JCMM
Available online at
www.jcmm.ro
www.jcmm.org
Epigenetic Review Series

J. Cell. Mol. Med. Vol 10, No 3, 2006 pp. 554-568

Guest Editor: J.J. Cassiman

Epigenetic boundaries of tumour suppressor gene promoters: the CTCF connection and its role in carcinogenesis

Felix Recillas-Targa [a] *, Inti A. De La Rosa-Velázquez [a], Ernesto Soto-Reyes [a], Luis Benítez-Bribiesca [b]

[a] Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, México
[b] Hospital de Oncología, Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, México

Received: May 10, 2006; Accepted: May 23, 2006

Abstract

Genetic and epigenetic regulations are essential mechanisms that ensure proper early and subsequent mammalian programming of diverse cellular processes. These mechanisms affect transcriptional regulation, stem cell determination and cell cycle control, including senescence and aging. It is not surprising that perturbation of the exquisite balance between genetic and epigenetic regulation can lead to diverse diseases, including cancer. Histone covalent modifications and DNA methylation do not explain all epigenetic phenomena. We describe a previously unsuspected epigenetic factor and propose the incorporation of the 11-zinc finger CCCTC-binding factor, known as CTCF as a novel and multifunctional epigenetic regulator.

Keywords: epigenetic • CTCF and BORIS nuclear factors • intergenic transcripts • chromatin • DNA methylation • histone modification • insulator • carcinogenesis • tumour suppressor gene

Introduction

A large number of recent reviews have provided numerous details concerning epigenetic mechanisms [1–3]. Epigenetics has been defined too loosely. Some consider epigenetics only to refer to DNA methylation. But the complete definition should include histone modification, diverse chromatin processes and genomic imprinting. Firstly, epigenetics should be understood through the

* Correspondence to: Félix RECILLAS-TARGA
Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510, México.
Tel.: (52 55) 56 22 56 74; Fax: (52 55) 56 22 56 30
E-mail: frecilla@ifc.unam.mx

Reprinted from:
Journal of Cellular and Molecular Medicine
doi:10.2755/jcmm010.003.02

genotype, which will give substrate to the epigenotype. Secondly, epigenetics needs to be viewed in the context of the epigenome, with its capacity to transmit heritable regulatory patterns without affecting the primary sequence of DNA [4]. As anticipated, the epigenome resides in a chromatinized environment in which histones, which form part of the nucleosomes, represent a key target for epigenetic regulation. Here we briefly describe the cellular processes that should be considered as part of mammalian epigenetic regulation [5–8]. Furthermore, we present a range of evidence leading to the proposal of a mechanism that protects a CpG-island against abnormal DNA methylation in the context of epigenetic silencing of tumour suppressor genes in cancer.

It is generally accepted that organization of the eukaryotic genome through chromatin is not the sole epigenetic component. One of the most studied aspects of epigenetics, in particular in vertebrate organisms, is DNA methylation that is understood as the incorporation of a methyl-group at the C-5 position of cytosine when it occurs in CpG dinucleotides [4]. Methylation is carried out by a family of DNA methyl-transferases (Dnmt's) [9]. DNA methylation is synonymous with epigenetic silencing and has been implicated in two main normal cellular processes: *i*) maintenance of a stable state of "parasitic" repetitive sequences (like retrotransposons, retroviral sequences, repetitive sequences among others) [10] and, *ii*) in early developmental gene expression through imprinting mechanisms [11]. Additionally, abnormal DNA methylation associated with human diseases has been clearly demonstrated [3, 12]. Such effect is thought to be mainly due to the abnormal and paradoxical DNA methylation of CpG-islands corresponding to gene promoters and differentialy methylated regions on imprinted genes [12–14]. The paradox resides in the fact that a CpG-island, a relatively short genomic sequence (200 to 1000 bp) rich in CpG dinucleotides, in normal conditions should never be methylated [15]. DNA hypermethylation over tumour suppressor gene promoters and initial exons has been now extensively demonstrated as a contributing factor for cancer development [12, 14]. But, in addition, interdependent contribution of genetic and epigenetic events has been clearly associated to cancer, for instance, one tumour suppressor allele mutation and additional epigenetic silencing of a second tumour suppressor gene through DNA hypermethylation could lead to cell transformation [16–18]. Interestingly, the first clue of the involvement of DNA methylation abnormalities in disease and development was the identification of generalized DNA hypomethylation in cancer cells [19, 20]. DNA hypomethylation causes global epigenome instability leading to the release of different types of genomic mobile elements from DNA organized in stable and heterochromatin-like conformation [10]. Furthermore, hypomethylation has been found to correlate with activation of several oncogenes [21]. DNA methylation contributes to different syndromes through altering genomic imprinting regulation [13, 22, 23]. All these evidences led to the emerging of an attractive concept suggesting the existence of a "*methylator phenotype*", which could be useful in the classification of tumour-types, by analysis of the genes inactivated through several epigenetic factors [24]. Moreover, due to the development of highly sensitive methodologies to identify DNA methylation it is feasible that in the near future this approach could constitute a useful tool for the detection and diagnosis of some diseases.

Histone modifications: an overview

It is not the purpose of this review to systematically describe all the histone covalent modifications. For extensive review see Hake *et al.* 2004 [25]. Today there is no doubt that histone modifications and the incorporation of histone variants contribute to the epigenetic regulation of gene expression. Modifications of histones like, mono-, di- and trimethylation of lysine or arginine residues, acetylation and deacetylation, phosphorylation, ubiquitination and sumoylation create diverse binding interfaces for epigenetic effectors [2]. The combinatorial modifications of amino-terminal and the core domains of histones reveal an unsuspected complex regulatory scenario that is far from being understood [25]. In addition to the varied effects on cellular processes, like chromosome segregation, DNA repair, chromatin assembly, gene expression,

among others, histone modifications also contribute to defective epigenetic regulation [25]. As one of the most important aspects emerging from epigenetic studies is the place which histone modifications occupy in the hierarchy of the epigenetic processes. This is a complex issue since, as the unified theory of gene regulation states [26], there are apparently no general rules governing epigenetic regulation, especially when considering histone post-translational modifications.

Other epigenetic processes

Even though researchers had placed abnormal DNA methylation on a prime line of epigenetic miss-regulation in human cancer, there is growing experimental evidence of additional epigenetic processes involved in distinct neoplasias [1, 2, 18]. Unfortunately, the order of appearance of the genetic and epigenetic events, as well as their combinatorial effects, leading to abnormal cell function remains undetermined [17, 24]. There is little doubt concerning the involvement of several epigenetic mechanisms, in addition to DNA methylation, in abnormal gene silencing with consequences in cancer. It is not the purpose of the present review to show in detail those processes but rather to provide a general view of them [for review see 2, 18]. Chromatin remodeling is at the center of all the epigenetic events. One of the principal chromatin remodeling activities is represented by the establishment, reading and translation of the wide variety of combinatorial covalent histone modifications [25]. But histone modifications are not the sole epigenetic source of regulation based on chromatin organization of eukaryotic genome. Moreover, histone modification, histone variants, ATP-dependent chromatin remodeling complexes, Polycomb and Trithorax group of proteins, non-coding RNA's and more recently nuclear structure and dynamics, all participate with certain interdependency on epigenetic regulation, and when altered they have a role on malignant cell transformation [2].

Here we review evidence about the involvement of novel molecules, in particular nuclear factors, in epigenetic regulation but also we underscore the fact that many other yet undiscovered factors might be involved.

