Science Advances
AAAS

Supplementary Materials for

Transcription-coupled changes in genomic region proximities during transcriptional bursting

Hiroaki Ohishi *et al.*

Corresponding author: Hiroshi Ochiai, ochiai.hiroshi.403@m.kyushu-u.ac.jp

Sci. Adv. **10**, eadn0020 (2024)
DOI: 10.1126/sciadv.adn0020

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

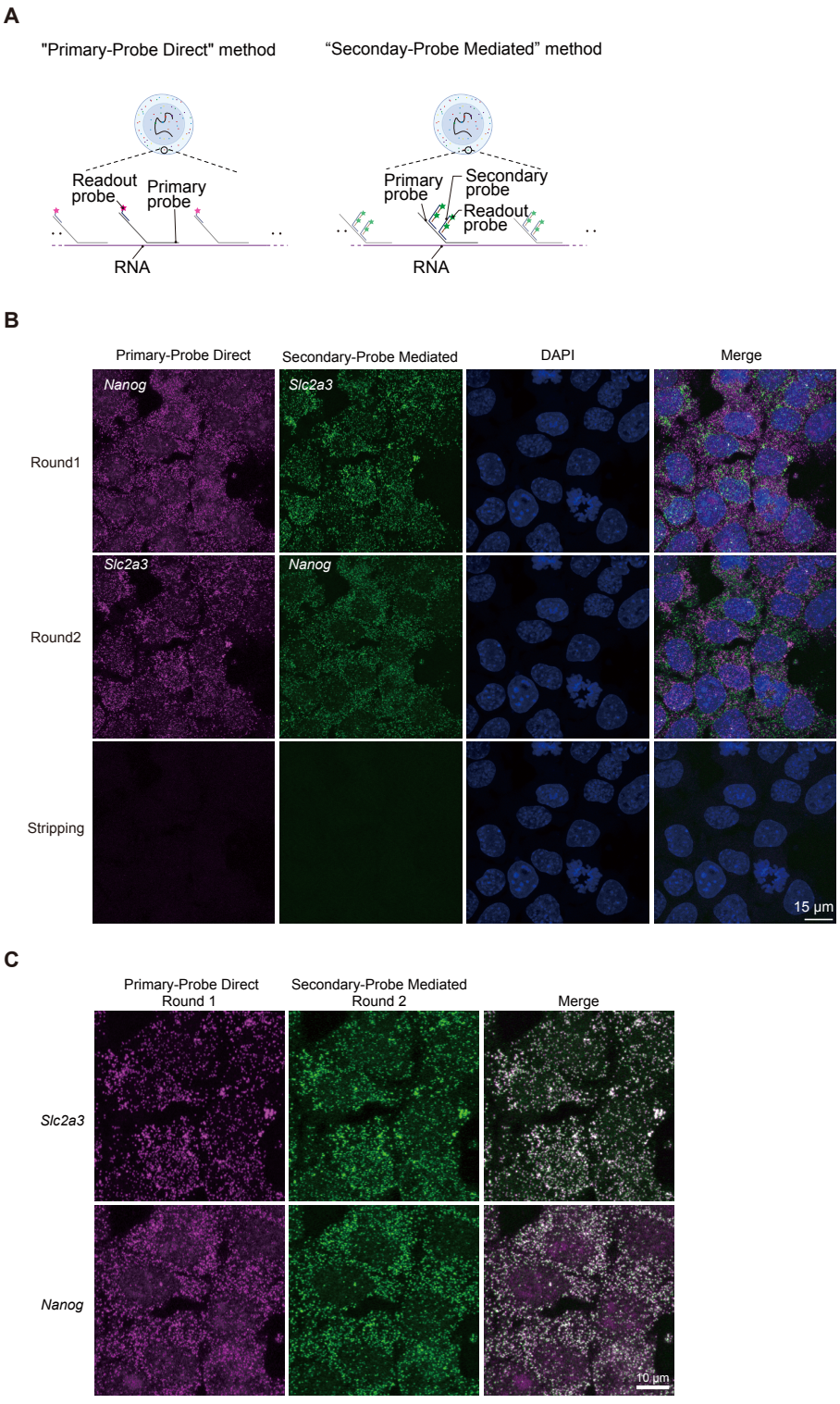
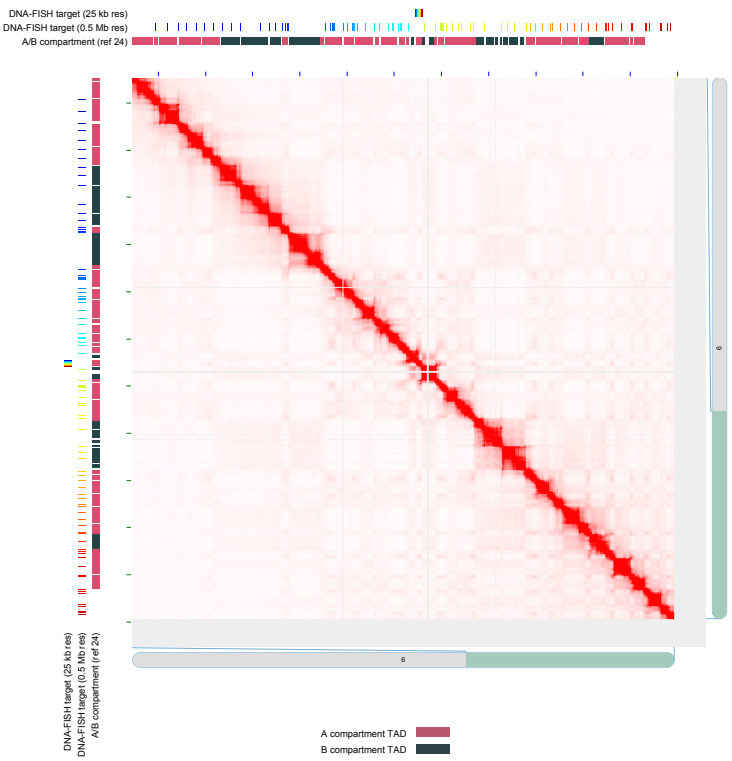


Fig. S1. Validation of spot detection method in seq-RNA-FISH. (A) Schematic representation of the fluorescent labeling strategies employed in seq-DNA/RNA/IF-FISH, specifically the "Primary-Probe Direct" and "Secondary-Probe Mediated" methods, with an example focusing on RNA detection. By simultaneously adding and hybridizing both the secondary and readout probes to the cellular samples, the readout probes become localized at the site of the primary

probes. This strategy requires preparing readout probes corresponding to the number of fluorescent dye types utilized, thereby streamlining the overall process and reducing cost. **(B)** Validation of both "Primary-Probe Direct" and "Secondary-Probe Mediated" seq-RNA-FISH methodologies. C57BL6/J mouse ES cells cultured in 2i medium were subjected to seq-RNA-FISH. We designed either 24 or 48 primary probes corresponding to the "Primary-Probe Direct" or "Secondary-Probe Mediated" seq-RNA-FISH methods for the detection of *Nanog* and *Slc2a3* mRNA. These primary probes were initially hybridized. Subsequently, the appropriate readout and secondary probes were hybridized to visualize the RNA (Round 1). Following imaging via microscopy, the secondary and readout probes were stripped off, and a different set of suitable readout and secondary probes were hybridized to re-visualize the RNA (Round 2). Signals were observed using both "Primary-Probe Direct" labeling via readout probes and "Secondary-Probe Mediated" labeling via secondary probes. Finally, imaging was conducted after stripping the readout and secondary probes to confirm their successful removal as intended. **(C)** Fluorescent spots labeled by the "Primary-Probe Direct" method were found to match fluorescent spots labeled by the "Secondary-Probe Mediated" method. This figure was created using the same image as in (B) but with slight magnification to illustrate the overlap of fluorescent spots detected by both methods, clarifying that the fluorescence spots detected by both methods overlap.

A


DNA-FISH target (25 kb res)
DNA-FISH target (0.5 Mb res)
A/B compartment (ref 24)
6
6
DNA-FISH target (25 kb res)
DNA-FISH target (0.5 Mb res)
A/B compartment (ref 24)
A compartment TAD
B compartment TAD

B

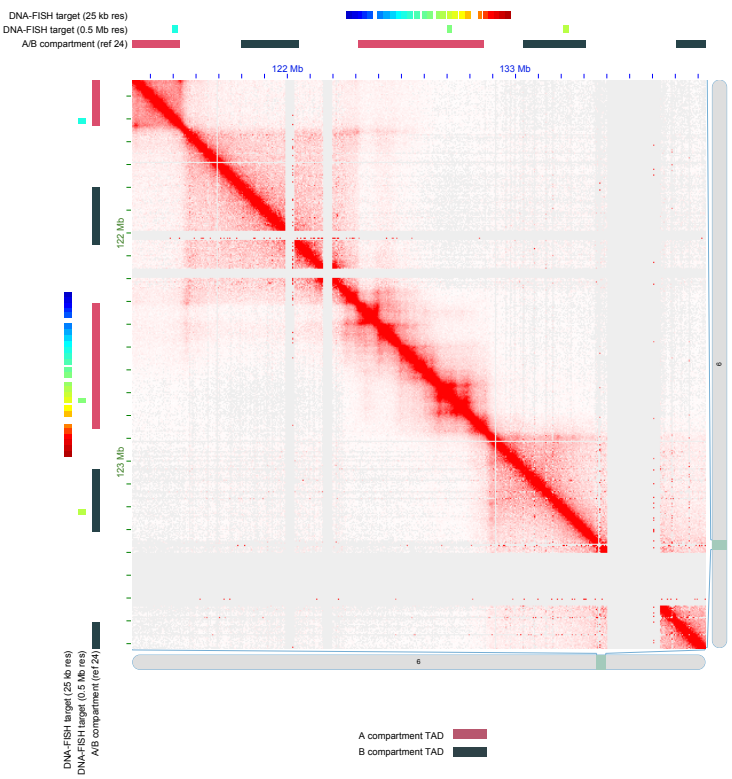

DNA-FISH target (25 kb res)
DNA-FISH target (0.5 Mb res)
A/B compartment (ref 24)
122 Mb
133 Mb
122 Mb
123 Mb
6
6
DNA-FISH target (25 kb res)
DNA-FISH target (0.5 Mb res)
A/B compartment (ref 24)
A compartment TAD
B compartment TAD

Fig. S2. Hi-C data of the chromosome 6 regions targeted in seq-DNA-FISH. (**A-B**) Hi-C data (ref *24*) within genomic regions where probes were designed at approximately 0.5 Mb intervals **(A)** or 25 kb intervals **(B)**. The A/B compartments are denoted as annotated in (ref *24*). The regions designed at 25 kb intervals roughly correspond to a single Topologically Associating Domain (TAD).

