HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2020 May 11.

Published in final edited form as:

Cell Rep. 2020 April 14; 31(2): 107503. doi:10.1016/j.celrep.2020.03.067.

A WIZ/Cohesin/CTCF Complex Anchors DNA Loops to Define Gene Expression and Cell Identity

Megan Justice[1,2], **Zachary M. Carico**[3], **Holden C. Stefan**[2], **Jill M. Dowen**[1,2,3,4,5,6,7,*]

[1]Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[2]Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[3]Cancer Epigenetics Training Program, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[4]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[5]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[6]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[7]Lead Contact

SUMMARY

Chromosome structure is a key regulator of gene expression. CTCF and cohesin play critical roles in structuring chromosomes by mediating physical interactions between distant genomic sites. The resulting DNA loops often contain genes and their *cis*-regulatory elements. Despite the importance of DNA loops in maintaining proper transcriptional regulation and cell identity, there is limited understanding of the molecular mechanisms that regulate their dynamics and function. We report a previously unrecognized role for WIZ (widely interspaced zinc finger-containing protein) in DNA loop architecture and regulation of gene expression. WIZ forms a complex with cohesin and CTCF that occupies enhancers, promoters, insulators, and anchors of DNA loops. Aberrant WIZ function alters cohesin occupancy and increases the number of DNA loop structures in the genome. WIZ is required for proper gene expression and transcriptional insulation. Our results uncover an unexpected role for WIZ in DNA loop architecture, transcriptional control, and maintenance of cell identity.

*Correspondence: jilldowen@unc.edu.

AUTHOR CONTRIBUTIONS

M.J., Z.M.C., and J.M.D. conceived of the studies and designed experiments. M.J., Z.M.C., H.C.S., and J.M.D. performed experiments. M.J., Z.M.C., and J.M.D. analyzed data. M.J. and J.M.D. wrote the manuscript with input from all of the authors.

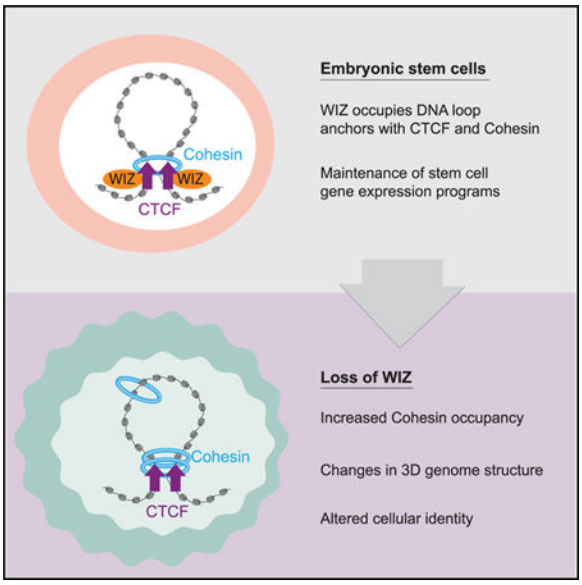
SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2020.03.067.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Graphical Abstract



In Brief

Justice et al. show that WIZ functions with CTCF and cohesin in the structural regulation of DNA loops. Aberrant WIZ function causes many changes in gene expression, including at DNA loops important for regulating stem cell identity genes.

INTRODUCTION

Eukaryotic genomes are organized into DNA loops that play important roles in gene expression control. DNA binding proteins and transcriptional cofactors can facilitate interactions between enhancers and promoters, or spatially constrain such interactions, to ensure proper transcriptional regulation of genes. Despite the importance of DNA looping to genome structure and function, the molecular mechanisms that control dynamic DNA loops and gene expression are poorly understood.

Two structural regulators of DNA loops include cohesin and CTCF. CTCF is a zinc finger-containing protein that binds to a specific DNA sequence motif that is often found at the anchors of DNA loops (Rowley and Corces, 2018). Cohesin is a ringshaped structural maintenance of chromosomes (SMC) protein complex that is thought to facilitate formation or stabilization of a DNA loop. The DNA loops occupied by cohesin and CTCF have been termed topologically associating domains (TADs), loop domains, CTCF contact domains, and insulated neighborhoods (Dixon et al., 2012; Dowen et al., 2014; Gibcus and Dekker, 2013; Gorkin et al., 2014; Merkenschlager and Nora, 2016; Nora et al., 2012). These DNA loops can influence the targeting of enhancers to specific genes at a locus and prevent enhancers from acting on other nearby genes (Dowen et al., 2014). DNA loop anchor sites have also been described as insulator elements for their ability to block other potential DNA loops and serve as barriers to the spread of a chromatin state (Bonev and Cavalli, 2016; Eagen, 2018). Recently, several other proteins have been reported to occupy DNA loop

anchors, including YY1, BRD2, TOP2B, and ZNF143 (Hsu et al., 2017; Uusküla-Reimand et al., 2016; Weintraub et al., 2017; Wen et al., 2018). How these proteins regulate the formation and/or dissolution of DNA loops, and how such activities affect gene expression, remains unclear.

Recently, WIZ (widely interspaced zinc finger-containing protein) was reported to localize to CTCF binding sites and therefore represents a candidate structural regulator of long-range DNA interactions (Isbel et al., 2016). WIZ is a zinc finger-containing protein that occupies promoters and CTCF binding sites in the mouse adult cerebellum (Isbel et al., 2016). Previous studies implicate WIZ in heterochromatin formation via the recruitment and stabilization of G9a and GLP histone methyltransferases to DNA, thereby directing the deposition of H3K9me1, H3K9me2, and H3K27me1 at specific sites in the genome (Bian et al., 2015; Mozzetta et al., 2014; Simon et al., 2015; Ueda et al., 2006; Wu et al., 2011). These histone modifications are associated with HP1 binding and Polycomb-mediated transcriptional repression of genes (Bannister et al., 2001; Lachner et al., 2001; Mozzetta et al., 2014). Like several other DNA loop structuring factors, the homozygous loss of *Wiz* results in embryonic lethality (Daxinger et al., 2013; Isbel et al., 2016). Heterozygous loss of *Wiz* results in improper expression of protocadherin genes in the brain and causes decreased activity and increased anxiety-like behavior in mice (Isbel et al., 2016). Here, we investigate the role of WIZ at CTCF binding sites in the genome and report a role for WIZ distinct from that with G9a and GLP in heterochromatin formation. We identify a function for WIZ in DNA loop architecture, regulation of gene expression, and maintenance of stem cell identity.

RESULTS

WIZ Binds CTCF Sites across the Mammalian Genome

To investigate the chromosomal localization of WIZ relative to other proteins that contribute to long-range DNA interactions, we performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) in mouse embryonic stem cells (mESCs) for WIZ, CTCF, and the cohesin subunit SMC1A (Table S1). We found that WIZ occupies 44,018 sites in the genome, including many sites occupied by CTCF and SMC1A (Figures 1A, 1B, and S1A). WIZ and CTCF signals were highly correlated across the genome and also within peaks (Figure 1C). WIZ binding was enriched at *cis*-regulatory elements, including enhancers, promoters, CTCF sites, cohesin-mediated DNA loop anchors, and super-enhancers (Figures 1D and S1B). Like CTCF and cohesin, WIZ was enriched at the boundaries of insulated neighborhood structures throughout the genome. WIZ binding sites were enriched for the CTCF consensus sequence motif, with its being the top motif represented in WIZ peaks (Figure 1E; Table S2). Motif discovery on (1) WIZ peaks that overlap CTCF peaks and (2) WIZ peaks that do not overlap CTCF peaks revealed mostly overlapping results, with enriched motifs for CTCF, ZIC1, and ZIC4 identified in both lists (Figure S1C). This suggests that WIZ is not recruited to a class of sites independent of CTCF, in a manner that is DNA sequence specific. Importantly, our CTCF ChIP-seq data are similar to other published datasets in terms of both peak number and overlap (Figure S1D) (Nora et al., 2017). We confirmed the specificity of the WIZ antibody using a Myc-tagged version of human WIZ, which is highly conserved with mouse WIZ (Figure S1E). Patterns

of CTCF binding across the genome are strikingly consistent among different cell types (Cuddapah et al., 2009; de Wit et al., 2015; Dixon et al., 2012; Kim et al., 2007; Nora et al., 2012). Similarly, WIZ shows a moderate overlap in binding between mESCs and adult mouse cerebellum, and 17,313 of 21,817 conserved WIZ peaks (79%) across these tissues overlap with conserved CTCF peaks in these tissues (Figure S1F) (Isbel et al., 2016; Shen et al., 2012). Additionally, WIZ and CTCF transcript levels are generally correlated across many different human cell types, consistent with a possible widespread role for WIZ with CTCF (Figure S1G). Together, these data show that WIZ occupies many sites across the genome with cohesin and CTCF, including the anchors of DNA loops and insulated neighborhoods.

WIZ Interacts with CTCF and the Cohesin Complex

To determine if WIZ forms a complex with CTCF and cohesin, we performed co-immunoprecipitations (coIPs) followed by western blots. Pull-downs using antibodies targeting either WIZ or CTCF co-purified CTCF and WIZ, respectively, suggesting that WIZ and CTCF are in a complex with each other (Figure 2A). Additionally, SMC1A was also co-purified using either WIZ or CTCF antibodies. These interactions appear to be independent of DNA and RNA, as nuclear extracts for the coIPs were prepared in the presence of a nuclease. To investigate co-occupancy of CTCF and WIZ on chromatin, a sequential ChIP experiment (re-ChIP) was performed in which a CTCF or control IgG antibody was used in a first ChIP reaction (Figure S2A). From the CTCF ChIP eluate, a second ChIP experiment was performed using CTCF, WIZ, IgG, or no antibody as a control. Both CTCF and WIZ antibodies showed enrichment in the second ChIP, demonstrating that CTCF and WIZ co-occupy chromatin sites. Together these results suggest that WIZ physically interacts, either directly or indirectly, with both CTCF and cohesin.

We next considered whether WIZ directly binds DNA at CTCF sites. WIZ and CTCF both contain multiple C2H2-type zinc finger motifs, but the two proteins share minimal amino acid sequence identity. WIZ has 6 zinc fingers that are widely spaced compared with those of other proteins, including CTCF. CTCF has 11 zinc fingers, of which zinc fingers 4–7 bind the core DNA consensus motif (Nakahashi et al., 2013). Importantly, the amino acids within zinc fingers 4–7 of CTCF that are responsible for the recognition of the CTCF DNA sequence motif are not conserved in WIZ. Because CTCF-occupied sites are known to have low nucleosome density (Nora et al., 2017), we examined nucleosome density at sites co-occupied by WIZ and CTCF and compared them with WIZ-occupied sites that are not CTCF peaks. WIZ/CTCF peaks showed decreased nucleosome occupancy, while WIZ peaks not overlapping CTCF peaks tend to be of low-amplitude signal and do not show well-positioned nucleosomes (Figure S2B) (Mullen et al., 2011). Taken together with the coIP experiments, our data suggest that WIZ exists in a complex with CTCF and cohesin at sites at which CTCF specifically binds to its consensus motif in DNA.

