

## Video Article

# Generation of Genome-wide Chromatin Conformation Capture Libraries from Tightly Staged Early *Drosophila* Embryos

Clemens B. Hug<sup>1</sup>, Juan M. Vaquerizas<sup>1</sup><sup>1</sup>Max Planck Institute for Molecular BiomedicineCorrespondence to: Juan M. Vaquerizas at [jmv@mpi-muenster.mpg.de](mailto:jmv@mpi-muenster.mpg.de)URL: <https://www.jove.com/video/57001>DOI: [doi:10.3791/57001](https://doi.org/10.3791/57001)Keywords: Genetics, Issue 140, *Drosophila melanogaster*, fly, chromatin architecture, Hi-C, embryo, topologically associating domain, chromatin conformation capture

Date Published: 10/3/2018

Citation: Hug, C.B., Vaquerizas, J.M. Generation of Genome-wide Chromatin Conformation Capture Libraries from Tightly Staged Early *Drosophila* Embryos. *J. Vis. Exp.* (140), e57001, doi:10.3791/57001 (2018).

## Abstract

Investigating the three-dimensional architecture of chromatin offers invaluable insight into the mechanisms of gene regulation. Here, we describe a protocol for performing the chromatin conformation capture technique *in situ* Hi-C on staged *Drosophila melanogaster* embryo populations. The result is a sequencing library that allows the mapping of all chromatin interactions that occur in the nucleus in a single experiment. Embryo sorting is done manually using a fluorescent stereo microscope and a transgenic fly line containing a nuclear marker. Using this technique, embryo populations from each nuclear division cycle, and with defined cell cycle status, can be obtained with very high purity. The protocol may also be adapted to sort older embryos beyond gastrulation. Sorted embryos are used as inputs for *in situ* Hi-C. All experiments, including sequencing library preparation, can be completed in five days. The protocol has low input requirements and works reliably using 20 blastoderm stage embryos as input material. The end result is a sequencing library for next generation sequencing. After sequencing, the data can be processed into genome-wide chromatin interaction maps that can be analyzed using a wide range of available tools to gain information about topologically associating domain (TAD) structure, chromatin loops, and chromatin compartments during *Drosophila* development.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/57001/>

## Introduction

Chromatin conformation capture (3C) has emerged as an exceptionally useful method to study the topology of chromatin in the nucleus<sup>1</sup>. The 3C variant Hi-C allows measuring the contact frequencies of all chromatin interactions that occur in the nucleus in a single experiment<sup>2</sup>. Application of Hi-C has played an important role in the discovery and characterization of many fundamental principles of chromatin organization, such as TADs, compartments, and loops<sup>3,4,5</sup>.

Studies of chromatin architecture in the context of developmental transitions and cell differentiation are increasingly used to unravel the mechanisms of gene regulation during these processes<sup>6,7,8,9</sup>. One of the model organisms of great interest is *Drosophila melanogaster*, whose development and genome are well characterized. However, few studies that investigate chromatin architecture in *Drosophila* outside of *in vitro* tissue culture settings have been conducted<sup>10,11</sup>. In embryos 16–18 h post fertilization, TADs and compartments reminiscent of similar structures in mammals were identified<sup>10</sup>, which raises the question of which role they are playing in gene regulation during *Drosophila* embryo development. Especially in the early stages of development, prior to gastrulation, such studies are technically challenging. Before gastrulation, *Drosophila* embryos undergo 13 synchronous nuclear divisions that proceed at an extremely rapid pace of 8–60 min per cycle<sup>12,13</sup>. In addition to this, the lack of visual features to distinguish the different stages make it difficult to obtain tightly staged embryo material in sufficient quantities.

In order to develop a protocol that allows studying chromatin architecture in early *Drosophila* development at nuclear cycle resolution, we combined two existing techniques: *in situ* Hi-C, which allows the generation of high resolution whole genome contact maps<sup>5</sup>, and embryo staging using a transgenic *Drosophila* line expressing a eGFP-PCNA transgene<sup>13,14</sup>. This transgene localizes to the nucleus during interphase and disperses throughout the syncytial blastoderm during mitosis. Using this property, it is possible to easily distinguish different stages by their nuclear density and mitotic embryos by the dispersion of the GFP signal.

Together, these techniques enable studying the three-dimensional structure of chromatin in high resolution from as few as 20 *Drosophila* embryos. This protocol includes the instructions for harvesting and sorting *Drosophila* embryos to obtain populations of embryos from a single nuclear division cycle. It further describes how the obtained embryos are used to perform *in situ* Hi-C. The end result is a nucleotide library suitable for sequencing on next generation sequencing machines. The resulting sequencing reads can then be processed into detailed chromatin interaction maps covering the entire *Drosophila* genome.

## Protocol

### 1. *Drosophila* Embryo Collection

NOTE: An equivalent embryo collection can be performed as shown in a previous publication<sup>15</sup>.

1. Transfer young eGFP-PCNA flies (<1 week old) into egg collection cages with yeasted collection plates<sup>16</sup> (1% ethanol, 1% acetic acid, and 4% agar).
2. Move collection cages to an incubator set at 25 °C. Incubation for 1–2 days prior to egg collection improves the egg yield significantly. Change collection plates twice a day.
3. Remove plates containing embryos from the collection chamber in 30–60 min intervals. Smaller intervals result in fewer embryos, but tighter distribution of developmental stages. Collect from multiple cages in parallel so that ideally >200 eggs are laid every 30–60 min.
4. Store the plates at 25 °C until the embryos reach the desired age. For blastoderm stage embryos (nuclear cycle 14), incubate for approximately 2 h.
5. After 2 h of incubation, add tap water from a squirt bottle to the collection plate so that the entire surface is covered with water. Suspend embryos and yeast using a soft brush.
6. Pour resuspended embryos from the collection plate into an embryo collection basket (commercial cell strainers with 100 µm pore size or homemade baskets<sup>17</sup> work well), adding additional tap water from a squirt bottle, if necessary. At this stage, combine embryos from all plates that were collected in parallel. The pooled sample represents a single batch.
7. Wash embryos well by rinsing the basket with tap water from a squirt bottle for 30 s until all yeast residue is washed away.
8. Dechorionate embryos by placing the collection basket into a 2.5% sodium hypochlorite solution in water. Light agitation by swirling facilitates removal of the chorion. Continue until embryos are sufficiently hydrophobic so that they float at the surface of the solution when the basket is lifted out and submerged again, which should take ~1.75–2 min.  
Caution: Sodium hypochlorite is corrosive. Wear appropriate personal protective equipment. Solutions containing <10% sodium hypochlorite can usually be disposed in the sink, make sure to check the regulations of the host institute.
9. Remove the basket from the solution and rinse thoroughly with tap water from a squirt bottle until the bleach smell is no longer noticeable.