CTCF and its contribution to epigenetics and imprinting control

It is well established that transcription factors regulate gene expression through their DNA interacting capacity [27]. One of such nuclear factors that clearly differ from this general definition is the CCCTC-binding factor known as CTCF [28, 29]. CTCF is a multifunctional 11-zinc-finger nuclear factor which is highly conserved among distinct organisms and can be post-translationally modified through phosphorylation and poly(ADP-ribosyl)ation [29, 30]. Until now no clear functional consequences have been attributed to these two modifications over CTCF even though it has been proposed that poly(ADP-ribosyl)ation may have a role in modulating chromatin insulator activity [31]. CTCF was initially characterized as a negative and positive regulator of different genes like c-myc and apolipoprotein [28]. Furthermore, through screening enhancer-blocking activity (one of the two functional properties of chromatin insulators) it has been demonstrated that such activity is dependent on CTCF in the great majority of vertebrate insulators [28]. An orthologue factor named BORIS, expressing specifically in testis, has been isolated, [32, 33]. More recently CTCF has been implicated in the control of allele-specific gene expression on imprinted loci and of a growing list of promoters involved in cell cycle control, differentiation and apoptosis (see below).

Genomic imprinting has long been understood as an epigenetic way of regulation of specific loci where only one parental allele is expressed, while the other is silenced [11, 22, 23]. It has been clearly established that chromatin structure and DNA methylation play a central role in the differential regulation of allele-specific expression [13]. A large number of imprinted loci, in addition to standard regulatory elements (enhancers and promoters), are controlled by differentially methylated domains (DMD), also known as imprinting control regions (ICR) [13, 34–36]. Recent investigations have demonstrated that ICR elements are recognized *in vivo* by CTCF and that such binding is DNA methylation-sensitive [37, 38]. Moreover, CTCF interacts with most of the imprinted loci studied. In the case of the human and mice *Igf2/H19* locus, CTCF interacts with the maternal ICR inducing the blockage of activation of the *Igf2* gene by an upstream enhancer

(Fig. 1), allowing maternal expression of the proximal *H19* gene [13, 23, 34]. In contrast, the paternal allele ICR turns to be hypermethylated, obstructing CTCF binding and the silencing of *H19* gene expression. In such a context, downstream enhancers are not any more blocked and the paternal *Igf2* allele is exclusively activated (Fig. 1). Thus, it has been demonstrated that CTCF protects the maternal allele ICR against DNA methylation [34, 39]. Novel studies using the chromosome conformation capture, also known as the 3C assay, demonstrate that the allele-specific enhancer-blocking activity of CTCF is achieved at least in part by chromosome looping [23]. Distinct ICR's were shown to establish differential contacts, interestingly; all are contacted by CTCF in their unmethylated state. The maternal unmethylated *H19* ICR, which is contacted by CTCF at multiple sites, is able to attract another ICR located in close proximity to *Igf2*, resulting in the formation of two independent chromatin domains, in which *H19* can be activated by the downstream enhancers [23].

In summary, CTCF represents an active player in the determination of allele-specific gene expression over distinct imprinted loci. Its participation in long-distance communication between regulatory elements and chromosomal loop formation contributes to topological organization, what partially explains how CTCF acts not only as a chromatin insulator (enhancer-blocker) in the regulation of some imprinted loci, but also protecting a specific allele from local chromosome topology-dependent DNA methylation (see below).

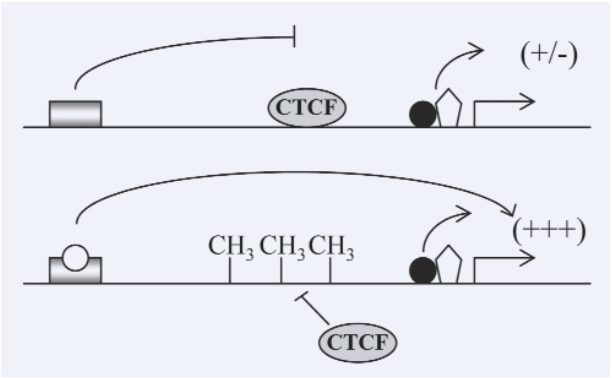

Fig. 1 Model of CTCF enhancer-blocking activity. Communication interference between enhancer (rectangle) and promoter (symbols close to the transcription initiation site) can be blocked when CTCF is located between them though its insulator enhancer-blocking properties. Such mechanism can be regulated through DNA methylation (CH_3) which interferes with CTCF binding eliminating the CTCF-dependent enhancer-blocking activity. This effect will promote enhancer action over promoter elements inducing regulated gene expression activation (lower panel). More over, the enhancer can adopt specific conformation (rectangle + open circle) that complements its function. This feature had been observed for several imprinted loci where CTCF-sensitive DNA methylation controls part of the choice of the expressed allele [13, 23]. More recently, the CTCF-dependent enhancer-blocking was defined to be regulated in a hormone-dependent way [56]. +/- indicate the transcription activation strength.

CTCF and BORIS

In recent studies focused in finding CTCF binding sites using nuclear extracts from testicular cells, Lobanenkov and collaborators identified a novel factor that binds to the CTCF target sequence. Such factor is the testis-specific expressing *CTCF* paralogue, named *Brother of the Regulator of Imprinted Site*s (BORIS), which encodes a protein highly conserved on the central 11-zinc-finger portion but with divergent amino- and carboxy-terminal domains [40, 41]. *BORIS* expression is restricted to germ cells during spermatogenesis [41]. The expression of *BORIS* in male germ cells correlates with a drastic decrease in *CTCF* expression, erasure of specific methylation patterns and up-regulation of a specific class of genes encoding testis cancer antigens [40, 41]. Once *CTCF* expression pattern is established BORIS is down-regulated with the concomitant incorporation of maternal-specific methylation marks [41]. Then, it has been proposed that BORIS could play a role in the early stages of male germ line epigenetic programming.

The identity between BORIS and CTCF over the 11-zinc-finger domain is consistent with the fact that both factors recognize the same DNA sites *in vitro* [41]. The current model points to a correlation between *BORIS* expression and a generalized erasure of DNA methylation, suggesting that BORIS could be associated with a hypothetical DNA demethylase [41]. When *BORIS* expression is experimentally silenced in normal male germ cells CTCF is activated, confirming their mutually

exclusive expression pattern. Thus, in such a context BORIS and CTCF could perform complementary epigenetic functions. Under this scenario, when CTCF is active it could protect CpG-dinucleotides against *de novo* DNA methylation, like in the *Igf2/H19* ICR maternal allele, however, the possibility that CTCF could directly or indirectly recruit *de novo* DNA methyltransferase activity to the paternal allele of imprinted loci remains to be tested. This question could be addressed by looking for CTCF interaction with other chromatin modifiers, such as Suv39h1, Suv39h2 or Polycomb group of proteins like EZH2, which has been recently shown to recruit DNA methyltransferase activity [42, 43].

Two alternative possibilities can not be ruled out. First, BORIS pattern of expression and spectrum of activities could not be restricted to testis cells. Second, CTCF and BORIS could represent pioneer members of a larger family of 11-zinc-finger factors with epigenetic regulatory properties. Such scenarios are currently under investigation.

The relevance of CTCF and BORIS epigenetic functions in cancer is highlighted by two recent reports showing that BORIS is abnormally expressed in multiple cancer cells, including human fibroblasts and lung cancer cells [32, 33]. Such aberrant BORIS expression in non-testis cells causes derepression and demethylation of at least two well-known testis-specific cancer genes: *MAGE-A1* and *NY-ESO-1* [32, 33]. In addition, and due to the high similarity of the 11-zinc-finger domain between BORIS and CTCF, BORIS competes with CTCF for previously identified binding sequences in non-testis cells [32, 33]. On the other hand, due to the divergence of the amino- and carboxy-termini between BORIS and CTCF, a logical prediction could be that BORIS establishes functional interactions with different co-factors from those normally associated with CTCF. Thus, the presence of BORIS could drastically affect the expression of CTCF gene targets in somatic cells causing epigenetic regulatory alterations leading to cell transformation. Furthermore, abnormal activation of BORIS in primary lung cancers coincides with DNA demethylation and histone hyperacetylation within a CpG-island located in the *BORIS* gene promoter [33]. Thus, we can predict that *BORIS* depression could be involved in a global hypomethylaiton-mediated feedback contributing to somatic cell transformation.

