Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press



The genomic landscape of cohesin-associated chromatin interactions

Laura E. DeMare, Jing Leng, Justin Cotney, et al.

Genome Res. published online May 23, 2013
Access the most recent version at doi:10.1101/gr.156570.113

P<P Published online May 23, 2013 in advance of the print journal.

Open Access Freely available online through the *Genome Research* Open Access option.

Creative Commons License This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at http://creativecommons.org/licenses/by-nc/3.0/.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**



To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

© 2013, Published by Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

Research

The genomic landscape of cohesin-associated chromatin interactions

Laura E. DeMare,[1] Jing Leng,[2] Justin Cotney,[1] Steven K. Reilly,[1] Jun Yin,[1] Richard Sarro,[1] and James P. Noonan[1,2,3,4]

[1]*Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06510, USA;* [2]*Program in Computational Biology and Bioinformatics, Yale University, New Haven, Connecticut 06520, USA;* [3]*Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA*

Cohesin is implicated in establishing tissue-specific DNA loops that target enhancers to promoters, and also localizes to sites bound by the insulator protein CTCF, which blocks enhancer-promoter communication. However, cohesin-associated interactions have not been characterized on a genome-wide scale. Here we performed chromatin interaction analysis with paired-end tag sequencing (ChIA-PET) of the cohesin subunit SMC1A in developing mouse limb. We identified 2264 SMC1A interactions, of which 1491 (65%) involved sites co-occupied by CTCF. SMC1A participates in tissue-specific enhancer-promoter interactions and interactions that demarcate regions of correlated regulatory output. In contrast to previous studies, we also identified interactions between promoters and distal sites that are maintained in multiple tissues but are poised in embryonic stem cells and resolve to tissue-specific activated or repressed chromatin states in the mouse embryo. Our results reveal the diversity of cohesin-associated interactions in the genome and highlight their role in establishing the regulatory architecture of development.

[Supplemental material is available for this article.]

Mammalian development requires the precise spatial and temporal control of gene expression. Much of this regulatory information is encoded in thousands of *cis*-acting elements that are distributed across the genome, often at great distances from their target genes (Visel et al. 2007, 2009; Blow et al. 2010; Cotney et al. 2012). The mechanisms that connect *cis*-regulatory elements to their specific targets and that prevent them from inappropriately influencing other genes are not well defined. Recent studies suggest cohesin stabilizes DNA loops between distant-acting enhancers and their target promoters (Kagey et al. 2010). Cohesin is a ring-shaped complex consisting of the core subunits SMC1A, SMC3, SCC1 (also known as RAD21), and SA1/SA2 (Nasmyth and Haering 2009). Although cohesin does not bind DNA directly, it colocalizes with tissue-specific transcription factors on chromatin (Schmidt et al. 2010; Nitzsche et al. 2011) and is thought to stabilize binding of transcription factors at enhancers (Faure et al. 2012). In embryonic stem cells, cohesin shows cell-type-specific binding at enhancers and promoters that engage in cell-type-specific interactions, and knockdown of cohesin results in aberrant gene expression and loss of pluripotency (Kagey et al. 2010). These studies suggest tissue-specific gene activation is the result of tissue-specific cohesin-mediated DNA looping events.

Additionally, cohesin has been shown to associate with the insulator factor CTCF (Rubio et al. 2008), which targets cohesin to specific sites in the genome (Parelho et al. 2008; Stedman et al. 2008; Wendt et al. 2008). Cohesin is required for the enhancer-blocking functions of CTCF binding (Parelho et al. 2008; Wendt et al. 2008). CTCF also establishes chromatin barriers to prevent the spread of heterochromatin (Cuddapah et al. 2009; Kim et al. 2011). In embryonic stem (ES) cells, chromatin loops mediated by CTCF interactions show correlated patterns of active or repressed histone modifications contained or excluded by the loop (Handoko et al. 2011). CTCF shows largely invariant binding patterns across tissues (Kim et al. 2007; Jothi et al. 2008) and, in conjunction with cohesin, may establish constitutive chromatin topologies in the nucleus (Dixon et al. 2012; Nora et al. 2012).

Despite these findings, global insight into the role of cohesin in gene regulation remains limited because cohesin-mediated interactions have yet to be mapped at a genome-wide scale. Here, we use chromatin interaction analysis with paired-end tag sequencing (ChIA-PET) to detect putative regulatory interactions involving the cohesin subunit SMC1A in the embryonic mouse limb. The limb is particularly well suited for this purpose. Distant-acting enhancers are essential for limb development (Lettice et al. 2002; Amano et al. 2009). Moreover, a large number of distant-acting enhancers in the limb have been experimentally characterized by chromatin mapping and mouse transgenic assays, providing a functional basis for interpreting cohesin-mediated interactions (Visel et al. 2007, 2009; Cotney et al. 2012). Previous ChIA-PET studies have been performed in mouse and human cell culture to capture interactions involving transcription factor estrogen receptor-α, CTCF, and RNAPII (Fullwood et al. 2009; Handoko et al. 2011; Li et al. 2012). However, cohesin is recruited to both insulators and enhancers, suggesting it is involved in diverse regulatory interactions (Kagey et al. 2010; Kim et al. 2011; Majumder and Boss 2011; Seitan et al. 2011; Guo et al. 2012). Our ChIA-PET analysis of cohesin-associated interactions in developing limb revealed tissue-specific enhancer-promoter interactions, as well as interactions involving both cohesin and CTCF that potentially establish constitutive chromatin domains across tissues. Surprisingly, we also identified interactions that are maintained in multiple tissues between promoters and distal regulatory elements that show tissue-specific activation or repression during development.

[4]Corresponding author
E-mail james.noonan@yale.edu
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.156570.113. Freely available online through the *Genome Research* Open Access option.

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

Results

Cohesin occupies developmental enhancers

We first identified cohesin binding sites using ChIP-seq analysis of the cohesin subunit SMC1A in mouse embryonic (E11.5) limb bud. To functionally annotate SMC1A binding sites, we also mapped CTCF sites and three histone modifications: H3K27ac, which marks active promoters and enhancers; H3K4me2, which marks both active and inactive enhancers; and H3K27me3, which marks sites repressed by the Polycomb repressive complex PRC2 (Supplemental Table S1; Heintzman et al. 2009; Creyghton et al. 2010; Ernst et al. 2011). We identified 41,114 SMC1A binding sites, most of which are located at intergenic or intronic regions (68%) (Fig. 1A). CTCF is present at 42% of intergenic or intronic SMC1A sites, consistent with previous studies (Schmidt et al. 2010; Faure et al. 2012). SMC1A sites that do not recruit CTCF show enrichment for H3K27ac and H3K4me2, suggesting they include enhancers (Supplemental Fig. S1A; Cotney et al. 2012). These sites also enrich for experimentally validated limb enhancers and limb-specific subsets of the enhancer-associated factor EP300 and H3K27ac (Fig. 1B,C; Supplemental Fig. S1B; Visel et al. 2009; Cotney et al. 2012) and are strongly associated with genes implicated in embryonic morphogenesis and limb development (Supplemental Fig. S1C; McLean et al. 2010). We did not detect significant evidence of SMC1A binding at 35% of known limb enhancers (Supplemental Fig. S1B). These may be false negatives in our ChIP-seq experiments, especially for enhancers active in a restricted population of cells in the limb. Alternatively, it may indicate a subset of enhancers do not recruit SMC1A at sufficient levels for detection. The dip in H3K27ac signal at known enhancers and EP300 sites co-occupied by cohesin suggests that cohesin binding displaces nucleosomes, similar to binding of a transcription factor (Fig. 1C,D; Cotney et al. 2012; Rada-Iglesias et al. 2012). Although most putative SMC1A-bound enhancers do not appear to recruit CTCF, there are a small number of regions that are marked by SMC1A, CTCF, and H3K27ac or H3K4me2 (1626 and 2369, respectively) (Fig. 1E; Supplemental Fig. S1D).

