

Deciphering High-Resolution 3D Chromatin Organization via Capture Hi-C

Antonia Hauth¹, Rafael Galupa¹, Nicolas Servant², Laura Villacorta¹, Kai Hauschulz³, Joke Gerarda van Bommel⁴, Agnese Loda¹, Edith Heard^{1,5}

¹EMBL: European Molecular Biology Laboratory ²Institut Curie ³Agilent Technologies ⁴Genmab BV ⁵Collège de France

Corresponding Authors

Agnese Loda

agnese.loda@embl.de

Edith Heard

edith.heard@embl.org

Citation

Hauth, A., Galupa, R., Servant, N., Villacorta, L., Hauschulz, K., van Bommel, J.G., Loda, A., Heard, E. Deciphering High-Resolution 3D Chromatin Organization via Capture Hi-C. *J. Vis. Exp.* (188), e64166, doi:10.3791/64166 (2022).

Date Published

October 14, 2022

DOI

10.3791/64166

URL

jove.com/video/64166

Introduction

The linear genome holds all the information necessary for an organism to undergo embryonic development and survive throughout adulthood. However, instructing genetically identical cells to perform different functions is fundamental for accurately controlling which information is used in specific contexts, including different tissues and/or developmental stages. The three-dimensional organization of the genome

Abstract

The spatial organization of the genome contributes to its function and regulation in many contexts, including transcription, replication, recombination, and repair. Understanding the exact causality between genome topology and function is therefore crucial and increasingly the subject of intensive research. Chromosome conformation capture technologies (3C) allow inferring the 3D structure of chromatin by measuring the frequency of interactions between any region of the genome. Here we describe a fast and simple protocol to perform Capture Hi-C, a 3C-based target enrichment method that characterizes the allele-specific 3D organization of megabased-sized genomic targets at high-resolution. In Capture Hi-C, target regions are captured by an array of biotinylated probes before downstream high-throughput sequencing. Thus, higher resolution and allele-specificity are achieved while improving the time-effectiveness and affordability of the technology. To demonstrate its strengths, the Capture Hi-C protocol was applied to the mouse X-inactivation center (*Xic*), the master regulatory locus of X-chromosome inactivation (XCI).

is thought to participate in this accurate spatio-temporal regulation of gene activity by facilitating or preventing the physical interaction between regulatory elements that can be separated by several hundred kilobases in the linear genome (for reviews^{1,2,3}). In the last 20 years, our understanding of the interplay between genome folding and activity has rapidly increased, largely owing to the development of

chromosome conformation capture technologies (3C) (for review^{4,5,6,7}). These methods measure the frequency of interactions between any regions of the genome and rely on the ligation of DNA sequences that are in close 3D proximity within the nucleus. The most common 3C protocols start with the fixation of cell populations with a cross-linking agent such as formaldehyde. The cross-linked chromatin is then digested with a restriction enzyme, although MNase digestion has also been used^{8,9}. After digestion, free DNA ends in close spatial proximity are re-ligated, and cross-linking is reversed. This step gives rise to the 3C 'library' or 'template', a mixed pool of hybrid fragments in which sequences that were in 3D proximity to the nucleus have higher chances of getting ligated in the same DNA fragment. The downstream quantification of these hybrid fragments enables inferring the 3D conformation of genomic regions that are located thousands of base pairs apart in the linear genome but might interact in the 3D space.

Many different approaches have been developed to characterize the 3C library, differing both in terms of which subsets of ligation fragments are analyzed and which technology is used for their downstream quantification. The original 3C protocol relied on the selection of two regions of interest and the quantification of their 'one versus one' interaction frequency by PCR^{10,11}. The 4C approach (circular chromosome conformation capture) measures the interactions between a single locus of interest (i.e., the 'view-point') and the rest of the genome ('one versus all')^{12,13,14}. In 4C, the 3C library undergoes a second round of digestion and re-ligation to generate small circular DNA molecules that are PCR amplified by view-point specific primers¹⁵. 5C (chromosome conformation capture carbon copy) enables the characterization of 3D interactions across larger regions of interest, providing insights into higher-

order chromatin folding within that region ('many versus many')¹⁶. In 5C, the 3C library is hybridized to a pool of oligonucleotides overlapping restriction sites that can be subsequently amplified by multiplex PCR with universal primers¹⁵. In both 4C and 5C, the informative DNA fragments were initially quantified by microarrays and later by next-generation sequencing (NGS)^{17,18,19}. These strategies characterize targeted regions of interest but cannot be applied to map genome-wide interactions. This latter goal is achieved with Hi-C, a 3C-based high-throughput strategy in which massively parallel sequencing of the 3C template allows the unbiased characterization of chromatin folding at the genome-wide level ('all versus all')²⁰. The Hi-C protocol includes the incorporation of a biotinylated residue at the digested fragments' ends, which is followed by pull-down of ligation fragments with streptavidin beads to increase the recovery of ligated fragments²⁰.

Hi-C revealed that mammalian genomes are structurally organized at multiple scales in the 3D nucleus. At the megabase scale, the genome is divided into regions of active and inactive chromatin, the A and B compartments, respectively^{20,21}. The existence of further subcompartments represented by different chromatin and activity states was also subsequently shown²². At higher resolution, the genome is further partitioned into sub-megabase self-interacting domains called topologically associating domains (TADs), first revealed by Hi-C and 5C analysis of the human and mouse genomes^{23,24}. Unlike compartments which vary in a tissue-specific manner, TADs tend to be constant (although there are many exceptions). Importantly, TAD boundaries are conserved across species²⁵. In mammalian cells, TADs frequently encompass genes sharing the same regulatory landscape and have been shown to represent a structural framework that facilitates gene co-regulation while limiting

the interactions with neighboring regulatory domains (for review^{3,26,27,28}). Furthermore, within TADs, interactions due to CTCF sites at the base of cohesin-extruded loops may increase the probability of promoter-enhancer or enhancer-enhancer interactions (for review²⁹).

In Hi-C, compartments and TADs can be detected at 1 Mb to 40 kb resolution, but higher resolution can be achieved to characterize smaller scale contacts such as looping interactions between distal elements at the scale of 5-10 kb. However, increasing the resolution to be able to detect such loops efficiently by HiC requires a significant increase in sequencing depth and, therefore, sequencing costs. This is exacerbated if the analysis needs to be allele-specific. Indeed, an X-fold increase in resolution requires an X^2 increase in sequencing depth, meaning that high-resolution and allele-specific genome-wide approaches can be prohibitively expensive³⁰.

To improve cost-effectiveness and affordability while maintaining high-resolution, target regions of interest can be physically pulled down from genome-wide 3C or Hi-C libraries following their hybridization with complementary biotin-labeled oligonucleotide probes before downstream sequencing. These target enrichment strategies are referred to as Capture-C methods and allow the interrogation of interactions of hundreds of target loci scattered across the genome (i.e., Promoter Capture (PC) Hi-C; Next Generation (NG) Capture-C; Low Input (LI) Capture-C; Nuclear Titrated (NuTi) Capture-C; Tri-C)^{31,32,33,34,35,36,37,38,39,40}, or across regions spanning up to several megabases (i.e., Capture HiC; HYbrid Capture Hi-C (Hi-C²); Tiled-C)^{41,42,43}. Two aspects can vary in capture-based methods: (1) the nature and design of biotinylated oligonucleotides (i.e., RNA or DNA, single oligos capturing dispersed genomic targets or

multiple oligos tiling a region of interest); and (2) the template that is used for pulling down targets which can be the 3C or Hi-C library, the latter consisting of biotinylated restriction fragments pulled down from the 3C library.

Here, a Capture Hi-C protocol based on the enrichment of target contacts from the 3C library is described. The protocol relies on the design of a custom-tailored tiling array of biotinylated RNA probes and can be performed in 1 week from the 3C library preparation to the NGS sequencing. The protocol is fast, simple, and allows characterizing the higher order 3D organization of megabase-sized regions of interest at 5 kb resolution while improving time effectiveness and affordability in comparison with other 3C methods. The Capture Hi-C protocol was applied to the master regulatory locus of X-chromosome inactivation (XCI), the X-inactivation center (*Xic*), which hosts the *Xist* noncoding RNA. The *Xic* has previously been the subject of extensive structural and functional analyses (for review^{44,45}). In mammals, XCI compensates for the dosage of X-linked genes between females (XX) and males (XY) and involves the transcriptional silencing of almost the entirety of one of the two X chromosomes in female cells. The *Xic* has represented a powerful, gold standard locus for studies in 3D genome topology and the interplay with gene regulation⁴⁴. 5C analysis of the *Xic* in mouse embryonic stem cells (mESCs) led to the discovery and naming of TADs, providing the first insights into the functional relevance of topological partitioning and gene co-regulation²⁴. The topological organization of the *Xic* was subsequently shown to be critically involved in the appropriate developmental timing of *Xist* upregulation and XCI⁴⁶, and unsuspected cis-regulatory elements that can influence gene activity within and between TADs were also recently discovered within the *Xic*^{47,48,49}. Applying Capture Hi-C to 3 Mb of the mouse X chromosome spanning the *Xic*

demonstrates the power of this approach in dissecting large-scale chromatin folding at high resolution. A detailed and easy-to-follow protocol is provided, starting from the design of the array of biotinylated probes across every DpnII restriction site within the region of interest to the generation of the genome-wide 3C library, the hybridization and capture of target contacts, and downstream data analysis. An overview of the appropriate quality controls and expected results is also included, and both the strengths and limitations of the approach are discussed in light of similar existing methods.

Protocol

The mouse embryonic stem cells (mESCs) used in this study were derived from a cross of a TX/TX R26^{rtTA/rtTA} female⁵⁰ with a *Mus musculus castaneus* male according to the animal care guidelines of Institut Curie (Paris)⁵¹.