In summary, there is an apparently sensitive balance between BORIS and CTCF that when altered leads to abnormal regulatory consequences involving DNA methylation. However, the *in vitro* and *in vivo* data still do not clarify the capacity of BORIS and CTCF to bind methylated DNA. Thus, the evidence strongly suggests the involvement of CTCF and BORIS in the epigenetic regulation of tumour related genes.

CTCF and escape genes

Another aspect of CTCF has to do with the creation of particular chromosomal boundaries on the inactive X chromosome allowing the expression of some genes that escape to X chromosome inactivation known as escape genes (between 10 to 20% of human genes located on the X chromosome escape inactivation) [44]. Recently published data revealed the presence of CTCF at the vicinity of some of the escape genes located in mouse and human X inactivated chromosomes [45]. CTCF has been found at the 5' non-coding region of the mouse *Jarid1C* and *Eif2s2x* genes as well as the human *EIF2S3* gene. Interestingly, the incorporation of CTCF over such escape genes correlates with some chromatin marks presumably affecting gene expression [45]. There is an unusual high level of histone H3 acetylation at the transition regions between distinct domains, what is consistent with the proposal that CTCF acts as a chromatin insulator [45]. Furthermore, newly defined CTCF binding sites, integrated in the chromatin context, showed enhancer-blocking properties in a colony assay. With the available evidence we cannot rule out the possibility that CTCF binding sites at escape domains protect, through enhancer-blocking properties, against inappropriate activation of silent genes adjacent or in proximity to the active escape domains. The most attractive observation has to do with the fact that CTCF influences the DNA methylation status of such particular domains embedded in the inactivated X chromosomes [45]. In fact there are low DNA methylation levels over the escape domain and basically no-methylation at the CTCF binding sites [45]. As mentioned before CTCF can be a positive or negative regulator of transcription but such regulatory functions does not seem to fully explain the CTCF

role on escape genes expression [28]. Apparently, CTCF plays a more structural role contributing to the formation of an escape gene domain, thus facilitating a particular topological organization (see below). What remains intriguing is the fact that not all the escape genes are associated with CTCF sites [45]. The possibility that those CTCF binding sites constitute regulatory elements or that topological contribution of CTCF is carried in a larger genomic scale remains difficult to visualize. But the main point from this series of findings is that CTCF sites in escape domains remain unmethylated during development. It is important to recall that initiation and propagation of chromosome X inactivation occurs in the absence of DNA methylation. Thus we cannot discard the possibility that CTCF and associated factors may contribute to shield escape domains not only against DNA methylation but also repressive histone marks (Fig. 2), or even counteract the repressive action of Polycomb protein members [46, 47].

In conclusion, the evidence provide an example in which CTCF protects a genomic domain, and probably some nearby genes, against DNA methylation besides the addition of epigenetic repressive marks in a very adverse environment for transcriptional activation like the inactive X chromosome.

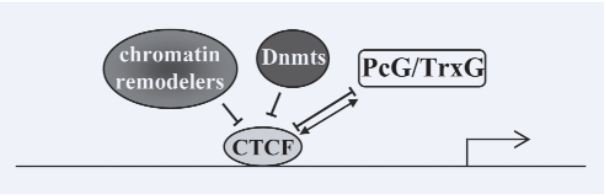

Fig. 2 CTCF and epigenetic components. In this model, CTCF could be responsible to interfere with the action of epigenetic repressive components and/or the recruitment of active chromatin remodeling activities. This activity will be in addition to the promoter regulation which is in charge of a combination of specific a ubiquitous transcription factors, and their associated cofactors. At the present time there are no clear demonstrations thus CTCF interacts directly with those proteins, but we cannot discard that it could recruit different chromatin remodeling activities like Trithorax (TrxG) family members to protect a promoter against repressive epigenetic components like HDAC's, HMTases, DNA methylatransferases (Dnmt's) or even members of the Polycomb repressive complexes (PcG).

CTCF and non-coding transcripts

The epigenetic regulation of CTCF is exerted at domain scale also by interacting with intergenic transcripts. CTCF has been found to be associated to abnormal epigenetic regulation of an intergenic domain implicated in myotonic dystrophy (DM), which is one of the most prevalent disease forms of adult muscular dystrophy [48, 49]. DM is in part caused by the expansion of a CTG repeat in the 3' non-coding region of *DMPL* at the *DM1* locus on chromosome 19q13.3. Normal individuals have between 5 to 38 CTG repeats whereas in pathological conditions the *DM1* locus harbors up to thousands of repeats and the length of such amplified region correlates with the severity of disease. An outstanding observation showed that on wild-type (wt) alleles the CTG repeat region is flanked by CTCF binding sites which confine heterochromatin conformation strictly to the CTG repeats [50]. In aberrant conditions, when there is expansion of the CTG repeats, the heterochromatin spreads and silence adjacent genes like *SIX5* [51]. Such abnormal chromatin spreading involves DNA methylation that prevents binding of CTCF thereby disrupting insulator function what may contribute to CTG repeat expansion. More recently, genes expressing non-coding transcripts have been identified across the *DM1* intergenic domain including the CTG repeats subdomain [52]. RNase protection assays revealed the expression of 21 nucleotide transcripts across the CTG repeats. Those transcripts seem to be specific to the CTG repeat cluster since no-intergenic transcription can be detected upstream and downstream to this region which is delimited by CTCF binding sites (Fig. 3) [52]. Such observation further supports a boundary role of CTCF that is DNA methylation sensitive, and let us speculates that CTCF contributes to protect the *DM1* domain against CTG expansion. Using an elegant experimental strategy the same group has demonstrated that extension of antisense transcription is restricted to CTG repeats and delimited by CTCF binding. In contrast, the expression of sense transcripts starting at the *SIX5* gene is not restricted and extends all across the intergenic region [52]. Unexpectedly, when CTCF binding sequences where mutated a drastic increase on transcription

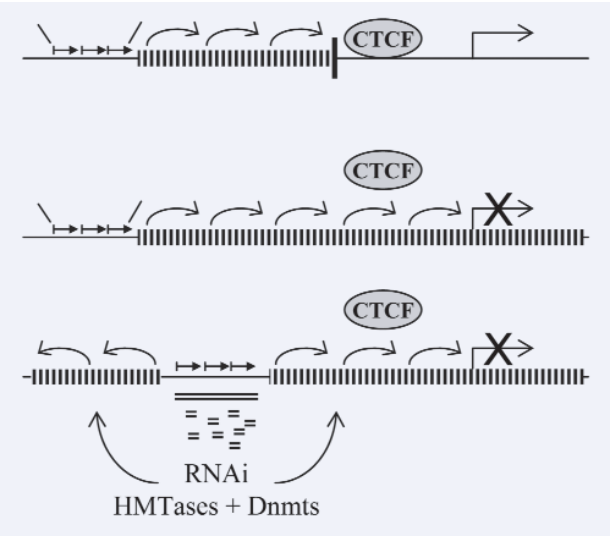

Fig. 3 Model of a promoter protection against spreading of epigenetic silencing by CTCF. In this model we propose that CTCF could have the capacity to block the spreading of repressive chromatin marks (vertical lines) that originate in non-coding regions of the genome corresponding to repetitive elements (horizontal arrows). Such elements could represent attraction "centres" for repressive or open chromatin marks like DNA methylation, histone deacetylation and PcG association [87, 88]. In this model, displacement of CTCF will cause the expansion of abnormal epigenetic silencing over the promoter and beyond the transcription initiation site (middle panel). The spreading will cause the epigenetic silencing of gene expression without presence of any genetic defects [14]. A complementary mechanism could come from the fact that those repetitive elements (horizontal arrows) could be transcribed [49], promoting the RNA interference (RNAi) machinery to induce the formation of heterochromatin (lower panel) [24]. In any case the signals for undesired epigenetic silencing can be situated in proximity or distant form promoters.

across the region was observed, consistent with an *in vivo* increase of incorporation of the phosphorylated form of RNA polymerase II [52].

