Cohesin is required for long-range enhancer action at the Shh locus

Cohesin is required for long-range enhancer action at the *Shh* locus

Lauren Kane[1], Iain Williamson[1], Ilya M. Flyamer[1#], Yatendra Kumar[1], Robert E. Hill[1], Laura A. Lettice[1], Wendy A. Bickmore*[1]

[1]MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK

[#]Present address: Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

*Correspondence to:
W.A.B: MRC Human Genetics Unit, IGC, Crewe Road, Edinburgh EH4 2XU, UK
Tel: +44 131 651 8570
Email: Wendy.Bickmore@ed.ac.uk

Abstract

The regulatory landscapes of developmental genes in mammals can be complex, with enhancers spread over many hundreds of kilobases. It has been suggested that three-dimensional genome organisation, particularly topologically associating domains formed by cohesin-mediated loop extrusion, are important for enhancers to act over such large genomic distances. By coupling acute protein degradation with synthetic activation by targeted transcription factor recruitment, here we show that cohesin, but not CTCF, is required for activation of a target gene – *Shh* - by distant enhancers in mouse embryonic stem cells. Cohesin is not required for activation directly at the promoter or from an enhancer located closer to the *Shh* gene. Our findings support the hypothesis that chromatin compaction mediated by cohesin-mediated loop extrusion allows for genes to be activated by enhancers that are located many hundreds of kilobases away in the linear genome but suggests that cohesin is dispensable for more genomically close enhancers.

Introduction

The mammalian genome is organised into topologically associating domains (TADs) that are formed through the process of cohesin-driven loop extrusion[1-3] and whose extent is constrained at TAD boundaries by orientation-dependent CTCF binding[4-7]. The large regulatory landscapes of developmental genes frequently correspond to TADs, leading to the hypothesis that TADs and/or loop extrusion are important for enhancers to act on their cognate gene[8,9].

However, it has proven hard to interpret the consequences of experimental disruption of TADs or loop-extrusion on gene regulation[3,6,10], in part because of the difficulty in distinguishing direct from indirect effects on enhancer-driven gene expression. CTCF null mice show early embryonic lethality[11] and conditional knockout of CTCF in the developing mouse limb results in extensive cell death[12]. Cohesin is also essential for cell proliferation, limiting study in vivo[13]. In vitro removal of cohesin does not seem to have very substantial effects on specific gene regulation[3], but it is required for inducible gene regulation in primary haematopoietic cells[14]. Conversely, conditional removal of cohesin in post-mitotic neurons has been reported to perturb gene expression but not to affect enhancer-driven inducible immediate early gene activation[15]. Depletion of the cohesin loading factor NIPBL in non-dividing liver cells in vivo could not attribute transcriptional effects to systematically altered enhancer function[16].

Here, we exploit synthetic transcriptional activators, coupled with the acute degradation of CTCF or cohesin, to investigate mechanisms of enhancer action at a distance, using the *Shh* locus as a paradigm. In mouse embryonic stem cells (mESCs) we show that cohesin, but not CTCF, is required for activation of *Shh* by enhancers located many hundreds of kilobases upstream, but is dispensable for an enhancer located closer to the target gene.

Results

Synthetic enhancer activation at long distance

Shh acts as a concentration-dependent morphogen during vertebrate embryonic development and the complex *Shh* regulatory domain is a paradigm for long-range enhancer regulation. Many of the tissue-specific enhancers of *Shh* operate over large genomic distances with the regulatory landscape extending over approximately 1 Mb (Fig. 1a). The limits of this regulatory landscape, defined using transposon-based regulatory sensors[9,17], correspond with a TAD which contains *Shh* and all of its enhancers that have been defined so far. One of the

TAD boundaries lies in an intergenic region 3' of *Shh* whereas the other is near the *Lmbr1* promoter. The murine *Shh* regulatory landscape contains at least five CTCF binding sites (Fig. 1a), including two strongly interacting convergent sites which may form the *Shh* TAD boundaries and block loop extrusion[18].

ZRS, located 849 kb upstream of *Shh* in intron 5 of the widely expressed *Lmbr1,* is the most distal *Shh* enhancer. *ZRS* is both necessary and sufficient for *Shh* expression in the zone of polarising activity (ZPA) in distal posterior mesenchymal cells of the developing limb bud[19,20]. Increased *Shh*-*ZRS* colocalization, observed in the ZPA, may be consistent with a gene-enhancer interaction[21]. Large inversions that encompass the *Shh* TAD boundaries disrupt *Shh*-ZRS interactions and *Shh* regulation in limb buds[9] but small deletions of CTCF sites at the *Shh* TAD boundaries, though disrupting TAD structure and reducing *Shh* - *ZRS* colocalization, do not alter the developmental pattern of *Shh* expression or cause a mutant phenotype[18].

Enhancers are activated by binding of the appropriate transcription factors (TFs) which can be mimicked by targeting of artificial TFs. Previously, we demonstrated synthetic activation of *Shh* in mESCs using transcription activator-like (TAL) effectors (TALEs) fused to multimers of VP16[22]. *Shh* expression could be induced by activator binding at the *Shh* promoter (tShh) and at the neural enhancers *SBE6* (100 kb upstream) and *SBE2* (410 kb upstream) (tSBE2-VP64 and tSBE2-VP64, respectively). Local peaks of H3K27 acetylation were also induced at the site of TALE binding and at the activated *Shh* in both cases[22]. To determine if *Shh* transcription could also be triggered by synthetic activator binding at the far end of the TAD, we designed a TALE for *ZRS* (tZRS) (Fig. 1a, b).

Previously we used qRT-PCR to assay the steady state level of *Shh* mRNA induced by synthetic activators averaged across the transfected cell population[22]. To detect *Shh* nascent transcripts at a single cell/allele level, here we used RNA FISH in mESCs 48hrs after transfections with tShh-VP64, tSBE2-VP64 and tZRS-VP64 (Fig. 1c, d, e) and with control constructs lacking the activation domain (-Δ) (Extended Data Fig.1a,b). A probe set detecting *Lmbr1* nascent transcripts was used as a positive control as TALE binding was not thought to be able to affect this broadly expressed gene.

Consistent with our previous demonstration of *Shh* activation by TALE-Vp64s, *Shh* nascent RNA FISH signals were detected in mESCs transfected with tShh-VP64 (9-11% of *Shh* alleles) or tSBE2-VP64 (4-5% of alleles)(Fig. 1f, Extended Data Fig. 1b,c). Cells transfected with TALE-Δ, and non-transfected cells (ntc) showed a very low signal levels. tZRS-VP64 also activated *Shh* (4-6% of alleles detected) indicating that *Shh* can be expressed following activator binding 850kb away. Induced *Shh* expression levels from all TALE constructs fused to VP64 were significantly greater than the equivalent TALE-Δ transfected cells (Extended Data Fig. 1d). *Lmbr1* transcripts were detected at approximately 60% of alleles and these levels were similar in cells transfected with either tShh, tSBE2 or tZRS with or without fusion to Vp16 (Fig. 1f; Extended Data Fig. 1c).

Synthetic activation in the absence of CTCF

The *Shh* TAD contains a number of CTCF binding sites (Fig. 1a) important for TAD structure but that individually are not necessary for *Shh* regulation in vivo[18]. Combinatorial deletion suggests that loss of more than one CTCF site within the *Shh* TAD may have a more marked effect on expression[23]. Genome-wide depletion of CTCF in mESCs dramatically alters TAD insulation with rather minimal effects on ongoing gene expression[6]. However, those studies did not address where the complete loss of CTCF affects the induction of gene activation, and particularly via enhancers.

To investigate whether synthetic activation of *Shh* was dependent on CTCF we used mESCs in which the degradation of CTCF can be induced via an auxin-inducible degron (AID) (CTCF-AID)[6]. FACS (for GFP) indicated that CTCF depletion occurred as early as 6 hours after auxin addition and persisted for up to 48hrs of auxin treatment (Fig. 2a). CTCF-AID auxin-treated cells appeared to divide for at least 1-2 cell cycles, maintained a normal colony morphology and did not show significant levels of cell death up to 48 hours of auxin treatment (Extended Data Fig. 1e). Immunofluorescence indicated a very small proportion of GFP positive cells in CTCF-AID cells treated for 6 hours and so 24 hour auxin treatment, when GFP^{+ve} cells were completely absent, was used for subsequent experiments using CTCF-AID cells (Fig. 2b). To ensure that auxin addition did not impact TALE activity *per se*, wild type mESCs were transfected with TALE-VP64/-Δ targeting the *Shh* promoter, *SBE2* and *ZRS* and auxin added to the media on the day after transfection for 24 hours. Targeting the *Shh* promoter or distal enhancers (*SBE2/ZRS*) with TALE-VP64, but not with TALE-Δ, led to activation of *Shh* expression both in the absence and presence of auxin with no significant differences in the proportion of expressing alleles caused by the addition of auxin (Fig. 1f and Extended Data Fig. 1c,d). *Lmbr1* expression was also consistent across all conditions and unaffected by the TALEs.

It has been reported that CTCF-AID molecules bound at different CTCF sites have different susceptibilities to auxin-dependent degradation[24], and ChIP for CTCF following auxin treatment of CTCF-AID mESCs[6] shows a complete absence of CTCF at sites 2 and 3 within the *Shh* TAD and a substantial reduction, but not complete absence, of CTCF sites 1, 4 and 5 at the *Shh* TAD boundaries (Fig. 2c). However, analysis of Hi-C data[6] from auxin-treated CTCF-AID cells shows that insulation at the *Shh* TAD boundaries, and particularly that at the *Lmbr1* end, are weakened (Fig. 2c), and more inter-TAD interactions between the *Shh* and neighbouring TADs are detected in the absence of CTCF. Intra-TAD interactions were also affected by CTCF depletion, confirmed by virtual 4C display of the Hi-C data (Extended Data Fig. 2a). Using the *Shh* promoter as a viewpoint, proteolytic degradation of CTCF leads to decreased interactions of *Shh* with sequences within its own TAD and increased interactions with sequences in the adjacent *En2* containing TAD. Of note, our previous 5C analysis of the *Shh* TAD did not detect the formation of specific enhancer-promoter loops as a consequence of TALE-Vp64 enhancer activation[22].

