Observing the natural and heteromorphic organization of the human genome in nanoscale resolution

Jeremy C. Horrell

B.A., Minot State University, 2011

Thesis

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Division of Biology and Medicine at Brown University

Providence, Rhode Island

May 2022

This dissertation by Jeremy C. Horrell is accepted in its present form by the Division of Biology and Medicine as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date _____ _____

Nicola Neretti, Ph.D.
(Primary Advisor)

Date _____ _____

John M. Sedivy, Ph.D.

Recommended to the Graduate Council

Date _____ _____

Erica Larschan, Ph.D.

Date _____ _____

Richard N. Freiman, Ph.D.

Date _____ _____

Robbert Creton, Ph.D.

Approved by the Graduate Council

Date _____ _____

Andrew G. Campbell, Ph.D.
Dean of the Graduate School

Jeremy Christian Horrell

Curriculum Vitae

Education

Sept. 2016 – Present — **Brown University, Dept. of Molecular Biology, Cell Biology and Biochemistry**
Providence, Rhode Island
Ph.D. in Molecular Biology, Cell Biology and Biochemistry
Advisors: Nicola Neretti, Ph.D and John M. Sedivy, Ph.D.

Aug. 2006 – May 2011 — **Minot State University, Dept. of Biology**
Minot, North Dakota
B.A. in Biology and Chemistry

Research Experience

May 2017 – Present — **Doctoral Thesis Research, Brown University**
Advisors: Nicola Neretti, Ph.D and John M. Sedivy, Ph.D.
Thesis: Observing the natural and heteromorphic organization of the human genome in nanoscale resolution.

Jan. 2019 – Nov. 2021 — **Visiting Doctoral Scholar, Dept. of Genetics, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts**
Advisor: Chao-ting Wu
Collaboration as part of the 4D-Nucleome Consortium (NIH) with the purpose of implementing sequential and multiplexed imaging of the human genome using Oligopaints-based technology and super-resolution 3D STORM.

June 2014 – Aug. 2016 — **Research Technician C, Dept. of Dermatology, Perelman School of Medicine, University of Pennsylvania Philadelphia, Pennsylvania**
Supervisor: Sarah E. Millar, Ph.D.
Characterized the effects of WNT10A in adult ectodermally-derived stem cell, proliferating cell, and differentiating cell populations in human subjects and murine genetic models.

Mar. 2008 – June 2010 — **Research Assistant, INBRE, Dept. of Biology, Minot State University, Minot, North Dakota**
Advisor: Heidi Super, Ph.D.
Investigated the mechanism of chromosomal rearrangements involved in the myeloid-lymphoid leukemia gene (*MLL*) and surveyed its association with acute lymphoblastic and acute myeloid leukemias.

Publications

1. Horrell, J.C., Chattoraj, S., Evans, S.A., Clark, K., Nguyen, H., Alkhatib, A., Wu. C-Ting, Neretti, N. Sequential OligoSTORM reveals the hierarchical topology of human chromosome 9. (2022). (*Manuscript in preparation*).
2. Evans, S. A., Horrell, J. & Neretti, N. The three-dimensional organization of the genome in cellular senescence and age-associated diseases. Semin Cell Dev Biol 90, 154–160 (2018).
3. M. Xu, J. Horrell, M. Snitow, J. Cui, H. Gochnauer, C. M. Syrett, S. Kallish, J. T. Seykora, F. Liu, D. Gaillard, J. P. Katz, K. H. Kaestner, B. Levin, C. Mansfield, J. E. Douglas, B. J. Cowart, M. Tordoff, F. Liu, X. Zhu, L. A. Barlow, A. I. Rubin, J. A. McGrath, E. E. Morrisey, E. Y. Chu, S. E. Millar, WNT10A mutation causes ectodermal dysplasia by impairing progenitor cell proliferation and KLF4-mediated differentiation. *Nat Commun*. 8, 15397 (2017).

