nature methods

ARTICLES

https://doi.org/10.1038/s41592-020-0890-0

3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing

Huy Q. Nguyen[1,14], Shyamtanu Chattoraj[1,14], David Castillo[2,14], Son C. Nguyen[1,12], Guy Nir[1,3], Antonios Lioutas[1], Elliot A. Hershberg[4], Nuno M. C. Martins[1], Paul L. Reginato[1,3,5], Mohammed Hannan[1], Brian J. Beliveau[4,6], George M. Church[1,3], Evan R. Daugharthy[1,3,7,8,13], Marc A. Marti-Renom[2,9,10,11] ✉ and C.-ting Wu[1,3] ✉

There is a need for methods that can image chromosomes with genome-wide coverage, as well as greater genomic and optical resolution. We introduce OligoFISSEQ, a suite of three methods that leverage fluorescence in situ sequencing (FISSEQ) of barcoded Oligopaint probes to enable the rapid visualization of many targeted genomic regions. Applying OligoFISSEQ to human diploid fibroblast cells, we show how four rounds of sequencing are sufficient to produce 3D maps of 36 genomic targets across six chromosomes in hundreds to thousands of cells, implying a potential to image thousands of targets in only five to eight rounds of sequencing. We also use OligoFISSEQ to trace chromosomes at finer resolution, following the path of the X chromosome through 46 regions, with separate studies showing compatibility of OligoFISSEQ with immunocytochemistry. Finally, we combined OligoFISSEQ with OligoSTORM, laying the foundation for accelerated single-molecule super-resolution imaging of large swaths of, if not entire, human genomes.

A capacity to view genomes in situ, in their entirety and at high genomic resolution is becoming increasingly important, with one potentially enabling class of methods being fluorescence in situ hybridization (FISH)[1]. Indeed, it was FISH that enabled the pioneering work demonstrating chromosome territories in interphase cells[2,3]. Of the several methods for FISH, a number are oligomer (oligo) based[1]; one such method is Oligopaints[4] (see Supplementary Note 1 for additional examples), which appends nongenomic sequences (Mainstreet and Backstreet) to enable multiple functionalities, including amplification, indirect visualization via fluorophore-conjugated (secondary) oligonucleotides, barcode-based multiplexing and sequential and combinatorial labeling of DNA or RNA[4–21]. In the context of megabase-level coverage, some studies have used these functionalities to walk along contiguous megabases of the genome[13,14], with others labeling up to 40 regions on single chromosomes to reveal chromosomal paths[9,21], and still other studies visualizing entire, or nearly entire, genomes, one chromosome or one chromosome arm at a time[15,19]. Here we demonstrate how streets enable a new technology, OligoFISSEQ, which vastly increases the number of targets that can be visualized, putting us within reach of genome-wide imaging via the visualization of a multitude of subchromosomal regions. As OligoFISSEQ is compatible with the single-molecule localization method OligoSTORM[5,10], it also accelerates the speed with which genomic regions can be visualized at super-resolution.

OligoFISSEQ is based on FISSEQ technologies that have been honed for in situ detection of transcripts[22,23] (see Supplementary Note 2 for recent iterations and earlier studies). Here we present three strategies that direct the sequencing to barcodes embedded in Oligopaint streets, wherein one strategy uses sequencing by ligation (SBL), another uses sequencing by synthesis (SBS) and a third strategy uses sequencing by hybridization (SBH). Focusing on OligoFISSEQ with SBL, we map 66 genomic regions in human diploid PGP1 skin fibroblast cells (XY; PGP1f) using only four rounds of sequencing. We next introduce a method to improve barcode detection and, in conjunction with OligoFISSEQ, trace the human X chromosome by mapping 46 regions along its length. We demonstrate that OligoFISSEQ is compatible with immunofluorescence (IF) and then conclude by combining OligoFISSEQ with OligoSTORM to achieve a much accelerated rate at which multiple genomic regions (ranging in size from tens of kilobases to megabases) can be visualized simultaneously at super-resolution.

Results

Principle and validation of OligoFISSEQ. FISSEQ technologies[22,23] leverage next-generation sequencing methods[24,25] to provide in situ 3D spatial maps of transcripts that have been reverse transcribed and then amplified. As FISSEQ can also be used for in situ decoding of barcodes introduced during the generation of cDNA, we reasoned that it might be possible for FISSEQ to read barcoded

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [3]Wyss Institute, Harvard Medical School, Boston, MA, USA. [4]Department of Genome Sciences, University of Washington, Seattle, WA, USA. [5]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Brotman Baty Institute for Precision Medicine, Seattle, WA, USA. [7]Department of Systems Biology, Harvard Medical School, Boston, MA, USA. [8]ReadCoor, Cambridge, MA, USA. [9]CRG, BIST, Barcelona, Spain. [10]Pompeu Fabra University, Barcelona, Spain. [11]ICREA, Barcelona, Spain. [12]Present address: Department of Genetics, University of Pennsylvania, Philadelphia, PA, USA. [13]Present address: ReadCoor, Cambridge, MA, USA. [14]These authors contributed equally: Huy Q. Nguyen, Shyamtanu Chattoraj, David Castillo. ✉e-mail: martirenom@cnag.crg.eu; twu@genetics.med.harvard.edu

Oligopaints. Furthermore, by targeting hundreds to thousands of identically barcoded Oligopaints to a genomic region, the combination of Oligopaints with FISSEQ, which we call OligoFISSEQ, could both obviate the need for target amplification, typically required by FISSEQ, and render the targeted chromosomal structure amenable to imaging. Finally, as FISSEQ is carried out using diffraction-limited microscopy, we anticipated a capacity of OligoFISSEQ to image the same genomic regions in hundreds to thousands of cells and thus provide the computational and statistical power necessary for addressing cell-to-cell variability.

We began by designing an Oligopaint library that targeted 18,536 oligonucleotides to a 4.8-Mb single-copy region on human chromosome 19 (Chr19-20K; Extended Data Fig. 1a) and then tested whether it could be sequenced in situ, focusing first on SBL to effect ligation-based interrogation of targets (LIT) and then on SBS to effect synthesis-based interrogation of targets (SIT), implementing hybridization-based interrogation of targets (HIT) only later (Fig. 1a–e). Importantly, as Oligopaint streets can accommodate multiple barcodes, we were able to design a single library to accommodate the sequencing chemistries of both LIT and SIT, with the primer binding site and barcode for LIT embedded on Mainstreet (5′ end of the Oligopaint oligonucleotide) and the primer binding site and barcode for SIT embedded on Backstreet (3′ end of the Oligopaint oligonucleotide; Fig. 1a). We use LIT and SIT to refer to the steps of sequencing per se, and OligoFISSEQ-LIT (O-LIT) and OligoFISSEQ-SIT (O-SIT) to refer to the use of LIT and SIT, respectively, in the context of OligoFISSEQ.

With O-LIT (Fig. 1c and Extended Data Fig. 1b), the barcode was read with SOLiD chemistry[24], wherein each barcode digit (defined as the smallest unit of a barcode; five nucleotides per digit) was read by cleavable 8-mers carrying one of four fluorophores. In brief, a sequencing primer was hybridized to the street, and a subsequent barcode readout began by binding of the first barcode digit by a labeled 8-mer, which was then ligated and imaged. The 8-mer was then cleaved between nucleotides five and six, leaving the first five nucleotides and removing the label, allowing the next digit to be read. Excluding the primer binding site, barcodes were 23 nucleotides in length and sufficient to accommodate four rounds of sequencing ((four rounds of sequencing × five nucleotides per digit) + three nucleotides uncleaved after the fourth round of sequencing); when fully utilized, four- or eight-digit barcodes have the potential to distinguish 256 (4^4) or 65,536 (4^8) targets, respectively. Using O-LIT on Chr19-20K, we recovered four-digit barcodes from 92.1% ± 5.7% of PGP1f cells ($n = 85$ cells from four replicates; Fig. 1f).

In the case of O-SIT (Fig. 1d and Extended Data Fig. 1b), barcodes were sequenced using Illumina NextSeq chemistry[24] via the extension of primers one base at a time and using only two fluorophores; one fluorophore was assigned to deoxycytidine (C), the other was assigned to deoxythymidine (T), both fluorophores were assigned simultaneously to deoxyadenosine (A), and deoxyguanosine (G) was left unlabeled (Fig. 1d,f). With each digit of the barcode being only a single nucleotide, SIT barcodes are compact, with an eight-nucleotide-long barcode theoretically able to identify 65,536 targets (4^8). Following the application of O-SIT to Chr19-20K, we recovered four-digit barcodes from 90.8% ± 5.6% of PGP1f cells ($n = 66$ cells from four replicates; Fig. 1f).

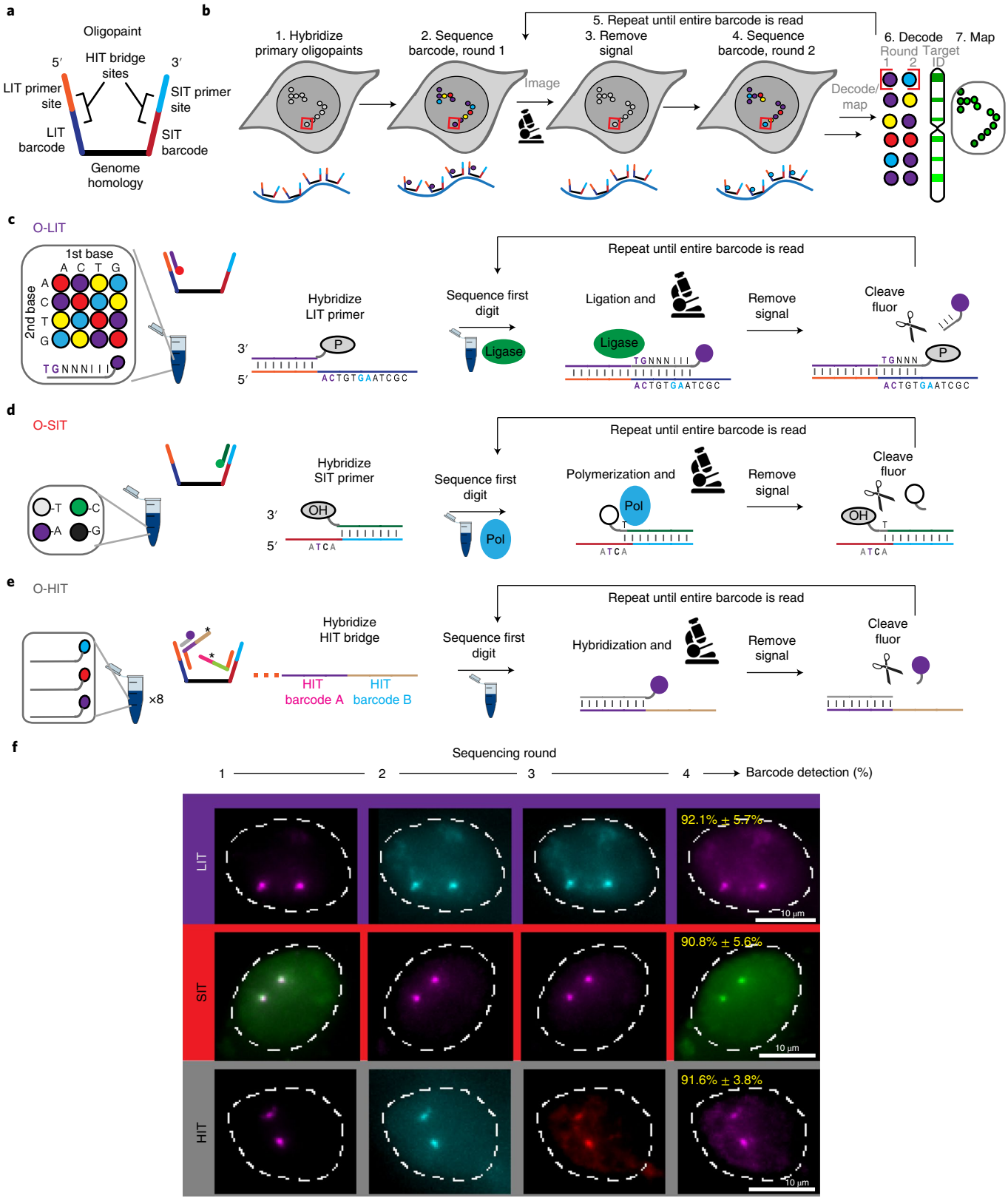
Chr19-20K can also be co-opted for HIT through SBH (Fig. 1a), reminiscent of strategies that have enabled Oligopaints to facilitate transcriptome profiling[6,8,12,18]. Here, we introduce SBH for 3D spatial mapping of chromosomal DNA. In particular, we implemented OligoFISSEQ-HIT (O-HIT) by appending SBH barcodes via two bridge oligonucleotides[14,19,20,26]—one hybridizing to the junction of the LIT barcode and its primer sequence on Mainstreet and the other hybridizing to the junction of the SIT barcode and its primer sequence on Backstreet; SBH barcodes can also be embedded directly into the streets. As each bridge carries two 20-nucleotide barcode positions, each position encoding one of six possible barcodes, the resulting 24 (4 × 6) barcodes had the potential to identify 1,296 (6^4) targets (Fig. 1e). Each barcode was identified via complementary labeled secondary oligonucleotides, and thus, using three fluorophore species, eight rounds of hybridization (8 × 3) were sufficient to identify all 24 barcodes in this iteration of O-HIT, with the option to increase target capacity through additional barcode positions, barcode sequences and/or fluorophore species. By using O-HIT on Chr19-20K, we successfully recovered four-digit barcodes from 91.6% ± 3.8% of PGP1f cells ($n = 79$ cells from four replicates; Fig. 1f and Extended Data Fig. 1b).

Mapping 66 genomic regions with O-LIT. We next assessed the potential of OligoFISSEQ to address multiple regions on multiple chromosomes. We chose to work with O-LIT because it is expected to scale without the increased costs predicted to accompany the scaling up of purely hybridization-based technologies, such as O-HIT, for which the number of species of labeled oligonucleotides, and thus their cost, would increase as the number of targets increases. In contrast, O-LIT reagents would remain the same regardless of whether they target one region or hundreds or thousands of regions. Furthermore, because the five-nucleotide O-LIT barcode digits are relatively compact, they decrease the requisite length of Oligopaint oligonucleotides, further reducing costs. In addition, because O-LIT delivers a positive signal at each round of sequencing, its barcoding is more robust, in contrast to O-SIT and O-HIT, which contain 'blank' readouts.

To assess the scalability of O-LIT, we designed an Oligopaint library (36plex-5K; Fig. 2a) targeting six regions along each of six chromosomes: chromosome 2 (Chr2; 242 Mb), Chr3 (198 Mb), Chr5 (181 Mb), Chr16 (90 Mb), Chr19 (58 Mb) and ChrX (156 Mb), with a unique barcode for each of the 36 targets. Thus, 36plex-5K targeted a total of 66 regions in PGP1f cells (six targets for each of two homologs of the five autosomes and six targets on the single X chromosome), each represented by 5,000 Oligopaint oligonucleotides and, together, encompassing 31.6 Mb, with targeted regions ranging in size between 642 kb and 1.22 Mb (876 kb average). We chose gene-poor chromosomes (5.4–6.1 genes per Mb; Chr2, Chr3,

Fig. 1 | Using OligoFISSEQ to sequence barcoded Oligopaints in situ. a, Oligopaint oligonucleotides used for OligoFISSEQ. Portions of the LIT and SIT primer sites and barcodes can function as binding sites for HIT bridges (**e**), as well as priming sites to amplify the Oligopaint library. **b**, OligoFISSEQ workflow. **c**, O-LIT workflow. After the phosphorylated LIT primer (P) is hybridized, it is ligated to an 8-mer (TGNNNIII), the first two nucleotides of which correspond to a specific fluorophore; as Oligopaint barcodes are predefined, each fluorophore corresponds to only a single barcode digit. N denotes a mixture of A, C, T or G; I denotes deoxyinosine[32], a universal base. **d**, O-SIT workflow. SIT primers contain 3′ hydroxyls (OH). A (purple) and C (green) are conjugated to distinct fluorophores and T (gray) is conjugated to two fluorophores, with G (black) remaining unlabeled. **e**, O-HIT workflow. In this iteration, two bridge oligonucleotides (asterisks) bring in four barcode positions, for each of which there are six possible barcode sequences. As each round of hybridization brings in three fluorophore-conjugated secondary oligonucleotides, each corresponding to one barcode sequence, eight rounds of hybridization (24 labeled oligonucleotides) are sufficient in this case to determine the sequence at each barcode position. **f**, Representative images after four rounds of O-LIT, O-SIT and O-HIT using Chr19-20K on PGP1f cells. Images are representative of maximum-intensity *z*-projections. The first round of SIT identified deoxyadenosine (labeled by a combination of purple and green and thus appearing white). Mean barcode detection efficiencies with s.d. values are shown from four replicates for LIT, SIT and HIT representing 85, 66 and 79 total cells, respectively.

Chr5 and ChrX) and gene-rich chromosomes (10.8 and 23 genes per Mb; Chr16 and Chr19, respectively), as well as large chromosomes (242 Mb; Chr2) and small chromosomes (58 Mb; Chr19). We positioned three targets along each chromosome arm—one target as close as possible to the telomere, one in the center of the arm and one as close as possible to the centromere, with intertarget distances ranging from 7 Mb to 74.9 Mb (average of 28.8 Mb). The number of Oligopaint oligonucleotides per target (5,000) was kept constant to assess the robustness of LIT with respect to target size and different densities of oligonucleotide binding sites (4–7.7 binding sites per



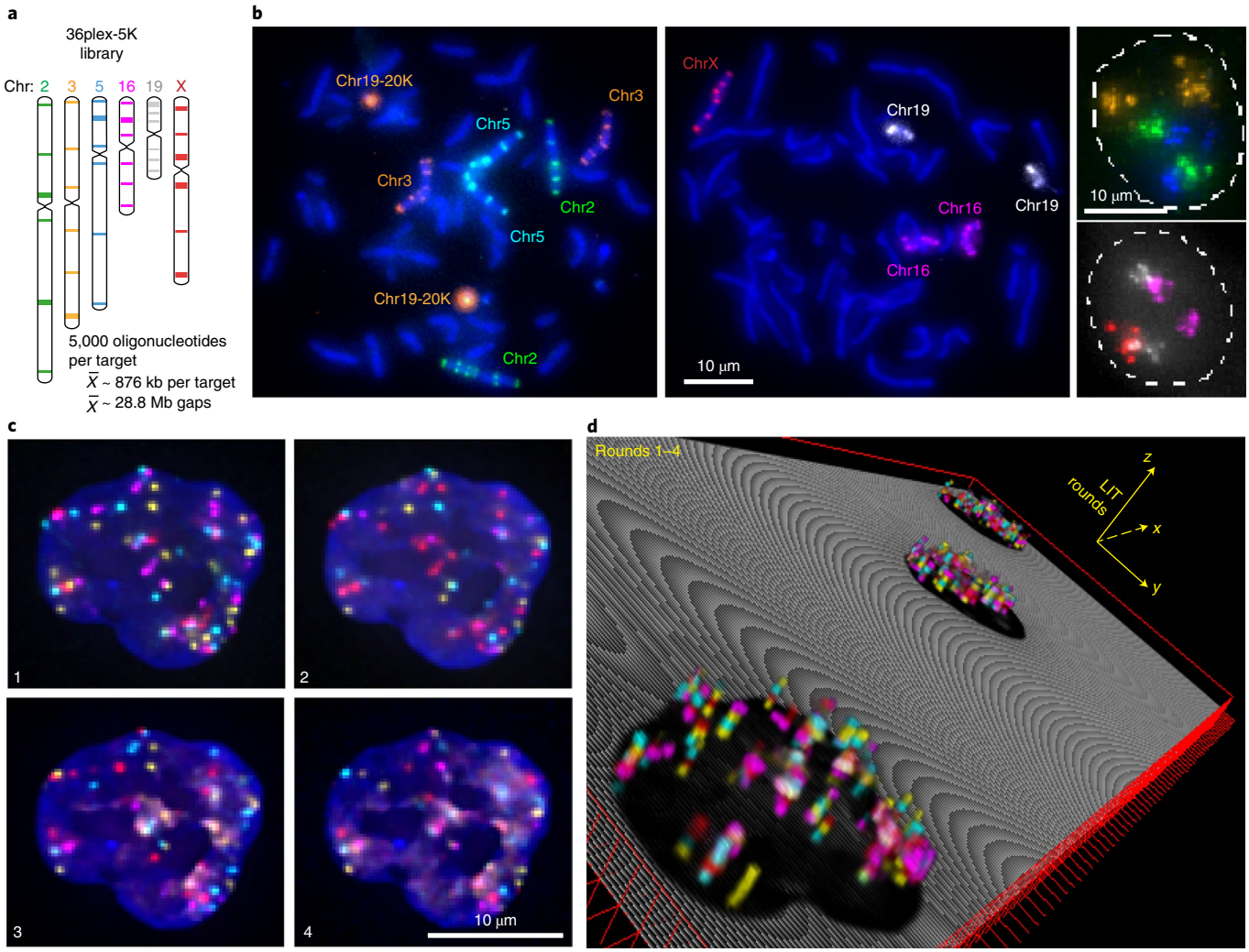


Fig. 2 | OligoFISSEQ-LIT on 36plex-5K. a, Chromosome numbers are color coded to correspond with images in **b**. Each target corresponds to a unique barcode. **b**, Metaphase chromosome spreads of male lymphoblast cells (left; cells from Applied Genetics; Methods) and interphase nuclei from PGP1f cells (right) are representative of four replicates. All six targets on any single chromosome were labeled with secondary oligonucleotides carrying the same species (color) of fluorophore. Chr19-20K was used as a positive control in metaphase chromosome spreads. Images are representative of maximum-intensity *z*-projections. **c**, Four rounds of O-LIT off both streets of 36plex-5K. Images were deconvolved and represent five-color merged maximum-intensity *z*-projections; $n=1$. **d**, 3D representation of the field of view (FOV) containing three cells sequenced with four rounds of O-LIT. Sequencing rounds are represented on the *z* axis, with the first round being closest to the DAPI-determined nuclear outline (black). The maximum-intensity *z*-projection of the sequencing signal from each round was taken, duplicated (a total of two images for better visualization) and then stacked on top of each other. The lower left cell corresponds to the cell in **c**.

kb, average of 5.8). In addition, because all 36plex-5K Oligopaint oligonucleotides targeting the same chromosome shared the same reverse primer sequence, it was possible to use indirect labeling to produce a six-banded pattern along all targeted chromosomes in metaphase and distinctly colored territories in interphase cells (Fig. 2b). This outcome confirmed the accuracy of the library.

