

Supplemental Methods

Cell culture

K562 cells were cultured in RPMI-1640 medium supplemented with 10% FBS (Sigma-Aldrich F2442) and 1% 100X Antibiotic-Antimycotic (GIBCO). Cells were seeded at 0.1 million/mL density and passaged or harvested at around 0.75 million/mL density. H1 cells were cultured in Matrigel (Corning 354277, Lot #7128002)- coated flasks in mTeSR medium (STEMCELL Tech 85850). Cells were passaged (or harvested) at optimal density and colony size (see SOP) every 4 or 5 days with 1 in 10 to 1 in 20 splits by digesting colonies into 50-200 μm aggregates using ReLeSR (STEMCELL Tech 05872). HFFc6 cells were cultured in DMEM medium supplemented with 20% heat-inactivated FBS (VWR 97068-091, Lot #035B15). Cells were passaged (or harvested) at 70%-80% confluency without significant drop of mitotic cell ratio with 1 in 2 splits using Trypsin-EDTA (0.05%) (Fisher Sci 25300054). HCT116 cells were cultured in McCoy's 5A Medium supplemented with 10% heat-inactivated FBS (VWR 97068-091, Lot #035B15). Cells were passaged (or harvested) at 70%-80% confluency (around 0.4 million cells/cm²) and seeded at $4 - 5 \times 10^4$ cells/cm² using Trypsin-EDTA (0.05%) (Fisher Sci 25300054).

Coverslip TSA staining

K562 cells were plated on poly-L-lysine (Sigma-Aldrich P4707, 70,000-150,000 M.W., 0.01% w/v) coated coverslips (Fisher Sci 12-545-81) with 0.3-0.5 mL at 0.4-0.7 million/mL cell density. Cells were cultured for 30 mins for attachment. HFFc6 cells were plated on coverslips 1 or 2 days before experiments and harvested at ~80% confluency. Cells were fixed with 1.6% freshly-made paraformaldehyde (PFA) (Sigma-Aldrich P6148) in PBS at room temperature (RT) for 20 mins. Cells were then permeabilized with 0.5% Triton X-100 (Sigma-Aldrich T8787) in PBS (0.5% PBST) at RT for 30 mins, treated with 1.5% H₂O₂ in PBS at RT for 1 hr to quench

endogenous peroxidases, and rinsed with 0.1% Triton X-100 (Sigma-Aldrich T8787) in PBS (0.1% PBST) 3x at RT. Cells were blocked with 5% normal goat serum (Sigma-Aldrich G9023) in 0.1% PBST (GS blocking buffer) at RT for 1 hr and then incubated with rabbit anti-SON polyclonal antibody (Chen et al. 2018) (Pacific Immunology Corp, custom-raised) 1:2000 in GS blocking buffer at RT for 5 hrs. Cells were then washed with 0.1% PBST at RT 3 x 5 mins and incubated with HRP conjugated goat anti-rabbit polyclonal antibody (Jackson Immuno 111-035-144) 1:1000 in GS blocking buffer at RT for 5 hrs or at 4 °C for 10-12 hrs. Cells were then washed with 0.1% PBST at RT for 3 x 5 mins and subject to TSA labeling.

For Condition A, the reaction solution is 50% sucrose (w/v), 1/10000 tyramide-biotin or tyramide-FITC (v/v), and 0.0015% H₂O₂ (v/v) in PBS. Tyramide-biotin was prepared as previously described (Chen et al. 2018). Tyramide-FITC was prepared according to an online protocol (http://wiki.xenbase.org/xenwiki/index.php/Flourescin_Tyramide_Synthesis). Labeling is done at RT for 10 mins. For Condition E, the reaction solution is 50% sucrose, 1/300 tyramide-biotin and 0.0015% H₂O₂ in PBS. Labeling is done at RT for 30 mins. For both conditions, 500 µL reaction solution was applied per coverslip.

For “two-rounds” of TSA labeling, after the 1st-round of TSA, cells were washed with 0.1% PBST at RT for 3 x 5 mins and then subject to a 2nd-round of TSA. After TSA labeling, cells were washed with 0.1% PBST at RT for 3 x 5 mins, stained with Streptavidin-Alexa Fluor 594 (Invitrogen) 1:200 and goat anti-rabbit – Alexa Fluor 647 (Jackson Immuno) 1:200 in GS blocking buffer at RT for 2 hrs or at 4 °C for 10-12 hrs, and then washed with 0.1% PBST at RT for 3 x 5 mins. Coverslips were mounted in DAPI containing, anti-fading medium (0.3 µg/ml DAPI (Sigma-Aldrich)/10% w/v Mowiol 4-88(EMD Millipore)/1% w/v DABCO (Sigma-Aldrich)/25% glycerol/0.1 M Tris, pH 8.5).

TSA-seq

The TSA-seq procedure was modified from our previous publication (Chen et al. 2018).

