



Published in final edited form as:

*Curr Protoc Mol Biol.* 2010 January ; CHAPTER 21: Unit–21.1525. doi:  
10.1002/0471142727.mb2115s89.

## Chromatin Interaction Analysis using Paired-End Tag Sequencing

Melissa J. Fullwood<sup>1,2</sup>, Yuyuan Han<sup>1,2</sup>, Chia-Lin Wei<sup>1</sup>, Xiaoan Ruan<sup>1</sup>, Yijun Ruan<sup>1,3</sup>

<sup>1</sup>Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore

### Abstract

Chromatin Interaction Analysis using Paired-End Tag sequencing (ChIA-PET) is a technique developed for large-scale, *de novo* analysis of higher-order chromatin structures. Cells are treated with formaldehyde to cross-link chromatin interactions, DNA segments bound by protein factors are enriched by chromatin immunoprecipitation, and interacting DNA fragments are then captured by proximity ligation. The Paired-End Tag (PET) strategy is applied to the construction of ChIA-PET libraries, which are sequenced by high-throughput next-generation sequencing technologies. Finally, raw PET sequences are subjected to bioinformatics analysis, resulting in a genome-wide map of binding sites and chromatin interactions mediated by the protein factor under study. This unit describes ChIA-PET for genome-wide analysis of chromatin interactions in mammalian cells, with the application of Roche/454 and Illumina sequencing technologies.

### Keywords

PET; Paired-End; Tag; Mate-pair; SAGE; DNA sequencing; ChIA-PET; 454 sequencing; Illumina sequencing; Transcription factor binding sites; Chromatin interactions; Chromosome Conformation Capture; Chromatin immunoprecipitation

## INTRODUCTION

Paired-End Tag sequencing (PET) is the process by which paired short tags are extracted from linear DNA fragments for high-throughput sequencing (Fullwood et al., 2009).

Subsequently, PET sequences are mapped onto the appropriate reference genome, so as to accurately define the identities of target DNA elements. As described in *UNIT 21.12*, the PET strategy has been developed and implemented as GIS-PET for transcriptome characterization and genome annotation (Ng et al., 2005), as well as ChIP-PET for the mapping of transcription factor binding sites (Wei et al., 2006).

PET technology has also been applied to the large-scale analysis of chromatin interaction networks through the development of the Chromatin Interaction Analysis using Paired-End Tag sequencing (ChIA-PET) method (Fig. 1). Essentially, ChIA-PET utilizes the proximity ligation concept, pioneered by the 3C method (Dekker et al., 2002), to capture interacting

<sup>3</sup>Corresponding author: Yijun Ruan, PhD, Genome Institute of Singapore, 60 Biopolis, Singapore 138672, T. 65 6478 8073; F. 65 6478 9059; ruanyj@gis.a-star.edu.sg.

<sup>2</sup>These authors made equal contributions to the manuscript.

DNA segments within DNA-protein complexes. Chromatin immunoprecipitation (ChIP) is used for the enrichment of specific chromatin fragments, while the PET strategy and high-throughput sequencing technologies permit the deep sequencing coverage necessary for analysis of complex proximity ligation mixtures. Unlike 3C and other 3C-based techniques such as 4C (Gondor et al., 2008) or 5C (Dostie et al., 2006), ChIA-PET does not depend on specific sites for detection, and thus provides a genome-wide and unbiased approach for the discovery of long-range chromatin interactions.

### Chromatin immunoprecipitation (ChIP)

The high complexity of substrate for proximity ligation inevitably leads to excessive non-specific noise, making the cost of sequencing such material to the required depth to find true proximity ligation products prohibitive even for the most advanced sequencing technology currently available. To reduce the level of complexity and background noise, ChIP is used against specific protein factors to enrich specific chromatin fragments of interest before proximity ligation (Fullwood and Ruan, 2009). This enrichment approach would not only make the ChIA-PET sequencing approach practical by reducing the complexity of the system to be examined, but would also add specificity to the identified interaction points. Furthermore, as in ChIP-PET, binding sites mediated by the protein factor will be detected (Wei et al., 2006). Hence, with the use of ChIP, ChIA-PET may potentially be applied to the identification of all chromatin interactions involved in a particular nuclear process. For instance, by targeting general transcription factors or RNA Polymerase II components for ChIP enrichment, ChIA-PET analysis may be used to uncover chromatin interactions involved in transcription regulation.

### Linker ligation and ditagging

Half-linker oligonucleotides are designed with a 5' overhang consisting of 4 nucleotides (GGCC) and a recognition site for the type II restriction enzyme MmeI (TCCAAC) (see Fig. 2). After ChIP enrichment, tethered DNA fragments in chromatin complexes are first ligated with an excess of half-linkers, and then circularized under dilute proximity ligation conditions such that interacting DNA fragments are connected by a complete linker sequence. Flanking restriction enzyme sites in the linker sequence allow for the extraction of tag-linker-tag constructs upon digestion by MmeI. To enable the purification of PET constructs by streptavidin-coated magnetic beads, each half-linker is modified with biotin. This biotin moiety is attached to the internal C6 of the 9<sup>th</sup> base (T) from the 5' end, so as to avoid steric hindrance in ligation. The purified ChIA-PET constructs can then be analyzed by high-throughput PET sequencing (Fig. 1a). When mapped to the reference genome, the ChIA-PET sequences are read out to detect the relationship of the two DNA fragments in chromatin interactions captured by proximity ligation (Fig. 1b).

ChIA-PET also allows for barcoding using two or more linker sequences with different nucleotide barcodes (Fig. 2, boxed sequences). As a linker sequence can include a unique nucleotide barcode, multiple linkers with distinctive nucleotide barcode sequences can be used to specify different experiments or replicates, as well as to monitor the non-specific chimeric ligation rate between different ChIP complexes (Fig. 1a, 1b). Hence, different

biological samples or replicates may be analyzed under similar experimental conditions in a time and cost-effective manner while reducing technical variations.

### High-throughput DNA sequencing

ChIA-PET constructs have been successfully sequenced using next-generation sequencing platforms such as Roche/454 GSFLX (Margulies et al., 2005) and the Illumina Genome Analyzer (GA) II. 454 GSFLX has long read lengths and good accuracies, but is fairly expensive and has a lower throughput as compared to the Illumina GAI, which has shorter read lengths and slightly lower base accuracies, but is cheaper and has a higher throughput (Holt and Jones, 2008).

Today, Illumina Paired-End Sequencing can be used to sequence PET constructs from both directions, hence allowing both tags to be read. This unit describes a library construction protocol for the generation of ChIA-PET constructs that can be sequenced on the 454 GSFLX and Illumina platforms; however, in principle, it is possible to adapt the protocol for other next-generation sequencing technologies, such as ABI SOLiD (Shendure et al., 2005) or Helicos single-molecule sequencing (Harris et al., 2008).

## BASIC PROTOCOL 1

### CONSTRUCTION OF A CHIA-PET LIBRARY

The protocol for ChIA-PET library construction is outlined in Fig. 1a. ChIP DNA is prepared before the start of library construction, as per standard procedures (*UNIT 21.3*). To ensure that there will be enough material for ChIA-PET library construction, and that the resulting library will be of sufficient complexity, it is recommended that ChIP DNA be generated from at least  $10^8$  cells.

First, ChIP DNA is end-blunted and ligated to biotinylated half-linkers containing flanking MmeI restriction sites. Intact chromatin complexes are then eluted off the beads and subjected to proximity ligation under extremely dilute conditions, such that interacting DNA fragments are preferentially ligated to one another. The ligation of two half-linkers creates a complete linker sequence at the interaction junction. After reverse cross-linking to remove DNA-associated proteins, MmeI digestion is carried out to release tag-linker-tag (PET) constructs, which are then purified by selective binding to streptavidin beads. The PET constructs are ligated with adapters for high-throughput sequencing. ChIA-PET libraries constructed using this method can be sequenced by either the 454 GSFLX or Illumina GAI platforms.

Details for oligonucleotide sequences used in this protocol are shown in Fig. 2.

#### Materials

ChIP DNA, bound to sepharose beads

Linkers A and B (Fig. 2)

Adapters A and B (Fig. 2)

Primers (Fig. 2)

T4 DNA Polymerase, with 10× reaction buffer (Promega)

10 mM dNTPs

Nuclease-free water

T4 DNA ligase (30 U/μl) (Fermentas)

5× T4 DNA Ligase Buffer with PEG (Invitrogen)

1 M Tris, pH 7.0 (*APPENDIX 2*)

0.5 M EDTA, pH 8.0 (*APPENDIX 2*)

5 M Sodium chloride

10% SDS (*APPENDIX 2*)

Wash buffer (see recipe)

T4 DNA Polynucleotide Kinase (10 U/μl) (NEB)

10× T4 DNA Ligase Buffer (NEB)

Elution buffer (see recipe)

EB buffer (Qiagen)

20% Triton X-100 (v/v)

Proteinase K (20 mg/ml) (Invitrogen)

GlycoBlue (15 mg/ml) (Ambion)

Absolute ethanol/ 75% ethanol (v/v)

MmeI (2 U/μl) (NEB)

S-adenosylmethionine (SAM) (NEB)

10× NEBuffer 4 (NEB)

Dynabeads M-280 Streptavidin (Invitrogen)

1× and 2× B&W buffer (see recipe)

*E. coli* DNA polymerase I (10 U/μl) (NEB)

10× NEBuffer 2 (NEB)

HotStarTaq Master Mix (Qiagen)

Phusion™ High-Fidelity PCR Master Mix with HF buffer (Finnzymes)

25 bp DNA ladder (Invitrogen)

4–20% TBE PAGE gel (10 wells) (Invitrogen)

10× TBE buffer (*APPENDIX 2*)

DNA loading dye

SYBR® Green I (Molecular Probes) (Invitrogen)  
QIAquick PCR Purification Kit (Qiagen)  
6% TBE PAGE gel (5 wells) (Invitrogen, by special request only)  
TE buffer, pH 8.0 (*APPENDIX 2*)  
Agilent DNA 1000 Kit (Agilent Technologies)  
0.2-ml PCR tubes  
0.6-ml tubes  
1.7-ml microcentrifuge tubes  
1.5-ml screw cap tubes  
DNA LoBind Tubes, 1.5-ml (Eppendorf)  
Microspin plastic centrifuge tube filter units (0.45- $\mu$ m, e.g., Costar Spin X)  
50-ml polypropylene conical tubes (e.g., Falcon)  
50-ml MaXtract High Density tubes (Qiagen)  
50-ml polypropylene copolymer (PPCO) centrifuge tubes (Nalgene)  
21-G needles  
Stainless steel sterile surgical blades  
96-well microtiter plates  
Novex Mini-Cell (Invitrogen)  
Intelli-Mixer RM-2L (Palico Biotech)  
Magnetic Particle Collector (Invitrogen)  
Dark Reader Transilluminator (Clare Chemical Research)  
Agilent 2100 Bioanalyzer (Agilent Technologies)  
Additional reagents and equipment for phenol/chloroform extraction and isopropanol precipitation (*UNIT 2.1A*), PicoGreen dsDNA quantitation (*APPENDIX 3D*) and PCR (*UNIT 15.1*).  
*Reagents and chemicals used should all be molecular biology grade (DNase-free).*

### Quantitate ChIP-DNA

- 1 Store ChIP DNA at 4°C in TE buffer until ready to begin library construction.  
*Take care to avoid protease contamination of ChIP material.*
- 2 Pellet beads by centrifuging for 3 min at 100  $\times g$  (800 rpm), 4°C. Remove and discard the storage buffer and top up to 50% in fresh, ice-cold TE buffer.
- 3 To 10% of beads (by volume), add 200  $\mu$ l elution buffer and incubate at 37°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) for 30 min.

