DNA-binding domain [44]. This assay offers advantages in terms of speed and a knowledge of the form of repression, but orientation effects were observed for the highly characterized HS4 insulator element that have not been reported in standard assays [44–47], suggesting that episomes may not be a suitable template for studying barrier elements. An alternative assay involves the tethering of a reporter transgene to the transcriptionally repressive nuclear lamina compartment via the binding of lexA-emerin fusion proteins to lexA operators [48]. It was recently reported that the inclusion of a putative heterochromatin barrier element can reverse the silencing that results from tethering [49]. However, this study did not determine whether this was as a result of heterochromatin barrier activity or whether the element simply directed relocalization away from the nuclear periphery. Further studies are required to determine whether this will be a generally useful heterochromatin barrier assay.

Lessons from chicken chromatin boundaries

The chicken *β-globin* gene cluster has served as a model for studying coordinated gene regulation and chromatin domain formation for many years [6,50–57]. The upstream *FOLR1* gene is separated from the *β-globin* locus by a 16-kb condensed chromatin region that bears all the hallmarks of heterochromatin [51,54,57,58] (FIGURE 2A). The 5′ boundary of DNaseI sensitivity [52] and histone modifications [50,54,57] across the *β-globin* domain is marked by a constitutive DNaseI hypersensitive site called HS4. HS4 is a very well characterized insulator element and has served as a very useful paradigm for

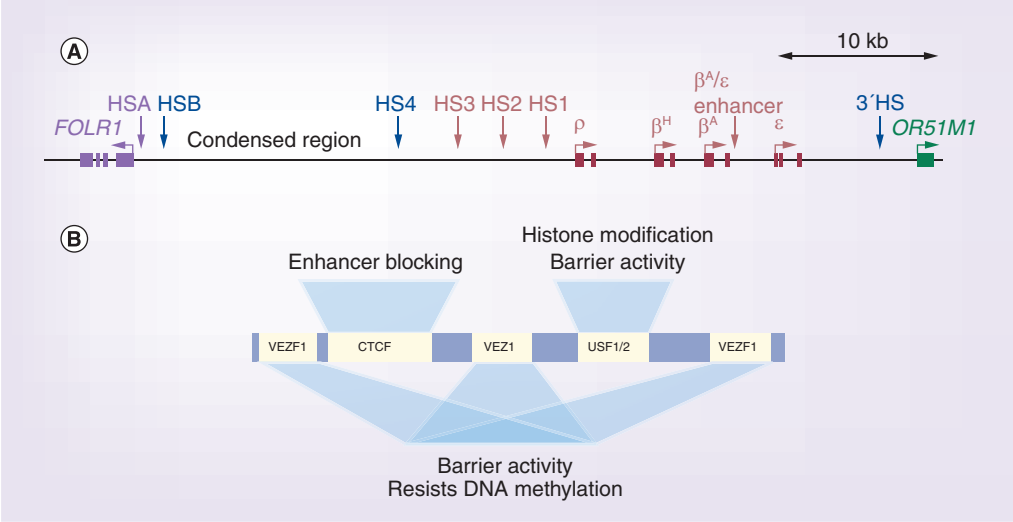

Figure 2. Heterochromatin boundaries partition the chicken *FOLR1* and *β-globin* gene loci.
(A) Scale map of the *FOLR1* and *β-globin* gene loci. The exons of the *FOLR1*, *β-globin* (*ρ*, β^H, β^A and $^\varepsilon$) and *OR51M1* genes are depicted by purple, red and green boxes, respectively. Vertical arrows show early/late erythroid (purple/red) or constitutive (blue) DNaseI hypersensitive regulatory elements. **(B)** Expanded view of the 275-bp 'core' of the HS4 insulator element showing the proteins and activities associated with the five DNaseI footprint elements.

the study of insulators in vertebrates. A 275-bp HS4 element functions in both enhancer blocking and barrier assays [34,41,42,59]. Functional dissection of the HS4 element found that its enhancer blocking and barrier activities involve different proteins and mechanisms and are separable in assay systems. It was found that a single binding site for the zinc finger protein CTCF was both necessary and sufficient for HS4's enhancer-blocking activity [34]. CTCF remains the only insulator-binding protein known to direct enhancer blocking in vertebrates. Despite CTCF's pivotal role in enhancer blocking, removal of the CTCF binding site has no effect on the barrier activity of HS4 [42,60]. The enhancer blocking and barrier to silencing properties are separable functions, and therefore employ different mechanisms.

The barrier activity of HS4 requires a binding site for the bHLH transcription factor proteins USF1/USF2 and three binding sites for the zinc finger protein VEZF1 (**Figure 2B**) [42,61,62]. It is now well-established that HS4 manipulates histone modification signatures to counteract gene silencing [54,62–64]. HS4 is persistently enriched in high levels of H3 and H4 acetylation, H3-lysine 4 methylation, H4-arginine 3 methylation, acetylated histone variant H2A.Z and ubiquitinated H2B, regardless of neighboring gene expression [50,53,54,57,64]. The binding site for USF proteins has been found to be required for the majority of these modifications [54,62]. Intriguingly, while this chromatin state is reminiscent of highly active gene promoters, HS4 lacks promoter activity [65] and is not bound by RNA polymerase [54]. Rather, experimental evidence supports a model where the active histone modifications at HS4 collectively act as a chain terminator to heterochromatin assembly by interfering with the propagation of repressive histone modifications [54,62–64]. The role of H2B ubiquitination at chromatin boundaries is particularly striking. Depletion of RNF20, an E3 ligase required for this modification, results in a loss of H3K4 methylation and multiple acetylation of H3, H4 and H2A.Z at HS4 [54]. The same RNF20-dependent active chromatin signature is found at the HSA/HSB elements that mark the boundary between the *FOLR1* gene and the opposite end of the heterochromatin domain bounded by HS4 (**Figure 2A**). The HSA/HSB elements are bound by USF1 and VEZF1, but lack CTCF. Loss of active modifications from the HSA/HSB and HS4 boundary elements is followed by pervasive spreading of H3K9 and H4K20 methylation from the heterochromatin domain into the *FOLR1* and *β-globin* gene loci, resulting in gene silencing [54].