For suspension cells (K562), cells were fixed by adding 8% freshly made PFA in PBS to reach a final concentration of 1.6% and incubated at RT for 20 mins. Aldehyde groups were quenched by adding 1.25M (10x) glycine in PBS and mixing at RT for 5mins. Cells were permeabilized with 0.5% PBST at RT for 30 mins, centrifuged at 116 g, and re-suspended in PBS. H₂O₂/PBS was added to reach a final H₂O₂ 1.5% concentration to quench endogenous peroxidases in a volume of 1 mL per 3 million cells; the cell suspension was incubated by slowly nutating at RT for 1 hr (open tubes 2 or 3 times during the incubation to release the generated gas). Cells were rinsed 3x with 0.1% PBST, blocked with 5% normal goat serum (Sigma-Aldrich G9023) in 0.1% PBST (GS blocking buffer) in a volume of 1 mL / 10 million cells at RT for 1 hr, and then incubated with rabbit anti-SON polyclonal antibody (Chen et al. 2018) (Pacific Immunology Corp, custom-raised) 1:2000 in GS blocking buffer at 1mL / 10 million cells at 4 °C for 20-24 hrs. Cells were then washed with 0.1% PBST at RT for 3 x 5 mins and incubated with HRP conjugated goat anti-rabbit polyclonal antibody (Jackson Immuno 111-035-144) 1:1000 in blocking buffer in a volume of 1 mL / 10 million cells at 4 °C for 20-24 hrs. Cells were washed with 0.1% PBST for 3 x 5 mins, washed with PBS for 5 mins at RT, and then subjected to TSA labeling. Cells were resuspended in 50% sucrose/PBS and then the same volume of 50% sucrose/PBS containing tyramide-biotin and hydrogen peroxide was added to reach the final concentrations for a specific labeling condition (Supplemental Fig S1A). The final volume of reaction solution was 1 mL per 10 million. Cells were gently nutated at RT during the TSA labeling time specific to each condition (Supplemental Fig S1A). Cells were then washed at RT with 0.1% PBST 3 x 5 mins and with PBS for 5 mins. For each sample, a small portion of cells were attached to a coverslip for

anti-biotin and anti-SON staining to visualize the TSA labeling. Remaining cells were pelleted and either immediately subjected to genomic DNA isolation or stored at -80 °C for later DNA isolation. K562 cells were lysed with high T-E buffer (10 mM Tris and 10 mM EDTA, pH 8.0) containing 0.5% SDS and 0.2 mg/mL Proteinase K (NEB P8107S). All centrifugations prior to TSA labeling were low-speed at 116 – 130 g for 5 – 10 mins to preserve cell structure.

For the TSA-seq mapping of K562 cells upon heat shock, cell flasks were incubated in the 37 °C incubator or in 42 °C water bath for heat shock. One replicate was done with TSA-seq 2.0 for control and heat shock of 30 mins, 1hr, or 2hrs. A second replicate was done with TSA-seq 1.0 for control and heat shock of 30 mins. The heat-shocked or control cells were immediately fixed and proceeded for TSA-seq procedure as described above.

For attached cells (H1, HFFc6, HCT116), cells were grown in tissue culture flasks and fixed by quickly pouring away growth media, adding freshly made 1.6% PFA in PBS, and incubating at RT for 20 mins. Cells were rinsed with PBS, and then washed/permeabilized with 0.5% PBST at RT for 3 x 5 mins. Free aldehyde groups were quenched with 20 mM glycine in PBS at RT for 3 x 5 mins. Cells were then washed with PBS and incubated with 1.5% H₂O₂ in PBS at RT for 1 hr. Cells were rinsed 3x with PBS, blocked with 5% normal goat serum (Sigma-Aldrich G9023) in PBS in a volume of 1mL per 25 cm² flask surface area at RT for 1 hr, and then incubated with rabbit anti-SON polyclonal antibody (Chen et al. 2018) (Pacific Immunology Corp, custom-raised) 1:2000 in 0.1% PBST in a volume of 1 mL per 25 cm² at 4 °C for 10-12 hrs. Cells were then washed with 0.1% PBST at RT for 3 x 5 mins and incubated with HRP conjugated goat anti-rabbit polyclonal antibody (Jackson Immuno 111-035-144) 1:1000 in 0.1% PBST in a volume of 1 mL per 25 cm² at 4 °C for 10-12 hrs. Cells were washed at RT with 0.1% PBST for 3 x 5 mins and with PBS for 5 mins. Cells were TSA-labeled with a working solution of 50% sucrose/PBS

containing 0.0015% H₂O₂ and the condition-specific tyramide-biotin concentration (Supplemental Fig S1A) in a volume of 0.8mL per 10 cm² surface area. Cells were incubated at RT for the condition-specific time (Supplemental Fig S1A) and then washed at RT with 0.1% PBST 3 x 5 mins and with PBS for 5 mins. For each staining, a small portion of attached cells were scraped off and loaded onto coverslips for anti-biotin and anti-SON immunostaining to visualize the TSA-labeling. Remaining cells were washed with high T-E buffer (10 mM Tris and 10 mM EDTA, pH 8.0) at RT for 5 mins and lysed with high T-E buffer containing 1% SDS and 0.2 mg/mL Proteinase K. Cell lysates were collected and immediately subjected to genomic DNA extraction. All incubation and washing steps before cell-lysing were done by gently shaking the original flasks.