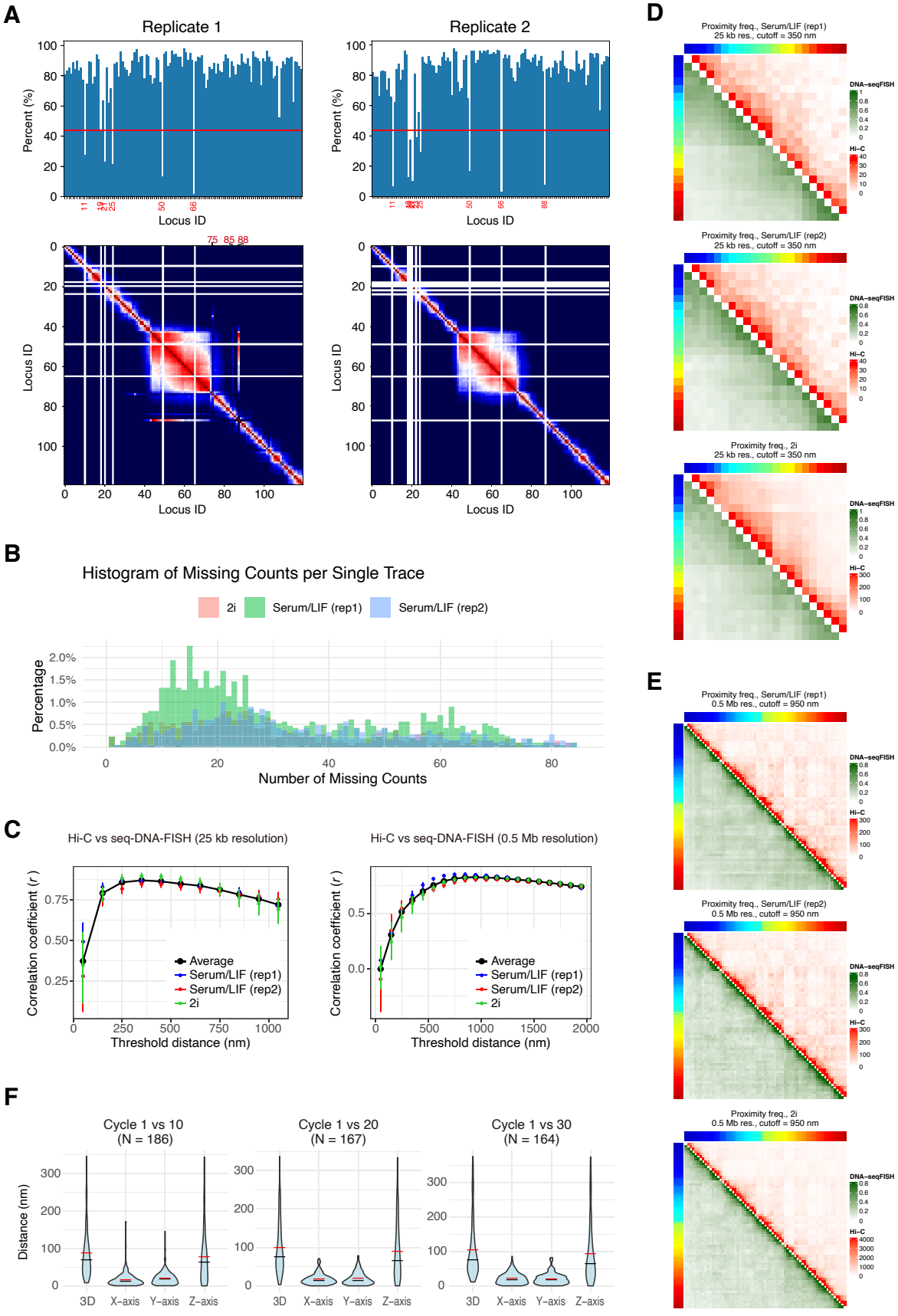

A
Replicate 1
Replicate 2
Percent (%)
Locus ID
B
Histogram of Missing Counts per Single Trace
2i
Serum/LIF (rep1)
Serum/LIF (rep2)
Percentage
Number of Missing Counts
C
Hi-C vs seq-DNA-FISH (25 kb resolution)
Hi-C vs seq-DNA-FISH (0.5 Mb resolution)
Correlation coefficient (r)
Threshold distance (nm)
Average
Serum/LIF (rep1)
Serum/LIF (rep2)
2i
D
Proximity freq., Serum/LIF (rep1)
25 kb res., cutoff = 350 nm
Proximity freq., Serum/LIF (rep2)
25 kb res., cutoff = 350 nm
Proximity freq., 2i
25 kb res., cutoff = 350 nm
DNA−seqFISH
Hi−C
E
Proximity freq., Serum/LIF (rep1)
0.5 Mb res., cutoff = 950 nm
Proximity freq., Serum/LIF (rep2)
0.5 Mb res., cutoff = 950 nm
Proximity freq., 2i
0.5 Mb res., cutoff = 950 nm
F
Cycle 1 vs 10
(N = 186)
Cycle 1 vs 20
(N = 167)
Cycle 1 vs 30
(N = 164)
Distance (nm)
3D
X−axis
Y−axis
Z−axis

Fig. S3. Quality assessment **of seq-DNA-FISH experiments (1). (A)** Efficacy of fluorescence signal detection for each targeted genomic region (Upper Panel) and corresponding median distance matrices in seq-DNA-FISH experiments (Lower Panel). The red line on the upper panel represents half of the average detection efficiency across all targeted regions in each experiment. Regions falling below this threshold **(48.9%)** were excluded from further analysis. In the median distance matrices, regions that are proximal in terms of 1D genomic distance are anticipated to exhibit spatial proximity in the 3D median distance as well. However, regions ID 74, 84, 87 in Replicate 1 deviated from this intuition and substantially differed from Replicate 2 data; thus, they were also excluded from the analysis. Replicate 1: $N = 1{,}100$ alleles; Replicate 2: $N = 580$ alleles. Unless otherwise stated, data from Replicate 1 are depicted in the subsequent Figures. **(B)** Distribution of the number of missing values per allele. Red, green, and blue bars represent data from cells cultured in for 2i, serum/LIF replicate 1, and serum/LIF replicate 2, respectively. **(C)** Correlation between seq-DNA-FISH and Hi-C data. Analysis of the correlation between seq-DNA-FISH proximity frequencies at set thresholds and Hi-C data for mouse ES cells in serum/LIF (ref *24*) and 2i media (ref *30*). Error bars represent the 95% confidence interval from 1000 bootstrap samples. Black dots show average values for serum/LIF and 2i replicates, with the highest average indicating the optimal proximity distance threshold—350 nm for 25 kb resolution (left panel) and 950 nm for 0.5 Mb resolution (right panel). **(D)** Comparison of proximity frequency matrices between Hi-C and seq-DNA-FISH data at 25 kb resolution. Proximity frequency for seq-DNA-FISH data was calculated using a 350 nm cutoff. Hi-C data from mouse ES cells in serum/LIF (ref 24) and 2i media (ref *30*) were used. Hi-C data are presented as count data. The color bar indicates the coordinates of the target regions as shown in Fig. 1B. **(E)** Comparison of proximity frequency matrices between Hi-C and seq-DNA-FISH data at 0.5 Mb resolution. Proximity frequency for seq-DNA-FISH data was calculated using a 950 nm cutoff. Hi-C data from mouse ES cells in serum/LIF (ref 24) and 2i media (ref *30*) were used. Hi-C data are presented as count data. The color bar indicates the coordinates of the target regions as shown in Fig. 1B. **(F)** Localization precision of seq-DNA-FISH. We determined the localization precision of the imaging system by imaging the same locus in different cycles and calculating the distance distributions between them. Specifically, we calculated the x, y, and z displacements, and the 3D distance between the initial fiducial marker positions and those from the 10th, 20th, and 30th imaging cycles. The red and black horizontal lines indicate the mean and median values, respectively.