Given this altered 3D chromatin landscape, we tested whether CTCF depletion affected the ability to synthetically activate *Shh,* including from distal enhancers. CTCF-AID cells were transfected with tShh-VP64, tSBE2-VP64 and tZRS-VP64 and with the corresponding TALE-Δs controls. Auxin was added the day after transfection for 24 hours. *Shh* expression could still be induced in CTCF-depleted cells when targeting either the *Shh* promoter or the enhancers with TALE-VP64 (Fig. 3a, Extended Data Fig. 3a, c). These data suggest that activation of *Shh* expression by targeting its distal enhancers does not require CTCF.

Consistent both with the Hi-C/ virtual 4C data (Fig. 2c, Extended Data Fig. 2a) from auxin-treated CTCF-AID cells, and with our previous analysis deleting specific CTCF sites at the *Shh* locus[18], DNA FISH on CTCF-AID cells transfected with tSBE2-VP64 or tZRS-VP64 revealed some decompaction within the *Shh* TAD caused by CTCF loss (Fig. 3b,c and Extended Data Fig. 3e). Notably, after CTCF loss in tSBE2-VP64 transfected cells we saw no alleles where *Shh* and *SBE2* were spatially co-localised (within 200nm) (Fig. 3c) despite no effect of CTCF loss on *Shh* nascent transcription by tSBE2-VP64 (Fig. 3a). This is consistent with previous observations that enhancer-gene co-localisation is not required for enhancer-driven gene-activation[22], though we cannot exclude extremely transient colocalization events that are undetectable by FISH.

Synthetic activation of Shh from a distance is cohesin dependant

To examine effects of cohesin loss on synthetic *Shh* activation we used auxin to acutely deplete SCC1 (RAD21) from mESCs[25]. Cohesin is required for sister chromatid cohesion during mitosis[13] and in its absence SCC1-AID cells fail to divide and die. FACS and immunofluorescence indicated that SCC1 depletion occurs as early as 6 hours after auxin addition (Fig. 2a, b) and we detected substantial cell death following 24 hours of auxin treatment of SCC1-AID cells (Extended Data Fig. 1d). Therefore, 6hrs of auxin treatment were used for subsequent experiments.

Genome-wide depletion of cohesin by auxin treatment of SCC1-AID is reported to erase TADs with minimal effects on steady-state gene expression[3]. Hi-C reveals a pronounced effect of SCC1 depletion on *Shh* TAD structure[25] (Fig. 2d). Both *Shh* TAD boundaries were abrogated, and intra-TAD interactions severely depleted. Virtual 4C analysis revealed the profound loss of long-range interactions of *Shh* both within its own TAD but also with the adjacent En2 TAD (Extended Data Fig. 2b).

To assess if distal enhancers can still activate *Shh* expression in the absence of cohesin, we transfected SCC1-AID cells with TALE-VP64/-Δ proteins targeting the *Shh* promoter, SBE2 and ZRS. Similar to results for CTCF-AID, *Shh* was activated in auxin-treated SCC1-AID cells by tShh-VP64 targeting the *Shh* promoter (Fig. 3d, Extended Data Fig. 3b). *Shh* was also activated from distal sites using tSBE2-VP64 and tZRS-VP64 in SCC1-AID cells in the absence of auxin. However, synthetic *Shh* activation from these two distal sites was drastically curtailed in auxin-treated SCC1-AID cells, to the extent that there was no significant difference between the TALE-VP64 and TALE-Δ transfected cells (Fig. 3d, Extended Data Fig. 3b, d). *Lmbr1* expression was unaffected by the depletion of SCC1. These data suggest that SCC1/cohesin is necessary for distal activation of *Shh* from its enhancers.

As expected, given the Hi-C and virtual 4C data, in the absence of cohesin-mediated loop extrusion (SCC1 degradation) DNA FISH confirmed significant decompaction across the *Shh* TAD that was more dramatic than that seen after CTCF depletion. Significantly increased physical distances were measured between *Shh* and the distal *SBE2* and *ZRS* enhancers (Fig. 3e,f and Extended Data Fig 3e).

Cohesin is not required for activation of Shh from a close enhancer

These data might indicate that cohesin is required for activation from all enhancers or may reflect a requirement for activation from large genomic distances. We previously demonstrated synthetic activation of *Shh* in mESCs by a TALE-VP64 targeting *SBE6* (tSBE6-VP64)[22], a *Shh* enhancer active in the developing brain and neural tube in vivo and neuronal progenitor cells ex vivo, and located only 100kb 5' of *Shh* (Fig. 1a)[26].

Cohesin degradation in SCC1-AID ESCs following auxin treatment did not impact on the ability of tSBE6-VP64 to activate expression from *Shh* (Fig. 4a and Extended Data Fig. 4a, b). Therefore, cohesin is not essential for enhancer-driven *Shh* activation *per se*, but it may be required for the function of enhancers located at relatively large genomic distances (>100kb) from their target promoter.

As we have previously reported[22], targeting of an activator to *SBE6* (in the absence of auxin) led to increased spatial separation between *Shh* and this enhancer (Fig. 4b, c). Cohesin depletion also led to decompaction between *Shh* and *SBE6* in the absence of activator (tSBE6-Δ). Notably, in the presence of an activator (tSBE6-VP64) cohesin depletion had no further effect on *Shh*-*SBE6* distances (Fig. 4b,c and Extended Data Fig. 4c).

Discussion

Our data support the observations that TAD boundaries formed by CTCF sites are not absolutely essential for an enhancer to activate its target gene located within the same TAD[18,23,27]. We find no role for either CTCF or cohesin in gene activation driven directly from the endogenous *Shh* promoter, but our data do indicate that cohesin-mediated loop extrusion is essential for activation from enhancers located at large genomic distances (>400kb) from their target gene. However, we show that cohesin is dispensable for activation from an enhancer (*SBE6*) that is located closer (100kb) to *Shh*. We note that this is consistent with a recent report showing that immediate early neuronal genes are still inducible after conditional ablation of cohesin in mouse neurons[15]. The enhancers in these cases appear to be close (~100kb) to their target gene promoters, whereas neuronal genes with longer chromatin loops had compromised expression.

One caveat of our study described here is that we are activating enhancers using a single synthetic transcription factor (TALE-VP64) and achieve levels of activation (% alleles expressing) that are far lower than those seen in some *Shh*-expressing tissues in vivo[18] where activation is driven by endogenous enhancers that likely recruit a much more complex cocktail of transcription factors and co-activators. However, we have shown that TALE-VP64s induce robust H3K27 acetylation at the enhancer they are targeted to and at the *Shh* gene[22]. Moreover, our recruitment of the same activator domain at different position within the *Shh* TAD may minimise some of the intrinsic differences between the enhancers and the differing cocktail of transcription factors they would each normally recruit during development.

Our data suggest that the process of cohesin-mediated loop-extrusion per se is not required for enhancer function – at least at the *Shh* locus. Rather, DNA-FISH suggests that it is the chromatin compaction brought about by loop-extrusion[28] that may be the important factor to consider. Whereas median inter-probe distances measured at ~400kb intervals across the *Shh* TAD were modestly affected by CTCF degradation (increases of between 10 to 140nm, Fig. 3c, Extended Data Fig. 3c), cohesin (SCC1)-loss leads to more extensive decompaction (median distance increases in the range of 130 – 330nm); Figure 3f, Extended Data Fig. 4b). In contrast, cohesin loss has no effect on chromatin decompaction between *Shh* and *SBE6* (100kb away) when *SBE6* is targeted by an activator (median distances 420nm with and without cohesin). At this stage, we cannot exclude that the differential requirement for cohesin we observe is a consequence of the inevitable cell cycle change that results from cohesin depletion in mESCs, with cells accumulating in late G2[25], though it is difficult to understand why this would affect long-range but not a relatively more closely located enhancer.

In contrast to a recent study examining the effect on reporter expression of genomic distance between promoters and enhancers inserted ectopically in mouse ESCs[29], here we find no decrease in the efficiency of nascent transcription (RNA-FISH) from an endogenous promoter driven by targeted activator (VP64) binding to different endogenous enhancer sites 100, 450 or 850kb away. One significant difference in the present study is that the activation signal has to overcome the repressive local and higher-order polycomb-mediated chromatin environment of the *Shh* locus in ESCs[30].

The molecular mechanisms by which activating signals seeded at an enhancer transmit triggers for transcriptional activation at a distant promoter remain unclear. They might involve direct translocation of regulatory information along the chromatin fibre driven by the forces of cohesin-mediated loop extrusion, but our finding that activation from an enhancer

255 located 100kb away from a promoter is cohesin independent argues against this model.
256 Rather we suggest that cohesin-mediated loop extrusion acts to maintain the entire regulatory
257 domain in a compact conformation[28]. This could then enable random close encounters
258 between enhancers and promoters to initiate molecular interactions between them, or could
259 facilitate both loci engage, for example, with the same transcriptional hub[31]. Our hypothesis
260 is also compatible with the transcription factor activity gradient model in which enhancers act
261 as nucleation sites to create diffusion gradients of activating signals that decay rapidly with
262 physical distance[32]. The size of an enhancer's sphere of influence remains to be determined
263 but our data examining the loss of enhancer-proximity caused by cohesin loss and the ability
264 of targeted enhancers to activate transcription suggest that this may be <500nm, compatible
265 with the observed distances seen between active enhancers and genes in vivo[21].

