Visualization of loop extrusion by DNA nanoscale tracing in single human cells

Beckwith KS[† 1], Ødegård-Fougner Ø[† 1,2], Morero NR [1], Barton C [3#], Schueder F [4,5], Alexander S [1], Jungmann R [4,5], Birney E [3] and Ellenberg J* [1]

[†] Contributed equally to this work.

* Corresponding author.

Affiliations:

1 Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Heidelberg, Germany.

2 Department of Molecular Cell Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway.

3 European Bioinformatics Institute, European Molecular Biology Laboratory, Hinxton, United Kingdom.

4 Faculty of Physics and Center for Nanoscience, Ludwig Maximilian University, Munich, Germany.

5 Max Planck Institute of Biochemistry, Martinsried, Germany.

Present address:

Department of Computer Science and Information Systems, Birkbeck, University of London, London, WC1E 7HX, UK.

Summary

The spatial organization of the genome is essential for its functions, including gene expression, DNA replication and repair, as well as chromosome segregation[1]. Biomolecular condensates and loop extrusion have been proposed as the principal driving forces that underlie the formation of non-random structures such as chromatin compartments and topologically associating domains[2,3]. However, if the actual 3D-folding of DNA in single cells is consistent with these mechanisms has been difficult to address in situ. Here, we present LoopTrace, a FISH workflow for high-resolution reconstruction of 3D genome architecture without DNA denaturation. Classical fluorescence in situ hybridization approaches can link chromatin architecture to DNA sequence but disrupt chromatin structure at the critical nanoscale of individual loops. Our method conserves chromatin structure and can resolve the 3D-fold of chromosomal DNA with

better than 5-kb-resolution in single human cells. Our results show that the chromatin fiber behaves as a random coil that can be further structured in a manner consistent with loop formation, explaining the emergence of topologically associated domain-like features in cell population averages. Mining a large amount of single-cell data computationally, we reveal chromatin folding intermediates consistent with progressive loop extrusion and stabilized loops, highlighting the power of our method to visualize the nanoscale features of genome organization in situ.

Main

The spatial organization of the eukaryotic genome at hierarchical scales, from the assembly of nucleosomes to supra-chromosomal compartments, plays a central role in regulating genomic functions, and thus is tightly regulated during cell development and differentiation[1,4]. Biochemical crosslinking and DNA sequencing-based chromosome conformation capture methods (3C/Hi-C) have revealed compartmentalization at the megabase (Mb) scale and topologically-associating domains (TADs) of several hundred kilobases (kb) as conserved features of chromatin organization[3,5]. Boundaries of such TADs have been strongly linked to the position of CCCTC-binding factor (CTCF) binding sites[3] and the activity of the SMC protein cohesin[6,7], which has been shown to function as a molecular motor that can extrude DNA loops in vitro[8]. It has been proposed that in cells, chromatin-bound cohesin would promiscuously extrude loops in chromosomal DNA until encountering CTCF-binding sites[9]. Computational models that integrate affinity interactions such proposed loop extrusion activity into biophysical polymer models have indeed been able to simulate chromatin architectures consistent with Hi-C data[10]. However, direct evidence for the extrusion and structure of loops in single cells is lacking. Whereas genome-wide maps of contact frequencies by Hi-C average from large cell populations[11], single-cell Hi-C methods suffer from lower resolution and poor genomic sampling[12,13]. In general, biochemical crosslinking and DNA sequencing-based methods do not preserve the native physical structure of chromatin in situ.

By contrast, microscopy-based methods such as fluorescence in situ hybridization (FISH) in principle offer direct, single-cell visualization of genomic structures. The combination of the oligopaint technology, labeling many small genomic loci, and super-resolution microscopy techniques enabled the visualization of genomic features with genomic resolution down to several kb[14–24]. To explore the limits of the obtainable genomic resolution, we performed in vitro oligoDNA-PAINT on linearized single-stranded M13 bacteriophage DNA, showing that individual sub-kb targets (64 bp) apart can be resolved (Supp. Fig. S1A,B). In human cells however, the success

of oligopaint labeling relies on opening the genomic DNA double helix to ensure efficient hybridization of a small number of short FISH probes at each target locus. In most currently employed methods, DNA denaturation is achieved by rather harsh conditions, including hydrochloric acid treatment and heating to high temperatures (80-95°C for 3-10 min)[18,19,21–23]. This results in significant perturbations of nuclear morphology and chromatin structure[25,26 27], including the loss of fine details below 1 Mb as well as large scale DNA redistribution such as loss of condensed chromatin regions[27], nuclear shrinkage and leakage of nuclear DNA into the cytoplasm[16]. Standard FISH protocols thus undermine the value of the high-resolution structural information that could be gained by oligopaint labeling and subdiffraction imaging.

To investigate the nanoscale 3D-structure of DNA in situ, we developed a high-resolution chromatin tracing workflow extending previous work on non-denaturing FISH after strand-specific enzymatic digestion of DNA[28–30] combined with state-of-the-art oligopaint FISH probes. Using this new method, which we term "LoopTrace", we systematically investigate the 3D-folding of chromatin near the proto-oncogene MYC from the few kb to Mb-scale in single human diploid cells. We demonstrate that our approach offers superior preservation of nuclear and chromatin structure and yet enables efficient probe hybridization and thus very high-resolution 3D chromatin tracing at the scale of single loops. Our 3D traces from many hundreds of single cells show that in the absence of structuring elements such as CCCTC, the 3D-fold of the chromatin fiber is fully consistent with a random coil. Mining our data for structural similarity between single-cell-folds using non-supervised and supervised clustering approaches revealed that genomic regions containing CTCF-binding sites display structures fully consistent with dynamic loop extrusion, which explains the emergence of TAD-like features at the population level when averaging data from many individual cells.

Non-denaturing FISH preserves nuclear structure

Individual DNA loops constituting Mb-scale TADs are predicted to be on the 100-kb-scale. To resolve them, chromosomal DNA therefore needs to be traced at a very high genomic resolution (<10 kb), which requires high-efficiency hybridization of the primary FISH probes to the genomic DNA. Harsh denaturing conditions of current tracing protocols cause significant structural perturbations of nuclear architecture at the micrometer scale[25] that would massively distort or even destroy the native structure of single loops inside an approximately 150-nm-diameter TAD[31].

To overcome this limitation, we set out to quantitatively compare state-of-the-art oligopaint FISH protocols regarding hybridization efficiency and preservation of nuclear structure. We imaged nuclei of diploid human cells (retinal pigment epithelial cells, RPE-1) at each step in the protocol using structured illumination microscopy (SIM). We visualized general genome structure with the

DNA-binding dye Hoechst as well as the TAD-level structure of co-replicating domains using co-replicatively incorporated fluorescent dUTPs[31]. While the initial fixation and permeabilization steps did not substantially alter neither genome nor TAD level structure (Supp. Fig. S1C,D), the standard denaturation steps with hydrochloric acid and/or denaturation at 86°C[19,22–24] led to significant structural changes at both levels (Fig. 1A,D,E, Supp. Fig. S1E,F), including fragmentation of heterochromatin areas, loss of integrity at the nuclear rim, spilling of DNA into the cytoplasm (Fig. 1A, Supp. Fig. S1F), and displacement/loss of co-replicating domains. Although the denaturing protocol achieved robust FISH signal (Fig. 1F), we therefore did not consider it further for nanoscale tracing of single DNA loops. A lower temperature (75°C) denaturing FISH protocol developed previously (3D-FISH)[27] showed milder but still significant perturbations of genome architecture (Fig. 1B,D,E; Supp. Fig. S1G), and the hybridization efficiency measured by FISH-signal intensity was already significantly reduced (Fig. 1F).

As further lowering of hybridization temperature diminished the FISH signal to unacceptable levels, we next compared an alternative FISH approach that has not been used for genome tracing so far[29]. Strand excision FISH (SE-FISH) uses enzymatic strand excision to render long stretches of genomic DNA single-stranded and thus accessible for hybridization without denaturation. After optimizing for hybridization efficiency (Supp. Fig. S1I), SE-FISH preserved overall genome structure substantially better than the heat/acid-denaturing FISH protocols, while maintaining probe hybridization efficiency superior to 3D-FISH (Fig. 1C-F, Supp. Fig. S1F-H).

Consequently, non-denaturing SE-FISH combines high-efficiency binding of primary oligopaint probes to genomic DNA with near-native preservation of nuclear structure at the whole genome and TAD level, which makes it a suitable method to investigate the folding of chromosomal DNA at the nanoscale in situ in human cells.

Precise automated 3D-chromatin tracing at single-loop-scale

Having a non-denaturing probe-binding protocol at hand, we next developed a robust sequential DNA-FISH imaging workflow (LoopTrace) to achieve high-resolution 3D-chromatin tracing at the scale of single DNA loops. We chose a non-coding euchromatic genomic region upstream of the MYC gene on the right arm of chromosome 8 as target (Fig. 2A), predicted to have little conserved structure (Supp. Fig. 2A) and containing no CTCF-binding site guiding loop extrusion (Fig. 2A). We targeted ten discrete 5-kb-loci spaced 10 kb apart with a set of oligopaint probes containing loci-specific barcodes. The primary probes were detected by sequentially hybridizing and imaging ten different fluorescent 12-bp imager strands matching each locus barcode[32]. The fully automated workflow (see methods), and relatively rapid hybridization and washes made the approach

scalable; e.g. imaging 200 cells in 20 fields of view with 13 sequential hybridizations (ten loci plus internal standards) required only ~4 h, making the method scalable to over 1000 cells/day.

We found that our workflow produced reliable FISH signal and completely removed signal from the previous locus with the washing step, enabling reliable reconstruction of chromatin paths (Fig. 2B-D, Supp. Fig. S2B). To estimate the tracing precision we could achieve, we re-localized the first locus after ten rounds of hybridization, having continuously drift-corrected the data using fluorescent beads in a separate channel (Fig. 2E,F). The bead-based drift-correction had a median 3D-distance error of <15 nm as determined by 3D-Gaussian fitting. Re-localization of the first locus showed a median 3D-distance deviation of <25 nm (Fig. 2F), or 7, 8 and 16-nm deviations along x, y and z, respectively (Supp. Fig. S2C), demonstrating that our approach is suitable for high-precision 3D-chromatin tracing well below the diffraction limit. LoopTrace was also highly efficient, as at least 90% of sequential hybridizations were detected in the majority of all cells imaged (Supp. Fig. S2D,E).

