

## MIT Open Access Articles

*Transcriptional Dysregulation of MYC Reveals  
Common Enhancer-Docking Mechanism*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Schuijers, Jurian, John Colonnese Manteiga, Abraham Selby Weintraub, Daniel Sindt Day, Alicia Viridiana Zamudio, Denes Hnisz, Tong Ihn Lee, and Richard Allen Young. "Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism." *Cell Reports* 23, no. 2 (April 2018): 349–360. <sup>a</sup> 2018 The Author(s).

**As Published:** <http://dx.doi.org/10.1016/j.celrep.2018.03.056>

**Publisher:** Elsevier

**Persistent URL:** <http://hdl.handle.net/1721.1/117042>

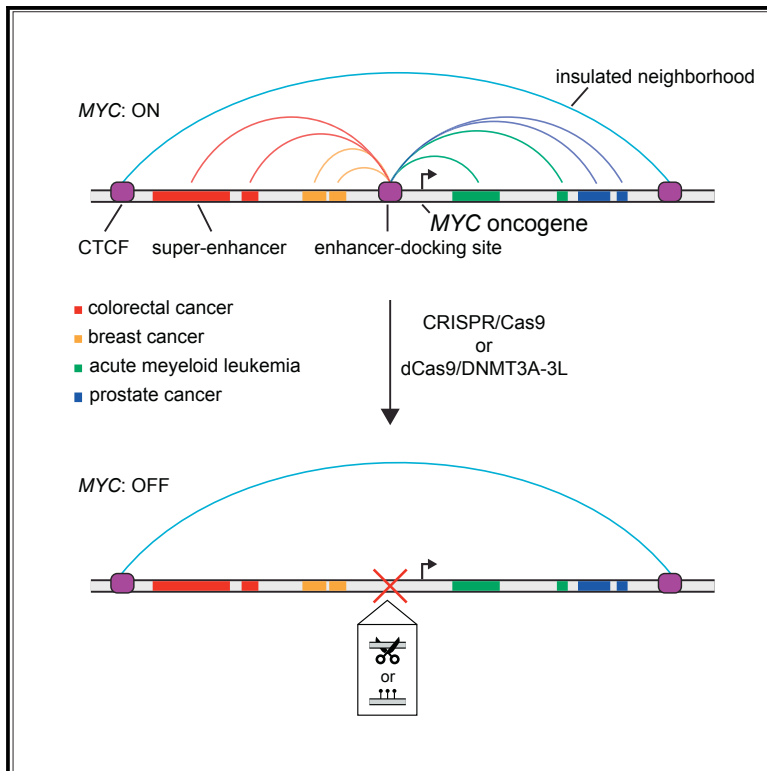
**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** Creative Commons Attribution 4.0 International License



## Transcriptional Dysregulation of *MYC* Reveals Common Enhancer-Docking Mechanism

### Graphical Abstract



### Authors

Jurian Schuijers,  
John Colonnese Manteiga,  
Abraham Selby Weintraub, ...,  
Denes Hnisz, Tong Ihn Lee,  
Richard Allen Young

### Correspondence

young@wi.mit.edu

### In Brief

Schuijers et al. show that a conserved CTCF site at the promoter of the *MYC* oncogene plays an important role in enhancer-promoter looping with tumor-specific super-enhancers. Perturbation of this site provides a potential therapeutic vulnerability.

### Highlights

- Tumor-specific super-enhancers loop to an enhancer-docking site at the *MYC* oncogene
- Deletion or methylation of *MYC* enhancer-docking site reduces enhancer-promoter looping
- *MYC* enhancer-docking site provides common vulnerability in diverse cancer cells
- Similar docking sites for multiple enhancers occur at additional genes

### Data and Software Availability

GSE92881



# Transcriptional Dysregulation of *MYC* Reveals Common Enhancer-Docking Mechanism

Jurian Schuijers,<sup>1,3</sup> John Colonnese Manteiga,<sup>1,2,3</sup> Abraham Selby Weintraub,<sup>1,2</sup> Daniel Sindt Day,<sup>1</sup> Alicia Viridiana Zamudio,<sup>1,2</sup> Denes Hnisz,<sup>1</sup> Tong Ihn Lee,<sup>1</sup> and Richard Allen Young<sup>1,2,4,\*</sup>

<sup>1</sup>Whitehead Institute for Biomedical Research, 455 Main Street, Cambridge, MA 02142, USA

<sup>2</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

<sup>3</sup>These authors contributed equally

<sup>4</sup>Lead Contact

\*Correspondence: [young@wi.mit.edu](mailto:young@wi.mit.edu)

<https://doi.org/10.1016/j.celrep.2018.03.056>

## SUMMARY

Transcriptional dysregulation of the *MYC* oncogene is among the most frequent events in aggressive tumor cells, and this is generally accomplished by acquisition of a super-enhancer somewhere within the 2.8 Mb TAD where *MYC* resides. We find that these diverse cancer-specific super-enhancers, differing in size and location, interact with the *MYC* gene through a common and conserved CTCF binding site located 2 kb upstream of the *MYC* promoter. Genetic perturbation of this enhancer-docking site in tumor cells reduces CTCF binding, super-enhancer interaction, *MYC* gene expression, and cell proliferation. CTCF binding is highly sensitive to DNA methylation, and this enhancer-docking site, which is hypomethylated in diverse cancers, can be inactivated through epigenetic editing with dCas9-DNMT. Similar enhancer-docking sites occur at other genes, including genes with prominent roles in multiple cancers, suggesting a mechanism by which tumor cell oncogenes can generally hijack enhancers. These results provide insights into mechanisms that allow a single target gene to be regulated by diverse enhancer elements in different cell types.

## INTRODUCTION

Elevated expression of the c-MYC transcription factor occurs in a broad spectrum of human cancers and is associated with tumor aggression and poor clinical outcome (Berns et al., 1992; Dang, 2012; Gabay et al., 2014; Grotzer et al., 2001). Activation of the *MYC* gene, which encodes c-MYC, is a hallmark of cancer initiation and maintenance. Dysregulation of *MYC* is often achieved through the formation of large tumor-specific super-enhancers in the region surrounding the *MYC* gene (Chapuy et al., 2013; Fulco et al., 2016; Herranz et al., 2014; Hnisz et al., 2013; Lin et al., 2016; Liu et al., 2015; Lovén et al., 2013; Shi et al., 2013; Whyte et al., 2013; Zhang et al., 2016). These large enhancer clusters differ in size, composition, and distance from the *MYC* promoter, yet all accomplish the same task of

stimulating *MYC* overexpression across a broad spectrum of tumors.

Selective gene activation is essential to the gene expression programs that define both normal and cancer cells. During gene activation, transcription factors (TFs) bind enhancer elements and regulate transcription from the promoters of nearby or distant genes through physical contacts that involve looping of DNA between enhancers and promoters (Bonev and Cavalli, 2016; Buecker and Wysocka, 2012; Bulger and Groudine, 2011; Fraser et al., 2015; Müller-Storm et al., 1989; Spitz, 2016; de Wit et al., 2013). The mechanisms that ensure that specific enhancers interact with specific promoters are not fully understood. Some enhancer-promoter interactions are likely determined by the nature of TFs bound at the two sites (Muerdter and Stark, 2016; Weintraub et al., 2017).

Recent studies have revealed that specific chromosome structures play important roles in gene control. Enhancer-promoter interactions generally occur within larger chromosomal loop structures formed by the interaction of CTCF proteins bound to each of the loop anchors (Dekker and Mirny, 2016; Fraser et al., 2015; Gibcus and Dekker, 2013; Gorkin et al., 2014a; Hnisz et al., 2016a, 2018; Ji et al., 2016). These loop structures, variously called topologically associated domains (TADs), sub-TADs, loop domains, CTCF contact domains, and insulated neighborhoods, tend to insulate enhancers and genes within the CTCF-CTCF loops from elements outside those loops (Dixon et al., 2012; Downen et al., 2014; Franke et al., 2016; Hnisz et al., 2016a, 2016b; Ji et al., 2016; Narendra et al., 2015; Nora et al., 2012; Phillips-Cremins et al., 2013; Rao et al., 2014). Constraining DNA interactions within CTCF-CTCF loop structures in this manner may facilitate proper enhancer-promoter contacts.

CTCF does not generally occupy enhancer and promoter elements (Cuddapah et al., 2009; Dixon et al., 2012; Downen et al., 2014; Handoko et al., 2011; Ji et al., 2016; Kim et al., 2007; Phillips-Cremins et al., 2013; Rao et al., 2014; Rubio et al., 2008; Tang et al., 2015). Another TF, YY1, generally binds to enhancers and promoters and facilitates their interaction through YY1 dimerization (Weintraub et al., 2017). However, when CTCF does bind these regulatory elements, it can also contribute to enhancer-promoter interactions (Banani et al., 2017; Nora et al., 2017; Splinter et al., 2006; Zuin et al., 2014).