1. Probe design

1. Design an array of biotinylated probes (120-mer RNA oligonucleotides) covering the target region of interest.
 1. Tile the region of interest with overlapping oligonucleotides so that on average each sequence within the target is covered by two unique probes (2x coverage) (**Figure 1**).
 2. Exclude repetitive sequences from probe coverage to avoid enrichment of unspecific interactions.

NOTE: To maximize the enrichment of informative ligation fragments, regions spanning 300 bp upstream and downstream of each DpnII restriction site across the target were defined (ChrX: 102,475,000-105,475,000), and 28,913 biotinylated probes were designed according to the SureSelect DNA target enrichment technology through the Sure

Design platform⁵². According to this strategy, up to a maximum of 40 bases of repetitive sequences are allowed in each oligonucleotide to minimize the enrichment of unspecific interactions. The probe array was synthesized by Agilent. Here, DpnII is used as a restriction enzyme for two reasons: (1) it is a four-cutter routinely used in several 3C-based methods⁵³; and (2) it maximizes the chances of capturing informative single nucleotide polymorphisms (SNPs) in the proximity of the cutting sites in comparison with other restriction enzymes that were tested *in silico* in F1 hybrid lines used in this study (C57BL/6J x CASTeI/J).

2. Experimental procedure

1. Cell preparation
 1. Seed the appropriate number of cells on one or multiple cell culture plates to achieve a total cell number of $\geq 5 \times 10^7$ cells on the day of fixation.

NOTE: Mouse embryonic stem cells (mESCs) were used in this study. mESCs are plated on gelatinized (0.1% gelatine in 1x PBS - o/n at 37 °C, 5% CO₂ incubator) cell culture plates in mESCs medium containing 2i + LIF and batch-tested fetal calf serum (DMEM, 15% FBS, 0.1 mM β -mercaptoethanol, 1,000 U/mL⁻¹ leukaemia inhibitory factor (LIF), CHIR99021 (3 μ M), and PD0325901 (1 μ M)). For this cell type, one 80% confluent 10 cm plate contains approximately 2×10^7 cells.
 2. Prepare one additional cell culture plate for cell counting.

NOTE: A smaller cell culture plate can be used to reduce media usage. In this case, the number

of cells to seed on the smaller plate needs to be adjusted accordingly (e.g., 3x fewer cells on a 10 cm plate compared to a 15 cm plate).

2. Formaldehyde fixation

1. Estimate the total number of cells to be cross-linked.

1. Before starting the cross-linking reaction, trypsinize and count cells from the control plate prepared specifically for cell counting using an automated cell counter according to the manufacturer's instructions.

2. Include a viability staining (e.g., Trypan Blue) to determine the percentage of viable cells⁵⁴. From this cell count, estimate the total number of cells in the plate(s) prepared for cross-linking.

2. Remove the culture medium from the plates prepared for cross-linking and replace it with the appropriate amount of fixation solution (2% formaldehyde in cell culture medium). Use 10 mL on a 10 cm plate (e.g., ~20 mL for a 15 cm plate).

NOTE: Add an exact volume of fixation solution. If fixing adherent cells is not possible, this step can be adapted to trypsinized cells and performed in 30 mL of fixation solution in 50 mL conical centrifuge tubes. Formaldehyde must not be more than 1 year old. It is preferred to use single-use vials. The fixation solution must be brought to room temperature (RT) before use.

CAUTION: Formaldehyde is hazardous and needs to be handled according to appropriate health and safety regulations.

3. Fix for 10 min at RT under gentle mixing on a shaker.

4. Quench the fixation reaction by addition of 2.5 M glycine-1x PBS to a final concentration of 0.125 M.

Add 530 μ L of 2.5 M glycine-1x PBS to 10 mL on a 10 cm plate (e.g., 1060 μ L to 20 mL on a 15 cm plate).

NOTE: If the cells were fixed in solution, quench the fixation reaction with 1590 μ L of 2.5 M glycine-1x PBS.

5. Incubate for 5 min at RT, gently mixing on a shaker.

6. Transfer the plates to ice and incubate for an additional 15 min on ice while gently mixing on a shaker.

NOTE: From now on, the cells must be kept on ice, and buffers must be pre-chilled to avoid further cross-linking. Move to a cold room if many plates need to be processed.

7. Remove the fixation solution from the cells by pouring it into a beaker to ensure fast handling.

NOTE: Make sure to dispose of the formaldehyde-containing liquid waste according to the appropriate health and safety regulations.

8. Rinse the 10 cm plate quickly two times with 5 mL of cold 0.125 M glycine-1x PBS (8 mL for a 15 cm plate) to wash off the debris and dead cells. Remove the liquid from the plate by pouring it into a beaker to ensure fast handling.

9. Add 5 mL of cold 0.125 M glycine-1x PBS to the 10 cm plate (10 mL for a 15 cm plate) and quickly scrape the cells from the plate using a plastic cell scraper.

10. Transfer the cell suspension into a pre-cooled 50 mL conical centrifuge tube using a serological pipette.

11. Rinse the plate twice with 5 mL of cold 0.125 M glycine-1x PBS and add the cell suspension to the conical centrifuge tube.

12. Spin down at 480 x g for 10 min at 4 °C.

NOTE: If the cells were fixed in solution, transfer the cell to a pre-chilled conical centrifuge tube and spin down at 480 x g for 10 min at 4 °C. Remove the fixation solution by pouring it into a beaker and wash three times in 10 mL of cold 0.125 M glycine-1x PBS. Make sure to resuspend the cells at each washing step.

13. Remove the supernatant by aspirating with a benchtop aspiration system. Resuspend the cells in 500 µL of 1x PBS per 1×10^7 cells by carefully pipetting up and down with a P1000 pipette. To resuspend cells in the accurate volume, refer to the total cell number estimation obtained in 2.2.1.
14. Aliquot 500 µL of the cell suspension into the calculated number of 1.5 mL microcentrifuge tubes (1×10^7 cells/tube).
15. Spin down at 480 x g for 10 min at 4 °C.
16. Remove the supernatant with a benchtop aspiration system and snap freeze the cell pellets in liquid nitrogen. Store the dry cell pellets at -80 °C.

NOTE: Samples can be stored for at least 1 year.

3. Cell lysis

1. Thaw the frozen pellet(s) on ice.
2. Prepare 1.5 mL of lysis buffer in H₂O per sample: Add 10 mM Tris-HCl, pH 8.0, 10 mM NaCl, and 0.2% NP40.
3. Add 600 µL of the cold lysis buffer and resuspend well on ice.
4. Incubate on ice for 15 min to let the cells swell.
5. Spin down at 2655 x g for 5 min at 4 °C and remove the supernatant using a benchtop aspiration system.

6. To remove debris, resuspend the pellet in 1 mL of the cold lysis buffer, spin down at 2655 x g for 5 min at 4 °C, and remove the supernatant.
7. Spin again shortly at 2655 x g and 4 °C and remove as much of the remaining supernatant as possible using a benchtop aspiration system fitted with a P200 tip.
8. Resuspend in 100 µL of 0.5% (vol/vol) SDS.
9. Incubate in a thermomixer at 62 °C, swirling at 1400 rpm for 10 min.
10. Add 290 µL of H₂O + 50 µL of 10% TritonX-100 and mix well, avoiding air bubbles.
11. Incubate in a thermomixer at 37 °C, swirling at 1400 rpm for 15 min.
12. Add 50 µL of 10x DpnII buffer and invert the tube to mix.
13. Take 50 µL of undigested DNA for quality control into a separate tube. Do not forget to take the undigested control sample.

4. DpnII digestion

1. Add 10 µL of DpnII high concentration (500 U total) and mix by inverting.
2. Incubate the samples and the undigested control in a thermomixer at 37 °C, swirling at 1400 rpm for >4 h.
3. Add 10 µL of DpnII high concentration (500 U total) at the end of the day.
4. Incubate the samples and the undigested control at 37 °C, swirling at 1400 rpm overnight.
5. Add 10 µL of DpnII high concentration (500 U total) at the beginning of the next day to the samples.

6. Incubate the samples and the undigested control in a thermomixer at 37 °C, swirling at 1400 rpm for 4 h.
5. Ligation and reversal of cross-linking
 1. Incubate the tubes at 65 °C for 20 min at 1400 rpm.
NOTE: Do not add SDS at this point. The idea is to preserve nuclear integrity, so the ligation is carried out inside the nuclei, circumventing the need for extreme dilution.
 2. Cool down the samples on ice for a maximum of 5-10 min. To avoid SDS precipitation, do not leave the samples on ice longer than this.
 3. Take 50 µL of the unligated digested DNA for quality control in a separate tube. Store the undigested and unligated controls at -20 °C.
NOTE: Do not forget to take the unligated control sample.
 4. Add 800 µL of ligation cocktail: 122 µL of 10x ligase buffer, 8 µL of T4 ligase (30 U/µL), and 670 µL of H₂O.
 5. Incubate at 16 °C, swirling at 1000 rpm overnight.
 6. Add 7.5 µL of Proteinase K (20 mg/mL) to the samples and 2 µL to the controls.
 7. Incubate at 65 °C for 4 h at 1000 rpm.
 6. DNA purification
 1. Transfer the samples on ice to pre-cooled 15 mL conical centrifuge tubes and add 2 mL of water, 10.5 mL of ice-cold EtOH, and 583 µL of 3 M NaAC.
NOTE: Additional water aims to prevent carry-over of DTT into the pellet.
 2. Add 200 µL of ice cold EtOH, 10.8 µL of NaAC, and 1 µL of the coprecipitant to the undigested and unligated quality controls.
 3. Incubate at -80 °C for at least 4 h up to overnight.
 4. Spin the 15 mL tubes at 2200 x *g* at 4 °C for 45 min.
 5. Spin the 1.5 mL control tubes at 20,500 x *g* at 4 °C for 30 min.
 6. Wash once with 3 mL (samples) and 1 mL (controls) of ice-cold 70% EtOH.
 7. Spin at 2200 x *g* (samples) or 20,500 x *g* (controls) at 4 °C for 10 min.
 8. Carefully remove EtOH and air dry at RT for 10-15 min; do not overdry.
 9. Resuspend the samples and controls in 100 µL and 20 µL of H₂O, respectively.
 10. Add 1 µL of RNaseA and incubate at 37 °C, swirling at 1400 rpm for 30 min.
 7. Quality control of 3C template preparation
 1. Quantify each sample and control using a fluorometer kit for high-sensitivity DNA concentration measurements.
 2. Load 100-200 ng of each sample and each control on a 1% agarose/1x TBE gel.
 3. Verify that the gel image shows the expected result by comparing the differences in DNA fragment sizes of the controls and the 3C template as shown in **Figure 2A**.
 4. Store the samples and controls at -20 °C.
 8. Hybridization, capture, and sample processing for multiplexed sequencing

