

SURVEY AND SUMMARY

CTCF as a regulator of alternative splicing: new tricks for an old player

Adel B. Alharbi ^{1,2,3,4,5}, Ulf Schmitz ^{1,3,4}, Charles G. Bailey ^{1,4,5,*} and John E. J. Rasko ^{1,4,6,*}

¹Gene & Stem Cell Therapy Program Centenary Institute, The University of Sydney, Camperdown, NSW 2050, Australia, ²Department of Laboratory Medicine, Faculty of Applied Medical Sciences, Umm Al-Qura University, Makkah, Saudi Arabia, ³Computational BioMedicine Laboratory Centenary Institute, The University of Sydney, Camperdown, NSW 2050, Australia, ⁴Faculty of Medicine & Health, The University of Sydney, NSW 2006, Australia, ⁵Cancer & Gene Regulation Laboratory Centenary Institute, The University of Sydney, Camperdown, NSW 2050, Australia and ⁶Cell & Molecular Therapies, Royal Prince Alfred Hospital, Camperdown, NSW 2050, Australia

Received January 13, 2021; Revised June 01, 2021; Editorial Decision June 02, 2021; Accepted June 10, 2021

ABSTRACT

Three decades of research have established the CCCTC-binding factor (CTCF) as a ubiquitously expressed chromatin organizing factor and master regulator of gene expression. A new role for CTCF as a regulator of alternative splicing (AS) has now emerged. CTCF has been directly and indirectly linked to the modulation of AS at the individual transcript and at the transcriptome-wide level. The emerging role of CTCF-mediated regulation of AS involves diverse mechanisms; including transcriptional elongation, DNA methylation, chromatin architecture, histone modifications, and regulation of splicing factor expression and assembly. CTCF thereby appears to not only co-ordinate gene expression regulation but contributes to the modulation of transcriptomic complexity. In this review, we highlight previous discoveries regarding the role of CTCF in AS. In addition, we summarize detailed mechanisms by which CTCF mediates AS regulation. We propose opportunities for further research designed to examine the possible fate of CTCF-mediated alternatively spliced genes and associated biological consequences. CTCF has been widely acknowledged as the ‘master weaver of the genome’. Given its multiple connections, further characterization of CTCF’s emerging role in splicing regulation might extend

its functional repertoire towards a ‘conductor of the splicing orchestra’.

INTRODUCTION

CTCF – an old player

CCCTC-binding factor (CTCF) was first identified as a nuclear DNA-binding protein, which negatively regulates chicken *c-myc* expression by interacting with three regularly spaced repeats of the CCCTC DNA motif (1). Since then, considerable attention has been focused on CTCF leading to pivotal discoveries related to its structure, binding activities and functions. CTCF is a multivalent 11-zinc finger DNA-binding protein that is ubiquitously expressed in most tissues of vertebrate species (1–4). Combinatorial usage of CTCF zinc fingers (4,5) allows its binding to tens of thousands of conserved DNA sites as revealed by chromatin immunoprecipitation sequencing (ChIP-seq) experiments (6–8). More than 30% and 50% of these binding sites are located in intronic and intergenic regions, respectively. These sites are mostly ubiquitous; however, cell-type-specific CTCF binding is observed (6–8). Ubiquitous CTCF target sites are predominantly localized within intergenic regions and are highly conserved compared to cell-type-specific ones, which are mostly located within introns (8).

CTCF has also been characterized as an RNA-binding protein and these interactions are essential for facilitating CTCF-mediated chromatin architecture (9–13). Such interactions are involved in some CTCF activities including CTCF dimerization, distal genomic binding sites, chromatin looping and gene regulation (9–12). Indeed, locus- or

*To whom correspondence should be addressed. Tel: +61 2 95656116; Fax: +61 2 95656101; Email: j.rasko@centenary.org.au
Correspondence may also be addressed to Charles G. Bailey. Email: c.bailey@centenary.org.au

transcript-specific functions for CTCF-RNA interactions have been clearly demonstrated (9,13,14). CTCF can also undergo multiple post-translational modifications which could affect interactions with binding partners. These interactors include transcription factors, chromatin remodelers, methylation regulators, histone modifiers and splicing factors (reviewed in (15–17)).

The complexity of CTCF interactions is further expanded by the ability of CTCF to homodimerize and mediate contacts with multiple distal intra- and inter-chromosomal target sites across the genome. Such interactions can facilitate genome folding into topologically associating domains (TADs) and chromatin loops, encompassing those interaction domains (18–21). This allows CTCF to coordinate chromatin and genome architecture, which reinforce CTCF's eminence as 'the master weaver of the genome' (22). Chromatin looping occurs when a DNA-bound CTCF molecule interacts with another CTCF molecule located at a proximal or distal genomic site (21,23). These CTCF–CTCF interactions mostly initiate chromatin loops when the CTCF-bound sites are in a convergent (i.e. forward and reverse) orientation (23,24) rather than divergent orientation, which has been mostly observed at non-looped TAD boundaries (20,24). CTCF also coordinates chromatin architecture in concert with the cohesin complex to help define and stabilize chromatin loops. While CTCF associates with over half of the cohesin binding sites genome-wide, both CTCF and cohesin also have their own specific binding sites and independent roles in coordinating chromatin organization (25–28).

Given its unique role in higher-order chromatin architecture and the multiplicity and specificity of its binding sites, CTCF is able to regulate diverse molecular and epigenetic functions, summarized herein. CTCF is a unique regulator of gene expression, which can function as a transcriptional repressor and activator as well as a chromatin insulator, the latter occurring through interference of the contact between enhancers/silencers and promoters (5,22). Furthermore, it epigenetically regulates gene expression through gene imprinting, X-chromosome inactivation, and by preserving methylation-free regions throughout the genome (15,29). Moreover, CTCF-mediated chromatin looping modulates gene expression by bringing genomic loci into spatial proximity (30–32).

Compelling evidence has linked CTCF to modulation of alternative splicing (AS) at both the individual transcript and transcriptome-wide level (33–46). AS is a complex biological process, which affects over 95% of human multi-exonic genes and enriches transcriptome and protein diversity by facilitating the production of multiple mRNA and protein isoforms from individual genes (47,48). There are four major forms of AS including exon skipping, mutually exclusive exons, alternative use of 5' or 3' splice sites and intron retention. With more than 40% prevalence, the most common form of AS is exon skipping, which involves the exclusion of one or more exons from the mature mRNA transcript. Splicing can occur at a site alternative to the canonical splice site, depending on whether the new splice site replaces the canonical donor or acceptor site. Mutually exclusive exons are exons from the same gene that can be differentially included in the spliced gene transcript, i.e. they are

rarely found together in the same mature mRNA. Finally, intron retention is a process by which an intron is not excised from the pre-mRNA and therefore retained in the mature mRNA transcript, often leading to nonsense-mediated mRNA decay (47,48).

CTCF's involvement with AS, identified over the past decade, has expanded the list of mechanisms by which CTCF can regulate gene expression and transcriptomic complexity (33–46); however, its direct impact on biological functions has not been fully elucidated. Herein, we highlight the CTCF-mediated mechanisms linked to AS regulation. While the majority of these mechanisms have not been fully characterized, we emphasize the extent to which CTCF has been experimentally verified as a regulator of AS. Furthermore, we track the putative roles of CTCF in modulating AS, placing them in context of the regulatory mechanisms governing the transcriptional and splicing machinery. This review provides a roadmap toward understanding and further studying the under-recognized role of CTCF in AS regulation.

NEW TRICKS: CTCF IN ALTERNATIVE SPLICING REGULATION

It is widely accepted that AS re-configures the transcriptome and proteome to facilitate certain biological processes. Perturbations to this highly calibrated process can cause diseases such as Hutchinson-Gilford Progeria syndrome, Duchenne muscular dystrophy, melanoma and breast cancer (49,50). Thus, maintaining tight control of AS requires a complex regulatory network integrating co-transcriptional, genomic and epigenetic factors such as transcriptional elongation, DNA methylation, chromatin architecture, histone modifications, and splicing factor expression and assembly (51,52). Over the past decade, CTCF has been characterized as a direct and indirect modulator of AS decisions. This function is achieved through various mechanisms involving multiple factors including co-transcriptional regulation, genomic features and epigenetic factors (33–46), as summarized (Figure 1) and discussed henceforth.

CO-TRANSCRIPTIONAL REGULATION

RNA polymerase II elongation and archetypal *CD45* locus

RNA polymerase II (RNAPII) elongation is fundamental to transcription and constitutive splicing processes, which require an optimal RNAPII elongation rate for producing canonical mature mRNA transcripts. This means that slowing or accelerating transcription elongation can both inhibit or enhance splice site recognition depending on additional factors. Alteration of the elongation rate influences inclusion of exons or introns particularly those having weak splice sites and short flanking introns (53–56). In a pivotal discovery, Shukla *et al.* revealed the first direct link 'road-block model' between CTCF and AS regulation involving RNAPII elongation and methylation of CTCF binding sites. Binding of CTCF to *CD45* exon 5 impedes RNAPII elongation which allows the spliceosome to assemble at the weak upstream splice site and eventually leads to the inclusion of exon 5 in human Burkitt lymphoma B cells (33) (Figure 2A). AS of *CD45* exons 4, 5 and 6 leads to the generation