Here, we investigate DNA looping structures in the *MYC* locus in multiple cancers and identify a CTCF-occupied site at the *MYC* promoter that facilitates docking with essentially any



enhancers that are formed within the 2.8 Mb *MYC* locus. The CTCF-occupied site at the *MYC* promoter, which we call the *MYC* enhancer-docking site, can be abrogated by genetic and epigenetic editing. Similar enhancer-docking sites occur at other oncogenes. This suggests a mechanism by which tumor cells can generally hijack enhancers and, with editing, a potential therapeutic vulnerability.

## RESULTS

### Cell-Type-Specific *MYC* Enhancers Loop to a Common Upstream CTCF Site

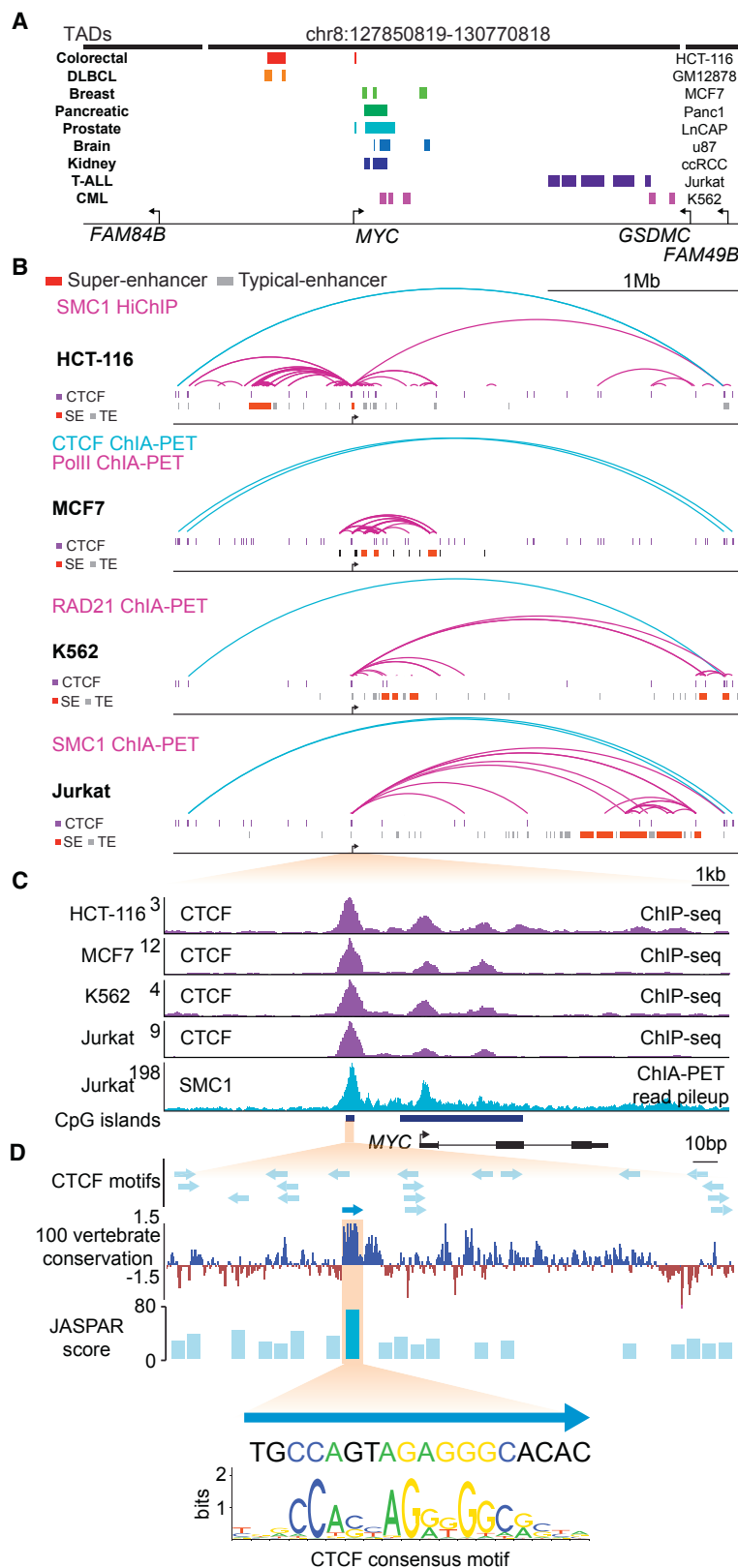
Previous studies have established that tumor cells acquire tumor-specific super-enhancers at various sites throughout the *MYC* locus (Figures 1A and S1A) (Bradner et al., 2017; Chapuy et al., 2013; Gabay et al., 2014; Gröschel et al., 2014; Herranz et al., 2014; Hnisz et al., 2013; Lin et al., 2016; Lovén et al., 2013; Parker et al., 2013; Zhang et al., 2016; Shi et al., 2013), but the mechanisms by which these diverse enhancer structures control *MYC* are not clear. In one case, for example, a super-enhancer located ~2 Mb downstream of the *MYC* gene has been shown to physically interact with *MYC*, but the mechanisms responsible for this specific interaction are unclear (Shi et al., 2013). To gain insights into the potential role of DNA loop structures in gene control at the *MYC* locus, we generated cohesin HiChIP data for HCT-116 cells and collected published DNA interaction data for three other cancer cell types for comparison (Figure 1B; Tables S1 and S5) (Hnisz et al., 2016a; Pope et al., 2014). Among the DNA loop structures observed in these datasets, a large 2.8 Mb DNA loop was evident in all four cell types. This loop connects CTCF sites encompassing the *MYC* gene and qualifies as an insulated neighborhood. The DNA anchor sites of this 2.8 Mb DNA loop occur at the boundaries of a TAD found in all cells (Figure S1B). The *MYC* TAD encompasses a region previously described as a “gene desert,” because this large span of DNA contains no other annotated protein-coding genes (Montavon and Duboule, 2012; Ovcharenko et al., 2005).

While all cells examined appear to share the TAD-spanning 2.8 Mb loop encompassing *MYC*, the loop structures within the neighborhood were found to be markedly different among the tumor types. The internal loops were dominated by interactions between a *MYC* promoter-proximal CTCF site and the diverse super-enhancers (Figures 1B and 1C). The major differences between these internal structures in the different tumor cells involved the different positions of the tumor-specific super-enhancer elements. Examination of Hi-C data for a broader spectrum of tumor cell types suggests that tumor cells generally have DNA contacts between the *MYC* promoter-proximal site and other sites within the 2.8 Mb *MYC* TAD (Figure S1B). This looping was not limited to cancer cells, because examination of enhancer and promoter-capture Hi-C data in a variety of normal cell types that express *MYC* (Javierre et al., 2016) revealed that cell-type-specific enhancers do indeed loop to the *MYC* proximal CTCF site (Figures S1C and S1D). This indicated that this CTCF site is also used during normal development by cell-type-specific enhancers to facilitate *MYC* expression and cellular proliferation.

Further examination of the *MYC* promoter-proximal region revealed three constitutive CTCF binding sites (Figure 1C). All three sites were found to be occupied by CTCF in a wide variety of normal cells and tumor cells, and this binding pattern is shared across species (Figure S1C). Previous studies have examined the role of CTCF binding at all three sites (Filippova et al., 1996; Gombert and Krumm, 2009; Gombert et al., 2003; Klenova et al., 1993; Rubio et al., 2008). The two sites located within the *MYC* gene have been shown to play roles in *MYC* transcript start site selection and in promoter-proximal pausing of RNA polymerase II (Filippova et al., 1996). The CTCF binding site located 2 kb upstream of the major transcription start site has been reported to protect the promoter from methylation and to be an insulator element (Gombert and Krumm, 2009; Gombert et al., 2003). The DNA interaction data described here, however, suggests that this upstream site dominates connections with distal enhancer elements, as the majority of reads in the DNA interaction data are associated with this site in all tumor cells examined (Figures 1C and S1E). The –2 kb CTCF binding site contains a number of putative CTCF binding motifs; one of these most closely matches the canonical CTCF motif in the JASPAR database (Sandelin et al., 2004) and occurs within a highly conserved sequence (Figure 1D). These features, the presence of CTCF sites in tumor super-enhancers, and the ability of two CTCF-bound sites to be brought together through CTCF homodimerization (Saldaña-Meyer et al., 2014; Yusufzai et al., 2004) led us to further study the possibility that the –2 kb site has an enhancer-docking function critical to *MYC* expression.