### 2. Embryo Fixation

NOTE: Optimal fixation conditions, primarily the concentration of detergent, formaldehyde and the duration of fixation, need to be empirically determined to fit the stage of the embryos. For stages around the syncytial blastoderm, a final concentration of 0.5% Triton X-100 and 1.8% formaldehyde in the aqueous phase work well. For later stages beyond embryo stage 9, further optimization of these parameters may be necessary. All solutions used during fixation and sorting should contain protease inhibitors.

1. Invert collection basket and place it over a 15 mL conical centrifugation tube. Flush embryos from the basket into the tube using a Pasteur pipette dispensing PBS-T (PBS, 0.5% Triton X-100).
2. Let embryos settle at the bottom and adjust the total volume to 2 mL with PBS-T.
3. Add 6 mL of heptane and 100 µL of 37% formaldehyde in water.  
Caution: Heptane and formaldehyde are toxic when inhaled or after skin contact. Wear appropriate personal protective equipment and work in a fume hood. Waste containing heptane or formaldehyde has to be disposed separately according to the host institute's regulations.
4. After the addition of the formaldehyde, start a 15 min timer and vigorously shake the tube up and down for 1 min by hand. The aqueous and organic phase will combine to form a shampoo-like consistency.
5. Agitate on a rotary mixer until 10 min after the addition of formaldehyde.
6. Centrifuge at 500 x g for 1 min at room temperature to collect embryos at the bottom of the tube.
7. Aspirate the entire shampoo-like liquid and discard it, taking care not to aspirate any embryos. Small remaining quantities of shampoo-like supernatant do not cause problems.
8. 15 min after the addition of formaldehyde, resuspend the embryos in 5 mL of PBS-T with 125 mM glycine to quench the formaldehyde. Mix vigorously by shaking up and down for 1 min.
9. Centrifuge at 500 x g at room temperature for 1 min and aspirate supernatant.
10. Wash embryos by resuspending them in 5 mL of ice-cold PBS-T. Let embryos settle and aspirate all supernatant.
11. Repeat the wash in step 2.10 two more times.
12. Keep embryos on ice until sorting. Usually, it is a good idea to collect 3–4 batches of fly embryos before proceeding to sorting. However, embryos should be sorted the same day. Extended storage on ice or in the fridge leads to altered embryo morphology.

### 3. Embryo Sorting

NOTE: Sorting can be done on any fluorescent stereo microscope equipped with a GFP filter at 60–80X magnification.

1. Using a 1,000 µL pipette, transfer a batch of approximately 100 embryos to a small glass vessel suitable for sorting, preferably of a dark color, and place it on ice.
2. **Sort embryos by nuclear density and cell cycle status (Figure 1) by pushing desirable embryos into a separate pile using a needle or syringe tip.**
  1. Remove all embryos with dispersed, non-nuclear distribution of eGFP-PCNA (Figure 1E). Also, embryos that partially show a non-nuclear GFP signal should be removed.

2. To aid in the sorting, assemble a line-up of reference embryos at nuclear cycle 12, 13, and 14 in each batch using the pictures in **Figure 1** as a guide. Use this line-up to match embryos of an unknown stage with one of the reference embryos in order to determine their stage.
3. To verify the developmental stage for reference embryos, measure the nuclear density by imaging the embryo and counting the number of nuclei at the surface of the embryo in an area of  $2,500 \mu\text{m}^2$  using imaging software that provides distance information.  
NOTE: The expected number of nuclei for an area of  $2,500 \mu\text{m}^2$  is 12 to 16 nuclei at nuclear cycle 12, and 20 to 30 nuclei at nuclear cycle 13<sup>13</sup>.
3. Once all embryos at the appropriate stage are separated, take pictures of the embryos for documentation and quality control. If the stereo microscope is not itself equipped with a camera module, any epifluorescence microscope with GFP filters may be used.
4. Pipette up the desired embryos using a 1,000  $\mu\text{L}$  pipette, transfer to a fresh tube, and place on ice.
5. Continue until enough embryos are sorted for the planned experiment. For embryos older than stage 9, generally 20 embryos are sufficient for one *in situ* Hi-C experiment. At nuclear cycle 12, 80 embryos are a good starting point. In earlier cycles, the number of embryos should approximately be doubled for every cycle.
6. Pool and split embryos into 1.5 mL tubes in such a way that one tube contains enough embryos for a single *in situ* Hi-C experiment. It is advisable to use tubes with low DNA binding characteristics, since the same tube will be used for the entire protocol and adsorption of DNA can lead to significant losses at low DNA concentrations.
7. Spin tubes briefly at 100 x g at room temperature and remove supernatant. The embryos should be as dry as possible for freezing.
8. Flash freeze embryos by submerging the tubes in liquid nitrogen and store at  $-80^\circ\text{C}$ .

