Topology regulatory elements: from shaping genome architecture to gene regulation

Current Opinion in Structural Biology

Topology regulatory elements: from shaping genome architecture to gene regulation

--Manuscript Draft--

Manuscript Number:	COSTBI-D-23-00091R1
Full Title:	Topology regulatory elements: from shaping genome architecture to gene regulation
Article Type:	VSI: 3D Genome Chromatin Organization and Regulation (2023)
Short Title:	Topological regulatory element function
Keywords:	Topological regulatory elements; 3D genome folding; chromatin fibre structure; gene expression; Transcription
Corresponding Author:	Hannah Katherine Long, DPhil The University of Edinburgh Edinburgh, Edinburgh UNITED KINGDOM
Corresponding Author's Institution:	The University of Edinburgh
Corresponding Author E-Mail:	hannah.long@ed.ac.uk
First Author:	Liang-Fu Chen, PhD
Order of Authors:	Liang-Fu Chen, PhD
	Hannah Katherine Long, DPhil
Abstract:	The importance of 3D genome topology in the control of gene expression is becoming increasingly apparent, while regulatory mechanisms remain incompletely understood. Several recent studies have identified architectural elements that influence developmental gene expression by shaping locus topology. We refer to these elements as topological regulatory elements (TopoREs) to reflect their dual roles in genome organisation and gene expression. Importantly, these elements do not harbour autonomous transcriptional activation capacity, and instead appear to facilitate enhancer-promoter interactions, contributing to robust and precise timing of transcription. We discuss examples of TopoREs from two classes that are either dependent or independent of CTCF binding. Importantly, identification and interpretation of TopoRE function may shed light on multiple aspects of gene regulation, including the relationship between enhancer-promoter proximity and transcription, and enhancer-promoter specificity. Ultimately, understanding TopoRE diversity and function will aid interpretation of how human sequence variation can impact transcription and contribute to disease phenotypes.
Author Comments:	

Cover Letter

MRC Human Genetics Unit
University of Edinburgh
Edinburgh
EH4 2XU
15 September 2023

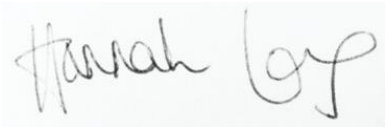
To Professors Geneviève Almouzni and Juanma Vaquerizas,

My co-author Liang-Fu Chen and I are pleased to submit our revised Review article entitled “Topology regulatory elements: from shaping genome architecture to gene regulation”. We were pleased that the article is accepted pending minor revisions. We thank the reviewers and editorial team for their helpful and insightful reviews and for their handling of our manuscript. We believe that incorporation of the comments and suggestions has improved the clarity and quality of discussion in the article.

Our revised submission includes line-by-line replies to reviewer comments, a text editable (docx) manuscript file, two updated figures (Figures 1 and 3), a compiled manuscript PDF with 3 main figures and 1 box figure embedded inline, individual PDFs of high-resolution versions of each figure and box figure.

We thank you both again for the invitation to participate in the **3D Genome Chromatin Organization and regulation** section (2023) of Current Opinion in Structural Biology.

Sincerely,



Hannah Long
Programme Leader Track Group Leader
T: +44 (0)131 651 8558
E: hannah.long@ed.ac.uk

Response to comments for article "Topology regulatory elements: from shaping genome architecture to gene regulation"

We thank the reviewers for their comments and suggested edits which we believe have improved the manuscript. We have updated the text to incorporate these suggestions and have outlined our changes below.

Reviewer 1: I read with great interest the manuscript by Chen et al. entitled "Topology regulatory elements: from shaping genome architecture to gene regulation". In this manuscript and based on recent publications, the authors proposed the existence of a set of regulatory elements (topology regulatory elements (TopoRE)) that contribute to gene regulation by controlling 3D chromatin architecture and facilitating enhancer-gene communication. The manuscript is very well written and illustrated with clear and informative figures. The concept of TopoRE is timely and helps understanding a recent body of work that has changed the current understanding of gene regulation and enhancer function. Overall, I fully support publication of the manuscript and I only have a couple of comments/suggestions:

- The authors should consider commenting on the fact that there are many enhancers, even distal ones, that seem to be able to faithfully control the expression of their target genes independently of both CTCF and tethering elements (e.g. CGI). The authors could comment on how these enhancers might be able to execute their regulatory function: (i) binding of TF (e.g. YY1) or cofactors (LDB1) that promoter enhancer-gene contacts (Deng et al, 2012; Weinstraub et al, 2017); (ii) non-linear relationship between gene expression and enhancer-gene contact frequency enabling high gene expression despite low enhancer-gene contact frequency (Zuin et al, 2022); (iii) enhancers controlling gene expression without physical contacts, etc (Karr et al., 2022).

We agree that it is valuable to comment upon other non-TopoRE based mechanisms for enhancer function. We have noted this point in the final section of the manuscript:

"Of note, a number of mechanisms have been proposed for enhancer-promoter communication which are independent of CTCF or TopoRE function, including transcription factor mediated interactions or diffusion of modified factors from an active enhancer to a promoter (Deng et al., 2012; Weintraub et al., 2017; Furling and Levine, 2018; Karr et al., 2022].

- From a disease point of view, the authors focused on the potential involvement of TopoRE in loss-of-function mechanisms. However, it is also conceivable that TopoRE might play an important role if gain-of-function pathomechanisms by enabling the establishment of strong contacts between enhancers and non-target genes upon loss of TAD boundaries/insulators (for example due to structural variants).

We agree that this is an interesting consequence to discuss, which we have incorporated into the manuscript, highlighting an example from Pachano et al.

"In the context of TAD boundary perturbations (due to a structural variant for example), a TopoRE could drive a gain-of-function pathological phenotype by forming novel contacts between enhancers and a non-target gene promoter (e.g. Figure 2A, lower)"

Reviewer 2: In the short Review "Topology regulatory elements: from shaping genome architecture to gene regulation" Chen and Long discuss recent advances in our understanding of a class of cis-regulatory elements whose primary role is shaping the 3D folding of the genome and what their influence on gene expression might be.
This is a timely and interesting topic as such type of regulatory elements might be of some importance for genome function/gene regulation. However, they have not been extensively

investigated, likely due to the difficulty to assess their function with common methods developed for enhancer/promoter function.

I find the proposal of a distinct term, "TopoRE," for a class of elements that might assist enhancers quite appealing. Nonetheless, caution is needed, particularly due to their strong connection with CTCF-binding. There's a risk of blurring the lines between TopoREs and TAD boundaries, a concern that is particularly evident in the discussion of the Pitx1 paragraph.

The dual role of CTCF, acting as both an insulator and a facilitator of enhancer-promoter interactions, further complicates the distinction between TopoREs and boundaries. For instance, a TopoRE might serve as a boundary, and vice versa. To bolster the clarity of their manuscript, the authors should make a concerted effort to differentiate between these concepts and avoid muddling the terminology.

Major:

TAD boundaries/borders and TopREs at the Pitx1 locus
The authors adopt Hung et al.'s perspective, suggesting that loop stacking at the Pitx1 locus enables the Pen enhancer to transcend TAD boundaries. However, this interpretation is problematic, as the features in question are not true TAD boundaries. In fact, aligning with Hung et al.'s terminology contradicts the very essence of the TopoRE concept presented in this review.
Based on my understanding of the Pitx1 locus, the work of Kragesteen et al indicates that the locus does "not show the presence of tissue- or species-invariant TADs or contact domains encompassing the entire Pitx1 regulatory landscape." (I double-checked). Yet the locus has sub-domains with regulatory anchors (which are the "TAD boundaries" Hung et al refer to).

I think it is very important for the logic of TopoREs not to use TAD boundary terminology when it is not appropriate. CTCF sites can insulate enhancers from promoters (as they would do at a TAD boundary) or they can facilitate enhancer-promoter interaction (as they would do at a TopoRE). The Regulatory Anchors of the Pitx1 locus act more like CTCF-sites (TopoREs!) within the Sox9 domain that facilitate enhancer activity and less like CTCF-boundaries at the flanks of a TAD (such as at the Sox9 locus). I think emphasizing such similarities would strengthen the argumentation instead of confusing less expert readers with ambivalent/imprecise terminology.

We thank the reviewer for their careful argumentation surrounding the concept of domain-stacking at the *Pitx1* locus. Our intention with this section of the manuscript was to discuss the proposal that some boundary elements can facilitate long-range interactions, while also insulating inter-domain interactions.

In our view, our ability to define TADs versus sub-domains is somewhat challenging, and the variability among TAD-calling algorithms has proven to be substantial (**Zufferey et al.**, **2018**). TADs have been broadly defined as self-interacting domains of chromatin, comprised of regions of greater intra-region compared to inter-region contact. We agree that there is ambiguity at the *Pitx1* locus whether the ‘triangles on the diagonal’ domains of self-interacting chromatin constitute a TAD or a sub-domain. There do appear to be regions of insulation between the three self-interacting domains of chromatin spanning the *Pitx1*-Pen locus however ((**Kragesteen et al.**, **2018**) – see Figure 4). Therefore, the CTCF sites between these domains appear to confer both insulation and long-range enhancer-promoter interactions via the stacking mechanisms proposed in (**Hung et al.**, **2023**).