A loop-scale region upstream of the MYC promoter behaves like a random coil

Comparing the 3D-traces of the 100-kb-region upstream of the MYC promoter in individual cells showed substantial cell-to-cell variability of chromatin structure (Fig. 2G). Combining the data from hundreds of cells into consensus chromatin traces by general Procrustes analysis (see methods) or pairwise median distance maps of all loci showed, highly reproducibly, no clear substructure (Fig. 2H; Supp. Fig. S2F, average Pearson's correlation coefficient r = 0.99 between median distance matrices).

As a polymer, the chromatin fiber is predicted to fold like a random coil in the absence of structuring elements. We therefore compared our experimental data of this region with the 3D-fold calculated from a random coil model of chromatin with an equidistant step size of 90 nm, matching the median physical distance of our FISH loci (Fig. 2H). Comparing the consensus traces or pairwise distance maps between experiment and model showed almost perfect agreement (Fig. 2H,I) and only little residual deviation could be detected by a z-score normalized difference, or by comparing radii of gyration or pairwise similarity of 3D-aligned traces with the random coils (Fig. 2J, Supp. Fig. S2G,H). Interestingly, our experimental data allowed us to determine the relationship between genomic and physical distance for this non-coding region of euchromatin. Again, as expected for a random coil, we found that the spatial distance scaled with the square root of the genomic distance in this region with a scaling parameter of $24.7\,nm/kb^{\frac{1}{2}}$ (r^2 = 0.94), a very valuable parameter to constrain theoretical models of chromatin folding.

The TAD-scale regulatory region upstream of the MYC gene shows non-random substructures

Given that our method could capture the highly variable and therefore likely dynamic structures of a relatively short stretch of euchromatin, we tested next if it can also resolve TAD boundaries and loops predicted by biochemical approaches. To this end, we probed a larger, TAD-scale euchromatic region spanning the regulatory domain upstream of the MYC gene with ten probes situated every 50-200 kb across this 1-Mb-region (Fig. 3A, Supp. Fig. S3A). This region is predicted to contain a number of clusters of frequently occupied CTCF binding sites, the three most prominent we targeted directly with our probes (Fig. 3A, H2, H4, H9).

Sequential imager strand exchange and high-resolution imaging of the ten loci (H1-10) resulted in 3D-localization data of very similar quality as for the shorter region (Fig. 3B, C), with an overall median 3D-tracing precision of ~25 nm. 3D-traces from single cells showed clear indications of loop-like features (Fig. 3B,C). Integrating the 3D traces from ~800 cells into consensus traces or median pairwise distance maps further confirmed the presence of looping conformations distinct from the random coil model (Fig. 3D), and could clearly identify the TAD boundaries (Supp. Fig. S3B) observed in population Hi-C experiments (e.g. at H2), validating our high-resolution 3D-imaging approach.

In contrast to the smaller non-structured region studied previously, the median distance matrix of the TAD-scale region deviated substantially from the predictions of the random coil model with strong residual z-score differences (Supp. Fig. S3C-D, Pearson's r=0.89) and reduced pairwise Pearson correlation (Supp. Fig. S3E). The significant non-randomness of our experimental traces is consistent with the idea of an active structuring process. We therefore hypothesized that a 3D-similarity analysis of all our single-cell traces might reveal intermediate folds produced by the structuring process. To test this, we clustered the individual traces based on their pairwise similarity after 3D-alignment (Fig. 3E). The consensus traces for each cluster visualized clear differences in the structural features, including random coil-like folds in cluster 1, intra-TAD compaction in cluster 2 and long-range looping in cluster 3 (Fig. 3E). These structural features were confirmed by median distance matrices, which especially highlighted the increased proximity between the predicted CTCF-bound loci H2, H4 and H9 in cluster 3 (Supp. Fig. S3F).

Unsupervised clustering of single-cell traces reveals common structural features of 3D-folds

The large number of high-resolution 3D-snapshots from single cells should allow us to obtain insight into the cell-to-cell variability of loops and more compact structures and potentially

reconstruct the dynamic process by which they form. Each of our three structurally different primary clusters contains several hundreds of individual single-cell traces. To explore substructure in the data, we plotted the similarity of all single-cell 3D-traces after reducing dimensionality by multidimensional scaling (MDS). The traces of the three primary clusters combined mapped to different regions of the MDS plot (Supp. Fig. S3I) validating this approach of exploring the data. Therefore, we performed MDS for the two clusters indicative of looped or compact structures, assigning colors to the two secondary clusters at the next tree hierarchy. This MDS also showed distinguishable secondary clusters, albeit with significant overlap at their boundaries, suggesting a continuum of dynamic structures rather than sharp boundaries between different stable conformations (Supp. Fig. S3J-K).

We thus further explored these secondary clusters structurally to see if we could observe distinct conformations of looping within primary cluster 3 or compaction in primary cluster 2 (Fig. 3F,G; Supp. Fig. S3G,H). For cluster 3, the consensus 3D-traces as well as the median distance maps showed that both secondary clusters contained its signature loop between H2 (or H4) and H9. Rather, the conformation of the most upstream part of the MYC regulatory region at position H1 either maintained larger distances from H2 (Fig. 3G, subcluster 3-1), or showed more variable and shorter distances to downstream neighbors (Fig. 3G, subcluster 3-2), which led to blurring of the TAD boundary (Supp. Fig. S3H). In cluster 2, the strong TAD-boundary between position H1 and H2 was present in both secondary clusters (Supp. Fig. S3G). The secondary clusters mainly differed further downstream at the beginning of the MYC gene (H9/H10), which was either completely excluded from the fold of the upstream region (Fig. 3F, subcluster 2-1), or in fact rather close to it (Fig. 3F, subcluster 2-2).

As the MDS plots had already indicated (Supp. Fig. S3J,K), the single-cell traces even in secondary clusters showed significant variability. Examining single 3D-traces representative of the clusters (Fig. 3E) highlights that all traces have common structural features such as a loop, explaining their clustering. Other parts of the chromatin fiber however were different in every cell and appear to behave rather as a random coil, explaining the broad distribution of traces within subclusters and the overlaps between them.

Supervised clustering reveals a series of 3D-folds consistent with progressive loop extrusion

Unsupervised clustering was sufficiently sensitive to detect the major structural features in our single-cell 3D-traces, but might perform poorly on the specific variability of a particular structural feature, such as single loops. In principle our high-resolution single-cell 3D-traces should contain

such information. Extracting it quantitatively could provide novel insights into the hypothesized process of dynamic loop formation, which is not possible with non-spatial approaches. We therefore decided to mine our 3D-single-cell data deeper using a supervised approach. We focused on the major long-range interaction identified by our unsupervised analysis, i.e. between the two predicted CTCF-bound loci H2 and H9. To this end, we sorted all our traces by a distance criterion of ≤150 nm in 3D between H2 and each of the downstream loci, from the direct neighbor H3 to the last CTCF-bound loci H9. Each of the generated seven clusters contained 50-100 single-cell traces. Their corresponding consensus traces indeed provide clear evidence for the existence of looping interactions anchored in H2 that proceed to more and more downstream loci, which is fully consistent with progressive loop extrusion (Fig. 3H, Supp. Fig. S3L).

Encouraged by this, we asked if our data contained evidence for stable intermediates of loop extrusion, which would be expected to be more abundant. To hypothesize such intermediates, we examined available ChIP-seq data for our RPE cells, in which a total of 11 CTCF binding sites have been mapped in the regulatory region 1 Mb upstream of the MYC gene that our FISH probes overlap with (Fig. 3A). The upstream probe H2 overlaps with five, probe H4 with two and probe H9 with four CTCF-binding sites, respectively. This would predict that the larger H2-H9 loop should be more stable than the H4-H9 loop, which would rely on fewer CTCF sites for upstream anchoring. Our unsupervised clustering picked up both loops (Fig. 3G), but could not quantify their abundance. By contrast, our supervised analysis (Fig. 3H) now showed the largest H2-H9 loop, bounded at both ends by strong CTCF-binding clusters, indeed occurred significantly more frequently (15% of cells) than the H4-H9 (9% of cells) (Supp. Fig. S3M). The less stable H4-H9 loop occurred at frequencies similar to the other intermediate looping interactions anchored at H2 (Fig. 3H, 8-10% of cells). As a control, we performed the same supervised clustering in traces simulated by a random coil model. In a random coil-fold, only 0.9% of traces had a distance <150 nm between the second and ninth positions (corresponding to H2-H9), indicating that reproducible loops are extremely rare in non-structured fibers. Our data is thus fully consistent with a progressive loop extrusion model from H2 towards H9 as well as with the available CTCF-site occupancy data, which illustrates the power of single-cell 3D genomic DNA tracing.

Discussion

Here we show that an enzymatic strand excision strategy offers superior preservation of chromatin structure compared to denaturing FISH methods, while providing a sufficiently robust FISH signal to allow reliable high-resolution 3D-chromatin tracing at the nanoscale. We demonstrate that the LoopTrace method can resolve the physical structure of the DNA at a genomic precision of ~3-5 kb

and a spatial precision of ~25 nm in single cells. The obtainable genomic resolution will be limited by unique primary probe density, which varies within the genome[33], and in our experience at least ~50 probes are necessary for a high-precision localization of a locus. We note that even the non-denaturing SE-FISH protocol still shows some alterations of general nuclear structure, such as slight swelling of nucleoli and moderate changes in nuclear shape. Hopefully in the future even more native methods of making genomic DNA available for efficient probe binding can be developed.

We have previously computationally estimated the degree of multiplexing (≥10 sequential DNA-PAINT "colors") and physical spatial precision (≥30 nm for each locus) that an oligopaint approach would need to achieve, in order to trace chromosome-scale DNA molecules in the nucleus with a genomic resolution ≤10 kb[34]. In this study we achieved the required spatial precision and validated 20 colors with distinct 12-nt-long imager strands that can be scaled to allow chromatin tracing of whole chromosomes (≥10 colors) or the whole genome (≥20 colors)[34]. It should be noted however, that an increased spatial multiplexing, which is desirable to maintain reasonable throughput in sequential locus detection for genome-scale libraries, would likely require even higher optical precision to distinguish physically close loci imaged simultaneously.

In LoopTrace we use 12-mer imager strands with an intermediate binding strength to allow relatively rapid cycles of hybridization and imaging without strand displacement or chemical cleavage/bleaching[22,23] and thereby good throughput and robustness. This makes our approach scalable, e.g. to compare different biological conditions and/or increase the number of "colors" in sequential hybridizations, e.g. ~40 hybridizations in ~300 cells is possible in a single overnight experiment. 12-mers require direct labeling and are not compatible with a bridging approach to a universal imager strand barcode[22,23], which is more cost effective. Given the small amount of reagents used in single cell imaging experiments and considering the declining cost of DNA synthesis, the gain in robustness outweighs the minor cost increase, in our opinion.

