Split-Pool Recognition of Interactions by Tag Extension (SPRITE) for DNA: Experimental Protocols

Sofia Quinodoz
squinodo@lncrna.caltech.edu
Mitchell Guttman Laboratory
Last edited: August 9, 2018

This document describes the experimental procedures for our SPRITE (Split-Pool Recognition of Interactions by Tag Extension) method to be used for mapping genome-wide higher order interactions between DNA molecules.

Table of Contents:

1 Materials
- 1.1 Equipment
- 1.2 Solutions
- 1.3 Additional Materials and Reagents

2 Cell Culture

3 Adaptor and Barcode Design
- 3.1 DNA Phosphate Modified (DPM) Adaptor
- 3.2 Odd and Even Tags
- 3.3 Terminal Tag
- 3.4 Final Library Amplification
- 3.5 DPM primers for QC of DPM ligation

4 Sample Preparation
- 4.1 Formaldehyde-DSG Crosslinking
- 4.2 Cell Lysis
- 4.3 DNA Fragmentation

5 Library Preparation Pt. 1
- 5.1 NHS Coupling
- 5.2 End Repair
- 5.3 DPM Adaptor Ligation
- 5.4 QC: Check to Determine Ligation Efficiency of the DPM Adaptor

6 SPRITE and Library Preparation Pt. 2
- 6.1 SPRITE
- 6.2 Library Preparation Pt. 2

7 Sequencing and Data Analysis
- 7.1 Computational Pipeline
- 7.2 DNA contact maps

1 Materials

1.1 Solutions

<u>DSG Crosslinking Solution</u>
1 x PBS
2mM DSG in DMSO

<u>Scraping Buffer</u>
1 x PBS pH 7.5
0.5% BSA
Store at 4 °C

<u>Cell Lysis Buffer A</u>
50mM Hepes pH 7.4
1mM EDTA
1mM EGTA
140mM NaCl
0.25% Triton-X
0.5% NP-40
10% Glycerol

<u>Cell Lysis Buffer B</u>
50mM Hepes pH 7.4
1.5mM EDTA
1.5mM EGTA
200mM NaCl

<u>Cell Lysis Buffer C</u>
50mM Hepes pH 7.4
1.5mM EDTA
1.5mM EGTA
100mM NaCl
0.1% DOC
0.5% NLS

<u>10x SPRITE DNase Buffer</u>
200mM Hepes pH 7.4
1M NaCl
0.5% NP-40
5mM CaCl2
25mM MnCl2

<u>25x DNase Stop Solution</u>
250mM EDTA
125mM EGTA

<u>MyRNK Buffer</u>
20mM Tris pH 7.5
100mM NaCl
10mM EDTA
10mM EGTA
0.5% Triton-X
0.2% SDS

<u>Coupling Buffer</u>
1 x PBS
0.1% SDS

<u>Modified RLT Buffer</u>
1 x Buffer RLT supplied by Qiagen
10mM Tris pH 7.5
1mM EDTA
1mM EGTA
0.2% NLS
0.1% Triton-X
0.1% NP-40

<u>SPRITE Wash Buffer</u>
20mM Tris pH 7.5
50mM NaCl
0.2% Triton-X
0.2% NP-40
0.2% DOC

<u>10X Annealing Buffer</u>
100 mM Tris-HCl pH 7.5
2 M LiCl
1.5 mM EDTA
0.5 mM EGTA

Note 1: Modified RLT Buffer contains guanidine thiocyanate which when mixed with bleach produces hydrogen cyanide gas and hydrogen chloride gas. Be careful to ensure that all liquid Modified RLT Buffer waste is disposed of in its own waste container. Solids that have touched Modified RLT Buffer such as tips and reservoirs should also be discarded in a separate solid Modified RLT Buffer container.

Note 2: DTT has a short half-life at pH 7.4 at 20C. It is important to keep PBLSD+ Buffer on ice during the procedure and frozen at -20C if not in use.

1.2 Equipment

Microcentrifuge
Plate Centrifuge
Sonication instrument and chiller
Gel Electrophoresis Equipment
Qubit Fluorometer
Eppendorf Thermomixer
Eppendorf SmartBlock 1.5mL thermoblock
Eppendorf SmartBloack PCR 96 thermoblock
Magnetic rack for 1.5mL tubes (*e.g. Invitrogen DynaMag-2)*
Magnetic rack for 15mL conical tubes
Magnetic rack for 96 well plate
PCR machine
Agilent Bioanalyzer

1.3 Additional Materials and Reagents

Barcodes and adaptors
Low-retention pipette tips
Low-bind 96-well plate
Reservoirs
PCR Strip tubes
Protein Lo-bind 1.5 mL eppendorf tubes
15 mL Falcon conicals
50 mL Falcon conicals
Trypsin Versene Phosphate Buffer (TVP), (1 mM EDTA, 0.025% Trypsin, 1% Sigma Chicken Serum)
PBS Sterile
Disuccinimidyl glutarate (DSG), 50 mg bottle from Pierce, brought to 0.5 M with the addition of DMSO
16% Formaldehyde Solution Ampules from Pierce
2.5 M Glycine
Protease Cocktail Inhibitor
TURBO DNase from ThermoFisher Scientific
Proteinase K from New England Biolabs
RNA Clean and Concentrator-5 Kit with Capped Columns from Zymo Research
Gel Electrophoresis System
Pierce NHS-Activated Magnetic Beads
NEBNext End Repair Enzyme Mix from New England Biolabs
NEBNext End Repair Reaction Buffer from New England Biolabs
Kelnow Fragment (3' to 5' exo-) From New England Biolabs
NEBNext dA-Tailing Reaction Buffer From New England Biolabs
Instant Sticky-end Ligase Master Mix from New England Biolabs
NEBNext Quick Ligation Reaction Buffer
Sigma 1,2-Propanediol #398039-25mL
Q5 Hot Start High-Fidelity 2X Master Mix from New England Biolabs
Agencourt AMPure XP Magnetic Beads from Beckman Coulter
Qubit dsDNA HS Assay Kit
High Sensitivity DNA Kit for the Agilent Bioanalyzer
D1000 Screentape for the Agilent 2200 TapeStation

2 Cell Culture

Materials and Methods

Mouse ES cell culture and Xist induction

All mouse ES cell lines were cultured in serum-free 2i/LIF medium as previously described (*7*, *11, 13*). Female ES cells (F1 2-1 line, generously provided by K. Plath) are an F1 hybrid wild-type mouse ES cell line derived from a 129 × CAST (*castaneous*) cross. Maintenance of 2 X chromosomes in this line was monitored by X chromosome paint imaging, restriction length polymorphism analysis, as well as Sanger sequencing of SNPs on the X chromosome. The pSM33 ES cell line (kindly provided by K. Plath) is a male ES cell line, derived from the V6.5 ES cell line, expressing the lncRNA Xist from the endogenous locus under the transcriptional control of a tet-inducible promoter and the Tet transactivator (SPRITE WASH BUFFERrtTA) from the Rosa26 locus. To induce Xist, doxycycline (Sigma, D9891) was added to cultures at a final concentration of 2ug/ml for 6-24hrs.