In summary, the CTG repeat cluster wt allele is associated with bidirectional transcription, expression of 21 nt RNA fragments, H3-K9 trimethylation and HP1γ recruitment to such region. The current model suggests that in the abnormal expanded allele CTCF is not able to bind its recognition sequence, mainly because there is an increase on DNA methylation, causing the expansion of CTG repeats and associated heterochromatin marks. Consequently, we can speculate that CTCF might prevent the propagation of heterochromatin and/or CTG repeats as well as DNA methylation to the surrounding regions in the *DM1* locus and other loci.

CTCF and long-distance regulation of gene expression

In addition to its regulatory role CTCF has been described as playing a critical functional role on the great majority of vertebrate insulators [53]. CTCF participates in one of the two functional properties of insulators [54]. It is needed for enhancer-blocking activity of insulators [54], whereas it is dispensable for the protection against chromosomal position effects, which is the other insulator property [54, 55]. A growing number of laboratories are now interested on the understanding of chromatin insulators mode of action. Among several distinct possibilities CTCF seems to contribute to gene regulation through long distance regulatory mechanisms. It has been suggested that such contribution can occur through its participation on higher order chromatin organization, genome topology and probably, nuclear dynamics.

Regulated CTCF-dependent enhancer-blocking activity

One of the first clear evidences of long-distance action of insulators was the discovery of a CTCF-dependent enhancer blocker element that is hormone regulated [56]. CTCF binding sites are often flanked by thyroid hormone response elements (TREs), as it occurs in the chicken *lysozyme* and human *c-myc* regulatory regions [56]. In both loci, CTCF binding sites, located in close proximity to TREs, are found around 2 kb upstream of the transcription initiation site and more importantly, between an enhancer and its corresponding promoter. Most importantly, in the absence of thyroid hormone, an enhancer-blocking activity is turned on. In contrast, in the presence of thyroid hormone, CTCF-dependent enhancer-blocking function is interrupted [56], leading to gene expression. Thus, this is the first evidence showing a thyroid hormone regulated CTCF-dependent enhancer-block-

ing activity, with consequences in cellular differentiation (Fig. 1) [56]. To illustrate this novel property of CTCF we describe the chicken *lysozyme* gene case in which under undifferentiated conditions the enhancer (located at -2.7 kb) is blocked, allowing low levels of *lysozyme* gene expression. Conversely, when myeloid precursor cells are induced to differentiate the enhancer is no longer shielded, leading to maximal gene expression levels. In this case, the enhancer-blocking activity is abolished in a T3 thyroid hormone dependent way, with a concomitant remarkable increase of histone acetylation over the CTCF binding site, but also over the upstream intergenic sequences, including the enhancer [56]. Therefore, T3 thyroid hormone and CTCF seem to collaborate creating a molecular complex, probably including various co-factors, which in turn could influence the recruitment of histone acetylase activities, allowing the CTCF-dependent enhancer-blocking activity to be controlled (Fig. 1). The loss of enhancer-blocking activity in the presence of thyroid hormone is not caused by the dissociation of CTCF from chromatin [56]. Thus, the combinatorial of factors and co-factors directly associated, or in close proximity, to CTCF clearly dictates its enhancer-blocking capabilities. In our opinion this is extremely relevant since such scenario could be applied to the regulatory role of CTCF on tumour suppressor gene regulation (see below).

CTCF epigenetics and long-range gene regulation

An interesting and unexplored aspect of CTCF epigenetics is its distribution over the genome and possible role on long-distance regulation of gene expression. One of the initial discoveries of such a role was the distribution of CTCF over the chicken, mouse and human β-globin loci, and more recently the chicken α-globin domain [54, 57–59]. The chicken β-globin group of genes is framed *in vivo* by binding of CTCF from a 16 kb region of condensed chromatin on the 5'-side and an olfactory receptor gene on the 3'-side [60]. By chromatin immunoprecipitation, no internal CTCF sites have been found along the β-globin domain [54, 61]. In the case of the mouse and human loci CTCF has been located coinciding with DNase I hypersensitive sites, in particular, 5'HS5 of the LCR and 3'HS1 at both loci [57]. CTCF has also been found in novel sites far upstream of the mouse locus, namely HS-62.5 [62]. The attractive aspect of all those findings comes from a series of studies in which using the chromosome conformation capture assay it has been demonstrated that the DNase I hypersensitive sites come together and associate between them to create spatial clustering of chromatin, forming through multiple loop formation the so-called active chromatin hub (ACH), with regulatory consequences like differential expression during cellular differentiation [63–65]. Such ACH is even able to acquire different conformations depending on the erythroid differentiation stage [65]. The convergent point is that the great majority of the DNase I hypersensitive sites implicated in the ACH formation bind CTCF [57, 62, 65]. Such presence suggests that CTCF may be an active contributor to the long-distance regulation of genes through spatial and physical association to specific genomic locations, facilitating chromatin loop and ACH formation in a regulated manner. This possibility seems viable, since it has also been demonstrated that CTCF can multimerize, probably through some of the zinc-fingers that are not involved in contacting DNA [53, 58, 61], allowing multiple interactions between CTCF factors located at distinct sites, thus potentially facilitating the spatial clustering of distal genomic sites leading to the creation of an ACH.

CTCF, epigenome and nuclear dynamics

Investigations by Corces and collaborators have demonstrated that the *Drosophila gypsy* insulators, present at least on 500 sites all over the *Drosophila* polytene chromosomes, are grouped preferentially at the nuclear periphery forming what they have called the "insulator bodies" [66, 67]. Interestingly, in a recent genome-scale survey, CTCF has been found in more than 200 sites [68]. This suggests that CTCF is not present in every domain but, from our point of view, for unknown reasons, some loci may need CTCF for their topological organization. At this point we cannot discard that CTCF contributes to the formation of "insulator bodies", influencing nuclear dynamics and compartmentalization. Such model is supported by additional

CTCF properties that lay beyond to its multimerization capacities. CTCF interacts with the nuclear matrix, in particular, associates to the nucleolar periphery through nucleophosmin suggesting that those interactions contribute to the formation of chromatin loop structures similar to those created by the *Drosophila gypsy* insulator [61, 69, 70].

Inter- and intrachromosomal physical association through loop formation seems to occur with consequences in gene expression coordination [64, 71, 72]. CTCF-dependent interchromasomal colocalization between two different imprinted loci has been recently demonstrated [73]. This observation supports the concept of three-dimensional chromosomal territories organization of the epigenome in interphase nuclei [74]. Such hypothesis postulates that active genes would be located on the periphery of chromosomal territories, in contact with the inter-chromosomal channels where transcription factories and post-transcriptional processing machinery would be located. Experimental evidences suggest that there are fewer transcriptional factories than transcribed genes in each cell; consequently several genes probably share, in time and space, common factories even if they are located in distal chromosomes [64, 73]. One prediction is the requirement of molecules that should contribute to attract distinct genomic regions and/or stabilize their interaction in association with a factory in a three-dimensional context. We favor the possibility that CTCF may act as a topological higher-order chromatin organizer through its multimerization capabilities, interaction with other proteins and association with the nuclear matrix (Fig. 2C) [58, 70, 73]. We speculate that such interactions may contribute to determine the three-dimensional location of CTCF-associated genes, or chromatin domains, inside the nucleus. This model should allow a tight regulation of, for example, tumour suppressor genes during cell cycle progression or even during senescence, process in which characteristic heterochromatin foci are formed [6, 73]. Under this scenario, loose of CTCF epigenetic regulatory functions may be translated into destabilization of transcription factories and dissociation of topological inter- and intrachromosomal physical contacts, what could contribute an alternative epigenetic phenomenon linked to tumorigenesis.

