ChIA-PET Protocol Standards for ENCODE4

Ruan Lab, The Jackson Laboratory for Genomic Medicine (July 28th, 2017)

SUMMARY

Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) is a robust method to capture genome-wide chromatin interactions. ChIA-PET includes a chromatin immunoprecipitation (ChIP) step that enriches for interactions mediated by specific protein factors. This unique feature allows ChIA-PET to provide the functional specificity and higher resolution of chromatin interactions in a highly efficient manner. The original ChIA-PET protocol generates short paired-end tags (2x20 bp) to detect two genomic loci that are far apart on linear chromosomes but are in spatial proximity in the folded genome. We have improved the original approach by developing long-read ChIA-PET (Figure 1), in which the length of the paired-end-tags is increased (up to 2x250 bp). The longer PET reads not only improve the tag mapping efficiency over the short tags in the original ChIA-PET, but also increase the probability of covering phased SNPs, which allows haplotype-specific chromatin interactions identification (Table 1). Comparing to Hi-C, each ChIA-PET experiment generate enriched chromatin interaction data in association with specific protein factors that mediated the interactions, thus provide high resolution mapping of chromatin contacts with considerable less sequencing reads (Table 2). Furthermore, each ChIA-PET experiment also include abundant unenriched chromatin contact data, similar as the data in Hi-C experiments, for plotting the higher-order neighborhood of topological proximity. Hence, each ChIA-PET experiment can produce three related genomic datasets for 3D genome analysis, a) the protein binding sites, b) the enriched chromatin contacts between the binding sites, and c) unenriched chromatin contact "Hi-C like" data (Figure 3). Here, we provide the detailed protocol for long-read ChIA-PET that includes cell fixation and lysis, chromatin fragmentation by sonication, ChIP, proximity ligation with bridge linker, Tn5 tagmentation, PCR amplification, high-throughput sequencing, and an outline of the data processing pipeline. Additional information of the long-read ChIA-PET protocol and data processing pipeline can be found in Nature Protocols (Li et al 2017). To a well-trained molecular biologist, it typically takes six days from cell harvesting to the completion of library construction, up to a further 36 hours for DNA sequencing, and less than 20 hours for processing raw sequencing reads to identify protein binding sites and chromatin interactions.

REAGENTS

- 5x T4 DNA ligase buffer (ThermoFisher Scientific, cat. no. 46300-018)
- 20x SSC (Ambion, cat. no. 1557044)
- Absolute ethanol (500 ml, Sigma, cat.no. E7023)
- Agilent DNA high sensitivity kit (Agilent Technologies, cat.no. 5067-4626)
- Agencourt Ampure XP beads-PCR purification (60 ml, Beckman Coulter, cat. no. A63881)
- Antibody against protein of interest, e.g. Monoclonal Antibody against RNA Polymerase II (8WG16), 0.5 ml (Biolegend, cat. no. 920102, MMS-126R) and Polyclonal Anti-CTCF Antibody (Abcam, cat. no. ab70303)
- EB buffer (250 ml, Qiagen, cat. no. 19086)
- TE buffer (pH8.0, Ambion, cat. no. AM9849)
- dATP solution (100 mM) (NEB, cat. no. N0440S)

- Dimethyl sulfoxide (DMSO, Sigma, cat. no. D2650-100ML)
- DNA clean & concentrator-5-capped columns (200) (Zymo Research, cat. no. D4014)
- 25bp DNA ladder (ThermoFisher Scientific, cat. no. 10597011)
- DNA LoBind tubes (1.5 ml, Eppendorf, cat. no. 022431021)
- dNTPs (10 mM, ThermoFisher Scientific, cat. no. 18427-088)
- DPBS, calcium- and magnesium-free (Gibco, cat. no. 14190-250)
- Dynabeads M-280 Streptavidin (ThermoFisher Scientific, cat. no.11205D)
- Dynabeads Protein G for immunoprecipitation (50 mL) (ThermoFisher Scientific, cat. no. 10009D)
- EDTA (pH 8.0, 0.5 M, 500 ml, Ambion, cat. no. AM9261)
- EGS (Ethylene glycol bis[succinimidylsuccinate], Thermo Fisher Scientific, cat. no. 21565)
- Formaldehyde (37% [vol/vol], EMD MILLIPORE, cat. no. 344198-250ML)
  CAUTION: Formaldehyde is toxic; always operate it in a fume hood.
- Glycine (Sigma, cat. no. G8898-500G)
- GlycoBlue (ThermoFisher Scientific, cat. no. AM9516)
- HEPES buffer (1 M, pH7.3, Fisher Scientific, cat. no. BP299-1)
- I-Block protein-based blocking reagent (ThermoFisher Scientific, cat. no. T2015)
- IGEPAL CA-630 (Sigma, cat. no. I8896)
- Isopropanol (Sigma, cat. no. I-9516-500ml)
- Klenow fragment (3'-->5' exo-) (NEB, cat. no. M0212L)
- MiSeq reagent kit v2 (300-cycles) (Illumina, cat. no. MS-102-2002)
- Maxtract high density (2 ml, QIAGEN, cat. no. 129056)
- NaCl Solution (500 ml, 5.0 M, Ambion, cat. no. AM9759)
- NEB buffer 2 (NEB, cat. no. B7002S)
- NextSeq 500 high output v2 kit (300 cycles) (Illumina, cat. no. FC-404-2004)
- Nextera DNA sample prep kit (Illumina, cat. no. FC-121-1031)
- Novex precast TBE gels (4-20% [wt/vol], ThermoFisher Scientific, cat. no. EC6225BOX)
- Nuclease-free water (50 ml, Ambion, cat. no. AM9937)
- PBS (Phosphate-Buffered Saline) buffer (10X) pH 7.4 (ThermoFisher Scientific, cat. no. AM9624)
- Phenol: Chloroform: IAA (Ambion, cat. no. AM9730) CAUTION Phenol and Chloroform are toxic; operate them in a fume hood.
- Pre-cast agarose gel cassettes for DNA size selection (2% Agarose, dye-free, w/ internal standards, BluePippin.100-600 bp.10/pkg, cat. no. BDF2010, Sage science)
- Proteinase inhibitors (Roche, cat. no. 04693116001)
- Proteinase K solution (20mg/ml, ThermoFisher Scientific, cat. no. AM2548)
- Sodium acetate (Ambion, cat. no. AM9740)
- Sodium dodecyl sulfate (10% SDS solution [wt/vol], Ambion, cat. no. AM9822)
- Sodium deoxycholate (Sigma, cat. no. 30970-100G)
- LightCycler® 480 SYBR Green I Master (Roche, cat. no. 04707516001)
- T4 DNA ligase, HC (30U/µl, ThermoFisher Scientific, cat. no. EL0013)
- T4 DNA polymerase (Promega, cat. no. M4215)

- Tris-HCl, pH 7.0 (1 M, 500ml, Ambion, cat. no. AM9856)
- Triton® x-100, Molecular Biology Grade (Promega, cat. no. H5141)
- TWEEN® 20, viscous liquid (Sigma, cat. no. P9416-100ML)
- Cell sample of choice. For example, we have successfully used GM12878 cell line (Coriell Institute, cat. no. GM12878), HeLa S3 cell line (ATCC, cat. no. CCL-2.2), K562 cell line (ATCC, cat. no. CCL-243) and MCF7 cell line (ATCC, cat. no. HTB-22)
- Bridge linker oligos for proximity ligation: Bridge linker-F: 5'-/5Phos/CGCGATATC/iBIOdT/TATCTGACT -3' and Bridge linker-R: 5'-/5Phos/GTCAGATAAGATATCGCGT -3'. HPLC purified (250nmole) from IDT (Integrated DNA Technologies). See Box 1 for use as 10 μl aliquots (200 ng/ μl) for preparation.