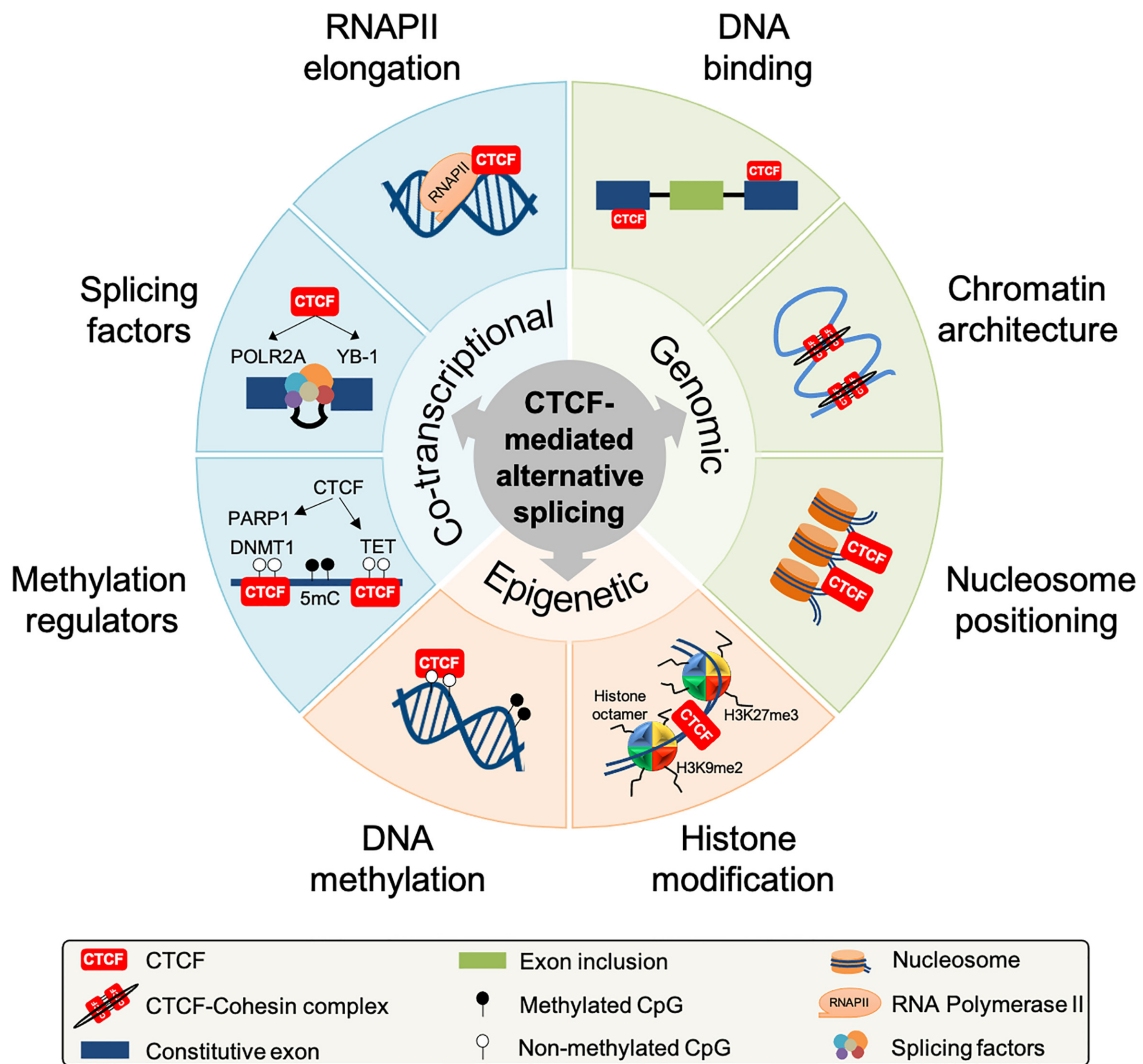


Figure 1. Mechanisms of CTCF-mediated regulation of alternative splicing. CTCF has been linked to key determinants of AS regulation which can be broadly categorized into co-transcriptional, genomic and epigenetic mechanisms. During co-transcriptional regulation, CTCF-mediated AS is regulated by stalling RNAPII transcriptional elongation (33–35,42), controlling DNA methylation (60–62) and recruitment of splicing factors (42,58,66). While the co-transcriptional factors have been mostly experimentally verified, the genomic and epigenetic regulatory roles of CTCF in modulating AS are putative. These involve localization of CTCF binding sites proximal to splice sites (33–38,40–42,44), chromatin architecture (35,36), nucleosome enrichment (7,85–87), histone modification (35,46) and regulation of CTCF binding via DNA methylation patterns (33,34,36,38,40,41).

of multiple isoforms, which have been tightly linked to lymphocyte development (57). This putative new role for CTCF in the regulation of lymphocyte development demands further study.

The decision for CTCF-mediated *CD45* exon 5 inclusion is tightly regulated by the methylation status of a CTCF binding site in *cis*. CTCF depletion or loss of CTCF binding due to methylation within its cognate binding site supports faster RNAPII elongation and subsequently exon exclusion (33,34). While 5-methylcytosine (5mC) evicts CTCF and promotes exon exclusion, DNMT1 depletion decreases 5mC levels at exon 5 and promotes CTCF binding and exon inclusion (33). In addition, members of the TET (Ten-eleven translocation) family of methylcytosine dioxygenases (TET1 and TET2) oxidize 5mC to 5-hydroxymethylcytosine (5hmC) and its downstream oxidative intermediates. This

can subsequently promote CTCF binding and exon inclusion, which is reversed by depleting TET1 and TET2 expression (34).

In line with this context-dependent regulatory mechanism, a global correlation between CTCF, RNAPII occupancy and exon inclusion was detected when CTCF binding occurs within ~1 kb downstream of alternatively spliced exons in BJAB and BL41 cells (33). This has been independently supported by a stronger association between CTCF and RNAPII sites mostly downstream (~0.6 kb) of differentially included exons in MCF7 breast cancer cells compared to non-tumorigenic MCF10A mammary cells (46). Furthermore, there were global reciprocal exchanges of 5mC and 5hmC at CTCF binding sites downstream (~1.5 kb) of differentially alternatively spliced exons in naive and activated CD4⁺ T cells, where 5mC and 5hmC favor up-

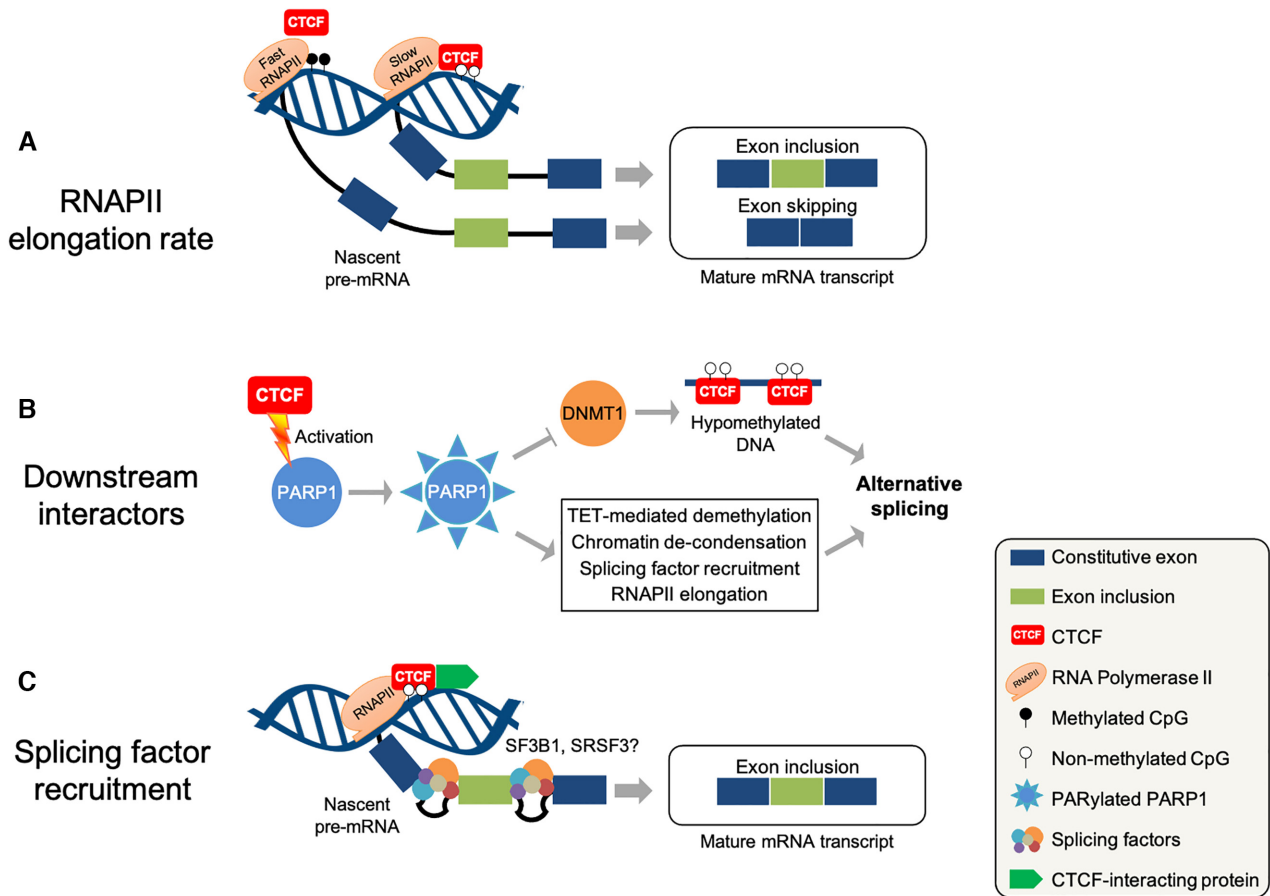


Figure 2. CTCF-mediated co-transcriptional regulation of alternative splicing. Schematic representation of co-transcriptional mechanisms by which CTCF modulates AS. (A) The ‘roadblock’ mechanism involves altering RNAPII elongation rate in a methylation-dependent manner (33,34). (B) CTCF regulation of DNA methylation via activation of PARP1 and PARylation (60,61), which are also involved in other AS-related regulatory activities (66–69). (C) Splicing factor recruitment might also take part in CTCF-mediated AS regulation through direct interaction with RNA binding proteins (42,58,70) or other transcription factors involved in AS such as PARP1 (66), MeCP2 (38), YB-1 (108,109) and HP1 α (46,124,127). SF3B1 and SRSF3 are examples of RNA-binding proteins.

stream exon exclusion and inclusion, respectively. Nevertheless, the highest enrichment of CTCF binding was observed upstream of alternatively spliced exons (34).

Beyond the roadblock model

The ‘roadblock’ model describes the phenomenon of CTCF-mediated RNAPII elongation stalling downstream of included exons (33). However, this role does not seem applicable to CTCF binding sites that are highly enriched upstream of alternatively spliced exons (34). In addition, it also does not fully explain how CTCF occupancy exerts its effects in an upstream or downstream direction. Moreover, the optimal distance between a CTCF binding site and the proximal alternatively spliced exon where RNAPII accumulation occurs remains elusive.

The association of CTCF binding with RNAPII stalling at specific sites has been previously observed genome-wide in multiple cell lines (58), particularly at CTCF sites located immediately downstream to promoters of specific genes (59) including the p53 target genes *PUMA* and *p21* (35). Within these promoters, a higher RNAPII pausing index was associated with more proximal CTCF binding sites (59). In-

terestingly, CTCF-mediated RNAPII stalling was also detected at sites where CTCF binds upstream to promoters (59), which is consistent with Marina *et al.* (34). Given that RNAPII elongation is sterically hindered by CTCF bound at downstream regions, alternative mechanism(s) regulating CTCF-mediated RNAPII pausing at upstream sites could be involved. For instance, chromatin architecture influences splicing decisions via modulation of RNAPII elongation rate and recruitment of splicing factors (reviewed in (56), more details provided in *GENOMIC FEATURES* section).

