

RESEARCH ARTICLE

# Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling

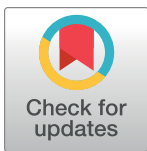
Guy Nir<sup>1</sup>, Irene Farabella<sup>2</sup>, Cynthia Pérez Estrada<sup>1,3,4</sup>, Carl G. Ebeling<sup>5</sup>, Brian J. Beliveau<sup>1,6,7,8</sup>, Hiroshi M. Sasaki<sup>6,7</sup>, S. Dean Lee<sup>1</sup>, Son C. Nguyen<sup>12</sup>, Ruth B. McCole<sup>1</sup>, Shyamtanu Chattoraj<sup>1</sup>, Jelena Erceg<sup>1</sup>, Jumana AlHaj Abed<sup>1</sup>, Nuno M. C. Martins<sup>1</sup>, Huy Q. Nguyen<sup>1</sup>, Mohammed A. Hannan<sup>1</sup>, Sheikh Russell<sup>3</sup>, Neva C. Durand<sup>3,9</sup>, Suhas S. P. Rao<sup>3,4,10</sup>, Jocelyn Y. Kishi<sup>6,7</sup>, Paula Soler-Vila<sup>2</sup>, Michele Di Pierro<sup>4</sup>, José N. Onuchic<sup>4</sup>, Steven P. Callahan<sup>11</sup>, John M. Schreiner<sup>11</sup>, Jeff A. Stuckey<sup>12\*</sup>, Peng Yin<sup>6,7\*</sup>, Erez Lieberman Aiden<sup>3,4,9,13\*</sup>, Marc A. Marti-Renom<sup>2,14,15,16\*</sup>, C.-ting Wu<sup>1,6\*</sup>

**1** Department of Genetics, Harvard Medical School, Boston, Massachusetts, United States of America, **2** CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain, **3** Center for Genome Architecture, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, United States of America, **4** Center for Theoretical Biological Physics, Rice University, Houston, Texas, United States of America, **5** Bruker Nano Inc., Salt Lake City, Utah, United States of America, **6** Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts, United States of America, **7** Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, United States of America, **8** Department of Genome Sciences, University of Washington, Seattle, Washington, United States of America, **9** Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, United States of America, **10** Department of Structural Biology, Stanford University School of Medicine, Stanford, California, United States of America, **11** Zero Epsilon, LLC, Salt Lake City, Utah, United States of America, **12** Bruker Nano Inc., Middleton, Wisconsin, United States of America, **13** Departments of Computer Science and Computational and Applied Mathematics, Rice University, Houston, Texas, United States of America, **14** Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona, Spain, **15** Universitat Pompeu Fabra (UPF), Barcelona, Spain, **16** ICREA, Barcelona, Spain

☞ These authors contributed equally to this work.

✉ Current address: Department of Genetics, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

\* [Jeff.Stuckey@bruker.com](mailto:Jeff.Stuckey@bruker.com) (JAS); [Peng\\_Yin@hms.harvard.edu](mailto:Yin@hms.harvard.edu) (PY); [erez@erez.com](mailto:erez@erez.com) (ELA); [martirenom@cnag.crg.eu](mailto:martirenom@cnag.crg.eu) (MAMR); [twu@genetics.med.harvard.edu](mailto:twu@genetics.med.harvard.edu) (CtW)



OPEN ACCESS

**Citation:** Nir G, Farabella I, Pérez Estrada C, Ebeling CG, Beliveau BJ, Sasaki HM, et al. (2018) Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLoS Genet* 14(12): e1007872. <https://doi.org/10.1371/journal.pgen.1007872>

**Editor:** Gregory P. Copenhaver, The University of North Carolina at Chapel Hill, UNITED STATES

**Received:** November 2, 2018

**Accepted:** December 4, 2018

**Published:** December 26, 2018

**Copyright:** © 2018 Nir et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Streets and MATLAB code used for the generation and appending of streets can be downloaded at <https://github.com/gnir/OligoLego>. PGP1f cells are also available from the Personal Genome Project (<https://www.personalgenomes.org/us>) upon acceptance of a standard MTA. Hi-C data will be placed in the Gene Expression Omnibus (GEO), where they are associated with accession number GSE123552. All data pertaining to modeling will be available through <http://3DGenomes.org/datasets>.

## Abstract

Chromosome organization is crucial for genome function. Here, we present a method for visualizing chromosomal DNA at super-resolution and then integrating Hi-C data to produce three-dimensional models of chromosome organization. Using the super-resolution microscopy methods of OligoSTORM and OligoDNA-PAINT, we trace 8 megabases of human chromosome 19, visualizing structures ranging in size from a few kilobases to over a megabase. Focusing on chromosomal regions that contribute to compartments, we discover distinct structures that, in spite of considerable variability, can predict whether such regions correspond to active (A-type) or inactive (B-type) compartments. Imaging through the depths of entire nuclei, we capture pairs of homologous regions in diploid cells, obtaining evidence that maternal and paternal homologous regions can be differentially organized. Finally, using restraint-based modeling to integrate imaging and Hi-C data, we implement a

**Funding:** This work was supported by funds from Ministerio de Ciencia, Innovación y Universidades of Spain (<http://www.ciencia.gob.es/>) (IJCI-2015-23352) to IF, Damon Runyon Cancer Research Foundation (<https://www.damonrunyon.org/>) and Howard Hughes Medical Institute (<https://www.hhmi.org/>) to BJB, Uehara Memorial Foundation Research (<https://www.taisho-holdings.co.jp/en/environment/social/sciences/>) to HMS, William Randolph Hearst Foundation (<https://www.hearstfdn.org/>) to RBM, EMBO (Long-Term fellowship) (<https://www.embo.org/>) to JE, NSF (Center for Theoretical Biological Physics, Rice University) (<https://www.nsf.gov/>) to MDP and JNO, NSF (CCF-1054898, CCF-1317291) (<https://www.nsf.gov/>), NIH (1R01EB018659-01, 1-U01-MH106011-01) (<https://www.nih.gov/>), and Office of Naval Research (N00014-13-1-0593, N00014-14-1-0610, N00014-16-1-2182, N00014-16-1-2410) (<https://www.onr.navy.mil/>) to PY, NIH (1DP2OD008540, U01HL130010, UM1HG009375, 4DP2OD008540) (<https://www.nih.gov/>), NSF (PHY-1427654) (<https://www.nsf.gov/>), USDA (2017-05741) (<https://www.usda.gov/>), Welch Foundation (Q-1866) (<http://www.welch1.org/>), NVIDIA (<https://www.nvidia.com/en-us/>), IBM (<https://www.ibm.com/us-en/?lnk=m>), Google (<https://www.google.com/>), Cancer Prevention Research Institute of Texas (R1304) (<http://www.cprit.state.tx.us/>), and McNair Medical Institute (<http://www.mcnaifoundation.org/what-we-fund/mcnair-medical-institute/>) to E.L.A., Horizon 2020 Research and Innovation Programme (676556) (<https://ec.europa.eu/programmes/horizon2020/en/>), European Research Council (609989) (<https://erc.europa.eu/>), Ministerio de Ciencia, Innovación y Universidades of Spain (BFU2017-85926-P) (<http://www.ciencia.gob.es/>), CERCA, and AGAUR Programme of the Generalitat de Catalunya and Centros de Excelencia Severo Ochoa (SEV-2012-0208) (<http://www.ciencia.gob.es/portal/site/MICINN/menuitem.7eac5cd345b4f34f09dfd1001432ea0/?vgnextoid=cba733a6368c2310VgnVCM1000001d04140aRCRD>) to M.A.M.R., and NIH (5DP1GM106412, R01HD091797, R01GM123289) (<https://www.nih.gov/>) to C-tW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** I have read the journal's policy and the authors of this manuscript have the following competing interests: C-tW, BJB, RBM, SCN, SC, and HQN hold or have patent filings pertaining to Oligopaints and related technologies, including other oligo-based methods for imaging. These technologies may be licensed to ReadCoor, a company in which C-tW holds equity. PY and

method—integrative modeling of genomic regions (IMGR)—to increase the genomic resolution of our traces to 10 kb.

## Author summary

Questions regarding the impact of chromosome structure on genome function are focusing increasingly on the manner in which chromosomes are organized within the nucleus. In fact, studies of processes as diverse as gene activation and repression as well as genome repair and stability are all querying how the 3D organization of chromosomal DNA may be a major player. Here, we apply our strategy for tracing chromosomes at super-resolution, traversing over 8 megabases of human chromosome 19 while visualizing genomic features ranging in size from kilobases to megabases. This technology has enabled exploration of the physical nature of a genomic feature called the compartment; compartments are widely hypothesized to reflect the partitioning of genomes into relatively more and less active regions. Excitingly, we find that compartments are, indeed, physical structures, that they are sometimes distinct and other times entangled. We also find that the maternally-derived and the paternally-derived homologous regions can differ more than would be expected by chance. Finally, by integrating image data with information regarding the frequency with which genomic segments contact each other, we produce a 3D model of how the 8 megabases we have imaged may be organized within a single nucleus, achieving 10 kb genomic resolution.

## Introduction

In terms of organization, chromosomal DNA is among the most challenging of biological molecules to study. It is massive, differs in sequence and structure from one region to the next, and often comes in pairs of homologs that defy easy distinction. Nevertheless, much progress has been made via genetic, cytogenetic, and epigenetic analyses, including DNA-DNA proximity ligation assays, such as Hi-C and other Chromosome Conformation Capture technologies [1–8], as well as three-dimensional (3D) modeling (reviewed by [9, 10]). Recently, researchers have used diffraction-limited microscopy and fluorescent *in situ* hybridization (FISH) probes to image up to 40 locations along single human chromosomes in sequential rounds of hybridization [11], each round targeting the central 100 kilobase (kb) portion of a Hi-C defined self-interacting contact domain, also known as a topologically associating domain, or TAD [3–7]. By connecting the resulting 40 positions, this strategy revealed the overall organization of four chromosomes, including the 3D segregation of compartmental intervals by epigenetic state; compartmental intervals are chromosomal regions whose loci share long-range contact patterns contributing to the relatively active (A-type) and inactive (B-type) compartments of the genome and appear in Hi-C maps as a plaid pattern extending beyond the Hi-C diagonal [3]. This study has since been used as a guideline for Hi-C contact-based modeling (bioRxiv [12]). Similarly, up to 2.5 Mbs of a human chromosome has been traced using contiguous steps of only 30 kb, thus enabling diffraction-limited imaging to provide a super-resolved path and reveal the dynamic nature of contact domains and their relationship to compartments [13]. Super-resolution [13–20] and electron [21] microscopy have also been used to dissect genomic features, including sequence-specific super-resolution imaging of contact domains [13, 15, 17, 20], bioRxiv [22], the most recent tracing 1.2 Mbs of a human chromosome in contiguous 30

BJB have filed patents related to DNA-PAINT. DNA-PAINT technology has been licensed to Ultivue Inc., a company in which PY is an equity holder and cofounder. PY is also a co-founder of NuProbe Global.

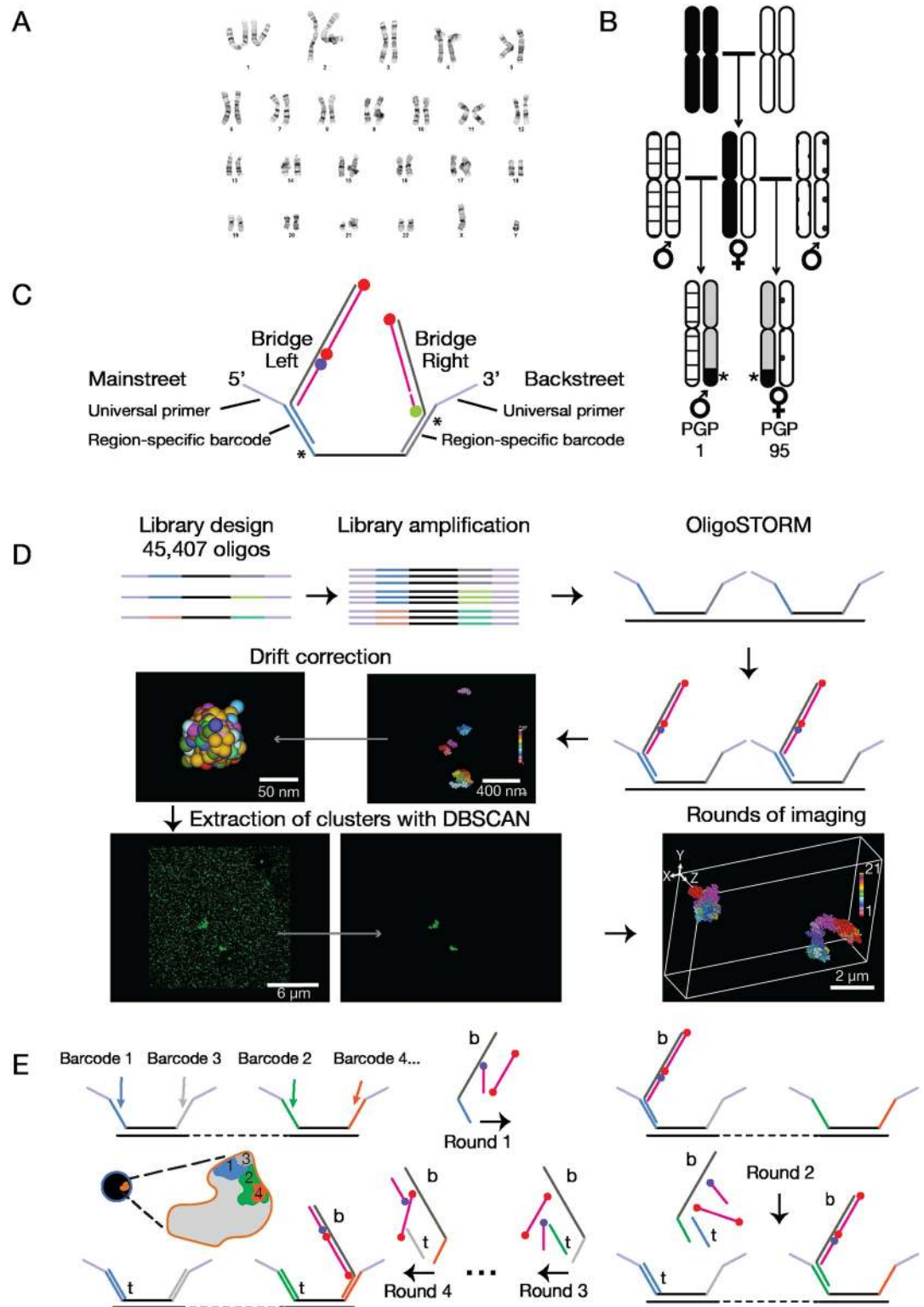
kb steps [13]. Such super-resolution studies have provided first glimpses of the physical nature of contact domains, such as their volumes and shapes.

What remains relatively unexplored at super-resolution is the physical nature of compartmental intervals. Here, we present a method that can scan the entire depths of nuclei using super-resolution microscopy and sequential hybridization and, thus, capture genomic regions of the size of compartmental intervals as well as all homologous copies of those regions in a single nucleus. Indeed, using this technology, we have produced an 8.16 Mb continuous trace of multiple, contiguous compartmental intervals, including both homologous copies in diploid nuclei, providing measurements of distance, entanglement, surface area, volume, and sphericity as well as gaining insights into inter-homolog differences. We then use a new data-driven protocol that, for the first time, integrates imaging with Hi-C interaction matrices in order to generate 3D models of chromosome organization, achieving 10 kb genomic resolution. We begin by describing our strategy for walking along human chromosome 19 in primary skin fibroblasts of donor PGP1 (XY, PGP1f; [Methods](#)), using oligonucleotide-based Oligopaint FISH probes [23, 24] and imaging in sequential steps using the single-molecule localization microscopy methods of OligoSTORM and OligoDNA-PAINT to achieve super-resolution [14, 24]. The resulting images resolved the physical nature of individual compartmental intervals, showing considerable variation from cell to cell yet matching, nevertheless, a newly generated Hi-C map of PGP1f cells, including the assignment of A and B compartments. We then integrate our image and Hi-C data to produce a single-cell, 10-kb resolution 3D model for the entire 8.16 Mb walk. As the PGP1 genome has been haplotyped and its pedigree is known, ([Fig 1A and 1B](#); ref. [25]; [Methods](#)) we also describe a strategy whereby the maternal and paternal homologs can be distinguished, discovering that homologous regions within the same nucleus differ in ellipticity more than would be expected at random.

## Results and discussion

We considered two approaches for walking along the chromosome—one in which the sizes of the steps are in accordance with genomic features and another in which step sizes are uniform—initiating our studies with the first. Thus, we designed Oligopaint probes corresponding to the Hi-C defined features of compartments as well as smaller features, including contact domains, loops, and genes. Because a Hi-C map of PGP1f cells was not available when we initiated our study, we designed our walk based on the Hi-C data of IMR-90 lung fibroblasts ([S1 Fig](#); ref. [8]; [Methods](#)). Focusing on an 8.16 Mb region extending from chr19:7,400,000 to chr19:15,560,000, we then generated Oligopaint probes to image the region in 9 segments ranging in size from 360 kb to 1.8 Mb, including probes that would permit visualization of features as small as 2.9 kb ([S1 Table](#)).