EQUIPMENT

- Agilent 2100 bioanalyzer instrument (Agilent techcologies, cat. no. G2940CA)
- Bio-Rad C1000 Thermal Cycler (Bio-rad, cat. no. 185-1148EDU)
- Biological safety cabinet classII, type A2 (Nuaire, model no. NU-S480-600)
- Bluepippin instrument, Targeted Size Selection for Next-Gen Sequencing (100bp-50kb) (Sage science)
- Eppendorf 5810R centrifuge (Eppendorf, cat. no. 22628180)
- DynaMag™-2 Magnet (ThermoFisher Scientific, cat. no. 12321D)
- DynaMag™-15ml Magnet (ThermoFisher Scientific, cat. no. 12301D)
- Dark Reader Transilluminator (Clare Chemical Research, Inc. Model: DR89X)
- Eppendorf 5424 Ventilated Microcentrifuge (Eppendorf, cat. no. 022620401)
- Eppendorf Thermomixer (Eppendorf, cat. no. 5382000023)
- Freezer (Fisher Scientific, cat. no. 3752-DB)
- Illumina MiSeq and HiSeq 2500 or NextSeq 500 high-throughput sequencing machine (Illumina)
- Incubators (Heratherm, cat. no. 50125590)
- Intelli-mixer RM-2L (Rose scientific, cat. no. MX1000)
- LightCycler®480 Instrument (Roche, cat. no. 05015243001)
- NanoDrop 8000 UV-Vis Spectrophotometers (ThermoFisher Scientific)
- Nikon TS100-F Microscope (Nikon, cat. no. TS100-F)
- Qubit 2.0 Fluorometer Specifications (ThermoFisher Scientific, cat. no. Q32866)
- Refrigerator (Fisher Scientific, cat. no. 13-994-201)
- Sonicator (Sonics & Materials Inc, cat. no. VCX130)
- TOMY MV-100 MicroVac Concentrator (CS Bio Co. cat. no. MV-100)
- Vortex-Genie® Mixers (VWR, cat. no. 58816-121)
- Water bath (Fisher Scientific, cat. no. 15-462-20Q)

COMPUTATIONAL PROGRAMS AND SOURCE CODES USED IN CHIA-PET TOOL V2

- Cutadapt package (version 1.6) (https://cutadapt.readthedocs.io/en/stable/)
- BWA software package (version 0.7.12-r1039 or above) (http://bio-bwa.sourceforge.net)
- SAMtools (version 0.1.19 and above) (http://samtools.sourceforge.net)

- Picard tools (version 1.107 and above) (https://broadinstitute.github.io/picard/)
- BEDTools (version 2.17.0 and above) (http://bedtools.readthedocs.io/en/latest/)
- Pysam python module (version 0.8.1) (http://pysam.readthedocs.io/en/latest/api.html)
- Perl scripts prepare_reads_single.pl (Supplementary Software 1) and prepare_reads_pair.pl (Supplementary Software 2) for bridge-linker trimming.
- Perl script clustering_bridge.pl for PET cluster identification.
- Python script fetch_base_snp_v081.py for assigning PET haplotype.

REAGENT SETUP

10% (wt/vol) sodium deoxycholate (100 ml): Add 10 g of Sodium deoxycholate to 80 ml of water and mix it well until the solution is clear; finalize the solution volume to 100 ml. The buffer can be stored at room temperature (20 to 22 °C) for up to six months.

2.5M glycine solution (100 ml): Add 27.89 g Glycine to 80 ml of water and mix it well until the solution is clear; finalize the solution volume to 100 ml. The buffer can be stored at 4 °C for up to six months.

0.1%(wt/vol) SDS FA cell lysis buffer (500 ml): Add 444 ml of water to a 500 ml beaker first, and then add sequentially 25 ml HEPES (1M, pH 7.5), 15 ml NaCl (5 M), 1 ml EDTA (0.5 M), 5 ml Triton X-100, 5 ml sodium deoxycholate (10%(wt/vol)) and 5 ml SDS (10% (wt/vol)), mix them well. The buffer can be stored at 4 °C for up to six months. CRITICAL 0.1% (wt/vol) SDS FA Cell Lysis buffer can be stored at 4 °C for up two months. However, because protease inhibitors are unstable in solution, dissolve 1 tablet of complete protease inhibitor into 50 ml buffer just before use.

1% (wt/vol) SDS FA nuclear lysis buffer (500 ml): Add 200 ml of water to a 500 ml beaker first, and then add sequentially 25 ml HEPES (1 M, pH 7.5), 15 ml NaCl (5 M), 1 ml EDTA (0.5 M), 5 ml Triton X-100, 5 ml sodium deoxycholate (10% [wt/vol]) and 50 ml SDS (10% [wt/vol]), mix them well and finalize the solution volume to 500 ml. The buffer can be stored at room temperature for up six months.

CRITICAL: SDS might precipitate out of the solution when the temperature is low and warm the buffer to re-dissolve SDS before use if necessary. Dissolve 1 tablet of complete protease inhibitor into 50 ml of the buffer immediately before use.

High salt ChIP buffer (500 ml): Add 324 ml of water to a 500 ml beaker first, and then add sequentially 25 ml HEPES-KOH (1M, pH 7.5), 35 ml NaCl (5 M), 1 ml EDTA (0.5 M), 5 ml Triton X-100, 5 ml sodium deoxycholate (10% [wt/vol]) and 5 ml SDS (10% (wt/vol)), mix them well. The buffer can be stored at 4 °C for up two months.

ChIP wash buffer (250 ml): Add 220.45 ml of water to a 250 ml beaker first, and then add sequentially 2.5 ml Tris•HCl (1 M, pH 8.0), 7.8 ml LiCl (8 M), 0.5 ml EDTA (0.5 M), 1.25 ml IGEPAL CA-630 and 12.5 ml Sodium deoxycholate (10% (wt/vol)), mix them well. The buffer can be stored at 4 °C for up two months.

ChIP elution buffer (2 ml): Add 1.66 ml of water to a 2 ml tube first and then add sequentially 100 µl of Tris•HCl (1 M, pH 7.5), 40 µl of EDTA (0.5 M) and 200 µl of SDS (10% wt/vol), using pipette to mix them well. The buffer should be prepared freshly before use.

ChIA-PET wash buffer (1000ml): Add 888 ml of water to a 1 L beaker first, and then add sequentially 10 ml Tris-Cl (1M, pH 7.5), 2 ml EDTA (0.5 M), 100 ml NaCl (5 M) and mix them well. The buffer can be stored at 4 °C for up two months.

TNE buffer (50ml): Add 48.99 ml of water to a 50 ml Falcon tube first, and then add sequentially 1 ml Tris-HCl (1 M, pH 8.0), 0.5 ml NaCl (5 M), 10 µl of EDTA (0.5 M) and mix them well. The buffer should be prepared freshly before use.