Although active histone modifications play an essential role in HS4's barrier activity, they are not sufficient. Deletion of VEZF1 binding sites at the HS4 insulator results in transgene silencing despite the recruitment of active histone modifications via USF binding to the mutant insulators that flank the transgene [42,61,62]. The loss of VEZF1 binding sites was marked by *de novo* DNA methylation of the mutant HS4 insulator and the silenced transgene promoter [61]. In striking contrast to the results with VEZF1 mutants, the transgene promoter remains hypomethylated when the USF site is deleted from HS4 insulators, despite the loss of active histone modifications and transgene silencing [27,62,65]. These results, albeit in a heterologous system, demonstrate that the apparent protection from DNA methylation mediated by VEZF1 elements is separable from USF-mediated histone modification. In further support for a role in controlling DNA methylation states, it was also found that VEZF1 binding sites can mediate both the protection and removal of DNA methylation from the *Aprt* promoter CpG island in murine ES cells [27]. Paradoxically, loss of *Vezf1* from mouse ES cells results in substantial reductions in methylation at specific elements, rather than the predicted gain, but this appears to be due to an essential role for Vezf1 in maintaining expression of the *Dnmt3b* DNA methyltransferase gene [66]. Further studies are required to determine the mechanisms VEZF1 employs to regulate DNA methylation states.

The dissection of the HS4 insulator's functions indicates that heterochromatin barrier elements in vertebrates may need to prevent DNA methylation in addition to resisting the propagation of repressive histone modifications, as either process is sufficient to direct the establishment of an epigenetically stable silent chromatin state [67]. The study of barrier elements across vertebrate genomes is at an early stage. It remains to be determined whether USF1/2 and VEZF1 are widely employed to create heterochromatin boundaries across vertebrate genomes. It should be noted that these proteins also have general roles at gene promoters [61,68]. It will be interesting to determine whether other promoter-associated *trans* factors are co-opted at other autonomous barrier elements. It is likely that any collection of factors capable of establishing a dominant active

histone modification state in addition to countering DNA methylation can form a barrier to heterochromatin. In this view, there is no absolute requirement for specialized barrier elements to form all heterochromatin boundaries. Specific classes of gene promoter (e.g., housekeeping) and some dominant enhancer regions such as locus control regions may act as *de facto* barriers if they can effectively resist the propagation of heterochromatin.

Do chromosomal loops formed by CTCF mediate chromatin domain boundaries?

CTCF is a transcription factor that has been linked to a variety of gene regulatory functions, including the enhancer-blocking activity of insulator elements, which is associated with chromosomal looping interactions [69]. A number of recent studies have reported significant enrichments of CTCF sites at the boundaries of repressive chromatin domains. Taken together, these observations have led to the speculation that CTCF may be a primary determinant of chromosomal domain formation through the formation of barrier insulators. In this section, we discuss these reports and compare them with functional analyses which show that CTCF sites are not sufficient for barrier activity.

CTCF is a ubiquitously expressed protein containing a large DNA-binding domain comprising 11 zinc fingers. CTCF is required for cell survival and is essential for early embryonic development in the mouse [70–73]. A number of groups have mapped the genomic binding of CTCF in several vertebrate cell types, with up to 50,000 sites reported [74–84]. Approximately half of all CTCF binding events are shared between different cell types [77]. A systematic analysis of conserved noncoding elements (CNEs) in mammals found nearly 15,000 CNEs that contain CTCF binding motifs [76]. These CTCF-associated CNEs are mostly located between genes of differential expression programs, suggesting a general role for CTCF in partitioning gene expression domains [76,84]. CTCF binding sites from many genomic elements have been found to mediate enhancer-blocking activity in transgene reporter assays, a partial list of which can be found at the CTCF binding site database website [85,201]. Several CTCF-binding factors have been identified, including p68 and the cohesin STAG2 (SA2) [86–88]. Several studies have found a substantial overlap between the genomic sites of CTCF and cohesin binding. Depending on the cell type, approximately 55–80% of CTCF sites are bound by cohesin and 65–90% of cohesin sites are bound by CTCF [52,89,90]. These findings correlate well with the finding that elements bound by CTCF mediate chromosomal looping interactions over many kilobases and that cohesins are required for these loops [69,86,91–95].