An every-pixel automated analysis pipeline. To improve target detection, we sequenced simultaneously off Mainstreet and Backstreet (Fig. 2c,d and Extended Data Fig. 1c–f), which, in the case of 36plex-5K, carried the same barcode. Indeed, this strategy identified 100% of the 66 targeted regions in PGP1f cells via manual decoding ($n=2$ from two replicates; Extended Data Fig. 1f). However, as manual decoding does not scale well, we developed an automated pipeline to address a range of signal intensities and sizes by interrogating every pixel individually (Fig. 3a); a centroid-based pipeline did not perform as well as the every-pixel pipeline (29.93% ± 4.9% versus 62.8% ± 4.8%, $n=111$ cells from three replicates; Extended Data Fig. 1g).

The every-pixel pipeline detected 95% ± 5.15% of 36plex-5K targeted regions but with many false positives (FPs; 574.86 ± 325.38 FPs per nucleus; $n=611$ cells from 15 replicates; Extended Data Fig. 2a,b). Thus, we developed a two-tier system (Fig. 3a) in which tier 1 filtered out pixels below a minimum signal intensity and/or patch size, reducing FPs 165-fold (3.49 ± 1.36 FPs per nucleus; 5.29% ± 2.06%) while detecting 62.2% ± 6.68% of the targeted regions (~41/66) in each nucleus ($n=611$ cells from 15 replicates; Extended Data Fig. 2c,d). In tier 2, the requirements for pixel intensity and patch size were lowered, after which barcode subsampling was applied, and all newly detected signals from the same chromosome were required to be within 4.5 μm of tier 1 detected regions. This proximity-based filtering reflects the propensity of chromosomes to occupy distinct territories[2], as well as measurements of distances between consecutive tier 1 regions along a chromosome (Methods; Supplementary

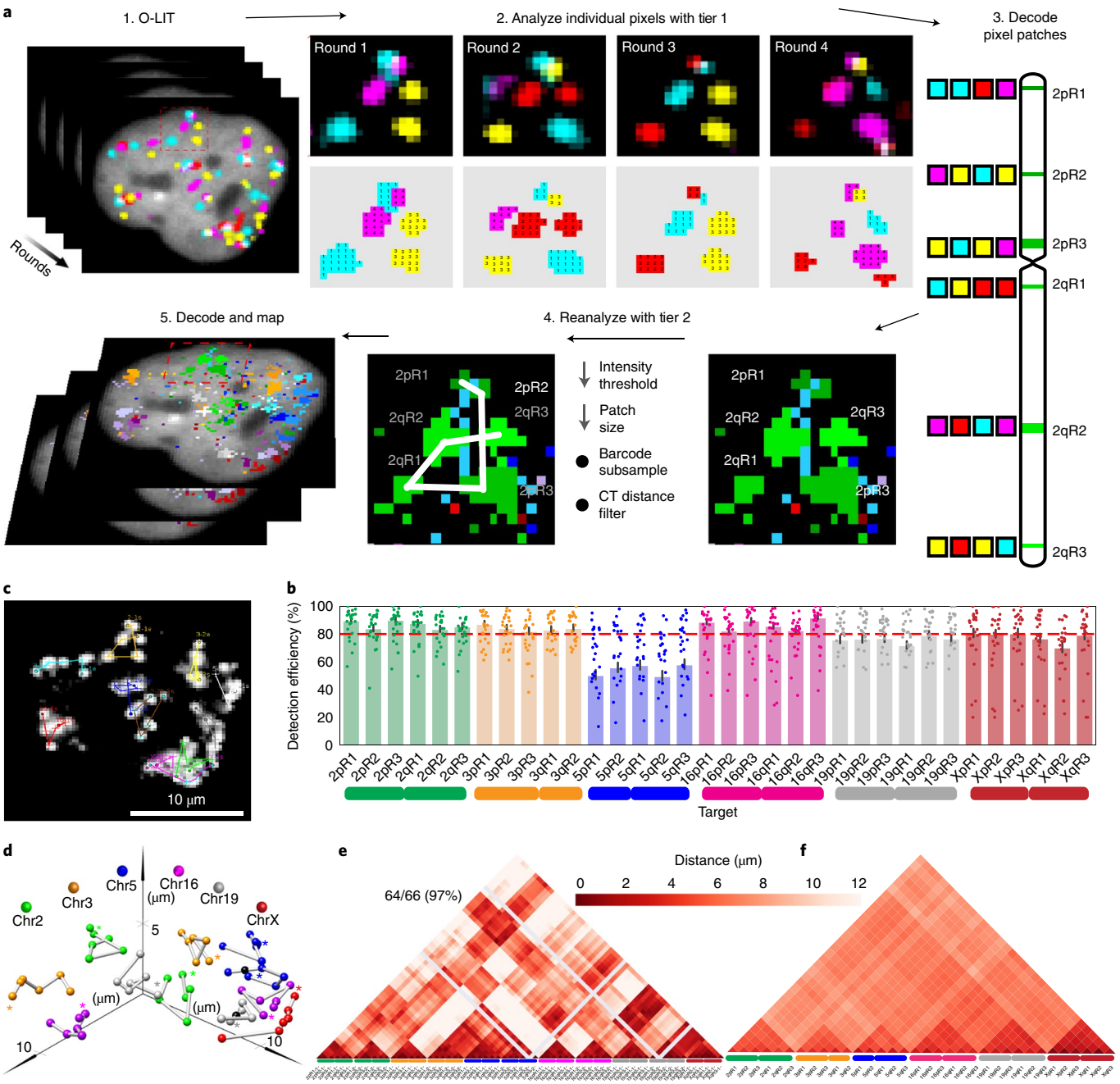


Fig. 3 | Every-pixel analysis pipeline on 36plex-5K. a, Sequencing rounds (step 1) were analyzed at the level of individual pixels using tier 1 parameters with thresholds for signal intensity and pixel patch size (step 2), and pixel patches were then decoded (step 3). Missing targets and FPs were filtered by reanalyzing images with tier 2 parameters (step 4) to produce traces (step 5). Tier 2 decreased the thresholds for signal intensity and pixel patch size, subsampled barcodes and applied filters for chromosome territories. Barcodes and color codes were designated as follows: 1, FITC; 2, Cy3; 3, TxRed; 4, Cy5. **b**, Tier 2 detection efficiency of 36plex-5K after sequencing off both streets; 80.2% ± 7.3% of targeted regions were detected in 611 cells from 15 replicates. **c**, Detection efficiencies from individual replicates are shown, with chromosomal targets on the *x* axis. The dashed red line marks the mean of all chromosomal targets. 3qR3 and 5pR3 shared a barcode and were not included. Error bars represent the 95% bootstrap confidence interval (CI) of the mean. **c**, Chromosome traces of Fig. 2c nucleus after tier 2. In total, 64 of 66 (97%) targeted regions were detected; $n = 1$. **d**, Ball-and-stick traces of the nucleus referred to in **c**. Colored spheres represent targets; black spheres represent undetected targets and were positioned by calculating the median proportionate distance between flanking detected spheres. Gray lines between signals denote extrapolations. The asterisks mark the beginning of chromosomes. **e**, Single-cell pairwise spatial distance matrix after tier 2 detection of the nucleus referred to in **b**. Homologs are displayed separately. Centroids of targets were used for this and all subsequent spatial distance matrices. Gray lines denote undetected targets. **f**, 36plex-5K population pairwise spatial distance measurements after tier 1 detection ($n = 611$ cells from 15 replicates). Homologous target measurements were combined.

Fig. 1), although in the context of chromosome rearrangements it would need to be modified. Tier 2 eliminated all FPs while detecting 80.2% ± 7.3% (~52/66) of targeted regions in each nucleus with at least 70% (~46/66) of targeted regions recovered in ~70% of cells (Fig. 3b and Extended Data Fig. 2e–g). The centroids of all detected targets were then conceptually connected to produce ball-and-stick

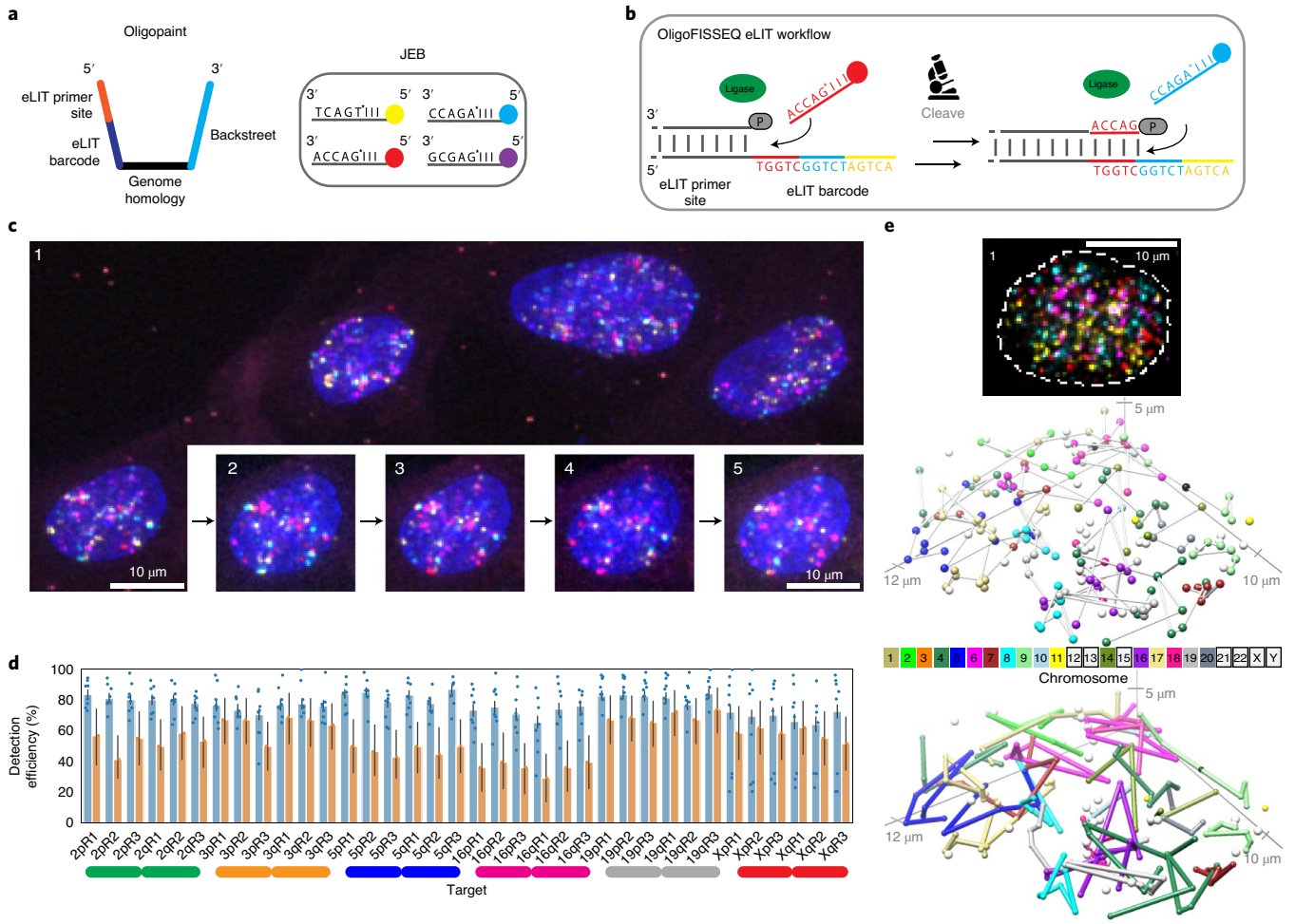


Fig. 4 | Improving O-LIT by using JEB. a, Design of Oligopaint oligonucleotides that use eLIT (left) and JEB-labeled 8-mers complementary to the five-nucleotide eLIT barcode digit (right). eLIT is compatible with a variety of barcode configurations; our strategy used barcodes consisting of five digits each, in which each digit was one of only four distinct five-nucleotide sequences. To further reduce the complexity of the pool of eight-nucleotide oligonucleotides, we also used deoxyinosine[32] in positions 6, 7 and 8. In short, JEB technology reduced the pool of labeled oligonucleotides from 1,024 to four (Extended Data Fig. 5a,b). **b**, eLIT workflow with JEB. **c**, Five rounds of sequencing with O-eLIT. PGP1f cells after the first round of sequencing (cropped field of view) and images from five rounds (1–5) of sequencing of one nucleus (inset). Extranuclear puncta are fiducial TetraSpeck beads (Thermo Fisher). Images are deconvolved maximum-intensity *z*-projections; $n=1$. **d**, Tier 2 target detection efficiency of 36plex-1K after five rounds of O-LIT with SOLiD reagents (orange; average of 54.6%; $n=41$) or O-eLIT with JEB (blue; average of 74% ± 11.2%; $n=440$ from nine replicates). Detection efficiencies from individual replicates are plotted. Error bars represent the 95% bootstrap CI of the mean. **e**, First O-eLIT round of 129-plex (top; deconvolved maximum-intensity *z* projection; $n=1$). Tier 2 tracings (middle; white spheres are tier 1 duplicated barcodes that did not move to tier 2, with untraced chromosomes boxed in color key). Sticks color-coded to facilitate visualization (bottom). Oligonucleotide target density was 5.8 to 11.9 per kb.

renditions of chromosomes, with undetected targets positioned by calculating the median distance between flanking centroids (Fig. 3c,d); ball-and-stick strategies have been used in other studies to trace chromosome paths and are useful when assessing chromosome structure and positioning[9,13,17,20,21]. Note that targets 3qR3 and 5pR3, which were designed to share barcodes, were both detected at 69% efficiency, boding well for the consistency and robustness of barcode recovery. Similarly, 15 replicates using PGP1f cells produced similar ranges of barcode recovery, with no remarkable batch effects as shown in the principal-component analysis (Extended Data Fig. 2h).

Development of eLIT to interrogate fine-scale genome organization. O-LIT mapping of 36plex-5K revealed the paths of all six chromosomes (Fig. 3c,d and Extended Data Fig. 3a,b), producing single-cell spatial genomics data (Fig. 3e,f and Extended Data Fig. 3c–e) that align with previous studies and thus argue the potential of OligoFISSEQ to be informative. First, the chromosomes fell into different territories[3], with the smaller chromosomes (Chr16 and Chr19) and larger chromosomes (Chr2, Chr3, Chr5 and ChrX) positioned toward the center and periphery of the nucleus, respectively (Extended Data Fig. 3f), in line with observations of a radial positioning of chromosomes that places smaller chromosomes more centrally[2,27]. Consistent with this, median inter-homolog distances for the smaller chromosomes were less than those for the larger chromosomes across hundreds of cells (Extended Data Fig. 4a; $P=4.3\times10^{-37}$). These robust sample sizes also enabled consideration of suggestions that diploid genomes can, under some circumstances, separate into two spatially distinct haploid sets[28–30]. Here, cluster analyses of 36plex-5K maps revealed that the five targeted PGP1f autosomes spatially separated into two haploid sets in 6.9% (18/258) of cells (Extended Data Fig. 4b–e), which, however, was statistically similar to proportions expected from randomized controls (5% and 5.4% for directed random and completely random). While definitive descriptions await the analysis of complete genomes, this observation, compounded with studies of homolog pairing and anti-pairing[31], highlights the possibility that it is in

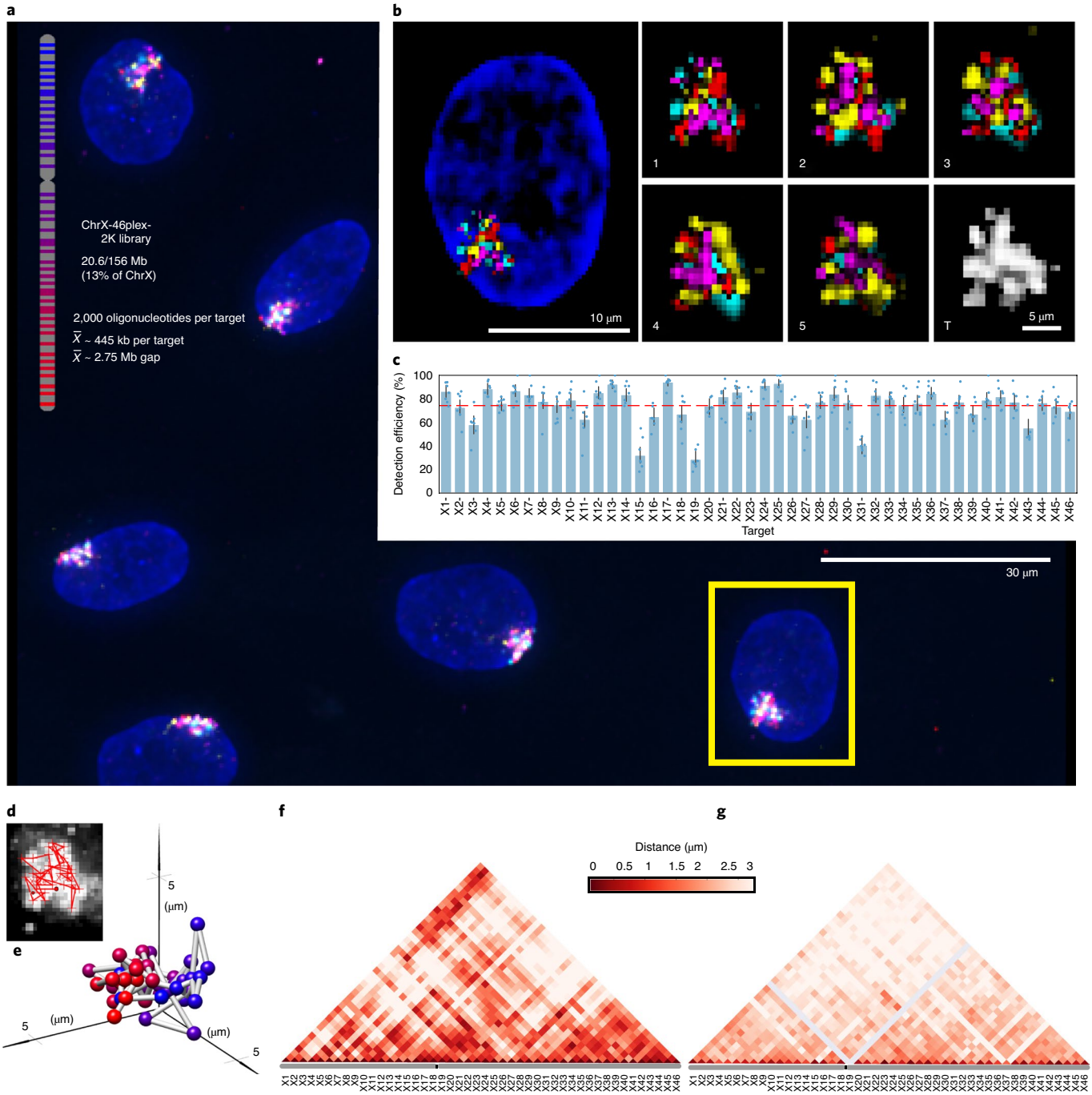


Fig. 5 | Tracing 46 regions along the X chromosome. a, Targets of ChrX-46plex-2K and nuclei after the first round of O-eLIT sequencing off both streets in PGP1f cells. Images are from deconvolved maximum-intensity *z*-projections. $n=1$. **b**, Five rounds of sequencing with O-eLIT off both streets; nucleus from **a** (yellow square). DAPI-stained nucleus after the first round of sequencing (left). T, totality of targets labeled simultaneously with a secondary oligonucleotide complementary to a barcode present on all oligonucleotides. Images are from deconvolved maximum-intensity *z*-projections; $n=1$. **c**, Tier 2 target detection efficiency after five rounds of O-eLIT off both streets in PGP1f cells. The mean detection efficiency (red dashed line) was 74.29% ± 2.5% ($n=177$ from seven replicates), and the average detection efficiency off one street was 73.7% ± 2.97% ($n=122$ from five replicates) and off both streets was 75.3% ± 1.97% ($n=55$ from two replicates). Detection efficiencies from individual replicates are plotted. Error bars represent the 95% bootstrap CI of the mean. **d**,**e**, Chromosome traces (**d**) and 3D visualization (**e**) of the nucleus from **b** after tier 2 analysis and interpolation of missing targets. Sphere color corresponds to chromosome cartoon in **a**; $n=1$. **f**, Single-cell pairwise spatial distances after interpolation of missing targets from the nucleus in **b**. **g**, Population pairwise spatial distances ($n=177$ from seven replicates) after tier 1 detection (combining reads off Mainstreet with reads off both streets).

cell types that do not segregate the genome into haploid sets that inter-homolog interactions will prevail.

