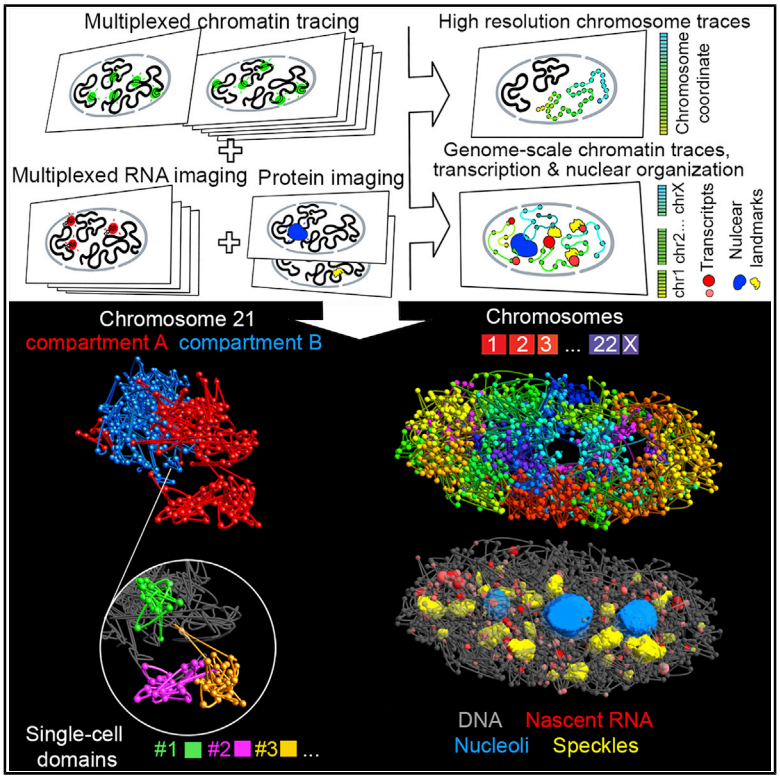
Cell

Genome-Scale Imaging of the 3D Organization and Transcriptional Activity of Chromatin

Graphical Abstract



Highlights

- Massively multiplexed FISH enables mapping chromatin structure at genome scale
- Multimodal high-throughput imaging places chromatin structure in functional context
- *Trans*-chromosome or long-range interactions occur preferentially among active chromatin
- Transcription activity correlates with local enrichment of compartment A chromatin

Authors

Jun-Han Su, Pu Zheng, Seon S. Kinrot, Bogdan Bintu, Xiaowei Zhuang

Correspondence

bbintu@g.harvard.edu (B.B.), zhuang@chemistry.harvard.edu (X.Z.)

In Brief

Su et al. present a high-throughput, multimodal imaging platform to assay the 3D organization of chromatin at genome scale in its functional context in thousands of cells. This platform enables integrated measurements of more than 1,000 genomic loci together with the transcription activity of more than 1,000 genes in the same cells with landmark nuclear structures, making it possible to directly link gene activity to chromosome folding and to answer interesting biological questions.

Su et al., 2020, Cell *182*, 1641–1659
September 17, 2020 © 2020 Elsevier Inc.
https://doi.org/10.1016/j.cell.2020.07.032

CellPress

10x Genomics EX2009
NanoString v. Harvard
IPR2024-00526

Cell

CellPress

Resource

Genome-Scale Imaging of the 3D Organization and Transcriptional Activity of Chromatin

Jun-Han Su,[1,2,4] Pu Zheng,[1,2,4] Seon S. Kinrot,[1,3,4] Bogdan Bintu,[1,4,*] and Xiaowei Zhuang[1,5,*]

[1]Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology, and Department of Physics, Harvard University, Cambridge, MA 02138, USA

[2]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA

[3]Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA

[4]These authors contributed equally

[5]Lead Contact

*Correspondence: bbintu@g.harvard.edu (B.B.), zhuang@chemistry.harvard.edu (X.Z.)

https://doi.org/10.1016/j.cell.2020.07.032

SUMMARY

The 3D organization of chromatin regulates many genome functions. Our understanding of 3D genome organization requires tools to directly visualize chromatin conformation in its native context. Here we report an imaging technology for visualizing chromatin organization across multiple scales in single cells with high genomic throughput. First we demonstrate multiplexed imaging of hundreds of genomic loci by sequential hybridization, which allows high-resolution conformation tracing of whole chromosomes. Next we report a multiplexed error-robust fluorescence *in situ* hybridization (MERFISH)-based method for genome-scale chromatin tracing and demonstrate simultaneous imaging of more than 1,000 genomic loci and nascent transcripts of more than 1,000 genes together with landmark nuclear structures. Using this technology, we characterize chromatin domains, compartments, and *trans*-chromosomal interactions and their relationship to transcription in single cells. We envision broad application of this high-throughput, multi-scale, and multimodal imaging technology, which provides an integrated view of chromatin organization in its native structural and functional context.

INTRODUCTION

The 3D organization of the genome regulates many essential cellular functions ranging from gene expression to DNA replication (Cavalli and Heard, 2019; Dekker and Mirny, 2016; Yu and Ren, 2017). Biochemical and imaging measurements have unveiled complex chromatin structures across a wide range of scales (Bickmore, 2013; Cavalli and Heard, 2019; Dekker and Mirny, 2016; Hnisz et al., 2017; Krijger and de Laat, 2016; Lakadamyali and Cosma, 2020; Levine et al., 2014; Yu and Ren, 2017). In particular, high-throughput chromosome conformation capture methods, such as Hi-C (Lieberman-Aiden et al., 2009), genome architecture mapping (GAM) (Beagrie et al., 2017), and other sequencing-based methods (Dekker et al., 2017; Yu and Ren, 2017), have revealed chromatin structures, such as domains and compartments, with a genome-wide view (Cavalli and Heard, 2019; Dekker and Mirny, 2016; Krijger and de Laat, 2016; Yu and Ren, 2017). Notably, chromatin is partitioned into genomic regions with enhanced self-interaction, called topologically associated domains (TADs) (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012), that often harbor co-regulated genes and have boundaries coinciding with regulatory epigenetic elements (Dekker and Mirny, 2016; Krijger and de Laat, 2016; Yu and Ren, 2017). At a larger scale, chromatin is partitioned into A and B compartments that are enriched for active and inactive chromatin, respectively (Lieberman-Aiden et al., 2009). The physiological significance of A/B compartmentalization is implicated by its changes during development and between cell types (Yu and Ren, 2017).

Although these high-throughput sequencing-based approaches are powerful and have greatly enriched our knowledge of 3D genome organization, they provide contact information for pairs of chromatin loci but not direct spatial position information of individual loci. Furthermore, most of the genome-wide insights into chromatin organization are built on population-averaged contact maps across millions of cells. Investigation of 3D genome organization in single cells remains a challenging task despite recent improvements in single-cell Hi-C methods (Flyamer et al., 2017; Nagano et al., 2017; Ramani et al., 2017; Stevens et al., 2017; Tan et al., 2018). In addition, although methods have emerged to combine Hi-C with other measurement modalities—for example, to provide characterization of chromatin contacts in the context of interacting proteins (Fang et al., 2016; Fullwood et al., 2009; Mumbach et al., 2016; Zheng et al., 2019), nuclear structures (Chen et al., 2018; Girelli et al., 2020; Quinodoz et al., 2018), or DNA modifications (Lee et al.,

2019; Li et al., 2019)—multi-modal measurements by sequencing remain challenging. Notably, a method that allows genome-scale measurements of chromatin organization and transcriptional activity in the same cells has not emerged, despite the demand for such a method to further our understanding of how chromatin organization regulates transcription and vice versa.

Imaging-based approaches, on the other hand, provide a direct measure of the spatial positions of chromatin loci in individual cells with high detection efficiency. In particular, fluorescence *in situ* hybridization (FISH) provides highly specific detection of chromatin loci in fixed cells (Levsky and Singer, 2003), and the clustered regularly interspaced short palindromic repeats (CRISPR) system substantially enhances our ability to image specific chromatin loci in live cells (Chen et al., 2016a; Knight et al., 2018; Xu and Qi, 2019). However, current imaging methods have limited throughput in sequence space, traditionally allowing the study of only a few genomic loci at a time. We recently developed a chromatin tracing approach by sequential rounds of FISH imaging that enabled imaging of tens of distinct chromatin loci in single cells (Wang et al., 2016). This approach provides insights into chromatin structures and their relationship with transcription (Bintu et al., 2018; Cardozo Gizzi et al., 2019; Liu et al., 2020; Mateo et al., 2019; Nir et al., 2018; Sawh et al., 2020; Wang et al., 2016). However, because the number of genomic loci that can be imaged in individual cells remains limited, a high-resolution view of whole chromosomes in single cells is still missing, not to mention a genome-scale view of chromatin organization in individual cells.

Here we report a multi-scale multiplexed FISH imaging technology that enables simultaneous imaging of hundreds to more than 1,000 distinct genomic loci at various resolutions and genomic coverages in single cells. Using this approach, we further demonstrate simultaneous imaging of more than 1,000 genomic loci with nascent RNA transcripts of more than 1,000 genes residing in these loci and landmark nuclear structures, including nuclear speckles and nucleoli. This approach allows us to place chromatin organization in its native structural and functional context and, thus, to explore the relationship between chromatin organization, transcriptional activity, and nuclear structures in single cells.

RESULTS

To enable a systematic view of chromatin structures across multiple scales, we developed an imaging platform (STAR Methods; Table S1) for direct visualization of chromatin with exceptionally high throughput in sequence space up to the genome scale. This platform includes two complementary approaches (Figure 1A). First, for imaging of chromatin structures that are relatively small so that different loci therein are difficult to resolve in any single image, we expanded our previously reported sequential imaging strategy (Wang et al., 2016) to allow tracing of hundreds of chromatin loci in single cells. In this approach, chromatin is imaged one locus at a time (or M loci at a time with M-color imaging) across many imaging rounds (Figure 1A, left). We demonstrated this approach by using it to trace the conformation of whole chromosomes in single cells at high resolution. Second, to image chromatin structures that occupy an area substantially larger than the diffraction-limited resolution, we developed a more efficient, combinatorial imaging strategy based on multiplexed error-robust FISH (MERFISH) (Chen et al., 2015), where many chromatin loci are imaged simultaneously in each round, and their distinct identities are determined based on the combinations of rounds in which they appear (Figure 1A, right). The latter approach allows more genomic loci to be imaged in fewer imaging rounds. We used this approach to provide a genome-scale view of chromatin organization in the context of transcriptional activity and important nuclear structures in single cells.

High-Resolution Chromatin Tracing of Whole Chromosomes

In this section, we describe high-resolution whole-chromosome tracing using the sequential imaging approach (Figure 1A, left; Figure S1A). We first focused on human chromosome 21 (Chr21) and partitioned the non-repetitive portion of the chromosome (Chr21, 10.4–46.7 Mb, from the hg38 assembly) into ~650 segments (i.e., ~650 genomic loci), each 50 kb in length (see STAR Methods for details). We then designed a library of primary oligonucleotide probes (Table S2), each containing a variable target sequence for hybridizing to the different regions of the chromosome and a readout sequence that is unique to each of the 50-kb loci and can be detected by complementary fluorescent readout probes (Figure S1A). However, identifying these loci with ~650 distinct, dye-labeled readout probes would be prohibitively expensive, given the high cost of dye-labeled oligonucleotides. To overcome this challenge, we devised a two-step labeling strategy to detect the distinct readout sequences with a common set of three dye-labeled readout probes (one for each color channel) mediated by unlabeled adaptor probes that convert each locus-specific readout sequence into one of the three common readout sequences (Figure S1A; Table S3). Using this strategy, we sequentially imaged ~650 genomic loci in Chr21 in human lung fibroblast (IMR-90) cells using more than 200 rounds of hybridization with three-color imaging. To enable stable imaging over such a large number of rounds, we optimized our imaging protocol to (1) maintain sample integrity and primary probe binding stability and (2) ensure complete removal of the fluorescence signal after each imaging round and minimize the accumulation of residual signal across hundreds of labeling rounds (see STAR Methods for a detailed experimental protocol).

After imaging, we determined the centroid position of each chromatin locus in 3D and reconstructed the conformation of each copy of Chr21 in each cell (Figure 1B). To estimate how stable the sample and imaging instruments were over many hybridization rounds, we re-imaged a subset of loci after different numbers of rounds and used the displacement between the originally imaged locations and re-imaged locations of the loci to assess measurement accuracy. The median displacement increased from ~70 nm to ~120 nm when the number of hybridization rounds separating the original and re-imaging instances increased from ~10 to ~250 (Figures S1B and S1C). Loci that displayed greater displacement upon re-imaging tended to have a lower fluorescence signal intensity (Figure S1D). Among the loci that had large spatial distances from neighboring loci, a fraction of them also exhibited large re-imaging displacement

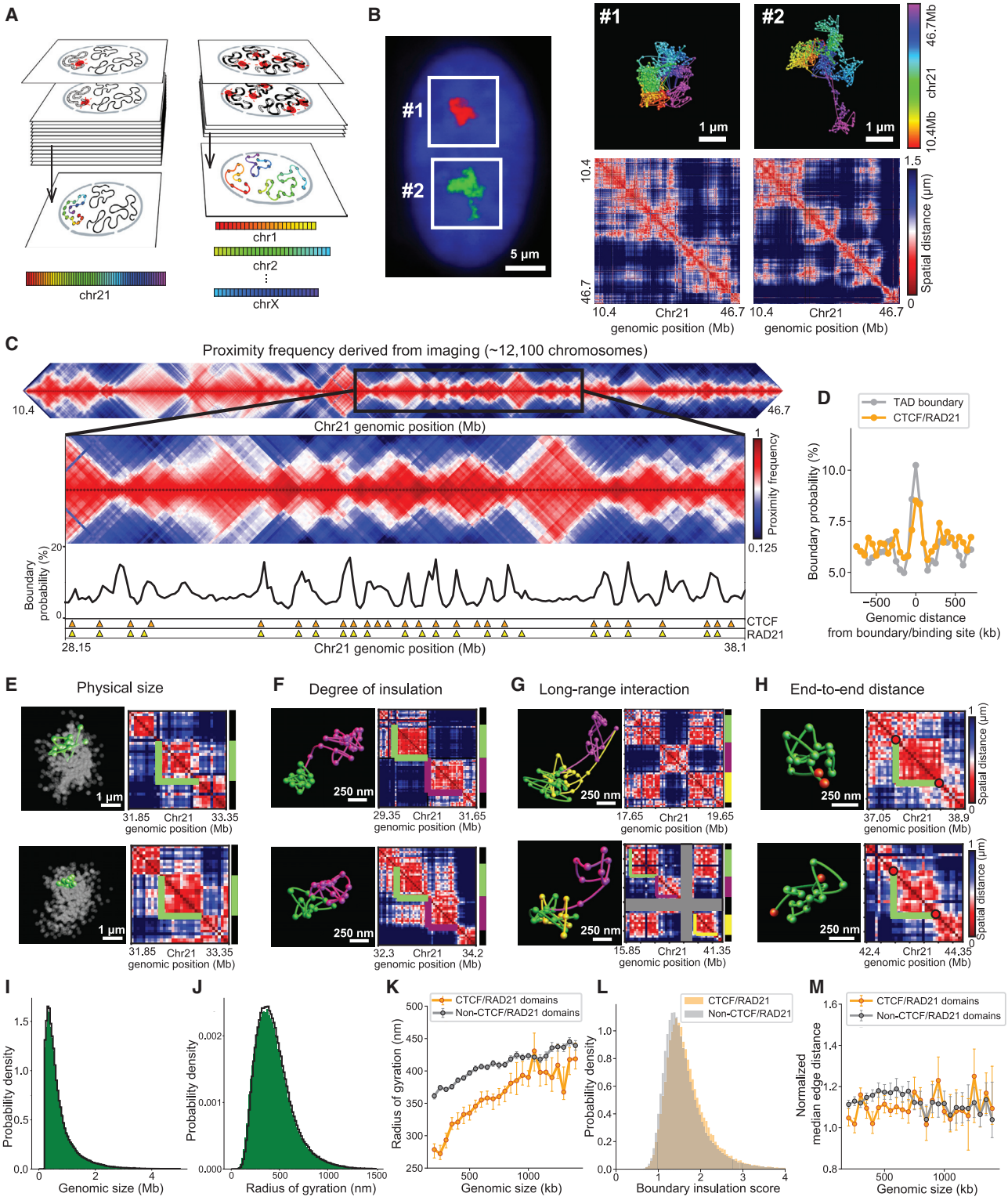


Figure 1. High-Resolution Whole-Chromosome Tracing by Sequential Hybridization and Characterization of Chromatin Domains in Single Cells

(A) Schematics of the multi-scale chromatin tracing platform. Left: schematic of chromosome tracing by sequential hybridization and imaging. A single chromatin locus is imaged in each color channel per imaging round. After many rounds of imaging, a 3D chromatin trace can be generated for each copy of the targeted chromosome. Right: schematic of genome-scale imaging by DNA-MERFISH. Multiple loci are imaged and resolved in each round, and the identity of each locus

(legend continued on next page)

errors and low brightness, suggesting relatively low confidence in the localizations of these loci that are far away from both of their genomic neighbors (see STAR Methods for details). Overall, the median displacement errors for individual loci upon re-imaging were substantially smaller than the median distances between neighboring chromatin loci (~250 nm) (Figure S1B). In addition, median pairwise distances between imaged loci were highly reproducible between biological replicates (Figure S1E). The detection efficiency of the chromatin loci in these experiments was more than 90%.

To obtain a population-averaged view of the conformation of Chr21, we quantified the pairwise interaction between imaged loci by calculating their median spatial distance and their proximity frequency (i.e., the probability that the loci come into proximity) across all imaged chromosome copies (~12,100 copies from the two replicates) (Figure 1C; Figures S1F–S1I). We observed a high correlation between our median pairwise distances from the imaging data and previously published Hi-C data (Figures S1F–S1I). To choose the cutoff distance below which we consider two loci to be in proximity, we calculated the Pearson correlation coefficient between the Hi-C data (Rao et al., 2014) and the proximity frequencies derived from our imaging data across a range of cutoff distances. The Pearson correlation coefficient remained high for a wide range of cutoff distances but peaked at 0.87 when the cutoff distance was ~400–600 nm (Figure S1J). We thus chose 500 nm as the cutoff distance to generate proximity frequency maps throughout this work.

The median distance and the proximity frequency maps showed block-like TAD structures (Figure 1C; Figures S1F and S1H). TAD boundaries identified from the distance and proximity frequency maps from our imaging data were highly similar to those determined from Hi-C data and were not sensitive to the localization error (~100 nm) in our chromatin traces (Figure S1K).

Chromatin Domains in Single Chromosomes

Chromosomes were partitioned into domains that manifested as on-diagonal block-like features in single-cell spatial distance matrices, which also exhibited complex off-diagonal features indicating longer-range interactions (Figure S1L). These domains, long-range interactions, and the inter-locus distances

can be derived from a barcode based on the combination of rounds in which the locus is detected. This approach significantly reduces the number of rounds required to image the same number of loci compared with the sequential imaging approach.

(B) 3D structural rendering and spatial distance matrices of the two copies of Chr21 in a single IMR90 cell imaged using the sequential hybridization approach. Left: the two copies of Chr21 in a single cell (red and green) are overlaid on a DAPI image of the nucleus (blue). Top right: 3D rendering of all detected chromatin loci (colored spheres) in the two Chr21 copies, color-coded according to their genomic coordinates along the chromosome (color index shown on the right). Flexible lines connect adjacent chromatin loci. Bottom right: pairwise spatial distance matrices corresponding to the chromosome copies displayed above (genomic regions without proper reference genomes or containing highly repetitive sequences are not imaged). The genomic positions of the start of the first targeted locus and the end of the last targeted locus are indicated. If a targetted locus was not detected, the matrix elements representing its spatial distances to other loci are shown in black.

(C) Ensemble proximity frequency matrix of Chr21 and preferential positioning of single-cell domain boundaries at CTCF/cohesin-binding sites. Top: proximity frequency matrix for Chr21. Each matrix element is defined as the frequency with which the measured distance between that pair of loci is shorter than a cutoff distance of 500 nm. Center: magnified version of the proximity frequency matrix for a 10-Mb portion of Chr21. Bottom: the probability of single-cell domain boundary formation at each of the imaged 50-kb segments. Orange and yellow triangles show CTCF and RAD21 (a subunit of cohesin complex) chromatin immunoprecipitation sequencing (ChIP-seq) peaks (Dunham et al., 2012), respectively.

(D) Average probability of domain boundary formation in single cells as a function of the genomic distance to CTCF/Rad21-binding sites (sites bound by both proteins, yellow) or to ensemble TAD boundaries (gray).

(E) Example of two single-cell chromatin domains with identical genomic coordinates occupying large (top) or small (bottom) volumes in physical space. Left: 3D rendering of the chromatin domains, with green balls representing imaged genomic loci within the domain. Grey spots represent imaged loci in the rest of the chromosome. Right: pairwise distance matrix for the chromatin domain shown on the left (marked with green lines) with flanking regions.

(F) Example of two pairs of chromatin domains with high (top) and low (bottom) insulation scores. Left: 3D rendering of the chromatin domains. The neighboring domains are shown in green and magenta. Right: pairwise distance matrix for the chromatin domains shown on the left, with rendered domains marked in corresponding colors.

(G) Two examples of long-range contact between chromatin domains. Left: 3D rendering of the chromatin domains, with different colors representing different domains. Right: pairwise distance matrix for the chromatin domains, with rendered domains marked in corresponding colors. The gray space in the bottom right panel indicates a gap in genomic distance of 22.85 Mb.

(H) Example of chromatin domains (green) flanked by CTCF binding sites (red), showing small (top) and large (bottom) distances between the CTCF sites. Left: 3D rendering of the chromatin domain. Right: pairwise distance matrix for the chromatin domain, with domain and border CTCF marked correspondingly.

(I) Distribution of genomic sizes of domains in Chr21 in single cells (green). Shown in the black outline is the distribution of genomic sizes of domains in Chr21 derived from data simulating a localization error of 100 nm. In this simulation, the positions of the imaged loci are perturbed with a 3D Gaussian noise with a standard deviation of 100 nm, similar to our measurement error.

(J) Distribution of physical sizes, as defined by the radii of gyration, of domains in Chr21 in single cells (green). Shown in the black outline is the distribution of physical sizes of domains in Chr21, derived from data simulating a localization error of 100 nm, as in (I).

(K) Median radius of gyration as a function of genomic size for chromatin domains with boundary loci containing interacting CTCF/Rad21 sites (orange) and with neither boundary locus containing CTCF/Rad21 sites (gray). Interacting CTCF/Rad21 sites are defined as those that formed loops in Hi-C maps (Rao et al., 2014). Error bars indicate 95% confidence intervals derived by resampling.

(L) Distribution of insulation scores between neighboring domains with domain boundaries occurring at CTCF/Rad21-binding sites (orange) and non-CTCF/Rad21-binding sites (gray).

(M) Median of normalized distance between the boundary loci of domains as a function of genomic size for chromatin domains with boundary loci containing interacting CTCF/Rad21 sites (orange), as defined in (K), and with neither boundary locus containing CTCF/Rad21 sites (gray). The normalized distance between boundary loci is defined as the distance between the domain's boundary loci divided by the median distance between the same type of locus pairs separated by the same genomic distance but lying in the interior of a single domain. Error bars indicate 95% confidence intervals derived by resampling.

showed high cell-to-cell variability (Figures S1L and S1M), consistent with the substantial variability in chromatin contacts observed in single-cell Hi-C data (Flyamer et al., 2017; Tan et al., 2018). We previously observed similar single-cell domains and cell-to-cell variability when imaging small (~2-Mb) regions of the chromosome (Bintu et al., 2018). However, within these small regions, a sizable fraction of cells did not display clear domain boundaries, and it remained unclear whether domains did not form within those cells or whether the entire imaged regions were within a single domain. Furthermore, the small size of the imaged regions prohibited accurate characterization of certain domain properties, such as their physical and genomic sizes. The high genomic throughput in this study provides a whole-chromosome view of these single-cell domain structures, revealing their presence in most of the imaged chromosome copies, allowing us to characterize their properties in a more systematic manner.

We first identified the genomic locations of the boundaries of these single-cell domains. Although we observed a non-zero probability of boundary formation at nearly all imaged genomic loci, the domain boundaries were preferentially positioned near the binding sites of CTCF and cohesin (Figures 1C and 1D).

In addition to the cell-to-cell variations in domain boundary positions, we also observed substantial heterogeneity in other features of these single-cell domains, ranging from their physical size to the degree of insulation or interaction between domains (Figures 1E–1H). Single-cell domains were variable in their genomic size (Figure 1I) and physical size (Figures 1E and 1J), neither of which was sensitive to our estimated locus localization error (Figures 1I and 1J). Domains bounded by the same genomic regions or having the same genomic size varied considerably in physical size from cell to cell (Figure 1E; Figure S1N). Interestingly, domains bounded by pairs of CTCF/cohesin binding sites that form loops in Hi-C maps (Rao et al., 2014) tended to be smaller in physical size than domains not bounded by such genomic loci (Figure 1K). In addition, the degree of physical segregation between neighboring domains also varied substantially (Figures 1F and 1L), with some neighboring domains completely segregated in space whereas others displayed partial overlap (Figure 1F). Moreover, domains could partially overlap in space with non-neighboring domains, separated by small or large genomic distances (Figure 1G). Last, we observed that the two chromatin loci at the ends of these single-cell domains also exhibited a variable distance from each other and did not exhibit a substantial tendency to be closer to each other compared with chromatin loci separated by a similar genomic distance in the interior of a domain (Figures 1H and 1M).

