Methods 170 (2020) 38–47

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

# BAT Hi-C maps global chromatin interactions in an efficient and economical way

Jie Huang[a,*,1], Yongpeng Jiang[a,1], Haonan Zheng[a], Xiong Ji[a,*]

[a] Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, School of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

ARTICLE INFO

*Keywords:*
Hi-C
3D genomics
Chromatin loops
Chromatin interactions

ABSTRACT

Chromosome Conformation Capture (3C)-based technologies, such as Hi-C, have represented a significant breakthrough in investigating the structure and function of higher-order genome architecture. However, the mapping of global chromatin interactions remains challenging across many biological conditions due to high background noise and financial constraints, especially for small laboratories. Here, we describe the Bridge linker-AluI-Tn5 Hi-C (BAT Hi-C) method, which is a simple and efficient method for delineating chromatin conformational features of mouse embryonic stem (mES) cells and uncover DNA loops. This protocol combines AluI fragmentation and biotinylated linker-mediated proximity ligation to obtain kilobase (kb) resolution with a marked increase in the amount of unique read pairs. The protocol also includes chromatin isolation to reduce background noise and Tn5 tagmentation to cut down on preparation time. Importantly, with only one-third sequencing depth, our method revealed the same spectrum of chromatin contacts as *in situ* Hi-C. BAT Hi-C is an economical (i.e., approximately $40 for library preparation) and straightforward (total hands-on time of 3 days) tool that is ideal for the in-depth analysis of long-range chromatin looping events in a genome-wide fashion.

## 1. Introduction

In eukaryotic cells, chromatin is packed and organized to form dynamic three-dimensional (3D) structures [1,2]. These higher-order structures have profound effects on gene expression and other genomic activities [3–6]. During the past decade, chromosome conformation capture assays have been widely adapted to assess the hierarchy of chromatin topology, such as 3C [7], 4C [8–10] and 5C [11]. In combination with next-generation sequencing, recent advances in protein-centered interaction capture has intimately linked 3D genome organization to its biological function, including ChIA-PET [12–14] and HiChIP/PLAC-seq [13,15]. In particular, as the most powerful of these 3C-based techniques, Hi-C has revolutionized our understanding of spatial genome architecture at unprecedented resolution with full detail, i.e., A/B compartments [16], topologically associating domains (TADs) [17,18] and long-range chromatin loops [19–23]. However, Hi-C has two major limitations. First, due to low signal-to-noise ratios, Hi-C requires, in the case of mammalian genomes, more than three billion reads to achieve sufficient resolution for the detection of chromatin conformational features beyond a few kilobases [22,24]. The economic consideration for obtaining high coverage of mammalian genomes makes Hi-C unfeasible for a small laboratory to conduct mechanistic studies requiring multiple perturbations. Second, Hi-C is sophisticated and produces high-quality libraries due to experimental noise and intrinsic biases [25]. Thus, further modifications to improve the overall efficiency, simplify the implementation, and reduce experimental and sequencing costs are highly desirable.

To accomplish these goals, we introduce four technical improvements to standard *in situ* Hi-C protocols to simplify experimental procedures, reduce library preparation costs, and attain sufficient spatial resolution for depicting global chromatin structural features at the loop level with high signal-to-noise ratios using lower sequencing depths than those of canonical Hi-C protocols.

## 2. Description of the method

### 2.1. Overview

Here, we describe a protocol termed Bridge linker-AluI-Tn5 Hi-C (BAT Hi-C), a modified *in situ* Hi-C method with a four-step optimization: 1) fragment chromatin with a blunt 4 cutter restriction enzyme, AluI; 2) ligate A-tailed DNA fragments with a biotinylated bridge linker

* Corresponding authors.
*E-mail addresses:* jih052@pku.edu.cn (J. Huang), xiongji@pku.edu.cn (X. Ji).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ymeth.2019.08.004
Received 15 March 2019; Accepted 15 August 2019
Available online 20 August 2019

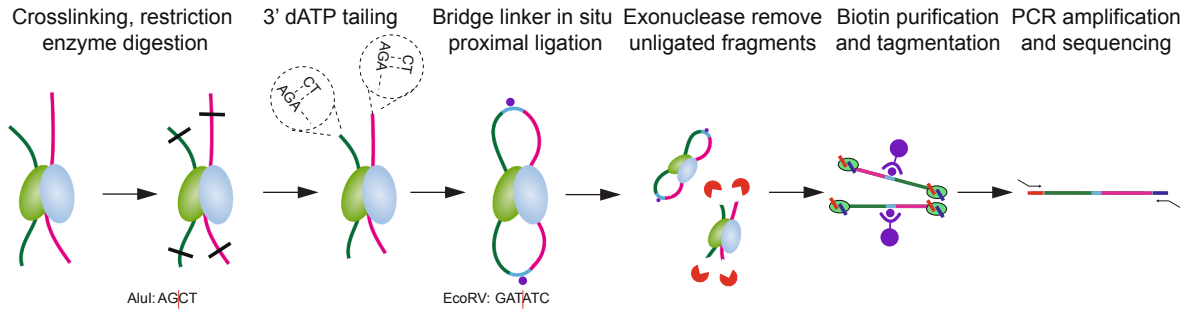



**Fig. 1.** Flowchart of the BAT Hi-C method. Chromatin are cross-linked and digested with AluI restriction enzyme. After chromatin blunt ends 3′A-tailing, the proximal DNA fragments are ligated by a 5′ T overhang biotinylated bridge linker then purified after chromatin isolation. The resulting DNA ligation products are enriched on streptavidin beads, followed by Tn5 transposase-mediated tagmentation, in the meantime, Illumina sequencing adaptors can be added to both ends of DNA fragments. Finally, PCR amplification products with lengths of 300–600 bp are selected and subjected to pair-end high-throughput sequencing.

[26,27]; 3) isolate chromatin to remove untethered random ligation products; and 4) construct the library with Tn5 tagmentation [14]. Briefly, mES cell nuclei were cross-linked *in situ* by formaldehyde. Then, A-tailing was directly performed after AluI digestion without end repair. Inspired by the well-established ChIA-PET method, we employed an exogenous 'bridge' linker instead of two half-linkers to reduce the number of enzymatic reactions, thereby improving the overall efficiency. The free ends of the chromatin were ligated with T4 DNA ligase, and optionally, chromatin isolation was performed to reduce background noise. Subsequently, the ligated fragments were purified using streptavidin-coated beads and tagmentated by Tn5 transposase, which is used to randomly fragment DNA and simultaneously add adaptors to cleavage sites. Finally, ligation products were PCR-amplified to generate sequencing libraries (Fig. 1).

Overall, BAT Hi-C is an efficient and robust method that can be completed in 3 days, with two quality-control measurements. Moreover, the library preparation costs are approximately $40 (Table B.1), i.e., less than canonical Hi-C protocols.

### 2.2. Detailed protocol

The whole procedure can be divided into six major components following the main steps of *in situ* ligation: (1) cell fixation, (2) AluI digestion, (3) *in situ* proximity ligation with bridge linkers, (4) chromatin isolation and exonuclease digestion, (5) Tn5 tagmentation and PCR amplification, and (6) high-throughput sequencing and bioinformatic analysis. Herein, we describe the rationale and design of a BAT Hi-C experiment, provide a step-by-step library generation protocol, and illustrate how the quality of our libraries can be monitored at various steps in the protocol to yield high-quality data.

#### 2.2.1. Bridge linker annealing

For the purpose of increasing the efficiency of the Hi-C ligation step, an optimized design of the bridge linker is required. To test whether the length of the linker may be an influencing factor, two different lengths were examined according to the ChIA-PET protocol [14]: a short-length linker (linker A) was 8 bp, and a long-length linker (linker B) was 20 bp. Both linkers were capable of mediating efficient ligation.