Discovery and analysis of cohesin-associated interactions

To identify chromatin interactions associated with SMC1A bound sites, we used chromatin interaction analysis with paired-end sequencing (ChIA-PET) in developing limb. From two ChIA-PET

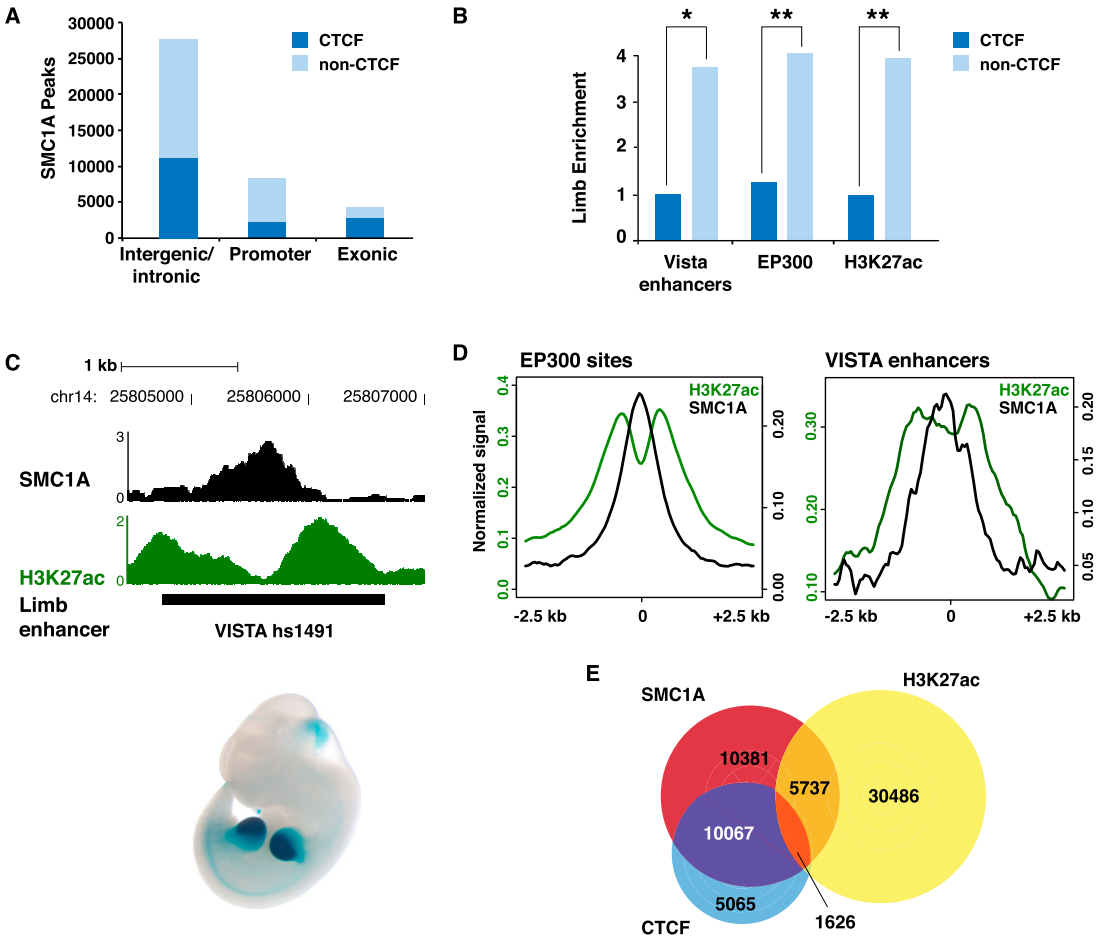


Figure 1. Cohesin binding in embryonic E11.5 limb bud. (*A*) Classification of SMC1A ChIP-seq peaks, partitioned by CTCF co-occupancy. (*B*) Intergenic and intronic SMC1A sites lacking CTCF (light blue bars) are enriched for limb enhancers in the VISTA Enhancer Browser (Visel et al. 2007) ([*] Fisher exact test *P*-value = 0.001) and putative enhancer marks, including the coactivator EP300 (Visel et al. 2009) and E11.5 limb-specific H3K27ac marking (Cotney et al. 2012) ([**] Fisher exact test *P*-value < 2×10^{-16}), compared to sites co-occupied by CTCF (dark blue bars). (*C*) SMC1A and H3K27ac ChIP-seq signal profiles at a known limb enhancer, VISTA hs1491 (Visel et al. 2007). (*D*) SMC1A (black) and H3K27ac (green) normalized ChIP-seq signal aggregated at intergenic or intronic limb EP300 sites (*left*; $n = 3613$) (Visel et al. 2009) and known VISTA limb enhancers (*right*; $n = 165$) (Visel et al. 2007). (*E*) Overlap of intergenic and intronic sites occupied by SMC1A, CTCF, and H3K27ac.

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

libraries, we generated 338,907,326 read pairs and identified 28,320 intrachromosomal ligation products, using a minimum size limit of 5 kb to exclude self-ligation events (Supplemental Table S2). To be considered further, we required each interaction to show SMC1A binding at one or both anchors. We also limited our study to interactions spanning <1 Mb, as the number of observed interactions at or below this length significantly exceeds the number of interactions expected from stochastic interaction events or random ligation (false discovery rate [FDR] ≤ 0.10) (Supplemental Fig. S2). Using this approach, we identified 2264 SMC1A interactions (Supplemental Table S3). We captured two major classes of interactions: those between two intergenic or intronic regions (1330), of which 68% are co-occupied by CTCF at either or both anchors; and interactions between promoters and distal intergenic or intronic regions (680), of which 56% are co-occupied by CTCF (Fig. 2A). The distribution of ChIA-PET interactions on chromosome 2, relative to chromatin modification patterns and gene expression levels, is shown in Figure 2B. We captured previously unknown interactions involving known developmental regulators, including *Snai1* (Figs. 2B, 3A). However, our ChIA-PET experiments likely suffer from a high false negative rate, due to regulatory heterogeneity in the limb bud and compounded inefficiencies in ChIP enrichment and subsequent ligation of interacting sites (Supplemental Note).

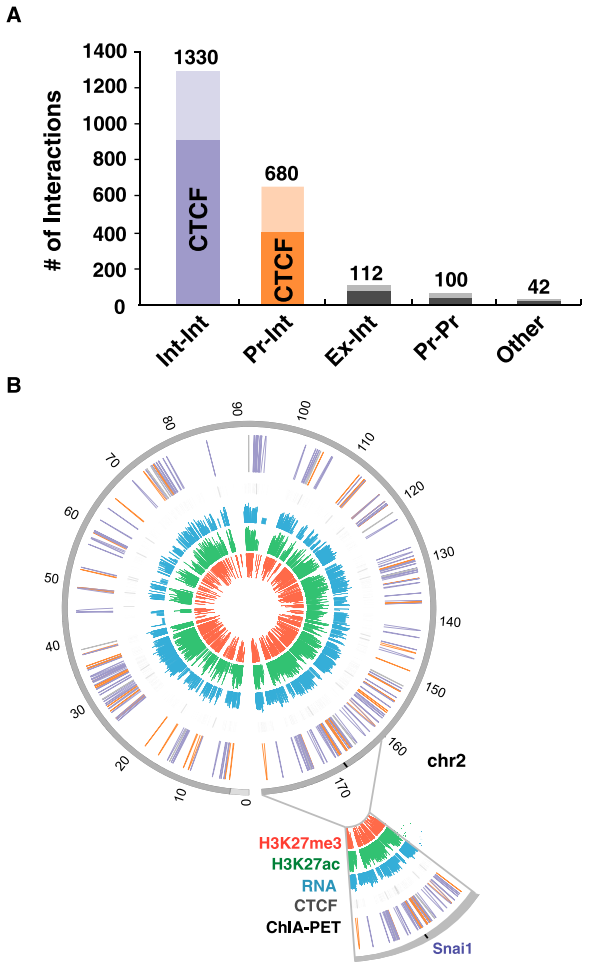


Figure 2. Cohesin interactions in the genome. (*A*) Classification of SMC1A ChIA-PET interactions, partitioned by CTCF co-occupancy. (Int) Intergenic or intronic sites; (Pr) promoter; (Ex) exonic. (*B*) Circos map of ChIA-PET interactions on chromosome 2. The *outermost* to *innermost* tracks are chromosome 2 (dark gray; centromere location indicated in light gray), ChIA-PET interactions (Int–Int interactions shown in blue, Pr–Int interactions in orange, and all others in gray), CTCF peaks (gray), *z*-scores of $\log_2$ transformed RPKM values of RNA-seq (blue), H3K27ac (green), and H3K27me3 (red). An expanded view of a 20-Mb region on chromosome 2 is shown with the location of the *Snai1* locus (Fig. 3A).

