SUPPLEMENTARY MATERIALS

SUPPLEMENTARY NOTE

Digestion, cross-linking and copolymer architecture and hetero/euchromatin phase separation – theoretical considerations

Chromatin in the G1 nucleus can be considered as a set of blocks of euchromatin and heterochromatin (the A and B compartments consisting of regions of predominantly euchromatin vs heterchromatin, respectively), which are constrained to be near each other by being part of the same linear chromosomes, i.e., effectively being long many-block copolymers. We suppose that the A and B heterochromatin/euchromatin monomers have a weak tendency to repel one another (or equivalently that A-A or B-B attract one another, for example via protein-mediated nucleosome-nucleosome interactions acting preferentially on euchromatin or heterochromatin, or even via physio-chemical effects such as relative hydrophobicity of more methylated nucleosomes).

If we suppose the A and B blocks to be on average N monomers long (roughly nucleosomes for the sake of this discussion), then under melt-like conditions the standard Flory theory of polymer phase separation predicts that if we were to cut the polymers into pure A and B blocks at the block boundaries (i.e., at a spacing of N monomers commensurate with the block sizes), they would phase separate for a segment-segment interaction strength stronger than $\chi^* = 2/N$ [1]. Note that this level of interaction (given approximately in k_BT units) is proportional to 1/N where N is crudely in nucleosome units; for 200 kilobase blocks, we have approximately N=1000, indicating that small fractions of a k_BT in effective A/B repulsion or A-A or B-B attraction is sufficient to drive strong euchromatin/heterochromatin phase separation [2].

Now if we were to instead cut less frequently than this, say at every second block boundary, so as to arrive at a system of AB linear diblock copolymers each of length 2N (N monomers of A followed by N monomers of B), the constraint that the A and B blocks be connected suppresses phase separation, increasing the critical interaction (all other factors held constant) to $\chi^* = 5.3/N$ [3]. In this case bulk phase separation cannot occur, but instead local, or "microphase separation" occurs, with formation of micelle-like or layered phase-separated structures. Nevertheless, for $\chi >> \chi^*$, strong segregation of the A and B monomers can still occur.

If we were to not cut at all, but rather to suppose that the chromosomes are very long multiblock copolymers, with many blocks each of N monomers alternating between A and B ("ABABABAB... multiblock copolymers"), the critical interaction strength will rise with increasing number of blocks, approaching the limit $\chi^* = 7.5/N$ for many blocks [4]. Therefore, starting from this limit, the tendency for chromosome domains to phase separate will be enhanced by cutting the chromosomes up into successively smaller pieces: as chromatin cutting increases from no cutting, we expect to see intensification of A/B compartment contrast in the Hi-C map.

Now, if we cut too frequently, when the cuts become spaced smaller than the block size (cut spacing M < N monomers), we will have the situation that the critical interaction strength will

become $\chi^* = 2/M > 2/N$, i.e., the cuts are frequent enough to suppress phase separation by decreasing the amount of interaction enthalpy per polymer "molecule". Therefore we expect that overly frequent cutting will cause a reduction in A/B compartment Hi-C map contrast, i.e., for some intermediate level of cutting similar to the sizes of the A and B blocks, one will see a maximum level of A/B compartment contrast.

There is also likely an effect of “crosslinking” ("chromatin cross-bridging"), which provides an additional level of constraint suppressing phase separation, above the linear-multiblock architecture of chromosomes. For example, taking linear diblock copolymers (N A monomers followed by N B monomers) and circularizating them raises the critical interaction for microphase separation from 5.3/N to 8.9/N, nearly a factor of 2 [5].

Similarly, if we start with A and B homopolymers each of length N, constraining them to have their ends at a flat surface, thus forcing them to mix at the surface, increases the critical interaction for phase separation from 2/N up to 4.5/N [6], with microphase separation again occurring in the constrained case. Releasing chromatin crosslinking/cross-bridging constraints (which also will occur for chromatin cutting) will in general also reduce the interaction strength needed to drive phase separation, increasing A/B compartment contrast in Hi-C maps.

In conclusion, basic polymer phase separation theory predicts that gradually increasing the cleavage of chromatin will gradually increase the intensity of A/B compartment contrast in Hi-C maps until the cuts are spaced by approximately one A or B block; further cutting will reduce the intensity of phase separation and A/B compartment contrast. Notably, the nature of the segregation can be expected to be "microphase segregation" rather than bulk phase separation, until the number of cuts is sufficient to liberate A or B "homopolymer" segments.

Estimation of the value of Chi parameter (χ)

Our results allow a crude estimate of the Flory-Huggins χ parameter for A/B segregation of chromatin in the nucleus. Given that HindIII and DpnII cut chromatin into segments of approximately 17±7 kb and 3±1 kb respectively (with FatI also generating approximately 3 kb average sized fragments), a reasonable estimate of the minimum length of fragments necessary to drive A/B segregation is N* = 10±4 kb. Given that these fragments are small compared to the A/B compartmentalization scale of a few Mb, the fragments will be essentially A- or B-type (euchromatin or heterochromatin) homopolymers. The similar ~1.2 Mb average size of the A and B compartments, as well as the remarkably similar ~3 kb fragment sizes observed for DpnII digestion of A and B regions (Extended Data Figure 4F) indicate the use of the simple Flory-Huggins model applied to a concentrated solution of symmetrical (equal length and concentration) A and B homopolymers.

For homopolymers, the Flory-Huggins model predicts a critical length needed for phase separation of $N^*=2/\chi$ [1], indicating $\chi = 0.20 \pm 0.07$/kb ($\chi = 0.036 \pm 0.013$ /nucleosome). When we assume 17±7 kb is the smallest fragment size that allows phase separation the lower limit would be $\chi = 0.12 \pm 0.04$/kb ($\chi = 0.01 \pm 0.004$ / /nucleosome).Given that χ is in k_BT units (1 k_BT = 0.6 kcal/mol at physiological temperatures), its small value indicates that the effective demixing

interaction between nucleosomes is weak, consistent with a liquid-like phase-separation picture for A and B compartments, where regulation of chromatin organization and compartmentalization is possible by relatively small changes in nucleosome interactions (e.g., via histone modifications). While extremely crude (e.g., we have used the χ estimate for a dense polymer melt rather than for a concentrated solution) our data clearly indicate a weak value for χ, and that the interactions that are driving compartmentalization are a small fraction of a kcal/(mol·nucleosome). Our result is within a factor of 2 of a recent theoretical estimate (critical size for demixing of 20 kb, [7]). We note that we have been able to rather use a simple application of Flory-Huggins theory by virtue of the similar A/B fragment and compartment sizes. We emphasize that these properties do need to be examined before similar estimates can be done for other cell types or species.

We also note that our estimated χ assumes that the fragments are able to equilibrate their positions in both the DpnII and HindIII cases. It is conceivable that entanglements, cross-bridging or other effects may strongly retard or preclude demixing in the HindIII case. Such effects would be a confounding factor in estimation of χ, but the DpnII data do solidly constrain the critical size for demixing to be larger than 3 kb (i.e., $\chi < 0.66$/kb). Further experiments on the kinetics of fragment demixing would be very interesting in this regard, as they would shed further light on the physical processes underlying A/B compartment formation in vivo.

EXPERIMENTAL METHODS

3C (Chromosome Conformation Capture) Protocol

Crosslinking*:* 1.25 mL of 37% formaldehyde was added to 40 mL of HBSS. 50 million cells or nuclei were washed twice using 20 mL of HBSS and then pelleted at 500 g for 10 min. The pellet was resuspended in 5 mL HBSS and then added to 41.25 mL of HBSS and formaldehyde (final formaldehyde concentration was 1%). The sample was incubated at RT for 10 min on a rocking platform. Afterward, to stop cross-linking 2.5 mL of 2.5 M glycine was added and samples were incubated at RT for 5 min on a rotating platform. To pellet the crosslinked cells or nuclei the sample was centrifuged at 800 g for 10 min at 4°C. After discarding the supernatant the pellet was washed twice using HBSS. Next, the pellet was either processed immediately as described below or was stored at -80 ºC after flash freezing using liquid nitrogen.

Cell lysis: (This step was included when cells are used, but was skipped for 3C with purified nuclei).Cells were lysed by adding 2 mL of cold lysis buffer [10 mM Tris-HCl (pH=8.0), 10 mM NaCl, 0.2% Igepal CA-630 (NP40)] and 20 µL of 100x Protease inhibitors. The sample was incubated on ice for 15 min to let the cells swell. The cells were lysed on ice using the homogenizer with pestle A (KIMBLE Kontes 885300-0002) by moving the pestle slowly up and down 30 times and incubating on ice for 1 min followed by another 30 strokes. The sample was transferred to two 1.5 mL microcentrifuge tubes, spun at 5,000 g at RT for 5 min using a benchtop centrifuge.

Digestion: each pellet was washed using 1 mL cold 1X NEBuffer 2.1, then spun at 5,000 g for 5 min at RT using a benchtop centrifuge, afterward each pellet was resuspended in 250 µL of 1X NEB2.1 buffer, and the two pellets were pooled (~ 500 µL). 50 µL aliquots of the suspension

were transferred to 10 new 1.5 mL microfuge tube and 292 µL of 1x NEBuffer 2.1 was added to each tube. Next, 38 µL of 1% SDS was added per tube and mixed well, the samples were incubated at 65ºC for 10 min, then placed on ice. 44 µL of 10% Triton X-100 was added to each tube to quench SDS. Finally, 400 U of EcoRI (NEB R0101L) was added per tube and incubated at 37ºC overnight on a thermocycler (with 900 rpm for 30 sec every 4 min).

