

Getting an A with the 3Cs: Chromosome Conformation Capture for Undergraduates

Acadia Joniec^{*1}, Joseph Leszczynski^{*1}, Sokhna Ndoye^{*1}, Jillian Sylvia^{*1}, Steven E. Weicksel^{1,2}

¹ Department of Biological and Biomedical Sciences, Bryant University ² Center of Health and Behavioral Sciences, Bryant University

*These authors contributed equally

Corresponding Author

Steven E. Weicksel
sweicksel@bryant.edu

Citation

Joniec, A., Leszczynski, J., Ndoye, S., Sylvia, J., Weicksel, S.E. Getting an A with the 3Cs: Chromosome Conformation Capture for Undergraduates. *J. Vis. Exp.* (195), e65213, doi:10.3791/65213 (2023).

Date Published

May 12, 2023

DOI

10.3791/65213

URL

jove.com/video/65213

Abstract

Chromosome conformation capture (3C) is a powerful tool that has spawned a family of similar techniques (e.g., Hi-C, 4C, and 5C, referred to here as 3C techniques) that provide detailed information of the three-dimensional organization of chromatin. The 3C techniques have been used in a wide range of studies, from monitoring the changes in chromatin organization in cancer cells to identifying enhancer contacts made with gene promoters. While many of the studies using these techniques are asking big genome-wide questions with intricate sample types (i.e., single-cell analysis), what is often lost is that the 3C techniques are grounded in basic molecular biology methods that are applicable to a broad range of studies. By addressing tightly focused questions of chromatin organization, this cutting-edge technique can be used to enhance the undergraduate research and teaching lab experience. This paper presents a 3C protocol and provides adaptations and points of emphasis for implementation at primarily undergraduate institutions in undergraduate research and teaching experiences.

Introduction

An organism's genome not only holds all the genes required for function but also all the instructions on how and when to use them. This makes regulating access to the genome one of the most important functions of the cell. There are many mechanisms to control gene function; however, at its base level, gene regulation comes down to the ability of regulatory transcription factors (*trans*-factors) to bind to their specific DNA sequences (*cis*-regulatory sequences). This is not an innate ability; instead, it is governed by the organization/

structure of the genome in the nucleus, which controls the availability/exposure of the *cis*-regulatory sequences to the *trans*-factors^{1,2,3}. If the *trans*-factors cannot find their *cis*-regulatory sequences, then the *trans*-factors cannot perform their regulatory tasks. This has made understanding how genomes are organized in the nucleus an important source of inquiry.

It is widely accepted that during interphase, eukaryotic chromosomes in the nucleus occupy their own domain

anchored to the nuclear lamina and nuclear matrix (**Figure 1**), thus making the chromosome more like a slice of pizza, rather than a noodle on a plate of spaghetti. Chromosomes are partially condensed by protein-DNA interactions (chromatin) that twist and loop portions of the chromosome. Through electron microscopy, three-dimensional DNA fluorescence *in situ* hybridization (FISH), and DNA tagging techniques (i.e., fluorescent and artificial DNA methylation), inactive domains of chromatin have been found to be packed tightly along the nuclear periphery^{4,5,6}, while portions of active, less condensed chromatin are found in the interior of the nucleus^{7,8,9,10}. These experiments provide a wide-angle view of chromosome dynamics but do little to capture the changes that occur locally around the gene promoters observed in DNase^{11,12} and nucleosome^{13,14,15} studies.

The key to unlocking higher-resolution chromatin dynamics was the formulation of the 3D chromosome mapping technique, 3C. The 3C technique itself comprises four main steps: crosslinking of chromatin, chromatin digestion by restriction enzymes, chromatin ligation, and DNA purification (**Figure 2**). The new artificial DNA fragments generated by this process can then be characterized to reveal the close physical association between linearly distant pieces of DNA¹⁶. The 3C technique became the basis for the creation of multiple spin-off techniques that utilize the initial steps of 3C to ask broader genome-wide questions (e.g., Hi-C, 4C, ChIP-C). This family of 3C techniques has identified that chromosomes are organized into multiple discrete units termed topologically associated domains (TADs). TADs are encoded in the genome and are defined by chromatin loops flanked by unlooped boundaries^{16,17,18,19}. The TAD boundaries are maintained by two evolutionarily conserved and ubiquitous factors, including CCCT binding factor (CTCF) and cohesion, which prevent loops within separate TADs from

interacting^{16,20}. The loops are mediated by the interaction of *trans*-factors with their regulatory sequences, as well as CTCF and cohesion²¹.

Though many studies using 3C technologies ask broad genome-wide questions and employ complicated sample collection techniques, the formulation of the 3C technique is based on basic molecular biology techniques. This makes 3C intriguing for deployment in both undergraduate research and teaching labs. The 3C technique can be employed for smaller focused questions and is inherently flexible to scaling up or down (single genes²², chromosomes¹⁶, and/or genomes¹⁸) depending on the focus and direction of the questions asked. This technique has also been applied to a wide range of model systems^{7,16,19,23} and has been proven to be versatile in its use. This makes 3C an excellent technique for undergraduates in that students can gain experience in common molecular biology techniques while also gaining valuable experience in answering directed questions.

Presented here is an adapted protocol for 3C library preparation based on previously published protocols^{24,25,26,27}. This protocol has been optimized for approximately 1×10^7 cells, though it has generated 3C libraries with as little as 1×10^5 cells. This protocol has proven to be versatile and has been used to generate 3C libraries from zebrafish embryos, zebrafish cell lines, and young-adult (YA) *Caenorhabditis elegans* (roundworm). The protocol should also be appropriate for mammalian cell lines and, with further adaptation, yeast.

The goal of these adaptations is to make 3C more accessible for undergraduates. Care has been taken to use techniques that are similar to those that can be accomplished in an undergraduate teaching laboratory. The 3C technique provides many learning opportunities for undergraduates to

learn basic molecular biology techniques that will benefit their development at the bench, in the classroom, and in their endeavors after graduation.

Protocol

1. Primer design

NOTE: The 3C primers design tools are available online²⁸. Alternatively, custom primers can be designed by the students (see below).

1. Identification of primer locations
 1. Open the UCSC Genome Browser (<http://genome.ucsc.edu/>), select the organism of study, and search the region of the genome to be assessed using 3C.
 2. In the next window, activate the enzyme track in the **Mapping and Sequencing** tab below the browser by clicking on **Restr Enzymes**.
 3. Enter the **restriction enzyme(s)** to be used, and set the **display mode** to **pack**. Click on **Submit**.
 4. In the **Variation and Repeats**, make sure **RepeatMasker** is set to **dense**.
 5. Using the restriction sites as guides, identify sites of interest not within the masked repeat regions (black bars). Highlight 300 base pairs (bp) flanking both the upstream and downstream sequence surrounding the restriction site by clicking on the **position (topmost track)** and dragging to the desired length (~600 bp).
 6. Release the mouse, and a pop-up box will appear. Take note of the genomic location, click on **Zoom in**, and let the browser readjust to the selection.
 7. Hover the mouse over the **View** tab on the upper browser ribbon, and in the drop-down menu, select **DNA**.
 8. Leave the default option, note the designation of masked sequences (there is an option to make them **N's**). Click on get **DNA**.
 9. The highlighted sequence will then be displayed in the window. Copy and paste this sequence into Primer3 (<https://primer3.ut.ee/>).
 10. Using the Primer3 default settings, generate primers by clicking on **Pick primers**.
 11. In the **results**, take note of the location of the restriction enzyme, and select primers that create a **200-500 bp PCR product** with the **restriction site** located in the **middle**.
 12. Following these steps, design test primers 1 kilobase (kb), 2 kb, 5 kb, 10 kb, and 20 kb from the site(s) of interest.
 13. To design the input control primers, repeat these steps for a site 1-2 kB from the site of interest that lacks a restriction site, meaning it is never cut.
2. Functional primer validation
 1. To validate the primer functionality, set up PCR reactions using titrated primer concentrations and purified genomic DNA. Either through quantitative or semi-quantitative means, determine if the primers create the expected product.
 2. Redesign primers that fail validation.

