Table S1. CRISPR primers

Name	Deletion/ inversion size	Vector primers	Genotyping primers
ΔCTCF1 mm9 chr5: 28,777,560- 28,778,320	1.2kb	1F-CACCGGCTAGCCATGAAGACAAGCA 1R- AAACTGCTTGTCTTCATGGCTAGCC 2F- CACCGGGCCCACTATTGTCAGAAAAT 2R- AAACATTTCTGACAATAGTGGGCCC	Deletion F- ATGAATGCCTGCAGTGGTTC Wild Type F- CTAATAGCAGCTGACCACGAAC R- TAGACTCAGAGGTTCGTAAG
ΔCTCF2 mm9 chr5: 28,795,867- 28,796,937	1.35kb	1F-CACCGATCACATCTGCAGGTATTCC 1R- AAACGGAATACCTGCAGATGTGATC 2F- CACCGAAATGATTTCCGTCCTCTAC 2R- AAACGTAGAGGACGGAAATCATTTC	Deletion F-GAATATAGACTGGTGAATGGATC WildType F- AATCACATGGCTGTGAGATAC R- CTTCAGCAGCTCCAAGACTG
ΔCTCF3 mm9 chr5: 29,619,575- 29,620,326	750bp	1F-CACCGATAGCCAGCTTTATGCTTTC 1R- AAACGAAAGCATAAAGCTGGCTATC 2F- CACCGTGTGCATCTCATACTGAGAA 2R- AAACTTCTCAGTATGAGATGCACAC	Deletion F- GCATACTGGCAGCATCACTG Wild Type F- AGACCAGACTTGTCAACTTGG R- CCTGACCCTCAGGTGTTAGC
ΔCTCF4 mm9 chr5: 29,704,457- 29,705,661	1.3kb	1F-CACCGCACTTGTCGCCCGTCATCAT 1R- AAACATGATGACGGGCTACAAGTGC 2F- CACCGTAAACAACGCCCTTGATATG 2R- AAACCATATCAAGGGCGTTGTTTAC	Deletion F- CCTCAGTTAGCAGGAGATG R- TCAGACACGTTGTTGCTGAC Wild Type F- GCAGCTTAGCACTCTGAAGTC
ΔCTCF5 mm9 chr5: 29,715,776- 29,716,437	1.2kb	1F-CACCGAAGATGCCTGCTTAGTGCCG 1R- AAACCGGCACTAAGCAGGCATCTTC 2F- CACCGCATGGAAGTGGACGACATGC 2R- AAACGCATGTCGTCCACTTCCATGC	Deletion F- TTACAAGACAAACCATCAGAC Wild Type F- TCAACTTGGATCGGAGAATC R- GTGATCCTGCTGTGGATTAG
Δ/Inv35 mm9 chr5: 29,684,041- 29,718,587	35kb	1F-CACCGAGCCTCTCACTTTATGCCCC 1R- AAACGGGGCATAAAGTGAGAGGCTC 2F- CACCGAGGGCTTCATGAATTACACC 2R- AAACGGTGTAATTCATGAAGCCCTC	Deletion F- CACATATCACACATATCCAGG Wild Type F- GATTTCCTCTGTGTGAAGATC R- AATCATGATTACAGATGAGTG Inversion F- CACATATCACACATATCCAGG R- GATTTCCTCTGTGTGAAGATC

Table S2. Median interprobe distances for *Cnpy1-Shh*, *Shh*-SBE2, SBE2-ZRS, *Shh*-ZRS, ZRS-*Lmbr1*, ZRS-*Mnx1*, ZRS-*Ube3c*, *Lmbr1-Mnx1* probes in wild type and ΔCTCF ESCs

ESCs	WT	ΔCTCF1	ΔCTCF2	ΔCTCF3	ΔCTCF4	ΔCTCF5
Fosmids	**Interprobe distance (nm)**					
Cnpy1-Shh	453	376 (p = 0.03)	439			
Shh-SBE2	396	473 (p = 0.003)	474 (p = 0.0008)	443 (p = 0.01)	397	390
SBE2-ZRS	390	473 (p < 0.0001)	522 (p < 0.0001)	551 (p < 0.0001)	449 (p = 0.009)	401
Shh-ZRS	341	449 (p = 0.0005)	449 (p < 0.0002)	453 (p < 0.0001)	324	434 (p = 0.0001)
ZRS-Lmbr1	533				473	427 (p = 0.01)
ZRS-Mnx1	503				427 (p = 0.04)	453
Lmbr1-Mnx1	411				406	413

Statistical analysis of data for Figures 3C, S3B, 4C & S4B. Interprobe distances are median values, p-values from Mann-Whitney U Tests.

Table S3. Median interprobe distances for *Shh*-SBE2, SBE2-ZRS and *Shh*-ZRS probes in wild type and ΔCTCF E11.5 limb distal posterior tissue

Posterior	WT	ΔCTCF1	ΔCTCF2	ΔCTCF3	ΔCTCF4	ΔCTCF5
Fosmids	**Interprobe distance (nm)**					
Shh-SBE2	406	429 (p = 0.005)	428 (p = 0.005)	507 (p < 0.0001)	418 (p = 0.04)	338
SBE2-ZRS	427	428	617 (p < 0.0001)	406	446	406
Shh-ZRS	241	276 (p = 0.002)	268 (p = 0.006)	276 (p = 0.004)	242	221

Statistical analysis of data for Figures 5B. Interprobe distances are median values, p-values from Mann-Whitney U Tests.