The results of the chromatin tracing we present here are in line with recent polymer physics predictions[10,35] that chromatin structure follows a random coil in the absence of structuring elements such as CTCF or engaged enhancer/promoter sites that confine the random coil. Importantly, when investigating regions that do contain such sites, we identify TAD boundaries that are predicted from Chip-seq and Hi-C data, demonstrating that our imaging approach recapitulates chromatin architecture seen with orthogonal approaches. Uniquely, our high-resolution single-cell traces provide direct measurements of the in situ 3D-fold and allowed us to explore the heterogeneity of single-cell structures. Their unbiased clustering revealed distinct groups with recurring structural conformations such as highly compact and looping features.

Maybe most excitingly, our supervised clustering of single-cell 3D-DNA folds retrieved features consistent with progressive loop extrusion bounded by the two strongest CTCF-binding sites in the region we examined, which are highly unlikely in a random coil. Beyond ~100 kb genomic distance, loop extrusion only minimally compacted neighboring regions, as exemplified by comparing H2-H8 distances in the H2-H9 loop cluster (Supp. Fig. S3N). Nearly half of all the chromatin traces contained loops anchored in the strongest upstream site H2, but double CTCF-site-bounded loops (e.g. H2-H9) were only ~2-fold more abundant than single CTCF-site-bounded loops (e.g. H2-H8) that most likely reflect active extrusion. These results are consistent with reports that once bound to CTCF on chromatin, cohesin has a long residence time and is predicted to act progressively[36].

We anticipate that the combination of a near-native FISH protocol and scalable high-resolution chromatin tracing we present here will be a valuable tool to further directly investigate the structure-function relationship of the genome at the nanoscale in single cells. Scaled to whole chromosomes and the whole genome, such methods will enable a deep understanding of the link between genome architecture and the functional state of individual cells in health and disease.

References

1. Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* **164**, 1110–1121 (2016).
2. Shin, Y. *et al.* Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. *Cell* **175**, 1481-1491.e13 (2018).
3. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
4. Gorkin, D. U., Leung, D. & Ren, B. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* **14**, 762–775 (2014).
5. Lieberman-Aiden, E. *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**, 289–293 (2009).
6. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* vol. 171 (2017).
7. Wendt, K. S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
8. Davidson, I. F. *et al.* DNA loop extrusion by human cohesin. *Science* **366**, 1338–1345 (2019).
9. Hansen, A. S. CTCF as a boundary factor for cohesin-mediated loop extrusion: evidence for a multi-step mechanism. *Nucleus* **11**, 132–148 (2020).
10. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci.* **115**, E6697–E6706 (2018).
11. Sati, S. & Cavalli, G. Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* **126**, 33–44 (2017).
12. Nagano, T. *et al.* Single cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, (2013).
13. Stevens, T. J. *et al.* 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017).
14. Beliveau, B. J. *et al.* Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci.* **109**, 21301–21306 (2012).
15. Beliveau, B. J. *et al.* Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, (2015).
16. Markaki, Y. *et al.* The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture. *BioEssays* **34**, 412–426 (2012).
17. Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
18. Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).
19. Bintu, B. *et al.* Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, (2018).
20. Nir, G. *et al.* Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLOS Genet.* **14**, e1007872 (2018).
21. Cardozo Gizzi, A. M. *et al.* Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization of Genome Organization and Transcription in Intact Organisms. *Mol. Cell* **74**, 212-222.e5 (2019).
22. Mateo, L. J. *et al.* Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
23. Su, J.-H., Zheng, P., Kinrot, S. S., Bintu, B. & Zhuang, X. Genome-Scale Imaging of the 3D Organization and Transcriptional Activity of Chromatin. *Cell* **182**, 1641-1659.e26 (2020).
24. Takei, Y. *et al.* Integrated spatial genomics reveals global architecture of single nuclei. *Nature* **590**, 344–350 (2021).

25. Robinett, C. C. *et al.* In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J. Cell Biol.* **135**, 1685–1700 (1996).
26. Mongelard, F., Vourc'h, C., Robert-Nicoud, M. & Usson, Y. Quantitative assessment of the alteration of chromatin during the course of FISH procedures. *Cytometry* **36**, 96–101 (1999).
27. Solovei, I. *et al.* Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). *Exp. Cell Res.* **276**, 10–23 (2002).
28. Goodwin, E. & Meyne, J. Strand-specific FISH reveals orientation of chromosome 18 alphoid DNA. *Cytogenet. Genome Res.* **63**, 126–127 (1993).
29. Brown, J. M. *et al.* A tissue-specific self-interacting chromatin domain forms independently of enhancer-promoter interactions. *Nat. Commun.* **9**, 3849 (2018).
30. Miron, E. *et al.* Chromatin arranges in chains of mesoscale domains with nanoscale functional topography independent of cohesin. *Sci. Adv.* **6**, eaba8811 (2020).
31. Xiang, W. *et al.* Correlative live and super-resolution imaging reveals the dynamic structure of replication domains. *J. Cell Biol.* **217**, 1973–1984 (2018).
32. Schueder, F. *et al.* Universal Super-Resolution Multiplexing by DNA Exchange. *Angew. Chem. Int. Ed.* **56**, 4052–4055 (2017).
33. Beliveau, B. J. *et al.* OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proc. Natl. Acad. Sci.* **115**, E2183–E2192 (2018).
34. Barton, C. *et al.* ChromoTrace: Computational reconstruction of 3D chromosome configurations for super-resolution microscopy. *PLOS Comput. Biol.* **14**, e1006002 (2018).
35. Giorgetti, L. *et al.* Predictive Polymer Modeling Reveals Coupled Fluctuations in Chromosome Conformation and Transcription. *Cell* **157**, 950–963 (2014).
36. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).

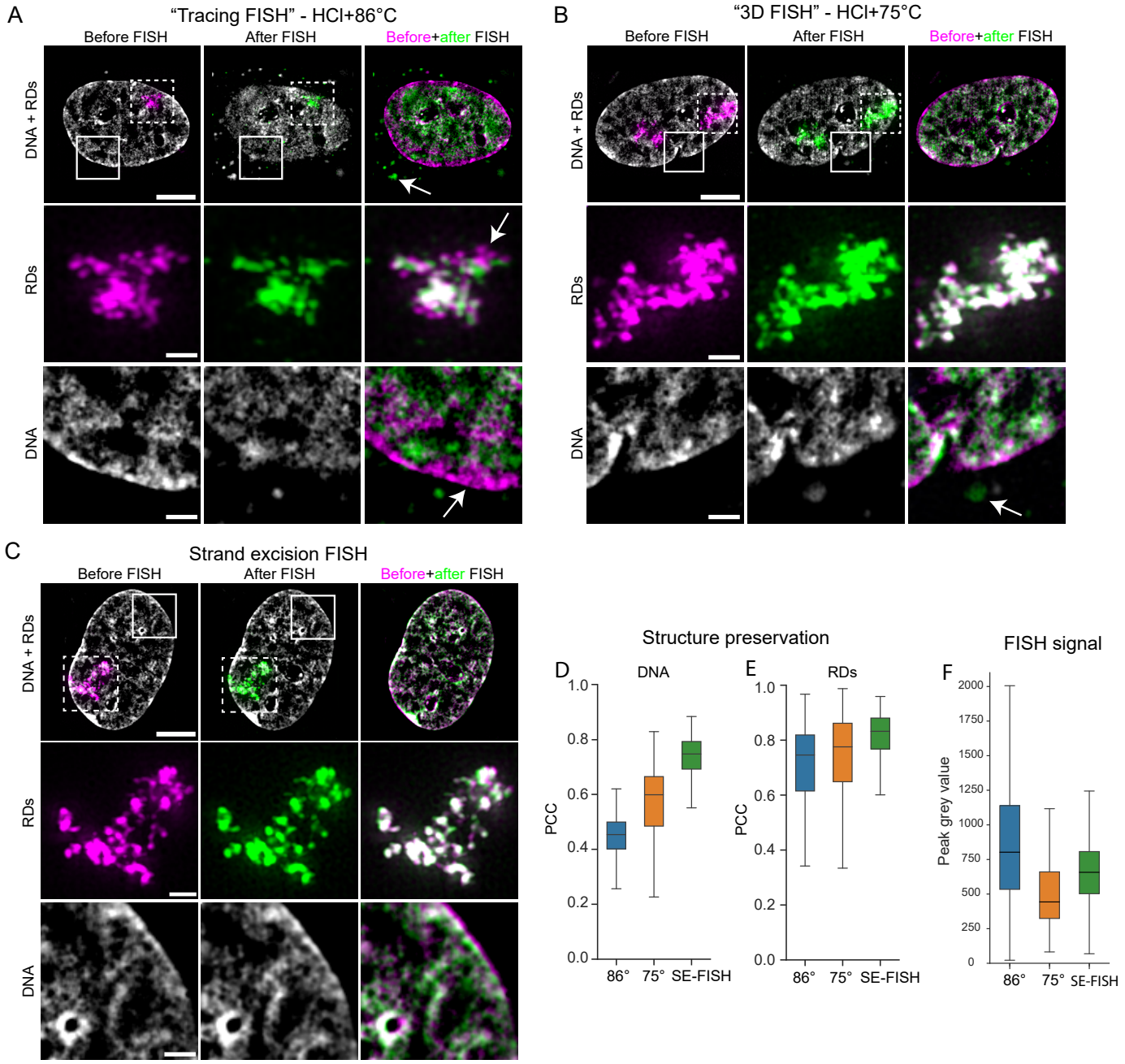
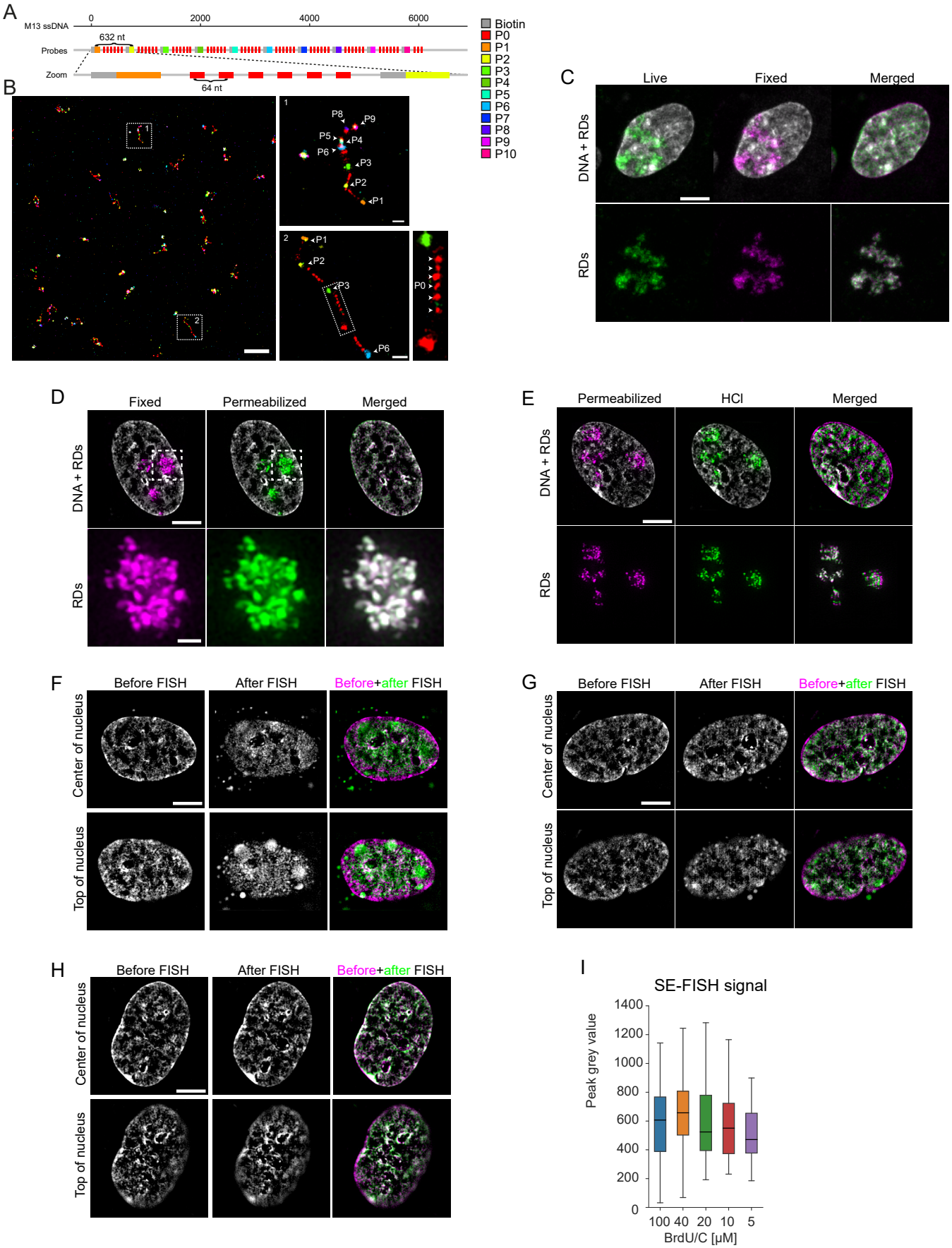