The first whole-genome analysis to show an enrichment of CTCF binding at chromatin boundaries resulted from the mapping of nuclear LADs in human cells [22]. It was found that 13% of LAD boundaries had a CTCF site located within the 10 kb region immediately outside of the LAD. However, CTCF binding in itself is not sufficient to establish a LAD boundary, as 19% of CTCF sites are found within LADs. Only 2.5% of all CTCF sites locate to LAD boundaries, whereas 78% locate away from LADs. The same study found a substantial enrichment of gene activity at LAD boundaries. 21.3% of LAD boundaries are marked by CpG islands or gene promoters, which include half of the boundary-associated CTCF sites [22]. LAD boundaries are also highly enriched for many transcription factor motifs, RNA polymerase II and transcribed genes. It is clear that a variety of factors can direct chromatin opening and nuclear localization away from the periphery. While the association with CTCF binding at LAD boundaries is significant, further experiments are required to determine whether the position of these boundaries is CTCF-dependent and what relationship exists with chromosomal loop interactions and insulator activities. It has been noted that one of the LAD boundaries lies close to a CTCF element, MIME, which marks a chromatin boundary upstream of the human *MYC* gene [69,96]. However, it was recently shown that this CTCF element can be deleted from human chromosomes without affecting this chromatin boundary [97].

An independent whole-genome study provided evidence in support of a role for CTCF elements in forming the boundaries of heterochromatin domains enriched in H3K27me3. It was found that 2–4% of H3K27me3 domain boundaries are marked by a CTCF site [77]. These boundary-associated sites represent only 4–6% of the total CTCF sites, but this finding suggests that some CTCF elements may act as barriers to the potential spreading of repressive chromatin. It was noted that approximately 97% of the CTCF-associated boundaries are cell-type specific. Whether these boundaries

overlap with cell-type-specific gene promoters and whether CTCF contributes any barrier activity at these boundaries remains to be determined. A recent study found that a putative barrier to H3K27me3 spreading located between the human *HOXA7* and *HOXA9* genes could reverse the silencing of a transgene that results from tethering to the nuclear lamina [49]. However, this study did not determine whether this activity is CTCF-dependent, whether the *HOXA* boundary element functioned as a position-dependent barrier to limit the spread of heterochromatin, or if the element simply directed relocalizations away from the nuclear periphery. Another recent study has reported a striking switch in chromatin states at the murine *Wnt4* locus following the loss of WT1 transcription factor function [98]. These altered states are restricted to a region flanked by CTCF/cohesin sites, which has lead to a model where CTCF/cohesin may form the boundaries of a *Wnt4* chromatin domain. Depletion of CTCF/cohesin expression leads to changes in chromatin states outside of the proposed *Wnt4* domain, which coincide with changes in expression of the flanking *Zbt40* and *Cdc42* genes. However, it remains unclear whether *Wnt4* resides within a chromatin domain that is bordered by the CTCF sites, or whether the changes following CTCF/cohesin depletion are as a result of spreading of one domain into another. Nevertheless, these findings lend weight to the model that CTCF/cohesin can define heterochromatin boundaries. Mutation of the CTCF sites at the *Wnt4* locus in combination with barrier assays will be useful when addressing this model.

Epigenomic profiling and heterologous reporter assays are powerful tools to screen for CTCF elements and characterize their potential regulatory activity. However, transgene assays cannot recapitulate the **endogenous chromosomal** environments that these elements have evolved to operate in. A small number of studies have now investigated the role of CTCF elements *in situ* or have employed a conditional *Ctcf* knockout. The outcome of these studies indicates that the current models of CTCF function require reconsideration. It is well established that CTCF sites at the *H19* imprinting control region (ICR) function to regulate *Igf2* expression in a manner that is entirely consistent with an enhancer-blocking mechanism *in vivo*, although it should be noted that CTCF has other regulatory contributions at this locus [99,100]. The ICR element forms CTCF- and cohesin-dependent chromosomal loops, although the exact purpose of the loops remains to be elucidated [99]. Chromosomal loops are also formed by multiple CTCF sites that flank the murine *β-globin* locus, suggestive of a chromosomal domain organization that both facilitates and partitions *β-globin* expression [101]. However, deletion of *Ctcf* or the CTCF binding sites, which disrupt loop formation, have little or no effect on gene expression or heterochromatin encroachment at this locus [70,102–104].

The homeotic gene clusters are arguably the best models to study CTCF-mediated domain organization in mammals, as there must be clear partitioning between neighboring genes with distinct programs of expression that drive correct body patterning [71]. The *HOX* gene clusters are regulated by long-range enhancers and CTCF sites are found to locate between and around these genes, which suggests that CTCF-mediated enhancer blocking may play a crucial role in *HOX* gene expression and body patterning. A recent study used a conditional *Ctcf* knockout in developing mouse forelimbs and found little evidence for an enhancer blocking or domain boundary role for murine CTCF [73]. *Ctcf* disruption was found to affect the expression of multiple apoptotic and mitochondrial genes, resulting in enhanced apoptosis and dramatic truncation of the forelimbs. CTCF was found to bind at, or close to the promoters of these genes. While *Ctcf* was found to be required for cell survival, there was no evidence for ectopic expression of the *Hoxa*, *Hoxb*, *Hoxc* or *Hoxd* genes. Furthermore, there were no significant changes in the expression of cell-type specific genes that are separated by CTCF sites from neighboring highly expressed genes on the two chromosomes studied. Previous studies have led to the proposal that CTCF can form heterochromatin boundaries at gene loci that escape X-inactivation [72,105]. However, there is no significant change in expression for the genes that escape X-inactivation in female *Ctcf* mutant limbs [73]. This study did not determine whether any heterochromatin redistribution occurs following *Ctcf* mutation.