We conclude that CTCF seems to possess very specific properties over distinct genomic locations and, in addition to its epigenetic regulatory role; it could participate not only in epigenome organization, nuclear compartmentalization and loop formation but also in the topology of the epigenome probably contributing to nuclear dynamics. At the present time more work is necessary to confirm some of the proposed models but an open question, very attractive for us, is how all those epigenetic properties of CTCF operate in tumour cells?

CTCF, tumour suppressor gene regulation and cancer

CTCF participates in the regulation of genes that are involved in cell cycle control since ectopic expression of CTCF causes cell growth inhibition [75]. Thus it has been postulated that CTCF may participate in the regulation of cell cycle progression at multiple levels, including perhaps, its action over human retinoblastoma gene. Moreover, in human erythroleukemic K562 cells CTCF over-expression promotes differentiation into the erythroid lineage in addition to cell growth retardation [76]. Conversely, CTCF knockdown inhibits such erythroid differentiation. Interestingly, CTCF participation in the control of the cell cycle could include an anti-apoptotic function [77] affecting varied target genes, suggesting the involvement of CTCF in distinct cellular pathways, what further supports its possible involvement in neoplasic processes.

In human cells, CTCF gene has been mapped to the chromosome band 16q22.1 [78]. Such chromosomal region frequently contains deletions found in sporadic breast and prostate tumours. For that reason loss of heterozygosity (LOH) at chromosome 16q22.1 and CTCF integrity has attracted the attention of several groups. It is thought that genetic loss of CTCF may be associated with miss-regulation of a large number of genes involved in the control of cell cycle progression. In a first series of studies, mutations or loss of CTCF transcript in Wilm's tumours were not detected [79, 80]. Three more recent studies surveyed the status of CTCF in breast cancer. The first one analyzed 17 tumour samples from invasive ductal breast carcinomas [81]. Sequence analysis revealed only one miss-sense mutation that is interpreted as a possible polymorphism, and a 14 bp insertion creating a premature

stop codon in the amino-terminus [81]. In addition, 153 patients with familial non-*BRCA1*/*BRCA2* breast cancer were analyzed. The results showed the existence of sequence variant (G240A) at the 5' untranslated region besides variant (C1455T) in two and five cases, respectively. Both variants were found to occur together only in three cases. This study suggests for the first time that germ line mutations in the *CTCF* gene do not represent a critical risk factor in familial breast cancer [82]. Finally the analysis of 344 samples of invasive breast carcinoma by differential immunocytolocalization revealed that CTCF abnormally locates in the cytoplasm in 77% of the cases [83]. Despite this fact the authors do not detect any correlation between the intensity and localization of the signal corresponding to CTCF protein with the tumour type. Thus, the potential role of CTCF in breast carcinogenesis and other neoplasias through regulating tumour suppressor genes must be clarified. It is important to emphasize that no systematic studies have been conducted, in addition to genetic surveys, to address the possibility of CTCF inactivation in cancer through epigenetic mechanisms (see below).

In the previous sections we have discussed a wide spectrum of CTCF activities and, to a certain extent we have provided evidence sustaining a role of CTCF in protecting against DNA methylation. Now we would like to extend our analysis taking into consideration several facts from the literature and some from our laboratory that suggest a role of CTCF in the control of many genes that regulate cell proliferation [40]. Previous investigations established that CTCF influences transcription positively as well as negatively [28]. More recently, CTCF has been found to regulate transcription of multiple genes including *p19*^{ARF}, *p16*^{INK4a}, *PIM-1*, *PLK*, *BRCA1*, *p53*, *p27*, *E-cadherine*, *E2F1*, *TERT* and *IGF2* among others [40, 84]. What seems very attractive is that the promoter sequence of the majority of these genes is abnormally methylated in different tumours [12]. Therefore, the list of aberrantly methylated CTCF DNA targets in different human cancers is not restricted to imprinted genes or the *Igf2/H19* ICR sites (Fig. 4) [13, 34, 59]. This view has now been strengthened by observations from our laboratory (De La Rosa-Velázquez *et al. submitted*), demonstrating that the CTCF-DNA interaction is methylation-sensitive and that the loss of CTCF from the *Rb* promoter correlates with the incorporation of a methyl-CpG-binding protein, with its cognate chromatin repressing components, which in turn could induce epigenetic silencing of this cell cycle regulator. We propose that CTCF is a critical component of a growing list of tumour suppressor genes that, in addition to its transcriptional positive and negative regulatory role, may protect such genes against undesired DNA methylation. Although our results strongly support the notion of the protective role of CTCF in several tumour suppressor gene promoters, it is clear that further investigation is needed to clarify this point.

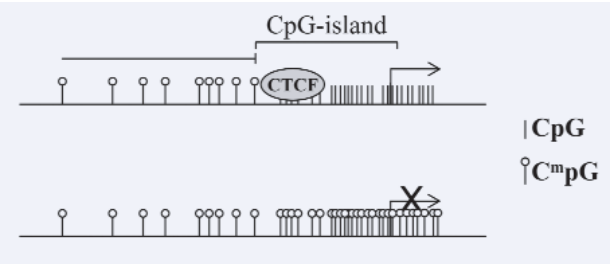

Fig. 4 Protection against DNA methylation of CpG-island by CTCF. The CpG-dinucleotide embedded in CTCF binding sites may participate in keeping the CpG-island unmethylated state. CTCF may be creating a boundary that delimits a DNA methylated domain from a hypomethylated CpG-island. Such model is supported by our observation suggesting that CTCF is located at the 5' boundary of the human retinoblastoma gene promoter but also by the identification of two CTCF sites at the human *BRCA1* gene promoter [84, De La Rosa-Velázquez *et al. submitted*]. Since it has been demonstrated that CTCF binding is DNA methylation-sensitive, its displacement will cause the spreading of DNA methylation with the subsequent epigenetic gene expression silencing (lower panel).

In agreement with our proposal for a protecting role of CTCF against DNA methylation it has been recently demonstrated that two CTCF sites, in coordination with Sp1 sites demarcate a boundary between methylated and unmethylated genomic domains in the *BRCA1* promoter (Fig. 4) [84]. Today it is generally accepted that genetic mutations are not the only way for the loss of *BRCA1* expression observed in breast tumours. In fact several studies have suggested the implication of altered patterns of DNA methylation in breast cancer [84, 85]. In fact, *BRCA1* tumour suppressor gene is abnormally methylated in 15 to 20% of sporadic breast cancer cases [84 and references therein]. The existence of CTCF binding sites creating a transition site