Table S4. Median interprobe distances for *Shh*-SBE2, SBE2-ZRS and *Shh*-ZRS probes in wild type and ΔCTCF E11.5 limb distal anterior tissue

Anterior	WT	ΔCTCF1	ΔCTCF2	ΔCTCF3	ΔCTCF4	ΔCTCF5
Fosmids	**Interprobe distance (nm)**					
Shh-SBE2	422	587 (p < 0.0001)	486 (p = 0.0003)	607 (p < 0.0001)	486	422
SBE2-ZRS	425	486 (p = 0.02)	436	427	411	314 (p = 0.03)
Shh-ZRS	250	342 (p = 0.0001)	314 (p = 0.02)	341 (p = 0.001)	268	276

Statistical analysis of data for Figures 5C. Interprobe distances are median values, p-values from Mann-Whitney U Tests.

Table S5. Co-localisation frequency (<200 nm) of *Shh* and ZRS probes in wild type and ΔCTCF E11.5 distal anterior and posterior limb tissue

	WT	ΔCTCF1	ΔCTCF2	ΔCTCF3	ΔCTCF4	ΔCTCF5
Tissue	**Co-localisation Frequency (%)**					
Anterior	20	10.9	14.9	11	15.7	17
Posterior	36	22 (p = 0.008)	17 (p = 0.0003)	22 (p = 0.008)	26 (p = 0.04)	32

Statistical analysis of data for Fig. 5D. p-values from Fisher's Exact Tests.

Table S6. Median interprobe distances for *Shh*-SBE2, SBE2-ZRS, *Shh*-ZRS & ZRS-*Mnx1* probes in wild type, Δ35 and Inv35 ESCs

ESCs	WT	35kbΔ	35kbinv
Fosmids	**Interprobe distance (nm)**		
Shh-SBE2	405	347 (p = 0.04)	
SBE2-ZRS	351	286 (p = 0.049)	
Shh-ZRS	336	275 (p = 0.01)	
ZRS-*Mnx1*	421	349 (p = 0.002)	376 (p = 0.04)

Statistical analysis of data for Figures 7D. Interprobe distances are median values, p-values from Mann-Whitney U Tests.

Table S7. RT-PCR primers

Region	Primers
qRT Shh	F- ACCCCGACATCATATTTAAGGA R- TTAACTTGTCTTTGCACCTCTGA
qRTMnx1	F- GATGCCGGACTTCAGCTC R- AGCTGCTGGCTGGTGAAG
RT Lmbr1	ex1 F- ACAGCCAAGTGCGAGAGTCC ex5 F- TTCTTTCTGGAATCAGAAGG ex6 R- CCATACTGGCAGCATCACTGT
RT Hprt	F- CACAGGACTAGAACACCTGC R- GCTGGTGAAAAGGACCTC

Table S8. Fosmid Probes

Region	Whitehead (Sanger) Name	Ensembl name	Coordinates Start	End	Size (bp)
Cnpy1	WI1-2816N11		28567538	28605641 End 27975636	38104
Shh	WI1-0574O18	G135P64333A4	28754458	28795879	41421
SBE2	WI1-1275C09	G135P603171G8	29195832	29239355	43523
ZRS	WI1-1047E14	G135P600929F6	29611727	29653695	41968
Lmbr1 pr.	WI1-2285K14		29692859	29736053	43195
Mnx1	WI1-1204B06		29791124	29827491	36368

Names are Ensembl (r 45) (http://jun2007.archive.ensembl.org/Mus_musculus/index.html). Mouse genome assembly number: NCBI m37

Table S9. Mouse 5C primers for *Shh* and *USP22* regions

Click here to Download Table S9

Table S10. 5C sequencing reads

E11.5 tissues

Sample	Number of reads	Number of used reads
E11.5 Body	2810910	1090908
E11.5 700kbΔ rep 1i	1562060	1082158
E11.5 700kbΔ rep 1ii	1415296	963894
E11.5 700kbΔ rep 2i	2566066	1853170
E11.5 700kbΔ rep 2ii	3898600	2847490
E11.5 Anterior limb rep 1	3302406	1925282
E11.5 Anterior limb rep 2	4664488	2579458
E11.5 Posterior limb rep 1	2340108	1432302
E11.5 Posterior limb rep 2	4721304	2662420
E14 ESCs wild type rep 1	2999100	1850978
E14 ESCs wild type rep 2	3842910	2447598
E14 ESCs ΔCTCF1 rep1	4324812	3370342
E14 ESCs ΔCTCF1 rep 2	4803792	3744989
E14 ESCs ΔCTCF2 rep1	5037248	2652966
E14 ESCs ΔCTCF2 rep2	4197292	2230540
E14 ESCs ΔCTCF3 rep1	4530282	2634184
E14 ESCs ΔCTCF3 rep2	3623840	1803502
E14 ESCs ΔCTCF4 rep1	4123544	2876952
E14 ESCs ΔCTCF4 rep2	2760100	1839494
E14 ESCs ΔCTCF5 rep1	3982104	1852298
E14 ESCs ΔCTCF5 rep2	3957266	1706900
E14 ESCs Δ35kb rep 1	3027284	2464184
E14 ESCs Δ35kb rep 2	3352334	2732898