1) For the linker preparation, two reverse complementary single strand linkers were initially dissolved (Table B.2) in 1 × TNE annealing buffer at a concentration of 100 μM, a dTTP overhang was held at the 3′ end, and an annealing mixture with an optimized F/R molecular ratio (normally 1:1) was prepared.
2) An annealing program was run on a PCR machine as follows:
   hold 2 min at 95 °C for total denaturation, ramp from 95 °C to 80 °C (rate: 0.1 °C/second);
   hold 2 min at 80 °C; ramp from 80 °C to 70 °C (rate: 0.1 °C/s);
   hold 2 min at 70 °C; ramp from 70 °C to 55 °C (rate: 0.1 °C/s);
   hold 2 min at 55 °C; ramp from 55 °C to 40 °C (rate: 0.1 °C/s);
   hold 2 min at 40 °C; ramp from 40 °C to 25 °C (rate: 0.1 °C/s);
   hold 2 min at 25 °C; ramp from 25 °C to 4 °C (rate: 0.1 °C/s); and
   hold at 4 °C indefinitely until collection.
3) Add 1 × TNE buffer to dilute the annealed bridge linker at a concentration of 200 ng/μl for subsequent experiments (annealed linker quality can be checked by DNA PAGE electrophoresis).

#### 2.2.2. Cell preparation and crosslinking

To ensure the complexity of the BAT Hi-C library, we recommend starting with 5 million cells. The cells typically need to be well-fixed to carry out a successful Hi-C experiment; insufficient crosslinking would degrade blunt ends, whereas overfixation may suppress enzyme-associated reactions and induce unnecessary background noise. For mouse ES cells that grow in aggregated states, higher formaldehyde concentrations may be required. Here, we used a final concentration of 1% formaldehyde to crosslink DNA-DNA interactions bridged by proteins.

1) For cells grown to proper confluency, harvest the cells in 10 ml volume of fresh culture medium with 1 million cells per milliliter. Add 278 μl of 37% methanol-free formaldehyde solution to obtain a final concentration of 1% (vol/vol), and gently shake at room temperature for 10 min.
2) To quench the fixation, add 500 μl of 2.5 M glycine to obtain a final concentration of 0.125 M. Gently shake for 5 min at room temperature.
3) Centrifuge the tube at 2500 rpm for 5 min at 4 °C (variant between cell types), and gently wash the pellet twice with ice-cold 1 × PBS.

#### 2.2.3. Chromatin digestion with AluI

Because 3C-based methods rely on restriction enzyme (RE) digestion to fragment chromatin, the ultimate resolution of fine-scale chromatin folding paradigms revealed by these methods is limited by the local distribution of RE sites. Previous high-resolution Hi-C libraries have used a four-cutter restriction enzyme, e.g., MboI or DpnII, to fragment DNA [22]. Our protocol substitutes these commonly used enzymes with AluI for two reasons. First, this enzyme specifically recognizes the 'AGCT' sequence and targets 12,630,761 sites in the mouse genome, which is far more than those of MboI (7,033,143) and DpnII (6,684,538). Thus, AluI would capture more chromatin contacts imposed by the occurrence of restriction sites. Second, blunt ends after AluI digestion could be directly A tailed and avoid bridge linker induced self-ligation.

1) Resuspend one formaldehyde-fixed cell aliquot ($5 \times 10^6$ mES cells) in 1 ml lysis buffer and gently rotate for 20 min at 4 °C.
2) Centrifuge at 3000 rpm (variant between cell types) for 2 min at 4 °C and discard supernatant.
3) Wash nuclei twice with 1xNEB Cutsmart Buffer and discard

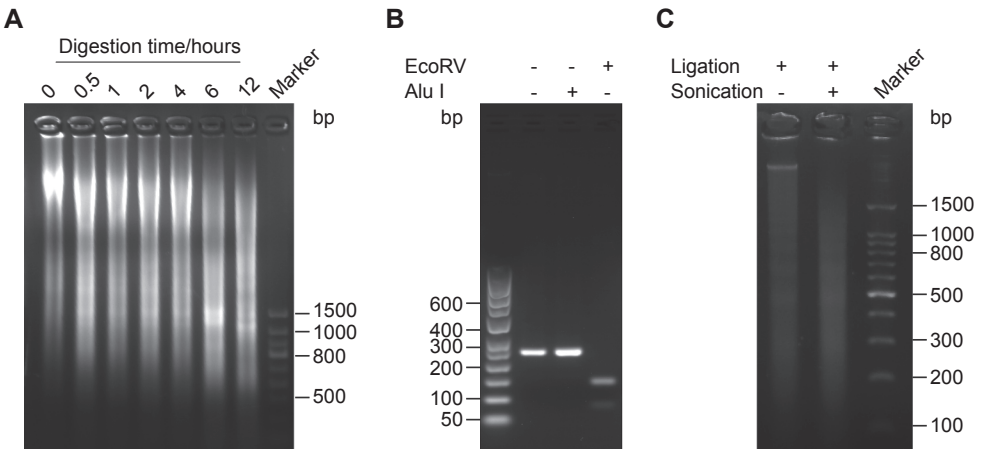



**Fig. 2.** Quality control of experimental steps. (A) Approximately 5 million formaldehyde fixed mES cells were digested by 100 U AluI at different timepoints, a representative example showing titration of AluI digestion of crosslinked DNA, and 6 h is enough for efficient digestion. (B) A specific ligation product with designed primer pairs was amplified to assess the efficiency of 3′dATP tailing and proximal ligation. PCR product (lane 1), PCR product digested with AluI (lane 2), EcoRV (lane 3); only 3′dATP tailing ligation products were specifically digested. (C) 50 μg DNA were sonicated with two cycles on Biotuptor, well-sonicated DNA forms a smear of 0.2–2 kb on agarose gel with a median length of 750 bp.

supernatant.

4) Gently resuspend pellet with 50 μl 0.5% SDS and incubate at 62 °C for 10 min.
5) After heating, add 140 μl $H_2O$ and 25 μl 10% Triton X-100 to quench SDS, mix gently and incubate at 37 °C for 15 min.
6) Add 25 μl 10× NEB Cutsmart Buffer and 100 U AluI (NEB) and digest the chromatin for no more than 12 h at 37 °C on a thermomixer at 700 rpm (the reaction can be scaled up according to cell numbers) (Fig. 2A-2B).

(Note: the amount of starting material varies between cell types, ranging from 1 to 10 million. We began with 5 million mES cells.)

*2.2.4. 3′dATP tailing*

A recent ChIA-PET experiment observed that long-range chromatin interactions are preferentially ligated to biotin-labeled bridge linkers; longer PET reads contribute toward a better understanding of gene control by uncovering long-range interacting regions with putative regulatory functions [14,27]. Inspired by this finding, we use T: A ligation guided by a biotinylated bridge linker to reduce random DNA collisions and to increase the informative fraction of valid long pair reads.

1) Add dATP on both 3′ blunt ends created by AluI with the incorporation of DNA polymerase I large (Klenow) fragment to mediate proximal ligation.
2) Collect AluI-digested nuclei, wash twice with 1× NEB buffer2 and discard supernatant.
3) Gently resuspend nuclei in 400 μl of Klenow (3′-5′exo-) solution (40 μl NEB buffer2, 8 μl 10 mM dATP, 40 μl 10% TritonX-100, 304 μl $H_2O$, and 8 μl Klenow (3′-5′exo-)), and incubate at 37 °C for 1 h at 700 rpm on a thermomixer.

*2.2.5. In situ chromatin ligation with bridge linker*

To increase the efficiency of proximity ligation, the sequence length, composition and secondary structure are carefully considered when designing the bridge linker. We selected a 20-bp linker with a 3′-protruding T nucleotide at both ends and a biotin group in the middle. An additional benefit of using this biotinylated bridge linker instead of biotin-14-dCTP was the significantly reduced experimental costs.

1) Collect nuclei and wash twice with 1× T4 DNA ligase buffer and discard supernatant.
2) Gently resuspend nuclei in 1 ml of proximity ligation solution (100 μl T4 DNA ligase buffer, 100 μl 10% TritonX-100, 782 μl $H_2O$, 6 μl T4 DNA ligase, 10 μl 10 mg/ml BSA, 2 μl bridge linker (200 ng/μl)), and rotate at 700 rpm at room temperature for at least 6 h. Centrifuge the tube at 5000 rpm for 3 min at 4 °C to collect nuclei.