## 4. In Situ Hi-C

### 1. Lysis

1. Place tubes with frozen embryos on ice.
2. Resuspend embryos in 500  $\mu\text{L}$  of ice-cold lysis buffer (10 mM Tris-Cl pH 8.0, 10 mM NaCl, 0.2% IGEPAL CA-630, protease inhibitors; dissolved in water). Then wait 1 min to let embryos settle at the bottom of the tube.
3. Grind embryos using a metal micro pestle, pre-cooled on ice, that is designed to tightly fit a 1.5 mL microcentrifuge tube.
  1. To avoid agitating the embryos, insert the pestle slowly until it touches the bottom of the tube, push down, and then grind by rotating the pestle twice in both directions.
  2. Lift the pestle very slightly, push to the bottom of the tube again, and repeat grinding.
  3. Repeat 4.1.3.2 10 times, or until the embryos are completely lysed. The solution should be homogenous, and no residual large pieces of embryos should remain.
4. Incubate the homogenized suspension on ice for 15 min. Spin at 1,000 x g,  $4^\circ\text{C}$  for 5 min, and discard supernatant.
5. Wash pellet by resuspending in 500  $\mu\text{L}$  ice-cold lysis buffer, pipetting up and down.
6. Spin again as in 4.1.4, and discard supernatant.
7. Resuspend washed pellet in 100  $\mu\text{L}$  of 0.5% sodium dodecyl sulfate (SDS), pipetting up and down. Permeabilize nuclei by incubating for 10 min at  $65^\circ\text{C}$  in a heating block. Quench SDS by adding 50  $\mu\text{L}$  of 10% Triton X-100 and 120  $\mu\text{L}$  water. Mix by flicking the tube.
8. Incubate at  $37^\circ\text{C}$  for 15 min in heat block.

### 2. Restriction enzyme digestion

1. Add 25  $\mu\text{L}$  of 10x restriction enzyme buffer and 20 U of 5 U/ $\mu\text{L}$  Mbol. Mix by flicking the tube.
2. Digest DNA by incubating for 90 min at  $37^\circ\text{C}$  in heat block under slight agitation (750 rpm).
3. Add another 20 U of Mbol and continue incubation for 90 min.
4. Heat-inactivate Mbol by incubating at  $62^\circ\text{C}$  for 20 min.

### 3. Overhang fill-in

NOTE: Filling in the overhang with biotinylated dATP allows selection of specific ligated fragments. Biotin-dATP at ligation junctions is protected from the exonuclease activity of T4 DNA Polymerase (section 4.6), whereas biotin-dATP at unligated blunt ends is efficiently removed. The pulldown using streptavidin-coated beads in section 4.7 therefore specifically enriches for ligated, chimeric DNA fragments.

1. Add 18  $\mu\text{L}$  of 0.4 mM biotin-14-dATP, 2.25  $\mu\text{L}$  of an unmodified dCTP/dGTP/dTTP mix (3.3 mM each), and 8  $\mu\text{L}$  of 5 U/ $\mu\text{L}$  DNA Polymerase I Klenow Fragment.
2. Mix by flicking the tube and incubate at  $37^\circ\text{C}$  for 90 min in heat block.

### 4. Ligation

1. Add 657  $\mu\text{L}$  of water, 120  $\mu\text{L}$  of 10x T4 DNA Ligase Buffer, 100  $\mu\text{L}$  of 10% Triton X-100, 6  $\mu\text{L}$  of 20 mg/mL bovine serum albumin (BSA), and mix by flicking the tube. Finally add 5  $\mu\text{L}$  of 5 U/ $\mu\text{L}$  T4 DNA Ligase and mix by flicking the tube.
2. Rotate tube gently (20 rpm) at room temperature for 2 h.
3. Add a second installment of 5  $\mu\text{L}$  of 5 U/ $\mu\text{L}$  T4 DNA Ligase and continue rotating for 2 more h.
4. Spin down nuclei at 2,500 x g for 5 min and discard supernatant.

### 5. DNA extraction

1. Resuspend pellet in 500  $\mu\text{L}$  of extraction buffer (50 mM Tris-Cl pH 8.0, 50 mM NaCl, 1 mM Ethylenediaminetetraacetic acid (EDTA), 1% SDS; dissolved in water) and add 20  $\mu\text{L}$  of 20 mg/mL proteinase K. Mix by flicking the tube.
2. Digest protein by incubating at  $55^\circ\text{C}$  for 30 min, shaking at 1,000 rpm.
3. To de-crosslink, add 130  $\mu\text{L}$  of 5 M NaCl and incubate overnight at  $68^\circ\text{C}$ , shaking at 1,000 rpm.
4. Pipette sample into a new 2 mL tube, preferentially with low DNA binding characteristics.
5. Add 0.1x volumes (63  $\mu\text{L}$ ) of 3 M sodium acetate pH 5.2 and 2  $\mu\text{L}$  of 15 mg/mL GlycoBlue. Mix well by inverting. Add 1.6x volumes (1,008  $\mu\text{L}$ ) of pure absolute ethanol and mix by inverting.