Cells on coverslips were stained with Streptavidin-Alexa Fluor 594 (Invitrogen) 1:200 and goat anti-rabbit – FITC (Jackson Immuno) 1:500 in GS blocking buffer at RT for 2 hrs or at 4 °C for 10-12 hrs. Cells were then washed with 0.1% PBST at RT for 3 x 5 mins. Coverslips were mounted in DAPI containing, anti-fading medium (0.3 µg/ml DAPI [Sigma-Aldrich]/10% w/v Mowiol 4-88 [EMD Millipore]/1% w/v DABCO [Sigma-Aldrich]/25% glycerol/0.1 M Tris, pH 8.5).

Genomic DNA was extracted by phenol/chloroform as previously described (Chen et al. 2018). For samples with low DNA concentration, glycogen (Roche) was added to a final concentration of 0.05 mg/mL to facilitate ethanol precipitation. Isolated DNA was fragmented to 100-600 bp using a Bioruptor Pico (Diagenode) machine with a mode of 30 sec ON – 30 sec OFF. Inspection of DNA gels determined the number of required sonication cycles. Biotin labelled DNA fragments were isolated with streptavidin beads as previously described (Chen et al. 2018).

Sequencing libraries were constructed using the TruSeq ChIP Sample Prep Kit (Illumina, IP-202-1012) for H1, HCT116, HFFc6 samples, or the Hyper Library Construction Kit (Kapa Bio) for K562 Condition A, B, C samples, or a protocol developed in the John Lis laboratory (Cornell

University) for K562 Condition D, E samples. The John Lis lab protocol includes steps of end repair (End-It DNA End-Repair Kit: Epicentre Biotech ER0720), 3'-end A tailing (Klenow Fragment: NEB M0212S), and indexed adaptor ligation (NEXTflex ChIP-seq Barcodes-6: Bioo Sci 514120; T4 DNA Ligase [Rapid]: Enzymatics Inc L6030-HC-F). Libraries were amplified by 8-12 PCR cycles. qPCR was used to measure the concentrations of each library which were then pooled at equimolar concentration for each lane and sequenced for 101 cycles from one end of the fragments on a HiSeq4000 using a HiSeq4000 sequencing kit version 1 (Illumina). Fastq files were generated and demultiplexed with the bcl2fastq Conversion Software (Illumina).

Dot blot and DNA biotinylation estimation

Biotinylation levels of DNA were assayed by dot blot prior to biotin-avidin pulldown of DNA. 5-fold serial dilutions of a biotin end-labeled 250 bp fragment were used as biotin concentration standards. The PCR product was produced from *Drosophila* DNA cloned within a BAC (BACR48E12) using primers: GAAACATCGC/iBiodT/GCCCATAAT (forward) and AGAAGCAGCTACGCTCCTCA (reverse) resulting in 1 biotin per fragment. These PCR standards were combined with control, unbiotinylated, sonicated K562 genomic DNA (100-600 bp) to reach a final DNA concentration the same as the test DNA sample (400 or 600 ng/ μ L) to eliminate concentration effects on DNA crosslinking to membranes.

DNA was spotted with 1 or 1.5 μ L (same volume for all standards and samples in a specific experiment) onto a nitrocellulose membrane (0.45 μ m; Bio-Rad), and were UV cross-linked to the membrane (0.24J; UV Stratalinker 2400; Agilent Technologies). The membrane was blocked with SuperBlock (TBS) blocking buffer (Thermo) with 0.05% Tween-20 (Fisher) at RT for 1 hr, and then incubated with streptavidin-HRP (Invitrogen 43-4323) diluted 1:10000 in blocking buffer at 4 °C for 1-3 hrs or overnight. The membrane was then washed with 0.05% Tween-20 in

TBS at RT for 6 x 5 mins with rigorous shaking. The membrane was treated with SuperSignal West Femto chemiluminescent substrate (Thermo) and developed with HyBlot CL film (Denville) or in an iBright machine (Invitrogen). Sample biotinylation level (kilobases of DNA with one biotin labeling on average) was calculated following the equation:

$$\frac{1}{l_{std}} \times c_{std} = \frac{1}{l_{smp}} \times c_{smp}$$

l_{std} : standard DNA length per biotin (0.25 kb). l_{smp} : average sample DNA length (kb) per biotin (to be calculated). c_{std} : concentration of the 0.25kb standard DNA that has the same signal intensity with the tested sample DNA. c_{smp} : concentration of the tested sample DNA.

TSA-seq data processing

We used a similar pipeline as in a previous paper (Chen et al. 2018) to process TSA-seq data. Briefly, we mapped raw sequencing reads to the human reference genome (hg38, Chromosome Y excluded for female cell line K562) using Bowtie2 (Langmead and Salzberg 2012) (version 2.1.0) with default parameters. We applied the rmdup command from SAMtools (Li et al. 2009) (version 1.5) to remove potential PCR duplicates in the alignment. We then used these alignment files as the input files for TSA-seq normalization.

TSA-seq normalization followed a similar scheme as used in the previous paper (Chen et al. 2018), using matched pulldown and input data with the following updates: First, instead of using sliding windows, we used separate genomic bins of 20kb. Each mapped read is exclusively assigned to a window according to its largest aligned position in the reference genome. Second, we modified the calculation of TSA-seq enrichment score, as described in the next paragraph.