The substantial overlap we observed between cohesin-associated interactions and CTCF binding suggests these interactions may partition chromatin into discrete domains (Handoko et al. 2011). Sites located within the interaction we detected at *Snai1* show correlated presence or absence of H3K27ac enrichment in E11.5 limb and E14.5 cortex, respectively (Fig. 3A). Analysis of all SMC1A interactions suggest they show highly correlated patterns of H3K27ac marking across 19 embryonic and adult mouse tissues (Fig. 3B; Shen et al. 2012). H3K27me3 marking is also highly correlated across six tissues (Fig. 3B; ENCODE Project Consortium 2011). The highest correlations of H3K27ac and H3K27me3 occur within SMC1A loops compared to flanking intervals (Wilcoxon P-value $< 2 \times 10^{-16}$) (Fig. 3C). Additionally, the expression levels of genes contained within SMC1A loops are significantly correlated across tissues (Wilcoxon P-value $< 2 \times 10^{-16}$) (Fig. 3C). These results suggest SMC1A interactions establish discrete domains with correlated histone modification states and gene expression. SMC1A interactions still significantly partition H3K27ac and H3K27me3 when we exclude interactions involving a CTCF binding event or motif from the analysis (Wilcoxon P-value $< 2 \times 10^{-16}$) (Supplemental Fig. S3A,B). If promoter-distal site interactions are considered separately from all interactions, they show correlated chromatin marking within the loop, suggesting that promoter-distal site interactions may also demarcate regulatory domains (Supplemental Fig. S3C).

We next considered putative enhancer-promoter interactions. One such interaction occurs at *Pitx1*, which is required for hindlimb development (Lanctôt et al. 1999). *Pitx1* shows hindlimb-specific expression at E11.5 and is looped to a previously uncharacterized distal site 133 kb away that has hindlimb-specific H3K27ac marking (Cotney et al. 2012). In forelimb and embryonic cortex this distal site is marked by H3K27me3 (Fig. 4A). Chromosome conformation capture (3C) analysis suggests this interaction is specific to the embryonic hindlimb (Fig. 4B; Supplemental Fig. S4A). The distal interacting site is bound by CTCF in limb and cortex (Fig. 4A), suggesting that constitutive CTCF binding events participate in tissue-specific enhancer-promoter interactions. For all SMC1A interactions we detected that involve a promoter and distal site, the interacting promoters show significantly higher gene expression compared to genes contained within the loop (Wilcoxon P-value $< 2 \times 10^{-16}$) (Fig. 4C). The distal sites show enrichment for H3K27ac and H3K4me2 (Fig. 4D). These observations suggest a subset of SMC1A interactions between promoters and distal sites result in tissue-specific transcriptional activation.

Our results also suggest cohesin is involved in promoter–promoter and enhancer–enhancer interactions. We captured a known interaction between the *Irx3* and *Irx5* promoters (Tena et al. 2011), which are both bound by SMC1A and may facilitate shared transcriptional regulation (Supplemental Fig. S4C). Additionally, we identify previously characterized interactions between enhancers at the *HoxD* locus (Supplemental Fig. S4B; Gonzalez et al. 2007; Montavon et al. 2011). Overall, we identified 258 SMC1A interactions in limb that involve two H3K27ac-marked intergenic or intronic sites. These sites may participate in larger "regulatory archipelagos," which have been shown to bring together multiple

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

regulatory elements to robustly drive transcription of target genes (Montavon et al. 2011). We also detected 41 interactions that involve a known developmental enhancer, only six of which are enhancer-promoter interactions (Supplemental Table S4; Visel et al. 2007). The remaining interactions occur between known enhancers and distal sites that are not promoters; 19 of these sites are marked

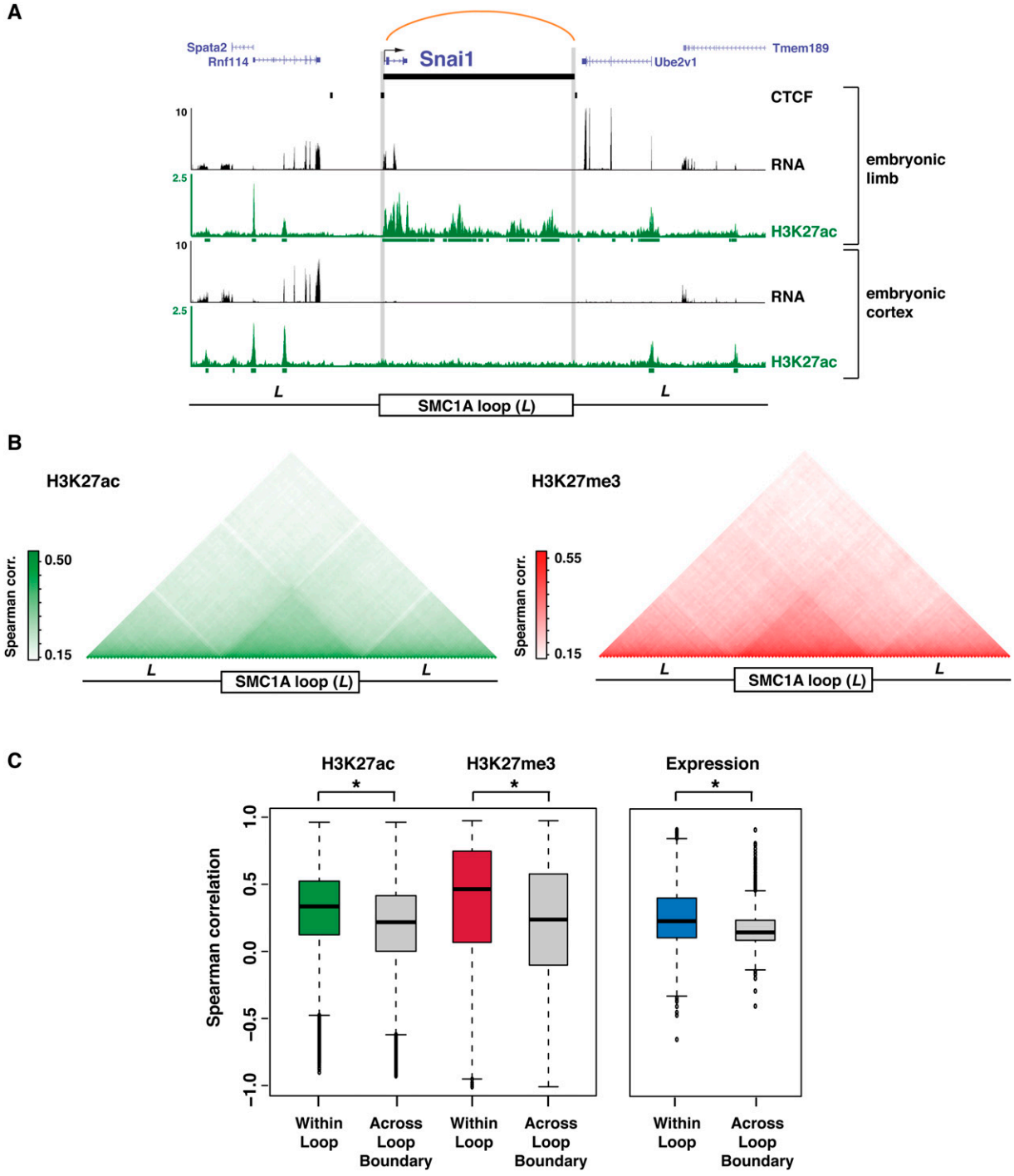


Figure 3. Cohesin interactions partition chromatin states and gene expression levels. (*A*) SMC1A ChIA-PET interaction at the *Snai1* locus in E11.5 limb, with associated H3K27ac ChIP-seq and RNA-seq profiles in limb and E14.5 cortex (Ayoub et al. 2011). The upstream and downstream flanking regions are also shown (+/− one loop distance, *L*). (*B, left*) Mean pairwise Spearman correlations of H3K27ac signal across 19 cell types (Shen et al. 2012) using sites binned across all SMC1A interactions, spanning from one loop distance upstream (*L*) to one loop distance downstream (*L*). (*Right*) Mean pairwise Spearman correlations of H3K27me3 signal across six cell types (The ENCODE Project Consortium 2011) across all SMC1A interactions. (*C, left*) Distribution of pairwise Spearman correlations of H3K27ac and H3K27me3 signal within SMC1A interactions (within loop) or other pairings in the *3L* interval (across loop boundary) ([*] Wilcoxon *P*-value < 2×10^{-16}). (*Right*) Distribution of pairwise Spearman correlations of gene expression for SMC1A interactions containing two or more genes (within loop) compared to other pairings in the *3L* interval (across loop boundary) ([*] Wilcoxon *P*-value < 2×10^{-16}).

