Shh and ZRS enhancer co-localisation is specific to the zone of polarizing activity

Citation for published version:
Williamson, I, Lettice, L, Hill, R & Bickmore, W 2016, 'Shh and ZRS enhancer co-localisation is specific to the zone of polarizing activity' Development. DOI: 10.1242/dev.139188

Digital Object Identifier (DOI):
10.1242/dev.139188

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Development

Publisher Rights Statement:
© 2016. Published by The Company of Biologists Ltd This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Shh and ZRS enhancer co-localisation is specific to the zone of polarizing activity

Iain Williamson, Laura A. Lettice, Robert E. Hill*, Wendy A. Bickmore*

MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, Crewe Road, Edinburgh EH4 2XU, UK

* Correspondence to: Wendy.Bickmore@igmm.ed.ac.uk or Bob.Hill@igmm.ed.ac.uk

6 key words: 5C, chromosome loop, enhancer, limb development, super-resolution microscopy,
Abstract

Limb-specific Shh expression is regulated by the (~1 Mb distant) ZRS enhancer. In the mouse, limb bud restricted spatiotemporal Shh expression occurs from ~E10-E11.5 at the distal posterior margin and is essential for correct autopod formation. Here, we have analysed the higher-order chromatin conformation of Shh in expressing and non-expressing tissues, both by fluorescence in situ hybridisation (FISH) and by chromosome conformation capture (5C). Conventional and super-resolution light microscopy identified significantly elevated frequencies of Shh/ZRS co-localisation only in the Shh expressing regions of the limb bud, in a conformation consistent with enhancer-promoter loop formation. However, in all tissues and developmental stages analysed, Shh-ZRS spatial distances were still consistently shorter than those to a neural enhancer located between Shh and ZRS in the genome. 5C identified a topologically associating domain (TAD) over the Shh/ZRS genomic region and enriched interactions between Shh and ZRS throughout E11.5 embryos. Shh/ZRS co-localisation, therefore, correlates with the spatiotemporal domain of limb bud-specific Shh expression, but close Shh/ZRS proximity in the nucleus occurs regardless of whether the gene or enhancer is active. We suggest that this constrained chromatin configuration optimises the opportunity for the active enhancer to locate and instigate Shh expression.
Introduction

Chromatin-looping is a popular model by which very long-range enhancers can communicate with their target gene promoter (Benabdallah and Bickmore, 2015), however the relationship of loop formation and gene activation remains unclear. It has been suggested that enhancer-target gene contacts are preformed and present in tissues even where the target gene is not activated (Montavon et al., 2011; Ghavi-Helm et al., 2014). However, other reports indicate enhancer-gene looping is spatially and temporally restricted to cells where the target gene is active. This includes in the developing mouse limb, where elevated levels of co-localisation of the global control region (GCR) and its target 5′HoxD genes is only seen in the cells of the distal posterior portion of the E10.5 limb bud (Williamson et al., 2012).

The complex spatiotemporal gene regulatory circuit in the developing limb is a rich system in which to study the activity of distal regulatory elements and their mechanisms of action. The sonic hedgehog gene (Shh), encodes a morphogen that directs cell fate during organogenesis. Limb-specific expression of Shh is regulated by the ZRS enhancer positioned within an intron of Lmbr1 ~1 Mb away at the opposite end of a large gene desert (Lettice et al., 2002; Lettice et al., 2003) (Figure 1A). The ZRS has a functional role in directing spatiotemporal Shh expression restricted to a region of the distal posterior mesenchyme of the limb bud known as the zone of polarizing activity (ZPA) (Saunders and Gasseling, 1968; Riddle et al., 1993). Limb-specific Shh expression is abrogated upon deletion of ZRS (Sagai et al., 2005), whereas point mutations across the 780-bp conserved sequence of the enhancer can induce anterior, ectopic Shh expression and can cause preaxial polydactyly (Lettice et al., 2003; Sagai et al., 2004; Lettice et al., 2008), triphalangeal thumb (Furniss et al., 2008) or Werner mesomelic syndrome (VanderMeer et al., 2014). Duplications, and even triplication, of the ZRS have been associated with severe forms of polysyndactyly: triphalangeal thumb-polysyndactyly syndrome and Haas type (syndactyly type IV) polysyndactyly (Klopopcki et al., 2008; Sun et al., 2008; Wieczorek et al., 2010).

Previously, fluorescence in situ hybridisation (FISH) and chromosome conformation capture (3C) (Amano et al., 2009) have reported increased associations between Shh and ZRS in E10.5 limb buds compared with other tissues. However, no significant difference in gene/enhancer co-localisation was detected between the ZPA and distal anterior tissue – where Shh is not normally expressed, or indeed in ZPA cells between wild-type and embryos with a deletion of the ZRS. This would be consistent with a model of pre-formed enhancer-
gene contacts. In contrast, FISH has revealed a significant decrease in Shh/ZRS co-localisation in E11.5 ZPA tissue from mouse embryos with a ZRS mutation which decreases ZRS long-range activity (Lettice et al., 2014), suggesting that ZRS/Shh juxtaposition is directly linked to Shh activation.

We have previously combined FISH and 3C carbon copy (5C) to elucidate the role of chromatin conformation in the long-range regulation of the 5′ Hoxd genes during distal limb bud development (Williamson et al., 2012; Williamson et al., 2014). Here we combined these methods to characterise the Shh locus in tissue sections – including those derived from three discrete developmental stages of mouse limb bud development. Spatial proximity of Shh and ZRS, as inferred indirectly from enriched 5C interactions, was identified throughout E11.5 embryos, and 5C data confirmed that Shh and its known enhancers form a compact regulatory chromatin domain. However, using super-resolution microscopy we show that, despite Shh and ZRS being proximal to one another in the nucleus in all tissue types and temporal stages analysed, high levels of Shh/ZRS co-localisation occurs only in ZPA cells at the time of Shh activation. Comparison between Shh/ZRS distances and those between either Shh or ZRS and an intervening genomic locus are consistent with the formation of a chromatin loop between the active gene and enhancer.

Materials and Methods

FISH

For 3D FISH, E10.5, E11.5 and E14.5 embryos from CD1 mice were collected, fixed, embedded, sectioned and processed as previously described (Morey et al., 2007), except that sections were cut at 6 μm. Fosmid clones (Figure 1A, Table S1) were prepared and labelled as previously described (Morey et al. 2007). Between 160-240 ng of biotin- and digoxigenin-labeled fosmid probes were used per slide, with 16-24 μg of mouse Cot1 DNA (Invitrogen) and 10 μg salmon sperm DNA.
**Image analysis**

For 3D analysis of tissue sections by conventional microscopy, slides were imaged with a Hamamatsu Orca AG CCD camera (Hamamatsu Photonics (UK) Ltd, Welwyn Garden City, UK), Zeiss Axioplan II fluorescence microscope with Plan-neofluar or Plan apochromat objectives, a Lumen 200W metal halide light source (Prior Scientific Instruments, Cambridge, UK) and Chroma #89014ET 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. Hardware control, image capture and analysis were performed using Volocity (Perkinelmer Inc, Waltham, MA). Images were deconvolved using a calculated point spread function with the constrained iterative algorithm of Volocity (Perkinelmer Inc, Waltham, MA). Image analysis was carried out using the Quantitation module of Volocity (Perkinelmer Inc, Waltham, MA).

**SIM imaging**

Images were acquired using Structured Illumination Microscopy (SIM) performed on an Eclipse Ti inverted microscope equipped with a Nikon Plan Apo TIRF objective (NA 1.49, oil immersion) and an Andor DU-897X-5254 camera. Laser lines 405, 488 and 561 nm were used. Step size for z stacks was set to 0.120 μm, which is well within the Nyquist criterion. For each focal plane, 15 images (5 phases, 3 angles) were captured with the NIS-Elements software. SIM image processing and reconstruction were carried out using the N-SIM module of the NIS-Element Advanced Research software. Image analysis was carried out using the Quantitation module of Volocity (Perkinelmer Inc, Waltham, MA) with x and y binning resolution of 32 nm.
3C library preparation

Limbs from ~70 E11.5 embryos, 3 E11.5 embryos with the limbs and heads removed, and the heads of 3 E11.5 embryos were collected in 15 ml tubes with enough PBS to cover them and to dissociate the cells by repeated pipetting with enlarged tip ends. Cells were fixed with 1% formaldehyde for 10 min at room temperature (r.t.). Crosslinking was stopped with 125 mM glycine, for 5 min at r.t. followed by 15 min on ice. Cells were centrifuged at 400 g for 10 min at 4°C, supernatants removed and cell pellets flash frozen on dry ice.