We also aggregated single-cell 36plex-5K data from 611 cells to produce an average distance matrix, but this time combining data for homologous chromosomes (Fig. 3f). The comparison of this matrix to a Hi-C map of PGP1f cells[14] revealed a strong correlation ($r=0.705$, $P=1.77\times10^{-174}$; Extended Data Fig. 3e), once more indicating the robustness of O-LIT. Nevertheless, the matrices also

differed, with O-LIT producing subchromosomal stripes of greater or lesser distance, and the Hi-C matrix being more mottled. While stripes may reflect discontinuities along a chromosome, they may also suggest chromosome-specific[9,21] and interchromosome-specific signatures. For example, chromosomal regions that are overall further from other regions may be relatively more buried within a chromosome territory or nearer the nuclear membrane, while chromosomal regions that are closer to other regions may be nearer to the surface of chromosomal territories or less constrained to the nuclear membrane. As for the mottled appearance of the Hi-C matrix, it suggests that, at the scale of whole chromosomes, distances on the order of microns may not always correlate with interaction frequencies and distances amenable to Hi-C; indeed, an absence of correlation may indicate that proximity and interaction are distinct features. Thus, O-LIT matrices of distance and Hi-C matrices of interaction frequency may, together, provide layers of information that neither matrix alone can provide.

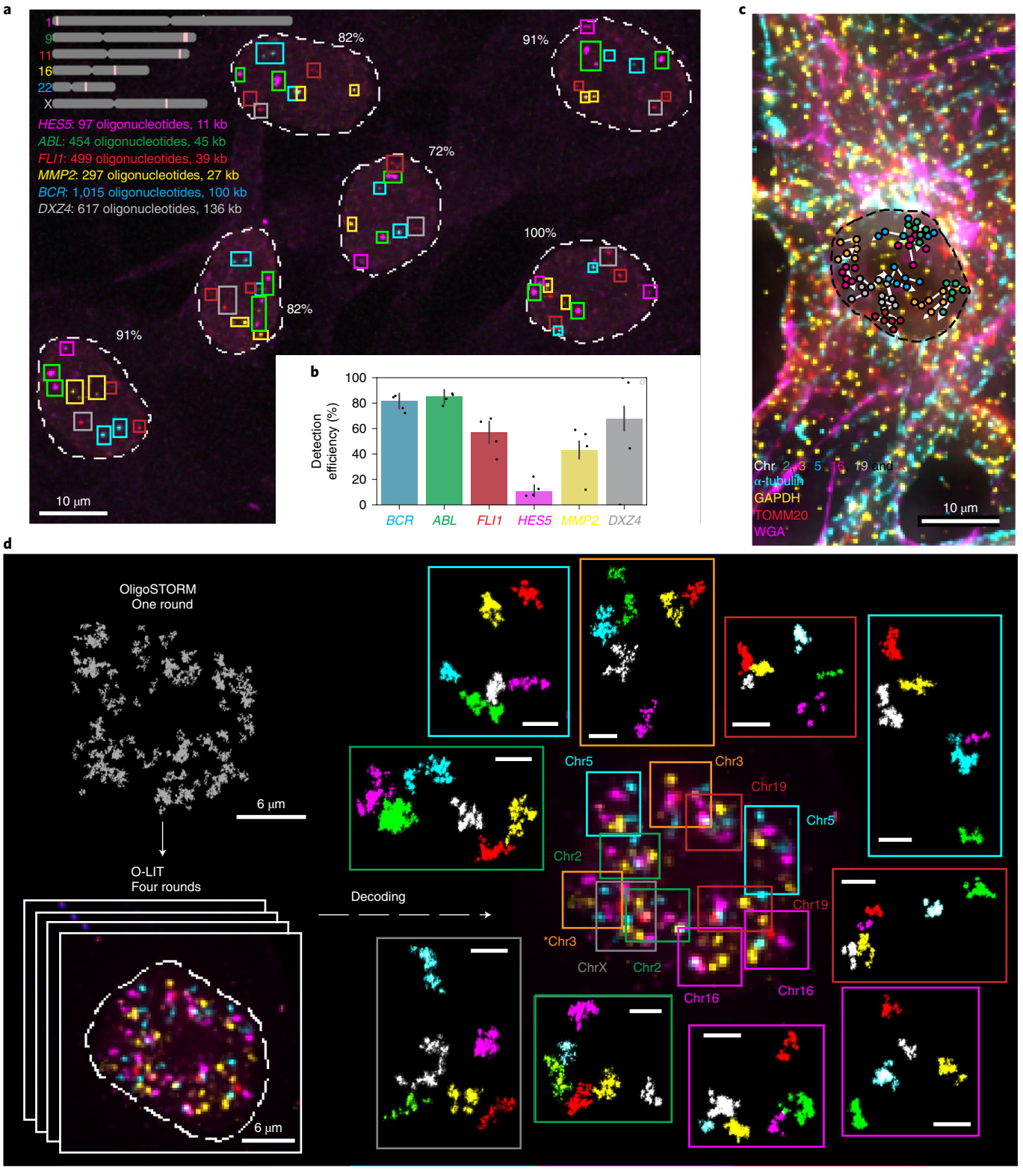
We next refined O-LIT so that it could target smaller genomic regions, as well as trace chromosomes at higher genomic resolution. However, because the commercial production of SOLiD reagents was discontinued at this juncture in our studies, we focused first on developing an alternative to the SOLiD reagents, the outcome of which was a method that ultimately improved signal detection. SOLiD chemistry reads sequences as dinucleotides using labeled eight-nucleotide oligonucleotides (TGNNNIII, where the first two positions represent all 16 dinucleotide combinations, positions 3–5 are degenerate and positions 6–8 are universal), thus entailing 1,024 (16×4^3) oligonucleotide species[24]. Because this level of complexity is excessive for O-LIT, where barcodes are defined by the user, we aimed to reduce the complexity of the oligonucleotide pool to the minimum necessary for decoding O-LIT barcodes, reasoning further that a minimally complex oligonucleotide pool might increase signal over background measurements. Thus, taking advantage of the universal base deoxyinosine[32], we reduced the complexity of the oligonucleotide pool from 1,024 to 4, referring to this strategy as 'just enough barcodes' (JEB) and the LIT chemistry using this strategy as eLIT (Fig. 4a,b). Application of OligoFISSEQ using eLIT (O-eLIT) to a library targeting 9,267 Oligopaint oligonucleotides to Chr19 (Chr19-9K) proved successful, yielding a 3.3-fold brighter signal-to-nuclear-background ratio as compared to the application of LIT to the same library using SOLiD oligonucleotides ($n=55$ cells for SOLiD and 57 cells for JEB from two replicates; Extended Data Fig. 5a,b).

Anticipating that the improved signal-to-nuclear background ratio would improve genomic resolution, we generated a library identifying smaller genomic regions (average of 173 kb) by directing Oligopaint oligonucleotides to only the first 1,000 of the 5,000 oligonucleotide targets defined by 36plex-5K for each designated genomic region (Extended Data Fig. 5c). Then, to benchmark this library, called 36plex-1K, against 36plex-5K, we adopted the same barcodes for 35 of 36 targets, with the exception being 5pR3, which was given a new barcode; 5pR3 had previously shared a barcode with 3qR3 to enable assessment of barcode detection across different regions. Five rounds of O-eLIT using only Mainstreet of 36plex-1K yielded a tier 2 barcode recovery efficiency of 74% ± 11.2% (48 of 66; $n=440$ cells from nine replicates), which was higher than that obtained with five rounds of O-LIT (54.6%, $n=41$; Fig. 4c,d and Extended Data Fig. 5c,d). O-eLIT of 36plex-1K gave homolog-resolved data (Extended Data Fig. 5e–i). These findings argued that O-eLIT would be useful genome wide. We recently imaged 249 regions with a genome-wide library (129plex) corresponding to 129 100-kb targets spanning all the autosomes (120 targets), ChrX (6 targets) and ChrY (3 targets). Five rounds of sequencing confirmed genome-wide capacity (Methods); although inadvertent barcode duplications complicated analyses, tier 2 can nevertheless detect 95% (165 of 174) of unique barcodes, while tier 1 can detect 44% (33 of 75) of duplicated barcodes (Fig. 4e and Supplementary Table 12).

Fine ChrX tracing and suggestions of chromosome signatures. To test the potential of O-eLIT to achieve finer genomic resolution, we applied an Oligopaint library, ChrX-46plex-2K, targeting 2,000 oligonucleotides to each of the 46 regions along the human X chromosome, the number of targets aligning with a previous study that used a hybridization-based Oligopaint strategy to image 40 regions of this chromosome[9]. The targets ranged in size from 253 kb to 1.22 Mb (average of 445 kb), with an average distance between targets of 2.75 Mb and total coverage of 20.6 Mb or 13.3% of the chromosome (Fig. 5a). As such, ChrX-46plex-2K served as an informative proxy for assessing the capacity of OligoFISSEQ to accommodate all other chromosomes. Here we applied O-eLIT to both streets and achieved a tier 2 barcode recovery efficiency of 74.3% ± 2.5% in PGP1f cells (~34/46 targeted regions, $n=177$ from seven replicates; Supplementary Fig. 2, Fig. 5b,c and Extended Data Fig. 6a,b), interpolating the positions of any target that had escaped detection (Methods). Although three targets were difficult to recover (X15, X19 and X31), the quality of the data nevertheless permitted 176 traces spanning the entirety of the X chromosome, single-cell spatial distance matrices and a population-based spatial distance matrix that was strongly correlated with a corresponding Hi-C map ($r=0.641$, $P=7.074\times10^{-245}$) and inversely correlated with Hi-C interaction frequencies ($r=-0.84$, $P=5.08\times10^{-275}$), the latter producing an exponential factor of 0.18 (Fig. 5d–g and Extended Data Fig. 6c–j), similarly to that observed previously[9]. Furthermore, the chromosome traces revealed two major clusters (Extended Data Fig. 7a–c; Calinski–Harabasz index of 213.71) that differed in their radii of gyration ($t=-10.1$; $P=3.9\times10^{-19}$; Extended Data Fig. 7d), one cluster consisting of 20 chromosomes (11%) and the other comprising 156 (89%) chromosomes. While the basis for this heterogeneity will require additional study, whether it is the cell cycle, chromatin accessibility and/or overall chromosome activity, these findings emphasize the potential of O-eLIT to advance understanding of the manner in which chromosomal material can be packaged and whether that packaging correlates with function.

Fig. 6 | OligoFISSEQ extensions and applications. a, O-eLIT detection of single-gene targets after sequencing off both streets. Colored squares mark gene targets identified after five rounds of sequencing. Values reflect the percentage of targets detected out of 11 (5 autosomal genes × 2, in addition to *DXZ4* on ChrX). Images are from deconvolved maximum-intensity *z*-projections and are representative of two replicates. **b**, Tier 1 target detection efficiency from the experiment in **a** ($n=61$ cells from two replicates). Tier 2 is inapplicable due to a lack of targets from the same chromosome. Detection efficiencies from individual replicates are plotted. Error bars represent the 95% bootstrap CI of the mean. **c**, Combining O-LIT and IF. 36plex-5K was sequenced for four rounds with O-LIT off both streets, followed by IF and staining with wheat germ agglutinin (WGA). Images are from deconvolved maximum-intensity *z*-projections with chromosome traces overlaid. $n=1$. **d**, 36plex-5K was hybridized to PGP1f cells and imaged with one round of OligoSTORM (2 h) to visualize all 66 regions simultaneously, followed by four rounds of O-LIT (2–3 h per round) to decode targets. OligoSTORM image showing the entire FOV with all unidentified targets (top left). Micrograph from the first round of O-LIT; image from deconvolved maximum-intensity *z*-projection (bottom left). All six chromosomes were identified and arrayed, in super-resolution, around the central nucleus (right; central image decorated with colored squares, color coded by chromosome). All 66 regions except for one region on Chr16 were detected and identified by O-LIT, with one homolog of Chr3 (asterisk) not captured by OligoSTORM because it fell outside the FOV. All scale bars for OligoSTORM images represent 1 μm.

Oligopaint libraries 36plex-5K and 36plex-1K have also enabled analyses of chromosome folding. Combining the two datasets (for 36plex-5K, $n=611$ cells from 15 replicates; for 36plex-1K, $n=440$ cells from 9 replicates), we evaluated the angles formed by the chromosomal segments flanking the centromeres (Extended Data Fig. 8a) and observed that only a minority, if any, of the chromosomes extend their p and q arms in polar opposite directions or are folded into a hairpin; median values for the angles ranged from 74° to 94° (Extended Data Fig. 8b,c). Furthermore, assessment of the angles formed by the two contiguous chromosomal segments lying within each arm (Extended Data Fig. 8a) showed that the p and q arm angles were significantly different for Chr2, Chr3, Chr16 and Chr19 ($n=686$, 668, 586 and 760, respectively; $P=4.15\times10^{-16}$, 0.004, 1.36×10^{-14} and 3.33×10^{-11}, respectively; Extended Data Fig. 8c). As



the larger angle was associated with the p (shorter) arm of Chr2 and Chr19 and with the q (longer) arm of Chr3 and Chr16, these findings cannot be explained solely by relative arm lengths. Consistent with this, arm angle and arm length were not significantly correlated ($r=0.26$, $P=0.42$; Extended Data Fig. 8d), leaving open the possibility that arm angles reflect the impact of centromere structure on flanking genomic regions and/or interdependence of the p and q arms, the constraints of chromosomal territories or other intrinsic organizational principles, Rabl configurations resulting from the last cell division and/or the state of gene activity, such as accessibilities underlying allelic skewing. Regardless of the reasons, these observations of X-chromosome conformations (Extended Data Fig. 7a–d) and arm angles (Extended Data Fig. 8a–d) demonstrate the potential of chromosome-wide imaging to address whether there are chromosome-level structural signatures, such as may be indicative of cell type, cell state and/or cellular health or age, with evidence from a recent study of two chromosomes in *Caenorhabditis elegans* aligning with these possibilities[21]. Chromosome organization may also reflect the evolutionary history of a chromosome[33,34]. The capacity of OligoFISSEQ to generate large datasets will facilitate the study of these potential paradigms of genome organization.

Single-gene identification, IF and acceleration of super-resolution imaging. OligoFISSEQ has proven versatile, capable of imaging single regions in the size range of tens of kilobases and accommodating IF, as well as accelerating super-resolution imaging (Fig. 6a,d, Extended Data Figs. 9a,b and 10a,b and Supplementary Fig. 4a). With respect to single regions, we applied O-eLIT to six genes ranging in size from 11 kb to 136 kb (Fig. 6a,b): *HES5* (11 kb, Chr1), *MMP2* (27 kb, Chr16), *FL11* (39 kb, Chr11), *ABL* (45 kb, Chr9), *BCR* (100 kb, Chr22) and *DXZ4* (136 kb, ChrX). Detection of the larger targets hovered between 43% and 80%, reaching as high as 83.7% ± 4.38% for *ABL* ($n=61$ cells from two replicates; Fig. 6b), and although detection of the smallest target *HES5* was low (9.82% ± 3.79%), with the incorporation of amplification strategies[35,36] we expect that detection of targets as small as, or even smaller than, *HES5* should become robust. Regarding IF, we conducted four rounds of O-LIT using 36plex-5K and sequencing off both streets, followed by immunocytochemical detection of antibodies directed against α-tubulin, GAPDH and TOMM20, and we were able to trace all six chromosomes, as well as obtain strong signals for all three proteins (Fig. 6c and Extended Data Fig. 10a). We have also applied ChrX-46plex-2K to IMR-90 human fibroblast cells (XX) and then distinguished the active X (Xa) from the inactive X (Xi) chromosome through IF to macroH2A.1, which preferentially binds the latter (Extended Data Fig. 9a–l). Xi displayed a lower radius of gyration ($P=9.07\times10^{-5}$; Extended Data Fig. 9h) and megadomain structures (Extended Data Fig. 9k,l), consistent with Hi-C and FISH studies[9,37–43] and further validating the use of O-eLIT for high-resolution chromosome tracing. Taken together, these findings confirm the potential of OligoFISSEQ to enable discoveries regarding the genome-wide spatial relationship between genes and their epigenetic partners.

Lastly, we demonstrated the capacity of OligoFISSEQ to improve the speed with which genomic regions can be imaged using single-molecule localization microscopy. Here we focused on OligoSTORM[5,10], which combines Oligopaints[4] with stochastic optical reconstruction microscopy[44], to provide super-resolution images of genomic regions in a space-filling fashion and thus reveal detailed volumetric structures[5,7,13,14,16]. The throughput of OligoSTORM, however, hovers at ten to a few hundred cells per experiment, with imaging times of up to 2 h. In contrast, because OligoFISSEQ can be carried out with diffraction-limited microscopy, it has the capacity to image hundreds to thousands of cells per experiment, with relatively negligible imaging times. Thus, we explored the possibility of accelerating super-resolution genome imaging by combining O-LIT with OligoSTORM (Fig. 6d, Extended Data Fig. 10b and Supplementary Fig. 4a).

First, using 36plex-5K and bridge oligonucleotides containing binding sites for secondary oligonucleotides conjugated with a fluorophore suitable for OligoSTORM (Alexa Fluor 647), we captured all 66 targets simultaneously in a single 2-h round of OligoSTORM (Fig. 6d and Extended Data Fig. 10b; see also Chr2-6plex in Supplementary Fig. 4a). Then, with only four rounds of O-LIT, we identified all 66 regions. Thus, by combining OligoSTORM with OligoFISSEQ, we enabled a 36-fold reduction in imaging time and data storage demands (from ~2.73 TB to ~76 GB; Fig. 6d), while achieving 17 nm ± 5 nm of lateral precision and 50 nm ± 10 nm of axial precision, and 40 nm ± 5 nm of lateral resolution and 60 nm ± 5 nm of axial resolution. Extrapolating to all 46 chromosomes of a diploid human nucleus and anticipating many more than six targets per chromosome, this study demonstrates the feasibility of simultaneously 'OligoSTORMing' hundreds of regions of the genome. O-LIT should also permit OligoSTORM walking along the genome, with many walks per nucleus. Previously, we accomplished multi-walk imaging through temporal barcoding[14]. Here, multiple rounds of OligoSTORM could produce super-resolved walks in multiple regions of the genome, simultaneously, after which all regions could be identified with O-LIT. In summary, given the potential of O-LIT to identify hundreds to perhaps thousands of regions, OligoSTORM should scale similarly.

Discussion

There is a growing need for methods that will enable the imaging of entire genomes at high genomic and optical resolution while also supporting the levels of throughput and reproducibility that are becoming increasingly essential for understanding biological entities as dynamic as the genome. To this end, we have described OligoFISSEQ, a set of three methods for in situ genome mapping, demonstrating the potential of these methods to scale toward whole-genome imaging. OligoFISSEQ also has the capacity to meld with other technologies and thus extend its usefulness further. For example, when combined with homolog-specific Oligopaints (HOPs)[5], it should enable genome-wide studies in the context of parent-of-origin and, with adjustments to the barcodes, OligoFISSEQ could also enable multiplexed and/or multicolor visualization of chromosome folding in combination with other technologies, such as OligoDNA-PAINT[5], Hi-M[17] and optical reconstruction of chromatin architecture (ORCA; ref. [20]). In terms of scaling, our capacity to map 46 regions on ChrX at ~1 genomic target per 2.75 Mb predicts that OligoFISSEQ could accommodate a thousand or more targets in human nuclei, with the potential to increase that number through a reduction in target size, temporal barcoding to better resolve targets, additional rounds of sequencing and incorporation of expansion microscopy[45]; preliminary studies show that Chr19-9K can support eight rounds of O-LIT (Extended Data Fig. 10c) and that OligoFISSEQ is feasible in the context of hydrogels (Extended Data Fig. 10d,e). Scaling could also be enhanced via microfluidics, which would significantly reduce the time required for each round of sequencing by 15–20%. Indeed, with the advent of improved enzymatics, methods for amplifying signal (for example, SABER[35] and ClampFISH[36]) and superior imaging, OligoFISSEQ should become applicable to the study of smaller targets, such as enhancers and promoters. As important will be improvements in image analysis. For example, implementation of point spread function-fitting algorithms should improve spatial resolution and thus scalability[46], while a reduction in the dependence on the proximity of signals to affirm true signal would permit better detection of chromosome rearrangements, where targets that are expected to be near each other are instead widely separated. Finally, OligoFISSEQ should interface beautifully with other FISSEQ-based technologies to achieve multi-omic views of the genome, with each round of sequencing visualizing DNA, RNA[22,23] and protein[47] simultaneously.

We note that, as OligoFISSEQ has the capacity for significant genome coverage and the potential to consistently identify the same targets across thousands of cells, it is well suited for studying variability at a handful of regions as well as addressing this challenging topic at the level of the entire genome. Structural variability of specific genomic features has now been widely observed[7,9,11,13,14,16,17,20,21,48,49], and, while often thought of locally, the impact of this structural variability may reach globally[14]. Even a minor, seemingly inconsequential change in one part of the nucleus may have a profound 'butterfly effect' (ref. [50]) on the global scale, with its impact potentially contributing to and/or propagating gene regulatory states and phase separations, perhaps even constituting essential, potentially heritable signatures of the genome. Thus, although variability may appear random at the local level, a genome-wide perspective may reveal that apparent randomness actually reflects global responsiveness and an exquisitely controlled regulatory program that directs structural conformations across the entire nucleus, as much the outcome of evolution as any other honed genetic function.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-020-0890-0.