2. Day 1

NOTE: The protocol can be paused (frozen at $-20\text{ }^{\circ}\text{C}$) after chromatin cross-linking and after the nuclei collection. The steps take, on average, 5-6 h with undergraduates.

1. Young-adult (YA) *C. elegans* nuclei collection (adapted from Han et al.²⁹)

1. Chromatin crosslinking

1. Collect 5,000 YA worms in 30 mL of M9 medium in a 50 mL conical tube, and spin at $400 \times g$ for 2 min at room temperature (RT).
2. Wash the worm pellet 3x more with M9 to remove bacteria.
3. Remove the supernatant, resuspend the worm pellet in 47.3 mL of M9 containing 2.7 mL of 37% formaldehyde (2% final), and incubate with agitation (rocking or nutating) for 30 min at RT.
CAUTION: Handle formaldehyde with care, and work under a fume hood with proper personal protective equipment (PPE-lab coat, proper gloves, and eye protection). Formaldehyde is an irritant that affects the eyes, nose, throat, and lungs.
4. Spin the worms at $400 \times g$ for 2 min at RT.
5. Remove the supernatant, and resuspend the worm pellet in 50 mL of 1 M glycine. Spin the worms at $400 \times g$ for 2 min at RT.

2. Nuclei collection

1. Resuspend the worm pellet in 6 mL of chilled NP buffer (50 mM HEPES at pH 7.5, 40 mM NaCl, 90 mM KCl, 2 mM EDTA, 0.5mM EGTA, 0.1% Tween 20, 0.2 mM DTT, 0.5 mM

spermidine, 0.25 mM spermine, 1x complete protease inhibitor).

2. Transfer the worm suspension to a 7 mL loose-fitting Dounce on ice. Dounce the sample 15x, and hold on ice for 5 min.
3. Transfer the worm suspension to a 7 mL tight-fitting Dounce on ice. Dounce the sample 20x, and hold on ice for 5 min.
4. Transfer the worm suspension to a clean 15 mL conical tube, and add NP buffer to 10 mL in total (approximately 4 mL).
5. Vortex the worm suspension on a high setting for 30 s. Incubate the sample on ice for 5 min.
6. Repeat the previous step.
7. Spin the worm suspension at $100 \times g$ for 5 min at $4\text{ }^{\circ}\text{C}$.
8. Transfer the supernatant to a fresh 15 mL conical tube.
9. Check the sample for worm debris. Visualize 10 μL of the sample with a light microscope. If worm debris is present, spin the sample at $2,000 \times g$ for 5 min at $4\text{ }^{\circ}\text{C}$, resuspend the pellet in a fresh 10 mL of NP buffer, and spin the sample again at $100 \times g$ for 5 min at $4\text{ }^{\circ}\text{C}$. Discard the pellet, check the supernatant for worm debris, and repeat until the sample is free of worm debris.
NOTE: Alternatively, the worm debris can also be removed by straining the sample through a 40 μm cell strainer (6x) followed by a 20 μm cell strainer (6x).

10. If the sample is clear of worm debris, take 5 μL of the sample, add 5 μL of methyl green pyronin (the nuclei will be blue), and count the nuclei with a hemocytometer.
11. Spin the sample at $2,000 \times g$ for 5 min at 4 $^{\circ}\text{C}$
12. Remove the supernatant, place the samples on ice, and proceed; alternatively, snap-freeze the nuclei, and store at -80°C .

2. Chromatin digestion

1. Resuspend the nuclei in 450 μL of clean water. Transfer the nuclei to a clean 1.5 mL microcentrifuge tube.
2. To the cross-linked sample, add 60 μL of 10x DpnII restriction enzyme buffer, and mix well.

NOTE: Other restriction enzymes that still cut crosslinked DNA can be used.

3. Add 15 μL of 10% sodium dodecyl sulfate (SDS) to permeabilize the nuclei, and incubate with agitation (rocking or nutating) at 37 $^{\circ}\text{C}$ for 1 h.

CAUTION: SDS is an irritant and is toxic if ingested or absorbed through the skin. Handle with care, and use proper PPE (lab coat, proper gloves, and eye protection).

4. Quench the SDS by adding 75 μL of 20% Triton X-100 and incubating with agitation at 37 $^{\circ}\text{C}$ for 1 h.
5. Take 10 μL from the sample as an undigested control. Store at 4 $^{\circ}\text{C}$.
6. Add 400 U of DpnII, and incubate at 37 $^{\circ}\text{C}$ overnight with agitation.

3. Day 2

NOTE: On average, it takes undergraduates 5 h to complete these steps.

1. Chromatin digestion

1. Add an additional 200 U of DpnII, and incubate the samples for 4 h at 37 $^{\circ}\text{C}$ with agitation to ensure the complete digestion of the crosslinked sample.
2. Take a 10 μL aliquot from the samples as a digestion control.
3. Heat-inactivate the restriction enzyme by incubating the sample for 20 min at 65 $^{\circ}\text{C}$ (or according to the manufacturer's instructions). Continue to step 2.1.3.
4. If the enzyme cannot be inactivated by heat, add 80 μL of 10% SDS, and incubate the sample for 30 min at 65 $^{\circ}\text{C}$. Then, add 375 μL of 20% Triton X-100, and mix by swirling. Incubate the sample for 1 h at 37 $^{\circ}\text{C}$, and continue to step 2.2.

2. Chromatin ligation

1. Transfer the sample to a clean 50 mL conical tube, adjust the sample volume to 5.7 mL with molecular-grade H_2O , and mix by swirling.
2. Add 700 μL of 10x T4 Ligase buffer, and mix by swirling.
3. Add 60 U of T4 DNA Ligase, and mix by swirling.
4. Incubate overnight at 16 $^{\circ}\text{C}$.

4. Day 3

NOTE: On average, it takes undergraduates 15-30 min to complete these steps. After the overnight incubation, the samples can be frozen.

1. Protein digestion and reverse cross-linking

1. Add 30 μL of proteinase K (10 mg/mL) to the 3C sample, and incubate at 65 $^{\circ}\text{C}$ overnight with agitation. Add 5 μL of proteinase K (10 mg/mL) to the undigested and digestion control, and incubate at 65 $^{\circ}\text{C}$ overnight with agitation.