Chromatin Compartments in Single Chromosomes

Next we took advantage of our high-resolution view of whole chromosomes to examine how chromatin loci in A and B compartments are arranged in single cells. First, we determined the ensemble A/B compartment boundaries using principal-component analysis (PCA) of the Pearson correlation matrix of the proximity frequency map for Chr21 using an algorithm described previously (Figure 2A; Figure S2A; Lieberman-Aiden et al., 2009; Wang et al., 2016). The compartment boundaries obtained from our imaging data were similar to those determined from Hi-C data (Rao et al., 2014; Figure S2A). Below, we use the compartment boundaries obtained from the ensemble proximity frequency map to assign an A/B identity to individual loci in our single-cell analysis.

Our previous imaging work showed that compartment A and compartment B loci tend to spatially segregate in single cells (Wang et al., 2016). The more than 10-fold increase in resolution achieved in this study allowed us to obtain a more detailed view of the organization of A and B loci in single chromosomes. We observed a high degree of variation in the arrangement of A and B loci between individual chromosome copies (Figure 2B). Although, in some chromosome copies, A and B loci were segregated into essentially non-overlapping spatial territories, other chromosome copies exhibited substantial spatial overlap between A and B loci. Intriguingly, compartment A loci in the same chromosome were sometimes separated into multiple "micro-compartments" (Figure 2B).

To quantify the degree of spatial segregation of A and B loci in individual chromosomes, we computed, for each imaged locus, the local density of other A and B loci (Figure S2B). As expected, compartment A loci, on average, tended to be surrounded by A loci, and the same was true for B loci (Figure S2C). We further defined an A/B segregation score for each individual chromosome copy based on the purity of loci observed in the spatial volumes harboring the majority of A or B loci (Figure 2C; STAR Methods). We expected complete physical segregation of A and B loci to result in a segregation score of 1 and thorough mixing of A and B loci to result in a segregation score of ~0.5. We observed that, for most of the imaged Chr21 copies, the segregation scores were considerably higher than the scores of the randomization controls (Figure 2C), demonstrating a tendency for A and B loci to spatially segregate in single cells. It is also worth noting that the spatial segregation of A and B loci was often incomplete (Figure 2C). This potentially reflects incomplete spatial segregation of active and inactive chromatin but could also be caused, in part, by cell-to-cell variability in epigenetic modifications, making the ensemble A/B compartment identity an imperfect proxy for active/inactive chromatin delineation in single cells. To examine the cell cycle dependence, we imaged ~4,600 chromosome copies together with cell cycle markers (STAR Methods). The degree of A/B segregation was cell cycle dependent, with A/B segregation being on average stronger for cells in G2/S phase than in G1 phase (Figure S2D), which is consistent with previous observations of gradual establishment of A/B compartments during the cell cycle (Abramo et al., 2019; Nagano et al., 2017).

Because Chr21 is one of the smallest chromosomes and is partitioned into only a small number of contiguous A and B regions, we next imaged one of the largest chromosomes, Chr2, which exhibits ~50 transitions between A and B compartments along its genomic sequence. Specifically, we traced Chr2 by labeling and imaging 50-kb segments at intervals of 250 kb along its genomic sequence (primary probes are listed in Table S2). We used the same approach described above to separately call A and B compartments in the p and q arms of the chromosome, based on our imaging data (Figure 2D; Figure S2E), and observed quantitative agreement with the A and B

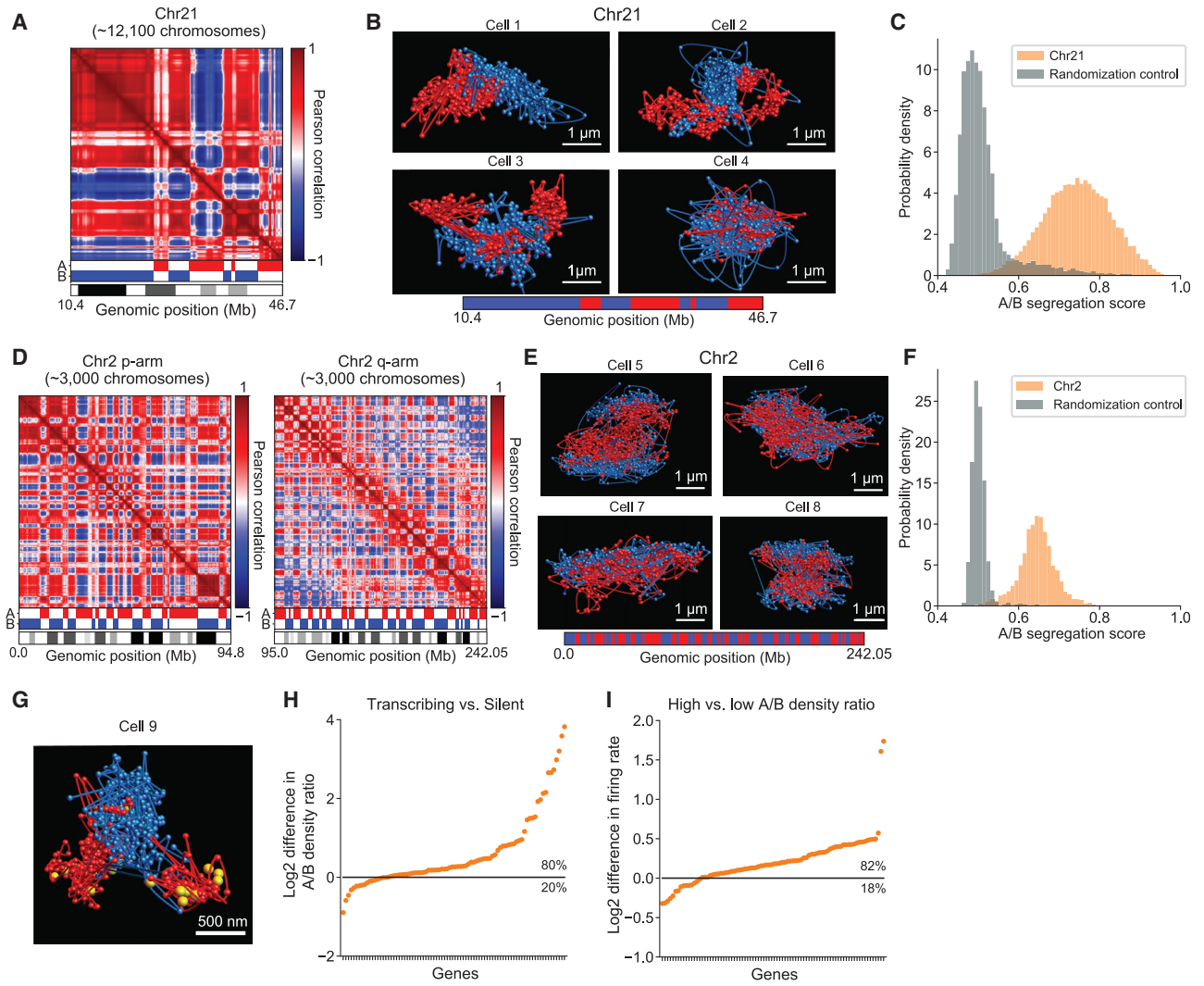


Figure 2. Compartment Structure in Single Chromosomes and Relationship between Transcription Activity and Local Chromatin Content

(A) Pearson correlation matrix for genomic distance-normalized proximity frequencies of Chr21, derived from our imaging data. Two loci are considered in proximity when their distance is smaller than a cutoff distance of 500 nm. Two bars at the bottom shows the A/B calling derived from the proximity frequency matrix (shown in red and blue for A and B compartments, respectively) and the G-banding (National Center for Biotechnology Information, 2020; shown in grayscale) of each genomic locus in the chromosome.

(B) 3D rendering of individual copies of Chr21, with A and B loci shown as red and blue balls, respectively. The bar at the bottom shows the A/B calling of each genomic locus in the chromosome, derived from proximity frequency matrix.

(C) Distribution of the A/B segregation score for individual copies of Chr21, shown in orange. To calculate the A/B segregation score, we defined an A (or B) dense volume for each chromosome copy by thresholding the local A (or B) density so that two-thirds of A (or B) loci with the highest A (or B) density score were contained within the volume (note that the A and B dense volumes can overlap). We further defined the purity of loci in the A (or B) dense volume of the chromosome copy as the fraction of all loci within the volume being A (or B) loci and defined the A/B segregation score of the chromosome copy as the mean purity of the A and B volumes. The gray histogram represents the distribution of A/B segregation scores for a randomization control, where the boundaries between contiguous A and B regions are shifted along the genomic sequence by a random amount for each imaged chromosome copy (in a cyclic fashion after joining the imaged chromosome ends), which keeps the number of A/B boundaries and the sizes of A and B regions largely unchanged, except for the regions at the chromosome ends.

(D) Pearson correlation matrix for genomic distance-normalized proximity frequencies for the p and q arms of Chr2 and corresponding A/B calling and G-banding as in (A).

(E) 3D renderings of individual copies of Chr2 as in (B).

(F) Distribution of A/B segregation scores for Chr2 in single cells as in (C).

(G) 3D rendering of a single copy of Chr21 shown together with transcriptional bursts of the measured genes. Chromatin is rendered in red (A loci) and blue (B loci) as in (B). Yellow balls represent the locations of all detected nascent RNA foci in this chromosome.

(H) Change (measured in log difference) in the A/B density ratios at the transcription start sites (TSSs) of the imaged genes between actively firing and non-firing states. For each gene, we compute the median A/B density ratio at its TSS in chromosomes where the gene is firing and in chromosomes where it is not firing. The

(legend continued on next page)

compartments determined from Hi-C data (Rao et al., 2014; Figures S2F and S2G). At the single-chromosome level, we again observed a variety of spatial arrangements of A and B loci, ranging from strong spatial segregation to substantial spatial overlap between A and B loci (Figure 2E). Some chromosomes showed a "sandwich" configuration in which A loci were placed between two layers of B loci, possibly because of the preferential association of the B loci with nuclear lamina near the top and bottom of the cell nucleus. Quantitatively, the A/B segregation score distribution again showed a tendency for A and B loci to spatially segregate compared with the randomization control (Figure 2F), but the degree of spatial segregation observed for Chr2 appeared to be smaller than that observed for Chr21.

Relationship of Transcription and Local A/B Chromatin Content

To study the relationship between chromatin compartmentalization and transcription, we designed oligonucleotide probes targeting the first intron of ~80 genes harbored in Chr21 (Table S2) and performed sequential hybridization to image the nascent RNA transcripts of these genes followed by chromatin tracing. Furthermore, to more accurately detect the spatial position of these genes, we additionally imaged the 5-kb genomic loci centered around the transcription start site (TSS) of each target gene (Table S2). To prevent false binding of RNA probes to DNA and DNA probes to RNA, we performed RNA probe hybridization without heat-denaturing the double-stranded genomic DNA, imaged the nascent RNA, digested the RNA molecules using RNase, and photobleached any residual signal from RNA probes before performing chromatin tracing (see STAR Methods for a detailed protocol). We confirmed that crosstalk between the RNA and DNA signals was negligible using this strategy (Figure S3).

Typically, a subset of the imaged genes showed transcriptional activity in any individual cell (Figure 2G). We examined how transcription was correlated with the local chromatin environment. To characterize the local A/B chromatin content, we calculated, for each gene, the local densities of A and B loci (excluding the self-locus harboring the gene) and used their ratio (referred to hereafter as the A/B density ratio) as a metric for local enrichment of active chromatin. We found that, for 80% of the genes studied, the median A/B density ratio at its TSS was higher when the gene was actively transcribing than when it was not firing (Figure 2H). As a corollary, the firing rate of a gene also tended to be higher in cells where the TSS of the gene had a higher A/B density ratio (Figure 2I). This increase in transcriptional activity with A/B density ratio could potentially be due to local enrichment of transcription machinery and/or de-enrichment of silencing factors. Alternatively, it is also possible that actively transcribing chromatin, with transcriptional machinery associated, has a stronger tendency to interact with other active chromatin, given that transcription machinery and co-factors can form condensates (Cho et al., 2018; Chong et al., 2018; Gibson et al., 2019; Lu et al., 2018; Sabari et al., 2018). These two possible mechanisms are not mutually exclusive and may work cooperatively to reinforce each other.

Relationship between Domains and Compartments in Single Chromosomes

Next we investigated the interactions between single-cell chromatin domains and how these interactions correlate with compartment identities. Because of the large size of Chr2 and the large number of compartment partitions therein, we expected analyses on Chr2 to provide more insights and, thus, focused on this chromosome.

Although the majority of the single-cell domains in Ch2 were "pure" domains comprised of either A or B loci, we noticed that a substantial fraction of single-cell domains crossed ensemble A/B boundaries and contained both A and B loci (Figures 3A–3C). The presence of these "mixed" domains suggests that domain formation in single cells may not be strongly coupled to the chromatin properties determining compartment identity, but it is also possible that cell-to-cell variability in epigenetic modifications made ensemble A/B boundaries an imperfect proxy for active/inactive chromatin boundaries in single cells.

We next examined how inter-domain interactions depend on the A and B composition of the domains as well as the genomic distances between the domains. Domains came into contact at short and long genomic separations, and such contact patterns varied from cell to cell (Figure 3D). Domain-domain interactions appeared to be modulated by the A/B composition of their underlying chromatin (Figure 3E): the contact frequencies between domains containing mainly B loci were, on average, higher than those between domains containing mainly A loci, which, in turn, were higher than the contact frequencies between domains dominated by chromatin of different A/B identities. This average picture is consistent with the hierarchy of A and B chromatin interaction strengths suggested based on Hi-C and imaging data (Falk et al., 2019).

Further examination of domain contact frequencies as a function of genomic distance revealed a more complex picture. For simplicity, we focused on pure A and pure B domains that contained loci of a single compartment identity. We observed that the contact frequencies between pairs of B domains (B-B) were higher than those between pairs of A domains (A-A) at shorter genomic distances (up to ~75 Mb), whereas A-A domain contacts formed with a greater tendency than B-B domain contacts at larger genomic separation (Figure 3F), consistent with the genomic distance dependence between A-A and B-B chromatin interactions reported in a recent Hi-C study (Abramo et al., 2019). Notably, at relatively large genomic distances, B-B domain contact probability decayed to a level similar to the contact probability between A and B domains (A-B), whereas A-A domain contact probability remained higher than A-B domain

log difference of these median values for the ~80 genes imaged on Chr21 is rank ordered according to the magnitude of change in their median A/B density ratios. Log difference stands for the difference in log values.

(I) Change (measured in log difference) in the firing rates of imaged genes as the environment of the gene's TSS is changed from low (bottom quartile) to high (top quartile) A/B density ratios. The log difference in firing rate for the ~80 genes imaged on Chr21 is rank ordered according to the magnitude of firing rate change.

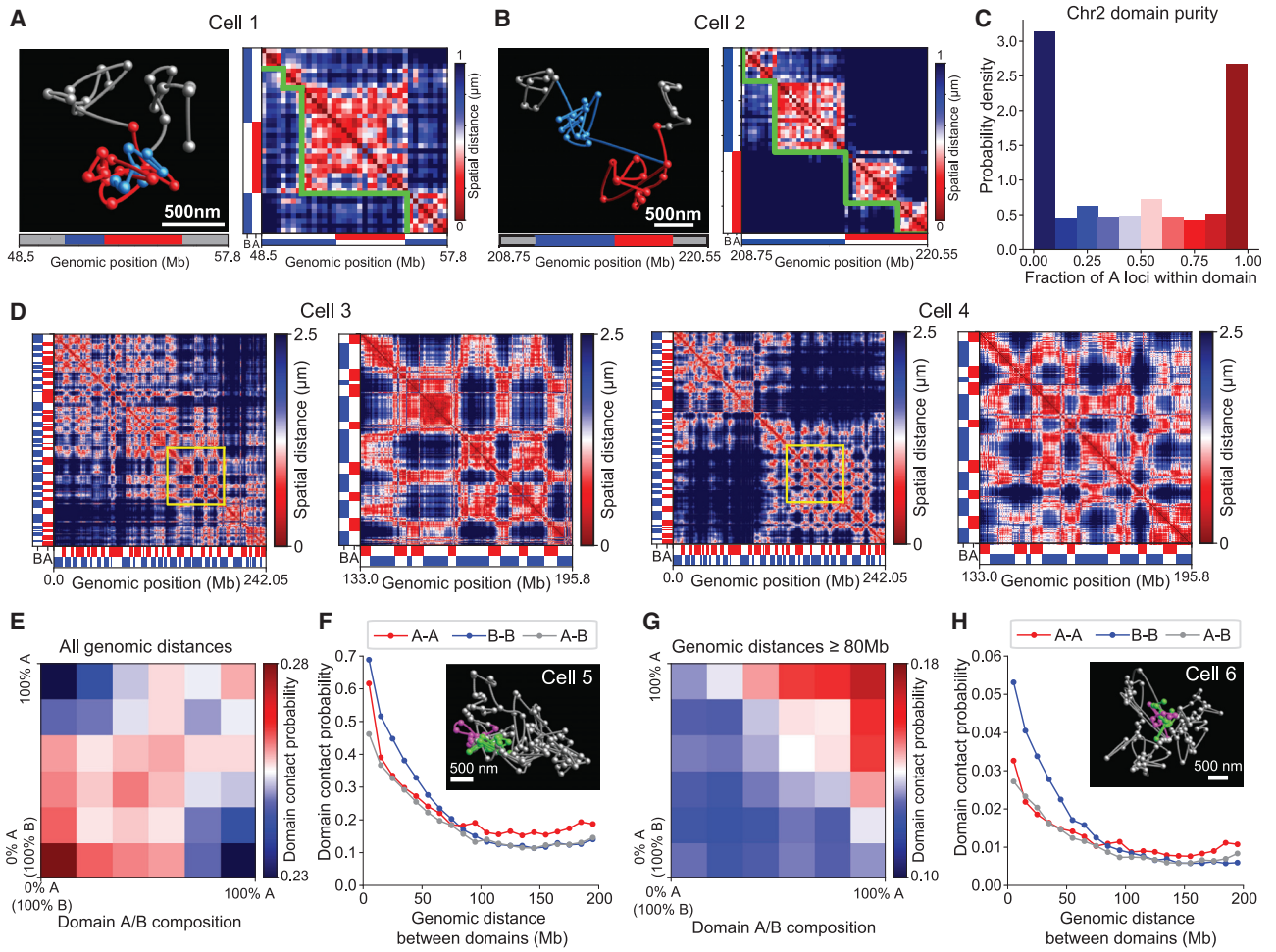


Figure 3. Dependence of Domain-Domain Interaction on A/B Composition and Genomic Distances

(A) Left: 3D rendering of a "mixed" chromatin domain containing A and B loci (displayed as red and blue balls, respectively), flanked by pure domains comprised of only B loci (displayed as gray balls) in a single copy of Chr2. Right: pairwise distance matrix for the region displayed on the left. The red and blue bars on the bottom and left display the A (red) and B (blue) calling of loci, and the green outline highlights the boundaries of the chromatin domains.

(B) Same as in (A) but for two pure domains, one comprised entirely of A loci and one of B loci, instead of a mixed domain.

(C) Distribution of the fraction of loci being A loci in single-cell chromatin domains in Chr2.

(D) Single-cell spatial distance matrices of two example copies of Chr2. The first and third panels show the matrix for two whole chromosomes, whereas the second and last panels show a magnified matrix for the region highlighted in yellow in the first and third panels, respectively. The side bars show the A (red)/B (blue) compartment calling.

(E) Domain contact probabilities for domains of different A/B compositions in Chr2. The x and y axes represent the fraction of loci within a domain being A loci (0% corresponds to pure B domain and 100% corresponds to pure A domain). Two domains are defined to be in contact when their insulation score is less than 2. See STAR Methods for calculation of the insulation score.

(F) Domain contact probability between pure A domains (A-A), between pure B domains (B-B), and between pure A and pure B domains (A-B) in Chr2, plotted as a function of the genomic distance between the two domains. The inset shows a 3D rendering of an example pair of domains displaying long-range interaction with an insulation score of ~2. The two interacting domains are displayed in green and magenta, whereas the chromatin in between is displayed in gray.

(G) Same as (E) but for domain pairs with genomic distances larger than 80 Mb.

(H) Same as (F), but restricted to domain pairs with a high degree of intermixing (as defined by a low insulation score of less than 1). The inset shows a 3D rendering of an example pair of domains displaying long-range interaction with a high degree of intermixing (insulation score being ~1).

contact probability at large genomic separations (Figure 3F). This led to clear domination of A-A domain interactions at large genomic distances (Figure 3G). In addition, we observed that contacting domain pairs also showed different degrees of spatial overlap, with some pairs of domains showing relatively superficial contact (Figure 3F, inset), whereas other pairs showed strong intermixing (Figure 3H, inset). Interestingly, B-B domain pairs showed a substantially stronger tendency to form such intermingled globules than A-A domain pairs at relatively short genomic distances (Figure 3H).

Overall, these results suggest that the nature of A-A and B-B domain interactions are distinct, potentially because of different

molecular factors involved in A-A and B-B association. For example, heterochromatin factors such as HP1 are thought to be involved in B-B interactions (Larson et al., 2017; Strom et al., 2017), whereas transcription machinery, activating factors, and co-activators may be involved in active chromatin interactions (Cho et al., 2018; Chong et al., 2018; Gibson et al., 2019; Lu et al., 2018; Sabari et al., 2018). Whether these different molecular factors are responsible for the observed differences in A-A and B-B domain interactions awaits further investigation.

Genome-Scale Chromatin Imaging

The sequential imaging approach described above is well suited for imaging chromatin structures that are comparable with or smaller than the diffraction-limited resolution. However, the number of genomic loci imaged increases only linearly with the number of imaging rounds in this approach. For genome-scale chromatin imaging, because many genomic loci could be simultaneously resolved and localized in the nucleus, we reasoned that much more efficient, non-linear scaling of the number of imaged loci with the number of imaging rounds would be possible.

To achieve this goal, we devised a combinatorial FISH approach, based on the MERFISH method we developed previously for transcriptome imaging (Chen et al., 2015), but with modifications specifically designed for chromatin imaging by considering the polymeric nature of chromatin and the territorial organization of chromosomes (Cremer and Cremer, 2010). We termed this approach DNA-MERFISH. Specifically, we assigned to each genomic locus a unique 100-bit binary barcode with a Hamming weight of 2 (HW2); i.e., each barcode containing two "1" bits and 98 "0" bits (Figure 4A). The bit values (1 or 0) in these barcodes determine the presence or absence of signal for each locus across sequential rounds of imaging. To avoid imaging spatially close chromatin loci in the same bit, we selected a subset of these barcodes to encode the targeted genomic loci and optimized barcode assignment so that loci with a 1 bit in the same barcode position were maximally separated in genomic space (STAR Methods). This strategy allowed us to minimize detection errors caused by overlapping signals from nearby loci. Moreover, because the vast majority of the possible 100-bit HW2 binary codes were invalid (i.e., not assigned to any targeted locus), this design allowed us to identify and discard detection errors and further improve measurement accuracy.

We physically imprinted the barcodes onto the targeted genomic loci using a library of encoding probes (Table S4), each containing a target region for binding to one of the targeted loci and a readout sequence (in multiple copies) chosen from 100 pre-designed readout sequences (Figure 4A). Each readout sequence corresponds to one of the 100 bits, and the encoding probe set for each genomic locus contains two distinct readout sequences corresponding to the two bits that read 1 in the barcode assigned to that locus. After encoding probe binding, we detected the barcodes imprinted on the chromatin loci by sequential hybridization of fluorescently labeled readout probes (or adaptor and readout probes; Table S3), each binding to one of the 100 readout sequences (Figure 4A). This allowed us to image and identify ~1,000 genomic loci with only 50 rounds of hybridization with 2-color imaging (Figures 4A–4C). This represents approximately 10-fold fewer rounds of hybridization compared with sequential imaging of the same number of loci with the same number of color channels. Because each chromosome in diploid cells has two homologs, we further assigned the homolog identities of the imaged loci using a clustering algorithm, exploiting the tendency of chromosomes to occupy distinct territories in each nucleus (Cremer and Cremer, 2010), although this could introduce error when the two homologs overlap in space.

We selected 1,041 genomic loci for imaging, each ~30 kb in size, more or less uniformly covering the 22 autosomes and the X chromosome in IMR-90 cells, with each chromosome containing 30–80 targeted genomic loci. We imaged these 1,041 genomic loci in ~5,400 cells across 5 biological replicates with a detection efficiency of ~80% for each locus, yielding on average ~1,700 chromatin loci detected in each cell (Figures 4D and 4E). At the end of the combinatorial imaging process, we re-imaged a small subset of the genomic loci with sequential imaging, one locus at a time. The displacement between the locus positions determined by combinatorial imaging and the re-imaged positions determined by sequential imaging was only ~60 nm (median value; Figure S4A), indicating a high decoding accuracy of the DNA-MERFISH approach and minimal sample degradation/deformation during the course of imaging.