This contrasts with the *SOX9* locus in CNCCs, where we see uniform domain-spanning interactions between the two stripe-associated structural elements (SSEs), in addition to the two strong stripes emanating from the SSEs ((**Chen et al.**, **2023**) – see Figure 1E, lower).

Therefore, the TopoRE elements at the *SOX9* locus appear to behave differently from those at the *Pitx1* locus. It is an interesting question to follow up why some CTCF sites act as TopoREs while some exhibit boundary/insulator function.

We have revisited the *Pitx1* paragraph and reworded to avoid ambiguity between TADs and sub-domains, using self-interacting domain instead. We believe our edits have addressed the reviewer's concerns about conflating TAD boundaries with proposed TopoRE function.

Changes to the text in this paragraph are highlighted in green in the revised manuscript. We have also updated Figure 1Di to replace "TAD boundaries" with "CTCF binding sites", and "Contacts across multiple TADs" to "Contacts across multiple domain boundaries".

Minor:

"These single-fibre topologies further allowed us to determine that SSEs promote the central positioning of both, the SOX9 gene and extremely distal ECs, thus facilitating the interaction of the promoter with the entire TAD."
Consider revising as this is unclear if one is not deep into the project/locus. Maybe split the sentence; refer to the geometric center; are ECs enhancers?

We have split and re-worded the sentence below following this suggestion. We believe this improves readability and clarity.

Original: "These single-fibre topologies further allowed us to determine that SSEs promote the central positioning of both the *SOX9* gene and extremely distal ECs, thus facilitating interaction of the promoter with the entire TAD."

Updated: "The ORCA single-fibre topologies further allowed us to determine that SSEs promote the positioning of the *SOX9* gene in the geometric centre of the domain. We propose that this locus topology facilitates interaction of the promoter with the entire TAD."

While the author's focus on their previous work on the Sox9 locus is logical and delivers a fine line of argumentation, it might be beneficial be more inclusive of other research. For instance, we (Despang et al (2019)) deleted CTCF sites (likely TopoRE-CTCF sites) at exactly this locus, with little effect on gene expression in the limb bud. Exploring this discrepancy, which can possibly be explained by the differential importance of TopoRE for NCC-specific enhancers versus limb enhancers, would poride a more well-rounded discussion.

We agree that it will be interesting to the reader to be aware of dissection of CTCF function at the *Sox9* locus during mouse limb development. We agree that it's currently not clear what drives the disparity in phenotypes between human CNCC single CTCF site deletion and mouse combinatorial CTCF site and boundary deletion. In CNCCs, we only observed a 20% reduction in *SOX9* expression upon single CTCF site ablation, which was detectable by allele-specific ddPCR. This level of perturbation may not see easily detectable by LacZ reporter assays, and qRT-PCR. Another possibility for the disparity may be due differences in cell-type tested, or evolutionary divergence between human and mouse.

To highlight the work in mouse limb buds, we have added the following text:

"Interestingly, similar stripe-like features were previously noted from Capture Hi-C in E12.5 mouse limb buds (**Despang et al.**, **2019**). Future singular deletion of the orthologous SSE1.35 element in mouse development will help to reveal cell-type specificity and evolutionary conservation of SSE1.35 function."

Of note, we have also updated Figure 3C to remove the unnecessary heading “LOF mutation”, and to add a “?” to “Sub-phenotypic?” to indicate that this is speculative that loss of function mutations in the distal *SOX9* TopoRE (SSE1.35) may not show an overt morphological phenotype.

References

Chen, L.-F., Long, H. K., Park, M., Swigut, T., Boettiger, A. N., & Wysocka, J. (2023). Structural elements promote architectural stripe formation and facilitate ultra-long-range gene regulation at a human disease locus. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2023.03.009

Despang, A., Schöpflin, R., Franke, M., Ali, S., Jerković, I., Paliou, C., Chan, W.-L., Timmermann, B., Wittler, L., Vingron, M., Mundlos, S., & Ibrahim, D. M. (2019). Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nature Genetics*, *51*(8), 1263–1271. https://doi.org/10.1038/s41588-019-0466-z

Hung, T.-C., Kingsley, D. M., & Boettiger, A. (2023). Boundary stacking interactions enable cross-TAD enhancer-promoter communication during limb development. *bioRxiv*, 2023.02.06.527380. https://doi.org/10.1101/2023.02.06.527380

Kragesteen, B. K., Spielmann, M., Paliou, C., Heinrich, V., Schöpflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A. M., Jerković, I., Harabula, I., Guckelberger, P., Pechstein, M., Wittler, L., Chan, W.-L., Franke, M., Lupiáñez, D. G., Kraft, K., Timmermann, B., … Andrey, G. (2018). Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nature Genetics*, *50*(10), 1463–1473. https://doi.org/10.1038/s41588-018-0221-x

Zufferey, M., Tavernari, D., Oricchio, E., & Ciriello, G. (2018). Comparison of computational methods for the identification of topologically associating domains. *Genome Biology*, *19*(1), 217. https://doi.org/10.1186/s13059-018-1596-9

Topology regulatory elements: from shaping genome architecture to gene regulation

Liang-Fu Chen[a], Hannah Katherine Long[b]

[a] Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

[b] MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Crewe Road, Edinburgh, UK

Correspondence: hannah.long@ed.ac.uk

Short title

Topological regulatory element function

Keywords

Topological regulatory elements, 3D genome folding, chromatin fibre structure, gene expression, transcription

Abstract

The importance of 3D genome topology in the control of gene expression is becoming increasingly apparent, while regulatory mechanisms remain incompletely understood. Several recent studies have identified architectural elements that influence developmental gene expression by shaping locus topology. We refer to these elements as topological regulatory elements (TopoREs) to reflect their dual roles in genome organisation and gene expression. Importantly, these elements do not harbour autonomous transcriptional activation capacity, and instead appear to facilitate enhancer-promoter interactions, contributing to robust and precise timing of transcription. We discuss examples of TopoREs from two classes that are either dependent or independent of CTCF binding. Importantly, identification and interpretation of TopoRE function may shed light on multiple aspects of gene regulation, including the relationship between enhancer-promoter proximity and transcription, and enhancer-promoter specificity. Ultimately, understanding TopoRE diversity and function will aid interpretation of how human sequence variation can impact transcription and contribute to disease phenotypes.

Diverse non-coding functional elements regulate gene expression

Regulatory elements within the non-coding portion of the genome play important roles in mediating the timing, location and levels of transcription from target genes. The importance of these elements is highlighted both by Mendelian genetic disorders driven by non-coding mutations that disrupt regulatory elements [1,2] and genome-wide association studies for complex traits which find that most contributing variants are in the non-coding genome [3,4]. One of the most well-studied classes of regulatory elements are enhancers, autonomous cis-regulatory sequences that encode for clusters of transcription factor binding sites that can activate gene expression at a distance in a tissue-specific manner [5–7]. Enhancers can function outside of their native context to drive expression of a reporter gene, therefore one strategy for identifying enhancers has been to use episomal reporter assays including luciferase assays, *in vivo* transgenic reporters or massively parallel reporter assays (MPRAs) [8,9]. However, it is becoming increasingly apparent that regulatory elements lacking autonomous activity, that would be overlooked in these assays, can also act to facilitate or boost classical enhancers (for example, [10–13]), or can regulate gene expression by altering 3D genome topology.

In this review, we will focus on an emerging class of architectural elements with a proposed direct function in facilitating enhancer-promoter interactions and gene expression. We refer to these regions as topological regulatory elements (TopoREs), genetically encoded sequences that facilitate regulation of gene expression without having autonomous enhancer activity, by supporting enhancer-promoter communication or otherwise impacting 3D chromatin folding (see **Box 1**). These TopoREs can be within, close to, or 10s of kilobases away from enhancers or promoters, and their function can span across topologically associated domains (TADs). We will explore in detail several recent examples of TopoREs that shape 3D genome architecture, and impact gene expression, categorised by their dependence on one particular trans-acting factor, CTCF. Finally, we discuss how these elements update our view of the importance of genome structure for gene expression and human disease.

CTCF-dependent topological regulatory elements facilitate long-range gene regulation

CCCTC-binding factor (CTCF) is an 11-zinc finger DNA-binding protein that plays a number of important functions including in VDJ recombination and organising 3D chromatin architecture [14,15]. CTCF modulates chromatin organisation together with cohesin, a ring-shaped complex that compacts chromatin through the active process of loop extrusion [16] (**Figure 1A**). In this process, cohesin is loaded onto chromatin and processively extrudes loops of DNA until stalled through collision with a barrier, such as CTCF bound to DNA in a convergent orientation [17,18]. Through its interactions with cohesin, CTCF has been implicated in a range of topological functions, for example in the formation of topologically associating domains (TADs) and insulated neighbourhoods, which broadly act to constrain regulatory activity to within a given locus [19].