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

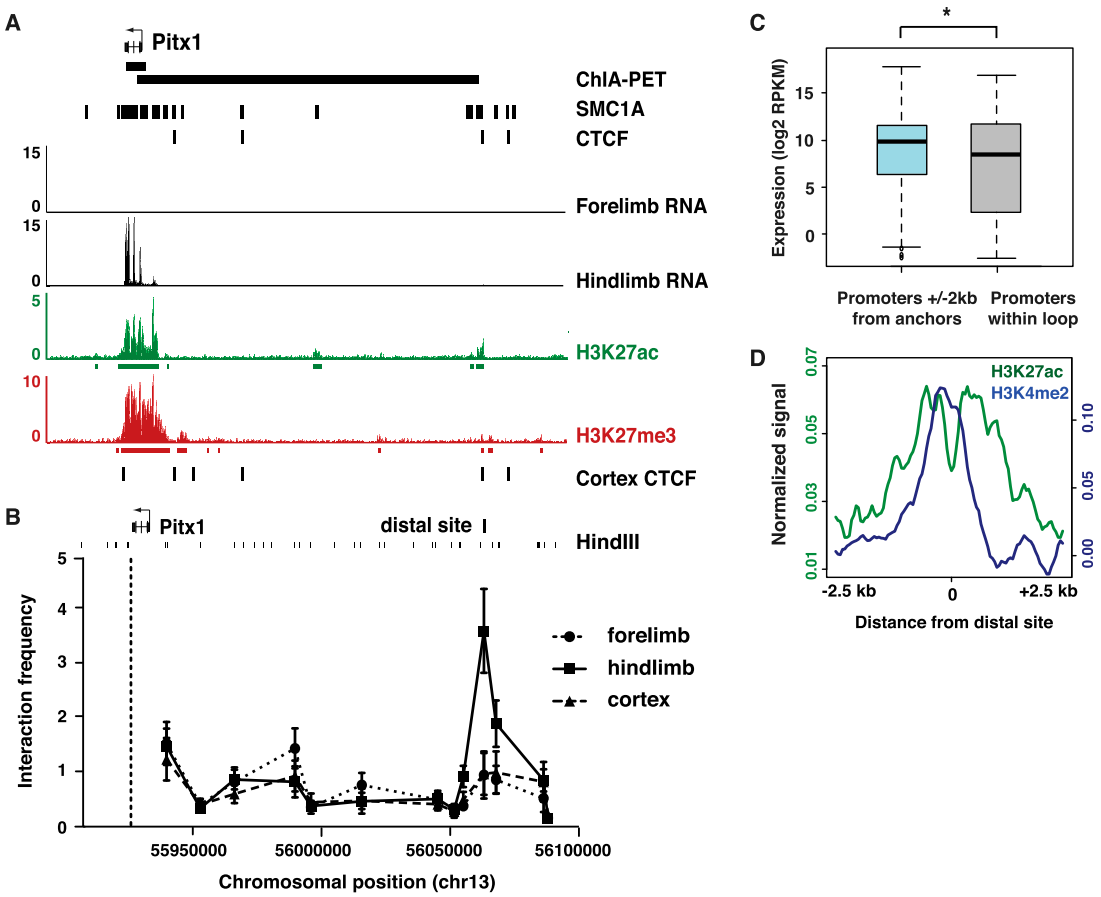


Figure 4. Cohesin participates in enhancer-promoter interactions. (*A*) SMC1A ChIA-PET interaction between *Pitx1* and a distal site 133 kb upstream, bound by CTCF and SMC1A. Although E11.5 forelimb and hindlimb tissue were combined for ChIP, it has previously been shown that both *Pitx1* and the distal site are marked only by H3K27ac in hindlimb and only by H3K27me3 in forelimb (Cotney et al. 2012). CTCF binding sites in E14.5 cortex are also shown (Shen et al. 2012). (*B*) 3C analysis of interactions at the *Pitx1* locus in E11.5 forelimb, E11.5 hindlimb, and E14.5 cortex, using an anchor primer located in the gene (dashed line) and primers tiled across the locus. Interactions are normalized to a crosslinking control at *Ercc3*. Each data point represents average interaction frequency and error bars represent standard error from three independent qPCR reactions (see Supplemental Fig. S4A for biological replicate). (*C*) Promoters within 2 kb of an interaction anchor that loop to a distal site show significantly higher expression than promoters contained within the loop ([*] Wilcoxon *P*-value < 2×10^{-16}). (*D*) E11.5 limb H3K27ac (green) and H3K4me2 (blue) normalized ChIP-seq signal aggregated at all distal sites looping to a promoter.

by H3K27ac in limb, suggesting potential enhancer–enhancer interactions. Notably, 22 of the known enhancers in our interaction set are only active in nonlimb embryonic tissues, suggesting inactive regulatory elements may also participate in long-range looping events.

A subset of cohesin interactions are present in multiple tissues and involve active, repressed, and poised loci

The correlated patterns of chromatin modification and gene expression we observe, coupled with the interactions we detected that involve enhancers not active in limb, suggest a subset of SMC1A interactions may be maintained across multiple tissues. One such interaction occurs at *Wnt7a*, a regulator of dorsal-ventral patterning (Parr and McMahon 1995). In the limb, *Wnt7a* expression is restricted to the dorsal ectoderm, whereas it is not expressed in limb bud mesenchyme (Parr et al. 1993). Since the mesenchyme comprises most of the limb bud at the time point we interrogated, most of our ChIP-seq signal and ChIA-PET interactions are likely derived from it rather than ectoderm. However, in the limb bud we find that the *Wnt7a* promoter interacts with a distal site 125 kb upstream. Both the promoter and the distal site are marked by H3K27me3 repression and bound by CTCF in limb (Fig. 5A). In embryonic cortex, both the promoter and distal site are marked by H3K27ac and Wnt7a is highly expressed (Parr et al. 1993; Ayoub et al. 2011). 3C analysis detects the interaction in both limb and cortex (Fig. 5B; Supplemental Fig. S5B), and the correlated H3K27ac and H3K27me3 chromatin states suggest the interaction is maintained across multiple tissues (Supplemental Fig. S5A). These results support a model in which the interaction we detect in limb serves to recruit tissue-specific regulatory factors that repress *Wnt7a* in the limb bud mesenchyme. Conversely, the same interaction at *Wnt7a* in embryonic cortex may recruit factors that enhance transcription. To address the prevalence of interactions showing different chromatin states in limb and cortex, we examined patterns of H3K27ac and H3K27me3 in both tissues at SMC1A limb interactions. We find 12 of the 61 promoter-distal site interactions marked at both anchors by H3K27me3 in limb show H3K27ac in cortex, and 27 of the 196 promoter-distal site interactions marked at both anchors H3K27ac in limb show H3K27me3 in cortex. These