Acknowledgements: We thank Elphege Nora (University of San Francisco, USA) and James Rhodes and Rob Klose (University of Oxford, UK) for their generous gifts of CTCF and SCC1-AID cell lines. We are grateful to the IGC FACs and Advanced Imaging facilities for their expert support**.** We acknowledge the following funding sources; UK Medical Research Council (MRC) University Unit programmes MC_UU_00007/2 (IW, YK and WAB) and MC_UU_00007/8 (REH and LAL), LK was supported by an MRC PhD studentship.

Author Contributions: W.A.B and R.E.H conceived the study. W.A.B, LK, I.W, YK, L.A.L designed experiments. L.K and I.W performed experiments. I.M.F. analysed Hi-C data, L.K, I.W and W.A.B wrote the manuscript.

Competing interests: The authors declare no competing interests.

Figure legends

Figure 1. Synthetic Shh activation

(**a**) Hi-C heatmap of the *Shh* TAD from wild type mESCs at 16kb resolution. Data are from ref 33 and were created using HiGlass. Genes, positions of TALE target sequences and the CTCF ChIP-seq track – including CTCF motif orientation - are shown below. Genome co-ordinates: mm9 assembly of the mouse genome. (**b**) Schematic of TALE-VP64 constructs used to target the *Shh* promoter (tShh-VP64), *SBE6* (tSBE6-VP64) *SBE2* (tSBE2-VP64) or *ZRS* (tZRS-VP64) enhancers. NLS: nuclear localisation sequence; 2A: self-cleaving 2A peptide. Repeat variable diresidue (RVD) code is displayed below using the one letter amino acid abbreviations. Equivalent TALE-Δ constructs lack the Vp64 module (**c-e**) Representative images of nuclei from mESCs transfected with (**c**) tShh-VP64, (**d**) tSBE2-VP64 and (**e**) tZRS-VP64 showing RNA FISH signals for *Shh* (white) and *Lmbr1* (red). Scale bars = 5 μm. (**f**) Timecourse of TALE transfection and auxin treatment is shown above. The percent of *Shh* (left) and *Lmbr1* (right) expressing alleles in wild-type mESCs transfected with TALE-Vp64 or TALE-Δ constructs assayed by RNA FISH in the absence or presence of auxin. The data were compared using a two-sided Fisher's Exact Test, n = numer of alleles scored, ns – not significant ($p>0.05$). Source Data Fig1. The biological replicate for these data, and the full statistical evaluation of all comparisons for *Shh* are in Extended Data Fig. 1c. and d.

Figure 2. Auxin mediated degradation of CTCF and SCC1

(a) Flow cytometric analysis of GFP fluorescence (a.u.) in wild type, CTCF-AID[6] and SCC1-AID[25] cells after 6, 24 and 48 hours of auxin treatment. (**b**) GFP fluorescence of untreated (-auxin) and treated (+ auxin) CTCF-AID and SCC1-AID cells after 24 and 6 hours of growth in auxin, respectively (scale bars = 100 μm). (**c**) Hi-C heatmaps of the *Shh* TAD from untreated (- auxin) and 48 hour treated (+ auxin) CTCF-AID mESCs (16-kb resolution, datafrom ref 6). Genes, CTCF ChIP-seq tracks and CTCF sites 1-5 from ref 18 are shown above. CTCF ChIP-seq data from untreated (left) and auxin-treated (right) CTCF-AID cells are from ref 6. Insulation scores are shown below the Hi-C heatmaps (**d**) Hi-C heatmaps of the *Shh* TAD from 6 hour auxin treated TIR1 control or SCC1-AID mESCs at 20-kb resolution (data from ref 25).

Figure 3. Depletion of cohesin, but not of CTCF, inhibits distal enhancer driven gene activation

(**a**) Timecourse of TALE transfection and auxin treatment is shown above. Percentage of (left axis) *Shh* and (right axis) *Lmbr1* expressing alleles, assayed by RNA FISH, in TALE-transfected CTCF-AID cells either untreated (- auxin) or treated with 24 hours of auxin (+ auxin). Cells were transfected with tShh-VP64, tSBE2-VP64 and tZRS-VP64 and equivalent TALE-Δ controls. Data shown are from one biological replicate. Data from an independent biological replicate are shown in Extended Data Fig. 3a. The data were compared using a two-sided Fisher's Exact Test, n = numer of alleles scored, ns – not significant (p>0.05). Source data Fig3. (**b**) Images from representative nuclei from - & + auxin CTCF-AID cells showing DNA FISH signals for *Shh*/SBE2/ZRS probes. Scale bars: 5 μm. (**c**) Violin plots showing the distribution of interprobe distances (μm) between *Shh*/*SBE2*, *SBE2*/*ZRS*, *Shh*/*ZRS* probes in tSBE2-Vp64- and tZRS-Vp64-transfected CTCF-AID cells - & + auxin. *p<0.05, **p<0.01, ****p<0.0001 (two-sided Mann-Whitney U-tests). Values for number of alleles scored, mean and inter-quartile distances are in Extended Data Fig 3e. **d**) As for (**a**) but for SCC1-AID with 6 hours of auxin (+ auxin). Source data Fig3. Biological replicate for those data in Extended Data Fig 3b. ns p>0.05 *p<0.05 **p<0.01 (Fisher's Exact tests). (**e**) and (**f**) As for (**b**) and (**c**) but for SCC1-AID cells.

Figure 4. Gene activation from a close enhancer is not affected by cohesin depletion

(**a**) Percentage of (left axis) *Shh* and (right axis) *Lmbr1* expressing alleles, assayed by RNA FISH, in TALE-transfected SCC1-AID cells either untreated (- auxin) or treated with 6 hours of auxin (+ auxin). Cells were transfected with tSBE6-VP64 or tSBE6-VP64 -Δ. Data shown are from one biological replicate. Data from an independent biological replicate are shown in Extended Data Fig. 4a. The data were compared using a two-sided Fisher's Exact Test, n = numer of alleles scored, ns – not significant (p>0.05). Source data Fig4. (**b**) Images from representative nuclei from tSBE6-Vp64- and tSBE6-Δ-transfected SCC1-AID cells - & + auxin showing DNA FISH signals for *Shh*/*SBE6* probes. Scale bars: 5 μm. (**c**) Violin plots showing the distribution of interprobe distances (μm) between *Shh*/*SBE6* probes in tSBE6-Vp64- and tSBE6-Δ-transfected SCC1-AID cells - & + auxin. **p<0.01, ***p<0.001, ****p<0.0001 (two-sided Mann-Whitney U-tests). Values for number of alleles scored, mean and inter-quartile distances are in Extended Data Fig 4c.

References

1. Sanborn, A.L., et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6456–E6465 (2015).
2. Fudenberg, G., et al. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
3. Rao, S.S.P., et al. Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320.e24 (2017).
4. Guo, Y., et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900–910 (2015).
5. de Wit, E., et al. CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* **60**, 676–684 (2015).
6. Nora, E.P., et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-933.e22 (2017).
7. Wutz, G., et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J*. **36**, 3573–3599 (2017).
8. Lupiáñez D.G., Spielmann, M. & Mundlos S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet*. **32**:225-237 (2016).
9. Symmons, O., et al. The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Dev. Cell* **39**, 529–543 (2016).
10. Haarhuis, J.H.I., et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**, 693-707.e14 (2017).
11. Moore, J.M., et al. Loss of Maternal CTCF Is Associated with Peri-Implantation Lethality of Ctcf Null Embryos. *PLoS One* **7**, e34915 (2012).
12. Soshnikova, N., Montavon, T., Leleu, M., Galjart, N. & Duboule, D. Functional Analysis of CTCF During Mammalian Limb Development. *Dev. Cell* **19**, 819–830 (2010).
13. Merkenschlager, M., Ege, A. & Nora, P. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annu. Rev. Genom. Hum. Genet* **17**, 17–43 (2016).
14. Cuartero, S., et al. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat Immunol*. **19**, 932-941 (2018).
15. Calderon, L., et al. Cohesin-dependence of neuronal gene expression relates to chromatin loop length. *Elife* **11**:e76539 (2022).
16. Schwarzer W., et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature*. **551**:51-56 (2017).
17. Anderson, E., Devenney, P.S., Hill, R.E. & Lettice, L.A. Mapping the Shh long-range regulatory domain. *Development* **141**, 3934-3943 (2014).
18. Williamson, I., et al. Developmentally regulated Shh expression is robust to TAD perturbations. *Development*. *146*, dev179523 (2019).
19. Lettice, L.A., et al. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet*. **12**, 1725–1735 (2003).
20. Sagai, T., et al. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797–803 (2005).
21. Williamson, I., Lettice, L.A., Hill, R.E. & Bickmore, W.A. Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity. *Development* **143**, 2994–3001 (2016).

22. Benabdallah, N.S., et al. Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Mol. Cell* **76**, 473-484.e7. (2019).
23. Paliou, C., et al. Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 12390–12399 (2019).
24. Luan, J., et al. Distinct properties and functions of CTCF revealed by a rapidly inducible degron system. *Cell Rep*. 34:108783 (2021).
25. Rhodes, J.D.P., et al. Cohesin Disrupts Polycomb-Dependent Chromosome Interactions in Embryonic Stem Cells. *Cell Rep*. **30**, 820-835.e10. (2020)
26. Benabdallah, N.S., et al. SBE6: a novel long-range enhancer involved in driving sonic hedgehog expression in neural progenitor cells. *Open Biol*. ***6***, 160197 (2016).
27. Despang A., et al. Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat Genet*. **51**,1263-1271 (2019).
28. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science* **366**,1345-1349 (2019).
29. Zuin, J., et al. Nonlinear control of transcription through enhancer-promoter interactions. *Nature* 604:572-577 (2022)
30. Boyle, S., et al. A central role for canonical PRC1 in shaping the 3D nuclear landscape. *Genes Dev*. **34**, 931-949. (2020).
31. Lim, B. & Levine, M.S. Enhancer-promoter communication: hubs or loops? *Curr Opin Genet Dev*. **67**, 5-9 (2021).
32. Karr, J.P., Ferrie, J.J., Tjian, R. & Darzacq, X. The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer-promoter communication. *Genes Dev*. **36**, 7-16. (2022).
33. Bonev, B., et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572.e24 (2017).

