Mapping 3D genome architecture through in situ DNase Hi-C

Vijay Ramani¹, Darren A Cusanovich¹, Ronald J Hause¹, Wenxiu Ma¹, Ruolan Qiu¹, Xinxiang Deng², C Anthony Blau³,⁴, Christine M Disteche²,⁵, William S Noble¹, Jay Shendure¹,⁶ & Zhijun Duan³,⁴

¹Department of Genome Sciences, University of Washington, Seattle, Washington, USA. ²Department of Pathology, University of Washington, Seattle, Washington, USA. ³Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington, USA. ⁴Division of Hematology, University of Washington, Seattle, Washington, USA. ⁵Department of Medicine, University of Washington, Seattle, Washington, USA. ⁶Howard Hughes Medical Institute, Seattle, Washington, USA.

Correspondence should be addressed to J.S. (shendure@uw.edu) or Z.D. (zjduan@uw.edu).

Published online 29 September 2016; doi:10.1038/nprot.2016.126

With the advent of massively parallel sequencing, considerable work has gone into adapting chromosome conformation capture (3C) techniques to study chromosomal architecture at a genome-wide scale. We recently demonstrated that the inactive murine X chromosome adopts a bipartite structure using a novel 3C protocol, termed in situ DNase Hi-C. Like traditional Hi-C protocols, in situ DNase Hi-C requires that chromatin be chemically cross-linked, digested, end-repaired, and proximity-ligated with a biotinylated bridge adaptor. The resulting ligation products are optionally sheared, affinity-purified via streptavidin bead immobilization, and subjected to traditional next-generation library preparation for Illumina paired-end sequencing. Importantly, in situ DNase Hi-C obviates the dependence on a restriction enzyme to digest chromatin, instead relying on the endonuclease DNase I. Libraries generated by in situ DNase Hi-C have a higher effective resolution than traditional Hi-C libraries, which makes them valuable in cases in which high sequencing depth is allowed for, or when hybrid capture technologies are expected to be used. The protocol described here, which involves ~4 d of bench work, is optimized for the study of mammalian cells, but it can be broadly applicable to any cell or tissue of interest, given experimental parameter optimization.

INTRODUCTION

The manner in which an incredibly long DNA polymer topologically organizes itself within a cell or nucleus is crucially linked to higher-order cellular function¹–². This form–function relationship, first discovered through early light microscopic studies of higher-order structures such as mitotic chromosomes³, the inactive X Barr body⁴, and polytene chromosomes⁵, has only become clearer in the face of advancing technologies. Techniques such as fluorescence in situ hybridization (FISH) of chromatin⁶–⁸ have provided clear evidence that chromosomes occupy compartments within the nucleus, ultimately leading to the development of correlative models associating biological function (i.e., transcription, splicing, silencing) with particular nuclear locales⁹–¹⁰.

With the advent of genome-scale technologies, high-throughput assays have been developed to characterize nuclear architecture at both increasing scale and increasing resolution. Techniques such as DNA adenine methyltransferase identification¹¹,¹², typically used to map protein–DNA interactions¹³–¹⁵, have been modified to map genome-wide associations between primary sequences and the nuclear lamina¹⁶ (i.e., lamina-associated domains or LADs), where silenced domains typically reside. Methods involving the proximity ligation of chromatin, now termed 3C (ref. 17), have also gained popularity. 3C techniques represent matured versions of early methods that used T4 DNA ligase to quantify the physical proximity of DNA sequences brought together by proteins¹⁸,¹⁹, and all share a common paradigm: fixation of chromatin within the nucleus via formaldehyde, endonucleolytic digestion of chromatin (normally via restriction enzyme (RE) digestion), and ligation of physically proximal fragments. The first 3C variants (e.g., 4C, 5C) used specific primers or sets of primers to determine contact frequencies between predefined sites in the genome²⁰,²¹.

Later, massively parallel versions of 3C, generally termed ‘Hi-C’, were developed²²–²⁴ that leverage paired-end sequencing to generate contact-frequency estimates between sequence windows across entire genomes.

Since the advent of 3C techniques, much work has gone into characterizing 3D genome architecture in a wide variety of biological contexts²⁵–²⁹, including mitotic cell division³⁰, the life cycle of a parasite³¹, and mammalian dosage compensation³²–³⁵. The vast amount of available Hi-C data has also enabled the discovery of novel ‘units’ of genome topology, including topologically associating domains (TADs)³³,³⁶ and chromosomal interacting domains²⁷,³⁷, genomic domains that predominantly self-associate in 3D space. Although the ultimate significance of these domains remains unknown, strong correlations between 1D epigenomic features (e.g., histone marks, DNA methylation and transcription factor binding) and sequence, both within and at the borders of these domains, suggest that they may have a gene regulatory role.

Although current Hi-C techniques generally allow us to visualize genome-scale chromosome architecture at a resolution of 100 kb to 1 Mb, methodological resolution limitations imposed by incomplete sequencing depth and genome-wide restriction site density have typically precluded the identification of topological units at smaller scales, in which local interactions may have crucial gene regulatory roles. The need for fine-scale resolution of these higher-order interactions has only become clearer in the wake of the immense amount of high-resolution, 1D epigenomic data generated by consortia such as ENCODE³⁸ and Roadmap Epigenomics³⁹.

Given the availability of such data, one crucial interest of the gene regulatory field is the potential link between complex gene...
regulatory programs and dynamic long-range ‘looping’ interactions between distal regulatory elements, features at a scale even smaller than those of TADs and LADs\(^40\). Since the earliest realizations that long-range interactions are effectors of gene expression\(^{41,42}\), the gene regulatory field has worked toward completely cataloging functional DNA looping interactions. In the realm of proximity ligation protocol development, considerable work has gone toward improving the resolution of the Hi-C protocol to the scale of kilobases, at which specific regulatory contacts (i.e., enhancer–promoter interactions, CCCTC-binding factor–mediated loops) might be identified.

The protocol presented here complements existing high-resolution Hi-C approaches\(^{37,43}\) by providing another flexible, convenient, and scalable methodology that eschews the use of restriction enzymes. Our approach therefore avoids the theoretical limit in resolution of the standard Hi-C protocol imposed by the occurrence of restriction sites in the genome, given enough sequencing depth and library complexity.

### Moving toward fine-scale resolution of 3D contacts

Core methodological improvements to the Hi-C protocol to improve resolution have broadly spanned three primary areas: deeper sequencing\(^{38}\), simplified library preparation protocols\(^{43,44}\), and the use of hybridization capture to enrich for sets of desired loci in a massively parallel manner\(^{45–47}\). We recently developed a method that unites many of these improvements with additional empirical changes to further increase the effective resolution of Hi-C libraries\(^{48}\). Our method, termed DNase Hi-C, eliminates the reliance on restriction enzymes associated with Hi-C by digesting fixed chromatin with the endonuclease DNase I in the presence of divalent manganese. We demonstrated that DNase Hi-C libraries mitigate many of the biases associated with traditional Hi-C, reducing the effective distance between fragments imposed by 4- and 6-cutter restriction enzymes while improving robustness with respect to G–C content, mappability, and genomic coverage. Furthermore, we also showed that DNase Hi-C may be paired with commercially available hybridization capture kits to visualize long intergenic noncoding RNA promoters at a previously unprecedented scale of 1 kb without the gross sequencing depth requirements typically associated with high-resolution contact maps.