Received: 15 December 2019; Accepted: 8 June 2020;
Published online: 27 July 2020

References

1. Hu, Q., Maurais, E. G. & Ly, P. Cellular and genomic approaches for exploring structural chromosomal rearrangements. *Chromosome Res.* **28**, 19–30 (2020).
2. Bolzer, A. et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* **3**, e157 (2005).
3. Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889 (2010).
4. Beliveau, B. J. et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl Acad. Sci. USA* **109**, 21301–21306 (2012).
5. Beliveau, B. J. et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147 (2015).
6. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
7. Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
8. Shah, S., Lubeck, E., Zhou, W. & Cai, L. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* **92**, 342–357 (2016).
9. Wang, S. et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).
10. Beliveau, B. J. et al. In situ super-resolution imaging of genomic DNA with OligoSTORM and OligoDNA-PAINT. *Methods Mol. Biol.* **1663**, 231–252 (2017).
11. Cattoni, D. I. et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat. Commun.* **8**, 1753 (2017).
12. Eng, C.-H. L., Shah, S., Thomassie, J. & Cai, L. Profiling the transcriptome with RNA SPOTs. *Nat. Methods* **14**, 1153–1155 (2017).
13. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
14. Nir, G. et al. Walking along chromosomes with super-resolution imaging, contact maps and integrative modeling. *PLoS Genet.* **14**, e1007872 (2018).
15. Rosin, L. F., Nguyen, S. C. & Joyce, E. F. Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in *Drosophila* nuclei. *PLoS Genet.* **14**, e1007393 (2018).
16. Szabo, Q. et al. TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci. Adv.* **4**, eaar8082 (2018).
17. 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 (2019).
18. Eng, C.-H. L. et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019).
19. Fields, B. D., Nguyen, S. C., Nir, G. & Kennedy, S. A multiplexed DNA FISH strategy for assessing genome architecture in *Caenorhabditis elegans*. *eLife* **8**, e42823 (2019).
20. Mateo, L. J. et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
21. Sawh, A. N. et al. Lamina-dependent stretching and unconventional chromosome compartments in early *C. elegans* embryos. *Mol. Cell* **78**, 96–111 (2020).
22. Ke, R. et al. In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013).
23. Lee, J. H. et al. Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014).
24. Metzker, M. L. Sequencing technologies—the next generation. *Nat. Rev. Genet.* **11**, 31–46 (2010).
25. Shendure, J. et al. DNA sequencing at 40: past, present and future. *Nature* **550**, 345–353 (2017).
26. Player, A. N., Shen, L.-P., Kenny, D., Antao, V. P. & Kolberg, J. A. Single-copy gene detection using branched DNA (bDNA) in situ hybridization. *J. Histochem. Cytochem.* **49**, 603–612 (2001).
27. Heride, C. et al. Distance between homologous chromosomes results from chromosome positioning constraints. *J. Cell Sci.* **123**, 4063–4075 (2010).
28. Mayer, W., Smith, A., Fundele, R. & Haaf, T. Spatial separation of parental genomes in preimplantation mouse embryos. *J. Cell Biol.* **148**, 629–634 (2000).
29. Hua, L. L. & Mikawa, T. Mitotic antipairing of homologous and sex chromosomes via spatial restriction of two haploid sets. *Proc. Natl Acad. Sci. USA* **115**, E12235–E12244 (2018).
30. Reichmann, J. et al. Dual-spindle formation in zygotes keeps parental genomes apart in early mammalian embryos. *Science* **361**, 189–193 (2018).
31. Joyce, E. F., Erceg, J. & Wu, C.-t. Pairing and anti-pairing: a balancing act in the diploid genome. *Curr. Opin. Genet. Dev.* **37**, 119–128 (2016).
32. Watkins, N. E. & SantaLucia, J. Jr. Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. *Nucleic Acids Res.* **33**, 6258–6267 (2005).
33. Yunis, J. J. & Prakash, O. The origin of man: a chromosomal pictorial legacy. *Science* **215**, 1525–1530 (1982).
34. Tjong, H. et al. Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. *Proc. Natl Acad. Sci. USA* **113**, E1663–E1672 (2016).
35. Kishi, J. Y. et al. SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. *Nat. Methods* **16**, 533–544 (2019).
36. Rouhanifard, S. H. et al. ClampFISH detects individual nucleic acid molecules using click chemistry–based amplification. *Nat. Biotechnol.* **37**, 84–89 (2019).
37. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
38. Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
39. Deng, X. et al. Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* **16**, 152 (2015).
40. Minajigi, A. et al. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276 (2015).
41. Darrow, E. M. et al. Deletion of *DXZ4* on the human inactive X chromosome alters higher-order genome architecture. *Proc. Natl Acad. Sci. USA* **113**, E4504–E4512 (2016).
42. Giorgetti, L. et al. Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).
43. Wang, C.-Y., Jégu, T., Chu, H.-P., Oh, H. J. & Lee, J. T. SMCHD1 merges chromosome compartments and assists formation of super-structures on the inactive X. *Cell* **174**, 406–421 (2018).
44. Rust, M. J., Bates, M. & Zhuang, X. W. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).
45. Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* **347**, 543–548 (2015).
46. Sage, D. et al. Super-resolution fight club: assessment of 2D and 3D single-molecule localization microscopy software. *Nat. Methods* **16**, 387–395 (2019).
47. Kohman, R. E. & Church, G. M. Fluorescent in situ sequencing of DNA barcoded antibodies. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.27.060624 (2020).
48. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515 (2019).
49. Finn, E. H. & Misteli, T. Molecular basis and biological function of variability in spatial genome organization. *Science* **365**, eaaw9498 (2019).
50. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).

Methods

Materials. Lists of reagents and catalog numbers (Supplementary Table 1), oligonucleotide sequences (Supplementary Tables 2–8 and 16) and library information such as coordinates, barcodes and density (Supplementary Table 12) are presented as supplementary information.

Oligopaint library design. All Oligopaint oligonucleotide sequences and coordinates for libraries used in this study can be found in Supplementary Tables 2–6. Oligopaints[4] leverages the ability to computationally design and synthesize sequence-specific oligonucleotide probes for FISH[4] (see Supplementary Note 1 for additional examples). Oligopaint FISH probes were computationally designed for optimal hybridization and high specificity. Oligopaint genome-binding sequences were obtained from the Oligopaints website (https://oligopaints.hms.harvard.edu/; ref. [51]), using the hg19 genome with 'Balanced' settings. 129plex sequences were obtained using OligoMiner on soft-masked hg38 sequence using a Tm window of 42–47 °C and a length range of 30–37 nucleotides[51]. Genome homology sequences of other libraries ranged from 35–41 nucleotides. Universal forward- and reverse-priming sequences were appended to each Oligopaint oligonucleotide using OligoLEGO (https://github.com/gnir/OligoLego/), allowing the libraries to be PCR amplified and renewable. The universal priming sequences also served as various OligoFISSEQ primer and bridge sites. Each library used in this study was designed with specific features and is described in detail in the supplementary file specific for each set.

LIT. For the Chr19-20K library, a portion of the universal forward-priming sequence was used as the LIT primer binding site, followed by the LIT barcode. Barcode and color-code designation was as follows: 4, Cy5/Alexa Fluor 647; 3, TxRd; 2, Cy3; 1, FITC/Alexa Fluor 488.

The 36plex-5K library shared the same universal forward-priming sequence among all oligonucleotides and contained chromosome-specific universal reverse-priming sequences. Individual chromosome targets could be amplified, hybridized and detected by using the universal reverse-priming sequence. Universal forward-priming sequences were used as LIT primer binding sites for 18-nucleotide primers. In cases where O-LIT was performed off both Mainstreet and Backstreet, a LIT primer binding site was hybridized to the Backstreet. Barcodes were specified using sequences from OligoLego (https://github.com/gnir/OligoLego/). Candidate barcode sequences were decoded to reveal color codes using a MATLAB script (https://www.mathworks.com/). To maintain color-code diversity between neighboring targets, barcodes were manually assigned to targets (for example, barcodes were specified so that neighboring targets would have different colors in the first round). Each LIT barcode digit required a five-nucleotide sequence, while the last barcode digit required eight nucleotides to allow adequate space for 8-mer binding. Thus, a four-digit barcode required 23 nucleotides in total. For 36plex-5K, the targets 3qR3 and 5pR2 contained the same barcode sequences to assess barcode recovery from separate genomic targets.

JEB/O-eLIT barcodes. The 36plex-1K library selected a subregion of 36plex-5K targets, with 1,000 Oligopaint oligonucleotides per target instead of 5,000 oligonucleotides. Additionally, 36plex-1K targets contained JEB-compatible barcode digits. The 36plex-1K targets contained the same barcode digit color coding as for 36plex-5K, with the exception of 5pR3. 36plex-1K could only be sequenced using Mainstreet and not both streets.

The ChrX-46plex library was designed to span the entire human X chromosome with 2,000 Oligopaint oligonucleotides per target. The library was divided into two sublibraries, ChrX-23plex-odd and ChrX-23plex-even, with each sublibrary targeting either odd (X1, X3, X5, ...) or even (X2, X4, X6, ...) targets. Each sublibrary contained the same universal forward-priming sequences and different universal reverse-priming sequences. ChrX-46plex barcodes contained JEB digits and were also manually assigned to maintain color-code diversity between neighboring targets. ChrX-46plex is compatible with sequencing off both streets.

The six-gene library shared the same universal forward-priming sequence and different universal reverse-priming sequences. Barcodes were manually specified using JEB digits. The six-gene library is compatible with sequencing off both streets.

The 129plex genome-wide library aims at imaging each chromosome arm of the human genome using OligoFISSEQ. We selected the regions based on the density of Oligopaint oligonucleotides that could be targeted (average, 8.6 oligonucleotide targets per kb) and position on the chromosome arm. First, using a custom-curated R script, we used a sliding window of 100 kb along all chromosomes to calculate oligonucleotide target densities. Then, wherever possible, we selected three regions for each chromosome arm: one near the telomere, another near the centromere, and a third more centrally located, selecting regions where the density of oligonucleotide targets would be above 6 per kb. For some chromosome arms, we selected fewer than three regions owing to the constraints of oligonucleotide target density. Each region corresponded to a 5-digit barcode. The 129plex was sequenced off both streets. Due to 21 inadvertently duplicated barcodes, 42 of the targets could not be assigned (Supplementary Table 12).

SIT barcode. For the Chr19-20K library, the universal reverse-priming sequence was used as the SIT primer binding site, followed by the SIT barcode sequence. Barcode and color-code designation was as follows: 4, Cy5; 3, Cy5 + Cy3; 2, Cy3; 1, blank. For 36plex-1K, the universal reverse-priming sequence was used as the SIT primer binding site, followed by SIT barcodes. Target color coding was designed to be the same as for 36plex-5K but with SIT reagents.

HIT barcode. For the Chr19-20K library, bridging oligonucleotides (HIT bridges) were designed to hybridize to Mainstreet and Backstreet. HIT bridges contained binding sites for HIT readout oligonucleotides. HIT readout oligonucleotide sequences were derived from OligoLego. Barcode and color-code designation was as follows: 0, blank; 1, Alexa Fluor 647/Cy5; 2, Cy3B/Cy3; 3, FAM/Alexa Fluor 488.

For the 36plex-5K library, HIT bridges were designed to hybridize to street-specific sequences for each target. This was done by designing bridges flanking universal priming sites (forward and reverse), as well as the 5′ or 3′ ends of LIT barcodes, due to similar LIT barcodes being present on both streets. HIT bridges contained binding sites for HIT readout oligonucleotides derived from OligoLego.

Oligopaint probe synthesis. Oligopaint oligonucleotides were purchased as single-stranded oligonucleotide pools from CustomArray (http://www.customarrayinc.com/oligos_main.htm/) or Twist Bioscience (https://www.twistbioscience.com/) in 12,000 and 92,000 chip formats. Oligonucleotide pools were amplified as previously described[10,14] with minor modifications (a step-by-step protocol can be found in Supplementary Protocol 1). Briefly, PCR conditions for each library and sublibrary were optimized using real-time PCR to obtain optimal template concentration, primer concentration and annealing temperature. Next, libraries were linearly amplified with low-cycle PCR using Kapa Taq reagents. dsDNA PCR products were purified using Zymo columns and eluted with ultra-pure water (UPW). T7 RNA promoter sequence was then appended to Oligopaints using REV primers containing the T7RNAP on the 5′ end. Note that some users may opt to add the T7RNAP straight from the raw library. dsDNA PCR products were purified using Zymo columns and eluted with UPW. PCR products were then in vitro transcribed using HiScribe (NEB, E2040S) overnight at 37 °C to make RNA.

RNA products were reverse transcribed with Thermo Maxima H Minus Reverse Transcriptase (Thermo Fisher, EP0753) to make cDNA. RNA was then digested to leave single-standed DNA. This product was purified using Zymo columns. Final single-standed DNA Oligopaint oligonucleotides were resuspended at 100 µM in UPW and stored at −20 °C until use. Linear PCR, touched-up PCR and single-standed DNA Oligopaint oligonucleotides were quality checked by running on 2% agarose DNA gels to confirm single bands were migrating at the expected sizes during synthesis.

Other oligonucleotides. Sequences for all other oligonucleotides can be found in Supplementary Tables 7 and 8. Primers, secondary fluorophore-labeled oligonucleotides, LIT sequencing primers, SIT sequencing primers, JEB oligonucleotides and molecular inversion probes were purchased from IDT (https://www.idtdna.com/). HIT secondary oligonucleotides were purchased from Bio-Synthesis (https://www.biosyn.com/). Alexa Fluor 405 activator fluorophore was purchased from Thermo Fisher (https://www.thermofisher.com/).

Cell culture. Our study used two human cell lines: PGP1f and IMR-90. PGP1f cells are primary human fibroblasts taken from the PGP1 male donor from the personal Genome Project (Coriell, GM23248; ref. [52]). They were previously found to be of normal karyotype[14,53]. PGP1f cells were cultured in DMEM (Gibco) supplemented with 10% FBS (Thermo Fisher; A3160401), 1× penicillin–streptomycin (Thermo Fisher, 15140122) and 1× nonessential amino acids (Thermo Fisher, 11140050). PGP1f cells were cultured for no more than five passages before thawing new cultures. IMR-90 cells were cultured in DMEM supplemented with 10% FBS and 1× penicillin–streptomycin. Cells were cultured at 37 °C in a 5% CO_2 incubator.

Sample preparation for OligoFISSEQ. Ibidi Sticky Slide VI (https://ibidi.com/, 80608) was used for all experiments except for metaphase spreads (Fig. 2b) and hydrogels (Extended Data Fig. 10d,e). Ibidi slides were assembled and allowed to cure overnight at 37 °C before use. Each well required 100–200 µl of reagent, and we generally designated one hole as the inlet and the other hole as the outlet. PGP1f cells from ~70% confluent 10-cm dishes were detached from the dishes using 1 ml of trypsin (Thermo Fisher, 25-200-056), neutralized with 2–3 ml of fresh medium. Next, 100 µl of cells in suspension was added to each Ibidi well and allowed to adhere and recover overnight at 37 °C in an incubator. The following day, the medium was aspirated and cells were washed with 1× PBS and fixed for 10 min with 4% formaldehyde (Electron Microscopy Sciences, 15710) in a final concentration of 1× PBS (Thermo Fisher, 10010-023). Fixative was removed and cells were rinsed with 1× PBS. Cells were then permeabilized with 0.5% Triton X-100 (Sigma-Aldrich, T8787-250ML) in a final concentration of 1× PBS for 15 min on a rotator. Permeabilization reagent was aspirated and cells were rinsed in 0.1% Triton/1× PBS and stored in either this or PBS at 4 °C until use. Samples were used within 2–3 weeks of fixation.

Cell samples for the molecular inversion probe and hydrogel experiments were grown on rectangular glass microscope slides. Cells were plated similarly to the Ibidi slides, except 150 µl of cells in suspension was plated onto discrete areas on rectangular slides (previously etched with a glass etching pen to note the region) and incubated overnight at 37 °C in a 10-cm petri dish. The following day, the same steps were performed as with Ibidi slides but in 50-ml Coplin jars. Cells were stored in 1× PBT in Coplin jars until use. Metaphase spreads were purchased from Applied Genetics (product: HMM).

DNA FISH. Step-by-step protocols can be found in Supplementary Protocols 2 and 3, which were adapted from Beliveau et al. [4] and based on previous studies[54,55]. All OligoFISSEQ methods begin with hybridization of primary Oligopaint libraries overnight and then deviate. The following steps are common to LIT, SIT and HIT with Ibidi slides (all steps were completed on a rotator unless specified otherwise). Ibidi wells were washed with 0.1% PBT at room temperature for 5 min and incubated with 0.1 N HCl for 8 min. Two SSCT washes were performed. Cellular RNA was digested with 50 µl of 2 µg ml^{-1} RNase A (Thermo Fisher, EN0531) in 2× SSCT for each well. Slides were incubated in 37 °C in a humidified chamber for 1 h. RNase A was washed out by adding 2× SSCT. Prehybridization began by adding 50% formamide/2× SSCT for 10 min at room temperature. Prehybridization continued with prewarmed (60 °C) 50% formamide/2× SSCT added, and the slide was placed on top of the heat block set in a 60 °C water bath for 20 min. Next, the primary Oligopaint library was added, the samples were aspirated and 50 µl total of primary Oligopaint oligonucleotide library (2 µM final concentration) was added in hybridization mix (50% formamide, 2× SSCT and 10% dextran sulfate). Samples with primary Oligopaint oligonucleotide libraries were then denatured, wells were sealed with parafilm to prevent evaporation and the slide was placed on a preheated hot block in an 80 °C water bath for 3 min under the weight of a rubber plug. Oligopaint oligonucleotide library hybridization to samples was performed by placing samples in a humidified chamber at 42 °C to incubate for >16 h. The next day, probes that did not hybridize were washed out by adding prewarmed (60 °C) 2× SSCT directly to each well containing primary hybe mix and were then aspirated. New prewarmed 2× SSCT was added and samples were incubated on a hot block for 15 min. This was repeated once and then again at room temperature. After this wash, the protocol deviates for the techniques (see below). Note that cellular DNA was stained after every two rounds of sequencing to maintain adequate DAPI signal.

For detection of Oligopaints via secondary hybridization, samples were then prepared for secondary oligonucleotide hybridization to primary oligonucleotide streets for detection. Samples were washed with 30% formamide/2× SSCT for 8 min and 50 µl in total of secondary oligonucleotides and/or bridge oligonucleotides was added at 1.2 µM in 30% formamide/2× SSCT to each well. Samples were incubated in a humidified chamber for 45 min at room temperature in darkness. Nonhybridized secondary oligonucleotides were washed out with 30% formamide/2× SSCT added directly to the samples, which were then aspirated and incubated twice for 15 min on a rotator. Samples were washed twice with 2× SSCT for 5 min. In some experiments, DNA was counterstained with DAPI (Thermo Fisher, D1306) in PBS for 10 min. Samples were then washed with 1× PBS twice for 5 min and imaged in 1× PBS or imaging buffer containing PBS, PCD, PCA and Trolox (Supplementary Protocol 3).

For cells on rectangular slides, the same overall protocol as above was performed but in Coplin jars, and wash volumes were scaled accordingly (25-µl volumes for primary and secondary hybridizations). The protocol was modified as follows: RNase was added directly to cells on a rectangular slide, which was covered with a 22 × 22 mm^2 coverslip. Post-RNase washes were performed by transferring the slide and coverslip to a Coplin jar and 'sliding' the coverslip off. The same approach was used for secondary hybridization. Primary Oligopaint hybridization was performed by adding primary Oligopaint mix directly to cells on a rectangular slide, covering with a 22 × 22 mm^2 coverslip and sealing the edges with rubber cement (Elmer's). Rubber cement was allowed to dry for 3 min and the sample was denatured on a heat block, similar to the process for Ibidi slides.

LIT. LIT is built upon Oligopaint[4], SBL[56] and FISSEQ technologies[22,23,57] (see Supplementary Note 2 and Supplementary Protocol 3 for recent iterations and the step-by-step protocol). After hybridization of the primary Oligopaint library, samples for O-LIT required treatment with phosphatase to deplete endogenous phosphates that could prime ligation, contributing to background and poor signal. The samples were washed with 50 µl of 1× NEB CutSmart buffer for 8 min. Next, 50 µl of shrimp alkaline phosphatase (rSAP; NEB, M0371L; 7.5 µl rSAP in 1× CutSmart) was added to each well followed by incubation at 37 °C with humidity for 1 h. To inactivate phosphatase, the sample was then transferred to a preheated heat block in a 65 °C water bath for 5 min and washed twice with preheated (65 °C) 2× SSCT on the heat block for 5 min each. The slides were washed for 5 min in 2× SSCT at room temperature. Samples were then prepared for LIT primer binding by washing with 30% formamide/2× SSCT for 8 min, and 50 µl of LIT sequencing primer was added at 1.2 µM in 30% formamide/2× SSCT to each well. Samples were incubated in humidified chambers for 45 min. Nonhybridized LIT primers were washed out with 30% formamide/2× SSCT being washed directly in, aspirated and incubated twice for 15 min on a rotator. Samples were washed with 2× SSCT twice for 5 min. Next, samples were prepared for the first round of LIT by adding 100 µl of 1× Quick Ligation buffer (NEB, B6058S) for 8 min and aspirated. LIT reaction mix (see Supplementary Protocol 3 for the recipe) was prepared on ice. Before adding ligases, vigorous vortexing was performed on the LIT reaction mix. After vortexing, ligases were added and mixed thoroughly by pipetting. O-eLIT reagent was performed similarly but, instead of SOLiD purple reagent mix, 40 pmol of each JEB oligonucleotide was added to each sample and UPW was adjusted accordingly. Next, 100 µl of this mix was added to each well and samples were incubated in a humidified chamber at 25 °C for 55 min. LIT reaction mix was then aspirated and samples were rinsed with 1 M guanadine hydrochloride (GHCL; Sigma-Aldrich, G3273) and washed twice for 15 min on a rotator at room temperature. Samples were washed in 1× PBS for 5 min. Cellular background fluorescence was reduced by treating the samples with 100 µl True Black (Biotum, 23007) in 70% ethanol for 2 min. Three 1× PBS quick rinses and a 10-min wash were performed. Samples were then imaged in 1× PBS or imaging buffer (see Supplementary Protocol 3 for recipe). Before proceeding to the next LIT round, nonligated phosphates were treated with phosphatase (Quick CIP; NEB, M0508L) for 30 min at 37 °C. Quick CIP was then washed out with three GHCL washes for 5 min. The previous LIT round was cleaved to release the fluorophore and regenerate the 5′ phosphate by rinsing and incubation for 15 min at room temperature on a rotator with cleave 1, followed by the same for cleave 2. Samples were then rinsed three times with GHCL and washed twice for 5 min. The next round of LIT could proceed with the pre-ligation step. After the last barcode digit was read, the fluorophore was cleaved and all targets were detected by hybridizing specific bridges and fluorophores as described above.