- 4 Centrifuge at  $100 \times g$  (800 rpm), 4°C for 3 min and transfer the supernatant to a fresh 1.5-ml screw cap tube. Add another 200  $\mu$ l elution buffer to the beads and repeat Step 2. Transfer the second wash to the same tube.
- 5 Add 6  $\mu$ l Proteinase K to the tube and incubate at 37°C overnight.
- 6 Transfer reaction mix to a 2-ml Eppendorf Phase Gel Lock tube and purify by phenol/chloroform extraction, followed by isopropanol precipitation (see *UNIT 2.1A*). Resuspend DNA in 20  $\mu$ l EB buffer and quantitate by PicoGreen fluorometry (see *UNIT 21.12*, Support Protocol 3).

#### End-blunting of ChIP-DNA

- 7 Resuspend the beads by pipetting and aliquot a volume of bead suspension containing 50 – 100 ng DNA. Split beads equally into at least two tubes and pellet the beads by centrifuging for 3 min at  $100 \times g$  (800 rpm), 4°C. Leave tubes on ice to allow the beads to settle completely. Remove supernatant carefully without disturbing the beads.

*The final volume of beads in each tube should be 100 – 150  $\mu$ l. If the bead volume is less than 100  $\mu$ l, top up to 100  $\mu$ l with similarly treated blank sepharose beads to minimize loss of DNA-bearing beads in subsequent steps. Sawed-off tips or big bore tips should be used for pipetting beads.*

*With the appropriate protocol modifications (e.g. using a magnetic particle collector instead of centrifugation to wash the beads), magnetic beads may be substituted for sepharose beads.*

- 8 Resuspend the beads in the following reaction mix:
  - 100  $\mu$ l 10 $\times$  Buffer for T4 DNA polymerase
  - 10  $\mu$ l 10 mM dNTPs
  - 880  $\mu$ l nuclease-free water
  - 10.4  $\mu$ l 9.7 U/ $\mu$ l T4 DNA polymerase.Incubate at 37°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) for 20 min.

#### Ligate biotinylated half-linkers to ChIP-DNA

- 9 Pellet beads by centrifuging at  $100 \times g$  (800 rpm), 3 min at 4°C. Carefully remove and discard the supernatant. Wash the beads 3 times with 1 ml ice-cold wash buffer. Mix well by inverting tubes. Pellet the beads by centrifugation after every wash.
- 10 Set up the following reaction mix, adding reagents in order:
  - 786  $\mu$ l nuclease-free water
  - 10  $\mu$ l 200 ng/ $\mu$ l biotinylated half-linker (A or B)

200  $\mu$ l 5 $\times$  T4 DNA ligase buffer with PEG

4  $\mu$ l 30 U/ $\mu$ l T4 DNA ligase.

*Thaw linkers gently on ice. First mix linkers well with water, and then with PEG buffer, before adding ligase.*

Resuspend beads in the above reaction mix. Incubate at 16°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) overnight (~16 h).

### Add phosphate groups to 5' ends

**11** Combine the two tubes of linker-ligated DNA in a 1.5-ml screw cap tube. Remove excess linkers by washing 3 times with ice-cold wash buffer. Pellet the beads by centrifuging for 3 min at 100  $\times g$  (800 rpm), 4°C after every wash.

**12** Set up the following reaction mix:

100  $\mu$ l 10 $\times$  T4 DNA Ligase Buffer

880  $\mu$ l nuclease-free water

20  $\mu$ l 10 U/ $\mu$ l T4 DNA polynucleotide kinase.

Remove wash buffer and resuspend the beads with the above reaction mix. Incubate at 37°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) for 30 min.

### Elute chromatin complex off beads

**13** Pellet the beads by centrifuging for 5 min at 100  $\times g$  (800 rpm), 4°C. Leave the tube on ice to allow the beads to completely settle, and discard the supernatant.

**14** Add 200  $\mu$ l of elution buffer to the beads. Incubate at 37°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) for 30 min.

**15** Centrifuge at 100  $\times g$  (800 rpm), 4°C for 5 min and transfer the supernatant to a fresh tube. Wash the beads with 900  $\mu$ l of buffer EB and centrifuge again. Transfer the supernatant to the same tube.

**16** Transfer the collected eluate to the filter cup of 2 SpinX columns and centrifuge for 1 min at 16000  $\times g$  (13000 rpm), 4°C.

**17** Transfer the filtrate to a 1.5-ml screw cap tube and add 90  $\mu$ l of 20% Triton X-100. Incubate for 1 h at 37°C with rotation on the Intelli-Mixer (Program F8, 30 rpm).

### Circularize DNA fragments

**18** Prepare ligation mixture on ice as follows:

5 ml 10 $\times$  T4 DNA Ligase Buffer

43 ml deionized water

167  $\mu$ l 30 U/ $\mu$ l T4 DNA ligase.

Add DNA to ligation mixture and invert to mix well. Incubate at 22°C for 20 – 24 h without rotating.

*This is the critical proximity ligation step and must be carried out under extremely dilute conditions (eg. 50-ml ligation mixture for 50 – 100 ng starting ChIP DNA) to minimize ligations between different DNA-protein complexes.*

### Reverse cross-link and purify DNA

- 19** Add 500 µl Proteinase K (20 mg/ml) to ligation mixture. Incubate at 37°C overnight (~16 h) without rotation.

*It is critical that chromatin is fully reverse cross-linked and DNA-associated proteins are completely digested, or else DNA would be lost together with protein during phenol/chloroform extraction.*

- 20** Split each 50-ml ligation mix into 3 portions (~17 ml each) and transfer each portion into a 50-ml MaXtract High Density tube. Add an equal volume of phenol to each of the tubes, invert to mix well, and centrifuge at 1800 × *g* (3000 rpm) for 5 min, room temperature.

- 21** Transfer the upper aqueous phase to a 50-ml PPCO tube and add the following:

1.8 ml 3M sodium acetate, pH 5.2

10 µl GlycoBlue

19 ml isopropanol.

Incubate at –80°C, for at least 1 h. Allow the frozen solution to thaw for ~10 min before centrifuging for 45 min at 38730 × *g* (18000 rpm), 4°C. Carefully decant the supernatant and wash DNA pellet twice with 30 ml 75% ethanol. Allow the pellet to air dry in a laminar flow hood until residual ethanol is completely removed.

*Due to the large size of the PPCO tubes, it may take between 30 min to a few hours to completely dry the DNA pellet. Take care to avoid introducing any contaminants during this time.*

Dissolve DNA pellet in 34 µl of EB buffer.

*Resuspend the DNA pellet very thoroughly and scrape the bottom of the tube with a pipette tip to recover as much DNA as possible.*

### Perform Mmel digestion (tagging)

- 22** Prepare the following reaction mix on ice:

34 µl DNA (in EB buffer)

5 µl 10× NEBuffer 4

5 µl 10× SAM (1 µl SAM + 63 µl dH<sub>2</sub>O)

5  $\mu$ l non-biotinylated linker (200 ng/ $\mu$ l)

1  $\mu$ l 2 U/ $\mu$ l MmeI.

Mix by pipetting and incubate at 37°C,  $\geq$ 2 hours without rotation.

*MmeI can be self-inhibitory in excess; hence, non-biotinylated linker, which contains the MmeI restriction site, is used to quench its activity. 10 $\times$  SAM should be prepared fresh, as it is unstable. The 10 $\times$  concentration should be 500 $\mu$ M.*

### Immobilize ChIA-PETs on Dynabeads

- 23** Transfer 50  $\mu$ l of resuspended Dynabeads to a 1.5-ml tube. Using a Magnetic Particle Collector (MPC), wash the beads twice with 150  $\mu$ l of 2 $\times$  B&W buffer. Resuspend beads in 50  $\mu$ l of 2 $\times$  B&W buffer.

*We recommend the use of Eppendorf LoBind tubes to minimize bead binding to the surfaces of tubes.*

- 24** Add 50  $\mu$ l digestion mix to the beads and mix well. Incubate at 22°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) for 30 min.
- 25** Remove the supernatant and wash the beads twice with 150  $\mu$ l of 1 $\times$  B&W buffer.

### Ligate 454 adapters

- 26** Prepare ligation mix on ice.
- 5  $\mu$ l 10 $\times$  T4 DNA ligase buffer
- 8  $\mu$ l GS20 Adapter A
- 8  $\mu$ l GS20 Adapter B
- 28  $\mu$ l nuclease-free water.
- 27** Remove 1 $\times$  B&W buffer from beads. Add ligation mix to beads, followed by 1  $\mu$ l 30 U/ $\mu$ l T4 DNA ligase and pipette to mix well. Incubate at 22°C overnight (~16 hours) with rotation on the Intelli-Mixer (Program F8, 30 rpm).

### Nick translation

- 28** Wash the beads twice with 150  $\mu$ l of 1 $\times$  B&W buffer.
- 29** Prepare the following reaction mix on ice:
- 5  $\mu$ l 10 $\times$  NEBuffer 2
- 2.5  $\mu$ l 10 mM dNTPs
- 38.5  $\mu$ l nuclease-free water
- 4  $\mu$ l 10 U/ $\mu$ l *E.coli* DNA polymerase I.