Figure 1: Chromatin preservation and FISH signal after different FISH procedures.
Structural integrity of chromatin during FISH procedures of cells labeled with Hoechst (DNA, grey) and Atto647-dUTP (replication domains, RDs, magenta/green). (A) "Tracing" FISH protocol with acid treatment and 86°C denaturation. (B) "3D FISH" protocol with acid treatment and 75°C denaturation. (C) SE-FISH protocol with enzymatic strand digestion. Images in A-C representative of >350 cells in 2 independent experiments. (D-E) Structural integrity of DNA (D) and RD (E) signal assessed by Pearson's correlation of drift-corrected images before and after FISH. Data from n>350 cells per condition from two independent experiments. (F) Peak intensity of FISH signal from 96 oligo probes targeting a single-copy region in the MYC locus with FISH protocols indicated. Data from n>50 cells per condition, representative of 2 independent experiments. Scale bars, 5 µm (DNA), and 1 µm (RDs/DNA zoom).



Supplementary Figure S1: Related to main Figure 1.

DNA (Hoechst, grey) and AF647-dUTP replication domain (magenta/green) preservation during different steps of the FISH protocols. (A) Overview of linearized single-stranded M13 bacteriophage DNA with biotin-associated docking handles (grey), 64-bp-spaced docking handles (P0, red) and ~600 bp spaced docking handles (P1-P10) in orange to magenta color scale. (B) 11-color-DNA-Exchange-PAINT image overview and zoom of indicated regions. Scale 500 nm for image, and 50 nm for inserts. (C-E) Comparison of preservation of chromatin structure during individual steps of the FISH protocol rom live to fixed cells (A), from fixed to permeabilized cells (B) and from permeabilization to acid treatment (C). (F-H) Comparison of central and apical planes of nucleus (DNA stained with Hoechst) before and after FISH treatment using “tracing FISH” with acid treatment and 86°C de-

naturation (F), “3D FISH” with acid and 75°C denaturation (G) and SE-FISH (H). Scale bars, 5 µm; RDs in (D), 1 µm. Images representative of >10 (A) or >50 cells (B-F) in two independent experiments. (G) FISH signal with different total BrdU + BrdC concentrations (3:1 ratio) used for SE-FISH. Data from n>50 cells per condition, representative of two independent experiments.

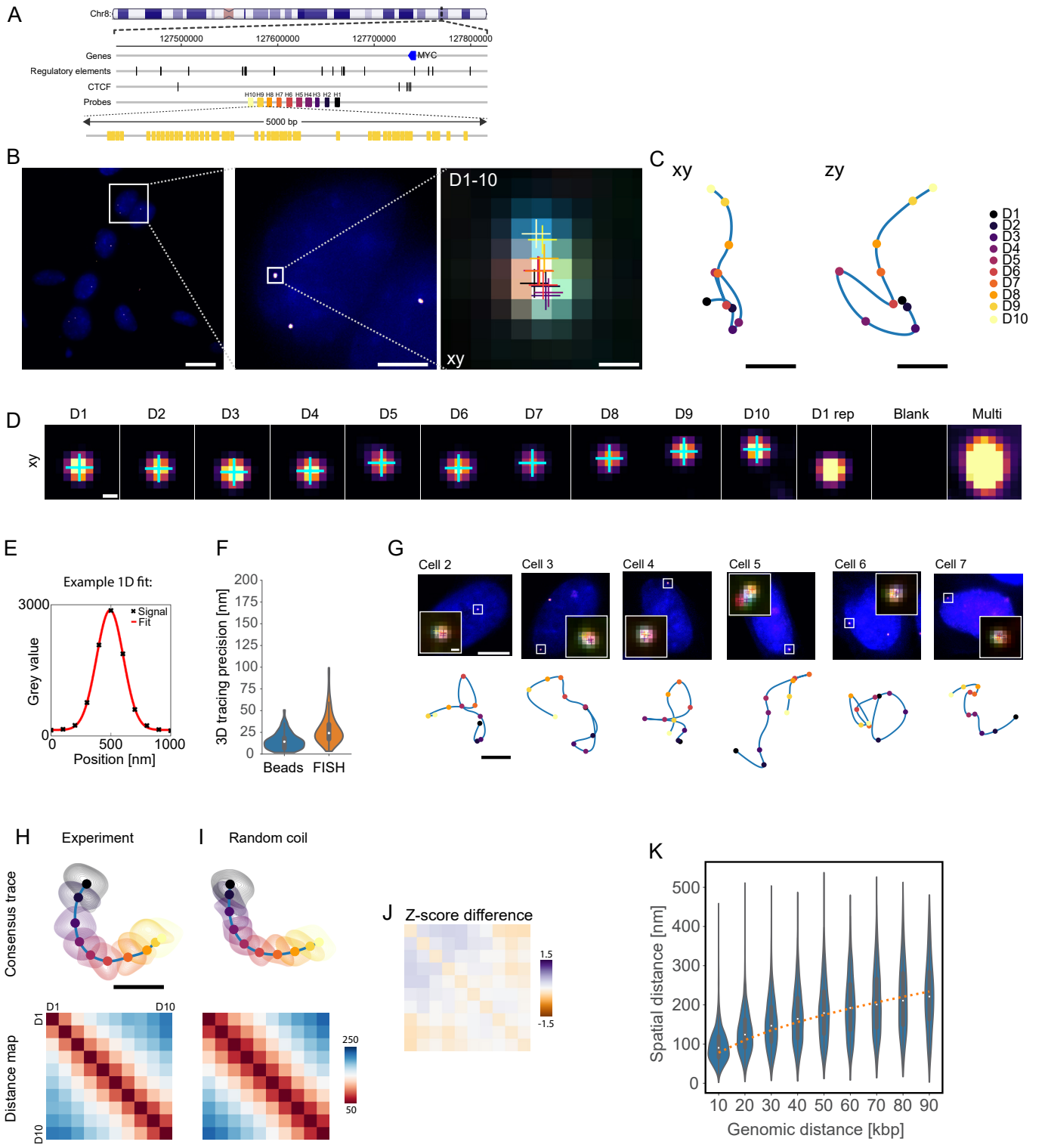
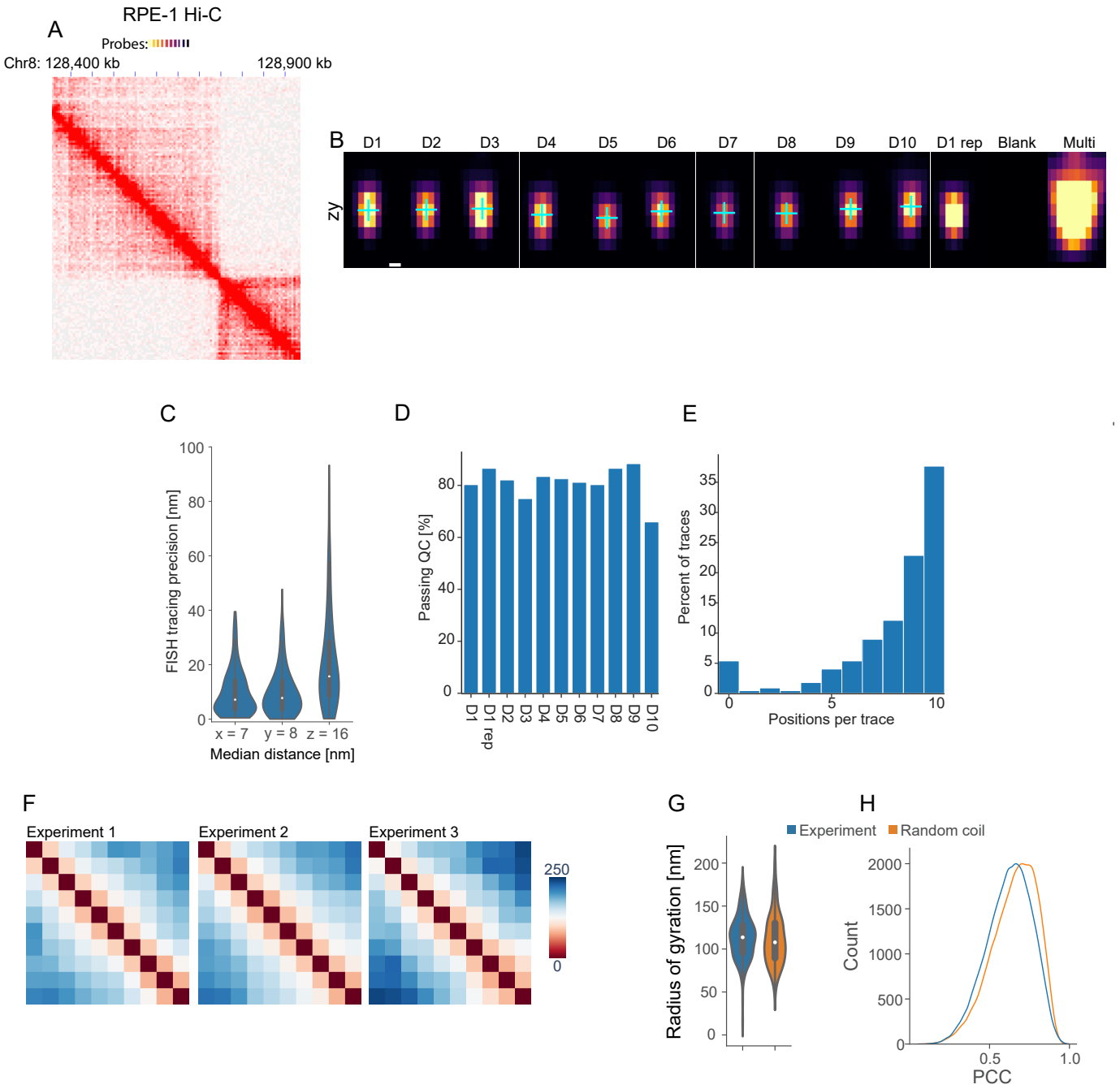