These studies have revealed that ectopic enhancer-blocking transgene assays can report a different function for CTCF binding elements as compared with *Ctcf* gene disruption or endogenous CTCF site deletion in an organismal context. While it is becoming increasingly clear that CTCF can direct the formation of chromosomal loops in different species and that practically all CTCF elements will block enhancers in reporter

assays, great care must be taken when assuming the functions of novel CTCF elements revealed by epigenomic profiling. It is likely that CTCF elements will have a wide array of functions, depending largely on their genomic context and the role of neighboring elements. A role for CTCF elements in heterochromatin boundary formation is frequently posited, but conclusive evidence in support of this particular function is not available at present. The CTCF binding site of the chicken HS4 insulator is not required for its barrier activity [42,60]. Furthermore, the well characterized CTCF-mediated human *H19* ICR and *TCRA/D* BEAD-1 enhancer-blocking elements also lack barrier activity [42]. These frequently overlooked results indicate that CTCF elements alone are not sufficient to form barriers to chromosomal silencing.

Barrier-like activity at gene regulatory elements

A chromatin domain boundary exists between the murine T-cell receptor α and δ gene cluster and the neighboring ubiquitous apoptosis gene *Dad1* [106]. This boundary coincides with a DHS cluster that forms the *Tcra* locus control region (LCR). Dissection of the LCR has revealed that HS1 and HS1′ encompass the *TCR*α enhancer (Eα) and a CTCF-dependent enhancer-blocking element (TAD1), respectively [107]. However, these elements alone are not sufficient to suppress position effect silencing in transgenic mice. Elements located within HS4 and HS6 are required to act as barrier-like elements, while the CTCF element HS1′ is largely dispensable [108–110]. The factors and mechanisms employed at these barrier elements remain to be determined.

Previous studies have found that a number of the heterochromatin boundaries in yeast are associated with gene promoters [111]. Epigenomic profiling is also beginning to reveal that a considerable fraction of heterochromatin boundaries in mammals are marked by active promoters [22]. This raises the possibility that not all gene promoters are sensitive to heterochromatin silencing from their chromosomal neighborhood. Some promoters may harbour barrier-like elements and thus determine their own chromosomal environment to ensure their activity. Recent studies illustrate how cell-type specific genes employ barrier-like activities to drive expression from a heterochromatin environment.

Human α-spectrin is encoded by the large (~120 kb) 52 exon *SPTA1* gene, which is surrounded by olfactory receptor family genes and associates with the nuclear lamina in nonexpressing cells [22]. Functional dissection of the human α-spectrin *SPTA1* gene promoter region has identified a proximal element within the noncoding exon 1′ that harbors barrier activity in erythroid cells [112]. This erythroid-specific element is bound by the erythroid regulator GATA-1 and by USF1/2, and lacks CTCF binding or enhancer-blocking activity [112,113]. Cooperation between GATA-1 and USF1/2 in erythroid cells may form a dominant chromatin opening element that allows localisation away from the nuclear periphery to facilitate *SPTA1* expression.

The human *ANK1* gene has also been shown to harbor a barrier-like activity [114]. *ANK1* is primarily expressed in erythrocytes and the gene is silenced within an H3K27me3 domain in nonexpressing cells [110]. It was recently demonstrated that a 181 bp 5′HS immediately upstream of the core *ANK1* promoter has barrier activity [114]. The *ANK1* 5′HS barrier resembles the chicken HS4 insulator in that it is bound by USF1/2, recruits several histone modifying enzymes and is enriched in H3K4me2 and acetylated H3 and H4. Mutations of *ANK1* have been identified in approximately half of all patients with hereditary spherocytosis, a group of disorders that cause anemia, jaundice and sphere-shaped erythrocytes. Two of the mutations, -108 and -153, locate to the 5′HS barrier element and result in reduced *ANK1* expression. It was found that these mutations reduced the binding of USF2 and BRG1 to 5′HS, resulting in diminished recruitment of active histone modifications. Consequently, the -108/-153 mutant *ANK1* 5′HS lacks barrier activity in transgenic mice. It will be interesting to determine whether the -108/-153 mutations cause the failure of the 5′HS barrier and heterochromatin silencing of *ANK1* in patients. This may represent the first example where disruption of a barrier element is a pathogenetic mechanism of human disease.

Conclusion & future perspective

The outcome of recent advances in epigenomics and the careful analysis of model gene loci have, by and large, provided support for the chromatin domain model of chromosome architecture for a sizable fraction of the vertebrate genome. However, our understanding of how insulator elements contribute to this architecture is still at an early stage. The recent studies of *Ctcf* gene mutation or endogenous CTCF site deletion indicate that transgene assays probably do

not accurately report the physiological activity of this paradigm insulator protein in many instances. It is clear that as we test models of insulator function, we must address the physiological roles of these elements at their endogenous chromosomal locations. The recent development of designer zinc finger nucleases and TALE nucleases, along with their commercial availability and increasing affordability, provides an opportunity for most groups to undertake efficient gene targeting in many cell types or species [115,116].

There are many questions about heterochromatin boundaries that remain unanswered. The wealth of epigenomic data available makes it possible to determine what proportion of heterochromatin domains have fixed boundaries and whether particular active chromatin marks or proteins such as VEZF1 or USF1 are frequently associated with heterochromatin boundaries. We need to look closely at the genes that locate to heterochromatin boundaries and determine whether, like *SPTA1* and *ANK1*, they employ promoter-proximal barrier-like elements to ensure their expression.