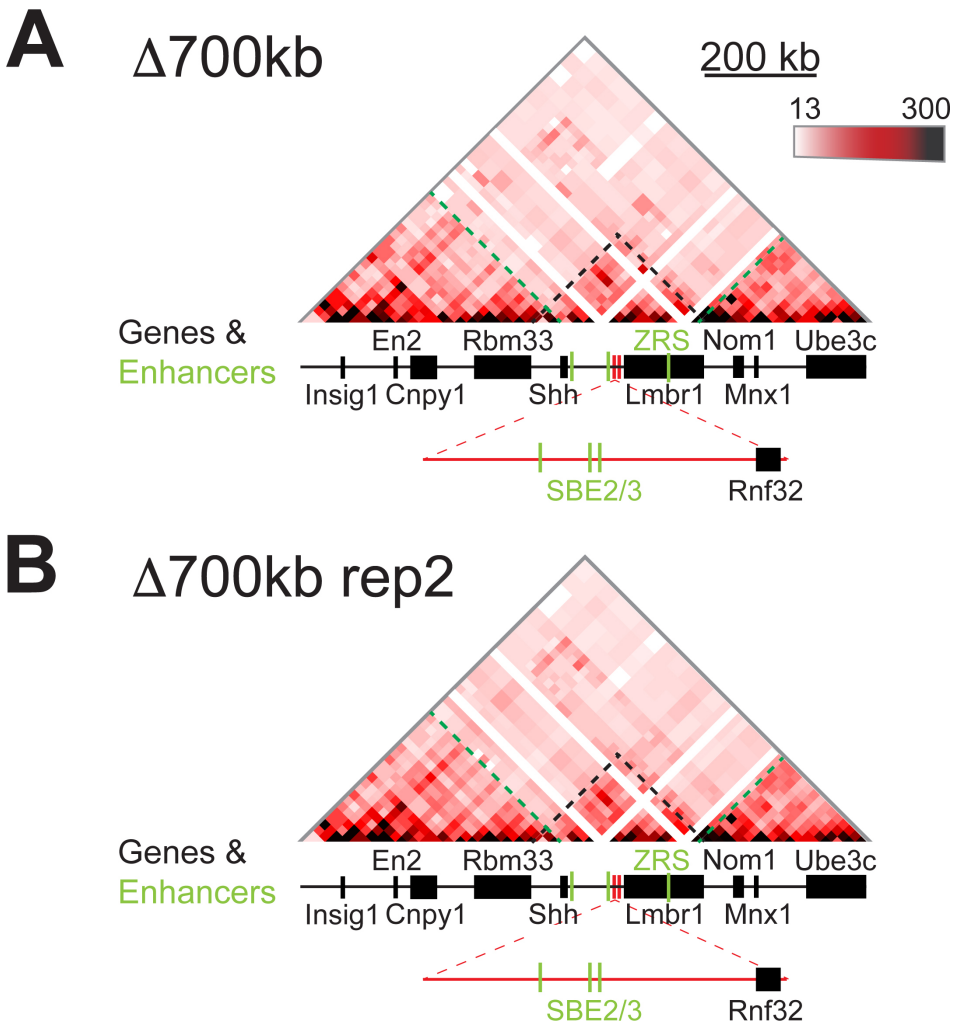
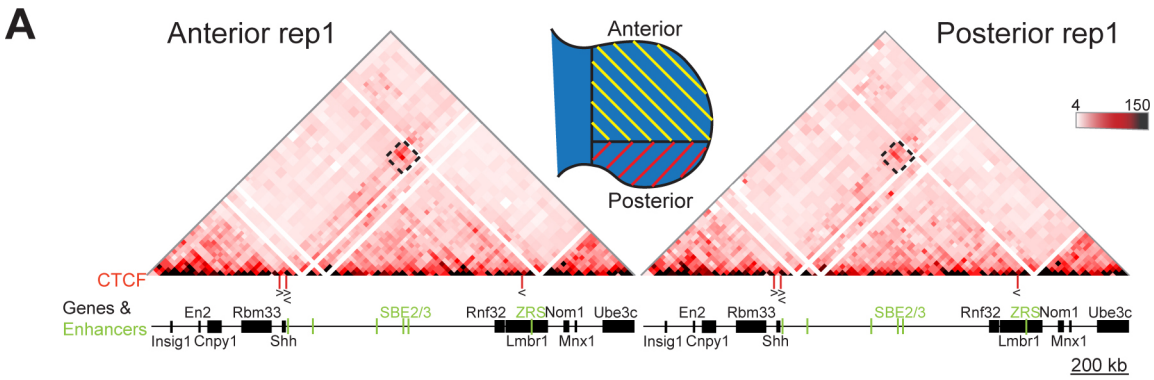
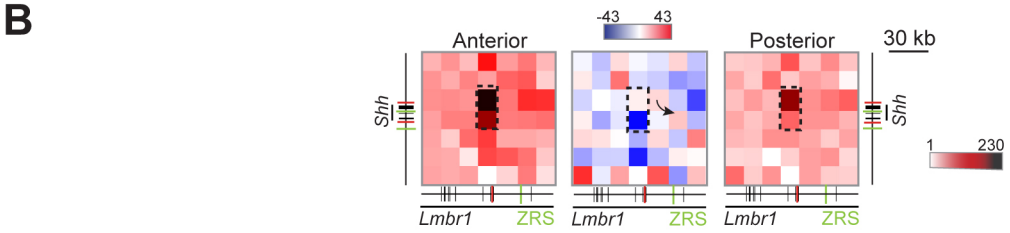
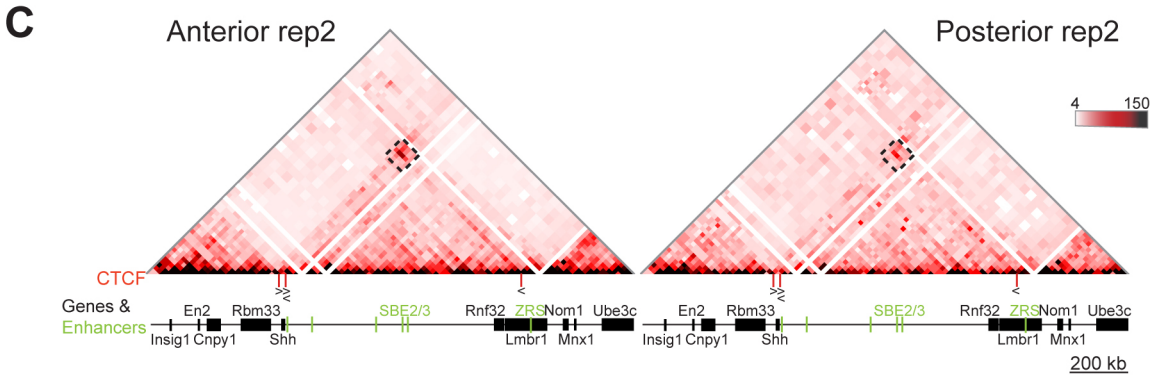