To investigate the function of WIZ in cohesin and CTCF occupancy on the genome, we generated *Wiz*^{del} cells using CRISPR/Cas9 genome editing (Figures S2C and S2D). These cells have a large in-frame deletion that removes 67% of the coding sequence of the *Wiz* gene (including zinc fingers 2–5), likely resulting in a null allele. Western blot analysis

confirmed that an epitope within the deleted region is not detected in *Wiz*^{del} cells. *Wiz*^{del} cells may have a slight reduction in CTCF levels. Likewise, we observed reduced WIZ levels following small interfering RNA (siRNA) depletion of CTCF in wild-type (WT) cells, possibly indicating that WIZ protein stability is sensitive to CTCF levels. Additionally, GAPDH siRNA treatment in *Wiz*^{del} cells may have reduced CTCF levels compared with WT cells. Importantly, the levels of cohesin subunits SMC1A, RAD21, and SMC3 are largely unaltered in *Wiz*^{del} cells (Figure S2E).

We performed ChIP-seq for CTCF and the cohesin subunit RAD21 in WT and *Wiz*^{del} cells (Figures 2B and 2C). Notably, the core cohesin complex members SMC1A, SMC3, and RAD21 are frequently used interchangeably as proxies for cohesin and their signals are correlated. We used a spike-in of human chromatin during the ChIPs in order to quantitatively measure relative levels of enrichment. *Wiz*^{del} cells showed increased cohesin signal, which could be detected in two distinct analyses. First, ChIP-seq for RAD21 revealed a striking increase in the number of cohesin peaks in *Wiz*^{del} cells (Figure 2D). Although most of the RAD21 peaks in WT cells were preserved in *Wiz*^{del} cells (84%), there was also a large increase in ectopic RAD21 peaks in *Wiz*^{del} cells (Figure 2E). The 25,549 ectopic RAD21 peaks represented 48% of the total RAD21 peaks detected in *Wiz*^{del} cells. The ectopic RAD21 peaks rarely overlapped CTCF sites (1,896 of 25,549), which is a strikingly different pattern from the shared RAD21 peaks that preferentially occur at CTCF sites. Motif discovery performed on the 25,549 ectopic RAD21 peaks identified the CTCF motif, as a subset of ectopic cohesin sites do overlap with CTCF, but did not reveal a striking relationship with other DNA-binding factors (Figure S2F). Importantly, the ectopic RAD21 peaks in *Wiz*^{del} cells did not strongly overlap SMC1A peaks in WT cells. Specifically, 13,729 of the 25,549 ectopic RAD21 peaks did not overlap SMC1 peaks in WT cells, supporting the notion that these are *de novo* peaks (Figure S2G). Importantly, the RAD21 ChIP-seq datasets showed similar IP efficiencies in WT and *Wiz*^{del} cells (Figure S2H).

A second distinct analysis that identifies differential ChIP-seq signal at a largely conserved peak set was used and revealed increased cohesin signal at cohesin-occupied sites in *Wiz*^{del} cells. We observed 23,056 sites of differential cohesin enrichment in *Wiz*^{del} cells, with the majority (97.3%) of these sites showing stronger RAD21 signal in *Wiz*^{del} cells compared with WT (Figure 2F). Although the analyses used to identify ectopic RAD21 peaks and sites with differential RAD21 signal are distinct, they likely measure different aspects of the same phenomenon of altered cohesin occupancy across the genome.

CTCF binding was largely unchanged in *Wiz*^{del} cells, as 26,498 of 29,082 peaks (91%) were preserved (Figure S2I). Quantitative analysis revealed that only 1,969 sites (7%) exhibited differential CTCF enrichment between WT and *Wiz*^{del} cells, of which most represented increased signal within weak CTCF peaks. The majority of sites with differential CTCF signal also showed differential cohesin signal; however, there was a large class of sites that also displayed only differential cohesin signal (Figure S2J). These data demonstrate that *Wiz*^{del} cells gain a large number of ectopic cohesin peaks across the genome at sites that are rarely CTCF occupied. Additionally, *Wiz*^{del} cells show increased cohesin enrichment at many sites normally occupied by cohesin and CTCF.

WIZ Is Required for Proper Gene Expression and Maintenance of Cell Identity

Proteins involved in structuring DNA loops are required for proper regulation of gene expression, as their loss can cause mis-targeting of enhancers to inappropriate genes and alter expression of cell identity genes (Dowen et al., 2014; Sun et al., 2019). Previous studies have shown that depletion of cohesin or CTCF leads to misregulation of genes, with fewer transcriptional changes seen following acute and partial depletions, and complete or long-term loss causing hundreds of misexpressed genes (Dowen et al., 2013; Ing-Simmons et al., 2015; Kagey et al., 2010; Nora et al., 2017; Rao et al., 2017; Seitan et al., 2013; Sofueva et al., 2013; Viny et al., 2015; Zuin et al., 2014). In order to examine the role of WIZ in gene regulation, we performed RNA sequencing (RNA-seq) in Wiz^{del} cells. Overall, 3,683 genes were differentially expressed in Wiz^{del} cells, with 1,519 genes downregulated and 2,164 genes upregulated compared with WT (10% false discovery rate [FDR]; Figure 3A; Table S3). Gene Ontology analysis revealed that some of the biological processes most affected in Wiz^{del} cells include system development, anatomic structure morphogenesis, and regulation of cell differentiation (Figure 3B; Table S4). Similarly, gene set enrichment analysis (GSEA) revealed that differentially expressed genes (DEGs) are enriched in the ‘‘regulation of embryonic development’’ gene set (Figure S3A; Table S4). Among the DEGs, the stem cell identity genes *Nanog* and *Pou5f1(Oct4)* were downregulated in Wiz^{del} cells, whereas endodermal transcription factors *Sox17* and *Gata6* were among the most upregulated (Figure 3C). These results suggest a widespread role for WIZ in proper transcriptional regulation and maintenance of embryonic stem cell identity.

Given that WIZ co-occupies the genome with CTCF and cohesin, we next evaluated genes whose expression is controlled by genome architecture (Dowen et al., 2014). Insulated neighborhoods are DNA loops formed by cohesin and CTCF. Some insulated neighborhoods focus the activity of strong enhancers on highly expressed target genes inside of loops and prevent enhancers from accessing genes outside of loops (Dowen et al., 2014). Importantly, genes within insulated neighborhoods often encode master transcription factors and other regulators of cell identity. Therefore, to assess whether WIZ supports transcriptional insulation at DNA loops, we examined expression of genes within and outside of insulated neighborhoods previously identified from cohesin ChIA-PET data (Dowen et al., 2014). We focused on insulated neighborhoods that contain super-enhancers, termed super-enhancer domains (SDs), and their highly expressed target genes (Figure 3D) (Dowen et al., 2014). At ten example SDs, we observed decreased expression of super-enhancer target genes and increased expression of genes outside of the DNA loop in Wiz^{del} cells relative to WT cells (Figure 3E). At eight of these ten SDs, WIZ peaks overlapped the anchors of the SD or fell within 1 kb of the anchors of the SD. Moreover, 111 DEGs in Wiz^{del} cells lie within SDs and tend to decrease in expression compared with all DEGs (Figures 3F and S3B). Furthermore, 178 DEGs that are located outside of SDs tend to show increased expression, consistent with inappropriate enhancer targeting. RAD21 signal was increased both at the boundaries and inside of SDs, but CTCF signal was largely unchanged (Figure 3G). Importantly, the function of WIZ in supporting transcriptional insulation appears distinct from its previously reported role as a G9a cofactor in heterochromatin formation. Although more than 8,500 genes are differentially expressed in $G9a^{-/-}$ ESCs (Mozzetta et al., 2014), about 3,500 genes are differentially expressed in Wiz^{del} cells (Figure S3C). Only 2,252

genes are differentially expressed in both *G9a*$^{-/-}$ ESCs and *Wiz*^{del} cells, consistent with the notion that WIZ and G9a do not genocopy each other with regard to their roles in regulating gene expression. Furthermore, although *Wiz*^{del} cells showed a signature change in gene expression inside and immediately outside of SDs, *G9a*$^{-/-}$ cells did not show this pattern (Figure S3D). Together, these results suggest that WIZ is required for proper transcriptional insulation and control of embryonic stem cell gene expression programs.

WIZ Is a Regulator of DNA Loops

To directly investigate whether WIZ is required for DNA loop architecture, we performed Hi-C in WT and *Wiz*^{del} cells. We generated two biological replicates for WT and *Wiz*^{del} cells, totaling 783 million reads and 835 million reads, respectively (Table S5). Overall, WT and *Wiz*^{del} cells displayed similar patterns of DNA interactions, as measured by the distances between PETs in the datasets (Figure S4A). To assess whether specific features of genome organization were altered in *Wiz*^{del} cells, we examined DNA loops, contact domains, and genome compartmentalization using established analysis methods.

DNA loops, distinct contacts between pairs of specific distal loci, were identified using HiCCUPS (Rao et al., 2014). DNA loops were largely intact in *Wiz*^{del} cells; however, local changes in DNA loops were detected, such as at the *Dazl* locus (Figures 4A and 4B). This locus also displayed differential cohesin and CTCF occupancy, as well as differential expression of the *Dazl* and *Tbc1d5* genes between WT and *Wiz*^{del} cells (Figure S4B). Overall, there was an increase in total DNA loop number in *Wiz*^{del} cells (4,119) versus WT cells (3,094) (Figure 4C). By comparing the DNA loops detected in WT cells and *Wiz*^{del} cells, we identified 2,321 persistent loops, in which a DNA loop used identical anchor sites in both cell lines. Differential loops were also detected, in which one or both anchors were altered in one of the cell lines. There were 773 DNA loops specific to WT cells and 1,798 DNA loops specific to *Wiz*^{del} cells. The DNA loops in *Wiz*^{del} cells were smaller than those in WT cells, with the mean loop size decreased from ~606 kb in WT cells to ~521 kb (Figures 4D and S4C). We also assessed the strength of loops using aggregate peak analysis (APA) (Rao et al., 2014), which revealed that persistent DNA loops display stronger APA scores in *Wiz*^{del} cells than WT cells (Figure S4D). As expected, WT cells displayed stronger APA scores than *Wiz*^{del} cells at WT-specific DNA loops, and *Wiz*^{del} cells displayed stronger APA scores than WT cells at *Wiz*^{del}-specific DNA loops.

The increased number of DNA loops detected in *Wiz*^{del} cells is consistent with the increased cohesin occupancy observed by differential RAD21 ChIP-seq signal and the presence of ectopic peaks. Notably, 26.5% of shared RAD21 sites and a similar 23.8% of differential RAD21 sites (showing mostly increased signal in *Wiz*^{del} cells) are located in DNA loop anchors (Figure 4E). Only 9.6% of ectopic RAD21 peaks overlap a DNA loop anchor. This suggests that although there are many new ectopic cohesin sites across the genome of *Wiz*^{del} cells, they are less frequently engaged in a DNA loop than shared cohesin sites.