There is also a need to understand how the chromatin domains revealed by whole-genome analyses relate to the control of DNA replication timing, chromosome structure and three-dimensional organization within the nucleus. A recurring theme of chromosomal domain models is that the formation of chromatin loops by insulators and chromatin boundaries will favor the interactions of regulatory element within loops rather than between loops. It will be interesting to determine whether this model is supported by emerging data from high throughput chromosomal interaction studies [87].

Financial & competing interests disclosure

A West is supported by the Biotechnology and Biological Sciences Research Council, the UK Medical Research Council, the Engineering and Physical Sciences Research Council and the European Commission. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Heterochromatin domains in higher eukaryotes

- Modern epigenomic profiling studies reveal that extensive heterochromatin domains are prevalent in higher eukaryote genomes.
- Heterochromatin domains increase in size upon cellular differentiation.

Defining chromatin insulators

- A look at the terminology used to define heterochromatin barrier and enhancer-blocking insulator elements.
- A review of the different assays used to functionally define insulator elements.

Lessons from chicken chromatin boundaries

- The chicken β-globin HS4 element is the best characterized insulator to date.
- CTCF mediate enhancer blocking, while USF1/2 and VEZF1 mediate heterochromatin barrier activity.
- Heterochromatin barrier activity involves active histone modification and the prevention of DNA methylation.
- Some gene promoters may act as de facto heterochromatin barriers.

Do chromosomal loops formed by CTCF mediate chromatin domain boundaries?

- A fraction of CTCF sites are found at heterochromatin domain boundaries, suggestive of a role in barrier activity.
- CTCF sites are not sufficient for barrier activity.
- CTCF sites can be deleted without affecting heterochromatin boundaries.
- Results from reporter gene assays do not always correlate with *in vivo* experiments.

Barrier-like activity at gene regulatory elements

- Heterochromatin barrier activities are found at locus control region and gene promoter elements.
- SNPs that disrupt a heterochromatin barrier associate with hereditary spherocytosis.

Future perspective: chromatin domain models become reality

- Transgene assays may not always accurately report the physiological activity of insulator elements as they do not reproduce chromosomal context.
- Technological advancements make gene editing more accessible, allowing the study of the physiological roles of heterochromatin boundary and other insulator elements.
- What fraction of heterochromatin boundaries are formed by specialised barrier elements or gene promoters?
- Do chromosomal loops correlate with the definition of domains of heterochromatin and replication timing?

References

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1 Adolph KW. Organization of chromosomes in mitotic HeLa cells. *Exp. Cell Res.* 125(1), 95–103 (1980).

2 Belmont AS, Braunfeld MB, Sedat JW, Agard DA. Large-scale chromatin structural domains within mitotic and interphase chromosomes *in vivo* and *in vitro*. *Chromosoma* 98(2), 129–143 (1989).

3 Benyajati C, Worcel A. Isolation, characterization, and structure of the folded interphase genome of *Drosophila melanogaster*. *Cell* 9(3), 393–407 (1976).

4 Marsden MP, Laemmli UK. Metaphase chromosome structure. evidence for a radial loop model. *Cell* 17(4), 849–858 (1979).

5 Rattner JB, Lin CC. Radial loops and helical coils coexist in metaphase chromosomes. *Cell* 42(1), 291–296 (1985).

6 Stalder J, Larsen A, Engel JD, Dolan M, Groudine M, Weintraub H. Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNAase I. *Cell* 20(2), 451–460 (1980).

7 Hawkins RD, Hon GC, Ren B. Next-generation genomics. an integrative approach. *Nat. Rev. Genet.* 11(7), 476–486 (2010).

8 Beisel C, Paro R. Silencing chromatin. comparing modes and mechanisms. *Nat. Rev. Genet.* 12(2), 123–135 (2011).

9 Pauler FM, Sloane MA, Huang R *et al.* H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res.* 19(2), 221–233 (2009).

10 Thurman RE, Day N, Noble WS, Stamatoyannopoulos JA. Identification of higher-order functional domains in the human ENCODE regions. *Genome Res.* 17(6), 917–927 (2007).

11 Filion Gj, van Bemmel JG, Braunschweig U *et al.* Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143(2), 212–224 (2010).

12 Kharchenko PV, Alekseyenko AA, Schwartz YB *et al.* Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471(7339), 480–485 (2011).

13 Riddle NC, Minoda A, Kharchenko PV *et al.* Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Res.* 21(2), 147–163 (2011).

14 Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS ONE* 3(9), e3156 (2008).

15 Turck F, Roudier F, Farrona S *et al.* *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.* 3(6), e86 (2007).

16 Zhang X, Clarenz O, Cokus S *et al.* Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* 5(5), e129 (2007).

17 Gerstein MB, Lu ZJ, van Nostrand EL *et al.* Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. *Science* 330(6012), 1775–1787 (2010).

18 Liu T, Rechtsteiner A, Egelhofer TA *et al.* Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Res.* 21(2), 227–236 (2011).

19 Hawkins RD, Hon GC, Lee LK *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 6(5), 479–491 (2010).

▪▪ **Evidence that repressive chromosomal domains considerably expand in size upon cellular differentiation.**

20 Mikkelsen TS, Ku M, Jaffe DB *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153), 553–560 (2007).

21 Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* 41(2), 246–250 (2009).