Methods

Cell Lines

Mouse embryonic stem cells (mESCs) used include wild type E14 (parental line of the CTCF-AID cells provided by Elphege Nora), CTCF-AID[6] and SCC1-AID[23].

Cell Culture and Transfections

Feeder-free mESCs were cultured on 0.1% gelatin-coated (Sigma G1890) Corning flasks or 10 cm dishes in GMEM BHK-21 (Gibco 21710-025) supplemented with 15% fetal calf serum (FSC; Sigma F-7524), 1000 units/mL Leukemia inhibitory factor (LIF; produced in-house), 2 mM L-glutamine, 1 mM sodium pyruvate (Sigma 58636), 5X non-essential amino acids (Sigma M7145) and 50 mM 2-β-mercaptoethanol (Gibco 31350-010). Cells were passaged at 60-90% confluence and plated onto gelatin-coated flasks at a density of approximately 4 x 10^4 cells/cm^2. Cells were maintained at 37°C with 5% CO_2 and routinely tested for mycoplasma.

2-3 x 10^6 mESCs were transfected with 14.5 µg of TALE plasmid and 26 µL Lipofectamine 3000 Reagent (Invitrogen L3000015) and seeded onto 0.1% gelatin-coated 10 cm dishes containing autoclaved SuperFrost Plus Adhesion glass slides. Fresh media was added after 24 hours. After 48 hours of transfection, slides were washed, fixed in 4% paraformaldehyde (pFa) and permeabilised in 70% ethanol at 4°C for minimum of 24 hours (up to one week). For each TALE transfection, half the cells were treated with auxin and half left untreated to internally controlled for each +/- auxin comparison. We assessed transfection efficiency and levels of TALE expression from the levels of cytoplasmic eGFP - encoded on the TALE plasmids after a T2A self-cleaving peptide[22]. Only cells with good levels of cytoplasmic

eGFP were analysed for RNA FISH. Measured transfection efficiencies for the TAL-VP64 constructs in Figure 1 were: tShh-VP64 92%, tSBE2-VPp64 69%, tZRS-VP64 71%. For the data in Figs 3 and 4 they were; *tShh-VP64* 76%, tSBE6-VP64 82%, tSBE2VPp64 78%.

Auxin-inducible degron induction

Indole-3-acetic acid (auxin) (MP Biomedicals 102037) was added to the medium either 6 (SCC1-AID) or 24 (wild type or CTCF-AID) hours prior to cell collection. 500 μM of auxin (1000X stock diluted in DMSO) was used for all experiments and stored at 4°C for up to a month or at -20°C for long-term storage.

TALE Design and Assembly

TALEs targeting the *Shh* promoter, *SBE2* and *SBE6* had previously been designed and assembled[22]. TALE protein specific to the limb enhancer *ZRS* was designed using TAL Effector Nucleotide Targeter 2.0 software (https://tale-nt.cac.cornell.edu) and assembled by golden-gate assembly using a modified protocol[22,34]. In brief, a DNA binding domain specific for a 15 nucleotide sequence was generated by the modular assembly of 4 pre-assembled multimeric TALE repeat modules (three 4-mer and one 3-mer) into a modified TALE backbone vector containing VP64-2A-eGFP. The backbone vector used for assembly of the ZRS TALE was modified to replace the ampicillin resistance cassette with spectinomycin resistance. TALE modules were picked from glycerol stocks of module library plates (Addgene 1000000034), incubated overnight at 37°C in L-broth containing 50 ng/μL ampicillin and DNA isolated using the QIAprep Spin Miniprep kit (Qiagen 27104) according to the manufacturer's instructions. Miniprep DNA was quantified using the Quibit dsDNA broad range assay with the Quibit 4 fluorometer. TALE modules were assembled into backbone vector by setting up a 20 μL one-pot golden-gate reaction as follows: vector (100 ng), TALE modules (200ng each), 10X Tango buffer (ThermoFisher ER0451), 20 Units Esp3I (ThermoFisher ER0451), 10 Units T4 DNA ligase (New England Biosciences M0202M), 1mM ATP in ddH_2O. Golden-gate reaction was performed on a thermal cycler ((37^{o}C 10 mins, 16^{o}C 10 mins x12) 36^{o}C 15 mins, 80^{o}C 5 mins, 4^{o}C). Competent E. coli (Invitrogen LS18263012) were transformed with 5 μL of reaction. Colonies were screened by PCR for fully assembled TALEs by setting up a 30 μL reaction as follows: single colony, 2X DreamTaq Green PCR Master mix (Thermo Scientific K1082) and 0.5 μM forward (5'GGCCAGTTGCTGAAGATCG3') and reverse (5'CGCTACAAGATGATCATTAGTG3') primers in ddH_2O. Colony PCR was performed on a thermal cycler (95^{o}C 3 mins, (95^{o}C 30s, 55^{o}C 30s, 72^{o}C 120s) x30), Reaction products were run on a 1.2% agarose gel to identify positive colonies and these colonies were confirmed by Sanger sequencing. TALE-Δ constructs were made by removing the BamHI-Bg1II fragment containing VP64 from the fully assembled TALE-VP64 plasmid by restriction digest. All TALE-VP64 and TALE- Δ plasmids were either miniprepped using the QIAprep Spin Miniprep kit (Qiagen 27104) or maxi-prepped. Plasmid DNAs were quantified using the Quibit dsDNA broad range assay with the Quibit 4 fluorometer and then stored at -20°C prior to transfection.

RNA FISH

Custom Stellaris® RNA FISH Probes were designed against *Shh* and *Lmbr1* nascent mRNAs (pool of 48 unique 22-mer probes) using the Stellaris® RNA FISH Probe Designer (www.biosearchtech.com/stellarisdesigner (version 4.2)). Following permeabilization, slides were incubated in wash buffer (2X SSC, 10% deionised formamide) for 5 mins at room temperature. Slides were hybridized with the *Shh* and *Lmbr1* Stellaris FISH Probe set labelled with Quasar 670 and 570, respectively (Biosearch Technologies, Inc.), following the manufacturer's instructions (www.biosearchtech.com/stellarisprotocols). RNA FISH probes

were warmed to room temperature, diluted to 125 nM in Stellaris RNA FISH hybridisation buffer (#SMF-HB1-10) containing 10% formamide and hybridised to slides overnight in a humidified chamber at 37°C. Slides were washed twice for 30 minutes in wash buffer at 37°C and rinsed in PBS. Slides were stained with 0.5 µg/mL DAPI and mounted using Vectashield. PBS and ddH_2O used during RNA FISH were treated with DEPC and autoclaved to inactivate RNase enzymes. RNase free consumables were used throughout and glassware treated with RNaseZAP (Invitrogen AM9780).

DNA FISH

Following RNA FISH, slides were re-probed for DNA FISH. Following the removal of coverslips, slides were briefly washed in PBS and then for 10 mins in 2xSSC at 85°C followed by denaturation in 70% formamide/2xSSC at 85°C for 50 minutes before a series of alcohol washes (70% (ice-cold), 90% and 100%). 160-240 ng of biotin- and Green496-dUTP-labeled (Enzo Life Sciences) (2-colour) or biotin- and digoxigenin- and red-dUTP-labeled (Alexa Fluor™ 594-5-dUTP, Invitrogen) (4-colour) fosmid probes (Table S1) were used per slide, with 16-24 µg of mouse Cot1 DNA (Invitrogen) and 10 µg salmon sperm DNA. EtOH was added and the probe air dried. Hybridisation mix containing deionised formamide, 20 x SSC, 50% dextran sulphate and Tween 20 was added to the probes for ~1h at room temperature. The hybridisation mix containing the probes was added to the slides and incubated overnight at 37°C. Following a series of washes in 2X SSC (45°C) and 0.1X SSC (60°C) slides were blocked in blocking buffer (4 x SSC**,** 5% Marvel) for 5 min. The following antibody dilutions were made: fluorescein anti-dig FAB fragments (Roche cat. no. 11207741910) 1:20, fluorescein anti-sheep 1:100 (Vector, cat. no. FI-6000)/ streptavadin Cy5 1:10 (Amersham, cat. no. PA45001, lot 17037668), biotinylated anti-avidin (Vector, cat. no. BA-0300, lot ZF-0415) 1:100, and streptavidin Cy5 1:10 for 3-colour detection; Texas Red avidin (Vector, cat. no. A2016) 1:500, biotinylated anti-avidin (Vector) 1:100 for 2-colour detection. Slides were incubated with antibody in a humidified chamber at 37°C for 30-60 min in the following order with 4X SSC/0.1% Tween 20 washes in between: fluorescein anti-dig, fluorescein anti-sheep, biotinylated anti-avidin, streptavidin Cy5 for 3-colour; Texas Red avidin, biotinylated anti-avidin, Texas Red avidin for 2-colour detection. Slides were treated with 1:1000 dilution of DAPI (stock 50ug/ml) for 5min before mounting in Vectashield.