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

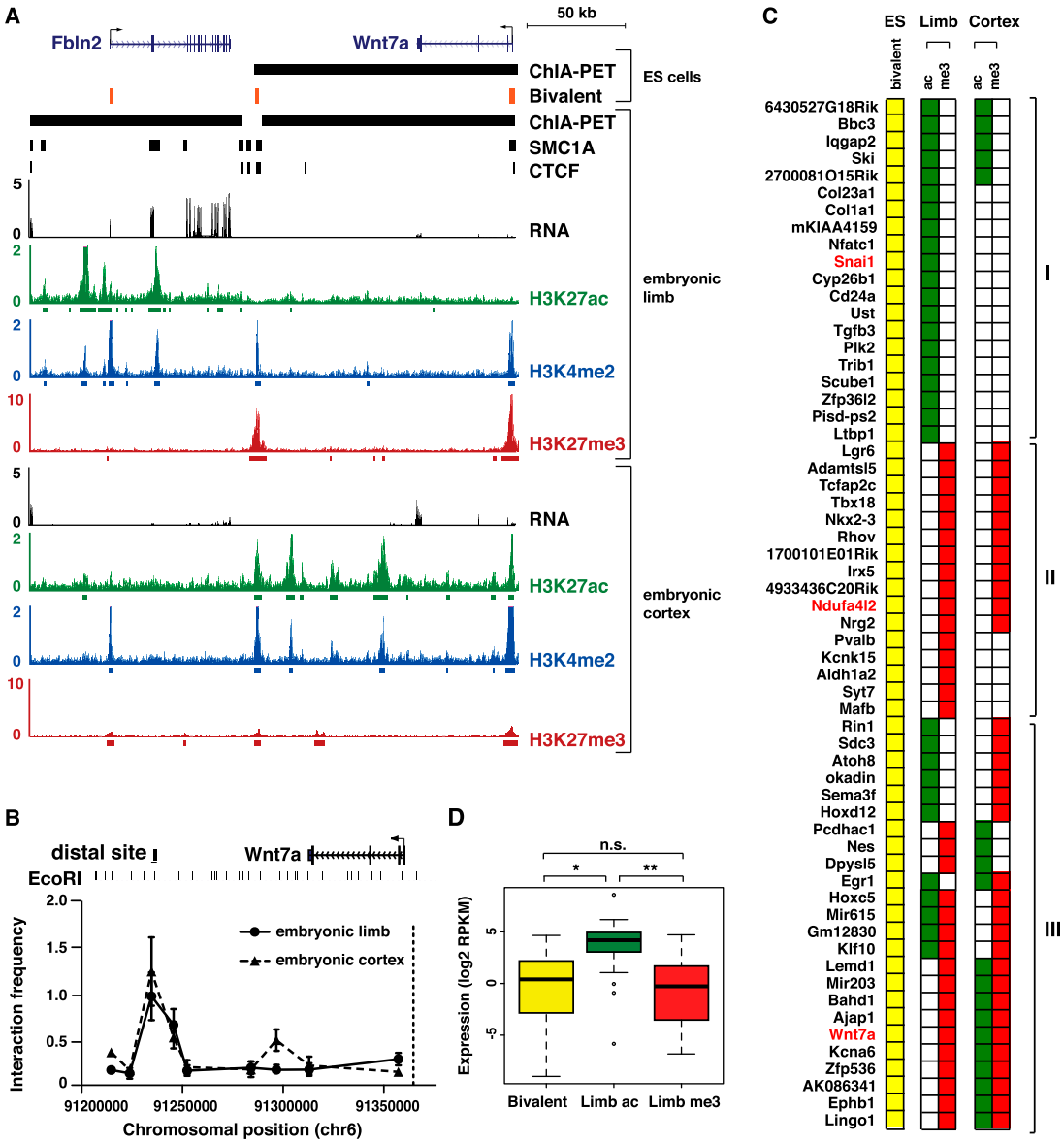


Figure 5. Tissue-dependent activation or repression of poised cohesin interactions. (*A*) SMC1A/CTCF interaction at *Wnt7a* in ES cells (Handoko et al. 2011) and limb. Bivalent sites (H3K4me3 and H3K27me3) in ES cells are shown in orange (Mikkelsen et al. 2007). RNA-seq and H3K27ac in E11.5 limb and E14.5 cortex, and H3K27me3 in E11.5 limb and E11.5 cortex are shown (E14.5 cortex RNA obtained from Ayoub et al. (2011). (*B*) 3C analysis of E11.5 limb and E14.5 cortex interactions between the *Wnt7a* promoter (dashed line) and distal sites across the locus, normalized to a crosslinking control at *Ercc3*. Each data point represents average interaction frequency and error bars represent standard error from three independent qPCR reactions (see Supplemental Fig. S5B for biological replicate). (*C*) Diagram of a subset of bivalent ES cell promoters involved in SMC1A or CTCF ChIA-PET interactions (Handoko et al. 2011). Only interactions with concordant chromatin states at both the promoter and distal site were considered. (I) Interactions resolving to active H3K27ac in limb; (II) interactions resolving to repressive H3K27me3 in limb; (III) interactions resolving to opposite chromatin states in limb and cortex. Gene names in red are involved in interactions detected in both ES cells (Handoko et al. 2011) and limb. The full diagram can be found in Supplemental Figure S5. (*D*) Gene expression of promoters involved in poised, activated, or repressed interactions. Promoters of H3K27ac marked limb interactions show significantly higher expression than bivalently marked promoters in ES cells ([*] Wilcoxon *P*-value = 1.6×10^{-11}) (Shen et al. 2012) and promoters in H3K27me3 marked interactions ([**] Wilcoxon *P*-value = 6.1×10^{-10}).

results suggest that a subset of promoter-regulatory element interactions may be maintained in multiple tissues and recruit tissue-specific regulatory factors that serve to activate or repress transcription of their target gene depending on the tissue context.

To identify additional SMC1A interactions present in multiple tissues, we compared our limb SMC1A interactions to a set of CTCF-mediated interactions previously identified in ES cells (Handoko et al. 2011). We observe 52 interactions that are present in both data sets, including the *Wnt7a* interaction. In ES cells, the *Wnt7a* promoter is poised, showing bivalent H3K4me3 and H3K27me3 marking (Fig. 5A). The distal site we identified that interacts with *Wnt7a* also exhibits a bivalent chromatin state in ES cells (Mikkelsen et al. 2007). Upon differentiation, the bivalent marks at both the promoter and the distal site resolve to H3K27me3 or H3K27ac in embryonic limb or cortex, respectively, while also maintaining H3K4me2, which marks both active and inactive

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

enhancers and promoters (Fig. 5A). The *Snai1* interaction we identified in limb is also present in ES cells (Handoko et al. 2011), and the *Snai1* promoter shows bivalent marking but resolves to H3K27ac and is transcriptionally active in limb (Mikkelsen et al. 2007).

To investigate whether maintenance of a consistent regulatory topology in both bivalent and resolved states may occur at other loci, we characterized all interactions that recruit CTCF in ES cells and/or embryonic limb and also involve bivalent promoters in ES cells (Mikkelsen et al. 2007). We found 83 bivalent interactions that resolve to an active or repressed chromatin state in limb. Of these, 40 resolve to H3K27ac in limb at the promoter and distal site, and 39 resolve to H3K27me3 (Fig. 5C; Supplemental Fig. S5C). The remaining four interactions are marked by both H3K27ac and H3K27me3 at the promoter and distal site, suggesting both elements exhibit a gradient of activation or repression in the limb bud (Cotney et al. 2012). Genes participating in these 83 interactions show low expression in mouse ES cells (Fig. 5D). In limb, genes that resolve to H3K27ac marking show significantly higher expression than in ES cells (Wilcoxon P-value = 1.69×10^{-11}), whereas the genes that resolve to H3K27me3 show no significant change in expression (Fig. 5D).

A previous study of the protocadherin alpha cluster suggests that removal of an enhancer participating in one such bivalent interaction results in both tissue-specific aberrant activation and reduced expression of the target gene. We detect an interaction in limb between the *Pcdhac1* promoter and the HS5-1 enhancer, both marked by H3K27me3 (Supplemental Fig. S6). This interaction also occurs in a cultured neural cell line that expresses *Pcdhac1* (Guo et al. 2012). The promoter is bivalent in ES cells (Mikkelsen et al. 2007). Previous studies indicate that HS5-1 is necessary for driving robust expression of protocadherin alpha genes, and deletion of HS5-1 not only results in threefold reduction of *Pcdhac1* expression in whole brain, but also a fivefold increase of *Pcdhac1* transcripts in kidney (Kehayova et al. 2011).

Discussion

Using ChIA-PET analysis of SMC1A, we obtained a direct view of cohesin-associated topology in the genome. Our results suggest that cohesin interactions facilitate tissue-specific regulatory outcomes through several mechanisms. Consistent with previous studies of individual loci, we find that cohesin is involved in tissue-specific looping between promoters and enhancers. In these cases, tissue-specific activation of gene expression is likely to depend on tissue-specific interaction events. However, cohesin is also associated with interactions between distal sites and promoters that are maintained across tissues, but show tissue-specific chromatin signatures and gene expression. Both the promoters and the distal sites in these interactions exhibit tissue-specific active or repressed chromatin states, suggesting in these cases that tissue-specific regulation is achieved by altering the activation state of a constitutive chromatin topology.

These "stable" cohesin-associated interactions appear to provide a mechanism for establishing tissue-specific promoter activation and repression through interaction with the same distal site. The interaction we observed at *Wnt7a* involves a distal site marked by H3K27ac in cortex and H3K27me3 in limb, suggesting it may act as an enhancer of *Wnt7a* expression in some tissue contexts and as a repressor in others. The *Wnt7a* promoter also interacts with the same distal site in embryonic stem cells, where it exhibits a bivalent chromatin state. This suggests the same interaction event may maintain a poised state in ES cells and serve to activate or repress the target gene in differentiated tissues (Fig. 6). One such distal site with dual functions is the HS5-1 enhancer at the *Pcdhac1* locus: Loss of the enhancer results in reduced *Pcdhac1* expression in brain and increased expression in tissues that normally exhibit very low levels of *Pcdhac1* (Kehayova et al. 2011).