N_{TSA} and N_{input} represent the original number of mapped reads in each 20kb bin in the pulldown and input samples, respectively. N_{TSA} in each bin (bins without mapped reads are skipped) was normalized by input read number, according to equation (1), where $Ave(N_{input})$ is the average value of N_{input} across genome-wide windows calculated by dividing the genome-wide sum of N_{input} by the number of bins containing mapped reads:

$$N'_{TSA} = \frac{N_{TSA} \times Ave(N_{input})}{N_{input}} \quad (1)$$

The TSA-seq enrichment score is then defined as the \log_2 ratio between N'_{TSA} and the genome-wide average of N'_{TSA} ($Ave(N'_{TSA})$), calculated using the number of bins containing non-zero numbers of mapped reads, as per equation 2:

$$\text{TSA-seq enrichment score} = \log_2 \left(\frac{N'_{TSA}}{Ave(N'_{TSA})} \right) \quad (2)$$

This normalized TSA-seq enrichment score can be considered as the \log_2 ratio of relative enrichment or depletion of DNA in a specific bin relative to the average pulldown value.

For subsequent analyses, the non-overlapping 20kb binned signals were smoothed by convolution using a Hanning window of length 21 (21 × 20 kb or 420 kb).

K562 pulldown data from all five conditions were normalized using one input constructed from fragmented K562 genomic DNA. H1, HCT116 and HFFc6 data were normalized using separate input libraries made from the DNA used for pulldown for each TSA-labeling experiment.

“Hybrid” method for TSA distance prediction

To convert TSA-seq scores to mean speckle distances, we previously (Chen et al. 2018) fit the TSA-seq enrichment ratios, y (the TSA-seq ratio prior to the \log_2 operation in the TSA-seq

normalization procedure), and mean speckle distances, x (measured by 3D immuno-FISH), corresponding to multiple genomic loci, to the calibration equation $y = y_0 + Ae^{R_0x}$, to obtain the constants y_0 , A and R_0 . Here we applied a new “hybrid” method, which instead estimates R_0 using FISH measurements and estimates y_0 and A using the minimum and maximum TSA-seq enrichment ratios.

To estimate the constants in the calibration equation, $y = y_0 + Ae^{R_0x}$, first we obtained the exponential decay parameter R_0 by fitting 16 FISH measurements from our previously published data (Chen et al. 2018) to the new TSA-seq data (Supplemental Table S1). For each FISH probe, we used the previously published mean cytological distance to speckles based on measurements of 100 alleles (Chen et al. 2018) (16 probes, Supplemental Table S1). We used the smoothed TSA-seq data and calculated the mean TSA-seq enrichment values over the genomic regions cloned within the BACs used to generate FISH probes. We fit the 16 TSA-seq fold-enrichment values, y , and their corresponding mean speckle distances, x , to the exponential function $y = y_0 + Ae^{R_0x}$ using OriginPro software (OriginLab) to obtain the exponential parameter, R_0 .

Next, we estimated y_0 and A based on the minimum and maximum TSA-seq fold-enrichment values, y_{min} and y_{max} : $\lim_{x \rightarrow \infty(x_{max})} (y_0 + Ae^{R_0x}) = y_0 = y_{min}$ ($R_0 < 0$) and $\lim_{x \rightarrow 0} (y_0 + Ae^{R_0x}) = y_0 + A = y_{max}$. Using these three estimated constants (R_0 , y_0 and A) (Supplemental Table S2), we converted TSA-seq enrichment values from all 20 kb bins of smoothed TSA-seq data into speckle distances by using the inverse equation, solving for x in the original exponential function: $x = \frac{1}{R_0} \ln \frac{y - y_0}{A}$.

Statistical method to identify “changed” genomic domains in pair-wise cell type comparisons

We adapted a previously published method (Peric-Hupkes et al. 2010) to test whether a genomic bin (20kb) is statistically different, or “changed”, from the values of the biological replicates of the two different cell lines.

For each dataset, we first rescaled TSA-seq enrichment scores (20kb bin) linearly between their min and max values to a new 1-100 scale based on equation (3) and rounded up to integers with min assigned as 1 instead of 0.

$$\text{Scaled enrichment score (bin } i) = \frac{\text{TSA-seq enrichment score (bin } i) - \text{min}}{\text{max} - \text{min}} \times 100 \quad (3)$$

To reduce the influence of outliers on this rescaling, we used a large and a small percentile of all ranked values as the max and min values, respectively (e.g. 99.95th and 0.05th percentile for HFFc6 and H1 comparison).

For a pair-wise comparison between two cell lines, we rescaled the SON TSA-seq scores for two biological replicates for each cell line:

$$\text{TSA}_{\text{H1}_{rep1}}, \text{TSA}_{\text{H1}_{rep2}}, \text{TSA}_{\text{HFFc6}_{rep1}}, \text{TSA}_{\text{HFFc6}_{rep2}}$$

For example, $\text{TSA}_{\text{H1}_{rep1}}$ denotes the H1 TSA-seq biological replicate 1 and is a row vector of N values from all 20kb non-overlapping bins with mapped reads in the genome. Scatterplots show near uniform data noise across the genome for both cell lines.