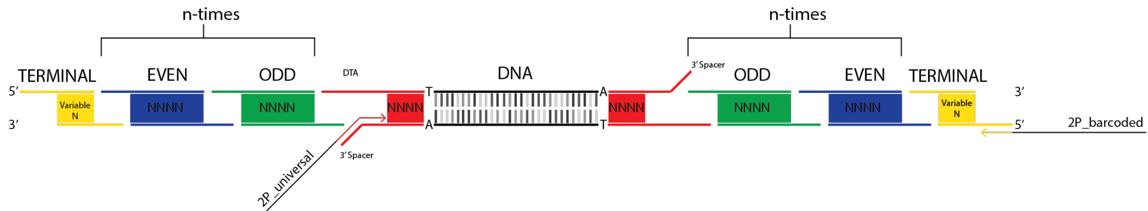
Human lymphoblast cell culture

GM12878 cells (Coriell Cell Repositories), a human lymphoblastoid cell line, was cultured in RPMI 1640 (Gibco, Life Technologies), 2mM L-glutamine, 15% fetal bovine serum, and 1 x penicillin-streptomycin and maintained at 37°C under 5% CO_2. Cells were seeded every 3-4 days at 200,000 cells/ml in T25 flasks and passaged or harvested before reaching 1,000,000 cells/ml.

References

1. R. Galupa, E. Heard, X-chromosome inactivation: new insights into cis and trans regulation. *Curr. Opin. Genet. Dev.* **31**, 57–66 (2015).
2. E. Splinter *et al.*, The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* **25**, 1371–1383 (2011).
3. S. S. Rao *et al.*, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. **159**, 1665–1680 (2014).
4. A. Rego, P. B. Sinclair, W. Tao, I. Kireev, A. S. Belmont, The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. *J Cell Sci.* **121**, 1119–1127 (2008).
5. A. Wutz, Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat. Rev. Genet.* **12**, 542–553 (2011).
6. C. M. Clemson, L. L. Hall, M. Byron, J. McNeil, J. B. Lawrence, The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7688–7693 (2006).
7. J. M. Engreitz *et al.*, The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science (80-89).* **341**, 1237973 (2013).
8. M. D. Simon *et al.*, High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature*. **504**, 465–469 (2013).
9. J. Chaumeil, P. Le Baccon, A. Wutz, E. Heard, A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**, 2223–2237 (2006).
10. A. Wutz, T. P. Rasmussen, R. Jaenisch, Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167–174 (2002).
11. C. A. McHugh *et al.*, The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*. **521**, 232–236 (2015).
12. C. Chu *et al.*, Systematic discovery of Xist RNA binding proteins. *Cell*. **161**, 404–416 (2015).
13. C. Chen *et al.,* Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science*. **354**, 468-472 (2016).

3 Adaptor and Barcode Design



The above figure demonstrates the adaptor and tag scheme that is central to the SPRITE process. SPRITE uses a split-and-pool strategy to uniquely barcode all molecules within a crosslinked complex by repeatedly splitting all complexes into a 96-well plate, ligating a specific tag sequence within each well, followed by pooling of these complexes such that the final product contains a series of tags ligated to each molecule, which we refer to as a barcode.

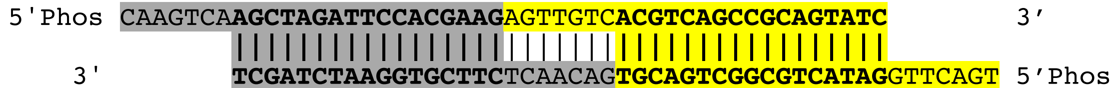
3.1 DNA Phosphate Modified (DPM) Adaptor

The above dsDNA molecule is an example of one of the 96 DPM adaptors used during our process. The 5' end of the molecule has a modified phosphate group that allows for the ligation between DPM and the target DNA molecules as well as the subsequent tag. The highlighted regions on DPM have the following functions:

1. The yellow T overhang is a sticky-end that ligates to our target DNA molecules, which are given a 5' A overhang following end repair.
2. The pink region is the 9-nucleotide sequence unique to each of the 96 DPM adaptors. These unique sequences help to identify post-sequencing DNA molecules that are in a complex.
3. The green sequence is a sticky end that ligates to the first tag.
4. The grey sequence is complementary to the First Primer used for library amplification. Part of the grey sequence makes up a 3' spacer to prevent the top strand of the Odd tag from ligating, and only the bottom 5'phosphorylated sticky end of the Odd tag will ligate to the green tag. Its purpose is discussed in section 3.4.

3.2 Odd and Even Tags

Odd and Even tags are so named because the Odd tag is ligated 1st, 3rd, 5th, etc… during the SPRITE process and the Even tag is ligated 2nd, 4th, 6th, etc… during SPRITE for however many rounds of tagging and pooling are completed. It is not necessary to ligate only an even number of tags or only an odd number of tags so long as there are two sets of Terminal tags; one that can ligate to Odd tags and one that can ligate to Even tags.

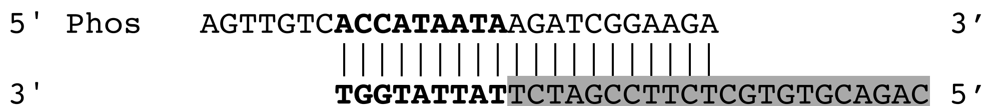


The above dsDNA molecule is an Odd tag and an Even tag ligated together. The following points are important to note:

1. The 5' overhang on the top strand ligates either to the DPM adaptor (green sequence in section 3.1) or the 5' overhang on the bottom strand of the Even tag.
2. Both the Odd tags and Even tags have modified 5' phosphate groups to allow for tag elongation.
3. The **bolded** regions of complementarity on each tag are the sequences unique to each of the 96 tags (192 total, accounting to Odd tags and Even tags).