The mechanisms by which these stable interactions produce tissue-specific transcriptional outcomes remain to be determined. In one potential model, stable interactions are maintained across many tissues by cohesin and CTCF, irrespective of their transcriptional output. Tissue-specific transcription factors would then serve to activate or repress this constitutive regulatory topology. Alternatively, apparent "stable" interactions may be independently specified in different tissues by CTCF in conjunction with cohesin, leading to activation or repression depending on the tissue context. In addition, although we explicitly focused on potentially stable cohesin-associated interactions that involve CTCF in our analysis, other factors besides CTCF may also be sufficient. Conditional deletion of *Ctcf* from the mouse embryonic limb has been shown to result in small changes in overall gene expression. However, critical limb development genes are down-regulated, including *Shh*, *Fgf4*, *Grem1*, and *Jag1*, whereas proapoptotic *Bbc3* is derepressed, potentially contributing to the degeneration of distal limb structures following loss of *Ctcf* (Soshnikova et al. 2010). Therefore, CTCF may only be required at a subset of stable interactions, or may not be necessary for the maintenance of previously established interactions.

Our results also suggest that cohesin generally establishes a stable chromatin topology in the nucleus, in addition to the specific examples we discuss here. Considered collectively, the cohesin-associated interactions we identified exhibit correlated H3K27ac and H3K27me3 chromatin states and gene expression across tissues. This is consistent with previous Hi-C and 5C studies

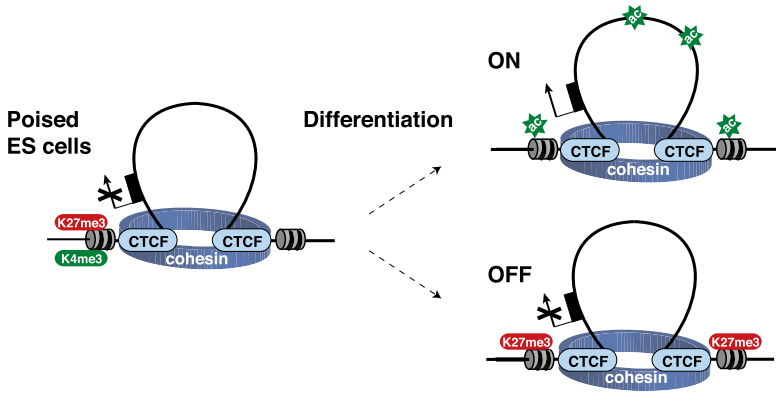


Figure 6. Distal regulatory sites can act as both enhancers and repressors in a tissue-dependent context. A model of tissue-specific gene regulation obtained via stable chromatin interactions. Bivalent promoters in embryonic stem cells are held in a poised, looped conformation with a distal site. Upon differentiation, the distal site is either activated by tissue-specific regulatory factors, leading to deposition H3K27ac and gene transcription, or repressed by recruitment of PRC2, as indicated by H3K27me3 marking.

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

that identified constitutive topological domains maintained across tissues and species (Lieberman-Aiden et al. 2009; Dixon et al. 2012; Nora et al. 2012). Stable cohesin-associated interactions may serve as a constitutive chromatin scaffold that delimits the activity of tissue-specific regulatory elements. For example, the stable interaction at *Wnt7a* encompasses several putative cortex enhancers identified by H3K27ac, and potentially restricts the activity of these enhancers to the *Wnt7a* promoter while excluding outside enhancers from influencing *Wnt7a* expression.

Investigating these hypotheses will require functional studies of *Wnt7a* and other loci that exhibit stable interaction events. For example, transgenic analysis of bacterial artificial chromosomes (BACs) spanning the *Wnt7a* locus, from which the distal interacting site has been removed using recombineering, may determine whether the long-range interaction is required for spatiotemporal regulation of *Wnt7a*. To conclusively establish that the stable interaction at *Wnt7a* maintains the fidelity of *Wnt7a* expression in vivo will ultimately require removal of the distal site from the mouse genome directly using homologous recombination in ES cells and generation of knockout mice. Such models would also potentially reveal developmental phenotypes arising from destabilization of specific long-range chromatin topologies.

Analyses of chromatin modification and transcription factor binding have produced two-dimensional regulatory maps of many mouse and human tissues, but lack the connectivity information required to assign regulatory elements to specific genes. Here, we obtained an initial, genome-wide view of three-dimensional cohesin-associated chromatin interactions. Our results highlight the diverse roles of cohesin in establishing chromatin topology and tissue-specific gene expression and provide insight into how regulatory functions are partitioned in the genome.

Methods

RNA-seq

All animal work was performed in accordance with approved Yale IACUC protocols. C57BL/6J mouse forelimb and hindlimb buds from 24 E11.5 embryos were dissected and total RNA extracted as described in Cotney et al. (2012). RNA-seq libraries were constructed with Illumina TruSeq RNA Sample Preparation Kit and sequenced on an Illumina HiSeq2000 (2 × 75-bp reads). RNA-seq reads from E11.5 forelimb and hindlimb were pooled together and aligned to the mouse reference genome (mm9) using TopHat (v1.4.1) (Trapnell et al. 2009) with a known transcriptome index (UCSC Known Gene annotation; downloaded 11/9/2012) (Dreszer et al. 2012). E14.5 cortical plate RNA-seq data (Ayoub et al. 2011) was downloaded from GEO (mapped reads in bigWig format).

ChIP-seq

SMC1A, CTCF, H3K27ac, H3K4me2, and H3K27me3 ChIP-seq was performed as described in Cotney et al. (2012) with modification. The detailed protocol is provided in the Supplemental Methods. Purified ChIP DNA was prepared for Illumina sequencing with NEBNext ChIP-Seq Library Prep (NEB) with multiplex adaptors and sequencing primers. Libraries were sequenced on an Illumina HiSeq2000 (1 or 2 × 75-bp reads). ChIP-seq reads were aligned to the mouse reference genome (mm9) using Bowtie (v0.12.7) (Langmead et al. 2009). SMC1A and CTCF peaks were called using MACS (Zhang et al. 2008; Feng et al. 2011). For histone modifications, peaks were called using a custom Perl script. Peak calling information is provided in the Supplemental Methods. Intergenic and intronic peaks were identified by filtering exons and regions within 1 kb from a transcription start site (TSS). ChIP-seq signal aggregation analyses of H3K27ac, H3K4me2, and H3K27me3 were performed as described in Cotney et al. (2012).

ChIA-PET

ChIA-PET was performed as described by Fullwood et al. (2010), with modification. Approximately 500 μg of soluble chromatin from E11.5 limb buds was immunoprecipitated in five parallel reactions with Dynabeads Protein G beads (Life Technologies) bound to SMC1A antibody. Custom ChIA-PET ligation adaptors (see Supplemental Table S5 for oligo sequences) containing 3′ T overhangs were ligated onto chromatin (which had complementary A overhangs). Following immobilization of purified ChIA-PET DNA on Dynabeads M-280 Streptavidin beads (Life Technologies), beads were added directly to Phusion High-Fidelity PCR Master Mix (NEB) and PCR amplified with standard Illumina PE primers for 18 cycles. Size-selected ChIA-PET libraries were sequenced on an Illumina HiSeq2000 (2 × 75-bp reads) using a modified denaturation protocol for low-concentration Illumina libraries (Quail et al. 2008). The detailed protocol is provided in the Supplemental Methods.