To investigate the relationship between differential DNA loops and differential gene expression, we identified classes of DNA loops across the genome and examined the expression of genes inside. The classes included persistent loops, differential loops, loops that are both differential and persistent (because of the nesting of multiple DNA loops), or

no loops. Overall, the proportion of DEGs was unchanged across these four classes of DNA loops (Figure S4E). However, there was a difference in the APA score of all DNA loops detected in each cell line, with the loops detected in *Wiz*^{del} cells showing stronger insulation scores than the DNA loops detected in WT cells (Figure S4F). Additionally, DNA loops containing DEGs showed stronger insulation scores in *Wiz*^{del} cells than the DNA loops containing DEGs in WT cells (Figure S4F). Furthermore, at seven of the ten example SDs (Figure 3E), either the gene inside or outside was located within 25 kb of a WT-specific DNA loop anchor or *Wiz*^{del}-specific DNA loop anchor.

Finally, we assessed changes in contact domains, defined as discrete regions of increased chromatin interactions above background, using the Arrowhead algorithm (Rao et al., 2014). Although the number and size of contact domains were not significantly altered between WT and *Wiz*^{del} cells, the overlap of contact domains revealed the presence of WT-specific and *Wiz*^{del}-specific structures (Figures S4G and S4H). Compartmentalization of the genome into A (active) and B (inactive) compartments was investigated using principal-component analysis (Heinz et al., 2010; Lieberman-Aiden et al., 2009) and revealed minimal instances of compartment switching between WT and *Wiz*^{del} cells (Figure 4F). Taken together, these results suggest that loss of WIZ causes changes in DNA loops by altering cohesin occupancy at specific sites across the genome.

DISCUSSION

Here, we demonstrate that WIZ is required for embryonic stem cell identity gene expression programs and represents a DNA loop structuring protein. WIZ forms a complex with CTCF and cohesin at the anchors of DNA loops across the mammalian genome. Aberrant WIZ function causes many changes in gene expression, including at DNA loops important for regulating stem cell identity genes. Like other proteins involved in structuring DNA loops, WIZ is essential for embryonic viability and is ubiquitously expressed across many cell and tissue types (Isbel et al., 2016; Uhlén et al., 2015). Although WIZ has previously been implicated in heterochromatin formation, this work reveals a distinct role for WIZ in transcriptional regulation and DNA loop architecture.

WIZ forms a complex with CTCF and cohesin at many sites across the genome, including CTCF binding sites, enhancers, promoters, DNA loop anchors, and insulated neighborhoods (Dowen et al., 2014). Recruitment of WIZ to these sites is likely mediated by interaction with CTCF and not direct binding to DNA, as coIP experiments revealed a physical interaction between CTCF and WIZ that is not dependent on DNA or RNA. Additionally, although both WIZ and CTCF contain zinc finger motifs, WIZ lacks the key residues that mediate CTCF binding to its consensus DNA sequence motif (Nakahashi et al., 2013). Specific aspects of how CTCF and WIZ interact remain to be investigated, including identifying the domains involved and determining whether the interaction is direct.

WIZ regulates both cohesin distribution on chromatin and DNA loop architecture. *Wiz*^{del} cells display >20,000 ectopic cohesin peaks that tend to not overlap CTCF sites, enhancers, or promoters of genes. The appearance of a large number of ectopic cohesin peaks in *Wiz*^{del} cells suggests that WIZ normally acts by negatively regulating cohesin occupancy on the

genome. The ectopic cohesin peaks are less likely to engage in a DNA loop than shared cohesin peaks. These findings suggest that aberrant cohesin localization alone is not sufficient for formation of DNA loops and that the gained cohesin occupancy and DNA loops in *Wiz*^{del} cells are not simply a consequence of altered gene expression programs, as they tend not to overlap promoters. Aberrant WIZ function caused an overall increase in DNA loop number and decrease in DNA loop size. Both persistent DNA loops and DNA loops specific to *Wiz*^{del} cells were stronger than those in WT cells. Although DNA loops were altered in *Wiz*^{del} cells, major changes in contact domains and compartmentalization of the genome into A and B compartments were not observed. This is consistent with evidence that contact domains and compartments are largely a product of transcriptional and chromatin state, while DNA loops are a product of the activity of cohesin and CTCF (Rao et al., 2017). We did not detect a significant global relationship between DEGs and either persistent or differential DNA loops in our analysis. Taken together, these results suggest that WIZ normally restricts cohesin levels and distribution across the genome limiting the number of DNA loops.

Wiz^{del} cells display gene expression signatures consistent with loss of pluripotency. In mESCs, the genes responsible for maintenance of stem cell identity and pluripotency are housed within DNA loops (Dowen et al., 2014; Sun et al., 2019). *Wiz*^{del} cells showed decreased expression of several of these genes, including *Nanog*, *Pou5f1(Oct4)*, and *Prdm14*. This is consistent with previous studies which found that DNA loop structuring proteins and complexes, such as cohesin, are required for maintenance of stem cell identity (Hu et al., 2009; Kagey et al., 2010). Furthermore, genes that direct changes in cell identity during differentiation, such as *Gata6* and *Sox17*, exist in insulated neighborhoods and show increased expression in *Wiz*^{del} cells. Although our study cannot distinguish between direct and indirect transcriptional effects of WIZ deletion, the altered expression of cell identity genes in *Wiz*^{del} cells likely contributes to broad transcriptional changes, affecting biological processes such as cellular differentiation, morphogenesis, and development. Thus, we conclude that WIZ is required for maintenance of embryonic stem cell identity, potentially through its regulation of DNA loop architecture.

Importantly, the phenotype of *Wiz*^{del} mESCs cannot be fully explained by loss of G9a/GLP-mediated heterochromatin formation. Previous work showed that a double knockout of *G9a* and *GLP* did not alter expression of *Pou5f1(Oct4)*, *Prdm14*, and *Gata6* (Mozzetta et al., 2014), which were identified as DEGs in *Wiz*^{del} cells. If WIZ solely functions in heterochromatin formation, then *Wiz*^{del} cells should largely genocopy loss of *G9a* and *GLP*, but they do not. Instead, we propose that WIZ can regulate gene expression through its role in mediating genome architecture.

Several recent reports have identified candidate DNA loop structuring factors that associate with cohesin and CTCF, including BRD2, ZNF143, YY1, and TOP2A/2B (Hsu et al., 2017; Uusküla-Reimand et al., 2016; Weintraub et al., 2017; Wen et al., 2018). The molecular mechanisms by which these proteins, and WIZ, regulate DNA loop architecture remain unclear. Notably, loss of the cohesin unloading factor WAPL has been shown to increase the number of DNA loops, similar to *Wiz*^{del} cells, but WAPL-deficient cells show larger DNA loops, while *Wiz*^{del} cells display smaller loops than WT cells (Haarhuis et al., 2017). It is

unclear how WIZ might support the lengthening of DNA loops, while also suppressing their number. As *Wiz*^{del} cells did not display an overall change in cohesin subunit protein levels, it is possible that WIZ regulates the ratio of DNA-associated versus free cohesin in the nucleus. Alternatively, WIZ could regulate the translocation of cohesin along DNA and/or cohesin stability at CTCF sites. Further studies are needed to elucidate the precise role of WIZ, along with other structural regulators, in DNA loop architecture.

In conclusion, WIZ is required for proper gene expression and maintenance of stem cell identity. WIZ co-occupies the genome with the DNA loop structuring proteins cohesin and CTCF. Aberrant WIZ function causes an increase in cohesin and, to a lesser to extent, CTCF occupancy across the genome. This is associated with an increase in the number of DNA loops, which tend to be smaller than those found in wild-type cells. This work identifies WIZ as a structural regulator of DNA loop architecture that is important for proper transcriptional regulation of cell identity genes.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jill Dowen (jilldowen@unc.edu). All unique/stable cell reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines—V6.5 murine embryonic stem cells were a gift from R. Young of the Whitehead Institute for Biomedical Research. V6.5 are male cells derived from a C57BL/6(F) x 129/sv(M) cross. HEK293T (female human embryonic kidney) cells were a gift from R. Young of the Whitehead Institute for Biomedical Research.

Cell culture—Naive V6.5 murine embryonic stem cells (mESCs) were grown on irradiated murine embryonic fibroblasts in serum +LIF standard conditions, as previously described (Dowen et al., 2014). Briefly, KnockOut DMEM(Thermo Fisher Scientific, 10829-018) was supplemented with 15% fetal bovine serum (VWR, 97068-085)). Cell counts were obtained on a Countess II FL Automated Cell Counter (Invitrogen). HEK293T cells were cultured in DMEM (GIBCO, 11995065) supplemented with 10% cosmic calf serum (Thermo Fisher Scientific, SH3008703), 1x GlutaMAX (Thermo Fisher Scientific, 35050-061), 100U/ml penicillin, 100ug/ml streptomycin (Thermo Fisher Scientific, 15140-122) and passaged similarly to mESCs.

Genome editing—mESCs were transfected with plasmids containing a sgRNA, Cas9 and a fluorescent gene (eGFP or mCherry) using Lipofectamine 2000 (Thermo Fisher Scientific 11-668-027). Two days later single cells were sorted by UNC Flow Cytometry Core Facility staff using a FACSAria II (BD Biosciences). 10^4 cells were collected, expanded, screened by PCR and DNA sequencing, and cryogenically stored. Individual allele sequences were determined by PCR of the region surrounding the mutated site, followed by TOPO-TA

cloning (Thermo Fisher Scientific, K4575J10) and Sanger sequencing. sgRNA sequences are provided below and were designed using the CRISPR tool (http://zlab.bio/guide-design-resources) (Cong and Zhang, 2015). The *Wiz* deletion allele (referred to in this text as Wiz^{del}) contains a homozygous deletion from exon 3 to exon 7. The official allele name according to the International Committee on Standardized Genetic Nomenclature for Mice is $Wiz^{em1Jdow}$.

sgRNA 1: 5′- CCATGCCCTTCCCGCCTACC —3′

sgRNA 2: 5′- TCCCTGGTTGGCCGAAGTGC 3′

Wiz^{del} murine embryonic stem cell line sequence at the site of genome editing:

CCACTGTCAGCTGCCCTTCCAGCTTCAAGCCCTGGTTCTCCCCCAG-deletion-GCGGGAAGGGCATGGTGAGAGGTAAGCGTG
GCCGTCTTGAAGTCAGGAAGCTTCTGGGTGGCCCCCCTGGCTCCCGGCCCAGGG
GACTTGGGGGCAACTCCGCTGCTGAGGTG
GCCAGCAGCTCTTGCAGGATGTTGATGGGTGAGATGGTGAGTTCCCAGTTGGTGA
TGCCAAAGTCACGAAGGTGGGCCCGGGCA
TGACTGGAGAGGCCAGCCCGAGTATCAAAAC

METHOD DETAILS

Chromatin immunoprecipitation—Chromatin immunoprecipitation (ChIP) was performed using the following antibodies: WIZ (Novus Biologicals, NBP1-80586), CTCF (Active Motif, 61311), SMC1 (Bethyl Laboratories, A500-055A), RAD21 (Bethyl Laboratories, A300-080A).