3.3 Terminal Tag

The terminal tag below ligates to Odd tags, though a terminal tag has also been made to ligate to Even tags. The key feature of the terminal tag is that there is no modified 5' phosphate on the bottom strand.



1. The grey sequence is complementary to the Second Primer used for library amplification.
2. However since DNA cannot be synthesized in a 3' to 5' direction, the Second Primer anneals to a daughter strand synthesized from the First Primer, as explained in section 3.4. The top strand is not primed because there is a break in the sequence generated by the 3'spacer on the DPM molecule and therefore priming the top strand of the terminal tag would terminate at the barcodes and would not PCR through to the gDNA sequence ligated to the barcodes.
3. The **bolded** sequence on the Terminal tag is unique to each of the 96 tags.

3.4 Final Library Amplification

The DPM adaptor is designed with a 3' spacer to aid in final library amplification. If the 3' spacer is absent, each strand will form a hairpin loop during the initial denaturation due to reverse complementarity of the sequences on either side of the target DNA molecule. Instead, the 3' spacer allows the barcodes to only ligate to the 5'end of each single-stranded DNA sequence, and not the 3'end, preventing these hairpin from forming.

2P_universal (F primer)
5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT 3'

2P_barcoded_85 (R primer)
5' CAAGCAGAAGACGGCATACGAGATGCCTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT 3'

Due to reverse complementarity of the sequences, only one primer amplifies the tagged DNA in the first PCR cycle. This First Primer anneals to a sequence in the DPM adaptor and extends, synthesizing two daughter strands with reverse sequences. This first primer serves as the Read1 primer during Illumina sequencing. To synthesize the complement, the Second Primer anneals to the daughter strand extended from the First Primer in the second PCR cycle.

The 2P_barcoded primer contains an 8 nucleotide barcode within the primer. This barcode is read from the illumina sequencer during the indexing priming step. This barcode effectively serves as an additional round of tag addition during SPRITE. Dilution of the sample into multiple wells is performed at the final step of SPRITE prior to proteinase K elution from NHS beads. Each dilution of the sample prior to proteinase K elution isolates a subset of the tagged complexes into different wells. Each dilution of complexes are amplified with a different 2P_barcoded primer.

Both the First and Second primers are around 30 nucleotides each. Yet the sequences they anneal to initially are ~20 nucleotides. For this reason, we set two different annealing temperatures during the final library PCR. The first annealing temperature is for the first four cycles until enough copies are made with fully extended primer regions. After these four cycles, the annealing temperature is raised for a remaining five cycles.

3.5 DPM primers for QC of DPM ligation

These primers are used to ensure that the DPM adaptor has been successfully ligated to DNA of the lysate. If no libraries are obtained at this step after 14-16 cycles of PCR, we strongly recommend to not proceed as subsequent ligation of tags and amplification of tagged DNA during the SPRITE protocol will be unsuccessful.

DPMQCprimerF 5' TACACGACGCTCTTCCGATCT 3'
DPMQCprimerR 5' TGACTTGTCATGTCTTCCGATCT 3'

The Forward and Reverse primers amplify the top strand and bottom strand of the DPM adaptor, respectively (section 3.1).

4 Sample Preparation

Goal: Crosslink cells to fix *in vivo* DNA complexes. Lyse cells and fragment DNA to appropriate sizes via sonication and DNase.

Optimization of lysis conditions (amount of sonication, amount/timing of DNase) is a critical step in establishing the protocol for the first time. The length of sonication might vary from 30 seconds to several minutes and DNase treatment might vary from 10-30% DNase concentration, depending on cell number, ploidy, crosslinking strength, and the desired DNA fragment size. To optimize DNase timing and conditions, DNase samples with varying enzyme concentration for 20 minutes, quench with EDTA and EGTA on ice, and assay DNA sizes for concentration as described in the protocol. If an appropriate combination of solubilization and DNA fragment sizes cannot be obtained by varying the amount of sonication or DNase, then reducing the strength of the crosslinking may be necessary. (1)

4.1 Formaldehyde-DSG Crosslinking

1. Grow adherent cells on 15 cm plates. Count the cells on one plate before proceeding with crosslinking. This protocol details crosslinking multiple plates of cells in one suspension, but it is important to maintain consistency in lysate batches. We typically store 5-10 million cells per pellet in a 1.7 mL microcentrifuge tube.

2. An hour before starting, warm TVP and wash solution at 37° C. Chill one bottle of 1 x PBS at 4° C, keep one bottle of 1 x PBS at room temperature. Store scraping buffer at 4° C.

3. **Lift cells from plate and wash**: Remove media from plates. Add 5 mL TVP to each 15 cm plate and rock gently for 3-4 minutes ✔. Afterwards, add 25 mL wash solution to each plate. Vigorously suspend cells in the wash solution and transfer from plate to a 50 mL conical tube. Rinse the plate with extra wash solution and add to the 50 mL conical. The same cells from different plates can be pooled into the same conical. Pellet in a centrifuge for 3 minutes at 3300 x G at room temperature. Wash cells by resuspending in 4 mL room temperature 1 x PBS per 10 million cells and transfer to a 15 mL conical. Pellet again.

   ✔ Note: Cells should lift from the plate within this time. If 5 minutes have passed and cells have not started to lift from the plate, quench the trypsin immediately with wash solution to prevent over-trypsinization and premature lysis of cells. Break cells off the plate by washing vigorously with wash solution.

4. **Crosslink cells in DSG**: Make a 0.5 M solution of DSG by adding 306 uL DMSO to one bottle containing 50 mg DSG. Make an appropriate amount of DSG crosslinking solution by adding 16 uL of DSG per 4 mL of room temperature PBS. Resuspend cells in DSG crosslinking solution (4 mL per 10 million cells) ✖. Rotate gently at room temperature for 45 minutes.

   ✖ Critical Point: It is vital that at the beginning of the crosslinking process the pellet is uniformly in suspension. To achieve this, completely resuspend the pellet in 1 mL of DSG crosslinking solution using a P-1000 micropipette. After the pellet is completely dissolved, add the remaining volume of DSG Crosslinking Solution. If you add the full volume of DSG Crosslinking Solution without first resuspending the pellet, it will be almost impossible to completely break up the pellet and will result in cell clumps being crosslinked together.

5. Pellet cells for 4 minutes at 1000 x G at room temperature. Discard supernatant.

6. Wash cells with 4 mL room temperature 1 x PBS per 10 million cells ✖. Pellet as before, discarding supernatant.

   ✖ Critical Point: As before, each time you resuspend the cell pellet, whether to wash or to put in formaldehyde or scraping buffer, ensure that the pellet is completely resuspend in the solution. Achieve this by first resuspending the pellet in 1 mL of the appropriate solution using a P-1000 micropipette, then adding the remaining volume. If you forget to do this and add the full volume without first resuspending, vortex the pellet until it is fully resuspended.