SIT. SIT is based on Oligopaint[4] and SBS[58] technologies using the Illumina NextSeq 500/550 TG Kit (Illumina, TG-160-2002). After hybridization of the primary Oligopaint library, samples were prepared for SIT primer binding by washing with 30% formamide/2× SSCT for 8 min, and 50 µl of LIT sequencing primer was added at 1.2 µM in 30% formamide/2× SSCT to each well. Samples were incubated in humidified chambers for 45 min. Nonhybridized SIT primers were washed out with 30% formamide/2× SSCT, which was added directly to the samples, aspirated and incubated twice for 15 min on a rotator. Samples were washed with 2× SSCT twice for 5 min. The first round of SIT proceeded by rinsing with 100 µl of prewarmed (60 °C) NextSeq polymerase solution (from reservoir 31) and then incubation on a 60 °C heat block in a water bath for 5 min. The samples were aspirated and washed with 2× SSCT three times for 10 min. The samples were washed in 1× PBS and then imaged in 1× PBS or imaging buffer. Before proceeding onto the next SIT round, samples were treated with NextSeq cleave solution (from reservoir 29) with a rinse and then incubated for 5 min on a 60 °C heat block in a water bath. Samples were then washed three times for 10 min in 2× SSCT. The next round of SIT could then proceed. For all target identification, SIT primers containing Alexa Fluor 488 were used, or secondary oligonucleotides with bridges were added.

HIT. HIT is based on Oligopaint[4] and SBH technologies[6,12,59]. After hybridization of the primary Oligopaint library, samples were prepared for HIT bridge oligonucleotide hybridization to primary oligonucleotide streets for detection. HIT bridges for 36plex-5K were designed to span the universal priming region and part of either the Mainstreet barcode or Backstreet barcode. Samples were washed with 30% formamide/2× SSCT for 8 min, and 50 µl of bridge oligonucleotides was added at 1.2 µM in 30% formamide/2× SSCT to each well. Samples were incubated in humidified chambers for 45 min at room temperature in darkness. Nonhybridized bridge oligonucleotides were washed out with 30% formamide/2× SSCT, which was added directly to the samples, aspirated and incubated twice for 15 min on a rotator. The first round of HIT commenced with the addition of 50 µl to each well with HIT secondary oligonucleotides specific to each round added at 1.2 µM in 30% formamide/2× SSCT for 45 min at room temperature in a dark humidified chamber. Nonhybridized HIT secondary oligonucleotides were washed out with 30% formamide/2× SSCT, which was added directly to the samples, aspirated and incubated twice for 15 min on a rotator. Samples were washed with 2× SSCT twice for 5 min and then with 1× PBS for 5 min. Samples were imaged in 1× PBS or imaging buffer. Before proceeding to the next round, the secondary oligonucleotide fluorophores from the previous HIT round were cleaved via rinsing and incubation for 15 min with 1 mM TCEP (Sigma-Aldrich, 646547-10×1ml). Samples were rinsed three times with PBS and the next HIT round commenced.

Immunofluorescence. To visualize proteins, samples were subjected to IF. After OligoFISSEQ, Oligopaint oligonucleotides were removed by washing with 80% formamide/2× SSCT twice for 7 min. Next, samples were washed with 2× SSCT for 3 min, rinsed with 1× PBS and fixed in 4% formaldehyde/PBS for 10 min. After PBS rinses and permeabilization in 0.5% Triton/PBS for 10 min, samples were blocked in 3% BSA/PBT for 1 h. Primary antibodies diluted in 1% BSA/PBT were then added to each well, and wells were sealed with parafilm and incubated overnight at 4 °C for >12 h. The next day, primary antibodies were removed and three PBT washes were performed. Secondary antibodies (Supplementary Table 1) diluted in 1% BSA/PBT were then added at a 1:500 dilution for each, for 1 h at room temperature on a shaker. WGA (Thermo Fisher, W11261; 1:20) was also added during the second incubation step. Three PBT washes for 5 min each were

performed, and samples were restained with DAPI (1:1,000) for 10 min and imaged in imaging buffer.

Hydrogel. Hydrogel embedding was based on work by Moffitt et al.[60] (see Supplementary Protocol 4 for the step-by-step protocol). Cells for hydrogel embedding were grown on rectangular glass slides. FISH was performed on these slides as described in 'DNA FISH'. After primary Oligopaint library hybridization, samples were washed at 60 °C in 2× SSCT for 20 min, then for 10 min at room temperature and then with 1× PBS for 5 min. In preparation for hydrogel embedding, slides were air-dried for 5 min and the area around cells was wiped dry with a Kimwipe. Hydrogel reagents were combined in Eppendorf tubes on ice and degassed on ice in a vacuum chamber (Thermo Fisher, 08-642-7) during incubations. Cells were then washed for 10 min at 4 °C with hydrogel mix without APS and TEMED. Hydrogel mix was then removed from samples, and ~20 µl of hydrogel solution (recipe in Supplementary Protocol 4) was spotted onto parafilm on a gelation chamber slide (rectangular slide wrapped in parafilm, using two 22 × 22 mm² coverslips as spacers on each end of the slide), then the slide sample was flipped onto hydrogel solution/gelation chamber, being careful to spread the hydrogel solution without forming bubbles. The sample was then incubated at 37 °C for 1 h in a vacuum chamber. After incubation, the gelation chamber was carefully removed. The edges of the hydrogel disc were trimmed, and a diamond etching pen was used to break the rectangular slide, preserving the gel/glass slide portion. The gel/glass slide portion was then transferred to a 35-mm petri dish and digested in 2 ml of digestion buffer (recipe in Supplementary Protocol 4; ref. [60]) overnight at 37 °C. After overnight digestion, the cell/hydrogel dissociates from the glass slide, so extra care was taken to avoid hydrogel damage. The digestion buffer and glass slide were removed, and the hydrogel was washed in 2× SSCT three times for 20 min each. The hydrogel was divided into smaller pieces for downstream applications. To note orientation, hydrogel pieces were cut into distinct shapes, to facilitate imaging and alignment downstream. After cutting, the hydrogel sample was transferred to 1.5-ml Eppendorf tubes for easier handling.

Metaphase FISH. Unless otherwise stated, all steps were performed using Coplin jars. Treatment commenced by adding 25 µl RNase A to the slides, sandwiching under a 22 × 22 mm² coverslip and then incubating in a humidified chamber. Primary Oligopaint hybridization was performed in the same way.

Diffraction-limited microscopy. OligoFISSEQ and diffraction-limited microscopy were carried out using a widefield epifluorescence setup. A Nikon Eclipse Ti body was equipped with a 60× 1.4-NA Plan Apo lambda objective (Nikon MRD01605), Andor iXon Ultra EMCCD camera (DU-897U: 512 × 512 pixel FOV, 16-µm pixel size), X-Cite 120 LED Boost light source, motorized stage and off-the-shelf filter sets from Chroma (~488 nm 49308 C191880; ~532 nm 49309 C191881; ~594 nm 49310 C191882; and ~647 nm 49009 C177216). Images were obtained with ND4 and ND8 filters in place. Microscope operation was handled by Nikon NIS Elements software. In general, *z*-stacks were obtained with 0.3-µm slices with an exposure time of 200–300 ms and LED intensity of 20–60%, depending on the library being imaged. *zxy* stage position was maintained within .nd2 metadata and was essential for returning to the same FOV. Orientation of the sample into the stage and sample holder was carefully maintained so as to enable returning to the same FOV. This was important as the sample was removed after imaging and between sequencing rounds.

OligoSTORM imaging. To combine OligoFISSEQ with OligoSTORM, we first performed one round of OligoSTORM imaging on all targets (Chr2-6plex or 36plex-5K) inside PGP1f male fibroblast cells by hybridizing Alexa Fluor 647-labeled secondary oligonucleotides that bind to the bridges (present in the Backstreet of individual Oligopaint oligonucleotides, with each chromosome containing a specific barcode), which contain a binding site for secondary oligonucleotides. OligoSTORM samples were imaged on a Vutara 352 biplane system with a ×60 1.3-NA silicone objective (UPLSAPO60XS2, Olympus). For single-molecule blinking, we used a switching buffer containing 2-mercaptoethanol and GLOXY[14]. The excitation laser power was set at 60% on the software (6.3 kW cm^{-2} at the objective) for the 640-nm laser and 0.5% on the software (0.08 kW cm^{-2} at the objective) for the photoactivation laser of 405 nm. We used 30–40 *z*-slices of 0.1-µm thickness for each *z*-slice. Approximately 10–12 photoswitching cycles of 250 frames per cycle were used for each *z*-slice.

The OligoSTORM images were analyzed using Vutara SRX software[14]. The DBSCAN clustering algorithm was used to identify the clusters from the raw image. Fifty particles within a 0.1-µm distance were used for clustering. The mean axial precision was 50 ± 10 nm in *z*, and the mean radial precision was 17 ± 5 nm in *xy*. The resolution of the super-resolved structures was calculated by Fourier ring correlation analysis (a built-up feature in SRX software). Resolution was 40 ± 5 nm in *xy* and 60 ± 5 nm in *z*.

Data visualization. Images were processed using either Nikon Elements or ImageJ/FIJI[61]. Image files (.nd2) were imported using the Bio-formats plugin[62]. Figure 2d was generated using ImageJ (under Plugins > 3D viewer)[63]. Chromosome schematics were generated using ChromoMap[64]. Figures were assembled in Adobe Illustrator. Micrograph images for publication figures were post-processed using brightness and contrast enhancement (ImageJ > Image > Adjust > Brightness/contrast). GraphPad Prism was also used for graphs. Molecular graphics and analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization and Informatics at the University of California and supported by the National Institutes of Health (NIH; grant no. P41-GM103311; ref. [65]).

Tier 1 detection. *Preprocessing.* Each round of OligoFISSEQ was imaged using five channels: Alexa Fluor 647, Texas Red, Cy3, Alexa Fluor 488 and DAPI and a series of *z*-slices. The *z*-stacks were deconvolved and background corrected using 20 iterations of the Richardson–Lucy algorithm using a theoretically calculated point spread function with Nikon software[66].

Rounds were compiled into hyperstacks composed of the five channels, a series of *z*-slices and one frame per round. If an image had all the puncta labeled, as with in toto images, it was included as a new additional frame. The hyperstacks were aligned using the Fiji plugin 'Correct 3D Drift' (ref. [67]). Images of DAPI-stained nuclei were used to perform threshold segmentation and extract each individual cell from the initial image as a separate region of interest. The segmentation provided information about the location and the envelope of the individual nuclei that made up each hyperstack. Nuclei with areas below 25 µm² were discarded.

Detection of barcodes. To compare intensities from different channels, images were normalized by dividing their intensities by the maximum intensity among the values of all the *z*-slices in the same round and channel.

For the detection of barcodes and for each round, the intensities of every pixel position were compared across different channels. A centroid-based pipeline using TrackMate[68] did not perform as well in our study; thus, we moved forward with this every-pixel approach. The channel with the highest value was kept as the prevalent channel. At every pixel position, the transition between channels along the different rounds was compared with the list of expected barcodes. A barcode was assigned to a pixel position if the set of transitions coincided with the one associated with the barcode. A maximum-intensity projection image was built by averaging the intensities of the prevalent channels from every round. Connected pixels that had the same barcode were grouped to form 3D patches. The following information was collected and saved for each patch:

- Barcode
- Center position
- Number of pixels forming part of the patch (size)
- Maximum intensity of the pixels of the patch
- Pixel position having the maximum intensity of the pixels of the patch

For an image with all puncta labeled, information on the intensity of each pixel position was stored in an additional file.

Tier 2 detection. *Chromosome tracing.* Patches composed of a single pixel location were discarded and the remainder were used in the tracing, disregarding the intensity or size of the patch.

Patches with high intensity values were selected as the most confident and were used to find the chromosome centers. We used an implementation of the constrained *k*-means algorithm[69] to find the center of the set of barcodes belonging to the same chromosome. To separate the homologs, we used a cannot-link constraint in the two copies of the same region to avoid having them in the same cluster. We used a sphere of radius 4.5 µm with origin in the centers to delimit the chromosome territory and filter out patches located outside.

The Domino sampler of the Integrative Modeling Platform[70] was the core element of the chromosome tracing. In Domino, each locus is represented by a particle with a finite set of different possible locations in the image. The locations are extracted from the list of patches having the same barcode as the one assigned to the locus. The remaining factors of the proposed problem are encoded in the system as restraints to the list of possible solutions. The following restraints are imposed by the system to filter compatible solutions:

- Two particles cannot share the same location or patch.
- Two consecutive particles of the same chromosome should be closer than a distance of 4 µm for the 36plex dataset and 1 µm for ChrX-46plex.
- Chromosomes must be confined in territories modeled as spheres of radius 4.5 µm.

Chromosome territory and the distance between consecutive regions were inferred as explained in 'Inferring chromosome territory and maximum distance between consecutive regions'. By applying these additional constraints to the barcodes, we were able to use patches that had intensities below, but not far from, the detection thresholds (Supplementary Table 14) and were likely to be true positives. Patches with higher intensities and sizes are most likely to be true-positive regions. Therefore, a score based on intensity and size was assigned to each patch as a measure of the likelihood of the patch being a true-positive detection. The list of patches was sorted by score and used as input data as an iterative process to find the most probable path of each chromosome (Supplementary Fig. 3).

The iterative process of tracing the chromosomes started by assigning patches with high scores to the corresponding regions. The process was executed once per

chromosome, considering all homologs at the same time because barcodes were not designed to distinguish them. Domino was used to list all possible solutions that were compatible with the imposed restraints. Each solution had a total score, obtained by summing the scores of the individual patches that were selected in that particular solution. We selected the conformation that had the highest total score. In the case where two or more solutions yielded an identical total score, we selected the solution that conformed to the shortest chromosome spatial length. An iterative process was performed for assignment of regions, whereby the threshold was lowered to allow more patches as input, and the previous approach was used to select the remaining unassigned regions. This iterative process was finished when all regions had been identified or there were no more input data to feed Domino.

Detection efficiency and false-positive ratios. To calculate the detection efficiency per barcode, the datasets were filtered using intensity thresholds (Supplementary Table 14) that were optimized for every experimental condition. Patches formed by a single pixel were also discarded regardless of the intensity of the patch.

For 36plex datasets, we calculated the mean of the barcodes detected per nuclei, excluding barcodes assigned to the X chromosome. In the ideal case, and due to the ploidy, we expected two barcodes per nucleus. In reality, the datasets may eventually include false positives or duplicates of patches that probably belong to the same oligonucleotide, which will increase the ratio. Nuclei with a mean of more than 2.5 barcodes were discarded because they were most likely in a mitotic process. For ChrX-46plex, we followed a similar procedure and discarded nuclei with a mean of detected barcodes that was greater than 1.5.

For each of the remaining nuclei, we computed the ratio of detected barcodes versus expected barcodes. We expected two barcodes per cell, except for the barcodes belonging to the X chromosome. The ratios per barcode and per cell were capped to 1 and averaged over all cells to produce the detection efficiency. For the false-positive ratio of the barcode, we instead calculated the excess of detections as the detected value minus the expected value in cases where the detected value was over the expected value, and we then computed the ratio of excess detections versus the expected values.

Distance heat maps and Hi-C maps. For every traced nucleus, we calculated all pairwise distances between the detected regions and averaged the results among all cells. For the average heat map of 36plex-5K, LIT dataset regions 3qR3 and 5pR3 were not taken into consideration because they shared the same barcode and were therefore indistinguishable. Hi-C maps of PGP1f cells were obtained from previous in situ Hi-C experiments[14]. The values of the interaction frequencies in the included Hi-C maps were extracted from the observed values of interaction matrices produced at a resolution of 5 Kb. The submatrices formed by the genomic regions of each pair of probes were aggregated to obtain the interregional observed interaction. Single-cell heat maps were built with the identification of homologous chromosomes. The list of barcodes was traced according to the procedure described in above in the 'Chromosome tracing' section of 'Tier 2 detection'. All pairwise distances of the traced regions were calculated. Non-identified regions appear as gray columns and rows.

Inferring chromosome territory and the maximum distance between consecutive regions. To infer the maximum distance between consecutive regions used in the chromosome tracing, the list of detected barcodes for all 36plex datasets was filtered to discard mitotic cells as explained in 'Detection efficiency and false-positive ratios'. Patches formed by a single pixel were also filtered out. After the filtering process, the 36plex dataset comprised 1,171 nuclei and 48,352 barcodes. Then, we calculated the distances between consecutive regions for each chromosome in each nucleus (Supplementary Fig. 1). The histograms of those distances show the expected bimodal distributions for the chromosomes, except for chromosome X as foreseen from male cells. Bimodality is more evident in bigger chromosomes because those tend to be in the periphery of the nucleus, while smaller chromosomes prefer the interior. After inspection of the histograms, we selected 4 µm as the general maximum distance between consecutive regions and a slightly higher value of 4.5 µm for the chromosome territory.

For the ChrX-46plex dataset, we followed a similar approach. After the filtering process, ChrX-46plex contained 189 nuclei and 7,752 barcodes. Based on the histograms of distances between consecutive regions, we selected 2.5 µm as the general maximum distance (Supplementary Fig. 2).

Clustering of 3D structures for ChrX-46plex. After tier 2 detection, we had 177 cells for the ChrX-46plex library, with an average of 34 detected regions. We discarded one of the cells that had fewer than 23 identified barcodes so as to meet the required 50% detection efficiency per cell in all the 3D structures. Next, we calculated the pairwise distances for each chromosome between all of their detected targets and used these as a measure of similarity to the built distance matrix. We used the coincident distances between structures to cluster them hierarchically using the Ward method. The Calinski–Harabasz criterion for clustering evaluation was used to evaluate the optimal number of clusters.

ChrX-46plex-2K tracing in IMR-90 cells. O-eLIT with the ChrX-46plex-2K library in IMR-90 cells was performed as in PGP1f cells. Five rounds of sequencing were performed off both streets, followed by immunostaining for MacroH2A.1 (Abcam, ab183041) at 1:250 dilution to mark the inactive X chromosome. For the every-pixel analysis, chromosome traces with fewer than 13 identified regions were filtered out.

MacroH2A.1 IF images were aligned and segmented with the DAPI channel of their OligoFISSEQ correspondence. For each nucleus, the position of maximum intensity of the IF image was compared with the geometric center of the traced X chromosomes. To filter out images without a clear IF signal, we only considered nuclei where their maximum IF intensity was greater than two times the average intensity inside. If the center was closer than 3.5 µm, the X chromosome was considered IF positive and annotated as inactive (Xi). The other X chromosome in the nucleus was annotated as active (Xa). In cases where both homologs were closer than 3.5 µm to the IF signal, the closest homolog was annotated as Xi and the farthest was annotated as Xa. The nuclei were manually checked to discard errors, mainly due to overlapping cells that resulted in 40 Xi chromosomes, for which we traced 31 homologs that were identified as Xa.

Generation of random nuclei for haploid separation. For the directed random nuclei, we first calculated the mean and s.d. of the distance to the nuclear envelope for every chromosome (Supplementary Table 15). We used this information to generate a set of random nuclei where the chromosomes were randomly placed following a normal distribution in which the mean and s.d. were equal to the values calculated in the observed data. The positions of the large chromosomes in the synthetic nuclei were biased toward the periphery, while the positions of the small chromosomes were biased toward the nuclear interior. No spatial bias was used for the completely random nuclei.

Histogram of split homologs by *k*-means. For the analysis, we selected 258 nuclei for which all centers of the 11 chromosomes were known. We used the conventional *k*-means algorithm to cluster the positions of the chromosomes into two groups and reported how many autosomes were split by the clustering, that is, how many autosomes had one copy in one of the groups and the homolog in the other group.

Method for the alignment of nuclei. For the analysis, we selected 258 nuclei for which all centers of the 11 chromosomes were known. We used an implementation of the constrained *k*-means algorithm[69] to cluster the chromosomes into two groups: one group contained one copy of each autosome and the other group contained the homolog. The X chromosome was assigned to the closest group. The geometric centers of the clusters were joined and the resulting segment, together with all the positions of the chromosomes, was rotated to be parallel to the x axis and moved to leave the middle point toward the origin; $x=0$, $y=0$. In the rotation of the nuclei, we kept the group containing the X chromosome at the left of the y axis.

Density plots. The density plots were built using the kernel density estimation of the projection to the xy plane of the position of the chromosomes.

Number of split homologs. We checked each aligned nucleus and reported how many autosomes could be split by a virtual line along the y axis, that is, the number of autosomes with one of the copies on the left of the y axis and the other on the right of the y axis.

Number of split homologs left to right. We checked each aligned nucleus and reported how many autosomes could be split by a virtual line parallel to the y axis at different distances from the origin, that is, the number of autosomes with one of their homologs on the left of the line and the other on the right of the line.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data are available in the main text or the supplementary materials, and materials are available upon request. Information regarding all datasets (for example, cells, replicates and filters) can be found in Supplementary Table 9. Source data are provided with this paper.