Similar to fluorophore- (dye-) conjugated oligonucleotides (oligos) pioneered over two decades ago [26] as well as other iterations ([Methods](#)), Oligopaint probes consist of computationally designed oligos ([Fig 1C](#); ref. [23]). In addition, they are single-stranded and carry a short region (e.g., ~32–42 nucleotides, nts) of genomic homology as well as nongenic sequences, called Mainstreet and Backstreet, at their 5' and 3' ends, respectively. Streets permit Oligopaints to be amplified, multiplexed, barcoded, and indirectly visualized via labeled secondary oligos [11, 13, 14, 16, 23, 24, 27–30]. Here, we use Streets to enable OligoSTORM and OligoDNA-PAINT [13, 14, 16, 24]. OligoSTORM melds Oligopaints with STORM (Stochastic Optical Reconstruction Microscopy; ref. [31]) or dSTORM (direct STORM; ref. [32]) and generates single events of fluorescence (localizations) via dye-pairs or single dyes. As for OligoDNA-PAINT, it combines Oligopaints with DNA-PAINT (DNA-points accumulation for imaging in nanoscale topography; ref. [33, 34]) and generates localizations via transient



**Fig 1. Tools.** (A) PGP1f karyotype (Methods). (B) Assignment of parental origin to homologs of chromosome 19 through pedigree analysis. Black and white, homologs of chromosome 19 in maternal parent of PGP1 and PGP95; patterned, homologs of paternal parents; asterisk and black, large SNV block shared between PGP1 and PGP95 identifying maternal homolog; grey, maternally derived regions without large SNV blocks shared by PGP1 and PGP95. (C) Oligopaint oligos (129–135 nts) consist of a segment of genomic homology (black; 35–41 nts) flanked by Mainstreet and Backstreet (47 nts), each consisting of

universal primers (violet; 20 nts), region-specific barcodes (blue and grey; 20 nt), and 7-nt sequences (asterisks) which, in combination with barcodes, generate binding sites for toehold oligos. A bridge oligo (left) bound to a Mainstreet barcode allows labeled secondary oligos (blue, Alexa Fluor 405; red, Alexa Fluor 647) to co-localize and enable OligoSTORM. A bridge (right) bound to a Backstreet barcode allows a labeled imager strand (green, Cy3B) to bind transiently, enabling OligoDNA-PAINT. (D) Workflow from library design through probe generation, OligoSTORM, drift correction using 90 nm fiducial markers, extraction of clusters with DBSCAN, and image assembly, as described in the text. (E) Walking along chromosomes using sequential hybridizations. All Oligopaint oligos are hybridized to the genome at once, and then rounds of hybridization and imaging are conducted. Each round brings in bridges ('b'), labeled secondary oligos, and, except for round 1, toehold oligos ('t') in order to label a new region of the walk while removing labeled oligos from the previously imaged region.

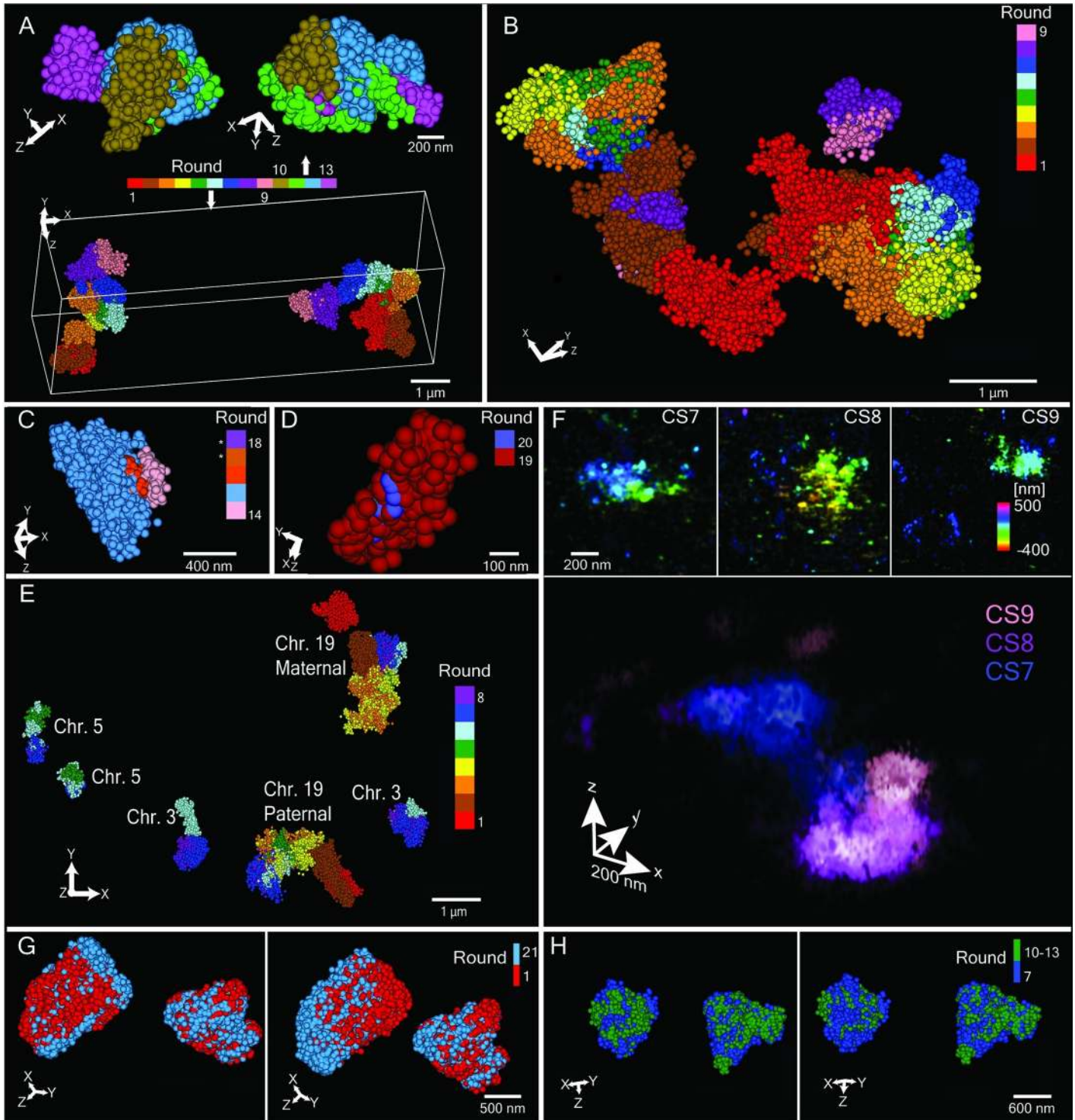
<https://doi.org/10.1371/journal.pgen.1007872.g001>

hybridizations of short (~9–10 nts) dye-labeled oligos, called imager strands, to complementary sequences, called DNA docking sites, located on the Streets; here the transient binding of imager strands is interpreted as localizations.

Consisting of 45,407 species of oligos, our Oligopaint library incorporated three key features (Fig 1C). First, all the oligos avoid sequences that differ between the parental genomes of PGP1f and, thus, they bind both homologs of chromosome 19. Second, both streets of each oligo carry 20-nt long barcodes corresponding to sub-regions of our 8.16 Mb walk; these barcodes confer multiple functionalities on the oligo. For example, when enabling OligoSTORM, these barcodes co-localize “bridge” oligos [13, 14] that recruit the labeled secondary oligos to be visualized by OligoSTORM. When enabling OligoDNA-PAINT, the bridges include a docking site for the imager strand. The barcodes also permit any targeted genomic region to be imaged twice, once via Mainstreet and then, again, via Backstreet. Finally, each street carries a 7-nt sequence (Fig 1C, asterisk) that facilitates sequential rounds of imaging; in conjunction with a flanking barcode, these 7 nts generate a region which, when bound by a 27-nt “toehold” oligo [35–37], dislodge any previously bound bridge.

Our workflow (Fig 1D; Methods) begins with T7 RNA polymerase-mediated amplification [24, 27, 38] of the entirety of an Oligopaint library that is to be used for a walk. The totality of the resulting single-stranded oligos is then hybridized to fixed cells in a single round of denaturation and hybridization [11, 13]. When implementing OligoSTORM, the sample is then placed on a 3D single-molecule localization microscope (Bruker Vutara 352) and scanned to identify nuclei to be imaged. Sequential rounds of imaging are then conducted with an automated microfluidic system (Fig 1E), each round targeting one step of the walk by hybridizing a specific bridge oligo (Fig 1E, labeled with 'b') along with its dye-conjugated oligos, after which images are acquired one nucleus at a time until all nuclei have been imaged; the next round is initiated through the use of toehold oligos (Fig 1E, labeled with 't'), which dislodge bridges bound in the previous round. To traverse the depth of PGP1f nuclei, imaging is conducted in 100 nm increments along the axial (Z) dimension for up to 4  $\mu\text{m}$ , well within the volumetric capability of the Vutara. This protocol has given an average localization precision of 11.14 +/- 1.47 nm and 11.02 +/- 1.22 nm along the lateral axes and 46.96 +/- 6.02 nm along the axial dimension, corresponding to average supported optical resolutions of  $26 \pm 3$  nm and  $112 \pm 11$  nm, respectively (Methods). All localizations are then subjected to drift correction with respect to fiducial beads, after which density-based spatial clustering of applications with noise (DBSCAN; ref. [39]) extracts clusters of localizations most likely to represent genomic regions (Fig 1E; Methods).

Using our workflow, we have walked the entirety of the 8.16 Mb region in over 60 chromosomes, separately imaging all 9 chromosomal segments via Mainstreet (CS1-9; Fig 2A and 2B; S1 Movie; S1 Table). These studies also demonstrated the effectiveness of using Backstreet in conjunction with Mainstreet to image smaller features lying within larger regions, all in one imaging run (S1 Table). For example, in a single imaging run that imaged CS1-9 using Mainstreet (Fig 2A and 2B, rounds 1–9), we used Backstreet to image four subregions comprising



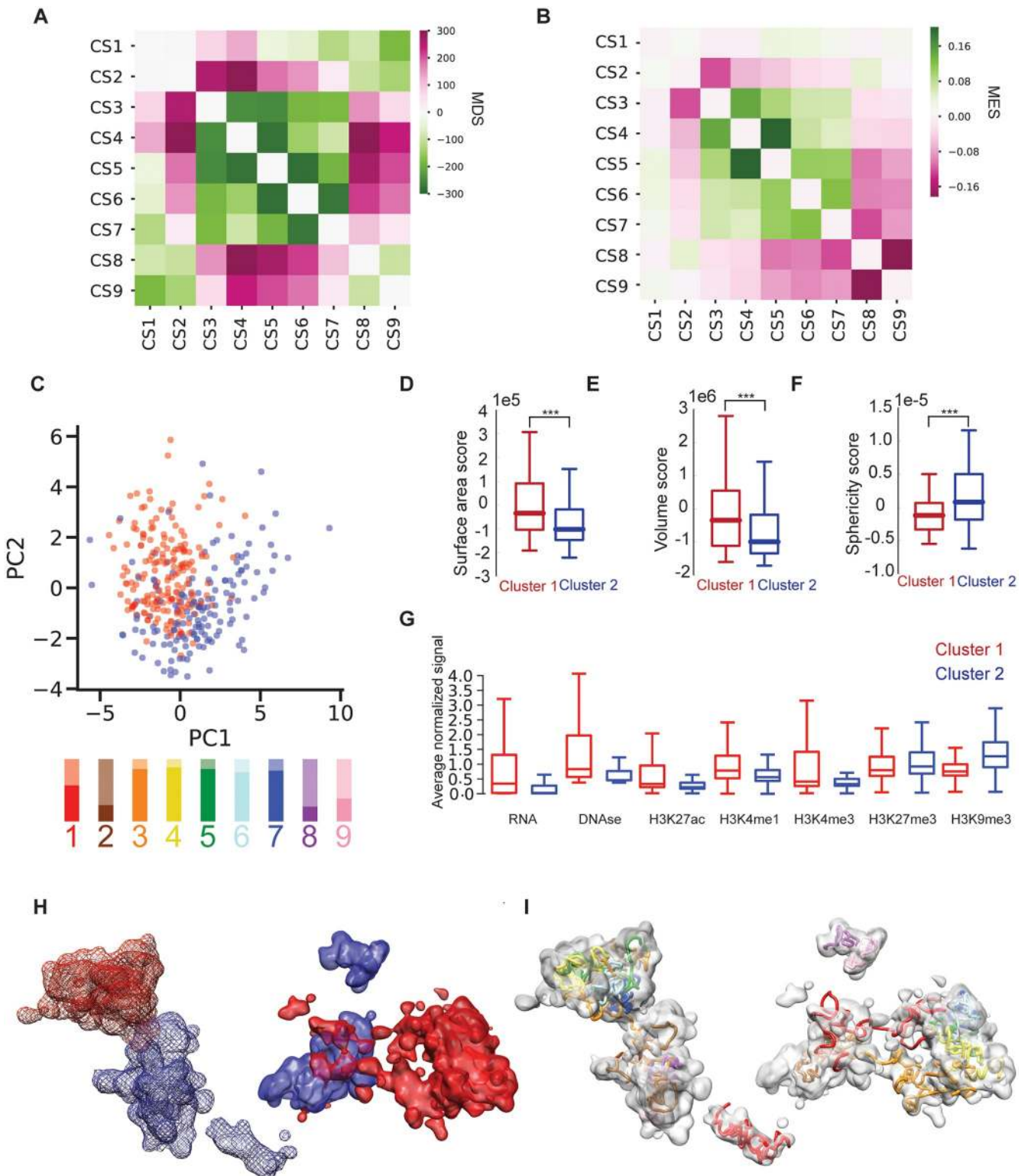
**Fig 2. Tracing PGP1f chromosomes.** (A) OligoSTORM image of a diploid nucleus showing CS1-9 (Bottom, rounds 1 through 9; 1.28, 1.24, 1.80, 1.04, 0.56, 0.52, 0.84, 0.52, and 0.36 Mb, respectively) and four subregions of CS7 (top, rounds 10 through 13; 140, 260, 350, and 90 kb, respectively) of both homologs. In this and all panels excepting F, radius of spheres represents the localization precision in the axial dimension. Note, the coordinates of the four subregions, selected based on the IMR-90 Hi-C map, correspond well to contact domains. Sequencing depth of the PGP1f Hi-C map was not, however, able to confirm these as contact domains in PGP1f, and thus they are referred to simply as subregions. (B) Same as in A with respect to CS1-9, but a different nucleus. (C) Loop (290 kb), including loop anchors (rounds 14 and 16; pink and orange; 10 kb each) and loop body (round 15; blue; 270 kb); \*, we also imaged the upstream and downstream regions flanking the loop (rounds 17 and 18; 20 and 80 kb, respectively), but these are not shown so that the structure of the loop can be more apparent. (D) *DNMT1* gene (round 19; 59.5 kb) and its DMR (round 20; 2.9 kb). (E) Walking along chromosome 19 in variable step sizes (see A) while also walking along chromosomes 3 and 5 in uniform step sizes (500 and 250

kb, respectively). (F) OligoDNA-PAINT images of CS7-9 of one homolog showing each chromosomal segment by itself (top; color scale represents depth in z) or merged (bottom; color scale represents different chromosomal segments). In contrast to other panels, localizations are blurred according to precision. (G) Superimposition of two images, taken 90 hours apart, each of both homologs of CS1 in (red, first image; blue, later image). Average overlap is  $80 \pm 9\%$  and  $68 \pm 12\%$  ( $n = 20$ ) when images are aligned based on their centers of mass (left) or not (right), respectively. (H) Superimposition of two images from one nucleus of both homologs of CS7, one image encompassing all of CS7 (blue) and the other being a composite of the four subregions (green; 140, 260, 350, and 90 kb) comprising CS7. With alignment based on centers of mass, 85% of the single image overlapped the composite image, and 76% of the latter overlapped the former (left). Without alignment, the analogous values were 84% and 75% (right).

<https://doi.org/10.1371/journal.pgen.1007872.g002>

CS7 (Fig 2A, rounds 10–13; S2 Fig). We have also imaged features as small as a loop (290 kb) in CS6, separately visualizing its anchors (10 kb), central body (270 kb), as well as flanking regions (Fig 2C), and the Dnmt1 gene (59.5 kb) as well as its differentially methylated region (DMR; 2.9 kb) [40] in CS3 (Fig 2D). As a further demonstration of the capacity of barcodes, Fig 2E illustrates a special use of them to enable simultaneous walks along multiple chromosomes in single nuclei. In particular, taking advantage of the propensity of different chromosomes to lie within their own territories [41] and thus for images of them to be by and large nonoverlapping, we have traced the 8.16 Mb region of chromosomes 19 while also walking along chromosomes 3 and 5 in uniform step sizes of 500 and 250 kb, respectively (Fig 2E; S2 Table; Methods). S3 Fig presents a larger field of view that also shows walks on both homologs of chromosome 8, for which step sizes were 100 kb (S3A Fig), as well as additional images of CS1-9 (S3B Fig), the four subregions of CS7 (S3C Fig), a loop (S3D Fig), and DNMT1 and its DMR (S3E Fig). We have also used our library to image CS7, 8, and 9 using OligoDNA-PAINT, achieving an exciting supported resolution of 18 nm (Fig 2F; S3 Table; Methods; ref. [42]). Finally, to assess the reproducibility of our images, we have conducted a number of studies using OligoSTORM (Methods). In one, we imaged CS1 using Mainstreet and then imaged it again 90 hours later using Backstreet, achieving an average spatial overlap of  $80 \pm 9\%$  ( $n = 20$ ; Methods) with minimal change in the average number of localizations ( $n = 4,8371$  for round 1 and 4,912 for round 21; Fig 2G, S4 Fig). Comparable results were obtained (75%–85%) when we imaged CS7 via Mainstreet and then reimaged it using barcodes on Backstreet to sequentially visualize four subregions comprising CS7 (Fig 2H).