Ligation*:* 86 µL of 10% SDS was added to the digested samples and the samples were then incubated at 65ºC for 30 min for EcoRI inactivation after which the tubes were placed on ice. Each sample was then transferred to a 15 mL conical tube and 7.69 mL of ligation mix was added [820 µL 10% Triton X-100, 820 µL 10x ligation buffer (500 mM Tris-HCl pH7.5, 100 mM MgCl2, 100mM DTT), 82 µL 10 mg/mL BSA, 82 µL 100 mM ATP and 5.886 µL ultrapure distilled water]. Finally, 10 U of T4 ligase (Invitrogen 15224090) was added per tube before incubation at 16ºC for 2 hr on a thermocycler (with 900 rpm for 30 sec every 4 min).

Reverse Crosslinking: 50 µL of 10 mg/mL proteinase K (Fisher BP1750I-400) was added per tube, the sample was incubated at 65°C for 4 hr followed by a second addition of 50 µL 10 mg/mL Proteinase K and overnight incubation at 65°C on a thermocycler (with 900 rpm for 30 sec every 4 min).

DNA purification: Tubes were cooled at room temperature, at this stage each tube contains ~ 8.21 mL final volume. The samples from every two tubes were combined to a 50 mL conical tube (~16,42 mL) to have five tubes in total. DNA was extracted by adding an equal volume of 17 mL of saturated phenol pH8.0: chloroform (1:1) (Fisher BP1750I-400) and vortexing for 3 min. Then the mix was transferred to a 15 mL phase-lock tube (Quiagen 129073) followed by spinning tubes at 5,000 g for 10 min. The upper phase was taken to a 50 mL tube to start the second extraction. We added an equal volume of 17 mL saturated phenol pH 8.0: chloroform (1:1), vortexing for 1 min. Then the upper phase was transferred to a 15 mL phase-lock tube, and tubes were centrifuged at 5,000 g for 10 min. We pooled all the upper phases from all 5 tubes ~ 85 mL into a single 300 mL high-speed centrifuge tube to precipitate the DNA. 8.5 mL (1/10 volume) of 3M sodium acetate pH 5.2 was added and brief vortexing was performed, then 212 mL (2.5 volumes) of ice-cold 100% ethanol was added, and the tube was inverted slowly several times and incubated at -80° C for 1 hr. Afterward, the DNA was pelleted at 16,000 g for 30 min at 4°C. The supernatant was discarded and the pellet was dissolved in 500 µL 1X TLE and transferred to a 0.5 mL AMICON Ultra Centrifuge filter (UFC5030BK EMD Millipore). The column was centrifuged for 5 min at 14,000 g and the flow-through was discarded. The column was washed 4 times using 450 µL of 1X TLE for desalting DNA. After the final wash, the library remaining in the column (~50 µL) was eluted in 30 µL of 1XTLE, the column was flipped upside down into a new tube to collect DNA by centrifugation for 3 min at 4,000 g. RNA was degraded by adding 1 µL of 10 mg/mL RNAase A and incubation for 30 min at 37°C.

Quality control assessment: to test the quality of the 3C library we used PCR to amplify a specific ligation product formed by two nearby restriction fragments, using the following primers:
GPF33: GACCTCTGCACTAGGAATGGAAGGTTAGCC
GPF23: GACTAATTCCTGACACTACTTGAGGGATAC
The amplicon was digested with EcoRI to assess the efficiency of 3C ligation. 3C primers used for analysis of the beta-globin locus are listed in Supplementary Table S1.

Chromosome Conformation Capture Carbon Copy (5C) Protocol

Starting with a 3C library generated with HindIII we performed 5C as follows:

Annealing: The 5C probes were pooled and combined with the 3C template each reaction contained 800,000 genome copies of 3C template and 0.2 fmol per 5C probe [800,000 genome copies of 3C template, 2 µL of 10X NEB4 (NEB B7004S), 2.75 µL of Salmon Sperm DNA (250 ng; (Invitrogen™ 15632011), 0.25 µL of 1 fmol/µL probes , up to 20 µL ultrapure distilled water]. We set up 8 annealing reactions for each library in a 96-well PCR plate. We then incubated the samples in a PCR machine and ran the following program [95°C for 9 min, Ramp 0.1°C/sec to 55 C, then keep at 55°C for 12 hr].

Ligation: We ligated 5C probe pairs, which represent a specific ligation junction in the 3C library, by adding 20 µL of ligation mix 2 µL of [10X *Taq* DNA ligase buffer (NEB B0208S), 0.25 µL Taq DNA ligase (NEB M0208S), 17.75 uL ultrapure distiller water] while the samples are kept in the PCR block at 55°C. We then incubated the reactions for 1 hr at 55°C followed by a 10 min incubation at 65°C; samples were then cooled to 4°C. Negative controls (no ligase, no template, no 5C oligonucleotide) were included to ensure the absence of any contamination.

PCR amplification: Universal emulsion primers were used for amplification of the ligated product by using 5C forward and reverse emulsion primers [Forward_primer: CCTCTCTATGGGCAGTCGGTGAT. Reverse_primer : CTGCCCCGGGTTCCTCATTCTCT] for 25 PCR cycles [6 µL of ligation product, 2.5 µL of 10XPCR (600 mM Tris-SO_4, pH 8.9, 180 mM $(NH_4)_2SO_4$), 1.8 mM MgCl2, 0.2 mM dNTP, 0.5 µL F-emulsion primer (80 µM), 0.5 µL R-emulsion primer (80 µM), 0.225 µL *AmpliTaq*Ò *Gold* DNA polymerase, ultrapure distilled water to bring volume up to 25 µL]. We then amplified DNA using this PCR program: [95° 9 min, 25 cycles (95°C 30 s, 65°C 30 s, 72° 30 s), 95°C 30 s, 65°C 30 s, 72°C 8 min, 4°C].

We pooled all the PCR reactions for the same library together and concentrated the DNA to 50 µL using 0.5 mL AMICON Ultra Centrifuge filter (UFC5030BK EMD Millipore). DNA was then loaded on a 2% agarose gel, along with a low molecular weight ladder, and the gel was run in a 4°C room at 200 volts for 90 min. The 150 bp DNA that corresponded to the ligated 5C probes was isolated from the gel using the QIAquick Gel Extraction Kit Protocol (QIAGEN 28115). DNA was finally eluted in 32 µL of 1XTLE.

A-tailing: A dATP was added to the 3' ends of the 5C library by adding 18 µL of A-tailing mix [5 µL NEB buffer 2.1, 10 µl of 1 mM dATP, 3 µL Klenow exo (NEB M0212S)] to the 32 µL of DNA sample from the previous step. The reaction was then incubated in a PCR machine [at 37°C for 30 min, then at 65°C for 20 min, and finally cooled down to 4°C]. Next, the tube was placed on ice immediately. 1:1 Ampure was used to remove unligated adaptors. The DNA was finally eluted in 40 µL 1X T4 DNA Ligase buffer (Invitrogen).

Illumina adapter ligation and paired-end PCR: For this step, we used the TruSeq DNA LT kit Set A (REF 15041757). 10 µL of ligation mix [5 µL Illumina paired-end adapters, 3 µL T4 DNA ligase Invitrogen, 2 µL 5x T4 DNA ligase buffer (Invitrogen 5X)] was added to the 40 µL sample from the previous step. The ligation sample was then incubated at RT for 2 hours on a tube rotator. Afterward, the sample was run on a 2% agarose gel in a cold room 4°C at 150 volts for 120 min along with a low molecular weight ladder. The 270 bp band that corresponds to 5C products (150 bp) ligated to the two adaptors (64 bp) was extracted from the gel and isolated using the QIAquick Gel Extraction Kit (QIAGEN 28115). DNA was finally eluted in 30 µL 1XTLE.

Hi-C 2.0 protocol for nuclei

Crosslinking: isolated, undigested, and pre-digested (with liquified chromatin) nuclei were not pelleted after the pre-digestion step above but were crosslinked immediately as follows: for each sample 1,250 µL volume of nuclei in the digestion buffer was transferred to a 21.875 mL mix [625 µL of 37% formaldehyde + 21.25 mL of HBSS]. For intact cells: 5 million K562 cells or nuclei were washed twice with 15 mL of HBSS and pelleted at 300 g for 10 min, then resuspended in 2.5 mL of HBSS. The sample was transferred to 20.625 mL crosslinking mix [625 µL of 37% formaldehyde + 20 mL of HBSS].

All samples were incubated at RT for 10 min on a rocking platform. Next, to stop cross-linking 1.25 mL of 2.5 M glycine was added to each sample and the mix was incubated at RT for 5 min on a rocking platform. To pellet the crosslinked cells/nuclei, the sample was centrifuged at 1,000 g for 10 min at 4°C. The supernatant was discarded and the pellet was washed twice with HBSS before going to the next step or storing samples at -80°C.