Motivated by the observation that the vast majority of proximity ligations occur in insoluble chromatin\(^49\), and the consequent improvements to traditional RE Hi-C based on this fact\(^{43,44,50}\), we recently published an improved version of our previously published DNase Hi-C termed in situ DNase Hi-C\(^{51}\). We applied this simplified and robust Hi-C protocol to study the inactive X chromosome in primary mouse brain tissue and an immortalized mouse embryonic kidney cell line, demonstrating for the first time that the murine inactive X chromosome adopts a bipartite conformation. In situ DNase Hi-C represents a considerable improvement over its parent protocol, requiring considerably less hands-on time and lower cellular input requirements\(^{51}\).

### Overview of in situ DNase Hi-C

A schematic of the in situ DNase Hi-C protocol is shown in Figure 1. Anywhere from 5 × 10^5 to 1 × 10^7 cells are fixated in formaldehyde to reversibly cross-link in vivo protein–DNA interactions. Fixed cells are then lysed to liberate nuclei, which are treated with the endonuclease DNase I to digest chromatin. Digested chromatin ends are end-repaired and dATP-tailed, facilitating the ligation of an exogenous, dTTP-tailed ‘bridge’ adaptor containing a single biotinylated thymidine, a half BamHI restriction site, and a four-base overhang. After clearing out excess adaptors, the free ends of the chromatin (now capped with bridge adaptors) are phosphorylated and in situ proximity ligation, after which DNA is purified and fragments containing ligation junctions are enriched via streptavidin beads for on-bead Illumina library preparation (optionally following sonication).

### After proximity ligation

Nuclei are lysed and cross-links are reversed with proteinase K treatment. DNA is then isolated with an isopropanol precipitation, after which fragments containing ligation junctions are amplified with PCR to generate sequencing libraries. Before sequencing, libraries may be treated with a simple BamHI digestion to assess the efficiency of proximity ligation.

### Traditional Hi-C versus in situ DNase Hi-C

In situ DNase Hi-C can be used in any situation in which traditional Hi-C would be used. Thanks to a reliance on the...
endonuclease DNase I, in situ DNase Hi-C eliminates the characteristic restriction enzyme biases that limit resolution in traditional Hi-C libraries while lowering the input cell requirements for library construction. Unlike other Hi-C protocols, in situ DNase Hi-C is the only protocol, to our knowledge, to use paramagnetic carboxylated beads as a tool to immobilize nuclei during in situ enzymatic treatments. This immobilization step not only reduces nuclei loss during the protocol, aiding low-input experiments, but also facilitates the removal of contaminating adaptors and free DNA. Finally, like traditional in situ Hi-C, in situ DNase Hi-C requires considerably less hands-on time for library prep, and it more efficiently generates cis (i.e., intrachromosomal) ligation products as compared with trans (i.e., interchromosomal) ligation products.

Considering the high sequencing depth required to generate high-resolution genome-wide contact maps, we note that, at low resolution, maps generated using in situ DNase Hi-C are practically very similar to those generated using other Hi-C protocols (except in cases in which loci may have particularly low resolution site density). However, in cases in which high-resolution (i.e., 1-kb resolution) maps are desired, we strongly believe that the relatively unbiased ligation junctions generated through DNase Hi-C present an important alternative to existing methods. This point is particularly relevant when hybrid capture techniques may be applied, as high-resolution, RE-independent maps can be generated for a fraction of the cost of genome-scale library sequencing.

Still, we acknowledge that in many cases cost may preclude the use of deep sequencing or hybrid capture. In cases such as these, we suggest more cost-effective solutions using more focused techniques (e.g., 3C, 4C and 5C), albeit at the price of interrogating interactions among only a set number of loci.

In situ DNase Hi-C is broadly applicable to any situation in which high-resolution chromatin conformation data or 3D maps are required. We have successfully carried out in situ DNase Hi-C in several immortalized cell lines and primary tissues, generating libraries for the human cell lines K562 and GM12878, as well as for mouse embryonic kidney cells and homogenized mouse brain tissue.

Limitations of the protocol

In situ DNase Hi-C is subject to the same limitations as any bulk Hi-C protocol. First, the protocol requires 5 × 10⁵ to 1 × 10⁷ cells to generate sequenceable libraries. Thus, in cases in which input might be particularly limited, or in which small populations of cells are sorted by FACS, in situ DNase Hi-C may not be appropriate. Second, it is also important to note that although the DNase enzyme is nonspecific when compared with restriction enzymes, it has been shown to exhibit mild sequence bias at cleavage sites. This must be considered when applying in situ DNase Hi-C to organisms with radical nucleotide content (i.e., low GC content) and when considering the inherent biases within in situ DNase Hi-C maps (as would be done with any Hi-C contact map).

Experimental design

The in situ DNase Hi-C protocol described here is relatively straightforward, and it can be completed over 4 d, allotting 3–6 h of bench work per day. Still, there are several experimental design parameters that should be considered before applying in situ DNase Hi-C to a new cell type of interest. These considerations primarily concern maintaining intact nuclei during the various in situ enzymatic treatments in the protocol. The in situ DNase Hi-C protocol also allows for sequencing-free quality control (QC) of libraries, due to the integration of half BamHI sites in the bridge adaptor. As discussed below, this allows for easy quantification of the efficiency of proximity ligation in the final in situ DNase Hi-C library.

Although the protocol presented here is robust to many different cell types, different immortalized cell lines may require optimization of formaldehyde cross-linking, DNase I digestion, and SDS concentration during digestion. Below we detail our process for optimizing these various parameters.

Formaldehyde concentration

As with other 3C methods and ChIP-seq protocols, formaldehyde fixation is an important component of the in situ DNase Hi-C protocol, promoting proximity ligation of long-range genomic contacts while maintaining the integrity of nuclei during in situ enzymatic steps. Incomplete cross-linking can lead to an under-representation of proximity ligation products in Hi-C libraries, and excessive breakage of nuclei can lead to considerable decreases in the ultimate molecular complexity of libraries and, at worst, can increase the degree of ‘spurious’ ligation formed. The guidelines for formaldehyde fixation of cells for in situ DNase Hi-C are the same as those for the other 3C-based techniques and ChIP-seq methods. In general, for single-cell suspension cultures (e.g., GM12878 and K562 cells) and monolayer adherent cells (e.g., HeLa cells), a standard condition of cross-linking, such as 1% (vol/vol) formaldehyde for 10 min at room temperature (RT, 25 °C), can be used. For other cell cultures (e.g., mouse and human embryonic stem cells (ESCs)) and primary-tissue cells (e.g., mouse brain cells and plant leaves), for which single-cell suspensions are difficult to obtain, increased formaldehyde concentrations or longer fixation times may be required to ensure efficient cross-linking. For example, both human and mouse ESCs often aggregate to form

### Box 1 | Assessment of nuclear lysis at various steps

**TIMING** ~3 d

To ascertain whether nuclei remain intact during the protocol, perform the following control experiment. After each enzymatic treatment step (Steps 15, 24, 29, 35 and 49), remove the supernatant and add 10 µl of proteinase K to it. Treat the supernatant overnight at 65 °C; then precipitate the DNA by adding 0.1 volumes of 3M sodium acetate, 3 µl of GlycoBlue, and 1 volume of 100% isopropanol, mixing, and then incubating the mixture at −80 °C for 1 h. Pellet the mixture at 4 °C at 16,000g for 30 min, carefully remove the supernatant, and resuspend the pellet in 100 µl of ddH₂O. Add 10 µl of RNase A to each sample, incubate at 37 °C for 10 min, and then purify the DNA using 1.2 volumes of AMPure XP beads. Resuspend the beads in 15 µl of ddH₂O, and run this out on a 6% TBE gel.
large clumps in culture. Higher concentrations of formaldehyde are generally used in these situations.48,54.