We calculated, in each cell, the spatial distance between every pair of imaged chromatin loci and then determined the median spatial distance and the proximity frequency between every pair of loci across all imaged cells (Figure 4F; Figure S4B). The proximity frequencies between locus pairs within the same chromosome showed high correlation with the contact frequencies detected by Hi-C (Rao et al., 2014; Figure S4C). Moreover, our imaging results showed high reproducibility between biological replicates (Figure S4D).

By exploring chromatin organization in individual cells, we noticed that chromosomes, although having a tendency to occupy distinct territories within each cell (Figures 4F and 4G), also displayed substantial overlap with each other (Figures 4G and 4H), suggesting substantial *trans*-chromosomal interactions. These results are consistent with and expand previous observations (Bolzer et al., 2005; Branco and Pombo, 2006; Cremer and Cremer, 2010).

Enrichment of *trans*-Chromosomal A-A Interactions

We next studied how *trans*-chromosomal interactions depend on the A/B compartment identity of chromatin. We classified each of our imaged genomic loci into compartments A and B based on Hi-C data (Rao et al., 2014). Although our genome-scale imaging data also allowed A/B compartment calling that was in ~87% agreement with Hi-C calling (Figure S4E), we used Hi-C calling because of the higher genomic resolution of the ensemble Hi-C data. 38% of our imaged loci belonged to compartment A, whereas 62% belonged to compartment B. To examine whether the extent of *trans*-chromosomal interactions differs for A and B loci, we rearranged the genomic loci in the *trans*-chromosomal proximity frequency matrix, placing all A loci next to each other, followed by all B loci. This matrix showed that a compartment A locus had, on average, a stronger tendency to interact *trans*-chromosomally with another A locus

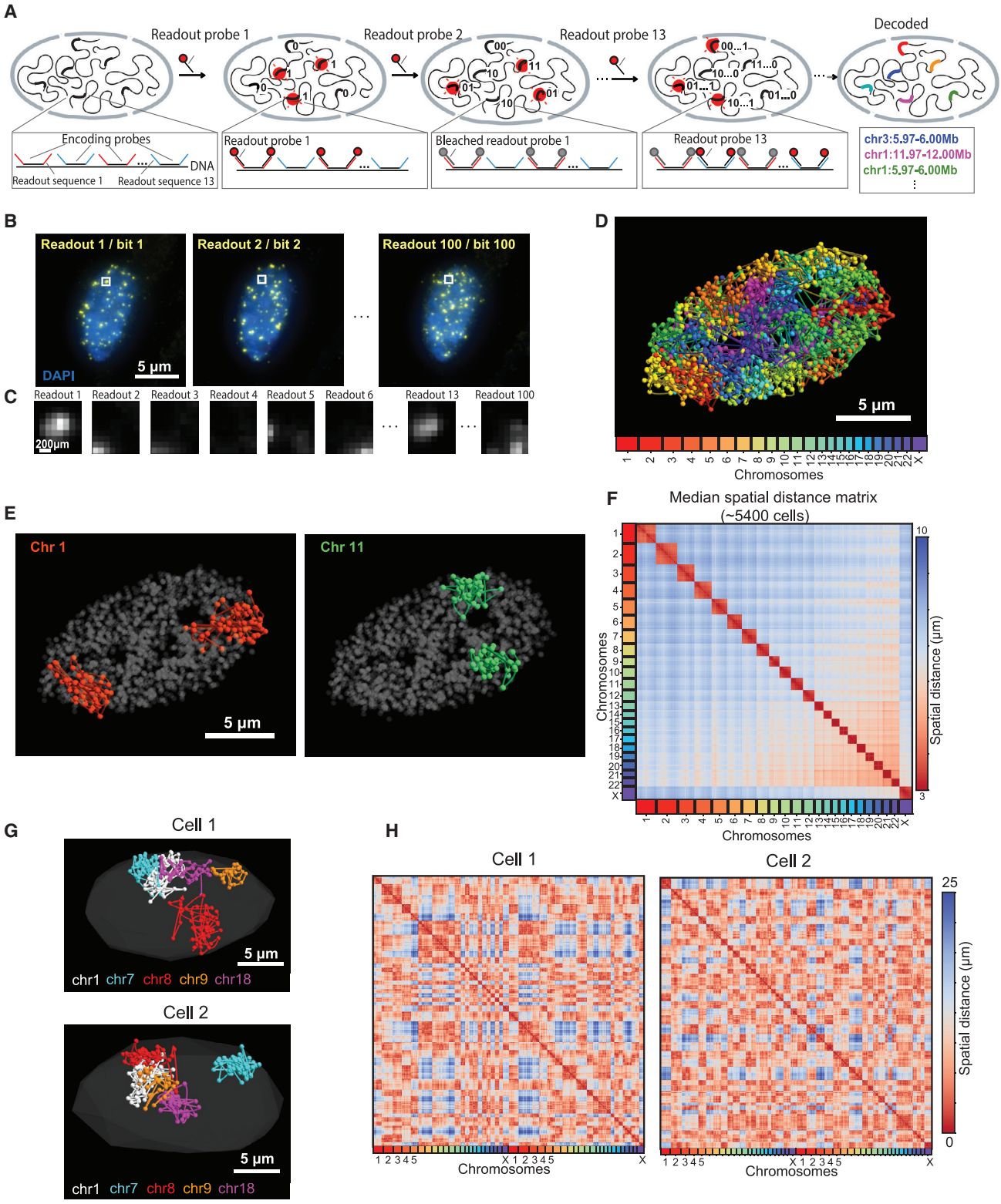


Figure 4. Genome-Scale Chromatin Imaging by DNA-MERFISH

(A) Imaging scheme of DNA-MERFISH. The targeted genomic loci are assigned error-robust barcodes; e.g., a subset of 100-bit HW2 binary barcodes. The barcodes are imprinted onto the genomic loci with encoding oligonucleotide probes that recognize the loci and associate two distinct readout sequences with

(legend continued on next page)

than with individual B loci (Figures 5A and 5B), consistent with previous observations of preferential *trans*-chromosomal interactions between active chromatin (Branco and Pombo, 2006; Rao et al., 2014; Yaffe and Tanay, 2011). In contrast, B loci showed comparable or lower *trans*-chromosomal affinity toward each other than toward A loci (Figures 5A and 5B). In other words, *trans*-chromosomal A-A interactions appeared with a stronger tendency than A-B interactions, which, in turn, appeared with a slightly stronger tendency than B-B interactions. A similar trend was observed for a wide range of cutoff distances used to construct the proximity frequency map (Figures S4F–S4H). Notably, this is in striking contrast with the overall hierarchy for *cis*-chromosomal interactions, where B-B interactions had a stronger tendency to form than A-A interactions, which, in turn, had a stronger tendency to form than A-B interactions (Figure 3E). Intriguingly, however, this observed trend for *trans*-chromosomal interactions (A-A > A-B ≈ B-B) was similar to that observed for *cis*-chromosomal interactions at large genomic distances in the high-resolution Chr2 data (Figures 3F and 3G). This trend was also observed for long-range *cis*-chromosomal interactions in the genome-scale data (Figure 5C), although the preference for B-B interactions at short genomic distances is less pronounced than that observed in high-resolution Chr2 tracing data, possibly because of the lower resolution here.

Next we examined the relationship of *trans*-chromosomal interactions with chromatin compartment identity at the single-cell level. In individual cells, A and B loci adopted different spatial distributions, with A loci exhibiting a tendency to be more centrally localized than B loci in the nucleus (Figure 5D; Figure S5), as expected (Solovei et al., 2016). To further characterize chromatin interactions in *trans,* we adopted a density-based approach similar to that presented earlier, except that here we only considered *trans*-chromosomal interactions. Briefly, for each imaged locus in each chromosome, we calculated the local densities of A loci and B loci from all other chromosomes in the same cell and determined the ratio of these two densities (referred to hereafter as the *trans* A/B density ratio) (Figures 5D and 5E). This quantity provides a measure of local enrichment of *trans*-chromosomal active chromatin near the locus. The majority (62%) of the imaged loci are B loci, creating an overall bias for the A/B ratio to be smaller than 1. To control for this, we compared distributions of the *trans* A/B density ratios observed for A and B loci, with the distribution obtained in a randomization control, where the A and B identities were randomly shuffled among the imaged loci while keeping the numbers of A and B loci unchanged. Notably, the *trans* A/B density ratios observed for A loci were substantially higher than the values observed for B loci, which, in turn, were higher than the values derived from the randomization control (Figure 5E). These analyses further support the notion that *trans*-chromosomal interactions are preferentially enriched for active chromatin.

We further tested whether enrichment of *trans*-chromosomal A-A interactions required transcription. We inhibited RNA polymerase II (Pol II) activity with α-amanitin treatment and found that this treatment did not lead to a substantial reduction in enrichment for *trans*-chromosomal A-A interactions (Figures S6A–S6C), consistent with a previous observation that the interactions of the activated β-globin locus with other active chromatin loci were not inhibited by transcription inhibition (Palstra et al., 2008) and suggesting that factors other than Pol II, such as transcription factors and co-factors, could stabilize these interactions.

Multi-modal Imaging of Chromatin, Nascent Transcripts, and Nuclear Structures

To place the 3D organization of chromatin in the context of its functional activity and other nuclear structures, we further extended the DNA-MERFISH method to enable simultaneous measurements of chromatin organization together with transcriptional activities of the imaged genomic loci as well as nuclear landmarks in single cells. Specifically, we imaged the aforementioned 1,041 genomic loci together with nascent RNAs transcribed from the 1,137 genes located at these loci and simultaneously with important nuclear structures, including nuclear speckles and nucleoli (Figure 6A).

To enable chromatin, RNA, and nuclear body imaging in the same cells, we sequentially stained and imaged each of these components, starting with RNA, followed by DNA and, finally, nuclear bodies (Figure 6A; see STAR Methods for a more detailed description). We imaged the intronic RNAs of the 1,137 genes by adopting a similar MERFISH strategy to the one described above for chromatin (Figure 6A; Table S4). Considering that not all genes would be simultaneously transcribed in individual cells and, hence, the density of transcription foci should not be as high as that of the chromatin loci, we encoded the RNAs with a 54-bit HW2 code, and selected 1,137 of the possible barcodes to encode the genes similar to how

each locus, corresponding to the two bits that read 1 in the barcode assigned to the locus. Each locus is labeled by a total of ~400 encoding probes, but only 4 are shown. Fluorescent readout probes complementary to the readout sequences are sequentially added and imaged, allowing the bits that read 1 at each locus (and, hence, the barcode identity of that locus) to be determined.

(B) Representative images from multiple imaging rounds in the nucleus of a single cell. The fluorescent signal of the chromatin loci from readout probes is shown in yellow, whereas the DAPI signal, used as a nuclear marker, is shown in blue.

(C) Magnified images of a small region (white box in B) centered around one chromatin locus across all imaging rounds, with two bits (1 and 13) on.

(D) A 3D rendering of all detected chromatin loci (colored spheres) in a single IMR-90 cell, color-coded according to the chromosomes to which they belong (color index for chromosomes shown below the image). Adjacent loci in genomic sequence are connected by a flexible line. 1,041 genomic loci are targeted.

(E) Chromatin loci of the same cell as in (D) but with two homologs of the indicated chromosomes shown in color and all other loci shown in gray.

(F) Median distance matrix computed from ~5,400 single cells. For each pair of loci, the median of observed 3D spatial distances between the loci across all cells is presented.

(G) Example images showing the positions of multiple chromosomes territories in single cells. Chromosomes are color-coded as indicated, and shaded areas represent the convex hull surrounding all imaged loci. Only one homolog is shown per chromosome for clarity.

(H) Spatial distance matrices for the cells shown in (G). Chromosome order is as noted beneath the matrices, with the two homologs of each chromosome shown separately.

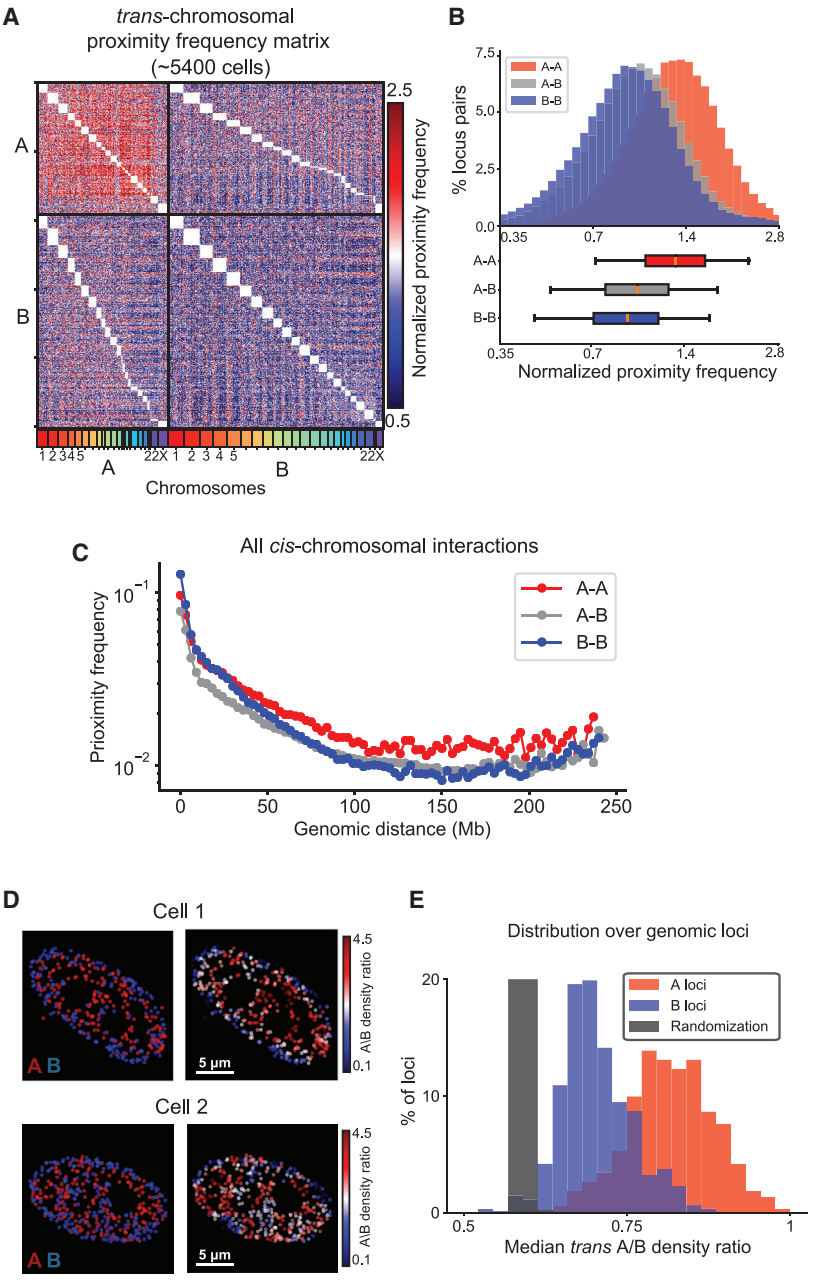


Figure 5. Enrichment of Active-Active Chromatin Interactions in *trans*-Chromosomal Interactions

(A) Normalized *trans*-chromosomal proximity frequency matrix. The normalized proximity frequency between each *trans*-chromosomal locus pair is shown, with pairs of loci considered to be in proximity when their distances are smaller than a cutoff distance of 500 nm. The loci are reordered so that compartment A loci appear first, followed by compartment B loci. Each entry in the matrix shows the proximity frequency between a *trans*-chromosomal locus pair normalized by the median proximity frequency of all locus pairs originating from the same pair of chromosomes to account for varying basal levels of interaction between pairs of chromosomes.

(B) Distributions of normalized *trans*-chromosomal proximity frequencies for pairs of A loci (A-A, red), pairs of B loci (B-B, blue), and pairs comprised of one A and one B locus (A-B, gray), derived from the matrix shown in (A). Distributions are represented in the top panel as histograms and in the bottom panel as boxplots, showing the median (center lines), 25th–75th percentiles (boxes) and 5th–95th percentiles (whiskers).

(C) The proximity frequency between pairs of *cis* chromatin loci within the same chromosomes as a function of their genomic distance, for pairs of loci separated by the same genomic distance across all chromosomes. The proximity frequencies are calculated across all chromosomes for pairs of loci separated by the same genome distance and are shown for pairs of A loci (A-A, red), pairs of B loci (B-B, blue), and mixed pairs of A and B loci (A-B, gray).

(D) Distributions of A and B loci in two single cells. The left panels show the locations of all detected loci within a single z plane in a nucleus, with A loci shown in red and B loci shown in blue. In the right panels, the color of each locus represents the ratio of the local densities of *trans*-chromosomal A and B loci (i.e., the *trans* A/B density ratio), in accordance with the color scale bar shown on the right.

(E) Distributions of the median *trans* A/B density ratio for imaged genomic loci. For each locus, the median *trans* A/B density ratio across all cells was determined, and the *trans* A/B density ratio distribution for A loci is shown in red and for B loci in blue. The dark gray histogram represents a randomization control where A and B compartment identity is shuffled randomly while keeping the total numbers of A loci and B loci unchanged. This histogram is cropped by the y-axis range, which is set to allow a clearer visualization of the histograms for the A and B loci.

we selected the barcodes for chromatin imaging to minimize the chance of imaging spatially proximal genes in the same bit. After RNA imaging was completed, we enzymatically digested the RNA transcripts and performed DNA-MERFISH as described above to image the 1,041 genomic loci (Figure 6A). Decoding of genomic loci and nascent RNA transcripts was performed largely independently, with the additional constraint of the transcripts being located close to their harboring genomic loci (STAR Methods). This procedure further improved detection accuracy for RNA transcripts. Finally, we imaged nuclear speckles and nucleoli using immunofluorescence against known molecular components of these structures (Figure 6A). The fluorescence signals for nuclear speckles and nucleoli displayed a high signal to noise ratio (>25), even with immunofluorescence staining performed after DNA-MERFISH. We additionally estimated the positions of nuclear lamina by computing a convex hull surface encompassing all imaged genomic loci. Together, these multi-modal measurements allowed an integrated single-cell view of 3D genome structure, transcriptional activity, and nuclear organization (Figure 6B). We performed these multi-modal imaging experiments on ~3,700 individual cells in two biological replicates. Chromatin imaging data from these multi-modal experiments were included in the 5 replicates and ~5,400 cells described above for 3D genome organization analyses.

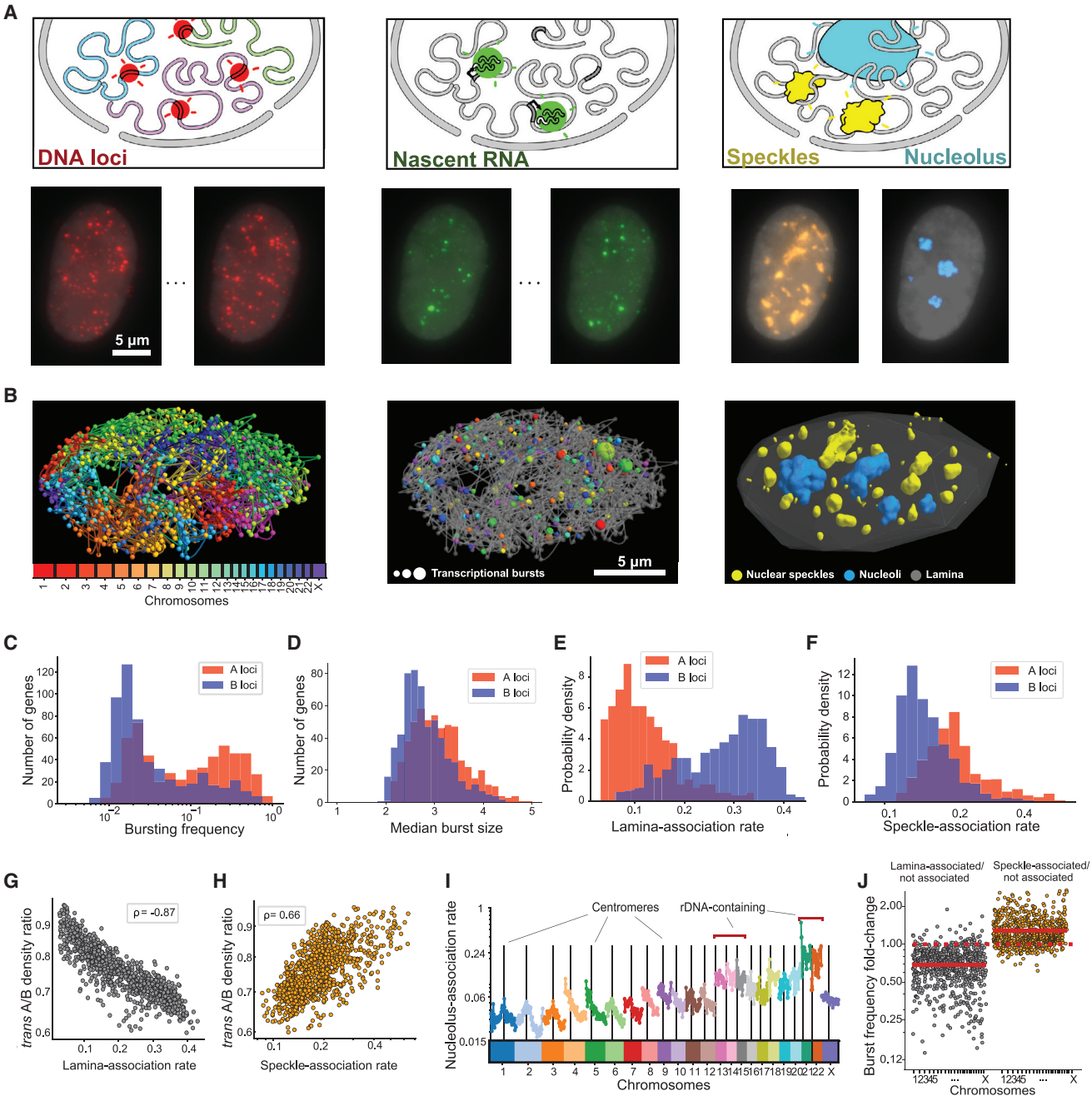


Figure 6. Multi-modal Genome-Scale Imaging of Chromatin and Transcription Activity in the Context of Nuclear Structures

(A) Top: illustration of the multi-modal imaging scheme combining chromatin (left panel), nascent RNA (center panel), and nuclear body (right panel) imaging. Approximately 1,000 genomic loci, nascent RNA transcripts of ~1,100 genes in the targeted loci, and two types of nuclear bodies (nuclear speckles and nucleoli) are imaged. Bottom: representative raw images for each imaging modality: chromatin loci across multiple imaging rounds (left), nascent RNA transcripts across multiple imaging rounds (center), and nuclear bodies (right; yellow, nuclear speckles imaged using an anti-SC35 antibody; blue, nucleoli imaged using an anti-fibrillarin antibody).

(B) 3D rendering of chromatin loci, transcriptional foci, and nuclear bodies in a single cell. Left: all detected chromatin loci, color-coded by chromosome (based on the chromosome index shown below). Center: all detected foci of nascent transcripts shown as colored spheres, with colors indicating the identities of the imaged genes and sphere size representing transcription burst size. Chromatin loci are shown in gray. Right: volume-filling representations of detected nuclear bodies. Nucleoli are shown in blue and nuclear speckles in yellow. The nuclear lamina is identified as the surface of the convex hull surrounding all detected chromatin loci (shaded gray area).

(C and D) Distributions of transcription burst frequencies (C) and burst sizes (D) for genes residing in the imaged A loci (red) and B loci (blue).

(legend continued on next page)

From the nascent RNA transcript measurements, we quantified, for each gene, the transcriptional burst frequency and the median burst size (Figures 6C and 6D). These measures showed high correlation across replicate experiments (Figures S4I and S4J). The burst frequency displayed bimodal behavior, with high burst frequency genes primarily harbored in compartment A, and low burst frequency genes present in compartments A and B (Figure 6C). Furthermore, we estimated whether specific chromatin loci were associated with nuclear bodies using a spatial distance cutoff of 250 nm (the approximate physical size of our imaged genomic loci) and observed a higher association frequency of B and A loci with the nuclear lamina and nuclear speckles, respectively (Figures 6E and 6F), consistent with previous observations (Chen and Belmont, 2019; van Steensel and Belmont, 2017). For individual loci, the median *trans* A/B density ratio had a negative correlation with lamina association frequency (Figure 6G) and a positive correlation with nuclear speckle association frequency (Figure 6H). Nucleoli additionally showed preferential association with centromeres, telomeres of certain chromosomes, and chromosomes containing ribosome-encoding genes (Figure 6I), as shown previously (Iarovaia et al., 2019) .

Notably, for most imaged loci, lamina association correlated with lower transcriptional activity, whereas nuclear speckle association correlated with higher transcription activity (Figure 6J), consistent with a recent study showing negative correlation between gene expression and lamina association in single cells (Rooijers et al., 2019). Moreover, we observed that treatment with the transcription inhibitor α-amanitin perturbed nuclear speckles, reduced speckle association rates, and increased lamina association rates for the imaged loci (Figures S6D–S6F). Together, these results expand upon previous imaging studies of the nuclear repositioning of genomic loci upon transcriptional activation or inhibition (Chen and Belmont, 2019; Shachar and Misteli, 2017; van Steensel and Belmont, 2017) and provide a genome-scale view of the relationship between transcriptional activity and interactions with nuclear structures.

trans-Chromosomal Interactions in Various Nuclear Contexts

Simultaneous imaging of chromatin organization and landmark nuclear structures in the same cells further allowed us to examine how the observed enrichment for *trans*-chromosomal A-A interactions depends on nuclear context. Because nuclear speckles are some of the most prominent nuclear bodies that concentrate actively transcribed loci (Chen and Belmont, 2019), we wondered whether our observed enrichment for *trans*-chromosomal A-A interactions was simply a result of this local concentration effect at the nuclear speckles. To address this question, we restricted analysis to loci that were not associated with nuclear speckles; i.e., for each locus pair, we only considered cells in which neither locus was associated with a nuclear speckle. Interestingly, we still observed the same trend for enrichment of *trans*-chromosomal A-A interactions over A-B and B-B interactions under this constraint (Figure S7A and S7B), indicating that association with nuclear speckles cannot fully account for the observed enrichment.