Intriguingly, we noticed early on that a number of consecutive steps of our walk appeared minimally entangled, suggesting that they correspond to distinct physical entities. To explore this, we further analyzed 38 OligoSTORM walks of CS1-9; these 38 represented the 19 nuclei with the highest image quality and for which we observed exactly two foci of signal per chromosomal segment and no evidence of aneuploidy (Methods). These nuclei were analyzed by convolving them with a Gaussian kernel [43], which enabled us to represent each chromosomal segment of each chromosome as a distribution of localization densities in 3D space (Methods). To assess the structural arrangement of chromosomal segments, we used our density maps and calculated pairwise distances between centers of mass (Distance score, DS) and overlap (Entanglement score, ES) (S5 Fig; Methods). The results revealed that chromosomal segments are not randomly placed with respect to each other. Although there was considerable variation in distances (S5A and S5B Fig), the observed distributions showed that the five central segments, CS3-7, were closer than expected in almost all pairwise combinations (Fig 3A, S5B Fig). Overall, these central segments were also spatially separated from the segments lying at the beginning (CS1 and CS2) and end (CS8 and CS9) of the walk. Importantly, even with only 38 walks, these differences from expectation were statistically significant for most pairwise comparisons (S5B Fig). Moreover, segments at the internal edges of the central region were further than expected from segments just beyond the central region; CS3 was further from CS2, and CS7 was further from CS8 than expected. Interestingly, the beginning (CS1 and CS2) and end (CS8 and CS9) of the walk trended towards being closer than expected. Altogether,



**Fig 3. Image analysis and integrative modeling.** (A, B) Matrices for (A) mean spatial distance scores (MDS) and (B) mean entanglement scores (MES) for CS1-9. (C) Unsupervised PCA of 3D features for the 342 visualized chromosomal segments suggests two major clusters. Score plot for chromosomal segments color-coded by cluster type: cluster 1 in red and cluster 2 in blue. Lower panel, proportion of images of each chromosomal segment in cluster 1 (opaque color) or cluster 2 (transparent color); the chromosomal segments are colored as in Fig 2A and 2B. (D-F) Variation and distribution of the two clusters in terms of (D) surface area, (E) volume, and (F) sphericity scores are shown as box plots, where boundaries represent 1st and 3rd quartiles, middle



line represents median, and whiskers extend to 1.5 times the interquartile range (Mann-Whitney rank test, \*\*\*:  $p < 10^{-3}$ ). (G) Mean normalized signal for RNA-seq (RNA), DNase I hypersensitive sites (DNase), and epigenetic modifications in PGP1f nuclei for the two identified clusters. Red, cluster 1; blue, cluster 2. (H) Density map of the 9 chromosomal segments for both copies spanning the entire 8.16 Mb region in a single PGP1f nucleus (same as that shown in Fig 2B) as imaged by OligoSTORM, color-coded by cluster type. Red, cluster 1; blue, cluster 2. One homolog is displayed with a solid surface, while the other homolog is displayed with a mesh surface. Note spatial segregation of clusters. (I) Density map (gray) representation of CS1-9 as in H except displayed, here, with the fitted 3D model obtained via IMGR (colored coded as in Figs 2A and 3B). Images were created with Chimera (Methods).

<https://doi.org/10.1371/journal.pgen.1007872.g003>

these observations suggested the center of the walk to be spatially separated from the ends, with the two ends being closer in space than expected.

Our analyses of entanglement also revealed a nonrandom pattern (Fig 3B, S5C and S5D Fig; Methods [43]). Although, again, there was considerable variation (S5C and S5D Fig), the distributions of entanglement scores revealed that the central segments, CS3-7, were more entangled than expected or trended as such, while the segments lying at the internal boundaries of this region (CS3 and CS7) entangled less than expected with the segments lying upstream and downstream, respectively (Fig 3B, S5D Fig). The contiguous pairs of segments lying at the ends of the walk were also minimally entangled; CS1 and CS2 showed insignificant entanglement, while CS8 and CS9 resulted in the lowest of all the entanglement scores. Thus, it is especially noteworthy that CS1 and CS2, lying at one end of the walk, nevertheless showed significant entanglement with CS9 and CS8, respectively, at the other end (Fig 3B, S5D Fig). These observations aligned with our analyses of distance, implying a 3D segregation of the 9 chromosomal segments into two groups, one consisting of the central segments (CS3-7) and the other consisting of the ends of the walk (CS1, 2, 8, and 9). This was confirmed by unsupervised clustering of the mean distance and entanglement matrices (S5E and S5F Fig).

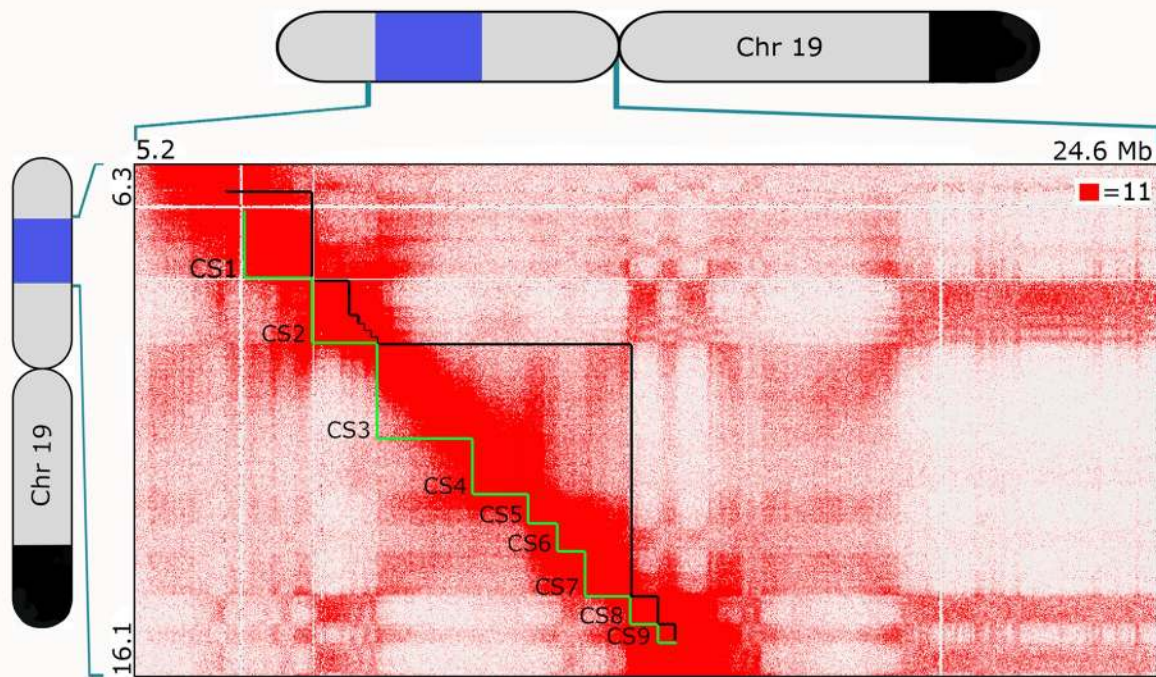
To further explore such segregation, we calculated the surface area, volume, and sphericity for each chromosomal segment, corrected for genomic size (S6 Fig; Methods) and then, combining these features with the 8 measures of distance (DS) and 8 measures of entanglement (ES), conducted unsupervised clustering via Principal Component Analyses (PCA) for all 342 imaged chromosomal segments (38 chromosomes x 9 chromosomal segments) (Methods; S7 Fig and Fig 3C). Despite the variation among them, two major clusters emerged, with the first two principal components accounting for 24.2% and 14.0% of the variability (Fig 3C, S7A Fig); the first cluster comprised 84.5% of all the images of CS1 and CS3-7, while the second cluster comprised 62.7% of all the images of CS2, 8 and 9. Considered together, the segments of cluster 1 had larger area and volume and were less spherical than the aggregate of the segments in cluster 2 (Fig 3D–3F). Interestingly, the segments that were primarily in cluster 1 harbored epigenetic markings of active chromatin, while those primarily in cluster 2 harbored markings of inactive chromatin (Fig 3G; Methods). These findings align with the differing degrees of chromatin compactness that have been observed for different epigenetic states in *Drosophila* [16].

Our PCA analyses also highlighted variability in the classifications for all chromosomal segments (S4 Table). For example, while CS3-7 fell primarily in cluster 1 and CS2 and CS8 fell primarily in cluster 2, CS1 and CS9 resulted in much more mixed classifications, with ratios of cluster 1:cluster 2 being 58:42 and 37:63, respectively (Fig 3C, bar graph). CS1 is particularly noteworthy in this regard; the distributions of PC1 values revealed that all segments were skewed towards one cluster type with the exception of CS1 (S7B–S7J Fig), which was also the most mixed in terms of intra-nuclear and inter-nuclear variation (S7K Fig). In brief, our data suggest that, while chromosomal segments can be broadly classified by their structural properties into categories reminiscent of active and/or inactive chromatin, they are also variable and dynamic in their character, potentially reflecting any number of attributes, including their state of activity and different stages of the cell cycle (Methods).

In sum, our analyses revealed organizational traits for the imaged region in single cells. First, CS3-7, can form an internally entangled entity. Second, this region can be distanced from, as well as less entangled with, its flanking segments, pointing to the potential of distance, *per se*, to be an organizational principle. Third, the entire region can fold back on itself such that its ends come into proximity and entangle to a modest degree (see example in Fig 2B). Fourth, CS1, 2, 8, and 9 may each correspond to an individual entity, suggesting that the imaged region may be composed of at least 5 structurally distinct sections. Fifth, the two clusters of segments revealed by PCA may correspond to spatially separated compartments (Figs 2B and 3H), active (A-type) and relatively inactive (B-type), consistent with observations from Hi-C [3] and diffraction-limited imaging [11] as well as models [44] for how phase separation [45, 46] may contribute to the physical distinction of compartment types. Finally, while some chromosomal segments were predominantly active or inactive, others could be classified as in either state or mixed, highlighting the capacity of super-resolution imaging to refine our understanding of epigenetic states by contributing measurements of the physical nature as well as variability of structures. Thus, even as our data align with observations that the 3D genome can be compartmentalized and organized by a variety of functions (e.g., refs. [3, 47–51]), they also emphasize the potential of chromosomal regions to slip from one milieu into another.

How well might our images correlate with population-based data from DNA-DNA proximity ligation based technologies? To address this, we generated an *in situ* Hi-C map for PGP1f cells (Fig 4; S5 Table; Methods) and assigned loci in the target region to A and B compartments based on GC content [52]. The annotation assigned CS1, CS3-7, and CS9 to the A compartment, whereas CS8 and most of CS2 were assigned to the B compartment (Fig 4, S8 Fig). Notably, the fact that CS3-7 fell in a single compartment meant that all loci in this extended interval exhibited enhanced contact frequencies with one another. Additionally, we observed a statistically significant correlation ( $\sim 0.6$  Pearson Correlation, S9 Fig) between the enrichment of Hi-C contacts between chromosomal segments and their mean distance score obtained by imaging. These findings recapitulate several features of the imaging data: [i] five physical entities (CS1, CS2, CS3-7, CS8, and CS9) with distinctive chromatin states, [ii] a single, extended structure spanning CS3-7, [iii] elevated contact frequency between the endpoints of the target region, and [iv] strong correspondence between the chromosomal segments assigned from Hi-C analysis to the A compartment (CS1, CS3-7, and CS9) and those that were most frequently assigned from image analysis to cluster 1 (CS1 and CS3-7) and, likewise, a strong correspondence between segments assigned by Hi-C to the B compartment (CS2 and CS8) and those most frequently assigned by imaging to cluster 2 (CS2, CS8, and CS9), with CS1 and CS9 being the segments that most straddle the two clusters (S4 Table). In brief, our images reveal that compartmental intervals have the capacity to form distinct structures. Moreover, the strong correspondence we observed between the outcomes of two very different technologies, FISH-based super-resolution imaging and Hi-C, lends confidence to our image-based measurements of distance, entanglement, size, volume, sphericity, and structural variability as well as to the overall potential of OligoSTORM to elucidate the organization of large swaths of, if not entire, chromosomes. It demonstrates, in particular, that key features of chromosome structure can be maintained to a reasonable extent through the various steps in which samples are prepared for FISH, including treatment of chromosomal DNA to heat-denaturation and binding of nucleic acid probes.

Next, we developed a method, inspired by the fitting of proteins to cryoEM density maps [53] and named *integrative modeling of genomic regions* (IMGR), that combines different modalities of data. Here, we integrated OligoSTORM images and 10-kb resolution Hi-C data to produce a 3D model of two homologous regions in a single diploid nucleus (Fig 3I, S10A Fig; S2 Movie; Methods; Supporting information S1 Text). Population-based 3D modeling of



**Fig 4. *In situ* Hi-C map of CS1-9 in PGP1f.** Hi-C map of region including CS1-9, displayed at 25 kb resolution. (See [Methods](#) for data processing and normalization.) Contact count indicated by the color of each pixel, ranging from 0 (white) to 11 (red); OligoSTORM steps (CS1-9) delineated in green below the diagonal, and compartmental intervals delineated in black above the diagonal. Cartoons of chromosome 19 show extent of the 8.16 Mb imaged region (blue) and block of SNVs shared between PGP1 and PGP95 (black; as in [Fig 1B](#)).

<https://doi.org/10.1371/journal.pgen.1007872.g004>

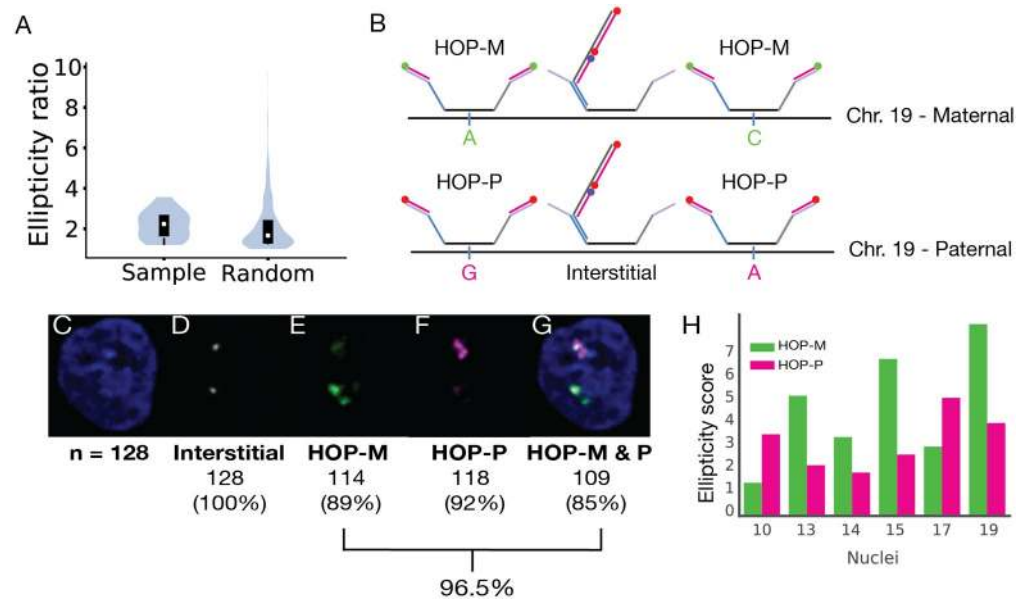
Hi-C data has been used previously to recapitulate diploid genomes [54], while single-cell Hi-C datasets have enabled haploid reconstruction [55, 56] as well as homolog-specific modeling of diploid human cells [57]. In this study, we addressed diploidy through a new, two-step integrative process that fits each conformation in an ensemble of Hi-C derived 3D models (built as previously described [58, 59] using 10-kb resolution PGP1f Hi-C data) into single-cell OligoSTORM density maps at the level of chromosomal segments (S10A Fig; [Methods](#); Supporting information [S1 Text](#)). First, the optimal positions and orientations of the 3D models of a chromosomal segment were found by maximizing the goodness-of-fit score with respect to the OligoSTORM density map. This step allowed us to filter out 3D models with conformations that were incompatible with the density map. Second, the resulting optimal rigid-fitted 3D models were optimized with 50 cycles of flexible fitting refinement using a simulated annealing molecular dynamics protocol [53]. Importantly, the scoring function used in this second step included terms (ECCC) assessing correlations with density maps as well as additional terms for satisfaction of the spatial restraints imposed by Hi-C data. In this way, we achieved a resolution of 10 kb for a genomic region that had been imaged in step sizes ranging from 0.36 to 1.8 Mb in size, the outcome being an integrative model that best satisfied both the Hi-C as well as the imaging data.

An exciting outcome of using modeling to decipher images at a fine resolution is the potential for gaining insights into the 3D organization of local structures. For example, chromosome 19 is extraordinarily rich in zinc-finger genes, which are clustered in 6 locations [60], and

[S10B Fig](#) models the arrangement of the two clusters located in the region we imaged; one lies at one end of CS2 and another extends across CS4. Similarly, all genes in the imaged region that are classified by OMIM as disease-causing [61] can now be positioned in 3D ([S10C Fig](#)). The models can also be used to map ChIP-seq data in 3D space [62]. For instance, it appears that one end of each copy of the region is decorated by a large patch of H3K4me1, H3K27ac, and H3K4me3 ([S10D–S10F Fig](#)). These patches suggest enhancer-promoter clusters and, consistent with this, RNA-seq and DNase accessibility data indicate that the patches co-localized with foci of expressed genes and greater accessibility ([S10G and S10H Fig](#)). In contrast, available ChIP-seq data for facultative (H3K27me3) and constitutive (H3K9me3) heterochromatin suggest multi-foci patches of heterochromatin throughout ([S10I and S10J Fig](#)).