Figure 2: Performance of chromatin tracing pipeline in loop-scale region in MYC locus.
(A) Overview of a non-coding region upstream of MYC gene selected for tracing with regulatory elements (FANTOM5) and CTCF (RPE Chip-seq) highlighted. The probes cover 10 regions of 5 kb with 5-kb-gaps, spanning a total of 95 kb. (B) Exemplary field of view and selection of one tracing region by FISH signal (orange) in a single nucleus (Hoechst, blue), an overlay of probe signals (maximum projection in z) after drift correction and centroid fitting (multi-colored crosses indicate fit positions). Scale bars, 20 µm, 5 µm and 200 nm. (C) 2D projections of fit FISH signal coordinates in xy or zy direction of the data shown in (B). Scale bars, 100 nm. (D) Individual FISH signals at the central z-plane of each fit from sequential hybridization of the 10 probed regions (D1-D10), as well as repeat labeling (D1 rep), wash buffer only (Blank) or simultaneous labeling with multiple probes for region identification (Multi). Fit centroid positions are indicated by blue crosses. Scale bar, 200 nm. (E) Example of 1D-signal (black) from a FISH spot, and a 1D-Gaussian fit (red line) to the signal that is used for centroid fitting. (F) Tracing precision judged by 3D-Euclidean distance across 10 hybridizations, measured on fiducial beads (n=66, median=14 nm) or by re-hybridizing the first position after tracing (n=161 traces, median=24 nm). Data from one representative of 8 experiments. (G) Examples of FISH signal, centroid fitting projection and resulting 2D-projected (see methods) chromatin traces from 6 different cells. Scale bars, 5 µm (cells), 200 nm (insets) and 100 nm (traces). (H) 2D-projected consensus traces of D1-10 probe positions generated by general Procrustes analysis, overlaid with kernel density estimates (KDE) of aligned single trace data (see methods), and median pairwise distance map for probe positions D1-10. Data from n=469 and n=523 traces (filtered for at least 70% or 50% completeness, respectively,

see methods for details) from 3 experiments. (I) 2D-projected consensus trace with KDE overlay and median distance matrix of random coil model with 10 equally spaced positions scaled by experimentally determined parameters (see methods). Data from n = 469 and n = 523 simulated traces. Scale bars, 100 nm (H,I), colormap scale is in [nm]. (J) The difference between the experimental and random coil distance matrixes after z-score normalization of the data, highlighting distances which are smaller (orange) or larger (purple) than expected compared to the model. (K) Scaling relation between the spatial distance measured between all pairs of probes and the genomic distance of probe hybridization targets. The data was fit to a square root dependence (orange dashed line). Data from n=523 traces in 3 experiments.



Supplementary Figure S2: Related to main Figure 2.
(A) Hi-C contact map of RPE-1 cells at 5-kb-binning 37 of the traced region in the vicinity of the MYC gene, with probe positions as indicated. (B) Slices along the zy axis at the central x position with fit positions overlaid for an exemplary D1-10 sequential hybridization, corresponding to main Fig. 2B. (C) Tracing precision in x, y and z estimated by relabeling the first position (D1) after completing D1-10 hybridizations, with median distances as indicated. Data from n=161 traces from one representative experiment. (D) The percentage of hybridizations detected and passing quality control (QC) in all traces. (E) Percentage of traces with indicated number of hybridizations detected and passing QC. Data in (D) and (E) from 223 traces in one representative experiment. (F) Distance maps for replicate experiments tracing the same region as in Figure 2. Experimental replicates show highly similar distance matrices (average Pearson's correlation of 0.99). n=138, 203 and 182 traces for experiments 1, 2, and 3, respectively. Scale bar in nm. (G) Distribution of the radius of gyration of experimental traces and traces from the random coil model. No difference was detected (two-sided Mann-Whitney test, p=0.11). n=220 traces per sample. Only full-length traces were included from the experiments to ensure that the radii of gyration are comparable. (H) Kernel density estimates of all pairwise Pearson' correlation values of traces aligned by 3D rigid registration in the experimental dataset and an equal number of random coil models.

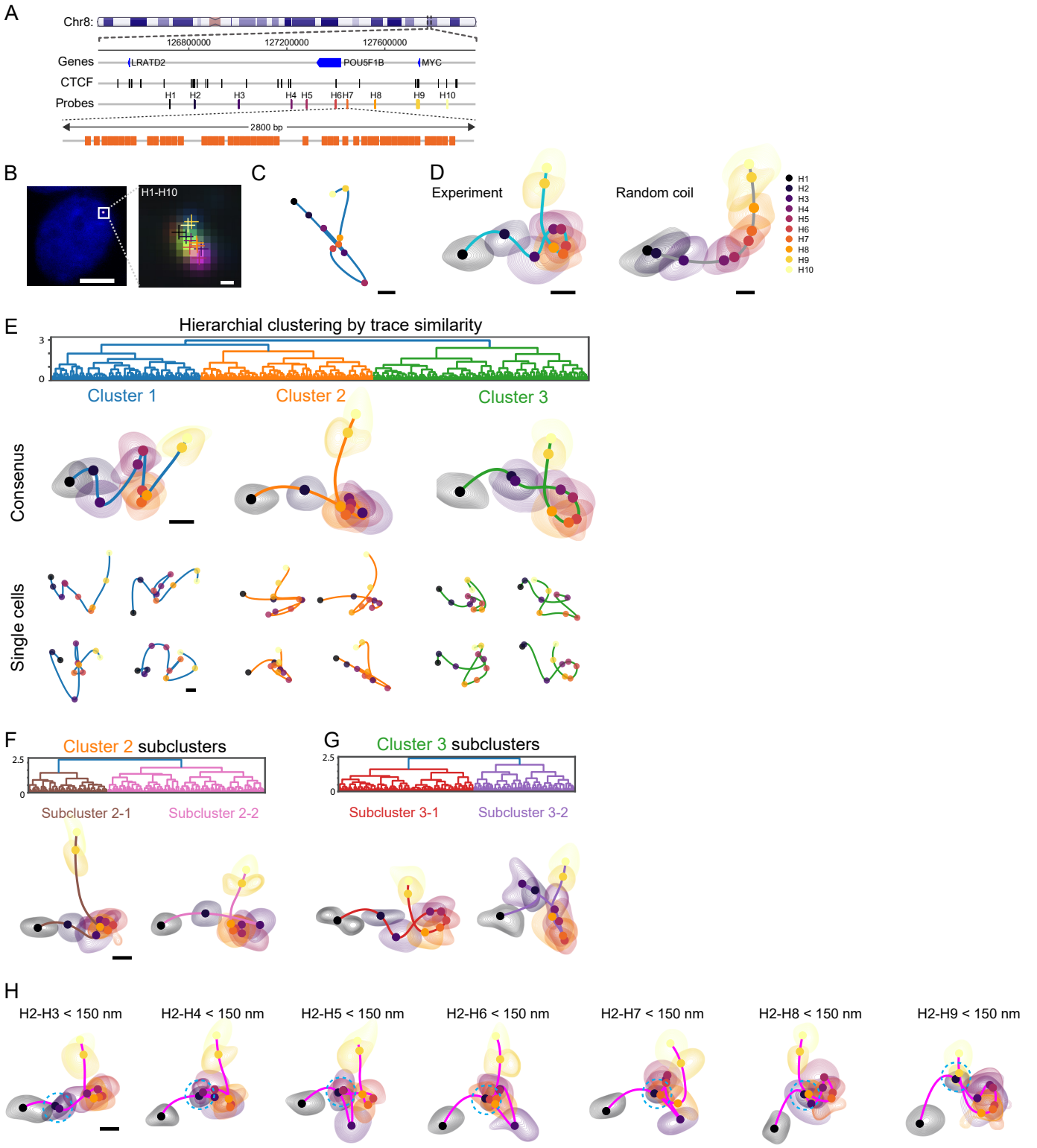
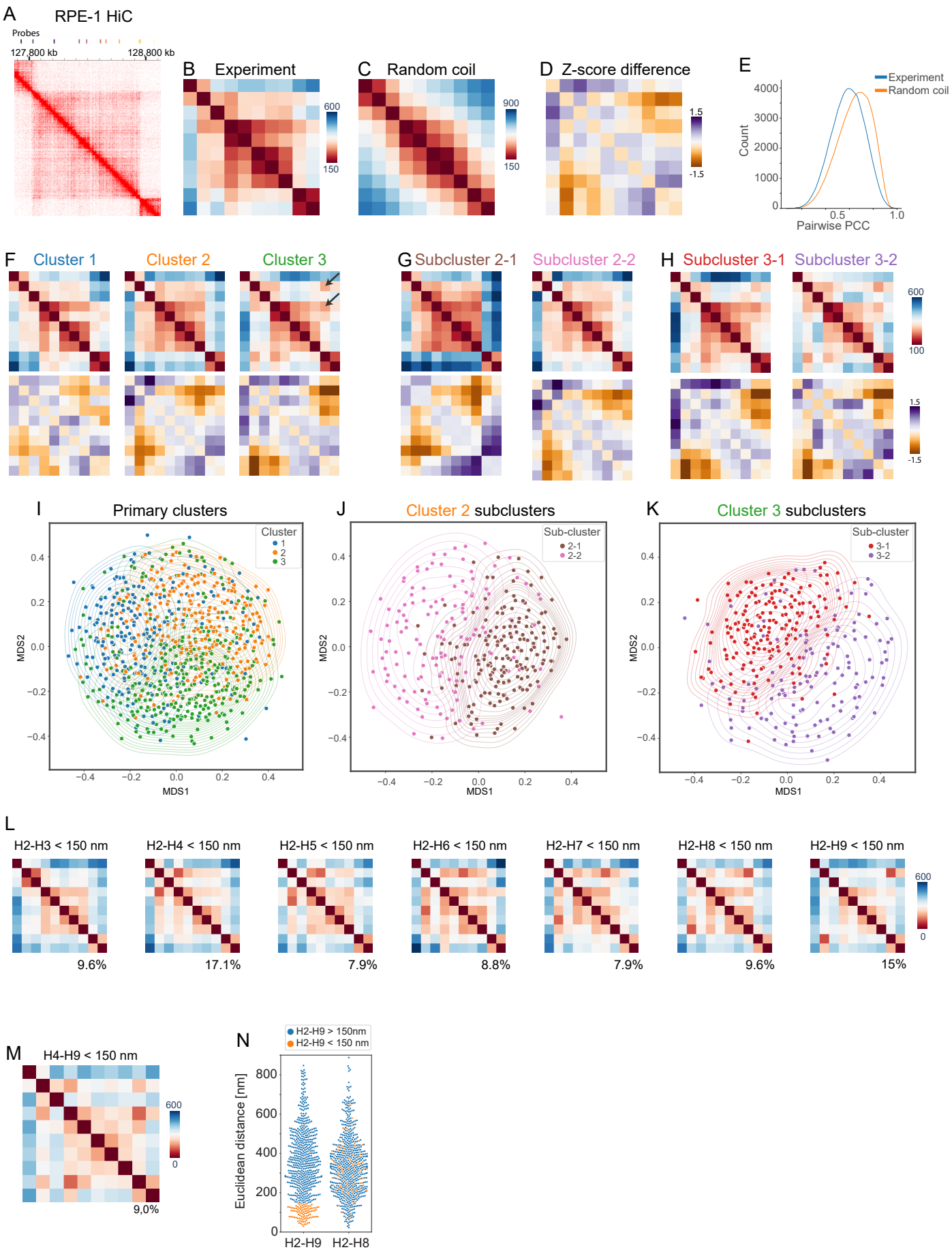