### *MYC* Promoter Proximal CTCF Site Is Necessary for Enhancer-Promoter Looping and High *MYC* Expression

To determine whether the putative enhancer-docking site plays a functional role in *MYC* expression through DNA loop formation, small perturbations of the CTCF binding site were generated in both alleles of the tumor cell lines K562, HCT-116, Jurkat, and MCF7 using clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 (Figures 2A and 2B). Attempts at genetic perturbation by transfection with constructs carrying CRISPR/Cas9 with a guide RNA specifically targeting the CTCF motif upstream of the *MYC* gene did not yield viable clones. To allow cells to continue to proliferate if the CTCF motif deletion was lethal, cells were virally transduced with an exogenous *MYC* gene driven by a phosphoglycerate kinase (PGK) promoter (Figure S2A). This construct contained sequence differences in the 3' UTR that allowed discrimination between the endogenous and exogenous *MYC* mRNAs. Cells expressing this exogenous *MYC* gene were then subjected to CRISPR/Cas9 perturbation. Clones were selected with small deletions or insertions disrupting the canonical CTCF motif (Figures 2B and S2B) and these cells were further characterized. CTCF chromatin immunoprecipitation quantitative polymerase chain reaction (ChIP-qPCR) showed complete loss of CTCF binding to this site in K562 and HCT-116 cells and a 60%–70% reduction in CTCF binding at this site in Jurkat and MCF7 cells (Figure 2C). RNA analysis revealed a 70%–80% reduction of endogenous *MYC* mRNA in the absence of the enhancer-docking site in all of these cell types (Figure 2D). Furthermore, an inducible CRISPR/Cas9



**Figure 1. Cell-Type-Specific Super-Enhancers in the MYC Locus Loop to a Common CTCF Site**

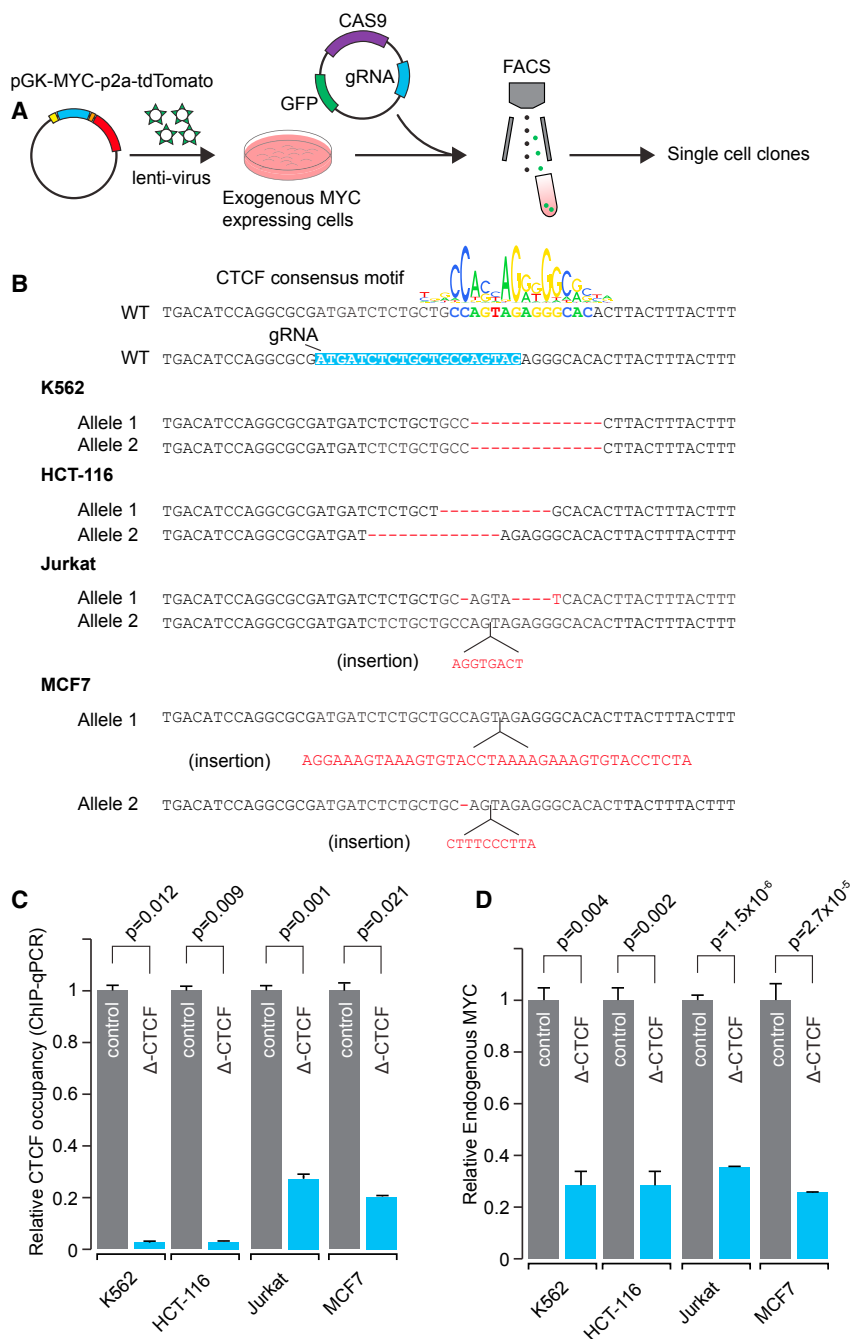
(A) The 4.5 Mb region surrounding the *MYC* gene. The 2.8 Mb TAD containing *MYC* and portions of the two adjacent TADs are indicated with thick black horizontal lines. Super-enhancers (data from Becket et al., 2016; Fietze et al., 2012; Lin et al., 2012; Pope et al., 2014; Wang et al., 2011) are shown in colored boxes for a panel of tumor cell lines that express *MYC*.

(B) Chromosome interaction data at the ~3 Mb *MYC* locus. For HCT-116, SMC1 HiChIP interactions with an origami score of at least 0.9 and a minimum PET count of 9 are shown as purple arcs; the insulated neighborhood spanning interaction, which encompasses the TAD, is shown as a blue arc (data from this study). For MCF7, Pol II ChIA-PET interactions with an origami score of at least 0.9 are shown as purple arcs; the insulated neighborhood spanning interactions from CTCF ChIA-PET data are shown in blue (data from ENCODE and Li et al., 2012). For K562, RAD21 ChIA-PET interactions with an origami score of at least 0.9 are shown as purple arcs; the insulated neighborhood spanning interaction is shown in blue and has an origami score of 0.44 (data from Heidari et al., 2014). For Jurkat, SMC1 ChIA-PET interactions with an origami score of at least 0.99 are shown as purple arcs; the insulated neighborhood spanning interactions are shown in blue (data from Hnisz et al., 2016a). CTCF ChIP-seq peaks are depicted as purple rectangles, super-enhancers are depicted as red rectangles, and typical enhancers are depicted as gray rectangles (data from this study; Hnisz et al., 2016a; Pope et al., 2014).

(C) CTCF ChIP-seq and SMC1 ChIA-PET read counts in the *MYC* promoter regions. Purple tracks display CTCF ChIP-seq signal in the four cell lines from (B). Light blue track displays the read counts from read pileup of Jurkat SMC1 ChIA-PET data, showing that the major peak of SMC1 ChIA-PET reads occurs at the -2 kb CTCF site. Dark blue bars indicate CpG islands. ChIP-seq read counts are shown in reads per million sequenced reads per base pair. ChIA-PET reads are shown as read counts per base pair.

(D) The top panel depicts all putative CTCF binding motifs as blue arrows, which indicate the orientation of the motif. The CTCF motif depicted in dark blue occurs in the most conserved region and shows the best match with consensus CTCF motif. 100 vertebrate conservation from the UCSC genome browser is depicted in the middle panel. The JASPAR score for all the motifs is indicated with blue bars. The position weight matrix for the canonical JASPAR CTCF motif and the actual sequence is shown below.

See also Figure S1.



**Figure 2. Perturbation of the Core CTCF Motif in the MYC CTCF Loop-Anchor Reduces CTCF Occupancy and MYC Expression**

(A) Schematic representation of the experiment. HCT-116, K562, Jurkat, and MCF7 cells were transduced with a construct expressing MYC under a PGK promoter and selected for successful integration. These cells were then transiently transfected with plasmid carrying Cas9 and a gRNA targeting the CTCF binding motif. Positive cells were identified and selected using fluorescence-activated cell sorting (FACS). These cells were multiplied, and clonal populations were characterized.

(B) The DNA sequences in the vicinity of the core CTCF motif and the mutations generated in clonal populations of K562, HCT-116, Jurkat, and MCF7 cell lines. The reference (WT, wild-type) sequence highlighted in blue is complementary to the gRNA sequence targeting the most prominent CTCF motif (shown here in bold colored sequence). For the aneuploid MCF7 cell line, the two most common mutations are depicted.