Cell pellets were treated as previously described (Dostie and Dekker 2007; Ferraiuolo et al., 2010; Williamson et al., 2014). HindIII-HF (NEB) was the restriction enzyme used to digest the crosslinked DNA.

5C primer and library design

5C primers covering the Usp22 (mm9, chr11: 60,917,307-61,003,268) and Shh regions (mm9, chr5: 28,317,087-30,005,000) were designed using ‘my5C.primer’ (Lajoie et al. 2009) and the following parameters: optimal primer length of 30 nt, optimal TM of 65°C, default primer quality parameters (mer:800, U-blast:3, S-blasr:50). Primers were not designed for large (>20 kb) and small (<100 bp) restriction fragments, for low complexity and repetitive sequences, or where there were sequence matches to >1 genomic target. The Usp22 region was used to assess the success of each 5C experiment but was not used for further data normalization or quantification.

The universal A-key (CCATCTCATCCCTGCGTGTCTCCGACTCAG-(5C-specific)) and the P1-key tails ((5C-specific)-ATCACCGACTGCCCATAGAGG) were added to the Forward and Reverse 5C primers, respectively. Reverse 5C primers were phosphorylated at their 5' ends. An alternating design consisting of 365 primers in the Shh region (182 Forward and 183 Reverse primers) was used. Primer sequences are listed in Table S6.
**5C library preparation**

5C libraries were prepared and amplified with the A-key and P1-key primers as previously described (Fraser et al. 2012). Briefly, 3C libraries were first titrated by PCR for quality control (single band, absence of primer dimers, etc.), and to verify that contacts were amplified at frequencies similar to those usually obtained from comparable libraries (same DNA amount from the same species and karyotype) (Dostie and Dekker 2007, Dostie, et al. 2007, Fraser, et al. 2010). We used 1 - 10 μg of 3C library per 5C ligation reaction.

5C primer stocks (20 μM) were diluted individually in water on ice, and mixed to a final concentration of 2 nM. Mixed diluted primers (1.7 μl) were combined with 1 μl of annealing buffer (10X NEBuffer 4, New England Biolabs Inc.) on ice in reaction tubes. 1.5 μg salmon testis DNA was added to each tube, followed by the 3C libraries and water to a final volume of 10 μl. Samples were denatured at 95°C for 5 min, and annealed at 55°C for 16 hours. Ligation with Taq DNA ligase (10 U) was performed at 55°C for one hour. One tenth (3 μl) of each ligation was then PCR-amplified individually with primers against the A-key and P1-key primer tails. We used 26 cycles based on dilution series showing linear PCR amplification within that cycle range. The products from 3 to 5 PCR reactions were pooled before purifying the DNA on MinElute columns (Qiagen).

5C libraries were quantified by bioanalyser (Agilent) and diluted to 26 pmol (for Ion PGM™ Sequencing 200 Kit v2.0). One microlitre of diluted 5C library was used for sequencing with an Ion PGM™ Sequencer. Samples were sequenced onto Ion 316™ Chips following the Ion PGM™ Sequencing 200 Kit v2.0 protocols as recommended by the manufacturer (Life Technologies™).

**5C data analysis**

Analysis of the 5C sequencing data was performed as previously described (Berlivet et al., 2013). Sequencing data was processed through a Torrent 5C data transformation pipeline on Galaxy (https://main.g2.bx.psu.edu/). Data was normalized by dividing the number of reads of each 5C contact by the total number of reads from the corresponding sequence run. All scales shown correspond to this ratio multiplied by 10³. For each experiment the number of
total reads, and of used reads, is provided in Table S7. The unprocessed heatmaps of the normalized 5C datasets can be found in Figure S4. 5C datasets are uploaded to the Gene Expression Omnibus (GEO) website (http://www.ncbi.nlm.nih.gov/geo/) under Accession number: GSE79947.

Results

Increased co-localisation of ZRS with Shh in the limb ZPA at E10.5 and E11.5

Previous analyses of the chromatin dynamics involved in the long-range regulation of Shh by ZRS has produced contradictory results, which could be due to the different temporal stages of development assayed (Amano et al., 2009; Lettice et al., 2014). To resolve this issue we carried out FISH on whole mouse embryo sections that include posterior and anterior forelimb tissue from E10.5, E11.5 and E14.5 developmental stages (Figure 1B). Shh is expressed within the ZPA at the two earlier stages but is switched off in the limb by E14.5 (Riddle et al., 1993). We compared inter-probe distances (Figure S1A shows representative images) and co-localisation frequencies (Figure 1C; left) between the gene and enhancer in mesenchymal tissue across the anterior-posterior axis of the distal forelimb bud. In addition proximal limb tissue and the adjacent flank where Shh is not expressed were compared.

By conventional wide-field deconvolution microscopy, the proportion of co-localised (<200 nm) Shh and ZRS probe pairs in ZPA cells at stages when Shh is expressed (E10.5, E11.5), was significantly higher (35%) than in inactive limb regions and the flank (distal anterior \( p < 0.05 \), proximal \( p < 0.01 \), flank \( p < 0.001 \) (Figure 1C, Table S2)). By E14.5 there is no Shh expression in the limb (expression ends between E11.5-E12.5), and the Shh-ZRS co-localisation frequency in distal posterior cells is significantly reduced, compared to E10.5 and E11.5 ZPA (E10.5 \( p = 0.006 \), E11.5 \( p = 0.01 \) (Figure 1D). At this later stage, differences in co-localisation frequencies between the distal posterior forelimb region (~20%) and the other limb regions and the flank mesoderm are also no longer detected (Figure 1C).

Dpp6 is located the same linear genomic distance away from Shh as the ZRS, but in the other direction and outside of the Shh regulatory domain (Figure 1A). In contrast to the spatial proximity of Shh-ZRS, Shh and Dpp6 are predominantly located > 400 nm apart (co-
localisation frequency < 5%) (Figure 1C; right) for all tissues and developmental stages examined.

The greater co-localisation of the active enhancer (ZRS) with its target gene (Shh) in the ZPA at E10.5 is similar to what we reported for these loci at E11.5 (Lettice et al. 2014) and to the co-localisation frequency of Hoxd13 and its GCR enhancer in distal posterior expressing limb tissue and cell lines at E10.5 (Williamson et al., 2012; Williamson et al., 2014). These differences in three-dimensional chromatin conformation between active and inactive tissues contradicts the previous report suggesting an equivalent rate of Shh-ZRS co-localisation on both sides of the distal limb field at this developmental stage (Amano et al., 2009).

Super-resolution imaging identifies Shh/ZRS co-localisation of most alleles in the ZPA

The data in Figure 1 are consistent with active gene-enhancer co-localisation during long-range regulation. From the images of tissue sections from the three developmental stages, acquired by conventional light-microscopy, it was apparent that Shh and ZRS are consistently very close in the nucleus, with differences in spatial distance frequently down to the signal centroids being in different layers of the z stack – the dimension with the lowest spatial resolution in the microscope. We therefore re-analysed the tissue sections containing E10.5 and E11.5 distal anterior and posterior (ZPA) cells by structured illumination microscopy (3D-SIM) (Figure 2A). This technique doubles the resolution limit in all dimensions (Toomre & Bewersdorf, 2010) and has previously been combined with 3D-FISH (Nora et al., 2012; Patel et al., 2013).

The greater resolution afforded by 3D-SIM, particularly for the z (depth) dimension (120 nm compared to 200 nm in conventional widefield microscopy), not only confirmed the difference in Shh/ZRS co-localisation frequency between ZPA and distal anterior limb bud but also suggests that conventional microscopy does not fully capture the proportion of co-localised Shh/ZRS probe pairs, especially in the Shh-expressing tissues where it now peaks at 79% (Figure 2B, Table S3). These data suggest that a substantial proportion of Shh/ZRS probe pairs with signal centroids not in the same plane of the z stack, that have been categorised as adjacent (between 200 nm and 400 nm apart (Figure S1)) due to the low z dimension resolution afforded by conventional widefield microscopy, are indeed co-localised
in ZPA cells. At both temporal stages the anterior/posterior differences in Shh/ZRS co-localisation frequency were highly significant (E10.5 \( p = 0.0002 \), E11.5 \( p = 0.0001 \)). Due to variation in fluorescent probe signal strength between alleles in the same nucleus, and the limited number of \( z \) stack planes imaged per nucleus by SIM to minimize fluorochrome bleaching, generally less than half of all probe pairs measured from each tissue in Figure 2 are from both alleles of the same nucleus. However, for cells where both alleles could be measured, ZRS/Shh co-localisation at both occurred in 33% (E10.5) and 59% (E11.5) of ZPA cells. The proportion of ZPA cells with only one co-localised allele was 56% (E10.5) and 31% (E11.5). Only around a tenth of ZPA cells at both temporal stages had no co-localising alleles, compared to a third of distal anterior cells. By comparing conventional and SIM data for the Shh/ZRS probe pair in anterior and posterior tissues at two developmental stages, we show that median inter-probe distances in distal anterior limb tissues are very similar when measured by either technique (conventional = 250 nm, SIM: E10.5 = 246 nm, E11.5 = 275 nm); whereas, these are significantly different for ZPA cells (E10.5: conventional = 221 nm, SIM = 160 nm, \( p = 0.01 \); E11.5: conventional = 241 nm, SIM = 136 nm, \( p < 0.0001 \)) (Figure 2C, Tables S4 & S5).