Image acquisition and deconvolution

Slides from RNA and DNA FISH were imaged using a Photometrics Coolsnap HQ2 CCD camera and a Zeiss AxioImager A1 fluorescence microscope with a Plan Apochromat 100x 1.4NA objective, a Nikon Intensilight Mercury based light source (Nikon UK Ltd, Kingston-on-Thames, UK) and Chroma #89014ET (3 colour) or #89000ET (4 colour) single excitation and emission filters (Chroma Technology Corp., Rockingham, VT) with the excitation and emission filters installed in Prior motorised filter wheels. A piezoelectrically driven objective mount (PIFOC model P-721, Physik Instrumente GmbH & Co, Karlsruhe) was used to control movement in the z dimension. Step size for z stacks was set to 0.2 µm. Hardware control and image capture were performed using Nikon Nis-Elements software (Nikon UK Ltd, Kingston-on-Thames, UK). Images were deconvolved using a calculated PSF with the constrained iterative algorithm in Volocity (PerkinElmer Inc, Waltham MA). RNA FISH signal quantification was carried out using the quantitation module of Volocity (PerkinElmer Inc, Waltham MA). Expressing alleles were calculated by segmenting the hybridisation signals and scoring each nuclei as containing 0, 1 or 2 RNA signals. Only cells expressing TALE constructs (cytoplasmic eGFP) were scored. DNA FISH measurements were carried out using the quantitation module of Volocity (PerkinElmer Inc, Waltham MA). For DNA FISH, only alleles

with single probe signals were analysed, to eliminate the possibility of measuring sister chromatids.

Hi-C data analysis and generation of virtual 4C profiles

Published data from ref. 6 (NCBI GEO: GSE98671 and ref. 25 (ArrayExpress: E-MTAB-7816) were re-analysed and ref. 22 was re-analysed using the distiller pipeline (https://github.com/open2c/distiller-nf). Balanced matrices at 10 kbp were used to extract the interaction profiles of the bin containing the Shh promoter with the rest of the genome in all conditions. Then these profiles, and $\log_2$-ratio of treatment over control, were saved as bigWig files and visualised using HiGlass.

Statistics & Reproducibility

The proportion of activated alleles in wild type, CTCF-AID, and SCC1-AID ESCs transfected with the various TALE constructs, as measured by RNA FISH, in the presence or absence of auxin were compared using two-sided Fisher's Exact Test, a categorical test that provides exact P-values and is suitable for small sample sizes. A biological replicate of each experiment was performed and these data are presented in Extended Data Figures.

DNA FISH interprobe distance data sets were compared using the two-tailed Mann-Whitney U test, a nonparametric test that compares two unpaired groups. Statistics and graphs were performed using Prism software (Graphpad).

No statistical method was used to predetermine sample size. No data were excluded from the analyses. Experiments were not randomized. Investigators were not blinded to allocation during experiments and outcome assessment.

Code availability:

Analysis of HiC data was performed using https://github.com/open2c/distiller-nf

Methods references

34. Therizols, P. et al. Chromatin decondensation is sufficient to alter nuclear re-organization in embryonic stem cells. *Science* **346**, 1238-1242 (2014).

Extended Data Figure legends

Extended Data Figure 1. Effect of auxin treatment on mESCs

(**a**) Schematic of *Shh* and *Lmbr1* genes showing the position of directly labelled Custom Stellaris® RNA FISH oligo probes used for RNA FISH. *Shh* probes were labelled with Quasar 670 and *Lmbr1* probe with Quasar 570 (**b**) Images of representative nuclei showing RNA FISH signals for *Shh* (white) and *Lmbr1* (red) probes from wild type mESCs transfected with tShh-VP64 or tShh-Δ, and either untreated (- auxin) or treated with 24 hours of auxin (+ auxin). *Shh* RNA FISH signal is indicated by white arrow. Scale bars = 5 μm. (**c**) Quantification of the percent of (left) *Shh* (pink and red bars) and (right) *Lmbr1 – intron 1* (white and grey bars) expressing alleles in mESCs transfected with tShh-VP64, tSBE2-VP64 and tZRS-VP64 and equivalent TALE-Δ controls. Cells were either untreated (- auxin) or treated with 24 hours of auxin (+ auxin). Biological replicate of the data shown in Fig. 1f. The data were compared using a two-sided Fisher's Exact Test, n = numer of alleles scored, ns – not significant. (**d**) Table showing two-sided Fisher's Exact Test *p*-values for differences in the percent of *Shh*-

expressing alleles in mESCs transfected with TALE-Vp64 or TALE-Δ constructs assayed by RNA FISH in the absence or presence of auxin. Data from Figure 1f and Extended Data Figure 1c. *p*-values in bold are significant (<0.05). (**e**) Quantification of live cells by DAPI staining during flow cytometry in untreated and auxin-treated wild type (WT), CTCF-AID and SCC1-AID cells after 6, 24 and 48 hours of growth in auxin. Source Data Extended Data Fig1.

Extended Data Figure 2. Virtual 4C following auxin mediated degradation of CTCF and SCC1

Virtual 4C plots obtained by extracting Hi-C interactions using the *Shh* promoter as a viewpoint (grey dashed line) from untreated (- auxin) and treated (+ auxin) (**a**) CTCF-AID mESCs (data are from ref 6) or (**b**) SCC1-AID mESCs (data are from ref 25). Gene track is shown above and yellow dashed lines indicate the position of enhancers *SBE6*, *SBE2* and *ZRS*. The lowest panel shows a subtraction of untreated and treated cells with gain of interactions indicated in red and loss of interactions indicated in blue.

Extended Data Figure 3. Replicate data for effect of CTCF or cohesin depletion on distal enhancer driven gene activation

(**a**) Quantification of the percentage of (left axis) *Shh* and (right axis) *Lmbr1* expressing alleles, assayed by RNA FISH, in TALE-transfected wild type mESCs (parental cell line used to generate the CTCF-AID cell line) and in CTCF-AID cells either untreated (- auxin) or treated with 24 hours of auxin (+ auxin). Cells were transfected with tShh-VP64, tSBE2-VP64 and tZRS-VP64 and equivalent TALE-Δ controls. Data shown are from an indepenent biological replicate of the experiment shown in Fig 3a. The data were compared using a two-sided Fisher's Exact Test, n = numer of alleles scored, ns – not significant. Source Data EDFig3. (**b**) As for (**a**) but for SCC1-AID with 6 hours of auxin (+ auxin). Data shown are from an indepenent biological replicate of the experiment shown in Fig 3b. **p<0.01. Source Data EDFig3. (**c**) Table showing two-sided Fisher's Exact Test *p*-values for differences in the percent of *Shh*-expressing alleles in TALE-transfected CTCF-AID cells assayed by RNA FISH in the absence or presence of auxin. Cells were transfected with Shh-VP64, SBE2-VP64, ZRS-VP64, and equivalent TALE-Δ controls. Data from Figure 3a and Extended Data Figure 3a. *p*-values in bold are significant (<0.05). (**d**) As (**c**) but for SCC1-AID cells. Data from Figure 3d and Extended Data Figure 3b. **(e)** Table showing the two-sided Mann-Whitney U *p*-values for differences in FISH inter-probe distances, for Shh-*SBE2*, *SBE2*-*ZRS* and Shh-*ZRS* probe pairs, between the data from TALE-Vp64 transfected CTCF-AID or SCC1-AID ESCs with or without the addition of auxin. No of alleles scored is indicated (in parentheses). Data are from Figures 3c and 3f. Median and inter-quartile distances are shown. *p*-values in bold are significant (<0.05).

Extended Data Figure 4. Replicate data showing gene activation from a close enhancer is not affected by cohesin depletion

(**a**) Percentage of (left axis) *Shh* and (right axis) *Lmbr1* expressing alleles, assayed by RNA FISH, in TALE-transfected SCC1-AID cells either untreated (- auxin) or treated with 6 hours of auxin (+ auxin). Cells were transfected with tSBE6-VP64 or tSBE6-VP64 -Δ. Data shown are from one biological replicate. Data from an independent biological replicate are shown in Fig. 4a. The data were compared using a two-sided Fisher's Exact Test, n = numer of alleles scored, ns – not significant. Source Data EDFig4. (**b**) Table showing two-sided Fisher's Exact Test *p*-values for differences in the percent of *Shh*-expressing alleles in TALE-transfected wild

type ESCs and SCC1-AID cells assayed by RNA FISH in the absence or presence of auxin. Cells were transfected with SBE6-VP64 and SBE6-Δ. Data from Figure 4a and Extended Data Figure 4a. *p*-values in bold are significant (<0.05). **(c)** Table showing the two-sided Mann-Whitney U *p*-values for differences in the Shh-SBE6 inter-probe distances between the data from SCC1-AID ESCs with or without the addition of auxin and transfected with either tSBE6-Vp64 or tSBE6-Δ. No of alleles scored is indicated (in parentheses). *p*-values in bold are significant (<0.05). Also shown are the median and interquartile distances of each data set. Data are from Figure 4c.

Data availability: The datasets generated during the current study are available from the corresponding author upon reasonable request.
Publicly accessible data used were:
Ensembl (r 45) (http://jun2007.archive.ensembl.org/Mus_musculus/index.html). Mouse genome assembly: NCBI m37 (mm9).
NCBI GEO: GSE98671

Additional Information: Correspondence and requests for materials should be addressed to WAB.