CTCF’s interplay with other factors

As a master regulator of gene expression, CTCF has the capacity to regulate the expression of numerous epigenetic modulators as well as genes involved in the transcriptional and splicing machinery. These include splicing factors, methylation regulators, chromatin remodelers, and histone modifiers (5,15,22,29). In regard to CTCF acting as a roadblock to AS of the archetypal *CD45* gene, CTCF has been directly linked to regulation of the three factors (RNAPII, DNMT-mediated methylation and TET-mediated demethylation) in different contexts (58,60–62).

For example, besides its role in pausing RNAPII elongation, CTCF recruits and activates RNAPII when CTCF directly interacts with its largest subunit, POLR2A, via its C-terminal domain (58). In addition, CTCF also binds to and activates poly(ADP-ribose) polymerase (PARP1), which inactivates DNA methyltransferase DNMT1 and thereby inhibits DNA methylation at CTCF binding sites (60,61). Moreover, CTCF physically interacts with active TET enzymes and promotes DNA demethylation at specific sites (62). The upstream regulatory effects of CTCF on these proteins could add further complexity to the ‘roadblock’ model (33,34), particularly in the context of their possible recruitment at a genome-wide level. Therefore, the consequences of this regulatory network on modulating AS demands further experimental verification.

The CTCF-PARP1 regulatory axis

An additional role for CTCF in modulating AS could be established from its known effect on activating PARP1 and its poly(ADP-ribosylation) (PARylation) activity (60,61). PARylation is a post-translational modification which is catalyzed by active PARP enzymes transferring ADP-ribose moieties from nicotinamide adenine dinucleotide (NAD⁺) to specific amino acid residues on the target substrate (63,64). It maintains various biological processes including genomic stability, chromatin architecture, transcriptional regulation, mitosis and cell death (64,65). While CTCF-regulated PARP1-PARylation activity has not been verified yet, inhibition of PARP1 expression or its PARylation activity has been linked to AS regulation. This has been demonstrated via PARP1-nucleosome interactions at target exon/intron boundaries or PARylation-mediated regulation of splicing factors including recruitment of spliceosomal factor 3B subunit 1 (SF3B1), a U2 snRNP (small nuclear ribonucleoprotein) spliceosomal member (66). A follow up study found that RNAPII elongation rate was also modulated at these sites depending on chromatin structure, which is disrupted upon PARP1 depletion (67). PARP1, in addition to its role as a PARylation activator, has been linked to regulation of transcription and AS via multiple mechanisms involving chromatin modulation, transcriptional modulation, post-transcriptional regulation of RNA-binding proteins and mRNA stabilization (reviewed in (68,69)). While this is true, possible roles for CTCF in the regulation of AS via activating PARP1 and PARylation will reveal new aspects of AS regulatory mechanisms, which require further investigation (Figure 2B).

Given its direct effect on PARP1 and its PARylation activity (60,61), CTCF might possibly regulate AS via RNA-binding proteins, whose binding, recruitment and expression are directly regulated by PARP1 and/or PARylation (68) (Figure 2B and C). This could be tested by examining the effect of CTCF on PARP1-mediated recruitment of SF3B1 and subsequently pre-mRNA splicing (66). Proteomics analysis of CTCF-interacting partners in MCF10A cells identified RNA-binding proteins including snRNPs, heterogeneous nuclear ribonucleoproteins (hnRNPs) and serine-arginine proteins (70), which are essential components of the splicing machinery (51). In addition, CTCF can transcriptionally regulate or directly interact with mul-

multiple splicing factors. For instance, the interaction between CTCF and POLR2A activates RNAPII (58), which provides a platform for factors regulating transcription initiation, elongation and termination as well as RNA processing (71). CTCF was also found to accumulate at specific RNAPII termination sites on protein-coding and small nuclear RNA (snRNA) genes (72).

A study focused on CTCF-RNAPII dynamics showed that CTCF regulated RNAPII elongation at the early elongation checkpoint of *c-myc*, and RNAPII termination at *U2* snRNA gene via regulating the recruitment or activation of key factors, such as DRB-sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) (42). Moreover, CTCF interacts with and recruits positive transcription elongation factor-b (P-TEFb) which stimulates transcriptional elongation by phosphorylating POLR2A, NELF and DSIF (42). The working model at the *c-myc* site depends on CTCF binding downstream of the *c-myc* promoter and recruitment of DSIF to stall RNAPII elongation followed by P-TEFb recruitment which elicits continuing RNAPII elongation and RNA processing (42). Whether CTCF modulates AS at this and other genomic sites via this working model still awaits investigation. Together, the direct effect of CTCF on RNAPII elongation, PARP1/PARylation activity and splicing factor recruitment support the notion that CTCF can modulate AS via co-transcriptional regulation. Characterization of more splicing factors, methylation regulators and other AS factors directly or indirectly regulated by CTCF will contribute to our understanding of the complex mechanisms by which CTCF modulates AS.

GENOMIC FEATURES

A regulatory framework for splicing provided by CTCF-mediated chromatin architecture

Intragenic CTCF binding sites, particularly those proximal to splice junctions, influence pre-mRNA splicing decisions, and thus mediate alternative exon or intron inclusion (33–38,40–42,44). Chromatin looping is a genome-wide function of CTCF placing it at the center of 3D chromatin architecture. Moreover, it brings genomic loci into spatial proximity allowing CTCF to modulate gene expression (30–32). The analysis of ChIP-seq, RNA-seq and high-throughput chromosome conformation capture (Hi-C) data have revealed the widespread formation of CTCF-mediated chromatin loops between promoters and intragenic regions (44). Interestingly, chromatin loops formed upstream of exons induced their inclusion by bringing these alternatively spliced exons into physical proximity of the gene promoter (Figure 3A). This particular study provides a putative genome-wide link between intragenic CTCF-mediated chromatin looping and AS regulation (44).

Such a phenomenon affects members of the protocadherin *Pcdh* gene family, which encode multiple isoforms responsible for neuronal cell surface diversity (36,37). CTCF in conjunction with the cohesin complex mediates DNA looping between *Pcdh* enhancers and active isoform-specific promoters which induces alternatively spliced isoforms in a methylation-dependent manner (36,37). Mutation or methylation of CTCF sites at the promoter proxi-

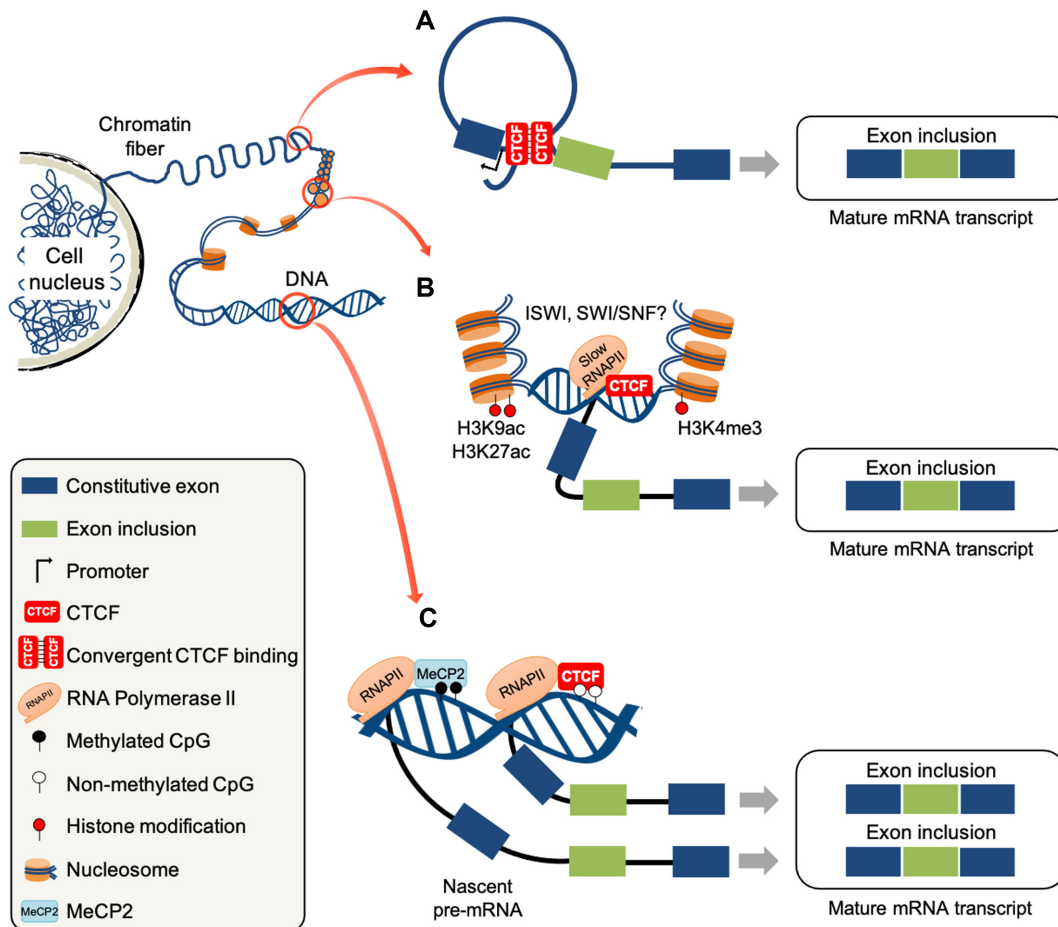


Figure 3. CTCF-mediated genomic and epigenetic regulation of alternative splicing. Schematic representation of putative genomic roles for CTCF in AS. (A) Formation of chromatin loops to bring alternatively spliced exons into physical proximity of the gene promoter (44). (B) Nucleosome occupancy-mediated altering of RNAPII elongation rate. (C) DNA methylation differentially regulates AS by CTCF and MeCP2. Unmethylated CpGs within CTCF sites promote exon inclusion via CTCF binding while DNA methylation favors exon inclusion via MeCP2 binding (38).

mal regions abolishes CTCF binding *in vitro*. Furthermore, CTCF knockdown leads to a reduction in CTCF binding as well as the cohesin complex subunits RAD21 and SMC1 recruitment/binding and subsequently chromatin looping-mediated enhancer/promoter communications (36). The latter was also observed by RAD21 knockdown, which showed no effect on CTCF binding (36,37). Despite the direct link between CTCF/cohesin-mediated DNA looping and modulation of *Pcdh* isoforms, the exact underlying AS regulatory mechanism and downstream physiological consequences are yet to be elucidated.