1× PBST buffer (100 ml): Add 89.9 ml of water to a 100 ml beaker first, and then add sequentially 10 ml PBS (10**×**), 100 µl Tween 20 and mix them well. The buffer can be stored at 4°C for up two months.

2× binding & washing buffer (2× B&W, 500ml): Add 294 ml of water to a 500 ml beaker first, and then add sequentially 5ml Tris-HCl (1 M, pH 8.0), 1 ml EDTA (0.5 M), 200 ml NaCl (5 M) and mix them well. The buffer can be stored at 4 °C for up six months.

1× binding & washing buffer (1× B&W, 500ml): Add 394 ml of water to a 500 ml beaker first, and then add sequentially 5ml Tris-HCl (1 M, pH 8.0), 1 ml EDTA (0.5 M), 100 ml NaCl (5 M) and mix them well. The buffer can be stored at 4 °C for up six months.

2× SSC/0.5%(wt/vol) SDS buffer (100 ml): Add 85 ml of water to a 100 ml beaker first, and then add sequentially 10 ml 20 **×** SSC, 5ml SDS (10% (wt/vol)) and mix them well. The buffer can be stored at room temperature for up three months. CRITICAL: lower room temperature (below 18 °C) may cause SDS to precipitate – if this occurs, place the buffer in a 25 °C water bath to solubilize SDS before use.

iBlock buffer (100 ml): Take 2 g I-Block Protein-Based Blocking Reagent into 95 ml water at 65 °C water bath until the powder was dissolved, then add 5 ml 10%(wt/vol) SDS to the solution. The buffer can be stored at room temperature for up six months.

PROCEDURE

Cell harvesting and dual cross-linking (TIMING 4-5 hours)

1 Freshly prepare EGS fixative solution in a fume hood by dissolving 0.02 g EGS (Ethylene glycol bis[succinimidylsuccinate]) in 200 µl DMSO at 37 °C for 5 minutes, then mix EGS/DMSO with 29.8 ml 1× PBS to a final concentration 1.5mM, and then place EGS mixture at 37 °C water bath until use.

2 Collect cells, cultured to ~80% confluency under recommended conditions, by centrifugation at 200 xg for 5 minutes at room temperature (18-22°C). Discard the medium and wash the cell pellet (~10^8 cells) with 10 ml 1× PBS buffer once. Centrifuge at 200 xg for 5 minutes at room temperature and discard 1× PBS buffer.

3 Re-suspend cell pellet in 30 ml freshly prepared EGS fixative solution and mix the cell with rotation for 45 minutes at room temperature.

4 Add 833 µl 37% (vol/vol) formaldehyde to the fixative solution to a final concentration of 1% (vol/vol) and mix with rotation for 20 minutes at room temperature.

5 Quench cross-linking reaction by adding 2.68 ml 2.5M glycine to the fixative solution to a final concentration of 0.2 M with rotation for 5min at room temperature. Spin at 400 xg for 5 minutes at room temperature and discard supernatant into an appropriate waste container.

6 Wash the cell pellet with 20 ml chilled 1× PBS buffer. Spin at 400 xg for 5 minutes at 4°C and discard 1× PBS buffer.
PAUSE POINT: The cell pellets can be frozen at -80 °C for up six months.

Chromatin preparation and immunoprecipitation (ChIP):

Preparing antibody-coated magnetic beads (TIMING 7 hours)

7 Transfer 600 µl resuspended protein G magnetic beads to 1.5ml Eppendorf tube and place the tube on magnetic rack for 30 seconds, discard the supernatant and wash the bead with 1× PBST buffer twice. Resuspend the beads in 900 µl of 1× PBST buffer.

8 Take 60 µg antibody of choice and mix with washed beads with rotation at 4 °C for 6 hours.
NOTE: Step 9-16 can be performed during this incubation.

ChIP: *Cell lysis and nuclear lysis* (TIMING 5-6 hours)

9 Prepare 0.1% (wt/vol) SDS FA cell lysis buffer containing proteinase inhibitor first. Resuspend crosslinked cell pellet (from step 6) with 30 ml of 0.1% (wt/vol) SDS FA cell lysis buffer.

10 Incubate the cells for 15minutes with rotation at 4 °C. Spin at 400 × g for 5 minutes at 4°C and discard cell lysis buffer.

11 Repeat steps 9-10 twice for a total of three times of cell lysis steps.

12 Prepare 1% (wt/vol) SDS FA cell lysis buffer containing proteinase inhibitor first. Resuspend cell pellet with 30 ml of 1% (wt/vol) SDS FA Nuclear Lysis buffer.

13 Incubate the cells for 15minutes with rotation at 4 °C. Spin at 2800 × g for 5 minutes at 4 °C and discard nuclear lysis buffer.

14 Repeat steps 12-13 once for a total two times of nuclear lysis steps.
TROUBLESHOOTING (Table 3)

15 Repeat steps 9-10 twice more for another two rounds of nuclei wash. NOTE: check cell and nuclear morphology after each lysis step to ensure the cell and nucleus are completely lysed.

16 Resuspend chromatin pellet in 4 ml of 0.1% (wt/vol) SDS FA cell lysis buffer containing proteinase inhibitor. Aliquot 1 ml chromatin solution into four 15 ml Falcon round-bottom tubes. Remove bubbles by pipet and place the chromatin solution on ice.
PAUSE POINT: The chromatin solution can be stored at -80 °C for up to one month.

ChIP: *Sonication, pre-clearing and immunoprecipitation* (TIMING 2 days)

17 Shear the chromatin with a sonication system using optimized parameters. NOTE: the sonication parameters and conditions should be optimized prior this step. For the Sonics sonicator, what we used parameter is: amplitude: 35%; time: 6 minutes; 30 seconds ON, 30 seconds OFF.

18 Spin the sonicated chromatin at 2800 **×** g for 10 minutes at 4 °C. Combine 4 tubes of supernatant to a 15 ml Falcon tube. Take 20 µl of chromatin as input control material (non-enriched) for quality check of ChIP (see details in Box 2). It can be stored at -20 °C for later use.

19 Wash 600 μl fresh magnetic beads with 1 ml 1 × PBST buffer twice. Add 4 ml of the collected chromatin solution to the tube to re-suspend the beads and transfer the mixture to the 15 ml Facol tube. Rotate the tube at 4 °C for at least one hour for pre-clearing of chromatin in order to remove non-specific binding of chromatin with magnetic beads.

20 Place the tube containing beads and chromatin on 15 ml magnetic rack for 30 seconds. Transfer chromatin supernatant into a fresh 15 ml tube.

21 Take the tube from step 8 and place the tube on magnetic rack for 30 seconds. Discard the supernatant and wash bead with 1× PBST buffer twice. Discard the buffer.

22 Re-suspend the antibody-coated beads with pre-cleared chromatin solution from step 20 for immunoprecipitation (IP), and rotate the tube at 4 °C for at least 8 hours or overnight.

23 Briefly spin down the tube, place the tube on magnetic rack for 30 seconds. Discard supernatant, and re-suspend beads with 5 ml 0.1% (wt/vol) SDS FA cell lysis buffer (same as used in step 9), and rotate at 4 °C for 5 minutes.