CTCF binding within gene regulatory elements has been further implicated in playing a more direct role in facilitating interactions between promoters and enhancers. For example, in mouse Th2 cells, CTCF-binding at enhancers conferred an increased tendency for interaction with promoters, and buffered transcriptional noise [20]. Furthermore, in mouse embryonic stem cells (mESCs), CTCF bound at promoters was shown to facilitate enhancer-mediated gene regulation, especially across long distances for genes without many enhancers in close proximity [21]. Mechanistically, CTCF sites adjacent to or within these regulatory elements may facilitate loop extrusion-dependent scanning across the regulatory domain. Transcription factors [22,23] or RNA polymerase II [24] at another distal regulatory element could then stall cohesin to facilitate long-range linking between the enhancer and promoter to promote gene expression. Together, for a subset of genes, CTCF binding at either the promoter or distal enhancer element can provide robustness to gene activation across long-range.

In our recent work, we took an in-depth single-locus approach to explore mechanisms of extreme long-range gene regulation at the *SOX9* locus where craniofacial enhancer clusters (ECs) lie over 1.2 megabases upstream of the *SOX9* gene [25,26]. Using optical reconstruction of chromatin architecture (ORCA) imaging [27] (**Figure 1B**), and plotting ensemble-average interaction frequencies across the domain, we identified two stripes of domain-spanning interactions emanating from the *SOX9* promoter and EC locus (**Figure 1C-i**). These elements, which we named “stripe associated structural elements” (SSEs), were dependent on CTCF binding sites for both topological function and for maintaining normal expression levels of *SOX9* [25]. The single chromatin fibre nature of the ORCA imaging enabled us to observe hugely dynamic and variable locus topologies for cell states across our *in vitro* differentiation time course, from human embryonic stem cells to cranial neural crest cells (**Figure 1C-ii**). We identified that the differences between averaged *SOX9* domain structures for these cell states arose from alterations of sampled frequencies in domain topologies rather a shift between two static preferred structures. The ORCA single-fibre topologies further allowed us to determine that SSEs promote the positioning of the *SOX9* gene in the geometric centre of the domain. We propose that this locus topology facilitates interaction of the promoter with the entire TAD. While a number of mechanisms may be at play to drive SSE function, we explored the role of loop extrusion through polymer simulations and determined that a multi-loop structure was consistent with the chromatin fibre topologies we

observed, where multiple extruded loops stack across the domain, bridging the long distance between the distal enhancers and the *SOX9* gene (**Figure 1C**). Ultimately, we propose that this conformation facilitates gene regulation by promoting sampling of the regulatory domain by the *SOX9* promoter. Interestingly, similar stripe-like features were previously noted from Capture Hi-C in E12.5 mouse limb buds [28]. Future singular deletion of the orthologous SSE1.35 element in mouse development will help to reveal cell-type specificity and evolutionary conservation of SSE1.35 function.

Stacking of loops as a mechanism for extreme long-range regulation has been further extended recently to span across multiple TADs or contact domains [29]. TADs or contact domains have been broadly defined as a domains of higher intra-region contact with reduced inter-region contact, seen as a triangle on the diagonal in HiC heatmaps [30,31]. ~~TAD~~ Boundaries between these domains have been considered to both facilitate intra-domain interactions and insulate genes from the regulatory influence of enhancers from adjacent domains. Indeed, patient mutations perturbing these boundary elements are associated with disease through the resultant mis-regulation of target genes [2,32,33]. Controverting this paradigm, there are examples of enhancer action spanning across ~~TAD~~ domain boundaries. For example, a distal super enhancer at the *Hoxa* locus important for ear development that functions across a TAD boundary [34]. Additionally, at the *Pitx1* locus, ~~or~~ the Pen enhancer ~~which~~ regulates *Pitx1* gene expression across ~~at~~ long-distance in mouse hindlimb development, traversing three self-interacting contact domains ~~over two TAD boundaries~~ [35]. In this second example, ORCA imaging revealed multi-way interactions between boundary elements in a single chromatin fibre. This boundary hub is thought to bring the boundary-proximal enhancer and promoter into close proximity. Therefore, ~~in this case~~ while the intervening ~~TAD~~ domain boundaries between *Pitx1* and the Pen enhancer insulate contacts between the self-interacting domains, they are also proposed to facilitate ~~rather than insulate~~ long-range enhancer function through ~~TAD~~ domain boundary-stacking ~~mediated by loop-extrusion~~ (**Figure 1D**). ~~This boundary hub is thought to facilitate border bypass such that enhancers and promoters in distal TADs are brought into close proximity. Counter-intuitively, the authors suggest that this function may be driven by stronger TAD boundaries in hindlimbs where increased CTCF binding is observed compared to forelimbs [29,36].~~ Given that ~~a reduction in CTCF binding at~~ these elements ~~reduces their capacity to~~ facilitate enhancer-promoter interaction through modulation of 3D genome folding, we propose that in this case ~~TAD~~ the *Pitx1* locus domain boundaries are behaving as TopoREs. While many ~~TAD~~ boundary elements ~~ies~~ may not meet our criteria as a TopoRE, a prediction of this model is that enhancers and promoters located near to ~~TAD~~ domain ~~borders~~ boundaries are more likely to be subject to this type of regulation. These observations therefore lay the ground for other extremely-distal enhancer-promoter pairs to be identified, and provides a framework for understanding other examples of ~~TAD~~ domain-spanning enhancer action.

CTCF plays diverse roles in the control of 3D chromosome topology and gene regulation, however it remains poorly understood how different CTCF sites act in a distinct manner to shape local chromosome topology. For example, at the *SOX9* locus there are many more sites bound by CTCF than those required for SSE function [25]. Possible contributing factors influencing CTCF topological function at distinct sites include co-binding of other trans-regulatory factors [37,38],

location of extruder loading, proximity of regulatory elements to the CTCF bound region, and the affinity of CTCF binding to the element itself. It is conceivable that these properties could then be regulated across diverse tissue-types and developmental stages to change the nature of a topological regulatory element, for example from an insulator to a structural element facilitating enhancer-promoter interactions. It remains to be seen therefore whether CTCF sites involved in mediating enhancer-promoter interactions play a pleiotropic regulatory role across all tissues where active gene regulation is occurring, or whether TopoRE elements exhibit cell-type specificity as is seen for enhancers.

Interplay between boundary elements and CTCF-independent topological regulatory elements provide specificity and precise timing for developmental gene expression

While many enhancer-promoter interactions have been shown to be mediated by CTCF and cohesin, there are several studies reporting that a distinct group of TopoREs can shape genome structure and regulate gene expression in a CTCF-independent manner. Indeed, the majority of genes are able to recruit enhancers and initiate transcription normally in the face of acute degradation of CTCF [21,39–42]. In mESCs, CpG islands (CGIs) have been shown to promote long-range communication between promoters with large CGIs and poised enhancers (PEs) associated with an orphan CGI (oCGI) [43]. In total, around 60–80% of PEs in mouse ESCs are located within 3 kilobases (kb) of an oCGI and deleting oCGIs at PEs reduces the expression of their target genes. However, these oCGIs do not increase the transcriptional activity of PEs. Instead, they facilitate PE interactions with target genes, and only promoters with large CGI clusters show a transcriptional responsiveness to PEs. These CGI-mediated interactions can be blocked by TAD boundaries and thus it was proposed that the combination of CGI-mediated long-range communication and the insulation from TAD boundaries provides specificity in the induction of certain genes during development [43] (**Figure 2A**).

A similar interplay between boundary elements and TopoREs has also been shown to shape the specificity and timing of developmental gene expression during *Drosophila* development. Leveraging high-resolution Micro-C data, Batut et al. identified two distinct classes of architectural elements that shape genome structure and regulate gene expression during a critical 60 minutes of development prior to gastrulation [44]. Insulators act to prevent spurious interactions, while distal tethering elements (DTEs) foster appropriate enhancer-promoter interactions. One third of all focal contacts detected by Micro-C connect promoters of protein-coding genes to DTEs, typically spanning tens of kilobases. Tethering elements identified at the *Scr-Antp* region overlapped with regions previously identified as facilitating enhancer-promoter selectivity for the *Scr* gene [45,46]. Overall, DTEs were observed at many critical developmental loci reflecting a potentially broad mechanism for mediating enhancer-promoter interactions important for transcriptional timing. DTE elements display no autonomous enhancer activity in the early embryo. However, using live cell imaging of transcription, the authors demonstrated that DTEs foster fast activation of transcriptional kinetics required for appropriate developmental progression, while boundaries prevent interference of cis-regulatory elements between neighbouring TADs. Therefore, DTEs meet our criteria as topological regulatory elements and the

interplay between boundaries and DTEs is proposed to confer the specificity and timing of developmental gene transcription in the developing *Drosophila* embryo (**Figure 2B**).