7. Resuspend cell pellet in 3% formaldehyde in room temperature PBS. Rock gently at room temperature for 10 minutes.

8. Add 200 uL of 2.5 M glycine stop solution per 1 mL of cell suspension. Rock gently at room temperature for 5 minutes.

9. Pellet cells at 4° C for 4 minutes at 1000 x G. Discard formaldehyde supernatant in an appropriate waste container. From here, keep cells at 4° C.

10. Resuspend cell pellet in 4° C scraping buffer and gently rock for 1-2 minutes.

11. Pellet cells at 4° C for 4 minutes at 1000 x G. Discard supernatant in formaldehyde waste container.

12. Resuspend cell pellet in cold scraping buffer again and gently rock for 1-2 minutes. Pellet as before and discard supernatant.

13. Resuspend pellet in 1 mL of scraping buffer per 10 million cells.

14. Aliquot 10 million cells each into 1.7 mL microcentrifuge tubes and pellet at 4° C for 5 minutes at 2000 x G. Remove supernatant.

15. Flash freeze pellets in liquid nitrogen and store pellets at -80C.

(1) Engreitz, Jesse "RNA Antisense Purification (RAP): Experimental Protocols"

4.2 Cell Lysis

1. Chill Lysis Buffers A, B, and C on ice.

2. If using an electronic chiller for the sonication chamber, pre-chill to 4° C.

3. Thaw a cell pellet (10 million cells) on ice for two minutes.

4. Add 700 uL of Lysis Buffer A supplemented with 1 x Proteinase Cocktail Inhibitor (PIC) to each pellet and resuspend fully ✖.

   ✖ Critical Point: Ensure that the pellet is fully resuspended in the lysis buffer. Any pellet that does not resuspend and remains clumpy, even after vigorous pipetting, may be over-crosslinked and/or may be difficult to DNase.

5. Incubate mixtures on ice for 10 minutes.

6. Pellet cells at 4° C for 8 minutes at 850 x G.

7. Discard the supernatant, taking care not to disturb the pellet.

8. Add 700 uL of Lysis Buffer B supplemented with 1 x PIC to each cell pellet and resuspend fully.

9. Incubate mixtures on ice for 10 minutes.

10. Pellet cells at 4° C for 8 minutes at 850 x G.

11. Discard the supernatant, taking care not to disturb the pellet.

12. Add 550uL of Lysis Buffer C supplemented with 1 x PIC to each 10 million nuclei pellet and resuspend.

13. Incubate mixture on ice for 8 minutes.

14. Sonicate each sample at 4-5 watts for 1 minute: 1 pulse for 0.7 seconds ON, 3.3 seconds OFF. During and after sonication, keep lysate at 4° C. A Branson needle-tip sonicator kept at 4° C was used for this protocol.

15. If sonicating more than one pellet of the same cell and conditions and you do not wish to keep them as biological replicates, pool all lysates together and split again into 10 million cell aliquots. This ensures that all samples in each tube are equally processed and DNased in the subsequent steps.

16. Flash freeze lysate and store at -80° C, or move directly into DNA fragmentation.

4.3 DNA Fragmentation

1. Thaw one tube of lysate on ice.

2. Fragment the DNA with DNase. To obtain a desired DNA size distribution, perform several reactions with varying DNase concentrations ✓.

   ✓ Note: We typically preform SPRITE on DNA that has a size distribution from 50-1000 base pairs with an average size between 200-300 base pairs. In general, a 20% DNase reaction achieves this.

Stock Solution	**Volume**
10X SPRITE DNase Buffer	2 uL
Lysate	10 uL
Turbo DNase from ThermoFisher	3 / 4 / 5 uL
H_2O	5 / 4 / 3 uL
Total	20 uL

   NOTE: Do NOT use 10x Turbo DNAse Buffer provided with the Enzyme. Use the 10x SPRITE DNAse Buffer components listed above (page 3).

3. Incubate at 37° C for 20 minutes.

4. Add 1 uL of 25X DNase Stop Solution to each sample to terminate the reaction.

5. Reverse the crosslinks for half of each sample. Flash freeze the remaining half of the DNased sample and store at -80° C ✓.

   ✓ Note: Reversing the crosslinks on only half the DNased sample ensures that there is remaining crosslinked sample ready for SPRITE. This negates the need to DNase the entire tube of lysate (10 million cells) which could yield a different size distribution than intended.

Stock Solution	Volume
DNased Lysate	10 uL
MyRNK Buffer	82 uL
Proteinase K	8 uL
Total	100 uL

6. Incubate for at 65° C for 2 hours at minimum.

7. Clean up samples by following the protocol provided in the RNA Clean and Concentrator-5 Kit, binding in 6 volumes of DNA Binding Buffer. Elute in 10 uL of H_2O.

8. Determine concentration of DNA in each sample by following directions provided with the Qubit dsDNA HS Assay Kit.

9. Determine the size distribution of DNA in each sample by following directions provided with either the High Sensitivity DNA Kit for the Agilent Bioanalyzer or the D1000 Screentape for the Agilent 2200 TapeStation.

10. If none of these concentrations of TURBO DNase led to ideal fragmentation, adjust concentrations and repeat the DNasing until optimal fragmentation is achieved. DNA for SPRITE generally has a size distribution from 50-1000 base pairs with an average size between 200-300 base pairs.

11. OPTIONAL: DNase the entire batch of crosslinked lysate at the identified optimal DNAase concentration ✓.

✓ Note: DNasing the batch of crosslinked lysate is not necessary for SPRITE if half of the DNased material was saved from step 5 in the DNA fragmentation process.

Stock Solution	Volume
10X DNase Buffer	110uL
Lysate	550uL
Turbo DNase from ThermoFisher	X uL
H_2O	X uL to reach final volume
Total	1100uL

a. Incubate at 37° C for 20 minutes.
b. Add 44 uL of 25 x DNase Stop Solution to each sample to terminate the reaction.
c. Flash freeze DNased lysate and store at -80° C.

5 Library Preparation Pt. 1

Goal: Lysate is coupled to Pierce NHS-Activated Magnetic Beads to allow for easy DNA library preparation. DNA overhangs caused by fragmentation are repaired and blunted the NEBNext End Repair Module, which contains enzymes with 5' to 3' polymerase activity as well as 3' to 5' exonuclease activity. Klenow fragment (-exo) is used to add an adenine dNTP to 3' ends of each DNA molecule. This aids in ligation of the DPM adaptor, which has a 3' thymine overhang, without creating spurious ligation products.