Code availability

All code is available at https://github.com/3DGenomes/OligoFISSEQ/.

References

51. 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. USA* **115**, E2183–E2192 (2018).
52. Ball, M. P. et al. A public resource facilitating clinical use of genomes. *Proc. Natl Acad. Sci. USA* **109**, 11920–11927 (2012).
53. Zhang, K. et al. Digital RNA allelotyping reveals tissue-specific and allele-specific gene expression in human. *Nat. Methods* **6**, 613–618 (2009).

54. Pardue, M. L. et al. Molecular hybridization of radioactive DNA to the DNA of cytological preparations. *Proc. Natl Acad. Sci. USA* **64**, 600–604 (1969).
55. Bauman, J. G., Wiegant, J., Borst, P. & van Duijn, P. A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labelled RNA. *Exp. Cell Res.* **128**, 485–490 (1980).
56. Shendure, J. et al. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732 (2005).
57. Valouev, A. et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* **18**, 1051–1063 (2008).
58. Guo, J. et al. Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc. Natl Acad. Sci. USA* **105**, 9145–9150 (2008).
59. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in situ RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014).
60. Moffitt, J. R. et al. High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl Acad. Sci. USA* **113**, 14456–14461 (2016).
61. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
62. Linkert, M. et al. Metadata matters: access to image data in the real world. *J. Cell Biol.* **189**, 777–782 (2010).
63. Schmid, B., Schindelin, J., Cardona, A., Longair, M. & Heisenberg, M. A high-level 3D visualization API for Java and ImageJ. *BMC Bioinformatics* **11**, 274 (2010).
64. Anand, L. ChromoMap: an R package for interactive visualization and annotation of chromosomes. Preprint at *bioRxiv* https://doi.org/10.1101/605600 (2019).
65. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
66. Richardson, W. H. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.* **62**, 55–59 (1972).
67. Parslow, A., Cardona, A. & Bryson-Richardson, R. J. Sample drift correction following 4D confocal time-lapse imaging. *J. Vis. Exp.* 51086 (2014).
68. Tinevez, J.-Y. et al. TrackMate: an open and extensible platform for single-particle tracking. *Methods* **115**, 80–90 (2017).
69. Wagstaff, K., Cardie, C., Rogers, S. & Schroedl, S. Constrained *k*-means clustering with background knowledge. In *Proc. 18th International Conference on Machine Learning*, 577–584 (2001).
70. Russel, D. et al. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol* **10**, e1001244 (2012).

Acknowledgements

We acknowledge members of the Marti-Renom and Wu laboratories for technical and conceptual support, especially T. Ryu, A. Lioutas and S. Aufmkolk as well as J. AlHaj Abed, S. D. Lee, J. Erceg and T. Hatkevich; B. Beliveau, H. Sasaki, J. Horrell, L. Cai, J. Kishi and P. Soler-Vila for discussion; D. Barclay, R. Kohman, E. Iyer, K. Rodgers, A. Skrynnyk, J. Tam and R. Terry for discussion about FISSEQ and sequencing reagents; S. Alon, F. Chen, Z. Chiang, D. Goodwin, A. Payne, A. Sinha and O. Wassie for discussion about FISSEQ; C. Ebeling, J. Rosenberg and J. Stuckey for discussion and technical assistance; F. Pan and A. Hutchinson for assistance in procuring SOLiD reagents; P. Montero-Llopis and the MicRoN imaging core at Harvard Medical School; the ImageJ discussion forum; and StackOverflow. This work was supported by a Damon Runyon Dale F. Frey Breakthrough Award (to B.J.B.) to support B.J.B. and E.A.H., awards from the NSERC of Canada (PGS D) to P.L.R., the NIH (HG005550 and HG008525) and NSF (DGE1144152) to E.R.D., the European Research Council under the Seventh Framework Program (FP7/2007–2013 609989), the European Union's Horizon 2020 Research and Innovation Program (676556) and the Spanish Ministerio de Ciencia, Innovación y Universidades (BFU2017-85926-P) to M.A.M.-R., the Centro de Excelencia Severo Ochoa 2013–2017 (SEV-2012-0208) and the CERCA Programme/Generalitat de Catalunya to the CRG, from the NIH to GMC (RM1HG008525-03) and the NIH (DP1GM106412, R01HD091797 and R01GM123289) to C.-t.W.

Author contributions

H.Q.N., S.C., D.C., S.C.N., G.M.C., E.R.D., M.A.M.R. and C.-t.W. conceived the study with the original conceptualization of OligoFISSEQ contributed by S.C.N and E.R.D.; G.N., A.L. and N.M.C.M. provided guidance for barcode design and angle analysis; A.L., E.A.H. and B.J.B. provided guidance for Oligopaint sequences and barcode design. P.L.R. supported early protocol development; M.H. provided technical support; H.Q.N. and S.C. designed and performed the experiments. H.Q.N., S.C., D.C., G.M.C., M.A.M.R. and C.-t.W. analyzed the data; H.Q.N. wrote the manuscript with S.C., D.C., M.A.M.R. and C.-t.W. with input from all authors; C.-t.W. oversaw the project.

Competing interests

Harvard University has filed patent applications on behalf of C.-t.W., H.Q.N. and S.C., pertaining to Oligopaints and related oligonucleotide-based methods for genome imaging. E.R.D. is currently an employee of ReadCoor and has an equity interest in ReadCoor. Potential conflicts of interest for G.M.C. are listed on http://arep.med.harvard.edu/gmc/tech.html/. C.-t.W. has an equity interest in ReadCoor and an active research collaboration with Bruker Nano in her laboratory at Harvard Medical School.

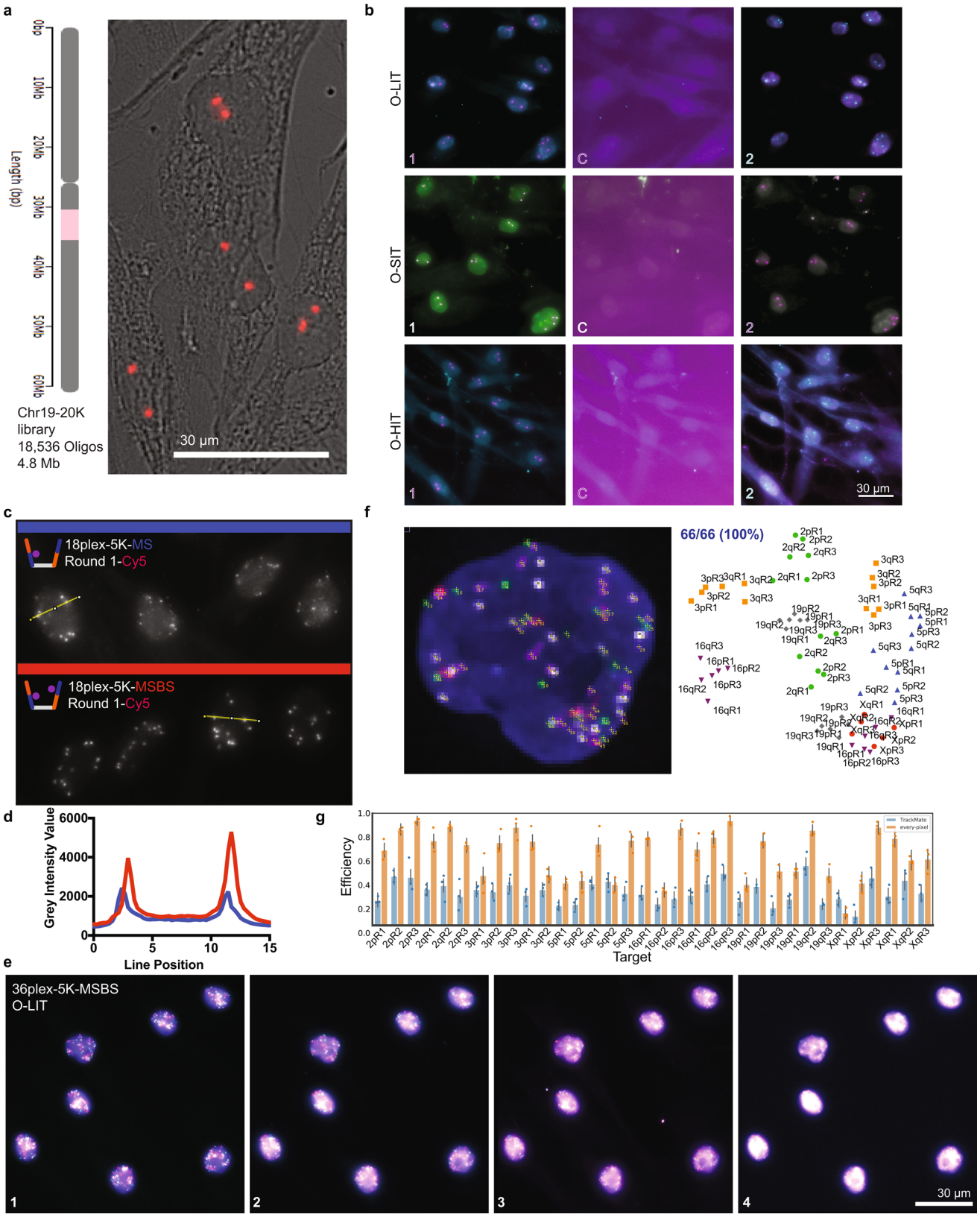
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41592-020-0890-0.

Supplementary information is available for this paper at https://doi.org/10.1038/s41592-020-0890-0.

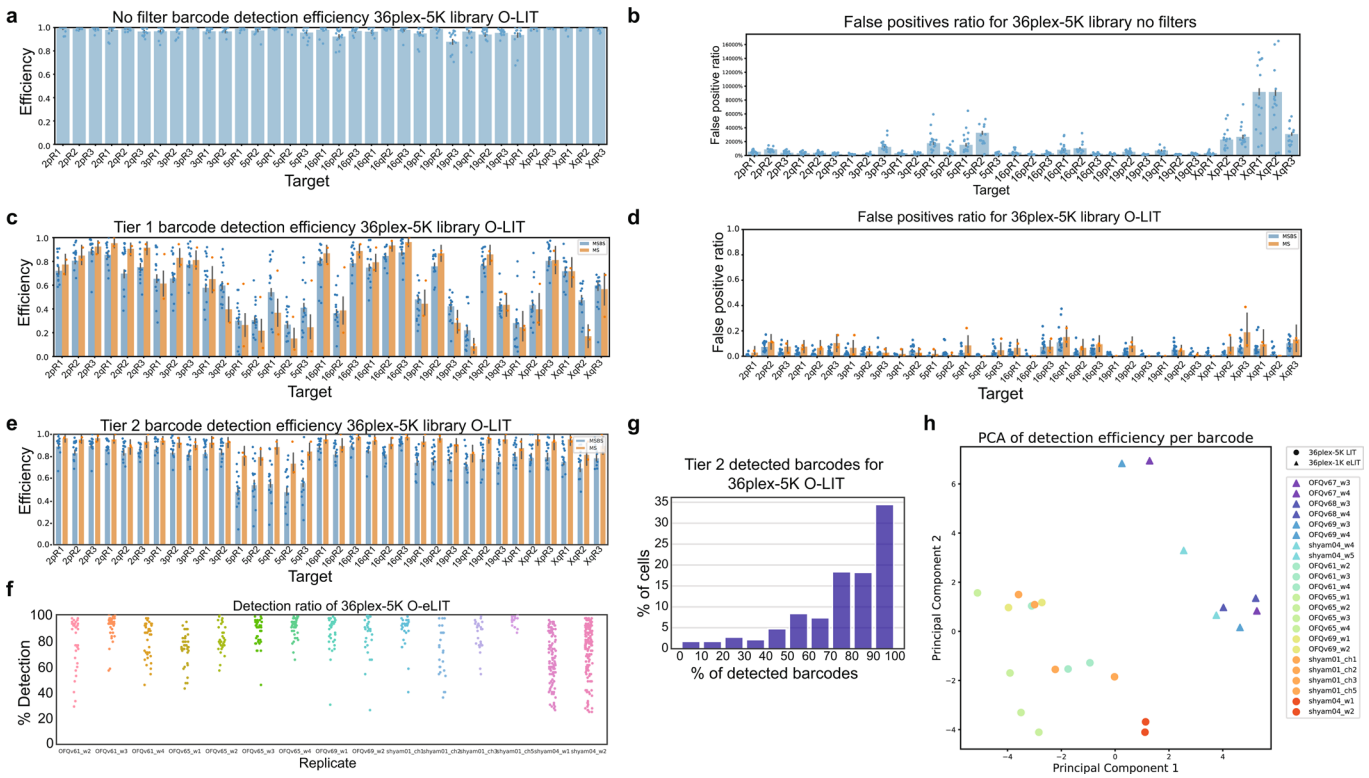
Correspondence and requests for materials should be addressed to M.A.M.-R. or C.-t.W.

Peer review information Lei Tang was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

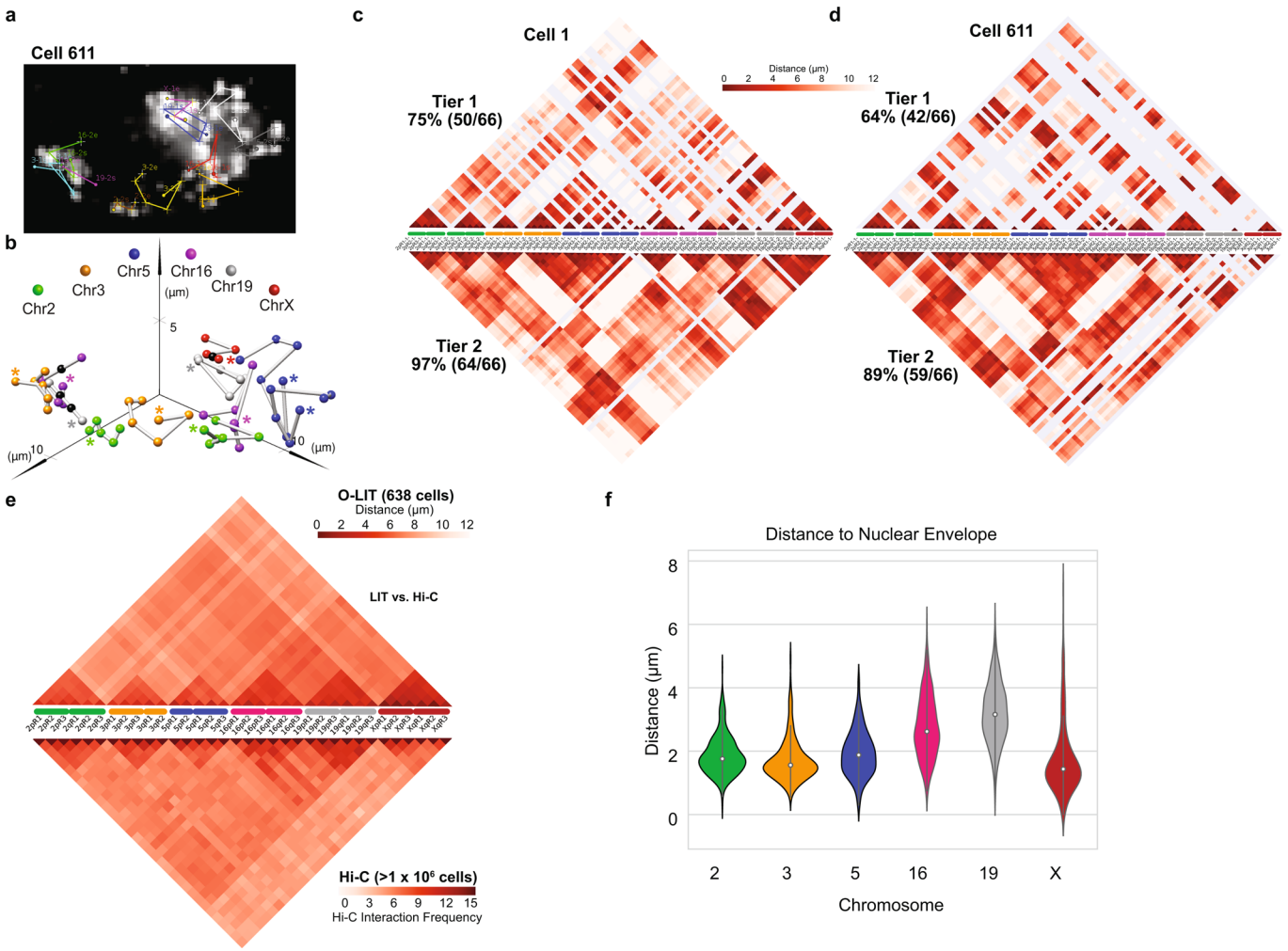


Extended Data Fig. 1 | See next page for caption.

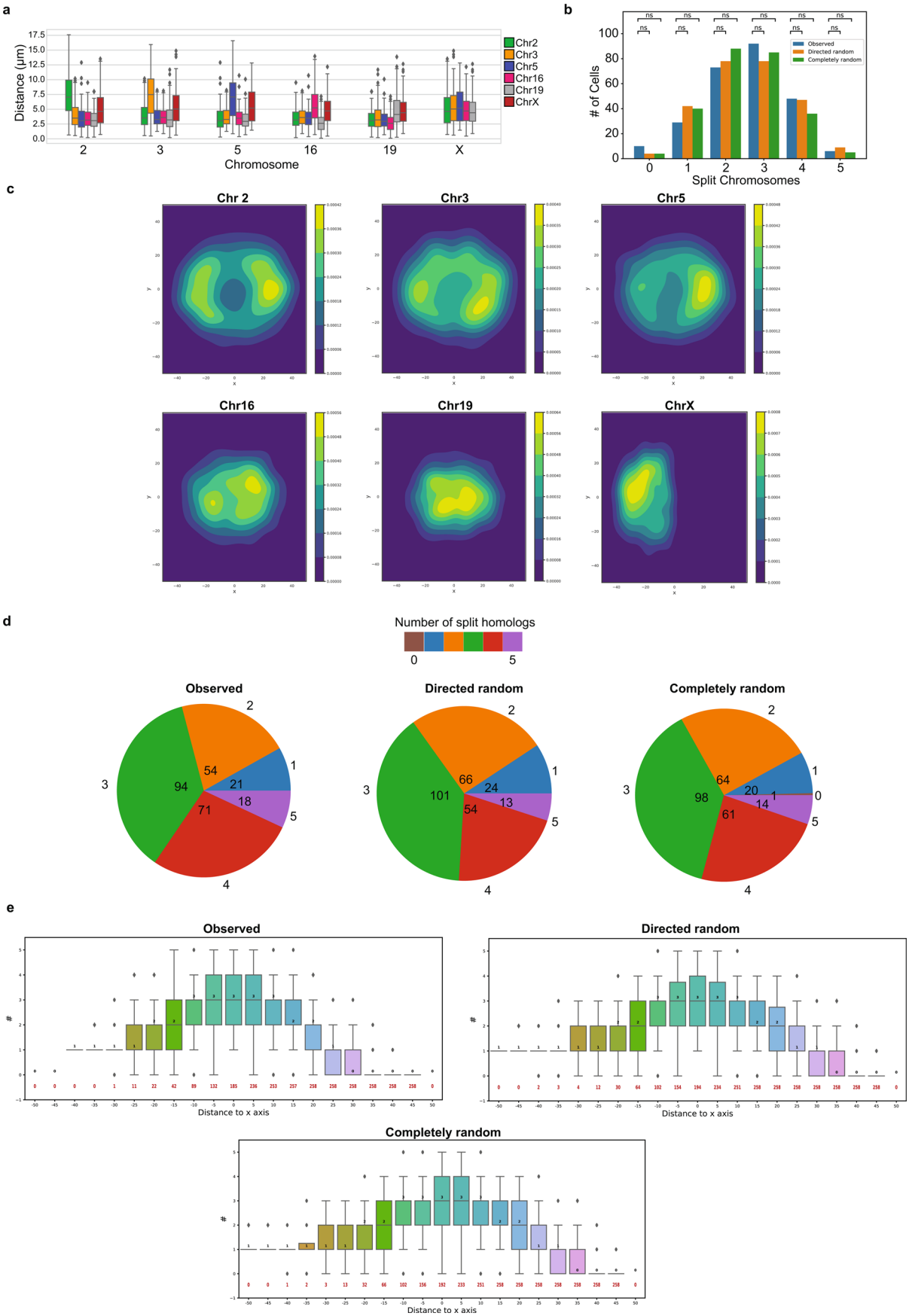
Extended Data Fig. 1 | Chr19-20K and 36plex-5K-O-LIT optimization. a, Chr19-20K targets 18,536 Oligopaint oligos to human chromosome 19. Right, Chr19-20K detection with secondary oligo (red) in PGP1f cells representative of 5 replicates. **b**, Signal is completely removed in each OligoFISSEQ method after cleavage. Images showing two rounds of sequencing with a cleavage step (C) and representative of 4 replicates. **c**, 36plex-5K O-LIT off of both Mainstreet and Backstreet (MSBS; bottom, red) produces stronger signal than off of Mainstreet (MS; top, blue). Cy5 channel from first round of O-LIT. n = 1. **d**, O-LIT off of both streets produces stronger signal than off of MS. Grey intensity value measurements from yellow lines in panel c. n = 1. **e**, Raw, non-deconvolved field of view of cell from Figs. 2c, d and 3a–c. Maximum z-projection. n = 1. **f**, Manual decoding of cell from panel c and Figs. 2c, d and 3a–c yields 100% target recovery. n = 1. **g**, Tier1 detection efficiency after 36plex-5K O-LIT off of both streets and detected with TrackMate (blue, 29.93 ± 4.9%) or Every-pixel (orange, 62.8% ± 4.8%). n = 111 cells from 3 replicates. Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean.