Cells lysis: This step is not needed for isolated, undigested, and pre-digested (with liquified chromatin) nuclei. For Hi-C with intact cells: the 5 million crosslinked cells were lysed by adding 1 mL cold lysis buffer [10 mM Tris-HCl (pH=8.0), 10 mM NaCl, 0.2% Igepal CA-630 (NP40)] and 10 µL of 100X Protease inhibitors. The sample was incubated on ice for 15 min to let the cells swell. The cells were lysed on ice using a dounce homogenizer with pestle A (KIMBLE Kontes 885300-0002) by moving the pestle slowly up and down 30 times and incubating on ice for 1 min followed by another 30 strokes. The sample was transferred to a new 1.5 mL microcentrifuge tube, and the sample was centrifuged at 5,000 g at RT for 5 min.

Digestion: from each sample (isolated undigested, and pre-digested (with liquified chromatin) nuclei and lysed cells) the pellet was resuspended in 500 µL of ice-cold 1X NEBuffer 3.1, and pelleted for 5 min at 4,000 g. The pellet was washed twice using 500 µL of ice-cold 1X NEBuffer 3.1. After the last wash, the pellet was resuspended in 350 µL 1X NEBuffer 3.1, and 8 µL was taken and kept at 4°C to assess the DNA integrity later. 38 µL of 1% SDS was added to 342 µL (380 µL total volume), and the mixture was resuspended and incubated for 10 min at 65°C. The tube was placed on ice immediately afterward. Next 43 µL of 10% Triton X-100 was added and the sample was mixed gently by pipetting. The tubes were placed at room temperature and 12 µL of 10X NEBuffer 3.1 was added. Then 400 U of DpnII (R0543L) was added and mixed gently before an overnight incubation at 37°C on a thermocycler (with 900 rpm for 30 sec every 4 min).

Biotin Fill-in: After overnight digestion, the sample was incubated at 65˚C for 20 min in order to inactivate the restriction enzyme. Then, 10 µL of the digested sample was taken and kept at 4°C to assess the digestion efficiency later. DNA ends were marked with biotin-14-dATP by adding 60 µL of biotin fill-in master mix [1XNEB 3.1, 0.25 mM dCTP, 0.25 mM dGTP, 0.25 mM dTTP, 0.25 mM biotin-dATP (ThermoFisher#19524016), 50U Klenow polymerase Polymerase I (NEB M0210L)]. Next, the sample was incubated for 4 h at 23°C on a thermocycler (with 900 rpm for 30 sec every 4 min). Finally, the sample was placed on ice immediately for 15 min before proceeding to the next step.

Ligation: After fill-in, the total sample volume was ~535 µL. Ligation was performed by adding 665 µL of ligation mix [240 µL of 5x ligation buffer (1.8X) (Invitrogen), 120 µL 10% Triton X-100, 12 µL of 10 mg/mL BSA, 50 µL T4 DNA ligase (Invitrogen 15224090), and 243 µL ultrapure distilled water (Invitrogen)], to make a total volume of 1,200 µL. The reaction was then incubated at 16°C for 4 hours in a Thermomixer with interval shake.

Reverse Crosslinking: 50 µL of 10 mg/mL **proteinase K (Fisher BP1750I-400) was added after ligation, the sample was incubated at 65°C for 4** hr followed by a second addition of 50 µL 10 mg/mL Proteinase K and overnight incubation 65°C

DNA purification: Reactions were cooled to room temperature and the 1.3 mL total volume was transferred to a 15 mL tube. The DNA was extracted by adding an equal volume of 1.3 mL of saturated phenol pH 8.0: chloroform (1:1) (Fisher BP1750I-400) and vortexing for 1 min. Then the total volume of 2.6 mL was transferred to a 15 ml phase-lock tube (Quiagen #129065) and tubes were centrifuged at 5,000 g for 10 min. The upper phase was transferred to a 15 mL tube to start the second extraction. An equal volume of 1.3 mL saturated phenol pH8.0: chloroform (1:1) was added and the sample was vortexed for 1 min. Then the mix was transferred to 15 ml phase-lock tube (Quiagen #129065) followed by spinning tubes at 5,000 g for 10 min. The upper phase of ~1.3 mL was transferred to a 15 mL tube (high speed) to precipitate the DNA. 1/10 volume (130 µL) of 3 M sodium acetate pH 5.2 was added and the sample was briefly vortexed. Then, 2.5 volumes of ice-cold 100% ethanol 3.25 mL was added, the tube was inverted slowly several times and then incubated at -80° C for 1hr. Next, the DNA was pelleted at 16,000 g for 30 min at 4°C. The supernatant was discarded and the pellet was dissolved in 500 µL 1X TLE and transferred to a 0.5 mL AMICON Ultra Centrifuge filter (UFC5030BK EMD Millipore). The column was spun for 5 min at 14,000 g and the flow-through was discarded. The column was washed 4 times using 450 µL of 1X TLE for desalting of DNA. After the final wash the DNA remaining in the column (~50 µL) was eluted in 52 µL of 1XTLE. The column flipped upside down into a new tube to collect DNA and spun for 3 min at 4,000 g, the volume was adjusted to 102 µL. RNA was degraded by adding 1 µL of 10 mg/mL RNAase A and incubation for 30 min at 37°C. To quantify the DNA concentration, 2 µL of the final DNA sample along with the first 8 µL sample taken before digestion, the 10 µL sample taken after digestion, and various amounts of the 1 kb ladder (NEB#N3232s) were run on 1% Agarose gel.

Removal of Biotin from unligated ends: To remove biotinylated nucleotides at DNA ends that did not ligate, the Hi-C sample was treated with T4 DNA polymerase. For each Hi-C sample, we assembled the following reaction: [up to 5 µg of Hi-C library, 5 µL 10x NEBuffer 3.1, 0.025 mM dATP, 0.025 mM dGTP and 15 U T4 DNA polymerase (NEB # M0203L). The samples were brought up to 50 µL total volume adding ultrapure distilled water . Reactions were incubated at 20°C for 4 hours, the enzyme was then inactivated by incubation of the reaction for 20 mins at 75°C and placed at 4°C. Next, the samples were pooled and the volume was brought up to 130 µL 1XTLE in preparation for sonication.

Sonication: the DNA was sheared to a size of 100-300 bp using a Covaris instrument [Duty cycle 10%, Intensity 5, Cycles per Burst 200, set Mode Frequency sweeping, continuous degassing, process time 60 sec, Number of cycles] for 3 cycles. The volume was brought up to 500 µL using TLE for Ampure fractionation.

Size fractionation using AMpure XP: 500 µL AMpure beads (Beckman Coulter A63881) were added to a 1.5 mL tube labeled as 1.1X. Then the tube was placed on the Magnetic Particle Separator (MPS) for 5 min, and the supernatant was removed. Beads were resuspended in 150 µL AMpure mixture in order to make the 1.1X solution. 400 µL of AMpure mixture was added to 500 µL of sonicated DNA from the previous step and the tube was labeled 0.8X. The sample was vortexed and spun down briefly followed by incubation at RT for 10 min on a rotating platform. Then the tube was placed on the MPS for 5 min at RT. The supernatants were

collected and added to the 1.1X tube, the tube was briefly vortexed and spun down followed by incubation at RT for 10 min on a rotating platform. Then the tube was placed on the MPS for 5 min at RT. The supernatant was discarded and the beads in 0.8X and 1.1X tubes were washed twice with 1 mL 70% ethanol. Beads were reclaimed by the MPS for 5 min. Beads were then air-dried on the MPS until ethanol had evaporated completely. Next, 51 µL of 1XTLE was added to the 0.8X and 1.1X tubes to resuspend the DNA from the beads. Tubes were incubated at RT on a rotating platform for 10 min. Then the tubes with AMpure beads from both 0.8X and 1.1X tubes were placed on the MPS for 5 min. Finally, the supernatants were transferred to 1.7 mL tubes labeled 0.8X and 1.1X. Our sample with DNA that ranges from 100-300 bp is in the 1.1X sample, the 0.8X sample was kept in case more DNA was needed. DNA from both samples 0.8X and 1.1X were quantified by running 1 µL on a 2% agarose gel along with different amounts of low molecular weight DNA ladder (100 ng, 200 ng, 400 ng).

End Repair*:* 50 µL of Hi-C sample was transferred to a PCR tube, then 20 µL of the end-repair mix [3.5X NEB ligation buffer (NEB B0202S), 17.5 mM dNTP mix, 7.5 U T4 DNA polymerase (NEBM0203L), 25 U T4 polynucleotide kinase (NEB M0201S), 2.5 U Klenow polymerase Polymerase I (NEB M0210L)] was added. The 70 µL total volume reaction was then incubated at 37°C for 30 min, followed by incubation at 65°C for 20 min to inactivate Klenow polymerase, and then the sample was put at 4°C. The volume was brought up to 400 µL using 1X TLE for the next step.

Biotin pull-down: All the following steps were performed with 1.5 mL loBind tubes (Eppendorf 22431021). 15 µL of MyOne streptavidin C1 beads mix (Thermo Fisher 65001) was transferred to a 1.5 mL tube. The beads were washed twice by adding 400 µL of TWB [5 mM Tris-HCl pH8.0, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween20] followed by incubation for 3 min at RT. The tube was then placed on an MPS for 1 min and the supernatant was removed. After the washes, the beads were resuspended in 400 µL of 2X Binding Buffer (BB) [10 mM Tris-HCl pH8, 1 mM EDTA, 2 M NaCl] and mixed with the 400 µL DNA from the previous step in a new 1.5 mL tube. The mixture was incubated for 15 min at RT with rotation, the tube was then placed on the MPS for 1 min and the supernatant was removed. The DNA bound to the beads was washed by adding 400 µL of 1X BB and transferred to a new tube. The beads were reclaimed against the MPS for 1 min, and the supernatant was discarded. The second wash used 100 µL of 1X TLE, beads were reclaimed against MPS for 1 min, and the supernatant was discarded. Finally, the DNA bound to the beads was eluted in 32 µL of 1X TLE.