For WIZ ChIP-seq replicate 1 in wild-type cells, mESCs were crosslinked in 1% formaldehyde (Sigma Aldrich, F1635) for 20 minutes, then quenched with 125 mM glycine. Cells were lysed first with Lysis Buffer 1 (50 mM HEPES-KOH pH7.5, 140 mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100) by incubating cells in the buffer for 10 minutes at 4C. Nuclei were next lysed with Lysis Buffer 2 (10mM Tris-HCl pH 8, 200mM NaCl, 1mM EDTA, and 0.5 mM EGTA) by incubating nuclei in the buffer for 10 minutes at room temp. Finally, nuclear extracts were resuspended in Sonication Buffer 1 (20mM Tris-HCl pH 8, 150mM NaCl, 2mM EDTA, 0.1% SDS, and 1% Triton X-100). Cells were sonicated using a Branson probe sonicator with the following settings: 18% amplitude, 30 s on, 60 s off, 17 cycles, 8.5 minutes total. WIZ antibody (4ug, NBP1-80586) was incubated with Protein G Dynabeads (150ul, Thermo Fisher Scientific, 10004D) for 6 hours at 4C. Unbound antibody was removed by washing beads three times with PBS before sonicated chromatin was added to antibody conjugated beads and incubated overnight at 4C. Beads were washed with sonication buffer, wash buffer 1 (20mM Tris-HCl pH 8, 500mM NaCl, 2mM EDTA, 0.1% SDS, and 1% Triton X-100), wash buffer 2 (10mM Tris-HCl pH 8, 250mM LiCl, 1mM EDTA, and 1% NP-40), and wash buffer 3 (10mM Tris pH 8, 1mM EDTA, and 50mM NaCl). Chromatin was eluted from beads by adding elution buffer (50mM Tris pH 8, 10mM EDTA, and 1% SDS) and incubating at 65C for 1 hour, spinning down the mixture, then moving the supernatant to a new tube. Supernatant was left at 65C

overnight to reverse crosslinks. RNA was degraded by adding TE and RNase A (Sigma Aldrich, R4642) to the tubes at 37C for 2 hours followed by protein degradation with CaCl2 and Proteinase K (NEB, P8107S) for 30 minutes at 55C. DNA was precipitated using phenol chloroform followed by NaCl, glycogen, and ethanol addition. Resulting DNA pellet was resuspended in 10 mM Tris HCl pH 8. ChIP-seq library was prepared using the ThruPLEX-FD Prep kit (Takara, R400428).

For WIZ ChIP-seq replicate 2, CTCF ChIP-seq replicate 1, and SMC1 ChIP-seq replicate 1, 50 million mESCs were crosslinked in 1% formaldehyde for 2 minutes before quenching with 125 mM glycine. Crosslinked cells were lysed using Lysis Buffer 1 and Lysis Buffer 2 before resuspension in Sonication Buffer 1. Human chromatin (from HEK293T cells) was spiked in (to final 5%) prior to sonication for the indicated experiments. Sonication of nuclei was performed on a Covaris E220 with the following settings: Duty Factor 8, PIP/W 210, and 200 cycles per burst for 12 minutes. Chromatin fragments of 200-1,000 base pair size were generated. Antibodies (WIZ, NBP1-80586, 4ug; CTCF, 61311, 10ug; and SMC1A, A300-055A, 10ug) were incubated with 100uL Protein G Dynabeads for 6 hours. Unbound antibody was removed via washing, as detailed above, before incubation of antibody bound beads with chromatin overnight. Beads were then washed with Sonication Buffer 1 and Wash Buffers 1, 2, and 3 as detailed above. Chromatin was eluted as described above. Crosslinks were reversed overnight via incubation at 65C and addition of 5ul Proteinase K. Zymo ChIP DNA Clean and Concentrate kit (Zymo Research, D5205) was used to purify DNA following Proteinase digestion. Sequencing libraries were prepared using NEBNext Ultra II DNA library prep kit for Illumina (NEB, E7645S).

For CTCF ChIP-seq replicate 2, 40 million cells were crosslinked in 1% formaldehyde for 2 minutes, lysed with Lysis buffers 1 and 2 as indicated above, before resuspension in Sonication Buffer 1 and addition of human spike-in chromatin. Sonication was performed on Covaris E220 as detailed above and ChIP protocol was completed as detailed above. Sequencing libraries were prepared using a Hyper Prep kit (Roche/Kapa Biosciences, KK8502) according to manufacturer’s instructions

For SMC1A ChIP-seq replicate 2, cells were crosslinked for 20 minutes in 1% formaldehyde (Sigma Aldrich, F1635) and quenched by adding 125 mM glycine. Cells were lysed using Lysis Buffers 1 and 2, and recovered nuclei were resuspended in Sonication Buffer 1 and sonicated using a Biorupter (Diagenode) water bath sonicator. Insoluble material was then cleared by spinning sonicated lysates for 10 minutes at 21,000 rcf. To perform the IP, 15mg of anti-SMC1A antibody (Bethyl Laboratories, A300-055A) was incubated with 100µl Dynabeads per IP for 8 hours, after which beads were washed twice with PBS to remove excess antibodies. An estimated $25x10^6$ cell equivalents of chromatin in 550ml then were incubated overnight at 4°C with the antibody-bound beads while rotating. Beads were collected using a magnet and the unbound fraction removed, followed by washes with Wash Buffers 1, 2, and 3 as detailed above. Crosslinks were reversed as described above, and phenol:chloroform extraction was used to collect DNA. Libraries were prepared using a ThruPlex DNaseq kit according to manufacturer’s instructions, and 50bp single-end sequencing reads were collected using the Illumina HiSeq 2500 platform.

For RAD21 ChIP-seq, cells were crosslinked in 1% formaldehyde (ThermoFisher, 28906) for 5 minutes in PBS and quenched by adding glycine. Cells were lysed using Lysis Buffers 1 and 2, and recovered nuclei resuspended in 1ml Covaris Shearing Buffer (50mM Tris pH 7.5, 10mM EDTA, 0.1% SDS). Human HEK293T nuclei prepared in the same manner were spiked-in to a final concentration of 5%, and nuclei were sheared for 12 minutes using a Covaris E220 water bath sonicator with device settings of duty factor 5, PIP/W of 140, and 200 cycles per burst. Insoluble material was cleared by spinning sonicated lysates for 10 minutes at 21,000rcf. 10mg anti-RAD21 antibody (Bethyl Laboratories, A300-080A) was incubated for 8 hours with 30μl of Dynabeads in PBS, after which excess antibodies were washed from the beads. 10^7 cell equivalents of sheared chromatin were incubated overnight with the antibody-bound beads in 1 mL Sonication Buffer 1, washed with Wash Buffers 1, 2, and 3, and crosslinks were reversed overnight as described above. DNA was purified using a ChIP DNA Clean and Concentrator Kit (Zymo Research, D5205) according to manufacturer's instructions. Sequencing libraries were prepared using a Hyper Prep kit (Kapa Biosciences, KK8502) according to manufacturer's instructions, and 150bp paired-end sequencing reads were collected using the Illumina NovaSeq SP sequencing platform and reagents.

Re-ChIP—For the Re-ChIP experiment, 120 million wild-type mESCs were crossliked in 1% formaldehyde for 5 minutes before lysing with Lysis Buffers 1 and 2 then resuspension in Covaris Shearing Buffer. Cells were sonicated using a Covaris E220 sonicator with the settings detailed above. Input material was saved post-sonication and the remaining chromatin was divided into three tubes for IP of CTCF (Active Motif, 61311, 10 ug), WIZ (Novus Biologicals, NBP1-80586, 4ug), and IgG (Bethyl Laboratories, P120-101, 10 ug). For the first ChIP, the Active Motif Re-ChIP-IT kit (53016) was used with the following modifications: 1) the antibodies and beads were incubated together for 6 hours prior to the addition of chromatin; 2) incubation of chromatin, antibodies/beads was performed overnight at 4C.

Following the first ChIP, additional input material was saved before each tube of chromatin was divided into four tubes for the second ChIP with CTCF (Active Motif, 61311, 10ug), WIZ (Novus Biologicals, NBP1-805864ug), IgG (Bethyl Laboratories, P120-101, 10 ug), or no antibody as a control and allowed to incubate overnight. Input material from the first ChIP was purified using the Zymo ChIP DNA Clean and Concentrator kit (Zymo Research, D5201) prior to qPCR analysis. qPCR was performed on an Applied Biosystems QuantStudio 6 qPCR machine using primers found in Table S6.

High-throughput sequencing—50bp or 100bp single-end or paired-end sequencing was performed on Illumina Hi-Seq 4000, Hi-Seq 2500, or NovaSeq 6000 platforms using Illumina reagents according to manufacturer's instructions.

Transfection—A pcDNA3.1 empty vector and vector containing human *Wiz* cDNA driven by a CMV promoter with a Myc-His tag was obtained from Dr. Samantha Pattenden. The plasmids were transformed into DH5-α competent *E. coli* and purified using the Zymo II Midiprep kit (Zymo Research, D4200). For transfection, 8 ug of isolated plasmid was mixed with 2 M CaCl2. The mixture was added to 2X HEPES Buffered Saline (HBS) with addition

of air bubbles with an empty pipette. Once entire CaCl2/DNA mixture was added to HBS, the DNA/HBS mixture was added dropwise to a 10 cm plate of HEK293T cells. Cells were incubated for 24 hours before receiving a change to fresh media. After an additional 24 hours, cells were scraped and pelleted for collection.

Co-immunoprecipitation—Co-immunoprecipitation studies were performed using a Nuclear Complex Co-IP Kit (Active Motif, 54001) and Protein G Dynabeads (Thermo Fisher Scientific, 10009D). Each immunoprecipitation was performed using 100ug of nuclear extract. Input material was loaded as a control, with 1X corresponding to 20ug of protein.

Western blotting—Cells were washed with PBS and collected via scraping. Pellets were resuspended in Lysis Buffer A (10mM HEPES pH 7.9, 10mM KCl, 0.1mM EDTA, and 0.1mM EGTA) with 1x protease inhibitor cocktail (Sigma Aldrich, 11697498001) and incubated for 15 minutes at 4C before addition of 1 mL 10% NP-40 and pelleting via centrifugation. The resulting pellet was resuspended in cold TEN250/0.1 buffer (50mM Tris-HCl pH 7.5, 250mM NaCl, 5mM EDTA, and 0.1mM NP-40) and incubated for a minimum 30 minutes at 4C. Following pelleting via centrifugation, the nuclear fraction (supernatant) was collected. Prior to western blotting, protein levels were quantified using the DC assay from BioRad (BioRad, 5000112). Samples were run in 4%-20% Tris-Glycine gels (BioRad, 4568094) and transferred to PVDF membranes (VWR, 29301-856). Membranes were blocked for at least 45 minutes with 5% Blotting-Grade Blocker (BioRad, 1706404) before overnight incubation at 4C with primary antibody. The membrane was then washed 3 × 10 minutes with TBS-T before incubation for 1 hour at room temperature with secondary antibody. After 3 × 10 minute washes with TBS-T, membranes were imaged using either Thermo SuperSignal West Pico (Thermo Fisher Scientific, 34577) or Thermo SuperSignal West Femto (Thermo Fisher Scientific, 34094) chemiluminescent substrate with an Amersham Imager 600. Primary antibodies included those used previously in ChIP-seq experiments and anti-H3 (Abcam, ab1791) and anti-MYC tag (Abcam, ab9106).