*2.2.6. Chromatin fraction isolation*

Typical Hi-C experiments enrich biotin-labeled DNA fragments with streptavidin beads, which also pull down non-proximity-ligated dangling ends, i.e., defined as “inward” pairs among invalid reads. These invalid data could be removed using an extra exonuclease digestion step [28,29], A previous study reported that chromatin fraction isolation could further reduce background noise via a method named “intrinsic 3C” (i3C). i3C provided evidence that centrifugation after digestion could remove more than 40% of total chromatin; this improved signal quality could be attributed due to reduced proportions of “bystander/baseline” ligation [30]. Motivated by this work, we proposed an attractive option that reduces background noise by taking advantage of chromatin fraction isolation. However, this non-essential step may result in a decrease of ligated products. We recommend omitting this step if the amount of starting material is limited.

1) Gently lyse the nuclei with 500 μl of ice-cold NP-40 lysis buffer on ice for 5 min. Transfer the cell lysate to the top of 2.5 volumes of a sucrose cushion comprising 24% (wt/vol) sucrose in NP-40 lysis buffer. Centrifuge at 12,000 rpm for 10 min at 4 °C to isolate nuclei.
2) Wash the nuclei pellet once with 1 × PBS/1 mM EDTA.
3) Gently resuspend the nuclei pellet with 0.5 ml glycerol buffer, followed by adding an equal volume of nuclei lysis buffer on ice for 2 min.
4) Centrifuge at 12,000 rpm for 2 min at 4 °C to isolate the chromatin pellet.
5) Wash the chromatin pellet twice with 1× PBS/1 mM EDTA.

*2.2.7. Exonuclease digestion and DNA purification*

1) Resuspend the isolated chromatin in 300 μl exonuclease mixture (30 μl 10× lambda exonuclease buffer, 3 μl lambda exonuclease, 3 μl exonuclease I, and 264 μl $H_2O$) and incubate at 37 °C on a thermomixer at 700 rpm for 1 h.
2) Centrifuge at 12,000 rpm for 2 min at 4 °C to collect chromatin and wash twice with 1× PBS/1 mM EDTA.
3) To reverse crosslink DNA, resuspend chromatin with 900 μl digestion buffer (10 mM Tris-HCl, pH 8, 2 mM EDTA, 1% SDS, and 1 mg/ml proteinase K), and incubate for 30 min at 55 °C. Then, add 100 μl 5 M NaCl and incubate overnight at 65 °C.
4) Thoroughly extract genomic DNA with an equal volume of phenol/chloroform/isoamyl alcohol and centrifuge at 12,000 rpm for 15 min at 4 °C.
5) Transfer the aqueous top layer to a new tube and add 1/10 vol of 3 M NaOAc, 2 vol of 100% ethanol and 1 μl GlycoBlue; store at −80 °C for 1 h. Recover DNA by centrifugation at 12,000 rpm for 20 min at 4 °C.
6) Wash DNA pellet with 75% ethanol twice. Decant ethanol and air-

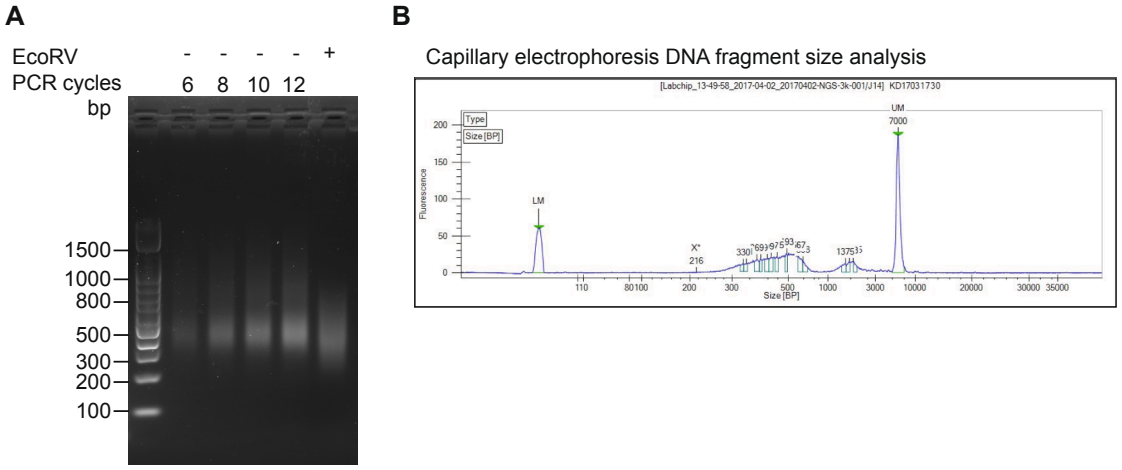



**Fig. 3.** PCR amplification and size selection. (A) PCR amplification was performed as a final quality control. PCR titration of the DNA library showed a downward shift of the last lane after digestion with EcoRV, indicating efficient 3′dATP tailing and bridge linker-mediated ligation. (B) Capillary electrophoresis of the size distribution of the DNA fragments.

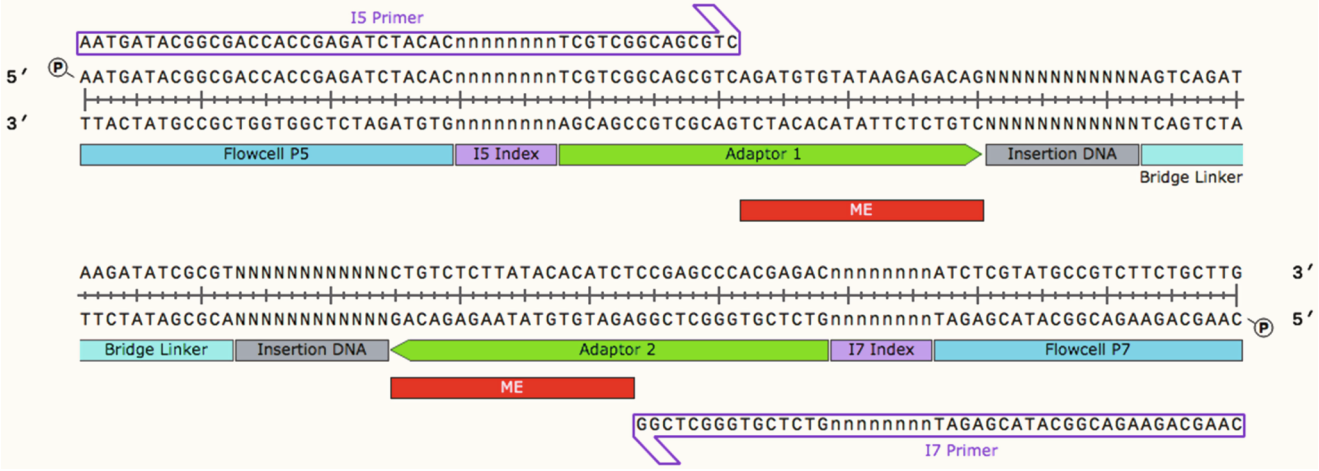



**Fig. 4.** Typical constitution of bridge linker-mediated hybrid DNA molecule, with paired Illumina sequencing adaptors on both ends and 20 bp bridge linker between proximal DNA fragments. Flowcell P5/P7 is the sequence binding with Illumina Sequencing Chip (flowcell); Adaptor 1/2 and ME are the sequence assembled with Tn5. BAT Hi-C library is PCR amplified with primer I5 and I7 primer with specific index, respectively.