**The Shh/ZRS regulatory domain is compact in expressing and non-expressing tissue**

Long-range gene/enhancer co-localisation is often depicted as a looping out of the intervening chromatin fibre (Williamson et al., 2011). Our previous work on the HoxD locus implicated a gross compaction of the regulatory region, rather than a simple loop with extrusion of the intervening chromatin, upon activation of *Hoxd13* by the long-range (~250-kb) limb-specific GCR enhancer (Williamson et al., 2012; Williamson et al., 2014). We therefore used 3D FISH and conventional wide-field deconvolution microscopy to measure the spatial distances between either *Shh* or the ZRS, and the SBE4 enhancer that drives *Shh* expression in the forebrain (Figure 3A) (Jeong et al., 2006). SBE4 is located midway through the gene desert separating Shh and ZRS (Figures 1A). If the entire genomic region between the gene and the limb enhancer forms a loop then *Shh*-SBE4 and SBE4-ZRS distances should be greater than those between *Shh* and ZRS.

At both temporal stages (E10.5 and E11.5) when *Shh* is active in the distal posterior limb mesenchyme, but not at E14.5, *Shh* is closer to SBE4, and Shh/SBE4 co-localisation
frequencies are higher, compared to the other tissues analyzed (Figure 3B & C, S2A & B, and Table S5). These data suggest that the genomic region between Shh and the ZRS is folded into a compact chromatin domain, which is at its most compact in distal posterior Shh-expressing cells. However, what is also apparent is that the spatial distances between Shh and the ZRS are less than those between either Shh-SBE4 or SBE4-ZRS in most expressing and non-expressing tissues (Figure S2C). These differences are significant for most tissues analysed and intriguingly is particularly apparent at E14.5, well past the stage of limb-specific Shh activity and therefore could be indicative of a constitutive chromatin conformation.

**Topography of the Shh regulatory domain is maintained throughout the E11.5 embryo**

Using FISH we could only infer the conformation of the Shh regulatory domain from the spatial relationships of three genomic loci across the Shh-ZRS region. To gain a more complete view of the locus, we used 5C to determine the frequency of cross-linked interactions captured between sequences in the ~1.7 Mb region from Irsig1 ~400kb 3′ of Shh to Ube3c ~350kb beyond the ZRS (Figure 1A) in dissected whole fore- and hindlimb buds from ~70 E11.5 embryos (x2 replicates) (Figures 4A, left-hand heatmap, S3A and S4). We were unable to dissect cells suitable for 5C specifically from the ZPA. Three interaction domains can be identified; with the middle topologically associated domain (TAD) containing Shh and its entire known regulatory elements with the boundaries located 3′ of Rbm33 and within the 5′ end of Lmbr1. This Shh regulatory TAD corresponds well with that identified by Hi-C in mouse ESCs (Dixon et al., 2012).

In limb cells 5C cross-linked interactions are enriched between genomic fragments across the Shh and ZRS loci (Figures 4A, left-hand heatmap, S3A and S4). The general spatial proximity of Shh and the ZRS detected by FISH and inferred from enriched 5C interaction frequencies in expressing and non-expressing tissues suggests that this conformation is constitutive. To determine whether the high cross-linking efficiency of Shh and ZRS identified in E11.5 limb buds can also be detected in tissues where the ZRS is not active we carried out 5C on cells derived from the bodies and heads of E11.5 embryos. Even with the vast majority of these cells not expressing Shh, high read frequencies between Shh and ZRS were captured (Figures 4A, middle and right-hand heatmaps, and S4) and the same TAD structures could be discerned as seen in limb tissue.
To examine more closely the regions probed by FISH (Shh, SBE4 and ZRS) we generated “virtual 4C” plots from the 5C data (Figures 4B, S3B) (Williamson et al., 2014). From the viewpoint of Shh, overall interaction frequencies with the rest of its regulatory domain is similar in limb-, body- and head-derived tissues, and are not substantially higher than those extending into the adjacent TAD 3’ of Shh (Figures 4B compare the top track with the track that profiles SBE4 located in the middle of a TAD). Highest interaction frequencies for Shh, apart from genomic regions immediately adjacent, are with regions within the neighbourhood of ZRS (limb-specific high interactions with a loci within the gene desert that does not contain any known regulatory elements was not identified in the limb replicate data (Figure S3B)). ZRS has reciprocal enriched interactions with the Shh region (Figures 4B, bottom track). However, these are not detectably higher in limb than in the embryonic body or head.

Discussion

**Activation of Shh in the limb bud is accompanied by co-localisation with the ZRS**

Using 3D-FISH and super-resolution imaging, we provide compelling evidence that co-localisation (<200 nm) between Shh and the ZRS enhancer is associated with Shh expression in the ZPA region of the distal posterior forelimb bud, to an extent not seen in control tissues, including the limb bud after Shh expression has ceased at E14.5 (Figures 1 and 2). The co-localisation frequencies detected by super-resolution microscopy rise to almost 80% at E11.5 – suggesting that the vast majority of Shh alleles in the ZPA are juxtaposed to the ZRS located 1Mb of genomic distance away. Analysis of the FISH images by either conventional wide-field, or structured illumination microscopy, showed a significantly higher gene-enhancer co-localisation frequency in the ZPA than in nuclei from the distal anterior region of the same limb buds (Figures 1 and 2). This anterior-posterior difference in chromain folding is consistent with our previous analysis for Shh-ZRS in E11.5 fore- and hindlimbs (Lettice et al. 2014) and is similar to the preferential co-localisation of Hoxd13-GCR in E10.5 distal posterior limb buds (Williamson et al., 2012). Like Shh, Hoxd13 expression is restricted to the posterior margin of the distal limb bud at this stage. These data, however, contradict previously published work that could not identify significant differences in Shh-
ZRS proximity between the Shh-expressing ZPA and distal anterior cells (Amano et al., 2009). Those data were derived from single cell suspensions of dissected tissue from specific points across the distal limb bud whereas our data are from sections cut through whole embryos; therefore cell/tissue preparation may be a factor in discrepancies between the data sets.

**Shh and its regulatory elements are located within a compact chromatin domain**

Long-range interactions between genes and cis-regulatory elements are usually described as loops, which should be visualized as a coming together of the two loci to the exclusion of the intervening chromatin (Williamson et al. 2011; Fraser et al. 2015). Indeed a looping mechanism in distal limb could be inferred from the shorter inter-probe distances between Shh and ZRS, than for either of these probes with the forebrain SBE4 enhancer – even though the latter is located midway between Shh and ZRS on the linear chromosome (Figure S2C). To our knowledge this apparent Shh/ZRS chromatin ‘loop’ is the first to be identified by FISH.

However, the Shh-ZRS distances are shorter than distances to SBE4 not only in the ZPA but also in anterior limb and in E14.5 tissues when the ZRS is no longer active. But, Shh-ZRS co-localisation frequencies are not significant in those tissues. Another interpretation of these data is that the Shh regulatory domain (Figures 4A, S3A, S4) is maintained in a tightly folded chromatin conformation where Shh and the ZRS are generally proximal in nuclear space. That the Shh-containing TAD is indeed compact can be discerned from the frequency distribution graphs which show that most Shh/ZRS, Shh/SBE4 and SBE4/ZRS probe pairs are adjacent (200 – 400 nm) or co-localised (<200 nm), with median interprobe distances of between 220 – 345 nm for most tissues and developmental stages analysed (Figures 1B, S2A & B; Table S5). This is consistent with our 5C analysis of E11.5 limb bud, body and head cells which suggests that the Shh regulatory region forms a constitutive self-interacting domain – the Shh TAD has also been identified in ES cells by Hi-C (Dixon et al. 2012). These data show somewhat enriched interactions between cross-linked DNA fragments from the genomic regions containing Shh and ZRS (Figures 4, S3, and S4), but in all analysed tissues/cell types. The very high co-localisation frequencies that we see by microscopy between ZRS and Shh in the distal posterior limb at stages of Shh expression are not reflected in elevated interactions captured by 5C. Similarly, the increased compaction of
the intervening genomic region in ZPA cells inferred from FISH analysis of distances to the neural SBE4 enhancer could not be identified by 5C. We do not know whether this is because the Shh expressing (ZPA) cells do not present at a high enough proportion of cells in the dissected limb buds, or because the spatial proximities of Shh and ZRS, and the ZPA-specific chromatin domain is not well captured by chromosome conformation methods (Belmont, 2014). Conversely, our previous analysis comparing 5C and FISH has highlighted that spatial proximity should not always be inferred from enriched cross-linked interactions between 3C fragments (Williamson et al., 2014).