Remove 1× B&W buffer from beads and add the enzyme mix. Mix by pipetting and incubate at 22°C, 2 hours with rotation on the Intelli-Mixer (Program F8, 30 rpm).

### PCR amplification of ChIA-PETs

- 30** Remove reaction mix and wash the beads twice with 150 µl of 1× B&W buffer using the MPC. Resuspend the beads in 50 µl of EB buffer and transfer to a fresh 1.5-ml tube.

*Beads in EB buffer may be stored at –20°C for several months.*

- 31** For quality checking purposes, set up the following 50-µl PCR reaction, varying the volume of beads or the number of cycles to find the optimum PCR conditions for your library.

1 µl – 4 µl beads suspension  
 1 µl 10 µM Primer A  
 1 µl 10 µM Primer B (biotin)  
 25 µl 2× HotStarTaq Master Mix  
 Nuclease-free water to 50 µl.

Carry out hot-start PCR using the following thermal cycling parameters:

Initial step:	15 min, 95°C (denaturation)
18 – 25 cycles:	30 sec, 95°C (denaturation) 1 min, 72°C (annealing) 45 sec, 72°C (extension)
Final step:	5 min, 72°C (final extension)

- 32** Add loading buffer to PCR products and analyze by electrophoresis on a TBE PAGE gel. Visualize by SYBR Green staining.

*If the PCR is successful, there should be a prominent, well-defined ChIA-PET band of approximately 166 bp. Refer to ANTICIPATED RESULTS for positive and negative examples of quality control data from this step.*

- 33** Using the appropriate conditions (as determined in Step 31), scale up the number of PCR reactions.

*Depending on the intensity of the ChIA-PET band observed in Step 32, we suggest scaling up to 8 – 16 reactions.*

## Purify ChIA-PET DNA

- 34** Purify the PCR products using QIAquick PCR Purification Kit. Run purified PCR products in a 6% TBE PAGE gel and visualize by SYBR Green staining. Carefully excise the 166 bp band from the gel.

*It is advisable to use the Dark Reader transilluminator for gel visualization, as exposure to UV light (especially short-wavelength UV) will damage DNA.*

- 35** Place excised gel fragments in 0.6-ml microcentrifuge tubes that have been pierced at the bottom with a 21-G needle. Place each tube inside a 1.5-ml screw cap tube and centrifuge at maximum speed for 5 min at 4°C.

*The gel pieces are shredded during centrifugation and collected at the bottom of each 1.5-ml tube.*

- 36** Add 200 µl of TE buffer to each tube. Make sure the gel pieces are fully immersed in the buffer, and then freeze at –80°C for ≥1 h, followed by incubation at 37°C overnight.
- 37** Transfer the gel pieces together with the buffer in each tube to the filter cup of a microspin centrifuge tube (eg. SpinX). Centrifuge at maximum speed for 10 min at 4°C. At the same time, rinse each 1.5-ml tube with 200 µl TE buffer and transfer the rinsing buffer to each filter unit upon completion of the first spin.

- 38** Pool the filter-through and carry out isopropanol precipitation as follows:

~430 µl DNA solution

43 µl 3 M sodium acetate, pH 5.2

2 µl GlycoBlue

430 µl isopropanol

Freeze at –80°C for 1 h, and then centrifuge for at least 30 min at 16000 × *g* (13000 rpm), 4°C.

- 39** Wash the DNA pellet twice with 500 µl 75% ethanol. Air dry the pellet and resuspend in 21 µl TE buffer.

*A phenol/chloroform extraction step prior to isopropanol precipitation is optional.*

- 40** Use 1 µl of sample to determine the quantity of ChIA-PET DNA by PicoGreen quantitation. Perform a quality control check using an Agilent 2100 Bioanalyzer with a DNA 1000 kit, according to manufacturer's instructions.

*There should be a well-defined, intense electropherogram peak corresponding to the fragment size of interest. See ANTICIPATED RESULTS for positive and negative examples of Agilent assay data.*

*At this point, the ChIA-PET DNA is ready to be processed for 454 sequencing on a GSFLX genome sequencer (Roche) according to the manufacturer's library*

*preparation protocol (not described here). Alternatively, the library can be converted for Illumina sequencing on Illumina's Genome Analyzer II (Steps 41 – 45). This is recommended as Illumina sequencing has lower per-unit costs and a higher throughput; also, Illumina Paired-End Sequencing is capable of reading tags from both ends of each ChIA-PET construct and thus is an ideal fit for the ChIA-PET method.*

### Convert 454 library for Illumina sequencing

- 41** Prepare 10 PCR reactions. Each reaction should contain the following:

2 ng purified DNA (from Step 40)

1  $\mu$ l 25  $\mu$ M Illumina 1–454 primer

1  $\mu$ l 25  $\mu$ M Illumina 2–454 primer

25  $\mu$ l 2 $\times$  Phusion Master Mix

Nuclease-free water to 50  $\mu$ l.

Carry out PCR using the following thermal cycling parameters:

Initial step:	30 sec, 98°C (denaturation)
15 cycles:	10 sec, 98°C (denaturation)
	30 sec, 65°C (annealing)
	30 sec, 72°C (extension)
Final step:	5 min, 72°C (final extension)

- 42** Pool together all reactions and purify PCR products using the QIAquick PCR Purification Kit. Separate PCR products on 6% TBE PAGE gel and visualize by SYBR Green staining.

*If the PCR is successful, there should be a prominent, well-defined ChIA-PET band of approximately 223 bp.*

- 43** Excise the 223 bp band from the gel and purify using the gel-crush method, as detailed in steps 35 – 39. Resuspend DNA in 15  $\mu$ l of TE buffer.
- 44** Quantitate purified DNA by PicoGreen fluorometry and check library quality using the Agilent 2100 Bioanalyzer DNA 1000 kit, according to manufacturer's instructions.

*There should be a single intense electropherogram peak corresponding to the fragment size of interest. Sample Agilent profiles can be found in ANTICIPATED RESULTS.*

- 45** Submit the library for Illumina Paired-End Sequencing.

*Important: Specific instructions for sample preparation should be obtained in consultation with your sequencing facility or service provider.*

## SUPPORT PROTOCOL 1

### PREPARATION OF LINKERS AND ADAPTERS

Linkers and adapters may be prepared beforehand and stored for several months at  $-20^{\circ}\text{C}$ . Single-stranded oligonucleotides are ordered desalted. Unmodified oligonucleotides may either be ordered PAGE- or HPLC-modified, while biotinylated oligonucleotides are ordered HPLC-purified.

1. Adjust oligonucleotides to a final concentration of  $100\ \mu\text{M}$  in  $1\times$  TNE prior to annealing. Refer to *UNIT 21.12*, Support Protocol 2 for the annealing procedure.
2. After annealing, run linkers (100 ng) together with their component ssDNA oligonucleotides (100 ng) and a 25 bp DNA ladder on a 4–20% TBE PAGE gel. A successfully annealed linker should show up as a single, well-defined band of the expected molecular weight, with no lower molecular weight bands indicating excess oligonucleotides.

*If excess oligonucleotides are present, repeat the annealing procedure with different ratios of oligonucleotides to determine the best ratio for maximum linker purity.*

## SUPPORT PROTOCOL 2

### VALIDATION OF LINKERS

It is important to evaluate the quality of half-linkers as they are critical for the success of ChIA-PET experiments. It is recommended that all new batches of annealed half-linkers be validated by this protocol before they are used in ChIA-PET library construction.

*This protocol is based on the main ChIA-PET protocol, the most important modification being that instead of ChIP DNA attached to beads, the substrate for linker ligation is a known DNA fragment in solution. Here, the use of only 1 DNA fragment is described, but multiple DNA fragments may be used to assess background chimerism levels in various volumes of ligation.*

### Materials

- Annealed half-linkers A and B (Fig. 2)
- DNA fragment X of known size (blunt-ended, 5'-phosphorylated,  $\sim 1 - 3$  kb)
- Adapters A and B (Fig. 2)
- Primers (Fig. 2)
- 10 mM dNTPs
- Nuclease-free water
- T4 DNA ligase (30 U/ $\mu\text{l}$ ) (Fermentas)
- 5 $\times$  T4 DNA Ligase Buffer with PEG (Invitrogen)
- T4 DNA Polynucleotide Kinase (10 U/ $\mu\text{l}$ ) (NEB)

10× T4 DNA Ligase Buffer (NEB)  
EB buffer (Qiagen)  
GlycoBlue (15 mg/ml) (Ambion)  
Absolute ethanol/ 75% ethanol (v/v)  
MmeI (2 U/μl) (NEB)  
S-adenosylmethionine (SAM) (NEB)  
10× NEBuffer 4 (NEB)  
Dynabeads M-280 Streptavidin (Invitrogen)  
1× and 2× B&W buffer (see recipe)  
*E. coli* DNA polymerase I (10 U/μl) (NEB)  
10× NEBuffer 2 (NEB)  
HotStarTaq Master Mix (Qiagen)  
25 bp DNA ladder (Invitrogen)  
4–20% TBE PAGE gel (10 wells) (Invitrogen)  
10× TBE buffer (*APPENDIX 2*)  
DNA loading dye  
QIAquick PCR Purification Kit (Qiagen)  
TE buffer, pH 8.0 (*APPENDIX 2*)  
0.2-ml PCR tubes  
0.6-ml tubes  
1.7-ml microcentrifuge tubes  
DNA LoBind Tubes, 1.5-ml (Eppendorf)  
Novex Mini-Cell (Invitrogen)  
Intelli-Mixer RM-2L (Palico Biotech)  
Magnetic Particle Collector (Invitrogen)

### **Ligate half-linkers to DNA fragment X**

- 1 Set up the following reaction mix, using separate tubes for each half-linker:
  - 100 ng DNA fragment X
  - 0.5 μl 200 ng/μl biotinylated half-linker
  - 20 μl 5× T4 DNA ligase buffer with PEG
  - 1 μl 30 U/μl T4 DNA ligase

Nuclease-free water to 100  $\mu$ l.

Incubate at 16°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) overnight (~16 h).

- 2 Purify the ligation mix and remove excess half-linkers by using the QIAquick PCR Purification Kit. Elute in 50  $\mu$ l of EB buffer.