RT-qPCR—Three replicates of the *Wiz*^{del} clonal cell line and wild-type cell line were resuspended in 1ml Trizol (Invitrogen, 15596018). Chloroform (Sigma Aldrich, C2432) was added for phase separation. RNA was purified using the Zymo RNA Clean and Concentrator Kit (Zymo Research, R1013). cDNA was prepared with Superscript IV (Thermo Fisher Scientific, 18091050). qPCR was performed on an Applied Biosystems QuantStudio 6 qPCR machine using primers found in Table S6.

RNA-seq—Three replicates of the *Wiz*^{del} clonal cell line and wild-type cell line were resuspended in 1ml Trizol (Invitrogen, 15596018). Chloroform (Sigma Aldrich, C2432) was added for phase separation. RNA was purified using the Zymo RNA Clean and Concentrator Kit (Zymo Research, R1013). Libraries were prepared using a TruSeq RNA Library Prep Kit v2 (Illumina, RS-122-2001) with indexes AR001-AR009. Library cleanup was performed using Agencourt AMPure XP beads (Beckman Coulter, A63881). Sequencing was performed on an Illumina Hi-Seq 4000 with 50 bp paried end reads.

Hi-C—2-5x10^6 cells were crosslinked with 1% formaldehyde in PBS for 2 minutes, then quenched by addition of 125 mM glycine. Hi-C library construction was performed using an Arima-HiC kit (Arima Genomics) and Hyper Prep DNA-seq library prep kit (Kapa Biosciences, KK8502) according to manufacturer's instructions, with the following modifications. Digested Hi-C libraries were treated with Arima ligase for 1 hour at room temperature instead of the recommended 15 minutes, and barcoded sequencing adapters were ligated for 1 hour at 20 °C instead of the recommended 15 minutes.

QUANTIFICATION AND STATISTICAL ANALYSIS

ChIP-seq analysis and normalization—Replicates were merged as raw fastq files before reads were aligned to a merged genome containing both mouse genome assembly mm10 and human genome assembly hg38 using bowtie (v 1.2.2) (parameters -v 2 -p 24 -S -m 1 –best –strata) (Langmead et al., 2009). Mouse chromosomes were denoted by Mchr prefix to allow for separating from human in future steps. Duplicate sequences were removed using samtools (v 1.9) markdup (-r -s) (Li et al., 2009). Reads mapping to mouse and human chromosomes were separated using samtools idxstats and counted with awk. A bam file containing only mouse reads was created using samtools view and converted to bed format using bedtools (v 2.26) bamtobed and reads were extended by 200 bp (Quinlan and Hall, 2010). Extended bed files were used to call peaks using MACS (v 2016-02-15) with a false discovery rate of 1% (macs2 callpeak -f BED -g mm -q 0.01) (Zhang et al., 2008). To obtain a high confidence peak set, called peak summits were expanded by 50 bp on either side using awk and any expanded peak overlapping a repeat element (defined using the Repeat Masker Track from UCSC genome browser) was removed prior to any peak-related analysis. A normalization factor was calculated for each sample using the formula 1/h where h is the number of human aligned reads in millions, as described previously (Orlando et al., 2014). Normalization to the provided reference (human genome), rather than the total read depth of the dataset, enables for the discovery and quantification of dynamic epigenomic changes. The bed file containing mouse reads was converted to bedgraph using bedtools genomecov (-bga -scale 1/h) before being converted to a bigwig file with bedGraphToBigWig from ucsctools (v 320) (Kent et al., 2002). Z-score normalization was performed where indicated using a custom R script from Spencer Nystrom of Dr. Daniel McKay's lab.

Overlap peak lists were generated by using bedtools intersect on summit files generated by MACS extended by 50 bp on either side. Average signal plots were generated using deeptools (v 3.0.1) computeMatrix (reference-point for CTCF sites, promoters, and enhancers; scale-regions for meta-loop anchors and insulated domains) followed by deeptools plotProfile (Ramírez et al., 2016). Enhancers were defined by merging ChIP-seq data from the master transcription factors *Nanog, Sox2,* and *Oct4* and calling peaks on the merged data (Whyte et al., 2013). Heatmaps were generated using deeptools computeMatrix reference-point or scaled-regions followed by deeptools plotHeatmap. Data in Figure 1B were subjected to k-means clustering with 2 clusters using deeptools plotHeatmap–kmeans 2 (chosen based on results from k-means clustering with a range of 2–6 clusters). In order to visualize different classes of binding sites in each group featured in Figure S2J, bedtools bigwigCompare was used to create a subtractive bigwig track in which WT RAD21 ChIP-

seq signal was subtracted from *Wiz*^{del} RAD21 ChIP-seq signal at each site. Each heatmap in Figure S2J was then plotted in order of descending signal based on the subtractive RAD21 heatmap values. Correlation plots were generated using deeptools multiBigwigSummary followed by plotCorrelation (–removeOutliers –skipZeros –corMethod pearson). Coverage tracks were visualized using the UCSC Genome Browser. Unbiased motif analysis was performed using MEME-ChIP (Bailey et al., 2009). Differentially bound CTCF and Cohesin sites were identified using DiffBind (Ross-Innes et al., 2012).

RNA-seq analysis—RNA sequence reads were aligned to genomic sequence using Star (version 2.6.0a) (Dobin et al., 2013). Differentially expressed genes were identified using DESeq2 from Bioconductor (Love et al., 2014) (Table S3). A PCA plot was generated using DESeq2 plotPCA and an MA plot was generated using DESeq2 plotMA. Locations of insulated neighborhoods (Super-enhancer Domains and Polycomb Domains) were obtained from (Dowen et al., 2014). Coordinates of insulated neighborhoods were converted from mm9 to mm10 using the UCSC LiftOver tool. Genes located within or near these neighborhoods were identified using bedtools intersect. Ranked list of DEGs by expression, barplot of GO terms, and barplot of pairs of genes exhibiting the loss of insulation signature were manually generated with Microsoft Excel (Table S4). Violin plots of genes within and near insulated neighborhoods were generated using ggplot2 geom_violin (Wickham, 2016). Significance of violin plots was computed using the Wilcoxan test via compare_means from R package ggpubr (Wickham, 2016). Gene Set Enrichment Analysis was performed to identify potential impacts on biological processes (Subramanian et al., 2005) (Table S4).

Hi-C analysis—Initial processing of Hi-C data was performed with the Juicer software package (Durand et al., 2016a). Reads were aligned using BWA-mem with default parameters, after which PCR duplicates, reads with Q % 30, and self-ligated fragments were filtered out before Hi-C matrices were assembled (Table S5) (Li and Durbin, 2009). Matrices were visualized using the Juicebox software package (Durand et al., 2016b). WT replicates were well correlated (Pearson R = 0.98) and *Wiz*^{del} replicates were well correlated (Pearson R = 0.93). Loops were called using the HiCCUPS algorithm (Rao et al., 2014) (parameters -p 8,4,2 -i 14,10,6) on Knight-Ruiz balanced matrices at resolutions of 5 kb, 10 kb, and 25 kb, and the resulting list of merged loops was used for subsequent analyses. Contact domains were called using the Arrowhead algorithm (Rao et al., 2014) with default parameters on Knight-Ruiz balanced matrices at 25 kb resolution. Eigenvectors for analysis and visualization of compartmentalization were calculated by passing aligned reads into the Homer Hi-C analysis software package (Heinz et al., 2010; Lieberman-Aiden et al., 2009).

Loops were classified as either dynamic or static by measuring the distance from a left anchor in one genotype to the nearest left anchor in the other genotype using bedtools closest and repeating this analysis for right anchors. Due to resolution limitations, a dynamic anchor was defined as being more than 25,001 bp away from the nearest same side anchor in the other genotype. To be considered static, both anchors of a loop must be within 25kb of an anchor in the opposite genotype. A dynamic loop may have either one or two altered anchors. Loop size was determined by measuring the distance from the start of the left anchor to the end of the right anchor. The same analyses were performed using domain

anchors. Genes were assigned to loop structures by performing bedtools intersect between the classes of looped regions and a list of gene promoters. Log fold change of genes inside various looped regions was plotted using pheatmap.

DATA AND CODE AVAILABILITY

The accession number for the raw and processed sequencing data reported in this paper is GEO: GSE137285. All datasets are summarized in Table S1. Oligos used are detailed in Table S6. Custom ChIP-seq processing script is available at GitHub: https://github.com/dowenlab.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank members of the Dowen lab and the labs of Dr. Daniel McKay and Dr. Douglas Phanstiel for helpful discussions and comments. We thank Dr. Jeremy Simon and Spencer Nystrom in the laboratory of Dr. Daniel McKay for discussion about ChIP-seq spike-in normalization. We thank Dr. Samantha Pattenden for sharing the WIZ overexpression plasmid. This work was supported by NIGMS grant R35GM124764. M.J. was supported in part by National Institutes of Health grant T32GM007092. Z.M.C. was supported in part by a grant from the National Cancer Institute under award T32CA217824. We thank the staff of the UNC High-Throughput Sequencing Facility and the UNC Flow Cytometry Core Facility for assistance with fluorescence-activated cell sorting (FACS) to create genome-edited cell lines.

REFERENCES

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, and Noble WS (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37, W202–W208. [PubMed: 19458158]

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, All-shire RC, and Kouzarides T (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120–124. [PubMed: 11242054]

Bian C, Chen Q, and Yu X (2015). The zinc finger proteins ZNF644 and WIZ regulate the G9a/GLP complex for gene repression. eLife 4, e05606.