1. To hybridize the array of biotinylated RNA probes to the 3C template, capture the targeted ligation fragments, and prepare the samples for multiplexed sequencing according to the target enrichment system used in this study for paired-end multiplexed sequencing (see **Table of Materials**). Follow the protocol according to the manufacturer's instructions while introducing the following minor modifications:

1. Section 2 of the manufacturer's protocol:
Sample preparation

1. Follow the instructions for target enrichment starting from 3 µg of gDNA input.
2. Shear the DNA in a sonicator using the following specifications: 10% duty cycle, 4 intensity, 200 cyc/burst, and 130 s. Start with 4 µg of 3C template resuspended in 130 µL of water for each capture reaction to ensure enough material to continue the sample preparation with 3 µg of the sheared DNA.
3. Assess the quality of sheared DNA. Run 1 µL of the sheared DNA on a DNA bioanalyzer according to the high-sensitivity protocol. Expect a distribution of fragment size between 150-700 bp (**Figure 2**).
4. Purify the sample using solid phase reversible immobilization (SPRI) beads. Add 124 µL of SPRI beads to 124 µL of the DNA sample to perform a 1:1 left side-size selection according to the manufacturer's instructions and elute in 25

µL of nuclease-free water. This purification step will remove shorter fragments to enrich fragments of around 300 bp (**Figure 2**).

NOTE: The amount of samples and SPRI beads used at this step take into account the volume loss that occurred while transferring the samples to new tubes and running the quality controls at the Bioanalyzer. All subsequent size selection steps are performed according to the ratios recommended by the manufacturer's protocol. DNA elution from SPRI beads is performed at RT throughout the protocol.

5. Assess the quality of size-selected sheared DNA. Run 1 µL of the sheared DNA on the DNA bioanalyzer according to the high-sensitivity (HS) protocol. Expect a distribution of fragment sizes with the highest enrichment at 300 bp (**Figure 2**). Go ahead with the quantification of the sheared DNA if shearing was successful.
6. Quantify the sheared DNA with a fluorometer kit for HS DNA concentration measurements.

NOTE: If DNA shearing results in a DNA yield of <3 µg, perform a second round of DNA shearing with another 4 µg of DNA and combine the sheared DNA samples after the first SPRI bead purification step to achieve a total of 3 µg of sheared DNA.

7. Add nuclease-free water to the size-selected cleaned DNA sample (3 µg total) to a final volume of 48 µL and proceed

with the end repair reaction according to the manufacturer's protocol.

8. After ligation of paired-end adaptors, amplify the library by performing five cycles of pre-capture PCR according to the manufacturer's instructions (the conditions for PCR and the primers are provided in the kit).
2. Section 4 of the manufacturer's protocol: Hybridization and capture
 1. To hybridize the prepared DNA samples to the target-specific RNA probes, dilute 750 ng of DNA samples in a final volume of 3.4 μ L, resulting in an initial concentration of 221 ng/ μ L. For DNA samples diluted in bigger volumes, use a speed-vacuum concentrator to reduce to the final volume. A speed-vacuum concentration (250 x g; \leq 45 °C) for 15-20 min is normally sufficient for the samples resuspended in 10 μ L. Make sure to have the same input volume for each sample before starting the speed-vacuum concentrator.
 2. Incubate the hybridization mixture for 16-18 h at 65 °C with a heated lid at 105 °C according to the manufacturer's instructions.
3. Section 5 of the manufacturer's protocol: Indexing and sample processing for multiplexed sequencing
 1. To amplify the captured libraries with indexing primers, perform 12 cycles of post-capture PCR according to the

manufacturer's instructions (the conditions for PCR and the primers are provided in the kit).

9. Next-generation sequencing

1. To run multiple capture Hi-C libraries on the same flow cell, prepare an equimolar mixture of the capture libraries and sequence 100-120 M reads per library.
2. If the allele-specific analysis is needed, sequence 150 bp paired-end to ensure sufficient SNP coverage.

3. Data analysis

1. Apply the HiC-Pro pipeline to perform Capture Hi-C data analysis⁵⁵. HiC-Pro provides quality controls at each step of the processing, including (**Figure 3**):
 - (i)The alignment rate on the reference genome specifying the fraction of reads spanning a ligation site, as well as the number of pairs and singletons.
 - (ii)The fraction of valid ligation products and non-informative read pairs (dangling-end, self-ligation, etc).
 - (iii)The fraction of short/long-range and intra/inter-chromosomal contacts.
 - (iv)The fraction of on-target contacts for Capture Hi-C.
 - (v)The fraction of allele-specific reads if specified.

NOTE: HiC-Pro supports a wide range of protocols including *in situ* Hi-C and Capture Hi-C. In the latter case, the user simply needs to specify the targeted region (BED format) in the configuration file. Once the data are processed, the HiC-Pro outputs can easily be converted into a cooler object for downstream analysis⁵⁶. At this step, the contact maps at various resolutions are normalized using the ICE method previously described

by Imakaev and colleagues⁵⁷. Several analyses can then be run to call chromosome compartments, TADs, or chromatin loops (for review⁵⁸). The workflow of the protocol is shown in **Figure 4**. Here, the 'cooltools' suite is applied to calculate the insulation score and TADs boundaries, as illustrated in **Figure 5** and **Figure 6**⁵⁹.

Representative Results

The described Capture Hi-C protocol is based on the preparation of the genome-wide 3C-template using a four-base cutter (DpnII). The subsequent enrichment of ligation fragments across the genomic region of interest is obtained by hybridization of an array of tiling RNA probes and their streptavidin-based capture according to the target enrichment system used in this study (**Figure 1**). Biotinylated RNA probes were selected as they show tighter binding affinity to their targets compared to DNA probes^{52,60}. Captured libraries are then indexed and pooled for multiplexed high-throughput sequencing. Capture Hi-C data can be visualized as high-resolution Hi-C interaction maps but also as 4C-like single view-point contact maps to specifically visualize the interactions of smaller sequences such as promoters or enhancers within the entire captured region. The workflow of the protocol is shown in **Figure 4**. Pre-sequencing quality controls are shown in **Figure 2** and include the assessment of proper digestion and re-ligation of the 3C template and its efficient shearing and purification across the different steps of the protocol. The sheared 3C template DNA is expected to run between 150 to 700 bp, and no enrichment of fragments >2 kb should be detected. During the following steps, several bead-based DNA cleanup and size selection steps are performed, first after the shearing, then after the pre-capture and post-capture PCRs. Cleaned libraries show a distinct fragment enrichment profile as visualized on a high

sensitivity DNA bioanalyzer (**Figure 2**). The mean fragment size increases over the course of the library preparation due to the ligation of adaptors, sequencing, and indexing primers. Post-sequencing quality controls are obtained via Hi-C Pro and shown in **Figure 3**. Many different bioinformatics software applications have been proposed for 3C-like data processing and analysis. Among them, the HiC-Pro pipeline is one of the most popular solutions, allowing the processing of raw sequencing data to the final contact maps at various resolutions⁵⁵. HiC-Pro uses a two-step mapping strategy to align the sequencing reads on the reference genome. The 3C products are then reconstructed and filtered out to remove non-informative pairs of contact and to generate the contact maps. In addition, it is able to use a list of known polymorphisms to perform allele-specific analysis and to separate the contacts coming from the two parental alleles in distinct contact maps. More recently, HiC-Pro has been included and extended into the nf-core framework (nf-core-hic), providing a highly scalable and reproducible community-driven pipeline^{61,62}.

To capture the mouse *Xic*, an array of 28,913 RNA probes tiling 3 Mb of the X chromosome was designed. This region includes the key player in XCI, the long noncoding gene *Xist*, and its known ~800 kb regulatory landscape (**Figure 5**). This ~800 kb region is partitioned into two TADs: one including the *Xist* promoter and its known positive regulators (i.e., the noncoding transcripts *Ftx*, *Jpx*, and *Xert* and the protein-coding gene *Rnf12*), and the neighboring TAD encompassing the negative cis-regulators of *Xist* (i.e., its antisense transcript *Tsix*, the enhancer element *Xite*, and the noncoding transcript *Linx*) (for review^{44,45}).

By applying the described Capture Hi-C protocol to the *Xic*, the topological organization of this locus was obtained at

unprecedented resolution (**Figure 6** and **Figure 7**). This is particularly clear when comparing the Capture Hi-C profile to previously published 5C⁴⁷ (**Figure 6** and **Figure 7**; **Supplementary Table 1**) and Hi-C⁶¹ (**Figure 6** and **Figure 7**; **Supplementary Table 1**) profiles. For instance, sub-TAD structures are more evident — the TAD containing the *Xist* promoter (*Xist*-TAD) is clearly subdivided into two smaller domains (**Figure 6A**, blue arrowhead). Previously, this could only be visually "guessed" from the 5C profile (**Figure 6B**), albeit the detection of a boundary in this region using the insulation score algorithm. Likewise, the resolution of the Capture Hi-C profile allows the identification of two smaller domains in the neighboring TAD (**Figure 6A, B**), which contains the promoter of the *Tsix* locus (*Tsix*-TAD); this was not previously achieved with 5C (**Figure 6B**). Of note, topological boundaries determined by the insulation score from the Capture Hi-C and 5C data are generally detected at slightly different locations and with different relative strengths.