#### Add phosphate groups to 5' ends

- 3 Set up the following reaction mix:

~48  $\mu$ l DNA

5.5  $\mu$ l 10 $\times$  T4 DNA Ligase Buffer

0.5  $\mu$ l nuclease-free water

1  $\mu$ l 10 U/ $\mu$ l T4 DNA polynucleotide kinase.

Incubate at 37°C for 30 min.

#### Self-circularize linker-ligated DNA fragments

- 4 Combine the two reaction mixes in a single 1.7-ml tube.

- 5 Prepare ligation mixture on ice as follows:

20  $\mu$ l 10 $\times$  T4 DNA Ligase Buffer

69  $\mu$ l nuclease-free water

1  $\mu$ l 30 U/ $\mu$ l T4 DNA ligase.

Add 110  $\mu$ l DNA (from Step 4) to ligation mixture and invert to mix well.

Incubate at 16°C overnight without rotating.

#### Purify DNA

- 6 Purify the reaction mix by phenol/chloroform extraction.

- 7 Transfer the upper aqueous phase (~200  $\mu$ l) to a 1.7-ml tube and add the following:

20  $\mu$ l 3M sodium acetate, pH 5.2

1  $\mu$ l GlycoBlue

~200  $\mu$ l isopropanol.

Incubate at -80°C for at least 1 h. Centrifuge at maximum speed for 30 min, 4°C. Carefully remove the supernatant and wash DNA pellet twice with 700  $\mu$ l 75% ethanol. Allow the pellet to air dry and resuspend in 20  $\mu$ l of EB buffer.

#### Perform Mmel digestion

- 8 Prepare the following reaction mix on ice:

20  $\mu$ l DNA (in EB buffer)

- 4  $\mu\text{l}$  10 $\times$  NEBuffer 4
- 4  $\mu\text{l}$  10 $\times$  SAM (1  $\mu\text{l}$  SAM + 63  $\mu\text{l}$  dH<sub>2</sub>O)
- 4  $\mu\text{l}$  non-biotinylated linker (200 ng/ $\mu\text{l}$ )
- 7  $\mu\text{l}$  nuclease-free water
- 1  $\mu\text{l}$  2 U/ $\mu\text{l}$  MmEI.

Mix by pipetting and incubate at 37°C,  $\geq$ 2 hours without rotation.

### Immobilize ChIA-PETs on Dynabeads

- 9 Transfer 50  $\mu\text{l}$  of resuspended Dynabeads to a 1.5-ml tube. Using a Magnetic Particle Collector (MPC), wash the beads twice with 150  $\mu\text{l}$  of 2 $\times$  B&W buffer. Resuspend beads in 40  $\mu\text{l}$  of 2 $\times$  B&W buffer.

*We recommend the use of Eppendorf LoBind tubes to minimize bead binding to the surfaces of tubes.*

- 10 Add 40  $\mu\text{l}$  of digestion mix to the beads and mix well. Incubate at 22°C with rotation on the Intelli-Mixer (Program F8, 30 rpm) for 30 min.
- 11 Remove the supernatant and wash the beads twice with 150  $\mu\text{l}$  of 1 $\times$  B&W buffer.

### Ligate 454 adapters

- 12 Prepare ligation mix on ice.
  - 5  $\mu\text{l}$  10 $\times$  T4 DNA ligase buffer
  - 8  $\mu\text{l}$  200 ng/ $\mu\text{l}$  Adapter A
  - 8  $\mu\text{l}$  200 ng/ $\mu\text{l}$  Adapter B
  - 28  $\mu\text{l}$  nuclease-free water.
- 13 Remove 1 $\times$  B&W buffer from beads. Add ligation mix to beads, followed by 1  $\mu\text{l}$  30 U/ $\mu\text{l}$  T4 DNA ligase and pipette to mix well. Incubate at 22°C overnight (~16 hours) with rotation on the Intelli-Mixer (Program F8, 30 rpm).

### Nick translation

- 14 Wash the beads twice with 150  $\mu\text{l}$  of 1 $\times$  B&W buffer and prepare the following reaction mix on ice:
  - 5  $\mu\text{l}$  10 $\times$  NEBuffer 2
  - 2.5  $\mu\text{l}$  10 mM dNTPs
  - 38.5  $\mu\text{l}$  nuclease-free water
  - 4  $\mu\text{l}$  10 U/ $\mu\text{l}$  *E.coli* DNA polymerase I.

Remove 1× B&W buffer from beads and add the nick translation reaction mix to the beads. Pipette well to mix and incubate at 22°C, 2 hours with rotation on the Intelli-Mixer (Program F8, 30 rpm).

### PCR amplification of ChIA-PETs

**15** Remove reaction mix and wash the beads twice with 150 µl of 1× B&W buffer using the MPC. Resuspend the beads in 50 µl of EB buffer and transfer to a fresh 1.5-ml tube.

**16** Set up the following PCR reaction:

- 2 µl beads suspension
- 1 µl 10 µM Primer A
- 1 µl 10 µM Primer B
- 25 µl 2× HotStarTaq Master Mix
- Nuclease-free water to 50 µl.

Carry out hot-start PCR using the following thermal cycling parameters:

Initial step:	15 min, 95°C (denaturation)
18 – 25 cycles:	30 sec, 95°C (denaturation)
	1 min, 72°C (annealing)
	45 sec, 72°C (extension)
Final step:	5 min, 72°C (final extension)

**17** Add loading buffer to PCR products and analyze by electrophoresis on a TBE PAGE gel.

*This is the most essential QC step. If the linkers are functional, there should be a prominent, well-defined ChIA-PET band of approximately 166 bp, similar to that of the main protocol (see ANTICIPATED RESULTS).*

**OPTIONAL: Capillary sequencing:** To further verify the fidelity of linker sequences, ChIA-PETs may be purified (*BASIC PROTOCOL 1*, Steps 34 – 40), cloned and processed for capillary sequencing according to standard procedures (*UNIT 7.4A*).

## REAGENTS AND SOLUTIONS

Use deionized, autoclaved water in all recipes. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

### 1× B&W buffer

- For 50 ml:
- 250 µl 1 M Tris-HCl, pH 7.5 (final 5 mM)

50  $\mu$ l 0.5 M EDTA, pH 8.0 (final 0.5 mM)  
10 ml 5 M sodium chloride (final 1 M)  
39.7 ml dH<sub>2</sub>O

### **2× B&W buffer**

For 50 ml:  
500  $\mu$ l 1 M Tris-HCl, pH 7.5 (final 10 mM)  
100  $\mu$ l 0.5 M EDTA, pH 8.0 (final 1 mM)  
20 ml 5 M sodium chloride (final 2 M)  
29.4 ml dH<sub>2</sub>O

### **1× TNE buffer**

For 50 ml:  
500  $\mu$ l 1 M Tris-HCl, pH 7.5 (final 10 mM)  
10  $\mu$ l 0.5 M EDTA, pH 8.0 (final 0.1 mM)  
500  $\mu$ l 5 M sodium chloride (final 50 mM)  
48.99 ml dH<sub>2</sub>O

### **Wash buffer**

For 50 ml:  
500  $\mu$ l 1 M Tris-HCl, pH 7.5 (final 10 mM)  
100  $\mu$ l 0.5 M EDTA, pH 8.0 (final 1 mM)  
5 ml 5 M sodium chloride (final 500 mM)  
44.4 ml dH<sub>2</sub>O

### **Elution buffer**

For 50 ml:  
49.5 ml TE buffer (*APPENDIX 2*)  
500  $\mu$ l SDS (1%)

*All buffers stable for years at room temperature.*

## **COMMENTARY**

### **Background Information**

The spatial conformation of chromatin plays an important role in genome regulation, and various methods have been developed to facilitate its study. Electron microscopy (Su et al., 1990) and atomic force microscopy (Yoshimura et al., 2004) have been used to visualize

DNA loop structures at high resolution. However, the harsh sample preparation requirements make such techniques unsuitable for the study of most *in vivo* interactions. Fluorescence in situ hybridization (FISH) approaches have been applied to the visualization of long-range interactions and chromosome structure (Branco and Pombo, 2006; Carter et al., 2002; Cremer and Cremer, 2001; Osborne et al., 2004); however, they are limited by low resolution and cannot be used to detect chromatin interactions on a kilobase scale. RNA TRAP, an advancement in FISH-based methods that allows high-resolution characterization of chromatin interactions, is limited to the study of transcriptionally active genes in the cell of interest and cannot give global information (Carter et al., 2002). In addition, all microscopy-based techniques have the disadvantage of being able to probe only one site at a time. The Chromosome Conformation Capture (3C) method (Dekker, 2006; Dekker et al., 2002; Simonis et al., 2007) and its variants such as ChIP-3C (Cai et al., 2006; Horike et al., 2005), 4C (Gondor et al., 2008; Simonis et al., 2006; Simonis et al., 2007; Zhao et al., 2006), 5C (Dostie et al., 2006; Simonis et al., 2007) and 6C (Tiwari et al., 2008) have been very useful for the high-resolution analysis of *in vivo* long-range chromatin interactions. However, 3C, ChIP-3C, 4C and 5C only provide a partial genome assay for known interactions and are incapable of *de novo* detection of interactions in a whole-genome context. While 6C is capable of *de novo* detection, it cannot be used to study chromatin interactions on a genome-wide scale due to its reliance on cloning, which remains a low-throughput and laborious technique.

The ChIA-PET technique offers a new strategy for mapping chromatin interaction networks on a global scale. The implementation of ChIP in ChIA-PET has the advantage of reducing library complexity and background noise, as well as adding specificity to chromatin interactions and enabling the examination of specific chromatin interactions mediated by particular transcription factors. Data generated from ChIA-PET analysis can be used to globally map protein factor binding sites as well as to construct a genome-wide chromatin interactome map associated with a specific protein factor of interest.

### Critical Parameters and Troubleshooting

**ChIP DNA**—For best results, it is critical that the starting ChIP material be of high quality. The reason for this is that interactions tend to occur rarely, the higher the ChIP enrichment, the less sequencing is required to detect the interactions. Poor ChIP enrichment usually results in low quality libraries that yield few binding sites and almost no interactions upon sequencing. Prior to the start of ChIA-PET library construction, ChIP enrichment should be verified by real-time quantitative PCR using primers specific for known binding sites and normalized against results obtained with background primers. Fold enrichment is calculated by comparing the normalized RT-qPCR results for both ChIP and input DNA derived from the same batch of chromatin. If enrichment levels are low, ChIP conditions must be optimized for the antibody and cell type used.