▪▪ **Evidence that repressive chromosomal domains considerably expand in size upon cellular differentiation.**

22 Guelen L, Pagie L, Brasset E *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453(7197), 948–951 (2008).

▪ **CTCF sites are found at the boundaries of chromosomal domains that locate to the repressive compartment at the nuclear periphery.**

23 Filion GJ, van Steensel B. Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells. *Nat. Genet.* 42(1), 4; author reply 5–6 (2010).

24 Dodge JE, Kang YK, Beppu H, Lei H, Li E. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.* 24(6), 2478–2486 (2004).

25 Faust C, Lawson KA, Schork NJ, Thiel B, Magnuson T. The Polycomb-group gene *eed* is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development* 125(22), 4495–4506 (1998).

26 O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. The Polycomb-group gene *Ezh2* is required for early mouse development. *Mol. Cell. Biol.* 21(13), 4330–4336 (2001).

27 Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K. The Polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol. Cell. Biol.* 27(10), 3769–3779 (2007).

28 Peters AH, O'Carroll D, Scherthan H *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 107(3), 323–337 (2001).

29 Tachibana M, Sugimoto K, Nozaki M *et al.* G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* 16(14), 1779–1791 (2002).

30 Negre N, Brown Cd, Shah PK *et al.* A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet.* 6(1), e1000814 (2010).

31 Negre N, Brown CD, Ma L *et al.* A *cis*-regulatory map of the *Drosophila* genome. *Nature* 471(7339), 527–531 (2011).

32 Gurudatta BV, Corces VG. Chromatin insulators: lessons from the fly. *Brief Funct. Genomic Proteomic* 8(4), 276–282 (2009).

33 Moshkovich N, Nisha P, Boyle PJ, Thompson BA, Dale RK, Lei EP. RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. *Genes Dev.* 25(16), 1686–1701 (2011).

34 Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98(3), 387–396 (1999).

35 West AG, Gaszner M, Felsenfeld G. Insulators: many functions, many mechanisms. *Genes Dev.* 16(3), 271–288 (2002).

36 Scott Kc, Taubman AD, Geyer PK. Enhancer blocking by the *Drosophila* gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* 153(2), 787–798 (1999).

37 Kellum R, Schedl P. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64(5), 941–950 (1991).

38 Roseman RR, Pirrotta V, Geyer Pk. The su(Hw) protein insulates expression of the *Drosophila melanogaster* white gene from chromosomal position-effects. *EMBO J.* 12(2), 435–442 (1993).

39 Mutskov V, Felsenfeld G. Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *EMBO J.* 23(1), 138–149 (2004).

40 Mutskov VJ, Farrell CM, Wade PA, Wolffe AP, Felsenfeld G. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev.* 16(12), 1540–1554 (2002).

41 Pikaart MJ, Recillas-Targa F, Felsenfeld G. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* 12(18), 2852–2862 (1998).

42 Recillas-Targa F, Pikaart MJ, Burgess-Beusse B *et al.* Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc. Natl Acad. Sci. USA* 99(10), 6883–6888 (2002).

43 Aker M, Tubb J, Groth AC *et al.* Extended core sequences from the cHS4 insulator are necessary for protecting retroviral vectors from silencing position effects. *Hum. Gene Ther.* 18(4), 333–343 (2007).

44 Van Der Vlag J, Den Blaauwen JL, Sewalt RG, Van Driel R, Otte AP. Transcriptional repression mediated by Polycomb group proteins and other chromatin-associated repressors is selectively blocked by insulators. *J. Biol. Chem.* 275(1), 697–704 (2000).

45 Moreno R, Martinez I, Petriz J, Gonzalez JR, Gratacos E, Aran JM. Boundary sequences stabilize transgene expression from subtle position effects in retroviral vectors. *Blood Cells Mol. Dis.* 43(2), 214–220 (2009).

46 Yannaki E, Tubb J, Aker M, Stamatoyannopoulos G, Emery Dw. Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. *Mol. Ther.* 5(5 Pt 1), 589–598 (2002).

47 Hanawa H, Yamamoto M, Zhao H, Shimada T, Persons DA. Optimized lentiviral vector design improves titer and transgene expression of vectors containing the chicken beta-globin locus HS4 insulator element. *Mol. Ther.* 17(4), 667–674 (2009).

48 Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452(7184), 243–247 (2008).

49 Kim YJ, Cecchini KR, Kim TH. Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc. Natl Acad. Sci. USA* 108(18), 7391–7396 (2011).

50 Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, Felsenfeld G. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.* 20(9), 2224–2235 (2001).

51 Prioleau MN, Nony P, Simpson M, Felsenfeld G. An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J.* 18(14), 4035–4048 (1999).

52 Hebbes TR, Clayton AL, Thorne AW, Crane-Robinson C. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* 13(8), 1823–1830 (1994).

53 Bruce K, Myers FA, Mantouvalou E *et al.* The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. *Nucleic Acids Res.* 33(17), 5633–5639 (2005).

54 Ma MK, Heath C, Hair A, West AG. Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity. *PLoS Genet.* 7(7), e1002175 (2011).

▪▪ **Barrier elements in chickens manipulate histone modification crosstalk to form the boundaries of a heterochromatin domain.**

55 Myers FA, Chong W, Evans DR, Thorne AW, Crane-Robinson C. Acetylation of histone H2B mirrors that of H4 and H3 at the chicken beta-globin locus but not at housekeeping genes. *J. Biol. Chem.* 278(38), 36315–36322 (2003).