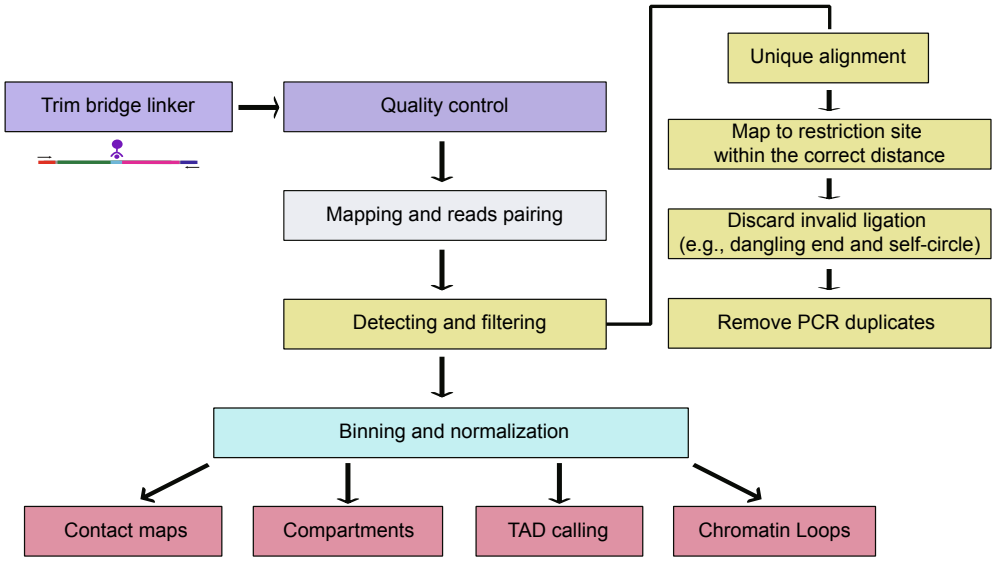



**Fig. 5.** BAT Hi-C data analysis workflow, organized into four modules: (1) linker trimming, (2) read alignment and pairing, (3) detection and filtering of valid interactions, (4) binning and contact map normalization.

dry the pellet.
7) Dissolve the DNA pellet in 200 µl 10 mM Tris-HCl, pH 8 for 15 min at 37 °C.

*2.2.8. DNA sonication*

Sonication is another critical step that fragments the chromatin to an ideal size for subsequent experiments. Therefore, depending on the specific instrument and sonication parameters, including strength, duration and volume, require optimization. We recommend sonicating 50 µg well-purified chimera DNA to ~750 bp (Fig. 2C).

1) Transfer 50 µg DNA solution into a Biorupter tube, and adjust the total volume to 200 µl. Sonicate DNA by Biorupter: high energy, 30 s working time, 60 s interval, 2 cycles.
2) Check DNA on 1.5% agarose gel (confirm the length to be 0.2–2 kb) (Fig. 2C).

*2.2.9. Biotin pull-down*

1) Transfer 30 µl of 10 mg/ml M280 streptavidin Dynabeads to a 1.5-ml tube, and wash three times with 500 µl of 1 × TWB buffer and one time with 1 × B&W buffer.
2) Mix the sonicated DNA solution with an equal volume of 2 × B&W buffer to the prewashed Dynabeads, Rotate at room temperature for 30 min.
3) Wash five times with 2 × SSC/0.5% SDS, three times with TE buffer, and transfer to a new tube after last wash.
4) Wash one time with 100 µl 1 × TD buffer.

(Note: due to the high affinity between streptavidin and biotin, the incubation time of purified biotinylated DNA and streptavidin beads should be limited; 30 min is sufficient.)

*2.2.10. Tn5 tagmentation and PCR amplification*

As mentioned in Section 2.1, Tn5 transposase can simultaneously fragment DNA ligation products and add adaptors to cleavage sites. Combining these two processes in one single tube reaction will increase the overall efficiency by reducing preparation time.

1) Prepare the following tagmentation mix:
   a. DNA & Dynabeads mixture
   b. 25 µl 2 × Tagmentation buffer
   c. 2.5 µl Transposase enzyme (TDE1)
   d. 22.5 µl $H_2O$
2) Adjust the total volume to 50 µl, incubate the reaction at 55 °C for 10 min, and then at 10 °C for 10 min in the PCR machine.
3) Place samples on a magnet and remove the supernatant. Wash twice with 2 × SSC/0.5% SDS, twice with TE buffer, and transfer to a new tube after last wash.
4) Resuspend beads with 30 µl 10 mM Tris HCl (pH 8.0).
5) Prepare following PCR mix:
   a. 15 µl DNA library-coated beads
   b. 5 µl $H_2O$
   c. 15 µl NPM mix
   d. 5 µl PPC PCR primer
   e. 5 µl Index Primer1 (i7)
   f. 5 µl Index Primer2 (i5)
6) Adjust the total volume to 50 µl, and run the following program: 72 °C 3 min; 98 °C 10 s, 63 °C 30 s, 72 °C 50 s, cycle 10, 72 °C 5 min (Fig. 3A).
7) Use 2% agarose gel electrophoresis size selection to purify 300–600 bp DNA with a gel purification kit (Fig. 3B).
8) Subject the purified DNA to 150 bp paired-end sequencing.

(Note: due to the high transposase activity of Tn5, we recommend titrating the amount of Tn5 used and the reaction time for tagmentation. For example, 2.5 µl of Tn5 from an Illumina Kit is suitable for preparing of 50 µg well-prepared hybrid DNA.)

*2.3. Quality control*

Several simple but important checks can provide information on whether the experiment was successful. To assess BAT Hi-C library quality during library preparation, we performed two checks after chromatin digestion and DNA ligation.

The first is to conduct an assay to determine the optimal digestion timepoint for mES cells; over digestion will lead to nuclei and chromatin fragmentation. Approximately 5 million crosslinked mES cells (fixed in 1% formaldehyde) were lysed and divided into seven equal aliquots, and each aliquot was digested with an equivalent amount of AluI for titration at each timepoint. Purified genomic DNA from each sample was subjected to gel electrophoresis. In general, efficient AluI digestion will result in many DNA fragments with sizes less than 1 kb. By checking the size distribution of DNA fragments, the sizes decreased globally after 6 h, indicating that 6 h was an appropriate choice for an efficient digestion in our experiment (Fig. 2A).

The second quality control check is to design specific primer pairs to amplify ligation products for two adjacent AluI digestion sites. The bridge linker contains an EcoRV digestion site. Because AluI requires efficient 3′dATP tailing and bridge linker-mediated T-A ligation, properly ligated chimeras would be digested by EcoRV but not AluI. Therefore, we carried out the following digestion assay with either AluI or EcoRV; as expected, only the latter was digested (Fig. 2B). For a valid library, there should be an apparent size shift of DNA fragments after enzyme digestion. Subsequently, DNA was sonicated to 0.2–2 kb with a median length of 750 bp (Fig. 2C).

After library preparation, as a final quality control before sequencing, we performed PCR titration to ensure sufficient library complexity. Given that PCR overamplification would reduce the complexity of BAT Hi-C libraries, we selected an optimal amount of PCR cycles by PCR titration. Normally, 8–10 PCR cycles will generate sufficient DNA for sequencing. For our starting material, 8 was the required minimum number of PCR cycles (Fig. 3A). Size selection of the PCR amplified product was performed with 2% agarose gel electrophoresis. Finally, the integrity and size distribution of the BAT Hi-C library was assessed by Bioanalyzer. The Bioanalyzer profile showed that the majority of the final DNA library was in the 300–600 bp range (Fig. 3B).

*2.4. High-throughput sequencing*

In principle, the BAT Hi-C library preparation protocol we proposed is applicable to any Illumina sequencing platform supporting Nextera adaptors. Samples were ligated using barcoded adaptors which allows many samples to be simultaneously multiplexed and sequenced in a single run and prepared according to manufacturer instructions. Next, qualified libraries were sequenced using an Illumina HiseqX10 sequencer with 150-bp paired-end mode to include bridge linker information, as demonstrated below (Fig. 4). The PCR primers used to

construct our libraries are shown in Supplementary Table B.2.

### 2.5. Bioinformatics analysis

#### 2.5.1. BAT Hi-C data processing

BAT Hi-C data processing comprised four modules (Fig. 5). To help other investigators reproduce our results, we provide command line examples for freely available software. Most utilities are adapted from the HiC-Pro pipeline without additional scripting [31].

Even a novice bioinformatician can perform the same analysis in Linux and Python environments with minimal programming skills.