It is crucial to have an optimal bead to molecule ratio for the library preparation and SPRITE processes. We aim to bind at a 1:4 to 3:4 ratio of DNA molecules to beads; generally we bind around 50 billion molecules to 75 billion beads. Assuming that we have 50% binding efficiency, coupling ratio is then 1:8 to 1:2.7 molecules per bead. We assume that there are far more molecules than crosslinked complexes making the complex to bead ratio even lower than stated above, but use molecules as an over-estimate to reduce noise per bead. To determine the microliter amount of lysate to couple we calculate the DNA molarity from the concentration and average size measurements obtained in the DNA Fragmentation step of the protocol. Molarity is multiplied by the lysate volume (10 uL of crosslinked, DNased lysate remains in the tube) to obtain DNA molecule number.

5.1 NHS Coupling

✔ Note: All wash steps at 4° C are performed in a cold room. All wash steps above room temperature are performed on an Eppendorf Thermomixer. If a temperature is not specified, it is at room temperature. To wash beads, place the tube containing the beads on a magnetic rack to capture the beads. Wait until the solution is clear and all beads are captured before removing the liquid. Add the wash solution to the beads and remove the tube from the magnet. Invert the tube until all beads are in suspension. If using an Eppendorf Thermomixer, set the thermomixer to shake at 1200 RPM. Briefly centrifuge the tube to remove beads from the lid, then place the tube back on the magnet to capture the beads again. Wait until the solution is clear and all beads are captured before removing the wash liquid. These steps are critical to avoid loss of beads throughout protocol

✔ Note: The protocol can be stopped at any point of the process. To ensure the integrity of the DNA, resuspend the beads in 1 mL Modified RLT Buffer and store at 4° C until you wish to resume. Wash three times with SPRITE Wash Buffer to remove all RLT before proceeding with the protocol to prevent enzyme denaturation in subsequent steps of the protocol.

✔ Note: All steps involving bead pipetting should use low-bind pipette tips.

1. Gently invert the bottle containing the Pierce NHS-activated beads in *N,N*-dimethylacetamide (DMAC) until there is a uniform suspension. Being careful not to introduce water into the bottle, transfer X mL of NHS beads into a clean 1.7 mL lo-bind tube. Place the tube on a magnetic rack to capture the beads.

2. Remove the DMAC and wash beads with 1 mL ice-cold 1mM HCl.

3. Wash beads with 1 mL ice-cold 1 x PBS.

4. Add 1 mL Coupling Buffer to the beads.

5. Add X uL of DNased lysate to the beads.

6. Incubate the lysate and beads overnight at 4° C on a mixer.

7. Place beads on a magnet and remove 500 uL of flowthrough. OPTIONAL: This flowthrough aliquot can be saved to determine coupling efficiency if the DPM QC fails.

8. Add 500uL 1M Tris pH 7.5 (3M ethanolamine pH 9.0 can also be used) to the beads and incubate on a mixer at 4° C for at least 45 minutes. This ensures that all NHS beads will be quenched with protein from bound lysate or Tris, and will not bind enzymes in the following steps.

9. Wash beads twice in Modified RLT Buffer.

10. Wash beads three times with SPRITE Wash Buffer.

11. Spin the beads down quickly in a microcentrifuge and place back on the magnet to remove any remaining liquid.

5.2 Phosphorylation and End Repair

1. Blunt the 5' and 3' ends of the DNA molecules to prevent unwanted ligation by adding the following mixture to the beads:

Stock Solution	**Volume**
H_2O	212.5uL
End Repair Reaction Buffer (10X)	25uL
End Repair Enzyme Mix	12.5uL
Total	250uL

2. Incubate on a thermomixer for 60 minutes at 24° C, 1200 RPM.

3. Wash once with Modified RLT Buffer.

4. Wash three times with SPRITE Wash Buffer.

5. Spin the beads down quickly in a microcentrifuge and place back on the magnet to remove any remaining liquid.

6. Add dATP to the 3' ends of each DNA molecule to allow for ligation of the DPM adaptor by adding the following mixture to the beads:

Stock Solution	**Volume**
H2O	215uL
dA-Tailing Reaction Buffer (10X)	25uL
Klenow Fragment (exo-)	10uL
Total	250uL

7. Incubate on a thermomixer for 60 minutes at 37° C, 1200RPM. If ligating the DPM adaptor barcode on the same day, set up the reaction during this incubation.

8. Wash once with Modified RLT Buffer.

9. Wash three times with SPRITE Wash Buffer.

10. Spin the beads down briefly in a microcentrifuge and place back on the magnet to remove any remaining liquid.

5.3 DPM Adaptor Ligation

✔ Note: There are 96 adaptors that are designed to ligate onto the DNA molecules. These DPM adaptors are kept in a 96-well stock plate at 4.5 uM. The ligation reaction between the adaptors and the DNA occurs in a 96-well plate. The following steps that detail set up are designed for optimum efficiency during the process.

✔ Note: All ligation steps include SPRITE Wash Buffer, which contains detergents to prevent beads from aggregating, sticking to the plastic tips and tubes, and for even distribution of the beads across a 96-well plate. We have verified that these detergents do not significantly inhibit ligation efficiency.

1. Create a dilute SPRITE Wash Buffer by mixing 1100 uL of SPRITE Wash Buffer with 792 uL of H_2O.

2. Accounting for bead volume, add the dilute SPRITE Wash Buffer to the beads to achieve a final volume of 1.075 mL. Ensure that the beads are equally resuspended in the buffer. Distribute the beads equally into a 12-well strip tube by aliquoting 89.6 uL of beads into each well.

3. Make Ligation Master Mix for five rounds of SPRITE (DPM + four extra tags). Split the master mix evenly into each well of a 12-well strip tube by pipetting 260 uL into each well. Keep on ice ✔.

Stock Solution	**Volume**
NEBNext Quick Ligation Reaction Buffer (5X)	1600 uL
Instant Sticky-end Ligation Master Mix (2X)	1000 uL
1,2-Propanediol	600 uL
Total	32000 uL

✔ Note: Ligation Master Mix can be stored overnight at -20° C.