ChIA-PET read alignment and interaction calling

Paired-end reads were trimmed of ChIA-PET adaptors using Cutadapt (v0.9.4) and the following parameters: -a AGTTGGAT ACCTGCAGTACTAGTCAGTGGGCCC –m 17 –M 18 -O 33 (Martin 2011). Trimmed mate pairs were aligned separately with ELAND (CASAVA v1.8.1) with the following parameters: -ub Y*–bam. Only reads with MAPQ ≥ 1 were retained using SAMtools (Li et al. 2009). Aligned reads were paired with mates and were filtered for PCR duplicates (>1 paired-end read with same start positions of both mates). Paired-end reads were categorized into interchromosomal, intrachromosomal (distance between mates >5 kb), and self (distance between mates <5 kb). Only intrachromosomal reads with mates <1 Mb apart and overlapping at least 1 SMC1A ChIP-seq peak by a minimum of 1 bp were considered for analysis (see "Simulation of ChIA-PET Size Distribution" for justification of the size restriction). Interactions in which both start positions were within 2 kb were joined and their midpoint used to specify their genomic location. The UCSC Known Gene and Known Alternative annotation (downloaded 11/9/2011) (Dreszer et al. 2012) was used to classify interactions. Interaction anchors were designated as a promoter if it was within 2 kb of a transcription start site. Interaction anchors were designated as an exon if it overlapped an exon by at least 1 bp and was not previously assigned to the promoter category. All other anchors were considered intergenic or intronic. To be classified as a CTCF co-occupied interaction, either one or both anchors were required to be within 500 bp of a CTCF ChIP-seq peak in E10.5 (Cotney et al. 2012) or E11.5 mouse limb bud. For correlation analyses of H3K27ac or H3K27me3 at interactions with or without CTCF, we also designated interactions within 500 bp of a predicted CTCF motif as CTCF interactions. See the Supplemental Methods for more information regarding eliciting and mapping the CTCF motif in the mouse genome.

Simulation of ChIA-PET size distribution

To determine a size range in our ChIA-PET data set in which authentic interactions are enriched over random ligation or stochastic collisions, we carried out 100 random ligation simulations. We randomly paired SMC1A ChIP-seq data from biological replicate 1 (reads were trimmed to 18 bp and each read was aligned separately from its pair with ELAND, as described for ChIA-PET)

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

using custom shell scripts and random number generation. The simulated ligations where then categorized into interchromosomal, intrachromosomal, and self. The intrachromosomal ligations were subsampled to equal the total number of observed ChIA-PET intrachromosomal interactions (n = 5469), and the resulting size distribution for each simulation was determined (Supplemental Fig. S2). We restricted our SMC1A ChIA-PET interaction set to those spanning <1 Mb to reduce the likelihood of spurious ligation products (FDR ≤ 0.10). FDR was calculated by dividing the frequency of random ligations <1 Mb by the observed frequency.

Chromosome conformation capture (3C)

3C was performed as described by Miele and Dekker (2009) with modification. Approximately 100 E11.5 limb buds or four E14.5 cortices were dissected in cold PBS. The detailed protocol is provided in the Supplemental Methods. 3C libraries were quantified by PicoGreen dsDNA assay (Invitrogen). Control 3C templates were generated from the following BACs: RP23-400A6 (*Pitx1*) and RP23-237E3 and RP23-446L3 (*Wnt7a*). Digestion and ligation efficiency were assessed using qPCR with Power SYBR Green PCR Master Mix (Life Technologies). Between 5- and 100-ng template DNA was loaded per reaction, and each reaction was performed in triplicate for each biological replicate (see Supplemental Table S5 for oligo sequences). The mean value for each ligation was normalized against an interaction at the *Ercc3* locus for each experiment (Tena et al. 2011), except for the *Pitx1* 3C replicate (Supplemental Fig. S3A), which was normalized by internal copy number (chr9.2_CTCF primers).

Identification of SMC1A sites with known or putative enhancer function

To identify SMC1A binding sites with known developmental enhancer activity, we obtained coordinates for all known human and mouse positive enhancers in the VISTA Enhancer Browser and intersected with SMC1A peaks using BEDTools, requiring at least 1-bp overlap (Visel et al. 2007; Quinlan and Hall 2010). To identify SMC1A binding sites overlapping limb EP300 binding sites, we used E11.5 limb EP300 peak calls from Visel et al. (2009) and intersected as described above. To identify SMC1A binding sites with limb-specific H3K27ac marking, we used the limb-specific E11.5 H3K27ac cluster (compared to mouse ES cells and neural progenitor cells) identified from Cotney et al. (2012) and filtered sites by overlapping (≥1 bp) H3K27ac peaks. Aggregation of H3K27ac and SMC1A ChIP-seq signals was performed as described in Cotney et al. (2012) at EP300 sites and VISTA limb enhancers >2 kb from a TSS.

Generation of Circos map

The map of ChIA-PET interactions on chromosome 2 was generated using the Circos software package (Krzywinski et al. 2009).

Correlation analysis of gene expression and histone modifications

E11.5 limb RNA-seq and H3K27ac ChIP-seq generated in this study was analyzed in conjunction with data from 18 other mouse tissues or cell types (bone marrow, cerebellum, cortex, embryonic brain, adult heart, embryonic heart, adult liver, embryonic liver, intestine, kidney, lung, MEF, mESC, olfactory bulb, placenta, spleen, testis, thymus) reported by Shen et al. (2012). Data was downloaded from GEO (mapped reads in BAM format). Of the 19 data sets reported, mouse limb E14.5 data was excluded due to its redundancy with the limb E11.5 data generated by this study. For RNA-seq data, a composite gene model was generated by combining all annotated transcripts from UCSC Known Gene annotation (downloaded 11/9/2011) (Dreszer et al. 2012), and RPKM (reads per kilobase per million mapped reads) values were calculated using these models. RSEQtools was used to construct the gene models and compute RPKM values (Mortazavi et al. 2008; Habegger et al. 2011). H3K27me3 data from E11.5 limb and E14.5 cortex generated in this study was analyzed in conjunction with data from four other mouse tissues or cell types (C2c12, G1E, mESC, mNPC). C2c12 and G1E data were downloaded from ENCODE (The ENCODE Project Consortium 2011) (mapped reads in BAM format). mESC and mNPC data were downloaded from GEO (Mikkelsen et al. 2007) (mapped reads in BAM format). For all histone modification data sets, only uniquely mapped reads were used for subsequent analyses. PCR duplicates were excluded.

For the histone modification analysis of SMC1A interactions (shown in Fig. 3; Supplemental Fig. S3), the loop and equivalently sized flanking regions upstream and downstream were each binned into 41 windows as follows: Starting from each end, 20 windows were generated of length equal to 1/41 of the size of the loop, and a final window covering any remainder. As a result, 123 bins were generated for each three-loop-size region. Each bin was represented by a vector of its RPKM values for each histone mark across all tissues/cell types (19 for H3K27ac and six for H3K27me3), and pairwise correlations between any two bins were calculated as the Spearman correlation coefficients of such vectors, using the "cor" function in R. To generate mean correlation coefficient heatmaps (shown in Fig. 3B; Supplemental Fig. S3A,B), the pairwise correlation coefficients for each bin were averaged across all loops. To statistically evaluate the correlation of histone modification signals at interactions (Fig. 3C; Supplemental Fig. S3), we considered the original, nonaveraged correlation coefficients and compared pairwise comparisons of bins contained within the loop versus all other pairings of bins in the three-loop size regions.

For the gene expression analysis of SMC1A interactions (shown in Fig. 3C), interactions were filtered using the following criteria: At least two genes were required to be present in the actual loop (995 loops passed filter), and at least one gene in either the upstream or downstream flanking region (973 loops passed filter). Pairwise Spearman correlation coefficients of RPKM values were calculated for each category: (1) all pairwise comparisons of genes within the loop; and (2) all other pairings. For each loop, values in each category were averaged, resulting in one value representing either within loop comparisons or all other pairings.

For the analysis of the *Wnt7a* locus shown in Supplemental Figure S5A, we only considered the region chr6: 91123700–91361700, which includes the ChIA-PET interactions at the Wnt7a locus and the adjacent Fbln2 locus. We used a 2-kb window in this analysis. Pairwise comparisons to calculate Spearman correlation coefficients of RPKM for H3K27ac and H3K27me3 were calculated as described above.

Shared interactions between ES and Limb ChIA-PET

To identify overlapping interactions between the limb SMC1A ChIA-PET and the ES cell CTCF ChIA-PET data reported in Handoko et al. (2011), we used the midpoint of each interaction anchor reported in Handoko et al. Interactions were considered shared between data sets if anchors at both sides were within 10 kb of each other.