Extended Data Fig. 2 | Detection efficiency after 36plex-5K O-LIT. a, Detection efficiency without filtering after 36plex-5K O-LIT off of both streets. 95 ± 5.15% of targeted regions are detected (n = 611 from 15 replicates). Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **b**, False positive (FP) discovery rate from panel a. FP discovery rate from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **c**, Tier 1 detection efficiency after 36plex-5K O-LIT off of Mainstreet (orange, 61.93 ± 12%, n = 53 from 2 replicates) versus off of both streets (blue, 62.17% ± 6.68%, n = 611 cells from 15 replicates). Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **d**, FP discovery rate from panel c. Using Mainstreet = 8.64% and using both streets = 5.29%. FP discovery rate from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **e**, Tier 2 detection efficiency after 36plex-5K off of Mainstreet (orange, 92.3% ± 3.42% from 53 cells from 2 replicates) versus off of both streets (blue, 80.19 ± 7.29%, n = 611 cells from 15 replicates). Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **f**, Detection efficiency after 36plex-5K O-LIT off of both streets for individual cells from 15 replicates in panel **e**. **g**, Percentage of cells displaying a range of efficiencies of barcode detection after 36plex-5K O-LIT off of both streets. Data taken from panel **e**. **h**, Principal component analysis showing lack of batch effect in 36plex datasets (n = 1171 cells from 15 36plex-5K O-LIT replicates and 8 36plex-1K O-eLIT replicates).

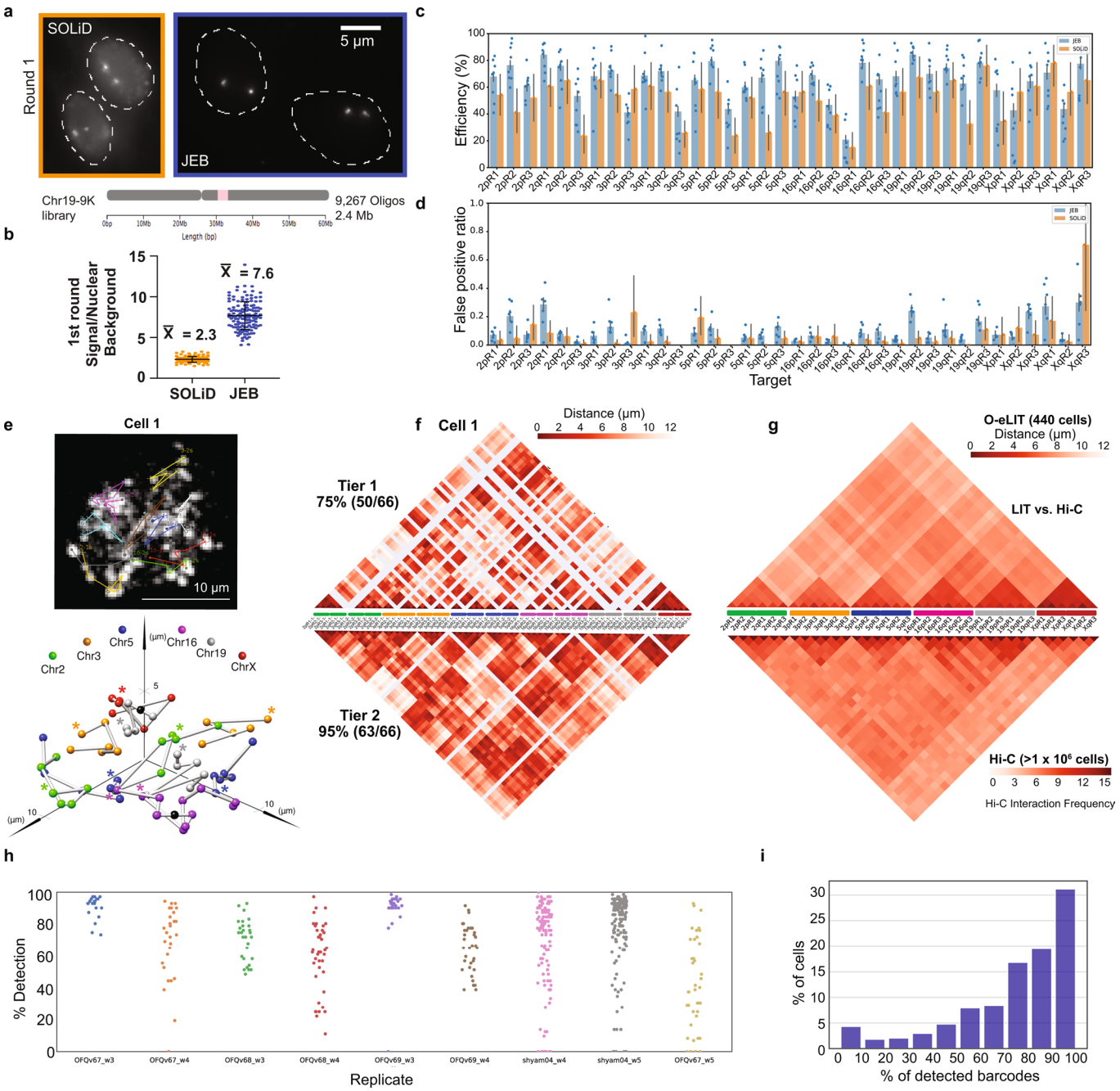


Extended Data Fig. 3 | O-LIT with 36plex-5K to interrogate genome organization. a, Chromosome traces of Cell 611 after Tier 2 detection of cell 611 after four rounds of O-LIT 36plex-5K off of both streets. 59/66 (89%) of 36plex-5K targets were detected. Image is from the first round of O-LIT with target identities. n = 1. **b**, Ball and stick of Cell 611. Colored spheres represent chromosomal targets, while black spheres represent targets that were not detected and, thus, were placed by calculating the median proportionate distance between flanking detected targets. Beginning of chromosome (for example 2pR1) marked by an asterisk. **c**, Single-cell pairwise spatial distance matrix after Tier 1 (top) and Tier 2 (bottom) detection of the nucleus in Fig. 3. Targets are represented on the x-axis with homologs separately displayed. Undetected targets are represented by grey lines. **d**, Single-cell pairwise spatial distance matrix after Tier 1 (top) and Tier 2 (bottom) detection of Cell 611. Targets are represented on the x-axis with homologs separately displayed. Undetected targets are represented by grey lines. **e**, 36plex-5K population pairwise spatial distances (top, from Fig. 3f). Average pairwise spatial distances from cell population after Tier 1 detection (n = 611 from 15 replicates). (Spearman's rank correlation 0.705, two-sided p-value for a hypothesis test whose null hypothesis is that two sets of data are uncorrelated = 1.77e-174). Measurements from homologous targets were combined. Bottom, Hi-C data of 36plex-5K targets obtained from (Nir et al. 2018). **f**, Average distances between the nuclear membrane and the closest of the six targets imaged for each chromosome. (n = 686, 668, 364, 586, 760, and 494 for Chr2, 3, 5, 16, 19, and X, respectively.) The thick line in each violin plot represents the Interquartile range (IQR), the white dot marks the median and the thin lines extend 1.5 times the IQR.

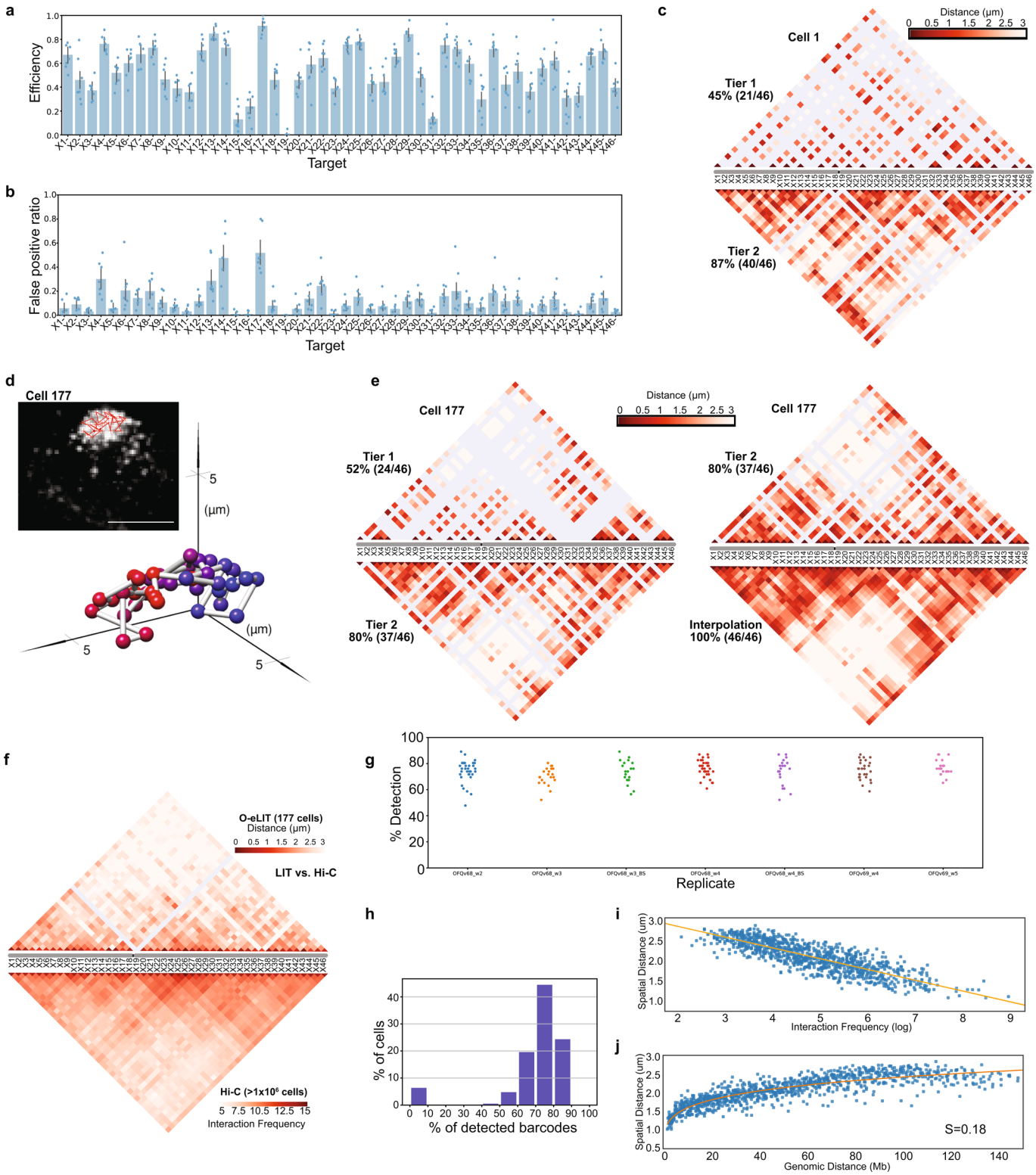


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | O-LIT with 36plex-5K to interrogate homolog organization. a, Minimum distances between heterologous and homologous chromosomes. All measurements represent distances between the geometric centers of chromosomes for which all six targets were imaged. Distances between a chromosome and a heterologous chromosome is the shorter of the two distances between that chromosome and the two homologous copies of the heterologous chromosome (n = 686, 668, 364, 586, 760, and 494 for Chr2, 3, 5, 16, 19, and X, respectively). Inter-homolog distances for Chr16 and 19 are less than those for Chr2, 3, and 5 (independent-samples t-test p = 4.28 ×10-37). Boxes represent the IQR (25th, 50th and 75th percentiles) and whiskers extend 1.5 times the IQR. **b**, Number of cells with varying numbers of homologs split by K-means clustering. The K-means algorithm was applied to 258 nuclei, individually, to cluster chromosomes into two groups based on proximity and then report the number of homolog pairs that were split by the clustering. A value of "5" indicates that the homologs from each five pairs of imaged autosomes in a single nucleus clustered into two spatially separate groups. Observed, PGP1f cells. Directed random, raw positions in Observed but with the chromosome identities of all positions randomized, with the larger chromosomes (2, 3, 5) biased towards the nuclear periphery and smaller chromosomes (16 and 19) biased towards the nuclear interior. Completely random category, randomization of the chromosome identities carried out with no spatial bias. The significance of each pair was evaluated from a two proportion z-test with n = 258 for each category with a null hypothesis of equal proportion and a significance level of 0.05. **c**, Density plots of homolog positions. Built by using Kernel density estimation (KDE) of nuclei projected and aligned along the x-y plane of the position of the chromosomes. **d**, Pie charts of total number of cells with homologs split by a virtual line along the y-axis. **e**, Number of aligned cells with homologs split by a virtual line parallel to the y-axis at different distances from the origin, that is, number of autosomes with one of their homologs on the left of the line and the other on the right (n = 258 for each category). Boxes represent the IQR (25th, 50th and 75th percentiles) and whiskers extend 1.5 times the IQR.

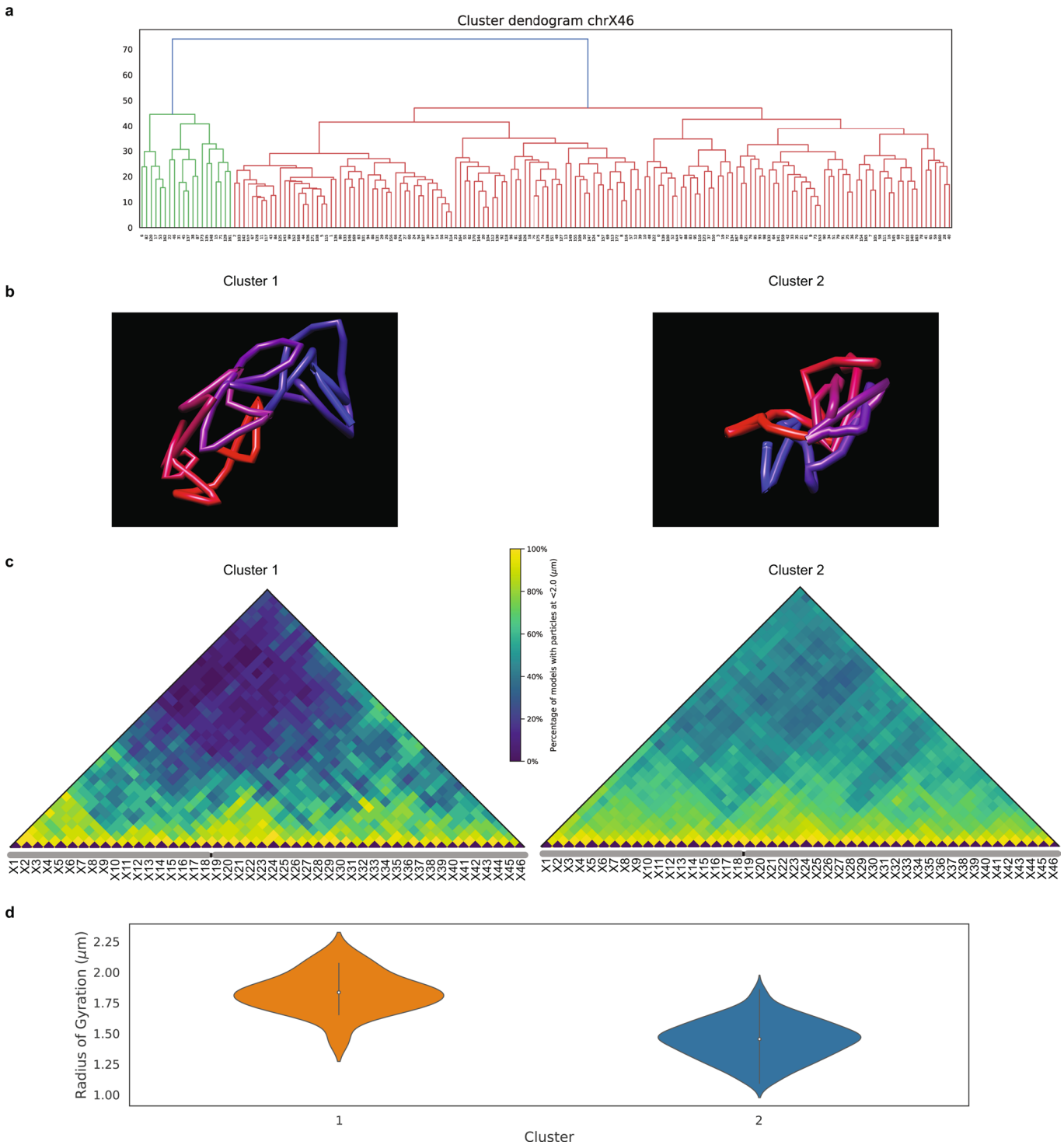


Extended Data Fig. 5 | O-eLIT with JEB. a, Chr19-9K. One round of O-LIT (SOLiD) or O-eLIT (JEB) off of Mainstreet. Maximum z-projections representative of 2 replicates. **b**, Chr19-9K signal over nuclear background measurements after one round of O-LIT (orange; n = 113 puncta from 55 cells from 2 replicates) or O-eLIT (blue; n = 136 puncta from 57 cells from 2 replicates). Bar is the mean and SD. **c**, Tier 1 detection of 36plex-1K after five rounds of O-LIT with SOLiD reagents (orange; average of 51.75%, n = 41) or O-eLIT with JEB (blue; average of 61.2 ± 10.2%, n = 440 from 9 replicates). Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. 36plex-1K library shares first 1,000 Oligopaint oligos of each target in 36plex-5K. For example, for target 2pR1, 36plex-5K spans the chromosomal region from nt position 1,002,895 to 1,660,898 (~658 kb), whereas 36plex-1K spans the region from nt 1,002,895 to 1,147,495 (~144 kb). **d**, FP discovery rate from panel c. SOLiD = 7.49% and JEB = 8.95%. FP discovery rate from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **e**, Chromosome traces and ball and stick of Fig. 4c cell after Tier 2 detection and five rounds of O-eLIT 36plex-1K. 63/66 (95%) targets were detected. Asterisks, beginning of chromosomes. n = 1. **f**, Single-cell pairwise spatial distance matrices of panel C cell. **g**, 36plex-1K population pairwise spatial distance measurements (top, from Fig. 3f). Average pairwise spatial distance from cell population after Tier 1 detection (n = 440 from 9 replicates). Measurements from homologous targets were combined. Bottom, Hi-C data of 36plex-5K targets obtained from (Nir et al. 2018). **h**, 36plex-1K detection rate for individual cells from 9 replicates. **i**, Percentage of cells displaying a range of efficiencies of barcode detection after 36plex-1K O-eLIT off of Mainstreet.

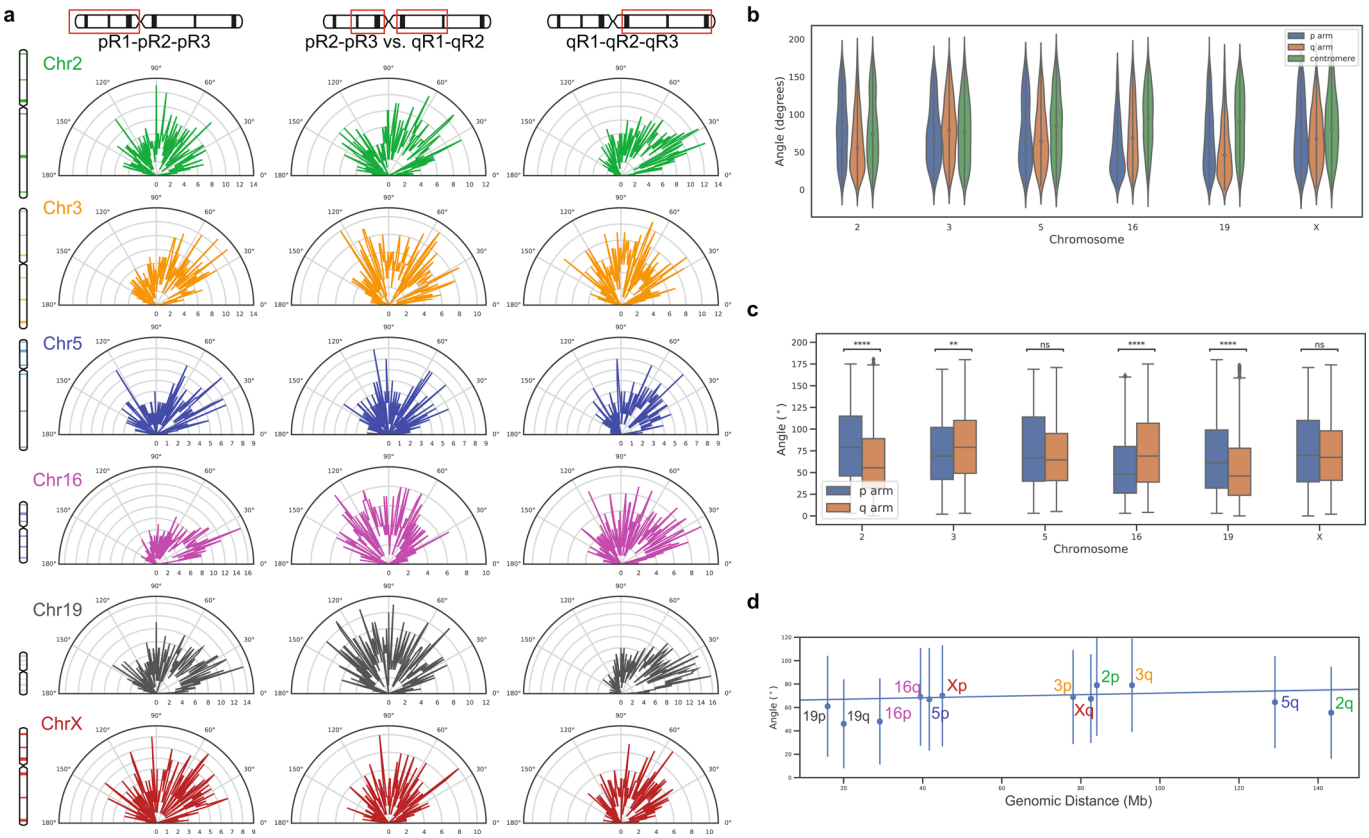


Extended Data Fig. 6 | See next page for caption.