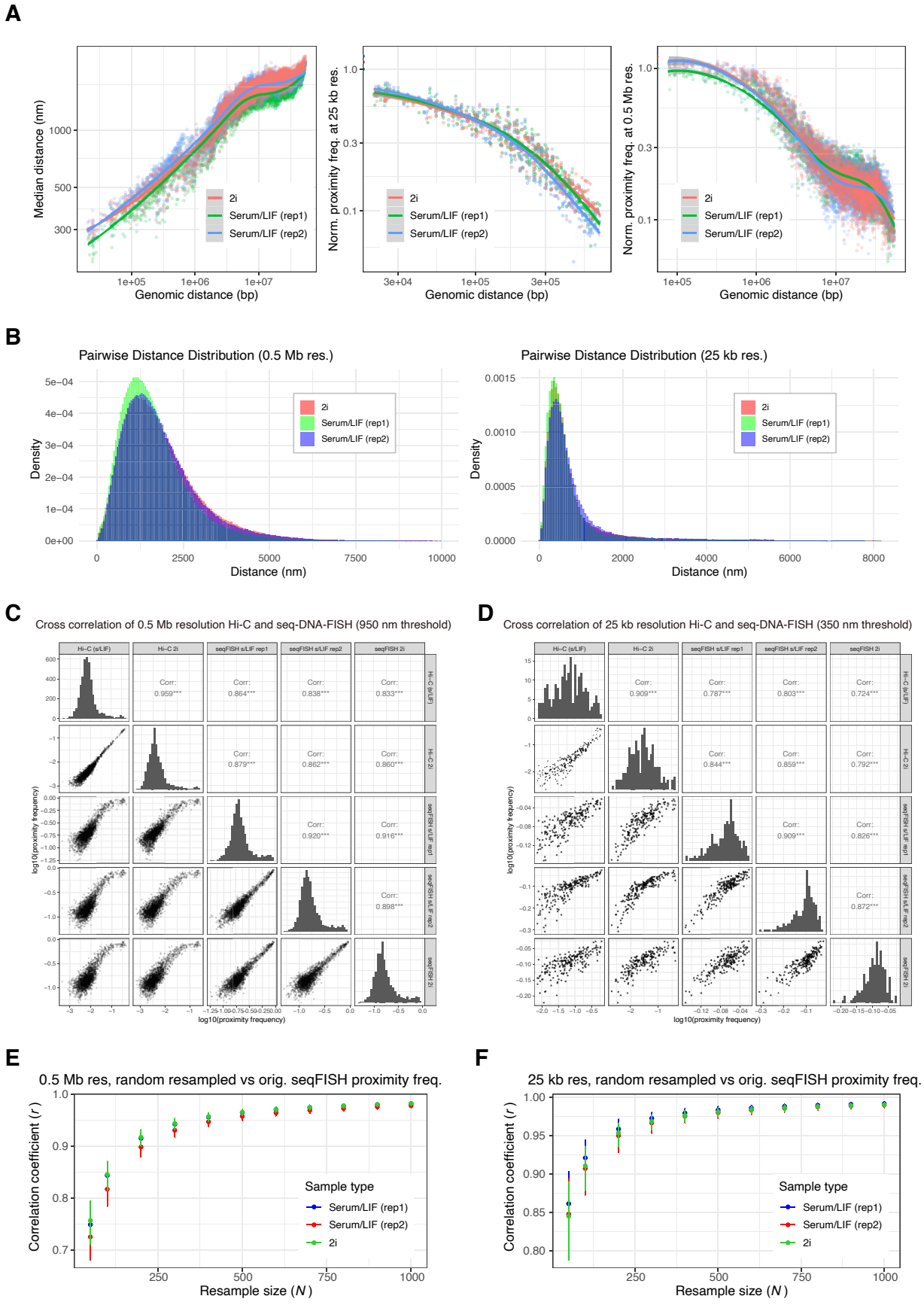

A
Median distance (nm)
Genomic distance (bp)
Norm. proximity freq. at 25 kb res.
Norm. proximity freq. at 0.5 Mb res.
2i
Serum/LIF (rep1)
Serum/LIF (rep2)
B
Pairwise Distance Distribution (0.5 Mb res.)
Pairwise Distance Distribution (25 kb res.)
Density
Distance (nm)
C
Cross correlation of 0.5 Mb resolution Hi-C and seq-DNA-FISH (950 nm threshold)
D
Cross correlation of 25 kb resolution Hi-C and seq-DNA-FISH (350 nm threshold)
log10(proximity frequency)
E
0.5 Mb res, random resampled vs orig. seqFISH proximity freq.
F
25 kb res, random resampled vs orig. seqFISH proximity freq.
Correlation coefficient (r)
Resample size (N)
Sample type

Fig. S4. Quality assessment of seq-DNA-FISH experiments (2). (A) Scatter plots of distance and proximity frequency. The upper panel displays median distances, the middle panel shows normalized proximity frequencies at 25 kb resolution (350 nm threshold), and the lower panel at 0.5 Mb resolution (950 nm threshold), against genomic distance. Data from two replicates in serum/LIF and 2i media are included, with LOESS regression curves depicted. (**B**) Pairwise distance distributions in seq-DNA-FISH data at 0.5 Mb and 25 kb resolutions. Data from two replicates in serum/LIF and 2i media are included. **(C-D)** Scatter plots (located in the lower left) and correlation coefficients (found in the upper right) for comparisons of proximity frequencies between Hi-C and seq-DNA-FISH data for mouse ES cells cultured under serum/LIF (s/LIF) and 2i conditions. Specifically, part (C) showcases the comparison at a 0.5 Mb resolution, while part (D) focuses on the comparison at a 25 kb resolution. **(E-F)** Sampling reflection of population proximity frequencies. This analysis evaluates how well the sample sizes reflect the overall population's proximity frequencies by employing random resampling at 950 nm for 0.5 Mb resolution (E), and at a 350 nm threshold for 25 kb resolution (F). Resampling was conducted X times from the data, allowing for duplicates within each resampling iteration, across serum/LIF replicate 1 (N = 1,100), serum/LIF replicate 2 (N = 580), and 2i medium (N = 484) conditions. This resampling process was repeated 1,000 times. Based on these iterations, the average correlation coefficient between the resampled proximity frequencies and the original dataset's proximity frequencies was calculated, along with its 95% confidence interval (CI). The 95% CI was determined from the distribution of correlation coefficients obtained across these 1,000 repetitions, rather than varying the resample size X, which ranged from 1,000 to 50. Error bars in the figures represent the 95% confidence interval.

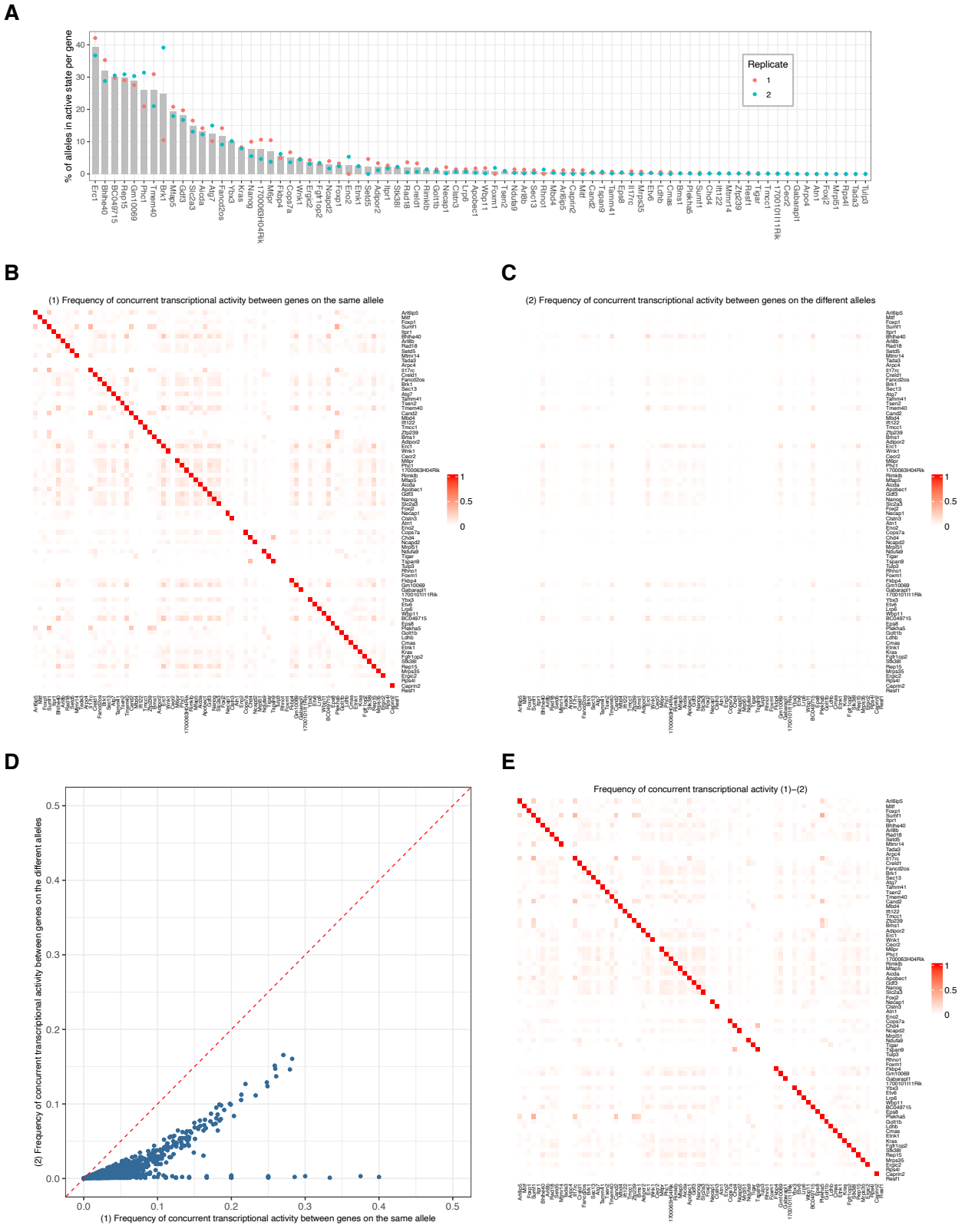


Fig. S5. Frequency of concurrent transcriptional activity between genes on the same allele. **(A)** Proportion of cells with transcription foci. This bar graph shows the average percentage of