ChIP DNA-bearing beads should preferably be used fresh for library construction, as long-term storage would likely result in degradation of chromatin proteins or a significant fraction of ChIP DNA detaching from the beads, making the sample less useful for library construction. If necessary, beads can be stored at 4°C for up to 4 weeks.

We recommend the use of 50 – 100 ng ChIP DNA as starting material. This is such that there would be sufficient sample for high ChIA-PET library complexity, while maintaining a low enough concentration of DNA during the proximity ligation step (*STEP 18*) to ensure adequate separation of chromatin complexes, hence minimizing chimeric ligations (see below).

**Analysis of chimerism**—During the proximity ligation step, a significant fraction of ligations occur in a non-specific and random manner between DNA fragments from different ChIP complexes. These are chimeric ligations and do not represent true *in vivo* chromatin interactions. Hence, to evaluate the quality of data from any ChIA-PET experiment, it is important to estimate the overall rate of chimerism. This can be done by designing two different half-linkers with specific nucleotide barcodes A and B. The linkers are ligated to two separate aliquots of the same ChIP DNA sample, which are then combined for proximity ligation and processed for PET sequencing. A portion of the PETs derived from chimeric ligation products can be identified by their A/B linker composition. By calculating the proportion of chimeras (number of PETs containing A/B linker sequences/ total number of uniquely mapped PETs × 100%), one can estimate the overall level of chimerism in a ChIA-PET library. A successful library should have an A/B chimerism ratio of less than 30%. However, since the ChIA-PET data analysis procedure detects chromatin interactions by identifying clusters of multiple overlapping PETs, it can be assumed that chimeric ligations, which distribute randomly throughout the genome and hence are present in the dataset mainly as singleton PETs, would not lead to high levels of false positives. Therefore, libraries with higher chimerism ratios can still be used for the analysis of chromatin interactions, albeit at lower sequencing efficiencies.

If linkers with different lengths are used, ChIA-PETs of different sizes will be generated, which may lead to biased gel excision of PETs and a subsequent imbalance in linker ratio. As the Illumina GAII sequencer has been observed to have a GC bias, barcoding linkers should be designed with the same GC content. The two linkers described in this protocol, A and B, are identical in length and GC content but differ in sequence at four nucleotides: CTTA for Linker A, and ACAT for Linker B (see Fig. 2).

**General notes on sample handling**—As ChIA-PET library construction involves extensive manipulation of very small amounts of DNA, care must be taken to minimize loss of nucleic acid. Hence, it is advisable to use siliconized or low-binding microcentrifuge tubes for sample handling, MaXtract High Density tubes to reduce loss of DNA-containing aqueous phase during phenol/chloroform extraction, and GlycoBlue to maximize recovery of DNA during isopropanol precipitation. Note that the presence of Glycoblue after DNA purification by isopropanol precipitation precludes spectrophotometric quantitation. Instead, the use of PicoGreen fluorometry (see *UNIT 21.12*, Support Protocol 3) is recommended.

Precautions should also be taken to prevent contamination of samples by nucleases and proteases.

**Proximity ligation**—Proximity ligation (*STEP 18*) must be carried out under extremely dilute conditions to minimize ligations between non-interacting DNA fragments from

different ChIP complexes. A ligation volume of 50 ml is strongly recommended, although smaller volumes may be used for testing purposes (e.g., validating new linkers).

**Quality control**—This is a crucial part of the protocol as it verifies that previous steps have been correctly performed and that a library is of sufficiently high quality for high-throughput sequencing. For this, a ChIA-PET library should meet 2 basic quality control criteria: 1) the initial diagnostic gel run of the PCR-amplified library (*STEP 32*) must reveal a prominent and well-defined ChIA-PET band of the correct size; 2) the library should yield a single, intense electropherogram peak upon Agilent DNA 1000 analysis (*STEP 44*).

PCR optimization should be carried out if the first diagnostic gel run shows a weak or absent ChIA-PET band. The number of cycles may be increased, although it is not recommended that more than 25 cycles be used as doing so may reduce the complexity of the resulting library. Annealing temperature or elongation time may also be adjusted. Alternatively, the amount of template DNA may be increased by using a larger volume of beads in each PCR reaction.

Sample impurity due to contamination with other DNA species during the gel excision step may result in multiple minor peaks in the Agilent QC assay. To minimize possible sources of contamination, take care to excise only the middle portion of the ChIA-PET band and use a fresh razor blade for each excision.

**Data analysis**—ChIA-PET sequences obtained from high-throughput sequencing are extracted and mapped to a reference genome using an automated pipeline. A software package called PET-TOOL has been developed for the processing of Sanger and Roche/454-sequenced PETs (Chiu et al., 2006), while data obtained by Illumina sequencing can be processed through Illumina's ELAND program (Chen et al., 2008).

To identify putative interactions and binding sites, data analysis is performed according to the following principles: first, as the chromatin is sonicated, the probability of generating exactly identical DNA fragments is low; hence any redundant PETs are considered to be copies amplified during PCR processes. Therefore, only nonredundant distinct PETs are used for further analysis. Secondly, non-specific antibody binding and ligation means that there is a substantial amount of noise in any given ChIA-PET library. To distinguish true signals from noise, we assume that inter-ligation PETs derived from a specific interaction between two DNA regions would be enriched by the ChIP procedure and hence would be sampled more frequently as compared to non-specific PETs, which occur by random chance and would be sampled much less frequently. Furthermore, specific interactions are expected to yield inter-ligation PETs that overlap with each other to form a cluster, whereas PETs derived from non-specific fragments are randomly distributed in the genome as background PETs. Hence, clusters of multiple overlapping inter-ligation PETs would signify real interaction signals, as opposed to random background noise, which is typically represented by singleton PETs (see Fig. 1c).

## Anticipated Results

Starting from approximately 50 ng of ChIP DNA, one should be able to obtain at least 100 ng of ChIA-PETs after 454 scale-up PCR (*STEP 40*), and at least 100 ng of ChIA-PETs after Illumina conversion (*STEP 44*). Technical variations during library construction and PCR amplification will result in variable yields, but the yields stated above are more than sufficient for Roche/454 and Illumina sequencing. Libraries that pass the quality control checks shown in Figures 3 – 5 can proceed to high-throughput sequencing. One full Illumina run on a successful ChIA-PET library should result in a whole genome map of strong binding sites and strong chromatin interactions between the binding sites (see Fig. 6), with a distribution of genomic spans similar to that shown in Figure 7a. Figure 8 shows different types of interaction and binding site data that can be obtained from ChIA-PET. If the ChIA-PET library is of sufficient complexity, deeper sequencing would reveal more binding sites and chromatin interactions. All tracks were viewed using the ChIA-PET genome browser (<http://cms1.gis.a-star.edu.sg>, username: “guest”, password: “gisimsgtb”), which was adapted based on the “generic genome browser” system (Stein et al., 2002).

The reagent cost of constructing and sequencing a ChIA-PET library for one full Illumina run is approximately US\$10,000.

## Time Considerations

For a typical ChIA-PET procedure starting from ChIP-enriched chromatin, it takes about 7 days to construct a library and a further 3 days to prepare the ChIA-PET library for Illumina sequencing. The time required for high-throughput sequencing depends on the availability of sequencing platforms. At several points during the procedure, experiments may be temporarily halted and samples may be stored at  $-20^{\circ}\text{C}$ . These pause points have been indicated in the protocol.

## Acknowledgments

The authors acknowledge the Genome Technology and Biology Group at the Genome Institute of Singapore for helpful comments on the manuscript, in particular Ms Hui Shan Sim, Ms Huay Mei Poh, Ms Yufen Goh and Dr Ziyad Jhumka. Y.R and C.L.W are supported for this work by A\*STAR of Singapore and NIH ENCODE grants (R01 HG004456-01 and R01HG003521-01). M.J.F. and Y.H. are supported by A\*STAR National Science Scholarships.

## Literature Cited

- Branco MR, Pombo A. 2006; Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* 4:e138. [PubMed: 16623600]
- Cai S, Lee CC, Kohwi-Shigematsu T. 2006; SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat Genet.* 38:1278–1288. [PubMed: 17057718]
- Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P. 2002; Long-range chromatin regulatory interactions in vivo. *Nat Genet.* 32:623–626. [PubMed: 12426570]
- Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, et al. 2008; Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell.* 133:1106–1117. [PubMed: 18555785]

- Chiu KP, Wong CH, Chen Q, Ariyaratne P, Ooi HS, Wei CL, Sung WK, Ruan Y. 2006; PET-Tool: a software suite for comprehensive processing and managing of Paired-End diTag (PET) sequence data. *BMC Bioinformatics*. 7:390. [PubMed: 16934139]
- Cremer T, Cremer C. 2001; Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet*. 2:292–301. [PubMed: 11283701]
- Dekker J. 2006; The three ‘C’ s of chromosome conformation capture: controls, controls, controls. *Nat Methods*. 3:17–21. [PubMed: 16369547]
- Dekker J, Rippe K, Dekker M, Kleckner N. 2002; Capturing chromosome conformation. *Science*. 295:1306–1311. [PubMed: 11847345]
- Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, et al. 2006; Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res*. 16:1299–1309. [PubMed: 16954542]
- Fullwood MJ, Ruan Y. 2009;ChIP-based methods for the identification of long-range chromatin interactions. *J Cell Biochem*.
- Fullwood MJ, Wei CL, Liu ET, Ruan Y. 2009; Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res*. 19:521–532. [PubMed: 19339662]
- Gondor A, Rougier C, Ohlsson R. 2008; High-resolution circular chromosome conformation capture assay. *Nat Protoc*. 3:303–313. [PubMed: 18274532]
- Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, Dimeo J, Efcavitch JW, et al. 2008; Single-molecule DNA sequencing of a viral genome. *Science*. 320:106–109. [PubMed: 18388294]
- Holt RA, Jones SJ. 2008; The new paradigm of flow cell sequencing. *Genome Res*. 18:839–846. [PubMed: 18519653]
- Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. 2005; Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat Genet*. 37:31–40. [PubMed: 15608638]
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al. 2005; Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 437:376–380. [PubMed: 16056220]
- Ng P, Wei CL, Sung WK, Chiu KP, Lipovich L, Ang CC, Gupta S, Shahab A, Ridwan A, Wong CH, et al. 2005; Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nat Methods*. 2:105–111. [PubMed: 15782207]
- Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, et al. 2004; Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*. 36:1065–1071. [PubMed: 15361872]
- Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM. 2005; Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*. 309:1728–1732. [PubMed: 16081699]
- Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W. 2006; Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet*. 38:1348–1354. [PubMed: 17033623]
- Simonis M, Kooren J, de Laat W. 2007; An evaluation of 3C-based methods to capture DNA interactions. *Nat Methods*. 4:895–901. [PubMed: 17971780]
- Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, et al. 2002; The generic genome browser: a building block for a model organism system database. *Genome Res*. 12:1599–1610. [PubMed: 12368253]
- Su W, Porter S, Kustu S, Echols H. 1990; DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. *Proc Natl Acad Sci U S A*. 87:5504–5508. [PubMed: 2164685]
- Tiwari VK, Cope L, McGarvey KM, Ohm JE, Baylin SB. 2008; A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. *Genome Res*. 18:1171–1179. [PubMed: 18502945]