Orthologous to CTCF-dependent TopoREs, diverse transcription factors/ cofactors bind to CTCF-independent TopoREs and can regulate their functions in different cell types or developmental stages. While the dynamics of locus topology, and constant re-establishment of TAD structure is becoming apparent [25,47,48], Pachano et al observed that PEs/oCGIs are already in close proximity on average to their target promoter/CGI before gene activation. These interactions are dependent on polycomb complexes in mESCs [49] and are proposed to be maintained by transcription factors and co-factors once the PEs become active in anterior neural progenitor cells [43]. Similarly, DTEs in Drosophila are bound by pioneer factors Trithorax-like (Trl), grainyhead (grh), and zelda (zld) [44] which appear to mediate enhancer-promoter interactions prior to gene activation. Indeed, zelda has been shown to mediate cis-regulatory chromatin interactions that arise before the formation of TADs and gene activation during early *Drosophila* development [50]. Importantly, this binding of CTCF-independent TopoREs by different transcription factors/cofactors appears to shape genome structure for subsequent gene expression during development. These TopoREs therefore promote the frequent sampling of permissive regulatory topologies whereby enhancers are already in proximity to their target genes prior to gene activation to ensure precise timing of developmental gene expression once the enhancer is turned on (**Figure 3A**). Of note, this is reminiscent of the observed proximity seen to link the Shh gene to the distal ZRS limb enhancer in non-expressing cell-types [51–53]. Together, TopoREs provide an extra layer for gene regulation, with the interplay between TopoREs and boundary/insulator elements providing specificity and precision of timing for developmental gene expression.

The importance of 3D chromatin organisation for gene expression and human disease

Above, we have described a breadth of topological regulatory elements that influence developmental and homeostatic gene expression through an influence on 3D genome folding and enhancer-promoter communication. At the *SOX9* and *Shh* regulatory loci, loss of CTCF binding sites leads to a general increase in pairwise distances across the domain [25,53]. Therefore, a key role of TopoREs may be also to compact a regulatory locus to promote frequency of enhancer-promoter interactions. This role of TopoREs may be most relevant for the activity of distal enhancers, as recent studies have revealed a differing requirement for cohesin for activation from proximal versus distal enhancers [54,55]. This suggests a distinct requirement for enhancer-promoter tethering or locus compaction for enhancer function across different genomic distances. In this context, TopoREs may play a greater role in facilitating gene regulation for more distal regulatory interactions, while enhancer-promoter proximity is less of a limiting feature for proximal regulatory elements. Of note, a number of mechanisms have been proposed for enhancer-promoter communication which are independent of CTCF or TopoRE function, including transcription factor mediated interactions or diffusion of modified factors from an active enhancer to a promoter [5,23,56,57].

An understanding of the role of TopoREs in development is of great importance in the context of a full appreciation of the impact of non-coding mutations on human disease. Increasing evidence

points to a critical role of 3D chromatin organisation during organismal development and cell differentiation, and gross deregulation of chromatin topology and TAD architecture is associated with development of human diseases [33,58]. In the context of TAD boundary perturbations (due to a structural variant for example), a TopoRE could drive a gain-of-function pathological phenotype by forming novel contacts between enhancers and a non-target gene promoter (e.g. Figure 2A, lower) [43]. Furthermore, it is likely that disruption of TopoRE function could also deregulate gene regulation to such an extent to cause developmental defects and disease. For example, the importance of CTCF for facilitating enhancer-promoter interactions is underlined by the discovery that patients with acheiropodia harbour mutations ablating a cluster of CTCF sites upstream of the Shh limb enhancer, ZRS, which facilitate interaction with the Shh gene [59]. In addition, deletion of tethering elements at the *Scr* locus in Drosophila caused a delay of precisely timed developmental gene expression. While the levels of transcription ultimately catch up, this delay impacts sex comb development proportional to the degree of transcriptional impact [44,60–62] (**Figure 3A and B**). At the *SOX9* locus, ablation of extreme long-range enhancers has a tissue-specific effect on lower jaw development likely due to a combination of tissue-specific dosage sensitivity to *SOX9* perturbation and spatially restricted domains of enhancer activity. While no patients have yet been identified with SSE perturbation alone, ablation of either of the *SOX9* locus structural elements perturbs expression levels in CNCC cell culture models. Loss of SSE function could therefore be sub-phenotypic but sensitise facial development to other genetic or environmental perturbations (**Figure 3C**). As discussed above, it remains to be seen whether TopoRE loss would have pleiotropic effects in disease or have tissue-specific functions. Future work will be required to determine the topology and SSE-status of other *SOX9* expressing cell types.

Due to the lack of autonomous regulatory activity, and in many cases a lack of uniquely bound trans-acting factors that distinguish them from non-regulatory elements, there are currently limited ways to identify TopoREs genome wide. Reporter assays are dependent on autonomous regulatory capacity, and without a unique molecular signature, ChIP-seq based methods cannot identify these elements in a high throughput manner. However, as our understanding and discovery of these elements increases, greater in-depth exploration of the genomic and trans-acting determinants of chromatin looping and topological regulatory element activity will further illuminate TopoRE function. As an example, the *Sox2* locus has been intensively studied, and highlights how focussing on a single locus can uncover fundamental features of cis-regulatory landscapes which can then be explored genome-wide. Both mutational screening or synthetic engineering of the *Sox2* locus in mESC have started to uncover the grammar of CTCF binding sites for CTCF function, as well as other factors that modulate chromatin architecture [12,40,63,64]. These single-locus studies together with CRISPR-based genome editing [65,66] and high-throughput screens for co-factors of CTCF [37] will greatly increase our understanding of TopoREs more broadly. Another way to identify architectural elements in the genome is *in silico* discovery of DNA sequence features that mediate distal interactions by deep learning (DL) models coupled with genome-wide 3C based sequencing data. Many recent studies have been able to apply DL models to train sequence-based predictors of chromatin looping and to identify specific sequence features that may facilitate physical contacts between distal genomic regions (see reviews from [67,68]). Extending this analysis to identify features unique to TopoREs, combined

with experimental validation, will be a powerful tool to study the DNA sequence grammar underlying TopoREs.

In concert with additional features of genome structure such as domain boundaries and insulators, TopoREs confer robustness and specificity in gene transcription. Ultimately, an improved understanding of how TopoREs are regulated during development will shed light on how alteration of these elements can impact gene expression and contribute to disease phenotypes.

Acknowledgments

We thank Professor Alistair Boettiger, Professor Joanna Wysocka and Dr Kirsty Uttley for their critical reading and comments on this manuscript. We would like to apologise to colleagues whose work we could not reference due to space constraints. HL is supported by a Medical Research Council (MRC) UK University Unit grant (MC_UU_00035/12).

Annotated references

Outstanding interest (**)

Chen, L.-F., Long, H.K., Park, M., Swigut, T., Boettiger, A.N., and Wysocka, J. (2023). Structural elements promote architectural stripe formation and facilitate ultra-long-range gene regulation at a human disease locus. Mol. Cell. 10.1016/j.molcel.2023.03.009.

Using optical reconstruction of chromatin architecture (ORCA) and polymer simulations, this work uncovered stripe-associated structural elements that mediate a 3D multi-loop topology and facilitate enhancer-promoter interactions at the *SOX9* locus in human cranial neural crest cells.

Batut, P.J., Bing, X.Y., Sisco, Z., Raimundo, J., Levo, M., and Levine, M.S. (2022). Genome organization controls transcriptional dynamics during development. Science *375*, 566–570. 10.1126/science.abi7178.

This work identified tethering elements that promote enhancer-promoter interactions and control the timing of developmental gene expression in the developing *Drosophila* embryo.

Kubo, N., Ishii, H., Xiong, X., Bianco, S., Meitinger, F., Hu, R., Hocker, J.D., Conte, M., Gorkin, D., Yu, M., et al. (2021). Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. Nat. Struct. Mol. Biol. *28*, 152–161. 10.1038/s41594-020-00539-5.

This work uncovered promoter-proximal CTCF binding and provided evidence that CTCF directly promotes enhancer-promoter and promoter-promoter interactions for lineage-specific genes.

Hung, T.-C., Kingsley, D.M., and Boettiger, A. (2023). Boundary stacking interactions enable cross-TAD enhancer-promoter communication during limb development. 2023.02.06.527380. 10.1101/2023.02.06.527380.

These authors revealed a TAD boundary stacking mechanism for cross-TAD enhancer-promoter communication and demonstrated that TAD boundaries can both facilitate and prevent cis-regulatory interactions.