6. Incubate at -80 °C for 15 min. Centrifuge at 20,000 x g at 4 °C for at least 30 min. The DNA pellet is often very small, almost invisible, and can only be spotted due to the blue color of GlycoBlue.
  7. Remove supernatant very carefully, moving the pipette tip into the tube along the opposite wall from where the DNA pellet is located. Small remaining droplets are often easily removed during this step and the following washes by pushing them out of the tubes using a P10 tip rather than pipetting them out.
  8. Wash pellet by adding 800 µL of 70% ethanol. Mix by inverting and centrifuge at 20,000 x g at room temperature for 5 min. Repeat this wash at least once.
  9. Remove all traces of ethanol and leave the tube standing with the lid open for up to 5 min to air-dry. Once no liquid is remaining, add 50 µL of 10 mM Tris-Cl pH 8.0. Repeatedly pipette the solution over the area on the wall of the tube where the pellet was located to solubilize the DNA.
  10. Add 1 µL of 20 mg/mL RNase A, mix by flicking the tube, and incubate at 37 °C for 15 min to digest RNA. The sample can now be stored in the fridge overnight or frozen at -20 °C indefinitely.
  11. Check the concentration of DNA using a fluorescent dye based assay according to the manufacturer's instructions. The total amount of DNA in the sample should be at least 10 ng, otherwise too little material is available for amplification and library complexity will likely be low. When this happens, the amount of starting material was probably not sufficient, or material was lost along the way, perhaps during lysis and precipitation.
- 6. Biotin removal and DNA shearing**
1. Add together 12 µL of 10x T4 DNA Polymerase buffer, 3 µL of 1 mM dATP, 3 µL of 1 mM dGTP, and 46 µL of water. Mix by flicking the tube. Add 5 µL of 3 U/mL T4 DNA Polymerase, mix by flicking the tube and incubate at 20 °C for 30 min.
  2. Add 3 µL of 0.5 M EDTA to stop the reaction, and use water to bring the sample to a volume of approximately 120 µL.
  3. Shear the DNA to a size of 200–400 bp using a sonication device according to the manufacturer's instructions. Using the sonicator mentioned in the **Table of Materials**, the following program is appropriate: 2 cycles each of 50 s, 10% duty, intensity 5, 200 cycles/burst.
- 7. Biotin pulldown**
1. Pipette 30 µL of 10 mg/mL streptavidin coated magnetic beads into a new tube, separate them on a magnetic stand, and discard supernatant.
  2. Resuspend beads in 1x B&W buffer (5 mM Tris-Cl pH 7.4, 0.5 mM EDTA, 1 M NaCl; dissolved in water) + 0.1% Triton X-100 and mix by vortexing. Place tube on a magnetic stand and wait for 1–5 min until the beads are separated, depending on the make and model.
  3. Aspirate and discard supernatant while sliding the pipette tip along the wall opposite of where the beads are located. Resuspend beads in 120 µL of 2x B&W buffer (10 mM Tris-Cl pH 7.4, 1 mM EDTA, and 2 M NaCl). Mix by vortexing.
  4. Transfer sheared DNA to a new low DNA binding tube, and mix with 120 µL of the bead suspension in 2X B&W buffer by vortexing. Rotate beads with the DNA sample at 20 rpm for 15 min.
  5. Separate beads on a magnetic stand and discard supernatant.
  6. Resuspend beads in 600 µL of 1x B&W + 0.1% Triton X-100, and incubate at 55 °C for 2 min, shaking at 1,000 rpm. After separation, discard supernatant. Repeat this wash once.
  7. Wash beads once with 600 µL of 10 mM Tris-Cl pH 8.0, and discard supernatant after separation.
  8. Resuspend beads in 50 µL of 10 mM Tris-Cl pH 8.0.

## 5. Sequencing Library Preparation

NOTE: All library steps are done using components from a commercial DNA library preparation kit (see **Table of Materials**). However, alternative kits or other reagents may be substituted. Precipitation tends to form in the library preparation agents during freezer storage. It is therefore important to make sure that all precipitation is dissolved before using the reagents.

- 1. End repair**
  1. Transfer the bead suspension in 50 µL of 10 mM Tris-Cl pH 8.0 into a new PCR tube.
  2. Add 3 µL of End Prep Enzyme Mix and 7 µL of End Prep Reaction Buffer. Mix by pipetting up and down.
  3. Transfer tube to a thermal cycler and run the following program: 20 °C for 30 min, 65 °C for 30 min, and hold at 4 °C.
- 2. Adapter ligation**
  1. Add 30 µL of Ligation Master Mix, 2.5 µL of 1.5 µM Sequencing Adaptor (dilute to 1.5 µM from stock), and 1 µL of Ligation Enhancer to the bead suspension. Mix by pipetting up and down.
  2. Incubate at 20 °C for 15 min in a thermal cycler.
  3. Add 3 µL of USER enzyme. Mix by pipetting up and down.
  4. Incubate at 37 °C for 15 min in a thermal cycler.
  5. Separate beads on a magnetic stand and remove supernatant.
  6. To wash beads, resuspend beads in 100 µL of 1x B&W buffer + 0.1% Triton X-100. Mix by vortexing, and transfer to a new microcentrifuge tube. Separate beads on a magnetic stand and remove supernatant.
  7. Repeat this wash once using 600 µL of the same buffer.
  8. Resuspend beads in 600 µL of 10 mM Tris-Cl pH 8.0, mix by vortexing, and transfer beads to a new tube.
  9. Separate beads on a magnetic stand, discard supernatant, and resuspend beads in 50 µL of 10 mM Tris-Cl pH 8.0.
- 3. PCR amplification**
  1. Prepare two PCR tubes and in each, mix 25 µL of Polymerase Master Mix, 1.5 µL of 10 µM Forward (unindexed) PCR primer, and 1.5 µL of 10 µM Reverse (indexed) PCR primer.  
NOTE: Forward (unindexed) PCR primer:  
5'-AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATC-T-3'.

Reverse (indexed) PCR primer:

5'-CAAGCAGAAGACGGCATACGAGATNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC\*T-3'. \* indicates phosphorothioate bonds and Ns in the indexed PCR primer.

2. In each tube, add 22  $\mu$ L of bead suspension and mix by pipetting up and down.
3. Run PCR using the following program: 98 °C for 1 min, (98 °C for 15 s, 65 °C for 75 s, ramping 1.5 °C/s) repeated 9-12 times, 65 °C for 5 min, and hold at 4 °C.

NOTE: The number of amplification cycles has to be determined empirically. However, we found that libraries that required more than 12 cycles were generally of low complexity and did not result in high quality Hi-C maps. On the other hand, libraries that required less than 12 cycles were not negatively affected by amplifying for a full 12 cycles. Therefore, it is possible to default to 12 cycles of amplification.

4. Pool the two PCR reactions in a single microcentrifuge tube, separate beads on a magnetic stand, and transfer the supernatant containing the library to a new tube.