4. Centrifuge the DPM adaptor stock plate before removing the foil seal. Aliquot 2.4 uL from the stock plate of DPM adaptors to a new low-bind 96-well plate ✔. Be careful to ensure that there is no mixing between wells at any point of the process to avoid cross-contamination of barcodes. Use a new pipette tip for each well. After transfer is complete, seal both plates with a new foil seal.

   ✔ Note: This step can be done in advanced, in bulk, so that these plates are ready-to-use.

5. Centrifuge the 96-well plate containing the aliquoted adaptors, and then remove the foil seal.

6. Aliquot 11.2 uL of beads into each well of the 96-well plate that contains 2.4 uL of the DPM adaptors. Be careful to ensure that there is no mixing between wells at any point of the process. Use a new pipette tip for each well. Also be careful to ensure that there are no beads remaining in the pipette tip.

7. Carefully add any remaining beads to individual wells on the plate in 1 uL aliquots.

8. Aliquot 6.4 uL of Ligation Master Mix into each well, mixing by pipetting up and down 10 times. Be careful to ensure that there is no mixing between wells at any point of the process. Use a new pipette tip for each well.

9. The final reaction components and volumes for each well should be as follows:

Stock Solution	Volume
Beads + SPRITE Wash Buffer + H_2O Mix	11.2 uL
DPM Adaptor (4.5 uM)	2.4 uL
Ligation Master Mix	6.4 uL
Total	20 uL

10. Seal the plate with a foil seal and incubate on a thermomixer for 60 minutes at 20° C, shaking for 30 seconds at 1600 RPM every five minutes to prevent beads from settling to the bottom of the plate ✓.

✓ Note: Ligation time is critical for high efficiency of ligation each round. We have tested ligation at 5, 15, 30, 45, and 60 minute reaction times and 60 minutes ligation time has significantly higher yields over the other times.

11. After incubation, centrifuge the plate before removing the foil seal.

12. Pour Modified RLT Buffer into a sterile plastic reservoir, and transfer 60 uL of Modified RLT Buffer into each well of the 96-well plate to stop the ligation reactions. It is not necessary to use new tips for each well.

13. Pool all 96 stopped ligation reactions into a second sterile plastic reservoir.

14. Place a 15 mL conical tube on an appropriately sized magnetic rack and transfer the ligation pool into the conical. Capture all beads on the magnet, disposing all Modified RLT Buffer in an appropriate waste receptacle.

15. Remove the 15 mL conical containing the beads from the magnet and resuspend beads in 1 mL SPRITE Wash Buffer. Transfer the bead solution to a microcentrifuge tube.

16. Wash three times with SPRITE Wash Buffer.

17. Spin the beads down briefly in a microcentrifuge and place back on the magnet to remove any remaining liquid.

5.4 QC: Check to Determine Ligation Efficiency of the DPM Adaptor

1. Resuspend the beads in MyRNK Buffer so that the final beads + buffer volume is 1 mL. Remove a 5% aliquot (50 uL) into a separate 1.7 mL microcentrifuge tube.

2. Place the remaining 95% of beads back on the magnetic rack, remove the MyRNK Buffer, and store beads in 1 mL of Modified RLT Buffer. Keep beads at 4° C overnight.

3. Elute DNA in the 5% aliquot by reversal of crosslinks through heating and Proteinase K.

Stock Solution	**Volume**
Sample on beads in MyRNK Buffer	50uL
MyRNK Buffer	42uL
Proteinase K	8uL
Total	100uL

4. Incubate at 65° C for two hours minimum.

5. Place the microcentrifuge tube on a magnet and capture the beads. Remove the flowthrough that contains the DNA ligated with DPM adaptor and place in a clean 1.7mL microcentrifuge tube.

6. Pipette 25 uL of H_2O into the tube containing the beads. Vortex, briefly centrifuge, and re-capture the beads. Remove the 25 uL of H2O that now contains any residual nucleic acid and add to the new sample tube. Discard the beads.

7. Clean the DNA by following the protocol provided in the RNA Clean and Concentrator Kit. Elute in 40 uL of H2O.

8. Amplify the DNA molecules that are ligated to the adaptors. The forward primer should prime off the 5' end of the DPM adaptor and the reverse primer should prime off the 3' end of the DPM adaptor.

Stock Solution	**Volume**
Sample (cleaned)	10 uL
DPMQCForward Primer (100uM)	1 uL
DPMQCReverse Primer (100uM)	1 uL
H2O	13 uL
Q5 Hot Start Master Mix	25 uL
Total	50 uL

PCR Program:

1. Initial denaturation: 98° C - 120 seconds
2. 12-16 cycles:
    a. 98° C -10 seconds
    b. 67° C - 30 seconds
    c. 72° C - 40 seconds

3. Final extension: 72° C - 120 seconds
4. Hold 4C

9. Clean the PCR reaction and size select for your target DNA molecules. Our DPM adaptors are 30 base pairs each and our target DNA molecules no less than 100 base pairs. Agencourt AMPure XP beads size select while cleaning the PCR reaction of unwanted products.

    a. Add 1.0 x volume (50 uL) of AMPure XP beads to the sample for a total volume of 100uL and mix thoroughly.
    b. Incubate for 10 minutes at room temperature.
    c. Place the beads on an appropriately sized magnet to capture the beads and the bound DNA. Wait a few minutes until all the beads are captured.
    d. Remove the supernatant and discard.
    e. Wash beads twice with 80% ethanol by pipetting ethanol into the tube while beads are captured, moving the tube to the opposite side of the magnet so that beads pass through the ethanol, and then removing the ethanol solution.
    f. Quickly spin down the beads in a microcentrifuge, re-capture on magnet, and remove any remaining ethanol.
    g. Air-dry beads while the tube is on the magnet.
    h. Elute the amplified DNA from the beads by resuspending the beads in 12 uL of H_2O. Place the solution back on the magnet to capture the beads. Remove the eluted amplified DNA to a clean microcentrifuge tube.

10. Determine concentration of DNA with the DPM adaptor in by following directions provided with the Qubit dsDNA HS Assay Kit.

11. Determine the size distribution of DNA with the DPM adaptor in each sample by following directions provided with either the High Sensitivity DNA Kit for the Agilent Bioanalyzer or the D1000 Screentape for the Agilent 2200 TapeStation. The average size should be roughly similar to the average size of the input lysate (around 200-400 base pairs).

12. Calculate the number of DNA molecules in the 5% aliquot by determining molarity from the concentration and average size✖.