5. Day 4

NOTE: On average, it takes undergraduates 4-5 h to complete these steps.

1. Purification of the 3C library

1. Add 30 μL of RNase A (10 mg/mL) to the 3C sample, and swirl to mix. Incubate the sample for 45 min at 37 $^{\circ}\text{C}$.
2. Add 7 mL of phenol-chloroform to the sample, and mix by shaking.

CAUTION: Phenol-chloroform is a skin and eye irritant and can cause burns if contact is not treated. Phenol-chloroform should be used in a fume hood with proper eye protection and gloves (nitrile).

3. Centrifuge the sample for 15 min at $3,270 \times g$ at RT. Collect the aqueous phase, and transfer to a clean 50 mL conical tube. Add equal volumes of chloroform, and mix the sample by shaking.
4. Centrifuge the sample for 15 min at $3,270 \times g$ at RT. Collect the aqueous phase, and transfer to a clean 50 mL conical tube. Add 7.5 mL of molecular-grade H_2O , 35 mL of 100% ethanol, and (optional) 7 μL of glycogen (1 mg/mL). Mix by shaking, and incubate at -80°C until the sample is frozen.

NOTE: It can take 1 h or more for the sample to freeze. The protocol can be paused here.

1. While the 3C sample is freezing, purify the control samples. To the control samples, adjust the volume to 500 μL using molecular-grade water, and add 2 μL of RNase A (10 mg/mL) mix by flicking the tube.
 2. Briefly spin to collect the sample at the bottom of the tube. Incubate the samples for 45 min at 37 $^{\circ}\text{C}$.
 3. To the control samples, add 1 mL of phenol-chloroform, and mix by shaking the tubes. Centrifuge the sample for 15 min at $3,270 \times g$ at RT. Collect the aqueous phase of the controls, and place in a clean 1.5 mL microcentrifuge tube.
 4. Add equal volumes of chloroform, and mix the sample by shaking. Centrifuge the sample for 15 min at $3,270 \times g$ at RT. Collect the aqueous phase of the controls, and place in a clean 1.5 mL microcentrifuge tube.
 5. Add 1 mL of 100% ethanol and 2 μL of glycogen (1 mg/mL). Mix by shaking, and incubate at -80°C for 30 min.
 6. Centrifuge the sample for 15 min at $3,270 \times g$ at 4 $^{\circ}\text{C}$. Remove the supernatant, and add 750 μL of chilled 70% ethanol.
 7. Centrifuge the sample for 10 min at $3,270 \times g$ at 4 $^{\circ}\text{C}$. Remove the supernatant, and air-dry the samples. Resuspend the pellet in 50 μL of molecular-grade H_2O . Freeze, or continue to step 6.
5. Centrifuge the sample for 60 min at $3,270 \times g$ at 4 $^{\circ}\text{C}$. Remove the supernatant, and add 10 mL of chilled

70% ethanol. Disrupt and break up the DNA pellet by shaking the sample to mix.

6. Centrifuge the sample for 30 min at $3,270 \times g$ at 4 °C. Remove the supernatant, and allow the sample to partially air-dry at RT. Resuspend the pellet in 150 μL of 10 mM Tris-HCl (pH 7.5) by pipetting up and down. This is the "3C Library."

NOTE: Freeze the sample, or continue to the analysis.

6. Day 5

NOTE: On average, it takes undergraduates 1-2 h to complete these steps.

1. Determining the sample quality using the standard curve of the control primer
 1. Quantify the DNA concentration for the 3C sample as well as all the controls, and adjust samples to 30 $\mu\text{g}/\mu\text{L}$ (if the concentrations permit). Serially dilute samples two-fold four times, resulting in five dilutions for each sample (1x, 0.5x, 0.25x, 0.125x, 0.0625x).

NOTE: If quantification cannot be easily done, serially dilute the samples as indicated above, and proceed.
 2. Set up PCR reactions for the 3C, genomic control, and digested control for each diluted sample with the control primers: 1 μL of DNA, 10 μL of 5x reaction buffer, 1 μL of 10 mM dNTPs, 1 μL of 10 mM control primer (forward and reverse mixed), 1 μL of Taq polymerase, and 36 μL of water.
 3. Following the instructions for the software specific to the PCR machine, set up the standard curve PCR program using cycling conditions as follows: 30 s at

98 °C; 30 cycles of 5 s at 98 °C, 5 s at 60 °C, and 10 s at 72 °C; 1 min at 72 °C; 4 °C hold.

4. Using the software, generate the standard curves for the samples. After the software generates the curve, take note of the PCR efficiency and R^2 value.
2. Determining the DNA concentration
 1. To determine the DNA concentration of the 3C samples, compare the samples to a genomic DNA sample of known concentration. Dilute the samples to 30 $\text{ng}/\mu\text{L}$, and serially dilute the genomic control from 10 $\text{ng}/\mu\text{L}$ to 0.01 $\text{ng}/\mu\text{L}$ in two-fold steps to create a standard curve.

NOTE: If quantification cannot be easily done, dilute the sample 1:10, and proceed.
 2. Set up the PCR reactions for the 3C, genomic control, and digested control for each diluted sample with the control primers: 1 μL of DNA, 10 μL of 5x reaction buffer, 1 μL of 10 mM dNTPs, 1 μL of 10 mM control primer (forward and reverse mixed), 1 μL of Taq polymerase, and 36 μL of water.
 3. Following the instructions for the software specific for the PCR machine, set up the standard curve PCR program using the cycling conditions as follows: 30 s at 98 °C; 30 cycles of 5 s at 98 °C, 5 s at 60 °C, and 10 s at 72 °C; 1 min at 72 °C; 4 °C hold.
 4. Using the software, generate the standard curves for the samples, plotting the 3C samples on the curve. Compare the position of the 3C sample abundance to the standard curve created by the reactions of the genomic DNA of known concentration, and adjust the 3C samples to 30 $\text{ng}/\mu\text{L}$.
3. Determining the presence or absence of chromatin interaction

1. Set up qPCR reactions with 30 ng of the 3C sample, genomic control sample, and digested control sample: 1 μL of DNA, 10 μL of 5x reaction buffer, 1 μL of 10 mM dNTPs, 1 μL of 10 mM primer (forward and reverse mixed), 1 μL of Taq polymerase, and 36 μL of water.

NOTE: The reactions should be set up using the control primer, the test primers for individual genomic loci, and desired combinations of the test primers (site "a" forward and site "b" reverse) being used to determine the chromatin interaction (**Figure 3**).