Moreover, other sub-TAD structures such as contact loops are clearly visible from the Capture Hi-C data, such as the

loop between *Xist* and *Ftx* (**Figure 7A**), previously identified with Capture-C⁶³, and the loop between *Xist* and *Xert* (**Figure 7B**), recently identified using a similar protocol for Capture Hi-C⁴⁸. Other contacts can also be mapped more precisely due to the increased resolution of the Capture Hi-C profiles, such as those forming the known contact hotspots within the *Tsix*-TAD between the *Linx*, *Chic1*, and *Xite* loci (**Figure 7A**).

In comparison with the Hi-C data shown in **Figure 7**, Capture Hi-C allowed for a fourfold increase in resolution, yet it required only one-fourth of sequencing depth (i.e., 126 M reads versus 571 M) (**Supplementary Table 1**). This increase in resolution allows for the detection of subTADs and looping interactions that could not be detected by Hi-C at the sequencing depth shown in **Figure 6** and **Figure 7**. The described protocol for Capture Hi-C thus allows for a much more detailed, high-resolution characterization of a large genomic region of interest, when compared to previous approaches.

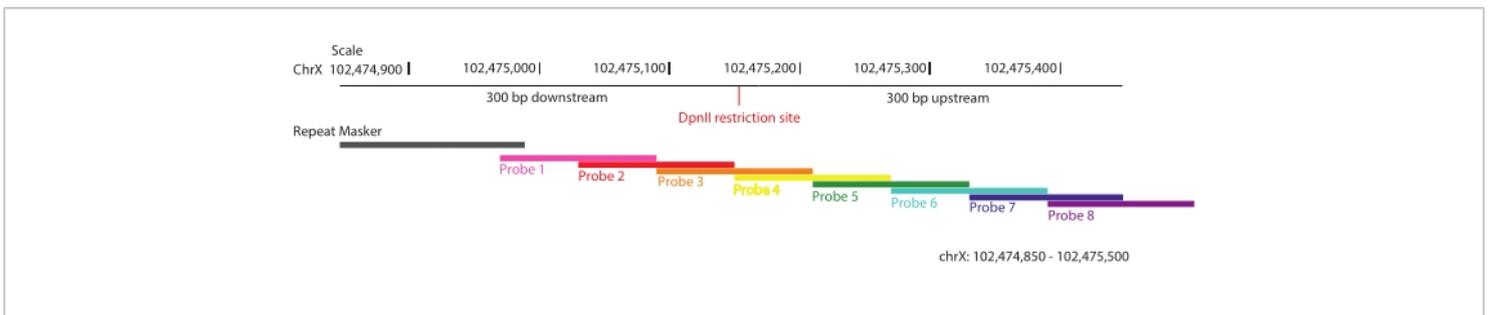


Figure 1: Probe design. Schematic representation of the strategy used for probe design. Regions of 300 bp upstream and downstream of each DpnII restriction site across the 3 Mb target region were selected and tiled with overlapping biotinylated RNA probes. One of these selected regions is shown, chrX: 102,474,805-102,475,500. No more than 40 bases of repetitive sequences are allowed in each probe. [Please click here to view a larger version of this figure.](#)

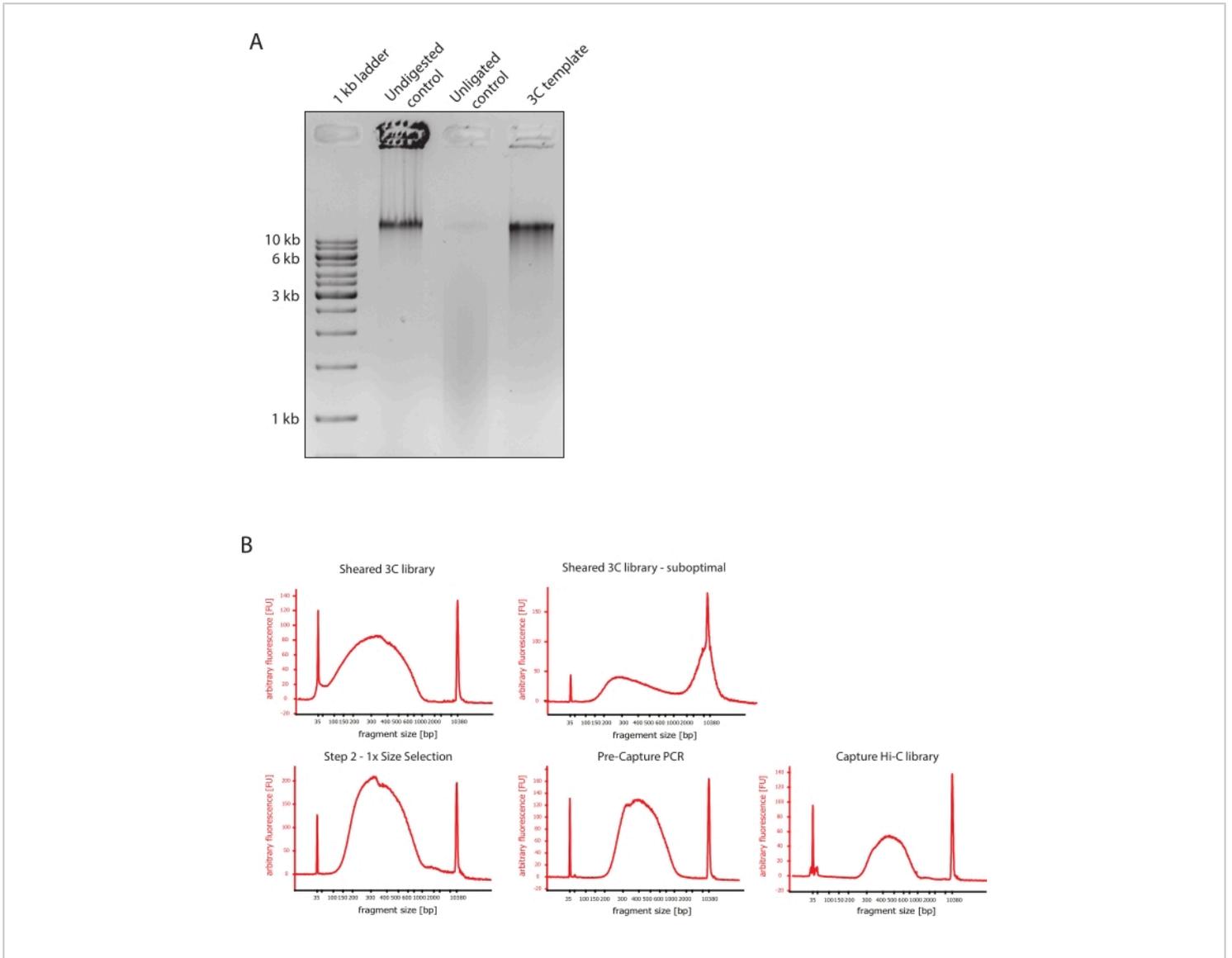


Figure 2: Capture Hi-C pre-sequencing quality controls. (A) Representative example of 3C template quality controls. 200 ng of DNA were loaded on a 1% agarose gel. Lane 1: 1 kb ladder. Lane 2: Undigested, cross-linked, and intact chromatin runs as a sharp band at >10 kb. Lane 3: DpnII-digested cross-linked chromatin runs as a smear between 1 kb to 3 kb in size. Lane 4: Final 3C library or template; free ends of digested cross-linked DNA fragments are re-ligated. The DNA smear of lower molecular size is almost undetectable, and the ligation product is detected as a band of >10 kb. (B) Representative examples of high sensitivity bioanalyzer DNA profiles. Top left: successfully sheared 3C library showing a distribution of fragment size between 150 bp and 700 bp. Top right: unsatisfactory sheared 3C library. Unsheared DNA is detected as broad enrichment of fragments >2 kb. (C) Bottom left: sheared DNA sample following a 1:1 left side-size selection using SPRI beads. Fragments of ~300 bp are enriched. Bottom middle: Pre-capture PCR profile after ligation of paired-end adaptors according to the manufacturer's protocol. Bottom right: final Capture Hi-C library including adaptors, sequencing,

and indexing primers for multiplexed sequencing. Abbreviations: bp = base pairs, FU = arbitrary fluorescence unit. [Please click here to view a larger version of this figure.](#)