Figure 3: Genomic architecture of a TAD-scale region in the MYC locus.
(A) Overview of the genomic region investigated, spanning 1.1 Mb of the genome in a regulatory region upstream of the MYC gene in chr8. 10 barcoded ~3 kb-sized probe sets were designed for this region targeting loci of interest such as clustered CTCF sites and the MYC promoter and gene. (B) Image of a representative cell traced with the probe library described in (A). The tracing region with a FISH signal (orange) is highlighted in the nucleus (blue). Single hybridizations and their centroid fits (colored crosses) are shown with a multi-color overlay of the different probe positions. Scale bars, 5 µm (left) and 200 nm (right). (C) 2D-projection in the xy-plane of the chromatin trace resulting from the fits shown in (B). (D) Best-fit 2D projections of the consensus trace of the experimental data and the corresponding random coil model, overlaid with kernel density estimates of the aligned individual traces. Data from n=775 traces in 3 experiments, or 775 simulated traces. (E) Dendrogram displaying the result of hierarchical clustering of traces (see methods). The consensus traces and representative single-cell example traces are shown under each cluster. A total of 775 traces from 3 experiments were clustered. Clusters consist of n=213, n=251, n=311 traces. (F-G) Secondary clustering of clusters 2 (F) and 3 (G) from (E) by setting a lower threshold in the dendrogram, with consensus traces and KDE shown. n=84 and n=167 traces in (F) and = 178 and n=133 traces in (G). (H) Consensus traces with KDE overlay filtered from all traces for a 3D euclidean distance between positions H2 and H3-H9 less than 150 nm. The position of the proximal interaction between H2 and H3-H9 are indicated with a blue ring of 150 nm in diameter. n = 50-100 traces per cluster. Scale bars, 100 nm (C–H).



Supplementary Figure S3: Related to main Figure 3.

(A) HiC contact map of RPE-1 cells at 5kb binning 37 of the TAD-scale region in vicinity of the MYC gene selected for chromatin tracing, with probe positions as indicated. (B) Median distance map of detected probe positions. n=884 traces from 3 experiments. (C) Median distance map of random coil model scaled according to the experimental data (see methods for details). n = 884 traces were simulated. Scale bar (B-C) in nm. (D) Difference between experimental data and random coil model, highlighting distances which are smaller (orange) or larger (purple) than expected compared to the model. Matrices were standardized to

z-scores before subtraction. Pearson's correlation between the experimental and random coil distance maps was 0.89. (E) Kernel density estimation of the distribution of pairwise Pearson's correlation coefficients for traces aligned by 3D-rigid registration in the experimental dataset and an equivalent number of random coil models. n=775 traces from 3 experiments or n=775 simulated traces. (F-H) Median distance maps and corresponding z-score normalized differences between experimental data and random coil models for the indicated primary and secondary clusters. n=213, 251 and 311 traces in the primary clusters, and n=84 and 167 traces and n=178 and n=133 traces in the secondary clusters, respectively. Pearson's correlations were 0.91, 0.86 and 0.82 for the primary clusters respectively. Arrows indicate indication of looping conformations in Cluster 3 in (F). Distance map scales are in [nm]. (I-K) Multidimensional scaling (MDS) plots showing relative trace similarity, overlaid with KDE plots indicating the point densities. Trace similarity was measured by Person's correlation between traces aligned by 3D-rigid registration and encoded into 2D-distances by MDS. The cluster each trace belongs to (identified in (F)-(H)) is represented by the marker color. The plots show the three primary clusters (I) and two secondary clusters of cluster 2 (J) and cluster 3 (K). (L) Median distance maps of traces filtered for H2 to H3-H9 distances less than 150 nm, with the percentage of traces included in the given cluster shown under the map. Scale in [nm]. (M) Median distance map of traces filtered for H4-H9 distance of less than 150 nm. The frequency of these traces was compared to the frequency of the H2-H9 loop, and found to be significantly different at p=0.009 by a two-sided t-test (n = 3 independent experiments). Scale in [nm]. (N) All pairwise Euclidean distances between the positions H2-H9 and H2-H8 respectively, with orange color indicating the traces with a distance less than 150 nm between H2-H9. n=775 traces in the dataset.

Methods

Cell culture

RPE-1 cells (hTERT-immortalized retinal pigment epithelial cell line, ATCC No. CRL-4000) were grown in Dulbecco's Modified Eagle Medium (DMEM/ F-12, Cat. No. 11320074, ThermoFisher) supplemented with 10% FBS (Cat.# 26140, ThermoFisher) and 1% Gibco Antibiotic-Antimycotic (Cat.# 15240096, ThermoFisher). Cells were incubated at 37 °C, 5% CO_2 in a humidified incubator. At 70-80% confluence cells were trypsinized with 0.05% Trypsin/EDTA (Cat.# 25300-054, ThermoFisher) and transferred to a new culture dish at appropriate dilutions every 2-3 days. For microslide seeding (µ-Slide VI 0.5 Glass Bottom, Cat.# 80607, Ibidi), trypsinized cells were counted and diluted to ~$5x10^5$ cells/ml in complete culture media.

In vitro DNA-Exchange-PAINT

To simulate a chromosome in DNA-Exchange-PAINT experiments, a single-stranded M13 bacteriophage DNA (Cat.# N4040S; New England Biolabs) was cut with the restriction enzymes BamHI (Cat.# R0136L; New England Biolabs) and BglII (Cat.# R0144S; New England Biolabs) leading to a 6566 nt long linear single stranded DNA molecule. Ten loci along this “mini-chromosome” were targeted with ten unique DNA barcodes (Supplementary Table 1) by extending three of the primary FISH probes (Supplementary Table 2) with a unique 3’ extension (P1-P10) for binding of a secondary probe containing the Cy3B fluorophore (MWG Eurofins). To facilitate binding to the streptavidin coated imaging chambers, two 28 nt long primary FISH probe sequences targeted the sequenced directly upstream of each of the ten loci. These had a 3’ extension (P_{biotin}) (Supplementary Table 1) for binding to a biotinylated DNA oligo. In addition, six primary FISH probes were distributed evenly between the loci with a generic 3’ extension (P0) (Supplementary Table 1), giving a traceable line between the locus-specific probes. As a reference for drift correction, DNA origami structures labelled in a 3-by-4 grid pattern with 20 nm between each point were assembled with a reference imager (P_{ref}, Supplementary Table 1). For details see [37]. Both the mini-chromosomes and the DNA origami structures were assembled in a 20 µl reaction as described in [37].

Prior to imaging, origami and mini-chromosomes were attached to glass bottom slides (Cat.# 80607; Ibidi GmbH) that were pre-cleaned with isopropanol. Each well was incubated with 40 µl BSA-biotin buffer [1 mg/ml BSA-biotin (Cat.# A8549; Sigma-Aldrich); 10mM Tris-HCl (pH 8); 100 mM NaCl] and washed 3 times with 180 µl of Buffer A+ [10 mM Tris-HCl (pH 8); 100 mM NaCl; 0.05% (v/v) Tween 20 (Tween 20; Cat.# P2287; Sigma-Aldrich) at pH 8.0]. Wells were incubated for 5 min with 40 µl streptavidin buffer [0.5 mg/ml streptavidin (Cat.# S888; Thermo Fisher Scientific); 10 mM Tris-HCl (pH 8); 100 mM NaCl; 0.05% (v/v) Tween 20], then washed 2 times with 180 µl Buffer A+ and 2 times

Buffer B+ [5 mM Tris-HCl (pH 8); 10 mM MgCl2; 1 mM EDTA, 0.05% (v/v) Tween 20]. Then, 40 µl of origami and mini-chromosomes mix [~0.1 nM origami; ~0.025 nM mini-chromosomes] was added and incubated for 8 min. Wells were washed 3 times with 180 µl Buffer B+ before starting the imaging session.