A-tailing: A dATP was added to the 3' ends by adding 18 µL of A-tailing mix [5 µL NEB buffer 3.1, 10 µL of 1 mM dATP, 3 U Klenow exo (NEB M0212S)] to the 32 µL of DNA sample from the previous step. The reaction was incubated in a PCR machine [at 37°C for 30 min, then at 65°C for 20 min, followed by cool down to 4°C]. Next, the tube was placed on ice immediately. The sample was transferred to a 1.5 mL loBind tube, the tube was placed on the MPS for 1 min and the supernatant was removed. The streptavidin beads bound to DNA were washed twice using 100 µL 1X T4 DNA Ligase Buffer (Invitrogen). Finally, streptavidin beads bound to DNA were resuspended in 40 µL 1X T4 DNA Ligase buffer (Invitrogen).

Illumina adapter ligation and paired-end PCR: For this step, the TruSeq DNA LT kit Set A (REF#15041757) was used. 10 µL of ligation mix [5 µL Illumina paired-end adapters, 3 µL T4 DNA ligase Invitrogen, 2 µL 5x T4 DNA ligase buffer (Invitrogen 5X)] was added to the 40 µL Hi-C sample from the previous step. The ligation sample was then incubated at RT for 2 hours on a

rotator. The sample was transferred to a 1.5 mL loBind tube, the tube was placed on the MPS for 1 min and the supernatant was removed. The streptavidin beads bound to DNA were washed twice with 400 µl of **TWB, then twice** using 100 µL 1X TLE. Finally, the sample was resuspended in 20 µL 0f 1XTLE.
Illumina Truseq Kit for PCR: We performed three trial PCR reactions as follows [2.5 µL DNA bound to beads, 2 µL of Primers mix (TruSeq DNA LT kit Set A 15041757)), 10 µL Master Mix (TruSeq DNA LT kit Set A 15041757), 10.5 µL of ultrapure distilled water (Invitrogen)]. We split the 25 µL over three PCR tubes (5 µL, 5 µL, 15 µL per tube). Each of the three samples was then amplified with different numbers of PCR cycles (6, 8, 10 respectively) to assess the Hi-C library quality: [30 sec at 98°C, n cycles of (30 sec at 98°C, 30 sec at 65°C, 30 sec at 72°C), 5 min at 72°C, hold at 10°C]. 10 µL was taken from the 15 µL sample (with 10 PCR cycles), the 10 µL sample was then digested with ClaI for 1 h by adding 10 µL of digestion mix [1.5 µL **10x NEB Cutsmart buffer, 1.5 µL ClaI** (NEB R0197S), 7 µL ultrapure distilled water]. The 5 µL of each PCR cycle sample along with the 20 µL digested sample, and titration of the low molecular ladder (100 ng, 200 ng, 400 ng) (NEB) were run on a 2% Agarose gel. After digestion with ClaI, a downward shift of the amplified DNA to smaller sizes is expected, which indicates DNA ends were correctly filled in and ligated (creating a ClaI site). The number of PCR cycles to generate the final Hi-C material for deep sequencing was chosen based the minimum number of PCR cycles in the PCR titration that was needed to obtain sufficient amounts of DNA for sequencing using the remaining 17.5 µL Hi-C sample.

DpnII-Seq and FatI-seq protocol
To assess positions of DpnII or FatI digestion during the pre-digestion steps we performed DpnII-seq and FatI-sq as described in detail below. Starting with pre-digested nuclei we performed the following protocol:

Proteinase K: 50 µL of 10 mg/mL proteinase K (ThermoFisher # 25530) was added to each 500 µL pre-digested nuclei sample (2 million nuclei) (See Methods: Pre-digestion) and the 5 tubes were incubated at 65°C for 3 hours**.**
DNA purification: Tubes were cooled to room temperature and all 5 samples were pooled in a single 15 mL tube (2.75 mL total volume). The DNA was extracted by adding an equal volume of 2.75 mL of saturated phenol pH8.0: chloroform (1:1) (Fisher BP1750I-400), followed by vortexing for 1 min. The sample (5.5 mL) was transferred to a 15 mL phase-lock tube (Quiagen #129065) followed by centriguation at 5,000 g for 10 min. The upper phase was transferred to a 15 mL tube to start the second extraction. An equal volume of 2.75 mL saturated phenol pH8.0: chloroform (1:1) was added, followed by vortexing for 1 min. Then the mix was transferred to a 15 mL phase-lock tube (Quiagen #129065) followed by centrifugation at 5,000 g for 10 min. The upper phase of ~ 2.75 mL was transferred to a 15 mL tube (high speed), 1/10 volume (275 µL) 3M sodium acetate pH 5.2 was added followed by brief vortexing and then 2.5 volumes of ice-cold 100% ethanol (6.875 mL) were added. The tube was inverted slowly several times, incubated at -80°C for 1 hr and then DNA was pelleted by centrifugation at 16,000 g for 30 min at 4°C. The supernatant was discarded and the pellet was dissolved in 500 µL 1X NEB3.1 and transferred to a 0.5 mL AMICON Ultra Centrifuge filter (UFC5030BK EMD Millipore). The column was centrifuged for 5 min at 14,000 g and the flow-through was discarded. The column

was washed 4 times using 450 µL of 1X NEB3.1 for desalting of DNA. After the final wash, the library remaining in the column (~50 µL) was eluted in 450 µL of 1XNEB3.1; the column was flipped upside down into a new tube to collect DNA and centrifuged for 3 min at 4,000 g. ~500 µL of DNA was recovered. RNA was degraded by adding 1 µL of 10 mg/mL RNAase A and incubation for 30 min at 37°C. The amount of DNA was estimated by running an aliquot on a 1% Agarose gel along with a 1kb ladder (NEB#N3232s).

Biotin Fill-in: 1XNEB3.1 was added the reaction to a final volume of 680 µL, and then the 680 µL was split over 2 1.5 mL tubes. DNA ends were filled in and marked with biotin-14-dATP. To each tube 60 µL of biotin fill-in master mix was added: [1xNEB2.1, 0.25 mM dCTP, 0.25 mM dGTP, 0.25 mM dTTP, 0.25 mM biotin-dATP (ThermoFisher#19524016), 50 U Klenow polymerase Polymerase I (NEB M0210L)]. Samples were incubated at 37°C in a Thermocycler for 75 mins. Next, the tubes were placed on ice immediately for 15 mins, and samples from the 2 tubes were combined to obtain a final volume ~800 µL. Amicon filters were used to reduce the volume of the final sample from 801 µL to 130 µL.

Sonication: DNA was sonicated to a size of 100 – 300 bp using a Covaris instrument (Duty Cycle 10%, Intensity 5, Cycles per Burst 200, set Mode Frequency sweeping, continuous degassing, process time 60 sec, Number of cycles) for 4 cycles. The 130 µL of sonicated DNA was transferred to a 1.5 mL tube and 1XTLE was added to a total volume of 500 µL. DNA fragment size was determined by running 2 µL of DNA along with low molecular ladder (NEB) on a 2% agarose gel.

Size fractionation using AMpure XP: 500 µL AMpure beads (Beckman Coulter A63881) were added to a 1.5 mL tube labeled as 1.1X. The tube was placed on the MPS for 5 min, and the supernatant was removed. Beads were resuspended in 150 µL AMpure mixture in order to make 1.1X. 400 µL of AMpure mixture was added to 500 µL of sonicated DNA from the previous step and the tube was labeled 0.8X. The sample was vortexed and centrifuged briefly using a tabletop small centrifuge followed by incubation at RT for 10 min on a rocking platform. Then the tube was placed on the MPS for 5 min at RT. The 0.8X supernatants were collected and added to the 1.1X tube, the tube was briefly vortexed and centrifuged followed by incubation at RT for 10 min on a rocking platform. The tube was placed on the MPS for 5 min at RT, and the supernatant discarded. Beads in the 0.8X and 1.1X tubes were washed twice with 1 mL 70% ethanol, reclaiming beads against the MPS for 5 min. Beads on the MPS were then dried until ethanol had evaporated completely. Next, 51 µL of 1XTLE was added to the 0.8X and 1.1X tubes to elute DNA from the beads. Tubes were incubated at RT on a rocking platform for 10 min. The 0.8X and 1.1X tubes were placed on the MPS for 5 min. Finally, the supernatants were transferred to 1.7 mL tubes labeled 0.8X and 1.1X. The 1.1X sample contains DNA that ranges in size from 100-300 bp. The DNA in the 0.8X sample was kept in case more DNA was required, in which case the DNA would be sonicated using 2 cycles followed by a similar round of size fractionation as described above. The amount of DNA from both samples 0.8X and 1.1X was quantified by running 1 µL on a 2% agarose gel along with a titration of low molecular weight DNA ladder (100 ng, 200 ng, 400 ng).

