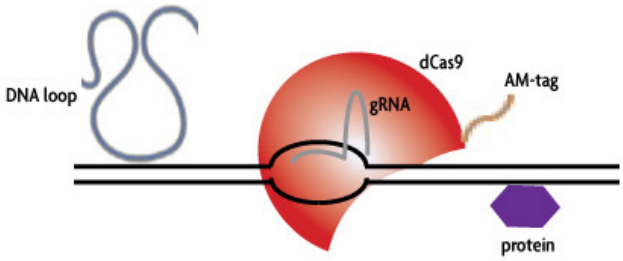
Guideline for Evaluating Chromosomal Looping using enChIP

Product Name	Format	Catalog No.
enChIP Kit	16 rxns	53125
pAM_gRNA Vector	10 µg	53121
pAM_dCas9 Vector	10 µg	53122
FuGENE® HD Transfection Reagent	0.2 ml	32042

The following guideline is designed for use with Active Motif's enChIP Kit (Catalog No. 53125) and Active Motif's deactivated Cas9 (dCas9) expression construct pAM_dCas9 Vector (Catalog No. 53122) for identification of *cis*- and *trans*-chromosomal interactions. The pAM_gRNA Vector (Catalog No. 53121) may be used for cloning the 20 bp target sequence, or researchers may utilize their own gRNA constructs. The target sequence must be expressed in combination with the AM-tagged dCas9 protein. Please refer to the vector data sheets and the enChIP Kit manual for detailed instructions on cloning and expression.



Introduction

Engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) was first described in 2013 by Toshitsugu Fujita and Hodaka Fujii[1, 2]. enChIP was designed to purify specific genomic regions using the CRISPR/Cas9 system. A 20 bp target sequence is cloned into a single guide RNA vector expressing a fusion of the crRNA-tracrRNA sequence (5´side to identify the DNA target sequence and a 3´duplex RNA structure to bind Cas9). When the gRNA is expressed in combination with a tagged deactivated *Streptococcus pyogenes* Cas9 (dCas9) protein, specific genomic regions can be immunoprecipitated using an antibody directed against the tag. By using a catalytically inactive form of Cas9 endonuclease, double-stranded breaks are not introduced and DNA, RNA and proteins associated with the target sequence can be recovered[1-6]. With Active Motif's enChIP Kit, simply clone your 20 nucleotide target sequence into Active Motif's gRNA or dCas9 expression vectors. Following transfection and expression of the gRNA and AM-tagged dCas9, cells are formaldehyde fixed, lysed and the chromatin is fragmented by sonication. An antibody directed against the AM-tag is used to enrich for genomic sequences bound by the gRNA/dCas9 complex. Following immunoprecipitation, DNA can be analyzed by qPCR or NGS to identify the enriched genomic regions.

Chromosomal looping is an epigenetic mechanism by which two (or more) genomic loci are brought into close proximity by DNA binding proteins and their cofactors for the purpose of long-range gene regulation. Looping events play a role in bringing distal regulatory elements such as enhancers into functional proximity of their target promoters, or to separate active regions of the genome from repressed regions via CTCF-mediated insulator elements[7]. Current methods to identify such looping events, namely chromatin conformation capture (3C) and its derivatives (4C, 5C, Hi-C)[8], have shaped the way we understand 3D organization of the nucleus, and how it regulates global gene expression.

The Fujii lab recently described a procedure for using enChIP to identify genomic regions interacting with the 5´HS5 locus in K562 chronic myelogenous leukemia cells[6]. Here we describe a parallel approach for using enChIP for the *de novo* identification of *cis*- and *trans*-chromosomal looping events with a single genomic locus.

Experimental Set-up

In order to confidently use enChIP to determine genomic regions interacting with a locus of interest, a minimum of two specific gRNAs must be designed against the target locus, with each gRNA being used independently in its own transfection. Additionally, a no gRNA control experiment should be performed to identify genomic regions that interact with dCas9 non-specifically. Detailed instructions for designing and cloning gRNAs into Active Motif's expression vectors can be found in the enChIP manual at www.activemotif.com/enchip.