Next we considered another relatively simple scenario because of concentration effects. Because compartment A loci were de-enriched from the lamina and more concentrated in the interior regions of the nucleus (Figure S5), we wondered whether the enrichment of *trans*-chromosomal A-A interactions was simply caused by local enrichment of A loci in the nuclear interior. To test this, we considered, for each locus pair, only cells in which both loci were associated with the nuclear lamina. Notably, enrichment for *trans*-chromosomal A-A interactions over A-B and B-B interactions was observed even for these lamina-associated loci (Figures S7C and S7D) despite the fact that the lamina is an environment enriched for inactive chromatin.

Overall, these results suggest a non-trivial molecular mechanism for the observed *trans*-chromosomal interactions between A loci. As noted earlier, enrichment for A-A over A-B and B-B interactions was also observed for *cis*-chromosomal interactions at large genomic distances (Figures 3F, 3G, and 5C). It is possible that long-range *cis* A-A interactions have an underlying mechanism in common with that of the *trans*-chromosomal A-A interactions. Which molecular factors give rise to these active chromatin interactions remains an open question.

Correlation between *Trans*-chromosomal Interactions and Transcriptional Activity

Next we used these multi-modal measurements to characterize the relationship between transcriptional activity of individual chromatin loci and their local chromatin environment defined by *trans*-chromosomal contributions. To this end, we calculated the *trans* A/B density ratios and determined the median values of this quantity for two populations of cells (determined independently for each genomic locus): (1) cells where the locus under consideration exhibited transcriptional activity and (2) cells where the locus appeared to be transcriptionally silent (Figure 7A). We observed a consistent trend for a higher *trans* A/B density ratio when the locus was actively transcribed: 86% of the imaged loci exhibited a greater median *trans* A/B density ratio in the actively transcribing state compared with the silent state (Figure 7B); similarly, 89% of the loci exhibited a greater median transcription firing rate when their *trans* A/B density ratios were in the top versus bottom quartile (Figure 7C). This positive correlation between transcription and local enrichment for A loci was observed across multiple distinct nuclear environments (Figure 7D).

(E and F) Distributions of association rates for A loci (red) and B loci (blue) with the nuclear lamina (E) and nuclear speckles (F). A chromatin locus is considered associated with a nuclear body when the distance of the locus to that body is less than 250 nm.

(G and H) Scatterplots of the median *trans* A/B density ratio for each imaged genomic locus as a function of the frequency with which the locus is associated with the nuclear lamina (G; Pearson correlation coefficient = −0.87) and nuclear speckles (H; Pearson correlation coefficient = 0.66).

(I) Association frequency with nucleoli for all imaged genomic loci, ordered by genomic position. Black vertical lines indicate the locations of centromeres, and red brackets highlight chromosomes containing ribosome-encoding genes (rDNAs).

(J) Correlation of transcription with nuclear structure association. Circles represent the fold change in transcriptional burst frequency for individual genomic loci when comparing the populations of cells in which the locus is lamina associated versus non-lamina-associated (left) and speckle-associated versus non-speckle associated (right). The dotted red line indicates no change, and the solid red lines represent the median fold change in each case.

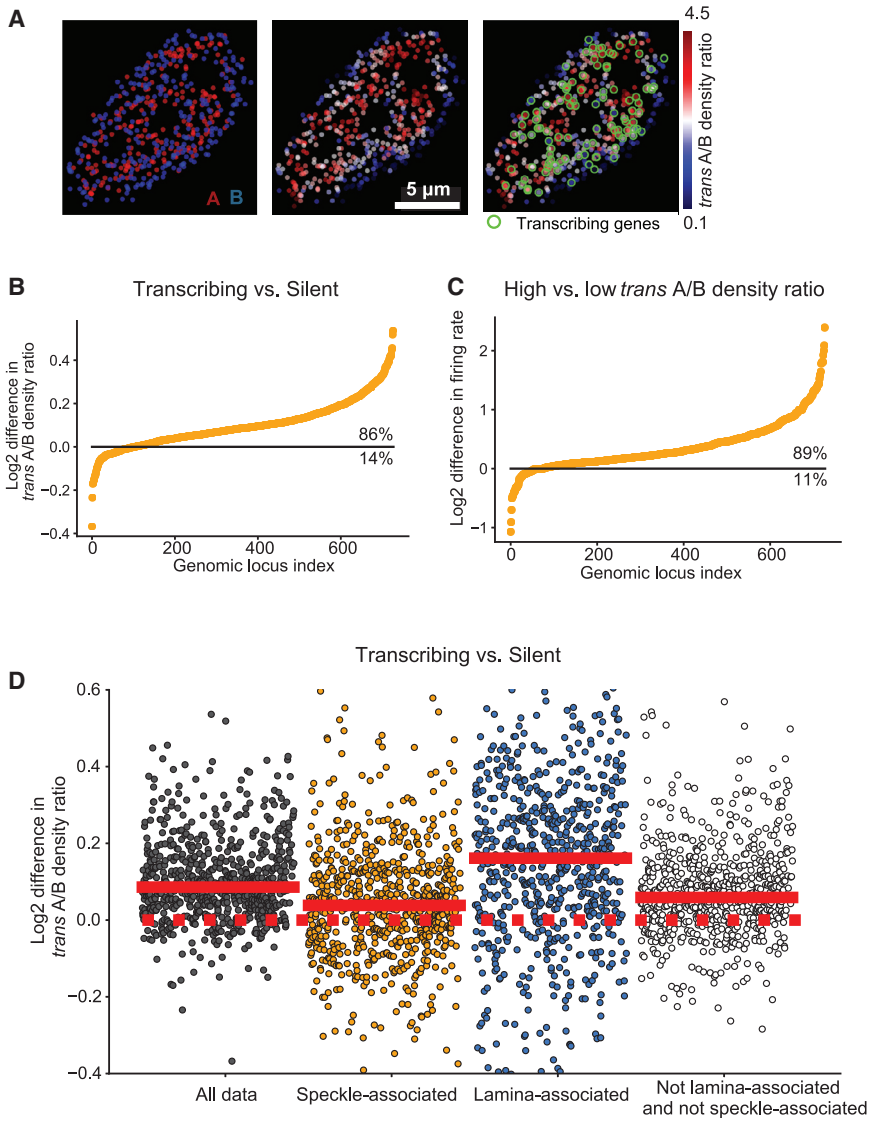


Figure 7. Correlation between Transcriptional Activity and Local Enrichment of *trans*-Chromosomal Active Chromatin

(A) Single-cell images of chromatin loci and transcriptional activities. Left: locations of all imaged A (red) and B (blue) loci in a single z plane from a nucleus. Center: *trans* A/B density ratios for the same loci, color-coded based on the color scale bar. Right: same as the center panel but with detected transcriptional bursts displayed as green circles.

(B) Change (measured in log difference) in the *trans* A/B density ratios for each imaged locus between actively firing and non-firing states. For each genomic locus containing at least one imaged gene, the median *trans* A/B density ratio was calculated for cells in which the genomic locus was actively transcribing and for cells in which it was not transcribing (silent). The log difference of these median values for each imaged locus was rank ordered according to the magnitude of change of these values.

(C) Change (measured in log difference) in the firing rates of the imaged gene-harboring loci between cells in which the *trans* A/B density ratio at the locus is low (bottom quartile) and cells in which this value is high (top quartile). The log-difference in firing rate for the imaged loci was rank-ordered according to the magnitude of the firing rate change.

(D) Change (measured in log difference) in *trans* A/B density ratios between transcribed and silent states for the imaged gene-containing loci, conditioned on their nuclear body association status. For each genomic locus, we computed the fold change in the *trans* A/B density ratio between transcribed and silent states of the locus, considering, from left to right, respectively, all cells (black circles), only cells in which the locus was associated with a nuclear speckle (yellow circles), only cells in which the locus was associated with the lamina (blue circles), and only cells in which the locus was not associated with a nuclear speckle or the lamina (hollow circles). The median *trans* A/B density ratio in each state (transcribed or silent) was determined for each locus, and each association condition and the log2 of the fold change between the two states are shown. The dotted red line represents no change, and the solid red lines represent the median log difference across all loci in each case. Some outliers were omitted to allow clearer visualization of the median log difference (14, 54, and 3 outlier loci for speckle-associated data, lamina-associated data, and not lamina-associated and not-speckle-associated data, respectively).

These observations expand our aforementioned results regarding the relationship between transcriptional activity and *cis* A/B density ratio within chromosomes (Figures 2G–2I) and together reveal a widespread positive correlation between the transcriptional activity of genes and enrichment for active chromatin in their local environment.

DISCUSSION

Chromosome-wide and Genome-Scale Chromatin Imaging

In this paper, we report massively multiplexed chromatin imaging to determine the 3D conformation of chromatin across multiple scales in single cells. Our approach further enabled placement of 3D chromatin organization in its native functional and structural context by combining chromatin tracing with nascent transcript and nuclear structure imaging.

Specifically, we demonstrate two complementary strategies for high-throughput chromatin tracing. First, by expanding the capability of our previously reported multiplexed FISH technology based on sequential hybridization and demonstrating stable imaging over hundreds of hybridization rounds, we provided a high-resolution view of the conformation of whole chromosomes and systematic characterizations of chromatin domains, compartments, and the relationship of transcription to chromatin organization in single cells. Second, for structures that span a

space substantially larger than the diffraction-limited resolution and, hence, allow many loci to be resolved in each imaging round, we developed a MERFISH-based combinatorial labeling strategy for chromatin imaging that allows a much more rapid, non-linear increase of the number of imaged loci with the number of imaging rounds. We demonstrated the power of this latter approach by performing genome-scale imaging of chromatin organization and transcription, simultaneously imaging more than 1,000 genomic loci and nascent transcripts of more than 1,000 genes in individual cells with only tens of hybridization rounds. These data revealed genome-wide *trans*-chromosome interactions and their relationships with nuclear structures and transcription. Putting together this DNA-MERFISH approach with our demonstrated ability to perform hundreds of rounds of hybridization and imaging, we envision that it should be possible to image more than 10,000 genomic loci to provide a high-resolution whole-genome view of chromatin structures in single cells.

Strengths, Limitations, and Future Improvements of the Method

Our high-throughput imaging technology has several key advantages for studying chromatin organization. First, compared with sequencing-based approaches, which rely on proximity information to infer chromatin structure, our method provides direct visualization of chromatin organization and direct measurements of the spatial locations of individual imaged loci in the nuclear context and the physical distances between pairs of imaged loci. Second, our method is intrinsically a single-cell approach and can probe chromatin structures across thousands of individual cells, revealing common features and cell-to-cell variability with robust statistics. Furthermore, the high detection efficiency of individual chromatin loci and pairwise and multiway interactions by our imaging methods provides a high-definition view of chromatin structures in single cells. Third, this chromatin tracing technology can be readily combined with other imaging modalities. This includes multiplexed transcriptional imaging and nuclear structure imaging, as demonstrated in this study, but could also be expanded further to include other modalities, such as imaging of epigenetic modifications or chromatin accessibility (Chen et al., 2016b). Such multi-modal imaging can provide key insights into the relationships between chromatin structure, nuclear organization, and transcription.

Our FISH-based chromatin tracing method also has its own limitations. First, unlike sequencing-based approaches, our method is limited to targeted genomic loci. However, as the genomic throughput of the method increases, we envision that this limitation will be mitigated. Based on the imaging capabilities already demonstrated in this work, we envision that simultaneous imaging of more than 10,000 loci should be possible. Future development could potentially lead to a further increase in the genomic throughput; for example, by exploring the space of combinatorial encoding and advanced image analysis approaches for more efficient, less sparse coding strategies. Second, the *in situ* hybridization process could potentially perturb chromatin structure. Such perturbation is likely small at the scales imaged in this study, which is supported by the quantitative agreement observed between our imaging results and the results from Hi-C. However, for probing DNA structures at smaller scales, this labeling technology might need to be adjusted to include hybridization approaches that do not require heat denaturation, such as CRISPR-based approaches (Deng et al., 2015). Third, the requirement for cell fixation also impedes the technology from imaging chromatin dynamics in live cells. This downside is partially compensated for by the ability to obtain snapshots of chromatin structure with drastically higher genomic throughput than what can be achieved with current live-cell imaging technologies. Furthermore, we expect that the gap between live-cell imaging and multiplexed, fixed-cell imaging could be partially bridged by a different form of multi-modal imaging in which CRISPR-based live imaging of chromatin dynamics of many unidentified chromatin loci is followed by multiplexed FISH to allow post hoc identification of each locus in the genomic space, as demonstrated for repetitive sequences (Guan et al., 2017; Takei et al., 2017). We anticipate that development of CRISPR-based imaging of non-repetitive sequences (Chen et al., 2016a; Knight et al., 2018; Xu and Qi, 2019) together with a CRISPR rainbow strategy to allow imaging of many colors (Ma et al., 2016) will allow many more genomic loci to be imaged in live cells and then identified post hoc by multiplexed FISH.

Applications of the Method

We envision many possible applications of the high-throughput chromatin imaging method reported here. Although, in the current work, we chose to target loci uniformly across chromosomes to provide an unbiased view of the overall 3D chromosome and genome organization, this method could also be used to target genomic loci with specific structural and functional properties. An interesting direction would be to target loci that contain specific genes or regulatory sequences or are bound by specific nuclear architecture proteins to study the interactions between these loci and their relationship with transcription. As a more specific example, we can target a large set of potential promoters and enhancers and study their interactions while simultaneously imaging the transcription activity of the genes governed by the promoters in the same cells. This would aid in understanding how networks of promoters and enhancers interact to regulate transcription. In another direction, many chromatin binding proteins, together with non-coding RNAs, have been reported to take part in physical condensates that organize chromatin in the nucleus (Cho et al., 2018; Chong et al., 2018; Gibson et al., 2019; Larson et al., 2017; Lu et al., 2018; Sabari et al., 2018; Strom et al., 2017), which, in turn, may be important for gene expression regulation. Imaging chromatin organization simultaneously with the structures formed by these protein and RNA factors and together with transcriptional output will be a promising avenue to decipher the relationship between chromatin structures, condensate formation, and transcription regulation. In yet another direction, different cell types exhibit different gene expression profiles that are regulated, in part, by 3D genome organization. Thus, imaging chromatin organization together with gene expression profiles of individual cells in tissues promises to provide critical insights into modes of chromatin organization that are important for cell-type-specific gene expression patterns.

We anticipate that this high-throughput, multiplexed imaging approach will enable study of many questions concerning genomic organization. The rich structural and functional datasets obtained by this imaging approach will help reveal the underlying mechanisms governing genome organization and its role in regulating various genomic functions as well as reveal exciting new questions to be addressed in chromatin biology.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Culture
- METHOD DETAILS
  - Oligonucleotide Probe Design
  - Experimental Setup
  - Experimental Procedures and Protocols
  - Image Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Estimation of the Nuclear Lamina Location
  - Spatial Distance between a Locus Pair or between Locus and Nuclear Body
  - Proximity Frequency Matrices from Imaging
  - A/B compartment calling
  - Local A/B Density Score
  - Insulation Score from Imaging Data
  - TAD Calling and Single-Cell Domain Calling
  - A/B Segregation Score within Chromosomes
  - Estimation of Detection Efficiency of Nascent RNAs in Genome-scale Imaging
  - Hi-C Data Analysis
  - CTCF and Rad21 ChIP-seq Data Analysis

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.07.032.

ACKNOWLEDGMENTS

We thank Stephen Eichhorn for help with RNA probe design. This work is supported in part by the National Institutes of Health (to X.Z.). J.-H.S. and P.Z. are in part supported by the Harvard Molecules, Cells and Organisms Training Program. S.S.K. is supported in part by the Harvard Biophysics Training Program. X.Z. is a Howard Hughes Medical Institute Investigator.

AUTHOR CONTRIBUTIONS

J.-H.S., P.Z., S.S.K., B.B., and X.Z. designed the experiments. J.-H.S., P.Z., S.S.K., and B.B. performed experiments and analyses. J.-H.S., P.Z., S.S.K., B.B., and X.Z. interpreted data and wrote the manuscript. X.Z. supervised the study.

DECLARATION OF INTERESTS

J.-H.S., P.Z., S.S.K., B.B., and X.Z. are inventors of a patent applied for by Harvard University related to the imaging technology. X.Z. is a co-founder and consultant of Vizgen, Inc.

Received: February 21, 2020
Revised: June 19, 2020
Accepted: July 21, 2020
Published: August 20, 2020

REFERENCES

Abramo, K., Valton, A.-L., Venev, S.V., Ozadam, H., Fox, A.N., and Dekker, J. (2019). A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. Nat. Cell Biol. *21*, 1393–1402.

Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C.A., Chotalia, M., Xie, S.Q., Barbieri, M., de Santiago, I., Lavitas, L.-M., Branco, M.R., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. Nature *543*, 519–524.

Bickmore, W.A. (2013). The spatial organization of the human genome. Annu. Rev. Genomics Hum. Genet. *14*, 67–84.

Bintu, B., Mateo, L.J., Su, J., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N., and Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science *362*, eaau1783.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R., and Cremer, T. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. *3*, e157.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. *4*, e138.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.

Cardozo Gizzi, A.M., Cattoni, D.I., Fiche, J.-B., Espinola, S.M., Gurgo, J., Messina, O., Houbron, C., Ogiyama, Y., Papadopoulos, G.L., Cavalli, G., et al. (2019). Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. Mol. Cell *74*, 212–222.e5.

Cavalli, G., and Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. Nature *571*, 489–499.

Chen, Y., and Belmont, A.S. (2019). Genome organization around nuclear speckles. Curr. Opin. Genet. Dev. *55*, 91–99.

Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. Science *348*, aaa6090.

Chen, B., Guan, J., and Huang, B. (2016a). Imaging Specific Genomic DNA in Living Cells. Annu. Rev. Biophys. *45*, 1–23.

Chen, X., Shen, Y., Draper, W., Buenrostro, J.D., Litzenburger, U., Cho, S.W., Satpathy, A.T., Carter, A.C., Ghosh, R.P., East-Seletsky, A., et al. (2016b). ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. Nat. Methods *13*, 1013–1020.

Chen, Y., Zhang, Y., Wang, Y., Zhang, L., Brinkman, E.K., Adam, S.A., Goldman, R., van Steensel, B., Ma, J., and Belmont, A.S. (2018). Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. J. Cell Biol. *217*, 4025–4048.

Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412–415.

Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic

and selective low-complexity domain interactions that control gene transcription. Science *361*, eaar2555.

Cordes, T., Vogelsang, J., and Tinnefeld, P. (2009). On the mechanism of Trolox as antiblinking and antibleaching reagent. J. Am. Chem. Soc. *131*, 5018–5019.

Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240–244.

Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harb. Perspect. Biol. *2*, a003889.

Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110–1121.

Dekker, J., Belmont, A.S., Guttman, M., Leshyk, V.O., Lis, J.T., Lomvardas, S., Mirny, L.A., O'Shea, C.C., Park, P.J., Ren, B., et al.; 4D Nucleome Network (2017). The 4D nucleome project. Nature *549*, 219–226.

Dempsey, G.T. (2013). A User's Guide to Localization-Based Super-Resolution Fluorescence Imaging. In Methods in Cell Biology, *114*, G. Sluder and D.E. Wolf, eds. (Elsevier Inc.), pp. 561–592.

Deng, W., Shi, X., Tjian, R., Lionnet, T., and Singer, R.H. (2015). CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. Proc. Natl. Acad. Sci. USA *112*, 11870–11875.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst. *3*, 99–101.

Falk, M., Feodorova, Y., Naumova, N., Imakaev, M., Lajoie, B.R., Leonhardt, H., Joffe, B., Dekker, J., Fudenberg, G., Solovei, I., and Mirny, L.A. (2019). Heterochromatin drives compartmentalization of inverted and conventional nuclei. Nature *570*, 395–399.

Fang, R., Yu, M., Li, G., Chee, S., Liu, T., Schmitt, A.D., and Ren, B. (2016). Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. Cell Res. *26*, 1345–1348.

Flyamer, I.M., Gassler, J., Imakaev, M., Brandão, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110–114.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor-α-bound human chromatin interactome. Nature *462*, 58–64.

National Center for Biotechnology Information. (2020). Genome Data Viewer. https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_000001405.39.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell *179*, 470–484.e21.

Girelli, G., Custodio, J., Kallas, T., Agostini, F., Wernersson, E., Spanjaard, B., Mota, A., Kolbeinsdottir, S., Gelali, E., Crosetto, N., and Bienko, M. (2020). GPSeq reveals the radial organization of chromatin in the cell nucleus. Nat. Biotechnol. Published online May 25, 2020. https://doi.org/10.1038/s41587-020-0519-y.

Guan, J., Liu, H., Shi, X., Feng, S., and Huang, B. (2017). Tracking Multiple Genomic Elements Using Correlative CRISPR Imaging and Sequential DNA FISH. Biophys. J. *112*, 1077–1084.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. Cell *169*, 13–23.

Iarovaia, O.V., Minina, E.P., Sheval, E.V., Onichtchouk, D., Dokudovskaya, S., Razin, S.V., and Vassetzky, Y.S. (2019). Nucleolus: A Central Hub for Nuclear Functions. Trends Cell Biol. *29*, 647–659.

Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204–2207.

Knight, S.C., Tjian, R., and Doudna, J.A. (2018). Genomes in Focus: Development and Applications of CRISPR-Cas9 Imaging Technologies. Angew. Chem. Int. Engl. *57*, 4329–4337.

Krijger, P.H.L., and de Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. Nat. Rev. Mol. Cell Biol. *17*, 771–782.

Lakadamyali, M., and Cosma, M.P. (2020). Visualizing the genome in high resolution challenges our textbook understanding. Nat. Methods *17*, 371–379.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature *547*, 236–240.

Lee, D.-S., Luo, C., Zhou, J., Chandran, S., Rivkin, A., Bartlett, A., Nery, J.R., Fitzpatrick, C., O'Connor, C., Dixon, J.R., and Ecker, J.R. (2019). Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. Nat. Methods *16*, 999–1006.

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13–25.

Levsky, J.M., and Singer, R.H. (2003). Fluorescence in situ hybridization: past, present and future. J. Cell Sci. *116*, 2833–2838.

Li, G., Liu, Y., Zhang, Y., Kubo, N., Yu, M., Fang, R., Kellis, M., and Ren, B. (2019). Joint profiling of DNA methylation and chromatin architecture in single cells. Nat. Methods *16*, 991–993.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293.

Liu, M., Lu, Y., Yang, B., Chen, Y., Radda, J.S.D., Hu, M., Katz, S.G., and Wang, S. (2020). Multiplexed imaging of nucleome architectures in single cells of mammalian tissue. Nature Communications *11*, 2907.

Lloyd, S.P. (1982). Least squares quantization in PCM. Special issue on quantization. IEEE Trans. Inf. Theory *28*, 129–137.

Lu, H., Yu, D., Hansen, A.S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., and Zhou, Q. (2018). Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature *558*, 318–323.

Ma, H., Tu, L.C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., and Pederson, T. (2016). Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. Nat. Biotechnol. *34*, 528–530.

Mateo, L.J., Murphy, S.E., Hafner, A., Cinquini, I.S., Walker, C.A., and Boettiger, A.N. (2019). Visualizing DNA folding and RNA in embryos at single-cell resolution. Nature *568*, 49–54.

Moffitt, J.R., Hao, J., Wang, G., Chen, K.H., Babcock, H.P., and Zhuang, X. (2016). High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. Proc. Natl. Acad. Sci. USA *113*, 11046–11051.

Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat. Methods *13*, 919–922.

Nagano, T., Lubling, Y., Várnai, C., Dudley, C., Leung, W., Baran, Y., Mendelson Cohen, N., Wingett, S., Fraser, P., and Tanay, A. (2017). Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature *547*, 61–67.

Nir, G., Farabella, I., Pérez Estrada, C., Ebeling, C.G., Beliveau, B.J., Sasaki, H.M., Lee, S.D., Nguyen, S.C., McCole, R.B., Chattoraj, S., et al. (2018). Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. PLoS Genet. *14*, e1007872.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Palstra, R.-J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de Laat, W. (2008). Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. PLoS ONE *3*, e1661.

Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., et al. (2018). Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. Cell *174*, 744–757.e24.

Ramani, V., Deng, X., Qiu, R., Gunderson, K.L., Steemers, F.J., Disteche, C.M., Noble, W.S., Duan, Z., and Shendure, J. (2017). Massively multiplex single-cell Hi-C. Nat. Methods *14*, 263–266.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Rasnik, I., McKinney, S.A., and Ha, T. (2006). Nonblinking and long-lasting single-molecule fluorescence imaging. Nat. Methods *3*, 891–893.