We averaged the replicates for the same cell line and used the residual, Δ , between two cell lines for comparison:

$$\Delta = \frac{1}{2} \{ (\text{TSA}_{\text{HFFc6}_{rep1}} + \text{TSA}_{\text{HFFc6}_{rep2}}) - (\text{TSA}_{\text{H1}_{rep1}} + \text{TSA}_{\text{H1}_{rep2}}) \} \quad (4)$$

We define data variance as the difference between biological replicates for the same cell line. We averaged the variance between the two cell lines to be compared to construct a vector Φ including all possible orderings with a length of $4N$, where N = number of genomic bins excluding unmapped regions from hg38 genome:

$$\begin{aligned} \Phi = \frac{1}{2} \{ & (TSA_{H1_{rep1}} - TSA_{H1_{rep2}}) + (TSA_{HFFc6_{rep1}} - TSA_{HFFc6_{rep2}}), \\ & (TSA_{H1_{rep1}} - TSA_{H1_{rep2}}) + (TSA_{HFFc6_{rep2}} - TSA_{HFFc6_{rep1}}), \\ & (TSA_{H1_{rep2}} - TSA_{H1_{rep1}}) + (TSA_{HFFc6_{rep1}} - TSA_{HFFc6_{rep2}}), \\ & (TSA_{H1_{rep2}} - TSA_{H1_{rep1}}) + (TSA_{HFFc6_{rep2}} - TSA_{HFFc6_{rep1}}) \} \quad (5) \end{aligned}$$

The vector Φ serves as the null distribution against which to test single bin value difference from Δ for statistical significance. Φ can be viewed as a set of observations with random variable ϕ , which is approximately Gaussian distributed with parameters μ and σ^2 .

We test if a single element in Δ (a 20kb bin score) shows a significant change between the two cell lines to be compared by comparing with ϕ . We calculated p-values for both tails of the distribution separately: for equation (4), we separate the bins with values above or below 0 to obtain bins with bigger TSA scores in either HFFc6 or H1. We calculated p-values for all 20kb bins and displayed $-\log_{10}$ (p-value) in the genomic tracks.

To identify genomic regions that showed statistically significant changes in position, we set a cutoff p-value of 0.01 and identified all 20kb bins with p-values < 0.01 for either ordering (bigger TSA-seq values in one cell line or the other). For each ordering, we merged the identified adjacent bins to call domains that changed location. We set a second cutoff to call only domains corresponding to 100 kb or larger. The p-value threshold of 0.01 was quite stringent considering

we were requiring 5 adjacent 20 kb bins, each with a p-value less than this threshold, to call regions that showed significant changes in speckle position.

To identify changes relative to SPADs, we defined SPADs as regions with mean TSA-seq score in the 96-100th genomic percentile. We then identified those SPAD regions with either significantly increased TSA-seq scores in one cell line within this top 5th percentile range or SPAD regions which showed statistically significant changes between cell lines but which were classified as non-SPAD in one of the two cell lines compared.

RNA-seq data processing

We applied the ENCODE RNA-seq processing pipeline (<https://github.com/ENCODE-DCC/rna-seq-pipeline>) to process RNA-seq data (Supplemental Table S5). Briefly, we used STAR (Dobin et al. 2013) (version 2.5.1b) to map raw sequencing reads to the human reference genome (<https://www.encodeproject.org/files/ENCFF742NER/>). Next, we used RSEM (Li and Dewey 2011) (version 1.2.26) to quantify gene expression using the ENCODE gene annotation file (URL: <https://www.encodeproject.org/files/ENCFF940AZB/>). We used the mRNA fragments per kilobase of transcript per million mapped reads (FPKM) for each gene reported by RSEM for downstream analysis. We generated Reads per million (RPM) tracks for each RNA-seq dataset using the STAR parameter “--outWigNorm RPM”.

Correlating RNA-seq with TSA-seq data

To correlate TSA-seq score with gene expression, we ranked TSA-seq enrichment scores of genome-wide 20kb bins (bins without mapped reads were removed) from largest to smallest, divided them into 20 equal sized groups using the cut function in R, and named them as vigintiles from vigintile 1 to vigintile 20 (smallest to largest). For each protein-coding gene

(based on the GENCODE annotation version 24), we calculated the average TSA-seq enrichment score across the whole gene region and assigned the gene to the corresponding TSA-seq vigintile group according to the TSA-seq enrichment score ranges for each vigintile. RNA-seq analysis and TSA-seq correlation results were summarized in Supplemental Table S6.

A housekeeping gene list was downloaded from https://www.tau.ac.il/~elieis/HKG/HK_genes.txt (Eisenberg and Levanon 2013) and processed as previously described (Chen et al. 2018) but using hg38 RefSeq gene annotation. 3791 protein-coding genes were identified as housekeeping genes, and the remaining 16541 protein-coding genes were determined as non-housekeeping genes.