Finally, we have explored the potential of our approach to address questions regarding the relationship between homologous chromosomal regions. Indeed, studies across a wide variety of species have shown that diploidy does more than provide an heir and a spare, with a burgeoning number of homology effects [63, 64] demonstrating that a chromosomal region can be far from indifferent to its homolog(s), that there can be molecular awareness, coordination, and even physical pairing (reviewed by [65, 66]). Structurally, nonrandom allelic differences in accessibility and volume, in some cases showing parent-of-origin effect, have been noted at the level of individual genes (e.g., refs. [67, 68]). Here, we asked whether structural differences between homologs can be detected across large chromosomal regions and, if so, whether they reflect parental origin. To address this, we calculated an ellipticity ratio to compare two homologous regions within each of the 19 nuclei examined and then ascertained whether this ratio differed from what might be expected at random. In particular, we fitted each imaged 8.16 Mb region to an ellipsoid, calculated an ellipticity score by dividing the length of the longest axis of the ellipsoid by that of the next longest axis, and then, for each nucleus, generated a ratio of the two ellipticity scores representing the two homologous regions, always dividing the larger by the smaller score ([Methods](#)). Intriguingly, the median ellipticity ratio was 2.5 ( $\pm 0.54$ ), which was significantly greater than the score for randomly selected pairs of homologous regions ( $1.3 \pm 0.48$ ; [Fig 5A](#), [S11A Fig](#), [S6 Table](#)), suggesting that megabase-sized homologous regions of a single nucleus may differ nonrandomly in ellipticity and that the difference is, on average, greater between homologous regions within a nucleus than between nuclei. Whether this difference reflects a regulatory function and, if so, whether that regulation is merely structural or also has functional implications for homolog-specific functions remains to be determined. These observations nevertheless argue that genomic regions can respond to the presence of homology, with one potential outcome being structural distinction from their homologs. Perhaps noteworthy in this context are phenomena that differentially affect homologous chromosomes in a chromosome-wide fashion, such as those underlying X-inactivation or asynchronous replication of random monoallelically expressed genes (reviewed by [69]). In light of these phenomena, our findings may indicate chromosome-wide structural anticorrelations between homologous chromosomes that could even extend genome-wide.

To ascertain whether greater ellipticity might correlate with parental origin, we turned to the most recent of our 38 walks as, for these, we had been successful in assigning parental origin. Here, we had implemented homolog-specific Oligopaints (HOPs; [Methods](#); ref. [14]), which target single nucleotide variants (SNVs; also see ref. [70, 71]) that differ between the haplotypes of a genome and, therefore, permit distinction of homologous genomic regions. While proven successful in *Drosophila* and mice with abundant SNVs ( $\sim 2\text{--}5$  SNVs per kb; ref. [14]), their applicability to humans, which have a lower frequency of SNVs ( $\sim 0.77$  SNVs per kb for PGP1), was unknown. Thus, we generated two libraries, HOP-M and HOP-P, the former targeting 11,259 maternal-specific SNVs extending across the entirety of chromosome 19, and the latter targeting the analogous 11,259 paternal-specific SNVs ([Fig 5B](#); [Methods](#)). These



**Fig 5. Homolog-specific Oligopaints for PGPIf.** (A) Violin plots of the ratios of 19 pairs of ellipticity scores, each pair representing the two homologs of one of the 19 imaged nuclei (Sample) or of 1,000 pairs of ellipticity scores, each pair representing two chromosomes chosen at random from the 38 representing the 19 imaged nuclei (Random). Boundaries of the black box-plot represent 1st and 3rd quartiles, white dot represents the median, and whiskers extend to 1.5 times the interquartile range (Mann-Whitney rank test,  $^* p < 10^{-2}$ ). (B) The HOP-M and HOP-P probes for chromosome 19 each encompass the entire chromosome and contain 11,259 oligos that cover  $\geq 1$  SNV per oligo. Thus, they are in contrast to Oligopaint oligos used to image CS1-9, these latter probes being “interstitial” in nature, as they avoid SNVs. HOP-M and HOP-P probes are visualized with secondary oligos labeled with different dyes, such that the two probes can be distinguished. (C-G) Images of a nucleus visualized with DAPI (C, blue), an interstitial probe targeting just CS3 (D, grey), HOP-M (E, green, Atto488N), HOP-P (F, magenta, Atto565N), and all probes (G), the latter demonstrating co-localization of all signals ( $n = 128$ ); percentages show efficiency of each probe configuration, with a combined efficiency of 96.5%. (H) Ellipticity for the maternal (green) and paternal (magenta) homologs of the 6 nuclei for which HOPs had been applied.

<https://doi.org/10.1371/journal.pgen.1007872.g005>

proved successful in differentially labeling homologs (Fig 5C–5G, S11B and S11C Fig; also, S3A and S3B Fig) and, hence, enabling parental origin to be assigned for 12 walks (6 nuclei). These studies demonstrated that ellipticity was not absolutely correlated with parental origin (Fig 5H), leaving open, however, the possibility of incomplete skewing. Intriguingly, absence of parental influence would imply that the distinction of homologs is established post-fertilization, perhaps as often as at each cell cycle, lending support to proposals that homologous chromosomal regions are not only molecularly aware of each other, but that that awareness can lead to inter-homolog communication and impact. Interestingly, analyses of even the small number of homologs in our dataset suggested differences in the 3D organization of maternal and paternal homologs (S11D–S11I Fig); of note, the combined set of maternal and paternal homologs used in our analyses produced matrices (S11F and S11I Fig) reminiscent of those of the entire set of 38 chromosomes (Fig 3A and 3B), arguing that the homologs we distinguished by HOPs approximate those of the larger population. While these studies must be greatly expanded to substantiate any homolog-specific trend, they nevertheless indicate that our approach should be useful for exploring the biology of homology.

In conclusion, we have described strategies for *in situ* super-resolution genome visualization through sequential rounds of imaging, followed by integration of the resulting images with Hi-C data to produce 3D models at 10 kb resolution in single nuclei. In conjunction with other efforts (e.g., refs. [11, 13–21], bioRxiv [22]), these strategies should advance our

understanding of genome organization and, when augmented with multichromosome walks and parent-of-origin information, may contribute to whole-genome clarity on the relationship between 3D genome configuration and genome function. Indeed, we have already observed patterns of chromosome organization that were otherwise unknown. In line with other studies, including a live-cell analysis correlating chromatin accessibility more with fluctuation than compaction (bioRxiv [72]), we have also observed significant structural variability and speculate, here, as to the source and/or potential of that variability. While such variability may indicate nonhomogeneity of the imaged samples with respect to, for example, the cell cycle, it may also suggest that some features of genome organization function only to enable the genome to be bundled into a small space and that, as long as this can be achieved cell cycle after cell cycle, the specific mechanism of bundling is unimportant, potentially even random. Alternatively, this variability may be intimately related to genomic function. In fact, variability, itself, may be a signature that distinguishes one genomic region, or chromatin state, from another. For instance, while some regulatory factors may recognize their binding regions by the conformation of their targets, it may be that others also, or solely, recognize their targets by the conformational variation of their targets, the number of conformations, and/or the speed with which their targets transition through conformations. If such dynamic variability can be transmitted along chromosomes or be of sufficient magnitude to generate a turbulence that is propagated through the nuclear milieu, it may function to convey information across nuclear distances, perhaps even perturb or promote phase separations. Finally, we note that, as our strategies should be applicable to entire chromosomes, it should be possible to implement them in whole genome studies of entire chromosomes as single, fully integrated units of structure and function, consistent with the unit of inheritance being as much the chromosome as it is the gene.

## Materials and methods

### Mining for Oligopaint targets

Akin to a growing category of oligo-based FISH probes (e.g., refs. [73–77]), Oligopaint FISH probes consist of oligos that have been computationally designed to target specific sequences of the genome [23] (see the Oligopaints website for preselected whole genome datasets of Oligopaint oligo targets for the human as well as mouse, zebrafish, Arabidopsis, Drosophila, and *C. elegans* genomes). We identified genomic sequences that would serve as good targets for Oligopaint oligos by applying Oligominer to a repeat-masked human hg19 assembly [78]. Specifically, we used the ‘balance’ settings of 35–41 nts of genome homology, 42–47°C for  $T_M$ , a simulated hybridization temperature of 42°C, and ‘LDA mode’ filtering. Candidate probes were further processed by the ‘kmerFilter’ script that calls the algorithm Jellyfish [79] to eliminate probes containing 18mers that occur in >5 times in hg19.

### Mining the genome for homolog-specific Oligopaint (HOP) targets on chromosome 19 of PGP1

As the genome of PGP1 has been fully sequenced and phased [25, 80] (Kun Zhang and Daniel Jacobsen, unpublished; Bing Ren and Anthony Schmitt, unpublished), we were also able to design HOP probes for the homologs of chromosome 19. First, genomic sequences that would be good targets for HOPs [14] were discovered using the approach described above, except that the input hg19 sequence was not repeat-masked. These probe sequences were then intersected with the locations of phased heterozygous SNVs for PGP1 according to Lo, *et al.* (2013, ref. [80]) using the ‘intersectBed’ utility from BEDTools [81] and then processed with in-house

software to generate two sets of haplotype-specific HOP probes in which each probe oligo targets  $\geq 1$  heterozygous SNV.

In order to ascertain which of the two HOP probes corresponds to maternal variants and which corresponds to paternal variants, we took advantage of the full genome sequence of PGP95 (courtesy of Rigel Chan and Elaine Lim), with whom PGP1 shares his maternal parent but not his paternal parent, and presumed that the homolog of chromosome 19 that carries one (or more) long blocks of variants in common with PGP95 would be the one that was maternally derived (Fig 1B). First, heterozygous variants that identify the two PGP1 haplotypes were matched to variants from the full genome sequence of PGP95, and any that were homozygous in the PGP95 genome were discarded. All remaining PGP95 variants were then assigned as matching one of the PGP1 haplotypes (arbitrarily designated as H0 or H1) or neither, permitting us to identify blocks of contiguous matches to H0 or H1. The H1 haplotype included significantly more of the longest blocks shared between PGP1 and PGP95, thus identifying the H1 haplotype as maternally derived (S12 Fig). Then, as many long blocks corresponding to H1 were discovered to occupy a segment of PGP1 chromosome 19 at coordinates chr19:48,932,903–59,087,560, we designated the chromosome 19 homolog that carried these blocks the maternal chromosome. Accordingly, we designated the HOP probe corresponding to H1 as HOP-Maternal (HOP-M) and that corresponding to H0 as HOP-Paternal (HOP-P). The probes are efficient. Up to 96.5% of nuclei producing two fluorescent foci when imaged with HOP-M or HOP-P showed one of the signals to be stronger than the other, with the focus labeled more strongly with one probe being the more weakly labeled focus when imaged with the other probe; dye-swap experiments suggested that the difference between ratios for HOP-M and HOP-P is due to dye chemistries (S11B and S11C Fig). Importantly, the corresponding ratio of signals obtained with probes targeting only the interstitial regions lying between SNVs approached 1.

### Multiplexing of the Oligopaint library and design of streets

We designed our Oligopaint libraries based on Hi-C maps that were available at the time our study was initiated. In particular, Geoffrey Fudenberg used IMR-90 Hi-C contact frequencies from Rao, *et al.*, (2014, ref. [8]) to identify compartmental intervals and contact domains in the 8.16 Mb region extending from chr19:7,400,000 (19p13.2) to chr19:15,560,000 (19p13.12), where the calling of compartmental intervals relied on ICE [82] and a 3-state Gaussian Hidden Markov Model (HMM) implemented in scikit-learn [83], while contact domain calling relied on [84]. Having segmented the 8.16 Mb region, we then incorporated Mainstreets and Backstreets into our library such that they would permit individual genomic features to be separately imaged. In particular, streets were generated using an in-house MATLAB (MathWorks, Natick, MA) script, 'MakingStreets', available via GitHub (<https://github.com/gnir/OligoLego>). Street sequences were vetted for predicted performance when serving as primers in PCR reactions using Primer3Plus [85] with default settings, except that the melting temperature was set to be 57–59°C, and the GC content was set to be 40–60% [85]. Street sequences were also checked for pairwise orthogonality using NUPACK [86] to estimate equilibrium hybridization yields between candidate streets and other potential target sites in 390 mM Na<sup>+</sup> and 50% formamide at room temperature (RT). Toehold sequences were also validated using NUPACK. Candidate streets passing all upstream checks were then examined for orthogonality to the human genome by using bowtie2 [87] to filter sequences aligning at least once to hg38 with '—very-sensitive-local' alignment mode. Mainstreet and Backstreet sequences were appended to Oligopaint oligos using an in-house MATLAB code available via GitHub (<https://github.com/gnir/OligoLego>).

## Multiplexing Oligopaint libraries to walk along multiple chromosomes simultaneously

Here, we designed Mainstreet barcodes so that a) they could be shared across multiple chromosomes, such that one round of imaging could target anywhere from one up to the maximum number of chromosomes to be imaged, and b) imaging different chromosomes would be initiated in different rounds, thus permitting each chromosome to be identified by the round in which it first appears. For example, in an imaging scheme that introduces a new chromosome in every round, one could use the following scheme, wherein the walk along Chromosomes 1, 2, 3, 4, 5, and N are initiated in Rounds 1, 2, 3, 4, 5, and N, respectively.

### Oligopaint probe synthesis

Oligopaint libraries were purchased from CustomArray (Bothell, WA) and amplified using *in vitro* transcription as previously described [24]. Briefly, for each sub-pool to be amplified, we first optimized the template and primer concentrations using real-time PCR. We next performed large-scale limited-cycle PCR, and then used the product as template for *in vitro* transcription by T7 RNA polymerase (NEB, E2040S). The resulting ssRNA was then reverse transcribed (EP0752, ThermoFisher Scientific). RNA was digested by adding a mixture of 0.5 M NaOH and 0.25 M EDTA of equal volume to the reverse transcription reaction at 95° C for 10 minutes. Oligopaint ssDNA oligo probes were purified using the DNA Clean & Concentrator-100 kit (Zymo Research, Irvine, CA) paired with an EZ-Vac Vacuum Manifold (Zymo Research).

### Culturing of cells

Human primary skin fibroblasts of donor PGP1 (GM23248, Coriell Institute; ref. [88]) were grown at 37° C + 5% CO<sub>2</sub> in serum-supplemented (10% v/v) Dulbecco's Modified Eagle Medium (DMEM) (serum Gibco 10437; media Gibco 10564). The cells were also supplemented with 1% (v/v) MEM Non-Essential Amino Acids Solution (Gibco 11140050). Penicillin and streptomycin (Gibco 15070) were also added to the cell culture media to final concentrations of 50 units/ml and 50 µg/ml, respectively.

|              | Round 1 | Round 2 | Round 3 | Round 4 | Round 5 | Round N |
|--------------|---------|---------|---------|---------|---------|---------|
| Chromosome 1 | ✓       | ✓       | ✓       | ✓       | ✓       | ✓       |
| Chromosome 2 |         | ✓       | ✓       | ✓       | ✓       | ✓       |
| Chromosome 3 |         |         | ✓       | ✓       | ✓       | ✓       |
| Chromosome 4 |         |         |         | ✓       | ✓       | ✓       |
| Chromosome 5 |         |         |         |         | ✓       | ✓       |
| Chromosome N |         |         |         |         |         | ✓       |

<https://doi.org/10.1371/journal.pgen.1007872.t001>

### Sample preparation

40 mm coverslips #1.5 (Bioprotechs, Butler, PA) were cleaned with a bath sonicator and stored in a 75% (v/v) ethanol solution at room temperature until used. In order to prepare PGP1f samples for imaging, the pre-cleaned coverslips were first allowed to fully dry in a tissue culture hood and then placed in sterile 150 mm tissue culture dishes (Falcon 353025). PGP1f cells were then deposited at ~15% confluency onto the 150 mm dishes and allowed to grow in a mammalian tissue culture incubator until ~85% confluent (~5 days). At this point, the PGP1f



cells were rinsed with 1x PBS, fixed using 4% (w/v) paraformaldehyde in 1x PBS for 10 minutes, rinsed with 1x PBS, permeabilized with 0.5% (v/v) Triton X-100 in 1x PBS for 10 minutes, rinsed with 1x PBS, and then stored in a cold room for up to 3 weeks.

### 3D DNA FISH

3D DNA FISH [89] was performed on fixed and permeabilized coverslips essentially as previously described [14, 23, 90]. Briefly, coverglass samples were placed in small glass-bottom tissue culture dishes (MatTek P50G-1.5-30-F) and incubated with PBST (1x PBS + 0.1% v/v Tween-20) for 2 minutes. Samples were then incubated once with 1N HCL for 5 minutes, twice with 2x SSCT (2x SSC + 0.1% v/v Tween-20) for 1 minute, once with 2x SSCT + 50% (v/v) formamide for 2 minutes, and then once for 20 minutes with a pre-heated solution of 2x SSCT + 50% formamide at 60°C. Coverslips were then air-dried and incubated with a hybridization cocktail consisting of 2x SSCT, 50% formamide, 10% (w/v) dextran sulfate, 0.4 µg/µL of RNase A (ThermoFisher EN0531), and Oligopaint probe whose concentration was adjusted such that the final amount of probe added was equivalent to ~1.4 µM per every 1 Mb targeted (e.g. ~2.8 µM for probe targeting 2 Mb). In cases where the parental origin of the homolog was to be determined, HOPs probes were added at a concentration of 2.85 µM each. The hybridization reaction was then sealed beneath a 22 x 30 mm, #1.5 coverslip (VWR, Randor, PA) using rubber cement (Elmer's, Westerville, OH), which was allowed to dry at 37°C for 7 minutes. Samples were then denatured on top of a heat block at 80°C (with the heat block standing in water in a water bath set at 80°C) for 3 minutes. Coverslips were then allowed to hybridize for 2–3 nights at 47°C in a humidified chamber. Following hybridization, coverslips were washed 4 times with 2x SSCT at 60°C for 5 minutes, then twice with 2x SSCT at RT. Coverslips were then rinsed with 1x PBS and allowed to air dry. For the purposes of drift correction, fiducial markers (gold nano-urchins, GNUs; d = 90 nm, 630 nm abs max) (Cytodiagnosics, Ontario, Canada) were sonicated for 10 minutes and then diluted 1:3 in PBS, pipetted onto the sample coverslips, and sandwiched with a blank, i.e. non-treated, 40 mm coverslip. GNU-coverslip sandwiches were centrifuged at 500 x g for 3 minutes. The sample coverslips were then washed for 2 minutes with 1x PBS.