Rooijers, K., Markodimitraki, C.M., Rang, F.J., de Vries, S.S., Chialastri, A., de Luca, K.L., Mooijman, D., Dey, S.S., and Kind, J. (2019). Simultaneous quantification of protein-DNA contacts and transcriptomes in single cells. Nat. Biotechnol. *37*, 766–772.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Co-activator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958.

Sawh, A.N., Shafer, M.E.R., Su, J., Zhuang, X., Wang, S., and Mango, S.E. (2020). Lamina-Dependent Stretching and Unconventional Chromosome Compartments in Early C. elegans Embryos. Molecular Cell *78*, 96–111.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458–472.

Shachar, S., and Misteli, T. (2017). Causes and consequences of nuclear gene positioning. J. Cell Sci. *130*, 1501–1508.

Solovei, I., Thanisch, K., and Feodorova, Y. (2016). How to rule the nucleus: divide et impera. Curr. Opin. Cell Biol. *40*, 47–59.

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A., et al. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature *544*, 59–64.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature *547*, 241–245.

Takei, Y., Shah, S., Harvey, S., Qi, L.S., and Cai, L. (2017). Multiplexed Dynamic Imaging of Genomic Loci by Combined CRISPR Imaging and DNA Sequential FISH. Biophys. J. *112*, 1773–1776.

Tan, L., Xing, D., Chang, C.-H., Li, H., and Xie, X.S. (2018). Three-dimensional genome structures of single diploid human cells. Science *361*, 924–928.

van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780–791.

Wang, S., Su, J.-H., Beliveau, B.J., Bintu, B., Moffitt, J.R., Wu, C., and Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. Science *353*, 598–602.

Xu, X., and Qi, L.S. (2019). A CRISPR-dCas Toolbox for Genetic Engineering and Synthetic Biology. J. Mol. Biol. *431*, 34–47.

Yaffe, E., and Tanay, A. (2011). Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat. Genet. *43*, 1059–1065.

Yu, M., and Ren, B. (2017). The Three-Dimensional Organization of Mammalian Genomes. Annu. Rev. Cell Dev. Biol. *33*, 265–289.

Zheng, M., Tian, S.Z., Capurso, D., Kim, M., Maurya, R., Lee, B., Piecuch, E., Gong, L., Zhu, J.J., Li, Z., et al. (2019). Multiplex chromatin interactions with single-molecule precision. Nature *566*, 558–562.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal [SC-35] to SC35 - Nuclear Speckle Marker	Abcam	Cat# ab11826; RRID:AB_298608
Anti-Geminin antibody - cell cycle marker	Abcam	Cat# ab195047; RRID:AB_2832993
Anti-Fibrillarin antibody - Nucleolar Marker	Abcam	Cat# ab5821; RRID:AB_2105785
Cy5 AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch Labs	Cat# 715-175-150; RRID:AB_2340819
Alexa Fluor® 647 AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch Labs	Cat# 711-605-152; RRID:AB_2492288
Chemicals, Peptides, and Recombinant Proteins		
1,6-Hexanediol	Sigma	Cat# 240117
α-amanitin	Sigma	Cat# A2263
Formamide	Ambion	Cat# AM9342
20xSSC	Ambion	Cat# AM9763
Triton-X	Sigma	Cat# T8787
Glucose oxidase	Sigma	Cat# G2133
EDTA	Boston Bioproducts	Cat# BM-150
EvaGreen Dye, 20x	Biotium	Cat# 31000
Phusion® Hot Start Flex 2X Master Mix	New England Biolabs	Cat# M0536
HiScribe T7 Quick High Yield RNA Synthesis Kit	New England Biolabs	Cat# E2050
Maxima H Minus Reverse Transcriptase	ThermoFisher	Cat# EP0752
dNTP mix	ThermoFisher	Cat# R1121
Fetal Bovine Serum, ESC qualified	ATCC	Cat# SCRR-30-2020
Eagle's Minimum Essential Media (EMEM)	ATCC	Cat# 30-2003
Fibronectin Bovine Plasma	Sigma	Cat# F1141
32% Paraformaldehyde	Fisher Scientific	Cat# 50-980-494
RNase A	ThermoFisher	Cat# EN0531
Dextran Sulfate	Sigma	Cat# D8906
Human Cot-1 DNA	ThermoFisher	Cat# 15279011
FluoSpheres Carboxylate-Modified Microspheres, 0.2 μm, blue fluorescent (365/415), 2% solids	ThermoFisher	Cat# F8805
FluoSpheres Carboxylate-Modified Microspheres, 0.04 μm, orange fluorescent (540/560), 5% solids, azide free	ThermoFisher	Cat# F8792
DAPI	ThermoFisher	Cat# D1306
RNase inhibitor, Murine	New England Biolabs	Cat# M0314
RNasin Plus RNase Inhibitor	Fisher Scientific	Cat# PRN2615
1M Tris, pH 8	ThermoFisher	Cat# 15568025
Catalase	Sigma	Cat# C3155
6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox)	Sigma	Cat# 238813
tris(2-carboxyethyl)phosphine (TCEP)	Sigma	Cat# C4706
Tween-20	Sigma	Cat# P9416
Bovine Serum Albumin	Jackson ImmunoResearch	Cat# 001-000-162

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
IMR90	ATCC	Cat# CCL-186, RRID:CVCL_0347
Oligonucleotides		
Primers, readout probes and adaptor probes	Integrated DNA Technologies, Bio-Synthesis Inc.	See Table S3
Primary/encoding oligonucleotide probe libraries	Twist Bioscience	See Tables S2 and S4
Software and Algorithms		
Software for analyzing multiplexed FISH data	This paper	GitHub: https://github.com/ZhuangLab/Chromatin_Analysis_2020_cell
Straw (juicebox)	Durand et al., 2016	GitHub: https://github.com/aidenlab/straw
UCSC bigWig and bigBed	Kent et al., 2010	http://hgdownload.soe.ucsc.edu/admin/exe
Deposited Data		
Multiplexed FISH data	This paper	Zenodo:https://doi.org/10.5281/zenodo.3928890
Hi-C data	Rao et al., 2014	GEO: GSE63525
RAD21 CHIP-seq data	Dunham et al., 2012	GEO: GSM935624
CTCF CHIP-seq data	Dunham et al., 2012	GEO: GSM935404

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Xiaowei Zhuang (zhuang@chemistry.harvard.edu).

Materials Availability

Oligonucleotide probe sequences used for chromatin and RNA imaging can be found in Tables S2, S3, and S4. These probes or templated for making these probes can be purchased from commercial sources, as detailed in the Key Resources Table.

Data and Code Availability

The datasets reported in this paper is available at Zenodo (https://doi.org/10.5281/zenodo.3928890). Analysis software is available at https://github.com/ZhuangLab/Chromatin_Analysis_2020_cell.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Culture

Cells were prepared similarly to our previous studies (Bintu et al., 2018; Wang et al., 2016). IMR-90 cells were purchased from American Type Culture Collection (ATCC, CCL-186) and grown in the recommended growth media (Eagle's Minimum Essential Media (EMEM; ATCC 30-2003), containing 10% w/v Bovine Serum Albumin (BSA; Jackson Immunoresearch 001-000-162) and 1% v/v Penicillin-Streptomycin (Sigma P0781)) at 37°C and 5% CO_2. Media was exchanged every 2 days and the cells were passaged whenever the culture reached confluency (every 2-5 days). To avoid potential alterations to chromatin structure, all chromatin tracing experiments in this study were performed on cells cultured for 6 weeks at most.

METHOD DETAILS

Oligonucleotide Probe Design

Choice of target genomic regions

For high-resolution whole-chromosome imaging by sequential hybridization, we first partitioned the target chromosome into 50-kb segments. After screening out repetitive elements and regions where < 100 unique probes can be designed per 50-kb segments, we kept a total of 651 target genomic loci for Chr21 and 4,844 target genomic loci for Chr2. We then designed primary probes for each 50-kb segment (~500 oligonucleotide probes) and kept the 250-360 most centrally positioned probes per segment for sequential imaging. For Chr21 imaging, all 651 genomic loci were imaged. For Chr2 imaging, we aimed for a 250-kb genomic resolution, and hence only designed primary probes for one in every five 50-kb segments.

For imaging nascent RNA transcripts on Chr21, we selected the genes for which we can design > 50 primary probes (see "Primary/encoding probe design" section, below) on their first introns from all the protein-coding genes on Chr21. We imaged ~80 genes that are interspersed across Chr21. In order to facilitate more accurate detection of the spatial positions of transcription initiation events, we also designed probes to target a 5-kb segment of DNA around the transcription start site (TSS) of each gene.

For genome-scale chromatin imaging by the combinatorial imaging strategy, DNA-MERFISH, we chose genomic loci for imaging in the following way: For each human chromosome (except the Y chromosome which is not present IMR-90 cells), we selected a 100-kb segment every ~3 Mb of spacing. If this spacing resulted in less than 30 selected loci on a given chromosome, the spacing was reduced for that chromosome, until all chromosomes had at least 30 loci selected. This resulted in a total of 1,041 target genomic loci for imaging, and the number of loci in individual chromosomes ranged from 30 to 80. We then designed encoding probes for an ~30-kb region within each 100-kb segment (~400 oligonucleotide probes) for the combinatorial FISH imaging.

For imaging of nascent RNA transcripts in the genome-scale imaging, we selected all intron-containing genes that completely or partially overlap with the 1,041 targeted genomic loci. We then designed encoding probes for the introns of all of these RNAs such that each RNA had ~20 encoding probes and that the targeting sequences of the encoding probes were kept as close as possible to the transcription start site. A total of 1,137 genes were targeted.

Barcode design for genome-scale imaging by combinatorial FISH (DNA-MERFISH)

Binary barcodes for imaging the 1,041 genomic loci were chosen in the following fashion. First, we generated all possible 100-bit binary barcodes with a Hamming weight of 2 (HW2, i.e., each barcode containing two "1" bits and 98 "0" bits) and randomly selected 1,041 barcodes from this list. The selected barcodes were then arbitrarily assigned to the 1,041 genomic loci first. Next, barcodes were exchanged randomly between the used and unused 100-bit HW2 code pool, as well as between loci from different chromosomes, in order to minimize, for each chromosome, the variance in the number of loci appearing (i.e., reading "1") across different bits. This resulted in an approximately equal number of loci imaged per bit for each chromosome. To optimize association of barcodes to loci within each chromosome, we allowed loci within the same chromosome to exchange barcodes and optimized for the largest minimal genomic distance between loci with barcodes reading "1" at the same code position. When comparing code assignments with identical minimal genomic distances, we selected the one that minimized the coefficient of variation of genomic distances (so that genomic distances have both larger means and smaller standard deviations).

Barcodes for imaging the nascent RNA transcripts of the 1,137 genes were chosen similarly, but using a 54-bit, HW2 code instead of a 100-bit, HW2 code.

Primary/encoding probe design

Primary/encoding probes for chromatin imaging were synthesized from a pool of oligonucleotides purchased from Twist Biosciences. Each oligo in this pool consisted of the following sub-sequences (from 5′ to 3′) (Tables S2 and S4):

1. A 20-nucleotide (nt) or 19-nt forward priming region for PCR amplification and reverse transcription (RT)
2. A 20-nt readout sequence corresponding to the genomic locus targeted by the probe in the case of sequential imaging or one of the bits in which the genomic locus targeted by the probe will be imaged in the case of combinatorial (DNA-MERFISH) imaging
3. A 42-nt or 40-nt target sequence (for sequential or combinatorial imaging, respectively), designed to bind uniquely to a single targeted genomic locus
4. An additional 1-2 copies of the 20-nt readout sequence as described above
5. A 20-nt or 19-nt reverse priming sequence for PCR amplification

Similar designs with minor modifications were used for nascent RNA imaging (see Tables S2 and S4).

The forward and reverse priming sequences were chosen from a previously generated list of random 20-nt sequences optimized for PCR, as described previously (Chen et al., 2015).

The readout sequences were chosen via the following process. First, a list of 30-nt sequences with minimal homology to the human genome was created, as previously described (Bintu et al., 2018; Chen et al., 2015; Wang et al., 2016). Then, we ranked a subset of these sequences by observed signal to noise ratio (SNR) and chose the best ones as readout probe sequences for DNA-MERFISH imaging (see Table S3 for readout probes). For sequential imaging, we needed substantially more readout sequences due to the larger number of hybridization rounds. Hence, we followed the same procedure outlined previously (Chen et al., 2015) to select ~1,200 candidate readout probe sequences. Then we filtered these candidates to ensure a GC content of 40%–60% and a melting temperature of 57-67°C. These sequences were further filtered using BLAST (Camacho et al., 2009) such that no readout probe sequence had hits with high-scoring segment pair (HSP) score larger or equal to 17. Lastly, the readout sequences were chosen by reverse-complementing the last 20-nt of each of these probe sequences generated above.

The 42-nt or 40-nt target sequences were chosen similarly to a procedure described previously (Bintu et al., 2018; Moffitt et al., 2016). Briefly, for each genomic region of interest (see the "Choice of target genomic regions" section above), we used a probe design algorithm encompassing the following steps:

1. We create a list of all 42-nt or 40-nt sequences complementary to the genomic region of interest (starting at each possible base in the targeted region).

2. Sequences were then filtered by requiring them to be within a defined range of melting temperatures and GC content.
3. The remaining sequences were then further filtered by limiting the allowed degree of homology to the human genome, the human transcriptome and a database containing repetitive sequences using the same procedure as previously (Bintu et al., 2018).
4. The final target sequences were then selected from the remaining sequences after the filtering step 3, such that no genomic overlap exists between the targets of any two target sequences.

For genome-scale imaging by DNA-MERFISH, we noticed in preliminary experiments that some rounds of hybridization displayed a relatively high background signal due to a small subset of probes with repetitive target sequences that exhibited off-target binding, which were not fully filtered in our initial implementation of step 3 above. We have since improved our implementation of the algorithm. This improved algorithm is provided at https://github.com/ZhuangLab/Chromatin_Analysis_2020_cell. Two options could be used to remove these remaining repetitive sequences from the encoding probe library or the effect of these repetitive sequences in imaging. Option 1: Use the improved algorithm to filter out these repetitive sequences in step 3 described above and then perform step 4 described above to re-generate the final target sequences with no genomic overlap. We note that the removal of these repetitive sequences would lead to a different total number and different starting positions of the final target sequences generated in step 4 as compared to the original encoding probe library. Option 2: Use the improved algorithm to identify the repetitive sequences, generate another library of competing probes complementary to these sequences, and use the competing probes together with the original encoding probe library to prevent binding of these repetitive target sequences. Since the original encoding probe library, which includes ~400,000 oligos, is expensive and only a small fraction of probes in the original encoding probe library had such repetitive target sequences, Option 2 was a more cost-effective solution and we thus used this approach. Experimentally, during the encoding probe hybridization step, we mixed this competing probe library with the original encoding probe library such that each of the competing probes was in > 5-fold excess of its complement in the encoding probe library. This resulted in successful removal of the high background signal. We list the original encoding probe library in Table S4, except for the probes that were competed out via this competition approach.

The target sequences used for whole-chromosome imaging by sequential hybridization were generated using the improved implementation of the algorithm to remove the repetitive sequences prior to generating the final target sequences with no genomic overlap in Step 4. These probes were also subjected to an additional computational filtering step using BLAST, to ensure high specificity to the intended genomic locus.

For whole-chromosome imaging by sequential hybridization, all of the 42-nt target sequences for each target genomic locus were matched with a unique readout sequence associated with that locus. To generate primary probe sequences, each target sequence was concatenated to one (for Chr21) or two (for Chr2) identical copies of the readout sequence assigned and then concatenated to the forward and reverse PCR primers. For the primary probes targeting TSS, we started with the targeting sequences for the designed 50-kb Chr21 tracing primary probes and kept only those sequences targeting the upstream and downstream 2.5-kb of eachTSS sites.

For generating the full-length encoding probes for genome-scale imaging by DNA-MERFISH, each of the chosen 40-nt target sequences for each target genomic locus was alternatingly assigned to 2 groups spanning the entire target locus. Each of these groups was associated with a single readout sequence, corresponding to one of the two bits in which the locus would be imaged. Then, each target sequence was concatenated to two identical copies of the readout sequence assigned to its group, and then concatenated to the forward and reverse PCR primers.

For a small subset of loci in Chr6 that were re-imaged after DNA-MERFISH by sequential hybridization, the target sequences were alternatingly assigned to three groups, two for generating encoding probes for DNA-MERFISH and one for generating primary probes used in re-imaging of these loci by sequential hybridization, as described in the "Sequential relabeling of genomic loci in combinatorial FISH imaging" section below. Each of these groups was associated with a single readout sequence, and each target sequence was concatenated to two identical copies of the readout sequence assigned to its group, and then concatenated to the forward and reverse PCR primers.

Probes for RNA imaging were designed similarly, with the following exceptions:

a. For genome-scale imaging, RNA-targeting probes contained 3 copies of an identical readout sequence on every probe, one at the 5′ end and two at the 3′ end of the target region.
b. For whole-chromosome imaging, RNA-targeting probes had 30-nt target sequences, and contained four copies of an identical readout sequence on every probe.

Readout sequences for RNA imaging were orthogonal to those used for DNA imaging and were selected from the same ranked list of tested readout sequences.

A complete list of all primary/encoding probes used in this study can be found in Table S2 (for whole-chromosome DNA, RNA and TSS imaging) and Table S4 (for genome-scale DNA and RNA imaging).

Experimental Setup

Overview of experimental system

The physical setup used for performing these experiments consists of several components. A custom-built fluorescence microscope was used to acquire images, while a custom-built fluidics system was used to automatically perform buffer exchanges on the microscope stage. Custom software was used to synchronize and control the various components, and to automate many experimental steps (https://github.com/ZhuangLab/storm-control). Below is a detailed description of each of these elements.

Microscope setup for Image acquisition

Image acquisition was performed using a custom-built microscope system (Bintu et al., 2018; Chen et al., 2015). The system was built around a Nikon Ti-U microscope body with a Nikon CFI Plan Apo Lambda 60x oil immersion objective with 1.4 NA.

Illumination was based on one of two alternatives:

1. Solid-state, single-mode lasers with the following wavelengths: 405 nm (Coherent, Obis 405 nm LX 200 mW), 560 nm (MPB Communications, 2RU-VFL-P-2000-560-B1R), 647 nm (MPB Communication, 2RU-VFL-P-1500-647-B1R) and 750 nm (MPB Communication, 2RU-VFL-P-500-750-B1R). In this case, the output of the 560-nm, 647-nm and 750-nm lasers were controlled by an acousto-optic tunable filter (AOTF) while the 405-nm laser was controlled directly via its laser control box. A custom dichroic (Chroma, zy405/488/561/647/752RP-UF1) and emission filter (Chroma, ZET405/488/461/647-656/752 m) were used to separate excitation and emission illuminations.
2. A Lumencor CELESTA light engine (a fiber-coupled solid-state laser-based illumination system) with the following wavelengths: 405 nm, 446 nm, 477 nm, 520 nm, 546 nm, 638 nm and 749 nm. This system was used with a penta-bandpass dichroic (IDEX, FF421/491/567/659/776-Di01-25x36) and a penta-bandpass filter (IDEX, FF01-441/511/593/684/817-25).

In most experiments, the illumination was flattened using a Refractive Beam Shaper (Newport Optics, GBS-AR14) or a vibrating optical fiber (Errol, custom Albedo unit).

A scientific CMOS camera (Hamamatsu FLASH4.0 or Hamamatsu C13440 with factory calibration for single-molecule imaging) was used for image acquisition. Sample position in three dimensions was controlled using a XYZ stage (Ludl). A custom-built auto-focus system (Dempsey, 2013) was used to maintain a constant focal plane over prolonged periods of time. This was achieved by comparing the relative position of two IR laser (Thorlabs, LP980-SF15) beams reflected from the glass-fluid interface and imaged on a separate CMOS camera (Thorlabs, uc480).

The different components were synchronized and controlled using a National Instruments Data Acquisition card (NI PCIe-6353) and custom software (see "Software for controlling experimental components" below).

Fluidics system configuration

The fluidics system consisted of several main components: a pump, a set of valves connected in series, a flow chamber in which the sample was mounted, and tubing and connectors. We used a peristaltic pump (Gilson, MINIPLUS 3) to generate flow in the system. The pump was connected to an array of 8-way valves (Hamilton, MVP and HVXM 8-5), connected in series. In this study, we used 3-5 valves connected in such a manner. Each valve's last connection was used as the input of the next valve in the series (except for the last one), while the rest were connected to a tube containing the buffer for a single round of hybridization. A fixed subset of the valves was used for imaging, bleaching and wash buffers (see "Experimental procedures and protocols" section). This valve system was used to flow the various buffers into the flow chamber (Bioptechs, 060319-2), in which the sample was placed. The chamber output was connected to a waste collection vessel, forming an open flow system. Components were connected using elastic plastic tubing, and connections were sealed using a pressure adessive (Blu-tack). The system was controlled using a custom software (see "Software for controlling experimental components" below). Overall, this system allowed for 20–34 rounds of hybridization (depending on the number of valves and the number of spots reserved for special buffers). In experiments where the number of hybridization rounds exceeded the capacity of the flow system, we replaced the buffers with new ones via the following procedure:

1. The output from the valve system was directly connected to the waste collection vessel, bypassing the sample-containing chamber.
2. All valves were washed using 30% formamide and double-distilled water.
3. The new set of buffers was introduced, and the chamber was reconnected to the flow system.
4. The experiment resumed for the next round of hybridization.

Software for controlling experimental components

All system components were controlled using custom-built software, available at https://github.com/ZhuangLab/storm-control. This software package is composed of several main modules, which work in concert:

1. "Hal" is the software package used to control and synchronize all illumination and microscope components. We note that in some cases it will be necessary to write drivers for components, which are not included in this package. Hal is also used to define imaging parameters, such as illumination strength, sequence of stage and illumination operations during imaging (e.g., during a z-scan), exposure time etc.

2. "Steve" is a module used to take mosaic images (i.e., a composite image made up of many individual fields of view) and select regions for imaging in experiments.
3. "Kilroy" is the software used to control the fluidics components, and to define pre-programmed sequences of operations to be performed as sets (e.g., the set of operations undertaken when a new round of hybridization is performed).
4. "Dave" can issue commands to both Hal and Kilroy, and is used to automate the performance of data collection by defining in advance a complete set of fluidics system and microscope operations, as well as the order and time-lag in which they are to be performed.

The general flow of an experiment is that, before the experiment starts, Hal and Kilroy are loaded with the parameters and specifications to be used. After the sample is loaded and the chamber is filled with imaging buffer, a mosaic image of the DAPI channel is taken using Steve, and regions of interest are selected. A file is then generated to specify the sequence of operations throughout the entire experiment and is loaded to Dave, together with the coordinates of the selected regions of interest. The rest of the experiment is run automatically, without manual intervention. If the number of rounds in the experiment exceeds the capacity of the flow system, the automatic sequence specifies actions up to the capacity of the system. The buffers are then replaced (see "Fluidics system configuration" section above), a new Dave file is created, and this is repeated until all rounds of imaging are completed.

Experimental Procedures and Protocols

Primary/encoding probe synthesis

Primary/encoding probes were amplified from the template library described above (see "Primary/encoding probe design" above). This was done using a previously described amplification protocol (Chen et al., 2015) involving the following steps:

1. The initial oligo pool was amplified using limited-cycle PCR for approximately 20 cycles. The reverse primer used in this step also introduced a T7 promoter sequence via primer extension.
2. The resulting product was purified via column purification and underwent further amplification and conversion to RNA by a high-yield *in-vitro* transcription reaction.
3. The RNA product was converted back to single-stranded DNA by a reverse transcription reaction.
4. The product of the previous step was subjected to alkaline hydrolysis (to remove residual RNA) and column purified (DNA Clean & Concentrator Kit, Zymo Research D4003 and D4033).
5. If necessary, the product of the previous step was dried in vacuum and resuspended in water to achieve the desired concentration of primary probe.

All primers were purchased from Integrated DNA Technologies (IDT).

Readout and adaptor probe preparation

All readout and adaptor probes were ordered from IDT (see Table S3) and were diluted directly from this stock, as described in the "Sequential hybridization of readout probes for sequential or combinatorial FISH imaging" section, below.

Cell culture preparation and primary/encoding probe hybridization

Cells were prepared similarly to our previous studies (Bintu et al., 2018; Wang et al., 2016). IMR-90 cells were purchased from American Type Culture Collection (ATCC, CCL-186) and cultured as described in the "Experimental Model and Subject Details" section above.