**Cell lysis and DNase I digestion.** After cross-linking chromatin interactions with formaldehyde, one must render fixed chromatin accessible to enable chromatin fragmentation and other downstream enzymatic reactions. As with restriction–digestion-based 3C methods, cell lysis in *in situ* DNase Hi-C is achieved primarily through SDS treatment. To ensure that nuclei remain intact throughout the multiple enzymatic reactions through the end of nuclear ligation (Step 48), the *in situ* DNase Hi-C protocol uses a relatively mild condition (0.3–0.5% (wt/vol) SDS treatment for 45 min at 37 °C). During this step, it is crucial to avoid overly lysing nuclei. A simple experiment to determine the extent of nuclear lysis is detailed in Box 1, with expected results shown in Figure 2a. We also note that overly lysed nuclei become apparent during any of the many centrifugation steps in the *in situ* DNase Hi-C protocol, as no pellet forms. Nuclei should remain intact through proximity ligation, as shown in Figure 2b.

We stress that the required SDS concentration for cell lysis and the amount of DNase I used during the DNase I digestion step can vary depending on the cell type being studied and the number of nuclei being processed. When attempting the protocol on new cell types, we recommend carrying out a DNase I and SDS optimization experiment using varying amounts of DNase I and varying concentrations of SDS, and then examining the DNase I fragmentation pattern following digestion. An example fragmentation pattern is shown in Figure 3a.

The role of paramagnetic carboxylated beads. Paramagnetic carboxylated beads (i.e., AMPure XP beads) have been used in both our standard and *in situ* DNase Hi-C protocols. As demonstrated in Figure 2, these beads appear to bind to intact nuclei and serve as carriers to pellet the nuclei by low-speed centrifugation. Here, we use these beads to efficiently remove DNase I and low-molecular-weight DNA that might escape the nucleus following chromatin digestion, as well as free unligated internal bridge adaptor following bridge adaptor ligation. Furthermore, the beads also aid with visualization of the nuclei pellet throughout the protocol when starting the protocol with fewer than a million cells.

**Nuclei treatment.** It is crucial that the fixed nuclei remain intact over the course of the DNase Hi-C protocol. To this end, pipetting should be carried out gently to minimize shear forces that may burst nuclei.

**BamHI digestion control.** A BamHI digestion test on the final PCR-amplified library can be used to quantify ligation efficiency of the reaction. Lack of a library ‘shift’ (properly digested products shown in Figure 3b) suggests inefficiency in the formation of proximity ligation products and can be indicative of suboptimal fixation conditions or defective reagents.
**MATERIALS**

**REAGENTS**

- Cell lines of interest (adherent, suspension or primary tissue): for example, we have used the human cell line GM12878 (Coriell GM12878) and the Patski cell line (provided by C. Distefo) in our previous study.\(^1\) **CAUTION**
  Cell lines should be regularly checked to ensure that they are authentic and that they are not infected with *Mycoplasma*.

- Penicillin–Streptomycin (Thermo Fisher Scientific, cat. no. 15140122)
- FBS (Thermo Fisher Scientific, cat. no. 10437-010)
- RPMI 1640 (Thermo Fisher Scientific, cat. no. 11875-093)
- DMEM (Thermo Fisher Scientific, cat. no. 11965118)
- Biotinylated bridge adaptor 5′/5′Phos/GCTGAGGA/GA/BiodT/C (Integrated DNA Technologies)
- Bridge adaptor 3′: CCTCAGCT (Integrated DNA Technologies)
- Bridge adaptor 5′: GCTGAGGAGAC (Integrated DNA Technologies)
- Blunt bridge adaptor 3′: CCTCAGC (Integrated DNA Technologies)
- SeqAdapt_F: AGACTCTTCTTCCCTACACGACGCCCTTCGGACGCCTC'T (Integrated DNA Technologies)
- SeqAdapt_R: 5′Phos/GCTGAGGAGA/BioTGCACTGACCTGAACGTCA (Integrated DNA Technologies)
- QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
- Protease Inhibitor Tablets (Roche, cat. no. 04693116001)
- 1× DPBS (Life Technologies, cat. no. 14190-250)
- Triton X-100 (Sigma-Aldrich, cat. no. X100-5ML)

**CAUTION**

Formaldehyde is flammable, can cause skin burns, and is toxic if inhaled. Formaldehyde should be handled in a chemical fume hood while using appropriate protective equipment.

**CRITICAL**

Formaldehyde has a limited shelf life; discard the solution if it is older than 1 year.

- Glycine (Sigma-Aldrich, cat. no. 50046)
- NEBuffer 2 (NEB B7002S)
- 10% (wt/vol) UltraPure SDS (Life Technologies, cat. no. 15553-027)
- DNsase I, RNase-free (supplied with MnCl₂ and 10× DNase I digestion buffer; 1 U/µl; Thermo Fisher Scientific, cat. no. EN0525)
- RNase A, DNase and protease-free (10 mg/ml; Thermo Fisher Scientific, cat. no. EK0032)
- Klenow fragment (10 U/µl; Thermo Fisher Scientific, cat. no. EP0052)
- Klenow fragment (exo-) (5 U/µl; Thermo Fisher Scientific, cat. no. EP0422)
- T4 DNA polymerase (5 U/µl; Thermo Fisher Scientific, cat. no. EP0062)
- T4 DNA Ligase (5 U/µl) provided with 50% PEG-4000 (Thermo Fisher Scientific, cat. no. EN0531)
- 10× DNase digestion buffer with 2% (vol/vol) Triton X-100 and RNase-free (10 mg/ml; Thermo Fisher Scientific, cat. no. EN0531)
- 20% (vol/vol) Ethanol
- 6× Stop solution
- 1× B&W buffer
- 0.5× DNase digestion buffer
- 8× Stop solution
- 1× Tube wash buffer
- 80% (vol/vol) Ethanol
- 0.5× B&W buffer
- 6× Stop solution
- 1× Tube wash buffer
- 0.5× DNase digestion buffer
- 8× Stop solution
- 1× Tube wash buffer

**REAGENT SETUP**

**Cell culture medium** Dilute RPMI 1640 with 15% (vol/vol) FBS (for GM12878 cells) or dilute DMEM with 10% (vol/vol) FBS (for Patski cells).

- 2.5 M Glycine Adjust the volume of 9.33 g of glycine to 50 ml with ddH₂O, and filter-sterilize the solution using a Steriflip filter. Store the solution at RT for up to 6 months.

- 10 mM dATP Mix 20 µl of 100 mM dATP from the dNTP Set with 80 µl of ddH₂O.