(C) ChIP-qPCR showing reduction of the CTCF occupancy in Δ-CTCF K562, HCT-116, Jurkat, and MCF7 cells. p values were generated with a Student's t test. Error bars represent the SD of the mean from three technical replicates.

(D) qPCR showing reduction of endogenous MYC mRNA levels in Δ-CTCF K562, HCT-116, Jurkat, and MCF7 cells. p values were generated with a Student's t test. Error bars represent the SD of the mean from three biological replicates.

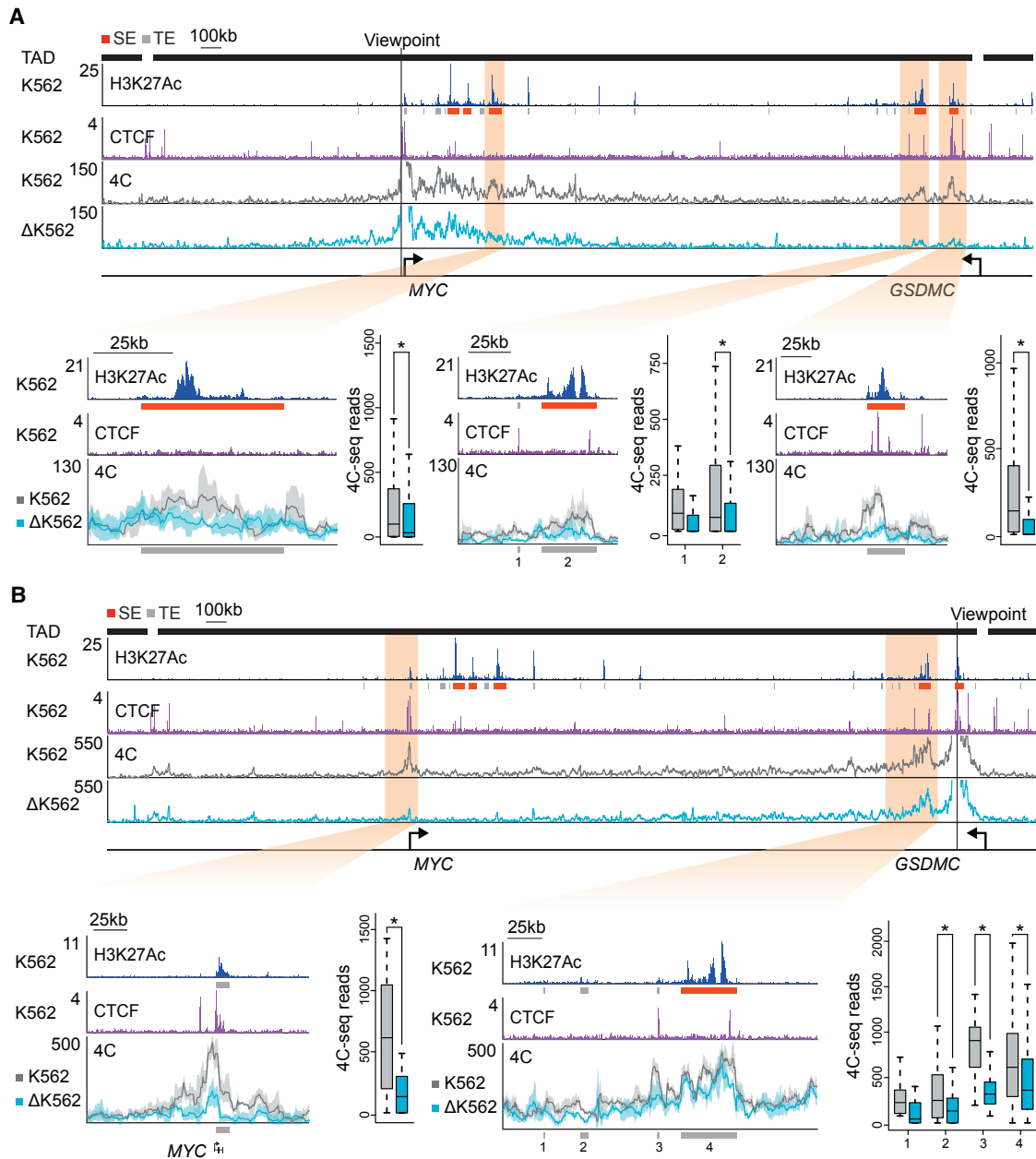
See also Figure S2.

perturbation model showed reduced proliferation for these four cell types upon induction of CTCF-site deletions (Figures S2C–S2G). These results indicate that the CTCF motif in the MYC enhancer-docking site is necessary for CTCF binding, for high levels of MYC expression and for cellular proliferation.

If the putative MYC enhancer-docking site contributes to looping interactions with distal enhancers, then the loss of this site should cause a decrease in DNA interactions between the MYC promoter and the distal super-enhancers. We used chromosome conformation capture combined with high-throughput

tions with the MYC enhancer-docking site as well as with the nearby super-enhancer, and these interactions were significantly reduced upon perturbation of the CTCF motif (Figures 3B and S3B). Similar results were obtained in HCT-116 cells, where the viewpoint was centered on the super-enhancer located ~0.4 Mb upstream of the MYC gene (Figure S3C). These results showed that the CTCF site in the promoter-proximal region of MYC is important for optimal interaction with distal enhancers and supports the idea that this CTCF site functions as an enhancer-docking site.



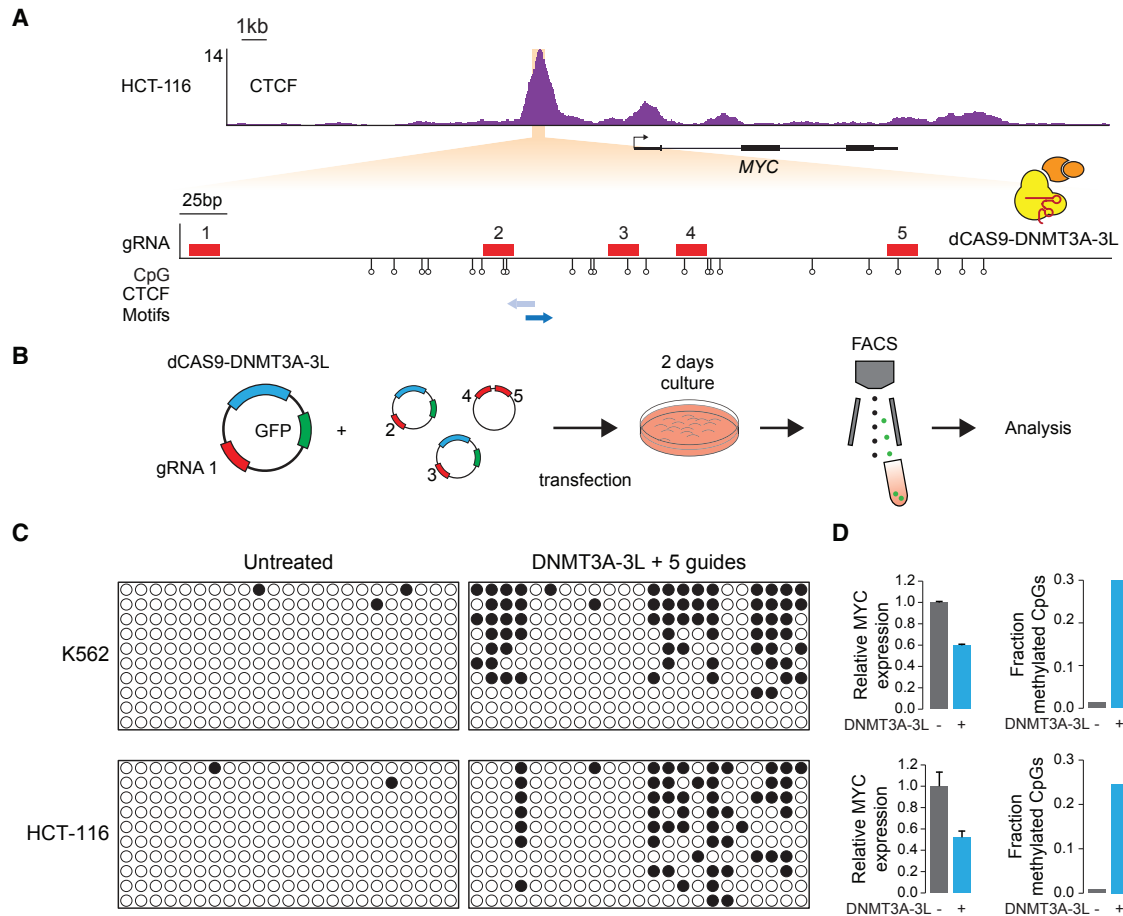


**Figure 3. Perturbation of the *MYC* Enhancer-Docking Site Reduces Looping to Super-Enhancers**

(A) 4C analysis showing reduced looping of the *MYC* promoter proximal CTCF site to super-enhancers in CTCF motif deletion cells ( $\Delta$ K562) versus unmodified cells (K562). H3K27Ac ChIP-seq and CTCF ChIP-seq are shown in blue and purple colors respectively. Blowups show the 4C interactions for three K562 specific super-enhancers. The 4C viewpoint is situated 112 base pairs upstream of the deleted loop-anchor region.