2. Following the instructions for the software of the PCR machine, set up the PCR program as follows: 30 s at 98 °C; 40 cycles of 5 s at 98 °C, 5 s at 60 °C, and 10 s at 72 °C; 1 min at 72 °C; 4 °C hold.
3. After the PCR has been run, inspect the amplification plot. Ensure that the PCR reactions display exponential amplification, doubling every cycle.
NOTE: Reactions that do not display exponential amplification cannot be analyzed further.
4. Following the instructions for the software of the PCR machine, define the threshold of the qPCR experiment.
5. Export the Ct values for all the samples.
6. Determine the digestion efficiency using the test primer (TP) and control primer (CP) Ct values from the undigested control (UC) and the digested control (DC) samples using equation (1).

$$\text{Percent digested} = 100 - \frac{100}{2^{((CtTP-DC - CtCP-DC) - (CtTP-UC - CtCP-UC))}} \quad (1)$$

7. These values should be in the range of 80%-90% digestion. Record these values (**Figure 4A**).
8. Determine the relative chromatin interaction using the test primer combination (TPC) and control primer (CP) Ct values from the undigested control (UC) and the 3C (3C) samples using equation (2).

$$\text{Chromatin interaction} = 100 - \frac{100}{2^{((CtTPC-3C - CtCP-3C) - (CtTPC-UC - CtCP-UC))}} \quad (2)$$

9. Plot the values for each sample on a bar graph for each test primer combination.
10. Compare the signal of the 3C samples to the control samples to determine if the 3C sample has enrichment of a particular chromatin interaction over the control samples. 3C samples that show enrichment over the control can be considered conditionally positive and require validation using Sanger sequencing (**Figure 4B**).

NOTE: When analyzing the graph data, it is important to remember that unless a "control template" is used (see discussion), the abundance between reactions (i.e., from one set of primers to the next) cannot be compared.

4. Identification of the 3C products
 1. Run the PCR products on a 1.5% agarose gel.
 2. Visualize the gel for the correct size of the PCR products using a UV light box or gel documentation system.
NOTE: Well-constructed 3C libraries will contain a diverse range of DNA fragments, and despite previous validation of the primers for specificity, PCR reactions from 3C libraries have the potential to

contain many bands (**Figure 5A**). This does not prevent the analysis of the 3C libraries.

- Excise the product bands corresponding to the expected fragment size, and perform gel extraction of the PCR products following the instructions of a commercial kit or the homemade protocol outlined below.

CAUTION: UV light is a known carcinogen, and great care should be taken to limit the time exposed to UV light. UV-resistant eye protection, a specimen shield, gloves, and a lab coat should be worn.

- To construct a homemade purification cartridge, poke a hole in the bottom of a 0.5 mL tube with a needle.
 - Pack a small amount of cotton from a cotton ball into the bottom of the 0.5 mL tube, filling no more than half of the tube.
 - Place the 0.5 mL tube into a 1.5 mL tube; ensure that the smaller tube rests on the lip of the bigger tube, not in the tube.
 - Carefully cut a gel fragment into smaller pieces, and place it in the 0.5 mL tube of the cartridge.
 - Place the assembly in a $-20\text{ }^{\circ}\text{C}$ freezer for 5 min. Spin the assembly for 3 min at $13,000 \times g$ at RT.
 - Keep the 1.5 mL tube with the extracted DNA in buffer, and dispose of the 0.5 mL tube containing the agarose debris.
 - Purified DNA can be sent for Sanger sequencing using forward and reverse primers for the 3C fragment.
- Identification of chromatin contacts using Blat

- Upon the completion of the sequencing, determine if the samples meet the quality control standards as indicated in the sequencing report.
- For sequences passing the QC, open .seq and .ab1 (trace) files in a sequencing editing program such as Another Plasmid Editor (ApE), inspect the .ab1 file for clear base calling peaks, and edit the peaks miscalled in the .seq file.
- Using the edited .seq file, search for the DpnII site. Depending on the sequence result, this should be halfway into the reported sequence.
- To determine if the sequence is the expected target sequence, highlight the DpnII site and an additional 30-50 bp of the upstream sequence corresponding to the forward primer of the genomic loci tested for 3C. Perform a Blat search of this sequence against the target species. Make sure the genomic loci match those of the primer.
- Repeat the step above for the reverse primer sequence, making sure the genomic locus returned matches that of the reverse primer.

Representative Results

This procedure will produce one experimental 3C sample and two control samples (undigested and digested). Using these three samples, qPCR was performed. From these results, the digestion efficiency was calculated (equation 1) and recorded (**Table 1**). From these calculations, it was determined that the 3C sample had an approximately 88% digestion efficiency (average of **Table 1**) across the seven genomic loci tested.

Next, the samples were tested for the presence of long-range chromatin contacts between the different genomic loci using combinations of loci-specific primers (**Table 2**) and qPCR. Using these results, the product abundances relative to a control primer set were calculated (equation 2) and graphed for comparison (**Figure 4**). These data indicated that 8 of the 10 reactions were conditional positives for long-range interactions.

The PCR reactions were then run on an agarose gel. The expected PCR product for the eight conditional positives and one negative reaction were gel-purified and sent for Sanger sequencing. The results for representative positive (blue arrow, blue box) and negative (red arrow, red box) reactions are shown (**Figure 5**).

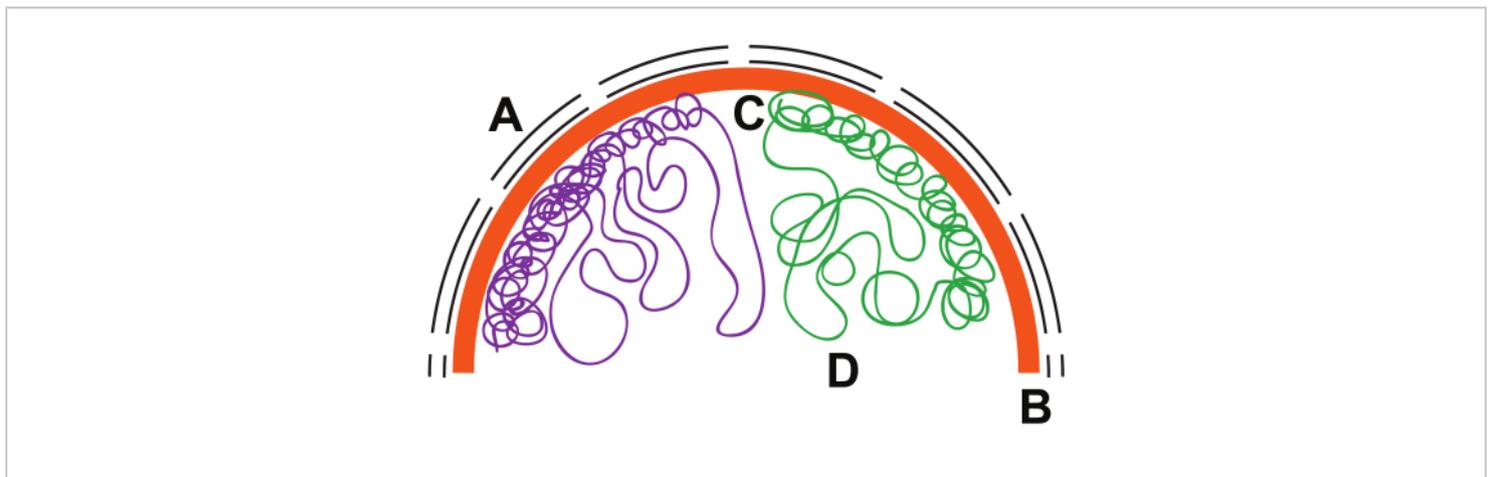


Figure 1: Chromosome structure in the nucleus. Hypothetical chromosome organization inside the nucleus. (A) Nuclear envelope, black lines; (B) nuclear lamina, orange; (C) heterochromatin, compacted lines; (D) euchromatin, loose loops.