**Facilitating gene regulation by enhancer – promoter proximity**

Here, we have shown that local chromatin conformation maintains spatial proximity of Shh with the regulatory domain containing its enhancers – including the limb-specific enhancer ZRS – in a variety of cell types – not just those expressing Shh. If the physical interaction of active enhancers and their target gene promoters is essentially a stochastic process, their constitutive relative proximity within the same chromatin domain could be advantageous – for example by reducing the search space of the enhancer for the promoter (Williamson et al., 2011; Benabdallah and Bickmore, 2015). Consistent with this model, we have previously shown that, in the limb, expression levels of a reporter gene inserted into several positions across the whole Shh regulatory domain, is highest when the reporter inserts close to either the ZRS or Shh compared to insertion sites within the intervening gene desert (Anderson et al. 2014). These data suggest that ZRS-induced expression requires direct or indirect interactions with the target gene and these interactions are optimised by minimising the search space within a constrained chromatin domain. Whether the actual co-localisation of the ZRS with Shh in the ZPA is a cause or consequence of limb-specific Shh activation remains to be determined.
Acknowledgements

We thank the staff of the IGMM imaging facility and technical services for their assistance with imaging and sequencing. We thank the Dostie lab at McGill University for access to and the use of their Torrent 5C data transformation pipeline on the McGill University local galaxy server. The Super-resolution imaging experiments were conducted using the facilities provided by the Edinburgh Super-Resolution Imaging Consortium (ESRIC). This work was supported by the Medical Research Council, UK.

Author Contributions

IW and LAL conducted the experiments, and contributed to both the experimental design and writing of the paper. REH and WAB contributed to the design of the project and the writing of the paper.
References

Amano, T., Sagai, T., Tanabe, H., Mizushima, Y., Nakazawa, H., Shiroishi, T., 2009. Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. Developmental cell. 16(1), pp. 47-57.

Anderson, E., Devenney, P. S., Hill, R. E., Lettice, L. A., 2014. Mapping the Shh long-range regulatory domain. Development. 141, pp. 3934-3943.

Belmont, A. 2014. Large-scale chromatin organisation: the good, the surprising , and the still perplexing. Curr. Opin Cell Biology. 26, pp. 69-78.

Benabdallah, N. S., Bickmore, W. A., 2015. Regulatory domains and their mechanisms. Cold Spring Harb. Symp. Quant. Biol. pii: 027268.

Berlivet, S., Paquette, D., Dumouchel, A., Langlais, D., Dostie, J., Kmita, M., 2013. Clustering of Tissue-Specific Sub-TADs Accompanies the Regulation of HoxA Genes in Developing Limbs. PLoS Genetics, 9(12).

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 485, pp. 376-380.

Dostie, J., Dekker, J., 2007. Mapping networks of physical interactions between genomic elements using 5C technology. Nat. Protocols. 2, pp. 988-1002.

Dostie, J., Zhan, Y., Dekker, J., 2007. Chromosome conformation capture carbon copy technology. Curr. Protoc. Mol. Biol. Chapter 21: Unit 21 14.

Ferraiuolo, M. A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X. Q. D., Nadler, M., Blanchette, M., Dostie, J., 2010. The three-dimensional architecture of Hox cluster silencing. Nucl. Acids Res. 38, pp. 7472-7484.

Fraser, J., Ethier, S. D., Miura, H., Dostie, J., 2012. A torrent of data: mapping chromatin organization using 5C and high-throughput sequencing. Methods Enzymol. 513, pp. 113-141

Fraser, J., Rousseau, M., Blanchette, M., Dostie, J., 2010. Computing chromosome conformation. Methods Mol. Biol. 674, pp. 251-268

Fraser, J., Williamson, I., Bickmore, W. A., Dostie, J., 2015. An Overview of Genome Organization and How We Got There: from FISH to Hi-C. Microbiology and Molecular Biology Reviews. 79(3). pp. 347-372.

Ghavi-Helm, Y., Klein, F. A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., Furlong, E. M., 2014. Enhancer loops appear stable during development and are associated with paused polymerase. Nature. 512, pp. 96-100.
Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M., Epstein, D. J., 2006. A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development*. 133. pp. 761-772.

Klopocki, E., Ott, C.-E., Benatar, N., Ullmann, R., Mundlos, S., Lehmann, K., 2008. A microduplication of the long range SHH limb regulator (ZRS) is associated with triphalangeal thumb- polysyndactyly syndrome. *J. Med. Genet.* 45, pp. 370-375

Lajoie, B. R., van Berkum, N. L., Sanyal, A., Dekker, J., 2009. My5C: web tools for chromosome conformation capture studies. *Nat. Methods*. 6, pp. 690-691

Lettice, L. A., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E., de Graaf, E., 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human Molecular Genetics*. 12(14). pp. 1725-1735.

Lettice, L. A., Hill, A. E., Devenney, P. S., Hill, R. E., 2008. Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly. *Human Molecular Genetics*. 17(7). pp. 978-985.

Lettice, L. A., Horikoshib, T., Heaney, S. J., van Barenb, M. J., van der Lindee, H. C., Breedvelde, G. J., Joossee, M., Akarsuf, N., Oostra, B. A., Endod, N., Shibatag, M., Suzukih, M., Takahashih, E., Shinkai, T., Nakahorii, Y., Ayusawaj, D., Nakabayashik, K., Schererk, S. W., Heutinke, P., Hill, R. E., Nojic, S., 2002. Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *PNAS*, 99(11). pp. 7548-7553.

Lettice, L. A., Williamson, I., Devenney, P. S., Kilanowski, F., Dorin J., Hill, R. E. 2014. Development of five digits is controlled by a bipartite long-range cis-regulator. *Development*. 141(8). pp. 1715-25.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Duboule, D., 2011. A Regulatory Archipelago Controls Hox Genes Transcription in Digits. *Cell*. 147. pp. 1132-1145.

Morey, C., Da Silva, N. R., Perry, P., Bickmore, W. A., 2007. Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development*. 134, pp. 909-19.

Nora, E.P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barilbot, E., Blu, N., Dekker, J., Heard, E., 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485, pp. 381-385.

Patel, N.S., Rhinn, M., Semprich, C. I., Halley, P. A., Dolle, P., Bickmore, W. A., Storey, K. G., 2013. FGF signalling regulates chromatin organisation during neural differentiation via mechanisms that can be uncoupled from transcription. *PLoS genetics*, 9(7), p.e1003614.
Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., Shiroishi, T., 2005. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development*. 132(4), pp. 797-803.

Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, T., Noda, T., Shiroishi, T., 2004. Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mamm. Genome*. 15 (1), pp. 23-34.

Sun, M., Ma, F., Zeng, X., Liu, Q., Zhao, X-L., Wu, F-X., Wu, G-P., Zhang, Z-F., Gu, B., Zhao, Y-F., Tian, S-H., Lin, B., Kong, X-Y., Zhang, X-L., Yang, W., Lo, W. H-Y., Zhang, X., 2008. Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are caused by genomic duplications involving the long range, limb-specific SHH enhancer. *J. Med. Genet.* 45, pp. 589-595.

Toomre, D. & Bewersdorf, J., 2010. A new wave of cellular imaging. *Annual review of cell and developmental biology*. 26, pp. 285-314.

Wieczorek, D., Pawlik, B., Li, Y., Akarsu, N. A., Caliebe, A., May, K. J. W., Schweiger, B., Vargas, F. R., Balci, S., Gillessen-Kaesbach, G., Wollnik, B., 2010. A specific mutation in the distant sonic hedgehog (SHH) cis-regulator (ZRS) causes Werner mesomelic syndrome (WMS) while complete ZRS duplications underlie Haas type polysyndactyly and preaxial polydactyly (PPD) with or without triphalangeal thumb. *Hum. Mutat.* 31, pp. 81-89.

Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R. S., Paquette, D., Dostie, J., Bickmore, W. A., 2014. Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes & development*. 28, pp. 2778-2791.

Williamson, I., Eskeland, R., Lettice, L. A., Hill, A. E., Boyle, S., Grimes, G. R., Hill, R. E., Bickmore, W. A., 2012. Anterior-posterior differences in HoxD chromatin topology in limb development. *Development*. 139(17), pp. 3157-67.

Williamson, I., Hill, R.E. & Bickmore, W. a., 2011. Enhancers: From Developmental Genetics to the Genetics of Common Human Disease. *Developmental Cell*. 21(1), pp. 17-19. Available at: http://dx.doi.org/10.1016/j.devcel.2011.06.008.
**Figures**

**Figure 1. ZRS-Shh proximity in the ZPA at E10.5 and E11.5.**

(A) (Top) Location of genes over a 2 Mb murine genomic locus containing *Shh*, with the position of tissue-specific *Shh* enhancers shown below in green. The bottom two tracks show the locations to which the fosmid probes used for FISH hybridize (blue) and the 3C fragments amplified for 5C (black).

(B) Schematic indicating the position and plane of the tissue sections taken through the anterior and posterior parts of the E11.5 forelimb bud. Distal and proximal parts of the posterior limb bud and the distal anterior limb bud are shown, as is the flank mesoderm. Below are images of nuclei from E11.5 ZPA tissue sections showing *Shh/ZRS* and *Shh/Dpp6* probe pairs. Scale bars = 5μm.

(C) Frequency distributions of FISH inter-probe distances (d) in 200 nm bins, between *Shh* and ZRS (left column), or *Shh* and *Dpp6* probes (right column) in proximal and distal (anterior and posterior) regions of the murine forelimb bud and adjacent flank at E10.5, E11.5 and E14.5 (n = 70-130 alleles). For E10.5 and E11.5 sections distal posterior limb = ZPA. Error bars represent SEM obtained from two or three different
tissue sections from 1-2 embryos. The statistical significance between data sets was examined by Fisher’s Exact Tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (D) Comparison of the proportion of co-localised Shh/ZRS probe pairs (<200nm) across the three temporal developmental stages for distal anterior and posterior and proximal forelimb tissue and flank tissue. Error bars represent SEM obtained from two or three different tissue sections. The statistical significance between data sets was examined by Fisher’s Exact Tests.
Figure 2. Super-resolution imaging identifies the majority of Shh-ZRS probes as co-localised in ZPA tissue. (A) Nuclei captured by super-resolution SIM imaging from the distal forelimb of E10.5 and E11.5 embryos after FISH with Shh and ZRS probe pairs. Scale bars = 1 µm. (B) Frequency distributions of Shh-ZRS inter-probe distances (d) measured from SIM images in 200 nm bins, in distal anterior and distal posterior regions of the murine forelimb at E10.5 and E11.5. n = 67-100 (alleles). Error bars represent SEM obtained from two different tissue sections from 1 embryo. The statistical significance between data sets was examined by Fisher’s Exact Tests. (C) Boxplots show the distribution of Shh-ZRS inter-probe distances (d in nm) in E10.5 and E11.5 distal anterior and distal posterior captured by
conventional (con) and structured illumination (sim) microscopy. Line = median, box = interquartile range, whiskers = 95% range. The statistical significance between data sets was examined by Mann-Whitney U Tests.
Figure 3. The *Shh*-ZRS regulatory domain is maintained in a compact chromatin conformation in expressing and non-expressing tissue. (A) Images of representative nuclei from E11.5 ZPA tissue showing FISH signals for *Shh/SBE4, SBE4/ZRS* probe pairs. Scale bars = 5 μm. (B) Comparison of the proportion of co-localised *Shh/SBE4* and *SBE4/ZRS* probe pairs (<200nm) across the three temporal developmental stages for proximal and distal anterior and posterior (ZPA in E10.5 and E11.5 sections) forelimb tissue (n = 70 – 100 alleles). Error bars represent SEM obtained from two or three different tissue sections from 1-2 embryos. The statistical significance between data sets was examined by Fisher’s Exact
Tests. (C) Boxplots showing the distribution of interprobe distances (d) in nanometres between Shh/SBE4 and SBE4/ZRS in E10.5, E11.5 and E14.5 proximal (p), distal anterior (da) and posterior (dp) forelimb. The statistical significance between data sets was examined by Mann-Whitney U Tests.
Figure 4. 5C-seq identifies enriched interactions between Shh and ZRS in E11.5 embryos. (A) Heatmaps showing 5C data from cells of the limbs, bodies and heads 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 based on the total number of sequence reads in the 5C data set and the data shown is binned over 28-kb windows. Arrows indicate interaction frequencies between windows containing Shh and ZRS. Data for biological replicates are in Supplemental Figure S3A and unprocessed normalized data are shown in Supplemental Figure S4. (B) Virtual 4C analysis obtained by extracting 5C interactions with viewpoints fixed at Shh, SBE4 and ZRS. Dashed lines indicate the position of the fixed viewpoint from the Shh genomic region (orange) or regulatory elements (green). Data from limbs are in closed black circles, bodies in closed grey circles and heads in open circles. Genome coordinates on Chr5 are from the mm9 assembly of the mouse genome.
Supplemental Tables and Figures

### Table S1. Fosmid Probes

| Region | Name          | Ensemble name                  | Coordinates  | Size (bp) |
|--------|---------------|--------------------------------|--------------|-----------|
|        | Dpp6 WI1-2157A11 | G135P600264D6                | 27932527 27975636 | 43109     |
|        | Shh WI1-482L15 | G135P64333A4                  | 28754458 28795879 | 41421     |
|        | SBE4 WI1-469P2 | G135P600205H10                | 29107140 29147593 | 40453     |
|        | ZRS WI1-121N10 | G135P600929F6                 | 29611727 29653695 | 41968     |

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

### Table S2. Co-localisation frequency (<200 nm) of Shh and ZRS probes in distal and proximal limb and adjacent flank tissue of E10.5, E11.5 and E14.5 embryos at normal resolution (x & y bins = 67 nm, z steps = 200 nm)

| Tissue           | E10.5 | E11.5 | E14.5 |
|------------------|-------|-------|-------|
| Distal posterior | 35    | 35    | 20    |
| Distal anterior  | 21 \((p = 0.02)\) | 18 \((p = 0.01)\) | 20    |
| Proximal         | 17 \((p = 0.003)\) | 17 \((p = 0.005)\) | 22    |
| Flank            | 13 \((p = 0.0002)\) | 10 \((p = 0.0002)\) | 22    |

Statistical analysis of data for Fig. 1C. \(p\)-values from Fisher’s Exact Tests.

### Table S3. Co-localisation frequency (<200 nm) of Shh and ZRS probes in distal anterior and posterior tissue of E10.5 and E11.5 embryos at super resolution (x & y bins = 32 nm, z steps = 120 nm)

| Tissue           | E10.5 | E11.5  |
|------------------|-------|--------|
| Distal posterior | 63    | 79     |
| Distal anterior  | 31 \((p = 0.0002)\) | 34 \((p < 0.0001)\) |

Statistical analysis of data for Fig. 2B. \(p\)-values from Fisher’s Exact Tests.

### Table S4. Median interprobe distances for Shh and ZRS probes in distal anterior and posterior tissue of E10.5 and E11.5 embryos at super resolution

| Tissue           | E10.5  | E11.5  |
|------------------|--------|--------|
| Distal posterior | 160    | 136    |
| Distal anterior  | 246 \((p < 0.0001)\) | 275 \((p < 0.0001)\) |

Statistical analysis of data for Fig. 2C. \(p\)-values from Mann-Whitney U Tests.
Table S5. Median interprobe distances for Shh-Dpp6, Shh-SBE4, SBE4-ZRS and Shh-ZRS probes in distal and proximal limb and adjacent flank tissue of E10.5, E11.5 and E14.5 embryos

| Tissue            | Shh-Dpp6  | Shh-SBE4 | SBE4-ZRS | Shh-ZRS |
|-------------------|-----------|----------|----------|---------|
| Distal posterior  | 593 (p = 0.005) | 276      | 250      | 221     |
| Distal anterior   | 493 (p = 0.03)  | 300      | 291      | 250 (p = 0.01) |
| Proximal          | 611       | 314 (p = 0.02) | 280     | 263 (p = 0.0001) |
| Flank             | 571       | 341 (p = 0.03) | 250     | 291 (p < 0.0001) |

E11.5

| Tissue            | Shh-Dpp6  | Shh-SBE4 | SBE4-ZRS | Shh-ZRS |
|-------------------|-----------|----------|----------|---------|
| Distal posterior  | 523       | 272      | 300      | 241     |
| Distal anterior   | 480       | 291 (p = 0.03) | 341     | 250 (p = 0.05) |
| Proximal          | 471       | 291 (p = 0.03) | 385 (p = 0.01) | 250 (p = 0.0004) |
| Flank             | 512       | 341 (p = 0.02) | 345     | 406 (p < 0.0001) |

E14.5

| Tissue            | Shh-Dpp6  | Shh-SBE4 | SBE4-ZRS | Shh-ZRS |
|-------------------|-----------|----------|----------|---------|
| Distal posterior  | 540       | 291      | 341      | 250     |
| Distal anterior   | 479       | 324      | 334      | 241     |
| Proximal          | 549       | 300      | 314      | 250     |
| Flank             | 474       | 295      | 287      | 241     |

Statistical analysis of data for Fig. S1B. Interprobe distances are median values, p-values from Mann-Whitney U Tests.