56 Felsenfeld G, Burgess-Beusse B, Farrell C *et al.* Chromatin boundaries and chromatin domains. *Cold Spring Harb. Symp. Quant. Biol.* 69, 245–250 (2004).

57 Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293(5539), 2453–2455 (2001).

58 Ghirlando R, Litt MD, Prioleau MN, Recillas-Targa F, Felsenfeld G. Physical properties of a genomic condensed chromatin fragment. *J. Mol. Biol.* 336(3), 597–605 (2004).

59 Chung JH, Whiteley M, Felsenfeld G. A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* 74(3), 505–514 (1993).

60 Yao S, Osborne CS, Bharadwaj RR *et al.* Retrovirus silencer blocking by the cHS4 insulator is CTCF independent. *Nucleic Acids Res.* 31(18), 5317–5323 (2003).

61 Dickson J, Gowher H, Strogantsev R *et al.* VEZF1 elements mediate protection from DNA methylation. *PLoS Genet.* 6(1), e1000804 (2010).

▪ **Barrier activity of the HS4 insulator mediates protection from DNA methylation in addition to histone modification recruitment.**

62 West AG, Huang S, Gaszner M, Litt MD, Felsenfeld G. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol. Cell* 16(3), 453–463 (2004).

63 Huang S, Li X, Yusufzai TM, Qiu Y, Felsenfeld G. USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier. *Mol. Cell. Biol.* 27(22), 7991–8002 (2007).

64 Huang S, Litt M, Felsenfeld G. Methylation of histone H4 by arginine methyltransferase PRMT1 is essential *in vivo* for many subsequent histone modifications. *Genes Dev.* 19(16), 1885–1893 (2005).

65 Chung JH, Bell AC, Felsenfeld G. Characterization of the chicken beta-globin insulator. *Proc. Natl Acad. Sci. USA* 94(2), 575–580 (1997).

66 Gowher H, Stuhlmann H, Felsenfeld G. Vezf1 regulates genomic DNA methylation through its effects on expression of DNA methyltransferase Dnmt3b. *Genes Dev.* 22(15), 2075–2084 (2008).

67 Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* 10(5), 295–304 (2009).

68 Heath H, Ribeiro de Almeida C, Sleutels F *et al.* CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J.* 27(21), 2839–2850 (2008).

69 Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell* 137(7), 1194–1211 (2009).

70 Chan PK, Wai A, Philipsen S, Tan-Un KC. 5′HS5 of the human beta-globin locus control region is dispensable for the formation of the beta-globin active chromatin hub. *PLoS ONE* 3(5), e2134 (2008).

71 Zakany J, Duboule D. The role of Hox genes during vertebrate limb development. *Curr. Opin. Genet. Dev.* 17(4), 359–366 (2007).

72 Goto Y, Kimura H. Inactive X chromosome-specific histone H3 modifications and CpG hypomethylation flank a chromatin boundary between an X-inactivated and an escape gene. *Nucleic Acids Res.* 37(22), 7416–7428 (2009).

73 Miller JC, Tan S, Qiao G *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29(2), 143–148 (2011).

74 Kim TH, Abdullaev ZK, Smith AD *et al.* Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128(6), 1231–1245 (2007).

75 Xi H, Shulha HP, Lin JM *et al.* Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet.* 3(8), e136 (2007).

76 Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES. Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc. Natl Acad. Sci. USA* 104(17), 7145–7150 (2007).

77 Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* 19(1), 24–32 (2009).

▪ **CTCF sites at the boundaries of chromosomal domains enriched in the Polycomb-associated mark H3K27me3.**

78 Barski A, Cuddapah S, Cui K *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* 129(4), 823–837 (2007).

79 Jothi R, Cuddapah S, Barski A, Cui K, Zhao K. Genome-wide identification of *in vivo* protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.* 36(16), 5221–5231 (2008).

80 Chen X, Xu H, Yuan P *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6), 1106–1117 (2008).

81 Schmidt D, Schwalie PC, Ross-Innes CS *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* 20(5), 578–588 (2010).

82 Bernstein BE, Stamatoyannopoulos JA, Costello JF *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* 28(10), 1045–1048 (2010).

83 Myers RM, Stamatoyannopoulos J, Snyder M *et al.* A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* 9(4), e1001046 (2011).

84 Martin D, Pantoja C, Fernandez Minan A *et al.* Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nat. Struct. Mol. Biol.* 18(6), 708–714 (2011).

85 Bao L, Zhou M, Cui Y. CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators. *Nucleic Acids Res.* 36(Database issue), D83–D87 (2008).

86 Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. *Curr. Opin. Genet. Dev.* 17(5), 400–407 (2007).

87 Lieberman-Aiden E, van Berkum NL, Williams L *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326(5950), 289–293 (2009).

88 Yao H, Brick K, Evrard Y, Xiao T, Camerini-Otero RD, Felsenfeld G. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev.* 24(22), 2543–2555 (2010).

89 Parelho V, Hadjur S, Spivakov M *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132(3), 422–433 (2008).

90 Wendt KS, Yoshida K, Itoh T *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451(7180), 796–801 (2008).

91 Chien R, Zeng W, Kawauchi S *et al.* Cohesin mediates chromatin interactions that regulate mammalian beta-globin expression. *J. Biol. Chem.* 286(20), 17870–17878 (2011).