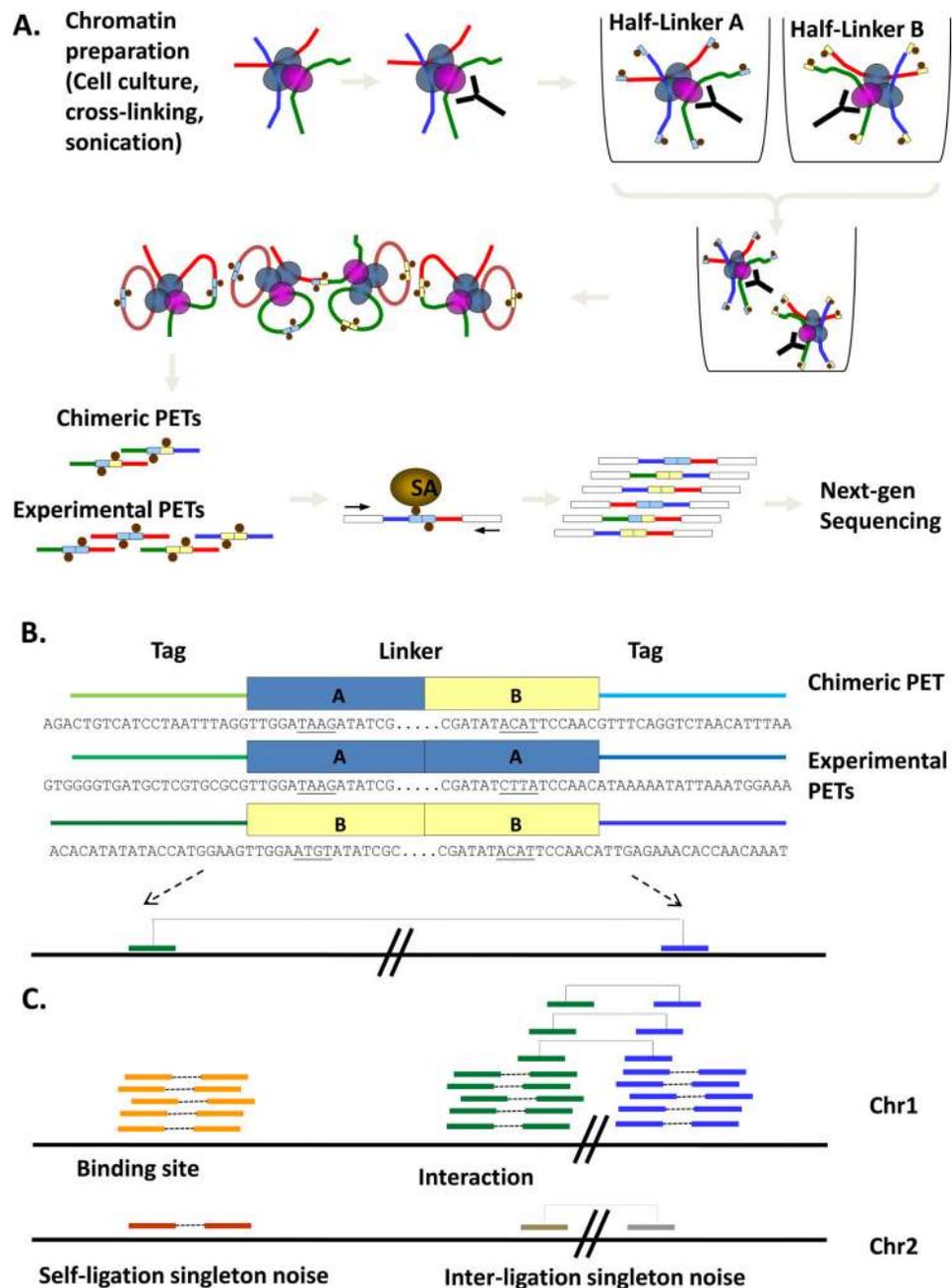
- Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, et al. 2006; A global map of p53 transcription-factor binding sites in the human genome. *Cell*. 124:207–219. [PubMed: 16413492]
- Yoshimura SH, Maruyama H, Ishikawa F, Ohki R, Takeyasu K. 2004; Molecular mechanisms of DNA end-loop formation by TRF2. *Genes Cells*. 9:205–218. [PubMed: 15005708]
- Zhao Z, Tavoosidana G, Sjolinder M, Gondor A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, et al. 2006; Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet*. 38:1341–1347. [PubMed: 17033624]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 1. Schematic overview of ChIA-PET method**

(A) Outline of ChIA-PET library construction procedure. Chromatin samples from cell cultures are cross-linked, sonicated and immunoprecipitated. Separate aliquots of ChIP DNA are ligated to barcoded half-linkers and proximity ligation is carried out. PETs are released by restriction digest, purified on streptavidin-coated magnetic beads and ligated to adapters for next-generation sequencing. (B) Examples of experimental and chimeric ChIA-PETs. Note the A/B linker composition of chimeric PETs. Tags flanking the central linker sequence are read out and mapped to the genome. (C) Binding sites and interactions are

identified by clusters of overlapping PETs; singleton PETs indicate non-specific ligations that do not represent true binding sites or interactions.