(B) 4C analysis showing reduced looping of the *MYC* promoter proximal CTCF site to super-enhancers in CTCF motif deletion cells ( $\Delta$ K562) versus unmodified cells (K562) using a viewpoint centered on the most distant super-enhancer downstream of the *MYC* gene. H3K27Ac ChIP-seq and CTCF ChIP-seq are shown in blue and purple, respectively. Blowups show the 4C interactions at the *MYC* promoter and distant super-enhancer near the viewpoint. Shading represents the 90% confidence interval based on three biological replicates. Peak calls from the H3K27Ac ChIP-seq were used to define the regions to be quantified and are indicated in gray boxes at the bottom of the panels. Boxplots show quantification of the reads per fragment for the indicated regions. p values were generated using Student's t test, and data pairs with a p value < 0.05 are indicated with an asterisk. Reads are shown in reads per million sequenced reads per base pair. Typical enhancers and super-enhancers are shown as gray boxes and red boxes, respectively.

See also Figure S3.



**Figure 4. dCas9-Mediated Methylation of the CTCF Loop-Anchor Site Reduces *MYC* Expression in Tumor Cells**

(A) Top panel shows CTCF ChIP-seq at the *MYC* gene region in HCT-116 cells. ChIP-seq reads are shown in reads per million sequenced reads per base pair. Bottom panels show a blowup of the ~700-bp region underneath the CTCF peak depicting the CTCF motifs (blue arrows) and the gRNAs (red rectangles) used to target dCas9-DNMT3A-3L to the enhancer anchor. Lollipop symbols indicate the location of CpGs that are assayed for methylation levels in (C). (B) Schematic representation of the experiment. HCT-116 or K562 cells were transfected with plasmids encoding the dCAS9-DNMT3A-3L, GFP, and a gRNA together with a plasmid encoding 2 additional gRNAs. HCT-116 or K562 cells were isolated by FACS after 2 days, and DNA and RNA were isolated. (C) Methylation at *MYC* promoter loop-anchor site in untreated cells and cells transfected with dCas9-DNMT3A-3L in conjunction with the 5 indicated gRNAs. (D) qPCR analysis of *MYC* mRNA levels and fraction of methylated CpGs for untreated and dCas9-DNMT3A-3L + 5 gRNA transfected cells. Error bars represent the SD of the mean for three biological replicates.

See also [Figure S4](#).

### Loss of *MYC* Expression with Methylation of Enhancer-Docking Site

CTCF binding is abrogated when its sequence motif is methylated (Bell and Felsenfeld, 2000; Maurano et al., 2015), and the *MYC* enhancer-docking site occurs within a CpG island that is consistently hypomethylated in different tumor types as well as in different normal tissues (Figures S4A and S4B). The recent development of tools that permit site-specific DNA methylation (Liu et al., 2016; Siddique et al., 2013) suggested a means to disrupt *MYC* expression by methylation of the enhancer-docking site. To achieve targeted methylation, we created a construct to express a dCas9 fusion protein consisting of the catalytic domain of DNMT3A and the interacting domain of DNMT3L. This dCas9-DNMT3A-3L protein was targeted to the *MYC* enhancer-docking site in K562 and HCT-116 cells using multiple

guide RNAs that span the region (Figures 4A and 4B). The targeting of dCas9-DNMT3A-3L resulted in robust local DNA methylation (Figure 4C) and a 40%–50% reduction in mRNA levels in both cell types (Figure 4D). The methylated region likely contains binding sites for additional TFs that may be sensitive to DNA methylation, so it is possible that the reduced mRNA levels are due to multiple factors. In order to test whether disruption of TFs other than CTCF contribute to the reduction in *MYC* mRNA levels, the dCas9-DNMT3A-3L was targeted to the *MYC* enhancer-docking site in CTCF-site deleted K562 cells. No further reduction of *MYC* mRNA levels was observed under these conditions (Figures S4C and S4D), indicating that loss of CTCF was a major contributor to the observed reduction of *MYC* expression upon targeted methylation of the *MYC* enhancer-docking site. These results demonstrate that



epigenetic editing of the enhancer-docking site can reduce *MYC* expression.

### CTCF Enhancer-Docking Sites at Additional Genes

Previous genomic studies have noted that CTCF might engender enhancer-promoter interactions at a minority of genes (Banani et al., 2017; Nora et al., 2017; Splinter et al., 2006; Zuin et al., 2014). We therefore identified the set of genes whose promoter-proximal regions contain CTCF-bound sites and that show evidence of enhancer interactions in K562, Jurkat, and HCT-116 cells. We identified all active transcription start sites (TSSs) that have at least one CTCF-bound site within 2.5 kb of the TSS that interacts with at least one enhancer. This yielded between 555 and 1,108 TSSs with a nearby CTCF site that loops to an active enhancer (Figure 5A; Table S2). We define these TSSs as having a putative CTCF enhancer-docking site. The majority of TSSs identified in this analysis were identified in only one cell type, with only 52 TSSs identified in all three cell types (Figure 5B). Nonetheless, these putative enhancer-docking sites tended to be constitutively bound by CTCF in all three cell types, and the CTCF motifs in these sites showed high sequence conservation (Figures 5C and 5D). This suggests that these putative enhancer-docking sites are occupied by CTCF regardless of interaction with active enhancers and that differences in cell-type-specific enhancers are largely responsible for differential use of enhancer-docking site genes in these cells.

Gene ontology analysis of the genes with putative enhancer-docking sites found different processes to be significantly enriched in each cell type, and these processes were dominated by the cellular identity of the cell lines (Figure S5A; Table S3). Common processes among the three cell types include cell cycle and other cancer-related processes such as gene expression and response to signaling (Figure S5A). A number of cancer-associated genes were found, including *TGIF1*, *VEGFA*, *RUNX1*, and *PIM1* (Figure 5E), as well as others (Figure S5B). We conclude that genes other than *MYC* are likely regulated by CTCF-bound enhancer-docking sites and that these include multiple cancer-associated genes.

### DISCUSSION

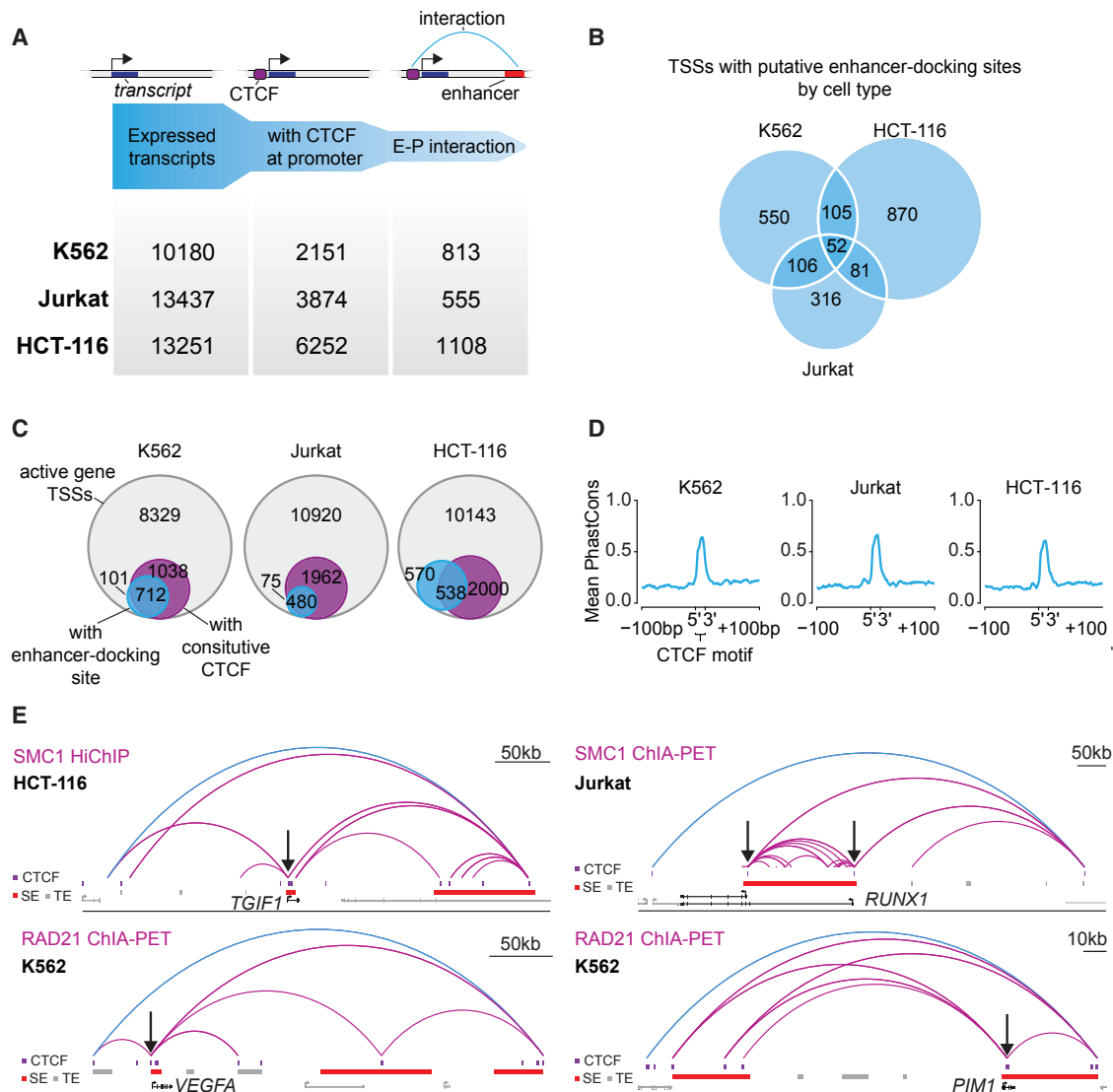
Aberrant transcriptional activation of the *MYC* oncogene occurs frequently in tumor cells and is associated with tumor aggression. *MYC* resides within a 2.8 Mb TAD and its aberrant activation is generally accomplished by acquisition of a super-enhancer somewhere within that TAD. How these diverse cancer-specific super-enhancers loop long distances to specifically interact with *MYC* has not been clear. We find that the diverse super-enhancers commonly interact with, and depend on, a conserved CTCF binding site located 2 kb upstream of the *MYC* promoter. Because tumor super-enhancers can encompass genomic regions as large as 200 kb, and CTCF occupies sites that occur on average every 10 kb, there is considerable opportunity for super-enhancers to adventitiously contain a CTCF-bound site, which in turn could serve to interact with the *MYC* CTCF site (Table S6). Thus, different tumor super-enhancers have the opportunity to form through diverse mecha-