Figure S1. Unchanged Shh TAD boundaries and 5C-seq enriched interactions between loci containing *Shh* and ZRS in biological replicates of cells derived from E11.5$^{\Delta 700kb/\Delta 700kb}$ embryos. (A & B) Heat maps shows the average interaction frequencies across *Shh* and its regulatory domain in E11.5 whole embryos homozygous for the 700kb deletion (replicate biological sample). Interaction frequencies were normalized to sum up to 50000 reads and adaptive coarsegraining of the matrices were performed to reduce noise with the three lowest coverage bins masked and the data shown is binned over 21-kb windows and are colour-coded according to the corresponding scales as described in Figure 1F. Green dashed lines highlight the TAD boundary locations, black dashed lines indicate the *Shh* TAD boundaries and the reduced size of the TAD.







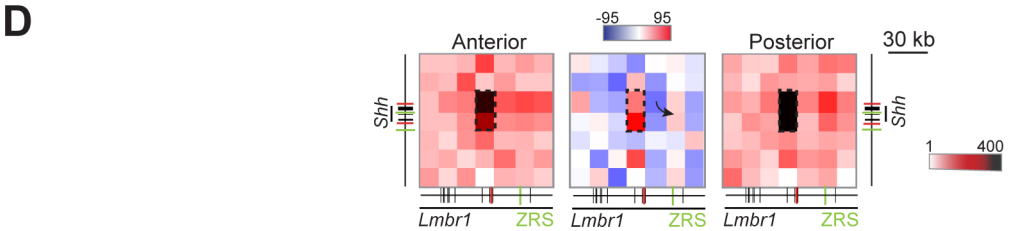


Figure S2. 5C-seq analysis in biological replicates of cells dissected from E11.5 distal anterior and posterior tissue. (A) Heat maps showing 5C data from distal anterior and posterior limb bud cells of E11.5 embryos, across the 1.7-Mb Shh region shown in Figure 1. Heat map intensities represent the average of interaction frequency for each window, colour-coded according to the scale shown. Interaction frequencies were normalized to sum up to 50000 reads and adaptive coarsegraining of the matrices were performed to reduce noise with the three lowest coverage bins masked and the data shown is binned over 21-kb windows. Green dashed lines indicate TAD boundaries, the interactions highlighted by the black dashed boxes locate the region of the heat maps shown in **(B)** at higher resolution. Schematic indicating the limb bud portions dissected for anterior and posterior cell populations. **(B)** Higher resolution (15-kb binning) heat maps from 5C data displayed in **(A)** showing interactions between 105kb genomic regions encompassing *Shh* and ZRS. Left-hand and right-hand heat maps from anterior and posterior tissues respectively with intensities representing the average of interaction frequency for each window, colour-coded according to the scale shown based on raw 5C data sets. The comparison heatmap (centre), with compared data sets normalised by read count, shows interactions enriched in posterior cells (red) and anterior cells (blue). Enriched interactions between loci containing CTCF binding sites are indicated by the black dashed boxes, the arrow in the comparison heat map highlights posterior enriched interactions between *Shh* and ZRS. **(C & D)** Second biological replicates showing 5C data for distal anterior and posterior limb bud cells of E11.5 embryos at 21-kb resolution **(C)** as in **(A)** and zoomed in higher resolution (15-kb) **(D)** as in **(B)**.