In designing gRNAs, take caution not to target genomic loci containing a conserved transcription factor binding site since that region may be blocked by DNA binding protein preventing gRNA binding and/or because blocking a protein binding site may prevent binding of that protein to biologically functional locus, which may result in functional dysregulation of cellular biology. It is also recommended to avoid designing gRNAs downstream of a transcription start site (TSS) as this may interfere with expression of that gene via a mechanism known as CRISPRi[9].

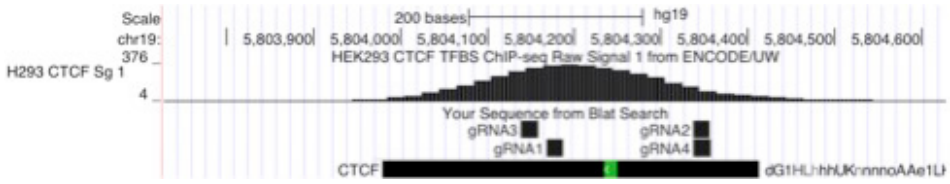


Figure 1: UCSC genome browser snapshot showing a CTCF peak in HEK293T cells.

The black and green bar at the bottom of the image represents a CTCF peak called by the ENCODE consortium, with the green segment representing the genomic region containing the consensus DNA motif recognized by CTCF. Immediately below the peak, four black boxes represent the locations of four gRNAs (gRNA1-4) targeting a 200 bp region surrounding the CTCF binding site on chromosome 19. To prevent interference of CTCF binding, all four gRNAs were designed not to overlap with the CTCF recognition motif or TSS.

To ensure that the experiment is designed to functionally characterize chromosomal looping of a single locus, design the two or more gRNA sequences within a relatively short genomic interval. It is recommended that gRNA target sites bind within the average fragment size of the chromatin, or within approximately 500-1,000 bp of each other. Each gRNA will represent a unique sequence targeting the same genomic interval. Each gRNA sequence is then independently cloned and expressed according to the recommendations provided in the enChIP Kit manual. **Figure 1,** shown above, displays the locations of four gRNAs designed to target a 200 bp region surrounding a CTCF binding site on chromosome 19 (5,804,115–5,804,209) in HEK293T cells.

Chromosomal Looping Experimental Design

- Design 2 or more gRNAs (within 0.5-1 kb of each other) targeting different 20 bp sequences of the same genomic locus
- Avoid sequences containing conserved transcription factor binding sites
- Avoid sequences downstream of a transcriptional start site
- Include a no gRNA negative control to identify and eliminate non-specific dCas9 protein binding during data analysis
- Following data analysis, design the reciprocal experiment to verify each potential chromosomal looping site

Data Analysis – Distinction Between Chromosomal Looping and Off-target Binding

After aligning raw sequencing reads to the appropriate genome, we recommend calling peaks using MACS2 (https://github.com/taoliu/MACS/) on each individual enChIP experiment using the no gRNA negative control reaction as the control or 'input' to remove any peaks that represent non-specific interactions of the dCas9 protein alone.

```
macs2 callpeak –t enChIP_experiment.bam –c enChIP_No_gRNA_Control.bam –g (genome
size of 'hs') ––outdir enChIP/MACS/ –n enChIP_experiment_name –p 1e-7
```

Peaks called on each experiment represent regions of the genome that interact with the dCas9/gRNA complex. Expected peak numbers range from several dozen to a couple of thousand peaks per experiment. In addition to the targeted genomic locus the peaks represent potential chromosomal interactions and off-target gRNA binding sites that are enriched relative to the no gRNA control. To distinguish between looping and off-target peaks, we take advantage of having targeted the locus of interest with multiple independent gRNAs. Each gRNA has a unique DNA binding context, so it is expected that any differences between experiments targeting the same genomic locus represent off-target binding, and that any peaks common between experiments represent potential *cis*- and *trans*-interacting regions.

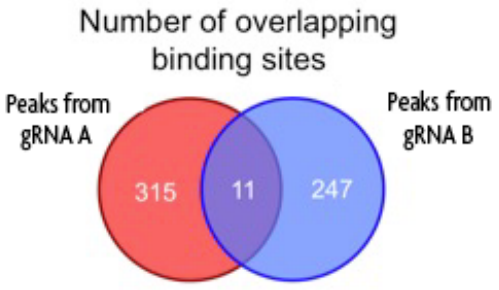


Figure 2: Identification of potential looping sites using an experimental overlap strategy.
Due to the fact that individual gRNAs recognize different 20 bp sequences in the genome, it is expected that off-target binding sites will differ between experiments. Each gRNA will have a unique set of off-target binding sites. Therefore, any peaks found in common between experiments performed with different gRNAs targeting the same genomic locus are expected to be potential looping sites rather than off-target binding.