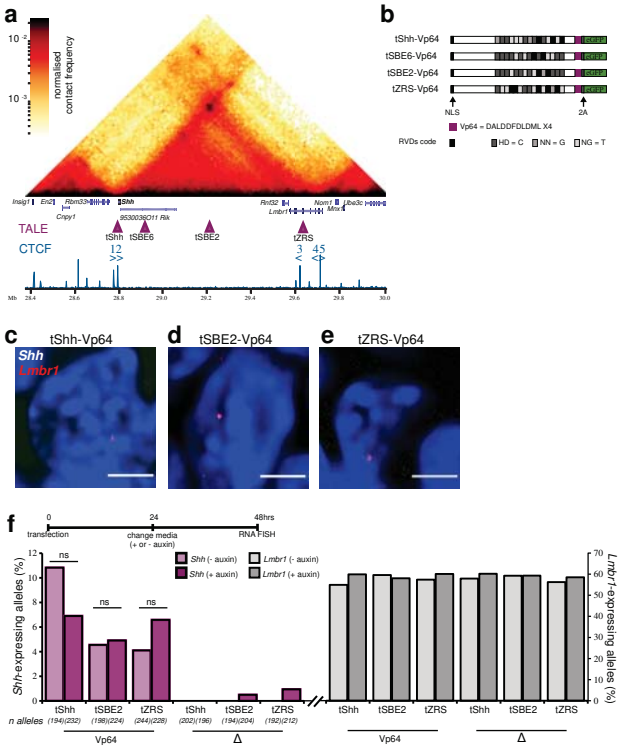

a
normalised contact frequency
TALE
CTCF
tShh
tSBE6
tSBE2
tZRS
b
tShh-Vp64
tSBE6-Vp64
tSBE2-Vp64
tZRS-Vp64
NLS
2A
Vp64 = DALDDFDLDML X4
RVDs code
HD = C
NN = G
NG = T
c
tShh-Vp64
Shh
Lmbr1
d
tSBE2-Vp64
e
tZRS-Vp64
f
transfection
change media
(+ or - auxin)
RNA FISH
Shh-expressing alleles (%)
Lmbr1-expressing alleles (%)
Shh (- auxin)
Lmbr1 (- auxin)
Shh (+ auxin)
Lmbr1 (+ auxin)
ns
n alleles
Vp64
Δ

Figure 1.

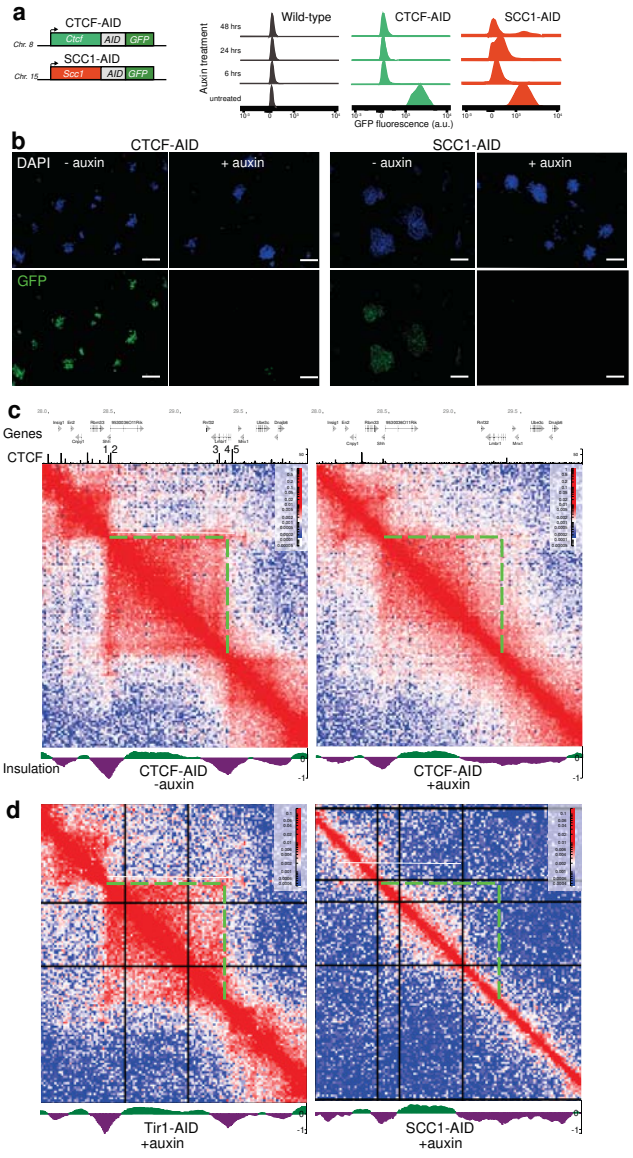

a
CTCF-AID
Ctcf
AID
GFP
Chr. 8
SCC1-AID
Scc1
AID
GFP
Chr. 15
Wild-type
CTCF-AID
SCC1-AID
Auxin treatment
48 hrs
24 hrs
6 hrs
untreated
GFP fluorescence (a.u.)
b
CTCF-AID
SCC1-AID
DAPI
- auxin
+ auxin
- auxin
+ auxin
GFP
c
Genes
CTCF
1
2
3
4
5
Insulation
CTCF-AID
-auxin
CTCF-AID
+auxin
d
Tir1-AID
+auxin
SCC1-AID
+auxin

Figure 2

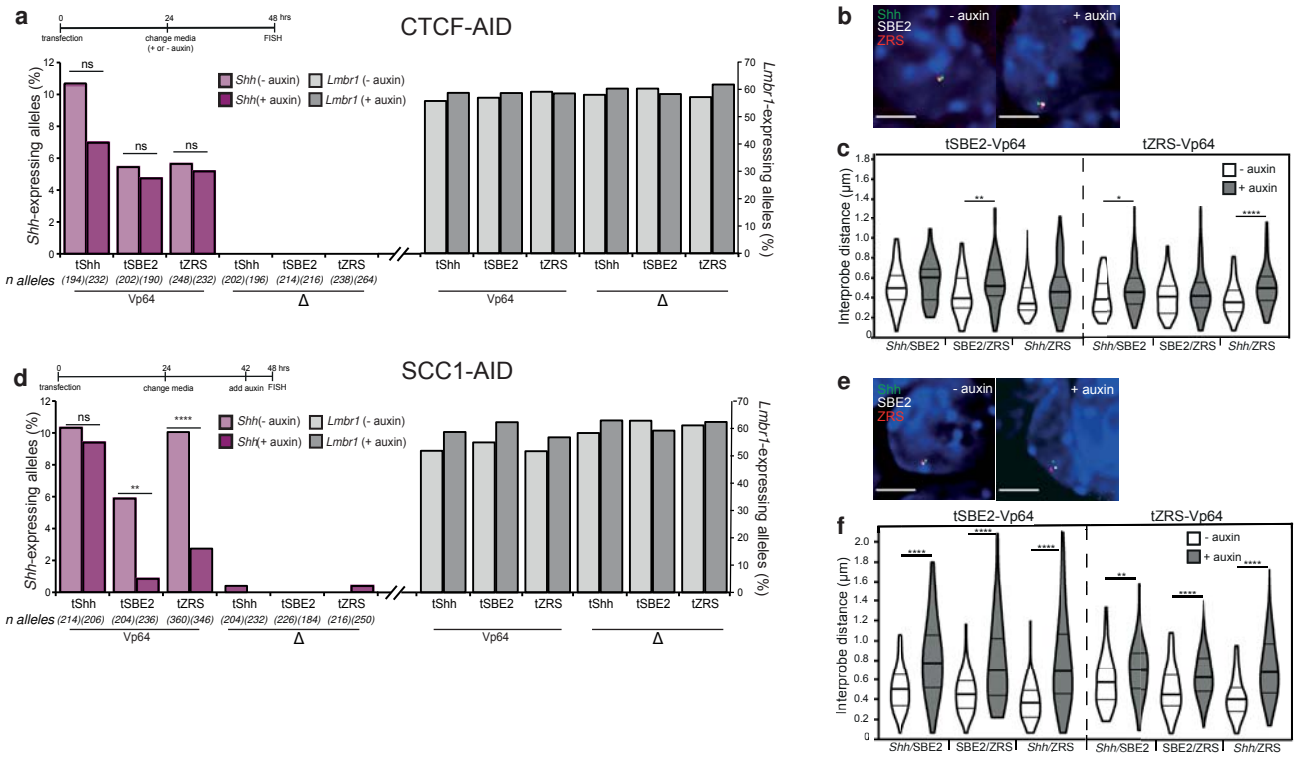

a
CTCF-AID
transfection
change media
(+ or - auxin)
FISH
Shh-expressing alleles (%)
Lmbr1-expressing alleles (%)
Shh(- auxin)
Lmbr1 (- auxin)
Shh(+ auxin)
Lmbr1 (+ auxin)
ns
tShh
tSBE2
tZRS
n alleles
(194)(232)
(202)(190)
(248)(232)
(202)(196)
(214)(216)
(238)(264)
Vp64
Δ
b
Shh
SBE2
ZRS
- auxin
+ auxin
c
tSBE2-Vp64
tZRS-Vp64
Interprobe distance (µm)
- auxin
+ auxin
Shh/SBE2
SBE2/ZRS
Shh/ZRS
d
SCC1-AID
add auxin
(214)(206)
(204)(236)
(360)(346)
(204)(232)
(226)(184)
(216)(250)
e
f