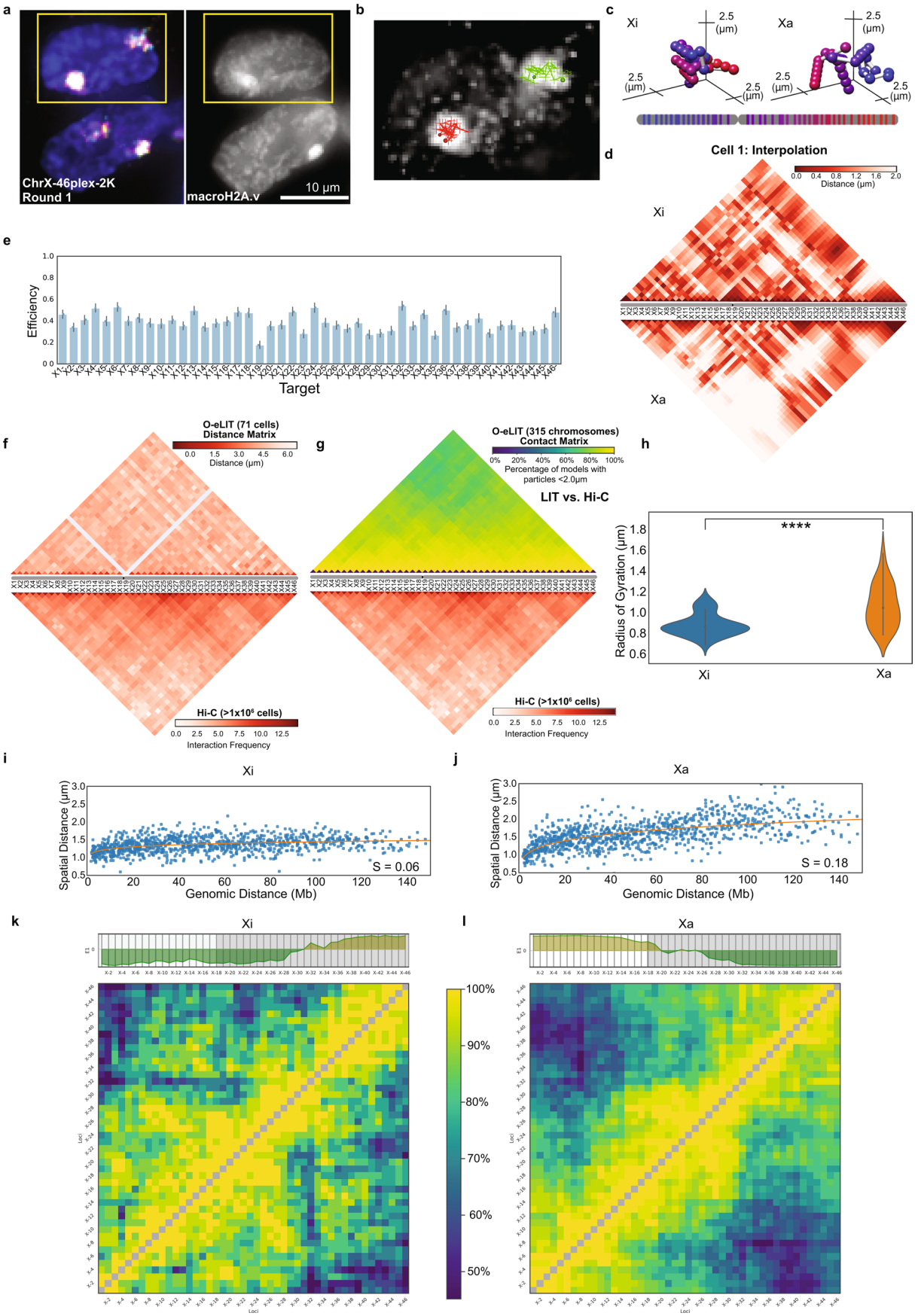
Extended Data Fig. 6 | O-eLIT with ChrX-46plex-2K. a, ChrX-46plex-2K O-eLIT Tier 1 detection off of one street and off of both streets combined (52.86 ± 5.78% from 177 cells from 7 replicates). Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **b**, FP discovery rate from panel a. Error bars represent 95% bootstrap confidence interval of the mean. **c**, Single-cell pairwise spatial distance matrix after Tier 1 (top) and Tier 2 (bottom) detection of Cell 1 from Fig. 5b. Undetected targets are represented by grey lines. **d**, Chromosome traces (top) and ball and stick representation (bottom) of Cell 177 after Tier 2 detection and interpolation and five rounds of O-eLIT on ChrX-46plex-2K off of both streets. Image is from the first round of O-eLIT with target identities. n = 1. **e**, Single-cell pairwise spatial distance matrix after Tier 1 (top), Tier 2 (bottom) of Cell 177 (left), and Tier 2 (top) and interpolation (bottom) of same cell (right). Undetected targets are represented by grey lines. **f**, ChrX-46plex-2K population pairwise spatial distances (top). Average pairwise spatial distances from cell population after Tier 1 detection (n = 177 from 7 replicates). Bottom, Hi-C (Nir et al. 2018) data of ChrX-46plex-2K targets. (Spearman's rank correlation 0.641, two-sided p-value for a hypothesis test whose null hypothesis is that two sets of data are uncorrelated = 7.074e-245). **g**, ChrX-46plex-2K detection rate for individual cells from 7 replicates. **h**, Percentage of cells displaying a range of efficiencies of barcode detection after ChrX-46plex-2K O-eLIT. **i**, Mean spatial distance versus Interaction frequency of Hi-C (Nir et al. 2018) of ChrX-46plex-2K targets. Pearson correlation coefficient (r = −0.84) and p-value = 5.08E-275 (two-sided, using slope = 0 for null hypothesis and Wald Test with t-distribution as test statistic) of the linear least-squares regression. **j**, Mean spatial distance versus genomic distance for all pairwise ChrX-46plex-2K targets (n = 177 from 7 replicates).



Extended Data Fig. 7 | O-eLIT identifies clusters after ChrX-46plex O-eLIT. a, Hierarchical clustering based on structure of ChrX traces from ChrX-46plex after 5 rounds of O-eLIT and Tier 2 detection yielded two clusters (Cluster 1 = 20; Cluster 2 = 156). See Methods for more details. **b**, ChrX representative models (existing traces that are closer to the virtual centroid) of the two clusters obtained after Hierarchical clustering in panel **a**. **c**, ChrX-46plex-2K population contact matrix of two clusters derived after Hierarchical clustering in panel a where pairwise spatial distances are considered to be in contact if less than 2 μm apart. **d**, Radius of gyration for the two clusters (Cluster 1 = 20; Cluster 2 = 156) derived after the hierarchical clustering shown in panel a. The thick line in each violin plot represents the Interquartile range (IQR), the white dot marks the median and the thin lines extend 1.5 times the IQR.

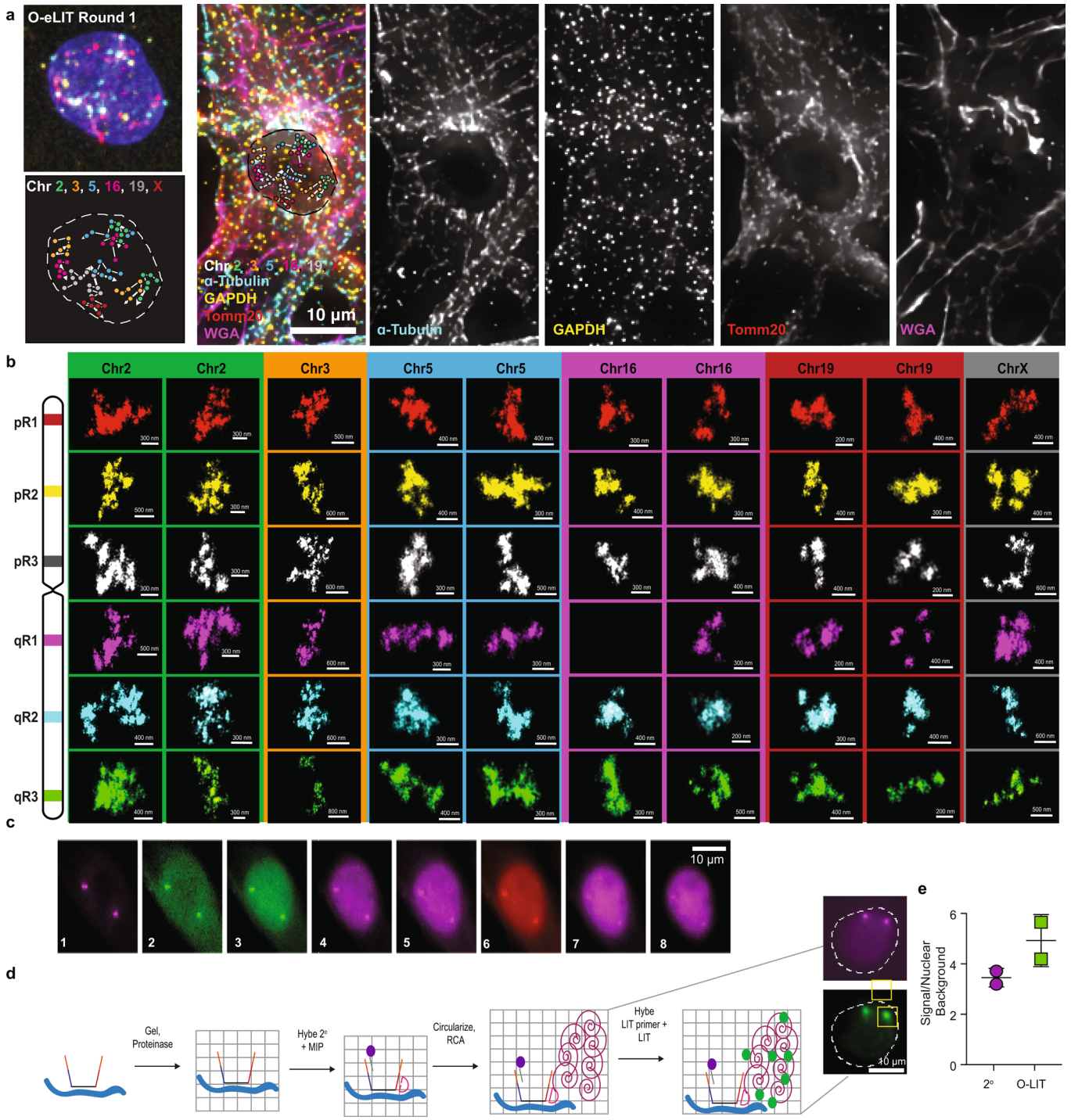


Extended Data Fig. 8 | Angles from 36plex. a, Measurements of angles formed by three points along the p arm (left), q arm (right), and intersection of vectors formed by pR2-pR3 and qR1-qR2 (middle) for each chromosome. Measurements were obtained by combining data from 36plex-5K and 36plex-1K analyses and selecting chromosomes that had all six targets identified. Chr2: n = 686, Chr3: n = 668, Chr5: n = 363, Chr16: n = 586, Chr19: n = 760, ChrX: n = 493 (n = 1,051 cells from 24 replicates; for 36plex-5K, n = 611 from 15 replicates; for 36plex-1K, n = 440 from 9 replicates). **b**, Distribution of angles formed by segments in panel a. The thick line in each violin plot represents the Interquartile range (IQR), the white dot marks the median and the thin lines extend 1.5 times the IQR. **c**, Box plots comparing p and q arm angles. Two-sided student's t-test with null hypothesis of equal mean was performed to compare arms, ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. Boxes represent the IQR (25th, 50th and 75th percentiles) and whiskers extend 1.5 times the IQR. Sample size information in a). Exact p-values for each chromosome: Chr.2 = 4.149e-16, Chr.3 = 0.004, Chr.5 = 0.093, Chr.16 = 1.357e-14, Chr.19 = 3.325e-11, Chr.X = 0.101. **d**, Linear least-squares regression between arm angle and arm length with Pearson correlation coefficient r = 0.26 and p-value = 0.42 (two-sided, using slope = 0 for null hypothesis and Wald Test with t-distribution as test statistic).



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | O-eLIT comparison of X chromosomes in female IMR-90 cells after ChrX-46plex-2K O-eLIT off of both streets. a, First round of O-eLIT sequencing. MacroH2A.1 immunostaining after five rounds of O-eLIT marks the Xi. n = 1. **b**, **c**, Xi and Xa traces (b) and ball and stick (c) of panel a nucleus after Tier 2 analysis and interpolation of missing targets. Sphere color corresponds to chromosome cartoon. n = 1. **d**, Single-cell pairwise spatial distances after interpolation of missing targets in panel **a**. **e**, Tier 2 target detection efficiency after five rounds of O-eLIT. 38.57% of targeted regions are detected in 71 cells. Detection efficiency from individual replicates are plotted. Error bars represent 95% bootstrap confidence interval of the mean. **f**, Population pairwise spatial distances after Tier 1 detection (n = 71 cells) and Hi-C data of IMR-90 cells (Rao et al. 2014). **g**, Population contact maps (top) where two targets are considered to be in contact if less than 2 μm apart (n = 315 chromosomes). Bottom, Hi-C data as in panel f. (Spearman's rank correlation with the Hi-C matrix is r = 0.733, two-sided p-value for a hypothesis test whose null hypothesis is that two sets of data are uncorrelated = 2.564 ×10e-175). **h**, Radius of gyration for the Xi (n = 40 chromosomes) and Xa (n = 31 chromosomes). The thick line in each violin plot represents the Interquartile range (IQR), the white dot marks the median and the thin lines extend 1.5 times the IQR. P-value = 7.08 ×10e-6 (two-sided t-test whose null hypothesis is equal means). **i**, **j**, Linear plot of the mean spatial distance versus the genomic distance for all pairwise targets for Xi (n = 40 chromosomes) and Xa (n = 31 chromosomes). **k–l**, Population contact maps for Xi (n = 40 chromosomes) and Xa (n = 31 chromosomes) with eigenvector analysis used to identify different domains. X1-X18 (white) and X19-X46 (grey) targets p and q arms, respectively.



Extended Data Fig. 10 | OligoFISSEQ applications. a, O-eLIT and immunofluorescence (IF). 36plex-1K was sequenced 5 rounds with O-eLIT off Mainstreet. Then, the same sample was prepared for IF and stained with antibodies. Samples were counterstained with wheat germ agglutinin (WGA) to stain membranes. Images are from deconvolved, maximum z-projections representative of 2 replicates. **b**, Chromosomal regions imaged with OligoSTORM from Fig. 6d enlarged and displayed separately. Orientation may differ from Fig. 6d. n = 1. **c**, 8 rounds of O-LIT sequencing of Chr19-9K off of Mainstreet. Images are maximum z-projections. Signal is detectable in all rounds even though the imaging was conducted without the advantage of eLIT, suggesting that 8 rounds of O-eLIT will produce even stronger signals. Images are representative of 2 replicates. **d**, O-LIT is compatible with gel embedding and target amplification via rolling circle amplification (RCA). Chr19-9K was hybridized to PGP1f cells, after which the sample was embedded in a hydrogel and then cleared of cellular background with proteinase. Next, a molecular inversion probe (MIP) was hybridized to a Chr19-9K specific barcode on Backstreet as well as a fluorophore labeled (purple) secondary oligo to Mainstreet to visualize Chr19-9K Oligopaint oligos. MIPs were circularized via ligation and RCA, after which the first digit of the barcode was sequenced using O-LIT (green). Images are representative of two replicates. **e**, Comparison of secondary fluorophore signal (2°) versus first round sequencing signal (LIT) from puncta in panel b images. Center values are mean values (3.4 for 2o and 4.9 for O-LIT) with SD.

nature research

Corresponding author(s): Dr. Chao-ting Wu (twu@genetics.med.harvard.edu) and Dr. Marc A. Marti-Renom (martirenom@cnag.crg.eu)

Last updated by author(s): 05/23/2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Nikon Elements (NIS ElementsAR ver. 5.02.01.) and Vutara SRX software were used to acquire images.
Data analysis	ImageJ/Fiji (version 2.0.0-rc-69/1.52p) were used to align, normalize, contrast, overlay, and measure images as described in the Methods section. Python (version 2.7) with custom scripts was used for image analysis. Constrained K-means algorithm (version 1.5) https://zenodo.org/record/831850 was used for clustering. Integrative Modelling Platform () was used for Tier 2 (https://www.ncbi.nlm.nih.gov/pubmed/22272186). Domino (version) was used for Tier 2 tracing. Seaborn package for Python was used to generate plots. R (version 3.6.1) and R-Studio (version 1.2.1335) was used for initial image analysis. GraphPad Prizm (version 8.2) was used to generate plots. Microsoft Excel (version 16.16.7) was used to generate tables. Adobe Illustrator (version 22.0.1) was used to assemble figures. ChromoMap package for R by Lakshay Anand was used to generate chromosome cartoons (https://doi.org/10.1101/605600). Nikon Elements (NIS ElementsAR ver. 5.02.01.) was used to process images and deconvolution. Chimera was used for ball and stick visualizations (https://doi.org/10.1002/jcc.20084). Python scripts will be available on GitHub (https://github.com/3DGenomes/OligoFISSEQ).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. All raw and processed data will be made available upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed as we aimed to obtain images of as many cells as possible given experimental constraints during technology optimization. For all experiments with analysis, a minimum of 3 technological replicates were performed to confirm reproducibility. We deemed this to be sufficient due to low observed variability between samples. Datasets were imaged to assess and compare the efficiency of the different versions of OligoFISSEQ. The samples were aggregated to study the structural variability of the cell population. We verified that the sample sizes were sufficient to capture such variability by comparing OligoFISSEQ distance matrices with interaction frequency matrices obtained with Hi-C experiments as an orthogonal method (see “Distance heat-maps and Hi-C maps” of the Material and Methods).
Data exclusions	Cells that did not pass initial quality control filtering were not included in downstream processing and analysis. Cells in mitotic process were discarded from the analysis following the procedure described in the section “Detection efficiency and False Positives ratios” of the Material and Methods. A second exclusion is applied for cells which total detection efficiency in Tier 1 is below 25%. Those cells are mainly presenting imaging distortions in one or more channels or are falling in the border of the images. Exclusion criteria was pre-established, as we focused on interphase cells that were entirely imaged.
Replication	All replication attempts were successful and detailed in Fig. S3, S6, S7, and Table S13. Preferential chromosome positioning as identified in our study (Fig. S4, S5), was in line with reported observations in the literature. Additionally, we found that chromosomes segregated into distinct regions (territories) in the nucleus, also in line with observations from the literature (Fluorescent In Situ Hybridization and Hi-C studies).
Randomization	Cells used for imaging were selected randomly and all imaged cells that passed quality filters were used for analysis, therefore, there was no requirement for randomization.
Blinding	Blinding was not performed as experimental conditions were evident from the image data. Analysis and quantifications were performed using computational pipeline applied equally to all conditions and replicates for a given Oligopaint oligo library. Thresholds for detecting puncta were chosen for each Oligopaint oligo library (see Table S14) on graphs with objective properties that appeared indistinguishable across conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Anti-Alpha Tubulin (Sigma-Aldrich: T9026) used at 1:500.
Anti-GAPDH (Abcam: ab9483) used at 1:200
Anti-TOMM20 (Abcam: ab78547) used at 1:500
Anti-macroH2A.1 (Abcam: ab183041) used at 1:250
Donkey Anti Mouse Cy5 (Jackson ImmunoResearch Laboratories: 715-175-150) used at 1:500 from 1.25mg/mL stock
Donkey Anti-Rabbit Cy3 (Jackson ImmunoResearch Laboratories: 711-165-152) used at 1:500 from 1.25mg/mL stock
Bovine Anti-Goat Alexa Fluor 594 (Jackson ImmunoResearch Laboratories: 805-585-180) used at 1:500 from 1.25mg/mL stock

Validation

Anti-Alpha Tubulin has been validated by Sigma-Aldrich to be specific in human cell lines (osteosarcoma and breast cancer) using western blotting and in HeLa cells by immunofluorescence microscopy (https://www.sigmaaldrich.com/catalog/product/sigma/t9026?lang=en&region=US).
Anti-GAPDH has been validated by Abcam to produce positive signal in whole cell lysates from HeLa as well as human brain tissue lysate as well as positive immunofluorescence signal in HeLa cells (https://www.abcam.com/gapdh-antibody-loading-control-ab9483.html).
Anti-TOMM20 has been validated by Abcam to produce positive signal in HEPG2 whole cell lysate and positive immunofluorescence signal in HEPG2 cells (https://www.abcam.com/tomm20-antibody-mitochondrial-marker-ab78547.html).
Anti-macroH2A.1 has been validated by Abcam to produce positive signal in HAP1 lysates and reduced signal in HAP1 m2A1 knockouts (https://www.abcam.com/mh2a1-antibody-epr93592-ab183041.html).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

PGP1f (Human male fibroblasts, Coriell: GM23248), IMR-90 (Human female fibroblasts, ATCC: CCL-186)

Authentication

None of the cell lines have been authenticated.

Mycoplasma contamination

Cell lines were not tested for mycoplasma contamination but no indication of contamination was observed.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.