[Please click here to view a larger version of this figure.](#)

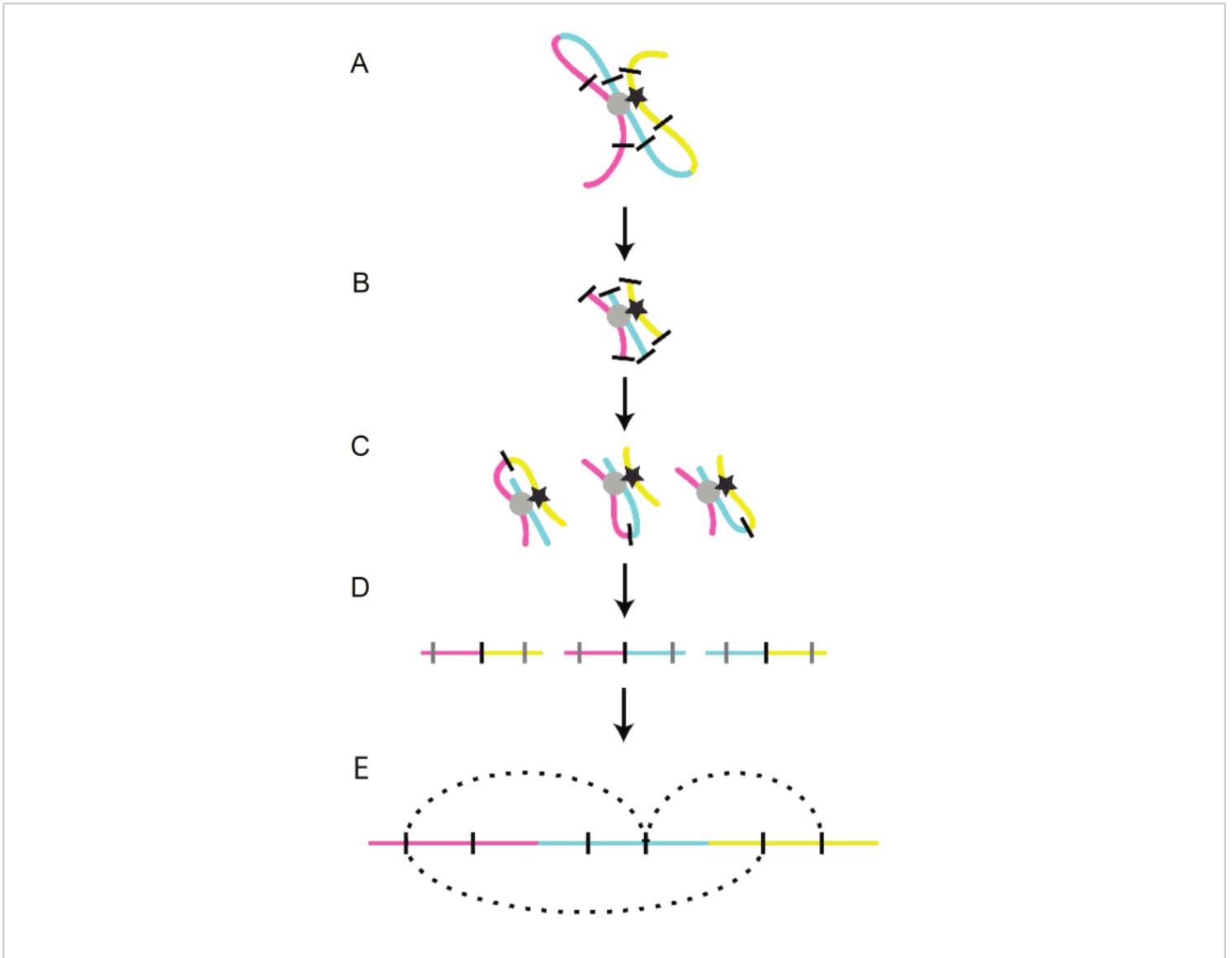


Figure 2: Schematic of the 3C protocol. Distant portions of linear chromosomes (blue, yellow, and pink) are brought close together within the nucleus through regulatory loops. **(A)** The loop structures are mediated by transcription factors (grey circle and black star); these interactions are preserved through chemical crosslinking. **(B)** The loops are broken through enzymatic digestion (black lines). **(C)** Distant chromatin pieces are ligated together through the sticking ends created by the digestion. **(D)** DNA is purified from protein. **(E)** The sequence within the fragments is identified and mapped back to the genome. [Please click here to view a larger version of this figure.](#)

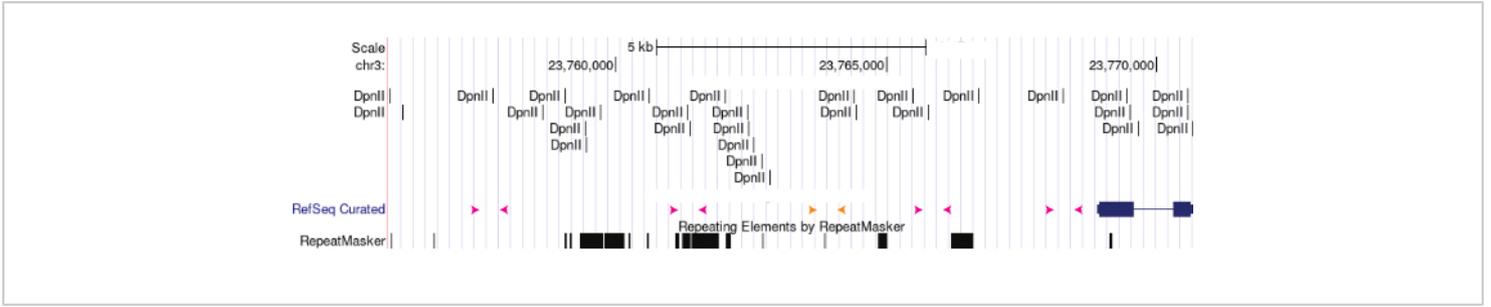


Figure 3: The 3C primer scheme. The 3C primers are designed around restriction sites at varying distances from the genomic location of interest. The pink arrows represent the experimental primer sets surrounding a DpnII site. The experimental primers can be mixed and matched to assess chromatin looping in the region. The orange primers represent a negative control without a DpnII site. [Please click here to view a larger version of this figure.](#)