    ✖ Critical Point: The 5% aliquot should contain, at the very minimum, 15 million unique DNA molecules in order to proceed with SPRITE. If the aliquot contains less than this number, there will not be enough unique reads to sequence the SPRITE library. If this is the case, troubleshoot what went wrong by assaying coupling efficiency from the flowthrough saved in Step 7 of the NHS Coupling protocol. If lysate was successfully coupled, consider whether a

mistake was made during ligation of the DPM adaptor or during one of the critical steps of crosslinking and lysis.

6 SPRITE and Library Preparation Pt. 2

Goal: The SPRITE method provides each DNA-DNA complex in the sample lysate with a unique nucleic acid barcode. When these complexes are decrosslinked, the individual DNA molecules that made up a single complex retain identical barcodes. These DNA libraries are sequenced on an Illumina Next-Generation sequencing platform and analyzed. Any DNA molecules found to have the same barcode interact *in-vivo*.

The SPRITE method works by splitting into a 96-well plate a pooled sample of crosslinked lysate where DNA molecules are ligated to the DPM adaptor. Each well of the 96-well plate contains a unique tag (Odd) to which the DNA molecules are ligated. The ligation reactions are stopped, pooled, and split again into a new 96-well plate containing different, unique tags than the first (Even). If n rounds of tag ligation are performed, 96^n unique barcodes are generated. We typically ligate 5 tags, creating over 8 billion unique barcodes. After all barcodes are ligated, the sample is split again into small m aliquots (100 wells of 1% aliquots up to 10 wells of 10% aliquots are typically used depending on the total material coupled) for PCR amplification. This final splitting of samples effectually sorts the DNA complexes once more, so that the chance that two different non-crosslinked complexes with the same barcode are amplified together is negligible. This last dilution into m wells effectively raises the number of unique tags to each molecule to $m*96^n$. For example, if the sample is aliquoted into 1% aliquots, then over 815 billion unique barcodes are generated.

The first round of SPRITE was already completed with the ligation of 96 unique DPM adaptors that allow for the subsequent ligation of new barcodes. As detailed in the Adaptor and Barcode Creation section, subsequent tag ligations are performed in the following order:

1. ODD Tag Ligation
2. EVEN Tag Ligation
3. ODD Tag Ligation
4. Terminal Tag Ligation

The four barcode ligations listed above are performed in the exact same manner with the only difference being the tag sequence. Thus, the following section will only detail one round of SPRITE.

6.1 SPRITE

1. Remove Modified RLT Buffer from the remaining 95% aliquot and wash three times with SPRITE Wash Buffer.

2. Spin the beads down briefly in a microcentrifuge and place back on the magnet to remove any remaining liquid.

3. Create a dilute SPRITE Wash Buffer by mixing 1100 uL of SPRITE Wash Buffer with 792 uL of H_2O.

4. Accounting for bead volume, add the dilute SPRITE Wash Buffer to the beads to achieve a final volume of 1.075 mL. Ensure that the beads are equally resuspended in the buffer. Distribute the beads equally into a 12-well strip tube by aliquoting 89.6 uL of beads into each well.

5. If frozen, thaw the strip tube containing the Ligation Master Mix made in Step 3 of the DPM Adaptor Ligation protocol. Keep on ice until ready to use.

6. Centrifuge the tag stock plate before removing the foil seal. Aliquot 2.4 uL from the stock plate of tags to a new low-bind 96-well plate ✓. Be careful to ensure that there is no mixing between wells at any point of the process to avoid cross-contamination of barcodes. Use a new pipette tip for each well. After transfer is complete, seal both plates with a new foil seal.

   ✓ Note: This step can be done in advanced, in bulk, so that these plates are ready-to-use.

7. Centrifuge the 96-well plate containing the aliquoted tags, and then remove the foil seal.

8. Aliquot 11.2 uL of beads into each well of the 96-well plate that contains 2.4 uL of the tags. Be careful to ensure that there is no mixing between wells at any point of the process. Use a new pipette tip for each well. Also be careful to ensure that there are no beads remaining in the pipette tip.

9. Carefully add any remaining beads to individual wells on the plate in 1 uL aliquots.

10. Aliquot 6.4 uL of Ligation Master Mix into each well, mixing by pipetting up and down 10 times. Be careful to ensure that there is no mixing between wells at any point of the process. Use a new pipette tip for each well.

11. The final reaction components and volumes for each well should be as follows:

Stock Solution	Volume
Beads + SPRITE Wash Buffer + H_2O Mix	11.2 uL
Tag (4.5 uM)	2.4 uL
Ligation Master Mix	6.4 uL
Total	20 uL

12. Seal the plate with a foil seal and incubate on a thermomixer for 60 minutes at 20° C, shaking for 30 seconds at 1600 RPM every five minutes to prevent beads from settling to the bottom of the plate ✓.

    ✓ Note: Ligation time is critical for high efficiency of ligation each round.

13. After incubation, centrifuge the plate before removing the foil seal.

14. Pour Modified RLT Buffer into a sterile plastic reservoir, and transfer 60 uL of Modified RLT Buffer into each well of the 96-well plate to stop the ligation reactions. It is not necessary to use new tips for each well.

15. Pool all 96 stopped ligation reactions into a second sterile plastic reservoir.

16. Place a 15 mL conical tube on an appropriately sized magnetic rack and transfer the ligation pool into the conical. Capture all beads on the magnet, disposing all Modified RLT Buffer in an appropriate waste receptacle.

17. Remove the 15 mL conical containing the beads from the magnet and resuspend beads in 1 mL SPRITE Wash Buffer. Transfer the bead solution to a microcentrifuge tube.

18. Wash three times with SPRITE Wash Buffer.

19. Repeat the process starting at Step 6 for the remaining three SPRITE rounds.

6.2 Final Library Preparation

1. Resuspend the beads in MyRNK Buffer so that the final beads + buffer volume is 950 uL.

2. Remove eight 5% aliquots into clean 1.7 mL microcentrifuge tubes and elute the barcoded DNA from the beads. Keep the remaining 55% of the lysate on beads in Modified RLT Buffer at 4° C.

Stock Solution	Volume

Beads in MyRNK Buffer	50 uL
MyRNK Buffer	42 uL
Proteinase K	8 uL
Total	100 uL

3. Incubate at 65° C overnight.

4. Place the microcentrifuge tubes on a magnet and capture the beads. Remove the flowthrough that contains the barcoded DNA and place in a clean microcentrifuge tube.

5. Pipette 25 uL of H_2O into the tube containing the beads. Vortex, and re-capture the beads. Remove the 25 uL of H_2O that now contains any residual nucleic acid and add to the new sample tube. Discard the beads.