Funding

Sept. 2018 – May 2020 **Initiative to Maximize Student Development (IMSD)**
Brown University, Providence, Rhode Island.

Selected talks

Oct. 2019 **Biology of Aging Retreat, Brown University, Providence Rhode Island**
A proposal: using Oligopaints and super-resolution microscopy to interrogate genomic loci in senescent cells.

Oct. 2019 **Chromosome Pairing Meeting, Cleveland Clinic Cleveland, Ohio**
Taking a TAD walk along human chromosome 9

Oct. 2017 **Chromosome Pairing Meeting, Bowdoin College Brunswick, Maine**
Cytoplasmic DNA in cellular senescence

Posters

Sept. 2020 **Epigenetics and Chromatin Meeting, Cold Spring Harbor Laboratories, Virtual**
Display of topologically associating domains along the span of human chromosome 9 at nano-scale resolution.

Jan. 2020 **4D Nucleome Annual Conference, Washington D.C.**
Taking a TAD walk along human chromosome 9

May 2018 **Biology of Human Aging Colloquium, Brown University**

	Providence, Rhode Island Uncovering mechanisms and physical features of senescence-associated cytosolic DNA
Mar. 2018	**Keystone Symposia: Chromatin Architecture and Chromosome Organization, Whistler, British Columbia** Senescence-associated chromatin instability leads to aberrant cytosolic DNA

Awards & Honors

2009-2010	President's Award (Minot State University)
2009	The Ruth & Charles Hoffman Botanical Scholarship (Minot State University)
2009	The Gertrude M. Eck Scholarship (Minot State University)
2001, 2005	Varsity Coach – Charter School of Wilmington, Wilmington, Delaware Delaware High School Soccer Championship

---

Teaching Experience

Sept. 2017 – Dec. 2017	**Graduate Teaching Assistant, Brown University** Course: Genetics
Oct. 2010 – Apr. 2011	**Tutor, Minot State University** Courses: biological sciences, Human Anatomy & Physiology

---

Skills

Programming languages	Bash-script, R statistical language
Data visualization	UCSC Genome Browser, IGV, WashU Epigenomic Browser Adobe Illustrator, Adobe Photoshop, Adobe InDesign, Prism Graphpad, Juicebox.js
Laboratory	Oligopaints library design; 3D FISH; Microfluidics automation; Super-resolution microscopy: STORM, SIM; Fluorescent Microscopy: confocal, widefield; Cell culture; Immunocyto/histochemistry

Acknowledgements

First and foremost, I would like to thank my dad Lester Horrell, my mom Ava Horrell, my brother Carlyle, and sister Roxanne for their unwavering support and love in my decade's long education journey. I thank them for being patient with me in my pursuit of using my curiosity as a career, even when it didn't seem as the most logical path. I also thank, from the bottom of my heart and my soul, the support and love of my fiancé. Heather, I could not have gone through with this process without you. Your many days of encouragement got me through so much of my Ph.D. that I often feel as though you should be awarded a degree as well. I also thank you for being an example for me, you showed me the way through your work ethic and your persistence. Words cannot express the gratitude that I have for you and the joy you bring me each day.

I would also like to thank my extended family and close friends, who always had words of inspiration for me. I am thankful that I have been given the opportunity to try to make all of you proud. My thoughts often centered around those of you, past and present, who have never been given the choice or the prospect to pursue your dreams. I can unequivocally say that my Ph.D. degree is dedicated to you. Special thanks to the MCB graduate program, to Mark Johnson, Judith Bender, Greg Valdez, Missy Berry, Ray Windsor, Ashley Walker, and Elaine Butler. To John Sedivy, I am thankful to you for inspiring me to continue to ask the big questions…in aging, always. In addition, I would like to acknowledge and thank all members of my thesis committee, Erica Larschan, Richard Freiman, and Robbert Creton for pointing me in the correct direction in all of my projects, and for your guidance and advice for the next steps in my career. Also, a special thanks goes out to my MCB cohort, for providing a web of support that made each of us stronger, together. I would also like to recognize and thank all members of the Wu lab and the Harvard Medical School community. I am especially grateful for Ting Wu's grace and

guidance during my time in her lab. A million thanks go out to Shyamtanu Chattoraj, with whom I have found a great mentor, colleague, and friend.