nisms throughout this large TAD and can exploit the *MYC* CTCF site to interact with and activate *MYC* expression.

The concept that enhancer-promoter interactions generally occur within larger chromosomal loop structures such as TADs, which are themselves often formed by the interaction of CTCF proteins bound to each of the TAD loop anchors (Dekker and Mirny, 2016; Fraser et al., 2015; Gorkin et al., 2014a; Hnisz et al., 2016a), is supported by the observations described here. These larger loop structures tend to insulate enhancers and genes within the CTCF-CTCF loops from elements outside those loops. Constraining DNA interactions within CTCF-CTCF loop structures in this manner may facilitate proper enhancer-promoter contacts.

The evidence described here argues that diverse human tumor cell super-enhancers depend on the *MYC* CTCF site for optimal levels of enhancer-promoter looping and mRNA expression. A recent independent study in K562 cells used a tiling CRISPR screen to systematically perturb the *MYC* locus and also found that full *MYC* expression and cell proliferation is dependent on this region (Fulco et al., 2016). However, deletion of the -2 kb CTCF site has limited effects on *MYC* expression in mice (Dave et al., 2017; Gombert and Krumm, 2009), and some translocated enhancers can drive *MYC* expression in the absence of this CTCF site (Shiramizu et al., 1991). There are several potential explanations for these diverse results. It is possible that the -2 kb CTCF site is important for optimal *MYC* expression levels in human cells, but not in mice. It is conceivable that the deletion of a region containing the CTCF site can be compensated by features of the new enhancer landscape in the deletion mutations. Furthermore, additional mechanisms normally involved in enhancer-promoter interactions, such as YY1-YY1 interactions, may mask the loss of the CTCF site *in vivo*; YY1 is present in the *MYC* promoter region and is thus likely to contribute to DNA looping and expression (Weintraub et al., 2017).

Our studies suggest that an additional set of human genes, beyond *MYC*, may utilize promoter-proximal enhancer-docking sites to mediate cell-type-specific enhancer-promoter interactions. Such CTCF-mediated enhancer-promoter interactions are generally nested within larger CTCF-mediated loops that would function as insulated neighborhoods. At these genes with CTCF-mediated enhancer docking, the promoter-proximal enhancer-docking sites tend to be constitutively bound by CTCF and these binding sites tend to be highly conserved. Indeed, two studies have reported that these genes tend to lose expression upon perturbation of CTCF (Nora et al., 2017; Zuin et al., 2014), consistent with a role for CTCF in enhancer-promoter looping. Among these genes are cancer-associated genes that likely employ this mechanism to engender interactions with tumor-specific enhancers. For example, at *CSNK1A1*, a drug target in acute myeloid leukemia (AML) tumor cells (Järås et al., 2014), *VEGFA*, which is upregulated in many cancers (Goel and Mercurio, 2013), and *RUNX1*, a well-defined oncogene in AML (Deltcheva and Nimmo, 2017), the evidence suggests that super-enhancers in these cancer cells use a CTCF enhancer-docking mechanism to interact with the oncogene. Thus, a CTCF-dependent enhancer-docking mechanism, which presumably facilitates interaction with different



**Figure 5. Putative Enhancer-Docking Sites Occur at Additional Genes with Prominent Roles in Cancer**

(A) Identification of genes with putative enhancer-docking sites. Genes were filtered for their expression status, presence of a CTCF binding site within 2.5 kb of the TSS and evidence of looping to an active enhancer, defined by H3K27Ac ChIP-seq.

(B) Venn-diagram showing the overlap of TSSs with putative CTCF enhancer docking in K562, HCT-116, and Jurkat cells.

(C) Venn-diagrams showing the number of TSSs from active genes, the number of these that exhibit putative CTCF enhancer-docking and how many of these have a constitutive CTCF site within 2.5 kb of the TSS.

(D) Conservation analysis of the CTCF motifs in the CTCF-bound elements in putative enhancer-docking sites. The mean 46-way PhastCons score of the highest JASPAR scoring motifs in CTCF peaks within putative CTCF enhancer docking and their flanking regions are shown.

(E) Examples of genes with putative CTCF enhancer-docking sites from the different cell types analyzed. CTCF ChIP-seq peaks are shown as purple rectangles, typical enhancers are shown as gray rectangles, and super-enhancers are shown as red rectangles. Black arrows indicate the CTCF sites that may facilitate enhancer docking. The insulated neighborhood loop is shown in blue and loops internal to it are shown in purple. HCT-116 HiChIP interactions internal to the neighborhood with an origami score of at least 0.97 are shown for the *RUNX1* locus. Jurkat SMC1 ChIA-PET interactions internal to the neighborhood with an origami score of at least 0.9 and a minimum PET count of 15 are shown for the *TGIF1* locus. K562 RAD21 ChIA-PET interactions internal to the neighborhood with an origami score of at least 0.9 and a minimum PET count of 30 are shown for the *VEGFA* locus. K562 RAD21 ChIA-PET interactions internal to the neighborhood with an origami score of at least 0.9 and a minimum PET count of 30 are shown for the *PIM1* locus. Data are from this study and two others (Hnisz et al., 2016a; Heidari et al., 2014).

See also Figure S5.

cell-specific enhancers during development, is exploited by cancer cells to dysregulate expression of prominent oncogenes.

*MYC* dysregulation is a hallmark of cancer (Bradner et al., 2017). The c-MYC TF is an attractive target for cancer therapy because of the role that excessive c-MYC levels play in a broad spectrum of aggressive cancers (Felsher and Bishop, 1999; Jain et al., 2002; Soucek et al., 2008), but direct pharmacologic inhibition of c-MYC remains an elusive challenge in drug discovery (Bradner et al., 2017). The *MYC* enhancer-docking site, and presumably those of other oncogenes, can be repressed by dCas9-DNMT-mediated DNA methylation. Oncogene enhancer-docking sites may thus represent a vulnerability in multiple human cancers.

## EXPERIMENTAL PROCEDURES

Further details and an outline for resources used in this work can be found in [Supplemental Experimental Procedures](#).

### CRISPR/Cas9 Genome Editing

Genome editing was performed using CRISPR/Cas9 essentially as described previously (Ran et al., 2013). The genomic sequences complementary to all guide RNAs are listed in [Table S4](#).