End Repair: 50 µL from the 1.1X sample was transferred to a PCR tube, and 20 µL of end repair mix was added: [3.5X NEB ligation buffer (NEB B0202S), 0.875 mM dNTP mix, 0.375 U/µL T4 DNA polymerase, 1.25 U/µL T4 polynucleotide kinase, 0.125 U/µL Klenow DNA polymerase]. The 70 µl total volume reaction was incubated for 30 min at 20°C in a PCR

machine and then placed on ice. The DNA was purified by 1:2 Ampure, by adding 140 µL 2X Ampure solution to the 70 µL DNA sample followed by incubation for 5 min at RT. The tube was placed on the MPS for 4 min to reclaim the beads and the supernatant was discarded. The beads were washed twice with 1 mL of 70% ethanol while on the MPS. After beads were dried DNA was eluted in 32 µL TLE (pH 8.0) and incubation for 10 min at RT. The supernatant was transferred to a 1.5 mL tube.

A-tailing: A dATP was added to the 3' ends by adding 18 µL of A-tailing mix [5 µL NEB buffer 3.1, 10 µL of 1 mM dATP, 3 U Klenow exo (NEB M0212S)] to the 32 µL of DNA sample from the previous step. The reaction was then incubated in a PCR machine at 37°C for 30 min followed by incubation 65°C for 20 min and cooling down to 4°C. The tube was placed on ice. The volume was brought to 100 µL by adding 1X NEB2.1. The DNA was then purified by adding 1:2 Ampure mix (200 µL of Ampure was added to the 100 µL final DNA volume). Finally, the DNA was eluted in 40 µL of 1X T4 DNA ligase buffer (Invitrogen 5X).

Illumina adapter ligation and paired-end PCR: For this step we used the TruSeq DNA LT kit Set A (REF#15041757). 50 µL of ligation mix [25 µL Illumina paired-end adapters, 15 µL T4 DNA ligase Invitrogen, 10 µL 5X T4 DNA ligase buffer (Invitrogen 5X)] was added to the 40 µL sample from the previous step. The ligation sample was then incubated at RT for 2 hours on a rotator. Next, the DNA was purified by adding 1:1 Ampure solution (180 µL of Ampure mix was added to the 90 µL sample), the supernatant was discarded and beads were washed twice with 1 mL of 70% ethanol. After the last wash step, the beads were resuspended in 400 µL of 1X TLE and incubated at RT on a rocking platform for 10 mins. The tube was placed on the MPS for 4 mins. Finally, the 400 µL supernatant was transferred to a new tube.

Biotin pull-down: All the following steps are done using 1.5 mL loBind tube (Eppendorf 22431021). 15 µL of MyOne streptavidin C1 beads mix (Thermo Fisher 65001) was transferred to a 1.5 mL tube. The beads were washed twice with 400 µL of TWB [5 mM Tris-HCl pH8.0, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween20] by incubation for 3 min at RT. After each wash, the tube was placed on the MPS for 1 min and the supernatant was removed. After the washes, the beads were resuspended in 400 µL of 2X Binding Buffer (BB) [10 mM Tris-HCl pH8.0, 1 mM EDTA, 2 M NaCl] and mixed with the 400 µL DNA from the previous step in a new 1.5 mL. The mixture was incubated for 15 min at RT with rotation. The tube was then placed on the MPS for 1 min and the supernatant was removed. The DNA bound to the beads was washed first by adding 400 µL of 1X BB and transferring to a new tube. The beads were reclaimed against the MPS for 1 min, and the supernatant discarded. 100 µL of 1X TLE was added and the beads were reclaimed against the MPS for 1 min, then the supernatant was discarded. Finally, the DNA bound to the beads was eluted in 32.5 µL of 1X TLE**.**

PCR optimization: The Illumina Truseq Kit (DNA LT kit Set A (REF#15041757)) was used for PCR amplification of DNA for DpnII-Seq. The trial PCR reaction was set up as follows: [2.5 µL DNA bound to beads, 2 µL of Primers mix (Truseq kit), 10 µL Master Mix (Truseq kit), 10.5 µL of ultrapure distilled water (Invitrogen)]. The 25 µL was split over four PCR tubes (5 µL/per tube). Each of the four samples was incubated for different PCR cycles (6, 8, 10, or 12 cycles): [30 sec at 98°C, n cycles of (30 sec at 98°C, 30 sec at 65°C, 30 sec at 72°C), 7 min at 72°C, hold at 10°C]. The optimal PCR cycle number needed to get enough DNA for sequencing was determined by running the 4 PCR reactions on a 2% agarose gel along with low molecular ladder titration (100 ng, 200 ng, 400 ng). Three PCR reactions of 50 µL volume were then

performed: [5 µL DNA bound to beads, 4 µL of Primers mix (Truseq kit), 20 µL Master Mix (Truseq kit), 21 µL of ultrapure distilled water (Invitrogen)]. The 3 PCR reactions were pooled together to obtain 150 µL total volume. The samples were reclaimed against the MPS for 1 min, then the PCR products (supernatant) were taken to new 1.5 mL tubes. 1:1 Ampure was performed for removal of primer dimers (150 µL of Ampure and 150 µL DNA sample). Finally, beads were resuspended in 35 µL of TLE to elute the DNA. DNA that remained bound to beads was saved after a first wash using TBW followed by two washes with 1X TLE and then resuspended in 30 µL of 1X TLE.

Fragment size determination after pre-digestion with DpnII

4 million cells were pre-digested for 4 hours using DpnII procedure described above (Pre-digestion of nuclei). The pre-digested nuclei were then treated as follows:

Proteinase K: 50 µL of 10 mg/mL proteinase K (ThermoFisher # 25530) was added to each 500 µL pre-digested nuclei sample (2 million nuclei) (See Methods: Pre-digestion) and the 2 tubes were incubated at 65°C for 3 hours**.**

DNA purification: Tubes were cooled to room temperature and all 2 samples were pooled in a single 15 mL tube (1.1 mL total volume). The DNA was extracted by adding an equal volume of 1.1 of saturated phenol pH8.0: chloroform (1:1) (Fisher BP1750I-400), followed by vortexing for 1 min. The sample (2.2 mL) was transferred to a 15 mL phase-lock tube (Quiagen #129065) followed by centriguation at 5,000 g for 10 min. The upper phase was transferred to a 15 mL tube to start the second extraction. An equal volume of 1.1 mL saturated phenol pH8.0: chloroform (1:1) was added, followed by vortexing for 1 min. Then the mix was transferred to a 15 mL phase-lock tube (Quiagen #129065) followed by centrifugation at 5,000 g for 10 min. The upper phase of ~ 1.1 mL was transferred to a 15 mL tube (high speed), 1/10 volume (110 µL) 3M sodium acetate pH 5.2 was added followed by brief vortexing and then 2.5 volumes of ice-cold 100% ethanol (2.75 mL) were added. The tube was inverted slowly several times, incubated at -80°C for 1 hr and then DNA was pelleted by centrifugation at 16,000 g for 30 min at 4°C. The supernatant was discarded, and the pellet was dissolved in 500 µL 1X NEB3.1 and transferred to a 0.5 mL AMICON Ultra Centrifuge filter (UFC5030BK EMD Millipore). The column was centrifuged for 5 min at 14,000 g and the flow-through was discarded. The column was washed 4 times using 450 µL of 1X NEB3.1 for desalting of DNA. After the final wash, the DNA remaining in the column (~50 µL) was eluted in 70 µL of 1XNEB3.1; the column was flipped upside down into a new tube to collect DNA and centrifuged for 3 min at 4,000 g. ~70 µL of DNA was recovered. RNA was degraded by adding 1 µL of 10 mg/mL RNAase A and incubation for 30 min at 37°C. The amount of DNA was estimated by running an aliquot on a 1% Agarose gel along with a 1kb ladder (NEB#N3232s).

Blunting overhang: DNA ends were filled in by adding 60 µL of fill-in master mix was added: [1xNEB2.1, 0.25 mM dCTP, 0.25 mM dGTP, 0.25 mM dTTP, 0.25 mM dATP (ThermoFisher#19524016), 50 U Klenow polymerase Polymerase I (NEB M0210L)]. Samples were incubated at 37°C in a Thermocycler for 75 mins. Next, the tubes were placed on ice immediately for 15 mins.

Gel Extraction for size-selection: a 1% gel was prepared. 25µL of loading dye (blue dark) was added to the 130µL DNA sample from the previous step. 25µL of 1kb DNA ladder was loaded in

one well and all sample in the remaining wells. After elecrophoresis the DNA was isolated from the gel in three size intervals: less than 1kb, 1kb-3kb, and larger than 3kb.

DNA purification from agarose gel: The DNA was extracted from the agarose gel using GFX™ PCR DNA and Gel Band Purification Kit (GE28-9034-70 Millipore Sigma). After DNA extraction the sample was washed five times with TLE buffer using a 1.5 Amicon column followed by elution with 50 µl TLE. Aliquots of each DNA samples was analyzed on a fragment analyzer.

Sonication: DNA from each size fraction was sonicated to a size of 100 – 300 bp using a Covaris instrument (Duty Cycle 10%, Intensity 5, Cycles per Burst 200, set Mode Frequency sweeping, continuous degassing, process time 60 sec, Number of cycles) for 4 cycles. The 130 µL of sonicated DNA was transferred to a 1.5 mL tube and 1XTLE was added to a total volume of 500 µL. DNA fragment size was determined by running 2 µL of DNA along with low molecular ladder (NEB) on a 2% agarose gel.