### HOPs imaging and processing

Images were obtained using the widefield mode on the Bruker Vutara 352 (Bruker Nano Inc.) microscope (described in the section in [Methods](#) on OligoSTORM imaging). *SlowFade* anti-fade reagent containing DAPI dye (ThermoFisher Scientific) was used as the imaging buffer. Images were acquired in the Z-dimension, approximating 2–3 µm across 4 different channels, with the following dyes: DAPI, Alexa Fluor 647, Alexa Fluor 488, and ATTO 561 with 405, 488, 561, and 639 nm lasers (Coherent), the emission filters (custom-made, Semrock) 465/30, 515/40, 600/50, respectively, and detected via an sCMOS camera (Hamamatsu Orca-flash 4.0 v3) with 100 ms exposure time. These four channels corresponded to nuclear staining, interstitial signals, and the two HOP probes, respectively. Images were analyzed using an in-house MATLAB (Mathworks) script, which was written to detect and then derive the intensity ratio of the foci in each channel. Detection of signal from foci was validated by eye.

### OligoSTORM imaging

OligoSTORM imaging was performed on a Bruker Vutara 352 commercial 3D biplane single-molecule localization microscope [91–93], equipped with a 60x water objective (Olympus) with a numerical aperture (NA) of 1.2. [Fig 2C and 2D](#) and [S3B–S3E Fig](#) were obtained using a 60X silicone objective (Olympus) with an NA of 1.3 which, akin to the water objective, does

not produce the degree of spherical aberrations common to oil objectives. For illuminating Oligopaint oligos with Alexa Fluor 647, we used a 639-nm Coherent Genesis laser with  $\sim 25$  kW/cm<sup>2</sup> at the back aperture of the objective, and for illuminating with Alexa Fluor 405, we used a 405-nm Coherent Obis laser for photo-activation with up to  $\sim 0.05$  kW/cm<sup>2</sup> at the back aperture (emission filters were described in the [Methods](#) section describing HOPs imaging and processing). Fluorescent detection was captured on an sCMOS camera (Hamamatsu Orca-flash 4.0 v2) using a 10 ms exposure time. At each imaging session, we imaged 5–10 nuclei at different X,Y stage locations, while taking up to 4  $\mu$ m z-scan using 100 nm Z-steps at each location, with up to 21 loci per nucleus, collecting 42,000–150,000 frames for each locus. Image data was collected by a Bruker-provided network attached storage (NAS) system.

### OligoDNA-PAINT

DNA-PAINT imaging was performed on a custom optical set-up based on a Nikon Ti Eclipse microscope featuring a Perfect Focus System and a custom-built TIRF illuminator. DNA-PAINT was performed using Highly Inclined and Laminated Optical sheet illumination (HILO; ref. [94]) with 10% of a 1000-mW, 532-nm laser (MPB Communications) using a CFI Apo TIRF 100 $\times$  oil (N.A. 1.49) objective at an effective power density of  $\sim 2$  kW/cm<sup>2</sup>. The 532-nm laser excitation light was passed through a quarter-wave plate (Thorlabs, WPQ05M-532), placed at 45° to the polarization axis and directed to the objective through an excitation filter (Chroma ZET532/10x) via a long-pass dichroic mirror (Chroma ZT532RDC\_UF2). Emission light was spectrally filtered (Chroma ET542LP and Chroma ET550 LP), directed into a 4f adaptive optics system containing a deformable mirror (MicAO 3DSR, Imagine Optic) and imaged on an sCMOS camera (Andor Zyla 4.2+) with 6.5- $\mu$ m pixels using rolling shutter readout at a bandwidth of 200 MHz at 16 bit, resulting in an effective pixel size of 65 nm. In order to estimate the axial positions of single molecule emitters, optical astigmatism was applied using the deformable mirror to cause an asymmetric distortion in the observed point spread functions [95, 96]. A total of 22,500–45,000 250-ms frames were acquired for each image by using 0.5–1 nM of Cy3B-labeled 10-mer oligos in 1 $\times$  PBS solution + 500 mM NaCl, 10 nM PCD, 2.5 mM PCA, and 2 mM Trolox. Gold nanoparticles (40 nm; no. 753637; Sigma-Aldrich) were used as fiducial markers to facilitate drift correction. The lateral positions of single-molecule localization events were first determined using Picasso [97] by applying the least-square fitting algorithm, and the axial positions were then also determined using Picasso by fitting to a pre-established calibration curve [95, 96]. Global lateral drift-independent localization precision was estimated by Nearest Neighbor based Analysis (NeNA) [98]. Global axial localization precision was estimated by empirical analysis of sub-diffraction sized fiducial markers after fitting to an empirically derived calibration curve with a mean precision of < 20 nm.

### OligoDNA-PAINT image reconstruction

Two-dimensional DNA-PAINT images were rendered using Picasso, with individual localization events being represented as single point sources blurred by an isotropic two-dimensional Gaussian function whose “sigma” parameter reflects the average X-Y NeNA localization precision of the localizations occurring in the presented field of view and colored according to inferred axial position [98]. Multicolor three-dimensional DNA-PAINT images were rendered using ViSP [99], with individual localization events being represented as single point sources blurred by an anisotropic three-dimensional Gaussian function whose X-Y “sigma” reflects the average NeNA localization precision of the localizations occurring in the presented field of

view and Z “sigma” was informed by the precision of the calibration curve used for axial fitting as determined by Picasso.

### Automated microfluidics

We exploited a fluidics system as previously described [11, 27]. Briefly, Biotech’s FCS2 flow chamber ( $v \sim 100 \mu\text{L}$ ) was integrated with a peristaltic pump (Gilson minipuls3) and 3 valves (Hamilton HVXM 8–5), with the resulting dead volume in our system being  $\sim 700 \mu\text{L}$ . Integration of the fluidics system and the Bruker Vutara 352 microscope was programmed into SRX, the commercially available microscopy software package that controls the Bruker Vutara 352.

### Sequential hybridization for multiple rounds of OligoSTORM

Secondary hybridizations were performed using  $1 \mu\text{M}$  of secondary probe in  $2\times$  SSC +30% (v/v) formamide, and allowed to hybridize for 30 minutes at RT. Following hybridization, a wash solution (40% v/v formamide +  $2\times$  SSC) was introduced for 12 minutes to the flow cell and incubated for 3 minutes without flow. Wash solution was then replaced with imaging buffer consisting of 10% (w/v) glucose,  $2\times$  SSC, 50 mM Tris, 1% (v/v)  $\beta$ -mercaptoethanol, and 2% (v/v) of a GLOX stock solution consisting of 75 mg/mL glucose oxidase (Sigma G2133-250KU), 7.5 mg/mL catalase (Sigma C40-500MG), 30 mM Tris, and 30 mM NaCl. A thin layer of mineral oil (Sigma M5904) was added at the top to prevent oxygen from penetrating the imaging buffer. Imaging was initiated after the imaging buffer was allowed to incubate for 2 minutes without flow.

### OligoSTORM data processing

During localization analysis, the precision of each localization event was determined by calculating the Cramér-Rao Lower Bound, namely the inverse of the Fisher information of the measured point-spread function [100, 101]. The Cramér-Rao Lower Bound is the lower bound of the variance of the estimation process, which is used to calculate the localization precision of each event.

For isolating localizations comprising structures of interest during each round of imaging (fluidic cycle), we first removed localizations of lesser quality according to the following criteria: initial filtering of the localization data consisted of applying a geometric mean goodness-of-fit filter calculated from three separate metrics from the localization data. The individual metrics consisted of calculating the ratio of the photons assigned to the Point Spread Function (PSF) in the localization routine to the total number of photons in the camera cutout around a localization event (first metric), total photons in the cutout around a localization event (second metric), and the offset in Z between the calculated localization position of the fluorophore and the physical position of the objective piezo (third metric). The goodness-of-fit metric of each localization event is calculated by sorting each metric through a self-normalization process (with a value of 0 and 1 being the lowest and highest, respectively, quality value to a given metric), multiplying the values of the three metrics together, and taking the cube-root. Localizations with a score closer to 1 are considered of higher quality as compared to localizations with a score closer to 0. An initial thresholding was set such that only localizations with a rank of 0.8 or higher were accepted for statistical analysis. Also, candidate localization events whose peak intensity varied no more than 2 pixels (pixel size  $\sim 100 \text{ nm}$ ) in up to 7 frames in a row were considered as one. Finally, we filtered out all localizations with an axial precision worse than 100 nm. We found that the number of localizations increases with genomic size (S13 Fig). Following localization filtering, we performed drift correction using a center-of-mass function

between subsets of fiducial markers across the data recording [102]. Finally, clusters of localizations were segmented using DBSCAN [39].

When determining the spatial overlap between two clusters, we used Bruker's SRX software to generate surfaces of the calculated clusters based upon alpha shapes, using an alpha radius of 150 nm, which reflects the maximum particle distance used for the DBSCAN analysis for CS1, CS7, and the merged four subregions that were imaged in rounds 10–13 (Fig 2A and 2H). Average spatial overlap between two clusters was expressed as the average of the fraction of the first cluster overlapping the second and the fraction of the second cluster overlapping the first (Fig 2G and 2H, S4 Fig). When comparing two clusters representing the same genomic region, the difference in number of localizations between the two clusters was assessed using a beeswarm plot ([103]; S4 Fig).

For comparison, we determined the spatial overlap of a cluster that was randomly divided into two clusters. In particular, we chose two nuclei as an example and divided the localizations obtained in rounds 1 and 21 of imaging into two, generating 8 clusters per diploid nucleus. We noticed that for the nucleus with the higher number of localizations (~6,500 localizations/cluster) the average spatial overlap was 90–92%, while for the nucleus with lower number of localizations (~3,380 localizations/cluster) the average spatial overlap was 85–89%. This implies that the spatial overlap between rounds is not expected to be more than 90% and is dependent upon the number of localizations.

### Selection of nuclei for analysis

For OligoSTORM analysis, we chose only diploid nuclei for which the homologs were not too close to be distinguished with a single dye (as we used only Alexa647N for OligoSTORM), did not show high levels of anisotropy in Z, and also did not show evidence of aneuploidy. Note, PGP1f cells were confirmed to be karyotypically normal (XY; n = 20; Cell Line Genetics, Madison, WI), and cell sorting experiments show that 72–86% of PGP1f cells are in G1.

### Generating OligoSTORM density maps

DBSCAN-extracted localizations for each chromosomal segment were convolved with a Gaussian kernel to match the mean precision [43] and to reduce the anisotropy of the image acquisition. The density map of each chromosomal segment was represented by intensities at points  $i$  ( $\rho_i$ ) on a cubic grid with fixed voxel size. For each localization, the density value of the nearest voxel was increased by a factor of 1. The obtained grid was convoluted with a Gaussian function using TEMPy [43], such that the density  $\rho_i$  was defined as:

$$\rho_i = \sum_N \frac{Z}{(\sigma\sqrt{2\pi})^3} e^{-\frac{(x-x_n)^2+(y-y_n)^2+(z-z_n)^2}{2\sigma^2}}$$

where  $x$ ,  $y$  and  $z$ , and  $x_n$ ,  $y_n$ ,  $z_n$ , are the Cartesian coordinates of grid point  $i$ , and localization  $n$  respectively.  $N$  is the total number of localizations,  $Z$  is the number of localizations per grid point, and  $\sigma$  is set to be proportional to the mean precision. The iso-contour threshold for each map was set to be one sigma after shifting of the background peak to zero using TEMPy [43]. The density map of the entire walk was also generated using the same approach.

### Analysis of OligoSTORM density maps

A total of five structural measures were obtained from each OligoSTORM density map obtained for each chromosomal segment. First, the distance score (DS) between two chromosomal segments was calculated as the spatial distance between the centers of mass of each

chromosomal segment of the OligoSTORM density maps. Next, the DS was corrected for linear genomic distance and expressed relative to the expected distance, as determined via a power-law fit, by subtracting the expected distance from the measured distance. This DS Z-score was calculated for each data point using the power-law function as expected value and is expressed as the value, only, without unit of measure, so as to avoid implications of actual distance. Second, the entanglement score (ES) between two chromosomal segments was calculated as the fraction of overlapping voxels within the optimal iso-contour threshold with respect to the smaller of the two density maps as implemented in TEMPy [43, 104]. As before, the ES was corrected for linear genomic distance using a power-law fit obtained by the ES Z-score. Third, the surface area of each OligoSTORM density map was calculated by summing over the area of the triangles on the surface on the convex hull using the “measure area” option in Chimera [105]. Next, the surface area was corrected for genomic size (S6A Fig) and transformed into a surface area Z-score. Fourth, volume was calculated within the optimal iso-contour threshold using the “measure volume” option in Chimera [105]. As before, the volume was corrected for genomic size (S6B Fig) and transformed into a volume Z-score. Fifth, a sphericity score ( $\psi$ ), which measures how closely the shape of an OligoSTORM density map approaches that of a mathematically perfect sphere, was calculated as:

$$\Psi = \frac{\pi^{\frac{1}{3}}(6V_p)^{\frac{2}{3}}}{A_p}$$

where  $V_p$  and  $A_p$  are the volume and surface area of the OligoSTORM density Map  $P$ . To take into account the bias due to genomic size, sphericity was corrected for genomic size (S6C Fig) and transformed into a sphericity Z-score.

For each homolog, a feature vector was created, which is a binary 1x9 vector encoding the cluster types of each chromosomal segment in the region. Next, a per-nucleus profile of the compartment state was made for comparing the feature vectors of each of the homologs in a single nucleus. The compartment state profile is represented by a 1x9 vector whose values are determined by a comparison score set as: (i) 1, both chromosomal segments belong to cluster 1, (ii) 0, the chromosomal segments belong to different clusters, or (iii) -1, both chromosomal segments belong to cluster 2. The resulting matrix of 19 compartment state profiles was hierarchically clustered, resulting in five clusters.

The resulting matrix of structural features for the 342 chromosomal segments (that is, 9 segments in 2 homologs in 19 cells) was analyzed using the PCA implemented in the Python library Scikit-learn. The clustering resulted best in two major clusters that were named cluster 1 and cluster 2 (Fig 3C, S7A Fig).

## ENCODE PGP1 data

Data available for RNA-seq, DNase-seq and five chromatin marks (H3K4me1, H3K27ac, H3K4me3, H3K27me3, H3K9me3) were downloaded from the ENCODE project [62]. Mapped reads were obtained from samples ENCFF043JNP and ENCFF280JPQ for H3K4me1, ENCFF612WGP and ENCFF345SWP for H3K27ac, ENCFF397XVM and ENCFF179TEP for H3K4me3, ENCFF378EQH for RNA-seq, ENCFF523JNF and ENCFF050OVJ for DNase-seq, ENCFF800CCA and ENCFF876WHO for H3K27me3, and ENCFF644QGC for H3K9me3 (<https://www.encodeproject.org/>). Next, coverage of the chr19:7 (19p13.2) to chr19:15.5 (19p13.12) visualized region was calculated using the SamTools bedcov command [106] at 10 kb resolution. All datasets, with the exception of RNA-seq and H3K9me3 ChIP-seq, had two replicates which were highly correlated ( $r > 0.99$ ) and, in such cases, the average coverage per each bin was used. The final number of mapped reads

for each dataset was 60,550,146 for H3K4m1, 52,286,921 for H3K27ac, 62,084,095 for H3K4m3, 200,776,708 for RNA-seq, 622,389,260 for DNase, and 65,357,125 for H3K27me3, and 26,947,714 for H3K9me3.

### Homolog ellipticity

The Cartesian coordinates of the location of each homolog in the cell population were fitted to an ellipsoid [107], and the ellipticity score for any 8.16 Mb region was then calculated to be the ratio between the two largest principal axes of the ellipsoid. When comparing homologous chromosomal regions in a single nucleus, we determined the ratio between the more elliptical homolog and the less elliptical one, calling that measure the ellipticity ratio. An ellipticity ratio equal to 1 indicates that the two homologs of a single nucleus had the same ellipticity score. For the evaluation of significance, a random ensemble of elongation ratios was constructed by calculating the ratio between two homologs randomly selected from the analyzed cell population.

### Modeling of genomes

**Integrative 3D modeling with TADbit.** The chromosome 19 normalized Hi-C interaction matrices of each chromosomal segment and the entire region (chr19: 7400000:15560000) were used for modeling at a resolution of 10 kb as previously described [58, 59]. The selected resolution was the highest possible with an MMP score predicted to result in reliable 3D models (ref. [108] and S7 Table). For each chromosomal segment, two independent ensembles of models were generated: (i) Isolated chromosomal segments and (ii) the chromosomal segments extracted from models of the entire region. Briefly, TADbit automatically generates 3D models [59] using a restraint-based modeling approach, where the experimental frequencies of interaction are transformed into a set of spatial restraints [58]. For each chromosomal segment, a total of 5,000 models were generated, with only the best 1,000 models (that is, those that best satisfy the input restraints) used for further analysis. For each of the modeled chromosomal segments, the contact map obtained from the final 1,000 models resulted in a good correlation with the input Hi-C interaction matrix (S7 Table), which is indicative of good model accuracy [108].