Identification of bivalent chromatin interactions

SMC1A ChIA-PET interactions generated in this study (2264) and CTCF ChIA-PET interactions <1 Mb in length (1787) from Handoko

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

et al. (2011) were merged into a single set. The midpoints of interaction anchors were calculated and extended +/− 2 kb. For the SMC1A ChIA-PET interactions, both the promoter and distal anchors were required to overlap a CTCF binding site in limb, mouse ES cells, or E14.5 cortex (Shen et al. 2012). Bivalent promoters were identified in mouse ES cells by intersecting H3K4me3 and H3K27me3 peaks using BEDTools (Quinlan and Hall 2010), called using the same histone modification peak-calling method described in the Supplemental Methods. Putative bivalent interactions were defined as interactions between an intergenic or intronic site and a bivalent promoter. To identify interactions acquiring H3K27ac or H3K27me3 in limb or cortex, H3K27ac peaks in E11.5 limb and E14.5 cortex or H3K27me3 peaks in E11.5 limb and E11.5 cortex were intersected with each anchor. Both the promoter and distal site for each interaction must overlap a given chromatin mark to be considered.

Data access

The data generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo; Edgar et al. 2002) and are accessible under accession number GSE42237.

Acknowledgments

This work was funded by NIH GM094780 (J.P.N.), a career award from the Edward J. Mallinckrodt, Jr. Foundation (J.P.N.), a Howard Hughes Medical Research Scholars Fellowship (L.E.D.), a Rudolph Anderson Fellowship (J.C.), and an NSF Graduate Research Fellowship (S.K.R.).

Author contributions: L.E.D. and J.P.N. conceived and designed the study; L.E.D. developed the experimental approach; L.E.D., J.C., S.K.R., and R.S. performed the experiments; L.E.D., J.L., and J.Y. analyzed the data; L.E.D. and J.P.N. wrote the paper with input from all authors.

References

Amano T, Sagai T, Tanabe H, Mizushina Y, Nakazawa H, Shiroishi T. 2009. Chromosomal dynamics at the *Shh* locus: Limb bud-specific differential regulation of competence and active transcription. *Dev Cell* **16:** 47–57.

Ayoub AE, Oh S, Xie Y, Leng J, Cotney J, Dominguez MH, Noonan JP, Rakic P. 2011. Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing. *Proc Natl Acad Sci* **108:** 14950–14955.

Blow MJ, McCulley DJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, et al. 2010. ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* **42:** 818–822.

Cotney J, Leng J, Oh S, DeMare LE, Reilly SK, Gerstein MB, Noonan JP. 2012. Chromatin state signatures associated with tissue-specific gene expression and enhancer activity in the embryonic limb. *Genome Res* **22:** 1069–1080.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci* **107:** 21931–21936.

Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* **19:** 24–32.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485:** 276–280.

Dreszer TR, Karolchik D, Zweig AS, Hinrichs AS, Raney BJ, Kuhn RM, Meyer LR, Wong M, Sloan CA, Rosenbloom KR, et al. 2012. The UCSC Genome Browser database: Extensions and updates 2011. *Nucleic Acids Res* **40:** D918–D923.

Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30:** 207–210.

The ENCODE Project Consortium. 2011. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol* **9:** e1001046.

Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, et al. 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473:** 43–49.

Faure AJ, Schmidt D, Watt S, Schwalie PC, Wilson MD, Xu H, Ramsay RG, Odom DT, Flicek P. 2012. Cohesin regulates tissue-specific expression by stabilizing highly occupied *cis*-regulatory modules. *Genome Res* **22:** 2163–2175.

Feng J, Liu T, Zhang Y. 2011. Using MACS to identify peaks from ChIP-Seq data. *Curr Protoc Bioinformatics* **34:** 2.14.1–2.14.14

Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, et al. 2009. An oestrogen-receptor-α-bound human chromatin interactome. *Nature* **462:** 58–64.

Fullwood MJ, Han Y, Wei CL, Ruan X, Ruan Y. 2010. Chromatin interaction analysis using paired-end tag sequencing. *Curr Protoc Mol Biol* **89:** 21.15.1–21.15.25

Gonzalez F, Duboule D, Spitz F. 2007. Transgenic analysis of *Hoxd* gene regulation during digit development. *Dev Biol* **306:** 847–859.

Guo Y, Monahan K, Wu H, Gertz J, Varley KE, Li W, Myers RM, Maniatis T, Wu Q. 2012. CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. *Proc Natl Acad Sci* **109:** 21081–21086.

Habegger L, Sboner A, Gianoulis TA, Rozowsky J, Agarwal A, Snyder M, Gerstein M. 2011. RSEQtools: A modular framework to analyze RNA-Seq data using compact, anonymized data summaries. *Bioinformatics* **27:** 281–283.

Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, Lee CW, Ye C, Ping JL, Mulawadi F, et al. 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* **43:** 630–638.

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, et al. 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459:** 108–112.

Jothi R, Cuddapah S, Barski A, Cui K, Zhao K. 2008. Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res* **36:** 5221–5231.

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467:** 430–435.

Kehayova P, Monahan K, Chen W, Maniatis T. 2011. Regulatory elements required for the activation and repression of the protocadherin-α gene cluster. *Proc Natl Acad Sci* **108:** 17195–17200.

Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128:** 1231–1245.

Kim YJ, Cecchini KR, Kim TH. 2011. Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc Natl Acad Sci* **108:** 7391–7396.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: An information aesthetic for comparative genomics. *Genome Res* **19:** 1639–1645.

Lanctôt C, Moreau A, Chamberland M, Tremblay ML, Drouin J. 1999. Hindlimb patterning and mandible development require the *Ptx1* gene. *Development* **126:** 1805–1810.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Lettice LA, Horikoshi T, Heaney SJH, van Baren MJ, van der Linde HC, Breedveld GJ, Joosse M, Akarsu N, Oostra BA, Endo N, et al. 2002. Disruption of a long-range *cis*-acting regulator for *Shh* causes preaxial polydactyly. *Proc Natl Acad Sci* **99:** 7548–7553.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25:** 2078–2079.

Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, et al. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148:** 84–98.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326:** 289–293.

Majumder P, Boss JM. 2011. Cohesin regulates *MHC class II* genes through interactions with MHC class II insulators. *J Immunol* **187:** 4236–4244.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17:** 10–12.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat Biotechnol* **28:** 1630–1639.

Downloaded from genome.cshlp.org on November 8, 2026 . Published by Cold Spring Harbor Laboratory Press

Miele A, Dekker J. 2009. Mapping *cis*- and *trans*- chromatin interaction networks using chromosome conformation capture (3C). *Methods Mol Biol* **464:** 105–121.
Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim T-K, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553–560.
Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D. 2011. A regulatory archipelago controls *Hox* genes transcription in digits. *Cell* **147:** 1132–1145.
Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5:** 621–628.
Nasmyth K, Haering CH. 2009. Cohesin: Its roles and mechanisms. *Annu Rev Genet* **43:** 525–558.
Nitzsche A, Paszkowski-Rogacz M, Matarese F, Janssen-Megens EM, Hubner NC, Schulz H, de Vries I, Ding L, Huebner N, Mann M, et al. 2011. RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. *PLoS ONE* **6:** e19470.
Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485:** 381–385.
Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132:** 422–433.
Parr BA, McMahon AP. 1995. Dorsalizing signal Wnt-7a required for normal polarity of D–V and A–P axes of mouse limb. *Nature* **374:** 350–353.
Parr BA, Shea MJ, Vassileva G, McMahon AP. 1993. Mouse *Wnt* genes exhibit discrete domains of expression in the early embryonic CNS and limb buds. *Development* **119:** 247–261.
Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ. 2008. A large genome center's improvements to the Illumina sequencing system. *Nat Methods* **5:** 1005–1010.
Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842.
Rada-Iglesias A, Bajpai R, Prescott S, Brugmann SA, Swigut T, Wysocka J. 2012. Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* **11:** 633–648.
Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. 2008. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci* **105:** 8309–8314.
Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT. 2010. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res* **20:** 578–588.
Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, et al. 2011. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476:** 467–471.
Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, et al. 2012. A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **488:** 116–120.
Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D. 2010. Functional analysis of CTCF during mammalian limb development. *Dev Cell* **19:** 819–830.
Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. 2008. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J* **27:** 654–666.
Tena JJ, Alonso ME, de la Calle-Mustienes E, Splinter E, de Laat W, Manzanares M, Gómez-Skarmeta JL. 2011. An evolutionarily conserved three-dimensional structure in the vertebrate *Irx* clusters facilitates enhancer sharing and coregulation. *Nat Commun* **2:** 310.
Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25:** 1105–1111.
Visel A, Minovitsky S, Dubchak I, Pennacchio LA. 2007. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res* **35:** D88–D92.
Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, et al. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457:** 854–858.
Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451:** 796–801.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137.

Received February 19, 2013; accepted in revised form May 20, 2013.