1) Download the raw reads as. *fastq* files from the Novogene Cloud server to local computational clusters.
2) Install the required software and obtain the reference genome (e.g. mm10) sequences from the UCSC Genome Browser.
3) Create Bowtie 2 index and genome annotation files.

Command example for mm10:

```
#bowtie2-build /DATA/genomes/Mmusculus/mm10/seq/mm10.fa mm10
#curl -s ftp://hgdownload.cse.ucsc.edu/goldenPath/mm10/database/chromInfo.txt.gz | gunzip -c
```

4) Establish appropriate digesting files for mm10.

Command example for genome digesting:

```
#/DATA/software/hi-c/HiC-Pro_2.9.0/bin/utils/digest_genome.py -r ACGT -o mm10_Alul.bed /DATA/genomes/Mmusculus/mm10/seq/mm10.fa
```

5) Scan and trim bridge linker of paired-end reads using the trimLinker program (part of ChIA-PET2 [32]) with options. Reads containing at least one instance of the bridge-linker sequence at either end are adapted for the following pipeline.

Command example for linker trimming:

```
#trimLinker -t 12 -m 1 -k 1 -l 16 -o./BAT_Hi-C_mESC_wt_rep1 -n BAT_Hi-C_mESC_wt_rep1 -A ACGCGATATCTTATC -B AGTCA GATAAGATAT s1215Rpb1-0 h-1_HCKYCCCXY_L2_1.fq.gz s1215R pb1-0 h-1_HCKYCCCXY_L2_2.fq.gz &
```

6) Link the trimmed fastq files into a given project or experiment folder, and initialize the folder structure as follows:

Folder organization example:

```
Pol II projects/
└── BAT_Hi-C_mESC_mm10
    └── DATA
        ├── BAT_Hi-C_mESC_wt_rep1
        └── BAT_Hi-C_mESC_wt_rep2
```

7) Edit the configuration file 'config-hicpro.txt' in local project folder to process trimmed reads of each replicate through HiC-Pro pipeline. Keep uniquely aligned reads and assign them to an Alul restriction fragment.

Configuration example for HiC-Pro mapping:

```
    ###################################
###################################
####
    ## Digestion Hi-C
    ###################################
###################################
#
    GENOME_FRAGMENT = mm10_Alul.bed
    LIGATION_SITE = AGCT
    MIN_FRAG_SIZE = 10
    MAX_FRAG_SIZE = 100000
    MIN_INSERT_SIZE = 70
    MAX_INSERT_SIZE = 600
```

8) Filter for valid interactions. Discard read pairs of 'dangling ends' and 'self-circles' incorrectly assigned to restriction enzyme digestion sites. Exclude duplicate pairs due to PCR artifacts for downstream analysis.
9) Parse the uniquely valid contacts pairs to build the raw Hi-C contact matrix at different resolutions.

Command example for bining:

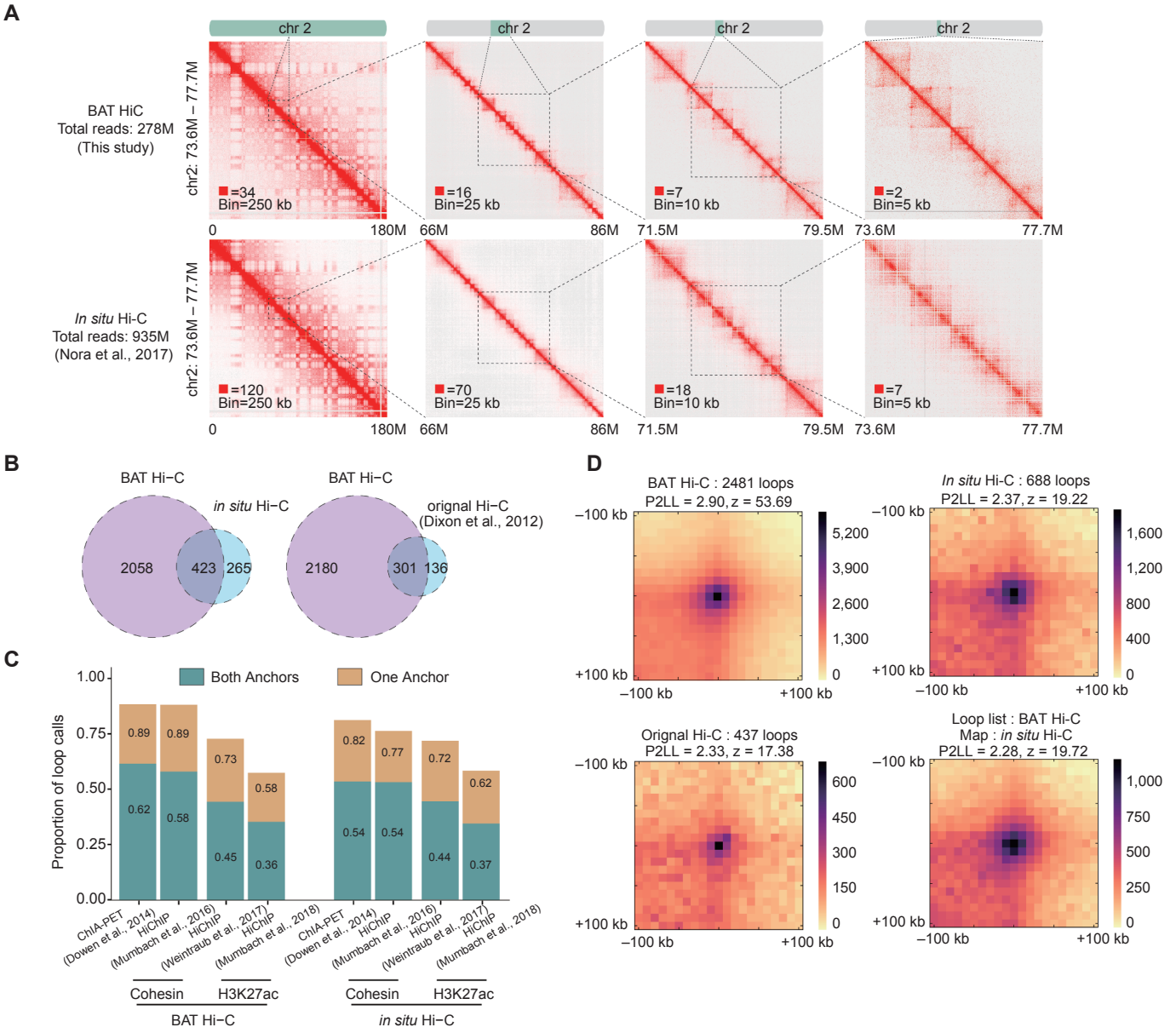
```
#/DATA/software/hi-c/HiC-Pro_2.9.0/scripts/build_matrix –binsize 25,000 –chrsizes /DATA/software/hi-c/HiC-Pro_2.9.0/annotation/chrom_mm10.sizes –ifile BAT_Hi-C_mESC_wt_rep1.allValidPairs –oprefix 0h_1_25kb
```

10) Use iterative correction method (ICE) to normalize the contact matrix [33].

Command example for normalization:

```
#python /DATA/software/hi-c/HiC-Pro_2.9.0/scripts/ice –results_filename BAT_Hi-C_mESC_wt_rep1_25kb_iced.matrix –filter_low_counts_perc 0.05 –filter_high_counts_perc 0.02 –max_iter 100 –eps 0.1 –remove-all-zeros-loci –verbose 1 BAT_Hi-C_mESC_wt_rep1_25kb.matrix
```

11) Convert HiC-pro output **allValidPairs* files into the Juicer compatible *.hic* format using the *hicpro2juicebox* function [34].