**Rigid fitting.** The resulting 3D models for each chromosomal segment ensemble were rigidly fitted in the corresponding OligoSTORM density maps using the ‘Fit\_in\_map’ global rigid-body fitting protocol in Chimera [105]. A total of 100 random placements of each chromosomal segment were generated by randomly rotating and translating the fit within the reference OligoSTORM density map, followed by a step of local optimization using steepest ascent. Next, two scores were computed for each chromosomal segment fit: (i) the cross-correlation coefficient (CCC) between the OligoSTORM density map and the chromosomal segment fit and (ii) the connectivity score (ConS) of the chromosomal segment fit with that of its sequentially neighboring chromosomal segment. The CCC score is expressed by the following formula [43, 109]:

$$CCC = \frac{\sum_{i \in M} (\rho_i^P - \bar{\rho}^P)(\rho_i^T - \bar{\rho}^T)}{\sqrt{[\sum_{i \in M} (\rho_i^P - \bar{\rho}^P)^2 \sum_{i \in M} (\rho_i^T - \bar{\rho}^T)^2]}}$$

where  $M$  is all the voxels in the density grid of the map;  $P$  and  $T$  are the chromosomal segment fit map and OligoSTORM density map, respectively;  $\rho_i^P$  and  $\rho_i^T$  are the intensity of the density maps at voxel  $i$ ; and  $\bar{\rho}^P$  and  $\bar{\rho}^T$  are the respective mean intensity values. The CCC ranges

from 0 (no match) to 1 (perfect match). The ConS, which ranges from 0 to 1 was defined as:

$$\text{ConS} = 1 - \frac{d_{P,COM}}{\max(d_{P,COM})}$$

where  $d_{P,COM}$  is the distance between the starting or end bin of the fit with the center of mass of the surface interface with its neighboring chromosomal segment, and  $\max(d_{P,COM})$  is the maximal  $d_{P,COM}$  in the ensemble of fits. The surface interface is defined in a multistep process: (i) identification of the surface area of the chromosomal segment in analysis that lies within a cut-off distance of 50 nm from its neighbor chromosomal segment and calculated by the “measure contactArea” option in Chimera [105], and (ii) selection of the density points of the chromosomal segment density map that are bounded by the identified surface area by using the “mask” command in Chimera [105]. Finally, a combined score (S14 Fig) integrating both measures was used for calculating the goodness of fit:

$$\text{CombScore} = \text{ConS} + 2 * \text{CCC}$$

For each chromosomal segment, the top-scoring fits (95<sup>th</sup> percentile) were chosen for further refinement.

**Flexible refinement.** To further improve the agreement between the fit and the OligoSTORM density maps, a flexible fitting method was implemented similarly to the protocols for conformational refinement of protein guided by cryoEM density [53]. The protocol is a heuristic optimization that relies on simulated annealing molecular dynamics applied to a series of subdivisions of the fit into progressively smaller rigid bodies ( $b_i$ ). A rigid bod  $b_i$  could be any set of particles in the fit (including loops, a sub-domain, a contact domain, chromosomal segment or larger region). A set of rigid bodies ( $B$ ) was defined for each fit using TADbit’s TAD caller and selecting borders with strength higher than 5. Finally, the exact rigid bodies were manually adjusted by visual inspection of the rigid fitted 3D starting model (S8 Table). During the refinement, the coordinates of each  $b_i \in B$  were displaced in the direction that maximized their cross-correlation with the OligoSTORM density map and avoided the violation of the imposed restraints from the population Hi-C interaction matrix. The final scoring function was defined as:

$$E = W_1 E^{CCC}(P) + W_2 E^{H-LB}(P) + W_2 E^{H-UB}(P) + W_3 E^{EV}(P) + W_2 E^C(P)$$

where  $E^{CCC}(P)$  quantified the fit of the probe map (map generated from the 3D fitted model) ( $\rho^P$ ) and the target OligoSTORM density map ( $\rho^T$ ), and is defined as the negative sum of cross-correlation coefficients between the OligoSTORM density map and the rigid bodies  $b_i$ ;  $E^{H-LB}(P)$  quantified the lower-bound harmonic restraints which prevents two particles from getting closer than a given equilibrium distance;  $E^{H-UB}(P)$  quantified the upper-bound harmonic restraints, which forces two particles to be closer than a given equilibrium distance;  $E^{EV}(P)$  is an hard sphere excluded volume term (lower bound harmonic oscillator dropping at zero at  $2r$ ); and  $E^C(P)$  is a harmonic bond that connects adjacent particles at a given equilibrium distance. The weights  $W_1$ ,  $W_2$ , and  $W_3$  determine the relative importance of the corresponding terms and were set as default, which proved optimal for fitting of sub-tomogram averaging [110]. The optimization protocol to identify the best fit consisted of simulated annealing rigid-body molecular dynamics followed by a final a conjugate-gradient minimization. A maximum of 50 cycles of simulated annealing molecular dynamics steps were performed, in each cycle the system was heated from 0K to 1000K for a maximum of 100 steps and cooled back to 0K for a maximum of 200 steps. At each step, the optimization was terminated if the change in CCC was  $< 0.001$ . Finally, a minimization step was performed comprising 200 conjugate-gradients steps with weights  $W_1$ ,  $W_2$  and  $W_3$  set to 1 and 200 conjugate-gradients steps with  $W_1$  set to 0 and  $W_2$  and  $W_3$  set to 1.

The goodness-of-fit was assessed using the CCC score and the clash score (CLS), although other goodness-of-fit measurements can also be used [43, 111]. The latter defined as the number of serious clashes (that is, distance  $d < 2r$ ) per 1000 beads. The refined model was chosen as the one that maximizes the CCC and minimizes the clash score (S8 Table).

### PGP1f Hi-C

PGP1f cells were cultured as described in the ‘Cell Culture’ section of Methods and pelleted at low speed centrifugation (185 x g) to maintain cellular and nuclear integrity. We generated 2 *in situ* Hi-C libraries with three million cells each, utilizing the MboI restriction enzyme (NEB, R0147M), following the protocol described in [8]. Briefly, the Hi-C protocol entails crosslinking cells with 1% formaldehyde (wt/vol), cell permeabilization with nuclei intact, DNA digestion with an appropriate 4-cutter restriction enzyme, and 5' -overhang filling with the incorporation of a biotinylated nucleotide. The resulting blunt end fragments are ligated, and the DNA is then sheared. Biotinylated ligation junctions are captured with streptavidin beads, and the resulting fragments are analyzed with paired-end sequencing.

### PGP1f Hi-C data processing

Ten Hi-C libraries were sequenced with 150 bp paired-end sequencing reads on an Illumina HiSeqX-10. The resulting sequencing data from the ten libraries (S9 Table) was processed separately using Juicer [112] and mapped to the GRCh37/hg19 (hg19) human genome assembly [8, 112]. To generate statistics and a Hi-C file for a combined map including all ten replicates (S5 Table; S15 Fig; Supporting information S2 Text), the mega.sh script was used (<https://github.com/theaidenlab/juicer/wiki/Usage>) [112], with the following command: `/gpfs0/juicer/scripts/mega.sh`.

The combined Hi-C map generated a total of 2,599,264,644 reads, achieving a resolution of 5 kb, in alignment with Hi-C map resolution parameters described in ref. [8]. This combined map was used with ‘KR normalization’ (ref. [112, 113]) for all subsequent analyses.

### Assignment of manually identified compartments as ‘A’ or ‘B’

To label the two manually identified compartments as ‘A’ or ‘B’, we began by partitioning chromosome 19 into 100kb loci. These loci were assigned to two compartments based on the intrachromosomal contact matrix for chromosome 19 using the eigenvector method (S8 Fig) as implemented in Juicer [3, 112]. Specifically, we used the following command: `java -jar /gpfs0/juicer/scripts/juicer_tools.jar eigenvector KR inter.hic 19 -p BP 100000`.

Visual inspection of the Hi-C map for chromosome 19 confirmed that the eigenvector corresponded reliably to the plaid pattern in the data [3] (S8 Fig and interactive S8 Fig <http://bit.ly/2zCG2X6>, for interactive figure documentation see: <https://igvteam.github.io/juicebox.js/>).

Next, we labeled the two compartments as ‘A’ and ‘B’ on the basis of GC content, which is known to be enriched in the ‘A’ compartment and in open chromatin more generally [52]. The GC content track for human hg19 chr19 was downloaded from the UCSC Genome Browser (<http://hgdownload.cse.ucsc.edu/goldenpath/hg19/gc5Base/>) and encodes the raw data for the gc5Base track on hg19 (site last updated on 24-Apr-2009 14:48). We binned the data at 100 kb resolution for chromosome 19, using the following awk command: `awk -v res = 100000 'BEGIN{fname = "temp"}NF>2{split($2,c," "); if (fname! = c[2]){print tot >> fname".txt"; close(fname".txt"); m = res; tot = 0} fname = c[2]; }NF == 2 && $1>m{while ($1>m){print tot >> fname".txt"; m = m+res; tot = 0;}}$1<= m{tot+= $2; }END{print tot >> fname".txt"; close(fname".txt");}' my_file.wig`.



Finally, we calculated the Spearman rank correlation coefficient between the compartment eigenvector and the GC content track. We obtained a Spearman rho value of 0.72, which implies that the ‘A’ compartment corresponds to loci whose eigenvector entry is negative and the ‘B’ compartment corresponds to loci whose eigenvector is positive (S8 Fig, Interactive Figure <http://bit.ly/2oZ3vcU>).

Accordingly, we labeled compartmental intervals 1, 3, 5, 7, and 9 (at coordinates 7–8.6 Mb, 9.3–9.5 Mb, 9.6–9.8 Mb, 9.9–14.6 Mb, and 15.2–15.5 Mb) as compartment ‘A’ and intervals 2, 4, 6, and 8 (with coordinates 8.7–9.3 Mb, 9.5–9.6 Mb, 9.8–9.9 Mb, and 14.6–15.2 Mb respectively), as compartment ‘B’. Note, compartmental intervals may differ from chromosomal segments by genomic coordinates and/or compartment classification.

## Supporting information

### S1 Text. Integrative modeling procedure.

(DOCX)

### S2 Text. Assessment of in-situ Hi-C data reproducibility.

(DOCX)

**S1 Table. Chromosomal segments imaged with OligoSTORM.** First 9 rounds were imaged through Mainstreet, while subsequent rounds were imaged through Backstreet. Number of frames per round of imaging were 42,500–109,200. \* As in [Fig 2A–2D](#).

(DOCX)

**S2 Table. Multiple chromosome walks.** (A) Walk along chromosome 5. (B) Walk along chromosome 3. \* As in [Fig 2E](#).

(DOCX)

**S3 Table. Chromosomal segments imaged with OligoDNA-PAINT.** \* As in [Fig 2F](#). Number of frames for each segment was 30,000.

(DOCX)

**S4 Table. Compartment classification.** Chromosomal segment (CS); Active, % active CS; Inactive, % inactive CS; Imaging classification; Hi-C classification ([Methods](#)).

(DOCX)

**S5 Table. Hi-C library statistics for PGP1f “mega map” (combined data from ten libraries).**

(DOCX)

**S6 Table. Homolog ellipticity.** Nucleus ID; ES1, ellipticity score for the first homolog; ES2, ellipticity score for the second homolog; ER, ellipticity ratio.

(DOCX)

**S7 Table. Assessment of potential for modeling of the analyzed regions.** Chromosomal segment (CS); Start and End, genomic location; MMPscore, the matrix modeling potential score; pSCC, predicted distance Spearman Cross Correlation for the predicted accuracy of the models; SCC, Spearman Cross Correlation between Hi-C contact frequencies and the model-based contact map.

(DOCX)

**S8 Table. Flexible fitting refinement: Assessment of the goodness-of-fit.** Chromosomal segment (CS); Cross Correlation Coefficient of the starting rigid fitted model (CCC<sub>s</sub>); Cross Correlation Coefficient of the final model after flexible fitting refinement (CCC<sub>f</sub>); Clash Score of

the starting rigid fitted model (CLSs); Clash Score of the final model after flexible fitting refinement (CLSf); Root Mean Square Deviation between starting rigid fitted model and final model after flexible fitting refinement (RMSDs-f).

(DOCX)

**S9 Table. Hi-C library statistics for PGP1f replicates.**

(DOCX)

**S1 Fig. *In situ* Hi-C map of CS1-9 in IMR-90.** Hi-C map of region including CS1-9, displayed at 25 kb resolution. (See [Methods](#) for data processing and normalization.) Contact count indicated by the color of each pixel, ranging from 0 (white) to 10 (red); OligoSTORM steps (CS1-9) delineated in green. Cartoons of chromosome 19 show extent of the 8.16 Mb imaged region (blue).

(TIF)

**S2 Fig. *In situ* Hi-C map of CS7 in PGP1f.** Hi-C map of region including CS7, displayed at 5 kb resolution. (See [Methods](#) for data processing and normalization.) Contact count indicated by the color of each pixel, ranging from 0 (white) to 39 (red); 4 subregions of CS7 delineated in green. Cartoons of chromosome 19 show extent of the 8.16 Mb imaged region (blue) and block of SNVs shared between PGP1 and PGP95 (black; as in [Fig 1B](#)).

(TIF)

**S3 Fig. Imaging chromosomes at multiple scales.** (A) Diploid nucleus showing walks along multiple chromosomes. Walks were initiated in rounds 4, 5, and 6 on Chromosomes 8, 5, and 3 in step sizes of 100, 250, and 500 kb, respectively, while simultaneously walking along Chromosome 19, commencing in round 1 and in step sizes ranging from 3 to 1,800 kb. This walk encompassed a total of 23 cycles (see color bar) and, although it was not successful in all rounds on all chromosomes, it nevertheless illustrates the potential of this approach to increase throughput. Parent-of-origin identified with HOPs, as described later in the text. (B, C) Another diploid nucleus showing 9 chromosomal segments (B, ranging in size from 0.36 to 1.8 Mb; parental origin identified with HOPs) and the four subregions comprising the 840 kb of chromosomal segment 7 (C, subregions ranging in size from 90 to 350 kb). (D) A loop (290 kb), including its anchors (pink and orange; 10 kb each) and loop body (blue; 270 kb). (E) *DNMT1* gene (59.5 kb) and its DMR (2.9 kb). For more details, refer to [S1](#) and [S2](#) Tables.

(TIF)

**S4 Fig. Comparison of number of localizations per cluster between fluidic cycles 1 and 21.**

After querying the change in fractional overlap for CS1 after 21 rounds of imaging with OligoSTORM ([Fig 2G](#)), we assessed whether the number of localizations had changed using the localizations detected from 20 homologs (total of 10 nuclei, colored as blue dots; [Methods](#)). The mean number of localizations per cluster (red cross) is similar between rounds 1 and 21, implying no significant change (two-sample Kolmogorov-Smirnov test,  $p = 0.9655$ ).