Size fractionation using AMpure XP: 500 µL AMpure beads (Beckman Coulter A63881) were added to a 1.5 mL tube labeled as 1.1X. The tube was placed on the MPS for 5 min, and the supernatant was removed. Beads were resuspended in 150 µL AMpure mixture in order to make 1.1X. 400 µL of AMpure mixture was added to 500 µL of sonicated DNA from the previous step and the tube was labeled 0.8X. The sample was vortexed and centrifuged briefly using a tabletop small centrifuge followed by incubation at RT for 10 min on a rocking platform. Then the tube was placed on the MPS for 5 min at RT. The 0.8X supernatants were collected and added to the 1.1X tube, the tube was briefly vortexed and centrifuged followed by incubation at RT for 10 min on a rocking platform. The tube was placed on the MPS for 5 min at RT, and the supernatant discarded. Beads in the 0.8X and 1.1X tubes were washed twice with 1 mL 70% ethanol, reclaiming beads against the MPS for 5 min. Beads on the MPS were then dried until ethanol had evaporated completely. Next, 51 µL of 1XTLE was added to the 0.8X and 1.1X tubes to elute DNA from the beads. Tubes were incubated at RT on a rocking platform for 10 min. The 0.8X and 1.1X tubes were placed on the MPS for 5 min. Finally, the supernatants were transferred to 1.7 mL tubes labeled 0.8X and 1.1X. The 1.1X sample contains DNA that ranges in size from 100-300 bp. The DNA in the 0.8X sample was kept in case more DNA was required, in which case the DNA would be sonicated using 2 cycles followed by a similar round of size fractionation as described above. The amount of DNA from both samples 0.8X and 1.1X was quantified by running 1 µL on a 2% agarose gel along with a titration of low molecular weight DNA ladder (100 ng, 200 ng, 400 ng).

End Repair: 50 µL from the 1.1X sample was transferred to a PCR tube, and 20 µL of end repair mix was added: [3.5X NEB ligation buffer (NEB B0202S), 0.875 mM dNTP mix, 0.375 U/µL T4 DNA polymerase, 1.25 U/µL T4 polynucleotide kinase, 0.125 U/µL Klenow DNA polymerase]. The 70 µl total volume reaction was incubated for 30 min at 20°C in a PCR machine and then placed on ice. The DNA was purified by 1:2 Ampure, by adding 140 µL 2X Ampure solution to the 70 µL DNA sample followed by incubation for 5 min at RT. The tube was placed on the MPS for 4 min to reclaim the beads and the supernatant was discarded. The beads were washed twice with 1 mL of 70% ethanol while on the MPS. After beads were dried DNA was eluted in 32 µL TLE (pH 8.0) and incubation for 10 min at RT. The supernatant was transferred to a 1.5 mL tube.

A-tailing: A dATP was added to the 3' ends by adding 18 µL of A-tailing mix [5 µL NEB buffer 3.1, 10 µL of 1 mM dATP, 3 U Klenow exo (NEB M0212S)] to the 32 µL of DNA sample from the previous step. The reaction was then incubated in a PCR machine at 37°C for 30 min followed by incubation 65°C for 20 min and cooling down to 4°C. The tube was placed on ice. The volume was brought to 100 µL by adding 1X NEB2.1. The DNA was then purified by adding 1:2 Ampure mix (200 µL of Ampure was added to the 100 µL final DNA volume). Finally, the DNA was eluted in 40 µL of 1X T4 DNA ligase buffer (Invitrogen 5X).
Illumina adapter ligation and paired-end PCR: For this step we used the TruSeq DNA LT kit Set A (REF#15041757). 50 µL of ligation mix [25 µL Illumina paired-end adapters, 15 µL T4 DNA ligase Invitrogen, 10 µL 5X T4 DNA ligase buffer (Invitrogen 5X)] was added to the 40 µL sample from the previous step. The ligation sample was then incubated at RT for 2 hours on a rotator. Next, the DNA was purified by adding 1:1 Ampure solution (180 µL of Ampure mix was added to the 90 µL sample), the supernatant was discarded and beads were washed twice with 1 mL of 70% ethanol. After the last wash step, the beads were resuspended in 400 µL of 1X TLE and incubated at RT on a rocking platform for 10 mins. The tube was placed on the MPS for 4 mins. Finally, the 400 µL supernatant was transferred to a new tube.
PCR optimization: The Illumina Truseq Kit (DNA LT kit Set A (REF#15041757)) was used for PCR amplification of DNA for DpnII-Seq. The trial PCR reaction was set up as follows: [2.5 µL DNA bound to beads, 2 µL of Primers mix (Truseq kit), 10 µL Master Mix (Truseq kit), 10.5 µL of ultrapure distilled water (Invitrogen)]. The 25 µL was split over four PCR tubes (5 µL/per tube). Each of the four samples was incubated for different PCR cycles (6, 8, 10, or 12 cycles): [30 sec at 98°C, n cycles of (30 sec at 98°C, 30 sec at 65°C, 30 sec at 72°C), 7 min at 72°C, hold at 10°C]. The optimal PCR cycle number needed to get enough DNA for sequencing was determined by running the 4 PCR reactions on a 2% agarose gel along with low molecular ladder titration (100 ng, 200 ng, 400 ng). Three PCR reactions of 50 µL volume were then performed: [5 µL DNA bound to beads, 4 µL of Primers mix (Truseq kit), 20 µL Master Mix (Truseq kit), 21 µL of ultrapure distilled water (Invitrogen)]. The 3 PCR reactions were pooled together to obtain 150 µL total volume. The samples were reclaimed against the MPS for 1 min, then the PCR products (supernatant) were taken to new 1.5 mL tubes. 1:1 Ampure was performed for removal of primer dimers (150 µL of Ampure and 150 µL DNA sample). Finally, beads were resuspended in 35 µL of TLE to elute the DNA. DNA that remained bound to beads was saved after a first wash using TBW followed by two washes with 1X TLE and then resuspended in 30 µL of 1X TLE.

DATA ANALYSIS, QUANTIFICATION AND STATISTICAL METHODS

3C-PCR analysis

The human β-globin locus is an ideal region to examine looping interactions between enhancers and genes because of the strong looping interactions between the LCR and HBG globin gene in the erythroleukemia cell line K562, which highly expresses the globin genes [8]. 3C libraries were generated from: (1) K562 cells that have an LCR-HBG interaction, (2) GM12878 cells in which the LCR-HBG looping interaction is absent, and (3) beta-globin BAC (ENm009) control to normalize for primer bias. To investigate the interaction between the LCR and HBG gene, 3C

primers from [8] were used. 16 forward primers of 28-33 bp length were designed 40-60 bp upstream of each EcoRI site throughout a 110 kb region around the Beta Globin locus (chr11: 5221788- 5337325). The EcoRI fragment overlapping with the LCR (HS3,4,5) was used as an anchor to detect the interaction frequencies between the LCR and EcoRI fragments throughout the β-globin locus. For each primer pair, triplicate PCR reactions were set up, and the mean of the three was normalized to the BAC signal for the same primer pair before plotting normalized interaction frequency in the y-axis, the distance from EcoRI fragment overlapping with LCR to neighboring EcoRI fragments is plotted in the x-axis. Error bars are the standard error of the mean (SEM). PCR primers are listed in SupplementaryTable S1.

5C data processing

The fastq files for 5C sequencing data were processed as described in https://github.com/dekkerlab/5C-CBFb-SMMHC-Inhib/blob/master/data_processing_steps.md The Fastq files were mapped using novoalign to a reference genome built from the pool of all 277 probes. After mapping, we combined the read-pairs. The results were then transferred to a matrix format, and interactions were filtered as previously described [9,10]. First, interactions that belong to the same EcoRI fragment were removed. Second, outliers that are overrepresented as a result of overamplification were also removed. Outliers were defined as the interactions with a Z-score greater than 20 in all datasets. Third, probes that strongly over or underperform which leads to strongly enriched or depleted interactions in a whole row of interactions, were also removed. The four matrices were then scaled to the same number of total reads. Finally, data were binned at 20 Kb (median) with a sliding window with 2.5 Kb steps.

Hi-C data processing

Hi-C read mapping, filtering, binning and matrix normalization were performed using the cMapping pipeline available at https://github.com/dekkerlab/cMapping [11]. In brief, Hi-C reads were mapped to reference human genome assembly hg19 using an iterative mapping strategy and Bowtie 2 [12]. Successfully mapped reads were then filtered to remove reads mapping to the same restriction fragment and to remove PCR duplicates. Interaction frequency versus distance plots displayed high variance for interactions below 1 kb for all samples. Hence, after mapping of valid pair, we removed all pairs with a genomic distance less than 1 kb. The remaining valid read pairs were then binned to 500 kb, 40 kb, and 10 kb resolution matrices. Outlier bins of these matrices with low signal were assigned values of NA. Then as a bias correction step, matrices were normalized such that the sum of interactions in each row/column are approximately equivalent via an iterative correction procedure (ICE) [13]. Lastly, for comparison between samples, matrices were scaled such that the total interactions for a genome-wide matrix equals one billion for each sample. These ICEd scaled matrices were used for subsequent analyses.