#### 4. Size selection

1. Bring Ampure XP bead suspension to room temperature and mix well by shaking.
2. Bring volume of the pooled PCR reaction to exactly 200  $\mu$ L with water. During PCR and the magnetic separation, some of the original volume is usually lost. Verify volume by setting the pipette to 200  $\mu$ L and aspirate the entire volume of the reaction. If air is aspirated, more water needs to be added. If the volume exceeds 200  $\mu$ L, adjust the volume of beads added in steps 5.4.3 and 5.4.6 proportionally.  
NOTE: The volumes in parentheses are valid if the total volume of the pooled PCR reactions is exactly 200  $\mu$ L.
3. Add 0.55x volumes (110  $\mu$ L) of Ampure XP bead suspension and mix by pipetting up and down at least 10 times.
4. Incubate at room temperature for 5 min, separate beads on a magnetic stand for 5 min.
5. Move supernatant to a new tube. Discard the tube containing the beads. The beads have bound DNA >700 bp, which is too large to be sequenced.
6. To the supernatant, add 0.2x volumes (40  $\mu$ L, resulting in a total of 0.75x Ampure buffer in the sample) of Ampure XP bead suspension and mix by pipetting up and down 10 times.
7. Incubate at room temperature for 5 min, separate beads on a magnetic stand for 5 min.
8. Discard supernatant which contains DNA <200 bp, which includes free primers, primer dimers, and fragments too small to be sequenced.
9. Leave the tube on the magnetic stand. To wash beads, add 700  $\mu$ L of 80% ethanol, taking care not to disturb the bead pellet, and incubate for 30 s.
10. Discard supernatant, then take the tube off the magnetic stand and resuspend beads in 100  $\mu$ L of 10 mM Tris-Cl pH 8.0. Mix by pipetting up and down 10 times, and incubate at room temperature for 1 min.
11. Add 0.8x volumes (80  $\mu$ L) of Ampure XP bead suspension. Mix by pipetting up and down 10 times and incubate at room temperature for 5 min. This second round of lower bound size selection ensures that the final library is completely free of primers and primer dimers.
12. Separate beads on a magnetic stand for 5 min and discard supernatant.
13. Wash the bead pellet twice with 700  $\mu$ L of 80% ethanol for 30 s each, while leaving the tube on the magnetic stand, as above.
14. With the tube still on the magnetic stand, remove all traces of ethanol. It helps to push droplets of ethanol out of the tube using a P10 pipette. Let residual ethanol evaporate for a maximum of 5 min.
15. Take tube off the magnetic stand and resuspend beads in 50  $\mu$ L of 10 mM Tris-Cl pH 8.0. Mix by pipetting up and down 10 times.
16. Incubate at room temperature for 5 min, then separate beads on a magnetic stand.
17. Transfer supernatant to a fresh tube. This is the final Hi-C library, ready to be quantified and sequenced on next generation sequencing machines, according to the manufacturer's instructions.

## Representative Results

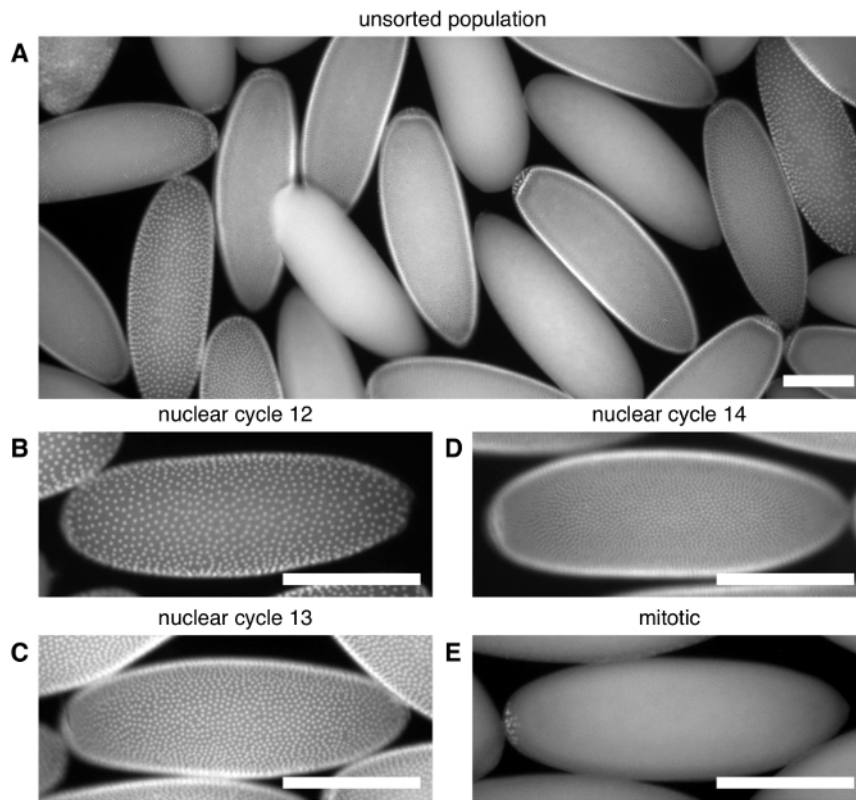
Sorted embryo populations at nuclear cycle 12, 13, and 14 (corresponding to 1:30, 1:45, and 2:10 hours post fertilization, respectively<sup>12</sup>) and 3–4 h post fertilization (hpf) were obtained according to the procedures described in the protocol. By taking pictures of the eGFP-PCNA signal of each sorted embryo batch, it is possible to document the precise stage and cell cycle status of every single embryo that is used in downstream experiments. Example pictures of embryos from sorted populations are shown in **Figure 1B-E**. The output of the *in situ* Hi-C protocol is a nucleotide library ready to be sequenced on next generation sequencing machines. For this purpose, a final library concentration of at least 2–4 nM is usually required. Using the recommended amounts of input material, this concentration is reliably achieved (**Table 1**).

The expected size distribution of DNA fragments after size selection is between 300–600 bp, with a maximum at around 500 bp (**Figure 2A**), depending on the exact shearing and size selection parameters. For sequencing, we recommend paired-end reads of at least 75 bp length to minimize the number of unmappable restriction fragments in the genome. High-resolution maps with 1–2 kb bin size can be obtained from 400 million reads. We recommend sequencing multiple biological replicates at a lower depth of ~150 million reads each, instead of sequencing a single replicate at very high depth. This allows assessment of the biological variation and leads to a lower number of discarded reads due to PCR duplication. For visual representation, the replicates can be combined. Before committing to sequencing a sample at high depth, we recommend running samples using shallow sequencing (a few million reads per sample) to determine basic library quality parameters as in **Figure 2B**.