cells exhibiting transcription foci across Replicates 1 and 2, illustrating the frequency of transcriptional activity observed within the cell population. **(B)** Data (1): Frequency of concurrent transcriptional activity between genes located on the same allele. The dataset comprises information from 1,100 alleles. Concurrent transcriptional activity is defined as the number of alleles where *Gene X* and *Gene Y* are simultaneously in the active state divided by the number of alleles where either *Gene X* or *Gene Y* is in the active state. Due to varying transcriptional frequencies for each gene, the sample size (*N*) differs between elements. **(C)** Data (2): Frequency of concurrent transcriptional activity between genes located on different alleles within the same cell. The dataset includes data from 489 cells. Concurrent transcriptional activity here is defined as the number of cells where *Gene X* on Allele 1 and *Gene Y* on Allele 2 are simultaneously in the active state, divided by the number of cells where either *Gene X* on Allele 1 or *Gene Y* on Allele 2 is in the active state. As with (B), due to different transcriptional frequencies for each gene, the sample size (*N*) varies between elements. **(D)** Scatter plot of Data (1) and Data (2). The red line represents the $y = x$ line. Genes that exhibit high frequency of concurrent transcriptional activity within the same allele also tend to have a high frequency of concurrent transcriptional activity between different alleles within the same cell. However, the frequency of concurrent transcriptional activity is markedly higher when both genes locate on the same allele. This suggests that while global nuclear factors (e.g., expression levels of transcription factors) might have a moderate influence on concurrent transcriptional activity, the genomic structure of locating on the same allele could partially induce concurrent transcriptional activity. **(E)** Heatmap of Data (1) - Data (2), derived from (B) and (C). This reveals what could be termed the apparent true frequency of concurrent transcriptional activity on the same allele, accounting for the influence of overall nuclear factors that contribute to concurrent transcriptional activity. It is important to note that this calculation is a simplified model and may not fully capture the complexities involved. Additional, more detailed analyses are necessary to comprehensively understand the true nature of concurrent transcriptional activity on the same allele. Several factors, such as varying transcriptional frequencies, sample size differences, and the influence of global nuclear factors, may introduce biases or limitations in this estimate.

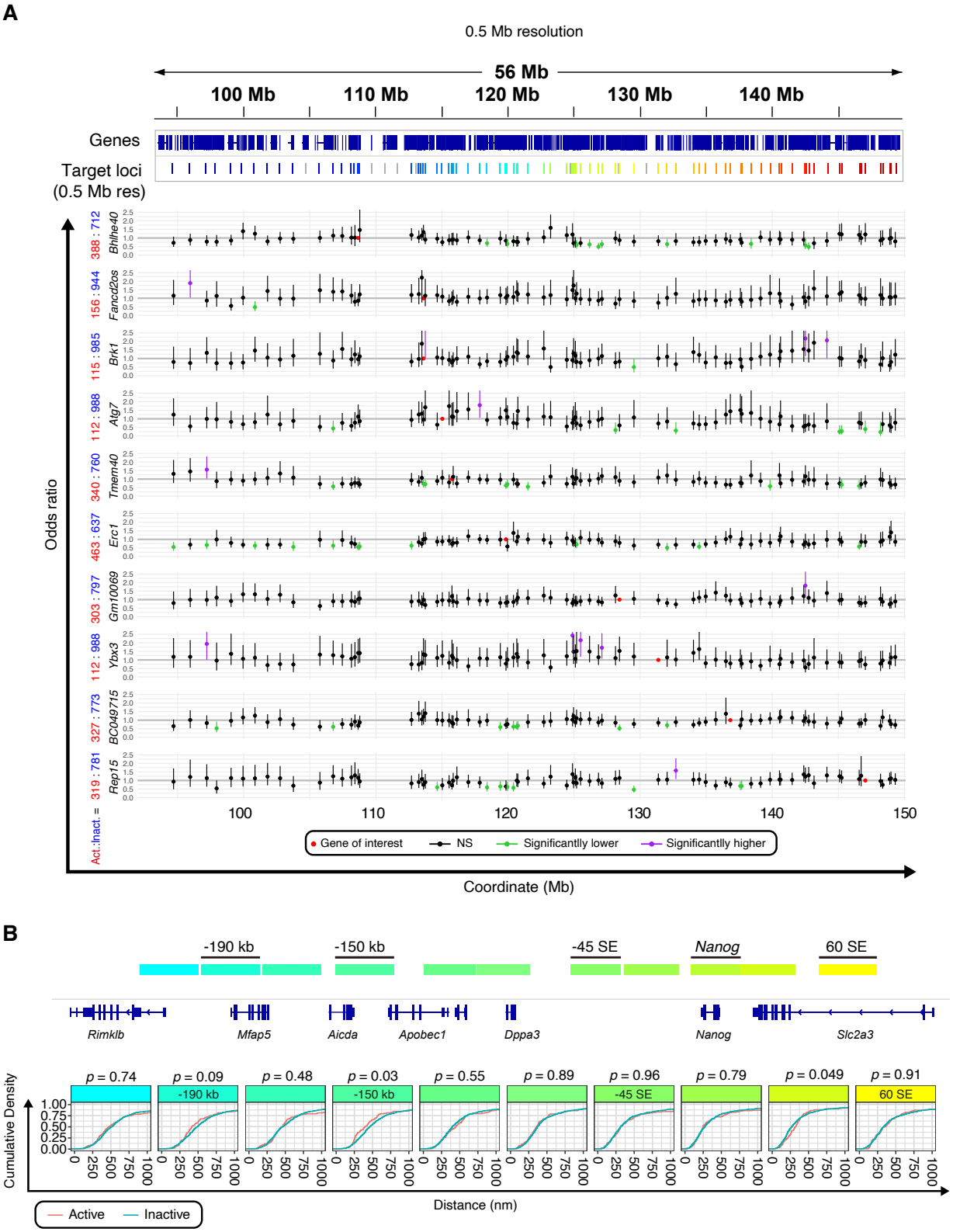


Fig. S6. Genomic proximity specific to transcriptionally active states. (A) Odds ratios for genomic proximities. This panel shows odds ratios calculated to compare proximity frequencies

with the region of interest for specific genes, categorized by their transcriptional activity states at a 0.5 Mb resolution. Error bars indicate the 95% confidence interval. Red, purple, green, and black dots respectively represent regions containing the gene of interest, regions with the lower limit of the 95% CI above 1, regions with the upper limit of the 95% CI below 1, and regions not classified within these criteria. The count of alleles where the gene of interest is in an active (Act.) or inactive (Inact.) state is specified beside the gene name. **(B)** Cumulative density distribution of 3D distances. Analysis of 3D distances between the *Nanog* gene and specific genomic regions, with sample sizes for active (inactive) states of *Nanog* denoted as $N = 110$ (990). *P*-values were calculated using a two-sample Kolmogorov-Smirnov test to assess the distribution differences.

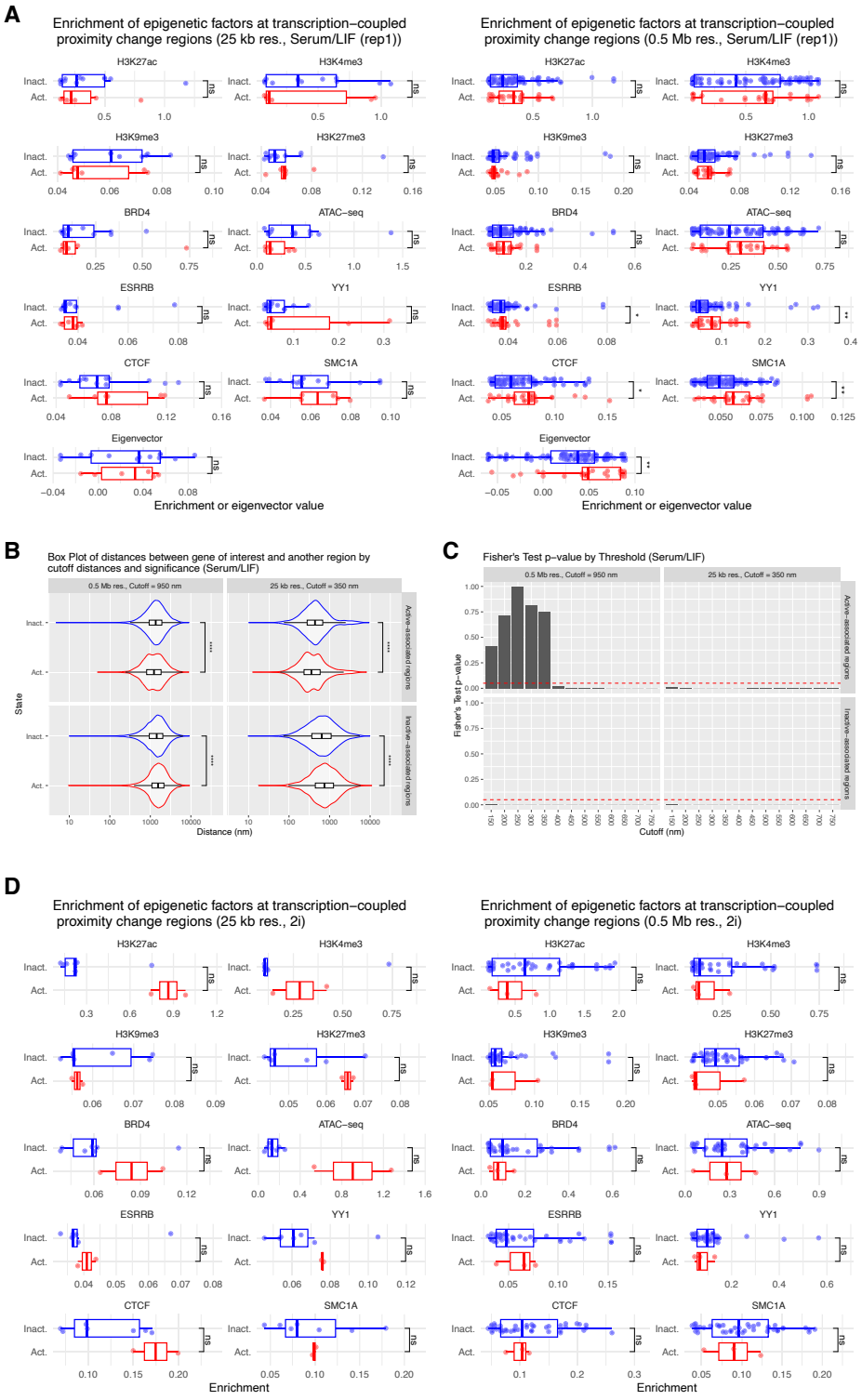


Fig. S7. Epigenomic characteristics of genomic regions proximally localizing specifically in active or inactive states. (A) Epigenomic characteristics of regions that proximally localize with specific gene regions in either active or inactive states in cells cultured under serum/LIF conditions. Proximity frequencies for specific gene transcriptionally active states were calculated at 25 kb resolution with a 350 nm cutoff or 0.5 Mb resolution with a 950 nm cutoff. The average enrichment of ATAC-seq, H3K27ac, H3K27me3, H3K4me3, H3K9me3, BRD4, ESRRB, YY1,