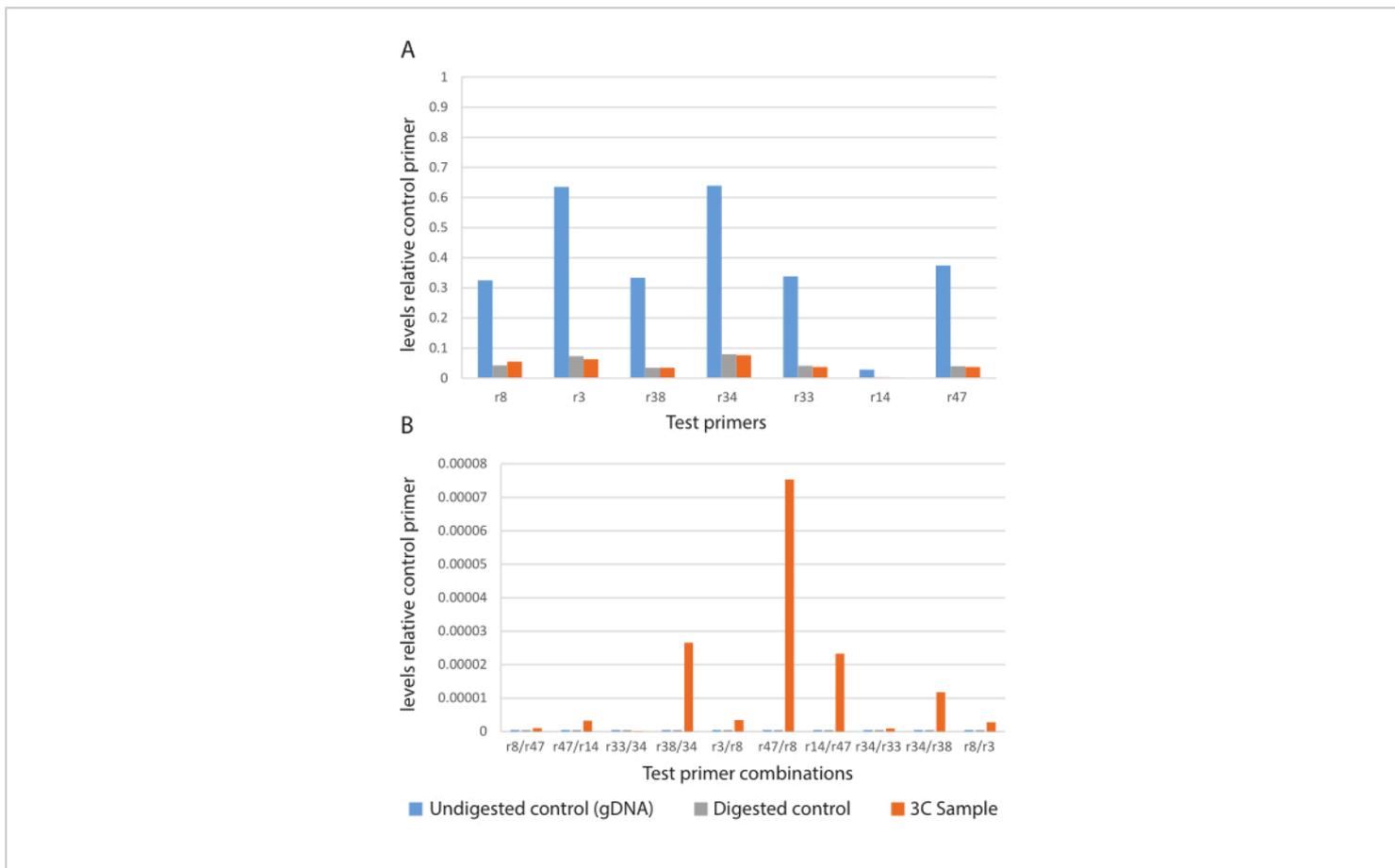


Figure 4: Representative qPCR data for a 3C experiment. Relative abundance of the 3C test primer sets. **(A)** Control graph showing the relative abundance of the product in the undigested control (blue), digested control (grey), and 3C sample (orange). **(B)** The 3C experiment; the relative abundance of the product from combinations of test primers in the undigested control (blue), digested control (grey), and 3C sample (orange). [Please click here to view a larger version of this figure.](#)

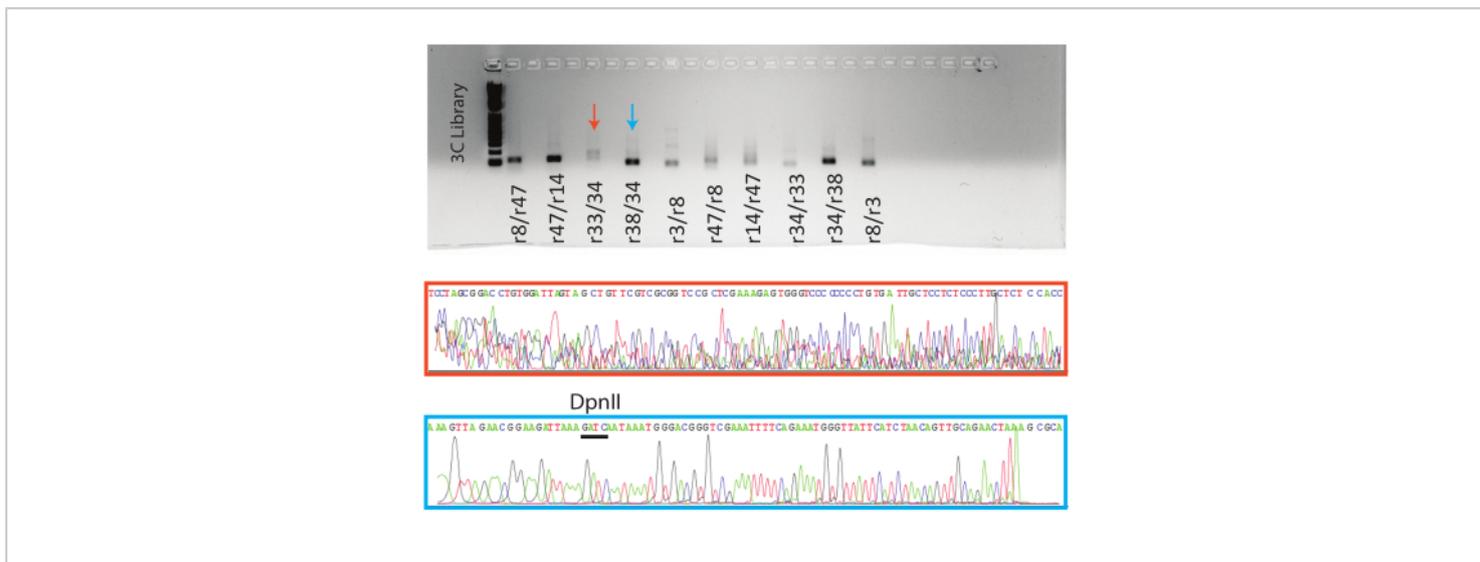


Figure 5: Visualization of the 3C qPCR products. Top, gel with the qPCR endpoint products with the samples indicated. Bottom, representative Sanger sequence trace files for the reactions indicated. The orange sample is an example of a false positive (see **Figure 4, r38/34**), and the blue sample is an example of a true positive with the DpnII site indicated above the trace. [Please click here to view a larger version of this figure.](#)

Site	% digestion
r8	86.98
r3	88.44
r38	89.64
r34	87.55
r33	87.85
r14	86.97
r47	89.45

Table 1: Calculated digestion efficiency.