To correlate repositioned regions with expression differences between two cell lines, first we identified all genes located within these genomic regions by overlapping gene and region coordinates (the whole gene must be located within the region to be called). Then we calculated the \log_2 ratio between the gene FPKM values in the two cell lines and plotted these \log_2 ratios against the region mean scaled TSA-seq score (max-min normalized, 1-100) change between the two cell lines.

For differential expression analysis between H1 and HFF (RNA-seq datasets summarized in Supplemental Table S5), we used STAR (Dobin et al. 2013) (version 2.5.3a) to map the raw sequencing reads using the index files from the ENCODE project (URL: <https://www.encodeproject.org/files/ENCFF742NER/>). Next, we used htseq-count function of HTSeq (Anders et al. 2014) (version 0.9.1) to count raw read number for each gene with the gencode.v24 annotation (https://www.encodeproject.org/files/gencode.v24.primary_assembly.annotation/). To identify differentially expressed (DE) protein-coding genes between H1 and HFF, we used DEseq2 (Love

et al. 2014) (version 1.24.0) with thresholds for adjusted P-value of <0.01 and for fold-expression change of >2 -fold (Results summarized in Supplemental Table S7). For DE genes with significantly higher expression in one cell line versus the other, we identified which of these genes located entirely within “repositioned” domains that are located closer to nuclear speckles in this cell line versus the other by overlapping gene genomic position with the coordinates of the domains that showed significantly higher scaled TSA-seq scores in this cell line. The remaining DE genes with significantly higher expression in this cell line were determined as differentially expressed but not located within domains that reposition closer to speckles (“non-repositioned”).

Next we calculated the mean scaled TSA-seq scores (max-min normalized, 1-100) across each of the genes and generated scatter plots to show the correlation of these TSA-seq scores between the two cell lines. We calculated \log_2 -fold expression changes for repositioned versus non-repositioned DE genes, and plotted boxplots of these expression changes for all genes in each of these two categories. We used DAVID (Huang da et al. 2009b; Huang da et al. 2009a) (version 6.8) to conduct Gene Ontology (GO) analysis using the “GOTERM_BP_DIRECT” category (Results summarized in Supplemental Table S8). We compared GO terms for the repositioned and non-repositioned DE genes by plotting bar plots of $-\log_{10}$ P-values of top 5 terms sorted by P-values.

CUT&RUN-seq data processing

CUT&RUN-seq data in H1 and HFF were downloaded from the 4DN data portal (<https://data.4dnucleome.org/>); these were generated by the Steven Henikoff laboratory (Seattle, WA) using an automated CUT&RUN platform (Janssens et al. 2018). First, we trimmed the paired end reads with trimmomatic (Bolger et al. 2014). The parameters used were

ILLUMINACLIP:\$adapter/TruSeq3-PE-2.fa:2:15:4:4:true LEADING:20 TRAILING:20

SLIDINGWINDOW:4:15 MINLEN:25. Then we used bowtie2 (Langmead and Salzberg 2012) to align the trimmed reads to reference genome hg38. We set the minimum fragment length to be 10 and maximum fragment length to be 700. Then we sorted and removed duplicate reads with SAMtools (Li et al. 2009). Finally, we used MACS2 (Zhang et al. 2008) to generate fold enrichment (FE) tracks and perform peak calling.

Profiling local TSA-seq signals and histone marks for changed domains

We used the deepTools suite (Ramirez et al. 2016) (version 3.2.1) to profile TSA-seq score percentiles and CUT&RUN-seq fold-enrichment for all the changed regions we identified comparing H1 vs HFFc6. The computeMatrix tool (deepTools) was used to compute signals 2-Mbp upstream and downstream of each region center (10-bp bin). The generated matrix file was used as input for the plotHeatmap tool (deepTools) to plot a heatmap and profile of mean scores for all domains.

PRO-seq data processing

We downloaded PRO-seq data for K562 cells cultured in optimal growth conditions (37 °C, NHS) or treated with heat shock at 42°C for 30 mins (HS30) (GEO: GSE89230), combining raw sequencing reads from two replicates for each condition. We mapped the reads to the human reference genome hg38 (excluded Chromosome Y) using Bowtie2 (Langmead and Salzberg 2012) (version 2.1.0) with default parameters and used the rmdup command from SAMtools (Li et al. 2009) (version 1.5) to remove potential PCR duplicates. We then used bamCoverage from the deepTools suite (Ramirez et al. 2016) (version 3.2.1) to normalize the alignment files using BPM as normalization method binning at 50 bp with strand separation: $\text{BPM (per bin)} = \text{number of reads per bin} / \text{sum of all reads per bin (in millions)}$.