Table S6. Mouse 5C primers for Shh and USP22 regions

| Fragment | Type | Genomic sequence (5’ to 3’) | HindIII position Start | HindIII position End |
|----------|------|-----------------------------|------------------------|----------------------|
| Shh      |      |                             |                        |                      |
| (chr.5; mm9) | |                             |                        |                      |
| 3 R      | CTTCCTACATGGTTTACAGTTAATGGAGT | 28317087               | 28319149              |
| 5 F      | AGGATAGGATTTGGTGTAGTGGTTAGCTCAAGTG | 28319935               | 28324734              |
| 7 R      | CTTCTTCATGCCCTACACTAACCAGGCCT | 28325862               | 28331585              |
| 8 F      | GGTGAGAATGCCAAAGAGACCTGTTAGTAG | 28331586               | 28333893              |
| 9 R      | CTGGAGACACTACCTACCTCTAGCATCAAAT | 28333894               | 28336649              |
| 10 F     | AACACACTGCCGATGATGACATTTTAGCAAG | 28336650               | 28340960              |
| 11 R     | CTTGAGGTGATGCGCTACTGTTGGGG | 28340961               | 28344675              |
| 12 F     | CAAACCTCAGAAACACAGGAGGACCAAG | 28344676               | 28347365              |
| 13 R     | CGACTCAATTTGTAGATGACCTCAGCAAG | 28347366               | 28348911              |
| 18 F     | GGAGACCCACACTAAGCCGCTCAAG | 28355883               | 28370234              |
| 19 R     | CTGGAGATGCTGGGTCCTGTAGTGCTGCTAGTAC | 28370235               | 28373716              |
| 20 F     | TTTGGTGTAGGGATGAGGGTGGATCTTTAAG | 28373717               | 28374948              |
| 25 R     | CTGCCTTCTGTAGATCCTATTGAGCATTTCCCT | 28385728               | 28388130              |
| 27 F     | TTGGTATAGTGTGCTGTTCTGGTGCTAGAAG | 28388373               | 28403099              |
| 30 R     | CCTCTCCGATAGTGGGAACTTTTTATTTATT | 28416843               | 28419608              |
| 31 F     | TCTCTAAATATACACAGGAAGAGGCTAAG | 28419609               | 28428926              |
| 32 R     | CTTGTCATCCACTAGTTGCTGCTGTAAG | 28428927               | 28447705              |
| 34 F     | CAATAAAGGTAGAAGTTGGGTCAGTAGAAG | 28452395               | 28454456              |
35 R CTTGAGTCATATGGGACACTCTTGCACA 28454457 28458325
38 F ATGGGCCCCGATTTAACTCAACAATCAAAG 28463308 28463573
41 R CTTCTCTAGCTAGCCAGCTAAATGTACCG 28469222 28469868
42 F AAAATCTCCCTGGAAGGTACAGTCAAG 28469869 28470959
44 R CTTTGCTCAAAATGTAGGAAATGGCCATTC 28476584 28484052
45 F AAAATCTCCCTGGAAGGTACAGTCAAG 28484053 28486216
46 R CTTGTGACACGTCACAGTCTCAGT 28486217 28491939
47 F CACAGGGCTCTTTCATAGCCTAAGAACAAG 28491940 28511814
49 R CTTATCCTCTTCTGTGTCTAGTTGAAGTGG 28511842 28514356
50 F CTGCAATATGACTCTGGTTTCTTGGCCCA 28514770 28515717
52 R CTTCCAGAAGATCTGCAGCAACTCTCTCTC 28541462 28542832
53 F GAGTTCAAGAGCCCCAAAATCCCTCTAAG 28570465 28576299
55 R CTTTAGCATGGACTCAGAAAACAAAATAGG 28566954 28566955
57 F GAGTTCAAGAGCCCCAAAATCCCTCTAAG 28570465 28576299
59 F CACATGACAGAATACATGATAAGAAATAG 28571001 28572084
61 R CTTTGCAAAATGAGTGTGACTCCCATTCTTCT 28577016 28580658
63 R CTTGGGGTGACTTGGCCTCCTATTCTTCT 28580659 28581218
65 R CTTGAAATTGAAGTATCTCTCAGCACCT 28581219 28585322
67 F CTTGACAGAGGAGCCTAAAAGGTGACTTAA 28585390 28593411
69 R CTTGCTTACAGTCTGCTAGCTAAGAACAAG 28593412 28595934
71 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28593412 28595934
73 F CTTCCAGAAGATCTGCAGCAACTCTCTCTC 28593412 28595934
75 F GGGAAAGGTACTCTGGGGTGCATCACAAAG 28604862 28605362
77 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
79 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
81 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
83 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
85 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
87 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
89 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
91 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
93 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
95 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
97 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
99 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362
101 F CTTGAGCTTCTTCTGGAAGGTACAGTCAAG 28604862 28605362