Bonev B, and Cavalli G (2016). Organization and function of the 3D genome. Nat. Rev. Genet 17, 661–678. [PubMed: 27739532]

Cong L, and Zhang F (2015). Genome engineering using CRISPR-Cas9 system. Methods Mol. Biol 1239, 197–217. [PubMed: 25408407]

Cuddapah S, Jothi R, Schones DE, Roh T-Y, Cui K, and 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. [PubMed: 19056695]

Daxinger L, Harten SK, Oey H, Epp T, Isbel L, Huang E, Whitelaw N, Apedaile A, Sorolla A, Yong J, et al. (2013). An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. Genome Biol 14, R96. [PubMed: 24025402]

de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, Splinter E, Wijchers PJ, Krijger PHL, and de Laat W (2015). CTCF binding polarity determines chromatin looping. Mol. Cell 60, 676–684. [PubMed: 26527277]

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, and Ren B (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380. [PubMed: 22495300]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Dowen JM, Bilodeau S, Orlando DA, Hübner MR, Abraham BJ, Spector DL, and Young RA (2013). Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. Stem Cell Reports 1, 371–378. [PubMed: 24286025]

Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schujiers J, Lee TI, Zhao K, and Young RA (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell 159, 374–387. [PubMed: 25303531]

Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, and Aiden EL (2016a). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst 3, 95–98. [PubMed: 27467249]

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, and Aiden EL (2016b). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst 3, 99–101. [PubMed: 27467250]

Eagen KP (2018). Principles of chromosome architecture revealed by Hi-C. Trends Biochem. Sci 43, 469–478. [PubMed: 29685368]

Gibcus JH, and Dekker J (2013). The hierarchy of the 3D genome. Mol. Cell 49, 773–782. [PubMed: 23473598]

Gorkin DU, Leung D, and Ren B (2014). The 3D genome in transcriptional regulation and pluripotency. Cell Stem Cell 14, 762–775. [PubMed: 24905166]

Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Medema RH, van Steensel B, et al. (2017). The Cohesin release factor WAPL restricts chromatin loop extension. Cell 169, 693–707.e14. [PubMed: 28475897]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. [PubMed: 20513432]

Hsu SC, Gilgenast TG, Bartman CR, Edwards CR, Stonestrom AJ, Huang P, Emerson DJ, Evans P, Werner MT, Keller CA, et al. (2017). The BET protein BRD2 cooperates with CTCF to enforce transcriptional and architectural boundaries. Mol. Cell 66, 102–116.e7. [PubMed: 28388437]

Hu G, Kim J, Xu Q, Leng Y, Orkin SH, and Elledge SJ (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev 23, 837–848. [PubMed: 19339689]

Ing-Simmons E, Seitan VC, Faure AJ, Flicek P, Carroll T, Dekker J, Fisher AG, Lenhard B, and Merkenschlager M (2015). Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. Genome Res 25, 504–513. [PubMed: 25677180]

Isbel L, Prokopuk L, Wu H, Daxinger L, Oey H, Spurling A, Lawther AJ, Hale MW, and Whitelaw E (2016). Wiz binds active promoters and CTCF-binding sites and is required for normal behaviour in the mouse. eLife 5, e15082. [PubMed: 27410475]

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. [PubMed: 20720539]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, and Haussler D (2002). The human genome browser at UCSC. Genome Res 12, 996–1006. [PubMed: 12045153]

Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RDD, Zhang MQ, Lobanenkov VV, and Ren B (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231–1245. [PubMed: 17382889]

Lachner M, O'Carroll D, Rea S, Mechtler K, and Jenuwein T (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410, 116–120. [PubMed: 11242053]

Langmead B, Trapnell C, Pop M, and Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25. [PubMed: 19261174]

Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. [PubMed: 19451168]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, and Durbin R; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. [PubMed: 19505943]

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. [PubMed: 19815776]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. [PubMed: 25516281]

Merkenschlager M, and Nora EP (2016). CTCF and Cohesin in genome folding and transcriptional gene regulation. Annu. Rev. Genomics Hum. Genet 17, 17–43. [PubMed: 27089971]

Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, and Ait-Si-Ali S (2014). The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Mol. Cell 53, 277–289. [PubMed: 24389103]

Mullen AC, Orlando DA, Newman JJ, Lové J, Kumar RM, Bilodeau S, Reddy J, Guenther MG, DeKoter RP, and Young RA (2011). Master transcription factors determine cell-type-specific responses to TGF-β signaling. Cell 147, 565–576. [PubMed: 22036565]

Nakahashi H, Kieffer Kwon KR, Resch W, Vian L, Dose M, Stavreva D, Hakim O, Pruett N, Nelson S, Yamane A, et al. (2013). A genome-wide map of CTCF multivalency redefines the CTCF code. Cell Rep 3, 1678–1689. [PubMed: 23707059]

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. [PubMed: 22495304]

Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, and Bruneau BG (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell 169, 930–944.e22. [PubMed: 28525758]

Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, and Guenther MG (2014). Quantitative ChIP-seq normalization reveals global modulation of the epigenome. Cell Rep 9, 1163–1170. [PubMed: 25437568]

Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. [PubMed: 20110278]

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–W165. [PubMed: 27079975]

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, and Aiden EL (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680. [PubMed: 25497547]

Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. (2017). Cohesin loss eliminates all loop domains. Cell 171, 305–320.e24. [PubMed: 28985562]

Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481, 389–393. [PubMed: 22217937]

Rowley MJ, and Corces VG (2018). Organizational principles of 3D genome architecture. Nat. Rev. Genet 19, 789–800. [PubMed: 30367165]

Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, Lenhard B, Giorgetti L, Heard E, Fisher AG, et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res 23, 2066–2077. [PubMed: 24002784]

Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, and Ren B (2012). A map of the cis-regulatory sequences in the mouse genome. Nature 488, 116–120. [PubMed: 22763441]

Simon JM, Parker JS, Liu F, Rothbart SB, Ait-Si-Ali S, Strahl BD, Jin J, Davis IJ, Mosley AL, and Pattenden SG (2015). A role for Widely Interspaced Zinc Finger (WIZ) in retention of the G9a methyltransferase on chromatin. J. Biol. Chem 290, 26088–26102. [PubMed: 26338712]

Sofueva S, Yaffe E, Chan WC, Georgopoulou D, Vietri Rudan M, Mira-Bontenbal H, Pollard SM, Schroth GP, Tanay A, and Hadjur S (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J 32, 3119–3129. [PubMed: 24185899]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A 102, 15545–15550. [PubMed: 16199517]

Sun F, Chronis C, Kronenberg M, Chen XF, Su T, Lay FD, Plath K, Kurdistani SK, and Carey MF (2019). Promoter-enhancer communication occurs primarily within insulated neighborhoods. Mol. Cell 73, 250–263.e5. [PubMed: 30527662]

Ueda J, Tachibana M, Ikura T, and Shinkai Y (2006). Zinc finger protein Wiz links G9a/GLP histone methyltransferases to the co-repressor molecule CtBP. J. Biol. Chem 281, 20120–20128. [PubMed: 16702210]

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. (2015). Tissue-based map of the human proteome. Science 347, 1260419. [PubMed: 25613900]

Uusküla-Reimand L, Hou H, Samavarchi-Tehrani P, Rudan MV, Liang M, Medina-Rivera A, Mohammed H, Schmidt D, Schwalie P, Young EJ, et al. (2016). Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. Genome Biol 17, 182. [PubMed: 27582050]

Viny AD, Ott CJ, Spitzer B, Rivas M, Meydan C, Papalexi E, Yelin D, Shank K, Reyes J, Chiu A, et al. (2015). Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. J. Exp. Med 212, 1819–1832. [PubMed: 26438361]

Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, et al. (2017). YY1 is a structural regulator of enhancer-promoter loops. Cell 171, 1573–1588.e28. [PubMed: 29224777]

Wen Z, Huang Z-T, Zhang R, and Peng C (2018). ZNF143 is a regulator of chromatin loop. Cell Biol. Toxicol 34, 471–478. [PubMed: 30120652]

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, and Young RA (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319. [PubMed: 23582322]

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis Using the Grammar of Graphics (Springer-Verlag).

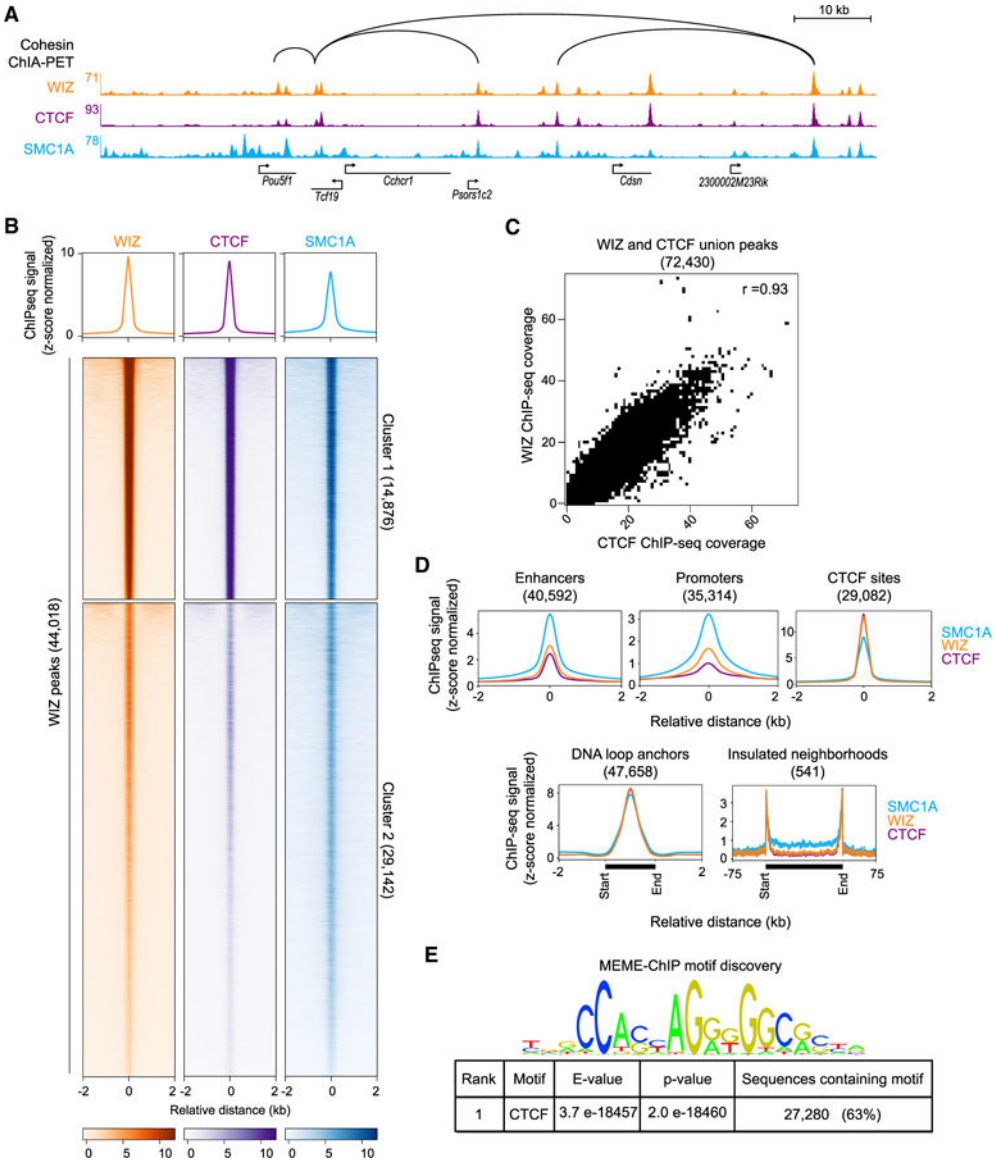
Wu H, Chen X, Xiong J, Li Y, Li H, Ding X, Liu S, Chen S, Gao S, and Zhu B (2011). Histone methyltransferase G9a contributes to H3K27 methylation in vivo. Cell Res 21, 365–367. [PubMed: 21079650]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Modelbased analysis of ChIP-Seq (MACS). Genome Biol 9, R137. [PubMed: 18798982]

Zuin J, Dixon JR, van der Reijden MIJA, Ye Z, Kolovos P, Brouwer RWW, van de Corput MPC, van de Werken HJG, Knoch TA, van IJcken WFJ, et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. U S A 111, 996–1001. [PubMed: 24335803]

Highlights

- WIZ generally colocalizes with CTCF and cohesin across the genome
- Loss of WIZ increases cohesin occupancy and DNA loops
- WIZ maintains proper gene expression and stem cell identity



Rank	Motif	E-value	p-value	Sequences containing motif
1	CTCF	3.7 e-18457	2.0 e-18460	27,280 (63%)

Figure 1. WIZ Occupies Enhancers, Promoters, Insulators, and DNA Loop Anchors across the Embryonic Stem Cell Genome

(A) Genome Browser tracks showing ChIP-seq signal for WIZ, CTCF, and SMC1A. High-confidence SMC1A ChIA-PET interactions are depicted as black lines (Dowen et al., 2014).