3D immuno-FISH

FISH probes were made from human BACs (ordered from Invitrogen, RP11-479113 for *HSPA1A* locus and RP11-173P16 for *HSPH1* locus), as previously described (Chen et al. 2018). K562 cells were plated on poly-L-lysine (Sigma-Aldrich P4707, 70,000-150,000 M.W., 0.01% w/v) coated coverslips (Fisher Sci 12-545-81) with 0.5 – 1 mL at 0.5 – 0.7 million/mL cell density. Cells were incubated in the 37 °C incubator for 10 mins for attachment. IMR-90, WI38, Tig3 and HCT116 cells were plated on coverslips (Fisher Sci 12-545-81) and cultured for 24 – 48 hrs to reach a confluence of 60% - 80%. Cells were either moved to water bath for heat shock treatment at 42 °C or kept in the 37 °C incubator as control. 3D immuno-FISH was done as previously described (Chen et al. 2018) with modifications for *HSPA1A* FISH in Supplemental Fig S12A-B and *HSPH1* FISH: 1) cells were permeabilized with 0.1% Triton X-100 in high magnesium buffer (5mM MgCl₂, 50 mM PIPES (pH 7.0)) instead of PBS before fixation; 2) rabbit anti-SON polyclonal antibody (Pacific Immunology Corp, custom-raised, 1:1000 diluted) was used as the primary antibody.

Combined (sequential) RNA- and DNA- FISH

Since we could not preserve the RNA-FISH signal during the high temperature denaturation in the DNA-FISH procedure, we applied a sequential FISH procedure for the *HSPH1* locus. We first performed single molecule RNA (smRNA)- FISH, recorded images, and then performed DNA-FISH, identified the same cells, and recorded images again.

To enable image acquisition of the same cells after the RNA-FISH and then again after the DNA-FISH, we recorded cell coordinates relative to etched grids either by using gridded glasses from Ibidi μ -Dishes (35 mm, high Glass Bottom, Grid-50; Ibidi 81158) OR relative to etched reference points on coverslips. For the latter approach, we etched three reference points

on a regular glass coverslip (Fisher Sci 12-545-81) with a diamond pencil. We used the point-marking function in the DeltaVision OMX microscope to record coordinates of these reference points and then the fields of view containing the RNA-FISH images of cells. We then stripped off coverslips from slides, conducted DNA-FISH, and re-mounted the coverslips. In the second-round of microscopy for DNA-FISH, we first measured locations of the original three reference points. We then established a transformation matrix, corresponding to a 2D rotation and translation, to transform the coordinates of our cell point list from the first microscopy round into the coordinate system of the reoriented coverslip in the second microscopy round. We then used these new coordinates to reposition the microscope stage to acquire images for the same cells imaged in the first round of microscopy. For the detailed protocol, R code, and example coordinate lists, see https://github.com/lgchang27/Find_cells).

The glasses/coverslips were autoclaved and coated with poly-L-lysine (Sigma-Aldrich P4707, 70,000-150,000 M.W., 0.01% w/v). K562 cells were plated on the coated glasses with 1 mL at 0.4 – 0.7 million/mL cell density and incubated in 37 °C incubator for 10 mins for attachment. Cells were then either moved to water bath for heat shock treatment at 42 °C for 1 hr or kept in the 37 °C incubator as control.

smRNA-FISH probes targeting the 5' UTR (398 bp) and the first intron (2612 bp) of *HSPH1* gene (Supplemental Table S4) were designed, obtained and labeled by Cy5 fluorophore as previously described (Kim et al. 2019). The smRNA-FISH procedure was modified from that previously described (Kim et al. 2019). Cells were fixed with 4% freshly-made paraformaldehyde (PFA) in PBS at room temperature (RT) for 12 mins. Cells were washed with PBS for 3 x 5 mins and then permeabilized with 0.5% Triton X-100 in PBS (0.5% PBST) containing 2 mM ribonucleoside vanadyl complex (NEB S1402) at RT for 10 mins. Cells were rinsed with PBS for 3

times and then equilibrated with freshly made wash buffer (10% formamide [Sigma-Aldrich F9037] and 2×SSC) at RT for 30 mins. Cells were then incubated with smRNA-FISH probes (final concentration: ~300 – 500 nM) diluted in hybridization buffer (2× SSC, 10% formamide, 10% wt/vol dextran sulfate [Sigma-Aldrich D8906], 1 mg/ml Escherichia coli tRNA [Sigma-Aldrich R8759], 2 mM ribonucleoside vanadyl complex [NEB S1402], and 0.2mg/mL RNase-free BSA [Ambion AM2618, 50 mg/mL] in water) at 37°C for 15 – 17 hrs. Cells were then washed with wash buffer at 37°C for 2 x 30 min. Cells were then post-fixed with 4% freshly-made paraformaldehyde (PFA) in PBS at room temperature (RT) for 12 mins and washed with PBS for 3 x 5 mins. Cells were blocked with 0.5 mg/mL RNase-free BSA in 0.5% PBST containing 2 mM ribonucleoside vanadyl complex at RT for 30 mins. Cells were incubated with mouse anti-Lamin B1/B2 monoclonal antibody [clone 2D8] (Chen et al. 2018) 1:1000 diluted in PBS containing 0.5 mg/mL RNase-free BSA and 2 mM ribonucleoside vanadyl complex at RT for 2hrs. Cells were washed with PBS for 3 x 5 mins and then incubated with goat anti-mouse – AMCA (Jackson Immuno 115-155-003) 1:50 diluted in PBS containing 0.5 mg/mL RNase-free BSA and 2 mM ribonucleoside vanadyl complex at 4 °C for 8 – 10 hrs. Cells were washed with PBS for 3 x 5 mins. Glasses/coverlips were mounted on slides with a non-solidifying glycerol-based medium (90% glycerol, 0.1×SSC, 0.1% p-phenylenediamine [pH 9]) and sealed with nail polish. All water, PBS and SSC buffer used in the smRNA-FISH and immunostaining procedures described above were pre-treated with diethyl pyrocarbonate (DEPC) to prevent RNase contamination. Images for smRNA-FISH were acquired using a DeltaVision OMX microscope system (GE Healthcare).