To prepare for whole-chromosome imaging, cells were plated on 40-mm, round #1.5 coverslips (Bioptechs, 0420-0323-2), at a density of ~500,000 cells per coverslip. Cells were allowed to grow for ~2 days until confluency at 37°C and 5% CO_2. In the Pol II transcription inhibition experiments, cell media was replaced with fresh media containing 100 μg/mL α-amanitin (Sigma-Aldrich, A2263) 6 hours prior to cell fixation. We then fixed the culture using 4% paraformaldehyde (PFA) in PBS for 10 minutes at room temperature and washed in PBS 2-3 times. Cells were then permeabilized in two steps: first, they were treated with 0.5% v/v Triton-X (Sigma-Aldrich, T8787) in PBS for 10 minutes at room temperature. Then, cells were treated with 0.1 M hydrochloric acid (HCl) for 5 minutes at room temperature and washed in PBS 2-3 times. Following HCl treatment, cells were treated with a solution of 0.1 mg/mL RNase A (ThermoFisher, EN0531) dissolved in PBS for 30-45 minutes at 37°C, to remove potential sources of off-target binding to RNA. Following this treatment, cells were incubated in pre-hybridization buffer, consisting of 2x saline-sodium citrate buffer (SSC; Ambion, AM9763), 50% formamide (Ambion, AM9342) and 0.1% of Tween-20 (Sigma-Aldrich, P4916) for approximately 30 minutes. Next, the cell coverslip was inverted and placed on a drop of 50 μL of hybridization buffer (2x SSC, 50% formamide, 10% dextran sulfate (Sigma-Aldrich, D8906) containing a mixture of primary/encoding probes at ~20 μM total concentration with or without 10 μg Human Cot-1 DNA (ThermoFisher, 15279011)) in a 60-mm Petri dish. The dish was partially submerged in a water bath at ~90°C for 3 minutes and incubated at 47°C in a humidified chamber for 16-36 hours. After incubation with primary/encoding probes, the sample was washed in 2x SSC and 40% formamide for 30 minutes and post-fixed with 4% PFA in 2x SSC for 10 minutes at room temperature. The sample was then incubated for 2-3 minutes with fiducial beads (either ThermoFisher F8805 or ThermoFisher F8792) in 2x SSC and stained with 1 μM 4',6-diamidino-2-phenylindole (DAPI; ThermoFisher D1306) in 2x SSC for 5-10 minutes, and then stored in 2x SSC until imaging. For genome-scale imaging, we followed a similar protocol except that pri-

mary/encoding probes were used at ~25 μM total concentration in the primary hybridization step, and in the incubation in pre-hybridization buffer step, Tween-20 was excluded in pre-hybridization buffer and the incubation was performed for 10 minutes instead of 30 minutes.

For whole-chromosome imaging experiments including RNA imaging, all buffers used from the point at which cells were fixed contained a 1:10-1:1,000 dilution of RNase inhibitor (either NEB M0314 or Fisher Scientific N2615). Treatment for RNA staining was identical to the above-described protocol up to treatment with HCl. After this step, cells were incubated in pre-hybridization buffer for 10 minutes, and the cell coverslip was then inverted and placed on a drop of hybridization buffer containing primary/encoding probes targeting the RNA introns at ~6 μM, as described for DNA staining. In this case, however, no 90°C heat denaturation was performed, and cells were immediately incubated at 47°C in a humidified chamber for 16-36 hours. After incubation with primary/encoding probes, the sample was washed in a formamide solution and post-fixed with PFA as described for DNA above. It was then incubated with fiducial beads and stained with 1 μM DAPI, before being stored in 2x SSC until imaging. After RNA imaging, RNA signal was removed using 50% formamide in 2x SSC before the sample was removed from the microscope, then the cells were treated with RNase A and the DNA hybridization proceeded in the same manner as described above for DNA imaging without RNA imaging. For genome-scale imaging experiments including RNA imaging, we followed a similar protocol except that the 50% formamide wash step before removing the sample from the microscope was not used.

Sequential hybridization of readout probes for sequential or combinatorial FISH imaging

All fluid exchanges in this part of the protocol were achieved via the use of a custom-built fluidics system (Chen et al., 2015), with the coverslip mounted in a FCS2 flow chamber (Bioptechs, 060319-2). The setup of this system is described in detail in the "Fluidics system configuration" section. Briefly, the fluidics system consisted of 3-5 computer-controlled eight-way valves (Hamilton, MVP and HVXM 8-5) and a computer-controlled peristaltic pump (Gilson, MINIPLUS 3). Put together, these components allow control of both the rate of fluid flow and of the type of fluid flowing at any given time.

Each round of hybridization consisted of the following general steps:

1. Flow in the hybridization buffer with a set of oligonucleotide probes specific to each round, as described below
2. Incubate for 10 minutes at room temperature.
3. Flow wash buffer
4. Incubate for ~200 s
5. Flow imaging buffer

Imaging buffer was prepared as described previously (Chen et al., 2015), and consisted of 60 mM Tris pH 8.0, 10% w/v glucose, 1% Glucose Oxidase Oxygen Scavenger Solution (containing ~100 mg/mL Glucose Oxidase (Sigma-Aldrich, G2133) and a 1:3 dilution of catalase (Sigma-Aldrich, C3155)), 0.5 mg/mL 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox; Sigma-Aldrich, 238813) and 50 μM Trolox Quinone (generated by UV irradiation of a Trolox solution) (Cordes et al., 2009; Rasnik et al., 2006). Trolox was dissolved in methanol before being added to the solution. After preparation, the imaging buffer was covered by a ~0.5 cm thick layer of mineral oil to prevent exposure to oxygen.

The hybridization buffer and wash buffer were made up of 35% and 30% formamide in 2x SSC, respectively, with the hybridization buffer also containing 0.01% v/v Triton-X. The hybridization buffer was kept separately for each hybridization round and contained two (for DNA-MERFISH) or three (for whole-chromosome imaging by sequential hybridization and for genome-scale RNA-MERFISH of nascent transcription) sets of readout probes. We introduced fluorescent signal in the following ways:

1. For whole-chromosome imaging, all experiments involved the sequential hybridization of oligonucleotide adaptors and fluorescent readout probe as described in Figure S1A. During each round of hybridization, a set of adaptor probes (each at 100 nM concentration) were first flowed in for the detection of three targeted genomic loci in three distinct color channels (Alexa750, Alexa647 or Cy5, and Cy3). Each adaptor probe comprised of a segment complementary to the readout sequence unique to one of the targeted loci and a segment containing a color-channel-specific common readout sequence (see Table S3). Next, three distinct dye-conjugated, readout probes, each complementary to one of the three color-channel-specific common readout sequences were flown in at 30 nM concentration for each probe. This procedure allows three genomic loci to be imaged, respectively in three color channels, during each round of hybridization. Fluorescent readout probes contained a disulfide bond linking the fluorophore to the oligonucleotide, as described previously (Moffitt et al., 2016), to enable efficient removal of signal between rounds by TCEP.
2. For genome-scale chromatin imaging by DNA-MERFISH, each round's hybridization buffer included two fluorescent readout probes, one labeled with Cy5 or Alexa647 and the other labeled with Alexa750. Fluorescent readout probes consisted of either: 1) a fluorescently labeled oligo complementary to a readout sequence common to all encoding probes imaged in a given bit, added at 100 nM concentration, or 2) a combination of an adaptor oligo consisting of the sequence complementary to a readout sequence, concatenated to an additional common readout sequence (common to all adaptors in each color channel), as described above, and a fluorescently labeled readout probe complementary to this common readout sequence. For some experiments, the adaptor and common readout probes were pre-mixed in a 1:1.5 ratio and added to a final concentration of ~100 nM. For other experiments the adaptor and readout probes were hybridized sequentially to the sample.

3. For RNA imaging by sequential hybridization or MERFISH, each round's hybridization buffer contained three adaptor probes, one for each color channel, as described above. Each round included two discrete hybridization steps – first the adaptors were flowed in, hybridized, and excess material was washed. Then three fluorescent readout probes, complementary to the common readout sequences on the adaptors, respectively labeled with Cy3, Cy5 (or Alexa647) and Alexa750, were flown in sequentially. After the fluorescent readouts were hybridized, imaging buffer was flowed in and signal was collected.

Before the next round of readout probe or adaptor probe hybridization, fluorescent signals from the readout probes in the current round were removed as described in the "Signal removal between rounds of hybridization" section below.

Before the first round of hybridization, we performed a round of imaging to acquire the DAPI signal and identify nuclear boundaries. For whole-chromosome imaging by sequential hybridization, we imaged 651 genomic loci on Chr21 over > 200 rounds or 935 genomic loci on Chr2 over > 300 rounds, all in 3 color channels. In each round, the genomic loci were imaged in 3D by stepping the stage in the z-dimension. Nascent RNA transcript and TSS imaging in Chr21 were performed similarly. For genome-scale chromatin imaging by combinatorial FISH (DNA-MERFISH), the entire set of 1,041 genomic loci were imaged in 50 rounds of hybridization and 2 color channels per round. The genome-scale imaging of the RNA transcripts of 1,137 genes were imaged in 3D in 18 rounds in 3 colors. Additional rounds were used to relabel subsets of genomic loci and assess chromatic aberration and bleedthrough between color channels, as well as stability of the sample and imaging instrument. Imaging of ~60 fields of view containing a total of ~1,000-3,000 cells took ~12-18 days for whole-chromosome imaging by sequential hybridization and 3 days for genome-scale chromatin imaging by MERFISH.

Our 3-5 valve system allowed loading up to 20–34 different hybridization solutions. As a result, after exhausting all the fluidic system's channels we bypassed the sample chamber and washed all the channels used for hybridization with 30% formamide in water. Next we reconnected the chamber and proceeded with the next set of hybridization and imaging rounds, as described in the "Fluidics system configuration" section above.

For whole-chromosome imaging by sequential hybridization, we performed a gentle post-fixation step with 2% PFA in 2x SSC for 5 minutes periodically (every ~4 days) to maintain the structural integrity of our sample over hundreds of rounds of hybridization.

Antibody labeling and imaging

Antibody imaging was performed immediately after RNA or DNA imaging, where the signals from the DNA or RNA have been quenched. Following completion of imaging via the protocols described above, samples underwent the following steps:

1. Samples were incubated with blocking solution (PBS with 0.1% v/v Tween-20 (Sigma-Aldrich P9416) and 1% w/v bovine serum albumin (BSA; Jackson Immunoresearch 001-000-162)) for 30 minutes.
2. Samples were incubated with primary antibody diluted in blocking solution for 1 hour.
3. Samples were washed 3 times in PBS with 0.1% Tween-20 for 5 minutes each.
4. Steps 2 and 3 were repeated for a fluorescently tagged secondary antibody.

All buffer exchanges were done on the microscope, using the microfluidic system described in the "Fluidics system setup" section below. We used the Cy5 color channel for imaging and extinguished the signal between sequential antibody labeling using photobleaching.

We used the following sets of primary and secondary antibodies:

1. For imaging nuclear speckles, we used a primary antibody against SC35 (Abcam, ab11826) – a splicing factor commonly used as a marker of nuclear speckles (Chen and Belmont, 2019) – at 1:200 dilution from stock and a donkey anti-mouse secondary antibody labeled by a Cy5 dye (Jackson Immunoresearch, 715-175-150) diluted 1:1,000 from stock concentration.
2. For imaging nucleoli, we used anti-fibrillarin antibody (Abcam, ab5821), at 1:200 dilution from stock, and a donkey anti-rabbit secondary antibody labeled by an Alexa 647 dye (Jackson Immunoresearch, 711-605-152), diluted 1:1,000 from stock concentration.
3. For cell-cycle state determination, we performed the immunofluorescence staining using the anti-geminin antibody (Abcam, ab195047), at 1:100 dilution from stock, and a donkey anti-rabbit secondary antibody labeled by an Alexa 647 dye (Jackson Immunoresearch, 711-605-152), diluted 1:1,000 from stock concentration.

Image acquisition

For each experiment, we selected approximately ~60 fields of view (FOVs) for imaging, avoiding regions where cells are sparse (we typically identified 10-50 cells per FOV). Each camera FOV consisted of either 1,000 × 1,000 pixels, with a camera pixel corresponding to 153 nm in each dimension in the imaging plane, or 2048 × 2048 pixels, with a camera pixel corresponding to 108 nm in each dimension in the imaging plane.

After each round of hybridization (see "Sequential hybridization of readout probes for sequential or combinatorial FISH imaging" above), we acquired z stack images of each FOV in 3 or 4 colors: 647 nm and 750 nm illumination (or 560 nm, 647 nm, and 750) were used to acquire FISH images, 560 nm illumination (or 405 nm illumination) was used to image fiducial beads. For the first round of imaging, we used 405 nm illumination to image the DAPI signal, while for antibody imaging, we used the 647 nm excitation channel.

Consecutive z sections were separated by 85, 100, 150 or 200 nm, covering the entirety of the nuclear volume for all imaged cells. At each z position, images were acquired in all channels before the stage was moved and images were acquired at a rate of ~10 Hz.

Signal removal between rounds of hybridization

Before each round of imaging, the signal from the previous round (or endogenous background, in the case of the first round) was extinguished. This was achieved via cleaving the disulfide bond connecting fluorophores to readout probes as described previously (Moffitt et al., 2016) followed by an optional photobleaching step. We note that when this step was performed before the start of the DNA or antibody imaging it also removed the potential residual signal from the previous imaging modality (i.e., the previous RNA signal or the previous DNA signal respectively). The buffer used for cleaving the fluorophores contained 50mM tris(2-carboxyethyl)phosphine (TCEP; Sigma-Aldrich, C4706) to reduce the disulfide bond connecting fluorophores to readout probes, as well as 1mM dye-free common readout probes in 35% formamide to block any unoccupied readout sequences from interfering with the next round of hybridization. In experiments where photobleaching was performed, it was accompanied by changing the buffer to 2x SSC with or without 35% formamide and illuminating each field of view with the maximum available power of the laser lines used for acquiring signal for 3-10 s. The photobleaching step in the sequential tracing experiments was done during the hybridization step with oligonucleotide adaptors, to minimize total experimental time. The DAPI signal was extinguished as a result of the high formamide concentration in the hybridization and wash buffers.

Relabeling of genomic loci in sequential DNA-FISH imaging

After finishing the sequential DNA-FISH imaging experiment for the chromosome tracing, we re-labeled and re-imaged a subset of regions. We treated the sample with 57% formamide in 2xSSC for 4 minutes and repeated this treatment 3 times to strip off the adaptor probes and common readout probes (whose fluorescence signal was removed after each round of imaging by first cleaving the dye from the oligonucleotide probe using TCEP and then by photobleaching during the original imaging rounds). After stripping off the adaptor and common readout probes, we added 1mM dye-free common readout probes in 35% formamide 2xSSC to block any unoccupied readout sequences on adaptor probes that were not stripped off. Next, relabeling of selected regions was achieved by following the standard readout probe hybridization protocol (described in section "Sequential hybridization of readout probes for sequential or combinatorial FISH imaging").

Sequential relabeling of genomic loci in combinatorial FISH imaging

For combinatorial imaging by DNA-MERFISH, the encoding probes for a subset of loci on Chr6 were designed such that these loci were associated with three readout sequences – two sequences for DNA-MERFISH and a third readout sequence unique to each of these loci and orthogonal to all 100 bit readout sequences used in DNA-MERFISH. To relabel these loci and image them sequentially, we performed imaging using the readout probes corresponding to these unique readout sequences (see Table S3), following the same protocol described in the "Sequential hybridization of readout probes for sequential or combinatorial FISH imaging" section.

Image Analysis

Overview of analysis pipeline

The image analysis pipeline used in this study was implemented in Python, and the code is available at https://github.com/ZhuangLab/Chromatin_Analysis_2020_cell. The overall pipeline consists of the following steps:

1. Identify and segment all imaged nuclei.
2. Fit 3D-Gaussians to all detected fluorescent spots in imaging channels used for DNA or RNA imaging, as well as for fiducial beads. DNA and RNA spots which did not overlap with identified nuclei were rejected.
3. Correct sample drift using the fiducial beads.
4. Correct chromatic effects between different color channels.
5. Assign identities to DNA loci and RNA molecules using custom algorithms and software (these are described separately below for DNA and RNA imaging, both for chromosome-wide and for genome-wide imaging).

Nucleus segmentation

DAPI images from the first round of imaging were used to segment individual nuclei using the maximum projection of the DAPI image onto the *xy* plane as input.

Spot fitting for DNA and RNA imaging

The following analysis pipeline was applied to each imaged FOV in order to obtain the three-dimensional (3D) positions of all loci of interest:

1. Fiducials were fitted in all rounds of imaging and used for image alignment (see "Drift correction" section below).
2. In the first imaging round (preceding the first round of hybridization), DAPI signal was used to identify the borders of individual nuclei, as well as for image registration between RNA and DNA imaging. See "Nucleus segmentation" section above and "Image registration between DNA and RNA imaging" section below for details.
3. Diffraction-limited spots within each identified nucleus were fitted to a 3D Gaussian function to identify their center of mass and brightness above local background (Bintu et al., 2018). To make analysis more manageable, we fixed the number of fitted spots per nucleus per image that will be retained for decoding to 125 or fewer in genome-scale imaging by DNA-MERFISH (~3-fold

greater than the number of distinct loci expected without noise). For whole-chromosome imaging by sequential hybridization, we fixed the number of fitted spots per chromosome per image to 6 or fewer.
4. The fitted spots from step 3 were then used for identifying DNA loci and transcription foci and determining their positions, as described in the corresponding sections below.

Drift correction

Fiducial bead spot fitting was performed in the same way as described above. The set of fiducial bead positions was then compared between rounds of hybridization and a rigid transformation was applied to minimize the sum of square difference of the relative position of beads. This transformation was used for drift correction between hybridization rounds.

Correction of chromatic effects

Bleedthrough and chromatic aberration between different color channels were determined and corrected for by labeling the same set of genomic loci in each imaging channel independently and comparing the signals of the same loci in the different color channels, respectively.

Image registration between DNA and RNA imaging

We first used DAPI signal for rough image-registration across the two sets of images (i.e., chromatin and RNA images) via 2D image correlation. All images within each set were aligned to the DAPI image acquired at the beginning of each set using fiducial beads. After an initial round of RNA decoding was performed (see "Identification of transcription foci from fitted RNA spots in sequential imaging" and "Decoding algorithm for fitted RNA spots in genome-scale imaging" sections below), we calculated a finer alignment by assuming that the displacement between nascent RNA localizations and their harboring DNA loci should average to zero when considered across all imaged genes and cells in a field of view. In accordance, we calculated an additional rigid transformation to minimize the mean displacement between imaged nascent RNA and their corresponding DNA loci and used this as our final alignment.

Identification of chromatin loci from fitted DNA spots in sequential imaging

Identification and 3D localization of each locus were achieved through the following steps:

1. A list was generated for the drift- and aberration-corrected locations of all fitted spots in each image. Because we allowed the spot finding algorithm to find up to 6 candidates for each chromosome in each sequential image corresponding to a specific color channel of a specific hybridization round (see "Spot fitting for DNA and RNA imaging"), we perform the following additional steps to identify the candidate spots that most likely originated from the imaged chromatin loci.
2. We generated an initial tentative chromatin trace by selecting the brightest spot corresponding to each chromosome copy in each cell in each color channel of each hybridization round.
3. For each fitted spot, regardless of whether it was selected for the initial tentative chromatin trace, we calculated three quality metrics:

i. The spot's brightness above local background
ii. The spot's distance to the local center of mass, which is calculated from five loci upstream and five loci downstream along the tentative chromosome trace
iii. The distance to the center of mass of the entire tentative chromosome trace

4. For each spot we combined the three quality metrics described above into a single measure by calculating the combined Fisher p value for every candidate spot against the distribution of quality metric values for spots included in the tentative chromatin trace (which we will term the "valid distribution"). This can be thought of as the overall quality score of each spot, and was calculated per spot in the following way: for each of the metrics we calculated the fraction of other spots in the "valid distribution" that had worse quality metric (i.e., lower brightness or greater distance) and multiplied these three fractions. We then used an expectation-maximization procedure to sequentially select the spot with the highest quality score corresponding to each targeted chromatin locus and reupdated the "valid distribution" based on the resulting chromatin trace. We repeated this optimization procedure until convergence. After convergence, the final sets of spots, each corresponding to a chromatin locus, were used to determine the 3D spatial positions of the targeted loci.
5. Finally, we filtered the spots from step 4 to remove those whose quality scores were below a set cutoff (and thus were of low confidence). To set a cutoff value for the combined quality score, we first computed the quality scores of loci that had been included in the re-imaging experiment to determine displacement error (see "Identification of re-imaged loci in sequential imaging" section, below). Then, we separately calculated the distribution of quality scores for spots with low displacement error (< 500nm) between the original and re-imaged loci and for spots with high displacement error (> 500 nm). Finally, we set the quality score threshold such that the fraction of loci in the final chromatin trace (after applying the threshold) expected to be in the high displacement error category is < 5%.

The remaining spots after step 5 were used to determine the final positions of the chromatin loci and trace the chromatin structure.

Identification of the transcription start sites (TSSs) from the fitted DNA spots in sequential imaging

Identification and 3D localization of each 5-kb TSS locus imaged were achieved following a similar procedure as outlined above for the 50-kb DNA loci with a few notable modifications:

A) In step 3 above for the quality metric number ii (the spot's distance to the local center of mass), we used the spot's distance to its harboring 50-kb DNA locus determined for each chromosome as described in the section above.
B) In step 5 above we set a quality score threshold similar to the one used for the 50-kb DNA loci to filter out the lower quality spots.

We note that the 5-kb loci are considerably dimmer than the 50-kb loci and thus have a reduced detection efficiency. Nonetheless, while the lower detection efficiency of these 5-kb loci could result in lower statistics for analysis in Figures 2H and 2I, it was compensated for by our imaging of a large number of chromosome copies.

Identification of transcription foci from fitted RNA spots in sequential imaging

Signal from RNA imaging rounds was analyzed using the following procedure:

1. The positions of fitted RNA spots for each cell was corrected for chromatic aberration and drift followed by a coarse RNA-DNA image registration using the DAPI signal and the brightest RNA spot within a distance of 1000 nm to the corresponding DNA locus was kept.
2. The registration between DNA and RNA imaging was refined based on the displacement between the initial selected RNA localizations from step 1 and the location of the DNA locus harboring them as described in the "Image registration between DNA and RNA imaging" section above.
3. The locations of all the candidate RNA spots, after the fine registration from step 2, were compared to the location of the 50kb DNA locus harboring the gene and the corresponding 5kb DNA transcription start site. At this stage, we applied a more stringent distance cutoff of 500 nm from either the 50-kb locus or the 5-kb transcription start site, as well as a signal-to-noise ratio threshold of 1. If the nascent RNA localization passed both thresholds, it was considered as a detected transcriptional burst.

We note that some transcriptional bursts could occur farther from their harboring genomic loci than the set cutoff distance. Because of this, a small fraction of transcriptionally active loci could be mis-classified as silent loci. Hence the difference observed between transcribing and silent loci could in fact be somewhat bigger than what is shown in Figures 2H and 2I.

Identification of re-imaged DNA loci in sequential imaging and estimate of displacement error

Identification of re-imaged loci was performed similarly to the description in the "Identification of chromatin loci from fitted DNA spots in sequential imaging" section, except that the re-imaged loci were used to replace the corresponding subset of loci in the original imaging rounds while also keeping the other loci that were not re-imaged in the analysis.

For computing the displacement error between the original and re-imaging rounds (Figures S1B–S1D), we considered only loci that passed a brightness threshold, based on the following observation. We noticed that the fluorescent signals in the re-imaging rounds were substantially dimmer than in the original imaging rounds. This is potentially due to the incomplete removal of the original adaptor/readout probes and/or due to the partial removal of the primary probes bound to the genomic DNA during the formamide treatment to strip off the adaptor/readout probes bound in the original imaging rounds. The lower brightness reduced the localization accuracy of the re-imaged loci and caused an artificial overestimation of the localization error of the initial imaging rounds. To mitigate this effect, in estimating the displacement error between the original and re-imaging rounds, we selected only those re-imaged loci that were > 20% in brightness compared to the signal in the original imaging round.

As an additional note, we also noticed that, when examining loci with > 1000nm from both neighboring loci, a fraction (~20%) of these exhibit a large re-imaging displacement error (> 500 nm) and have lower brightness. We thus caution that loci that are far away from both of their genomic neighbors may have a relatively low confidence.

Decoding algorithm for fitted DNA spots in genome-scale imaging by DNA-MERFISH

Identification and 3D localization of each locus were achieved through the following steps:

1. A list was generated for the drift- and aberration-corrected locations of all identified spots in each bit-image (corresponding a specific color channel in a specific round of imaging).
2. For each detected spot in every bit-image, we found all spots from other bit-images that were within a set cutoff distance from its location. We retained all such pairs of spots for further analysis, whether or not the barcode produced by a spot pair (based on which round and color channel they appeared in) corresponded to a valid barcode (i.e., a barcode that was assigned to a genomic locus).
3. For each pair of spots, we calculated three quality metrics:

i. The displacement between the 3D localizations of the two spots
ii. The difference in brightness between the two spots
iii. The mean brightness of the two spots

The brightness of each spot was normalized by the median brightness of all spots in the corresponding bit-image.

4. Spot pairs were then separated into two groups, based upon whether they correspond to a valid barcode (and hence potentially to a genomic locus) or not. Within each group, the distributions of the quality metrics were calculated. For convenience, we will refer to the distribution of spot-pair quality metrics from the invalid barcodes as the "invalid distribution" and from all valid barcodes as the "valid distribution."