24 Repeat washing step 23 twice for a total of three times wash.

25 Briefly spin down the tube, place the tube on magnetic rack for 30 seconds. Discard supernatant, and re-suspend beads with 5 ml High salt ChIP buffer. Rotate the tube at 4 °C for 5 minutes.

26 Repeat step 25 once.

27 Briefly spin down the tube, place the tube on magnetic rack for 30 seconds. Discard supernatant, and resuspend beads with 5 ml ChIP wash buffer. Rotate the tube at 4 °C for 5 minutes. Discard supernatant.

28 Wash beads with 5 ml 1× TE buffer, Rotate the tube at 4 °C for 5 minutes. Discard supernatant.

29 Repeat step 28 once.

30 Re-suspend the beads into 1 ml of 1× TE buffer, take 20-50 μl of beads to check the quality and quantity of ChIP-DNA following the procedures in Box 2. Store the remaining beads at 4 °C.

PAUSE POINT: the remaining beads can be stored at 4 °C (NOT -20°C) for up to two weeks.
CRITICAL STEP: This is quality control checkpoint QC1, and usually Bioanalyzer result with high sensitivity chip should have a peak around 2~3.5kb (Figure 2b).

Long read ChIA-PET library preparation:

End-repair, A-tailing and proximity ligation (TIMING ~24 hours)

31 Short spin the beads from step 30, wash the beads with TE buffer once, suspend beads in 693 ul of T4 DNA polymerase master mix:

Component	Amount (μl)
10 × Buffer for T4 DNA polymerase	70
10mM dNTPs	7
T4 DNA polymerase	7
ddH_2O	616
Total	700

Add 7 μl T4 DNA polymerase to the tube.

32 Mix and incubate at 37°C for 40 minutes with rotation on a Intelli-Mixer in a 37°C incubator (parameters: F8, 30 rpm; U = 50, u = 60).

33 Short spin the tube, place the tube on magnetic rack for 30 seconds. Discard the T4 DNA polymerase master mix (carefully without disturbing the beads). Wash the beads with 1ml of ice-cold ChIA-PET Wash Buffer for three times.

34 Place the tube on magnetic rack for 30 seconds. Discard the wash buffer and resuspend beads with 700 µl of Klenow (3'-5'exo$^-$) master mix:

Component	Amount (µl)
10 × NEB Buffer 2	70
10 mM dATP	7
ddH_2O	616
Total	700

Add 7 µl Klenow (3'-5'exo-) enzyme to the tube.

35 Incubate at 37 °C for 50 minutes with rotation on a Intelli-Mixer (F8, 30 rpm, U = 50, u = 60) in incubator.

36 Short spin the tube, place the tube on magnetic rack for 30 seconds. Discard the Klenow (3'-5'exo-) master mix. Wash the beads with 1ml of ice-cold ChIA-PET Wash Buffer for three times.

37 Short spin the tube, discard the wash buffer and re-suspend beads with 1.4ml of T4 DNA ligase master mix:

Component	Amount (µl)
5 × T4 DNA ligase buffer	280
Bridge linker (200 ng/µl)	4
ddH_2O	1110
Total	1394

Add 6 µl T4 DNA ligase to the tube.

38 Incubate at 16°C with rotation on a Intelli-Mixer (F8, 30 rpm, U=50, u=60) in incubator overnight.

Reverse cross-linking, DNA purification and tagmentation (TIMING ~20 hours)

39 Short spin the tube from step 38, place the tube on magnetic rack for 30 seconds. Wash beads with ChIA-PET wash buffer for 3 times, discard the supernatant and resuspend magnetic beads into 200 µl of ChIP Elution Buffer. Place the tube on the Thermomixer with rotation (900 rpm) at 65 °C for 15 minutes.

40 Place the tube on the magnetic rack and transfer the 200 µl ChIP Elution Buffer, which contains the protein-chromatin DNA complex, to a new 1.5 ml tube (Do not discard!).

41 Re-suspend magnetic beads in 400 µl of Buffer EB to wash beads. Place the tube on the magnetic rack and take 400 µl supernatant and combine with 200 µl Elution Buffer from step 40.

42 Add 10 µl Proteinase K to the 600 µl of eluted protein-DNA complex solution, mix and incubate at 55 °C overnight for reverse cross-linking.

43 Prepare a MaXtract High Density tube by centrifuging at 16000 × g for 2 minutes at room temperature.

44 Add equal volume of Phenol-Chloroform-Isoamyl alcohol (pH 7.9) to the tube containing reverse cross-linking reaction solution from step 42, and mix vigorously for 5 seconds with hands, transfer the mixture to the centrifuged MaXtract High Density tube, centrifuge at 16000 × g, at room temperature for 5 minutes.

45 Transfer upper aqueous phase above the gel matrix to a new 1.5 ml tube. Precipitate the DNA by adding the following reagents:

Component	Amount (µl)
3M Sodium Acetate	60
GlycoBlue	2
Isopropanol	650

46 Invert the tube to mix solution. Incubate at -80 ºC for at least 30mins.

47 Spin the DNA pellet at 16000 × g at 4 ºC for 20 minutes. Wash pellet with 1 ml of ice-cold 75% (vol/vol) ice-cold ethanol twice. Remove all the ethanol and dry the pellet using Micro vacuum concentrator for 2 minutes at room temperature.

48 Re-suspend the pellet in 20 µl Resuspension Buffer (Illumina) and make sure the DNA is dissolved into buffer completely at 4 ºC. Proceed to DNA quantitation on a Qubit 2.0 fluorometer.

PAUSE POINT: Proximity ligation DNA can be stored at -20 ºC for up to several months.
CRITICAL STEP: This is quality control checkpoint QC2 in Figure 2b and the Bio-analyzer trace should have a peak at ~5 kb.
TROUBLESHOOTING (Table 3)

49 Fragmented the proximity ligated DNA (from step 48) and add the adaptor by using Nextera Tn5 transposome as tabulated below. Generally, two to six tagmentation reactions (100 - 300 ng proximity ligated DNA) are recommended for each Longread ChIA-PET library preparation.

Component	Amount (µl)
Proximity ligation DNA	x
2xTagmentation buffer	25
Tagmentation Enzyme	8
ddH_2O	17-x
Total	50

CRITICAL STEP: The ratio between Tagmentation Enzyme and proximity ligation DNA is critical for optimal reactions and this ratio need to be optimized every time.

50 Gently pipette up and down 6 times to mix DNA and master solution. Incubate at 55 ºC for 5 minutes on a PCR machine.

51 Purify the tagmentated DNA using Zymo Genomic DNA Clean & Concentrator™ kit, add 350 µl of Zymo DNA Binding Buffer to each tagmentation reaction and mix thoroughly.

52 Transfer mixture to Zymo-Spin™ IC-XL column placed in a collection tube. Centrifuge at 16000 × g for 30 s at room temperature. Discard the flow-through.

53 Add 200 µl Zymo DNA Wash Buffer to the column. Centrifuge for 16000 × g for 30 s at room temperature. Discard the flow-through. Repeat this step once.

54 Centrifuge the empty column at 16000 × g for 1 minute at room temperature with lid open to ensure any residual ethanol is removed.

55 Transfer the column to a new 1.5 ml LoBind tube and add 15 µl Resuspension Buffer to membrane of the column. Incubate at room temperature for 1 minute.