SMC1A, and CTCF, or the average eigenvector values in regions with significantly higher odds ratios (significantly proximal in the active state, denoted as Act.) or lower odds ratios (significantly proximal in the inactive state, denoted as Inact.) were plotted. Statistical significance was determined using the Wilcoxon test. ns, *, ** indicate $p > 0.05$, $p <= 0.05$, and $p <= 0.01$, respectively. (**B**) The distribution of distances between regions with significantly increased (Active-associated regions) or decreased proximity frequencies (Inactive-associated regions) and the gene of interest, plotted for each transcriptional active (Active (Act.) or inactive states (Inact.)) at 25 kb resolution with a 350 nm cutoff or 0.5 Mb resolution with a 950 nm cutoff. Statistical significance was determined using the Wilcoxon test. **** indicates $p <= 0.0001$. (**C**) p-values calculated using the Fisher's exact test, based on the proportion of distances below specific cutoffs between regions that proximally localize specifically in active or inactive transcriptional states (Active-associated regions or Inactive-associated regions) and the gene of interest. The red dashed line indicates $p = 0.05$. (**D**) Epigenomic characteristics of regions that proximally localize with specific gene regions in either active or inactive states in cells cultured under 2i conditions. Proximity frequencies for specific gene transcriptional activation states were calculated at 25 kb resolution with a 350 nm cutoff or 0.5 Mb resolution with a 950 nm cutoff. The average enrichment of ATAC-seq, H3K27ac, H3K27me3, H3K4me3, H3K9me3, BRD4, ESRRB, YY1, SMC1A, and CTCF in regions with significantly higher odds ratios (significantly proximal in the active state, denoted as Act.) or lower odds ratios (significantly proximal in the inactive state, denoted as Inact.) were plotted. Statistical significance was determined using the Wilcoxon test. ns indicates $p > 0.05$.

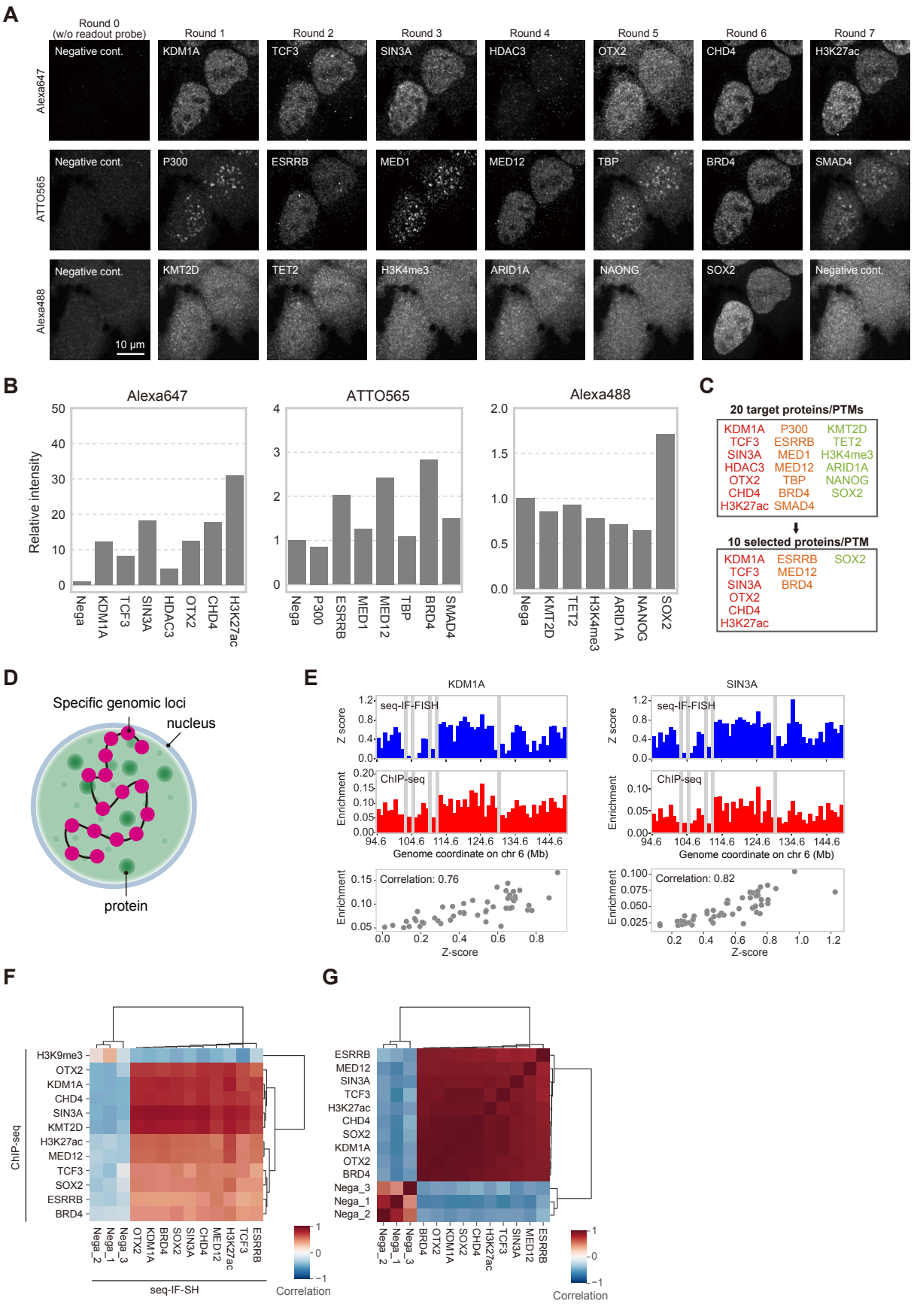

A
Round 0 (w/o readout probe)
Round 1
Round 2
Round 3
Round 4
Round 5
Round 6
Round 7
Alexa647
ATTO565
Alexa488
Negative cont.
KDM1A
TCF3
SIN3A
HDAC3
OTX2
CHD4
H3K27ac
P300
ESRRB
MED1
MED12
TBP
BRD4
SMAD4
KMT2D
TET2
H3K4me3
ARID1A
NANOG
SOX2
10 μm
B
Alexa647
ATTO565
Alexa488
Relative intensity
Nega
KDM1A
TCF3
SIN3A
HDAC3
OTX2
CHD4
H3K27ac
P300
ESRRB
MED1
MED12
TBP
BRD4
SMAD4
KMT2D
TET2
H3K4me3
ARID1A
NANOG
SOX2
C
20 target proteins/PTMs
KDM1A P300 KMT2D
TCF3 ESRRB TET2
SIN3A MED1 H3K4me3
HDAC3 MED12 ARID1A
OTX2 TBP NANOG
CHD4 BRD4 SOX2
H3K27ac SMAD4
10 selected proteins/PTM
KDM1A ESRRB SOX2
TCF3 MED12
SIN3A BRD4
OTX2
CHD4
H3K27ac
D
Specific genomic loci
nucleus
protein
E
KDM1A
SIN3A
seq-IF-FISH
Z score
ChIP-seq
Enrichment
Genome coordinate on chr 6 (Mb)
Correlation: 0.76
Correlation: 0.82
Z-score
F
ChIP-seq
H3K9me3
OTX2
KDM1A
CHD4
SIN3A
KMT2D
H3K27ac
MED12
TCF3
SOX2
ESRRB
BRD4
Nega_2
Nega_1
Nega_3
seq-IF-SH
Correlation
G
ESRRB
MED12
SIN3A
TCF3
H3K27ac
CHD4
SOX2
KDM1A
OTX2
BRD4
Nega_3
Nega_1
Nega_2
Correlation

Fig. S8. **Quality check of seq-IF-FISH data. (A)** Representative visualization of intranuclear proteins and post-translational modifications in seq-IF-FISH analysis. **(B)** Fluorescence intensity values were measured within the cell nucleus for each probe utilized in seq-IF-FISH. These values are presented as relative fluorescent intensities when compared to a negative control within the same fluorescence channel. Sample size: $N = 611$. **(C)** Initially targeted proteins and post-translational modifications (PTMs) in seq-IF-FISH are listed, along with those ultimately selected for further analysis. Proteins/PTMs with a relative fluorescent intensity value of 1.5 or below in b were excluded from further analysis. Additionally, HDAC3 and SMAD4, whose localization patterns substantially deviated from those obtained through general immunofluorescence, were also excluded. **(D)** Schematic representation of the localization of proteins/PTMs in specific genomic regions. **(E)** Correlation between fluorescence intensity data from seq-IF-FISH and ChIP-seq enrichment data, featuring KDM1A and SIN3A. The top graph displays fluorescent intensity Z-scores of KDM1A and SIN3A in seq-IF-FISH at specific genomic coordinates as determined by seq-DNA-FISH (details refer to Materials and Methods). $N = 611$. The middle graph represents ChIP-seq enrichment values for the same region (details refer to Materials and Methods). Gray areas denote locations excluded from analysis in seq-DNA-FISH. The bottom graph shows the scatter plot of fluorescent intensity Z-scores in seq-IF-FISH and ChIP-seq enrichment. **(F)** Heatmap of correlation values between fluorescent intensity Z-scores from seq-IF-FISH and ChIP-seq enrichment. "Nega_1," "Nega_2," and "Nega_3" represent the negative controls for the Alexa647, ATTO565, and Alexa488 fluorescent channels, respectively. **(G)** Heatmap of correlation values between fluorescent intensity Z-scores in seq-IF-FISH.