RNAPII elongation is prone to stalling at specific CTCF/cohesin sites (35,40,73). This leaves open the possibility that CTCF might regulate AS via integrating chromatin looping and co-transcriptional factors. While CTCF-mediated RNAPII stalling has been verified (33–35,42), cohesin has been found to be recruited to sites where RNAPII pausing occurs, particularly when associated with CTCF (40,74). The CTCF/cohesin complex downstream of the *PUMA* and *p21* promoters has been shown to stall RNAPII elongation and regulate their expression (35). Similar to CTCF-mediated *Pcdh* isoform regulation (36,37), knock-down of CTCF abrogates CTCF binding at the *PUMA* and

p21 promoters and reduces cohesin complex (RAD21 and SMC1) levels (35). Likewise, CTCF-mediated cohesin recruitment and chromatin looping hinders RNAPII elongation downstream of proximal poly(A) sites in a methylation-dependent manner. This links CTCF to pre-mRNA alternative polyadenylation (40), a process allowing usage of alternative poly(A) sites to produce multiple mRNA transcripts from an individual gene (reviewed in (75)). Knowing that CTCF recruits cohesin to sites where RNAPII pausing occurs (35,40,73), it is possible that CTCF-mediated RNAPII elongation stalling induces cohesin recruitment and thereby chromatin looping, which leads to AS. Indeed, these findings highlight a new role of CTCF-mediated 3D chromatin organization in AS regulation. Conversely, CTCF-mediated AS might also itself be involved in shaping the genome.

A role for CTCF-anchored nucleosomal occupancy

Nucleosome positioning is known to modulate RNAPII kinetics by acting as a physical barrier hindering RNAPII elongation and thus promoting the recognition of exon-intron boundaries by the splicing machinery (76). Nucleosomes are preferentially positioned in exons rather than

introns or alternatively spliced exons, highlighting a role for nucleosome positioning in exon definition (77–80). Furthermore, nucleosome occupancy has been correlated with regulating AS, specifically exon inclusion (78–80) and intron retention (81). Exon-associated nucleosomes are enriched in active H3K36me3 and repressive H3K27me2 histone modifications, which help demarcate exon-intron boundaries (77,82). H3K36me3 is particularly enriched at actively transcribed genes (82) and associated with different splicing decisions involving recruitment of splicing regulators in a context-dependent manner (83,84).

Interestingly, several studies have reported an enrichment of CTCF sites around nucleosomes genome-wide, indicating that CTCF acts as an anchor for positioning nucleosomes in a symmetrical distribution (7,85–87). Yet, it is still unknown whether CTCF binds proximal to well-positioned nucleosomes or whether these nucleosomes are positioned depending on CTCF binding. SNF2H, the ATPase subunit of the ISWI chromatin remodeler, was found to be essential for global organization of nucleosomes and TAD formation via promoting CTCF binding. Knockdown of SNF2H causes a genome-wide reduction in CTCF binding but without affecting its abundance (88). Furthermore, CTCF was identified as an interacting partner of BRG1, the ATPase subunit of the chromatin remodeling complex SWI/SNF (89). While BRG1 depletion has no effect on genomic CTCF binding (88), it results in reduced nucleosome occupancy around CTCF sites (90). Moreover, knowing that PARylation mediates chromatin decondensation (65), CTCF-activated PARylation can possibly alter RNAPII dynamics via chromatin relaxation (Figure 2B). Taken together, it would be interesting to examine how CTCF and nucleosome positioning alters RNAPII elongation and subsequently AS (Figure 3B). Nucleosome positioning is perhaps the least studied mechanism linked to CTCF-mediated AS, which warrants further experimental investigation.

EPIGENETIC FACTORS

Impact of DNA methylation on CTCF occupancy

Constitutive splicing depends on many factors including the guanine and cytosine (GC) content, which is higher in exons compared to introns. DNA methylation contributes to the definition of exon-intron splice sites depending on the enrichment of CpG dinucleotides within the spliced exons/introns (91,92). It has been shown that 40% of variable CTCF occupancy is linked to differential CpG methylation, concentrated at two main positions within the canonical CTCF consensus sequence (93). Methylation of CpGs within CTCF target sites can prevent CTCF binding (94–98), which has been found to be correlated with DNA methylation patterns in a tissue-specific manner (93). The interplay between CTCF and DNA methylation in modulating AS has been described at a genome-wide or individual transcript level (36–41,43,45).

Genome-wide methylation profiling of multiple tissues could reveal an association between DNA methylation and alternatively spliced transcripts (43). This is interesting, given that CTCF binding sites often occur in GC-rich regions, typically in CpG islands (8,99). These transcripts

were highly enriched in CTCF binding motifs and mutually exclusive exons in a position-dependent context (43). As an exemplar, CTCF was found to regulate a pair of mutually exclusive exons observed in synaptic calcium ion channel *Cacna1b* transcription, where CTCF binding to hypomethylated exon 37a (but not 37b) promotes its recognition and inclusion in neurons (41). Methylation of the exon 37a locus, on the other hand, leads to the inhibition of CTCF binding and exon 37b inclusion. The resulting *Cacna1b* isoform alters the calcium ion channel plasticity of the neuron (41). However, the mechanism by which CTCF modulates splicing at these sites has not been fully explored.

Learning from *BDNF*: an exemplar of CTCF-regulated AS

CTCF is important for exon inclusion by pausing RNAPII-mediated transcription in a methylation-dependent manner (33,34). However, in one archetypal example, CTCF binding is required to protect against mis-splicing in brain-derived neurotrophic factor (*BDNF*), in a seemingly RNAPII elongation-independent manner (38). Mammalian *BDNF*, which is essential for regulating neurogenesis, synaptic plasticity, learning, memory and cognition, encodes multiple transcripts generated by AS (100,101). In turtle brain, intra-exonic splicing of *BDNF* causes a 40 bp deletion wholly within the protein coding exon IV, generating the functionally distinct BDNF2a isoform which has a 13 amino acid deletion and a truncated out-of-frame 21 amino acid C-terminus (102). When methyl-CpG-binding protein MeCP2 binds to its cognate site upstream of the *BDNF* intra-exonic splice site, it recruits the splicing factor Y-box binding protein (YB-1), and generates this truncated BDNF2a isoform. Interestingly, TET1 binding at MeCP2 sites was also essential for *BDNF* splicing, which is inhibited by knockdown of MeCP2, YB1 or TET1 (38).

Moreover, suppression of the TET1-MeCP2-YB1 complex, using a neural correlate of a classical conditioning process in turtles, induces site-specific DNA demethylation which primes CTCF binding to *BDNF*, proximal to the alternatively spliced locus. This interaction leads to shielding of canonical *BDNF* splicing and inhibition of its truncation, without RNAPII elongation stalling (38). The methylation status within the *BDNF* promoter region appears to govern the *BDNF* splicing decision. While methylated CpG dinucleotides provide docking sites for MeCP2 binding, DNA demethylation dissociates MeCP2 and promotes CTCF binding (38). Similarly, DNMT inhibition-mediated hypomethylation in turtle or intracellular NAD biosynthesis inhibition-regulated DNA hypermethylation in mouse, both modulate *BDNF* transcription via triggering dissociation of MeCP2 or CTCF, respectively, from the *BDNF* promoter region (39,103). In addition, NAD-depleted mouse cortical neurons have decreased cohesin recruitment and binding to the *BDNF* promoter, indicating a relationship between chromatin structure and *BDNF* transcription (39). Using hippocampal-specific depletion of CTCF in mice, CTCF/cohesin binding proximal to the *BDNF* promoter was also observed upon fear-conditioning. This attenuated their learning and memory activities as well as expression of key learning-related genes including *BDNF*, *Arc*, *Reln* and *Ppp1c* (104). Altogether, the interplay between

CTCF and the TET1-MeCP2-YB1 complex in modulating *BDNF* splicing is influenced by chromatin structure and their functional impact remains to be elucidated.

DNA methylation governs CTCF/MeCP2/TET1 interplay

Since MeCP2 is known for its role in regulating methylation-dependent AS (38,105,106), the link between CTCF and MeCP2 may help reveal an important underlying mechanism of CTCF-mediated AS. DNA methylation mediates opposing effects on the role of CTCF and MeCP2 binding to DNA and subsequently regulation of pre-mRNA splicing (Figure 3C). Apart from MeCP2/CTCF-mediated *BDNF* splicing (38), both MeCP2 and CTCF can promote exon inclusion in a context-dependent manner by binding to methylated or unmethylated DNA, respectively (33,34,105). Conversely, exon exclusion could occur via DNA demethylation or methylation inhibiting MeCP2 or CTCF binding, respectively. Both CTCF- and MeCP2-mediated exon inclusion have been linked to modulating RNAPII kinetics (33,34,105). While CTCF physically blocks RNAPII elongation (33,34), MeCP2-mediated pausing of RNAPII elongation is linked to recruiting histone deacetylase (HDAC) activity and inducing histone hypoacetylation (105), the inhibition of which increases the RNAPII elongation rate and subsequently exon exclusion (107). In addition, the DNA- and RNA-binding protein YB-1 physically interacts with both CTCF (108,109) and MeCP2 (38,110,111). While MeCP2-mediated recruitment of YB-1 has been linked to AS (38,111), the impact of CTCF and YB-1 interaction on pre-mRNA splicing has not been verified yet.

TET1 is another mutual interacting partner of CTCF (62) and MeCP2 (103,112), through which it regulates AS (33,34,38). Knowing that MeCP2 protects methylated 5mC from TET1-mediated oxidation (113), it is not very clear how abolishing MeCP2 binding to *BDNF* reduces TET1-catalyzed demethylation during the conditioning process, yet promotes CTCF binding (38). This could be explained by the notion that MeCP2 and TET1 are binding partners (103,112), implying that dissociation of MeCP2 possibly removes the whole TET1-MeCP2-YB1 complex. Alternatively, TET1 dissociation might be compensated by TET3, whose binding to *BDNF* has been found to increase following conditioning (38), and important for learning-dependent gene expression and behavioral adaptation (114). Given the common factors acting in CTCF- and MeCP2-mediated AS, their combined or independent roles on modulating AS are still poorly understood and require further studies particularly at a genome-wide level.

Lessons from CTCF haploinsufficiency in human and mouse

While CTCF binding is inhibited by DNA methylation (94–98), CTCF itself regulates DNA methylation and preserves methylation-free regions throughout the genome (115–117). Moreover, CTCF knockdown induces global hypermethylation, particularly at CTCF binding sites, as well as loss of CTCF binding associated with altered expression of CTCF-regulated genes (118). Interestingly, two individuals with

an intellectual disability and diagnosed with *CTCF* heterozygous deletions exhibited hypermethylation at nearly 300 CTCF binding sites genome-wide (119). The regulation of genome-wide methylation by CTCF was also demonstrated in *Ctcf* hemizygous mice, which are an ideal model of *Ctcf* haploinsufficiency. *Ctcf* haploinsufficient mice exhibit aberrant hypermethylation and increased spontaneous tumor formation within various organs (120). Using a similar model, we recently examined the effect of *Ctcf* haploinsufficiency on differential AS in several mouse tissues including brain, kidney, liver, muscle and spleen (45). Our analysis showed an overall perturbation of the AS landscape, causing mostly exon skipping, in a tissue-specific manner. This was despite *Ctcf* dosage only decreasing less than 40% in most tissues examined (45). We also detected a significant tissue-specific increase in intron retention events in *Ctcf* haploinsufficient liver and kidney, primarily affecting short introns with high GC content. The intron retention events in the liver showed that *Ctcf*-mediated intron retention has a differing impact on the mRNA expression of the intron-retaining genes. Interestingly, there was enrichment of *Ctcf* binding sites proximal, particularly upstream, to the majority of the differentially retained introns in mouse liver (45).