56 Spin the column in a 1.5 ml LoBind tube at 16000 xg for 1 minute at room temperature to elute the DNA. Take 1 µl of sample to check tagmentation size distribution by High sensitive DNA chip with Bioanalyzer.
TROUBLESHOOTING (Table 3)
CRITICAL STEP: This is quality control QC3 in Figure 2b and the majority of DNA fragments should be distributed around 200 bp - 1 kb range.

Immobilization on beads and PCR amplification (TIMING 6-7 hours)

57 Mix the M280 Streptavidin dynabeads suspension well, and transfer 25 µl dynabeads suspension for each sample to a 1.5 ml LoBind tube.

58 Place the tube on the magnetic rack, discard the buffer and wash the dynabeads with 150 µl of 2×Binding & Washing buffer (2×B&W) twice.

59 Resuspend the beads in 100 µl iBlock buffer, mix and incubate at room temperature for 45 minutes with rotation on the Intelli-Mixer (UU, 50 rpm, U = 50, u = 60). Wash beads with 200 µl 1×B&W buffer for 3 times.

60 Discard the 1× B&W buffe, add 500 ng (in 50 µl water) sheared genomic DNA (fragment size range 300 to 500 bp) into 50 µl of 2 × B&W buffer in a new tube. The sheared genomic DNA is for non-specific blocking the surface of the M280 Streptavidin Dynabeads and can be prepared by sonication of mammalian genomic DNA.

61 Place the tube containing iBlock buffer and M280 beads on the magnetic rack and Discard the iBlock buffer and wash the M280 beads with 200 µl 1× B&W buffer twice.

62 Resuspend the M280 bead with 100 µl 1× B&W buffer containing blocking DNA (from step 60) and incubate at room temperature for 30 minutes with rotation on a Intelli-Mixer (UU, 50 rpm, U = 50, u = 60).

63 Discard the 1×B&W buffer containing blocking DNA and wash the dynabeads with 200 µl of 1× B&W buffer twice. Place the tube on the magnetic rack and discard the supernatant.

64 Mix 50 µl fragmented and tagged DNA library (from step 56 add resuspension buffer to 50 µl if original volume is not 50 µl) with equal volume of 2 × B&W buffer, then transfer the mixture to the tube containing blocked M280 beads (from step 63). Mix and incubate at room temperature for 45 minutes with rotation on a Intelli-Mixer (UU, 50 rpm, U = 50, u = 60).

65 Place the tube on the magnetic rack and discard supernatant, wash the M280 beads with 500 µl of 2×SSC/0.5% (wt/vol) SDS five times.

66 Place the tube on the magnetic rack and discard supernatant, wash the M280 beads with 1 × B&W buffer twice. Discard the 1× B&W buffer and resuspend the beads with 30µl ddH2O.

67 Using 10 µl M280 beads as template for each PCR amplification (in total three reactions) of library as follow:

Component	Amount (µl)
Beads (from step 66)	10
NPM mix	15
PPC PCR primer	5
Index 1 primer (i7)	5
Index 1 primer (i5)	5
Total	50

68 Perform PCR amplification with the following cycling program:

Temperature	Time	Cycles
72°C	3:00 min	
98°C	10 sec	
98°C	10 sec	11-13 cycles
63°C	30 sec	
72°C	40 sec	
72°C	5:00 min	
4°C	hold	

CRITICAL STEP: The PCR amplification cycle is important for final library quality, the more cycles used here, the lower complexity of the library. The optimal PCR amplification cycles need to be empirically determined. Generally, do not perform more than 13 cycles.

69 Clean-up of PCR products by Ampure XP beads and check the results following steps in Box 3.
TROUBLESHOOTING (Table 3)
CRITICAL STEP: This is quality control checkpoint QC4 in Figure 2b and the Bio-analyzer trace should show that the majority of DNA fragments fall in 200 bp - 1 kb range.

70 Combine 30 µl of cleaned PCR products with 10 µl of V1 maker (Sage science), mix and load them on 2% (wt/vol) Agarose gel cartridge (Sage science).

71 Select DNA fragment in size range from 300 to 600 bp using Blue pippin following manufacture's manual. Check library on Bio-analyzer with high sensitive chip and quantify on Qubit.
CRITICAL STEP: This is quality control checkpoint QC5 in Figure 2b and profile of the size distribution should show majority of final DNA library fall in 300 - 600 bp range from Bio-analyzer. In addition, 10 - 30 ng of library DNA should be obtained for sequencing from Qubit's results.

ChIA-PET library sequencing

72 A ChIA-PET library will be sequenced using MiSeq with 300 cycles sequencing kit first for 5-10 PET reads to assess the library quality. Each MiSeq run can generate approximately 20 million PET reads. We routinely combine 2-4 ChIA-PET libraries (with different library barcodes) for one Miseq run. Output statistics and browser-based visual examination of

ChIA-PET data are used to assess data quality (Table 4, Figure 3d). Usually, 5-10 million PET reads per library can provide sufficient QC assessment.

73 Qualified libraries will then proceed for data production sequencing by Hiseq2500 or Hiseq4000. Usually, 200 million PET reads were generated from a high-quality ChIA-PET library before hitting saturation.

Box1 Preparation of bridge linker (TIMING 5 hours)

1. Add 1× Tris-NaCl-EDTA (TNE) buffer to dissolve bridge linker top and bottom oligo to a concentration of 100 μM.
2. Vortex the oligos for 10 seconds and then leave the solution for 30 min at room temperature to ensure complete resuspension.
3. Prepare 5 different ratios of top and bottom bridge linker oligo (1:1, 1.5:1, 2:1, 1:1.5, 1:2), mix top oligonucleotide (100 μM) and bottom oligonucleotide (100 μM) as follow:

Ratios (vol/vol)	Top oligonucleotide (μl)	Bottom oligonucleotide (μl)
1:1	5	5
1:1.5	5	7.5
1.5:1	7.5	5
1:2	5	10
2:1	10	5

4. Run on PCR machine using the following program:

Cycle number	Temperature and duration
1	95 ºC, 2 min
2-71	Decrease by 1 ºC per cycle and hold for 1 min
72	25 ºC, 5 min
73	4 ºC, 5 min

5. Dilute the annealed bridge linkers to 200 ng/μl.
6. Run 200 ng of each single stranded oligo as control alongside 200 ng (10 μl) of annealed adapters from step 5 with 4-20% (wt/vol) TBE gel.
7. Immerse the gel to 1× TBE buffer with 0.01% (vol/vol) SYBR Green I for 10 minutes at room temperature, then check the gel on Dark Reader Transilluminator.
8. Choose the optional ratio between top oligo and bottom oligo (i.e. only double stranded bridge-linker is observed with no detectable unannealed top or bottom oligo in the lane) and mix the rest of the top and bottom oligo stocks with the optional ratio (The maximum volume of mixed oligos is 100 μl), and run PCR program following step 4 above. Quantify the annealing linker with NanoDrop 8000 UV-Vis Spectrophotometers. The annealed bridge linker should be diluted (final concentration: 200 ng/μl) and aliquoted for storage at -20 ºC for future use.