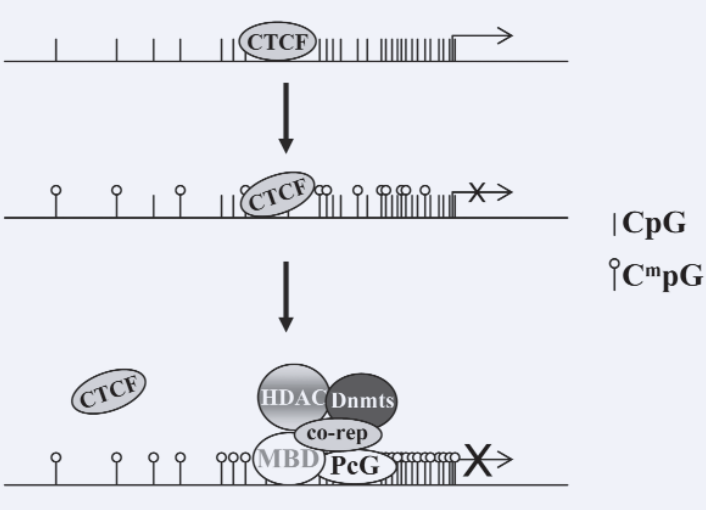

Fig. 5 Model in which CTCF can be displaced from a tumour suppressor gene promoter. CTCF binding can be weakened or blocked by DNA methylation (central panel). A complementary model to the previous figure considers that CTCF displacement could abnormally allow the incorporation of CpG-binding proteins (MBD's) that, through their associated co-repressors, may favor the assembly of a highly compacted chromatin structure (lower panel). Such a model is based on the fact that we have found, overlapping CpG-rich CTCF binding sequences, methylated CpG's that facilitate the abnormal association of a methyl-CpG-binding protein.

between methylated and unmethylated state in the *BRCA1* promoter strongly supports the model that we envision for the CTCF site identified at the human *Rb* promoter. We think that CTCF binding creates a protective barrier against the spreading of DNA methylation over the core promoter sequences of *BRAC1* gene. Consequently, one prediction would be that the interference on CTCF binding induces abnormal DNA methylation and epigenetic silencing of the human *BRCA1* gene expression.

The scenarios described here are in agreement with the proposal that abnormal DNA methylation may initiate far upstream from promoter regions on "methylation centres" that could correspond to repetitive sequences, from where DNA methylation spread and invade normally unmethylated CpG-islands (Fig. 3) [24]. To avoid such abnormal invasion the presence of some kind of barrier element that would stop the methylation spreading, maintaining the normal methylation-free status of CpG-islands like those regulating tumour suppressor genes, has been visualized [24].

Then, based on what we have discussed in the present manuscript, we propose that CTCF is an appealing candidate to protect tumour suppressor gene promoters and introns against DNA methylation, possibly in collaboration with other factors (Fig. 5), and that could represent a key epigenetic factor in carcinogenesis.

Conclusions and prospects

Modern biology is immersed in accelerated and dynamic changes in terms of understanding cell physiology and the very complex mechanisms of neoplasic transformation. The genetic view is clearly at the center of molecular medicine in the study of many normal and abnormal processes leading to disease but in the last few years it has become apparent that this geocentric view is insufficient. Side by side with genetics, epigenetics has emerged as a critical field of research to understand a growing list of normal molecular processes and the pathogenesis of human diseases. Epigenetics is rapidly evolving and new concepts are emerging such as the "epigenetic phenotype" that predicts that epigenetic changes start at stem or progenitor cells followed by secondary genetic and epigenetic disorders [18, 21, 24]. Due to the ever increasing information on epigenetic phenomena new discoveries reveal that this field is much more intricate than anticipated and that the future challenge will be to integrate all such information, through bioinformatics in interactomes to understand the fine regulation of gene expression in health and disease.

In the present review we focused our analysis the multivalent 11-zinc-finger CCCTC-binding factor, CTCF, as a novel epigenetic regulator. We discussed a series of findings from the literature that seem to demonstrate that CTCF participates to different extents in a varied spectrum of epigenetic processes, such as gene regulation, imprinting control, chromatin insulation, epigenome topology, tumour suppressor gene regulation and others. But the main aspect discussed here, with direct consequences on epigenetic origin and development of human cancer is the proposed capacity of CTCF to protect against DNA methylation in different con-

texts. The available data do not allow us to propose a single mechanism for such a protection. Even more, what seems more likely is that the mechanisms through which CTCF protects against DNA methylation are varied and probably reflect responses to distinct epigenome contexts. Of particular interest is the proposal that CTCF may contribute to keep CpG-islands, from tumour suppressor gene promoters, on an unmethylated state. We need to recall that all the evidences supporting such a model are indirect and that no definitive demonstration has been done. In any case, the basic concept has to do with the capacity of CTCF and associated chromatin remodelers to act as a barrier and block the spreading of epigenetic silencing effects of the surrounding chromatin over a CpG-island [24]. It is important to mention that probably such CTCF protection is not only directed against DNA methylation, since other epigenetic processes could be counteracted by CTCF, like histone modifications or even the expression of intergenic transcripts. An alternative possibility is the one proposed by Antequera in which the "boundary" of a CpG-island could be defined by a DNA replication origin and that such origin and their chromatin associated modifications contribute to maintain a CpG-island hypomethylated [86]. Complementary to those models, we have noticed that for some CpG-islands and their non-coding genomic milieu, like for the human retinoblastoma 5' non-coding region, there is a clear gradient on CpG-dinucleotides reaching their highest density over functional gene promoters or first introns (Fig. 4). We think that such CpG-distribution is not random and may hide some functional properties that contribute to define and protect a CpG-island against abnormal DNA methylation. This model does not seem so unlikely in terms of CTCF activity since, as we have discussed earlier, CTCF demarcates two CpG-rich zones, maintaining the promoter of the *BRCA1* tumour suppressor gene methylation-free [84]. Our results are in agreement with our proposed model since CTCF location delimits on its 5'-side the human *Rb* gene promoter (De La Rosa-Velázquez et al. *submitted*).

In view of the epigenetic role for CTCF that we support it appears reasonable to ask how general is the function of CTCF on epigenetics? At this point it is hard to imagine that all the responsibility is given to CTCF. In our opinion CTCF could be the primary member of a novel family or class of nuclear factors participating on epigenetic regulation, in other words CTCF may represent the tip of the iceberg. Therefore, other factors remain to be discovered that could contribute to the epigenetic regulation, maybe at distinct levels and with different specificities.

To conclude we can say that once the full information of the progenitor-epigenetic and epigenetic and genetic mechanisms become integrated we will probably be allowed to better understand cell physiology and cancer, accurately classify tumours, early detect and diagnose neoplasic malignancies and perhaps develop novel therapies.

Acknowledgements

We thank Paul Delgado-Olguín for constant discussions and critical reading of the manuscript. This work was supported by the Dirección General de Asuntos del Personal Académico-UNAM (IN203200 and IN209403), the Consejo Nacional de Ciencia y Tecnología, CONACyT (Grants 38168-M, 33863-N and 42653-Q), the Third World Academy of Sciences (TWAS, Grant 01-055 RG/BIO/LA), and Fundación Miguel Alemán. IADR-V is supported with a Ph.D. fellowship from IMSS and CONACyT and ES-R is supported with a Ph.D. fellowship from CONACyT.