(B) Average signal plots and clustered heatmaps displaying WIZ, CTCF, and SMC1A ChIP-seq signal (*Z* score normalized) at WIZ peaks.

(C) Correlation of WIZ and CTCF ChIP-seq signal (*Z* score normalized) at a union set of peaks (Pearson correlation r = 0.93).

(D) Average signal plots showing the occupancy of WIZ, CTCF, and SMC1A at enhancers, promoters, CTCF sites, DNA loop anchors from cohesin ChIA-PET, and insulated neighborhoods.

(E) MEME-ChIP motif discovery identifies the CTCF consensus motif as the top motif present within WIZ peaks.

See also Figure S1 and Table S2. See STAR Methods for detailed description of genomics analyses. Datasets used in this figure are listed in Table S1.

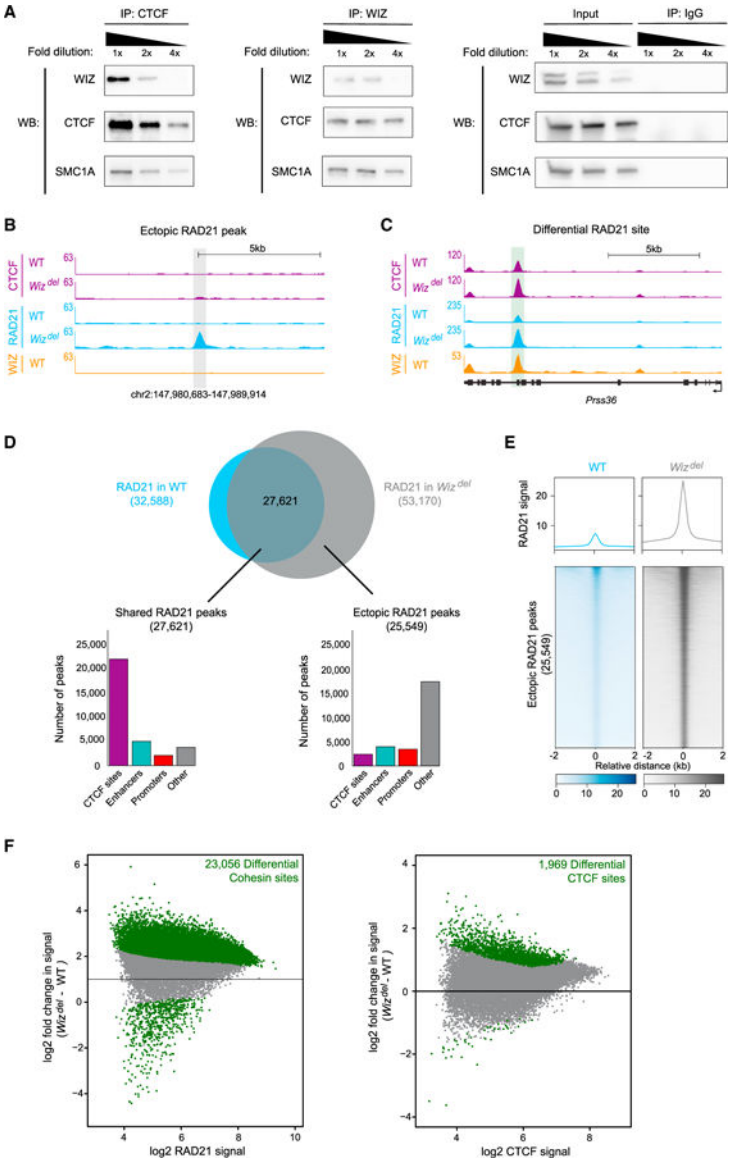


Figure 2. WIZ Forms a Complex with CTCF and Cohesin

(A) Western blot analysis showing co-immunoprecipitation of WIZ, CTCF, and SMC1A, as well as IgG controls from nuclear lysates.

(B) Genome Browser tracks showing CTCF and RAD21 occupancy in wild-type and *Wiz*^{del} cells at an ectopic RAD21 peak. WIZ occupancy in wild-type cells is shown.

(C) Genome Browser tracks showing CTCF and RAD21 occupancy in wild-type and *Wiz*^{del} cells at a differential RAD21 site. WIZ occupancy in wild-type cells is shown.

(D) Overlap of RAD21 peaks in wild-type and *Wiz*^{del} cells. For shared RAD21 peaks and ectopic RAD21 peaks, the overlap with functional elements in the genome is shown (CTCF sites, enhancers, promoters, other).

(E) Average signal plots and heatmaps of RAD21 signal in wild-type and *Wiz*^{del} cells at 25,549 ectopic RAD21 peaks in *Wiz*^{del} cells.

(F) MA plots showing differential enrichment of RAD21 and CTCF between wild-type and *Wiz*^{del} cells. Sites of significantly differential enrichment are shown in green.
See also Figure S2 and Table S2. See STAR Methods for detailed description of genomics analyses. Datasets used in this figure are listed in Table S1.

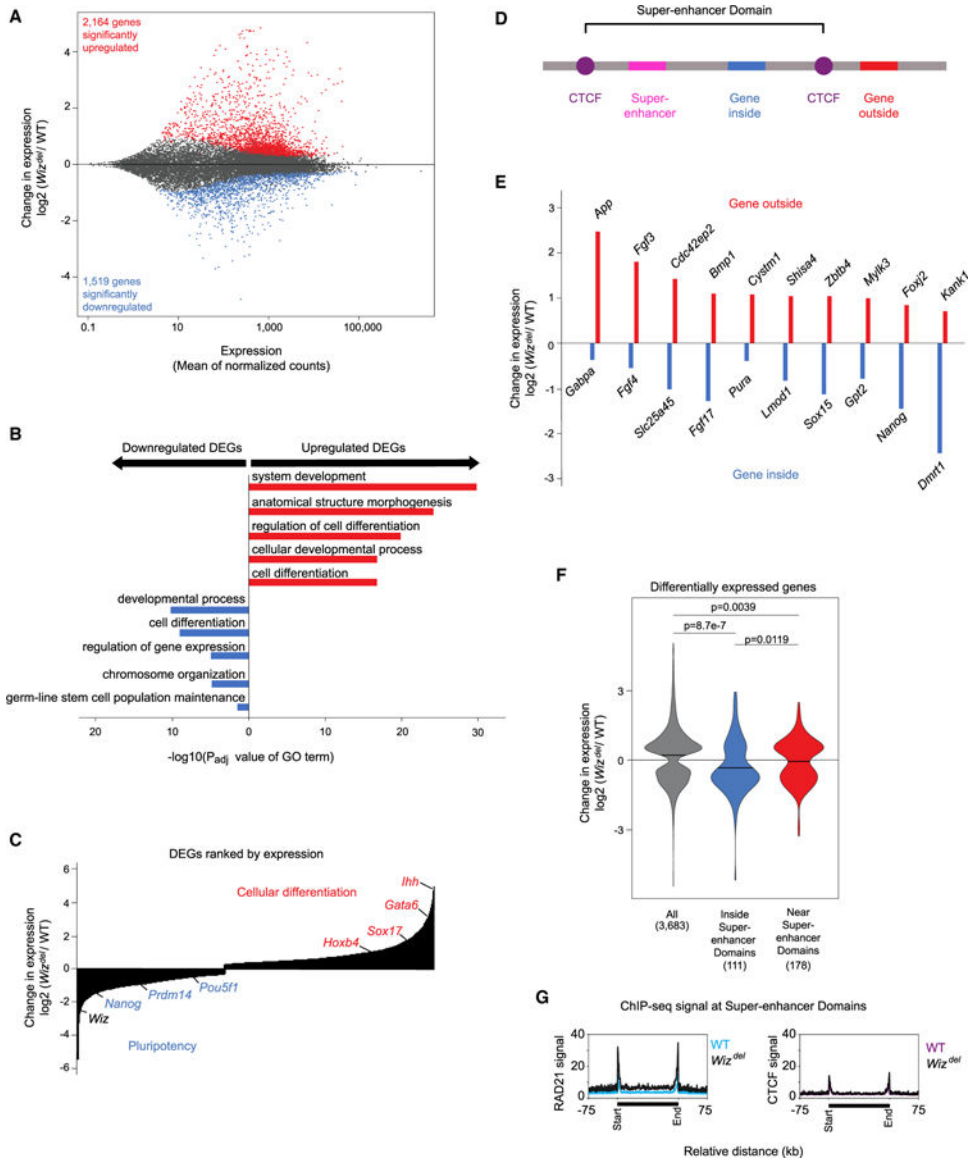


Figure 3. WIZ Is Required for Proper Gene Expression

(A) Changes in gene expression from RNA-seq in *Wiz*^{del} cells versus wild-type cells. Genes with significant changes in expression (false discovery rate [FDR]-adjusted $p < 0.1$) are colored with upregulated genes shown in red and downregulated genes in blue.

(B) Gene Ontology analysis identifies misregulated biological processes in *Wiz*^{del} cells that are involved in stem cell identity and differentiation.

(C) Differentially expressed genes ranked by $\log_2$ fold change with key pluripotency and cell identity markers indicated.

(D) Model depicting a Super-enhancer Domain where transcriptional insulation may occur.

(E) Change in gene expression at ten Super-enhancer Domains from RNA-seq in *Wiz*^{del} cells versus wild-type cells.

(F) Change in gene expression of DEGs located inside Super-enhancer Domains and within 150 kb of a Super-enhancer Domain.

(G) Average signal plots showing RAD21 and CTCF signal in wild-type and *Wiz*^{del} cells at Super-enhancer Domains.

See also Figure S3, Table S3, and Table S4. See STAR Methods for detailed description of genomics analyses. Datasets used in this figure are listed in Table S1.