5. For each spot pair we combined these quality metrics in step 3 into a single measure by calculating the combined Fisher p value for every candidate spot pair against the ''valid distributions'' in step 4. This can be thought of as the overall quality score of each spot-pair, and was calculated per pair in the following way: for each of the metrics we calculated the fraction of other spot-pairs in the ''valid distribution'' that had worse quality metric and multiplied these fractions for all metrics. We then used an expectation-maximization procedure to sequentially select the two spot-pairs with the highest quality score corresponding to each targeted chromatin locus and reupdated the ''valid distribution,'' and repeated this optimization procedure until convergence. After convergence, the final sets of spot pairs, each corresponding to a chromatin locus, were used to determine the 3D spatial positions of the loci.
6. After step 5, we used a modified K-means algorithm (Lloyd, 1982) to separate the chromatin loci belonging to the same chromosome into two homologs. As opposed to the standard K-means clustering algorithm which splits points into two groups and minimizes the radius of gyration within each group, we progressively switch points between the groups to first maximize the fraction of assigned points in each homolog and then minimize the radius of gyration of each homolog.
7. After separating the two homologs we calculated their center of mass and the distance of each spot-pair from step 2 to their parent chromosome's center of mass. We added the distance to the chromosome center as another quality metric in addition to the 3 metrics considered in step 3 and repeated steps 3-6.
8. Finally, we filtered the spot pairs from step 7 to remove the pairs whose quality scores remained similar to the ''invalid distribution.''

The remaining spot pairs after step 8 were used to determine the final positions of the chromatin loci and trace the chromatin structure.

Decoding algorithm for fitted RNA spots in genome-scale imaging

Signal from RNA imaging rounds was decoded using the following procedure:

1. A list was generated for the drift- and aberration-corrected locations of all identified spots in each bit-image (corresponding a specific color channel in a specific round of imaging).
2. For each detected spot in every bit-image, we found all spots from other bit-images that were within a set cutoff distance from its location and retained these spot pairs as candidate RNA bursts if they formed a valid barcode.
3. The location of each of these candidate RNA bursts was then compared to the location of the DNA locus harboring the relevant gene, after chromatic aberration and drift correction and a coarse RNA-DNA image registration based on DAPI signal, and the candidate was kept if it was within a set threshold distance to the DNA locus.
4. The registration between DNA and RNA imaging was refined based on the displacement between the initial decoded RNA localizations (from step 3) and the location of the DNA locus harboring them as described in the ''Image registration between DNA and RNA imaging'' section above.
5. Locations of all candidate RNA bursts were compared again to the locations of the DNA loci harboring the corresponding genes, this time with the refined image registration. If the nascent RNA localization was within a cutoff distance from its harboring DNA locus at this stage, it was considered as a detected transcriptional burst.

Identification of re-imaged DNA loci in genome-scale imaging

In the genome-scale imaging by DNA-MERFISH, a subset of the targeted genomic regions on Chr6 were assigned probes such that they could be re-imaged individually using sequential two-color FISH, after the combinatorial, MERFISH imaging. For each decoded instance of one of these loci in the combinatorial, MERFISH imaging, we then estimated the displacement error as the distance between the localization determined in combinatorial, MERFISH imaging and the nearest spot in the sequential re-imaging rounds.

Identification of nuclear bodies from immunofluorescence imaging

The location of nuclear bodies (nuclear speckles and nucleoli) was extracted from immunofluorescence signals by applying a threshold to the intensity of the immunofluorescence signals, resulting in a pixelated mask identifying high immunofluorescence signals. This was then treated as a pixelated set of locations ''containing'' nuclear bodies.

Cell-cycle stage determination from anti-geminin and DAPI images

First, cells undergoing mitosis were eliminated by visual inspection and not considered for our analysis. Next, using a combination of geminin immunofluorescence signal and nuclear signal measured using DAPI, cells were classified as G1 (low geminin signal, low DAPI signal) or G2/S (high geminin signal) similarly to our previous study (Bintu et al., 2018).

QUANTIFICATION AND STATISTICAL ANALYSIS

Estimation of the Nuclear Lamina Location

The position of the nuclear lamina was estimated by generating the minimal 3D convex hull surface (using Python's SciPy package) surrounding the locations of all decoded chromatin loci in a given cell.

Spatial Distance between a Locus Pair or between Locus and Nuclear Body

The spatial distance between any pair of loci was simply calculated as the Euclidean distance between their fitted 3D Gaussian centers. In the case of distance to nuclear bodies, we calculated the minimal Euclidean distance to all identified nuclear body "locations" or the minimal distance to the surface of the convex hull defining the nuclear lamina.

Proximity Frequency Matrices from Imaging

To calculate the proximity frequency between any given pair of loci, we first counted the number of measured distances between that locus-pair that was smaller than a set cut-off distance (500 nm in this study, unless otherwise mentioned). This number was then divided by the total number of distances measured for that pair of loci. The cut-off distance was determined by assessing the Pearson correlation between the proximity frequency matrices resulting from a range of cut-off thresholds and the Hi-C contact matrix (Rao et al., 2014). We noticed that the Pearson correlation coefficient between the proximity frequency map and Hi-C map remained high, in the range of 0.82 – 0.87, for cut-off distances between 200 nm to 800 nm and reaches a maximum at cut-off distances of ~400-600 nm. Hence, we selected a cut-off distance of 500 nm for all analyses.

A/B compartment calling

To call A/B compartments from the imaging data of Chr21 and Chr2, we first generated the normalized proximity frequency or spatial distance matrices by dividing the observed proximity frequency or spatial distance over the expected value given the genomic distance between a given pair of loci. The expected proximity frequency or spatial distance was derived from fitting the ensemble data against the genomic distance by linear regression in log scale. Then a Gaussian filter was applied to the normalized matrix to mitigate influences from regions with low quality. Next, Pearson correlation matrices are calculated from the normalized proximity frequency or spatial distance matrix and subject to principal component analysis. Two candidate compartment identities were called corresponding to the positive/negative value of the first principal component (PC1). After this initial determination of compartments, we removed very small compartment fragments (fragments < 6 regions in Chr21, < 3 regions in Chr2) to avoid over splitting compartments. Finally, the A and B identities of the two candidate compartments were assigned based on epigenetic markers (Dunham et al., 2012). For Chr2 data, p and q arms were independently analyzed.

The calling of A and B compartments for the genome-scale Hi-C data or proximity frequency data derived from imaging follows a similar algorithm but without the Gaussian smoothing step, and p and q arms were independently analyzed except for chromosomes with relatively small p arms.

Local A/B Density Score

To calculate the local density of compartment-A and compartment-B loci at each decoded location, we first calculated the spatial distances between each pair of chromatin loci for each cell. For each locus, we then computed the local A/B density ratio in the following way:

1. First, we placed a Gaussian probability density function centered around each A or B locus with a standard deviation of 100 nm (for Chr21 imaging), 125 nm (for Chr2 imaging), or 500 nm (for genomic-scale imaging).
2. The total A density score at the locus was then computed as the sum of this Gaussian probability density function values from all A loci excluding itself in whole-chromosome imaging. For genome-scale imaging, the total *trans* A density score at the locus were summed from all *trans*-chromosomal A loci (i.e., all A loci from other chromosomes). The total B density score was computed in an analogous way.
3. To determine the A/B density ratio at the locus, we divided the total density score of compartment-A loci by the total density score of compartment-B loci at the locus. The *trans* A/B density ratio was computed analogously.

Insulation Score from Imaging Data

Insulation score has been previously defined for ensemble Hi-C (Crane et al., 2015). We used an analogous definition for our imaging results and applied it to compute insulation scores of neighboring or non-neighboring domains in individual chromosomes in single cells.

To calculate the insulation score between two domains, we first calculated an intra-domain distance distribution by considering all spatial distances between each pair of loci within the first domain and all distances between each pair of loci within the second domain. We then calculated an inter-domain distance distribution by considering all distances between pairs of loci that reside in different domains. We then defined the insulation score as the median of all inter-domain distances divided by the median of intra-domain distances. Two highly intermixed domains would have insulation score close to 1, while domains that are just contacting will have an insulation score of ~2.

TAD Calling and Single-Cell Domain Calling

To enable comparison of TAD calling in imaging and Hi-C data, we slightly modified the insulation score definition above such that insulation scores from these data would fit into the same dynamic range. For TAD-calling in median pairwise distance

matrices from imaging data, for each genomic locus we selected upstream and downstream loci with a fixed window (i.e., a fixed number of loci on each side of selected locus). We treated these two chromatin regions, up- and down-stream of the selected locus, as two ''domains'' and computed insulation score, as described above. We then defined the normalized insulation score as the difference between the median of inter-region distances and the median of intra-region distances normalized by the sum of these two median values. Therefore, a normalized insulation score will always be between 0 and 1. With this definition of normalized insulation score, we applied a sliding window along the chromosome to calculate a vector of insulation scores corresponding to genomic loci, and found local maxima by standard peak-calling algorithm from Scipy and considered these positions as our TAD boundaries. TAD calling based on Hi-C data or on proximity frequency matrices derived from imaging were performed similarly.

To call chromatin domains in single chromosomes, for each genomic locus, we selected upstream and downstream loci within a fixed window size and calculated the insulation score between the upstream and downstream regions as described in the "Insulation Score from Imaging Data" section above. With this sliding window calculation, we identified loci exhibiting local maxima of insulation scores and considered these loci candidate domain boundaries. To minimize aberrant domain calling due to noise, we removed the boundaries separating domains exhibiting similar proximity patterns with the rest of chromosome. To do this, for each pair of neighboring candidate domains, we calculated their proximity scores to the other loci outside the domains, where the proximity score of a domain to a locus outside the domain represents the fraction of loci within the domain that are within 500nm of the locus outside the domain. We then merged neighboring candidate domains if the Pearson correlation coefficient between their proximity scores to the rest of the chromosome is above a certain threshold and iterated this process.

A/B Segregation Score within Chromosomes

A/B segregation score quantifies the level of spatial separation between A and B loci within a chromosome. To calculate this quantity, we first operationally defined the A-dense volume within each chromosome as the 3D space that contains all A loci with A density scores being in the top 2/3 range. The B-dense volume was operationally defined in an analogous fashion. We then defined a purity metric of the A- and B-dense volumes as the fraction of all loci within these volumes being A and B loci, respectively. Finally, we defined the A/B segregation score as the mean of purity values of the A-dense and the B-dense volumes. This segregation score would be 1 if A and B loci are entirely segregated, and a chromosome with A and B loci completely intermixed would have segregation score approximately 0.5.

Estimation of Detection Efficiency of Nascent RNAs in Genome-scale Imaging

Estimation of the detection efficiency of transcriptional burst events was performed in the following way:

1. We considered all targeted genomic loci that harbor a gene whose RNA introns were imaged. For any of these genomic loci, their corresponding RNA signal should appear in two pre-defined bits if the gene is transcribed. Knowing the rate with which each of these two bits is not detected (p) allows us to estimate the detection efficiency of the RNA. We thus identified the set of genomic loci that colocalized (within ~150nm) with RNA signal in at least one of the two expected bits of their corresponding genes.
2. From the total set of chromatin loci identified in step 1, we determined the fraction (f) of loci that colocalized with RNA signal from exactly one of its gene's corresponding bits (and not with both bits). From the measured f (8.4%), which should be equal to $2p(1-p)/(1-p^2)$, we estimated p (4.4%).
3. The overall detection efficiency for detecting a colocalized signal in both bits was calculated using the equation: $\eta = (1-p)^2$, and was found to be ~92%.

We note that this estimate assumes that the detection efficiency in all rounds of imaging and for all genes is identical. In reality, there might be variation in detection efficiency per gene. Furthermore, some bursts may occur farther from their harboring genomic loci than the colocalization cutoff used to detect them. Thus, the detection efficiency per gene could be lower than the reported value above, leading to a similar effect on the difference between transcribing and silent loci as described in the "Identification of transcription foci from fitted RNA spots in sequential imaging" section above, but for Figures 7B and 7C.

Hi-C Data Analysis

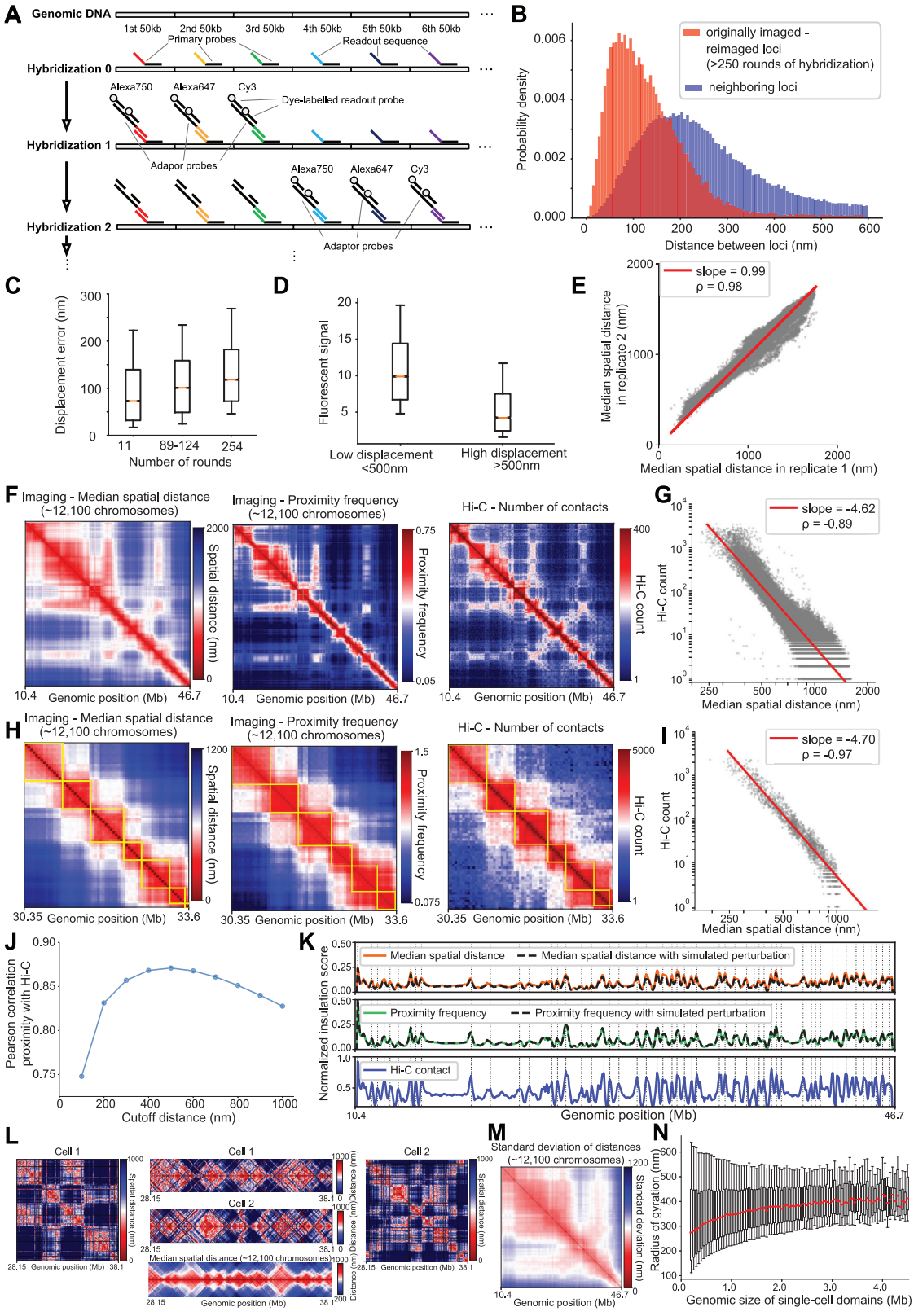
Hi-C data for IMR-90 cells was procured from Rao et al. (2014) and loaded using straw (Durand et al., 2016). For identification of A/B compartments in individual chromosomes, we followed established published protocols (Lieberman-Aiden et al., 2009). For identification of TADs, we used the method described in the "TAD calling and single-cell domain calling" section. For comparison of proximity frequencies derived from our imaging data to Hi-C number of contacts, we created bins centered around the regions we targeted and procured Hi-C data for these bins by summing the number of reads in higher resolution Hi-C data. The bin size used for the high-resolution Chr21 and Chr2 coordinates was 50 kb (same as the size of the targeted loci in imaging). The bin size used for the genomic-scale coordinates is 500 kb. We used a larger bin size in this case since the detected Hi-C contacts are relatively sparse for long-range interactions.

CTCF and Rad21 ChIP-seq Data Analysis

CTCF and Rad21 ChIP-seq data were downloaded from ENCODE dataset (Dunham et al., 2012) and converted to wig format by UCSC Genome Browser Utilities (Kent et al., 2010). Read counts for our targeted genomic segment were collected and normalized by input correspondingly. Local maxima of CTCF or Rad21 ChIP-seq signal enrichment over input along the chromosome were called by standard peak calling algorithm from Scipy.

CellPress

Supplemental Figures



(legend on next page)

Figure S1. High-Resolution Whole-Chromosome Tracing by Sequential Hybridization and Ensemble Statistics of Chr21 Structural Features in Comparison with Hi-C, Related to Figure 1

(A) Labeling and imaging scheme for sequential hybridization with adaptor probes. First, the sample is hybridized with primary probes, each containing a target sequence that allows specific binding to a targeted genomic locus and a readout sequence. Each locus is labeled by a total of 350-500 primary probes, but only one is shown. Each targeted genomic locus is assigned a unique readout sequence (shown in various colors), common to all primary probes that bind to the locus. The readout sequences are then detected using sequential rounds of hybridization. During each round of hybridization, the readout sequences corresponding to the targeted loci (one for each of the three color channels, Alexa750, Alexa647, and Cy3) are hybridized to oligonucleotide adaptor probes that each consist of two portions: a segment complementary to the locus-specific readout sequence (shown in color) and a segment containing a color-channel-specific common readout sequence (shown in black). Each color channel contains a unique common readout sequence that is shared by all the adaptors visualized in the same color channel. The common readout sequences are then hybridized to dye-conjugated, readout probes in corresponding color channels. This procedure allows three genomic loci to be imaged in three color channels during each round of hybridization. Following each round of imaging, fluorescent dyes, which are connected to the readout probes by a di-sulfide bond, are cleaved from the common readout probes by TCEP, and any unoccupied readout sequences on the adaptors are blocked with unlabeled common readout probes to prevent crosstalk between rounds of hybridization. The process is iterated over hundreds of rounds until the detection of all readout sequences, and hence all targeted genomic loci, is completed.

(B) Displacement of loci over the course of a single experiment. Consecutive 50-kb segments within a 900-kb region in Chr21 (chr21:32.45-33.35Mb) were imaged both at the beginning and at the end of the experiment, separated by > 250 rounds of hybridization. The distribution of the displacement between the re-imaged spot and its originally imaged counterpart was shown in red. For comparison, the distribution of the distances between adjacent 50-kb segments in the same 900-kb region measured in the original imaging rounds are shown in blue.

(C) Boxplots of the displacement of the measured locations of chromatin loci between the original and re-imaging rounds separated by different number of hybridization rounds. The medians (center lines), 25th – 75th percentiles (boxes) and 10th – 90th percentiles (whiskers) are shown.

(D) Boxplots of the fluorescent signal intensity of chromatin loci with low (< 500 nm) and high (> 500 nm) displacements errors between original and re-imaging experiments separated by > 250 rounds of hybridization. The medians (center lines), 25th – 75th percentiles (boxes) and 10th – 90th percentiles (whiskers) are shown.

(E) Comparison of median inter-loci spatial distances between two replicate experiments. Median spatial distances between pairs of imaged chromatin loci were calculated separately for two biological replicate experiments of Chr21 and plotted against each other. The Pearson correlation coefficient between data measured in the two replicates is $\rho = 0.98$. Red line represents linear regression of the data (slope = 0.99).

(F) Comparison of median spatial distance matrix derived from imaging (left), proximity frequency matrix derived from imaging (middle), and ensemble Hi-C contact matrix (right) (Rao et al., 2014) for Chr21. For the imaging data, two chromatin loci are considered to be in proximity when the spatial distance between the two loci is smaller than a cutoff distance of 500 nm. The Hi-C contact matrix is binned at 50 kb and centered around our target regions.

(G) Log-log scatterplot of number of contacts derived from Hi-C and median pairwise spatial distances derived from imaging (gray) for individual pairs of chromatin loci. The Pearson correlation coefficient between imaging and Hi-C data is $\rho = -0.89$. Red line represents linear regression of the data on the log-log plot (slope = -4.62).

(H) Same as (F) but for a 3-Mb region in Chr21 (chr21: 30.35-33.6 Mb). TAD boundaries are marked with yellow lines.

(I) Same as (G), but for the 3-Mb region shown in (H). $\rho = -0.97$; slope = -4.70.

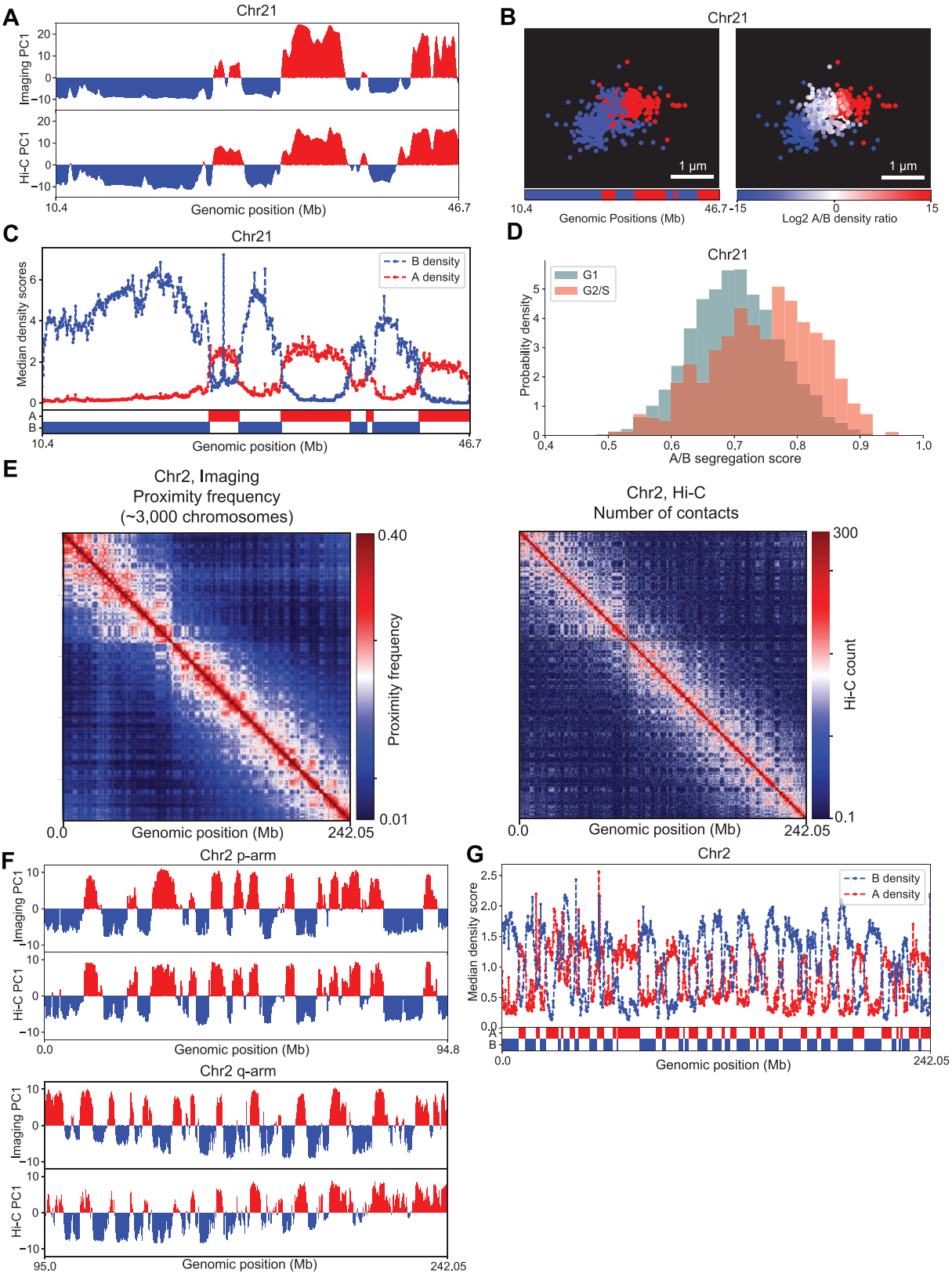
(J) Pearson correlation between Hi-C contact map and imaging-derived proximity frequency maps generated with various cutoff distances. To generate a proximity frequency map, a cutoff distance is chosen and two loci with a distance smaller this cutoff value are considered to be in proximity. Then, proximity frequency between a pair of loci was calculated as the number of incidences in which the measured distance between the loci is shorter than the cutoff distance divided by the total number of measured distances between the two loci. A log-log scattered plot of Hi-C contact number versus proximity frequency for all locus pairs were generated for each cut-off distance and the Pearson correlation coefficients were determined from these plots.

(K) Normalized insulation scores as a function of the genomic position on Chr21 calculated from: 1) the median pairwise distances from imaging (orange line, top panel), 2) proximity frequencies from imaging (green line, middle panel), and 3) Hi-C reads (bottom panel). TAD boundaries are defined as local maxima of the normalized insulation score along the chromosome (see STAR Methods). The vertical dotted lines are the ensemble TAD boundaries called from the Hi-C data. Also shown in the top and middle panels are the normalized insulation scores calculated from median distances (black line, top panel) and proximity frequencies (black dashed line, middle panel) after perturbing the loci positions with a 3D Gaussian noise term with a standard deviation of 100nm, comparable to our estimated localization measurement error.

(L) Chromatin domains in a 10-Mb region of Chr21 (chr21: 28.15-38.1Mb) in two example chromosome copies. Pairwise distance matrices of this region from two individual copies of Chr21 in single cells are shown in the left and right panels. The diagonal regions of these matrices are shown in the top two middle panels together with median pairwise distances derived from all imaged chromosome copies (bottom middle panel, reproduced from Figure 1C).