A/B compartments

All reads from Hi-C in control K562 samples were pooled to identify A (active) and B (inactive) compartments in K562 cells. A/B compartments were identified at 40 kb resolution following the procedure described in [14] using matrix2compartment.pl in https://github.com/dekkerlab/cworld-dekker. Briefly, each *cis* interaction matrix was first transformed into a z-score matrix followed

by transformation into a correlation matrix. PCA was performed on the correlation matrix and the first eigenvector (PC1) of the PCA analysis was used to identify compartments for each chromosome. A/B compartments were assigned based on gene density such that the A-compartment was more gene-dense than the B-compartment. Positive PC1 values indicate gene-rich A compartments and negative PC1 values indicate gene-poor B compartments. For chromosome 9 the compartments were called for each chromosome arm separately as PC1 captured preferences for interactions within the same arm as opposed to canonical compartment preferences.

LOS and half-life calculation

To measure the 3D structure changes resulting from DpnII, HindIII, or FatI pre-digestion we quantified the amount of cis interactions lost or gained in a 6 Mb window centered at every 40 kb bin genome wide. We note that we did not observe detectable amounts of DNA in the supernatant after chromatin fractionation indicating the large majority of liquefied DNA remains within the nuclei. Even if some DNA is lost, ICE balancing of Hi-C matrices ensures any biases in sequence coverage are removed. For each 40 kb bin, the percent of interactions occurring within its 6 Mb window (corresponding to interactions less than or equal to 3 Mb in distance either upstream or downstream from 40 kb bin) out of total interactions for the 40 kb bin (cis and trans) was calculated. These 6 Mb cis percentages were calculated for control, DpnII pre-digested, HindIII-pre-digested nuclei, and FatI pre-digested nuclei. The change in 3D structure relative to control using these cis percentages was given by the following loss of structure (LOS) metric:

$$LOS = \frac{Control_{Cis\%} - Predigest_{Cis\%}}{Control_{Cis\%}}$$

Hence, LOS values in the range (0, 1) represent a loss in short range contacts after pre-digestion; LOS values < 0 represent an increase in short range contacts after pre-digest, and an LOS equal to zero would indicate no change in structure after pre-digestion. A window of 6 Mb was chosen as we sought here to quantify interactions disrupted by pre-digestion. Many longer range interactions increased after pre-digestion, potentially due to random ligations of cut fragments that start to mix. Difference noted in A and B stability was preserved when LOS was calculated using cis percentages for entire chromosomes as opposed to a 6 Mb window, however the size of chromosomes did bias results by giving 40 kb bins in small chromosomes greater LOS. We note that any loci that may have been lost from the nuclei will not be included in the Hi-C dataset. LOS represents the relative redistribution of short-range interactions to longer-range and inter-chromosomal interactions for the set of loci that remained contained within the nucleus after pre-digestion and we assume this re-distribution would not be affected by any lost loci. FatI-pre digested libraries were of lower sequencing coverage and hence had a lower signal to noise ratio compared with DpnII and HindIII-pre digested libraries. To reduce noise, we applied a loess based smoothing with an α smoothing parameter of 0.01 to the signal track of LOS for nuclei pre-digested using FatI. Correlations between FatI LOS and PC1 were evident both before and after smoothing. Correlations also remained evident before and after

smoothing between FatI LOS residuals and PC1 residuals corrected for digestion efficiency by FatI-seq.

To quantify the timing of disrupted interactions we generated a half-life track utilizing the Hi-C matrices from the DpnII timecourses. For each 40 kb bin we fit a curve to the LOS of each timepoint following an exponential decay of the form (Extended Data Figure 7C):

$$LOS = a - (b \times e^{-c \times minutes})$$

such that a, b and c are parameters to fit.The half-life, $t_{1/2}$, was defined as the time required to reach half saturation, saturation being the 16 hour timepoint where maximal cis interactions have been lost. Half-life values were then computed for every 40 kb bin genome wide. To remove noisy and less reliable $t_{1/2}$ data, we first removed all extreme outliers bins where the sum of squared residuals (SSR) for the exponential fit was greater than 0.1. Then all bins with an SSR greater than two standard deviations from the mean were deemed as outliers and also removed from analyses.

As LOS and $t_{1/2}$ are both dependent on digestion efficiency we also generated residual LOS and $t_{1/2}$ tracks to account for bin to bin variation in digestion efficiency. We used a moving average approach to calculate residuals for LOS as a function of DpnII-seq signal and also $t_{1/2}$ as a function of DpnII-seq signal since the relationships between these variables were non-linear (Figure 3F left, Extended Data Figure 6C). For both stability metrics LOS and $t_{1/2}$, a sliding window of 200 DpnII-seq signal with a step size of one was used to calculate mean LOS or $t_{1/2}$ signal for each DpnII-seq signal increment (Figure 3F left, Extended Data Figure 6C). Window and step size were selected by manual inspection of moving averages and compromising between over and underfitting. These moving averages were used to calculate residuals such that a positive LOS residual indicates more structure loss than expected by given digestion efficiency and a negative LOS residual indicates less structure loss than expected. As $t_{1/2}$ is inversely related to LOS, positive $t_{1/2}$ residuals indicate less structure loss than expected and negative $t_{1/2}$ residuals indicate more structure loss than expected. Moving averages were also used to generate residuals for DpnII-seq as a function of PC1 and LOS as a function of PC1 (Figure 3G, right)

Similar to the digestion efficiency correction by DpnII-seq, we also used estimated average fragment size as an independent measure to correct LOS and $t_{1/2}$ for biases in digestion efficiency (Extended Data Figure 4). For both stability metrics LOS and $t_{1/2}$, a sliding window of 200 bp with a step size of one was used to calculate mean LOS or $t_{1/2}$ signal for each average fragment size bp increment (Extended Data Figure 4G). These moving averages were then used to calculate LOS residuals or $t_{1/2}$ residuals as in the DpnII-seq correction approach described previously.

DpnII-seq data analysis

Sequenced reads were mapped to the hg19 genome using the Bowtie read aligner [12] and reads mapping to multiple sites of the genome were removed. As expected, a high percentage of

reads mapped precisely to their associated restriction cut site (Extended Data Figure 3C). To remove potential artificial biases, we filtered out paired-end reads from fragments whose start or end coordinate was more than three nucleotides from an appropriate restriction cut site. Filtered reads were then binned to 500 kb or 40 kb resolutions. The K562 cell line has a primarily triploid karyotype with regions of the genome in diploid and tetraploid states. Copy number state assignments for each 500 kb or 40 kb bin were assigned using publicly available K562 copy number data from the Catalogue of Somatic Mutations In Cancer (COSMIC) database (https://cancer.sanger.ac.uk/cell_lines/download). Copy number segments in the COSMIC dataset were identified by PICNIC analysis of Affymetrix SNP6.0 array data (PMID:19837654). Read coverage files at 500 kb and 40 kb were corrected to a genome wide diploid state using the copy number state assignments and dividing coverage by appropriate correction value (diploid = 1, triploid =1.5, tetraploid = 2, etc.) per bin. (Extended Data Figure 3D,E). Final copy number corrected coverage files were used for all downstream analysis. DpnII-seq computational workflow is maintained at https://github.com/tborrman/DpnII-seq

FatI-seq data analysis

Computational workflow for FatI-seq analysis was identical to previously described DpnII-seq analysis, with the exception that FatI restriction sites were used in the filtering step as opposed to DpnII restriction sites. The DpnII-seq workflow maintained at https://github.com/tborrman/DpnII-seq has options for analyzing restriction enzyme-seq experiments using the following enzymes: DpnII, HindIII, and FatI.

DpnII pre-digestion average fragment size analysis

Pre-digestion by DpnII leads to variable DNA fragment sizes across the genome. To estimate the average fragment size for a genomic bin after a 4-hour DpnII pre-digestion, we first separated 4-hour DpnII pre-digestion DNA into three slices: less than 1kb, 1kb-3kb, and larger than 3kb (See previously described above: DpnII Pre-digestion size assessment, Extended Data Figure 4A,B). DNA fragments purified from these slices were sequenced and sequenced read pairs were mapped to the hg19 genome using the Bowtie read aligner [12]. Mapped reads were then binned to 40 kb resolution, normalized for sequencing depth, and corrected for copy number state as in the DpnII-seq workflow (See previously described above: DpnII-seq data analysis). This resulted in a coverage track for each of the three DpnII pre-digested slices: less than 1 kb, 1 kb – 3 kb, and larger than 3kb (Extended Data Figure 4C).

To estimate average fragment size for a given genomic bin we used the following formula:

$$Average\ fragment\ size = \frac{p_1 q_1 s_1 + p_2 q_2 s_2 + p_3 q_3 s_3}{p_1 q_1 + p_2 q_2 + p_3 q_3}$$

such that p_x equals the percent of normalized slice x reads that mapped to bin, q_x equals quantity of slice x fragments (ng/μL), and s_x equals mean size of fragments from slice x (bp). The variable x represents one of the three slice intervals (1: less than 1 kb, 2: 1 kb - 3 kb, and 3: larger than 3 kb). Hence, the average fragment size for a given bin estimates the quantity of fragments from each slice size mapping to the bin over the total quantity of fragments mapped

to the bin. The values for p_x are extracted from our coverage tracks and vary bin to bin, while the values for q_x and s_x are extracted from the Fragment Analyzer analysis and are constants (s_1 = 643 bp, s_2 = 2332 bp, s_3 = 5495 bp, q_1 = 1.6562 ng/μL, q_2 = 2.544 ng/μL, q_3 = 2.4632 ng/μL, Extended Data Figure 4B). The average fragment size track was then used as an independent metric for measuring 4-hour DpnII digestion efficiency as compared to the DpnII-seq signal track.