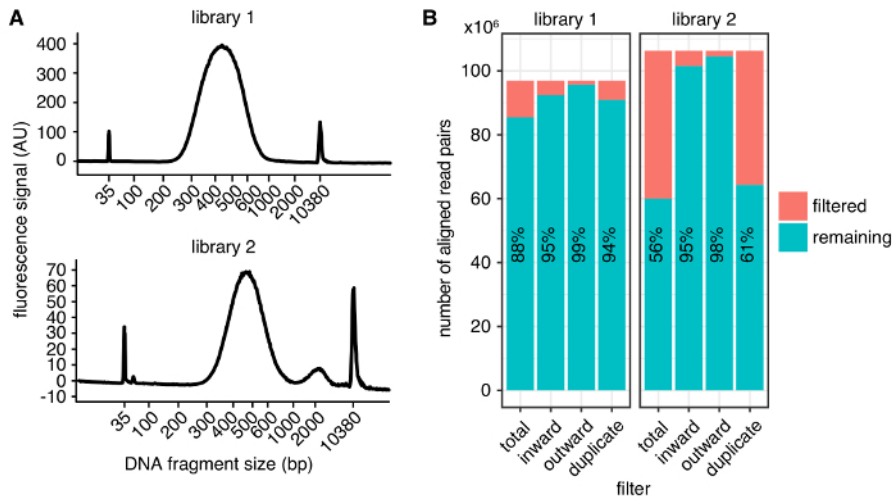
Analysis of Hi-C data requires significant computational resources and bioinformatics expertise. As a rough overview, the paired reads are mapped independently to the reference genome, the resulting alignments are filtered for quality and orientation, then a matrix of contacts at a given bin resolution or fragment level can be generated from the filtered alignments. The contact matrix is the basis for all further downstream analysis exploring TADs, loops, and compartments. For the initial analysis of the sequencing reads, several bioinformatics pipelines are available that enable processing of raw reads into contact matrices without much specialized bioinformatics knowledge<sup>18,19,20,21,22,23</sup>. How further analysis is carried out depends largely on the exact biological question under study and might require significant experience in programming and scripting in R or Python. However, several tools and algorithms to call TADs are available<sup>5,24,25,26,27,28</sup>, as well as software to analyze and explore Hi-C data in the web browser and as stand-alone desktop applications<sup>29,30,31,32</sup>.

Once processed, the quality of the library can be determined using different metrics (**Figure 2B**). First, the rate of PCR duplicates, which is the number of sequenced read pairs arising from the same original molecule, should be as low as possible to limit the amount of wasted sequence reads. However, even libraries with >40% PCR duplication can be processed into high-quality contact maps if the duplicates are filtered. Second, the rate of filtered reads due to their orientation, as described in<sup>4</sup>, should consistently be lower than 10% of aligned read pairs.

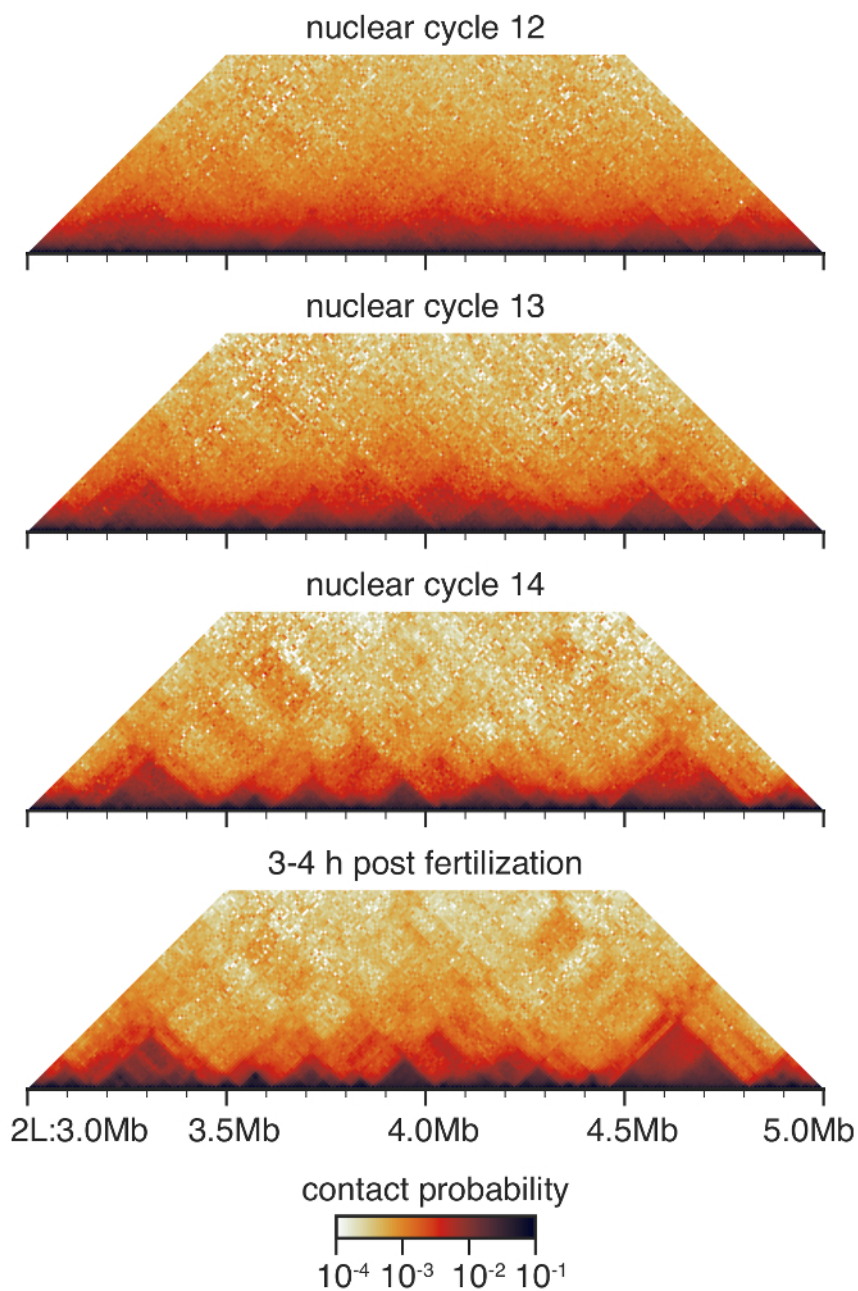
During pre-gastrular development of *Drosophila* between nuclear cycle 12 and 14, the nuclear architecture is drastically remodeled<sup>33</sup> (**Figure 3**). At nuclear cycle 12, few TADs are detected, and the overall distribution of contacts is very smooth without many discernable features. This is dramatically changed at nuclear cycle 13 and 14, when TADs are increasingly prominent and unspecific long-range contacts are depleted.