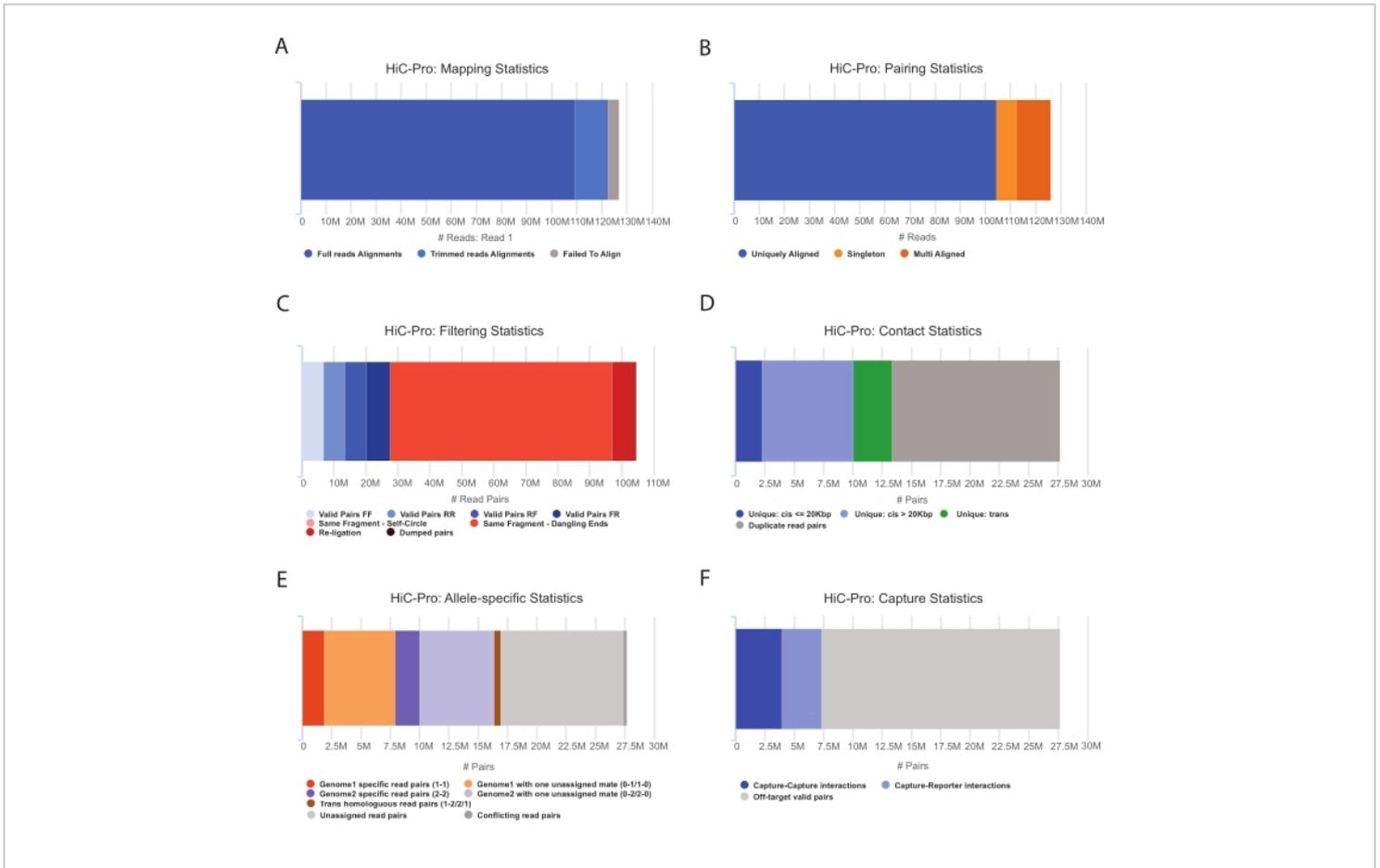


Figure 3: Capture Hi-C post-sequencing quality controls with HiC-Pro. (A) Example of mapping rate on the reference genome for the first mate of the sequencing pairs. The light blue fraction represents the reads aligned by HiC-Pro and spanning a ligation junction. This metric can thus be used to validate the experimental ligation step. (B) Once sequencing mates are aligned on the genome, only uniquely aligned read pairs are kept for analysis. (C) Non-valid pairs (in red) such as dangling-end, self-circle, or re-ligation are discarded from the analysis. The fraction of valid pairs is a good indicator of the ligation and pull-down efficiency. (D) The valid pairs can be further divided into intra/inter-chromosomal and short/long-range contacts. Duplicated read pairs that are likely to represent PCR artifacts are discarded from the analysis. (E) For allele-specific analysis, HiC-Pro reports the number of allelic reads supported by either one or two mates for each parental genome (i.e., C57BL/6J x CASTEi/J). The same fraction of reads assigned to the maternal and paternal allele are expected. (F) Finally, only valid pairs overlapping the capture region are selected to build the contact maps. Capture-capture pairs represent contacts within the targeted region, while capture-reporter pairs involve interaction between the targeted region and an off-target one. [Please click here to view a larger version of this figure.](#)

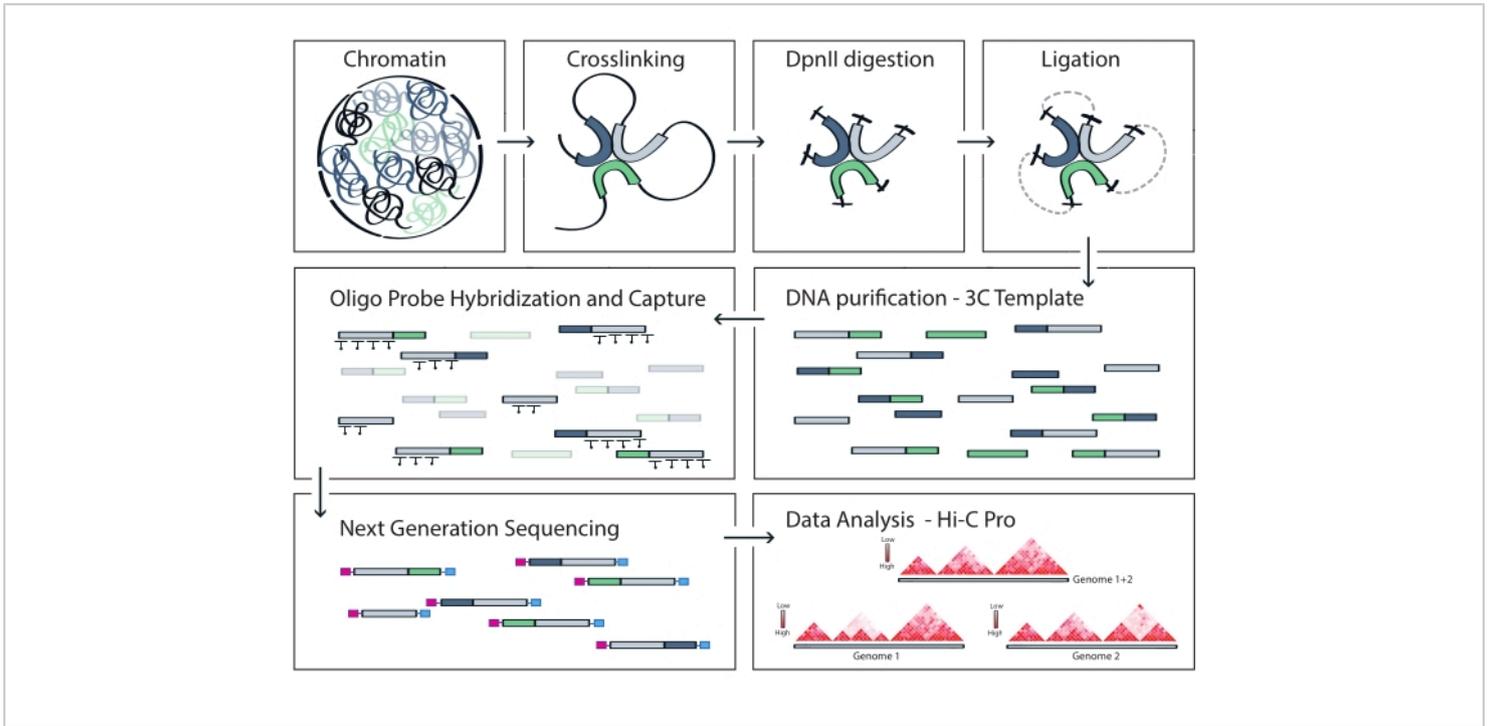


Figure 4: Workflow of Capture Hi-C protocol. Schematic representation of different protocol steps. To generate the genome-wide 3C template, chromatin is first cross-linked with formaldehyde and then digested with the DpnII restriction enzyme. Free DNA ends are then re-ligated, cross-linking is reversed, and DNA is purified. To enrich fragments encompassing the target region, an array of biotinylated RNA probes is hybridized to the 3C template and captured by streptavidin-mediated pull-down. Capture libraries are processed for multiplexed sequencing, and valid ligation fragments are quantified to infer the frequency of chromatin contacts across the target, which are visualized as high-resolution interaction maps. [Please click here to view a larger version of this figure.](#)

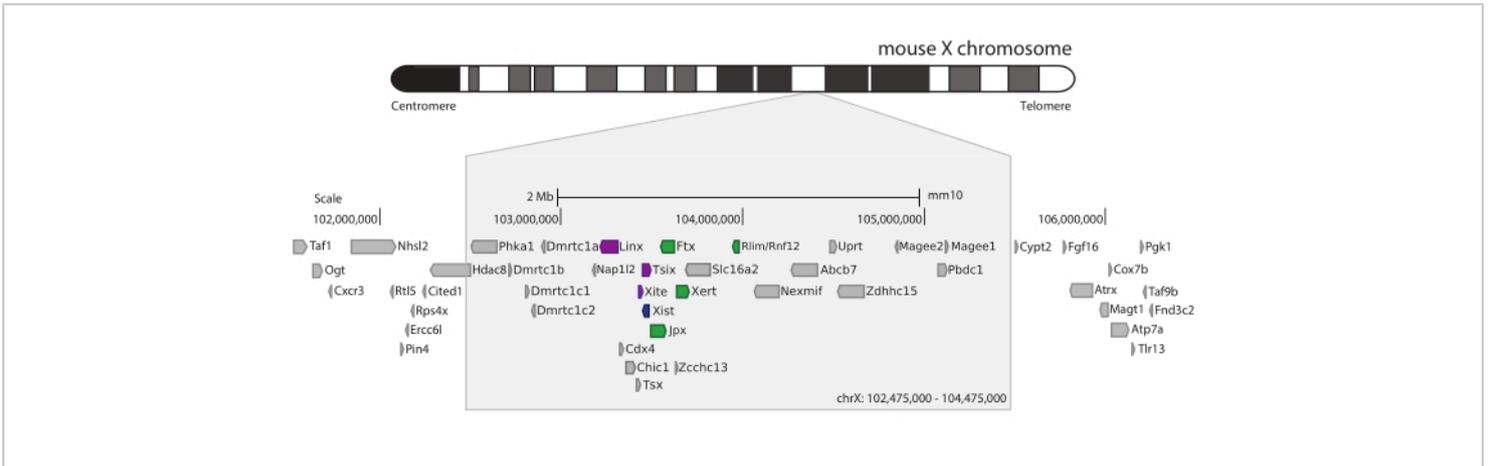


Figure 5: Overview of the region encompassing the *Xic* on the mouse X chromosome. Schematic representation of the mouse X chromosome and zoom in of the 3 Mb captured region (ChrX: 102,475,000-104,475,000). The targeted region includes ~800 kb of DNA corresponding to the *Xic*, the master regulatory locus of XCI. The *Xic* includes the long noncoding genes, *Xist*, a key player of XCI, and its regulatory landscape. Positive regulators of *Xist* are shown in green, and negative regulators in purple. [Please click here to view a larger version of this figure.](#)