(TIF)

**S5 Fig. Variability of spatial distances and entanglement observed for chromosomal segments.**

(A) Distance versus genomic distance for the 342 analyzed chromosomal segments using a power-law function fitted to the data (blue line). Data are colored by their distance value (rainbow color palette). (B) Distribution and variation of the distance score (DS) of the two chromosomal segments being compared. (C) Entanglement versus genomic distance for the 342 analyzed chromosomal segments with a power-law function fitted to the data (blue line). Data are colored by their entanglement value (rainbow color palette). (D) Distribution and variation of the entanglement score (ES) of the two chromosomal segments being

compared. (B, D) Box boundaries represent 1st and 3rd quartiles, middle line represents median, and whiskers extend to 1.5 times the interquartile range. (Mann-Whitney rank, \*:  $p < 5 \times 10^{-2}$ ). (E, F) Hierarchical clustering of (E) mean distance scores (MDS) and (F) mean entanglement scores (MES) reveal two major groups of chromosomal segments. (TIF)

**S6 Fig. 3D features of chromosomal segments.** (A-C) Surface area, volume and sphericity versus genomic distance for the 342 analyzed chromosomal segments (linear fit in blue). (D-F) Box plots showing distribution and variation of the 9 chromosomal segments in terms of (D) surface area, (E) volume, and (F) sphericity score. The chromosomal segments are color-coded as in Fig 2A. Box boundaries represent 1st and 3rd quartiles, middle line represents median, and whiskers extend to 1.5 times the interquartile range (Mann-Whitney rank test, \*:  $p < 5 \times 10^{-2}$ , \*\*:  $p < 10^{-2}$ , \*\*\*:  $p < 10^{-3}$ ). (TIF)

**S7 Fig. Clustering of chromosomal segments.** (A) Unsupervised PCA of 3D features for the 342 visualized chromosomal segments, the 3D features being: surface area, volume, sphericity, entanglement, and distance scores. Upper panel, PCA biplot (combination of score plot and loading plot) assessing the relationships between the loadings and the variables of the first two components. The lengths of the arrows indicate how each original variable contributes to the component (variables that have little contribution have almost zero length). The direction of the vectors indicates how two variables are correlated. Aligned vectors are strongly positively correlated between each other, opposite vectors indicate negative correlation, and vectors at right angles indicate no correlation. Each data point coloured in grey represents a chromosomal segment. Lower panel, PCA of 3D features for the 342 visualized chromosomal segments, each chromosomal segment colored as in Fig 2A and 2B, with relative abundances displayed as pie charts. (B-J) Distribution along PC1 for CS1-9, color coded as in Fig 2. Vertical black dotted line represents the partition of cluster 1 and cluster 2. (K) Hierarchical clustering of the compartment state profile vectors for the 19 analyzed nuclei results in 5 large clusters. The compartment state profile (Methods) was calculated comparing the chromatin state profile of each homolog within a single nucleus. Note, the two most populated clusters confirm that, in the majority of the analyzed nuclei, the compartment state profile is maintained with CS2, 8, and 9 mostly inactive, CS3-7 mostly active, and CS1 seeming to occupy both states. (TIF)

**S8 Fig. *In situ* Hi-C analysis of 8.16 Mb imaged region segregates into two compartments.** Map shows region including CS1-9, displayed at 100 kb resolution. (See Methods for data processing and normalization.) Contact count indicated by the color of each pixel, ranging from 0 (white) to 138 (red); manually curated intervals in compartments A and B are delineated below the diagonal in black and yellow, respectively, and reflect the positive values (blue) and negative values (red), respectively, of the compartment eigenvector at 100 kb resolution (above map). Cartoons of chromosome 19 show extent of the 8.16 Mb imaged region (blue) and block of SNVs shared between PGP1 and PGP95 (black; as in Fig 1B). (TIF)

**S9 Fig. Comparison of *in situ* Hi-C and OligoSTORM.** Correlation of Hi-C contact frequencies between chromosomal segments and corresponding mean distance scores. Hi-C contact frequency between chromosomal segments is expressed as  $\log_2(\text{observed/expected})$  at 25 kb resolution.  $R = \sim 0.6$  ( $p\text{-value: } 5.5e10^{-5}$ ). (TIF)

**S10 Fig. Genomic features of the modeled 8.16 Mb region.** (A) IMGR protocol for fitting and refinement of 3D chromatin models. The protocol inputs are a density map derived from a single PGP1f nucleus imaged with OligoSTORM and an ensemble of 3D models derived from a PGP1f population-based *in situ* Hi-C contact map. The protocol entails rigid fitting by a random 6D search (step1), followed by flexible fitting refinement via simulated annealing rigid-body molecular dynamics and conjugate gradient minimization (step 2). After each step, the goodness-of-fit is assessed as described using different scores (Supporting information [S1 Text](#)). (B) Mapping of Znf clusters onto the 3D trace. The trace is colored in grey, and the two clusters are colored in blue and red. Insets are zoomed-in views of one of the Znf clusters, highlighting the differences in conformation of the cluster in the model of the two homologs. (C) Annotated disease-causing genes mapped into the 3D trace. The trace is colored in grey. Shown in red are genes for which the molecular basis for a disorder is known and a mutation has been found, as classified by OMIM Genes track in the UCSC Genome Browser [[114](#)]. (D-J) Mapping of genome features. Genomic features are presented as mean value per 10kb. The color scale for each genomic feature ranges from white (mean value) to red (mean + 2 standard deviation). The figures in this panel were created with Chimera [[105](#)]. (TIF)

**S11 Fig. Ellipticity ratios and homolog-specific analyses.** (A) Ellipticity ratios of the 19 nuclei analyzed (see [S6 Table](#) for values of ellipticity ratios). Grey line, ellipticity score (1.0) expected when the two largest principal axes are equal. (B) Ratios of signal intensity using interstitial, HOP-M (Atto488N), and HOP-P (Atto565N) probes. (C) Same as B, except HOP-M is now labeled with Atto565N, and HOP-P is labeled with Atto488N. (D-I) The MDS and MES matrices for six nuclei showing only (D, G) maternal homologs, (E, H) only paternal homologs, and (F, I) maternal and paternal homologs, combined. (TIF)

**S12 Fig. Identification of the maternally derived haplotype of PGP1.** The longest blocks of SNVs on chromosome 19 of PGP1 that are shared with PGP95 correspond to the H1 haplotype, thus identifying homologs bearing long blocks of H1 SNVs as maternally derived (Mann-Whitney rank test, ns: non-significant, \*:  $p < 5 \times 10^{-2}$ , \*\*\*:  $p < 10^{-3}$ ). (TIF)

**S13 Fig.** (A) Number of localizations versus number of oligos and (B) genomic distance for the 342 analyzed chromosomal segments (linear fit in blue). (C) Number of localizations per round of imaging. Color coding as in [S7 Fig](#). (TIF)

**S14 Fig. Assessment of the rigid body step of the fitting protocol.** (A, B) Distribution of the Combined Score (CombScore) for each chromosomal segment of each of the two homologs traced. The chromosomal segments are color-coded as in [Fig 2A and 2B](#). (TIF)

**S15 Fig. *In situ* Hi-C results for PGP1f are reproducible.** (A) The outcomes of ten Hi-C experiments conducted on PGP1f cells are extremely similar, as shown here and in [S9 Table](#). The Hi-C maps represent the entirety of chromosome 19, displayed at 100 kb resolution. (See [Methods](#) for data processing and normalization.) Contact count indicated by the color of each pixel, ranging from 0 (white) to 34 (red). Cartoons of chromosome 19 show extent of the 8.16 Mb imaged region (blue) and block of SNVs shared between PGP1 and PGP95 (black; as in [Fig 1B](#)). (B) Correlation between *in-situ* Hi-C PGP1f libraries as a function of distance; Library 1 vs libraries 2–10 are shown at 100 and 500 kb resolutions. Hi-C replicates are highly

correlated as demonstrated by the Pearson's  $r$  value. At 500 kb resolution,  $r > 0.9$  for distances of tens of megabases and  $0.99 > r > 0.95$  at a distance of 1 Mb for all replicates. At 100 kb resolution,  $r > 0.85$  at a distance of 1 Mb for all replicates. Our results indicate that PGP1f libraries are highly similar.

(TIF)

**S1 Movie. OligoSTORM imaging of the 8.16 Mb region of human chromosome 19 encompassing CS1-9.** Nucleus is the same as that shown in [Fig 2B](#). Color coding as in [Fig 2A and 2B](#). (MP4)

**S2 Movie. Model obtained via IMGR of the 8.16 Mb region of human chromosome 19 encompassing CS1-9.** Nucleus is the same as that shown in [Fig 2B](#). Color coding as in [Fig 2A and 2B](#).

(MOV)

## Acknowledgments

We thank Geoffrey Fudenberg for calling contact domains and compartmental intervals in IMR-90 cells; Kun Zhang, Daniel Jacobsen, Bing Ren, and Anthony Schmitt for PGP1 haplotyping; Bogdan Bintu for help regarding microfluidics; Bogdan Bintu, Alistair Boettiger, Florian Schueder, and Xiaowei Zhuang for helpful input; and Jennifer Waters and the Nikon Imaging Center of Harvard Medical School. The authors also thank members of the Yin, Aiden, Marti-Renom, and Wu laboratories for countless valuable discussions. Finally, C.-t.W. thanks H.N., in particular, for many conversations regarding the implications of structural variation and turbulence and, for remarkable cells and years of discussion, she also thanks PGP1 and PGP95.

## Author Contributions

**Conceptualization:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Son C. Nguyen, Ruth B. McCole.

**Data curation:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Carl G. Ebeling, Steven P. Callahan, John M. Schreiner.

**Formal analysis:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Carl G. Ebeling, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Son C. Nguyen, Ruth B. McCole, Suhas S. P. Rao, Jocelyn Y. Kishi, Paula Soler-Vila.

**Funding acquisition:** Jeff A. Stuckey, Peng Yin, Erez Lieberman Aiden, Marc A. Marti-Renom, C.-ting Wu.

**Investigation:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Son C. Nguyen, Ruth B. McCole, Shyamtanu Chatteraj, Jumana AlHaj Abed, Mohammed A. Hannan, Sheikh Russell, Neva C. Durand.

**Methodology:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Carl G. Ebeling, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Son C. Nguyen, Ruth B. McCole.

**Project administration:** Jeff A. Stuckey, Peng Yin, Erez Lieberman Aiden, Marc A. Marti-Renom, C.-ting Wu.

**Resources:** Jelena Erceg, Jumana AlHaj Abed, Nuno M. C. Martins, Huy Q. Nguyen.

**Software:** Irene Farabella, Carl G. Ebeling, Steven P. Callahan, John M. Schreiner.

**Supervision:** Jeff A. Stuckey, Peng Yin, Erez Lieberman Aiden, Marc A. Marti-Renom, C.-ting Wu.

**Validation:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Carl G. Ebeling, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Ruth B. McCole, Michele Di Pierro, José N. Onuchic.

**Visualization:** Carl G. Ebeling, Steven P. Callahan, John M. Schreiner.

**Writing – original draft:** Guy Nir, Irene Farabella, Cynthia Pérez Estrada, Carl G. Ebeling, Brian J. Beliveau, Hiroshi M. Sasaki, S. Dean Lee, Ruth B. McCole, Jumana AlHaj Abed.

**Writing – review & editing:** Jeff A. Stuckey, Peng Yin, Erez Lieberman Aiden, Marc A. Marti-Renom, C.-ting Wu.

## References

1. Cullen KE, Kladde MP, Seyfred MA. Interaction between transcription regulatory regions of prolactin chromatin. *Science*. 1993; 261(5118):203–6. PMID: [8327891](#)
2. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002; 295(5558):1306–11. <https://doi.org/10.1126/science.1067799> PMID: [11847345](#)
3. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009; 326(5950):289–93. <https://doi.org/10.1126/science.1181369> PMID: [19815776](#)
4. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012; 485(7398):376–80. <https://doi.org/10.1038/nature11082> PMID: [22495300](#)
5. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nature Biotechnol*. 2012; 30(1):90.
6. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 2012; 485(7398):381–5. <https://doi.org/10.1038/nature11049> PMID: [22495304](#)
7. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, et al. Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell*. 2012; 148(3):458–72. <https://doi.org/10.1016/j.cell.2012.01.010> PMID: [22265598](#)
8. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*. 2014; 159(7):1665–80. <https://doi.org/10.1016/j.cell.2014.11.021> PMID: [25497547](#)
9. Nicodemi M, Pombo A. Models of chromosome structure. *Curr Opin Cell Biol*. 2014; 28:90–5. <https://doi.org/10.1016/j.ceb.2014.04.004> PMID: [24804566](#)
10. Serra F, Di Stefano M, Spill YG, Cuartero Y, Goodstadt M, Baù D, et al. Restraint-based three-dimensional modeling of genomes and genomic domains. *FEBS Lett*. 2015; 589(20):2987–95.
11. Wang S, Su J-H, Beliveau BJ, Bintu B, Moffitt JR, Wu CT, et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science*. 2016; 353(6299):598–602. <https://doi.org/10.1126/science.aaf8084> PMID: [27445307](#)
12. Abbas A, He X, Zhou B, Zhu G, Ma Z, Gao J-T, et al. Integrating Hi-C and FISH data for modeling 3D organizations of chromosomes. *bioRxiv*. 2018.
13. Bintu B, Mateo LJ, Su J-H, Sinnott-Armstrong NA, Parker M, Kinrot S, et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*. 2018; 362(6413).
14. Beliveau BJ, Boettiger AN, Avendaño MS, Jungmann R, McCole RB, Joyce EF, et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat Commun*. 2015; 6:7147. <https://doi.org/10.1038/ncomms8147> PMID: [25962338](#)
15. Fabre PJ, Benke A, Joye E, Huynh THN, Manley S, Duboule D. Nanoscale spatial organization of the HoxD gene cluster in distinct transcriptional states. *Proc Natl Acad Sci U S A*. 2015; 112(45):13964–9. <https://doi.org/10.1073/pnas.1517972112> PMID: [26504220](#)
16. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*. 2016; 529:418. <https://doi.org/10.1038/nature16496> PMID: [26760202](#)
17. Cattoni DI, Gizzi AMC, Georgieva M, Stefano M, Valeri A, Chamousset D, et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet

- specific interactions. *Nat Commun.* 2017; 8(1):1753. <https://doi.org/10.1038/s41467-017-01962-x> PMID: 29170434
18. Kundu S, Ji F, Sunwoo H, Jain G, Lee JT, Sadreyev RI, et al. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Mol Cell.* 2017; 65(3):432–46. e5. <https://doi.org/10.1016/j.molcel.2017.01.009> PMID: 28157505
  19. Trussart M, Yus E, Martinez S, Bau D, Tahara YO, Pengo T, et al. Defined chromosome structure in the genome-reduced bacterium *Mycoplasma pneumoniae*. *Nat Commun.* 2017; 8:14665. <https://doi.org/10.1038/ncomms14665> PMID: 28272414
  20. Szabo Q, Jost D, Chang J-M, Cattoni DI, Papadopoulos GL, Bonev B, et al. TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci Adv.* 2018; 4(2):eaar8082.
  21. Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH, O’Shea CC. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science.* 2017; 357(6349).
  22. Cardozo Gizzi AM, Cattoni DI, Fiche J-B, Espinola S, Gurgo J, Messina O, et al. Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *bioRxiv.* 2018.
  23. Beliveau BJ, Joyce EF, Apostolopoulos N, Yilmaz F, Fonseka CY, McCole RB, et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc Natl Acad Sci U S A.* 2012; 109(52):21301–6. <https://doi.org/10.1073/pnas.1213818110> PMID: 23236188
  24. Beliveau BJ, Boettiger AN, Nir G, Bintu B, Yin P, Zhuang X, et al. In Situ Super-Resolution Imaging of Genomic DNA with OligoSTORM and OligoDNA-PAINT. In: Erfle H, editor. *Super-Resolution Microscopy.* 1663: Springer; 2017. p. 231–52.
  25. Chu WK, Edge P, Lee HS, Bansal V, Bafna V, Huang X, et al. Ultraaccurate genome sequencing and haplotyping of single human cells. *Proc Natl Acad Sci U S A.* 2017; 114(47):12512–7. <https://doi.org/10.1073/pnas.1707609114> PMID: 29078313
  26. Kislauskis EH, Li Z, Singer RH, Taneja KL. Isoform-specific 3'-untranslated sequences sort alpha-cardiac and beta-cytoplasmic actin messenger RNAs to different cytoplasmic compartments. *J Cell Biol.* 1993; 123(1):165–72. PMID: 8408195
  27. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science.* 2015; 348(6233):aaa6090.
  28. 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.* 2016; 92(2):342–57. <https://doi.org/10.1016/j.neuron.2016.10.001> PMID: 27764670
  29. Eng C-HL, Shah S, Thomassie J, Cai L. Profiling the transcriptome with RNA SPOTs. *Nat Methods.* 2017; 14(12):1153. <https://doi.org/10.1038/nmeth.4500> PMID: 29131163
  30. Shah S, Takei Y, Zhou W, Lubeck E, Yun J, Eng C-HL, et al. Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. *Cell.* 2018:363–76. <https://doi.org/10.1016/j.cell.2018.05.035> PMID: 29887381
  31. Bates M, Huang B, Dempsey GT, Zhuang X. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science.* 2007; 317(5845):1749–53. <https://doi.org/10.1126/science.1146598> PMID: 17702910
  32. Heilemann M, Van De Linde S, Schüttelz M, Kasper R, Seefeldt B, Mukherjee A, et al. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed.* 2008; 47(33):6172–6.
  33. Jungmann R, Steinhauer C, Scheible M, Kuzyk A, Tinnefeld P, Simmel FC. Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami. *Nano Lett.* 2010; 10(11):4756–61. <https://doi.org/10.1021/nl103427w> PMID: 20957983
  34. Jungmann R, Avendaño MS, Woehrstein JB, Dai M, Shih WM, Yin P. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat Methods.* 2014; 11(3):313–8. <https://doi.org/10.1038/nmeth.2835> PMID: 24487583
  35. Yurke B, Turberfield AJ, Mills AP Jr, Simmel FC, Neumann JL. A DNA-fuelled molecular machine made of DNA. *Nature.* 2000; 406(6796):605. <https://doi.org/10.1038/35020524> PMID: 10949296
  36. Zhang DY, Chen SX, Yin P. Optimizing the specificity of nucleic acid hybridization. *Nat Chem.* 2012; 4(3):208–14. <https://doi.org/10.1038/nchem.1246> PMID: 22354435
  37. Rosin LF, Nguyen SC, Joyce EF. Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in *Drosophila* nuclei. *PLoS Genet.* 2018; 14(7):e1007393. <https://doi.org/10.1371/journal.pgen.1007393> PMID: 30001329
  38. Murgha Y, Beliveau B, Semrau K, Schwartz D, Wu CT, Gulari E, et al. Combined in vitro transcription and reverse transcription to amplify and label complex synthetic oligonucleotide probe libraries. *Bio-techniques.* 2015; 58(6):301–7. <https://doi.org/10.2144/000114298> PMID: 26054766