6. Follow the protocol provided in the RNA Clean and Concentrator Kit. Elute in 20 uL of H_2O.

7. Amplify the barcoded DNA. Refer to section 3.4 for details about the final library amplification step.

Stock Solution	**Volume**
Barcoded DNA (cleaned)	20 uL
First Primer (100uM)	1 uL
Second Primer (100uM)	1 uL
H_2O	3 uL
Q5 Hot Start Master Mix	25 uL
Total	50 uL

PCR Program:

1. Initial denaturation: 98° C - 180 seconds
2. 4 cycles:
   a. 98° C -10 seconds
   b. 68° C - 30 seconds
   c. 72° C - 40 seconds
3. 7 cycles:
   a. 98° C -10 seconds
   b. 70° C - 30 seconds
   c. 72° C - 40 seconds
4. Final extension: 72° C - 180 seconds
5. Hold 4° C

8. Clean the PCR reaction and size select for your target libraries. The total length of our barcode on one amplified product is around 160 base pairs and each target DNA molecules no less than 100 base pairs. Agencourt AMPure XP

beads are able to size select while cleaning the PCR reaction of unwanted products.

a. Add 0.7 x volume (35 uL) AMPure XP beads to the sample for a total volume of 85 uL and mix thoroughly.
b. Incubate for 10 minutes at room temperature.
c. Place the beads on an appropriately sized magnet to capture the beads and the bound DNA. Wait a few minutes until all the beads are captured.
d. Remove the supernatant and discard.
e. Wash beads twice with 80% ethanol by pipetting ethanol into the tube while beads are captured, moving the tube to the opposite side of the magnet so that beads pass through the ethanol, and then removing the ethanol solution.
f. Quickly spin down the beads in a microcentrifuge, re-capture on magnet, and remove any remaining ethanol.
g. Air-dry beads while the tube is on the magnet.
h. Elute the amplified DNA from the beads by resuspending the beads in 50 uL of H2O.
i. Repeat the size-select clean up with 0.7 x AMPure XP beads (add directly to the eluted DNA/bead mix), eluting finally in 12 uL H_2O ✔.

✔ Note: To ensure all library material is eluted from beads, elute twice with 6 uL H_2O. Most of the material will be removed in the first elution, and any remaining material will be removed in the second.

9. Determine concentration of the barcoded libraries by following directions provided with the Qubit dsDNA HS Assay Kit. Final libraries are generally between 0.3 ng/uL and 1.5 ng/uL.

10. Determine the size distribution and average size of the barcoded libraries by following directions provided with the High Sensitivity DNA Kit for the Agilent Bioanalyzer. Average sizes are generally around 350-450 base pairs.

11. Calculate the number of DNA molecules in each barcoded library by determining molarity from the concentration and average size. We typically pool together 300 million unique DNA molecules for sequencing.

7 Sequencing and Data Analysis

The Illumina, Inc. HiSeq v2500 platform was employed for next generation sequencing of the generated libraries using a TruSeq Rapid SBS v1 Kit – HS (200 cycle) and TruSeq Rapid Paired End Cluster Kit – HS.

7.1 Computational pipeline

The sequencing data is output as two FASTQ files: one file for read one (forward strand) that contains information about the genomic sequence, and one file for read two (reverse strand) that comprises information about the attached barcodes. In addition to the sequence information, in this case as well as in general, the FASTQ files also contain information about the sequencing quality and the flowcell position for each read, as well as a unique name for each read. This information is retained throughout the following process.

First, the tag sequences within the reads are identified. The process starts with a hashtable containing all barcodes used for the specific experiment. This hashtable maps nucleotide sequence to barcode name, allowing up to two mismatches for the EVEN and ODD tag sequences, but requiring no mismatches for the DPM and the Yshape adapter.

The engine queries the hashtable for subsequences of the reads, first looking at the beginning of read one for a DPM sequence, then "walking down" read two looking sequentially for the Yshape adapter, an ODD barcode, an EVEN barcode, and a final ODD barcode. The output of this process is two FASTQ files identical to the input files, but with each read name adjusted to additionally contain the five identified sequences. (If fewer than five sequences are found, the read name contains all of the found sequences, as well as a "NOT_FOUND" placeholder for each barcode not found.)

The first eleven bases, which should be the DPM sequence, are trimmed from the read-one sequences. The remainder should be wholly genomic DNA, and this is aligned by Bowtie2 (local alignment, default scoring, -D 15 -R 2), first to the mouse mm9 assembly (nucleotide sequences of reference chromosomes and scaffolds), then separately to the human hg19 assembly. Local alignment is chosen over global alignment in case the sequence "reads through" the genomic data into the barcodes on the opposite side. Some sequences may align to both mm9 and hg19, so the two resulting BAM files are compared against each other, and any reads sequences, which align to both assemblies, are removed. Finally, the BAM files are filtered to keep only those reads in which all barcodes were successfully identified. Ultimately, each BAM record contains the following information:

1. Read name
2. Location within the flowcell (lane, swath, tile)
3. Sequencing quality
4. Alignment coordinates (chromosome and nucleotide position)
5. The sequence of identified barcodes
6. The sample identifier (i.e molecule-to-bead ratio tracked indirectly, from the DPM)
7. Alignment score (how well the read aligned)
8. The identification of mismatches/indels

7.2 DNA contact maps

In case of data originating from DNA-DNA barcoding with a pooled sample, each condition was specifically labeled with a subset of DPM adapter barcodes. To de-convolute the different individual samples a script is run that identifies the DPM adapter barcodes in the BAM files resulting from the computational pipeline and generates a file for each DPM adapter barcode subset. This step is not necessary in case of RNA-DNA barcoding. Next, all reads within an individual sample possessing the same barcodes are grouped together and contact frequencies between bins of a given size (for instance 1 Mb) are calculated. Since the number of contacts that each bin participates in is not uniform, we must take this into account in order to accurately compare contact frequencies between different pairs of bins. To do this, we normalize the contact frequencies by calculating the expected number of contacts for each pair of bins based on a uniform distribution of contacts and then dividing the observed number of contacts by the expected number of contacts. This information is then translated into a matrix (contact frequency map) that can plot either the observed/expected ratio or a Z-score of this ratio. The plot can be made for single chromosome and/or a genome-wide contact frequency map. Iterative correction and eigenvector decomposition of Hi-C data can also be performed to normalize DNA contacts as described by Imakaev *et al*. Nature Methods (2012).

The signal to noise level is measured by calculating the percentage of observed over expected contacts containing reads from both specimens (human and mouse) whereby expected is the percent of human-mouse contacts at random.