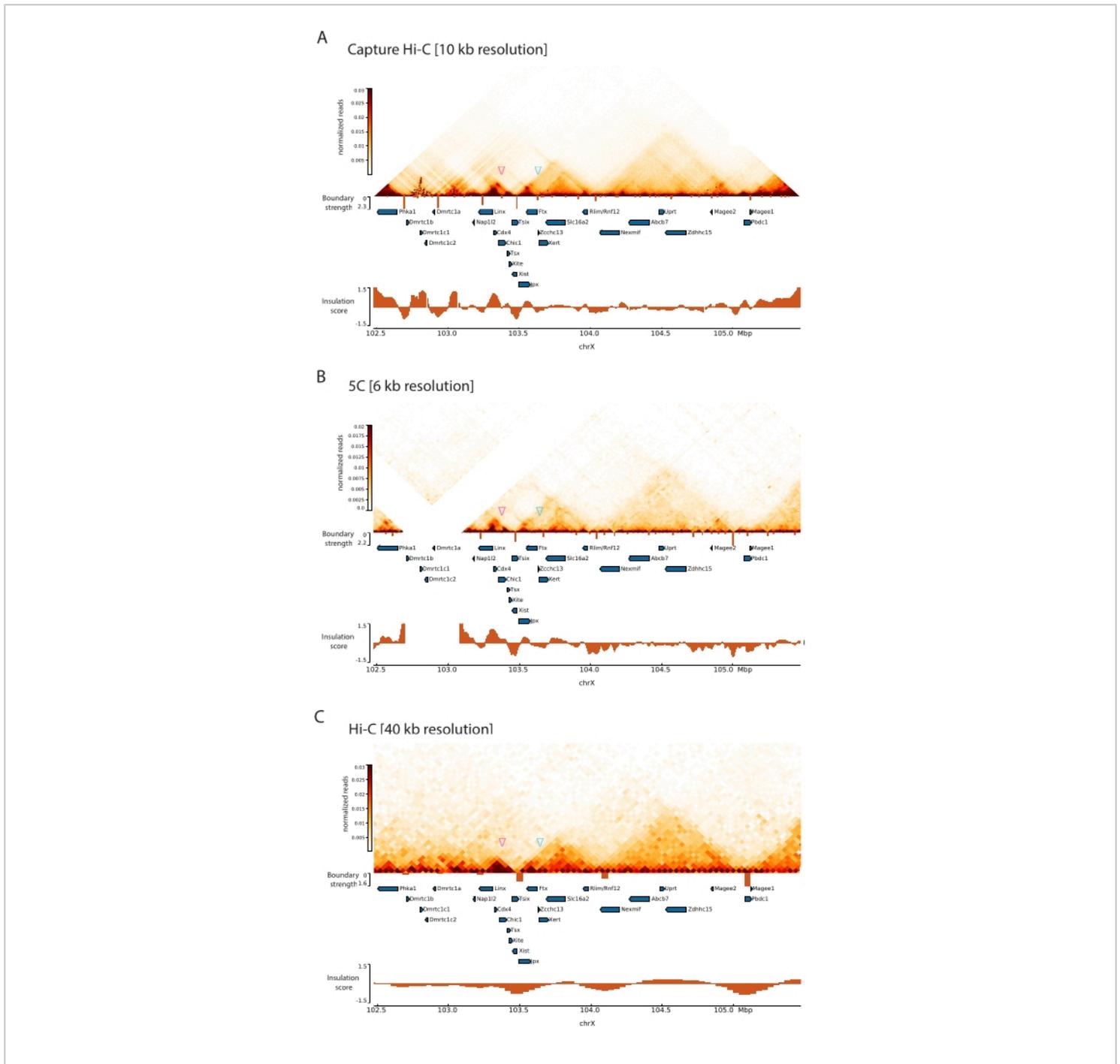


Figure 6: Capture Hi-C, 5C, and Hi-C interaction maps across the 3 Mb captured region. (A) Capture Hi-C interaction map of the 3 Mb target encompassing the mouse *Xic* at 10 kb resolution (this study). **(B)** 5C interaction map of the same target region as in **A** at 6 kb resolution (data reprocessed from⁴⁷). Repetitive regions not included in the analyses are masked in white. The 5C data require their own bioinformatics processing (see⁴⁷). After cleaning and alignment, the 5C maps at the primer resolution are binned using a running median (window = 30 kb, step = 5) to reach a final resolution of 6 kb. **(C)** Hi-C interaction map of the same genomic region as in **A** and **B** at 40 kb resolution (data reprocessed from⁶⁴). All

interaction maps were generated from mouse ESCs. The insulation score was calculated using cooltools and is represented as histograms with insulation minimas at TAD boundaries. TAD boundaries are shown as vertical lines below the map. The height of each line indicates boundary strength. Genes are shown as arrows pointing in the direction of transcription. Sub-TAD boundaries that are detected exclusively or more precisely in Capture Hi-C maps are indicated by magenta and blue arrowheads for sub-TADs in the Tsix and Xist TADs, respectively. [Please click here to view a larger version of this figure.](#)

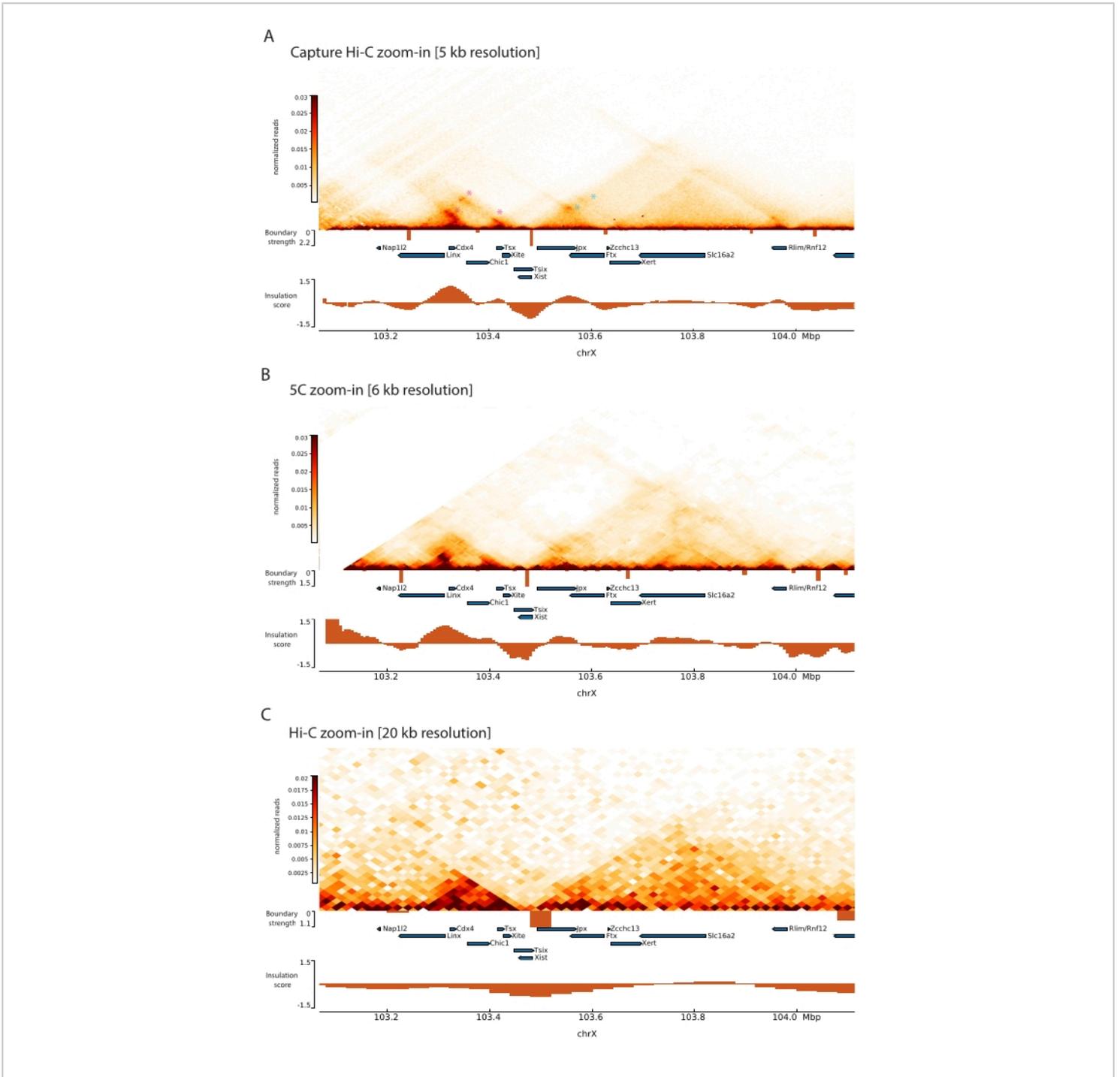


Figure 7: Capture Hi-C, 5C, and Hi-C interaction maps across 1 Mb within the captured region. (A) Capture Hi-C interaction map of the 1 Mb genomic region encompassing the mouse Xic at 5 kb resolution (this study). **(B)** 5C interaction map of the same genomic region as in **A**. at 6 kb resolution (data reprocessed from⁴⁷). Repetitive regions not included in the analyses are masked in white. Of note, the 5C data require their own bioinformatics processing (see⁴⁷). After cleaning and alignment, the 5C maps at the primer resolution are binned using a running median (window = 30 kb, step = 5) to reach a

final resolution of 6 kb. **(C)** Hi-C interaction map of the same genomic region as in **A** and **B** of Hi-C at 20 kb resolution (data reprocessed from⁶⁴). All interaction maps were generated from mESCs. The insulation score was calculated using cooltools and is represented as histograms with insulation minimas at TAD boundaries. TAD boundaries are shown as vertical lines below the map. The height of each line indicates boundary strength. Genes are shown as arrows pointing to the direction of transcription. Contact loops that are detected exclusively or more precisely in Capture Hi-C are indicated by magenta and blue asterisks for loops in the Tsix and Xist TADs, respectively. [Please click here to view a larger version of this figure.](#)