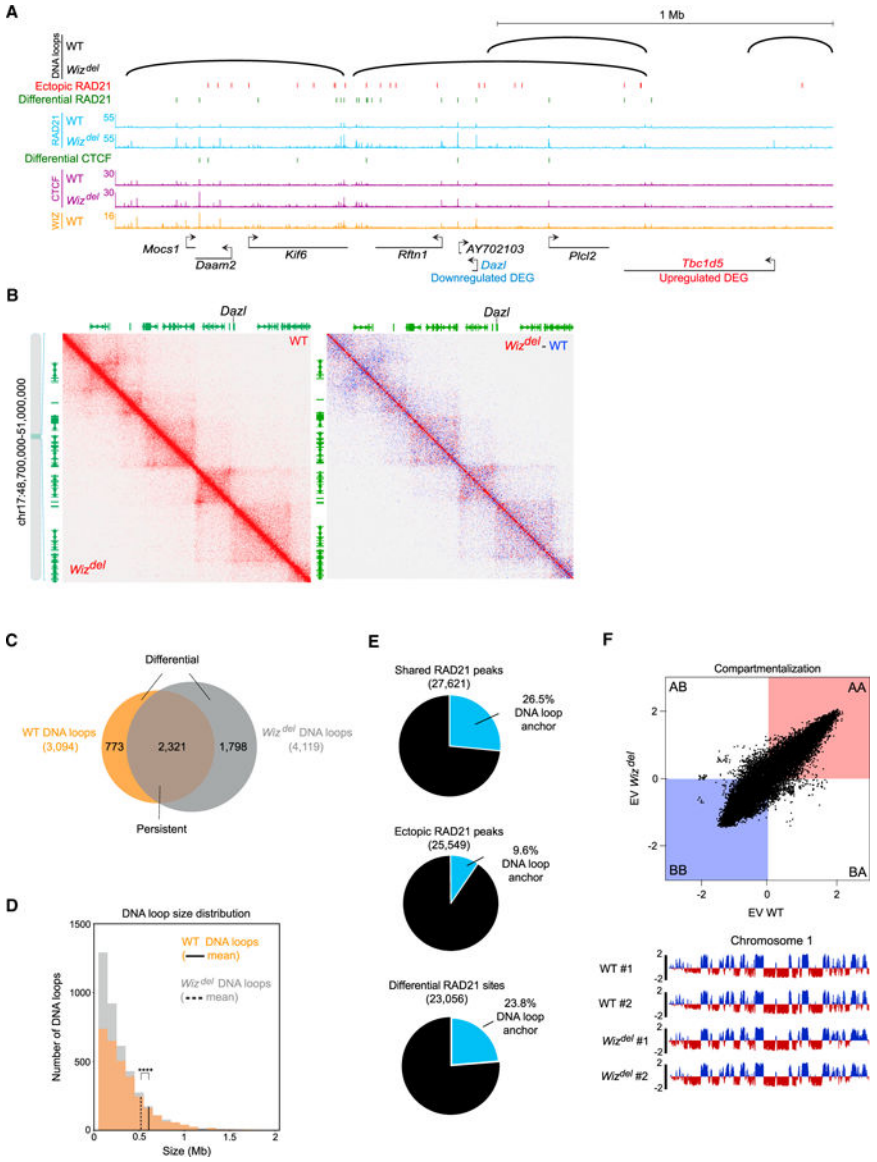


Figure 4. WIZ Is Important for DNA Loop Architecture of the Genome

(A) Genome Browser tracks showing DNA loops, cohesin (RAD21) occupancy, and CTCF occupancy in wild-type cells and *Wiz*^{del} cells. DNA loops were identified using HiCCUPS. WIZ occupancy in wild-type cells is shown.

(B) Hi-C maps showing signal in wild-type and *Wiz*^{del} cells (left) at the *Dazl* locus at 5 kb resolution. Differential signal between wild-type and *Wiz*^{del} cells shown on the right.

(C) Venn diagram showing both persistent and differential DNA loops between wild-type and *Wiz*^{del} cells.

(D) Size distribution of DNA loops in wild-type and *Wiz*^{del} cells. To compare the means of the distributions, a Wilcoxon rank-sum test was performed; **** represents an adjusted p value of 4e-18.

(E) Proportion of shared RAD21 peaks, ectopic RAD21 peaks and differential RAD21 sites that overlap the anchor of a DNA loop detected in wild-type or *Wiz*^{del} cells.

(F) Compartmentalization of the genome into A and B compartments on the basis of eigenvector (EV) score, represented as wild-type versus Wiz^{del} cells. Eigenvector track of chromosome 1 is shown for each replicate.

See also Figure S4 and Table S5. See STAR Methods for detailed description of genomics analyses. Datasets used in this figure are listed in Table S1.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
WIZ antibody	Novus Biologicals	Cat#NBP1-80586; RRID: AB_11011659
CTCF antibody	Active Motif	Cat#61311; RRID: AB_2614975
SMC1A antibody	Bethyl Laboratories, Inc	Cat#A300-055A; RRID: AB_2192467
RAD21 antibody	Bethyl Laboratories, Inc	Cat#A300-080A; RRID: AB_2176615
IgG antibody	Bethyl Laboratories, Inc	Cat#P120-101; RRID: AB_479829
H3 antibody	Abcam	Cat#ab1791; RRID: AB_302613
MYC tag antibody	Abcam	Cat#ab9106; RRID: AB_307014
Chemicals, Peptides, and Recombinant Proteins		
cOmplete Protease Inhibitor Cocktail	Sigma Aldrich	Cat#11697498001
DMEM, high glucose, pyruvate	GIBCO	Cat#11995065
HyClone Cosmic Calf Serum	Thermo Fisher Scientific	Cat#SH3008703
GlutaMAX Supplement	Thermo Fisher Scientific	Cat#35050-061
Penicillin-Streptomycin (10,000 U/mL)	Thermo Fisher Scientific	Cat#15140-122
Lipofectamine 2000 Transfection Reagent	Invitrogen	Cat#11-668-027
Formaldehyde Solution	Sigma Aldrich	Cat#F1635
Protein G Dynabeads	Thermo Fisher Scientific	Cat#10004D
Pierce 16% Formaldehyde, Methanol-free	Thermo Fisher Scientific	Cat#28906
DC Protein Assay Kit II	BioRad	Cat#5000112
Blotting-Grade Blocker	BioRad	Cat#1706404
SuperSignal West Pico PLUS Chemiluminescent Substrate	Thermo Fisher Scientific	Cat#34577
SuperSignal West Femto Maximum Sensitivity Substrate	Thermo Fisher Scientific	Cat#34094
TRIzol Reagent	Invitrogen	Cat#15596018
Chloroform	Sigma Aldrich	Cat#C2432
AMPure XP beads	Beckman Coulter	Cat#A63881
Proteinase K	New England BioLabs	Cat#P81075
Ribonuclease A (RNase A) from Bovine Pancreas	Sigma Aldrich	Cat#R4642
KnockOut DMEM	Thermo Fisher Scientific	Cat#10829-018
Premium Grade Fetal Bovine Serum (FBS)	VWR	Cat#97068-085 Lot#249B17
Critical Commercial Assays		
TOPO TA Cloning Kit for Sequencing	Invitrogen	Cat#K4575J10
ThruPLEX DNA-seq Kit	Takara Bio	Cat#R400428
Zymo ChIP DNA Clean & Concentrator	Zymo Research	Cat#D5205
NEBNext Ultra II DNA Library Prep Kit	New England BioLabs	Cat#E7645S

REAGENT or RESOURCE	SOURCE	IDENTIFIER
KAPA HyperPrep Kit	Roche/Kapa	Cat#KK8502
Re-ChIP-IT	Active Motif	Cat#53016
Nuclear Complex Co-IP Kit	Active Motif	Cat#54001
RNA Clean and Concentrator Kit	Zymo Research	Cat#R1013
SuperScript IV First Strand Synthesis System	Invitrogen	Cat#18091050
TruSeq RNA Library Prep Kit v2	Illumina	Cat#RS-122-2001
Arima-HiC Kit	Arima Genomics	N/A (This is the only product at of publication.) company's the time
ZymoPURE II Plasmid Midiprep Kit	Zymo Research	Cat#D4200
Deposited Data		
Calibrated (Spike-In) ChIP-seq	This study.	GEO: GSE137285
Hi-C	This study.	GEO: GSE137285
RNA-seq	This study.	GEO: GSE137285
OCT4 ChIP-seq	Whyte et al. (2013). PMID: 23582322	GEO: GSE44286
SOX2 ChIP-seq	Whyte et al. (2013). PMID: 23582322	GEO: GSE44286
NANOG ChIP-seq	Whyte et al. (2013). PMID: 23582322	GEO: GSE44286
WIZ ChIP-seq (Cerebellum)	Isbel et al. (2016). PMID: 27410475	GEO: GSE76909
CTCF ChIP-seq (Cerebellum)	Shen et al. (2012). PMID: 22763441	GEO: GSE23830
H3 ChIP-seq	Mullen et al. (2011). PMID: 22036565	GEO: GSE23830
CTCF ChIP-seq (mESC)	Nora et al. (2017). PMID: 28525758	GEO: GSE98671
SMC1 ChIA-PET	Dowen et al. (2014). PMID: 25303531	GEO: GSE57911
mESC RNA-seq	Mozzetta et al. (2014). PMID: 24389103	GEO: GSE49669
G9a$^{-/-}$ mESC RNA-seq	Mozzetta et al. (2014). PMID: 24389103	GEO: GSE49669
Experimental Models: Cell Lines		
Murine Embryonic Stem Cells (mESC) v6.5	Laboratory of Dr. Richard Young	N/A
Human Embryonic Kidney Cells (HEK293T)	Laboratory of Dr. Richard Young	N/A
Wiz^{del} Mouse Embryonic Stem Cells (mESC)	This study	N/A
Oligonucleotides		
RT-qPCR Primers	This Study	See Table S6
ChIP-qPCR Primers	This Study	See Table S6
Software and Algorithms		
Custom ChIP-seq Processing Script	This Study	https://github.com/dowenlab
Bowtie	Langmead et al. (2009). http://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-3-r25	v1.2.2
Samtools	Li et al. (2009)	v1.9
Bedtools	Quinlan and Hall, 2010	v2.26
MACS	Zhang et al. (2008)	v2016-02-15
UCSC Tools	Kent et al. (2002). https://doi.org/10.1101/gr.229102	v320
DeepTools	Ramírez et al. (2016)	v3.0.1

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MEME Suite	Bailey et al. (2009)	N/A
DiffBind	Ross-Innes et al. (2012)	N/A
STAR Aligner	Dobin et al. (2013)	v2.6.0a
DESeq2	Love et al. (2014). http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8	N/A
ggpubr, ggplot2	Wickham (2016)	N/A
Juicer/Juicebox	Durand et al., 2016b, 2016a. https://doi.org/10.1016/j.cels.2016.07.002; https://doi.org/10.1016/j.cels.2015.07.012	N/A
Arrowhead, HiCCUPs	Rao et al. (2014). https://doi.org/10.1016/j.cell.2014.11.021	N/A
HOMER	Heinz et al. (2010). https://doi.org/10.1016/j.molcel.2010.05.004	N/A
BWA (Burrows-Wheeler Aligner)	Li and Durbin (2009)	N/A
GSEA (Gene Set Enrichment Analysis)	Subramanian et al. (2005). https://doi.org/10.1073/pnas.0506580102	N/A
Other		
4-20% Mini-PROTEAN TGX Stain-Free Protein Gels	BioRad	Cat#4568094
FluoroTrans Transfer Membranes	VWR	Cat#29301-856