Knowing that *Ctcf* haploinsufficiency (119,120) or knockdown (118) induces hypermethylation, we proposed that *Ctcf* haploinsufficiency increased intron retention through a methylation-dependent mechanism (45), which may involve altering the kinetics of RNAPII elongation or chromatin looping. However, this hypothesis requires further direct experimental validation, considering possible indirect effects of decreased *Ctcf* levels. To account for these effects, it would be useful to examine the direct effect of CTCF depletion on global AS by using the auxin-inducible degron system. The degron system has been successfully used to study the acute and reversible effect of CTCF depletion on TAD boundaries and chromatin looping as well as gene expression (121–123). In addition, mutating CTCF binding sites using genome editing could provide another experimental approach to study CTCF-mediated AS at the individual transcript level. Given the cell-type-specific effects of CTCF on AS (45), further investigations into the specific mechanism are required.

A chromatin code enriching AS regulation

Genome-wide chromatin-dependent regulation of AS involves association of CTCF, HP1 α and histone modifications (46). The heterochromatin protein (HP1) family members (α , β and γ), which can co-localize with the heterochromatin-associated histone modification H3K9me3, have been shown to regulate the impact of DNA methylation on AS in a position-dependent manner. Binding of HP1 to an exon or its upstream intron was associated with diminishing or enhancing exon inclusion, respectively. This methylation-dependent binding modulates RNA splicing by recruiting splicing factors, such as SRSF3, to the methylated site (124). Yet, it remains to be determined how changing HP1 binding sites leads to different AS decisions. A proposed ‘chromatin code’ highlighted by enrichment of HP1 α , CTCF, the transcriptional

gene silencer Argonaute (AGO1), RNAPII and select histone marks around AS-regulated exons governs different modes of exon inclusion or skipping (46). The interaction of RNAPII with AGO1 (125) affects its elongation in the context of HP1 α -mediated chromatin condensation thus causing higher inclusion of alternative exons (126). Knowing that HP1 α recruits and interacts with CTCF (127), it would be reasonable to postulate that AS can result from crosstalk between HP1 and CTCF. In addition, given the association between CTCF and HP1 α proximal to included exons (46), it would be interesting to further explore whether CTCF binding contributes to altering exon splicing decisions. This might be facilitated by DNA methylation and HP1-mediated splicing factor recruitment (124) (Figure 2C). Alternatively, CTCF may modulate AS by slowing RNAPII elongation via HP1 α -mediated chromatin condensation at methylated sites (128).

The distribution of CTCF binding sites coincides with nucleosome positioning (7,85–87). These studies demonstrated that enrichment of CTCF binding sites is detected around nucleosomes associated with histone variants H2A.Z and H3.3 and multiple histone modifications. The co-localization of CTCF with these histone variants and modifications marks the boundaries of histone methylation domains and was linked to chromatin-mediated gene expression regulation (7,85–87). Consistent with the above, CTCF binding along with active histone modifications H3K9ac, H3K27ac or H3K4me3 have been shown to be highly enriched in ‘more-included’ exons compared to ‘less-included’ exons (129). However, correlations between the repressive histone modifications H3K27me3 and H3K9me2 and CTCF binding sites have been detected around alternatively skipped exons in MCF7 more than MCF10A cells (46). For example, the *PUMA* locus harbors an intragenic chromatin boundary consisting of the active marks H3K4me3 and H3K9ac and the repressive mark H3K9me3, important for CTCF-mediated regulation of *PUMA* (35). Interestingly, CTCF knockdown reduces H3K9me3 levels within *PUMA* as well as at *p21* loci (35). Moreover, a broad putative role for CTCF in regulating diverse histone modifications could be achieved via PARP1 and PARylation (65). These findings argue for a plausible link between CTCF-associated histone modifications and pre-mRNA splicing outcome, which might involve regulation of RNAPII kinetics and/or splicing factor recruitment via chromatin-binding proteins.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

Transcriptomic diversity is governed by AS, which is orchestrated by complex regulatory mechanisms comprising genomic, epigenetic, and transcriptional layers (51,130). The versatile and ubiquitous factor CTCF plays diverse roles in these regulatory layers through multimodal mechanisms. However, there are still gaps in our understanding of the interplay between these mechanisms. Further research is required to gain a clearer picture of the CTCF-mediated AS regulatory network. Combined computational and experimental efforts are required to foster a comprehensive mechanistic understanding of the experimentally verified and putative AS regulatory roles mediated by CTCF (Table 1).

Overall, most evidence of CTCF-mediated AS regulation has been found at the individual transcript level while the putative roles are mostly described at a genome-wide level.

Enrichment of CTCF binding sites proximal to alternatively spliced exons has been linked to methylation, RNAPII elongation, chromatin looping, histone modifications and splicing factor recruitment (33–46). To date, most studies have examined RNAPII elongation- and DNA methylation-related mechanisms. However, other CTCF-mediated mechanisms regulating AS, particularly nucleosome positioning, chromatin architecture and histone modifications, remain largely uncharacterized. While CTCF appears to spatio-temporally coordinate AS, it is still unknown whether these factors have cooperative, competitive, dependent, or independent relationships in CTCF-mediated AS. In addition, the extent of CTCF’s role in modulating AS remains uncertain in regard to its specificity to individual transcripts versus genome-wide levels, ubiquitous versus cell-type-specific CTCF target sites and cellular versus systemic functions. It is worth noting that the association of CTCF-mediated AS events and resulting biological functions have not been extensively studied. Some studies have already shed light on several functions influenced by this role such as lymphocyte development (33,34), p53-mediated apoptotic response (35), learning-dependent activities (38), neuronal cell surface diversity (36), calcium ion channel plasticity (41) and the impact of *Ctcf* haploinsufficiency (45).

Pursuing these avenues of research present many computational challenges including the complexities of analyzing the transcriptome and epigenome as well as current limitations of the available bioinformatic tools in detecting AS events, annotating protein isoforms and providing optimal data resolution and coverage (reviewed in (131,132)). There are several data resources, such as ENCODE (encodeproject.org), the Cancer Genome Atlas (TCGA: <https://www.cancer.gov/tcga>), and the Cancer Cell Line Encyclopedia (portals.broadinstitute.org/ccle), which are useful for comprehensive analysis of AS landscapes in various normal and malignant cells. Given frequent allelic variations and somatic mutations in CTCF (133–136), it would be important to establish links between aberrant CTCF-mediated AS and cancer. Knowing that CTCF and its testis-specific paralogue CTCF-like (CTCFL)/BORIS have some mutual genomic occupancy (137), a new role for BORIS in AS regulation may also warrant further investigation. Comprehensive data analyses should aim to resolve the reciprocal and mutually exclusive interplay between CTCF and key factors involved in AS regulation rather than studying CTCF-mediated effects on an individual factor.

Structural changes in the genome arising from genetic variation or introduced experimentally via deletion or inversion of CTCF sites using genome editing can have pathological consequences. Disrupting the highly conserved architecture of TADs has been shown to have a dramatic impact on development and normal spatio-temporal control of gene expression (138–140). While the impact of genetic variation on CTCF-mediated chromatin looping has been well established (138–140), the consequences for the AS landscape is only now being revealed. Allelic variation at CTCF sites adjacent to exons can both affect CTCF binding and exon inclusion (44). Future analyses will seek to

Table 1. Summary of the known mechanisms of CTCF in modulating alternative splicing. CTCF-mediated AS regulatory mechanisms are divided into verified or putative roles

Mechanism	Context	Model	Description	Biological relevance	Ref
Verified mechanisms					
RNAPII elongation*	<i>CD45</i>	Human Burkitt lymphoma B cells	CTCF binding to <i>CD45</i> exon 5 promotes exon inclusion via stalling RNAPII elongation	Lymphocyte development	(33,34)
RNAPII elongation	<i>PUMA, p21</i>	Human HCT116 colorectal carcinoma cells	CTCF binding recruits cohesin downstream to <i>PUMA</i> and <i>p21</i> promoters, which stalls RNAPII elongation and regulates their expression	p53-mediated apoptotic response	(35)
DNA methylation	<i>BDNF</i>	Turtle brain tissue, mouse embryonic cortical neurons	CTCF binding to unmethylated target sites proximal to <i>BDNF</i> promotes canonical <i>BDNF</i> splicing	Learning-dependent activities	(38,39)
DNA methylation	<i>Caena1b</i>	F11 (rat dorsal-root ganglion neurons/mouse neuroblastoma hybrid) cells	CTCF binding to hypomethylated sites at <i>Caena1b</i> promotes mutually exclusive exons	Calcium ion channel plasticity	(41)
Chromatin architecture*	<i>Pcdh</i>	Human SK-N-SH and mouse N2a neuroblastoma cells, mouse CAD catecholaminergic neuronal tumor cells	CTCF/cohesin-mediated DNA looping induces alternatively spliced <i>Pcdh</i> isoforms	Neuronal cell surface diversity	(36,37)
Chromatin architecture*	Multiple transcripts	Human HCT116 colorectal carcinoma cells	CTCF/cohesin-mediated DNA looping induces alternative polyadenylation	Altering cancer transcriptome	(40)
Splicing factor recruitment	<i>c-myc, U2 snRNA</i>	HeLa cells	CTCF controls RNAPII elongation and termination via recruitment of NELF, DSIF and P-TEFb	Gene expression	(42)
Putative mechanisms					
RNAPII elongation	Genome-wide	Human Burkitt lymphoma B cells	Association between CTCF and RNAPII elongation proximal to included exons	-	(33,34)
RNAPII elongation	Genome-wide	Various cells	Association between CTCF binding and RNAPII stalling at specific sites	-	(58,59)
DNA methylation	Genome-wide	Mouse retina and brain tissues	Association of CTCF binding sites and DNA methylation in alternatively spliced transcripts	-	(43)
Chromatin architecture*	Genome-wide	Human lymphoblastoid cells	CTCF-mediated chromatin loops bring alternatively spliced exons into physical proximity of gene promoter	-	(44)
DNA methylation	Genome-wide	Mouse brain, kidney, liver, muscle and spleen tissues	<i>Ctcf</i> haploinsufficiency mediates tissue-specific changes in AS events, notably an increase in intron retention in <i>Ctcf</i> haploinsufficient liver and kidney	-	(45)
Histone modification	Genome-wide	Various cells	Association between CTCF binding and histone modifications close to alternatively spliced exons	-	(35,46,129)
Miscellaneous	Genome-wide	Human MCF7 breast cancer cells Non-tumorigenic MCF10A mammary cells	Association between CTCF and HP1 α proximal to included exons which may involve splicing factor recruitment	-	(46)
Miscellaneous*	Genome-wide	Various cells	CTCF-mediated activation of PARP1 and PARylation govern AS via hypomethylation of CTCF binding sites, RNAPII elongation, chromatin relaxation and splicing factor recruitment	-	(60,61,66–69)

*denotes involvement of methylation-regulated CTCF sites.

explore correlations between transcriptomic and genomic data and cellular or developmental phenotypes. Moreover, given the extensive characterization of CTCF as a tumor suppressor gene (120,136,141–143), its frequent mutation in cancer (133,134,136) and loss- or gain-of-function phenotypes (135), causal links between CTCF-mediated aberrant AS and cancer are likely to emerge (144,145).