- 80% (vol/vol) Ethanol Mix 8 ml of 100% (vol/vol) ethanol with 2 ml of ddH₂O. Freshly prepare 80% (vol/vol) ethanol on the day of the experiment.

- Triton X-100 Mix 1 ml of 100% (vol/vol) Triton X-100 with 9 ml of ddH₂O. Store the solution at RT for up to 6 months.

- Cell lysis buffer Mix 500 µl of 1M Tris-HCl, pH 8.0, 100 µl of 5M NaCl, and 1 ml of 10% IGEpal CA-630 and bring the final volume to 50 ml with ddH₂O. Store the buffer at 4 °C for up to 6 months.

- TE lysis buffer Mix 2.5 ml of 1M Tris-HCl, pH 7.0, 100 µl of 0.5 M EDTA, and 5 ml of 10% (wt/vol) SDS and bring the final volume to 50 ml with ddH₂O. Store the buffer at RT for up to 6 months.

- B&W buffer Mix 500 µl of 1M Tris-HCl, pH 8.0, 100 µl of 0.5 M EDTA, and 20 ml of 5 M NaCl and bring the final volume to 50 ml with ddH₂O. Store the buffer at RT for up to 6 months.

**Software**

- Python 2.7 (http://www.python.org/)
- FastQC version 0.11.3 or higher (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)
- SAMtools version 0.1.18 or higher (http://samtools.sourceforge.net/)
- hiclib (library (http://mirnylab.cshl.edu/hiclib/)
- matplotlib library (http://matplotlib.org/)

**EQUIPMENT**

- Water bath (set to 60 °C, VWR 89501-464 or similar)
- Thermocycler (Eppendorf 6321000019 or similar)
- DynaMag Magnetic Rack (Life Technologies, cat. no. 12321D)
- DynaMag Spin Magnet (Thermo Fisher Scientific, cat. no. 12320D)
- Qubit Fluorometer (Life Technologies, cat. no. Q33216)
- 0.2-ml PCR tubes (Fisher, cat. no. 14-320-212)
- 1.5-ml Microcentrifuge tubes (Fisher, cat. no. 50480129)
- 6% TBE-PAGE gels (Life Technologies, cat. no. EC6265BOX)
- Cell scrapper (Fisher, cat. no. 08-100-241)
- 50-ml Tubef (Fisher, cat. no. 14-432-22)
- Cell culture plates (Sigma-Aldrich, cat. no. CLS430167-100EA)
- Microcentrifuge
- MilliPore Sterilip Filters (Millipore, cat. no. SCGP00525)
- Covaris Sonicator (Covaris, S220)
- Covaris Microtube (Covaris, cat. no. 520045)
- A computer running Unix/Linux distribution

**pnp**
PROCEDURE

Adaptor annealing and cross-linking of cells ● TIMING 3 h plus overnight incubation

1. Set up the reactions shown in Tables 1–3.

2. Anneal the mixtures by heating them to 98 °C for 6 min, and then allow the tubes to naturally cool to RT overnight. ■ PAUSE POINT Annealed adaptors can be kept at −20 °C indefinitely.

3. Cells should be grown in appropriate culture medium. 2–5 × 10^6 cells are sufficient for making one DNase Hi-C library. However, we suggest growing, cross-linking, and aliquotting many cells (i.e., 1–5 × 10^7 cells) to provide replicates, if necessary. Below are protocols for handling adherent monolayer cells (option A) or suspension cells (option B):

(A) Cross-linking of adherent monolayer cells
   (i) Aspirate out the medium and add 10 ml of serum-free medium per 10-cm plate.
   (ii) Cross-link the cells by adding 280 μl of 37% (vol/vol) formaldehyde to obtain a 1% final concentration. Mix gently, immediately after the addition of formaldehyde.
   (iii) Incubate the cells at RT for exactly 10 min, gently rocking the plates every 2 min.
   (iv) Quench the reaction by adding 0.5 ml of 2.5 M glycine and mixing the solution well.
   (v) Incubate the cells for 5 min at RT, and then on ice for 15 min to stop cross-linking completely.
   (vi) Wash the cells once with cold 1× PBS.

---

**TABLE 1** | Blunt bridge adaptor (40 μM final).

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 μM bridge adaptor 5’</td>
<td>80</td>
<td>40 μM</td>
</tr>
<tr>
<td>100 μM blunt bridge adaptor 3’</td>
<td>80</td>
<td>40 μM</td>
</tr>
<tr>
<td>10× NEBuffer 2</td>
<td>20</td>
<td>1×</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Total volume</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 2** | Biotinylated bridge adaptor (40 μM final).

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 μM biotinylated bridge adaptor 5’</td>
<td>80</td>
<td>40 μM</td>
</tr>
<tr>
<td>100 μM bridge adaptor 3’T</td>
<td>80</td>
<td>40 μM</td>
</tr>
<tr>
<td>10× NEBuffer 2</td>
<td>20</td>
<td>1×</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Total volume</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 3** | Y adaptor (25 μM final).

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 μM SeqAdapt_F</td>
<td>50</td>
<td>25 μM</td>
</tr>
<tr>
<td>100 μM SeqAdapt_R</td>
<td>50</td>
<td>25 μM</td>
</tr>
<tr>
<td>10× NEBuffer 2</td>
<td>20</td>
<td>1×</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Total volume</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>
(vii) Treat the cells with 3–5 ml of 0.25% trypsin per dish at 37 °C for 5 min.
(viii) Add 5 ml of fresh medium with serum.
(ix) Scrape the cells from the plates with a cell scraper and transfer them to a 50-ml tube (combine all the cells from all the dishes in one tube).
(x) Centrifuge the cross-linked cells at 800 g for 10 min at RT.
(xi) Discard the supernatant by aspiration, and wash the cross-linked cells with 1 × PBS once.
(xii) Divide the cells into aliquots in 1.5-ml microtubes (2.5 million cells per tube).

■ PAUSE POINT The cells can be snap-frozen in liquid nitrogen and stored for at least one year at −80 °C, or one can continue with cell lysis.

(B) Cross-linking of suspension cells

(i) Gently pellet the cells by spinning them at 300 g for 10 min at RT.
(ii) Discard the supernatant.
(iii) Resuspend the pellet in 10 ml of fresh culture medium without serum. Break the cell clumps by pipetting up and down.
(iv) Cross-link the cells by adding 280 µl of 37% (vol/vol) formaldehyde (1% final concentration). Mix quickly by inverting the tube several times.
(v) Incubate the cells at RT for exactly 10 min. Gently invert the tube every 2 min.
(vi) Add 0.5 ml of 2.5 M glycine to quench the cross-linking reaction, and mix well.
(vii) Incubate the mixture for 5 min at RT and then on ice for 15 min to stop cross-linking completely.
(viii) Centrifuge the cross-linked cells at 800 g for 10 min at 4 °C.
(ix) Aspirate and discard the supernatant, and wash the cross-linked cells with 1× PBS once.
(x) Split the cross-linked cell suspension into aliquots of 2.5 × 10^6 cells (in 1.5-ml microtubes).

■ PAUSE POINT Cells can be snap-frozen in liquid nitrogen and stored for up to 1.5 years at −80 °C, or one can continue with cell lysis.