Pachano, T., Sánchez-Gaya, V., Ealo, T., Mariner-Faulí, M., Bleckwehl, T., Asenjo, H.G., Respuela, P., Cruz-Molina, S., Muñoz-San Martín, M., Haro, E., et al. (2021). Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. Nat. Genet. *53*, 1036–1049. 10.1038/s41588-021-00888-x.

This study showed that there are orphan CpG islands adjacent to most of the poised enhancers in mouse embryonic stem cells (mESCs) and that these oCGIs promote permissive regulatory topologies for developmental genes in mESC.

Special interest (*)

Ortabozkoyun, H., Huang, P.-Y., Cho, H., Narendra, V., LeRoy, G., Gonzalez-Buendia, E., Skok, J.A., Tsirigos, A., Mazzoni, E.O., and Reinberg, D. (2022). CRISPR and biochemical screens identify MAZ as a cofactor in CTCF-mediated insulation at Hox clusters. Nat. Genet. *54*, 202–212. 10.1038/s41588-021-01008-5.

This study used CRISPR and biochemical screens to identify CTCF cofactors in mESCs and neurons.

Brosh, R., Coelho, C., Ribeiro-dos-Santos, A.M., Ellis, G., Hogan, M.S., Ashe, H.J., Somogyi, N., Ordoñez, R., Luther, R.D., Huang, E., et al. (2023). Synthetic regulatory genomics uncovers enhancer context dependence at the Sox2 locus. Mol. Cell *83*, 1140-1152.e7. 10.1016/j.molcel.2023.02.027.

This study used synthetic regulatory genomics to repeatedly rewrite the *Sox2* locus in mESC, dissecting its overall architecture and delineating roles for individual regulatory elements.

Taylor, T., Sikorska, N., Shchuka, V.M., Chahar, S., Ji, C., Macpherson, N.N., Moorthy, S.D., de Kort, M.A.C., Mullany, S., Khader, N., et al. (2022). Transcriptional regulation and chromatin architecture maintenance are decoupled functions at the Sox2 locus. Genes Dev. *36*, 699–717. 10.1101/gad.349489.122.

Parallel work to Charkraborty 2023 and Huang 2021, this study investigated the role of CTCF mediated loops in chromatin topology and transcriptional regulation at *Sox2* locus in mESCs.

Chakraborty, S., Kopitchinski, N., Zuo, Z., Eraso, A., Awasthi, P., Chari, R., Mitra, A., Tobias, I.C., Moorthy, S.D., Dale, R.K., et al. (2023). Enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness. Nat. Genet. *55*, 280–290. 10.1038/s41588-022-01295-6.

Parallel work to Taylor 2022 and Huang 2021, this study investigated the role of CTCF mediated loops in chromatin topology and transcriptional regulation at *Sox2* locus in mESCs.

Huang, H., Zhu, Q., Jussila, A., Han, Y., Bintu, B., Kern, C., Conte, M., Zhang, Y., Bianco, S., Chiariello, A.M., et al. (2021). CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. Nat. Genet. *53*, 1064–1074. 10.1038/s41588-021-00863-6.

Parallel work to Charkraborty 2023 and Taylor 2022, this study investigated the role of CTCF mediated loops in chromatin topology and transcriptional regulation at *Sox2* locus in mESCs.

Espinola, S.M., Götz, M., Bellec, M., Messina, O., Fiche, J.-B., Houbron, C., Dejean, M., Reim, I., Cardozo Gizzi, A.M., Lagha, M., et al. (2021). Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. Nat. Genet. 53, 477–486. 10.1038/s41588-021-00816-z.

This study demonstrated that many cis-regulatory chromatin interactions arise before the formation of TADs and gene activation in a zelda-dependent manner during early *Drosophila* development.

Box 1 - Proposed definition and criteria for a topological regulatory element (TopoRE)

We suggest the following features for defining a topological regulatory (TopoRE) element:

I. a region of DNA that lacks autonomous enhancer activity (by reporter assay in a relevant cell-type);
II. genetic ablation leads to an alteration of 3D genome topology (e.g. E-P distance);
III. genetic ablation impacts target gene expression either in the current cell-state or during developmental progression.

Figure Legends

Figure 1 - Loop extrusion, CTCF and loop-stacking promote enhancer function across extreme distances

A) Schematic of loading of cohesin onto chromatin followed by loop extrusion. Loop extrusion is halted when cohesin collides with loop extrusion barriers or another cohesin complex. B) Tracing chromatin conformation by optical reconstruction of chromatin architecture (ORCA). A locus of interest is labelled with primary probes, with each probe-set marking desired segments along the locus distinguished with a unique barcode. Each barcode is imaged by sequentially introducing a readout oligo carrying a fluorophore. The 3D structure of the locus then reconstructed after rounds of imaging. C) i) Chen et al imaged the *SOX9* locus in human cranial neural crest cells and observed stripes at two stripe-associated structural elements (SSEs). ii) These stripes are proposed to form through a multi-loop model whereby loops stack wherever extruding cohesins happen to collide into one another across the domain, anchored at the SSEs. SSEs compact the TAD, draw the *SOX9* promoter into the centre of the 3D domain, and thereby facilitate the promoter interacting with enhancers across the domain. D) i) At the *Pitx1* locus in mouse hindlimbs, Hung et al observed contacts of the *Pitx1* gene spanning across two TAD boundaries to the distal Pen enhancer. ii) This TAD-spanning interaction is proposed to occur through stacking of TAD boundaries that bring the distal enhancer and promoter that are adjacent to TAD boundaries into close proximity. EC = enhancer cluster.

Figure 2 – Interplay between topological regulatory elements and boundary elements drives gene expression specificity

A) In mouse embryonic stem cells, orphan CpG islands (oCGIs) help to bridge poised enhancers to target promoters embedded within CGIs, as detected by 4C-seq. This is mediated by polycomb in mESC and proposed to be bridged by transcription factors upon transcriptional initiation after differentiation to anterior neural progenitors (AntNPCs). Loss of oCGI elements cause a reduction in this interaction, and a reduction of target gene expression during differentiation to anterior

neural progenitors (AntNPCs). While loss of TopoRE (oCGI) leads to a reduction in gene expression, loss of the TAD boundary can drive mis-expression of another gene embedded within a CGI in the adjacent TAD due to interaction compatibility with the nearby oCGI/PE. B) During *Drosophila* development, a tethering element (TE) at the *Scr* gene interacts with a distal tethering element (DTE) near an enhancer, as detected by Micro-C, bypassing an intervening self-interacting domain. Ablation of the DTE leads to a delayed developmental expression of *Scr*, while ablation of an intervening insulator element enables a regulatory element (Rep, AE1) to interact with the *Scr* gene also leading to transcriptional downregulation of *Scr*.

Figure 3 - Transcriptional and phenotypic consequences of topological regulatory element perturbation

A) Perturbation of a TopoRE can impact absolute expression levels (left, example of *SOX9* locus Chen et al), transcriptional timing (middle, Batut et al, Pachano et al) or can lead to loss of transcriptional precision (right, Ren et al). B) In wildtype *Drosophila* embryos, developmental expression of *Scr* is required for a normal number of sex comb bristles (a mean of 9.5), and heterozygous loss of *Scr* reduces this to an average of 6.3. Heterozygous loss of the distal Scr enhancer element (EE) reduces *Scr* expression, leading to fewer sex comb bristles (6-8 on average). Perturbation of a distal tethering element (DTE) adjacent to the EE enhancer leads to a delay in Scr induction (see A) and a subtle reduction in sex comb number. C) At the *SOX9* locus, stripe-associated structural elements (SSEs) facilitate *SOX9* expression in cranial neural crest cells. Heterozygous loss of *SOX9* function impacts all *SOX9* expressing tissues (simplified here to show face and limb expression) leading to severe phenotypes in both tissues. Loss of craniofacial distal enhancer elements (for example EC1.45) reduces *SOX9* expression only in the face, leading to phenotypes in PRS patients restricted to the lower jaw. Mutation of the SSE elements have a milder impact on *SOX9* expression in CNCCs, and it is predicted this may have sub-phenotypic consequences during development but may sensitise embryonic development to other environmental or genetic perturbations. Whether the SSE elements are tissue-specific in their function remains to be determined.

Declaration of Interest

The authors declare no conflict of interest.

References

1. Zhang F, Lupski JR: **Non-coding genetic variants in human disease**. *Hum Mol Genet* 2015, **24**:R102–R110.

2. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al.: **Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions**. *Cell* 2015, **161**:1012–1025.

3. Claringbould A, Zaugg JB: **Enhancers in disease: molecular basis and emerging treatment strategies**. *Trends Mol Med* 2021, **27**:1060–1073.

4. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, et al.: **Systematic Localization of Common Disease-Associated Variation in Regulatory DNA**. *Science* 2012, **337**:1190–1195.

5. Furlong EEM, Levine M: **Developmental enhancers and chromosome topology**. *Science* 2018, **361**:1341–1345.

6. Kim S, Wysocka J: **Deciphering the multi-scale, quantitative cis-regulatory code**. *Mol Cell* 2023, **83**:373–392.

7. Long HK, Prescott SL, Wysocka J: **Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution**. *Cell* 2016, **167**:1170–1187.