Mini-chromosomes were imaged in eleven exchange rounds (P0-P10) by sequentially adding 60 µl imager strand mixture [5 nM exchange round-specific imager strand (P0-P10); 5 nM P_{ref}; 5 mM Tris-HCl (pH 8); 10 mM MgCl2; 1 mM EDTA, 0.05% (v/v) Tween 20; 1x Trolox solution; 1x PCA solution; 1x PCD solution]. 8000 frames were acquired with a 300 ms long exposure time in each exchange round. Imaging was performed on a Nikon Eclipse Ti microscope (Nikon Instruments) with a 100x oil-immersion objective (CFI Apo TIRF 100x, NA 1.49; Nikon Instruments) with a 160 nm pixel size. A 561 nm laser (200 mW nominal; Coherent Sapphire) was filtered (ZET561/10; Chroma Technology) and directed to the objective with a multi-band beam splitter (ZT561rdc; Chroma Technology). Light from the fluorescent molecules were filtered (ET600/50m; Chroma Technology) and collected with an EMCCD camera (iXon X3 DU-897; Andor Technologies). The optimal laser intensity was found by increasing the laser power until the duration of the blinking event decreased. The intensity was set below this value to obtain maximum number of photons per binding event while still not having excessive bleaching during the docking events.

Images were analysed with Picasso [37] by first localizing the binding events using the Localize module followed by filtering out localizations with higher than 0.05 nm lpx in the Filter module. Drift correction was applied using the Render module by first applying a global cross correlation followed by picking DNA origamis and subsequently drift correcting based on these picks. The third level of drift correction was performed by picking all the visible 3-by-4 points of the DNA origami and then drift correcting based on these fine grained picks.

FISH probe design

Unique 40 or 50-nt-long FISH probes were designed using the human reference genome GRCh38.p12. Sequences used are listed in Supplementary Table 4 and 6. First, sequences were excluded if they were highly repetitive, if they had high sequence similarity elsewhere in the genome and if their GC content was under 30% or over 70%. To test for any off-target binding sites of the remaining probes the genome was searched for areas which have an exact match of 7 nt or more with a probe. These possible off-target matches were then filtered using a sensitive filtering scheme based on spaced seeds to remove lower similarity matches. For the remaining off-target matches, the melting temperature was estimated using the algorithm presented in [38]. Probes for which the highest off-target melting-temperature was below 58 °C (calculated at 50mM [Na^+])

were selected until the desired number of probes in a region was achieved. This pre-computed probe library could be queried in a fast and flexible way to design FISH probes targeting any locus. Probes were ordered as pooled libraries (IDT, oPools).

Secondary imager design

The D sequences were designed using NUPACK [39]. Here, we aimed for a free energy of ~17 kcal/mol at an ion concentration of 50 mM NaCl, 10 mM MgCl and a temperature of 24 ºC. Inspired by the sequences used in [32], this free energy allows stable hybridization during image acquisition and, at the same time, efficient removal of the imaging probes with the washing buffer. To ensure orthogonality between different D sequences (no cross-reactivity), we checked that the free energy of a D sequence to every other D sequence is smaller than 10 kcal/mol. Atto-565 labelled imagers (D1-10, Supplementary Table 3) were purchased from MWG Eurofins. To extend the number of 12-mer imagers, sequences S32-S40 of the non-cross-hybridizing 20-mers orthogonal to the human genome [40] were used as templates and are referred to as the H series (H1-H10) (Supplementary Table 5). A BLAST+ (blast-2.7.1+) search optimized for short sequences with a word size of 8 was run on all 12-mer subsets of each of these 20-mers to ensure that the sequence subset remained orthogonal to the human genome. The 12-mer with the highest E-value to its off-targets was selected. For equally good matches, the 12-mer with the lowest number of off-targets followed by the shortest alignment length was selected. H-series Atto-565 labelled 12-mer oligos were ordered from Sigma.

Strand excision FISH

Strand excision FISH was adapted from previously described protocols[29] to be performed in 6-channel microslides (µ-Slide VI 0.5 Glass Bottom, Ibidi), while minimizing DNA denaturation and chromatin structure perturbations and optimizing FISH signal. Briefly, cells were seeded in microslide channels at a final concentration of ~$5x10^5$ cells/ml in 120 µL cell culture media containing 40 µM BrdU:BrdC mix (3:1, Cat.# B5002, Sigma and Cat.# 284555, Santa Cruz Biotec) and grown for 17-24 h. Afterwards, cells were washed once with PBS and fixed with 4% PFA (v/v, Cat.# 15710, EMS) in PBS for 15 min. Free aldehyde groups were quenched with 100 mM NH_4Cl (Cat.# 213330, Sigma) in PBS for 10 min and cells were permeabilised in 0.5% Triton X-100 (v/v, Cat.# T8787, Sigma) in PBS for 20 min. To further sensitize BrdU/C-labelled DNA to UV light, cells were treated with 0.5 µg/ml DAPI (Cat.# D9542, Sigma) in PBS for 15 min. Direct exposure of microslides to 254 nm UV light was then conducted for 15 min in a Stratalinker 2400 UV crosslinker, and cells were finally treated with 1U/µL Exonuclease III (Cat.# M0206, NEB) in 1X NEBuffer 1 at 37 °C for 15 min in a humid container. All incubations were performed at room temperature, unless stated differently. For primary probe hybridization, cells were first incubated at 37 °C for 1 h in buffer H1 (50% formamide (FA, Cat. # AM9342, ThermoFisher), 10% (w/v) dextran sulfate (D8906,

Sigma), 0.4 mg/ml RNAse A (Cat.# R6513, Sigma) in 2xSSC (Cat.# AM9763, ThermoFisher) and afterwards with primary probes (~2nM per primary probe) diluted in the same buffer. Probe annealing was conducted overnight at 42°C in a humid chamber. Primary probes were washed twice in 50% FA in 2xSSC buffer at 25 or 35°C (depending on probe library on-target melting temperature) for 5 min, followed by washes with 2xSSC containing 0.2% (v/v) Tween-20 (Cat.# P9416, Sigma). For non-sequential secondary probe hybridization, fluorescently labelled 20-mer imagers were diluted to a final concentration of 20 nM in buffer H2 (25% FA, 10% (w/v) dextran sulfate, 0.1% (v/v) Tween-20, 2xSSC buffer) and incubated for 2 h at 30 °C. Washing steps were then conducted with 25% FA in 2xSSC buffer for 5 min, followed by washes with 2xSSC containing 0.2% (v/v) Tween-20. Before imaging, DNA was stained with 0.2 µg/ml DAPI for 10 min.

Classical FISH

Standard FISH protocols including heat denaturation of DNA were performed in parallel to compare with SE-FISH procedures. RPE-1 cells were seeded in microslides, fixed and permeabilized as above and treated with 0.1M HCl (Cat.# 320331, Sigma) for 15 minutes. For primary probe hybridization, cells were incubated for 1h in buffer H1 at 37 °C, followed by 1h incubation in the same buffer containing the primary probes and a 3 min heat incubation step in a ThermoBrite slide hybridization system (Leica Biosystems), either at 75°C or 86°C. The temperature during incubation was verified with an additional external digital thermometer. Hybridization was then continued overnight at 42 °C. Washes and secondary hybridization steps were conducted as indicated for the strand excision FISH protocol.

Pulse DNA labelling with fluorescent nucleotides

Fluorescent labelling of TAD domains was adapted from[31]. Briefly, cells were subjected to cell cycle arrest with 1 µg/ml aphidicolin (Cat.# A0781, Sigma) for 9 h, and subsequently allowed to recover and enter S-phase by incubation in fresh DMEM media for 20 min. Afterwards, cells were trypsinized, washed with PBS and resuspended in Resuspension Buffer R (Neon™ Transfection System 10 µL Kit, Cat.# MPK1025, Invitrogen) at a final density of 5.5 x 10^6 cells/ml. dUTP-AF647 (Cat.# NU-803-XX-AF647-L, Jena Biosciences) was added at a final concentration of 60 µM and cells were electroporated with the following pulse parameters: 1 pulse, 1100 v, 20 ms with the Neon transfection system (Invitrogen). Labeled cells were grown in complete DMEM media, passaged after 1 day and seeded in microslide chambers 3 days after transfection. For SE-FISH experiments, 40 µM BrdU/C was added at the seeding step.

SIM image acquisition and analysis

To label DNA, samples were incubated with 1µg/mL Hoechst 33258 (Cat.# 14530, Sigma) in 1XPBS for 30 min. To avoid photobleaching of the replication domain signal, samples were imaged in an oxygen scavenger buffer system for SIM imaging (100 mM tris, 50 mM NaCl, 0.5 mg/mL glucose oxidase, 40 µg/mL catalase, 10% glucose, 1 mM trolox for imaging before FISH, while tris and NaCl were replaced with 2XSSC for imaging after FISH). SIM images were acquired on a Zeiss Elyra 7 with a 63X 1.4NA oil immersion objective using lattice SIM mode. Initial pixel size was 97 nm and z-spacing was 144 nm and the camera exposure time was 50 ms. DNA (Hoechst, 405 nm laser) and replication domain (AF647, 642 nm laser) channels were acquired sequentially. Raw SIM images were processed in Zen Black 3.0 using standard settings except sharpening was set to "weak", resulting in images with twofold increase in pixels in each dimension. To compare images acquired before and after steps in the FISH protocol a global 3D translational drift correction was first applied using SciPy[41]. Single nuclei were detected and segmented using Cellpose[42], and the central plane of the nucleus was identified as the plane with the highest integrated intensity, and the corresponding plane was found in the comparison image as the plane with highest Pearson's correlation. 2D slices of the central plane and the central plane + 8 slices were used for comparison. Maximum projections were used for the replication domains, and individual clusters of replication domains were segmented by Otsu thresholding. Images were cropped to the detected nuclei or replication domains, and a second pixel-level drift correction was applied to both rotation and translation (imreg-dft package). Finally, a subpixel drift correction was applied, before the pixel-wise Pearson correlation coefficient between the aligned images was calculated. Values for the two nuclear planes were averaged.