To identify common peaks, we recommend using the 'mergepeaks' function in HOMER[10]. (http://homer.salk.edu/homer/)

```
mergePeaks –venn outputFile.txt –prefix OverlappingRegions gRNA1peakFile.bed
gRNA2peakFile.bed
```

The resulting text file contains an all-by-all matrix for each peak file being compared enumerating peaks found in common-between or unique-to each data set. Using the '-prefix' option will result in the generation of new peak files containing overlapping and unique regions identified in the overlap, and the name provided after '-prefix' will be appended onto the beginning of each new peak file name. The resulting list of overlapping peaks represents a set of potential looping sites with your target locus.

Verification of Potential Looping Sites

Peaks identified in common between all independent gRNA experiments targeting the same locus (as described above) represent a condensed list of potential chromosomal interactions. enChIP-seq can now be employed to verify these interactions in a reciprocal experiment. If a genuine interaction between the two chromosomal loci exists, designing two or more gRNAs to any of these potential looping regions should yield a common peak at the original target site as depicted in **Figure 3**.

Figure 3: A strategy for validating potential chromosomal looping sites.
Top panel: Targeting a single genomic locus with several independent gRNAs yields potential looping sites.
Bottom panel: To verify these potential looping events, several independent gRNAs are designed for each potential looping locus. Identification of the original target site as a peak in this reciprocal experiment serves as verification of this chromosomal interaction.

CTCF plays a role in the 3D organization of chromatin by binding to insulator elements, which separate active regions of the genome from repressed regions. CTCF-bound insulator elements have been shown to form clusters of chromosomal looping events[7]. To determine the extent to which a specific CTCF binding site plays a role in chromosomal looping, four gRNAs were designed to target a binding site on chromosome 19 as shown in FIGURE 4. Each gRNA gave a specific peak at the target site (Bottom left panel) as well as several peaks distributed across the genome. An overlap of all four data sets yielded 520 peaks in common. The verification of one such regions, the promoter of the CARM1 gene on chromosome 19 is shown in the bottom right panel. Each gRNA targeting the CTCF binding site yielded a distinct peak at the CARM1 promoter, and when a gRNA was designed to target the CARM1 promoter, a peak was detected at the original CTCF binding site.

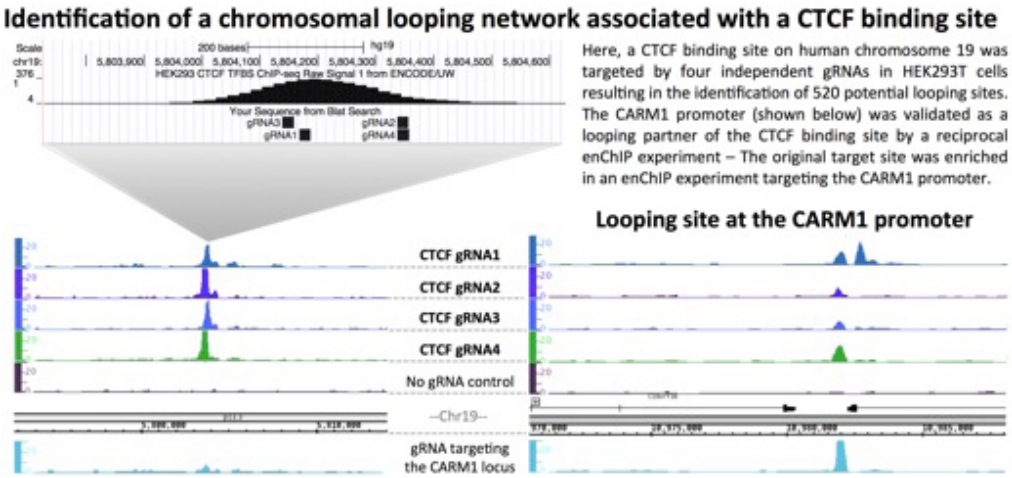


Figure 4: Identification and verification of looping events with a CTCF binding site on chromosome 19.
UCSC Genome Browser snapshot showing the CTCF peak in HEK293T cells (black mound). Four separate gRNAs were designed to target this locus within a 200 bp region. Below the UCSC track are the enChIP-seq peaks identified for each gRNA at the target locus compared to a no gRNA control. The right panel shows a looping site identified in each independent experiment targeting the CTCF binding site. A gRNA targeting this site yielded a peak at the original CTCF binding site, verifying the looping event (lower blue track).