8. Inoue F, Ahituv N: **Decoding enhancers using massively parallel reporter assays**. *Genomics* 2015, **106**:159–164.

9. Kvon EZ: **Using transgenic reporter assays to functionally characterize enhancers in animals**. *Genomics* 2015, **106**:185–192.

10. Thomas HF, Buecker C: **What is an enhancer?** *BioEssays* 2023, **n/a**:2300044.

11. Blayney J, Francis H, Camellato B, Mitchell L, Stolper R, Boeke J, Higgs D, Kassouf M: **Super-enhancers require a combination of classical enhancers and novel facilitator elements to drive high levels of gene expression**. *bioRxiv* 2022, doi:10.1101/2022.06.20.496856.

12. Brosh R, Coelho C, Ribeiro-dos-Santos AM, Ellis G, Hogan MS, Ashe HJ, Somogyi N, Ordoñez R, Luther RD, Huang E, et al.: **Synthetic regulatory genomics uncovers enhancer context dependence at the Sox2 locus**. *Mol Cell* 2023, **83**:1140-1152.e7.

13. Thomas HF, Kotova E, Jayaram S, Pilz A, Romeike M, Lackner A, Penz T, Bock C, Leeb M, Halbritter F, et al.: **Temporal dissection of an enhancer cluster reveals distinct temporal and functional contributions of individual elements**. *Mol Cell* 2021, **81**:969-982.e13.

14. Ong C-T, Corces VG: **CTCF: an architectural protein bridging genome topology and function**. *Nat Rev Genet* 2014, **15**:234–246.

15. Rowley MJ, Corces VG: **Organizational principles of 3D genome architecture**. *Nat Rev Genet* 2018, **19**:789–800.

16. Davidson IF, Peters J-M: **Genome folding through loop extrusion by SMC complexes**. *Nat Rev Mol Cell Biol* 2021, **22**:445–464.

17. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y, et al.: **CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function**. *Cell* 2015, **162**:900–910.

18. Nora EP, Caccianini L, Fudenberg G, So K, Kameswaran V, Nagle A, Uebersohn A, Hajj B, Saux AL, Coulon A, et al.: **Molecular basis of CTCF binding polarity in genome folding**. *Nat Commun* 2020, **11**:5612.

19. Beagan JA, Phillips-Cremins JE: **On the existence and functionality of topologically associating domains**. *Nat Genet* 2020, **52**:8–16.

20. Ren G, Jin W, Cui K, Rodrigez J, Hu G, Zhang Z, Larson DR, Zhao K: **CTCF-Mediated Enhancer-Promoter Interaction Is a Critical Regulator of Cell-to-Cell Variation of Gene Expression**. *Mol Cell* 2017, **67**:1049-1058.e6.

21. Kubo N, Ishii H, Xiong X, Bianco S, Meitinger F, Hu R, Hocker JD, Conte M, Gorkin D, Yu M, et al.: **Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation**. *Nat Struct Mol Biol* 2021, **28**:152–161.

22. Beagan JA, Duong MT, Titus KR, Zhou L, Cao Z, Ma J, Lachanski CV, Gillis DR, Phillips-Cremins JE: **YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment**. *Genome Res* 2017, **27**:1139–1152.

23. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, et al.: **YY1 Is a Structural Regulator of Enhancer-Promoter Loops**. *Cell* 2017, **171**:1573-1588.e28.

24. Zhang S, Übelmesser N, Barbieri M, Papantonis A: **Enhancer-promoter contact formation requires RNAPII and antagonizes loop extrusion**. *bioRxiv* 2022, doi:10.1101/2022.07.04.498738.

25. Chen L-F, Long HK, Park M, Swigut T, Boettiger AN, Wysocka J: **Structural elements promote architectural stripe formation and facilitate ultra-long-range gene regulation at a human disease locus**. *Mol Cell* 2023, doi:10.1016/j.molcel.2023.03.009.

26. Long HK, Osterwalder M, Welsh IC, Hansen K, Davies JOJ, Liu YE, Koska M, Adams AT, Aho R, Arora N, et al.: **Loss of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder**. *Cell Stem Cell* 2020, **27**:765-783.e14.

27. Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN: **Visualizing DNA folding and RNA in embryos at single-cell resolution**. *Nature* 2019, **568**:49–54.

28. Despang A, Schöpflin R, Franke M, Ali S, Jerković I, Paliou C, Chan W-L, Timmermann B, Wittler L, Vingron M, et al.: **Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture**. *Nat Genet* 2019, **51**:1263–1271.

29. Hung T-C, Kingsley DM, Boettiger A: **Boundary stacking interactions enable cross-TAD enhancer-promoter communication during limb development**. *bioRxiv* 2023, doi:10.1101/2023.02.06.527380.

30. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al.: **A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping**. *Cell* 2014, **159**:1665–1680.

31. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions**. *Nature* 2012, **485**:376–380.

32. Laugsch M, Bartusel M, Rehimi R, Alirzayeva H, Karaolidou A, Crispatzu G, Zentis P, Nikolic M, Bleckwehl T, Kolovos P, et al.: **Modeling the Pathological Long-Range Regulatory Effects of Human Structural Variation with Patient-Specific hiPSCs**. *Cell Stem Cell* 2019, **24**:736-752.e12.

33. Spielmann M, Lupiáñez DG, Mundlos S: **Structural variation in the 3D genome**. *Nat Rev Genet* 2018, **19**:453–467.

34. Kessler S, Minoux M, Joshi O, Ben Zouari Y, Ducret S, Ross F, Vilain N, Salvi A, Wolff J, Kohler H, et al.: **A multiple super-enhancer region establishes inter-TAD interactions and controls Hoxa function in cranial neural crest**. *Nat Commun* 2023, **14**:3242.

35. Kragesteen BK, Spielmann M, Paliou C, Heinrich V, Schöpflin R, Esposito A, Annunziatella C, Bianco S, Chiariello AM, Jerković I, et al.: **Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis**. *Nat Genet* 2018, **50**:1463–1473.

36. Andrey G, Schöpflin R, Jerković I, Heinrich V, Ibrahim DM, Paliou C, Hochradel M, Timmermann B, Haas S, Vingron M, et al.: **Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding**. *Genome Res* 2017, **27**:223–233.

37. Ortabozkoyun H, Huang P-Y, Cho H, Narendra V, LeRoy G, Gonzalez-Buendia E, Skok JA, Tsirigos A, Mazzoni EO, Reinberg D: **CRISPR and biochemical screens identify MAZ as a cofactor in CTCF-mediated insulation at Hox clusters**. *Nat Genet* 2022, **54**:202–212.

38. Ortabozkoyun H, Huang P-Y, Cho H, Tsirigos A, Mazzoni E, Reinberg D: **Novel Chromatin Insulating Activities Uncovered upon Eliminating Known Insulators in vivo**. *bioRxiv* 2023, doi:10.1101/2023.04.25.538167.

39. Aljahani A, Hua P, Karpinska MA, Quililan K, Davies JOJ, Oudelaar AM: **Analysis of sub-kilobase chromatin topology reveals nano-scale regulatory interactions with variable dependence on cohesin and CTCF**. *Nat Commun* 2022, **13**:2139.

40. Chakraborty S, Kopitchinski N, Zuo Z, Eraso A, Awasthi P, Chari R, Mitra A, Tobias IC, Moorthy SD, Dale RK, et al.: **Enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness**. *Nat Genet* 2023, **55**:280–290.

41. Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Darzacq X, Tjian R: **Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1**. *Nat Genet* 2022, **54**:1919–1932.

42. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG: **Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization**. *Cell* 2017, **169**:930-944.e22.

43. Pachano T, Sánchez-Gaya V, Ealo T, Mariner-Faulí M, Bleckwehl T, Asenjo HG, Respuela P, Cruz-Molina S, Muñoz-San Martín M, Haro E, et al.: **Orphan CpG islands amplify**

poised enhancer regulatory activity and determine target gene responsiveness. *Nat Genet* 2021, **53**:1036–1049.

44. Batut PJ, Bing XY, Sisco Z, Raimundo J, Levo M, Levine MS: **Genome organization controls transcriptional dynamics during development**. *Science* 2022, **375**:566–570.

45. Calhoun VC, Stathopoulos A, Levine M: **Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex**. *Proc Natl Acad Sci* 2002, **99**:9243–9247.

46. Calhoun VC, Levine M: **Long-range enhancer–promoter interactions in the *Scr-Antp* interval of the *Drosophila* Antennapedia complex**. *Proc Natl Acad Sci* 2003, **100**:9878–9883.

47. Gabriele M, Brandão HB, Grosse-Holz S, Jha A, Dailey GM, Cattoglio C, Hsieh T-HS, Mirny L, Zechner C, Hansen AS: **Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging**. *Science* 2022, **376**:496–501.