To date, the role of CTCF in AS has been mostly elucidated using computational approaches, while experimental validation remains challenging especially at a genome-wide

level. This is further complicated by the fact that CTCF is essential for somatic cell viability (143), genome architecture (30,32) and gene expression regulation (5,22). Integrating the known roles for CTCF in regulating chromatin architecture and gene expression with AS-mediated effects on transcriptome and proteome diversity would provide opportunities for future directions. For instance, it is still unknown how AS is impacted by CTCF-mediated TADs and compartmentalization of chromatin. In addition, CTCF is known to transcriptionally regulate many

key AS-regulating genes and interact directly with proteins and RNA molecules leading to direct or indirect effects on AS. Ultimately, unravelling the underlying mechanisms by which CTCF regulates AS may address the question as to what determines the fates of CTCF-mediated alternatively spliced transcripts and their biological functions.

FUNDING

National Health and Medical Research Council [Project grants 1128748 and 1128175 to J.E.J.R., Investigator grants to J.E.J.R. 1177305 and to U.S. 1196405]; Tour de Cure (Scott Canner Research Fellowship to C.G.B. and for research grants to C.G.B. and J.E.J.R.); Tour de Rocks project support (to C.G.B.); Cancer Council NSW project grants [RG11-12, RG14-09 to J.E.J.R. and C.G.B., (RG20-12) to U.S. and C.G.B.]; Cure the Future (to J.E.J.R.); U.S. was supported via an Early Career Researcher Fellowship from Cancer Institute New South Wales; A.B.A. is supported by a PhD Scholarship from Umm Al-Qura University in Saudi Arabia to conduct a PhD at The University of Sydney. Funding for open access charge was supported by an anonymous foundation.

Conflict of interest statement. None declared.

REFERENCES

- Lobanekov, V.V., Nicolas, R.H., Adler, V.V., Paterson, H., Klenova, E.M., Polotskaja, A.V. and Goodwin, G.H. (1990) A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene*, **5**, 1743–1753.
- Klenova, E.M., Nicolas, R.H., Paterson, H.F., Carne, A.F., Heath, C.M., Goodwin, G.H., Neiman, P.E. and Lobanekov, V.V. (1993) CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol. Cell. Biol.*, **13**, 7612–7624.
- Bell, A.C., West, A.G. and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, **98**, 387–396.
- Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J. and Lobanekov, V.V. (1996) An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol. Cell. Biol.*, **16**, 2802–2813.
- Ohlsson, R., Renkawitz, R. and Lobanekov, V. (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.*, **17**, 520–527.
- Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanekov, V.V. and Ren, B. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*, **128**, 1231–1245.
- Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K. and Zhao, K. (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.*, **19**, 24–32.
- Chen, H., Tian, Y., Shu, W., Bo, X. and Wang, S. (2012) Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PLoS One*, **7**, e41374.
- Saldana-Meyer, R., Gonzalez-Buendia, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F. and Reinberg, D. (2014) CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.*, **28**, 723–734.
- Hansen, A.S., Hsieh, T.S., Cattoglio, C., Pustova, L., Saldana-Meyer, R., Reinberg, D., Darzacq, X. and Tjian, R. (2019) Distinct classes of chromatin loops revealed by deletion of an RNA-binding region in CTCF. *Mol. Cell*, **76**, 395–411.
- Saldana-Meyer, R., Rodriguez-Hernandez, J., Escobar, T., Nishana, M., Jacome-Lopez, K., Nora, E.P., Bruneau, B.G., Tsirigos, A., Furlan-Magaril, M., Skok, J. et al. (2019) RNA interactions are essential for CTCF-mediated genome organization. *Mol. Cell*, **76**, 412–422.
- Caudron-Herger, M., Rusin, S.F., Adamo, M.E., Seiler, J., Schmid, V.K., Barreau, E., Kettenbach, A.N. and Diederichs, S. (2019) R-DeeP: proteome-wide and quantitative identification of RNA-dependent proteins by density gradient ultracentrifugation. *Mol. Cell*, **75**, 184–199.
- Kung, J.T., Kesner, B., An, J.Y., Ahn, J.Y., Cifuentes-Rojas, C., Colognori, D., Jeon, Y., Szanto, A., del Rosario, B.C., Pinter, S.F. et al. (2015) Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell*, **57**, 361–375.
- Sun, S., Del Rosario, B.C., Szanto, A., Ogawa, Y., Jeon, Y. and Lee, J.T. (2013) Jpx RNA activates Xist by evicting CTCF. *Cell*, **153**, 1537–1551.
- Ohlsson, R., Lobanekov, V. and Klenova, E. (2010) Does CTCF mediate between nuclear organization and gene expression? *Bioessays*, **32**, 37–50.
- Zlatanova, J. and Caiafa, P. (2009) CTCF and its protein partners: divide and rule? *J. Cell Sci.*, **122**, 1275–1284.
- Arzate-Mejia, R.G., Recillas-Targa, F. and Corces, V.G. (2018) Developing in 3D: the role of CTCF in cell differentiation. *Development*, **145**, dev137729.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376–380.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Pilot, T., van Berkum, N.L., Meisig, J., Sedat, J. et al. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, **485**, 381–385.
- Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A. and Hadjur, S. (2015) Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.*, **10**, 1297–1309.
- Redolfi, J., Zhan, Y., Valdes-Quezada, C., Kryzhanovska, M., Guerreiro, I., Iesmantavicius, V., Pollex, T., Grand, R.S., Mulugeta, E., Kind, J. et al. (2019) DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat. Struct. Mol. Biol.*, **26**, 471–480.
- Phillips, J.E. and Corces, V.G. (2009) CTCF: master weaver of the genome. *Cell*, **137**, 1194–1211.
- Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. et al. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
- Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y. et al. (2015) CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell*, **162**, 900–910.
- Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T. et al. (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, **132**, 422–433.
- Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y. et al. (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*, **153**, 1281–1295.
- Rubio, E.D., Reiss, D.J., Welch, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A. and Krumm, A. (2008) CTCF physically links cohesin to chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 8309–8314.
- Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T. et al. (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**, 796–801.
- Filippova, G.N. (2008) Genetics and epigenetics of the multifunctional protein CTCF. *Curr. Top. Dev. Biol.*, **80**, 337–360.
- Ong, C.T. and Corces, V.G. (2014) CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.*, **15**, 234–246.