Lastly, I would like to thank my lab for always being there on an intellectual and a personal level. To Shane Evans, thank you for being my partner in crime and for picking up some of the loose ends I would often leave behind. Nick Skvir, thank you for all the end of day talks we've had about life and for sometimes sharing your snacks with me. To Yee Voan Teo, thank you for being such a special person where there is never a dull moment. To Azu Rocha, thank you for being that person who was always there for me, no questions asked, whether it be in lab or out—no one on Earth can order crescent wrenches like you do. To Audrey Dalgarno, I thank you for truly understanding me even when my weirdness level is at 10 out of 10, and for your appreciation of 'microscoper' tacos, perhaps it is true what they say, great minds do you think alike. To my mentor, Nicola Neretti, I am eternally grateful for all of your advice, your patience, and your direction in helping me through my Ph.D. career. Thank you for recognizing my independent style of learning, and for giving me physical and intellectual space to try to figure things out on my own—for your openness to fielding new ideas. Finally, thank you for serving as a good model to which I can fashion my own career pursuits.

Table of Contents

List of Tables

List of Figures

List of Supplementary Data

CHAPTER 1: INTRODUCTION

DNA is life

DNA, it is the most complex molecule that everyone knows about. Over the past few decades DNA has made its claim to fame in pop culture, in movies, in music, and in art. One has to only mention Jurassic Park and their thoughts turn to DNA. Indeed, DNA has become commonplace from dinner table to benchtop table. It has also become familiar the realization by the general public the importance of DNA—from the expression of genes to heredity. In the 19th century the father of genetics, Gregor Mendel, was the first to unknowingly refer to DNA as a 'substance' that is passed from generation to generation(Strachan et al., 2019). The first to isolate DNA was Frederich Miescher also in the mid-19th century, and he termed this molecule, nuclein (Dahm, 2008). For the next 100 years, scientists learned more about this mysterious molecule of life. Further key discoveries included the identification of nucleic acids by Albrecht Kossel , "jumping genes" by Barbara McClintock (McClintock, 1953), and DNA structure discoveries by Roslind Franklin and later James Watson and Francis Crick. But why is there such a fuss about DNA? Why is it so important? How does it work?

First, we will have to explore what the physical properties of DNA are in terms of its chemistry and its structure. We will also explore what the mechanistic properties of DNA are, how it interacts with other molecules in the nucleus, and what do we know of DNA organization and three-dimensional structure. Lastly, we want to consider what techniques that are currently available to interrogate DNA structure in the nucleus, what means do we have to study DNA?

Intrinsic structure of DNA

The molecular structure of deoxynucleic acid (DNA) consists of the combination of four nucleotides, cytosine, guanine, adenine, thymine. These nucleotides are joined together by covalent bonds, cytosine and guanine form a pair while adenine and thymine form the other pair. The cytosine and guanine pair are held together by three hydrogen

bonds while the adenine and thymine pair are held together by 2 hydrogen bonds. This means the energy that is needed to form or to break these bonds between these unique pairs are different. These nucleotides are joined together between the sugar molecule of one nucleotide and the phosphate molecule of another (Gregory et al., 2006). Other important components of DNA are phosphate and sugar groups, which forms the backbone of DNA. The sugars are joined by phosphate groups that form phosphodiester bonds between the 3rd (3'-end) and 5th (5'-end) carbon atoms of the nearest sugar molecule. Therefore, any DNA strand normally has one end where a phosphate group is attached to a 5' carbon and the other end there's attached to a 3rd carbon. In a single strand, the way DNA is oriented confers polarity of the strand itself. This fact becomes critical because when two DNA strands come together, they position themselves in a manner where the 5'-end on one strand connects to the 3'-end of the opposing strand. DNA strands come together in an antiparallel manner, and together with intrinsic steric properties of the DNA macromolecule, it forms the infamous double helix discovered by x-ray crystallography by Rosalind Franklin (Franklin and Gosling, 1953; Gregory et al., 2006; Mandelkern et al., 1981, 1970). The DNA has two very essential physical properties that is worth remembering, the first is that DNA can form extremely long polymers with by repeating backbone and nucleotide units along its length, the other is that DNA it's quite flexible along its length, conferring and capable of taking on coiling shapes, looping shapes, and many other conformations (Mandelkern et al., 1981).