References

1. **Egger G, Liang G, Aparicio A, Jones PA**. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004; 429: 457–63.
2. **Lund AH, van Lohuizen M.** Epigenetics and cancer. *Genes Dev.* 2004; 18: 2315–35.
3. **Esteller M.** Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. *Br J Cancer* 2006; 94: 179–83.
4. **Bird A.** DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002; 10: 6–21.
5. **Jaenisch R, Bird A.** Epigenetic regulation of gene expression: how the genome integrates instrinsic and environmental signals. *Nat. Genet.* 2003; 33: 245–54.
6. **Narita M, Nunez S, Heard E, Narita M, Lin AW, Hearn SA, Spector DL, Hannon GJ, Lowe SW**. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* 2003; 113: 703–16.
7. **Santos F, Zakhartchenko V, Stojkovic M, Peters A, Jenuwein T, Wolf E, Reik W, Dean W.** Epigenetic mark-

ing correlates with development potential in cloned bovine preimplantation embryos. *Curr Biol.* 2003; 13: 1116–21.
8. **Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE, Eguia R, Dean DC, Esteller M, Jenuwein T, Blasco MA.** Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat Cell Biol.* 2005; 7: 420–8.
9. **Goll MG, Bestor TH.** Eukaryotic cytosine methyltransferases. *Annu Rev Biochem.* 2005; 74: 481–514.
10. **Wolffe AP, Matzke MA.** Epigenetics: regulation through repression. *Science* 1999; 286: 481–6.
11. **Li E.** Chromatin modifications and epigenetic reprogramming in mammalian development. *Nat Rev Genet.* 2002; 3: 662–73.
12. **Esteller M.** Dormant hypermethylated tumour suppressor genes: questions and answers. *J Pathol.* 2005; 205: 172–80.
13. **Recillas-Targa F.** DNA methylation, chromatin boundaries and mechanisms of genomic imprinting. *Arch Med Res.* 2002; 33: 428–38.
14. **Herman JG, Baylin SB.** Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med.* 2003; 349: 2042–54.
15. **Antequera F.** Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci.* 2003; 60: 1647–58.
16. **Grady WM, Willis J, Guilford PJ, Dundier AK, Toro TT, Lynch H, Wiesner G, Ferguson K, Eng C, Park JG, Kim SJ, Markowitz S.** Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. *Nat Genet.* 2000; 26: 16–7.
17. **Chen W, Cooper TK, Zahnow CA, Overhltser M, Zhao Z, Ladanyi M, Karp JE, Gokgoz N, Wunder JS, Andrulis IL, Levine AJ, Mankowski JB, Baylin SB.** Epigenetic and genetic loss of Hic function accentuates the role of p53 in tumorigenesis. *Cancer Cell* 2004; 6: 387–98.
18. **Baylin SB, Ohm JE.** Epigenetic gene silencing in cancer- a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer* 2006; 6: 107–16.
19. **Feinberg AP, Vogelstein B.** Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 1983; 301: 89–92.
20. **Goelz SE, Vogelstein B, Hamilton SR, Feinberg AP.** Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science* 1985; 228: 187–90.
21. **Feinberg AP, Ohlsson R, Henikoff S.** The epigenetic progenitor origin of human cancer. *Nat Rev Genet.* 2006; 7: 21–33.
22. **Tilghman SM.** The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell* 1999; 96: 185–93.
23. **Reik W, Murrell A, Lewis A, Mitsuya K, Umlauf D, Dean W, Higgins M, Feil R.** Chromosome loops, insulators and histone methylation: new insights into regulation of imprinting in clusters. *Cold Spring Harbor Symposia on Quantitative Biology*. 2004; LXIX: 29–37.
24. **Issa J-P.** CpG island methylator phenotype in cancer. *Nat Rev Cancer* 2004; 4: 988–93.
25. **Hake SB, Xiao A, Allis CD.** Linking the epigenetic "language" of covalent histone modifications to cancer. *Br J Cancer* 2004; 90: 761–9.
26. **Orphanides G, Reinberg D.** A unified theory of gene expression. *Cell* 2002; 108: 439–51.
27. **Hahn S.** Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol.* 2004; 11: 394–403.
28. **Ohlsson R, Renkawitz R, Lobanenkov VV.** CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.* 2001; 17: 520–7.
29. **Klenova E, Ohlsson R.** Poly(ADP-ribosyl)ation and epigenetics. Is CTCF part of the plot? *Cell Cycle* 2005; 4: 96–101.
30. **El-Kady A, Klenova E.** Regulation of the transcription factor, CTCF by phosphorylation with protein kinase CK2. *FEBS Lett.* 2005; 579: 1424–34.
31. **Yu W, Ginjala V, Pant V, Chernukhin I, Whitehead J, Docquier F, Farrar D, Tavoosidana G, Mukhopadhyay R, Kanduri C, Oshimura M, Feinberg AP, Lobanenkov V, Klenova E, Ohlsson R.** Poly(ADP-ribosyl)ation regulated CTCF-dependent chromatin insulation. *Nat Genet.* 2004; 36: 1105-1110.
32. **Vatolin S, Abdullaev Z, Pack SD, Flanagan PT, Custer M, Loukinov DI, Pugacheva E, Hong JA, Morse H 3rd, Schrump DS, Risinger JI, Barrett JC, Lobanenkov VV.** Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. *Cancer Res.* 2005; 65: 7751–62.
33. **Hong JA, Kang Y, Abdullaev Z, Flanagan PT, Pack SD, Fischette MR.** Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Res.* 2005; 65: 7763–74.
34. **Lewis A, Murrell A.** Genomic imprinting: CTCF protects the boundaries. *Curr Biol.* 2004; 14: R284–6.
35. **Yoon B, Herman H, Hu B, Park YJ, Lindroth A, Bell A, West AG, Chang Y, Stablewski A, Piel JC, Loukinov DI, Lobanenkov VV, Soloway PD.** Rasgrf1 imprinting is regulated by a CTCF-dependent methylation-sensitive enhancer blocker. *Mol Cell Biol.* 2005; 25: 11184–90.
36. **Rosa AL, Wu YQ, Kwabi-Addo B, Coveler KJ, Reid Sutton V, Shaffer LG.** Allele-specific methylation of a functional CTCF binding site upstream of MEG3 in the human imprinted domain of 14q32. *Chromosome Res.* 2005; 13: 809–18.
37. **Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM.** CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* 2000; 405: 486–9.
38. **Bell AC, Felsenfeld G.** Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 2000; 405: 482–5.
39. **Fedoriw AM, Stein P, Svoboda P, Schultz RM, Bartolomei MS.** Transgenic RNAi reveals essential function for CTCF in *H19* gene imprinting. *Science* 2004; 303: 238–40.
40. **Klenova EM, Morse HC 3rd, Ohlsson R, Lobanenkov VV.** The novel *BORIS+CTCF* gene family is uniquely involved in the epigenetics of normal biology and cancer. *Semin Cancer Biol.* 2002; 12: 399–414.
41. **Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, Chernukhin I, Mannan P, Larsson E, Kanduri C,**

Vostrov AA, Cui H, Niemitz EL, Rasko JE, Docquier FM, Kistler M, Breen JJ, Zhuang Z, Quitschke WW, Renkawitz R, Klenova EM, Feinberg AP, Ohlsson R, Morse HC 3rd, Lobanenkov VV. BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci USA.* 2002; 99: 6806–11.