As we have shown here, enChIP provides the ability to identify chromosomal looping events with a genomic locus of interest. We find this method to be analogous to existing methods for identifying chromosomal interactions and encourage researchers to cross-validate their experiments using these methods. For an in-depth guide on how to use 4C technology to identify chromosomal looping events, please refer to the Splinter, E. *et al*. *Methods* paper[11].

References

1. Fujita T and Fujii H: **Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR**. *Biochem. Biophys. Res. Commun.* 2013, **439(1)**: 132–136.
2. Fujita T, Asano Y, Ohtsuka J, Takada Y, Saito K, Ohki R, Fujii H: **Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP)**. *Sci. Rep.* 2013, **3: 3171**.
3. Fujita T and Fujii H: **Identification of Proteins Associated with an IFNγ-Responsive Promoter by a Retroviral Expression System for enChIP Using CRISPR**. *PLoS ONE* 2014, **9(7)**: e103084.
4. Fujita T, Yuno M, Okuzaki D, Ohki R, Fujii H: **Identification of Non-Coding RNAs Associated with Telomeres Using a Combination of enChIP and RNA Sequencing**. *PLoS ONE* 2015, **10(4)**: e0123387.
5. Fujii H and Fujita T: **Isolation of Specific Genomic Regions and Identification of Their Associated Molecules by Engineered DNA-Binding Molecule-Mediated Chromatin Immunoprecipitation (enChIP) Using the CRISPR System and TAL Proteins.** *Int. J. Mol. Sci.* 2015, **16(9)**: 21802–21812.

6. Fujita T, Yuno M, Suzuki Y, Sugano S, Fujii H: **Identification of physical interactions between genomic regions by enChIP-Seq**. *bioRxiv* 2016: 036160.

7. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.** *Cell* 2014, **159(7)**:1665–1680.

8. de Wit E, de Laat W: **A decade of 3C technologies: insights into nuclear organization**. *Genes Dev* 2012, **26(1)**: 11–24.

9. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression**. *Cell* 2013, **152(5)**: 1173–1183.

10. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK: **Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities**. *Molecular Cell* 2010, **38(4)**: 576–589.

11. Splinter E, de Wit E, van de Werken HJG, Klous P, de Laat W: **Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation**. *Methods* 2012, **58(3)**: 221–230.

Technical Services

If you need assistance at any time, please call Active Motif Technical Service at one of the numbers listed below.

Active Motif North America
1914 Palomar Oaks Way, Suite 150
Carlsbad, CA 92008
USA
Toll Free: 877 222 9543
Telephone: 760 431 1263
Fax: 760 431 1351
E-mail: tech_service@activemotif.com

Active Motif Europe
Avenue Reine Astrid, 92
B-1330 La Hulpe, Belgium
UK Free Phone: 0800 169 31 47
France Free Phone: 0800 90 99 79
Germany Free Phone: 0800 181 99 10
Telephone: +32 (0)2 653 0001
Fax: +32 (0)2 653 0050
E-mail: eurotech@activemotif.com

Active Motif Japan
Azuma Bldg, 7th Floor
2-21 Ageba-Cho, Shinjuku-Ku
Tokyo, 162-0824, Japan
Telephone: +81 3 5225 3638
Fax: +81 3 5261 8733
E-mail: japantech@activemotif.com

Active Motif China
787 Kangqiao Road
Building 10, Suite 202, Pudong District
Shanghai, 201315, China
Telephone: (86)-21-20926090
Hotline: 400-018-8123

Visit Active Motif on the worldwide web at **http://www.activemotif.com**

At this site:

- Read about who we are, where we are, and what we do
- Review data supporting our products and the latest updates
- Enter your name into our mailing list to receive our catalog, *MotifVations* newsletter and notification of our upcoming products
- Share your ideas and results with us
- View our job opportunities

Don't forget to bookmark our site for easy reference!