(M) Standard deviation matrix of the inter-loci spatial distances for Chr21. For each pair of regions, the standard deviation of the distances between the corresponding pair of loci across all single chromosome copies is shown.

(N) Boxplots of the physical sizes of chromatin domains of different genomic sizes in Chr21, measured by the radii of gyration. For each genomic size, the median (center lines), 25th – 75th percentiles (boxes) and 10th – 90th percentiles (whiskers) of the domain physical sizes across imaged chromosome copies are shown.



(legend on next page)

Figure S2. Ensemble A/B Compartment analyses for Chr21 and Chr2, Related to Figure 2

(A) Compartment calling based on principal component analysis for Chr21. First principal component (PC1) calculated for the Pearson correlation matrices from genomic-distance-normalized proximity frequency derived from imaging (top) and ensemble Hi-C (bottom) experiments are shown, with PC1 value > 0 (shown in red) corresponding to compartment A and PC1 value < 0 (shown in blue) corresponding to compartment B. To ensure a robust calling, the boundaries of very small compartments (with less than 6 loci for Chr21 and less than 3 loci for Chr2) were removed.

(B) 3D rendering of compartment-A loci (A loci) and compartment-B loci (B loci), and A/B density ratios in a single copy of Chr21. Left: A (red) and B loci (blue) of a representative copy of Chr21. A (red) and B (blue) compartment calling from the ensemble proximity frequency map derived from imaging is shown at the bottom bar. Right: the same chromosome but with each locus colored by its local A/B density ratio.

(C) Mean A and B density scores for each imaged locus in Chr21 averaged across all imaged chromosomal copies (see STAR Methods for details on how to calculate A and B density score at each locus). The bottom panel represents the A (red) or B (blue) compartment calling of each locus from the proximity frequency map derived from imaging.

(D) Distributions of A/B segregation scores for individual copies of Chr21 in cells in the G1 (green) and G2/S (red) phases of the cell cycle.

(E) Proximity frequency matrix derived from imaging (left) and Hi-C contact matrix (right) (Rao et al., 2014) for Chr2. The Hi-C contact matrix is binned at 50 kb, but only the contacts from the imaged segments (selected at 250-kb intervals) are shown.

(F) Same principal component analyses as in (A), but for the p-arm (top) and q-arm (bottom) of Chr2.

(G) Same mean A and B density analysis as in (C), but for Chr2.

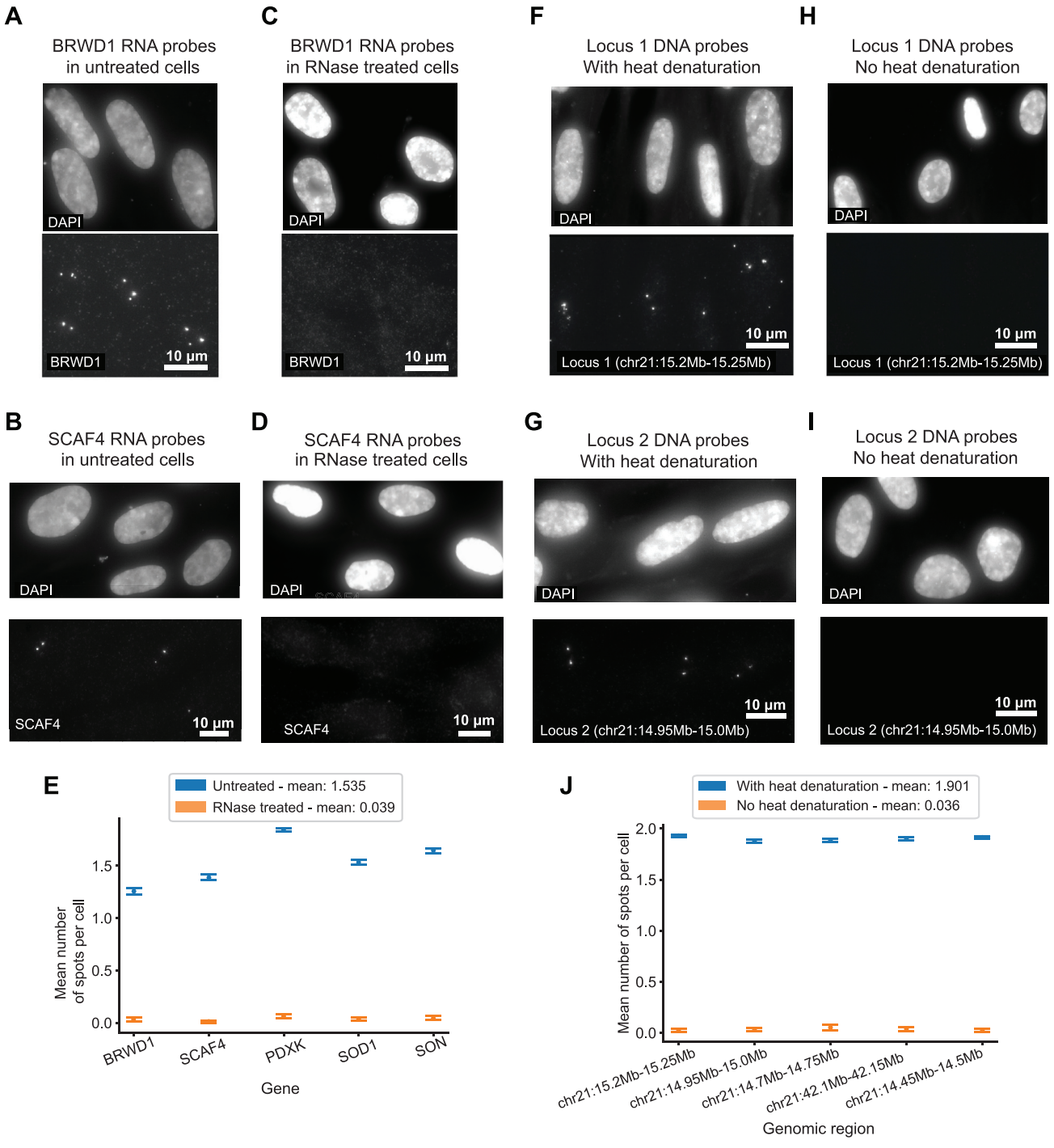


Figure S3. Measurements for RNA and DNA FISH Probe Crosstalk, Related to Figures 2 and 6

(A) Example cells with their nucleus marked by DAPI (top) and the fluorescent signal of the FISH probes targeting the nascent RNA of a gene (BRWD1) (bottom).

(B). Same as (A) but for a different gene (SCAF4). The staining of (A) and (B) follows the RNA FISH protocol described in the "Cell culture preparation and primary/encoding probe hybridization" section in STAR Methods.

(C)-(D) Same as (A) and (B), respectively, except that the RNA FISH protocol was modified to include an additional RNase treatment step to remove the cellular RNAs prior to addition of the RNA FISH probes. The cells in (C) and (D) were imaged under similar illumination conditions to (A) and (B) and their fluorescent signal is displayed with the same brightness range as in (A) and (B).

(E) The number of spots per cell with a signal-to-noise ratio > 3 for untreated (blue) and RNase-treated cells (orange) across 5 measured genes. Error bars indicate 95% confidence intervals derived by resampling.

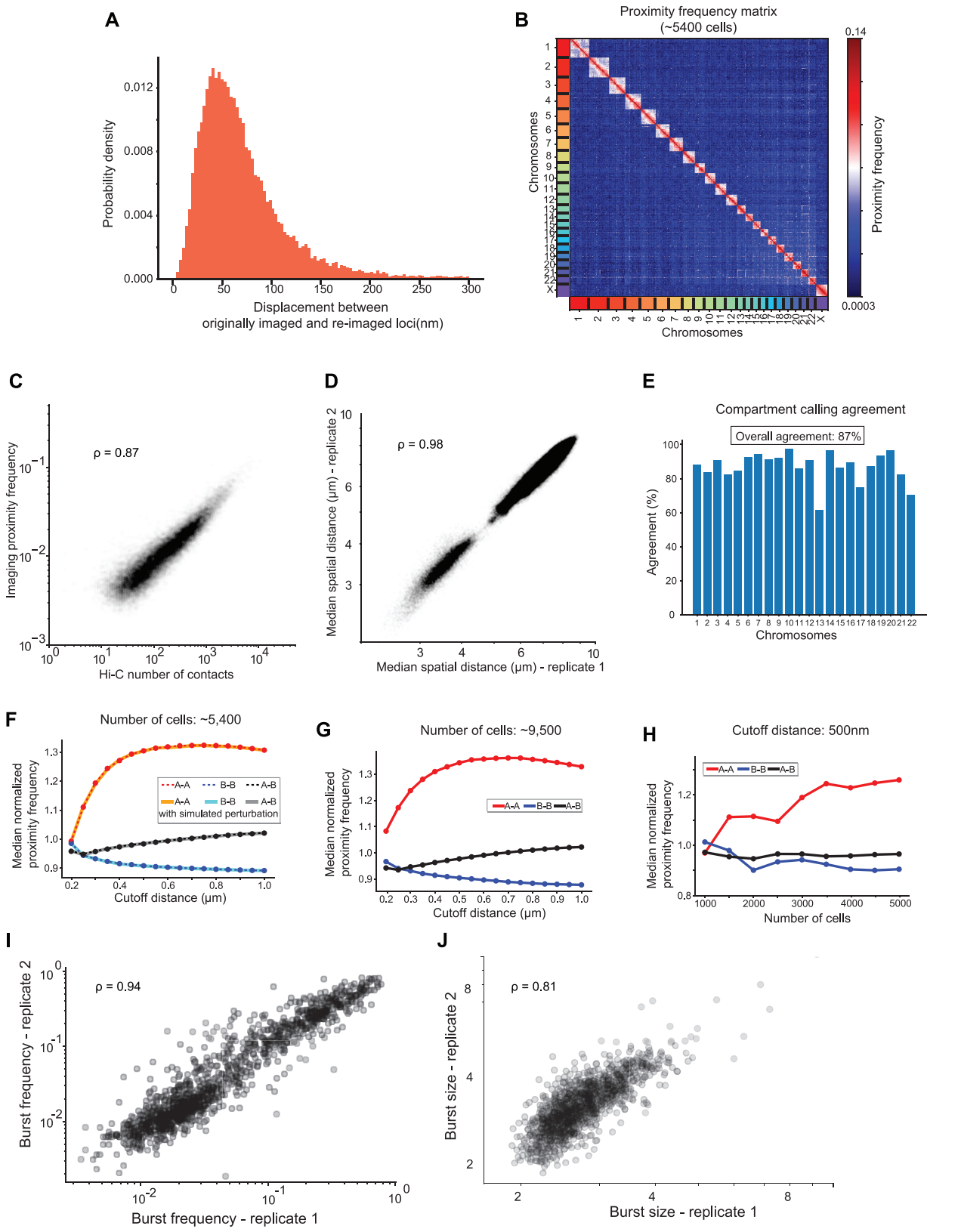
(F) Example cells with their nucleus marked by DAPI (top) and the fluorescent signal of the probes targeting a genomic locus (chr21:15.2Mb-15.25Mb) (bottom).

(legend continued on next page)

(G). Same as (F), but for a different locus (chr21:14.95Mb-15Mb). The staining of (F) and (G) follows the DNA FISH protocol described in the "Cell culture preparation and primary/encoding probe hybridization" section in STAR Methods.

(H)-(I) Same as (F) and (G), respectively, except that the DNA FISH protocol was modified to omit the heat denaturation step and hence to remove the accessible genomic DNA sites. The cells in (H) and (I) were imaged under similar illumination conditions to (F) and (G) and their fluorescent signal is displayed with the same brightness range as in (F) and (G).

(J) The number of spots per cell with a signal-to-noise ratio > 3 for the cells treated with the heat denaturation step (blue) and the cells for which this step was omitted (orange). Error bars indicate 95% confidence intervals derived by resampling.



(legend on next page)

Figure S4. Genome-Scale Imaging by DNA-MERFISH: Localization Error, Reproducibility, and Comparison with Hi-C, Related to Figures 4, 5, and 6

(A) Distribution of the displacement between the localizations of genomic loci measured in the combinatorial, DNA-MERFISH imaging run and those of the same loci re-imaged individually using sequential hybridization after completing the combinatorial imaging. 10 genomic regions in Chr6 were re-imaged across ~2000 cells. The median displacement is ~60 nm.

(B) Proximity frequency matrix for all 1,041 genomic loci imaged by DNA-MERFISH. The proximity frequency between a pair of loci was calculated as the number of incidences in which the measured distance between the loci is smaller than a cutoff distance of 500 nm divided by the total number of measured distances between the two loci.

(C) Correlation plot for the proximity frequencies between pairs of loci within chromosomes derived from our imaging data and the number of contacts derived from Hi-C experiments (Rao et al., 2014), binned at 500 kb and centered around our target loci. The Pearson correlation coefficient is 0.87. The available Hi-C data for IMR90 cells is sparse for trans-chromosomal contacts, precluding a reliable comparison of trans-chromosomal interactions between our imaging data and the Hi-C data.

(D) The correlation of pairwise distances between chromatin loci observed in two independent biological replicates of the genome-scale imaging experiments. The Pearson correlation coefficient between replicates is 0.98. The upper right cloud represents the *trans*-chromosomal pairwise distances and the lower-left cloud represents the intra-chromosomal pairwise distances.

(E) The percentage of loci whose A/B compartment assignment agree between our genome-scale imaging data and Hi-C data (Rao et al., 2014) for each human autosome. Overall, ~87% of all the imaged loci showed agreement in A/B assignment between our imaging data and the Hi-C data.

(F) The median normalized *trans*-chromosomal A-A (red), A-B (black) and B-B (blue) proximity frequencies (defined as in Figures 5A and 5B) as a function of the cutoff distance used to evaluate proximity. The normalized proximity frequencies are calculated from ~5,400 IMR-90 cells. Also shown are the median normalized trans-chromosomal A-A (orange), A-B (gray) and B-B (cyan) proximity frequencies after perturbing the loci positions with a 3D Gaussian noise term with a standard deviation of 60nm, comparable to our estimated localization measurement error, as shown in (A).

(G) Same as (F: red, black and blue), but when additional data from α-amanitin treated cells is pooled with the untreated cells (for a total of ~9,500 cells). α-amanitin treated cells showed similar enrichment of *trans*-chromosomal A-A over A-B and B-B proximity frequencies, as untreated cells (Figures S6A and S6B). This pooling result suggests that the lower enrichment for A-A interactions observed at lower cutoffs distance in (F) is likely a result of poorer statistics because for any pair of loci, a lower cutoff distance will result in a lower number of cells in which the loci are within the cutoff distance of each other.

(H) The median normalized *trans*-chromosomal A-A (red), A-B (black) and B-B (blue) proximity frequencies, as a function of the number of cells included in the analysis. Cells were subsampled randomly from the ~5,400 untreated IMR-90 cells imaged to the specified cell numbers and the proximity cutoff distance was fixed to 500 nm.

(I-J) Correlation for each gene's burst frequency (I) and burst size (J) between two biological replicates. Pearson correlation coefficients are 0.94 and 0.81, respectively.

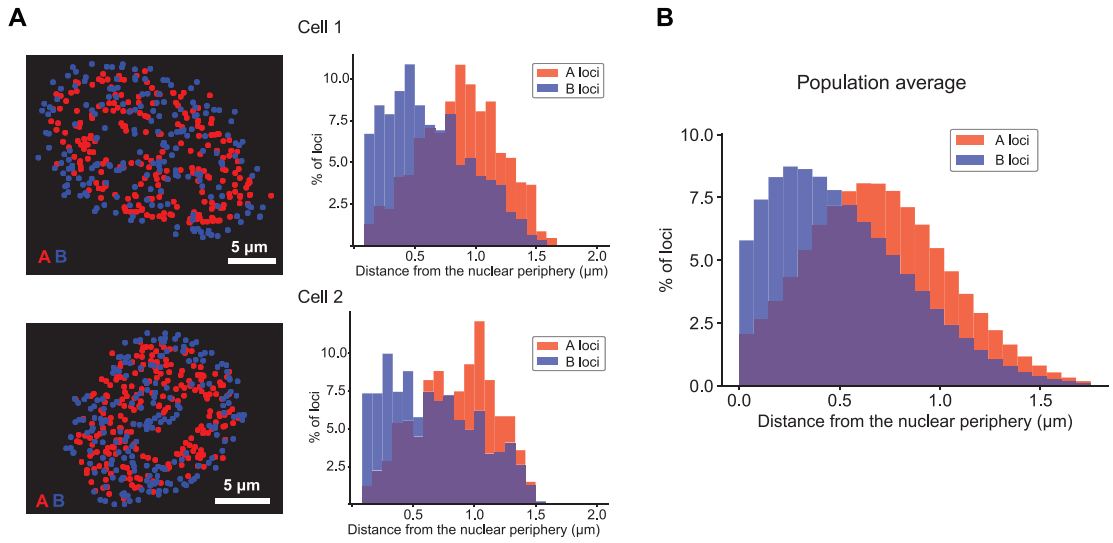


Figure S5. Compartment A and Compartment B loci Display Distinct Spatial Distributions in the Nucleus, Related to Figure 5

(A) The left panels show example images displaying A loci (red) and B loci (blue) in a single z-plane of single cells. The right panel shows the distribution of distances to the nuclear periphery for A loci (red) and B loci (blue) in these single cells. The nuclear periphery is identified as a convex hull surrounding all detected chromatin loci.

(B) Population-averaged distributions of the distance to nuclear periphery for all individual A loci (red) and B loci (blue).

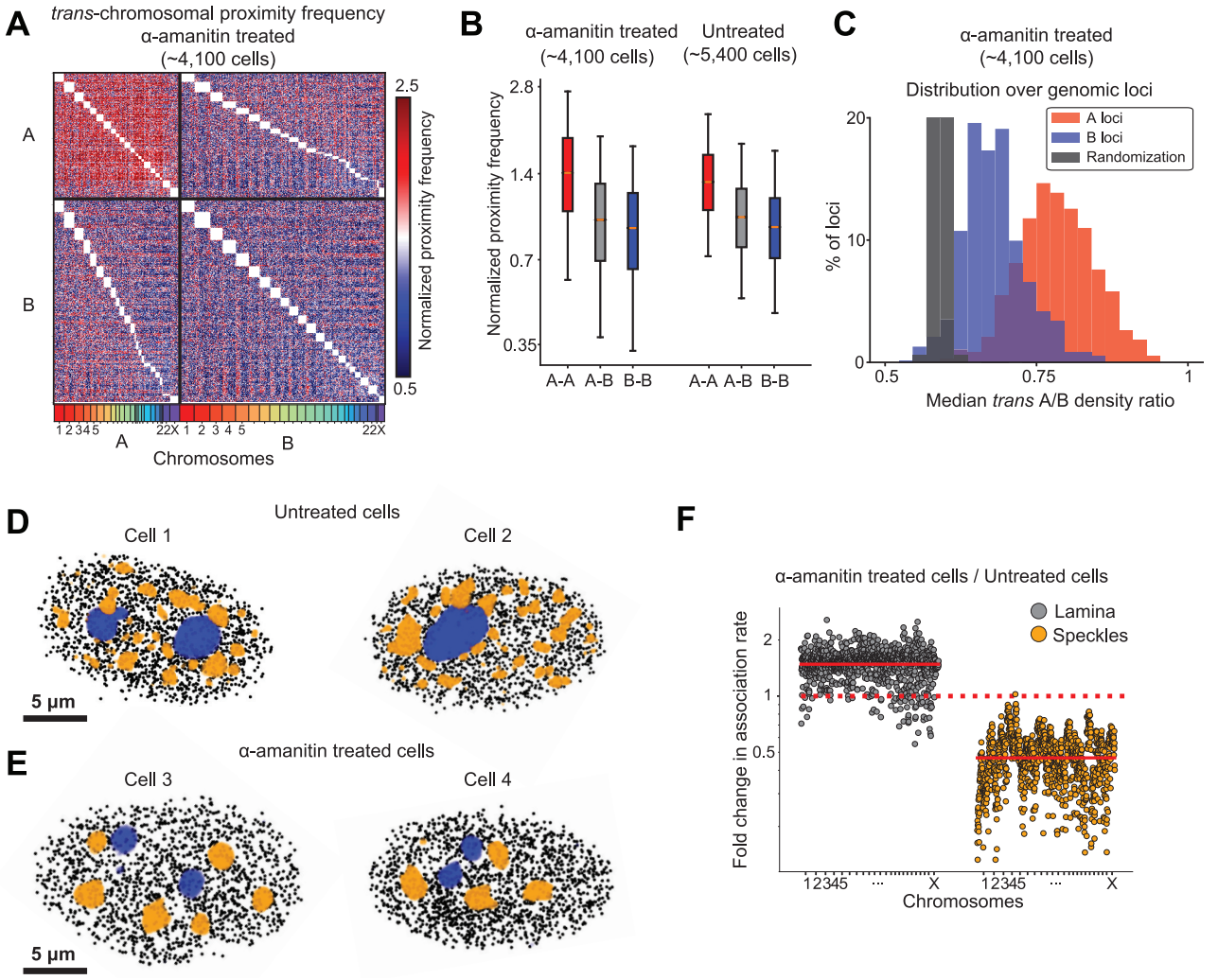


Figure S6. Effect of α-Amanitin Treatment on the *trans*-Chromosome Chromatin Interactions and the Nuclear Body Association Rates of Chromatin Loci, Related to Figures 5 and 6

(A) Normalized *trans*-chromosomal proximity frequency matrix, as in Figure 5A, but for cells treated with α-amanitin to inhibit Pol-II transcription.

(B) Distribution of normalized *trans*-chromosomal A-A, B-B and A-B proximity frequencies shown as boxplots, as in Figure 5B, but for cells treated with α-amanitin. For comparison, we also reproduced the normalized *trans*-chromosomal A-A, B-B and A-B proximity frequencies for untreated cells from Figure 5B.

(C) Distributions of the local *trans* A/B density ratio across imaged A (red) and B (blue) loci, as in Figure 5E, but for cells treated with α-amanitin. The dark gray histogram represents a randomization control where the A and B compartment identity is randomly shuffled, while keeping the total number of A loci and the total number of B loci unchanged.

(D-E) Representative images of individual nuclei with imaged chromatin loci (black spots), nucleoli (blue) and nuclear speckles (yellow) shown for untreated cells (D) and cells treated with α-amanitin (E).

(F) Fold change in the rate of associate of each locus with lamina (gray, left) and nuclear speckles (yellow, right) upon α-amanitin treatment. The data points for each genomic locus are shown in circles, the solid red lines are the median fold changes of all loci in each case, and the dotted red line represents no change. We note that the nuclear volume and the size and number of nuclear speckles also changed upon treatment with α-amanitin, and these might contribute to the observed changes in nuclear body association rates.

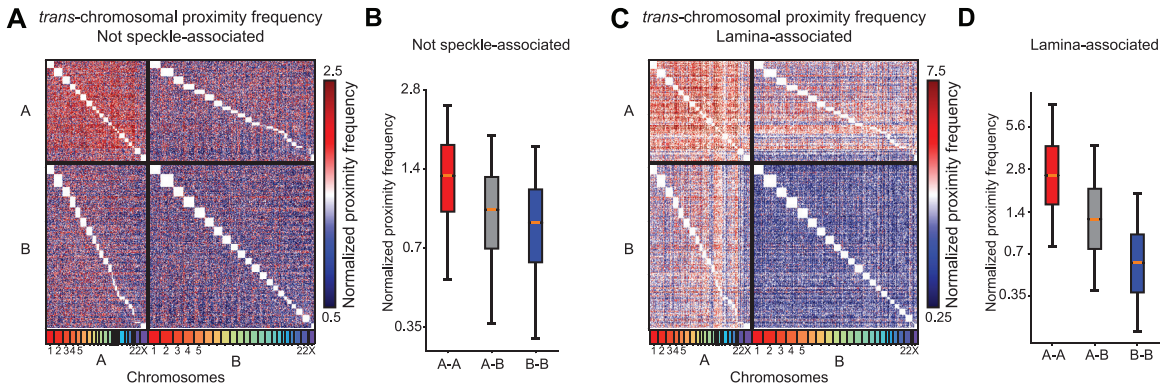


Figure S7. Enrichment of *trans*-Chromosomal Active Chromatin Interactions in Different Nuclear Environments, Related to Figures 5 and 7

(A) Normalized *trans*-chromosomal proximity frequency matrix, as in Figure 5A, but considering only loci that are not associated with nuclear speckles. For each locus pair, only cells in which neither locus is associated with nuclear speckles are considered.

(B) Distribution of *trans*-chromosomal proximity frequency for pairs of A loci (A-A), pairs of B loci (B-B), and pairs comprised of one A and one B locus (A-B), as in Figure 5B, but considering only the cells in which neither locus is associated with nuclear speckles.

(C) Normalized *trans*-chromosomal proximity frequency matrix, as in Figure 5A, but considering only loci that are associated with nuclear lamina. For each locus pair, only cells in which both loci are associated with nuclear lamina are considered.

(D) Distribution of *trans*-chromosomal proximity frequency for pairs of A loci (A-A), pairs of B loci (B-B), and pairs comprised of one A and one B locus (A-B), as in Figure 5B, but considering only the cells in which both loci are associated with nuclear lamina.