Supplementary Table 1: Post-sequencing statistics for the datasets used in this manuscript: Capture Hi-C (this study), Hi-C⁶⁴, and 5C⁴⁷. [Please click here to download this File.](#)

Discussion

Here we describe a relatively quick and easy Capture Hi-C protocol to characterize the higher-order organization of megabase-sized genomic regions at 5-10 kb resolution. Capture Hi-C belongs to the family of Capture-C technologies that are designed to enrich targeted chromatin interactions from genome-wide 3C or Hi-C templates. To date, the large majority of Capture-C applications have been exploited to map chromatin contacts of relatively small regulatory elements scattered across the entire genome. In the first Capture-C protocol, multiple overlapping RNA biotinylated probes were used to capture >400 pre-selected promoters in 3C libraries prepared from erythroid cells³¹. The same strategy was subsequently improved in Next Generation (NG) and Nuclear Titrated (NuTi) Capture-C to achieve high-resolution interaction profiles of >8,000 promoters by using single 120 bp DNA baits spanning single restriction sites and two sequential rounds of Capture to maximize the enrichment of informative ligation fragments^{32,40}. These strategies led to the functional dissection of cis-acting elements in many different contexts, including mouse embryonic development, cell differentiation,

X-chromosome inactivation, and gene mis-regulation in pathological conditions^{46,63,65,66,67,68,69,70,71}.

In Promoter Capture Hi-C (PCHi-C), >22,000 annotated promoters containing restriction fragments were pulled down from Hi-C libraries by hybridization of single RNA 120-mer biotinylated probes at either or both ends of the restriction fragment^{34,72}. This method allowed dissection of the interactome of thousands of promoters in a rapidly increasing number of cell types, including mouse embryonic stem cells, fetal liver cells, and adipocytes^{34,35,72,73}, but also human lymphoblastoid lines, hematopoietic progenitors, epidermal keratinocytes, and pluripotent cells^{37,74,75,76,77}.

In comparison with these target enrichment technologies, Capture Hi-C targets contiguous genomic regions up to the megabase scale, thereby spanning one or more TADs and encompassing regulatory landscapes of genes. The entire region of interest must be tiled with an array of biotinylated probes encompassing each DpnII restriction site within the target. The hybridization of the biotinylated array to the 3C template, its subsequent streptavidin-based capture, and processing for multiplexed sequencing is performed using a target enrichment system for Illumina Paired-End multiplexed sequencing. The entire protocol is fast, as it can be performed in 1 week from 3C library preparation through to NGS sequencing, and it requires only minor adaptations and/or custom-specific troubleshooting.

The protocol also provides advantages in comparison with other 3C-based methods. To obtain interaction maps at a resolution of 5-10 kb, we sequenced 100-120 M paired-end reads. As a comparison, we used here a Hi-C dataset of 571 M reads to reach a 20 kb resolution⁶⁴ (GSM2053973), and at least 1 billion reads would be required to reach a 5 kb resolution with chromosome-wide Hi-C²².

Capture Hi-C as used in the present study reaches a much higher resolution than the previously published 5C based on a 6-bp cutter restriction enzyme⁴⁷ (**Supplementary Table 1**). Importantly, the strategy designed to enrich and amplify targeted interactions in 5C does not allow for allele-specific analysis of chromatin interactions. On the contrary, Capture Hi-C data can be mapped allele-specifically, allowing the dissection of the 3D structural landscapes of pairs of homologous chromosomes, for example in human cells or in F1 hybrid cell lines derived by crossing genetically different mouse strains⁷⁸. To generate allele-specific Capture Hi-C interaction maps at 5 kb resolution, we sequenced 150 bp paired-end reads to increase SNP coverage. Similar allele-specific approaches can be applied to human cell lines, for which the annotation of SNPs is available²².

Importantly, although Capture Hi-C generally ensures high resolution while improving the affordability of sequencing costs, the production of custom-tailored biotinylated oligonucleotides does have an impact on the overall cost of this method. Therefore, the choice of the most suitable 3C method will differ for different applications, and will depend on the biological question that is being addressed and the resolution required, as well as the size of the region of interest. Other Capture Hi-C protocols developed share key features with the protocol described here. For example, a Capture Hi-C strategy was applied to characterize ~50

kb to 1 Mb genomic regions spanning noncoding variants associated with breast and colorectal cancer risk; in this protocol, target regions were pulled down from Hi-C libraries by hybridizing 120-mer RNA baits tiling the target regions at a 3x coverage^{33,38,79}. Similarly, HYbrid Capture Hi-C (Hi-C²) was used to target interactions within regions of interest up to 2 Mb⁸⁰. In both protocols, the use of a Hi-C template enriched for biotin pulled-down ligation fragments increased the percentage of total informative reads compared to our protocol. For example, in the Hi-C dataset we used here for comparison⁶⁴ (GSM2053973), the percentage of valid pairs after the removal of duplicates is 4.8-fold higher than the valid pairs obtained in Capture Hi-C as described in **Figure 3** and **Supplementary Table 1**. However, the consecutive pull-down of biotinylated ligated fragments and hybridized probes makes the protocol significantly more complex and time consuming while possibly decreasing the complexity of the captured region.

Another available method to enrich 3C templates with tiling probes is Tiled-C, which was applied to study chromatin architecture at high spatial and temporal resolution during mouse erythroid differentiation⁴³. In Tiled-C, a panel of 70 bp biotinylated probes is used to enrich contacts within large-scale regions in two consecutive rounds of capture to generate very high-resolution maps of targeted interactions^{43,81}. The double capture enrichment also makes the protocol longer and more complex when compared to Capture Hi-C. However, different to the Capture-C strategies targeting single restriction sites, in Tiled-C the second round of capture does not seem to significantly increase the capture efficiency, and therefore can probably be omitted⁴³. Finally, a similar tiling approach based on the same target enrichment strategy used in this study was applied to the dissection of regulatory landscapes encompassing structural variants

described in patients with congenital malformations and re-engineered in transgenic mice^{41,42}. In this case, the tiling array of probes was designed across the entire target rather than in the proximity of DpnII restriction sites⁴¹. Nevertheless, this work was seminal in highlighting the sensitivity and power of this strategy to achieve high-resolution characterization of large genomic regions in different contexts^{41,42,48}.

In conclusion, the protocol described here represents an easy, robust, and powerful strategy for the high-resolution 3D characterization of any genomic regions of interest. The application of this approach to different model systems, cell types, developmentally-regulated chromatin landscapes, and gene regulation in healthy and pathological conditions is likely to facilitate our understanding of the interplay and causality between genome topology and gene regulation, one of the fundamental open questions in the epigenetics field. Furthermore, applying Capture Hi-C to map long-range interactions and higher-order chromatin folding of risk variants identified by GWAS studies has the potential to reveal the functional relevance of noncoding genomic loci associated with human diseases in different contexts, thereby providing novel insights into the processes potentially underlying pathogenesis.

Disclosures

Kai Hauschulz is Field Application Scientist at Agilent Technologies - Diagnostic and Genomics Group. All other authors declare no competing interests.

Acknowledgments

Work in the Heard laboratory was supported by a European Research Council Advanced Investigator award (XPRESS - AdG671027). A.L. is supported by a European Union Marie Skłodowska-Curie Actions Individual Fellowship (IF-838408).

A.H. is supported by the ITN Innovative and Interdisciplinary Network ChromDesign, under the Marie Skłodowska-Curie Grant agreement 813327. The authors are thankful to Daniel Ibrahim (MPI for Molecular Genetics, Berlin) for helpful technical advice, to the NGS platform at Institut Curie (Paris), and to Vladimir Benes and the Genomics Core Facility at EMBL (Heidelberg), for support and assistance.

References

1. Ibrahim, D. M., Mundlos, S. The role of 3D chromatin domains in gene regulation: a multi-faceted view on genome organization. *Current Opinion in Genetics & Development*. **61**, 1-8 (2020).
2. Bolt, C. C., Duboule, D. The regulatory landscapes of developmental genes. *Development*. **147** (3), dev171736 (2020).
3. Glaser, J., Mundlos, S. 3D or not 3D: Shaping the genome during development. *Cold Spring Harbor Perspectives in Biology*. **14** (5), a040188 (2021).
4. Denker, A., De Laat, W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes & Development*. **30** (12), 1357-1382 (2016).
5. Kempfer, R., Pombo, A. Methods for mapping 3D chromosome architecture. *Nature Reviews Genetics*. **21** (4), 207-226 (2020).
6. McCord, R. P., Kaplan, N., Giorgetti, L. Chromosome conformation capture and beyond: Toward an integrative view of chromosome structure and function. *Molecular Cell*. **77** (4), 688-708 (2020).
7. Jerkovic, I., Cavalli, G. Understanding 3D genome organization by multidisciplinary methods. *Nature Reviews Molecular Cell Biology*. **22** (8), 511-528 (2021).