| Position | Letter | Sequence                              | Start   | End     |
|----------|--------|---------------------------------------|---------|---------|
| 102      | R      | CTTCCTATGGCTGAGAACTGCTTAGATAAT         | 2865142 | 2865243 |
| 103      | F      | GCTGATCTCTTGCTGACTGGAGTGTGAAG         | 2865244 | 2865608 |
| 106      | R      | CTTTGGTGTGTTTACACCATTTTTCCTTTA        | 2866207 | 2866894 |
| 107      | F      | GCTCTCTTAATCCCTGGTGAGAATGCTTAGTAAT    | 2866895 | 2866986 |
| 109      | R      | CTTCCTCTAGGAAACACAGCCTCTCCTGTAAG      | 2867654 | 2868126 |
| 111      | F      | CAGCATGGCTGTGAGGGAAAGTAGCTAAG         | 2868158 | 2868714 |
| 112      | R      | CTTTACAACAGTACATTATCCTAAGCGTCTA       | 2868296 | 2868409 |
| 114      | F      | CAATTTCAGTGCCAGCCTCTCGGGAAG           | 2868514 | 2868942 |
| 115      | R      | CTTCCTCTCTCAGGAAACCAGTCTTCTGAG        | 2868715 | 2868840 |
| 117      | F      | CTTCAACAGGCCCATTCTCTGAGAAACAT         | 2870577 | 2871008 |
| 118      | R      | CTATTCCAGGCCATTCTCTGAGGAAACAT         | 2870584 | 2871008 |
| 119      | F      | AAAAGAAAAAGAATCACCATGACTCTTAAG        | 2871008 | 2871086 |
| 120      | R      | CTTATACCACTGGGTGGAGGTCAATTCTGAGC      | 2871086 | 2871246 |
| 121      | F      | CATGGTAACATCGTGTGTAGATAGAAAAAG        | 2871087 | 2871159 |
| 123      | R      | CTTTATTGCCAGGCCATTCTCTCAAGGAAACAT     | 2871246 | 2871318 |
| 124      | F      | GTTCGAGACTGAGGGGCTCCAGAAG             | 2871340 | 2871544 |
| 125      | R      | CTTTAATGCTCCCTGGCTCTCTGAGAAACAT      | 2871545 | 2871887 |
| 126      | F      | CTTAACTGGCTGACTCTCTCTCTGCTTTAAG      | 2871887 | 2871974 |
| 129      | R      | CTTTTGAAGGAGACCCTTTTTCCTATTTAAG       | 2872350 | 2872496 |
| 130      | F      | CTTAATGTGAGCGGCGCTCACTCTAAGGAAAG     | 2872494 | 2872866 |
| 132      | F      | ATTTCTTGAGCATTAGACCGAGTAAG            | 2873818 | 2874346 |
| 133      | R      | CTATAATTGCTGTTGTTGTTGTTCTTCAA        | 2874346 | 2874805 |
| 134      | F      | CTTTCACTGGAAGGGCTCGGAGGAAGAAG       | 2874804 | 2875056 |
| 139      | R      | CTTTATTGCCAGGCCATTCTCTTACAA          | 2875910 | 2876397 |
| 140      | F      | ATATTGGAGATTGTGCAGGTGCTGAGAAG       | 2876397 | 2876708 |
| 143      | R      | CTTGTTCCGCTTACCACTAAAAGGCCC          | 2877067 | 2878056 |
| 144      | F      | CCAGAGACCCCTTGCCATCTCTGCAAGG         | 2878054 | 2878263 |
| 145      | R      | CTTTCCCTCACCACCTGAAAGAAGGAAAG       | 2878263 | 2878952 |
| 147      | F      | CATCTGATTGGCGAAGGGCGCAGAG             | 2878954 | 2879370 |
| 149      | R      | CTTAATACCGCTCTGTCTGCTCTAA            | 2879405 | 2879760 |
| 151      | F      | CTGGACTACCTGAGAACCCACTAGGTTTGTGGAG   | 2879854 | 2879928 |
| 152      | R      | TACGAAAGATGCTGGGAGAGCTTCTTCAAAG      | 2879929 | 2880154 |
| 154      | F      | CTTTGGACAGTTTCCCTCTCTCTGCTCTTCC      | 2880509 | 2880580 |
| 155      | R      | GTGGAGCCATCTGGAATGGTATGGGAGAAG       | 2880508 | 2880684 |
| 158      | F      | CTTCTCCTAGGAAACACTAGCCTAGCTATGCAAG   | 2882167 | 2882766 |
| 162      | F      | AGAACACAGGATAACCATAGCGCAGCAAG        | 2883190 | 2884328 |
| 163      | R      | CTTAAGACGTATGCTCTTTGAGATGTCTG       | 2884329 | 2884760 |
| 168      | F      | ATACTCACCACCTCTTCTAAATGGGAAAG       | 2887068 | 2887651 |
| 169      | R      | CTATATTGCCAGGTCAATGTGTTATTTAACAG    | 2887652 | 2887813 |
| 170      | F      | CTTGATCTGAGAGGTGTAAGCTGAGATAAG      | 2887813 | 2887874 |
| 171      | R      | CTTTGAGGAGACCCCTTTCTTCTCTGCTCTG    | 2887875 | 2887928 |
| 173      | F      | GTGTTGGAGCAGCTGGTAGCTAGCAAGAG       | 2887930 | 2888040 |
| 174      | R      | CTTAATACCTCTGACAGCGGCGACAAG         | 2888041 | 2888083 |
| 176      | F      | GGCTGCAAAGATGGGTCTCTATTTGTGAGA      | 2888235 | 2888382 |
| 177      | R      | CTTTGGAGGCTGGTGTGTCAGC              | 2888382 | 2888480 |
252 R CTAGGCCCATGAAGGAAGATGGCTTTGACA 29131878 29136802
254 F GTTTTTTTCTGGCAACAGCTACACCTAAG 29137043 29150444
255 R CTTCCAATCTCCTCCTGGCTCAAATGAAA 29154045 29155707
256 F TTTATCTAACACTTATCCCCATCTGGCAA 29155708 29159386
258 R CTGGGAGCATTAAAAATATGTTCTCAGAT 29159409 29160128
259 F CTAATTTATCTGAAGTACATGGCTCAAGAT 29160129 29166826
260 R CTGCCATGTAAATGATGGGATATCTGGCAA 29166827 29168466
261 F GCACCGAGACCTGTTCAAGCTACCTAAG 29168467 29173627
262 R CTTATTTATCAAGTACAGTTGCTCAGAT 29173628 29173865
263 F CATAAAACGTGTAACTTTTATTTAATAG 29173866 29174415
264 R CTGCGAGCCTCAAGCTTATCCTTGCAAG 29174416 29179395
265 F CTAGGCACGCTTTAATGAGCTTCAGTGAAG 29179396 29180591
266 R CTTCCTCTCCTCCGCCTGCTAATC 29180592 29184579
267 F AGTTTTTTCTTTTTATGTGTCAGCTCAGAAG 29184580 29185047
268 R CTGGATATTACAGTGATGAATTGATATGT 29185048 29185852
269 F TGGACACAGTATTCTTCTTTATTTGGAAG 29185853 29188276
271 R CTAGACCTTTTTAGTTAAGCTTCTGGTT 29188338 29191726
272 F TTGAAAGCTGATTTCAAACAATGATTAAAG 29191727 29196352
274 R CTATCAACAACCTGCACTTATTTAAGAAC 29198345 29199536
275 F AGCTATCATTTGGTTAAAAACTGTTAGAAG 29199537 29203806
276 R CTTCATGCTGGCAGACAAAGTAAATTCGGA 29203807 29203920
278 F GGCCCCATGCTGGCTCCCAGATAAG 29206157 29209241
279 R CTTCCCTGTAAATATCTGGAAATAGAAG 29209242 29216466
280 F CATGGGAGGTCAACAGGATTGGTGAAAG 29216469 29220327
281 R CTGGGAGCTTTCCAGTTGGAAGATGAAG 29220328 29223230
283 F GGCAGCATGCTGGGACCAAGGAGCATCAAAG 29230618 29233018
284 R CTTCACTACAGTCTGTGTTAATCAAGATC 29233019 29236380
285 F CAAAGGTCTATCTATGTGATGCTCAGAAG 29236381 29237108
287 R CTGGCTAACACCAGTTGAGGTGGAGATGC 29237129 29249342
288 F TGTCATCTCTATCTTCTGATTCTCTCAGAT 29240128 29243626
289 R CTTCCTCTGTACCTATCTGCTCAGGCTTTAAG 29243267 29246728
290 F GCATGTTAACACGGGGGTGTTAAAGCTAAG 29246729 29247568
291 R CTGCGTCACGTGCACTTCCCCCTGTATTATA 29247569 29250429
292 F GGGAAGACAAATCAGTATTGCCAGCTTTAAG 29250430 29254037
293 R CTGGGAAGATTTCCAAGAGGAGATCCAAG 29253038 29259186
296 F CTTTGTTTTCTGGAAGAGGGGGCTATAGAAG 29259187 29262463
297 R CTGGAAAAGTGAAAGATATGATATGCAAT 29262464 29264721
298 F GGGAGCAGCCAGTACCTGACCCAGAAG 29264722 29268159
299 R CTGGCTTTGCCCATTGGACCTTTGTCAG 29268160 29268267
300 F TGATGCTATCTCCTTCAAGGAGGAAGAAG 29268268 29270862
301 R CTTAAAAAAGCAATATGATATGACATC 29270863 29272035
303 F TGATGAAAGAAATGATAAGGAGTTCAAGAAG 29277402 29281527
304 R CTTCTCTTTTTCTTTTATGAAAGATCAGAAGC 29281528 29287941
309 F GATGGACACCTGGAAGAGGGACACCAAAG 29292444 29296257
313 R CTAAATTTCCTGCTGTAAAATTTGTATG 29305044 29309647
314 F CCACCGTGATGACAGCTTTGGAAG 29309648 29310307
| Start | End | Sequence | Start | End |
|-------|-----|----------|-------|-----|
| 316   | 319 | ACCTTTCCTTCCTAAGTCCATGCAAAG | 29310490 | 29311402 |
| 320   | 324 | CTTGACCTGCTCATCTGGGAAAG | 29311402 | 29315522 |
| 325   | 329 | GCCCCTTGACCCACAGTCTGGGAAAG | 29317749 | 29319160 |
| 330   | 334 | CTGAGTGCCATCTGGGAAAG | 29319161 | 29319272 |
| 335   | 339 | TAAATCCATCTGGGAAAG | 29319273 | 29320195 |
| 340   | 344 | CCCATCTCTGGGAAAG | 29321894 | 29331157 |
| 345   | 349 | TAAATCTAATAAGATGAAGGAAAATAAC | 29331158 | 29332122 |
| 350   | 354 | CAGGCTCCTGGGAAAG | 29331158 | 29332122 |
| 355   | 359 | CTTGGGGAAGTGGGAAAG | 29332123 | 29347103 |
| 360   | 364 | CAAGAAGAAGGCTACAAGAGGAGAGGCAAG | 29347104 | 29352141 |
| 365   | 369 | CCATCTCTGCCTTGAAGGAAATGGAAGGAAAG | 29356572 | 29358986 |
| 370   | 374 | CCCCATGCTCTGGGAAAG | 29358987 | 29363174 |
| 375   | 379 | TTCGTCCTTGGGGAAATGGAAGGAAAG | 29364016 | 29407971 |
| 380   | 384 | CAAGAAGAAGGCTACAAGAGGAGAGGCAAG | 29407972 | 29411033 |
| 385   | 389 | CTTGCCCCCACAGGGCAGGC | 29411528 | 29412802 |
| 390   | 394 | CTTGGCTGTCCTTTGCCACCAAACAAAG | 29411528 | 29412802 |
| 395   | 399 | CTTGGCTGTCCTTTGCCACCAAACAAAG | 29411528 | 29412802 |
| 400   | 404 | CTTGGCTGTCCTTTGCCACCAAACAAAG | 29411528 | 29412802 |
|    |    | DNA Sequence            | Start | End   |
|----|----|-------------------------|-------|-------|
| 478 | R  | CTTGAGTTTTCATATCAACAGGCTCAGT | 29769327 | 29770690 |
| 479 | F  | GGTGAGCTAGTCAGAGCAGTGCTGAAAAG | 29770691 | 2977235 |
| 481 | R  | CTTGCTTTGAGCCCGGGTGGCT | 29774284 | 29777533 |
| 482 | F  | TGCCATTTTATATCTACTTGGAGAAAAAG | 29777534 | 29781541 |
| 483 | R  | CTTTTCTCAAGTCAGGTGTTAGTAAAAGCAGA | 29781542 | 29782720 |
| 484 | F  | GACCCCAAGGACCCAACTTCAAG | 29787271 | 29784642 |
| 487 | F  | TTCTGAGCTCCTTCCCCCTCAGAGTAAG | 29786563 | 29787575 |
| 488 | R  | CTTGATTTCTCAAACTACTTATTGATTCGT | 29797272 | 29805362 |
| 489 | F  | GAGTTGGAGAGTTTGGAGGCTGAACACAAG | 29805363 | 29806188 |
| 491 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29807450 | 29809879 |
| 492 | F  | CTTACTAAAAAGGGAGGAAGGAAATCCAAG | 29807451 | 29815593 |
| 494 | R  | CTTCCCTGCTGTAGGGGAGAGCG | 29815594 | 29816516 |
| 495 | F  | GGTGCAGGGAAAGTTGATAAGGGCAAAG | 29816517 | 29820938 |
| 497 | R  | CTTGAGTTTTCATATCAACAGGCTCAGT | 29835164 | 29835981 |
| 498 | R  | CTTGCTTTGGAGAGTTTGGAGGCTGAACACAAG | 29844732 | 29845798 |
| 500 | F  | ATTTGCCACTCAAAATCTGCACTTTCCAAG | 29845799 | 29846398 |
| 501 | R  | CTTGAATGAGGAAACATAGGCTGAGAGGCC | 29846399 | 29848887 |
| 502 | F  | GTACTGAGCCTAGCAGAGGAAGCTCAAG | 29848888 | 29849213 |
| 503 | R  | CTTCTAAGGGAGGAAATCCAAG | 29849214 | 29850560 |
| 504 | F  | CTTGCAAAAGGGAGGCAGCTTTTTAATTCT | 29849215 | 29851352 |
| 505 | F  | TTCAGCTTTTGTCTTGTTGGTGACTAAG | 29849216 | 29851353 |
| 506 | F  | CTTGAATGAGGAAACATAGGCTGAGAGGCC | 29851354 | 29851355 |
| 507 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851356 | 29851357 |
| 508 | R  | CTTTACTGCTTTGCTCTTCAAG | 29851358 | 29851359 |
| 509 | F  | AGTACCTAGTCACACCTACCTATTTTATAATTTCT | 29851360 | 29851361 |
| 510 | R  | CTTGCGGCAGGAATGCTCTTCAAG | 29851362 | 29851363 |
| 511 | F  | CTTGCGGCAGGAATGCTCTTCAAG | 29851364 | 29851365 |
| 512 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851366 | 29851367 |
| 513 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851368 | 29851369 |
| 514 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851370 | 29851371 |
| 515 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851372 | 29851373 |
| 516 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851374 | 29851375 |
| 517 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851376 | 29851377 |
| 518 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851378 | 29851379 |
| 519 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851380 | 29851381 |
| 520 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851382 | 29851383 |
| 521 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851384 | 29851385 |
| 522 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851386 | 29851387 |
| 523 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851388 | 29851389 |
| 524 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851390 | 29851391 |
| 525 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851392 | 29851393 |
| 526 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851394 | 29851395 |
| 527 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851396 | 29851397 |
| 528 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851398 | 29851399 |
| 529 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851400 | 29851401 |
| 530 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851402 | 29851403 |
| 531 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851404 | 29851405 |
| 532 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851406 | 29851407 |
| 533 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851408 | 29851409 |
| 534 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851410 | 29851411 |
| 535 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851412 | 29851413 |
| 536 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851414 | 29851415 |
| 537 | F  | CTTATACCTTTATGACCCATAATGCACAGA | 29851416 | 29851417 |
| 538 | R  | CTTATACCTTTATGACCCATAATGCACAGA | 29851418 | 29851419 |
Table S7. 5C sequencing reads