### ChIP-Seq

ChIP was performed as described previously (Lee et al., 2006). Approximately 30 million cells were crosslinked for 10 min at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution to the growth media followed by 5 min quenching with 125 mM glycine. Cells were washed twice with PBS, and then the supernatant was aspirated and the cell pellet was flash frozen at  $-80^{\circ}\text{C}$ . 100  $\mu\text{L}$  Protein G Dynabeads (Thermo 10003D) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 40  $\mu\text{L}$  anti-CTCF antibody (Millipore 07-729). Nuclei were isolated as previously described (Lee et al., 2006) and sonicated in lysis buffer on a Misonix 3000 sonicator for 5 cycles at 30 s each on ice (18–21 W) with 60 s on ice between cycles. Sonicated lysates were cleared once by centrifugation and incubated overnight at  $4^{\circ}\text{C}$  with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed with wash buffers A, B, C, and D sequentially. DNA was eluted, cross-links were reversed, and DNA was purified with phenol chloroform extraction and ethanol precipitation. Libraries for Illumina sequencing were prepared following the Illumina TruSeq DNA Sample Preparation v2 kit and sequenced on the Illumina HiSeq 2500 for 40 bases in single-read mode.

### 4C-Seq

A modified version of 4C-seq (van de Werken et al., 2012) was developed ([Supplemental Experimental Procedures](#)). The major change was the ligation is performed in intact nuclei (*in situ*). This change was incorporated because previous work has noted that *in situ* ligation dramatically decreases the rate of chimeric ligations and background interactions (Rao et al., 2014).

### HiChIP

HiChIP was performed essentially as described (Mumbach et al., 2016). 10 million HCT116 cells were crosslinked for 10 min at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution to the growth media followed by 5-min quenching with glycine. Cells were washed twice with PBS, and then the cell pellet was flash frozen in liquid nitrogen. Frozen samples were processed according to protocol ([Supplemental Experimental Procedures](#)).

### Targeted Methylation and Bisulfite Sequencing

To perform targeted methylation, cells were transfected with a dCas9-DNMT3A-3L construct and five guides. To generate the dCas9-DNMT3A-3L

construct, dCas9 was isolated from pSQL1658 (Addgene 51023) by PCR. Cas9 was removed from pX330-Cas9 (Addgene 42230) and replaced by DNMT3A-3L (Siddique et al., 2013). Guide RNAs used for targeting can be found in [Table S4](#).

## Statistical Methods

### ChIP-Seq Data Analysis

ChIP-seq datasets were generated for this study as well as collated from previous studies ([Table S5](#)) and were aligned using Bowtie (version 0.12.2) to the human genome (build hg19, GRCh37) with parameter  $-k\ 1\ -m\ 1\ -n\ 2$ . We used MACS version 1.4.2 with the parameter “no-model-keep-dup=auto.” A p value threshold of enrichment of  $1e-09$  was used.

### 4C Analysis

4C-seq reads were trimmed and mapped using bowtie with options  $-k\ 1\ -m\ 1$  against the hg19 genome assembly. We only used the reads from non-blind fragments for further analysis. The normalized profile of each sample was smoothened using a 6-kb running mean at 500-bp steps across the genome. Quantification of the 4C signal counted the reads per fragment per million sequenced reads in the super-enhancers or the CTCF MACS peak calls.

### HiChIP and ChIA-PET Data Analysis

We developed a new software pipeline and analytical method called origami to process HiChIP and chromatin interaction analysis by paired-end tag sequencing (ChIA-PET). The software and releases can be found at <https://github.com/younglab/origami> using version alpha20160828. The ChIA-PET datasets analyzed along with their corresponding linker sequence and called interactions in and around the *MYC* TAD can be found in [Table S4](#). Each ChIA-PET datasets were processed as follows: the reads were first trimmed and aligned using origami-alignment. Each end of a paired end tag (PET) with a linker sequence were separately mapped to the hg19 genome assembly using bowtie with the following options:  $-v\ 1\ -k\ 1\ -m\ 1$ . After alignment, the separated PETs were re-paired in the final BAM output. After repairing, all duplicated PETs within the data were removed. Peaks were called on the re-paired ChIA-PET reads using MACS1 v1.4.2 with the following parameters:  $-nolambda\ -nomodel\ -p\ 1e-9$ .

## DATA AND SOFTWARE AVAILABILITY

The accession number for the sequencing data reported in this paper is GEO: GSE92881. The origami algorithm is accessible at <https://github.com/younglab/origami> (version alpha20160828).

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and six tables and can be found with this article online at <https://doi.org/10.1016/j.celrep.2018.03.056>.

## ACKNOWLEDGMENTS

We thank X. Shawn Liu and Rudolf Jaenisch for their help with the targeted methylation of the CTCF-bound element. We thank the Whitehead Institute Genome Technology Core and the FACS facility for their contribution to this work. This work is supported by NIH grant HG002668 (R.A.Y.), a Rubicon fellowship by NWO (J.S.), Ludwig Graduate Fellowship funds (A.S.W.), American Cancer Society fellowship PF-16-146-01-DMC (D.S.D.), a Margaret and Herman Sokol Postdoctoral Award (D.H.), and National Science Foundation GRFP funds (A.V.Z).

## AUTHOR CONTRIBUTIONS

J.S., J.C.M., and R.A.Y. wrote the manuscript. J.S. and J.C.M. performed methylation, perturbation, and other experiments. D.H. performed HiChIP. D.S.D. analyzed HiChIP and ChIA-PET data and wrote origami. J.S., J.C.M., and A.V.Z. performed 4C experiments. A.S.W. made inducible CTCF-site deletion lines. T.I.L. provided helpful feedback and R.A.Y. supervised.

## DECLARATION OF INTERESTS

The Whitehead Institute filed a patent application based on this study. R.A.Y. is a founder of Syros Pharmaceuticals, of Marauder Therapeutics and of Omega Therapeutics. The remaining authors declare no competing interests.

Received: December 8, 2017

Revised: February 9, 2018

Accepted: March 14, 2018

Published: April 10, 2018

## REFERENCES

- Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298.
- Becket, E., Chopra, S., Duymich, C.E., Lin, J.J., You, J.S., Pandiyan, K., Nichols, P.W., Siegmund, K.D., Charlet, J., Weisenberger, D.J., et al. (2016). Identification of DNA methylation-independent epigenetic events underlying clear cell renal cell carcinoma. *Cancer Res.* **76**, 1954–1964.
- Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482–485.
- Berns, E.M.J.J., Klijn, J.G.M., van Putten, W.L., van Staveren, I.L., Portengen, H., and Foekens, J.A. (1992). *c-myc* amplification is a better prognostic factor than *HER2/neu* amplification in primary breast cancer. *Cancer Res.* **52**, 1107–1113.
- Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678.
- Bradner, J.E., Hnisz, D., and Young, R.A. (2017). Transcriptional addiction in cancer. *Cell* **168**, 629–643.
- Buecker, C., and Wysocka, J. (2012). Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet.* **28**, 276–284.
- Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144**, 327–339.
- Chapuy, B., McKeown, M.R., Lin, C.Y., Monti, S., Roemer, M.G.M., Qi, J., Rahl, P.B., Sun, H.H., Yeda, K.T., Doench, J.G., et al. (2013). Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* **24**, 777–790.
- Cuddapah, S., Jothi, R., Schones, D.E., 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.
- Dang, C.V. (2012). *MYC* on the path to cancer. *Cell* **149**, 22–35.
- Dave, K., Sur, I., Yan, J., Zhang, J., Kaasinen, E., Zhong, F., Blaas, L., Li, X., Kharazi, S., Gustafsson, C., et al. (2017). Mice deficient of *Myc* super-enhancer region reveal differential control mechanism between normal and pathological growth. *eLife* **6**, 1–25.
- de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* **501**, 227–231.
- Dekker, J., and Mirny, L. (2016). The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110–1121.
- Deltcheva, E., and Nimmo, R. (2017). *RUNX* transcription factors at the interface of stem cells and cancer. *Biochem. J.* **474**, 1755–1768.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380.
- Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., and Young, R.A. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387.
- Felsher, D.W., and Bishop, J.M. (1999). Reversible tumorigenesis by *MYC* in hematopoietic lineages. *Mol. Cell* **4**, 199–207.
- Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenko, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian *c-myc* oncogenes. *Mol. Cell. Biol.* **16**, 2802–2813.
- Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L., et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269.
- Fraser, J., Ferrai, C., Chiariello, A.M., Schueler, M., Rito, T., Laudanno, G., Barbieri, M., Moore, B.L., Kraemer, D.C., Aitken, S., et al.; FANTOM Consortium (2015). Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol. Syst. Biol.* **11**, 852.
- Frietze, S., Wang, R., Yao, L., Tak, Y.G., Ye, Z., Gaddis, M., Witt, H., Farnham, P.J., and Jin, V.X. (2012). Cell type-specific binding patterns reveal that CTCF7L2 can be tethered to the genome by association with GATA3. *Genome Biol.* **13**, R52.
- Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**, 769–773.
- Gabay, M., Li, Y., and Felsher, D.W. (2014). *MYC* activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harb. Perspect. Med.* **4**, 1–14.
- Gibcus, J.H., and Dekker, J. (2013). The hierarchy of the 3D genome. *Mol. Cell* **49**, 773–782.
- Goel, H.L., and Mercurio, A.M. (2013). VEGF targets the tumour cell. *Nat. Rev. Cancer* **13**, 871–882.
- Gombert, W.M., and Krumm, A. (2009). Targeted deletion of multiple CTCF-binding elements in the human *C-MYC* gene reveals a requirement for CTCF in *C-MYC* expression. *PLoS ONE* **4**, e6109.
- Gombert, W.M., Farris, S.D., Rubio, E.D., Morey-Rosler, K.M., Schubach, W.H., and Krumm, A. (2003). The *c-myc* insulator element and matrix attachment regions define the *c-myc* chromosomal domain. *Mol. Cell. Biol.* **23**, 9338–9348.
- Gorkin, D.U., Leung, D., and Ren, B. (2014a). The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* **14**, 762–775.
- Gröschel, S., Sanders, M.A., Hoogenboezem, R., de Wit, E., Bouwman, B.A.M., Erpelinck, C., van der Velden, V.H.J., Havermans, M., Avellino, R., van Lom, K., et al. (2014). A single oncogenic enhancer rearrangement causes concomitant *EV11* and *GATA2* deregulation in leukemia. *Cell* **157**, 369–381.
- Grotzer, M.A., Hogarty, M.D., Janss, A.J., Liu, X., Zhao, H., Eggert, A., Sutton, L.N., Rorke, L.B., Brodeur, G.M., and Phillips, P.C. (2001). *MYC* messenger RNA expression predicts survival outcome in childhood primitive neuroectodermal tumor/medulloblastoma. *Clin. Cancer Res.* **7**, 2425–2433.
- Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W.H., Ye, C., Ping, J.L.H., Mulawadi, F., et al. (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* **43**, 630–638.
- Heidari, N., Phanstiel, D.H., He, C., Grubert, F., Jahanbani, F., Kasowski, M., Zhang, M.Q., and Snyder, M.P. (2014). Genome-wide map of regulatory interactions in the human genome. *Genome Res.* **24**, 1905–1917.
- Herranz, D., Ambesi-Impiombato, A., Palomero, T., Schnell, S.A., Belver, L., Wendorff, A.A., Xu, L., Castillo-Martin, M., Llobet-Navás, D., Cordon-Cardo, C., et al. (2014). A *NOTCH1*-driven *MYC* enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat. Med.* **20**, 1130–1137.
- Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947.
- Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.-L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., et al. (2016a). Activation of



- proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351, 1454–1458.
- Hnisz, D., Day, D.S., and Young, R.A. (2016b). Insulated Neighborhoods: structural and functional units of mammalian gene control. *Cell* 167, 1188–1200.
- Hnisz, D., Schuijers, J., Li, C.H., and Young, R.A. (2018). Regulation and dysregulation of chromosome structure in cancer. *Annu. Rev. Cancer Biol.* 2, 21–40.
- Jain, M., Arvanitis, C., Chu, K., Dewey, W., Leonhardt, E., Trinh, M., Sundberg, C.D., Bishop, J.M., and Felsher, D.W. (2002). Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science* 297, 102–104.
- Järås, M., Miller, P.G., Chu, L.P., Puram, R.V., Fink, E.C., Schneider, R.K., Al-Shahrour, F., Peña, P., Breyfogle, L.J., Hartwell, K.A., et al. (2014). Csnk1a1 inhibition has p53-dependent therapeutic efficacy in acute myeloid leukemia. *J. Exp. Med.* 211, 605–612.
- Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., et al.; BLUEPRINT Consortium (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* 167, 1369–1384.
- Ji, X., Dadon, D.B., Powell, B.E., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., Hnisz, D., Pegoraro, G., Lee, T.I., et al. (2016). 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* 18, 262–275.
- Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D.D., Zhang, M.Q., Lobanenko, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128, 1231–1245.
- Klenova, E.M., Nicolas, R.H., Paterson, H.F., Carne, A.F., Heath, C.M., Goodwin, G.H., Neiman, P.E., and Lobanenko, V.V. (1993). CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol. Cell. Biol.* 13, 7612–7624.
- Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat. Protoc.* 1, 729–748.
- Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., 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.
- Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 151, 56–67.
- Lin, C.Y., Erkek, S., Tong, Y., Yin, L., Federation, A.J., Zapotka, M., Haldipur, P., Kawachi, D., Risch, T., Warnatz, H.-J., et al. (2016). Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature* 530, 57–62.
- Liu, M., Maurano, M.T., Wang, H., Qi, H., Song, C.-Z., Navas, P.A., Emery, D.W., Stamatoyannopoulos, J.A., and Stamatoyannopoulos, G. (2015). Genomic discovery of potent chromatin insulators for human gene therapy. *Nat. Biotechnol.* 33, 198–203.
- Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A., and Jaenisch, R. (2016). Editing DNA methylation in the mammalian genome. *Cell* 167, 233–247.e17.
- Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153, 320–334.
- Maurano, M.T., Wang, H., John, S., Shafer, A., Canfield, T., Lee, K., and Stamatoyannopoulos, J.A. (2015). Role of DNA methylation in modulating transcription factor occupancy. *Cell Rep.* 12, 1184–1195.
- Montavon, T., and Duboule, D. (2012). Landscapes and archipelagos: spatial organization of gene regulation in vertebrates. *Trends Cell Biol.* 22, 347–354.
- Müller-Sturm, H.P., Sogo, J.M., and Schaffner, W. (1989). An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. *Cell* 58, 767–777.
- Muerdter, F., and Stark, A. (2016). Gene regulation: activation through space. *Curr. Biol.* 26, R895–R898.
- Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* 13, 919–922.
- Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* 347, 1017–1021.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385.
- Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 169, 930–944.
- Ovcharenko, I., Loots, G.G., Nobrega, M.A., Hardison, R.C., Miller, W., and Stubbs, L. (2005). Evolution and functional classification of vertebrate gene deserts. *Genome Res.* 15, 137–145.
- Parker, S.C.J., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., van Bueren, K.L., Chines, P.S., Narisu, N., Black, B.L., et al.; NISC Comparative Sequencing Program; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors (2013). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl. Acad. Sci. USA* 110, 17921–17926.
- Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.-T.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1281–1295.
- Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., et al. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature* 515, 402–405.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308.
- Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.
- Rubio, E.D., Reiss, D.J., Welch, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proc. Natl. Acad. Sci. USA* 105, 8309–8314.
- Saldaña-Meyer, R., González-Buendía, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F., and Reinberg, D. (2014). CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* 28, 723–734.
- Sandelin, A., Alkema, W., Engström, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* 32, D91–D94.
- Shi, J., Whyte, W.A., Zepeda-Mendoza, C.J., Milazzo, J.P., Shen, C., Roe, J.S., Minder, J.L., Mercan, F., Wang, E., Eckersley-Maslin, M.A., et al. (2013). Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev.* 27, 2648–2662.
- Shiramizu, B., Barriga, F., Neequaye, J., Jafri, A., Dalla-Favera, R., Neri, A., Gutierrez, M., Levine, P., and Magrath, I. (1991). Patterns of chromosomal breakpoint locations in Burkitt's lymphoma: relevance to geography and Epstein-Barr virus association. *Blood* 77, 1516–1526.
- Siddique, A.N., Nunna, S., Rajavelu, A., Zhang, Y., Jurkowska, R.Z., Reinhardt, R., Rots, M.G., Ragozin, S., Jurkowski, T.P., and Jeltsch, A. (2013). Targeted demethylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J. Mol. Biol.* 425, 479–491.

- Soucek, L., Whitfield, J., Martins, C.P., Finch, A.J., Murphy, D.J., Sodir, N.M., Karnezis, A.N., Swigart, L.B., Nasi, S., and Evan, G.I. (2008). Modelling Myc inhibition as a cancer therapy. *Nature* **455**, 679–683.
- Spitz, F. (2016). Gene regulation at a distance: From remote enhancers to 3D regulatory ensembles. *Semin. Cell Dev. Biol.* **57**, 57–67.
- Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* **20**, 2349–2354.
- Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Rusczycki, B., et al. (2015). CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611–1627.
- van de Werken, H.J.G., Landan, G., Holwerda, S.J.B., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Öz, Y., Bouwman, B.A.M., et al. (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972.
- Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A., et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **474**, 390–394.
- Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 is a structural regulator of enhancer-promoter loops. *Cell* **171**, 1573–1588.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319.
- Yusufzai, T.M., Tagami, H., Nakatani, Y., and Felsenfeld, G. (2004). CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell* **13**, 291–298.
- Zhang, X., Choi, P.S., Francis, J.M., Imielinski, M., Watanabe, H., Cherniack, A.D., and Meyerson, M. (2016). Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. *Nat. Genet.* **48**, 176–182.
- Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. USA* **111**, 996–1001.