Subcompartments

Rao et al. (2014) divided the canonical A/B compartments into five primary subcompartments A1, A2, B1, B2, B3 based on each subcompartment's preferential Hi-C interactions in GM12878 cells. Subcompartments were annotated using high resolution (~1 kb) Hi-C data and were shown to display unique genomic and epigenomic profiles. K562 subcompartments were annotated in [15] via the method SNIPER using lower resolution Hi-C data. In short, SNIPER infers subcompartments via a neural network approach to accurately annotate subcompartments using Hi-C datasets with moderate coverage (~500 million mapped read pairs). Xiong et al.'s K562 SNIPER subcompartments showed a substantial conservation with GM12878 annotations from Rao et al. [16] and were also enriched in similar epigenetic features, hence we utilized these SNIPER annotations to compare subcompartment status with chromatin stability. K562 SNIPER subcompartments were annotated at 100 kb resolution. To compare with our 40 kb resolution liquid chromatin Hi-C data, we binned the 100 kb subcompartment annotations to 40kb such that any 40 kb bin overlapping a boundary of two separate subcompartments was assigned a value of NA. Upon piling up K562 subcompartment boundaries, we also found enrichment and depletion of various chromatin features consistent with those described in both Rao et al. [16] and Xiong et al. [15].

Sub-nuclear structures

To assess the effect of sub-nuclear structures on chromatin stability we utilized the extensive genetic and epigenetic data publicly available for K562 cells (Supplementary Table S4).

Fold change over control ChIP-seq tracks for histone modifications, chromatin remodellers, and other various proteins were downloaded from the ENCODE Portal. To compare ChIP-seq data with $t_{1/2}$, or residuals of $t_{1/2}$ after correction for DpnII-signal, we binned the ChIP-seq signal tracks into 40 kb such that each 40 kb bin represented the mean signal found across the bin. Bins with no overlapping signal were designated a value of NA.

To examine the association between methylation state and $t_{1/2}$ or residuals of $t_{1/2}$ after correction for DpnII-signal, we downloaded methylation state at CpG Whole-Genome Bisulfite Sequencing (WGBS) tracks from ENCODE. As the methylation data was mapped to hg38, we used the UCSC LiftOver program to convert coordinates to hg19. Then percentage methylation at CpG sites was binned to 40 kb resolution using the mean.

As there is currently no nucleolus associated domains (NADs) data available for K562, we analyzed a binary NADs state track for the human embryonic fibroblast IMR90 cell line [17]. Dillinger et al. annotated NADs via a two-state hidden Markov model of aCGH data from DNA of

isolated nucleoli. Using these annotated NADs, coverage of each 40 kb bin for NADs was assessed and used for all our downstream analyses.

Mapping of nuclear speckle, nuclear lamina and PolII associated loci for K562 cells was accomplished recently via the TSA-seq protocol [18]. Signal tracks of log2(pull-down/input) were downloaded from GEO and binned to 40 kb as previously described for ChIP-seq files. Microarray data for LaminB1 associated domains identified through the DamID protocol was also available from that study. We used the UCSC LiftOver program to convert coordinates from hg18 to hg19. We then binned the log2(Dam-LaminB1/Dam) signal to 40 kb bins as previously described for ChIP-seq files.

To analyze cell cycle relationship with chromatin stability we downloaded Repli-seq data for K562 cells from ENCODE. Actively replicating regions are quantified as a percentage normalized signal for FACS sorted cells in G1 phase, four stages of S phase (S1-S4) and G2 phase. Signal tracks for Repli-seq data were binned to 40 kb as previously described for ChIP-seq files.

Binning of data was performed using the bedtools/v2.26.0 software. To assess the quality of the publicly downloaded data we generated the spearman correlation matrix of all binned signal tracks (Extended Data Figure S8A). Hierarchical clustering of rows of the correlation matrix position heterochromatic marks (H3K9me3, HP1α, HP1β, NADs, and LADs) near one another as expected. The majority active marks form a larger cluster, with the markers for polycomb regions (H3K27me3, CBX8, BMI1, RNF2, and SUZ12) representing facultative heterochromatin clustered together segregating active from inactive marks.

Gene Expression

To assess the effect of gene expression on chromatin stability we utilized processed gene expression quantifications of total RNA-seq for K562 cells available from ENCODE (Accession ID: ENCFF782PCD). Gene locations were mapped using the hg19 ensGene table from UCSC Table Browser. To compare expression values with 40 kb resolution $t_{1/2}$ or residuals of $t_{1/2}$ after correction for DpnII-signal tracks, fragments per kilobase million (FPKM) values for each gene were binned to 40 kb such that each 40 kb bin represented the mean FPKM for all genes overlapping that bin. Bins without any genes were assigned a value of NA. Binned FPKM >=1 was determined to be a reasonable cutoff for expression by inspection of the full distribution of FPKM values.

Compartmentalization saddle plots

Saddle plot calculations were performed using tools in the cooltools repository: (https://github.com/hms-dbmi/hic-data-analysis-bootcamp/tree/master/notebooks/04_analysis_cooltools-eigenvector-saddle.ipynb).
To measure the strength of compartments, intra-chromosomal interaction frequencies were first normalized by the average interaction frequency at a given genomic distance (observed/expected Hi-C maps) at a resolution of 40 kb. Then the distance corrected interaction frequencies were sorted based on PC1 values of a pair of bins that define a given interaction.

Finally, sorted frequencies were aggregated into 50 by 50 groups according to their PC1 values and averaged to obtain a compartmentalization saddle plot. In a compartmentalization saddle plot, preferential B-B interactions are in the upper left corner, and preferential A-A interactions are in the lower right corner.

Homotypic interaction saddle plots

Intra-chromosomal interactions frequencies between 40 kb bins were normalized by the average interaction frequency at a given genomic distance (observed/expected Hi-C maps). Then, the distance corrected interaction frequencies were sorted based on signal values (TSA-seq, DamID) of a pair of bins that define a given interaction, for a given factor (SON, Lamin). Finally, sorted frequencies were aggregated into 50 by 50 groups according to their signal values and averaged, to obtain homotypic interaction saddle plots. In these plots, pair-wise interactions between loci enriched in factor binding are shown in the lower right corner, and pair-wise interactions between loci not bound by the factor are shown in the upper left corner.

Scaling plot calculation

The script to generate scaling plots was adapted from cooltools (https://github.com/mirnylab/cooltools/tree/master/cooltools). Genome-wide curves of normalized contact frequency $P(s)$ is plotted as a function of genomic distance for all intra-chromosomal interactions. Each library was normalized by total number of valid interactions

Mean z-score heatmap

Each genome wide 40kb signal vector for a sub-nuclear structure was cleaned for outliers above three standard deviations of the vector's mean. Each cleaned vector was z-score transformed and then partitioned based on the different $t_{1/2}$ residual intervals for associated bins. The mean z-score for all bins within a given $t_{1/2}$ residual interval is plotted as a square in the heatmap.

References

1. de Gennes, P.-G. *Scaling theory of polymer physics*, (Cornell University Press, 1979).
2. Marko, J.F. & Siggia, E.D. Polymer models of meiotic and mitotic chromosomes. *Mol Biol Cell.* **8**, 2217-2231 (1997).
3. Leibler, L. Theory of microphase separation in block copolymers. *Macromolecules* **13**, 1602-1617 (1980).
4. Matsen, M.W. & Schick, M. Stable and unstable phases of a linear multiblcok copolymer melt. *Macromolecules* **27**, 7157-7163 (1994).
5. Marko, J.F. Microphase separation of block copolymer rings. *Macromolecules* **26**, 1442-1444 (1993).
6. Marko, J.F. & Witten, T.A. Phase separation in a grafted polymer layer. *Phys Rev Lett* **66**, 1541-1544 (1991).
7. MacPherson, Q., Beltran, B. & Spakowitz, A.J. Bottom-up modeling of chromatin segregation due to epigenetic modifications. *Proc Natl Acad Sci U S A* **115**, 12739-12744 (2018).
8. Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): A Massively Parallel Solution for Mapping Interactions between Genomic Elements. *Genome Res.* **16**, 1299-1309 (2006).

9. Lajoie, B.R., van Berkum, N.L., Sanyal, A. & Dekker, J. My5C: web tools for chromosome conformation capture studies. *Nat. Methods* **6**, 690-691 (2009).
10. Sanyal, A., Lajoie, B.R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109-113 (2012).
11. Lajoie, B.R., Dekker, J. & Kaplan, N. The Hitchhiker's guide to Hi-C analysis: Practical guidelines. *Methods* **72**, 65-75 (2015).
12. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
13. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nature Methods* **9**, 999-1003 (2012).
14. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293 (2009).
15. Xiong, K. & Ma, J. Revealing Hi-C subcompartments by imputing high-resolution inter-chromosomal interactions. *BioRxiv* **doi: https://doi.org/10.1101/505503**(2019).
16. Rao, S.S.P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).
17. Dillinger, S., Straub, T. & Nemeth, A. Nucleolus association of chromosomal domains is largely maintained in cellular senescence despite massive nuclear reorganisation. *PLoS One* **12**, e0178821 (2017).
18. Chen, Y. *et al.* Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *J Cell Biol* **217**, 4025-4048 (2018).