Cell lysis and chromatin digestion with DNase I ● TIMING 1.5 h

4 | Resuspend one cross-linked cell aliquot (0.5–2.5 × 10^6 cells) in 0.4 ml of ice-cold cell lysis buffer containing protease inhibitor. ▲ CRITICAL STEP Add 1 tablet of protease inhibitor to 10 ml of ice-cold lysis buffer immediately before lysis. We recommend using lysis buffer with freshly added protease inhibitor for all experiments.

5 | Incubate the mixture on ice for 10 min.

6 | Centrifuge the mixture for 60 s at 2,500g at RT.

7 | Discard the supernatant and resuspend the pellet in 100 µl of 0.5× DNase I digestion buffer containing 0.2% (wt/vol) SDS. ▲ CRITICAL STEP For larger cell inputs (i.e., 3–5 × 10^6), we recommend using 200 µl of 0.5× DNase I digestion buffer instead.

8 | Incubate the mixture at 37 °C for 30 min.

9 | Add 100 µl of 0.5× DNase I digestion buffer containing 2% (vol/vol) Triton X-100 and 4 µl of RNase A, and mix the solution well.

10 | Incubate the mixture at 37 °C for 10 min.

11 | Add 1.5 units of DNase I and mix well.

12 | Incubate the mixture at RT for 4 min.

13 | Add 40 µl of 6× stop solution, and mix well.

14 | (Optional) To determine the efficacy of DNase I digestion, take 20 µl of lysed cells from the previous step and add them to a new tube. Add 70 µl of 1× TE lysis buffer and 10 µl of proteinase K (20 mg/ml). Incubate for 30 min at 65 °C. Purify the DNA using a Qiagen PCR purification kit. Check the quality of chromatin digestion by running the samples out on a 6% TBE-PAGE gel. The sample is properly digested if one sees a large smear of DNA fragments between ~100 bp and 1 kb (Fig. 3a). We recommend characterizing DNase I digestion efficiency when performing the protocol on a new cell type.
In the event of overdigestion or underdigestion of chromatin, we recommend optimizing the concentration of SDS in the digestion reaction, the amount of DNase I used or the digestion time.

**Troubleshooting**

15| Centrifuge the cells for 60 s at 2,500g at RT.

16| Discard the supernatant and resuspend the pellet in 150 µl of water.

**Troubleshooting**

17| Add 300 µl of AMPure XP beads; mix them thoroughly by pipetting up and down.

18| Incubate the mixture at RT for 5 min and place the tube in a DynaMag magnet for 2 min.

19| Discard the supernatant and wash the beads twice with 1 ml of freshly prepared 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol.

20| Resuspend the beads in 169 µl of water, and proceed immediately to the next step.

**Chromatin end repair and dA-tailing**

**TIMING 2.5 h**

21| Prepare the end-repair reaction as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei w/beads</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>10× T4 ligase buffer with ATP</td>
<td>20</td>
<td>1×</td>
</tr>
<tr>
<td>10 mM dNTPs</td>
<td>5</td>
<td>0.25 mM</td>
</tr>
<tr>
<td>T4 DNA Polymerase (3 U/µl)</td>
<td>3</td>
<td>0.045 U/µl</td>
</tr>
<tr>
<td>Klenow fragment (10 U/µl)</td>
<td>3</td>
<td>0.15 U/µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

22| Incubate the mixture at RT for 1 h.

23| Add 5 µl of 10% (wt/vol) SDS to stop the reaction.

24| Centrifuge the mixture for 60 s at 2,500g at RT.

25| Aspirate and resuspend the pellet in 135 µl of water.

26| Prepare the dA-Tailing reaction as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei with beads</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>10× NEBuffer 2</td>
<td>20</td>
<td>1×</td>
</tr>
<tr>
<td>10 mM dATP</td>
<td>10</td>
<td>0.5 mM</td>
</tr>
<tr>
<td>10% Triton X-100</td>
<td>20</td>
<td>1%</td>
</tr>
<tr>
<td>Klenow fragment (exo−) (5 U/µl)</td>
<td>15</td>
<td>0.375 U/µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

27| Incubate the resulting mixture at 37 °C for 1 h.

28| Add 5 µl of 10% (wt/vol) SDS to stop the reaction.

29| Centrifuge the mixture for 60 s at 2,500g at RT.
30| Aspirate and resuspend the pellet in 30 µl of nuclease-free water.

**Ligation of biotin-labeled bridge adaptors** ● **TIMING** overnight, followed by 0.5 h

31| Prepare the adaptor ligation reaction as follows, using the annealed adaptors from Step 2:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei with beads</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Annealed bridge adaptor with biotin (40 µM)</td>
<td>20</td>
<td>8 µM</td>
</tr>
<tr>
<td>Annealed blunt adaptor w/o biotin (40 µM)</td>
<td>20</td>
<td>8 µM</td>
</tr>
<tr>
<td>10× T4 ligase buffer with ATP</td>
<td>10</td>
<td>1×</td>
</tr>
<tr>
<td>PEG-4000 (50%)</td>
<td>10</td>
<td>5%</td>
</tr>
<tr>
<td>10% Triton X-100</td>
<td>5</td>
<td>0.5%</td>
</tr>
<tr>
<td>T4 DNA Ligase (5 U/µl)</td>
<td>5</td>
<td>0.25 U/µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

32| Incubate the mixture at 16 °C overnight.

33| (Optional) To examine the efficacy of the above end-repair, dA-tailing, and adaptor ligation reactions, take 3 µl of nuclei from Step 30 to perform a control ligation reaction with the Illumina Y adaptor from Step 2 as given below:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei with beads</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Illumina Y adaptor (50 µM)</td>
<td>1</td>
<td>2.5 µM</td>
</tr>
<tr>
<td>Water</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>10× T4 ligase buffer with ATP</td>
<td>2</td>
<td>1×</td>
</tr>
<tr>
<td>PEG-4000 (50%)</td>
<td>2</td>
<td>5%</td>
</tr>
<tr>
<td>10% Triton X-100</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td>T4 DNA Ligase (5 U/µl)</td>
<td>1</td>
<td>0.25 U/µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

After incubation at 16 °C overnight, add 70 µl of 1× TE lysis buffer and 10 µl of proteinase K (20 mg/ml). Incubate the mixture for 30–60 min at 65 °C. Purify genomic DNA using a QiaQuick PCR purification kit. Check the ligation efficiency by carrying out qPCR with Illumina PCR primers. If upstream end-repair and dA-tailing steps are efficient, one should see amplification before 10 PCR cycles using 10 ng of genomic DNA as a template. We recommend this QC step when performing the protocol on a new cell type. In the event of inefficiency of these steps, we recommend optimizing the concentration of SDS in the cell lysis step or optimizing the amount of DNase I used for digestion.

34| Add 5 µl of 10% (wt/vol) SDS to stop the reaction.

35| Centrifuge the mixture for 60 s at 2,500g at RT.

36| Resuspend the pellet in 200 µl of nuclease-free water.

37| Add 165 µl of AMPure buffer; mix thoroughly by pipetting up and down.

38| Incubate the mixture at RT for 5 min, and place the tube in a DynaMag magnet for 2 min.

39| Discard the supernatant and wash the beads once with 1 ml of freshly prepared 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol.