Author Manuscript

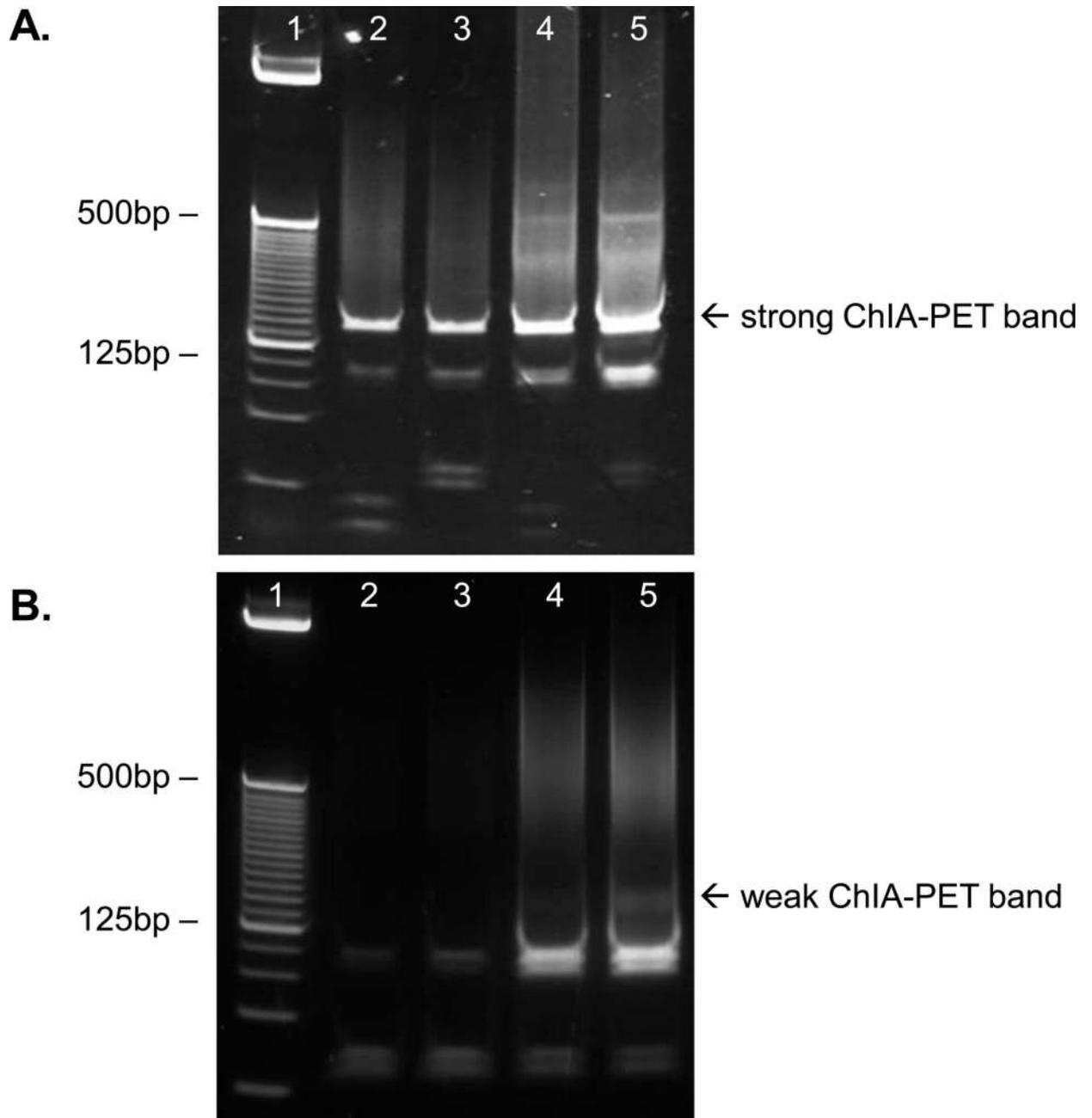
Author Manuscript

Author Manuscript

Author Manuscript

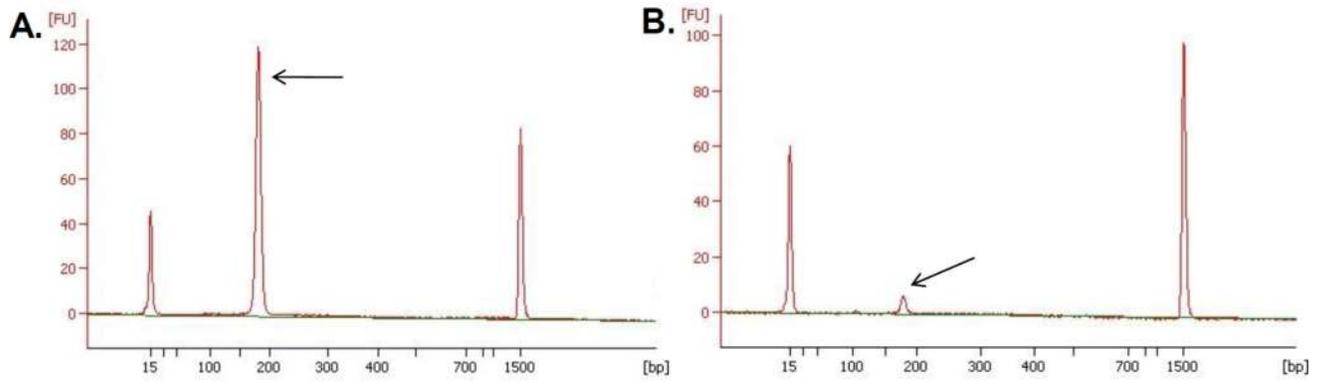


**Figure 2. Custom oligonucleotide sequences for ChIA-PET**



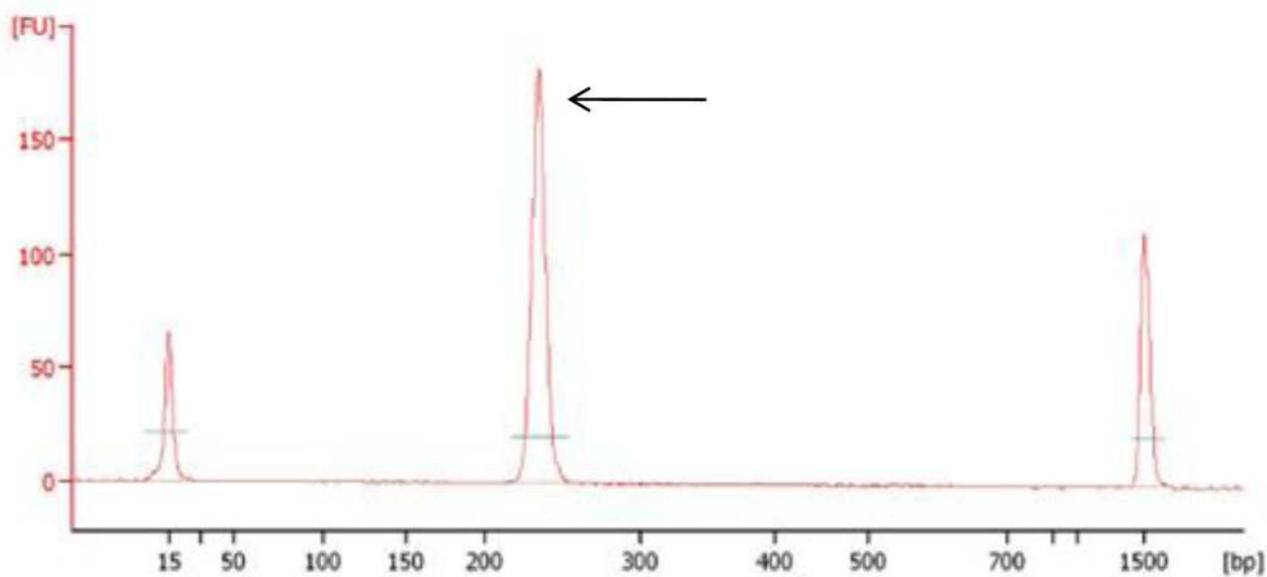
**Figure 3. Gel analysis of ChIA-PETs after PCR amplification (STEP 32)**

A 25 bp DNA ladder is shown in lane 1 for size reference. Lanes 2 and 3 are the PCR products generated after 20 cycles of PCR amplification from 1 µl and 2 µl of bead-immobilized template respectively. 25 cycles were used to generate the PCR products in lanes 4 and 5 program, from 1 µl and 2 µl of beads respectively. **(A)** This is a successful library, as indicated by the bright, well-defined bands at the expected size of 166 bp. **(B)** PCR amplification has failed to yield sufficient ChIA-PET DNA, as seen from the very weak band present in lane 5. This could indicate that PCR conditions need to be optimized, or that library construction has failed.

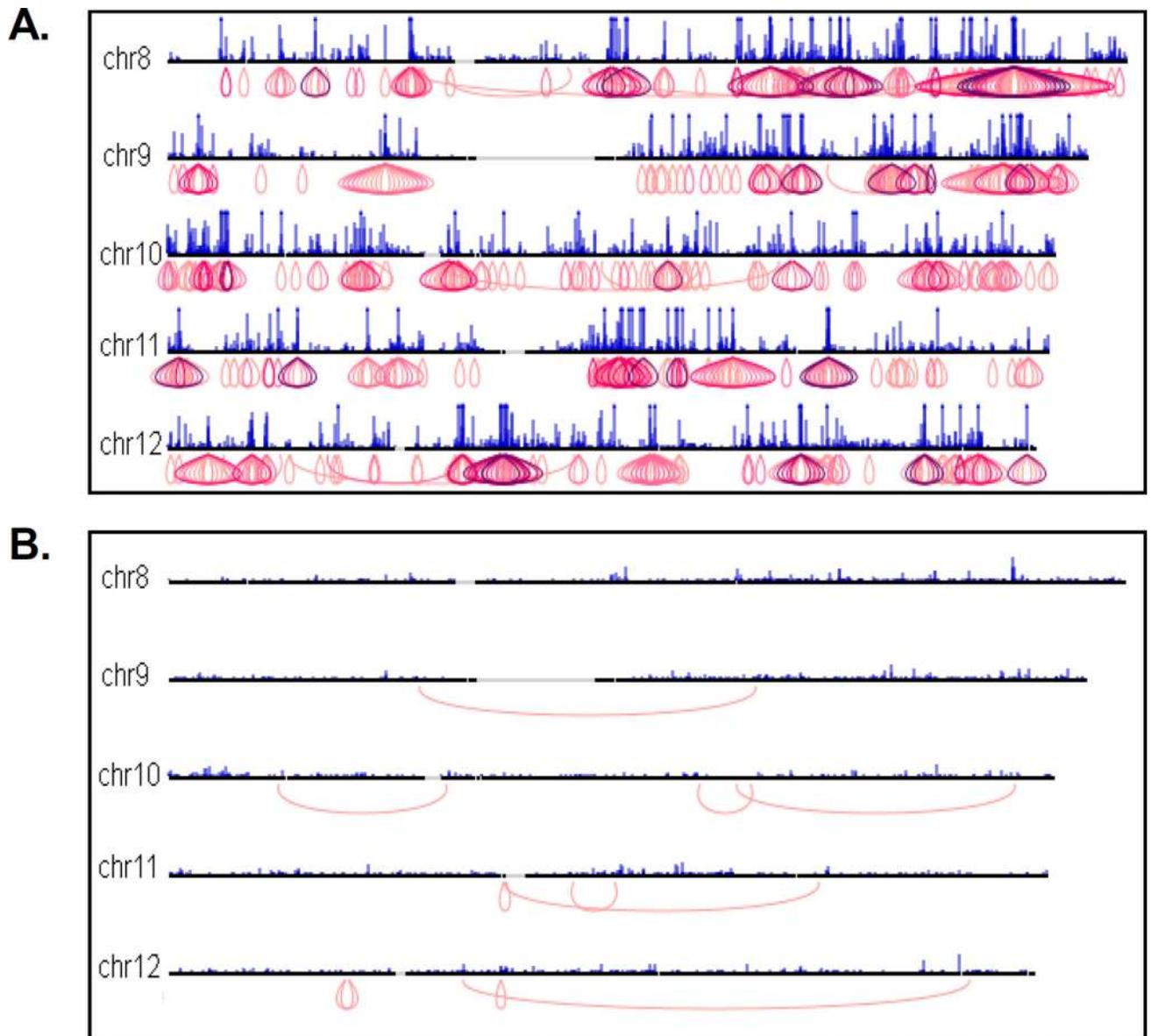


**Figure 4. Agilent 2100 Bioanalyzer analysis of purified 454 adapter-ligated ChIA-PETs (STEP 40)**

Screen capture of Agilent 2100 Bioanalyzer electropherograms profiling a successful library (A), with a single intense peak at the expected size of 166 bp, and an unsuccessful library (B), with a faint signal indicating insufficient ChIA-PET DNA. Peaks of interest are indicated by arrows. Note that the Agilent Bioanalyzer assay usually reports a slightly higher-than-expected size; in this case, the desired peak is displayed at 180 bp instead of 166 bp. This is within the 10% error range of the Agilent assay.