31. Zlatanova, J. and Caiafa, P. (2009) CCCTC-binding factor: to loop or to bridge. *Cell. Mol. Life Sci.*, **66**, 1647–1660.
32. Lee, B.K. and Iyer, V.R. (2012) Genome-wide studies of CCCTC-binding factor (CTCF) and cohesin provide insight into chromatin structure and regulation. *J. Biol. Chem.*, **287**, 30906–30913.
33. Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R. and Oberdoerffer, S. (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, **479**, 74–79.
34. Marina, R.J., Sturgill, D., Bailly, M.A., Thenoz, M., Varma, G., Prigge, M.F., Nanan, K.K., Shukla, S., Haque, N. and Oberdoerffer, S. (2016) TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing. *EMBO J.*, **35**, 335–355.
35. Gomes, N.P. and Espinosa, J.M. (2010) Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev.*, **24**, 1022–1034.
36. Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K.E., Li, W., Myers, R.M., Maniatis, T. and Wu, Q. (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 21081–21086.
37. Monahan, K., Rudnick, N.D., Kehayova, P.D., Pauli, F., Newberry, K.M., Myers, R.M. and Maniatis, T. (2012) Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin- α gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 9125–9130.
38. Zheng, Z., Ambigapathy, G. and Keifer, J. (2017) MeCP2 regulates Tet1-catalyzed demethylation, CTCF binding, and learning-dependent alternative splicing of the BDNF gene in Turtle. *Elife*, **6**, e25384.
39. Chang, J., Zhang, B., Heath, H., Galjart, N., Wang, X. and Milbrandt, J. (2010) Nicotinamide adenine dinucleotide (NAD)-regulated DNA methylation alters CCCTC-binding factor (CTCF)/cohesin binding and transcription at the BDNF locus. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21836–21841.
40. Nanavaty, V., Abrash, E.W., Hong, C., Park, S., Fink, E.E., Li, Z., Sweet, T.J., Bhasin, J.M., Singuri, S., Lee, B.H. *et al.* (2020) DNA methylation regulates alternative polyadenylation via CTCF and the cohesin complex. *Mol. Cell*, **78**, 752–764.
41. Lopez Soto, E.J. and Lipscombe, D. (2020) Cell-specific exon methylation and CTCF binding in neurons regulate calcium ion channel splicing and function. *Elife*, **9**, e54879.
42. Laitem, C., Zaborowska, J., Tellier, M., Yamaguchi, Y., Cao, Q., Egloff, S., Handa, H. and Murphy, S. (2015) CTCF regulates NELF, DSIF and P-TEFb recruitment during transcription. *Transcription*, **6**, 79–90.
43. Wan, J., Oliver, V.F., Zhu, H., Zack, D.J., Qian, J. and Merbs, S.L. (2013) Integrative analysis of tissue-specific methylation and alternative splicing identifies conserved transcription factor binding motifs. *Nucleic Acids Res.*, **41**, 8503–8514.
44. Ruiz-Velasco, M., Kumar, M., Lai, M.C., Bhat, P., Solis-Pinson, A.B., Reyes, A., Kleinsorg, S., Noh, K.M., Gibson, T.J. and Zaugg, J.B. (2017) CTCF-mediated chromatin loops between promoter and gene body regulate alternative splicing across individuals. *Cell Syst.*, **5**, 628–637.
45. Alharbi, A.B., Schmitz, U., Marshall, A.D., Vanichkina, D., Nagarajah, R., Vellozzi, M., Wong, J.J., Bailey, C.G. and Rasko, J.E. (2021) Ctf haploinsufficiency mediates intron retention in a tissue-specific manner. *RNA Biol*, **18**, 93–103.
46. Agirre, E., Bellora, N., Allo, M., Pages, A., Bertucci, P., Kornblihtt, A.R. and Eyra, E. (2015) A chromatin code for alternative splicing involving a putative association between CTCF and HP1 α proteins. *BMC Biol.*, **13**, 31.
47. Keren, H., Lev-Maor, G. and Ast, G. (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.*, **11**, 345–355.
48. Nilsen, T.W. and Graveley, B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
49. Montes, M., Sanford, B.L., Comiskey, D.F. and Chandler, D.S. (2019) RNA splicing and disease: animal models to therapies. *Trends Genet.*, **35**, 68–87.
50. Scotti, M.M. and Swanson, M.S. (2016) RNA mis-splicing in disease. *Nat. Rev. Genet.*, **17**, 19–32.
51. Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.*, **10**, 741–754.
52. Dvinge, H. (2018) Regulation of alternative mRNA splicing: old players and new perspectives. *FEBS Lett.*, **592**, 2987–3006.
53. de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A.R. (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell*, **12**, 525–532.
54. Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X.D. and Bentley, D.L. (2014) Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev.*, **28**, 2663–2676.
55. Ip, J.Y., Schmidt, D., Pan, Q., Ramani, A.K., Fraser, A.G., Odom, D.T. and Blencowe, B.J. (2011) Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res.*, **21**, 390–401.
56. Saldi, T., Cortazar, M.A., Sheridan, R.M. and Bentley, D.L. (2016) Coupling of RNA Polymerase II transcription elongation with pre-mRNA splicing. *J. Mol. Biol.*, **428**, 2623–2635.
57. Hermiston, M.L., Xu, Z. and Weiss, A. (2003) CD45: a critical regulator of signaling thresholds in immune cells. *Annu. Rev. Immunol.*, **21**, 107–137.
58. Chernukhin, I., Shamsuddin, S., Kang, S.Y., Bergstrom, R., Kwon, Y.W., Yu, W., Whitehead, J., Mukhopadhyay, R., Docquier, F., Farrar, D. *et al.* (2007) CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol. Cell Biol.*, **27**, 1631–1648.
59. Paredes, S.H., Melgar, M.F. and Sethupathy, P. (2013) Promoter-proximal CCCTC-factor binding is associated with an increase in the transcriptional pausing index. *Bioinformatics*, **29**, 1485–1487.
60. Guastafierro, T., Cecchinelli, B., Zampieri, M., Reale, A., Riggio, G., Sthandier, O., Zupi, G., Calabrese, L. and Caiafa, P. (2008) CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. *J. Biol. Chem.*, **283**, 21873–21880.
61. Zampieri, M., Guastafierro, T., Calabrese, R., Ciccarone, F., Bacalini, M.G., Reale, A., Perilli, M., Passananti, C. and Caiafa, P. (2012) ADP-ribose polymers localized on Ctf-Parp1-Dnmt1 complex prevent methylation of Ctf target sites. *Biochem. J.*, **441**, 645–652.
62. Dubois-Chevalier, J., Oger, F., Dehondt, H., Firmin, F.F., Gheeraert, C., Staels, B., Lefebvre, P. and Eeckhoutte, J. (2014) A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and transcriptional induction of adipocyte differentiation. *Nucleic Acids Res.*, **42**, 10943–10959.
63. Hottiger, M.O., Hassa, P.O., Lüscher, B., Schuler, H. and Koch-Nolte, F. (2010) Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem. Sci.*, **35**, 208–219.
64. Bürkle, A. and Virág, L. (2013) Poly(ADP-ribose): PARadigms and PARadoxes. *Mol. Aspects Med.*, **34**, 1046–1065.
65. Ciccarone, F., Zampieri, M. and Caiafa, P. (2017) PARP1 orchestrates epigenetic events setting up chromatin domains. *Semin. Cell Dev. Biol.*, **63**, 123–134.
66. Matveeva, E., Maiorano, J., Zhang, Q., Eteleeb, A.M., Convertini, P., Chen, J., Infantino, V., Stamm, S., Wang, J., Rouchka, E.C. *et al.* (2016) Involvement of PARP1 in the regulation of alternative splicing. *Cell Discov.*, **2**, 15046.
67. Matveeva, E.A., Al-Tinawi, Q.M.H., Rouchka, E.C. and Fondufe-Mittendorf, Y.N. (2019) Coupling of PARP1-mediated chromatin structural changes to transcriptional RNA polymerase II elongation and cotranscriptional splicing. *Epigenet. Chromatin*, **12**, 15.
68. Eleazer, R. and Fondufe-Mittendorf, Y.N. (2021) The multifaceted role of PARP1 in RNA biogenesis. *Wiley Interdiscip. Rev. RNA*, **12**, e1617.
69. Ke, Y., Zhang, J., Lv, X., Zeng, X. and Ba, X. (2019) Novel insights into PARPs in gene expression: regulation of RNA metabolism. *Cell. Mol. Life Sci.*, **76**, 3283–3299.
70. Lehman, B.J., Lopez-Diaz, F.J., Santisakultarm, T.P., Fang, L., Shokhiev, M.N., Diffenderfer, K.E., Manor, U. and Emerson, B.M.

- (2021) Dynamic regulation of CTCF stability and sub-nuclear localization in response to stress. *PLoS Genet.*, **17**, e1009277.
71. Hsin, J.P. and Manley, J.L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.*, **26**, 2119–2137.
 72. Egloff, S., Al-Rawaf, H., O'Reilly, D. and Murphy, S. (2009) Chromatin structure is implicated in "late" elongation checkpoints on the U2 snRNA and beta-actin genes. *Mol. Cell Biol.*, **29**, 4002–4013.
 73. Wada, Y., Ohta, Y., Xu, M., Tsutsumi, S., Minami, T., Inoue, K., Komura, D., Kitakami, J., Oshida, N., Papantonis, A. *et al.* (2009) A wave of nascent transcription on activated human genes. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 18357–18361.
 74. Fay, A., Misulovin, Z., Li, J., Schaaf, C.A., Gause, M., Gilmour, D.S. and Dorsett, D. (2011) Cohesin selectively binds and regulates genes with paused RNA polymerase. *Curr. Biol.*, **21**, 1624–1634.
 75. Di Giammartino, D.C., Nishida, K. and Manley, J.L. (2011) Mechanisms and consequences of alternative polyadenylation. *Mol. Cell*, **43**, 853–866.
 76. Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M. and Bustamante, C. (2009) Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science*, **325**, 626–628.
 77. Spies, N., Nielsen, C.B., Padgett, R.A. and Burge, C.B. (2009) Biased chromatin signatures around polyadenylation sites and exons. *Mol. Cell*, **36**, 245–254.
 78. Schwartz, S., Meshorer, E. and Ast, G. (2009) Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.*, **16**, 990–995.
 79. Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcarcel, J. and Guigo, R. (2009) Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.*, **16**, 996–1001.
 80. Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C. and Komorowski, J. (2009) Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.*, **19**, 1732–1741.
 81. Petrova, V., Song, R., Nordström, K.J.V., Walter, J., Wong, J.J.L., Armstrong, N.J., Rasko, J.E.J. and Schmitz, U. (2021) Chromatin accessibility determines intron retention in a cell type-specific manner. bioRxiv doi: <https://doi.org/10.1101/2021.02.17.431609>, 30 March 2021, preprint: not peer reviewed.
 82. Kolasinska-Zwiercz, P., Down, T., Latorre, I., Liu, T., Liu, X.S. and Ahringer, J. (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.*, **41**, 376–381.
 83. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli, T. (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
 84. Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R. and Bickmore, W.A. (2012) Psp1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet.*, **8**, e1002717.
 85. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
 86. Fu, Y., Sinha, M., Peterson, C.L. and Weng, Z. (2008) The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.*, **4**, e1000138.
 87. Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K. and Felsenfeld, G. (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.*, **41**, 941–945.
 88. Barisic, D., Stadler, M.B., Iurlaro, M. and Schubeler, D. (2019) Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature*, **569**, 136–140.
 89. Marino, M.M., Rega, C., Russo, R., Valletta, M., Gentile, M.T., Esposito, S., Baglivo, I., De Feis, I., Angelini, C., Xiao, T. *et al.* (2019) Interactome mapping defines BRG1, a component of the SWI/SNF chromatin remodeling complex, as a new partner of the transcriptional regulator CTCF. *J. Biol. Chem.*, **294**, 861–873.
 90. Barutcu, A.R., Lajoie, B.R., Fritz, A.J., McCord, R.P., Nickerson, J.A., van Wijnen, A.J., Lian, J.B., Stein, J.L., Dekker, J., Stein, G.S. *et al.* (2016) SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells. *Genome Res.*, **26**, 1188–1201.
 91. Gelfman, S. and Ast, G. (2013) When epigenetics meets alternative splicing: the roles of DNA methylation and GC architecture. *Epigenomics*, **5**, 351–353.
 92. Gelfman, S., Cohen, N., Yearim, A. and Ast, G. (2013) DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon-intron structure. *Genome Res.*, **23**, 789–799.
 93. Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R. *et al.* (2012) Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.*, **22**, 1680–1688.
 94. Bell, A.C. and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**, 482–485.
 95. Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanov, V.V. and Tapscott, S.J. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.*, **28**, 335–343.
 96. Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M. and Tilghman, S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**, 486–489.
 97. Szabo, P., Tang, S.H., Rentsendorj, A., Pfeifer, G.P. and Mann, J.R. (2000) Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. *Curr. Biol.*, **10**, 607–610.
 98. Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C.F., Wolfe, A., Ohlsson, R. and Lobanov, V.V. (2000) Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.*, **10**, 853–856.
 99. Wiehle, L., Thorn, G.J., Raddatz, G., Clarkson, C.T., Rippe, K., Lyko, F., Breiling, A. and Teif, V.B. (2019) DNA (de)methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. *Genome Res.*, **29**, 750–761.
 100. Chen, K.W. and Chen, L. (2017) Epigenetic regulation of BDNF gene during development and diseases. *Int. J. Mol. Sci.*, **18**, 571.
 101. Zuccato, C. and Cattaneo, E. (2009) Brain-derived neurotrophic factor in neurodegenerative diseases. *Nat. Rev. Neurol.*, **5**, 311–322.
 102. Ambigapathy, G., Zheng, Z., Li, W. and Keifer, J. (2013) Identification of a functionally distinct truncated BDNF mRNA splice variant and protein in *Trachemys scripta elegans*. *PLoS One*, **8**, e67141.
 103. Ambigapathy, G., Zheng, Z. and Keifer, J. (2015) Regulation of BDNF chromatin status and promoter accessibility in a neural correlate of associative learning. *Epigenetics*, **10**, 981–993.
 104. Sams, D.S., Nardone, S., Getselter, D., Raz, D., Tal, M., Rayi, P.R., Kaphzan, H., Hakim, O. and Elliott, E. (2016) Neuronal CTCF is necessary for basal and experience-dependent gene regulation, memory formation, and genomic structure of BDNF and Arc. *Cell Rep.*, **17**, 2418–2430.
 105. Maunakea, A.K., Chepelev, I., Cui, K. and Zhao, K. (2013) Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.*, **23**, 1256–1269.
 106. Wong, J.J., Gao, D., Nguyen, T.V., Kwok, C.T., van Geldermalsen, M., Middleton, R., Pinello, N., Thoeng, A., Nagarajah, R., Holst, J. *et al.* (2017) Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. *Nat. Commun.*, **8**, 15134.
 107. Hnilicová, J., Hozeifi, S., Dušková, E., Icha, J., Tománková, T. and Staněk, D. (2011) Histone deacetylase activity modulates alternative splicing. *PLoS One*, **6**, e16727.
 108. Chernukhin, I.V., Shamsuddin, S., Robinson, A.F., Carne, A.F., Paul, A., El-Kady, A.I., Lobanov, V.V. and Klenova, E.M. (2000) Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *J. Biol. Chem.*, **275**, 29915–29921.
 109. Klenova, E., Scott, A.C., Roberts, J., Shamsuddin, S., Lovejoy, E.A., Bergmann, S., Bubb, V.J., Royer, H.D. and Quinn, J.P. (2004) YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders. *J. Neurosci.*, **24**, 5966–5973.
 110. Gonzales, M.L., Adams, S., Dunaway, K.W. and LaSalle, J.M. (2012) Phosphorylation of distinct sites in MeCP2 modifies cofactor