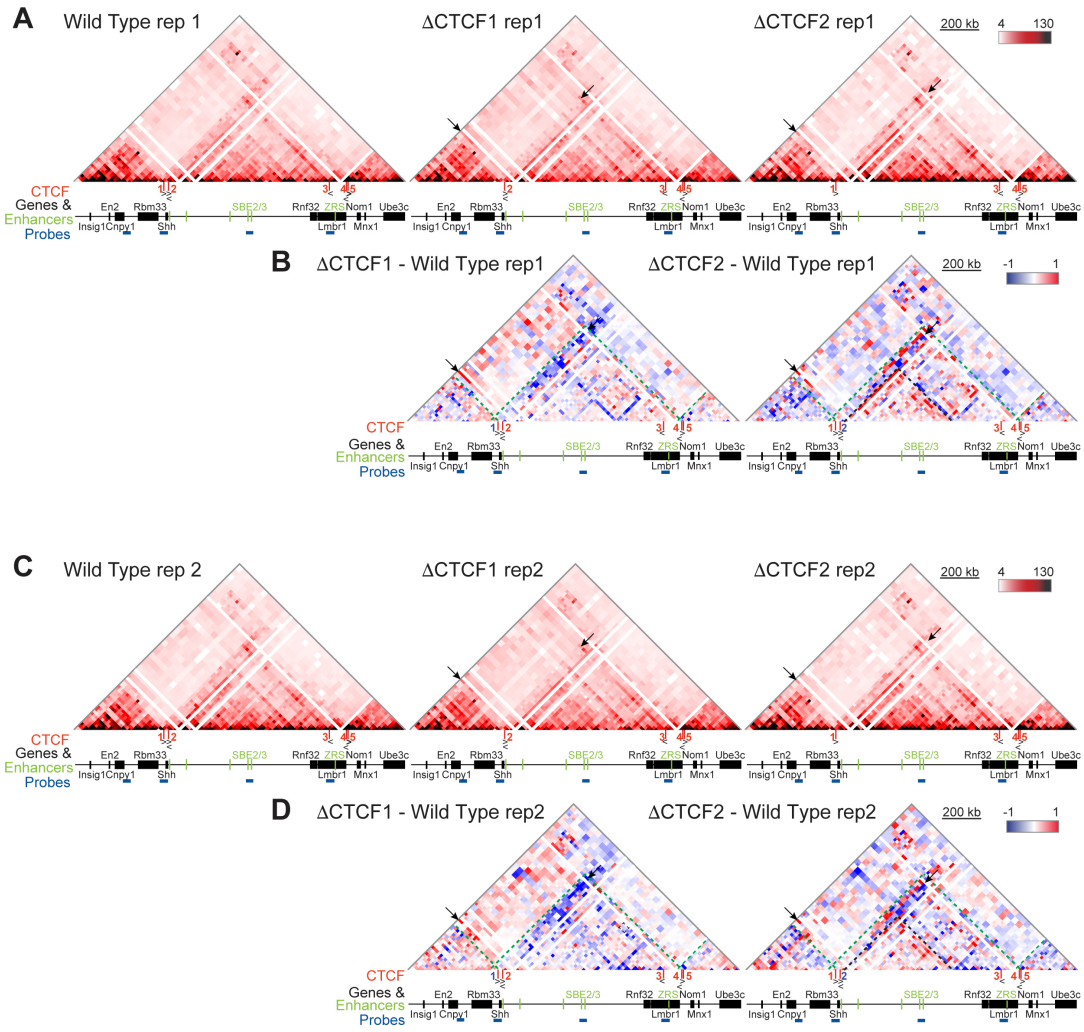


Figure S3. 5C-seq identifies perturbations to chromatin conformation in biological replicates of ΔCTCF1 and ΔCTCF2 ESCs. (A & C) Heat maps showing 5C data from two biological replicates of wild type, ΔCTCF1 and ΔCTCF2 ESCs across the 1.7-Mb Shh region shown in Figure 1. Heat map intensities represent the average of interaction frequency for each window, colour-coded according to the scale shown. Interaction frequencies were normalized to sum up to 50000 reads and adaptive coarsegraining of the matrices were performed to reduce noise with the three lowest coverage bins masked and the data shown is binned over 21-kb windows. **(B & D)** Heat maps comparing ΔCTCF1 or ΔCTCF2 enrichment (red) with wild type (blue). For comparison of 5C matrices across conditions, additional observed/expected normalization were performed by dividing each diagonal of the matrix by its mean. Green dashed lines indicate TAD boundaries, black dashed lines highlight enriched contacts within sub-TADs (ΔCTCF2), black arrows highlight loss of interactions in ΔCTCF1 or ΔCTCF2 cells and black arrowheads indicate enriched interactions in ΔCTCF1 or ΔCTCF2 cells.