42. **Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH.** Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol.* 2003; 13: 1192–220.
43. **Vire E, Brenner C, Deplus R, Blanchon L, Frga M, Didelot C, Morey L, van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, DiCroce L, deLamoit Y, Fuks F.** The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 2005; 38: 566–9.
44. **Brown CJ, Greally JM.** A stain upon the silence: genes escaping X inactivation. *Trends Genet.* 2003; 19: 432–8.
45. **Filippova GN, Cheng MK, Moore JM, Truong J-P, Hu YJ, Nguyen DK, Tsuchiya KD, Disteche CM.** Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Develop Cell* 2005; 8: 31–42.
46. **Sado T, Okano M, Li E, Sasaki H.** *De novo* DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* 2004; 131: 975–82.
47. **Heard E.** Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr Opin Genet Develop.* 2005; 15: 482–9.
48. **Machuca-Tzili L, Brook D, Hilton-Jones D.** Clinical and molecular aspects of the myotonic dystrophies. *Muscle Nerve* 2005; 32: 1–18.
49. **Gatchel JR, Zoghbi HY.** Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet.* 2005; 6: 743–55.
50. **Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, Klesert TR, Lobanenkov VV, Tapscott SJ.** CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet.* 2001; 28: 335–43.
51. **Otten AD, Tapscott SJ.** Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc Natl Acad Sci USA.* 1995; 92: 5465–9.
52. **Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ.** Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol Cell* 2005; 20, 483–9.
53. **West AG, Gaszner M, Felsenfeld G.** Insulators: many functions, many mechanisms. *Genes Dev.* 2002; 16: 271–88.
54. **Recillas-Targa F, Pikaart MJ, Burgess-Beusse B, Bell AC, Litt MD, West AG, Gaszner M, Felsenfeld G.** Position-effect protection and enhancer blocking by the chicken β-globin insulator are separable activities. *Proc Natl Acad Sci USA.* 2002; 99: 6883–8.
55. **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.* 1998; 12: 2852–62.
56. **Lutz M, Burke LJ, LeFevre P, Myers FA, Thorne AW, Crane-Robinson C, Bonifer C, Filippova GN, Lobanenkov V, Renkawitz R.** Thyroid hormone-regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor. *EMBO J.* 2003; 22: 1579–87.
57. **Farrell CM, West AG, Felsenfeld G.** Conserved CTCF insulator elements flank the mouse and human β-globin loci. *Mol Cell Biol.* 2002; 22: 3820–31.
58. **Valadez-Graham V, Razin SV, Recillas-Targa F.** CTCF-dependent enhancer blockers at the upstream region of the chicken α-globin gene domain. *Nucleic Acids Res.* 2004; 32: 1354–62.
59. **Klochkov D, Rincón-Arano H, Ioudinnkowa ES, Valadez-Graham V, Gavrilov A, Recillas-Targa F, Razin SV.** A CTCF-dependent silencer located in the differentially methylates area may regulate expression of a housekeeping gene overlapping a tissue-specific gene domain. *Mol Cell Biol.* 2006; 26: 1589–97.
60. **Burgess-Beusse B, Farrell C, Gaszner M, Litt M, Mutskov V, Recillas-Targa F, Simpson M, West A, Felsenfeld G.** The insulation of genes from external enhancers and silencing chromatin. *Proc Natl Acad Sci USA.* 2002; 99: 16433–7.
61. **Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G.** CTCF tethers and insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 2004; 13: 291–8.
62. **Bulger M, Schübeler D, Bender MA, Hamilton J, Farrell CM, Hardison RC, Groudine M.** A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse β-globin locus. *Mol Cell Biol.* 2003; 23: 5234–44.
63. **Tolhuis B, Palstra RJ, Splinter E, Grosveld F, deLaat W.** Looping and interaction between hypersensitive sites in the active β-globin locus. *Mol Cell* 2002; 10: 1453–65.
64. **Chakalova L, Debrand E, Mitchell JA, Osborne CS, Fraser P.** Replication and transcription: shaping the landscape of the genome. *Nat Rev Genet.* 2005; 6: 669–77.
65. **Paltra R-J, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F. deLaat W.** The β-globin nuclear compartment in development and erythroid differentiation. *Nat Genet.* 2003; 35: 190–4.
66. **Gerasimova TI, Byrd K. Corces VG.** A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* 2000; 6: 1025–35.
67. **Capelson M, Corces VG.** Boundary elements and nuclear organization. *Biol Cell.* 2004; 96: 617–29.
68. **Mukhopadhyay R, Yu W, Whitehead J, Xu J, Lezcano M, Pack S, Kanduri C, Kanduri M, Ginjala V, Vostrov A, Quitschke W, Chernukhin I, Klenova E, Lobanenkov V, Ohlsson R.** The binding site for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res.* 2004; 14: 1594–602.

69. **Dunn KL, Zhao H, Devie JR.** The insulator binding protein CTCF associated with the nuclear matrix. *Exp Cell Res.* 2003; 288: 218–23.
70. **Yusufzai TM, Felsenfeld G.** The 5'-HS4 chicken β-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci USA.* 2004; 101: 8620–4.
71. **Murrell A, Heeson S, Reik W.** Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet.* 2004; 36: 889–93.
72. **Spilianakis CG, Laliotis MD, Twon T, Lee GR, Flavell RA.** Interchromosomal associations between alternatively expressed loci. *Nature* 2005; 435: 637–45.
73. **Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qui XW, Cherry AM, Hoffman AR.** CTCF mediates interchromosomal colocalization between *Igf2*/*H19* and *Wsb1*/*Nf1*. *Science* 2006; 312: 269–72.
74. **Cremer T, Cremer C.** Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet.* 2001; 2: 292–301.
75. **Rasko JE, Klenova EM, Leon J, Filippova GN, Loukinov DI, Vatolin S, Robinson AF, Hu YJ, Ulmer J, Ward MD, Pugacheva EM, Neiman PE, Morse HC 3rd, Collins SJ, Lobanenkov VV.** Cell growth inhibition by the multifunctional multivalent zinc-finger factor CTCF. *Cancer Res.* 2001; 61: 6002–7.
76. **Torrano V, Chernukhin I, Docquier F, D'Arcy V, Leon J, Klenova E, Delgado MD.** CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *J Biol Chem.* 2005; 280: 28152–61.
77. **Docquier F, Farrar D, D'Arcy V, Chernukhin I, Robinson AF, Loukinov D, Vatolin S, Pack S, Mackay A, Harris RAN, Dorricott H, O'Hare MJ, Lobanenkov V, Klenova E.** Heightened expression of CTCF in breast cancer cells is associated with resistance to apoptosis. *Cancer Res.* 2005; 65: 5112–22.
78. **Filippova GN, Lindblom A, Meincke LJ, Klenova EM, Neiman PE, Collisn SJ, Doggett NA, Lobanenkov V.** A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. *Genes Chromosomes Cancer* 1998; 22: 26–36.
79. **Yeh A, Wei M, Golub SB, Yamashiro DJ, Murty VV, Tycko B.** Chromosome arm 16q in Wilms tumors: unbalanced chromosomal translocations, loss of heterozygosity, and assessment of the CTCF gene. *Genes Chromosomes Cancer* 2002; 35: 156–63.
80. **Cui H, Niemitz EL, Ravenel JD, Onyango P, Brandenburg SA, Lobanenkov VV, Feinberg AP.** Loss of imprinting of insulin-like growth factor-II in Wilm's tumor commonly involves altered methylation but not mutations of CTCF or its binding site. *Cancer Res.* 2001; 61: 4947–50.
81. **Aulmann S, Bläker H, Penzel R, Rieker RJ, Otto HF, Sinn HP.** CTCF gene mutations in invasive ductal breast cancer. *Breast Cancer Res.* 2003; 80: 347–52.
82. **Zhou X-L, Werelius B, Lindblom A.** A screen for germ line mutations in the gene encoding CCCTC-binding factor (CTCF) in familial non-*BRCA1*/*BRCA2* breast cancer. *Breast Cancer Res.* 2004; 6: R187–90.
83. **Rakha EA, Pinder SE, Paish CE, Ellis IO.** Expression of the transcription factor CTCF in invasive breast cancer: a candidate gene located at 16q22.1. *Br J Cancer* 2004; 91: 1591–6.
84. **Butcher DT, Mancini-DiNarfo DN, Archer TK, Rodenhiser DI.** DNA binding sites for putative methylation boundaries in the unmethylated region of the *BRCA1* promoter. *Int J Cancer* 2004; 111: 669–78.
85. **Dobrovic A, Simpfendorfer D.** Methylation of the *BRCA1* gene in sporadic breast cancer. *Cancer Res.* 1997; 57: 3347–50.
86. **Antequera F.** Genomic specification and epigenetic regulation of eukaryotic DNA replication origins. *EMBO J.* 2004; 23: 4365–70.
87. **Sanchez-Elsner T, Gou D, Kremmer E, Sauer F.** Noncoding RNAs of Trithorax response elements recruit *Drosophila* Ash1 to Ultrabithorax. *Science* 2006; 311: 1118–23.
88. **Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G.** RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* 2006; 124: 957–71.