Box2 Quality control of ChIP-DNA (TIMING 4 hours)

1. Spin down 50 µl of the beads from step 30 of the main Procedure and discard TE buffer.
2. Add 200 µl of ChIP Elution buffer.
3. Incubate at 65 °C with mixer for 10min with agitation at 900 rpm.
4. Transfer supernatant to new tube and add 400 µl of Buffer EB.
5. Add 10 µl proteinase K to the tube containing eluted ChIP-DNA and incubate at 55 °C for 2h for reverse-crosslinking of protein-DNA complex.
6. Prepare a MaxTract High Density tube for each sample by centrifuging at 16000 × g for 1 minute at room temperature.
7. Add sample into the tube, and then add equivalent volume of phenol:chloroform:IAA (pH 7.9) to the sample.
8. Gently mix the tube with hands and centrifuge the tube at 16000 × g for 5 minutes at room temperature.
9. Transfer the upper aqueous phase into a new tube, precipitate DNA by adding the following:

Component	Amount (µl)
3M Sodium Acetate	60
GlycoBlue	2
Isopropanol	650

10. Incubate the tube at -80 °C for at least 30 minutes.
11. Thaw the sample, mix well and centrifuge at 16000 × g for 20 minutes at 4 °C.
12. Wash the pellet twice with 1 ml 75% (vol/vol) ethanol by gently pipetting in and then pouring off. Make sure the pellet remains in the tube.
13. Dry the pellet using MicroVac Concentrator for 2 minutes at room temperature.
14. Dissolve DNA in 20 µl TE buffer.
15. Quantitate DNA using Qubit 2.0 and check the profile of ChIP-DNA on Bioanalyzer with High-sensitivity chip using 1 µl of library, according to the manufacturer's instructions.

Box 3 Clean-up of PCR products (TIMING 40 minutes)

1. Vortex AMPure XP beads to resuspend them completely and take 270µl AMPure XP beads to a new 1.5 ml LoBind tube (place AMPure XP beads at room temperature for 30 minutes before use).
2. Combine the products of the three PCR reactions (from step 68 of the main Procedure) and transfer the entire volume (~150 µl) to the tube with AMPure XP beads. Mix well by pipetting up and down at least 10 times.
3. Incubate for 5 minutes at room temperature.
4. Quickly spin the tube and place it on magnetic rack to separate beads from supernatant. After the solution is clear (about 5 minutes), carefully remove and discard the supernatant.
5. Keep the tube on the magnetic rack and add 400 µl of 80% (vol/vol) freshly prepared ethanol to the tube. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
6. Repeat Step 5 once, for a total of two washes.

7. Keep the tube on the magnetic rack and leave the lid open to air-dry AMPure XP beads for 3 - 5 minutes.
8. Elute the DNA from the beads by adding 35 μl of Buffer EB.
9. Check the quantity of DNA library and size distribution by Qubit and Agilent assay with 1 μl of sample, respectively.

TIMING

Steps1-6, Cell harvesting and dual-cross-linking: 4 - 5 hours
Steps7-8, ChIP: Preparing antibody-coated magnetic beads: 7 hours
Steps9-16, ChIP: Cell lysis and nuclear lysis: 5 - 6 hours
Steps17-30, ChIP: Sonication, pre-clearing and immunoprecipitation: 2 days
Steps31-38, Long read ChIA-PET library preparation: End-repair, A-tailing and proximity ligation: ~24 hours
Steps39-56, Long read ChIA-PET library preparation: Reverse cross-linking, DNA purification and tagmentation: ~20 hours
Steps57-71, Long read ChIA-PET library preparation: Immobilization on beads and PCR amplification: 6-7 hours
Step 72, Sequencing of long-read ChIA-PET library by MiSeq: ~ 36 hours
Step 73, Sequencing of long-read ChIA-PET library by HiSeq: ~ 6 days
Box 1, bridge linker preparation: 5 hours
Box 2, ChIP-DNA QC: 4 hours
Box 3, Cleanup of PCR DNA: 40 minutes

ANTICIPATED RESULTS

A quality ChIA-PET library dataset usually include ~40 million uniquely mapped non-redundant PETs from ~200-300 million raw reads. The sequencing cost on the current Illumina system is about $1,700. After sequencing and data processing, long-read ChIA-PET data is classified into two major categories based on the mapping span between two tags on either side of the bridge-linker in a single paired-end read: 1) mapping span < 8 kb. Most of the PET data less than 8 kb were derived from self-ligation of the same DNA fragment as previously characterized. The piling up of these reads at specific loci as peaks reflect DNA binding sites by target protein factor. This data is similar to ChIP-Seq datasets; and 2) mapping span > 8 kb. These PET data represent inter-ligation between different DNA fragments in the same chromosome with mapping span larger than 8 kb (intra-chromosomal interactions) and between different chromosomes (inter-chromosomal interactions). The inter-ligation data can be further classified as PET clusters and PET singletons based on our clustering scheme (direct overlaps of PET reads with 500bp extension). Together, a ChIA-PET dataset can identify enriched protein binding sites mostly by self-ligation PETs, enriched chromatin interactions by clustered PETs, and the higher-order chromosomal folding by singleton PETs (Figure 3a). The ChIA-PET data can be visualized using 2D heatmaps similar as Hi-C data (Figure 3b). With the increased read length of PETs (long-read ChIA-PET), the PET clusters are more likely to overlap with phased heterozygous SNPs, thus the binding peaks of target proteins and consequent interaction loops anchored to the binding peaks can be assigned to a specific allele for haplotype-specific chromatin interaction analysis (Figure 3c) and achieve single nucleotide

resolution for chromatin interaction mapping.

REFERENCES

Li, X. *et al.* Long-read ChIA-PET for base-pair resolution mapping of haplotype-specific chromatin interactions. *Nature Protocols* 12(5): 899-915 (2017).

Tang, Z. *et al.* CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* 163: 1611-27 (2015).

TABLES

Table 1. Comparison of long-read ChIA-PET with original ChIA-PET

GM12878 cells	Long-read ChIA-PET (ChIA-PET v2)		Original ChIA-PET (ChIA-PET v1)	
Length of sequencing tags	Up to 2x150 bp		2x20 bp	
Enzymatic reactions (steps)	4		7	
Ligation reactions (steps)	1		3	
Time (days)	7		10	
Uniquely mapped PETs	71.2% ± 4.0%		59.1% ± 1.3%	
Uniquely mapped inter-ligation PETs	*17,492,999 (RNAPII)	*2,649,901 (CTCF)	*17,492,999 (RNAPII)	*2,649,901 (CTCF)
Genome coverage (bp)	1,624,706,794	414,228,424	320,817,222	74,313,729
Genome coverage increment	5.1X	5.6X	1X	1X
SNP coverage	848,765	223,096	175,255	42,791
SNP coverage increment	4.8X	5.2X	1X	1X

*To manifest the improvement by ChIA-PET v2 over v1 protocol in terms of genome coverage and heterozygous SNP coverage, same numbers of PET reads from different RNAPII and CTCF ChIA-PET libraries were collected for the calculation of fold changes.