Figure 3

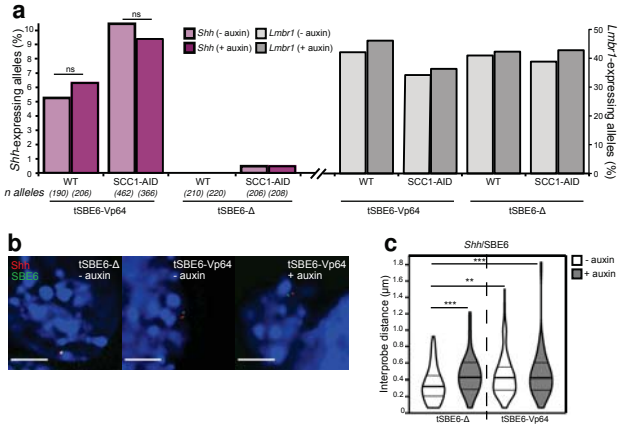
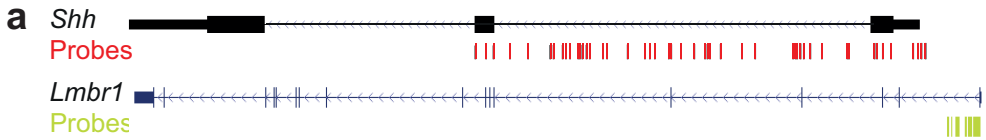
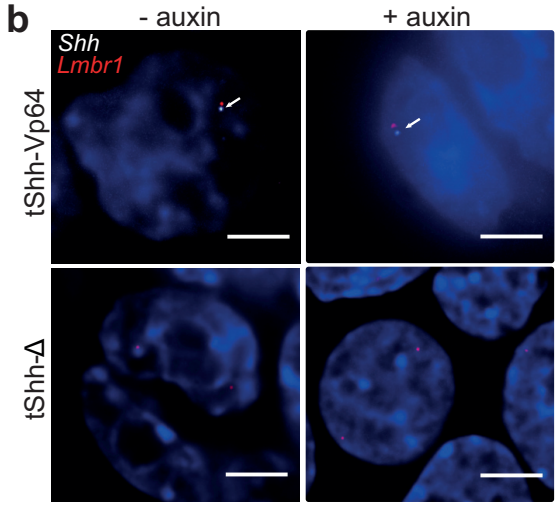
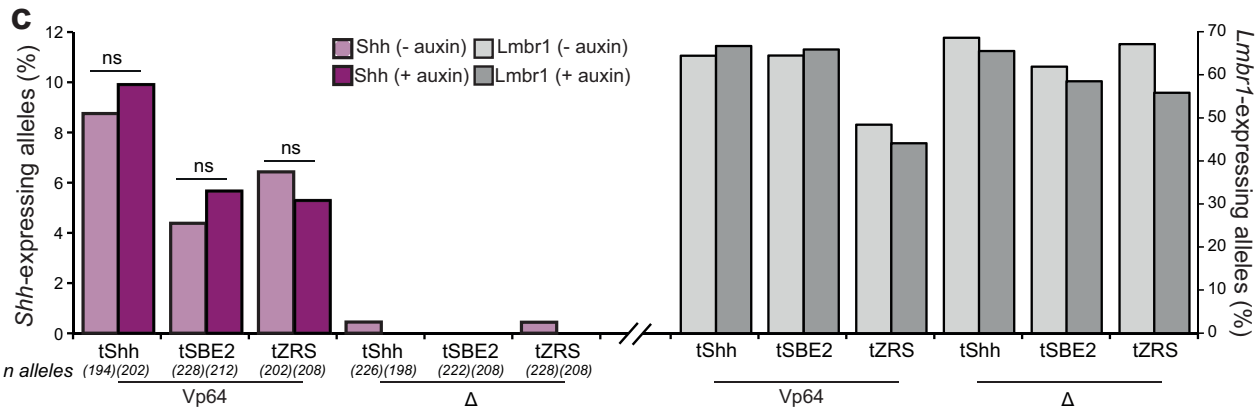

a
Shh-expressing alleles (%)
Lmbr1-expressing alleles (%)
Shh (- auxin)
Lmbr1 (- auxin)
Shh (+ auxin)
Lmbr1 (+ auxin)
ns
ns
WT
SCC1-AID
WT
SCC1-AID
WT
SCC1-AID
WT
SCC1-AID
n alleles (190) (206) (462) (366) (210) (220) (206) (208)
tSBE6-Vp64
tSBE6-Δ
tSBE6-Vp64
tSBE6-Δ
b
Shh
SBE6
tSBE6-Δ
- auxin
tSBE6-Vp64
- auxin
tSBE6-Vp64
+ auxin
c
Shh/SBE6
- auxin
+ auxin
***
**
***
Interprobe distance (μm)
tSBE6-Δ
tSBE6-Vp64

Figure 4

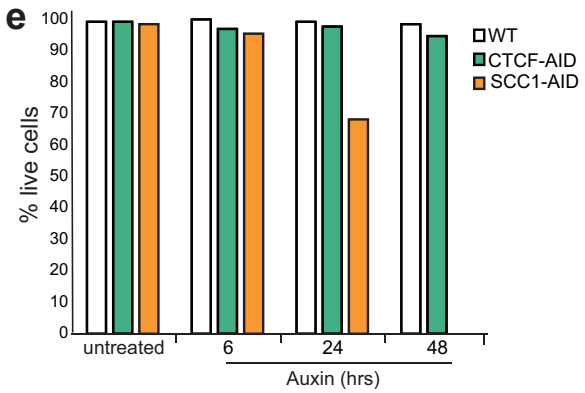




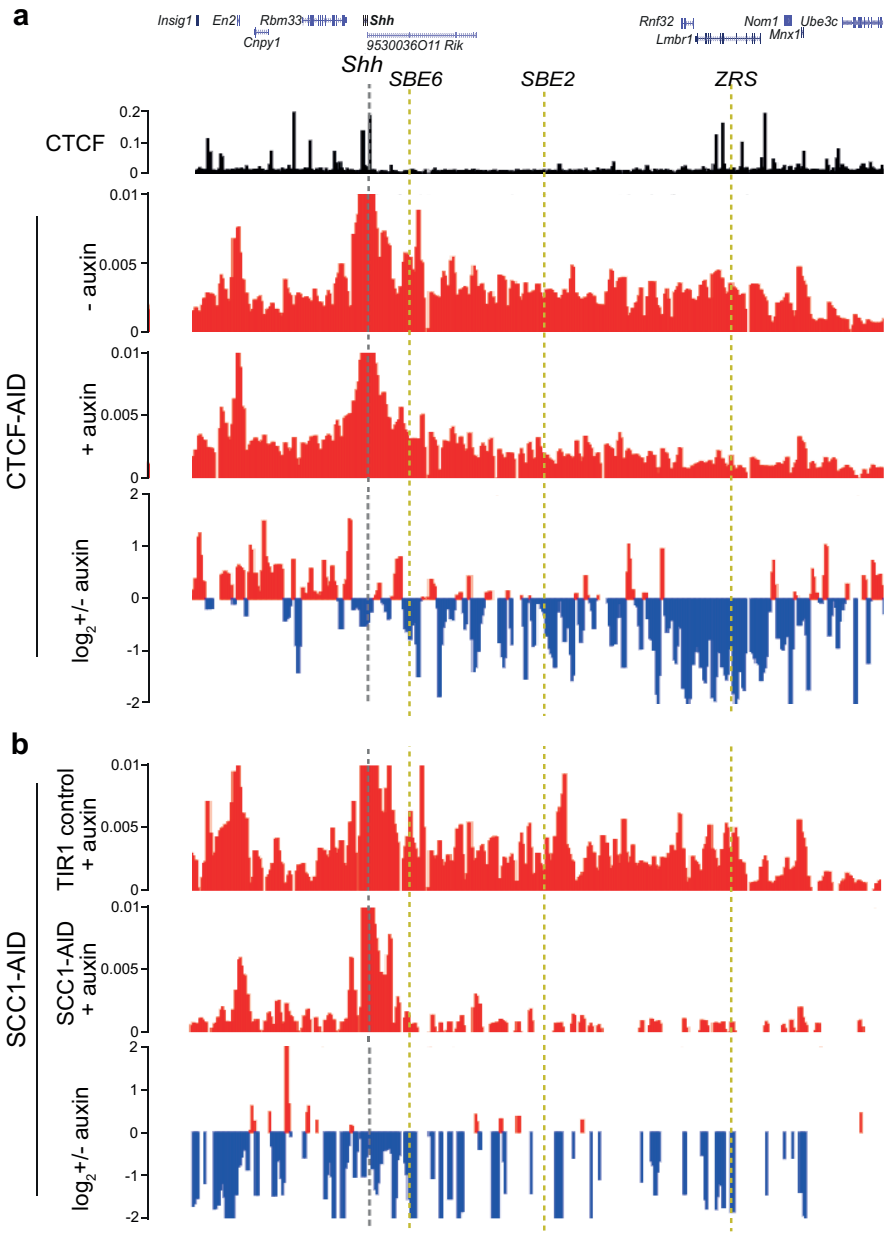


d

TALE constructs		*p* value	
		tShh-Vp64 + aux	tShh-Δ - aux
Rep1:	tShh-Vp64 - aux	0.17	< **0.0001**
	tShh-Δ + aux	< **0.0001**	
Rep2:	tShh-Vp64 - aux	0.73	< **0.0001**
	tShh-Δ + aux	< **0.0001**	
		tSBE2-Vp64 + aux	tSBE2-Δ - aux
Rep1:	tSBE2-Vp64 - aux	1	**0.003**
	tSBE2-Δ + aux	**0.006**	
Rep2:	tSBE2-Vp64 - aux	0.66	**0.002**
	tSBE2-Δ + aux	**0.0004**	
		tZRS-Vp64 + aux	tZRS-Δ - aux
Rep1:	tZRS-Vp64 - aux	0.3	**0.003**
	tZRS-Δ + aux	**0.002**	
Rep2:	tZRS-Vp64 - aux	0.68	**0.0005**
	tZRS-Δ + aux	**0.0009**	

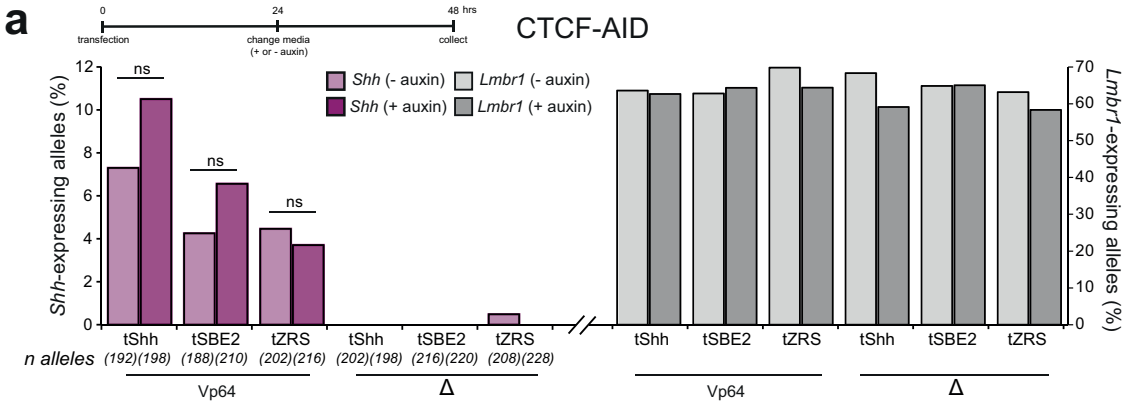


Extended Data Figure 1.