**Figure 1: Representative pictures of eGFP-PCNA embryos during sorting.** (A) eGFP-PCNA signal from an unsorted population of embryos after 60 min collection and 2 h incubation at 25 °C (B-E) Examples of embryos from sorted populations at nuclear cycle 12 (B), nuclear cycle 13 (C), nuclear cycle 14 (D), and from embryos undergoing synchronous mitosis (E). Scale bars = 200 μm. [Please click here to view a larger version of this figure.](#)



**Figure 2: Examples of *in situ* Hi-C library quality metrics.** (A) Bioanalyzer traces showing the distribution of DNA fragment sizes from a successful Hi-C library (Library 1, top) and from a library that displays a peak of fragments that are too large for sequencing (Library 2, bottom). Library 2 was successfully sequenced, but even larger amounts of undesired DNA fragments may lead to decreased sequencing yields. (B) Filtering statistics of two Hi-C libraries: displayed is the number of aligned read pairs that are excluded from further analysis due to read orientation and distance (inward, outward)<sup>4</sup> or PCR duplication (duplicate). In each bar, the number of reads passing the filter (remaining) and failing (filtered) are plotted. The percentage of reads passing the filter is additionally shown as text. [Please click here to view a larger version of this figure.](#)



**Figure 3: Hi-C interaction maps from staged embryos.** Hi-C interaction maps are binned at 10 kb resolution and balanced as described before<sup>33</sup>. Shown is a region on chromosome 2L. [Please click here to view a larger version of this figure.](#)



Library	Stage	Number of embryos	Amount DNA before shearing (ng)	PCR cycles	Final library concentration (nM)
1	nuclear cycle 12	71	46	12	28.2
2	nuclear cycle 12	46	40	12	22.2
3	nuclear cycle 12	60	13	13	12.3
4	nuclear cycle 13	36	39	12	22.2
5	nuclear cycle 13	35	10	12	5.0
6	nuclear cycle 13	48	18	12	8.7
7	nuclear cycle 14	33	30	12	39.8
8	nuclear cycle 14	24	36	12	20.4
9	nuclear cycle 14	14	8	12	4.2
10	3-4 hpf	17	30	12	24.0
11	3-4 hpf	18	42	11	19.1
12	3-4 hpf	22	63	11	48.4

**Table 1: List of representative sequencing library statistics.** For each library in the list, the number of embryos that were used for its generation, the amount of total DNA before biotin pulldown and shearing measured by Qubit, the number of PCR cycles used for amplification, and the final concentration of the sequencing library after purification and size selection are indicated.

## Discussion

The protocol presented here is very effective at generating high-quality maps of the chromatin architecture in early *Drosophila* embryos. Compared to an earlier protocol<sup>34</sup>, the approach described here uses an up-to-date *in situ* Hi-C procedure<sup>5</sup>, resulting in quicker processing, higher resolution, and less reagent usage. The overall procedure including the *in situ* Hi-C protocol is expected to work on a wide range of stages and experimental systems besides *Drosophila*. Since the protocol has a low input requirement, it could also be used on isolated cell populations. In *Drosophila*, when using the protocol for embryos outside the range described here, some parameters, in particular the fixation of the material, might need to be adjusted. Since older embryos develop a highly impermeable cuticle, raising the concentration of formaldehyde and prolonging fixation may be appropriate. For collection of embryos at stages other than nuclear cycle 14, the incubation times of embryos at 25 °C in step 1.4 need to be adjusted as follows: nuclear cycle 12, 70 min; nuclear cycle 13, 90 min; 3–4 hpf, 3:30 h.

During the 13 cleavage divisions (stage 1-4), the nuclei density roughly doubles with each division. The nuclei can easily be identified by their bright GFP fluorescence. During mitosis, eGFP-PCNA is not located in the nucleus, and its signal is dispersed throughout the embryo. This feature makes identifying embryos that are undergoing a synchronous cleavage division possible. For studying chromatin conformation, these mitotic embryos are usually not desirable, since the mitotic organization of chromatin is drastically different than the interphase organization<sup>35</sup>. It is possible to adapt the protocol to specifically select embryos undergoing a synchronous mitotic division. In this case, only embryos with dispersed, non-nuclear distribution of eGFP-PCNA should be kept, and all other embryos should be discarded. Since the nuclear density cannot be determined, alternative methods to stage embryos by their morphology viewed in transmitted light microscopy must be employed. Presence of pole cells and nuclei at the embryo periphery indicate that the embryo has completed at least nuclear cycle 9, whereas visible cellularization at the periphery indicates nuclear cycle 14<sup>12</sup>.

Hi-C experiments can be successfully performed using a wide selection of restriction enzymes<sup>5</sup>. Current approaches typically use enzymes that recognize either a 4-base sequence, such as MboI, or a 6-base recognition site, such as HindIII. The advantage of 4-base cutters over 6-base cutters is that they offer higher potential resolution, given enough sequencing depth, and a more even coverage of restriction sites across the genome. There is no clear advantage in choosing one 4-base cutter over another<sup>5,23,36,37</sup>. The two most commonly used enzymes, MboI and DpnII, both recognize the same GATC recognition site. DpnII is less sensitive to CpG methylation, which is of no concern in *Drosophila*. The protocol presented here can also be successfully completed using DpnII as a restriction enzyme. In section 4.2. restriction enzyme and buffer have to be adjusted for DpnII compatibility, according to the manufacturer's recommendations.

If the fragment size of the sequencing library deviates significantly from the range shown in **Figure 2A**, cluster formation during sequencing may be less efficient or fail completely. In this case, the size distribution after shearing should be checked and shearing parameters adjusted accordingly. Peaks in the distribution of DNA fragments of very small (<100 bp) or very large (>1,000 bp) sizes indicates problems with size selection, such as carry over of beads or supernatant that are supposed to be discarded. Often these libraries with small peaks at these undesirable sizes, such as the one pictured, are still sequenced successfully with only a minor decrease in clustering efficiency.