*(caption on next page)*

**Fig. 6.** Delineation of the 3D genome architecture of mES cells. (A) Interaction heatmaps comparing the KR-normalized contact matrices (250-kb bin, chr2) between BAT Hi-C (278 million raw reads) and *in situ* Hi-C [36] (935 million raw reads). Zoomed-in views (25-kb, 10-kb, and 5-kb bin) are also shown with Juicebox software. (B) Venn diagram showing the overlap number of two sets for HiCCUPS loop calls; left is between BAT Hi-C and *in situ* Hi-C, and right is between BAT Hi-C and original Hi-C [17]. (C) Bar plots depicting the overlap proportion of cohesin and H3K27ac high-confidence loops at either one or both BAT Hi-C loop anchors compared with *in situ* Hi-C. (D) Aggregate peak analysis of loops at 10-kb resolution in BAT Hi-C, *in situ* Hi-C, and original Hi-C datasets in mES cells. (E) Browser view presenting two representative loci: Pax6 (Left) and Zfp42 (Right). ChIP-seq profiles for H3K4me3, H3K27me3, H3K36me3, and H3K27ac. RNA-seq from mES cells are overlayed on the top, Loop calls of BAT Hi-C, *in situ* Hi-C, cohesin HiChIP [15] and H3K27ac HiChIP [37] as arcs displaying normalized interactions. The bottom track is the selected RefSeq gene positions of Pax6, Pax6os1, Elp4 and Zfp42. Gray box: focused interaction regions.

Command example for data visualization:

```
# /DATA/software/hi-c/HiC-Pro_2.9.0/bin/utils/hicpro2juicebox.sh -i BAT_Hi-C_mESC_wt_rep1.allValidPairs -g /DATA/genomes/mouse.mm10.genome -j /DATA/software/hi-c/juicer/juicer_tools_linux_0.8.jar
```

### 2.5.2. Validation of interaction matrices

BAT Hi-C is theoretically broadly applicable to any situation in which canonical Hi-C would be used. To benchmark our approach, we conducted the BAT Hi-C experiment in mouse embryonic stem cells and compared the results with an *in situ* Hi-C data set. Raw FASTQ data were first processed using the HiC-Pro pipeline to filter a total of 278 million paired-end reads into informative unique read pairs. Of the total reads, more than 65% were unique read pairs in our libraries. Of note, in canonical Hi-C, read pairs of the same restriction fragment sites are considered invalid, and the majority of invalid reads are “dangling ends” originating from nonligated DNA [31,35].

Conversely, we observed that the only type of invalid pairs from the BAT Hi-C libraries was ‘self-circles’ (Table B.3). This finding was consistent with the fact that bridge linkers would bias long-distance interactions. However, a general concern for Hi-C methods, regardless of library complexity and sequencing depth, is that long-range loops are likely missed. We reasoned that we could circumvent this limitation with improved T-A ligation efficiency to delineate chromatin loops. Because loop calling algorithms heavily rely on the number of unique read pairs observed and are distance dependent at a given sequencing depth, the proportion of long-range contact pairs ( > 20 kb) among total reads was considered a strong indicator of method efficiency. In support of this notion, a higher fraction of unique read pairs in our libraries were mapped as long *cis*-interactions ( > 20 kb) (Table B.3). Next, we broadly evaluated the performance of BAT Hi-C by examining its conformational features relative to published *in situ* Hi-C data at the coarsest level of genomic organization. The interactions between loci are shown on a 2D heatmap at four scales arranged from chromosome-scale (250 kb) to single-gene resolution (5 kb) using KR-normalized contact matrices (Fig. 6A). Visual inspections of the heatmaps of these two Hi-C methods revealed highly concordant patterns illuminating various features of chromatin folding, as previously reported in *in situ* Hi-C analyses [22,36]. Notably, the BAT Hi-C maps were capable of identifying focal interactions at 10-kb resolution with 278 million reads, which was comparable with the *in situ* Hi-C maps sequenced to a 3-fold greater depth.

### 2.5.3. Comparing chromatin loop calls

To comprehensively validate the interactions predicted by BAT Hi-C, we aimed to compare high-confidence loop calls of mES cells generated by our approach and other existing methods. For analyses pertaining to different chromatin conformation assays, raw data were obtained from previous studies and processed in-house. To provide more equitable comparisons, we randomly sampled equal numbers of long-range ( > 20 kb) intra-chromosomal read pairs (n = 75 million) from each library for subsequent analyses and applied the same processing steps described above. Significant loop calls were performed by the HiCCUPS algorithm from merged biological replicates using default parameters at 5-kb, 10-kb and 25-kb resolutions (with a false discovery rate filter at 0.1). Importantly, we identified a slightly larger number of loops for BAT Hi-C data (2481) in mES cells, compared with *in situ* Hi-C (688) and original Hi-C (437) downsampled sets. Furthermore, as anticipated, these three Hi-C methods displayed high degrees of loop consistency (Fig. 6B). With the *in situ* Hi-C method, 423 of 688 loops were shared with BAT Hi-C, and 301 of the 437 original Hi-C loops were detected by our approach.

We then sought to characterize the structural and regulatory features of our high-confidence loop calls. Recent high-resolution HiChIP or ChIA-PET experiments have revealed the cohesin and histone H3 lysine 27 acetylation (H3K27ac)-mediated genome architecture and enhancer-promoter interactions. Given the efficiency gains of BAT Hi-C compared with canonical Hi-C, we determined whether the enriched looping interactions of BAT Hi-C could also be detectable by these two protein-centric methods. Consequently, loop calls for cohesin and H3K27ac HiChIP were made via the same approach as that used for Hi-C. Independent HiChIP data sets derived from two different laboratories were used to avoid potential biases [37,38]. Significant interactions of cohesin ChIA-PET were directly downloaded from the supplementary material of Dowen et al. [21]. Loop overlap was measured by comparing loop anchors for all pairs of loops using pairToPair in bedtools with parameters ‘-slop 1000 –type both/either -is’. Because of the uneven numbers of interactions identified in the four individual data sets (Table B.5), we investigated the proportion of cohesin- and H3K27ac-enriched loops at either one or both ends of our loop anchors. Overall, the overlap was high, with 89% of our loops anchored by cohesin at either one end and 73% present in the H3K27acmediated looping events. These results were in close agreement with *in situ* Hi-C data (Fig. 6C).

Finally, we focused on two representative loci to conform the enrichment pattern [39,40]. The Pax6 gene region in the 11p13 locus was selected within which multiple regulatory elements reside. Pax6 is required for embryonic stem cell differentiation into neural lineages. Previous studies have suggested that the transcriptional upregulation of Pax6 during the conversion is accompanied by the formation of enhancer-promoter contacts and the nearby emergence of enhancer signature H3K27ac [41]. To scrutinize the functional relationships of our identified loops, we integrated cell-type specific epigenetic data to annotate the loop anchors connecting promoter, enhancer or insulators elements. As expected, both BAT and *in situ* Hi-C detected previously reported long-range interactions between Pax6 and distal regulatory enhancers. The chromatin fragments anchoring the base of detected loops were highly correlated with active histone markers, such as H3K27ac and H3K4me3 (Fig. 6E). Similarly, we observed that H3K27ac and H3K4me3 were insulated within loop anchors demarcated by CTCF at the Zfp42 locus in mES cells. Altogether, these results suggested that the combination of chromatin immunoprecipitation followed by sequencing (ChIP–seq) in BAT Hi-C may provide topological insights into gene regulation.

### 2.5.4. Aggregate peak analysis

To further evaluate the quality of BAT Hi-C and typical Hi-C approaches, genome-wide loop intensity was assessed by aggregate peak analysis (APA) to estimate signal-to-background ratios [34,42]. Loops were first piled up on the center of a 210-kb × 210-kb matrix with a 10-

kb resolution of KR-normalized data. Juicer APA with default parameters was applied to discover the enrichment of global mean profiles of all called loops with nearby regions. Generally, the peak-to-lower left (P2LL) score was used to measure the aggregate enrichment, which was calculated by dividing observed contact in a searching window by the expected bottom-left submatrix. Standard APA heatmap outputs were used to visualize in a global fashion. In line with expectations, BAT Hi-C data exhibited a relatively higher signal-to-background ratio at chromatin loops compared with the other two (the P2LL was 2.90 and the z score was 53.69). We analyzed HiCCUPS loop calls from the BAT Hi-C and canonical Hi-C contact maps, as shown in APA plots (Fig. 6D). The P2LL reached 2.28, suggesting most chromatin loops from ours were well matched with the signals from *in situ* Hi-C. These results indicated that BAT Hi-C provides greater efficiency with low spatial noise, making it a valuable alternative to canonical Hi-C.