Table 2. Comparison of long-read ChIA-PET to Hi-C

GM12878 cells	CTCF ChIA-PET (Tang et al., 2015)	in situ Hi-C (Rao et al., 2014)
No. of cells per library	100 million	5 million
No. of libraries used for final dataset	2	29
No. of cells used for final dataset	200 million	145 million
No. of raw sequencing reads	0.68 billion	6.5 billion
No. of non-redundant mapping reads	51 million	4.9 billion
Total high confidence loops	42,297	9,448
Total chromatin loop anchors	21,777	12,903
Loop anchor size (bp)	50-100 (CTCF binding sites)	1,000 (Clustering bin size)

Table 3. Troubleshooting

Step	Problem	Possible reason	Solution
14	Too much cytoplasm remains around the nucleus	The cell lysis conditions might be too gentle	Increase the temperature of cell lysis from 4 ºC to room temperature or even up to 37 ºC; or increase the incubation time to 30 mins.
48	Low yield of proximity ligated DNA	Low yield of ChIP-DNA	High quality antibodies are recommended for immunoprecipitation and at least 200 - 300 ng of ChIP DNA is required for library preparation.
56	Too many DNA fragments > 1 kb	Too much proximity ligation DNA or not enough Tn5 transposome are used for tagmentation reaction.	Increase the amount of Tn5 enzyme for digestion or decrease the amount of proximity ligated DNA in the reaction.
69	Low complexity of the library	Excessive PCR amplification or not enough DNA templates	Reduce the PCR amplification cycles or increase the amount of ChIA-PET DNA for library preparation

Table 4. Output statistics of ChIA-PET data (GM12878 cells)

CTCF ChIA-PET (LHG0018)	Miseq QC	HiSeq Production	Range of expected quality data (±20%)
Total PET	21,508,933	338,804,565	200,000,000
PET w/ linker	13,248,828	207,310,687	
Mappable PET (>Q30)	9,982,456	158,200,925	
Unequally non-Redundant PET	8,703,461	44,190,806	40,000,000
Redundancy	13%	72%	
Peak	37,294	81,544	50,000
Self-ligation PET (< 8Kb)	1,727,616	8,945,849	10,000,000
Inter-ligation PET	6,975,845	35,244,957	30,000,000
Intra-chr PET	2,193,679	11,040,703	10,000,000
Inter-chr PET	4,782,166	24,204,254	20,000,000
Intra-chr PET cluster			
PET count 1	1,950,088	7,806,282	
PET count 2	48,385	631,006	
PET count 3	6,512	121,915	
PET count 4	2,571	46,937	
PET count 5	(≥5) 8,561	22,200	
PET count 6		11,677	
PET count 7		6,553	
PET count 8		4,085	
PET count 9		2,928	
PET count ≥10		26,312	>20,000
Inter-chr PET cluster			
PET count 1	4,657,277	20,294,522	
PET count 2	60,424	1,311,915	
PET count 3	1,247	214,025	
PET count 4	56	77,929	
PET count 5	(≥5) 6	33,682	
PET count 6		14,425	
PET count 7		6,093	
PET count 8		2,430	
PET count 9		965	
PET count ≥10		549	<1,000

MiSeq data
Intra / inter PET ratio: 2,193,679 / 4,782,166 = **0.46**
Intra / inter PET singleton ratio: 1,950,088 / 4,657,277 = **0.42**
Intra / inter PET cluster (≥5) ratio: 8,561 / 6 = **1,425**

HiSeq data
Intra / inter PET ratio: 11,040,703 / 24,204,254 = **0.46**
Intra / inter PET singleton ratio: 7,806,282 / 20294,522 = **0.38**
Intra / inter PET cluster (≥10) ratio: 26,312 / 549 = **48**

FIGURES

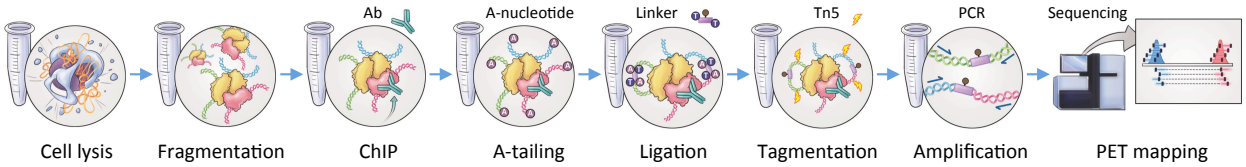

Figure 1. Schematic of long-read ChIA-PET protocol.

Cross-linked cells are lysed and chromatin is fragmented by sonication. The target proteins with its binding DNA fragments are immunoprecipitated (ChIP) with antibody in similar way as regular ChIP-Seq. After DNA end-repair and A-Tailing, a bridge-linker is used to perform proximity ligation. The proximity ligation products are released from protein-DNA by reverse-cross-linking, followed by Tn5 transposome digestion, in which sequencing adapters are added to the DNA ends simultaneously. The resulting DNA fragments are immobilized on Streptavidin beads for on-bead PCR library amplification. PCR products in length of 300 to 600 bp are selected and subjected to sequencing analysis.