▲ **CRITICAL STEP** We recommend diluting fresh 80% (vol/vol) ethanol before each experiment.
40| Resuspend the pellet in 200 µl of water.

41| Add 165 µl of AMPure bead buffer; mix the solution thoroughly by pipetting up and down.

42| Incubate the mixture at RT for 5 min, and then place the tube in a DynaMag magnet for 2 min.

43| Discard the supernatant and wash the beads twice with 500 µl of 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol as completely as possible, and then air-dry the beads for no more than 2 min.

44| Resuspend the nuclei–bead mixture in 80 µl of nuclease-free water.

**In situ phosphorylation ● TIMING 1.25 h**

45| Prepare the PNK reaction as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei with beads</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>10× T4 ligase buffer w/ATP</td>
<td>10</td>
<td>1×</td>
</tr>
<tr>
<td>PNK (10 U/µl)</td>
<td>10</td>
<td>1 U/µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

46| Incubate the mixture at 37 °C for 1 h.

**In situ ligation ● TIMING 4.25 h**

47| Add the following to the PNK mixture after incubation in Step 46:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>794</td>
<td></td>
</tr>
<tr>
<td>10× T4 ligase buffer</td>
<td>100</td>
<td>1×</td>
</tr>
<tr>
<td>T4 DNA Ligase (5 U/µl)</td>
<td>6</td>
<td>0.03 U/µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>1 ml</td>
<td></td>
</tr>
</tbody>
</table>

48| Incubate the mixture at RT for 4 hr. For a micrograph of nuclei after this stage, see Figure 2b.

**Cross-linking reversal, isopropanol precipitation, and DNA purification ● TIMING overnight, followed by 2.5 h**

49| Centrifuge the mixture for 60 s at 2,500g at RT.

50| Resuspend the pellet in 400 µl of 1× NEBuffer 2.

51| Add 40 µl of 10% (wt/vol) SDS.

52| Add 40 µl of 20 mg/ml proteinase K.

53| Incubate the mixture overnight at 60 °C.

54| Add 3 µl of GlycoBlue, 50 µl of 3 M sodium acetate, pH 5.2, and 550 µl of isopropanol.

55| Incubate the mixture at −80 °C for 2 h.

56| Centrifuge the mixture for 30 min at 4 °C at maximum speed in a microcentrifuge.

57| Resuspend the DNA pellets in each tube with 100 µl of nuclease-free water.

58| Add 100 µl of AMPure XP beads, and mix well.
59| Incubate the mixture at RT for 5 min, and place the tube in a DynaMag magnet for 2 min.

60| Discard the supernatant and wash the beads twice with 1 ml of 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol as completely as possible, and then air-dry the beads for no more than 2 min.

61| Resuspend the beads in 130 µl of nuclease-free water.

62| Incubate the beads at RT for 1 min. Collect the beads via DynaMag magnet, and transfer the eluent to a fresh 1.5-ml tube. At this point, determine the concentration of the recovered DNA with a spectrophotometer. A typical yield is 3–5 µg if one is starting with 2.5 × 10⁶ cells.


PAUSE POINT

Purified DNA can be stored indefinitely at −20 °C.

DNA sonication ● TIMING 0.5 h

CRITICAL At this point, purified DNA may be sonicated to shear large fragments to the 100- to 500-bp range or one may proceed directly to sequencing library preparation. Sonication promotes a less biased representation of fragment ends at the cost of additional preparation time and loss of material. The protocol here is suitable for Covaris sonicators. If sonication is not desired, skip to Step 66.

63| Transfer the DNA to a Covaris microtube.

64| Shear the DNA to a size of 100–500 bp using a sonicator. For a Covaris instrument, use the following parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duty cycle</td>
<td>15%</td>
</tr>
<tr>
<td>Peak incident power</td>
<td>450</td>
</tr>
<tr>
<td>Cycles per burst</td>
<td>200</td>
</tr>
<tr>
<td>Set mode</td>
<td>Frequency sweeping</td>
</tr>
<tr>
<td>Continuous degassing</td>
<td></td>
</tr>
<tr>
<td>Process time</td>
<td>80 s</td>
</tr>
<tr>
<td>Number of cycles</td>
<td>5</td>
</tr>
</tbody>
</table>

65| Transfer 130 µl of sonicated DNA to a 1.5-ml tube.

PAUSE POINT Eluted DNA may be stored indefinitely at −20 °C.

Biotin pull-down ● TIMING 0.5 h

66| Wash 30 µl of MyOne C1 beads twice with 100 µl of 1× B&W buffer, wash once with 100 µl of 2× B&W buffer, and then resuspend in 100 µl of 2× B&W buffer.

67| Add 100 µl of eluted DNA to resuspended streptavidin beads and mix well.

68| Incubate the sample for 20 min at RT on a rotator.

69| Place the tube in DynaMag magnet for 1 min, and discard the supernatant.

70| Wash the beads once with 300 µl of 0.5× TE lysis buffer plus 300 µl of 0.5× B&W buffer.

71| Wash the beads twice with 600 µl of 1× B&W buffer.

72| Wash the beads once with 600 µl of 1× NEBuffer 2.

73| Wash the beads once with 600 µl of Buffer EB.

74| Resuspend the beads in 170 µl of Buffer EB.

PAUSE POINT Resuspended beads may be stored at −20 °C indefinitely or at 4 °C for short-term storage.
End repair and dA-tailing ● TIMING 1.5 h

75| Set up the end-repair reaction with the Fast DNA End Repair Kit as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purified DNA</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>10x Reaction buffer</td>
<td>20</td>
<td>1×</td>
</tr>
<tr>
<td>End-repair enzyme mix</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td><strong>Total volume</strong></td>
<td><strong>200</strong></td>
<td></td>
</tr>
</tbody>
</table>

76| Incubate the mixture at 18 °C for 10 min.

77| Add 200 µl of AMPure buffer, and mix it thoroughly by pipetting up and down.

78| Incubate the mixture at RT for 5 min, and place the tube in a DynaMag. Spin magnet at 500g for 10s.

79| Discard the supernatant and wash the beads twice with 500 µl of 80% (vol/vol) ethanol. Briefly spin down the beads, remove the residual ethanol as completely as possible, and air-dry the beads for 5 min.

80| Resuspend the beads in 21.5 µl of water.

81| Set up the dA-tailing reaction as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-repaired DNA with beads</td>
<td>21.5</td>
<td></td>
</tr>
<tr>
<td>10× NEBuffer 2</td>
<td>3</td>
<td>1×</td>
</tr>
<tr>
<td>10 mM dATP</td>
<td>3</td>
<td>1 mM</td>
</tr>
<tr>
<td>Klenow fragment (exo⁻) (5 U/µl)</td>
<td>2.5</td>
<td>0.42 U/µl</td>
</tr>
<tr>
<td><strong>Total volume</strong></td>
<td><strong>30</strong></td>
<td></td>
</tr>
</tbody>
</table>

82| Incubate the mixture at 37 °C for 30 min.

83| Wash the beads twice with 400 µl of 1× B&W buffer.

84| Wash the beads twice with 400 µl of Buffer EB and resuspend them in 30 µl of Buffer EB. ▲ CRITICAL STEP Proceed immediately to adaptor ligation.