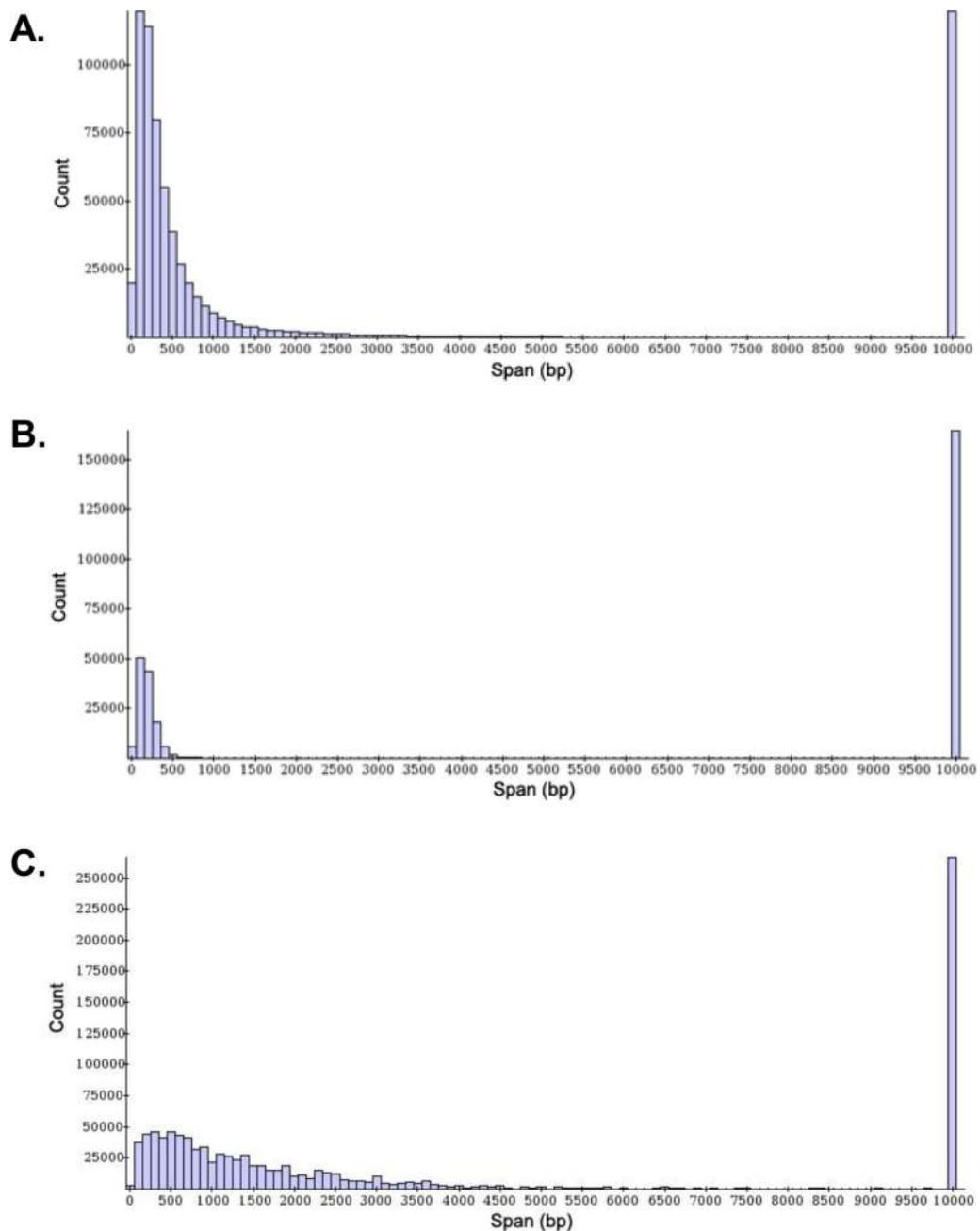


**Figure 5. Agilent 2100 Bioanalyzer analysis of purified ChIA-PETs after conversion for Illumina sequencing (*STEP 44*)**  
A successful library is characterized by a single intense peak (indicated by an arrow) with an expected size of 223 bp.



**Figure 6. Screen captures of ChIA-PET whole genome interaction views**

Chromatin interactions are displayed with purple loop structures where the interactions are located. Interactions are colored according to cluster size; darker purple loops represent more interaction PETs in an interaction unit, and hence, higher-confidence interactions. Regions with multiple unique interactions appear “onion-shaped”. An example of a high-quality ChIA-PET library containing many high-confidence interactions is shown in (A). In contrast, an unsuccessful library (B) has few interactions, and these interactions are mostly low-confidence.



### Figure 7. Example of PET genomic span histograms

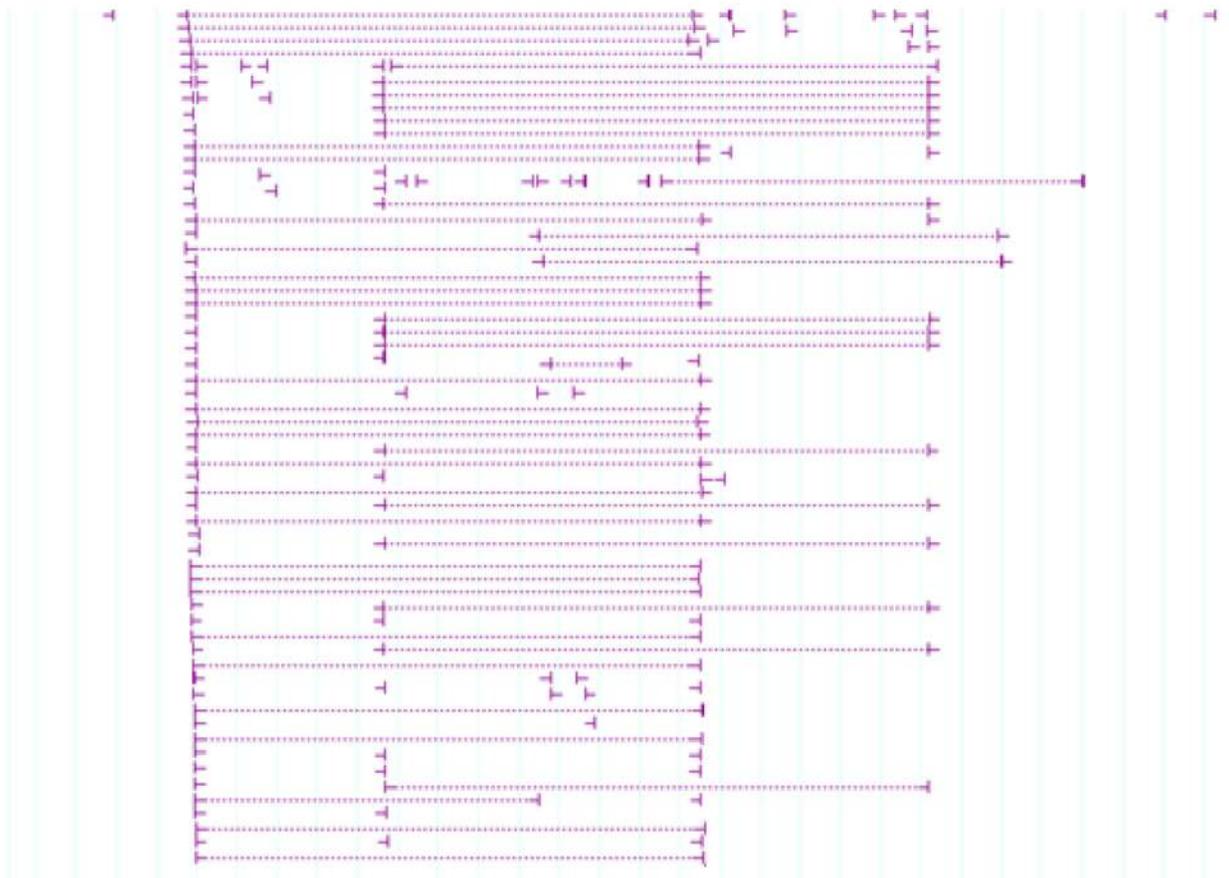
Genomic span is defined as the genomic distance between the two mapped tags within each PET sequence. The histogram in (A) represents the distribution of genomic spans for a successful ChIA-PET library: it shows a high number of self-ligation and inter-ligation PETs with an exponentially decreasing distribution of genomic spans. An example of a low quality library, shown in (B), has an abnormally low PET count and few intrachromosomal inter-ligation PETs with genomic spans exceeding 3kb. (C) is an example of a library with a broad distribution of genomic spans, most likely due to poor sonication quality of ChIP DNA.

**A.**

Density Histogram of Self-ligation PETs



Interaction PETs

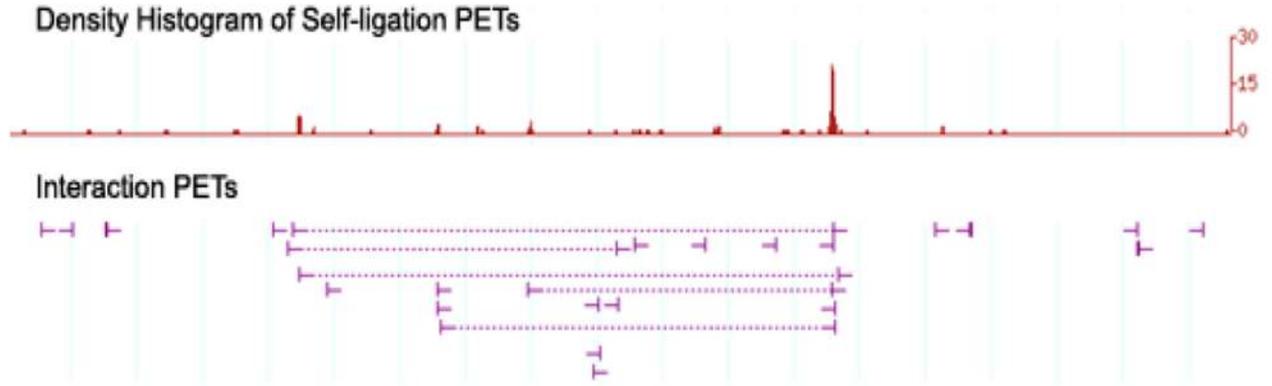
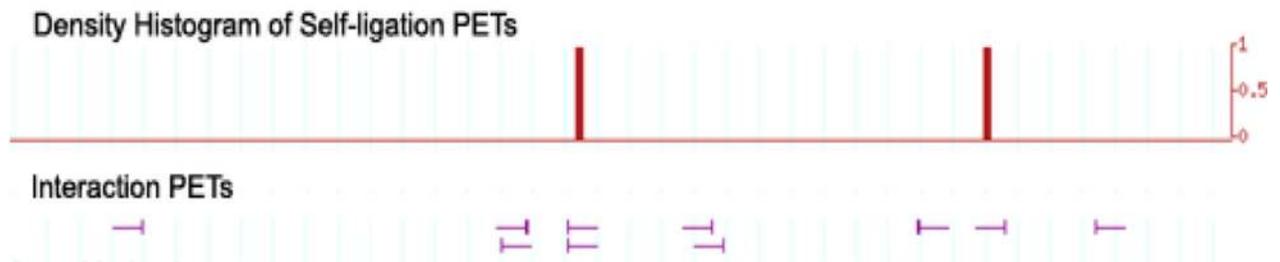


Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**B.****C.****Figure 8. Screenshots for ChIA-PET binding site and interaction data**

(A) High-confidence interaction cluster with good binding site peaks; (B) Weak interaction cluster with good binding site peaks; (C) Region with poor binding site peaks and no interactions. Self-ligation PETs show transcription factor binding sites and interaction clusters comprise multiple overlapping inter-ligation PETs. A successful ChIA-PET library should have an abundance of strong binding sites and a high number of interaction clusters, as shown in examples (A) and (B). An unsuccessful library primarily contains weak binding sites and few, if any, interactions, as in example (C). This problem is possibly due to poor ChIP enrichment and may be resolved by optimizing ChIP procedures.