| Sample            | Number of reads | Number of used reads |
|-------------------|-----------------|----------------------|
| Limb              | 1438064         | 677006               |
| Limb-replicate    | 2483452         | 1715790              |
| Body              | 1661514         | 1057871              |
| Head              | 1815022         | 1050062              |
| Head-technical replicate | 1771610   | 932274               |
Figure S1. Images of representative nuclei showing probe pairs at various distances apart. Shh/ZRS probes up to 400 nm, Shh/Dpp6 probes shown for distances greater than 400 nm. All nuclei are from E10.5 ZPA.
Figure S2. Spatiotemporal frequency distribution of Shh/SBE4 and SBE4/ZRS interprobe distances, and spatiotemporal interprobe distances of Shh/Dpp6, Shh/SBE4, SBE4/ZRS and Shh/ZRS. (A) Frequency distributions of interprobe distances (d) in 200 nm bins between Shh and SBE4 probes, and SBE4 and ZRS probes, in proximal and distal regions of the murine forelimb bud and adjacent flank at E10.5, E11.5 and E14.5 temporal stages. F: flank, P: proximal limb, DA: distal anterior limb, DP: distal posterior limb (ZPA in E10.5 and E11.5 sections). n = 70 – 100. Error bars represent SEM obtained from two or three different tissue sections. The statistical significance between data sets was examined by Fisher’s Exact Tests: * p < 0.05. (B) Boxplots show the distribution of interprobe distances (d) in micrometres between the four sets of probe pairs in E10.5, E11.5 and E14.5 distal anterior and posterior and proximal forelimb, and flank tissue. The statistical significance between data sets was examined by Mann-Whitney U Tests. (C) Boxplots comparing the distribution of interprobe distances (d) in nanometres of the three sets of probe pairs located across the Shh regulatory region in E10.5, E11.5 and E14.5 distal anterior and posterior and proximal forelimb, and flank tissue. S/SB: Shh/SBE4, SB/Z: SBE4/ZRS, S/Z: Shh/ZRS. The statistical significance between data sets was examined by Mann-Whitney U Tests.
Figure S3. 5C-seq enriched interactions between Shh and ZRS are recapitulated in a biological replicate. (A) 5C heatmap shows the average interaction frequencies (28-kb bins) across Shh and its regulatory domain in E11.5 limb bud cells (replicate biological sample). Arrows indicate interaction frequencies between windows containing Shh and ZRS. Interaction frequencies are colour-coded according to the corresponding scales as described in Figure 4A. (B) Virtual 4C analysis obtained by extracting 5C interactions with viewpoints fixed at Shh, SBE4, and ZRS. Dashed lines indicate the position of the fixed viewpoint from the Shh genomic region (orange) or regulatory/structural elements (green). Data from limb bud and limb bud replicate cells are in open and filled circles, respectively.
Figure S4. Unprocessed heatmaps of the normalised E11.5 limb bud, body and head cells 5C datasets. 5C-seq normalised data are presented in the heatmap form according to colour scales as described in Figure 4A. Genes are indicated in black, regulatory elements in green and fosmid probes in blue. Grey shading highlight the position of the genes in the 5C heatmaps. Blue arrows indicate enriched interactions between genomic fragments over the Shh locus and ZRS locus.