Ligation of sequencing adaptors ● TIMING 1 h

85| Immediately resuspend the beads in the following reaction mixture:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>dA-tailed DNA with beads</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>5× Rapid Ligation Buffer (from Rapid DNA Ligation Kit)</td>
<td>10</td>
<td>1×</td>
</tr>
<tr>
<td>Y adapter (2.5 µM)</td>
<td>6</td>
<td>0.3 µM</td>
</tr>
<tr>
<td>T4 DNA Ligase (5 U/µl)</td>
<td>4</td>
<td>0.4 U/µl</td>
</tr>
<tr>
<td><strong>Total volume</strong></td>
<td><strong>50</strong></td>
<td></td>
</tr>
</tbody>
</table>

86| Incubate the mixture at RT for 30 min.

■ PAUSE POINT The ligation reaction in Step 86 can also be performed at 16 °C overnight.

87| Add 5 µl of 0.5 M EDTA to stop the reaction. Add 145 µl of ddH₂O to bring the volume to 200 µl, and mix thoroughly by pipetting up and down.
Add 200 µl of AMPure buffer to each tube and mix thoroughly by pipetting up and down.

Incubate the mixture at RT for 5 min and then place the tubes in a DynaMag magnet for 2 min.

Discard the supernatant and wash the beads twice with 500 µl of 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol as completely as possible, and then air-dry the beads for no more than 2 min.

Resuspend the beads in 200 µl of ddH2O and add 165 µl of AMPure buffer.

Mix thoroughly by pipetting up and down.

Incubate the mixture at RT for 5 min, and place the three tubes in a DynaMag magnet for 2 min.

Discard the supernatant and wash the beads twice with 0.5 ml of 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol as completely as possible, and then air-dry the beads for 5 min.

Resuspend the beads in each tube with 50 µl of EB.

Library amplification ● TIMING 2.5 h

Optimization of input amount and PCR cycle number is integral to obtaining a sufficiently diverse in situ DNase Hi-C library. We recommend running several ‘pilot’ PCR reactions with various bead input amounts and various cycle numbers and running these ‘pilot’ libraries on a 6% TBE-PAGE gel to ensure that library overamplification is not occurring.

To determine the number of PCR cycles that are necessary to generate ample PCR products for sequencing—importantly, without overamplification—set up trial PCR reactions with 10, 12, or 14 cycles, and 2.5 or 5 µl of DNA-bound streptavidin beads as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-repaired DNA with beads</td>
<td>2.5/5</td>
</tr>
<tr>
<td>2× HotStart ReadyMix</td>
<td>10</td>
</tr>
<tr>
<td>10 µM SeqPrimer_F</td>
<td>1</td>
</tr>
<tr>
<td>10 µM SeqPrimer_R</td>
<td>1</td>
</tr>
<tr>
<td>ddH2O</td>
<td>Up to 20</td>
</tr>
<tr>
<td>Total volume</td>
<td>20</td>
</tr>
</tbody>
</table>

Use the following PCR program:

<table>
<thead>
<tr>
<th>Cycle number</th>
<th>Denature</th>
<th>Anneal</th>
<th>Extend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95 °C, 3 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2–6</td>
<td>98 °C, 20 s</td>
<td>60 °C, 20 s</td>
<td>72 °C, 1 min</td>
</tr>
<tr>
<td>7–17a</td>
<td>98 °C, 20 s</td>
<td>65 °C, 20 s</td>
<td>72 °C, 1 min</td>
</tr>
</tbody>
</table>

*Use optimized cycle number.

Run 2 µl from each PCR reaction on a 6% TBE-PAGE gel to determine the appropriate number of cycles and amount of input beads for each PCR reaction. PCR products should run from ~200 bp to ~1 kbp, with the majority of product running from 300–600 bp, as shown in Figure 3b. The presence of products much larger than 1 kbp (i.e., those that do not migrate on a 6% TBE-PAGE gel) indicates overamplification, and this should be avoided by reducing PCR cycle number or the volume of beads used.

Divide the remaining beads into 20-µl aliquots, and amplify the DNA using multiple PCR reactions at the optimized cycle and input parameters.

Pool all PCR reactions into one 1.5-ml microcentrifuge tube.

Purify the library by adding 0.8× volumes of AMPure XP beads.

Incubate the mixture at RT for 5 min and place the tube in a DynaMag magnet for 2 min.
Discard the supernatant and wash the beads twice with 1 ml of 80% (vol/vol) ethanol. Spin down the beads at 500g for 10 s to remove the residual ethanol as completely as possible, and then air-dry the beads for no more than 2 min.

Resuspend the beads in 25 µl of Buffer EB and incubate the mixture at RT for 1 min.

Place the resuspended beads on DynaMag magnet, and transfer the supernatant containing the eluted DNA to a fresh 1.5-ml tube.

**QC of DNase Hi-C library by BamHI digestion** ● **TIMING** 1.25 h

Quantify the amount of dsDNA in the library using the Qubit dsDNA HS kit per the manufacturer’s protocols.

Digest a small aliquot of the final DNase Hi-C library (50–100 ng) with BamHI to estimate the portion of molecules with valid biotinylated junctions as follows:

<table>
<thead>
<tr>
<th>Reagents (add in this order)</th>
<th>Digest</th>
<th>(-) Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>10× Fast digestion buffer</td>
<td>1 µl</td>
<td>1 µl</td>
</tr>
<tr>
<td>DNase Hi-C product</td>
<td>1–2 µl (50–100 ng)</td>
<td>1–2 µl (50–100 ng)</td>
</tr>
<tr>
<td>FastDigest BamHI</td>
<td>1 µl</td>
<td>0 µl</td>
</tr>
<tr>
<td>Water</td>
<td>To 10 µl</td>
<td>to 10 µl</td>
</tr>
</tbody>
</table>

Incubate the mixture at 37 °C for 30 min.

Run the entire volume of the mixture on a 6% TBE-PAGE gel. Digested libraries should demonstrate a marked shift in library size distribution, as shown in **Figure 3b**. If the libraries pass this QC metric, proceed to Illumina sequencing.

**TROUBLESHOOTING**

(Optional) Hybrid capture experiments may be carried out according to the manufacturer’s protocols provided with the Agilent SureSelect system.

**PAUSE POINT** Data can be analyzed at any point after sequencing.

Mapping, normalization, and visualization of Hi-C contact maps ● **TIMING** variable, dependent on the volume of data

Copy the output fastq sequencing files generated by the Illumina sequencer to the storage on the Linux computer.

Open a terminal on the computer and enter after the $ sign the commands described in the following steps. First, run FastQC to investigate the sequencing qualities, in which ‘L1_1’ and ‘L1_2’ correspond to the fastq sequence files for read 1 and read 2, respectively.

```
$ fastqc --extract -f fastq L1_1.fq L1_2.fq
```

**TROUBLESHOOTING**

Obtain reference genome sequences. For instance, the mouse mm9 reference sequences can be downloaded from the UCSC Genome Browser using the command below.

```
$ wget "http://hgdownload.cse.ucsc.edu/goldenPath/mm9/bigZips/chromFa.tar.gz"
$ tar -xzvf chromFa.tar.gz
$ gunzip -c chr*.fa.gz > mm9.fa
```

**CRITICAL STEP** If the *in situ* DNase Hi-C data are from female cells, do not include chrY (delete the chrY.fa.gz file prior to generating the mm9.fa file).