Name	Sequence	Genomic loci
r3 FWD	ACGCAAGTAAAATTCTGGTTTTTGACC	chrX:11361475
r3 RVS	TTTCCTGAGCTCTAACCATGTTTGC	chrX:11361561
r38 FWD	TTACTTCTGAAGTAATCTTTTCTTATCCCC	chrX:5859700
r38 RVS	AGACGAGCTGATTAAGTAGTTGAGAG	chrX:5859775
r34 FWD	ATTTGTGGATTGCGTGGAGACG	chrX:5429702
r34 RVS	AATAATCCTCTTAACAAACGTGGCC	chrX:5429777
r33 FWD	AAGAGTTGTCCAAAATAAATTGAGCTAAC	chrX:6296704
r33 RVS	TTCAGAAAAGTAACTTTGACTTGAACG	chrX:6296807
r14 FWD	AATTATCGATTTTTCCATCGCGCAG	chrX:8036367
r14 RVS	ATTTCAATGAAAATGTAAAAATGTTCTTC	chrX:8036427
r47 FWD	ATCTAGACTTGATAATATTTGTGTGCCTC	chrX:9464939
r47 RVS	AAGTTCTGCAACTGTTAGATGAATAACAC	chrX:9465064
r8 FWD	GAGAATGTTGTTCTGTAAGTAACTTG	chrX:11094257
r8 RVS	TTACGAAATTTGGTAGTTTTGGACC	chrX:11094362
Control primer FWD	CAATCGTCTCGCTCACTTGTC	chrX:7608049
Control primer RVS	GATGTGAGCAACAAGGCACC	chrX:7608166

Table 2: Primers for the representative 3C experiment.

Discussion

The 3C is a powerful technique that is rooted in basic molecular techniques. It is this foundation of fundamental tools that makes 3C such an intriguing technique to use with undergraduates. With so many recent studies observing chromatin dynamics on such a broad scale, using these results to devise a narrow-focused experiment on a single gene or genomic region has the potential to create a unique and impactful experiment in undergraduate research. Often, experiments like these are considered too advanced for

undergraduates, but with careful planning, they are easily achievable. It is important to note that the assays designed to probe the chromatin connections captured by the 3C library can vary from semi-quantitative endpoint PCR to whole-genome sequencing. In fact, data from the first 3C paper¹⁶ were generated from qPCR. This wide range of assays can all be used because all 3C technologies produce the same product—a library of DNA fragments representing 3D connections in the nucleus.

Presented here is an adaptation of a more flexible and accommodating protocol that is a better fit for undergraduate researchers. The pausing periods listed above imply overnight delays; however, these pauses can extend over weekends and, in the case of the cells and nuclei, for weeks. The most crucial consideration is when the work will get done. Often in protocols, there are time-sensitive steps when pausing is not an option. Outside of a few points (day 1 and day 2), there are many places to stop and freeze the sample. These are critical when working with undergraduates where the schedules and timings of lab work need to be flexible. In addition to engineering these stops into the protocol, undergraduates are encouraged to work in pairs or even small groups of three or four. Groups work well for this protocol, as the students can support each other and create a buddy system so everyone is working safely. Lab work is also more fun with others involved. With groups, students can also work on a variety of questions focused on the organization of chromatin while still performing the same protocol. Thus, even while students are working on separate projects, the protocol links their efforts, and because of this, they can support each other.

Other adaptations are meant to work around the fact that certain specialized tools and equipment are not necessarily found in all undergraduate institutions. These pieces of equipment include but are not limited to qPCR thermocyclers, gel documentation systems, and nano volume spectrometers. Indeed, these pieces of equipment are convenient but are not a requirement. Here, the classic method of 3C is also described in the primer design portion; this involves identifying a genomic locus of interest and, from that, assessing other genomic loci for chromatin contact points further away. This technique also works well if a published dataset is used, such as a dataset using Hi-C, where known

positive (connecting) and negative (non-connecting) loci are identified. Designing experiments using these published data sets is another great adaptation for teaching labs, as the chance for the successful identification of chromatin connections is usually greater. In addition, the research article can be discussed in class and be used as a reference.

This protocol uses a modified qPCR approach to visualize the 3C product formation. Controls are essential to the success of the 3C technique. Each experiment uses both sample controls and primer controls to determine the completion of the 3C procedure. The sample controls include an undigested control (genomic DNA) and digested control. The undigested control determines the baseline signal for the primer sets and is used with the cross-linked digestion control to determine the digestion efficiency. It is expected that there will be a drop in product for any primers directed across a restriction site. Comparing this value to the undigested control provides an indication of how well the sample was digested.

The primers for the PCR include a control primer and test primers. The control primer is a primer set that is near the genomic region being assayed and does not contain a restriction site. This provides the baseline for determining the abundance of the test primer PCR products. Test primers are forward and reverse primers that flank a restriction site for a particular genomic locus of interest (**Figure 3**). Reactions using these primer sets are compared to determine the digestion efficiency, as the product abundance should drop if the restriction site has been cut. In determining the chromatin organization, one test primer from one locus is paired with another test primer from a different genomic locus to determine if these two loci are close together in 3D space. In that case, the expectation is that a PCR product would only be found using the 3C sample as the template.

It is important to note that even validated primers have the tendency to fail (**Figure 4**: r14 primer set). In addition, PCR products are frequently identified in control reactions and in reactions in which a chromatin connection is not predicted (such as the digested control, since it is not ligated). These instances are sequenced and either fail Sanger QC or return without a defined sequence (**Figure 5**). Additionally, traditional 3C experiments generate a "control template," an uncross-linked, digested, and ligated DNA sample, which represents all the possible ligated fragments that can be produced with a given amount of DNA. The "control template" plays an important role in comparing the intensities of qPCR signals between two genomic loci to determine if the signal represents a true interaction or just a random association. Creating a "control template" can be problematic, as a large portion of the chromatin being assayed must be captured in the form of an artificial chromosome and processed along with the 3C samples. Securing such a construct may not be feasible, and creating one could be outside the scope of a semester project. Due to these difficulties, we suggest using a control primer. The control primer does not replace all of the functionality of the "control template" but still provides the opportunity to analyze the data to make "presence" or "absence" determinations.

When performing qPCR, using equal amounts of the sample is important. This should be determined, even if using a nano-spectrophotometer such as a nanodrop, by generating a standard curve from genomic DNA of a known concentration and fitting the 3C samples to that line. These amounts should be recorded and used in subsequent PCRs. The quality of the PCR reaction is also important. As the PCR runs in qPCR, the product abundance is measured using fluorescence and recorded. This recording is accessible in the amplification plot. Once the program has finished, it is important to check

the amplification plot and ensure that the reactions (except for the no template controls) have three phases: a baseline, an exponential, and a plateau/saturation phase. It is important to check that the reactions have an exponential phase, in particular for setting the threshold (see below). Additionally, for serially diluted samples, there should be a shift in Ct values consistent with the dilution of the sample (the highest concentration will have the lowest Ct values, while the lowest concentration will have the highest Ct values). Samples that do not reflect this change in the amplification plot require a new dilution or indicate a larger issue with the 3C sample formation. Finally, while generating the standard curve, the PCR software will compute the PCR efficiency and the R^2 value. The PCR efficiency should be greater than 90%, and the R^2 value should be greater than 0.99. If either of these conditions are not met, it is likely that something is wrong with the sample or the PCR primers.