- associations and the dynamics of transcriptional regulation. *Mol. Cell Biol.*, **32**, 2894–2903.
111. Young, J.L., Hong, E.P., Castle, J.C., Crespo-Barreto, J., Bowman, A.B., Rose, M.F., Kang, D., Richman, R., Johnson, J.M., Berget, S. *et al.* (2005) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 17551–17558.
 112. Cartron, P.F., Nadaradjane, A., Lepape, F., Lallier, L., Gardie, B. and Vallette, F.M. (2013) Identification of TET1 partners that control its DNA-demethylating function. *Genes Cancer*, **4**, 235–241.
 113. Ludwig, A.K., Zhang, P., Hastert, F.D., Meyer, S., Rausch, C., Herce, H.D., Müller, U., Lehmkuhl, A., Hellmann, I., Trummer, C. *et al.* (2017) Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. *Nucleic Acids Res.*, **45**, 2438–2457.
 114. Li, X., Wei, W., Zhao, Q.Y., Widagdo, J., Baker-Andresen, D., Flavell, C.R., D'Alessio, A., Zhang, Y. and Bredy, T.W. (2014) Neocortical Tet3-mediated accumulation of 5-hydroxymethylcytosine promotes rapid behavioral adaptation. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 7120–7125.
 115. Mukhopadhyay, R., Yu, W., Whitehead, J., Xu, J., Lezcano, M., Pack, S., Kanduri, C., Kanduri, M., Ginjala, V., Vostrov, A. *et al.* (2004) The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res.*, **14**, 1594–1602.
 116. Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanenkova, V. and Ohlsson, R. (2004) Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol. Cell Biol.*, **24**, 3497–3504.
 117. Schoenherr, C.J., Levorse, J.M. and Tilghman, S.M. (2003) CTCF maintains differential methylation at the Igf2/H19 locus. *Nat. Genet.*, **33**, 66–69.
 118. Damaschke, N.A., Gawdzik, J., Avilla, M., Yang, B., Svaren, J., Roopra, A., Luo, J.H., Yu, Y.P., Keles, S. and Jarrard, D.F. (2020) CTCF loss mediates unique DNA hypermethylation landscapes in human cancers. *Clin. Epigenet.*, **12**, 80.
 119. Hori, I., Kawamura, R., Nakabayashi, K., Watanabe, H., Higashimoto, K., Tomikawa, J., Ieda, D., Ohashi, K., Negishi, Y., Hattori, A. *et al.* (2017) CTCF deletion syndrome: clinical features and epigenetic delineation. *J. Med. Genet.*, **54**, 836–842.
 120. Kemp, C.J., Moore, J.M., Moser, R., Bernard, B., Teater, M., Smith, L.E., Rabaia, N.A., Gurley, K.E., Guinney, J., Busch, S.E. *et al.* (2014) CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.*, **7**, 1020–1029.
 121. Hyle, J., Zhang, Y., Wright, S., Xu, B., Shao, Y., Easton, J., Tian, L., Feng, R., Xu, P. and Li, C. (2019) Acute depletion of CTCF directly affects MYC regulation through loss of enhancer-promoter looping. *Nucleic Acids Res.*, **47**, 6699–6713.
 122. Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A. and Bruneau, B.G. (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, **169**, 930–944.
 123. Stik, G., Vidal, E., Barrero, M., Cuartero, S., Vila-Casadesus, M., Mendieta-Esteban, J., Tian, T.V., Choi, J., Berenguer, C., Abad, A. *et al.* (2020) CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response. *Nat. Genet.*, **52**, 655–661.
 124. Yearim, A., Gelfman, S., Shayevitch, R., Melcer, S., Glaich, O., Mallm, J.P., Nissim-Rafinia, M., Cohen, A.H., Rippe, K., Meshorer, E. *et al.* (2015) HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep.*, **10**, 1122–1134.
 125. Huang, V., Zheng, J., Qi, Z., Wang, J., Place, R.F., Yu, J., Li, H. and Li, L.C. (2013) Ago1 interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. *PLoS Genet.*, **9**, e1003821.
 126. Allo, M., Buggiano, V., Fededa, J.P., Petrillo, E., Schor, I., de la Mata, M., Agirre, E., Plass, M., Eyrales, E., Elela, S.A. *et al.* (2009) Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.*, **16**, 717–724.
 127. Bosch-Presegue, L., Raurell-Vila, H., Thackray, J.K., Gonzalez, J., Casal, C., Kane-Goldsmith, N., Vizoso, M., Brown, J.P., Gomez, A., Ausio, J. *et al.* (2017) Mammalian HP1 isoforms have specific roles in heterochromatin structure and organization. *Cell Rep.*, **21**, 2048–2057.
 128. Eissenberg, J.C. and Elgin, S.C. (2014) HP1a: a structural chromosomal protein regulating transcription. *Trends Genet.*, **30**, 103–110.
 129. Curado, J., Iannone, C., Tilgner, H., Valcarcel, J. and Guigo, R. (2015) Promoter-like epigenetic signatures in exons displaying cell type-specific splicing. *Genome Biol.*, **16**, 236.
 130. Shin, C. and Manley, J.L. (2004) Cell signalling and the control of pre-mRNA splicing. *Nat. Rev. Mol. Cell Biol.*, **5**, 727–738.
 131. Vanichkina, D.P., Schmitz, U., Wong, J.J. and Rasko, J.E.J. (2018) Challenges in defining the role of intron retention in normal biology and disease. *Semin. Cell Dev. Biol.*, **75**, 40–49.
 132. Park, E., Pan, Z., Zhang, Z., Lin, L. and Xing, Y. (2018) The expanding landscape of alternative splicing variation in human populations. *Am. J. Hum. Genet.*, **102**, 11–26.
 133. Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B. *et al.* (2018) Comprehensive characterization of cancer driver genes and mutations. *Cell*, **173**, 371–385.
 134. Kandath, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A. *et al.* (2013) Mutational landscape and significance across 12 major cancer types. *Nature*, **502**, 333–339.
 135. Bailey, C.G., Gupta, S., Metierre, C., Amarasekera, P.M.S., O'Young, P., Kyaw, W., Laletin, T., Francis, H., Semaan, C., Singh, K.P. *et al.* (2021) Somatic mutations in CTCF zinc fingers produce cellular phenotypes explained by structure-function relationships. bioRxiv doi: <https://doi.org/10.1101/2021.01.08.425848>, 08 January 2021, preprint: not peer reviewed.
 136. Marshall, A.D., Bailey, C.G., Champ, K., Vellozzi, M., O'Young, P., Metierre, C., Feng, Y., Thoeng, A., Richards, A.M., Schmitz, U. *et al.* (2017) CTCF genetic alterations in endometrial carcinoma are pro-tumorigenic. *Oncogene*, **36**, 4100–4110.
 137. Pugacheva, E.M., Rivero-Hinojosa, S., Espinoza, C.A., Mendez-Catala, C.F., Kang, S., Suzuki, T., Kosaka-Suzuki, N., Robinson, S., Nagarajan, V., Ye, Z. *et al.* (2015) Comparative analyses of CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions. *Genome Biol.*, **16**, 161.
 138. Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schopflin, R., Kraft, K., Kempfer, R., Jerkovic, I., Chan, W.L. *et al.* (2016) Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, **538**, 265–269.
 139. Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R. *et al.* (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, **161**, 1012–1025.
 140. Narendra, V., Bulajic, M., Dekker, J., Mazzoni, E.O. and Reinberg, D. (2016) CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes Dev.*, **30**, 2657–2662.
 141. Rasko, J.E., Klenova, E.M., Leon, J., Filippova, G.N., Loukinov, D.I., Vatolin, S., Robinson, A.F., Hu, Y.J., Ulmer, J., Ward, M.D. *et al.* (2001) Cell growth inhibition by the multifunctional multivalent zinc-finger factor CTCF. *Cancer Res.*, **61**, 6002–6007.
 142. Tiffen, J.C., Bailey, C.G., Marshall, A.D., Metierre, C., Feng, Y., Wang, Q., Watson, S.L., Holst, J. and Rasko, J.E. (2013) The cancer-testis antigen BORIS phenocopies the tumor suppressor CTCF in normal and neoplastic cells. *Int. J. Cancer*, **133**, 1603–1613.
 143. Bailey, C.G., Metierre, C., Feng, Y., Baidya, K., Filippova, G.N., Loukinov, D.I., Lobanenkova, V.V., Semaan, C. and Rasko, J.E. (2018) CTCF expression is essential for somatic cell viability and protection against cancer. *Int. J. Mol. Sci.*, **19**, 3832.
 144. Wong, A.C.H., Rasko, J.E.J. and Wong, J.J. (2018) We skip to work: alternative splicing in normal and malignant myelopoiesis. *Leukemia*, **32**, 1081–1093.
 145. Monteuis, G., Schmitz, U., Petrova, V., Kearney, P.S. and Rasko, J.E.J. (2021) Holding on to junk bonds: intron retention in cancer and therapy. *Cancer Res.*, **81**, 779–789.