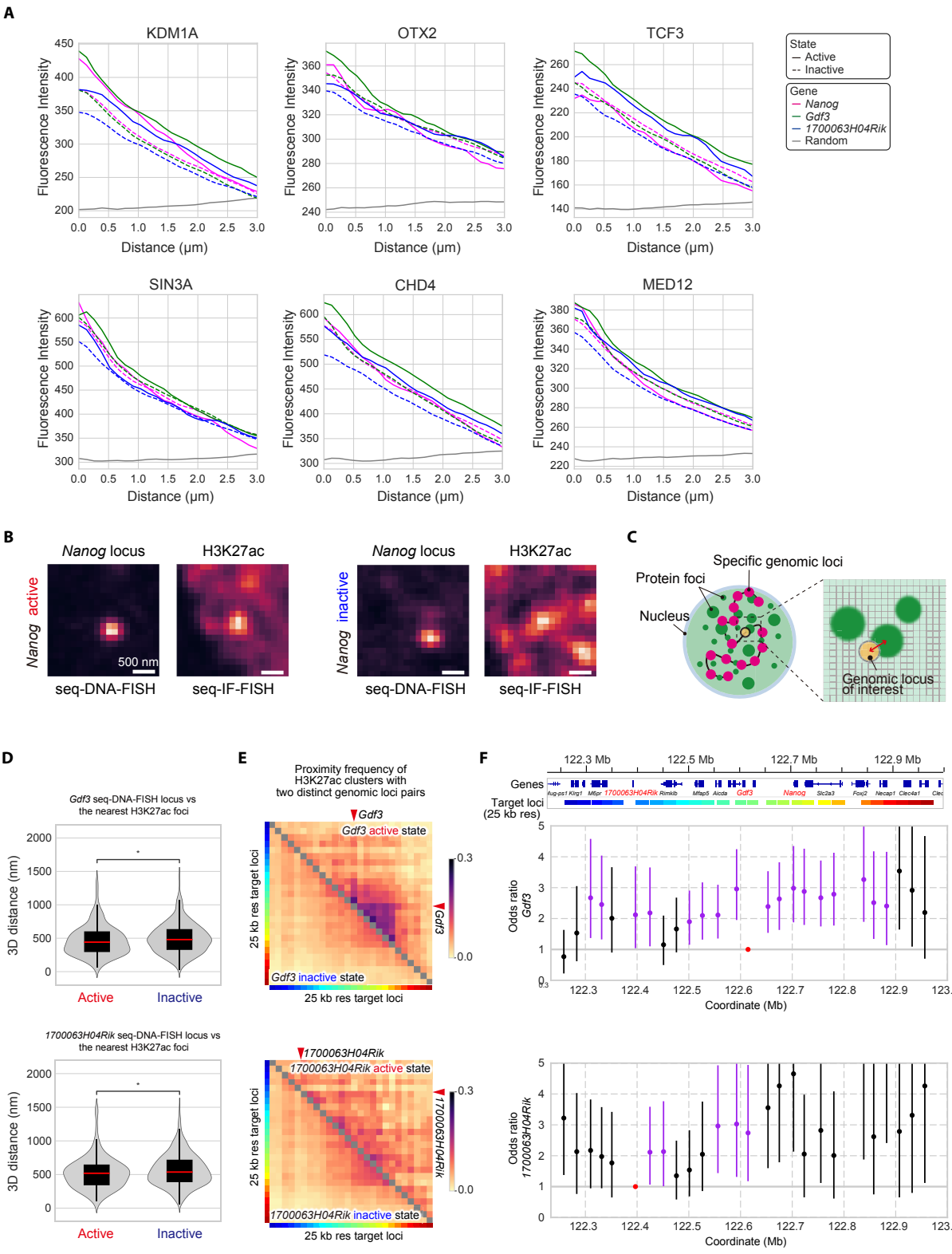


Fig. S9. Spatial distribution of transcriptional regulatory factors relative to transcriptional activity states. (A) Radial distribution function of KDM1A, OTX2, TCF3, SIN3A, CHD4, and MED12 seq-IF-FISH signals for the median IF signal centered at *Nanog*, *Gdf3*, and

1700063H04Rik seq-DNA-FISH foci. Data are shown separately for active and inactive states of each gene. Sample sizes for each gene in the active (inactive) states are as follows: *Nanog*: N = 110 (990); *Gdf3*: N = 217 (883); *1700063H04Rik*: N = 117 (983). **(B)** Example images of the of H3K27ac seq-IF-FISH, in relation to the Nanog seq-DNA-FISH focus. Transcription state of *Nanog* was determined by using seq-RNA-FISH data. **(C)** Schematic representation elucidating the spatial relationship between specific genomic regions and foci of proteins/PTMs. **(D)** 3D distance distribution between the *Gdf3 and 1700063H04Rik* seq-DNA-FISH foci and the nearest H3K27ac foci. Data are shown separately for active and inactive states of each gene. The red line within the box plot indicates the median. Statistical significance was determined using the Wilcoxon test. * indicates $p <= 0.05$. **(E)** Colocalization frequency matrix denoting the presence of H3K27ac clusters proximal to two distinct genomic loci pairs, with colocalization defined as being within 350 nm of each other. Color coding on the left or bottom of each heatmap corresponds to the seq-DNA-FISH target regions at a 25 kb resolution, linked to gene regions as detailed in (F). Red arrowheads indicate the location of the genes of interest. **(F)** Odds ratios comparing the frequency at which regions containing the gene of interest, a specific genomic region, and H3K27ac foci are within 350 nm proximity, based on the gene's activity state. Error bars denote the 95% confidence interval. Red, purple, and black dots indicate regions with the target gene, regions where the lower limit of the 95% CI exceeds 1, and regions not categorized as such, respectively.

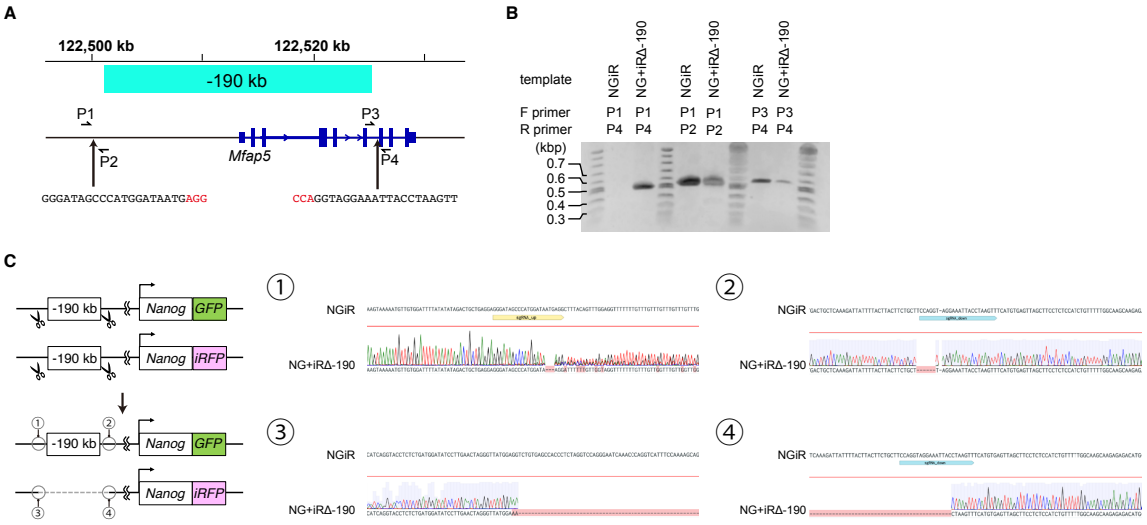


Fig. S10. Establishment of a cell line with deletion of the -190 kb region. (A) Schematic representation of the -190 kb region, which encompasses the transcriptional start site of *Mfap5*. The nucleotide sequences noted at the bottom of the Figure represent the CRISPR target sites, with the Protospacer Adjacent Motif (PAM) sequence indicated in red. Positions P1 to P4 denote the locations of the primers used in (B). **(B)** Genomic PCR of WT(NGiR) and deletion cell lines (NG+iRΔ-190). Genomic PCR was performed on both NGiR and NG+iRΔ-190, using PCR conditions optimized to allow amplification of sequences up to approximately 2 kb in length. Various combinations of the primers specified in a were used to perform the genomic PCR. In the out-out (P1 and P4 combination) primer set, amplification occurs in the NG+iRΔ-190 but not in the NGiR. However, amplification is observed in the NG+iRΔ-190 for both the P1-P2 and P3-P4 primer combinations, suggesting that only one allele is deleted in the NG+iRΔ-190 cell line. **(C)** Nucleotide sequence determination of the PCR-amplified fragments from genomic PCR. A deletion of 6 bases is observed in the downstream CRISPR target sequence of the non-deleted allele of NG+iRΔ-190. The upstream CRISPR target sequence of the non-deleted allele in NG+iRΔ-190 is unclear due to an ambiguous sequence chromatogram, although a deletion of a few bases is suspected. For the deleted allele of NG+iRΔ-190, a loss extends from the CRISPR target sequence up to approximately 190 bp upstream.