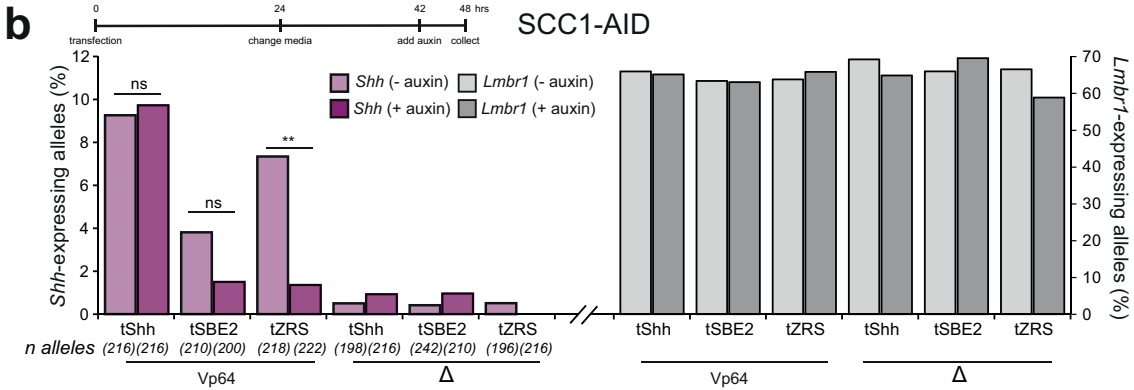


Extended Data Figure 2.

a



b



c

CTCF-AID

TALE constructs		*p* value	
		tShh-Vp64 + aux	tShh-Δ - aux
Rep1:	tShh-Vp64 - aux	0.17	< **0.0001**
	tShh-Δ + aux	< **0.0001**	
Rep2:	tShh-Vp64 - aux	0.29	< **0.0001**
	tShh-Δ + aux	< **0.0001**	
		tSBE2-Vp64 + aux	tSBE2-Δ - aux
Rep1:	tSBE2-Vp64 - aux	0.82	**0.0003**
	tSBE2-Δ + aux	**0.001**	
Rep2:	tSBE2-Vp64 - aux	0.38	**0.002**
	tSBE2-Δ + aux	<**0.0001**	
		tZRS-Vp64 + aux	tZRS-Δ - aux
Rep1:	tZRS-Vp64 - aux	0.84	**0.0001**
	tZRS-Δ + aux	**0.0001**	
Rep2:	tZRS-Vp64 - aux	0.80	**0.01**
	tZRS-Δ + aux	**0.003**	

d

SCC1-AID

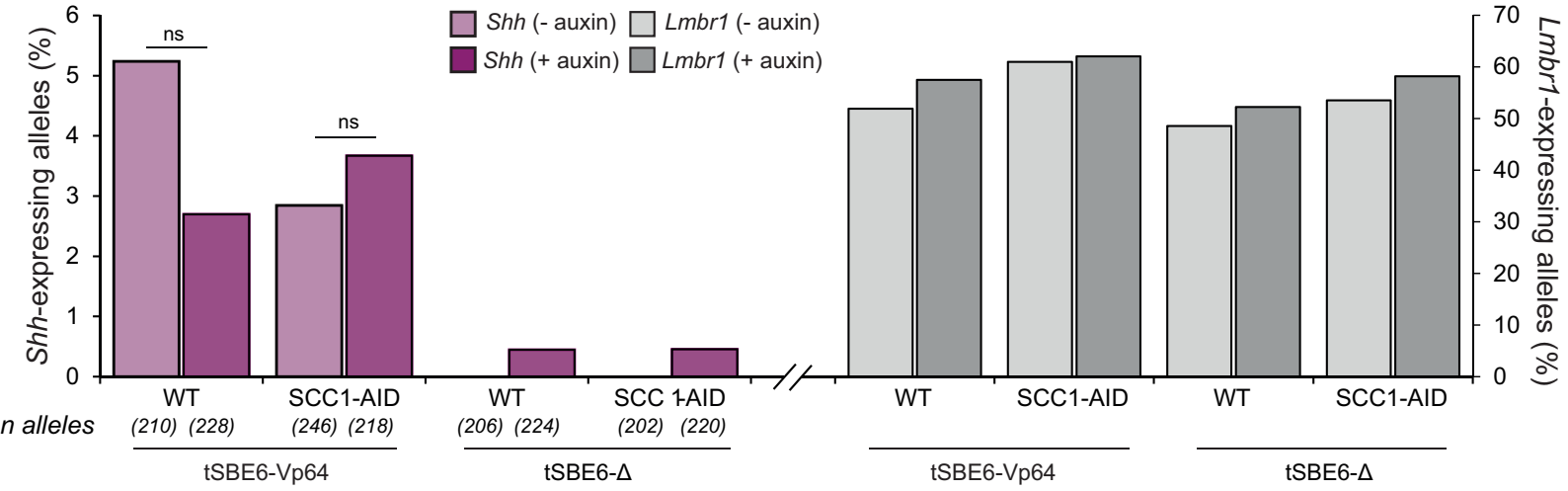
TALE constructs		*p* value	
		tShh-Vp64 + aux	tShh-Δ - aux
Rep1:	tShh-Vp64 - aux	0.87	< **0.0001**
	tShh-Δ + aux	< **0.0001**	
Rep2:	tShh-Vp64 - aux	1.00	< **0.0001**
	tShh-Δ + aux	< **0.0001**	
		tSBE2-Vp64 + aux	tSBE2-Δ - aux
Rep1:	tSBE2-Vp64 - aux	**0.005**	**0.0001**
	tSBE2-Δ + aux	0.50	
Rep2:	tSBE2-Vp64 - aux	0.22	**0.01**
	tSBE2-Δ + aux	0.68	
		tZRS-Vp64 + aux	tZRS-Δ - aux
Rep1:	tZRS-Vp64 - aux	< **0.0001**	< **0.0001**
	tZRS-Δ + aux	0.36	
Rep2:	tZRS-Vp64 - aux	**0.002**	**0.0003**
	tZRS-Δ + aux	0.25	

e

TALE-Vp64 constructs	Probe pairs	Cell lines - or + auxin (no. of alleles)	*p* value	Median distances (μm)	Interquartile distances (μm)
tSBE2-Vp64	**Shh-SBE2**	**CTCF-AID** – (72)		0.49	0.38-0.62
		CTCF-AID + (58)	0.12	0.61	0.38-0.68
		SCC1-AID – (89)		0.50	0.33-0.65
		SCC1-AID + (89)	< **0.0001**	0.77	0.52-1.06
	SBE2-ZRS	**CTCF-AID** – (72)		0.39	0.30-0.60
		CTCF-AID + (58)	**0.004**	0.52	0.42-0.68
		SCC1-AID – (89)		0.45	0.31-0.59
		SCC1-AID + (89)	< **0.0001**	0.69	0.44-1.02
	Shh-ZRS	**CTCF-AID** – (72)		0.34	0.28-0.50
		CTCF-AID + (58)	0.06	0.46	0.30-0.61
		SCC1-AID – (89)		0.36	0.22-0.49
		SCC1-AID + (89)	< **0.0001**	0.69	0.45-1.07
tZRS-Vp64	**Shh-SBE2**	**CTCF-AID** – (80)		0.39	0.26-0.55
		CTCF-AID + (66)	**0.04**	0.45	0.33-0.60
		SCC1-AID – (71)		0.57	0.40-0.71
		SCC1-AID + (77)	**0.003**	0.70	0.51-0.88
	SBE2-ZRS	**CTCF-AID** – (80)		0.41	0.24-0.52
		CTCF-AID + (66)	0.43	0.42	0.30-0.55
		SCC1-AID – (71)		0.46	0.34-0.66
		SCC1-AID + (77)	**0.0003**	0.63	0.49-0.82
	Shh-ZRS	**CTCF-AID** – (80)		0.35	0.25-0.47
		CTCF-AID + (66)	<**0.0001**	0.49	0.37-0.61
		SCC1-AID – (71)		0.41	0.29-0.53
		SCC1-AID + (77)	<**0.0001**	0.68	0.47-0.97

Extended Data Figure 3.

a



b

TALEs and cell lines		*p* value	
wild-type		**tSBE6-Vp64 + aux**	**tSBE6-Δ - aux**
Rep1:	**tSBE6-Vp64 - aux**	0.67	**0.0005**
	tSBE6-Δ + aux	**<0.0001**	
Rep2:	**tSBE6-Vp64 - aux**	0.22	**0.0009**
	tSBE6-Δ + aux	0.12	
SCC1-AID		**tSBE6-Vp64 + aux**	**tSBE6-Δ - aux**
Rep1:	**tSBE6-Vp64 - aux**	0.40	**<0.0001**
	tSBE6-Δ + aux	**<0.0001**	
Rep2:	**tSBE6-Vp64 - aux**	0.79	**0.02**
	tSBE6-Δ + aux	**0.02**	

c

Probe Pairs	TALE- constructs	Cell lines - or + auxin (no. of alleles)	*p* value	Median distances (µm)	Interquartile distances (µm)
Shh-SBE6	**tSBE6-Vp64**	**SCC1-AID** – (91)		0.42	0.28-0.55
		SCC1-AID + (72)	0.75	0.42	0.28-0.60
	tSBE6-Δ	**SCC1-AID** – (100)		0.32	0.21-0.45
		SCC1-AID + (50)	**0.0006**	0.43	0.29-0.60

Extended Data Figure 4.