After acquiring images, the glasses/coverlips were scratched off carefully from slides and the mounting medium was washed off in 4×SSC/0.1% Triton X-100 at 42 °C for 3 x 5 mins. Cells were permeabilized with 0.5% PBST at RT for 10 mins, blocked with 5% normal goat serum (Sigma-Aldrich G9023) in 0.1% PBST (GS blocking buffer) at RT for 1 hr, and then incubated with

rabbit anti-SON polyclonal antibody (Chen et al. 2018) (Pacific Immunology Corp, custom-raised) 1:1000 and mouse anti-Lamin B1/B2 monoclonal antibody [clone 2D8] (Chen et al. 2018) 1:1000 diluted in GS blocking buffer at 4 °C for 10 hrs. Cells were then washed with 0.1% PBST at RT for 3 x 5 mins and incubated with goat anti-rabbit – Texa Red (Jackson Immuno 111-075-003) 1:250 (or goat anti-rabbit – FITC [Jackson Immuno 111-095-144] 1:200) and goat anti-mouse – AMCA (Jackson Immuno 115-155-003) 1:50 diluted in GS blocking buffer at RT for 5 hrs. Cells were then washed with 0.1% PBST at RT for 3 x 5 mins and continued with 3D DNA-FISH for the *HSPH1* locus as described above. Glasses/coverslips were finally mounted on slides. The same cells were found as described above to acquire images for immunostaining and DNA-FISH.

FISH image analysis

To measure the distance between DNA-FISH dot signal and speckle, we took the shortest 3D distance from the center of each FISH signal to the edge (defined as 50% drop from highest speckle intensity to nuclear background intensity) of its nearest nuclear speckle using FIJI. For the elongated DNA-FISH signal, we consider the signal as a composition of multiple diffraction-limited dots that are connected. We took the shortest distance between these dots (center) and a nearby speckle (edge) as the signal's distance to speckle. To measure the elongated DNA-FISH signal length, we measured the length of the trajectory connecting the dot centers. All DNA-FISH distance and signal length measurements were done with deconvolved images. All OMX images used for distance or signal length measurement were aligned for channel correction by bead calibration using the SoftWoRx software.

To measure the smRNA-FISH intensity, we projected optical sections (sum) that contain the FISH signal in z with the “z projection” function in FIJI. We measured the sum of intensity of the FISH signal area in the z-projected image as the signal intensity. We also measured 5

surrounding areas and took the mean as background intensity. We subtracted the background intensity from the signal intensity to yield a normalized RNA-FISH signal intensity, then divided this normalized RNA-FISH signal intensity by the mean of the normalized signal intensity of the “speckle-associated with decondensation (asso dec)” group. All RNA-FISH images from the same experiment were using the same exposure time and % transmitted exciting light. All RNA-FISH intensity measurements were done with non-deconvolved raw images.

Plotting and statistical analysis for distance, length, and RNA signal intensity were done with R. All figure panel images were prepared using FIJI and Illustrator CC (Adobe). RNA-FISH images were shown with projection of 20 slices (sum) that contain the FISH signal in the middle of z (non-deconvolved). All other images were shown with one slice from 3D image stacks (deconvolved).

References

- Anders S, Pyl PT, Huber W. 2014. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166-169.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.
- Chen Y, Zhang Y, Wang Y, Zhang L, Brinkman EK, Adam SA, Goldman R, van Steensel B, Ma J, Belmont AS. 2018. Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *The Journal of cell biology* doi:10.1083/jcb.201807108.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.
- Eisenberg E, Levanon EY. 2013. Human housekeeping genes, revisited. *Trends in genetics : TIG* **29**: 569-574.
- http://wiki.xenbase.org/xenwiki/index.php/Flourescin_Tyramide_Synthesis.
- <https://github.com/ENCODE-DCC/rna-seq-pipeline>.
- Huang da W, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **37**: 1-13.
- Huang da W, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**: 44-57.
- Janssens DH, Wu SJ, Sarthy JF, Meers MP, Myers CH, Olson JM, Ahmad K, Henikoff S. 2018. Automated in situ chromatin profiling efficiently resolves cell types and gene regulatory programs. *Epigenetics Chromatin* **11**: 74.
- Kim J, Venkata NC, Hernandez Gonzalez GA, Khanna N, Belmont AS. 2019. Gene expression amplification by nuclear speckle association. *The Journal of cell biology* doi:10.1083/jcb.201904046: jcb.201904046.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357-359.
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**: 550.
- Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, Graf S, Flicek P, Kerkhoven RM, van Lohuizen M et al. 2010. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Molecular cell* **38**: 603-613.
- Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* **44**: W160-165.
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based Analysis of CHIP-Seq (MACS). *Genome biology* **9**: R137.