92 Dorsett D. Cohesin: genomic insights into controlling gene transcription and development. *Curr. Opin. Genet. Dev.* 21(2), 199–206 (2011).

93 Hadjur S, Williams LM, Ryan NK *et al.* Cohesins form chromosomal cis-interactions at the developmentally regulated *IFNG* locus. *Nature* 460(7253), 410–413 (2009).

94 Mishiro T, Ishihara K, Hino S *et al.* Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J.* 28(9), 1234–1245 (2009).

95 Nativio R, Wendt KS, Ito Y *et al.* Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* 5(11), e1000739 (2009).

96 Gombert WM, Farris SD, Rubio ED, Morey-Rosler KM, Schubach WH, Krumm A. The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. *Mol. Cell. Biol.* 23(24), 9338–9348 (2003).

97 Gombert WM, Krumm A. Targeted deletion of multiple CTCF-binding elements in the human C-MYC gene reveals a requirement for CTCF in C-MYC expression. *PLoS ONE* 4(7), e6109 (2009).

98 Essafi A, Webb A, Berry RL *et al.* A wt1-controlled chromatin switching mechanism underpins tissue-specific wnt4 activation and repression. *Dev. Cell* 21(3), 559–574 (2011).

▪ **CTCF/cohesin sites mark the boundaries of a dynamic chromatin domain at the Wnt4 locus.**

99 Bartolomei MS. Genomic imprinting: employing and avoiding epigenetic processes. *Genes Dev.* 23(18), 2124–2133 (2009).

100 Lin S, Ferguson-Smith AC, Schultz RM, Bartolomei MS. Nonallelic transcriptional roles of CTCF and cohesins at imprinted loci. *Mol. Cell. Biol.* 31(15), 3094–3104 (2011).

101 de Laat W, Grosveld F. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* 11(5), 447–459 (2003).

102 Splinter E, Heath H, Kooren J *et al.* CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* 20(17), 2349–2354 (2006).

103 Farrell CM, West AG, Felsenfeld G. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.* 22(11), 3820–3831 (2002).

104 Bender MA, Byron R, Ragoczy T, Telling A, Bulger M, Groudine M. Flanking HS-62.5 and 3′ HS1, and regions upstream of the LCR, are not required for beta-globin transcription. *Blood* 108(4), 1395–1401 (2006).

105 Filippova GN, Cheng MK, Moore JM *et al.* Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev. Cell* 8(1), 31–42 (2005).

106 Bernstein BE, Kamal M, Lindblad-Toh K *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120(2), 169–181 (2005).

107 Magdinier F, Yusufzai Tm, Felsenfeld G. Both CTCF-dependent and -independent insulators are found between the mouse T cell receptor alpha and Dad1 genes. *J. Biol. Chem.* 279(24), 25381–25389 (2004).

108 Gomos-Klein J, Harrow F, Alarcon J, Ortiz Bd. CTCF-independent, but not CTCF-dependent, elements significantly contribute to TCR-alpha locus control region activity. *J. Immunol.* 179(2), 1088–1095 (2007).

109 Harrow F, Amuta Ju, Hutchinson Sr, Akwaa F, Ortiz Bd. Factors binding a non-classical Cis-element prevent heterochromatin effects on locus control region activity. *J. Biol. Chem.* 279(17), 17842–17849 (2004).

110 Ortiz BD, Harrow F, Cado D, Santoso B, Winoto A. Function and factor interactions of a locus control region element in the mouse T cell receptor-alpha/Dad1 gene locus. *J. Immunol.* 167(7), 3836–3845 (2001).

111 Raab JR, Kamakaka RT. Insulators and promoters: closer than we think. *Nat. Rev. Genet.* 11(6), 439–446 (2010).

112 Gallagher PG, Nilson DG, Steiner LA, Maksimova YD, Lin JY, Bodine DM. An insulator with barrier-element activity promotes alpha-spectrin gene expression in erythroid cells. *Blood* 113(7), 1547–1554 (2009).

113 Wong EY, Lin J, Forget BG, Bodine DM, Gallagher PG. Sequences downstream of the erythroid promoter are required for high level expression of the human alpha-spectrin gene. *J. Biol. Chem.* 279(53), 55024–55033 (2004).

114 Gallagher PG, Steiner LA, Liem RI *et al.* Mutation of a barrier insulator in the human ankyrin-1 gene is associated with hereditary spherocytosis. *J. Clin. Invest.* 120(12), 4453–4465 (2010).

▪ **Mutations that disrupt a barrier element at the human ankyrin-1 gene are associated with hereditary anemias.**

115 Carroll D. Genome engineering with zinc-finger nucleases. *Genetics* 188(4), 773–782 (2011).

116 Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. *Science* 333(6051), 1843–1846 (2011).

117 van Bemmel JG, Pagie L, Braunschweig U *et al.* The insulator protein SU(HW) fine-tunes nuclear lamina interactions of the *Drosophila* genome. *PLoS ONE* 5(11), e15013 (2010).

118 Peric-Hupkes D, Meuleman W, Pagie L *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* 38(4), 603–613 (2010).

▪▪ **Chromosomal domains that associate with the repressive compartment at the nuclear periphery redistribute upon cellular differentiation.**

Website

201 CTCFBSDB: a CTCF binding site database for characterization of vertebrate genomic insulators. http://insulatordb.uthsc.edu