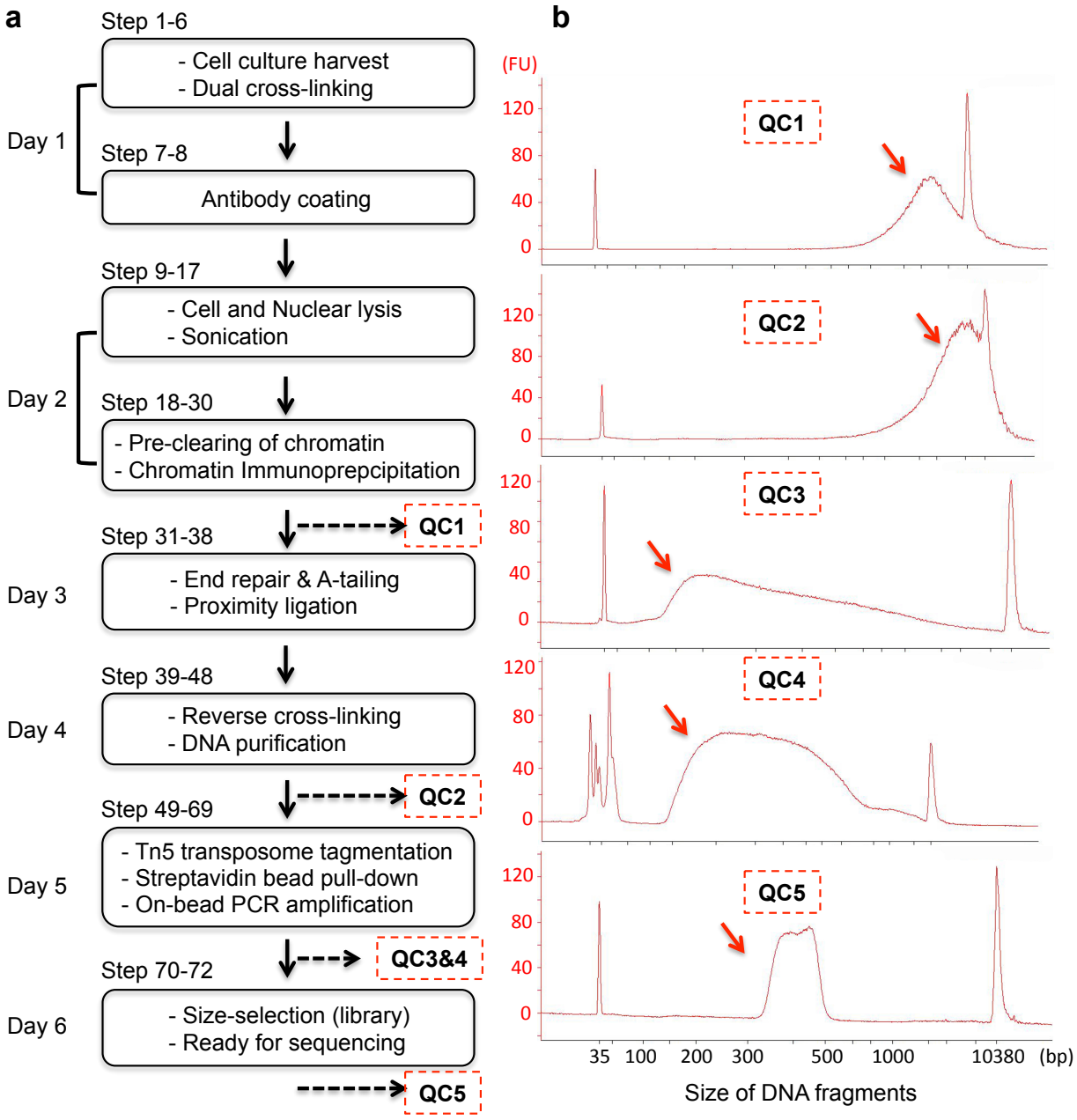

Figure 2. Flowchart of key steps and quality control in long-read ChIA-PET protocol.

a. Key steps for long-read ChIA-PET library preparation are shown in the boxes. The timing for relevant steps is included on the left and the corresponding steps for each major component of the workflow are included on the top-left of the boxes. **b.** Five quality control checkpoints are shown with DNA distribution profiles, QC1, CTCF ChIP DNA fragments distribution with peak at ~3.5 kb (arrow); QC2, proximity ligation DNA distribution with peak at ~4.3 kb (arrow); QC3, the size range of Tn5 Transposome tagmentation DNA product mainly falls between 140 bp to 1 kb with peak at ~200 bp (arrow); QC4, the DNA fragments size distribution of PCR product is 200 to 900 bp; QC5, final DNA sequencing library size is from 320 to 500 bp after size-selection.

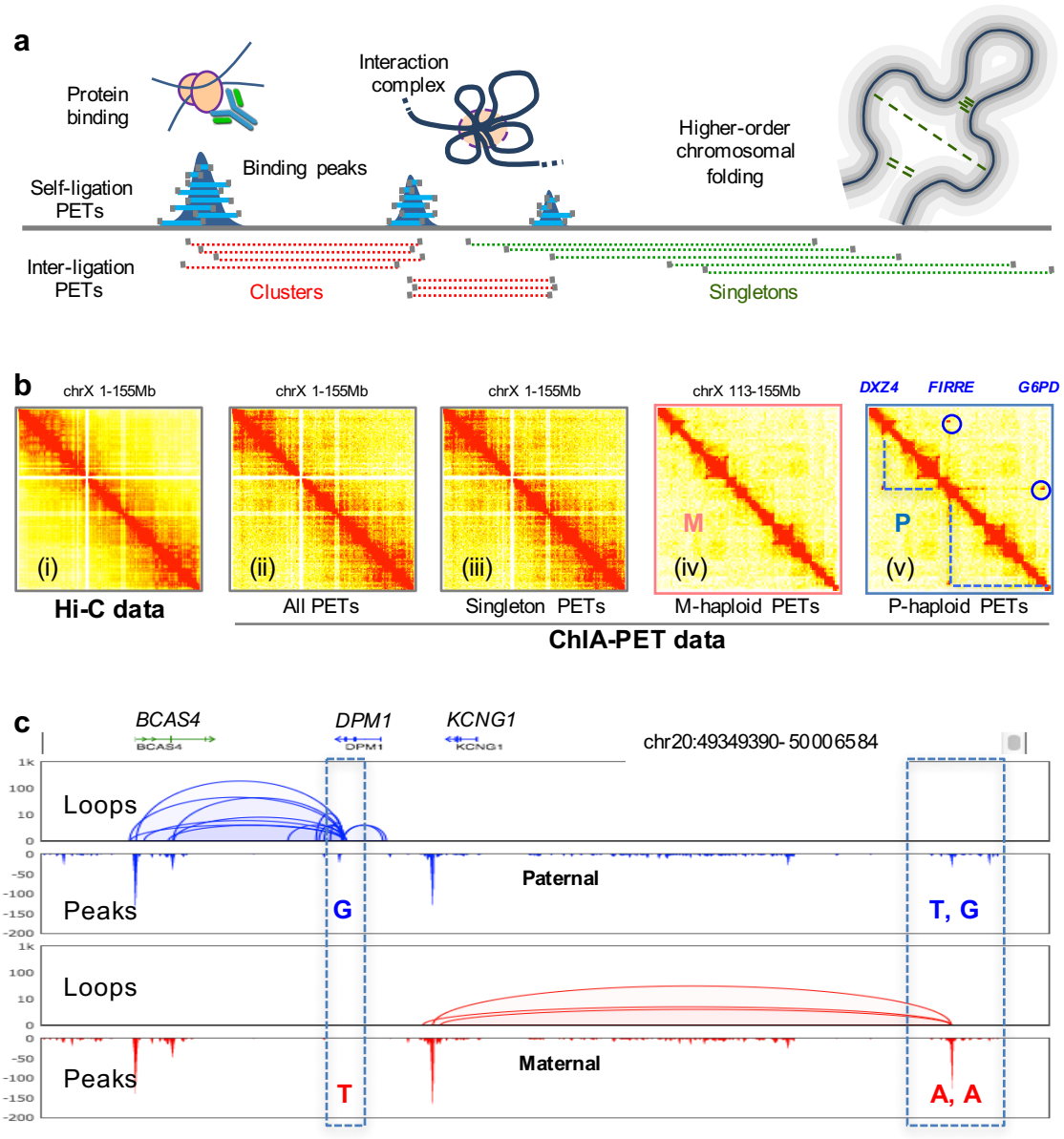

Figure 3. ChIA-PET data features. a. Graphical illustration of binding peaks, enriched chromatin interactions (clusters), and non-enriched singleton PETs for inferring topological neighborhood proximity produced from one long-read ChIA-PET experiment (adapted from [11]). **b.** Comparison of CTCF ChIA-PET (right four panels) and Hi-C (leftmost panel) data in 2D contact heat maps of chromosome X (chrX). Heat maps of the entire X chromosome (chrX:1-155Mb) are for Hi-C data (i), all ChIA-PET data (ii), and only singleton PET data (iii), respectively. Zoomed heat maps (iv and v) are for maternal (M) and paternal (P) haploid ChIA-PET data, respectively, at the same chrX 3' segment (chrX:133-155Mb), where haplotype-specific long-range chromatin interactions (blue dotted lines) between the three loci (DXZ4, FIRRE and G6PD) were detected only in the paternal haploid of the X chromosome. **c.** An example of haplotype-resolved chromatin interactions. The screenshot shows allele-specific chromatin interactions and binding peaks of CTCF identified by

long-read ChIA-PET in a given region on chr20 (chr20: 49349390-50006584). The top track shows the gene annotation. The paternal (blue) and maternal (red) specific chromatin loops and binding peaks are shown in the lower tracks. The dashed boxes indicate the phased SNPs that are located inside the CTCF binding sites.