8. Hsieh, T. -H. S. et al. Mapping nucleosome resolution chromosome folding in yeast by Micro-C. *Cell*. **162** (1), 108-119 (2015).
9. Krietenstein, N. et al. Ultrastructural details of mammalian chromosome architecture. *Molecular Cell*. **78** (3), 554-565 (2020).
10. Dekker, J., Rippe, K., Dekker, M., Kleckner, N. Capturing chromosome conformation. *Science*. **295** (5558), 1306-1311 (2002).
11. Naumova, N., Smith, E. M., Zhan, Y., Dekker, J. Analysis of long-range chromatin interactions using Chromosome Conformation Capture. *Methods*. **58** (3), 192-203 (2012).
12. Simonis, M. et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature Genetics*. **38** (11), 1348-1354 (2006).
13. Zhao, Z. et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra-and interchromosomal interactions. *Nature Genetics*. **38** (11), 1341-1347 (2006).
14. Würtele, H., Chartrand, P. Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Research*. **14** (5), 477-495 (2006).
15. De Wit, E., De Laat, W. A decade of 3C technologies: insights into nuclear organization. *Genes & Development*. **26** (1), 11-24 (2012).
16. Dostie, J. et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Research*. **16** (10), 1299-1309 (2006).
17. Splinter, E. et al. The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes & Development*. **25** (13), 1371-1383 (2011).
18. Ferraiuolo, M. A., Sanyal, A., Naumova, N., Dekker, J., Dostie, J. From cells to chromatin: capturing snapshots of genome organization with 5C technology. *Methods*. **58** (3), 255-267 (2012).
19. Kim, J. H. et al. 5C-ID: Increased resolution Chromosome-Conformation-Capture-Carbon-Copy with in situ 3C and double alternating primer design. *Methods*. **142**, 39-46 (2018).
20. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. **326** (5950), 289-293 (2009).
21. Zhang, Y. et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell*. **148** (5), 908-921 (2012).
22. Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. **159** (7), 1665-1680 (2014).
23. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. **485** (7398), 376-380 (2012).
24. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. **485** (7398), 381-385 (2012).
25. Krefting, J., Andrade-Navarro, M. A., Ibn-Salem, J. Evolutionary stability of topologically associating

- domains is associated with conserved gene regulation. *BMC Biology*. **16** (1), 87 (2018).
26. Galupa, R., Heard, E. Topologically associating domains in chromosome architecture and gene regulatory landscapes during development, disease, and evolution. *Cold Spring Harbor Symposia on Quantitative Biology*. **82**, 267-278 (2017).
 27. Tena, J. J., Santos-Pereira, J. M. Topologically associating domains and regulatory landscapes in development, evolution and disease. *Frontiers in Cell and Developmental Biology*. **9**, 702787 (2021).
 28. Lupiáñez, D. G., Spielmann, M., Mundlos, S. Breaking TADs: How alterations of chromatin domains result in disease. *Trends in Genetics*. **32** (4), 225-237 (2016).
 29. Davidson, I. F., Peters, J. -M. Genome folding through loop extrusion by SMC complexes. *Nature Reviews Molecular Cell Biology*. **22** (7), 445-464 (2021).
 30. Schmitt, A. D., Hu, M., Ren, B. Genome-wide mapping and analysis of chromosome architecture. *Nature Reviews Molecular Cell Biology*. **17** (12), 743-755 (2016).
 31. Hughes, J. R. et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nature Genetics*. **46** (2), 205-212 (2014).
 32. Davies, J. O. J. et al. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nature Methods*. **13** (1), 74-80 (2016).
 33. Jäger, R. et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nature Communications*. **6**, 6178 (2015).
 34. Schoenfelder, S. et al. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Research*. **25** (4), 582-597 (2015).
 35. Sahlén, P. et al. Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biology*. **16**, 156 (2015).
 36. Joshi, O. et al. Dynamic reorganization of extremely long-range promoter-promoter interactions between two states of pluripotency. *Cell Stem Cell*. **17** (6), 748-757 (2015).
 37. Mifsud, B. et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nature Genetics*. **47** (6), 598-606 (2015).
 38. Dryden, N. H. et al. Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. *Genome Research*. **24** (11), 1854-1868 (2014).
 39. Oudelaar, A. M., Davies, J. O. J., Downes, D. J., Higgs, D. R., Hughes, J. R. Robust detection of chromosomal interactions from small numbers of cells using low-input Capture-C. *Nucleic Acids Research*. **45** (22), e184 (2017).
 40. Oudelaar, A. M. et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nature Genetics*. **50** (12), 1744-1751 (2018).
 41. Franke, M. et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*. **538** (7624), 265-269 (2016).
 42. Despang, A. et al. Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nature Genetics*. **51** (8), 1263-1271 (2019).
 43. Oudelaar, A. M. et al. Dynamics of the 4D genome during in vivo lineage specification and differentiation. *Nature Communications*. **11** (1), 1-12 (2020).

44. Galupa, R., Heard, E. X-chromosome inactivation: A crossroads between chromosome architecture and gene regulation. *Annual Review of Genetics*. **52**, 535-566 (2018).
45. Loda, A., Collombet, S., Heard, E. Gene regulation in time and space during X-chromosome inactivation. *Nature Reviews. Molecular Cell Biology*. **23** (4), 231-249 (2022).
46. van Bommel, J. G. et al. The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nature Genetics*. **51** (6), 1024-1034 (2019).
47. Galupa, R. et al. A conserved noncoding locus regulates random monoallelic Xist expression across a topological boundary. *Molecular Cell*. **77** (2), 352-367 (2020).
48. Gjaltema, R. A. F. et al. Distal and proximal cis-regulatory elements sense X chromosome dosage and developmental state at the Xist locus. *Molecular Cell*. **82** (1), 190-208 (2022).
49. Galupa, R. et al. Inversion of a topological domain leads to restricted changes in its gene expression and affects inter-domain communication. *Development*. **149** (9), dev200568 (2022).
50. Savarese, F., Flahndorfer, K., Jaenisch, R., Busslinger, M., Wutz, A. Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. *Molecular and Cellular Biology*. **26** (19), 7167-7177 (2006).
51. Schulz, E. G. et al. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell*. **14** (2), 203-216 (2014).
52. Gnirke, A. et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology*. **27** (2), 182-189 (2009).
53. Akgol Oksuz, B. et al. Systematic evaluation of chromosome conformation capture assays. *Nature Methods*. **18** (9), 1046-1055 (2021).
54. Piccinini, F., Tesei, A., Arienti, C., Bevilacqua, A. Cell counting and viability assessment of 2D and 3D Cell cultures: Expected reliability of the trypan blue assay. *Biological Procedures Online*. **19** (1), 8 (2017).
55. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology*. **16**, 259 (2015).
56. Abdennur, N., Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics*. **36** (1), 311-316 (2020).
57. Imakaev, M. et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nature Methods*. **9** (10), 999-1003 (2012).
58. Forcato, M. et al. Comparison of computational methods for Hi-C data analysis. *Nature Methods*. **14** (7), 679-685 (2017).
59. Venev, S. et al. *open2c/cooltools: v0.4.1*. (2021).
60. Wages, J. M. *NUCLEIC ACIDS | Immunoassays. Encyclopedia of Analytical Science*. 408-417, Elsevier (2005).
61. Ewels, P. A. et al. The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*. **38** (3), 276-278 (2020).
62. Servant, N., Peltzer, A. nf-core/hic: Initial release of nf-core/hic. *Zenodo*. (2019).

63. Furlan, G. et al. The Ftx noncoding locus controls X chromosome inactivation independently of its RNA products. *Molecular Cell*. **70** (3), 462-472 (2018).
64. Giorgetti, L. et al. Structural organization of the inactive X chromosome in the mouse. *Nature*. **535** (7613), 575-579 (2016).
65. Simon, C. S. et al. Functional characterisation of cis-regulatory elements governing dynamic Eomes expression in the early mouse embryo. *Development*. **144** (7), 1249-1260 (2017).
66. Williams, R. M. et al. Reconstruction of the global neural crest gene regulatory network in vivo. *Developmental Cell*. **51** (2), 255-276.e7 (2019).
67. Godfrey, L. et al. DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. *Nature Communications*. **10** (1), 2803 (2019).
68. Hanssen, L. L. P. et al. Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nature Cell Biology*. **19** (8), 952-961 (2017).
69. Larke, M. S. C. et al. Enhancers predominantly regulate gene expression during differentiation via transcription initiation. *Molecular Cell*. **81** (5), 983-997 (2021).
70. Oudelaar, A. M. et al. A revised model for promoter competition based on multi-way chromatin interactions at the α -globin locus. *Nature Communications*. **10** (1), 5412 (2019).
71. Long, H. K. et al. Loss of extreme long-range enhancers in human neural crest drives a craniofacial disorder. *Cell Stem Cell*. **27** (5), 765-783 (2020).
72. Schoenfelder, S., Javierre, B. -M., Furlan-Magaril, M., Wingett, S. W., Fraser, P. Promoter Capture Hi-C: High-resolution, genome-wide profiling of promoter interactions. *Journal of Visualized Experiments*. (136), e57320 (2018).
73. Siersbæk, R. et al. Dynamic rewiring of promoter-anchored chromatin loops during adipocyte differentiation. *Molecular Cell*. **66** (3), 420-435 (2017).
74. Rubin, A. J. et al. Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. *Nature Genetics*. **49** (10), 1522-1528 (2017).
75. Freire-Pritchett, P. et al. Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells. *eLife*. **6**, e21926 (2017).
76. Javierre, B. M. et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell*. **167** (5), 1369-1384 (2016).
77. Miguel-Escalada, I. et al. Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nature Genetics*. **51** (7), 1137-1148 (2019).
78. Keane, T. M. et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. **477** (7364), 289-294 (2011).
79. Baxter, J. S. et al. Capture Hi-C identifies putative target genes at 33 breast cancer risk loci. *Nature Communications*. **9** (1), 1028 (2018).
80. Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences*. **112** (47), E6456-E6465 (2015).

81. Owens, D. D. G. et al. Dynamic Runx1 chromatin boundaries affect gene expression in hematopoietic development. *Nature Communications*. **13** (1), 773 (2022).