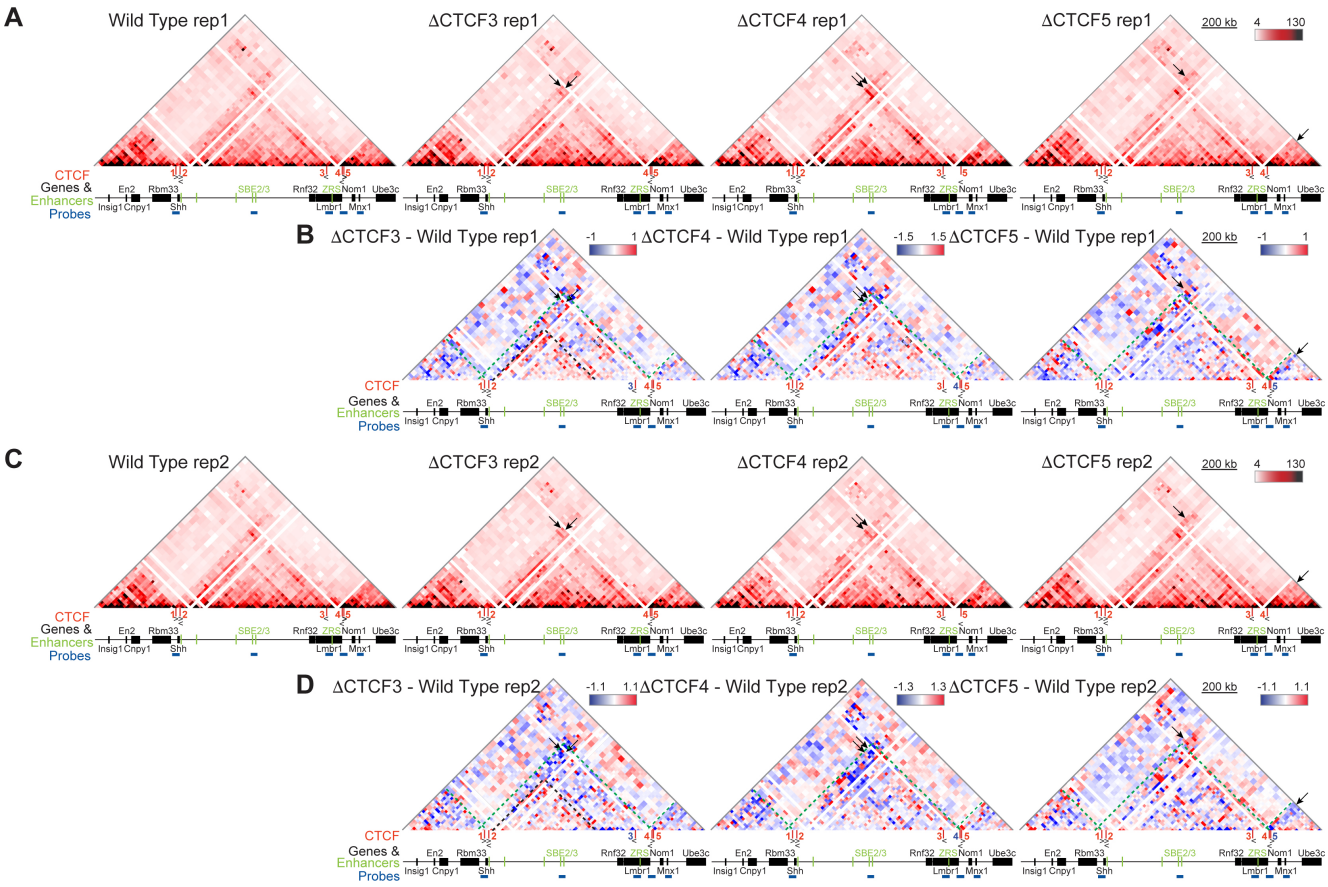


Figure S4. 5C-seq identifies perturbations to chromatin conformation in biological replicates of ΔCTCF3, ΔCTCF4 and ΔCTCF5 ESCs. (A & C) Heat maps showing 5C data from two biological replicates of wild type, ΔCTCF3, ΔCTCF4 and ΔCTCF5 ESCs, across the 1.7-Mb Shh region shown in Figure 1. Heat map intensities represent the average of interaction frequency for each window, colour-coded according to the scale shown. Interaction frequencies were normalized to sum up to 50000 reads and adaptive coarsegraining of the matrices were performed to reduce noise with the three lowest coverage bins masked and the data shown is binned over 21-kb windows. **(B & D)** Heat maps comparing ΔCTCF3, ΔCTCF4 or ΔCTCF5 enrichment (red) with wild type (blue). For comparison of 5C matrices across conditions, additional observed/expected normalization were performed by dividing each diagonal of the matrix by its mean. Green dashed lines indicate TAD boundaries, black dashed lines highlight enriched contacts within sub-TADs (ΔCTCF3), black arrows highlight loss of interactions in ΔCTCF3 or ΔCTCF5 cells and black arrowheads indicate enriched interactions in ΔCTCF3 or ΔCTCF5 cells.