48. Mach P, Kos PI, Zhan Y, Cramard J, Gaudin S, Tünnermann J, Marchi E, Eglinger J, Zuin J, Kryzhanovska M, et al.: **Cohesin and CTCF control the dynamics of chromosome folding**. *Nat Genet* 2022, **54**:1907–1918.

49. Cruz-Molina S, Respuela P, Tebartz C, Kolovos P, Nikolic M, Fueyo R, van Ijcken WFJ, Grosveld F, Frommolt P, Bazzi H, et al.: **PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation**. *Cell Stem Cell* 2017, **20**:689-705.e9.

50. Espinola SM, Götz M, Bellec M, Messina O, Fiche J-B, Houbron C, Dejean M, Reim I, Cardozo Gizzi AM, Lagha M, et al.: **Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development**. *Nat Genet* 2021, **53**:477–486.

51. Paliou C, Guckelberger P, Schöpflin R, Heinrich V, Esposito A, Chiariello AM, Bianco S, Annunziatella C, Helmuth J, Haas S, et al.: **Preformed chromatin topology assists transcriptional robustness of *Shh* during limb development**. *Proc Natl Acad Sci* 2019, **116**:12390–12399.

52. Williamson I, Lettice LA, Hill RE, Bickmore WA: **Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity**. *Development* 2016, **143**:2994–3001.

53. Williamson I, Kane L, Devenney PS, Flyamer IM, Anderson E, Kilanowski F, Hill RE, Bickmore WA, Lettice LA: **Developmentally regulated Shh expression is robust to TAD perturbations**. *Development* 2019, **146**:dev179523.

54. Calderon L, Weiss FD, Beagan JA, Oliveira MS, Georgieva R, Wang Y-F, Carroll TS, Dharmalingam G, Gong W, Tossell K, et al.: **Cohesin-dependence of neuronal gene expression relates to chromatin loop length**. *eLife* 2022, **11**:e76539.

55. Kane L, Williamson I, Flyamer IM, Kumar Y, Hill RE, Lettice LA, Bickmore WA: **Cohesin is required for long-range enhancer action**. *bioRxiv* 2021, doi:10.1101/2021.06.24.449812.

56. Karr JP, Ferrie JJ, Tjian R, Darzacq X: **The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer–promoter communication**. *Genes Dev* 2022, **36**:7–16.

57. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA: **Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor**. *Cell* 2012, **149**:1233–1244.

58. Sánchez-Gaya V, Mariner-Faulí M, Rada-Iglesias A: **Rare or Overlooked? Structural Disruption of Regulatory Domains in Human Neurocristopathies**. *Front Genet* 2020, **11**.

59. Ushiki A, Zhang Y, Xiong C, Zhao J, Georgakopoulos-Soares I, Kane L, Jamieson K, Bamshad MJ, Nickerson DA, Shen Y, et al.: **Deletion of CTCF sites in the SHH locus alters enhancer–promoter interactions and leads to acheiropodia**. *Nat Commun* 2021, **12**:2282.

60. Eksi SE, Barmina O, McCallough CL, Kopp A, Orenic TV: **A Distalless-responsive enhancer of the Hox gene Sex combs reduced is required for segment- and sex-specific sensory organ development in Drosophila**. *PLoS Genet* 2018, **14**:e1007320.

61. Pattatucci AM, Otteson DC, Kaufman TC: **A functional and structural analysis of the Sex combs reduced locus of Drosophila melanogaster.** *Genetics* 1991, **129**:423–441.

62. Sivanantharajah L, Percival-Smith A: **Analysis of the Sequence and Phenotype of Drosophila Sex combs reduced Alleles Reveals Potential Functions of Conserved Protein Motifs of the Sex combs reduced Protein**. *Genetics* 2009, **182**:191–203.

63. Huang H, Zhu Q, Jussila A, Han Y, Bintu B, Kern C, Conte M, Zhang Y, Bianco S, Chiariello AM, et al.: **CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains**. *Nat Genet* 2021, **53**:1064–1074.

64. Taylor T, Sikorska N, Shchuka VM, Chahar S, Ji C, Macpherson NN, Moorthy SD, de Kort MAC, Mullany S, Khader N, et al.: **Transcriptional regulation and chromatin architecture maintenance are decoupled functions at the Sox2 locus**. *Genes Dev* 2022, **36**:699–717.

65. Diao Y, Fang R, Li B, Meng Z, Yu J, Qiu Y, Lin KC, Huang H, Liu T, Marina RJ, et al.: **A tiling1deletion based genetic screen for cis-regulatory element identification in mammalian cells**. *Nat Methods* 2017, **14**:629–635.

66. Findlay GM, Boyle EA, Hause RJ, Klein JC, Shendure J: **Saturation editing of genomic regions by multiplex homology-directed repair**. *Nature* 2014, **513**:120–123.

67. Piecyk RS, Schlegel L, Johannes F: **Predicting 3D chromatin interactions from DNA sequence using Deep Learning**. *Comput Struct Biotechnol J* 2022, **20**:3439–3448.

68. Yang M, Ma J: **Machine Learning Methods for Exploring Sequence Determinants of 3D Genome Organization**. *J Mol Biol* 2022, **434**:167666.

CURRENT OPINION IN STRUCTURAL BIOLOGY

Author declaration

[Instructions: Please check all applicable boxes and provide additional information as requested.]

1. Conflict of Interest

Potential conflict of interest exists:

We wish to draw the attention of the Editor to the following facts, which may be considered as potential conflicts of interest, and to significant financial contributions to this work:

The nature of potential conflict of interest is described below:

☒No conflict of interest exists.

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

2. Funding

☒Funding was received for this work.

All of the sources of funding for the work described in this publication are acknowledged below:

[List funding sources and their role in study design, data analysis, and result interpretation]

HL is supported by a Medical Research Council (MRC) UK University Unit grant (MC_UU_00035/12).

☐No funding was received for this work.

3. Intellectual Property

☒We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

4. Research Ethics

☐We further confirm that any aspect of the work covered in this manuscript that has involved human patients has been conducted with the ethical approval of all relevant bodies and that such approvals are acknowledged within the manuscript.

☐IRB approval was obtained (required for studies and series of 3 or more cases)

☐Written consent to publish potentially identifying information, such as details or the case and photographs, was obtained from the patient(s) or their legal guardian(s).

5. Authorship

The International Committee of Medical Journal Editors (ICMJE) recommends that authorship be based on the following four criteria:

1. Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND

2. Drafting the work or revising it critically for important intellectual content; AND

3. Final approval of the version to be published; AND

4. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

All those designated as authors should meet all four criteria for authorship, and all who meet the four criteria should be identified as authors. For more information on

authorship, please see http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two.

☒All listed authors meet the ICMJE criteria. We attest that all authors contributed significantly to the creation of this manuscript, each having fulfilled criteria as established by the ICMJE.

☐One or more listed authors do(es) not meet the ICMJE criteria.

We believe these individuals should be listed as authors because:

[Please elaborate below]

☒We confirm that the manuscript has been read and approved by all named authors.

☒We confirm that the order of authors listed in the manuscript has been approved by all named authors.

6. Contact with the Editorial Office

The Corresponding Author declared on the title page of the manuscript is:

Hannah Katherine Long

☒This author submitted this manuscript using his/her account in Editorial Manager (EM).

☒We understand that this Corresponding Author is the sole contact for the Editorial process (including EM and direct communications with the office). He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

☒We confirm that the email address shown below is accessible by the Corresponding Author, is the address to which Corresponding Author's EM account is linked, and has been configured to accept email from the editorial office of *Current Opinion in Structural Biology*:

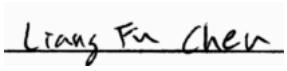
Hannah.long@ed.ac.uk

☐Someone other than the Corresponding Author declared above submitted this manuscript from his/her account in EM:

[Insert name below]

☐We understand that this author is the sole contact for the Editorial process (including EM and direct communications with the office). He/she is responsible for communicating with the other authors, including the Corresponding Author, about progress, submissions of revisions and final approval of proofs.

We the undersigned agree with all of the above.