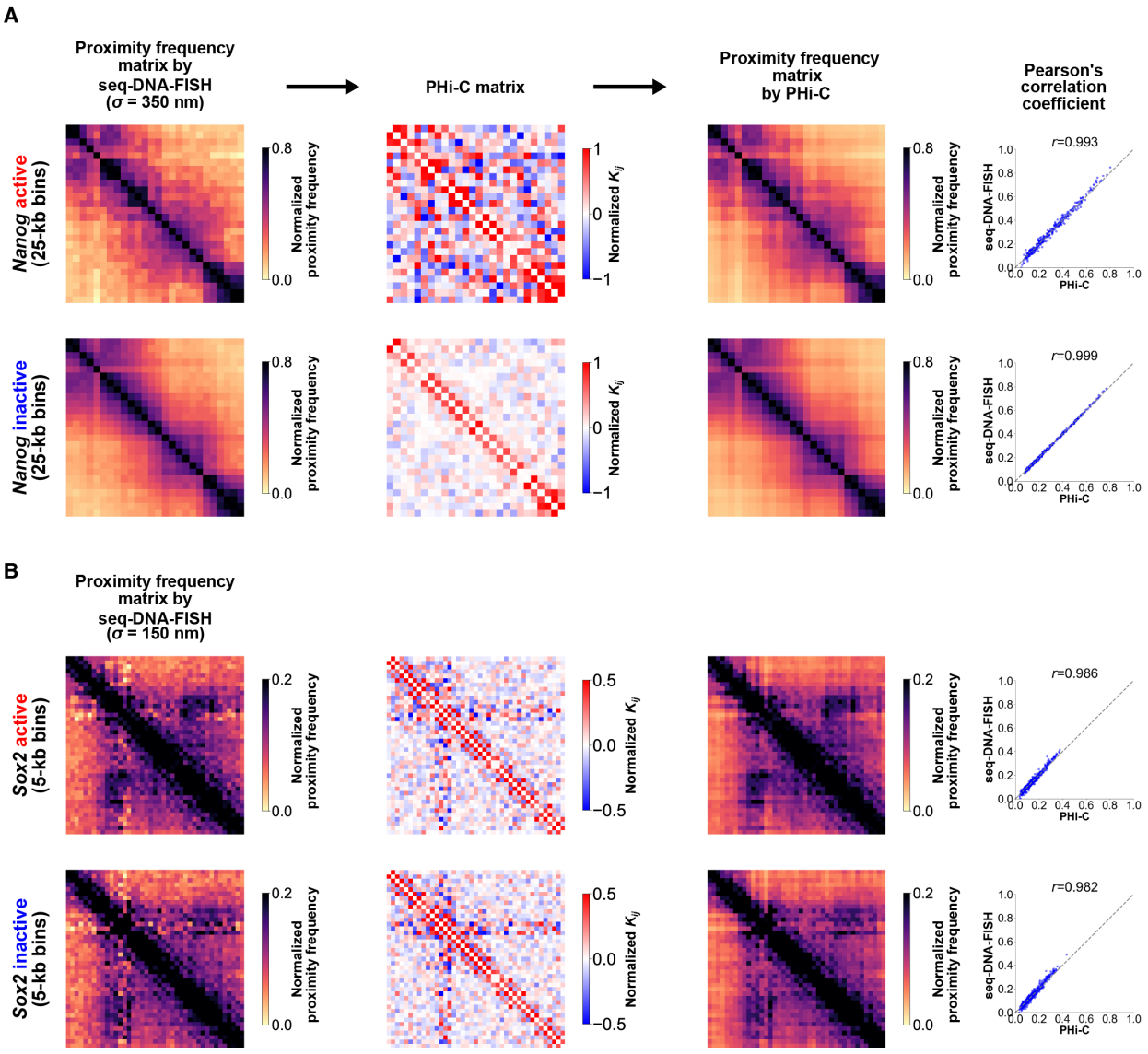


Fig. S11. Data processing in the PHi-C polymer model. (A-B) Processed data during the application of the PHi-C polymer model, focusing on the proximity frequency calculations derived from the transcriptional activity states of *Nanog* (A) and *Sox2* (B). These calculations serve as inputs for the PHi-C analysis. Specifically, (A) utilizes seq-DNA/RNA-FISH data from this study for *Nanog*, with a resolution of 25 kb and σ set to 350 nm. Conversely, (B) employs data from Huang et al. (*37*) for *Sox2*, with resolutions of 5 kb, and σ set to 150 nm. The methodology for data processing follows the flow described in Fig. 5A.

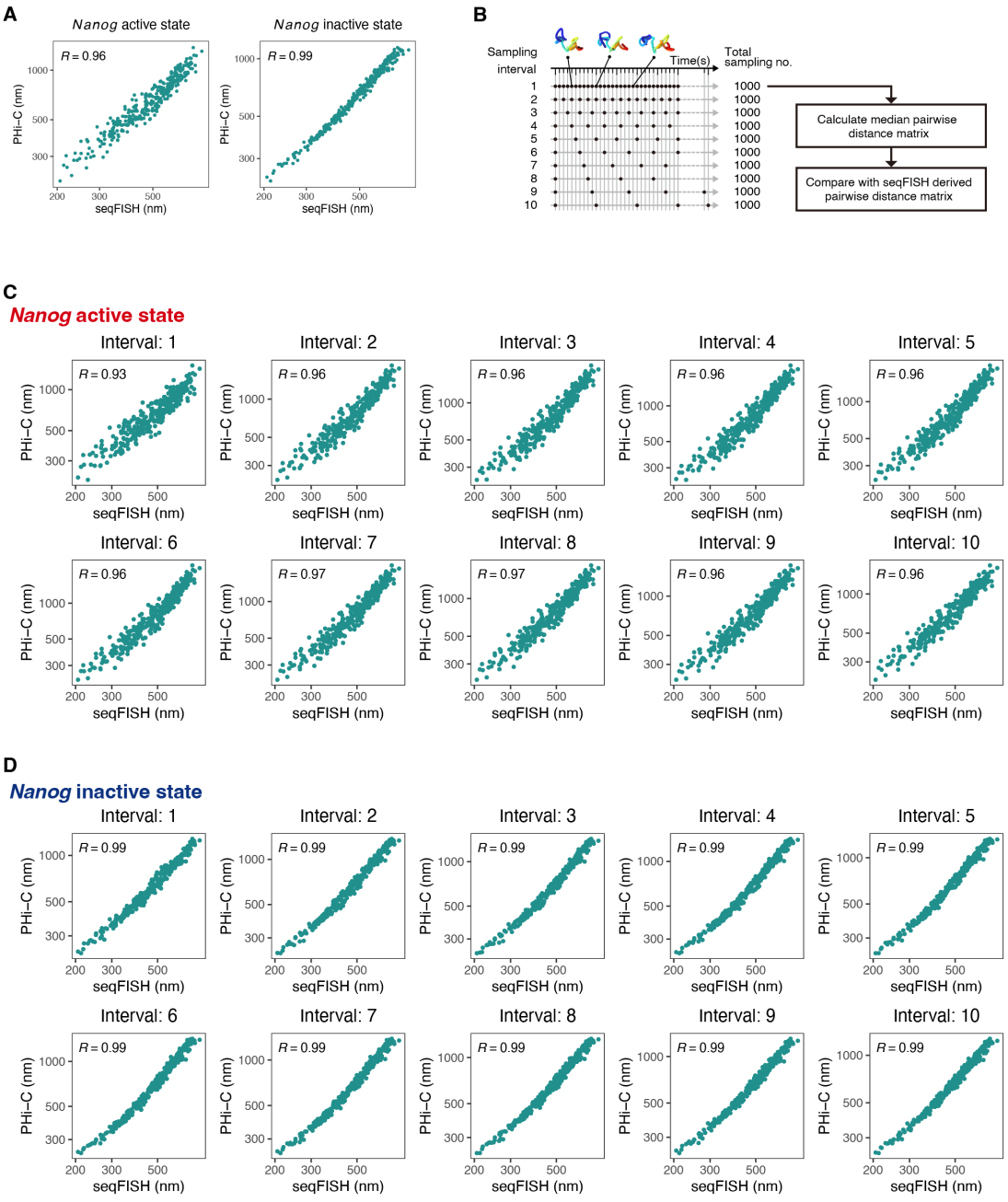


Fig. S12. Comparison of structures estimated by the PHi-C polymer model with seq-DNA-FISH pairwise distance data. (A) Comparison of median pairwise distances between seq-DNA-FISH data from mouse ES cells cultured under serum/LIF conditions and structures estimated using the PHi-C polymer model. A median distance matrix was calculated based on $N = 1000$ structures randomly extracted by the PHi-C polymer model. This was compared with the median distance matrix calculated from seq-DNA-FISH data, and Pearson correlation coefficients were calculated. (**B**) Extraction of time-series data at specific time intervals from the structures generated by the PHi-C polymer model. The PHi-C polymer model generated structures for 10,000 seconds at 1-second intervals from a specific initial structure. Subsequently, 1000 structures were extracted at 1 to 10-second intervals. Based on this, a median pairwise distance

matrix was calculated. (**C-D**) Comparison of time-series structural data from the PHi-C polymer model and seq-DNA-FISH structural data for the *Nanog* active state (C) or *Nanog* inactive state (D). Using the method described in (B), 1000 data points were obtained at intervals of 1 to 10 seconds from the time-series structural data of the PHi-C polymer model, and a median pairwise distance matrix was calculated. This was compared with the median pairwise distance matrix from seq-DNA-FISH data, and Pearson correlation coefficients were calculated.