Post qPCR, the percent digestion and the presence of 3C interactions can be calculated using the qPCR Ct for each reaction. To determine these, first, the threshold for the PCR reaction must be set. This is normally done using the software that comes with the qPCR machine. Setting the threshold will define the concentration of PCR product that will be used to compare the sample Ct values. The threshold should bisect the amplification curves of the PCR reactions in the exponential phase of amplification. Only PCR reactions with exponential amplification can be compared (in this case, the control primers and the test primer reactions), as this is the only way to ensure the reactions are amplifying DNA at the same rate and can be compared faithfully. When analyzing the 3C graphs, conditionally positive reactions are identified as those with more product over the control samples, genomic control, and the digestion control (**Figure 4B**). However, these samples must be further validated using

Sanger sequencing following the gel purification of the PCR product.

After Sanger sequencing, samples that pass the QC can be analyzed using Blat. The goal of this analysis is to determine if the sample has the sequence of both target genomic loci flanking the restriction site (DpnII in the case of this protocol). If both sequences are identified, then the 3C fragment can be considered validated. If the results of the Blat do not return the expected sequence, this may indicate that one or both primers are not optimal, resulting in a false positive qPCR result. The trace files for the false positive samples will have undefined base peaks, and the Seq reports will contain mostly "n" base calls.

Sanger validation is essential, as false positives from artifactual PCR product formation are possible. These false positives can be identified when the sequencing products do not have the expected target sequence or a DpnII site characteristic of a proper 3C fragment (**Figure 5**). The sequencing of the PCR fragments also provides another data point for the experiment and drives home to the students that the 3C technique is identifying distant genomic loci that are coming together in 3D space within the nuclei.

The 3C technique provides a wealth of foundational molecular techniques for undergraduates in a flexible, straightforward procedure. This 3C technique is also a launching point for the other 3C techniques that incorporate next-generation sequencing (NGS). These types of experiments can expose undergraduates to important aspects of bioinformatics and are rooted in the basic principles outlined here. Undergraduate experience and involvement are key to their success and development as young scientists. By providing these opportunities, undergraduates can strengthen their

understanding of basic principles while building their confidence to tackle cutting-edge techniques and questions.

Disclosures

The authors declare no conflicts of interest.

Acknowledgments

This work was supported in part by the Rhode Island Institutional Development Award (IDeA) Network of Biomedical Research Excellence from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103430 and the Bryant Center of Health and Behavioral Sciences.

References

1. McBryant, S. J., Adams, V. H., Hansen, J. C. Chromatin architectural proteins. *Chromosome Research*. **14** (1), 39-51 (2006).
2. Nalabothula, N. et al. The chromatin architectural proteins HMGD1 and H1 bind reciprocally and have opposite effects on chromatin structure and gene regulation. *BMC Genomics*. **15**, 92 (2014).
3. John, S. et al. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nature Genetics*. **43** (3), 264-268 (2011).
4. Fawcett, D. W. On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. *The American Journal of Anatomy*. **119** (1), 129-145 (1966).
5. Reddy, K. L., Zullo, J. M., Bertolino, E., Singh, H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. **452** (7184), 243-247 (2008).

6. Finlan, L. E. et al. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genetics*. **4** (3), e1000039 (2008).
7. Chambeyron, S., Da Silva, N. R., Lawson, K. A., Bickmore, W. A. Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development*. **132** (9), 2215-2223 (2005).
8. Chambeyron, S., Bickmore, W. A. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes and Development*. **18** (10), 1119-1130 (2004).
9. Mahy, N. L., Perry, P. E., Gilchrist, S., Baldock, R. A., Bickmore, W. A. Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. *The Journal of Cell Biology*. **157** (4), 579-589 (2002).
10. Mahy, N. L., Perry, P. E., Bickmore, W. A. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *The Journal of Cell Biology*. **159** (5), 753-763 (2002).
11. Li, B., Carey, M., Workman, J. L. The role of chromatin during transcription. *Cell*. **128** (4), 707-719 (2007).
12. Chen, A., Chen, D., Chen, Y. Advances of DNase-seq for mapping active gene regulatory elements across the genome in animals. *Gene*. **667**, 83-94 (2018).
13. Hughes, A. L., Rando, O. J. Mechanisms underlying nucleosome positioning in vivo. *Annual Review of Biophysics*. **43**, 41-63 (2014).
14. Weiner, A., Hughes, A., Yassour, M., Rando, O. J., Friedman, N. High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. *Genome Research*. **20** (1), 90-100 (2010).
15. Hughes, A. L., Jin, Y., Rando, O. J., Struhl, K. A functional evolutionary approach to identify determinants of nucleosome positioning: A unifying model for establishing the genome-wide pattern. *Molecular Cell*. **48** (1), 5-15 (2012).
16. Dekker, J., Rippe, K., Dekker, M., Kleckner, N. Capturing chromosome conformation. *Science*. **295** (5558), 1306-1311 (2002).
17. van Berkum, N. L., Dekker, J. Determining spatial chromatin organization of large genomic regions using 5C technology. *Methods in Molecular Biology*. **567**, 189-213 (2009).
18. Smith, E. M., Lajoie, B. R., Jain, G., Dekker, J. Invariant TAD boundaries constrain cell-type-specific looping interactions between promoters and distal elements around the CFTR locus. *American Journal of Human Genetics*. **98** (1), 185-201 (2016).
19. Crane, E. E. Two inputs into *C. elegans*. dosage compensation: Chromosome conformation and the miRNA-specific argonaute ALG-2. PhD thesis. University of California, Berkeley (2012).
20. Bell, A. C., West, A. G., Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*. **98** (3), 387-396 (1999).
21. Cuadrado, A. et al. Specific contributions of cohesin-SA1 and cohesin-SA2 to TADs and polycomb domains in embryonic stem cells. *Cell Reports*. **27** (12), 3500-3510.e4 (2019).
22. Sarro, R. et al. Disrupting the three-dimensional regulatory topology of the Pitx1 locus results in overtly normal development. *Development*. **145** (7), dev.158550 (2018).

23. Tolhuis, B. et al. Interactions among polycomb domains are guided by chromosome architecture. *PLoS Genetics*. **7** (3), e1001343 (2011).
24. Splinter, E., de Wit, E., van de Werken, H. J. G., Klous, P., de Laat, W. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. *Methods*. **58** (3), 221-230 (2012).
25. Fernández-Miñán, A., Bessa, J., Tena, J. J., Gómez-Skarmeta, J. L. Chapter 21 - Assay for transposase-accessible chromatin and circularized chromosome conformation capture, two methods to explore the regulatory landscapes of genes in zebrafish. *Methods in Cell Biology*. **135**, 413-430 (2016).
26. Dekker, J. The three "C" s of chromosome conformation capture: controls, controls, controls. *Nature Methods*. **3** (1), 17-21 (2006).
27. Hagège, H. et al. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature Protocols*. **2** (7), 1722-1733 (2007).
28. Lajoie, B. R., van Berkum, N. L., Sanyal, A., Dekker, J. My5C: Webtools for chromosome conformation capture studies. *Nature Methods*. **6** (10), 690-691 (2009).
29. Han, M., Wei, G., McManus, C. E., Hillier, L. W., Reinke, V. Isolated *C. elegans* germ nuclei exhibit distinct genomic profiles of histone modification and gene expression. *BMC Genomics*. **20** (1), 500 (2019).