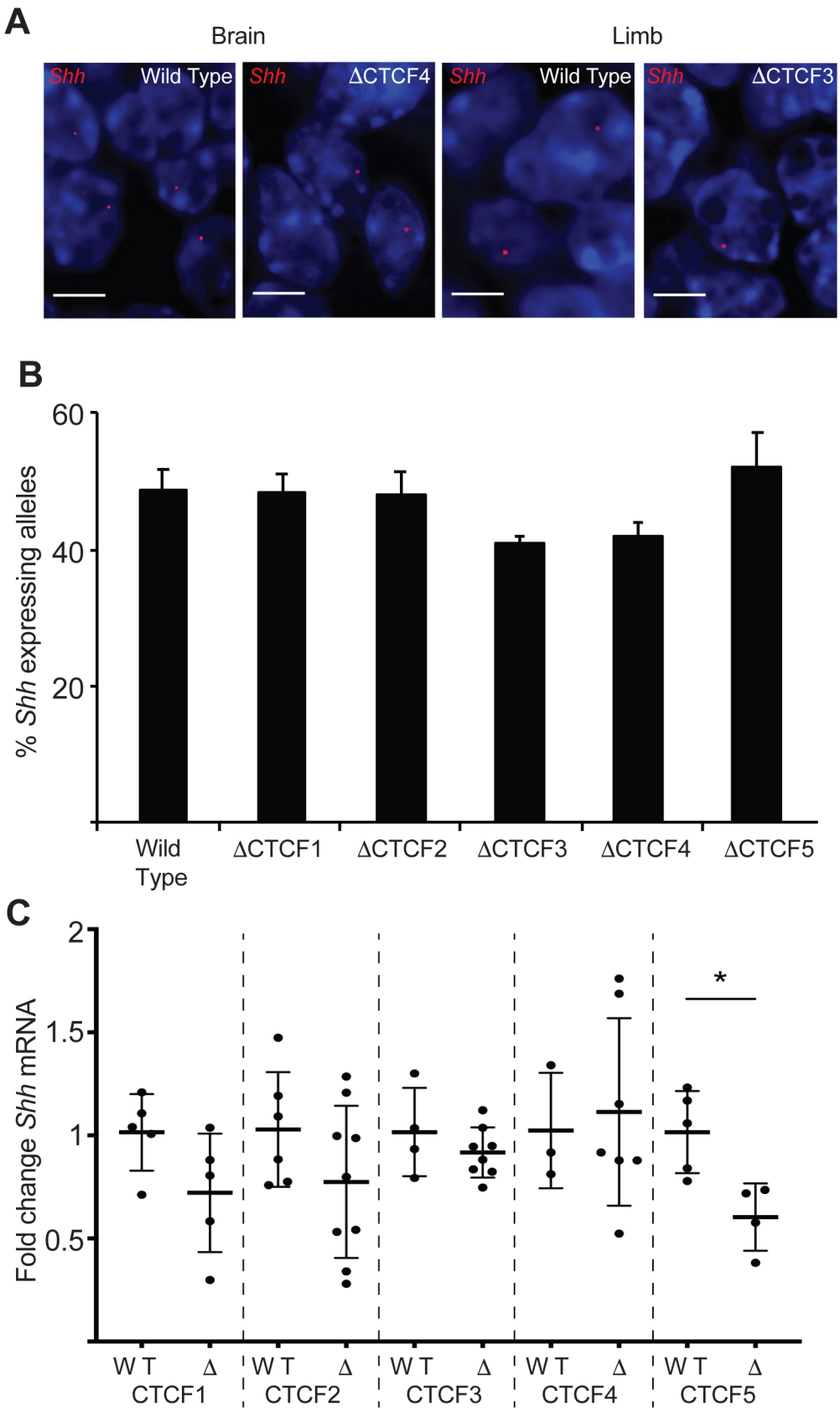


Figure S5. *Shh* RNA FISH and qRT-PCR for limb buds. (A) Images of representative RNA FISH signal (red) of *Shh*-expressing nuclei within the brain and ZPA of limb tissue of E11.5 wild type and ΔCTCF embryos. Scale bars = 5 μm. **(B)** Counts of expressing alleles detected by RNA FISH within the ZPA of the E11.5 limb bud. The statistical significance between data sets was examined by Fisher's exact test. **(C)** Graph of fold change of *Shh* expression determined by qRT-PCR of cDNA made from E11.5 limb buds. Mutants from each of ΔCTCF lines are compared to wild type litter mates. Each dot represents a single embryo and data is graphed as mean +/- SD. The statistical significance between data sets was examined using unpaired Student t-tests, * < 0.05.