Run BWA to generate index files for the reference genome.

```
$ bwa index -a bwtsw -p mm9 mm9.fa
```
**PROTOCOL**

114| Run BWA to map each end of the pair-ended reads to the reference genome separately.

```
bwa aln mm9 L1_1.fq > L1_1.sai
bwa samse mm9 L1_1.sai mm9.fa > L1_1.sam
bwa aln mm9 L1_2.fq > L1_2.sai
bwa samse mm9 L1_2.sai mm9.fa > L1_2.sam
```

▲ CRITICAL STEP The two ends of the reads should be mapped separately.

115| Run SAMtools to extract high-quality (mapping quality (MAPQ) score ≥ 30) and uniquely mapped reads.

```
samtools view -S -F 4 L1_1.sam | awk '$5 ≥ 30 && $12 == "XT:A:U"' | cut –f 1-4 | sort -k1,1 > L1_1.mapped
samtools view -S -F 4 L1_2.sam | awk '$5 ≥ 30 && $12 == "XT:A:U"' | cut –f 1-4 | sort -k1,1 > L1_2.mapped
```

116| Join mapped loci pairs if both ends are successfully mapped.

```
join L1_1.mapped L1_2.mapped > L1.mapped
```

117| Remove PCR duplicates.

```
cut -f 2-7 L1.mapped | awk 'BEGIN{OFS="\t"}; {if($2<$5){print $0;} else if($2>$5){print $4,$5,$6,$1,$2,$3;} else if($3<=$6){print $0;} else{print $1,$2,$6,$2,$5,$3;}}' | sort -u > L1.unique
```

118| Parse the mapped contacts loci pairs to generate the Hi-C contact map at a given resolution.

119| Run ICE (iterative correction and eigenvector decomposition)\(^5^3\) to normalize the contact matrix using the Mirny laboratory's hiclib library (https://bitbucket.org/mirnylab/hiclib).

120| Visualize the contact map.

**TROUBLESHOOTING**

**TROUBLESHOOTING**

Troubleshooting advice can be found in **Table 4**.

**TABLE 4 | Troubleshooting table.**

<table>
<thead>
<tr>
<th>Steps</th>
<th>Problem</th>
<th>Possible reasons</th>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>gDNA digestion efficiency is poor</td>
<td>Undertreatment of fixed nuclei with SDS; inadequate amount of DNase I used for digestion</td>
<td>Optimization of the appropriate SDS and DNase I amounts may be necessary. We recommend performing the protocol through Step 38 for a variety of SDS concentrations (i.e., 0.1–0.5%) and DNase I amounts (i.e., 1–8U)</td>
</tr>
<tr>
<td>16</td>
<td>Nuclear pellet disappears during <em>in situ</em> enzymatic treatments</td>
<td>Overtreatment of fixed nuclei with SDS in Step 7</td>
<td>Reduce the amount of SDS used in the cell lysis</td>
</tr>
<tr>
<td>108, 120</td>
<td>Low percentage of long-range contacts in sequencing library or BamHI digest does not shift the library</td>
<td>Inefficient or incomplete formaldehyde cross-linking in Step 3</td>
<td>For new cell types, optimizing the amount of formaldehyde used for cross-linking may be necessary</td>
</tr>
<tr>
<td>111</td>
<td>FastQC metrics are poor</td>
<td>High duplication rate in library (e.g., &lt;60% unique sequences); low-quality sequencing run (e.g., total percentage of bases with <em>q</em> &gt; 30 is less than 85%)</td>
<td>To maximize library complexity, make sure to set up several PCR reactions in Step 114. Issues with sequencing runs themselves may be difficult to diagnose and may require outside help</td>
</tr>
</tbody>
</table>
TIMING

Steps 1 and 2 (day 1), adaptor annealing: 1 h plus overnight
Steps 3–32 (day 2), fixation, cell lysis, chromatin digestion, end repair, and adaptor ligation: ~6 h plus overnight
Steps 33–53 (day 3), adaptor cleanup; in situ phosphorylation and ligation; cross-link reversal: ~8 h plus overnight
Steps 54–65 (day 4), DNA purification and sonication: 2.5–3.5 h
Steps 66–86 (day 5), biotin pull-down, end repair/dA tailing, and adaptor ligation of Hi-C fragments: ~3 h
Steps 87–109 (day 6), library amplification, BamHI quality check, and sequencing: ~4 h for amplification and quality check; up to several days/weeks for sequencing, depending on the instrumentation
Steps 110–120 (day 7 and beyond), data analysis time is variable and depends on the sequencing depth and available computer resources

Box 1, assessment of nuclear lysis at various steps: ~3 d

ANTICIPATED RESULTS

We recommend QCing all libraries that pass the BamHI digestion test (typical results, including a negative control EcoRI digest, shown in Fig. 3b) by sequencing at low depth first to ensure that the libraries are sufficiently complex for your desired application. We also recommend quantifying the length classes of sequenced ligation pairs in libraries; in situ DNase Hi-C libraries should demonstrate an enrichment for pairs mapping with long-range (i.e., >1 kb) distances between them (example distributions shown in Fig. 4a). Furthermore, we recommend quantifying the relative numbers of different ligation pairs (i.e., ‘in-facing,’ ‘out-facing,’ ‘left,’ and ‘right’) in libraries (a typical example is shown in Fig. 4b). Corrected matrices generated from valid in situ DNase Hi-C (DHC) libraries should be analogous to the example shown in Figure 4c, with large-scale structures (i.e., TADs) clearly visible even at 100-kb resolution.

We have observed that the relative fraction of interchromosomal ligation pairs in in situ DNase Hi-C libraries is largely cell type–specific but highly reproducible—in line with previously published in situ results. This is evident in Supplementary Figure 1, which compares fractions of various ligation pairs between the Patski cell line and three replicates of the human lymphoblastoid cell line GM12878. When considering gold standards for in situ DNase Hi-C experiments, we typically look to the abundance of ‘long-range’ ligation pairs in our libraries, which typically make up >40% of uniquely mapped read pairs.

Using this modified DHC protocol, we have shown that the inactive murine X chromosome adopts a bipartite structure.

Figure 4 In situ DNase Hi-C results for the mouse embryonic kidney Patski cell line. (a) In situ DNase Hi-C reads (950,206 downsampled reads from data published in Deng et al.51 (using the mouse Patski cell line, rather than the GM12878 cell line) demonstrate an enrichment for long-range (i.e., >1 kb) intrachromosomal read pairs expected of Hi-C libraries. (b) Expected breakdown of orientations for read pairs in in situ DNase Hi-C data. For intrafragment distances >1 kb, a roughly 25% split should be observed for each orientation class. (c) Normalized heatmap generated from data published in Deng et al. (GEO accession no. GSE68992) for mouse chromosome 18- at 100-kb resolution. The data set used to generate this heatmap contained 60,666,200 uniquely mapped, high-quality read pairs.
consistent with results obtained using traditional Hi-C, in both an analogous murine system\(^43\) and human lymphoblastoid cells\(^43\). These results suggest that the in situ DHC protocol produces a signal comparable to that of existing Hi-C protocols while ultimately providing a less-biased empirical method for generating higher-resolution 3D maps of chromatin structure.