## 3. Conclusions

In summary, we successfully developed a highly efficient and cost-effective technology to facilitate gene regulation and the exploration of the 3D genomic landscape. The method employs AluI enzyme digestion to fragment chromatin, a biotinylated bridge linker to guide proximity T-A ligation, chromatin isolation and exonuclease digestion, and Tn5 transposase-mediated tagmentation. Overall, this method enables the generation of high-quality DNA libraries and achieves loop-resolution interaction maps at a fraction of the cost of canonical Hi-C. For future improvements, miniaturized devices and more efficient enzymes may reduce the required amount of starting material for single cell library preparation [43,44]. Similar to all 3C-based procedures, the quality control experiments discussed in detail will also help validate and ensure the specificity of the observed long-range chromatin interactions. Additionally, BAT Hi-C could be coupled with results from complementary experiments (e.g., 4C and ChIP-qPCR) or with protein-centric chromatin conformation techniques [12,13,15] to determine the functional impact of chromatin looping.

Funding

This work was supported by funds from the Ministry of Science and Technology of China and National Natural Science Foundation of China (Grants 2017YFA0506600 and 31871309), the Young Thousand Talents Program of China, and grants from the Peking-Tsinghua Center for Life Sciences and Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education at Peking University School of Life Sciences to X. Ji. J. H. was supported by a grant from the China Postdoctoral Science Foundation.

## Appendix A

Software and Data availability

- cutadapt v1.18 (https://cutadapt.readthedocs.io/en/stable/)
- Bowtie 2 v2.3.4 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml)
- samtools v1.9 (https://github.com/samtools/samtools)
- BEDTools v2.17.0 (https://bedtools.readthedocs.io/)
- ChIA-PET2 v0.9.3 (*trimLinker* program: https://github.com/GuipengLi/ChIA-PET2)
- UCSC Genome Browser (https://genome.ucsc.edu/index.html)
- WashU Epigenome Browser (http://epigenomegateway.wustl.edu/browser/)
- HiC-Pro v2.9.0 (https://github.com/nservant/HiC-Pro)
- HiCPlotter v0.7.1 (https://github.com/kcakdemir/HiCPlotter)
- Juicer and JuicerBox v1.8.8 on Linux (https://github.com/theaidenlab/juicebox/wiki/Download)
- Bioconductor v3.6 (https://www.bioconductor.org/)
- R v3.5.0 (RColorBrewer, ggplot2 and HiTC packages: https://cran.r-project.org/)
- FastQC v0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)
- HICUP v0.7.0 (http://www.bioinformatics.babraham.ac.uk/projects/hicup/)
- Python v2.7 (pysam (> =0.8.3), bx-python (> =0.5.0), numpy (> =1.8.2), and scipy(> =0.15.1) libraries: https://www.python.org/)
- hiclib library (https://bitbucket.org/mirnylab/hiclib)

Raw and processed BAT Hi-C data are available at NCBI GEO with accession number GSE136756.

## Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymeth.2019.08.004.

## References

[1] W.A. Bickmore, The spatial organization of the human genome, Annu. Rev. Genomics Hum. Genet. 14 (14) (2013) 67–84.
[2] J. Dekker, L. Mirny, The 3D genome as moderator of chromosomal communication, Cell 164 (6) (2016) 1110–1121.
[3] D. Hnisz, D.S. Day, R.A. Young, Insulated neighborhoods: structural and functional units of mammalian gene control, Cell 167 (5) (2016) 1188–1200.
[4] M. Yu, B. Ren, The three-dimensional organization of mammalian genomes, Annu. Rev. Cell Dev. Biol. 33 (33) (2017) 265–289.
[5] M.J. Rowley, V.G. Corces, Organizational principles of 3D genome architecture, Nat. Rev. Genet. 19 (12) (2018) 789–800.
[6] B. Bonev, G. Cavalli, Organization and function of the 3D genome, Nat. Rev. Genet. 17 (11) (2016) 661–678.
[7] J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation, Science 295 (5558) (2002) 1306–1311.
[8] Z. Zhao, G. Tavoosidana, M. Sjolinder, A. Gondor, P. Mariano, S. Wang, C. Kanduri, M. Lezcano, K.S. Sandhu, U. Singh, V. Pant, V. Tiwari, S. Kurukuti, R. Ohlsson, Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions, Nat. Genet. 38 (11) (2006) 1341–1347.
[9] M. Simonis, P. Klous, E. Splinter, Y. Moshkin, R. Willemsen, E. de Wit, B. van Steensel, W. de Laat, Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C), Nat. Genet. 38 (11) (2006) 1348–1354.
[10] E. Splinter, E. de Wit, H.J. van de Werken, P. Klous, W. de Laat, Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation, Methods 58 (3) (2012) 221–230.
[11] J. Dostie, T.A. Richmond, R.A. Arnaout, R.R. Selzer, W.L. Lee, T.A. Honan, E.D. Rubio, A. Krumm, J. Lamb, C. Nusbaum, R.D. Green, J. Dekker, Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements, Genome Res. 16 (10) (2006) 1299–1309.
[12] M.J. Fullwood, M.H. Liu, Y.F. Pan, J. Liu, H. Xu, Y. Bin Mohamed, Y.L. Orlov, S. Velkov, A. Ho, P.H. Mei, E.G.Y. Chew, P.Y.H. Huang, W.J. Welboren, Y.Y. Han, H.S. Ooi, P.N. Ariyaratne, V.B. Vega, Y.Q. Luo, P.Y. Tan, P.Y. Choy, K.D.S.A. Wansa, B. Zhao, K.S. Lim, S.C. Leow, J.S. Yow, R. Joseph, H.X. Li, K.V. Desai, J.S. Thomsen, Y.K. Lee, R.K.M. Karuturi, T. Herve, G. Bourque, H.G. Stunnenberg, X.L. Ruan, V. Cacheux-Rataboul, W.K. Sung, E.T. Liu, C.L. Wei, E. Cheung, Y.J. Ruan, An oestrogen-receptor-alpha-bound human chromatin interactome, Nature 462 (7269) (2009) 58–64.
[13] R.X. Fang, M. Yu, G.Q. Li, S. Chee, T. Liu, A.D. Schmitt, B. Ren, Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq, Cell Res. 26 (12) (2016) 1345–1348.
[14] X. Li, O.J. Luo, P. Wang, M. Zheng, D. Wang, E. Piecuch, J.J. Zhu, S.Z. Tian, Z. Tang, G. Li, Y. Ruan, Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions, Nat. Protoc. 12 (5) (2017) 899–915.
[15] M.R. Mumbach, A.J. Rubin, R.A. Flynn, C. Dai, P.A. Khavari, W.J. Greenleaf, H.Y. Chang, HiChIP: efficient and sensitive analysis of protein-directed genome architecture, Nat. Methods 13 (11) (2016) 919–922.
[16] E. Lieberman-Aiden, N.L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B.R. Lajoie, P.J. Sabo, M.O. Dorschner, R. Sandstrom, B. Bernstein, M.A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L.A. Mirny, E.S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome, Science 326 (5950) (2009) 289–293.
[17] J.R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J.S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions, Nature 485 (7398) (2012) 376–380.
[18] E.P. Nora, B.R. Lajoie, E.G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N.L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Bluthgen, J. Dekker, E. Heard, Spatial partitioning of the regulatory landscape of the X-inactivation centre, Nature 485 (7398) (2012) 381–385.
[19] G.L. Li, X.A. Ruan, R.K. Auerbach, K.S. Sandhu, M.Z. Zheng, P. Wang, H.M. Poh,

Y. Goh, J. Lim, J.Y. Zhang, H.S. Sim, S.Q. Peh, F.H. Mulawadi, C.T. Ong, Y.L. Orlov, S.Z. Hong, Z.Z. Zhang, S. Landt, D. Raha, G. Euskirchen, C.L. Wei, W.H. Ge, H.E. Wang, C. Davis, K.I. Fisher-Aylor, A. Mortazavi, M. Gerstein, T. Gingeras, B. Wold, Y. Sun, M.J. Fullwood, E. Cheung, E. Liu, W.K. Sung, M. Snyder, Y.J. Ruan, Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation, Cell 148 (1–2) (2012) 84–98.