39. Ester M, Kriegel H-P, Sander J, Xu X, editors. A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*; 1996.
40. Das R, Lee YK, Strogantsev R, Jin S, Lim YC, Ng PY, et al. DNMT1 and AIM1 Imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation. *BMC Genomics*. 2013; 14:685. <https://doi.org/10.1186/1471-2164-14-685> PMID: [24094292](https://pubmed.ncbi.nlm.nih.gov/24094292/)
41. Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol*. 2005; 3(5): e157. <https://doi.org/10.1371/journal.pbio.0030157> PMID: [15839726](https://pubmed.ncbi.nlm.nih.gov/15839726/)
42. Dai M, Jungmann R, Yin P. Optical imaging of individual biomolecules in densely packed clusters. *Nat Nanotech*. 2016; 11(9):798.
43. Farabella I, Vasishtan D, Joseph AP, Pandurangan AP, Sahota H, Topf M. TEMPY: a Python library for assessment of three-dimensional electron microscopy density fits. *J Appl Crystallogr*. 2015; 48(4):1314–23.
44. Nuebler J, Fudenberg G, Imakaev M, Abdennur N, Mirny LA. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A*. 2018; 115(29):E6697–706. <https://doi.org/10.1073/pnas.1717730115> PMID: [29967174](https://pubmed.ncbi.nlm.nih.gov/29967174/)
45. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature*. 2017; 547(7662):236–40. <https://doi.org/10.1038/nature22822> PMID: [28636604](https://pubmed.ncbi.nlm.nih.gov/28636604/)
46. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. *Nature*. 2017; 547(7662):241–5. <https://doi.org/10.1038/nature22989> PMID: [28636597](https://pubmed.ncbi.nlm.nih.gov/28636597/)
47. Papantonis A, Cook PR. Transcription factories: genome organization and gene regulation. *Chem Rev*. 2013; 113(11):8683–705. <https://doi.org/10.1021/cr300513p> PMID: [23597155](https://pubmed.ncbi.nlm.nih.gov/23597155/)
48. van Steensel B, Belmont AS. Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell*. 2017; 169(5):780–91. <https://doi.org/10.1016/j.cell.2017.04.022> PMID: [28525751](https://pubmed.ncbi.nlm.nih.gov/28525751/)
49. Allshire RC, Madhani HD. Ten principles of heterochromatin formation and function. *Nat Rev*. 2018; 19(4):229–44.
50. Németh A, Grummt I. Dynamic regulation of nucleolar architecture. *Curr Opin Cell Biol*. 2018; 52:105–11. <https://doi.org/10.1016/j.ceb.2018.02.013> PMID: [29529563](https://pubmed.ncbi.nlm.nih.gov/29529563/)
51. Sawyer IA, Sturgill D, Dunder M. Membraneless nuclear organelles and the search for phases within phases. *Willy Interdiscip Rev RNA*. 2018; e1514.
52. Hartono SR, Korf IF, Chédin F. GC skew is a conserved property of unmethylated CpG island promoters across vertebrates. *Nucleic Acids Res*. 2015; 43(20):9729–41. <https://doi.org/10.1093/nar/gkv811> PMID: [26253743](https://pubmed.ncbi.nlm.nih.gov/26253743/)
53. Topf M, Lasker K, Webb B, Wolfson H, Chiu W, Sali A. Protein structure fitting and refinement guided by cryo-EM density. *Structure*. 2008; 16(2):295–307. <https://doi.org/10.1016/j.str.2007.11.016> PMID: [18275820](https://pubmed.ncbi.nlm.nih.gov/18275820/)
54. Tjong H, Li W, Kalhor R, Dai C, Hao S, Gong K, et al. Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. *Proc Natl Acad Sci U S A*. 2016; 113(12): E1663–E72. <https://doi.org/10.1073/pnas.1512577113> PMID: [26951677](https://pubmed.ncbi.nlm.nih.gov/26951677/)
55. Paulsen J, Gramstad O, Collas P. Manifold based optimization for single-cell 3D genome reconstruction. *PLOS Comput Biol*. 2015; 11(8):e1004396. <https://doi.org/10.1371/journal.pcbi.1004396> PMID: [26262780](https://pubmed.ncbi.nlm.nih.gov/26262780/)
56. Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*. 2017; 544(7648):59. <https://doi.org/10.1038/nature21429> PMID: [28289288](https://pubmed.ncbi.nlm.nih.gov/28289288/)
57. Tan L, Xing D, Chang C-H, Li H, Xie XS. Three-dimensional genome structures of single diploid human cells. *Science*. 2018; 361(6405):924–8. <https://doi.org/10.1126/science.aat5641> PMID: [30166492](https://pubmed.ncbi.nlm.nih.gov/30166492/)
58. Baù D, Marti-Renom MA. Genome structure determination via 3C-based data integration by the Integrative Modeling Platform. *Methods*. 2012; 58(3):300–6. <https://doi.org/10.1016/j.ymeth.2012.04.004> PMID: [22522224](https://pubmed.ncbi.nlm.nih.gov/22522224/)
59. Serra F, Baù D, Goodstadt M, Castillo D, Filion GJ, Marti-Renom MA. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. *PLOS Comput Biol*. 2017; 13(7):e1005665. <https://doi.org/10.1371/journal.pcbi.1005665> PMID: [28723903](https://pubmed.ncbi.nlm.nih.gov/28723903/)
60. Lukic S, Nicolas J, Levine A. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. *Cell Death Differ*. 2014; 21(3):381. <https://doi.org/10.1038/cdd.2013.150> PMID: [24162661](https://pubmed.ncbi.nlm.nih.gov/24162661/)



61. McKusick VA. Mendelian inheritance in man: a catalog of human genes and genetic disorders: JHU Press; 1998.
62. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489(7414):57. <https://doi.org/10.1038/nature11247> PMID: [22955616](https://pubmed.ncbi.nlm.nih.gov/22955616/)
63. Bestor TH, Chandler VL, Feinberg AP. Epigenetic effects in eukaryotic gene expression. *Dev Genet*. 1994; 15(6):458–62. <https://doi.org/10.1002/dvg.1020150603> PMID: [7834904](https://pubmed.ncbi.nlm.nih.gov/7834904/)
64. Wu CT, Morris JR. Transvection and other homology effects. *Curr Opin Genet Dev*. 1999; 9(2):237–46. [https://doi.org/10.1016/S0959-437X\(99\)80035-5](https://doi.org/10.1016/S0959-437X(99)80035-5) PMID: [10322135](https://pubmed.ncbi.nlm.nih.gov/10322135/)
65. Apte MS, Meller VH. Homologue pairing in flies and mammals: gene regulation when two are involved. *Genet Res Int*. 2012; 2012:430587. <https://doi.org/10.1155/2012/430587> PMID: [22567388](https://pubmed.ncbi.nlm.nih.gov/22567388/)
66. Joyce EF, Erceg J, Wu CT. Pairing and anti-pairing: a balancing act in the diploid genome. *Curr Opin Genet Dev*. 2016; 37:119–28. <https://doi.org/10.1016/j.gde.2016.03.002> PMID: [27065367](https://pubmed.ncbi.nlm.nih.gov/27065367/)
67. Khan WA, Rogan PK, Knoll JH. Localized, non-random differences in chromatin accessibility between homologous metaphase chromosomes. *Mol cytogenet*. 2014; 7(1):70. <https://doi.org/10.1186/s13039-014-0070-y> PMID: [25520753](https://pubmed.ncbi.nlm.nih.gov/25520753/)
68. Khan WA, Rogan PK, Knoll JH. Reversing chromatin accessibility differences that distinguish homologous mitotic metaphase chromosomes. *Mol Cytogenet*. 2015; 8(1):65.
69. Chess A. Random and non-random monoallelic expression. *Neuropsychopharmacology*. 2013; 38(1):55–61. <https://doi.org/10.1038/npp.2012.85> PMID: [22763620](https://pubmed.ncbi.nlm.nih.gov/22763620/)
70. Nilsson M, Krejci K, Koch J, Kwiatkowski M, Gustavsson P, Landegren U. Padlock probes reveal single-nucleotide differences, parent of origin and in situ distribution of centromeric sequences in human chromosomes 13 and 21. *Nat Genet*. 1997; 16(3):252. <https://doi.org/10.1038/ng0797-252> PMID: [9207789](https://pubmed.ncbi.nlm.nih.gov/9207789/)
71. O'Keefe CL, Warburton PE, Matera AG. Oligonucleotide probes for alpha satellite DNA variants can distinguish homologous chromosomes by FISH. *Hum Mol Gen*. 1996; 5(11):1793–9. PMID: [8923008](https://pubmed.ncbi.nlm.nih.gov/8923008/)
72. Auduge N, Padilla-Parra S, Tramier M, Borghi N, Coppey-Moisan M. Chromatin condensation fluctuations rather than steady-state predict chromatin accessibility. *bioRxiv*. 2018.
73. Rogan PK, Cazcarro PM, Knoll JH. Sequence-based design of single-copy genomic DNA probes for fluorescence in situ hybridization. *Genome Res*. 2001; 11(6):1086–94. <https://doi.org/10.1101/gr.171701> PMID: [11381034](https://pubmed.ncbi.nlm.nih.gov/11381034/)
74. Navin N, Grubor V, Hicks J, Leibu E, Thomas E, Troge J, et al. PROBER: oligonucleotide FISH probe design software. *Bioinformatics*. 2006; 22(19):2437–8. <https://doi.org/10.1093/bioinformatics/btl273> PMID: [16740623](https://pubmed.ncbi.nlm.nih.gov/16740623/)
75. Boyle S, Rodesch MJ, Halvensleben HA, Jeddelloh JA, Bickmore WA. Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Res*. 2011; 19(7):901–9. <https://doi.org/10.1007/s10577-011-9245-0> PMID: [22006037](https://pubmed.ncbi.nlm.nih.gov/22006037/)
76. Yamada N, Rector L, Tsang P, Carr E, Scheffer A, Sederberg M, et al. Visualization of fine-scale genomic structure by oligonucleotide-based high-resolution FISH. *Cytogenet Genome Res*. 2011; 132(4):248–54. <https://doi.org/10.1159/000322717> PMID: [21178330](https://pubmed.ncbi.nlm.nih.gov/21178330/)
77. Bienko M, Crossetto N, Teytelman L, Klemm S, Itzkovitz S, Van Oudenaarden A. A versatile genome-scale PCR-based pipeline for high-definition DNA FISH. *Nat Methods*. 2013; 10(2):122. <https://doi.org/10.1038/nmeth.2306> PMID: [23263692](https://pubmed.ncbi.nlm.nih.gov/23263692/)
78. Beliveau BJ, Kishi JY, Nir G, Sasaki HM, Saka SK, Nguyen SC, et al. OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proc Natl Acad Sci U S A*. 2018; 115(10):E2183–E92. <https://doi.org/10.1073/pnas.1714530115> PMID: [29463736](https://pubmed.ncbi.nlm.nih.gov/29463736/)
79. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*. 2011; 27(6):764–70. <https://doi.org/10.1093/bioinformatics/btr011> PMID: [21217122](https://pubmed.ncbi.nlm.nih.gov/21217122/)
80. Lo C, Liu R, Lee J, Robasky K, Byrne S, Lucchesi C, et al. On the design of clone-based haplotyping. *Genome Biol*. 2013; 14(9):100.
81. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010; 26(6):841–2. <https://doi.org/10.1093/bioinformatics/btq033> PMID: [20110278](https://pubmed.ncbi.nlm.nih.gov/20110278/)
82. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods*. 2012; 9(10):999. <https://doi.org/10.1038/nmeth.2148> PMID: [22941365](https://pubmed.ncbi.nlm.nih.gov/22941365/)

83. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res*. 2011; 12(Oct):2825–30.
84. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature*. 2017; 551(7678):51. <https://doi.org/10.1038/nature24281> PMID: 29094699
85. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012; 40(15):e115–e. <https://doi.org/10.1093/nar/gks596> PMID: 22730293
86. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, et al. NUPACK: analysis and design of nucleic acid systems. *J Comput Chem*. 2011; 32(1):170–3. <https://doi.org/10.1002/jcc.21596> PMID: 20645303
87. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9(4):357. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
88. Ball MP, Thakuria JV, Zaranek AW, Clegg T, Rosenbaum AM, Wu X, et al. A public resource facilitating clinical use of genomes. *Proc Natl Acad Sci U S A*. 2012; 109(30):11920–7. <https://doi.org/10.1073/pnas.1201904109> PMID: 22797899
89. Solovei I, Cavallo A, Schermelleh L, Jaunin F, Scasselati C, Cmarko D, et al. Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). *Exp Cell Res*. 2002; 276(1):10–23. <https://doi.org/10.1006/excr.2002.5513> PMID: 11978004
90. Beliveau BJ, Apostolopoulos N, Wu CT. Visualizing Genomes with Oligopaint FISH Probes. *Current Protocols in Molecular Biology*: John Wiley & Sons, Inc.; 2014.
91. Juette MF, Gould TJ, Lessard MD, Mlodzianoski MJ, Nagpure BS, Bennett BT, et al. Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. *Nat Methods*. 2008; 5(6):527. <https://doi.org/10.1038/nmeth.1211> PMID: 18469823
92. Mlodzianoski MJ, Juette MF, Beane GL, Bewersdorf J. Experimental characterization of 3D localization techniques for particle-tracking and super-resolution microscopy. *Opt Express*. 2009; 17(10):8264–77. PMID: 19434159
93. Deschout H, Zanicchi FC, Mlodzianoski M, Diaspro A, Bewersdorf J, Hess ST, et al. Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat Methods*. 2014; 11(3):253. <https://doi.org/10.1038/nmeth.2843> PMID: 24577276
94. Tokunaga M, Imamoto N, Sakata-Sogawa K. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat Methods*. 2008; 5(2):159. <https://doi.org/10.1038/nmeth1171> PMID: 18176568
95. Huang B, Wang W, Bates M, Zhuang X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*. 2008; 319(5864):810–3. <https://doi.org/10.1126/science.1153529> PMID: 18174397
96. Izeddin I, El Beheiry M, Andilla J, Ciepiewski D, Darzacq X, Dahan M. PSF shaping using adaptive optics for three-dimensional single-molecule super-resolution imaging and tracking. *Opt Express*. 2012; 20(5):4957–67. <https://doi.org/10.1364/OE.20.004957> PMID: 22418300
97. Schnitzbauer J, Strauss MT, Schlichthaerle T, Schueder F, Jungmann R. Super-resolution microscopy with DNA-PAINT. *Nat Protoc*. 2017; 12(6):1198. <https://doi.org/10.1038/nprot.2017.024> PMID: 28518172
98. Endesfelder U, Malkusch S, Fricke F, Heilemann M. A simple method to estimate the average localization precision of a single-molecule localization microscopy experiment. *Histochem Cell Biol*. 2014; 141(6):629–38. <https://doi.org/10.1007/s00418-014-1192-3> PMID: 24522395
99. El Beheiry M, Dahan M. ViSP: representing single-particle localizations in three dimensions. *Nat Methods*. 2013; 10(8):689. <https://doi.org/10.1038/nmeth.2566> PMID: 23900246
100. Chao J, Ram S, Abraham AV, Ward ES, Ober RJ. A resolution measure for three-dimensional microscopy. *Opt Commun*. 2009; 282(9):1751–61. <https://doi.org/10.1016/j.optcom.2009.01.062> PMID: 20161040
101. Tahmasbi A, Ward ES, Ober RJ. Determination of localization accuracy based on experimentally acquired image sets: applications to single molecule microscopy. *Opt Express*. 2015; 23(6):7630–52. <https://doi.org/10.1364/OE.23.007630> PMID: 25837101
102. Mlodzianoski MJ, Schreiner JM, Callahan SP, Smolková K, Dlasková A, Šantorová J, et al. Sample drift correction in 3D fluorescence photoactivation localization microscopy. *Opt Express*. 2011; 19(16):15009–19. <https://doi.org/10.1364/OE.19.015009> PMID: 21934862
103. Jonas. Plot spread points (beeswarm plot). 1.2.0.0 ed. MATLAB Central File Exchange: Mathworks; 2012.

104. Joseph AP, Lagerstedt I, Patwardhan A, Topf M, Winn M. Improved metrics for comparing structures of macromolecular assemblies determined by 3D electron-microscopy. *J Struct Biol.* 2017; 199(1):12–26. <https://doi.org/10.1016/j.jsb.2017.05.007> PMID: [28552721](https://pubmed.ncbi.nlm.nih.gov/28552721/)
105. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25(13):1605–12. <https://doi.org/10.1002/jcc.20084> PMID: [15264254](https://pubmed.ncbi.nlm.nih.gov/15264254/)
106. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: [19505943](https://pubmed.ncbi.nlm.nih.gov/19505943/)
107. Taylor W, Thornton Jt, Turnell W. An ellipsoidal approximation of protein shape. *J Mol Graph.* 1983; 1(2):30–8.
108. Trussart M, Serra F, Baù D, Junier I, Serrano L, Marti-Renom MA. Assessing the limits of restraint-based 3D modeling of genomes and genomic domains. *Nucleic Acids Res.* 2015; 43(7):3465–77. <https://doi.org/10.1093/nar/gkv221> PMID: [25800747](https://pubmed.ncbi.nlm.nih.gov/25800747/)
109. Roseman AM. Docking structures of domains into maps from cryo-electron microscopy using local correlation. *Acta Crystallogr D.* 2000; 56(10):1332–40.
110. Zeev-Ben-Mordehai T, Vasishtan D, Durán AH, Vollmer B, White P, Pandurangan AP, et al. Two distinct trimeric conformations of natively membrane-anchored full-length herpes simplex virus 1 glycoprotein B. *Proc Natl Acad Sci U S A.* 2016; 113(15):4176–81. <https://doi.org/10.1073/pnas.1523234113> PMID: [27035968](https://pubmed.ncbi.nlm.nih.gov/27035968/)
111. Vasishtan D, Topf M. Scoring functions for cryoEM density fitting. *J Struct Biol.* 2011; 174(2):333–43. <https://doi.org/10.1016/j.jsb.2011.01.012> PMID: [21296161](https://pubmed.ncbi.nlm.nih.gov/21296161/)
112. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* 2016; 3(1):95–8. <https://doi.org/10.1016/j.cels.2016.07.002> PMID: [27467249](https://pubmed.ncbi.nlm.nih.gov/27467249/)
113. Knight PA, Ruiz D. A fast algorithm for matrix balancing. *IMA J Numer Anal.* 2013; 33(3):1029–47.
114. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res.* 2002; 12(6):996–1006. <https://doi.org/10.1101/gr.229102> PMID: [12045153](https://pubmed.ncbi.nlm.nih.gov/12045153/)