Table S1. Oligonucleotide sequences used in this study.

Name	Sequence (5' - 3')
polyT LNA oligonucleotide	T+T T+TT +TT+T T+TT +TT+T TTTAACCGATACCGGCATCATCTTCCGATACCGGCATCA
seqDNA1F primer	ATGCGCTGCAACTGAGACCGA
seqDNA1R primer	CTCGACCAAGGCTGGCACAA
seqDNA2F primer	ATGCGCTGCAACTGAGACCGA
seqDNA2R primer	GAGCTAATACGACTCACTATAGCTCGACCAAGGCTGGCACAA
seqDNA_revT primer	Phos-ATGCGCTGCAACTGAGACCGA
seqRNA1F primer	CGGCAAACTACGCGAGCCATA
seqRNA1R primer	GATGATCGACCATGCAGAGCT
seqRNA2F primer	CGGCAAACTACGCGAGCCATA
seqRNA2R primer	GAGCTAATACGACTCACTATAGGGTAT
31-nt global ligation bridge	TCAGTTGCAGCGCATGCTCGACCAAGGCTGG
Δ-190_gPCR-F1 primer	CCTGAGGGAGGAACAAAGACTTGA
Δ-190_gPCR-R1 primer	TTTGCACGCTTCACATCTTAGCTC
Δ-190_gPCR-F2 primer	GACCTGAGGGAGGAACAAAGAC
Δ-190_gPCR-R2 primer	GGGATGGCTAAAAAGGTTACAC
Δ-190_gPCR-R3 primer	GCTCTACCTTTCCACTCCCCAATC
Δ-190_gPCR-F3 primer	TACCTGAGAATTCTCATTCTTCAG

Table S2. List of antibodies used in seq-IF-FISH in this study.

Antigen	Vender	Antibody name	Catalog no	RRID	Host
KDM1A (LSD1)	Abcam	Anti-KDM1/LSD1 antibody	ab17721	AB_443964	Rabbit
TCF3	R&D Systems	Mouse TCF-3/E2A Antibody	MAB7650	N/A	Rat
SIN3A	Genetex	Anti-SIN3A, Rabbit	GTX129156	AB_2885913	Rabbit
HDAC3	Genetex	Anti-HDAC3, Rabbit	GTX113303	AB_10721050	Rabbit
OTX2	Merk	Anti-OTX2 Antibody	AB9566-I	N/A	Rabbit
CHD4	Novus Biologicals	CHD4 Antibody (CHD4 3F2/4)	NBP2-50163	N/A	Mouse
H3K27ac	Clontech	Anti-Acetyl Histone H3 (Lys27), mouse monoclonal antibody	MA309B	N/A	Mouse
EP300	Millipore	Anti-p300 CT Antibody, clone RW128	05-257	AB_309670	Mouse
ESRRB	Perseus Proteomics	Anti human ERRbeta mouse monoclonal antibody	632-06161	AB_1142301	mouse
MED1	Abcam	Anti-TRAP220/MED1 antibody (ab64965)	ab64965	AB_669756	Rabbit
MED12	Bethyl Laboratories	MED12 Antibody, A300-774A	A300-774A	AB_306337	Rabbit
TBP	Abcam	Anti-TATA binding protein TBP antibody [1TBP18] - ChIP Grade	ab818	AB_306337	Rat
BRD4	Abcam	Anti-Brd4 antibody [EPR5150(2)] - BSA and Azide free (ab182446)	ab182446	AB_1951939	Rabbit
SMAD4	Genetex	Anti-SMAD4, Rabbit-Poly	GTX112980	AB_2633033	Rabbit
KMT2D (MLL4)	Thermo Fisher	KMT2D Recombinant Rabbit Monoclonal Antibody (9H13L17)	701869	AB_10734584	Rabbit
TET2	Proteintech	Anti TET2	21207-1-AP	AB_10716739	Rabbit
H3K4me3	Monoclonal Antibody Institute	Anti trimethyl Histone H3(Lys4), mouse monoclonal antibody	MABI0004	AB_2885993	Mouse
ARID1A	Genetex	Anti-ARID1A, Rabbit	GTX129433	AB_2150114	Rabbit
NANOG	Abcam	Anti-Nanog antibody	ab80892	AB_355110	Rabbit
SOX2	R&D Systems	Human/Mouse/Rat SOX2 Antibody, 50% glycerol (1 mg/ml)	AF2018	AB_355110	Goat

Table S3. List of public ChIP-seq data used in seq-IF-FISH quality check.

Target	URL
KDM1A	http://chip-atlas.org/view?id=SRX047141
TCF3	http://chip-atlas.org/view?id=SRX1080398
SIN3A	http://chip-atlas.org/view?id=SRX4403317
OTX2	http://chip-atlas.org/view?id=SRX499117
CHD4	https://chip-atlas.org/view?id=SRX2007463
H3K27ac	http://chip-atlas.org/view?id=SRX3141820
EERRB	http://chip-atlas.org/view?id=SRX4167130
MED12	http://chip-atlas.org/view?id=SRX022692
BRD4	http://chip-atlas.org/view?id=SRX3898502
SOX2	http://chip-atlas.org/view?id=SRX5023741
H3K9me3	http://chip-atlas.org/view?id=SRX305003

Movie S1. Representative images of seq-DNA/RNA/IF-FISH.

The movie is generated from maximum intensity projections of images obtained in each round using seq-DNA/RNA-IF-FISH. In each round, three types of both secondary and readout probe sets were used, enabling the acquisition of images across three channels. Each channel was utilized for observing the localization of a singular RNA, genomic region, protein, or post-translational modification. Another channel was designated for capturing images of DAPI-stained nuclei.

Movie S2. Genome dynamics in *Nanog* active state simulated by PHi-C.

The movie illustrates a simulation of the dynamic genomic structure of the *Nanog* active state, based on seq-DNA/RNA-FISH data and modeled using PHi-C. The left panel displays the genomic structural dynamics, while the right panel presents a real-time inter-region distance matrix corresponding to that structure. In the active state, the genome exhibits elevated viscosity, resulting in relatively slow motion.

Movie S3. Genome dynamics in *Nanog* inactive state simulated by PHi-C.

The movie illustrates a simulation of the dynamic genomic structure of the *Nanog* inactive state, leveraging seq-DNA/RNA-FISH data and modeled through PHi-C. The left panel reveals the dynamic genomic architecture, while the right panel displays a real-time inter-region distance matrix corresponding to that architecture. Characteristically, in the inactive state, the genomic viscosity is lower compared to the active state, resulting in relatively rapid motion.

Data S1. Oligonucleotides used in the evaluation of the fluorescent spot detection method in seq-RNA-FISH.

Data S2. Genomic regions targeted by seq-DNA-FISH analysis in this study.

Data S3. Complementary sequences of target-binding sites for primary probes used in the seq-DNA-FISH analysis in this study.

Data S4. Complementary sequences of target-binding sites for primary probes used in the seq-RNA-FISH analysis in this study.

Data S5. Spatial coordinates of each target region determined by seq-DNA-FISH analysis in this study.

The table displays the coordinates of each fluorescent spot determined through seq-DNA-FISH analysis. Each row in the table represents information about an individual fluorescent spot. The column headers indicate the following information: "replicate" refers to the replication experiment number; "Medium" specifies the culture conditions; "FOV" is the field of view

number; "IntraFOV_Cell_ID" represents the cell ID within each FOV; "Cell_ID" denotes the cell ID across the entire dataset; "Allele_ID" is the allele ID across the dataset; "IntraCellular_Allele_ID" refers to the allele ID within a single cell; "hyb" indicates the imaging round ID; "Locus ID" is the ID of the targeted genomic region; "x_nm" gives the x-coordinate (in nm) of the fluorescent spot; "y_nm" provides the y-coordinate (in nm); and "z_nm" notes the z-coordinate (in nm).

Data S6. Spatial coordinates of each transcription site determined by seq-RNA-FISH analysis in this study.

The table displays the coordinates of each transcription site determined through seq-RNA-FISH analysis. Each row in the table represents information about an individual transcription site. The column headers indicate the following information: "replicate" refers to the replication experiment number; "Medium" specifies the culture conditions; "FOV" is the field of view number; "IntraFOV_Cell_ID" represents the cell ID within each FOV; "Allele_ID" is the allele ID across the dataset; "IntraCellular_Allele_ID" refers to the allele ID within a single cell; "RNA_id" indicates the target RNA name; "z_px" notes the z-coordinate (in pixels) of the transcription site; "y_px" provides the y-coordinate (in pixels); "x_px" gives the x-coordinate (in pixels); "R_cluster_num" represents the estimated number of RNA contained in the transcription site; "Locus ID" is the ID of the targeted genomic region that includes the RNA target locus.

Data S7. Spatial coordinates of each H3K27ac focus determined by seq-IF-FISH analysis in this study.

The table displays the spatial coordinates of each H3K27ac focus as determined by seq-IF-FISH analysis. Each row in the table provides information about an individual focus. Data from replicate 1 are included in this table. The column headers indicate the following information: "FOV" refers to the field of view number; "IntraFOV_Cell_ID" represents the cell ID within each FOV; "Cell_ID" denotes the cell ID across the entire dataset; "Antibody_target" refers to the antibody target; "x_nm" provides the x-coordinate (in nm) of the focus; "y_nm" provides the y-coordinate (in nm); and "z_nm" provides the z-coordinate (in nm).