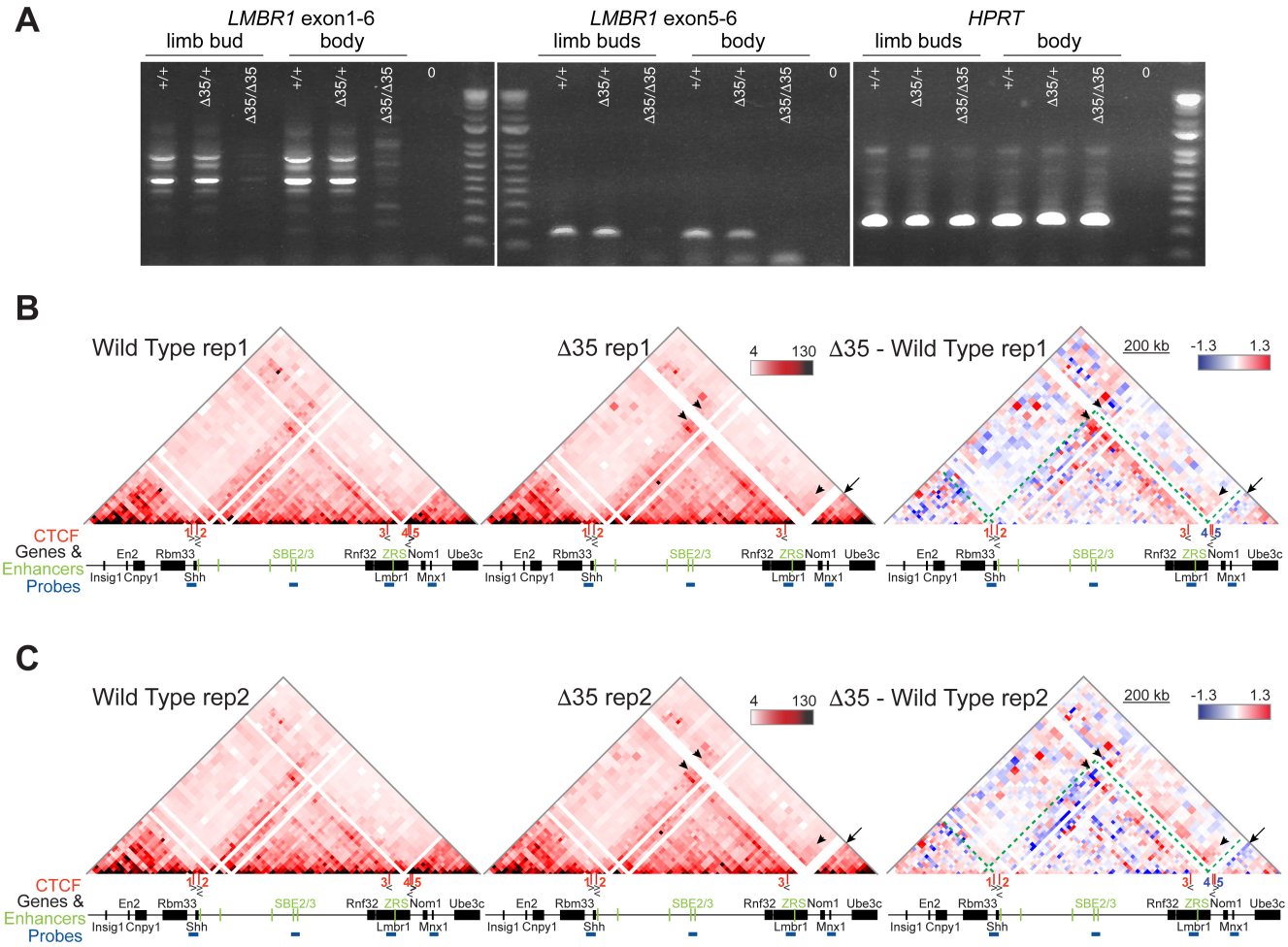


Figure S6. 5C-seq identifies perturbations to local chromatin conformation in biological replicates of Δ35 ESCs.

(A) RT-PCR analysis of gene expression in limb buds and bodies from E11.5 wild type, heterozygous and homozygous 35kb deletion (Δ35) embryos showing a loss of transcription through Lmbr1. **(B & C)**Heat maps showing 5C data from two biological replicates of wild type and Δ35 ESCs, across the 1.7-Mb Shh region shown in Figure 1. Heat map intensities represent the average of interaction frequency for each window, colour-coded according to the scale shown. Interaction frequencies were normalized to sum up to 50000 reads and adaptive coarsegraining of the matrices were performed to reduce noise with the three lowest coverage bins masked and the data shown is binned over 21-kb windows. Right-hand heat maps comparing Δ35 enrichment (red) with wild type (blue). For comparison of 5C matrices across conditions, additional observed/expected normalization were performed by dividing each diagonal of the matrix by its mean. Green dashed lines indicate TAD boundaries, black arrows highlight loss of interactions in Δ35 cells and black arrowheads indicate enriched interactions in Δ35 cells.