Microfluidics setup

A custom microfluidics setup was build using a commercial 3-axis 30x18 cm GRBL controlled CNC stage. A syringe needle was mounted in place of the CNC drill head and connected using 1 mm i.d. PEEK and silicone tubing (VWR) to a male Luer adapter, which connected to the sample microslide. The microslide output tubing was connected to a mp6-liquid piezo micropump (Bartel's mikroteknik), connected to an MP-X controller (Bartel's mikroteknik) running at 130 Hz and 180 Vpp when dispensing. This gave a flow rate of ~3.5 mL/min. Probe, wash and imaging buffers were kept in 96 well or 4 well deep plates covered in parafilm (Cat.# P7793, Sigma) and placed on the CNC bed, and the appropriate solution was chosen by moving the syringe needle into the liquid through the parafilm. The automated control software was written in Python and was run on a PC connected to the MP-X controller and GRBL CNC board.

Chromatin trace acquisition

Cells on microslides were prepared for FISH as above, omitting secondary hybridization. 100 nm infrared fluorescent fiducial beads (Cat.# F8799, ThermoFisher) were diluted 1:20 000 in 2XSSC

and applied to the sample 10 minutes before washing in 2XSSC. 12-mer Atto-565-labeled secondary imaging probes (sequences listed in Supplementary Table 2) were diluted to 20 nM in hybridization buffer (5% ethylene carbonate (EC, Cat.# E26258, Sigma) in 2XSSC/0.2% tween). Washing buffer was 10% FA in 2XSSC/0.2% tween, imaging buffer was 2XSSC with 1µg/mL Hoechst 33258 and probe stripping buffer was 25% FA in 2XSSC/0.2% tween. Probes were hybridized for 3 minutes and washed for 1 min prior before switching to imaging buffer and imaging. After imaging, probes were stripped for 2 minutes, then wash buffer was flowed through before the next round of probe hybridization. Imaging was performed on a Zeiss Elyra 7 widefield system using a 63X1.46NA oil immersion objective in HILO mode to reduce background signal while maintaining high dynamic range for accurate position determination. 3D stacks of Hoechst (405 nm laser), FISH probes (561 nm laser) and fiducial beads (642 nm laser) were acquired as sequential frames at each z-position. The pixel size was 97 nm and z-spacing 150-200 nm, and the camera exposure time was 50 ms. 15-40 fields of view were acquired per round of hybridization. Total acquisition time for a typical experiment with 20 fields of view and 14 hybridizations was about 4h. The microscope was controlled using Zen 3.0 software (Carl Zeiss), with the MyPic macro [43] and custom Python software for automation and synchronization with the microfluidic software.

Chromatin trace fitting

Custom python software built primarily on SciPy[41], pandas[44], scikit-image[45], scikit-learn[46], pysimplegui, dask[47] and napari[48] was developed to fit and analyse chromatin traces from the sequential FISH images. Raw CZI images were converted to the ZARR file format and loaded as virtual arrays. Subpixel drift correction was done for all hybridization rounds by cross-correlation of 50-100 segmented fiducial beads and averaged. Regions of interest (ROIs) with FISH signal were automatically identified based on enhancement of spot-like signals using a difference of gaussian filter and an empirical threshold. Nuclei were detected using Cellpose and only ROIs inside nuclei were included. Detected ROIs were manually quality controlled. Typically, the images used for detecting ROIs were from an additional hybridization with multiple imaging probes for a brighter signal. The ROIs with FISH signals were cropped and deconvolved using a theoretical microscope-appropriate point spread function and 10-30 iterations of a Richardson-Lucy iterative algorithm (FlowDec package[49]). The FISH signal from each hybridization was fit to a 3D Gaussian function by least squares minimization, with the center defining the coordinate in the corresponding chromatin trace. All fits were quality controlled by cut-offs on signal to background, fit standard deviation and distance to the multiple-probe signal to avoid fitting of spurious background signals. Chromatin traces were filtered for minimum percentage of quality-controlled fit loci of 50% for distance matrices, 70% for pairwise similarity measures and 100% for radius of gyration measures.

Chromatin trace analysis

Chromatin traces were analysed using SciPy[41] and scikit-learn[46] python packages. Similarity between traces was measured by first performing a 3D rigid alignment by singular value decomposition (SVD) without scaling to minimise the RMSE [50], then calculating the resulting Pearson correlation coefficient (PCC). Radii of gyration were calculated by $R_g^2 = \frac{1}{N} * \sum_{k=1}^{N}(r_k - r_{mean})^2$. Clustering of traces was done by hierarchical clustering with Ward's method and 1-PCC as a distance metric, the dendrogram cutoff was set to include 3 clusters. To rule out that the primary clusters we found could arise by chance, we compared them to randomly assigning traces to clusters of corresponding sizes. All random permutations showed no significant difference between intra- and inter-cluster dispersion (ratio 0.9998 +/- 0.0011, 100 permutations), while our actual clustering had a far larger ratio of 1.102. For sub-clustering a lower threshold was set to subdivide the selected primary cluster into two. Multi-dimensional scaling was performed using the same distance metric as in the clustering above, and scaling into two components. Consensus traces for each cluster were determined by a general Procrustes alignment using 3D rigid alignment as above, initially choosing a random trace in the cluster as the template. A quadratic spline interpolation was used to connect the fit loci coordinates for visualization. Unless otherwise noted, best fit 2D projections of the traces are shown. These were generated by SVD of the 3D point set to find the normal vector of the best fitting 2D plane, then projecting the points onto this plane. Consensus traces of multiple traces were overlaid with kernel density estimates (KDE) of single trace coordinates, aligned to and projected on the plane of the consensus trace. A cutoff at 75% of maximum density was used for legibility of the KDE plots.

Random coil model

Random coil models for the genomic regions studied were generated from an equal number of 3D random walkers as traces, where each step was in a random direction and had a length drawn from a normal distribution with a mean corresponding to the genomic length between each probe scaled by the empirical factor of 24.7 nm $kb^{-(1/2)}$ found in our tracing data, and a standard deviation of 50% of the mean, which is similar to what was observed in the tracing data. Random coil data distance matrices were compared with tracing data by z-score normalization of both matrices followed by subtraction. Random coils were visualized as the experimental traces above.

Data availability

Source image data and processed data from this study are available on the EMBL EBI Bioimage Archive/Biostudies at https://www.ebi.ac.uk/biostudies/studies/S-BIAD59. Publicly available datasets used in this study are GSE71831, ENCSR000DVI.

Code availability

All custom code for image acquisition, microfluidic control, image analysis and LoopTrace analysis is available online at https://git.embl.de/grp-ellenberg/looptrace. The primary oligo FISH probe design software is available online at https://gitlab.com/Chromotrace_utils/probe-design.

Methods references

37. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nat. Protoc.* **12**, 1198–1228 (2017).
38. Gans, J. D. & Wolinsky, M. Improved assay-dependent searching of nucleic acid sequence databases. *Nucleic Acids Res.* **36**, e74 (2008).
39. Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
40. Gelali, E. *et al.* iFISH is a publically available resource enabling versatile DNA FISH to study genome architecture. *Nat. Commun.* **10**, 1–15 (2019).
41. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
42. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
43. Politi, A. Z. *et al.* Quantitative mapping of fluorescently tagged cellular proteins using FCS-calibrated four-dimensional imaging. *Nat. Protoc.* **13**, 1445–1464 (2018).
44. The pandas development team. *pandas-dev/pandas: Pandas*. (Zenodo, 2020). doi:10.5281/zenodo.3509134.
45. Walt, S. van der *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).
46. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
47. Dask Development Team. *Dask: Library for dynamic task scheduling*. (2016).
48. napari contributors. *napari: a multi-dimensional image viewer for python.* (2019).
49. Czech, E., Aksoy, B. A., Aksoy, P. & Hammerbacher, J. Cytokit: a single-cell analysis toolkit for high dimensional fluorescent microscopy imaging. *BMC Bioinformatics* **20**, 448 (2019).
50. Arun, K. S., Huang, T. S. & Blostein, S. D. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Trans. Pattern Anal. Mach. Intell.* **9**, 698–700 (1987).

Acknowledgements

We thank Merle Hantsche-Grininger for developing the electroporation protocol for replication domain labelling and critical reading of the manuscript; Franziska Kundel, Jonas Ries and Virginie Uhlmann for discussions; Magda Bienko and Jill Brown for sharing of FISH protocols; Orsolya K. Wade, Maximilian S. Strauss, and Thomas Schlichthaerle for assistance with Exchange-PAINT; Marko Lampe and the EMBL Advanced Light Microscopy Facility (ALMF) for microscope support; Christian Tischer and Jean-Karim Hériché for help with data/metadata handling, and the electronic workshop and mechanical workshop at EMBL for assistance with the microfluidics setup. This work was supported by grants from the National Institutes of Health Common Fund 4D Nucleome Program (Grant U01 EB021223 / U01 DA047728) to J.E. and E.B., as well as by the The Paul G. Allen Frontiers Group through the Allen Distinguished Investigator Program to J.E. and R.J. K.S.B. was supported by the Alexander von Humboldt foundation and EMBL.

Author contributions

J.E. and E.B. conceived the project. J.E. and S.A planned and supervised the project. Ø.Ø-F. planned and performed FISH and DNA-PAINT experiments. Ø.Ø-F and C.B. optimized the FISH probe design algorithm and designed FISH probes. K.S.B. performed SIM and chromatin tracing experiments with the help of N.R.M. N.R.M. extensively validated reagents and protocols. K.S.B. wrote the microscope/fluidic automation and chromatin tracing and analysis code, and analysed the data. C.B. designed the random coil model. K.S.B., Ø.Ø-F. and J.E. interpreted the data. R.J. supervised all parts of the project that includes DNA-PAINT and DNA-Exchange and F.S. designed the D series of 12-mer imager strands. K.S.B., Ø.Ø-F. and J.E. wrote the manuscript. All co-authors critically read and edited the manuscript.

Authors declare no competing interests.

Additional information

The following supplementary tables are available:

Supplementary Table 1: DNA-exchange-PAINT imager sequences (P-series)

Supplementary Table 2: DNA-exchange-PAINT FISH probe sequences (P0-10)

Supplementary Table 3: 12-mer imager sequences (D-series)

Supplementary Table 4: FISH probe sequences MYC 100kb region (D1-10)

Supplementary Table 5: 12-mer imager sequences (H-series)

Supplementary Table 6: FISH probe sequences MYC TAD region (H1-10)