Author's name (Fist, Last)	Signature	Date
1. Hannah Long ____		18th June 2023___
2. Liang-Fu Chen		18th June 2023
3. _____	_____	_____
4. _____	_____	_____
5. _____	_____	_____
6. _____	_____	_____
7. _____	_____	_____
8. _____	_____	_____

9. _____ _____ _____

10. _____ _____ _____

Box 1

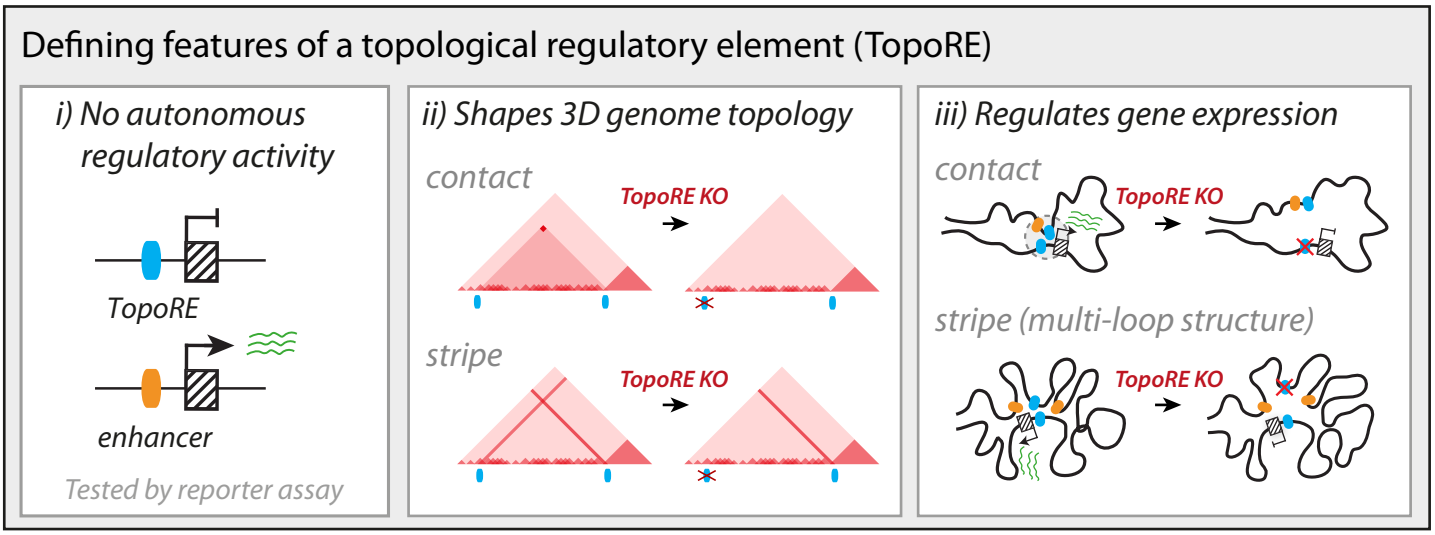


Figure 1

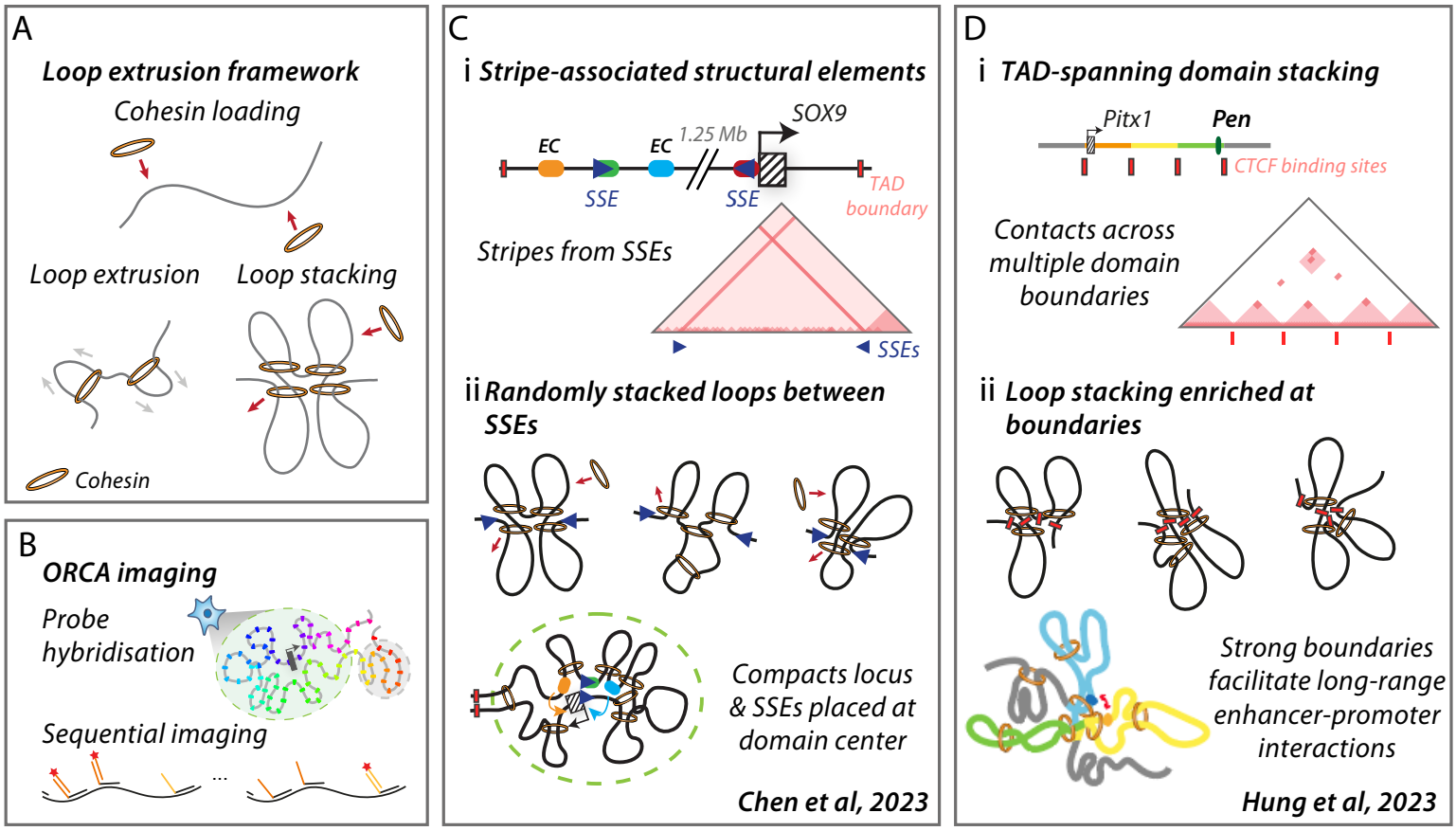


Figure 2

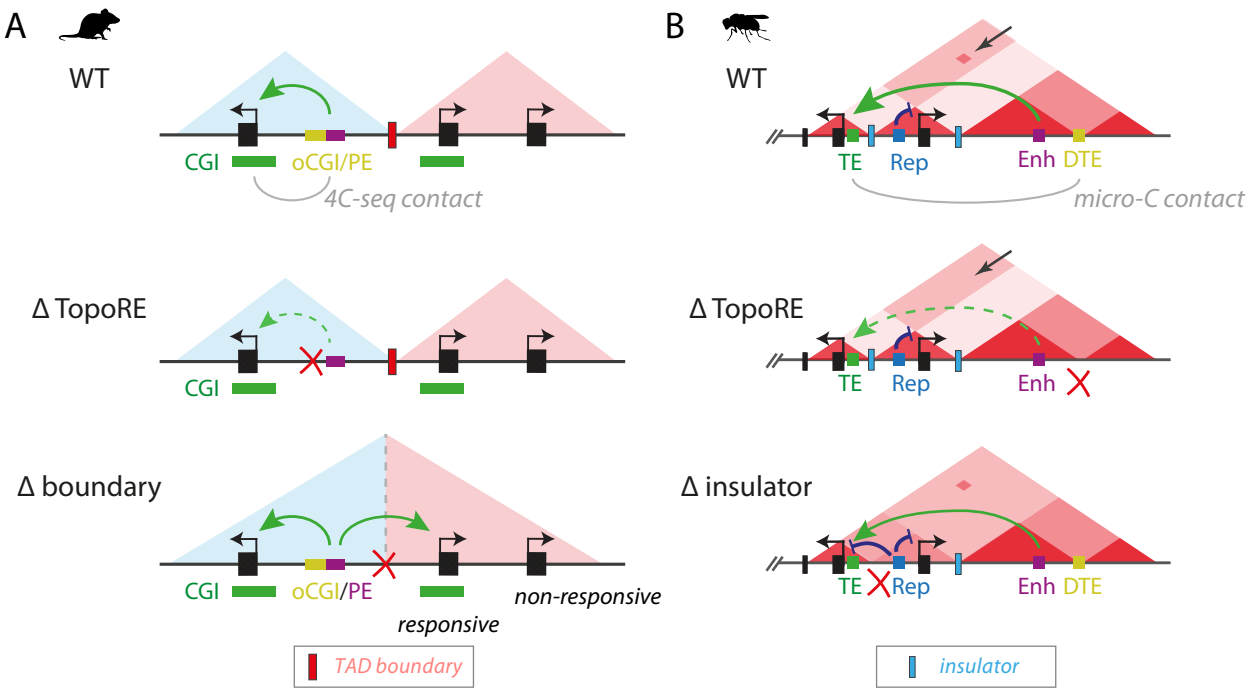


Figure 3