[20] X. Ji, D.B. Dadon, B.E. Powell, Z.P. Fan, D. Borges-Rivera, S. Shachar, A.S. Weintraub, D. Hnisz, G. Pegoraro, T.I. Lee, T. Misteli, R. Jaenisch, R.A. Young, 3D chromosome regulatory landscape of human pluripotent cells, Cell Stem Cell 18 (2) (2016) 262–275.

[21] J.M. Dowen, Z.P. Fan, D. Hnisz, G. Ren, B.J. Abraham, L.N. Zhang, A.S. Weintraub, J. Schujiers, T.I. Lee, K. Zhao, R.A. Young, Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes, Cell 159 (2) (2014) 374–387.

[22] S.S.P. Rao, M.H. Huntley, N.C. Durand, E.K. Stamenova, I.D. Bochkov, J.T. Robinson, A.L. Sanborn, I. Machol, A.D. Omer, E.S. Lander, E.L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping, Cell 159 (7) (2014) 1665–1680.

[23] D. Hnisz, A.S. Weintraub, D.S. Day, A.L. Valton, R.O. Bak, C.H. Li, J. Goldmann, B.R. Lajoie, Z.P. Fan, A.A. Sigova, J. Reddy, D. Borges-Rivera, T.I. Lee, R. Jaenisch, M.H. Porteus, J. Dekker, R.A. Young, Activation of proto-oncogenes by disruption of chromosome neighborhoods, Science 351 (6280) (2016) 1454–1458.

[24] F.L. Jin, Y. Li, J.R. Dixon, S. Selvaraj, Z. Ye, A.Y. Lee, C.A. Yen, A.D. Schmitt, C.A. Espinoza, B. Ren, A high-resolution map of the three-dimensional chromatin interactome in human cells, Nature 503 (7475) (2013) 290–294.

[25] T. Nagano, C. Varnai, S. Schoenfelder, B.M. Javierre, S.W. Wingett, P. Fraser, Comparison of Hi-C results using in-solution versus in-nucleus ligation, Genome Biol 16 (2015) 175.

[26] J. Zhang, H.M. Poh, S.Q. Peh, Y.Y. Sia, G. Li, F.H. Mulawadi, Y. Goh, M.J. Fullwood, W.K. Sung, X. Ruan, Y. Ruan, ChIA-PET analysis of transcriptional chromatin interactions, Methods 58 (3) (2012) 289–299.

[27] Z. Liang, G. Li, Z. Wang, M.N. Djekidel, Y. Li, M.P. Qian, M.Q. Zhang, Y. Chen, BL-Hi-C is an efficient and sensitive approach for capturing structural and regulatory chromatin interactions, Nat. Commun. 8 (1) (2017) 1622.

[28] J.M. Belton, R.P. McCord, J.H. Gibcus, N. Naumova, Y. Zhan, J. Dekker, Hi-C: a comprehensive technique to capture the conformation of genomes, Methods 58 (3) (2012) 268–276.

[29] H. Belaghzal, J. Dekker, J.H. Gibcus, Hi-C 2.0: an optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation, Methods 123 (2017) 56–65.

[30] L. Brant, T. Georgomanolis, M. Nikolic, C.A. Brackley, P. Kolovos, W. van Ijcken, F.G. Grosveld, D. Marenduzzo, A. Papantonis, Exploiting native forces to capture chromosome conformation in mammalian cell nuclei, Mol. Syst. Biol. 12 (12) (2016) 891.

[31] N. Servant, N. Varoquaux, B.R. Lajoie, E. Viara, C.J. Chen, J.P. Vert, E. Heard, J. Dekker, E. Barillot, HiC-Pro: an optimized and flexible pipeline for Hi-C data processing, Genome Biol. 16 (2015) 259.

[32] G.P. Li, Y. Chen, M.P. Snyder, M.Q. Zhang, ChIA-PET2: a versatile and flexible pipeline for ChIA-PET data analysis, Nucleic Acids Res. 45 (1) (2017).

[33] M. Imakaev, G. Fudenberg, R.P. McCord, N. Naumova, A. Goloborodko, B.R. Lajoie, J. Dekker, L.A. Mirny, Iterative correction of Hi-C data reveals hallmarks of chromosome organization, Nat. Methods 9 (10) (2012) 999–1003.

[34] N.C. Durand, M.S. Shamim, I. Machol, S.S. Rao, M.H. Huntley, E.S. Lander, E.L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments, Cell Syst. 3 (1) (2016) 95–98.

[35] B.R. Lajoie, J. Dekker, N. Kaplan, The Hitchhiker's guide to Hi-C analysis: practical guidelines, Methods 72 (2015) 65–75.

[36] E.P. Nora, A. Goloborodko, A.L. Valton, J.H. Gibcus, A. Uebersohn, N. Abdennur, J. Dekker, L.A. Mirny, B.G. Bruneau, Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization, Cell 169 (5) (2017) 930.

[37] M.R. Mumbach, A.T. Satpathy, E.A. Boyle, C. Dai, B.G. Gowen, S.W. Cho, M.L. Nguyen, A.J. Rubin, J.M. Granja, K.R. Kazane, Y. Wei, T. Nguyen, P.G. Greenside, M.R. Corces, J. Tycko, D.R. Simeonov, N. Suliman, R. Li, J. Xu, R.A. Flynn, A. Kundaje, P.A. Khavari, A. Marson, J.E. Corn, T. Quertermous, W.J. Greenleaf, H.Y. Chang, Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements, Nat. Genet. 49 (11) (2017) 1602–1612.

[38] A.S. Weintraub, C.H. Li, A.V. Zamudio, A.A. Sigova, N.M. Hannett, D.S. Day, B.J. Abraham, M.A. Cohen, B. Nabet, D.L. Buckley, Y.E. Guo, D. Hnisz, R. Jaenisch, J.E. Bradner, N.S. Gray, R.A. Young, YY1 is a structural regulator of enhancer-promoter loops, Cell 171 (7) (2017) 1573–1588 e28..

[39] S. Bhatia, H. Bengani, M. Fish, A. Brown, M.T. Divizia, R. de Marco, G. Damante, R. Grainger, V. van Heyningen, D.A. Kleinjan, Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia, Am. J. Hum. Genet. 93 (6) (2013) 1126–1134.

[40] N. Rajagopal, S. Srinivasan, K. Kooshesh, Y. Guo, M.D. Edwards, B. Banerjee, T. Syed, B.J. Emons, D.K. Gifford, R.I. Sherwood, High-throughput mapping of regulatory DNA, Nat. Biotechnol. 34 (2) (2016) 167–174.

[41] A. Buckle, R.S. Nozawa, D.A. Kleinjan, N. Gilbert, Functional characteristics of novel pancreatic Pax6 regulatory elements, Hum. Mol. Genet. 27 (19) (2018) 3434–3448.

[42] S.S.P. Rao, S.C. Huang, B. Glenn St Hilaire, J.M. Engreitz, E.M. Perez, K.R. Kieffer-Kwon, A.L. Sanborn, S.E. Johnstone, G.D. Bascom, I.D. Bochkov, X. Huang, M.S. Shamim, J. Shin, D. Turner, Z. Ye, A.D. Omer, J.T. Robinson, T. Schlick, B.E. Bernstein, R. Casellas, E.S. Lander, E.L. Aiden, Cohesin loss eliminates all loop domains, Cell 171 (2) (2017) 305–320.

[43] B. Bintu, L.J. Mateo, J.H. Su, N.A. Sinnott-Armstrong, M. Parker, S. Kinrot, K. Yamaya, A.N. Boettiger, X.W. Zhuang, Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells, Science 362 (6413) (2018) p. 419-+.

[44] L.Z. Tan, D. Xing, C.H. Chang, H. Li, S. Xie, Three-dimensional genome structures of single diploid human cells, Science 361 (6405) (2018) 924–928.