High rates of PCR duplication should be avoided because this drastically reduces the number of usable sequence reads. The rate of PCR duplicates is directly related to the amount of input material. Using more input therefore usually alleviates problems with PCR duplication.

Higher numbers of reads filtered due to read orientation (**Figure 2B**) indicate insufficient digestion, which can be the result of using too little enzyme, too much input material, or incomplete homogenization of the embryos.

## Disclosures

The authors have nothing to disclose.

## Acknowledgements

This research was funded by the Max Planck Society. C.B.H. was supported by a fellowship from the International Max Planck Research School – Molecular Biomedicine. We thank Shelby Blythe and Eric Wieschaus for kindly providing the eGFP-PCNA *Drosophila melanogaster* line.

## References

1. Bonev, B., & Cavalli, G. Organization and function of the 3D genome. *Nat Rev Genet.* **17** (11), 661-678 (2016).
2. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* **326** (5950), 289-293 (2009).
3. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* **485** (7398), 376-380 (2012).
4. Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature.* (2013).
5. 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).
6. Darbellay, F., & Duboule, D. Topological Domains, Metagenes, and the Emergence of Pleiotropic Regulations at Hox Loci. *Current topics in developmental biology.* **116**, 299-314 (2016).
7. Beagan, J. A. *et al.* Local Genome Topology Can Exhibit an Incompletely Rewired 3D-Folding State during Somatic Cell Reprogramming. *Cell stem cell.* **18** (5), 611-624 (2016).
8. Andrey, G. *et al.* Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. *Genome Res.* **27** (2), 223-233 (2017).
9. Krijger, P. H. L., & de Laat, W. Regulation of disease-associated gene expression in the 3D genome. *Nature Reviews. Molecular Cell Biology.* **17** (12), 771-782 (2016).
10. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell.* **148** (3), 458-472 (2012).
11. Ghavi-Helm, Y. *et al.* Enhancer loops appear stable during development and are associated with paused polymerase. *Nature.* **512** (7512), 96-100 (2014).
12. Foe, V. E., & Alberts, B. M. Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. *J Cell Sci.* **61**, 31-70 (1983).
13. Blythe, S. A., & Wieschaus, E. F. Zygotic Genome Activation Triggers the DNA Replication Checkpoint at the Midblastula Transition. *Cell.* **160** (6), 1169-1181 (2015).
14. Blythe, S. A., & Wieschaus, E. F. Establishment and maintenance of heritable chromatin structure during early *Drosophila* embryogenesis. *eLife.* **5**, e20148 (2016).
15. JoVE Science Education Database. *Biology I: yeast, Drosophila and C. elegans. Drosophila melanogaster.* Embryo and Larva Harvesting and Preparation. JoVE, Cambridge, MA, (2017).
16. Sicaeros, B., & O'Dowd, D. K. Preparation of Neuronal Cultures from Midgastrula Stage *Drosophila* Embryos. *Journal of Visualized Experiments.* (5) (2007).
17. Shermoen, A. W. Preparation of Baskets for *Drosophila* Egg Collections, Treatments, and Incubations. *Cold Spring Harbor Protocols.* **2008** (10), pdb.ip57-pdb.ip57 (2008).
18. Ay, F., & Noble, W. S. Analysis methods for studying the 3D architecture of the genome. *Genome biology.* **16** (1), 183 (2015).
19. Lazaris, C., Kelly, S., Ntziachristos, P., Aifantis, I., & Tsigirgos, A. HiC-bench: comprehensive and reproducible Hi-C data analysis designed for parameter exploration and benchmarking. *BMC Genomics.* **18** (1) (2017).
20. Servant, N. *et al.* HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology.* **16** (1) (2015).
21. Durand, N. C. *et al.* Juicebox Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell systems.* **3** (1), 95-98 (2016).
22. Lajoie, B. R., Dekker, J., & Kaplan, N. The Hitchhiker's guide to Hi-C analysis: Practical guidelines. *Methods.* **72**, 65-75 (2015).
23. 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).
24. Shin, H. *et al.* TopDom: an efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res.* **44** (7), e70-e70 (2016).
25. Kruse, K., Hug, C. B., Hernández-Rodríguez, B., & Vaquerizas, J. M. TADtool: visual parameter identification for TAD-calling algorithms. *Bioinformatics.* **32** (20), 3190-3192 (2016).
26. Lévy-Leduc, C., Delattre, M., Mary-Huard, T., & Robin, S. Two-dimensional segmentation for analyzing Hi-C data. *Bioinformatics (Oxford, England).* **30** (17), i386-392 (2014).
27. Filippova, D., Patro, R., Duggal, G., & Kingsford, C. Identification of alternative topological domains in chromatin. *Algorithms for molecular biology: AMB.* **9** (1), 14 (2014).
28. Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature.* **523** (7559), 240-244 (2015).
29. Durand, N. C. *et al.* Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell systems.* **3** (1), 99-101 (2016).
30. Zhou, X. *et al.* Exploring long-range genome interactions using the WashU Epigenome Browser. *Nature Methods.* **10** (5), 375-376 (2013).
31. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying genome organization in flies. *bioRxiv.* 115063 (2017).
32. Kerpedjiev, P. *et al.* HiGlass: Web-based Visual Comparison And Exploration Of Genome Interaction Maps. *bioRxiv.* 121889 (2017).
33. Hug, C. B., Grimaldi, A. G., Kruse, K., & Vaquerizas, J. M. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell.* **169** (2), 216-228.e19 (2017).

34. Berkum, N. L. *et al.* Hi-C: a method to study the three-dimensional architecture of genomes. *Journal of Visualized Experiments: JoVE*. (39) (2010).
35. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science*. **342** (6161), 948-953 (2013).
36. Denker, A., & de Laat, W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes & development*. **30** (12), 1357-1382 (2016).
37. Belaghal, H., Dekker, J., & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods (San Diego, Calif.)*. **123**, 56-65 (2017).