Mechanisms in DNA function

Still to this day, we are trying to figure out the many ways in which DNA functions as a key player in cellular functions. It is well established that the role of DNA is to serve as the blueprint for transcription and translation of genes--the unit of DNA that contain the specific instructions for the formation of functional proteins (Koonin et al., 1989). For decades scientists had a very one-dimensional view of how genes were expressed.

Indeed, this is how we formulated the central dogma of molecular biology where everything starts with the instruction manual of DNA then the code of DNA is transcribed to RNA, and then the code of RNA it's translated to form amino acid—the building blocks of proteins (CRICK, 1970). However, as the fields and molecular biology and genetics progress the central dogma has become more and more appended. We now must make space for modifications found in transcription as well translation within the process of making proteins (Franks et al., 2017; Struhl, 1999). This, in large part, is due to DNA's intimate relationship with many proteins that contribute to its function. Indeed, chromatin, the marriage of DNA, polymerases, histones, transcription factors, and others contribute to the whole functional unit that make proteins synthesis happen (Dame, 2005). Chromatin plays many pivotal roles, the way DNA functions, and the way that DNA is shaped. Within a diploid human cell, we can find approximately 2 meters of DNA in ~10 μm diameter nucleus. Thus, one of the primary functions of chromatin is to the package DNA into more compact structures by use of histones and other structural proteins (Hansen, 2012; Mandelkern et al., 1981).

Chromatin conformation, does form follow function?

Now that we know that DNA can take on many shapes with the help of it's structural protein partners, does this interplay of chromatin and its ability to change shape contribute to function? In other words, does DNA form follow function? It is truly difficult to imagine tightly packed DNA and the nucleus can be so dynamic, to take on many forms. Through many decades scientists have sought to decipher shapes that DNA can take on, probably one of the most striking shapes and most people are mitotic structures that were first characterized by Ito, et.al. (Hansen, 2002; Ito et al., 1994; Lieberman-Aiden et al., 2009; Maeshima et al., 2010; Mandelkern et al., 1981). This form of DNA is, perhaps, the most condensed and it is very tightly wrapped. As molecular techniques and microscopy have improved, we now know that the spatial arrangement of chromatin in the nucleus is not

random (Nicodemi and Pombo, 2014). This includes the elusive view within interphase nuclei, which gives us insight into the conformations that take place—the part of the cell cycle a cell spends the most time within. If we take a zoomed out to a zoomed in view of chromatin organization in the nucleus, there is an established hierarchy or model that exists in the form of chromosome territories, functional transcriptional compartments, self-interacting domains, and chromatin loops of active transcription and transcription regulatory sites (Gilbert et al., 2004; Lieberman-Aiden et al., 2009; Misteli and Finn, 2021).

Methods into viewing distinct territories in the genome

In a diploid cell, there are two copies of each chromosome. Humans have 23 unique chromosomes and all somatic cells, one copy is maternal and the other copy is paternal. Viewing interphase chromosomes has been a technical challenge, even to this day. Indeed, scientists were able to view mitotic events with electron and light microscopy (Brack and Griffith, 2008; Paweletz, 2001). But the decondensed individual chromosomal structures in between cell divisions were elusive. A key experiment by Cremer and Cremer confirmed that single chromosomes occupy distinct territories within the confines of the nucleus. They did this by implementing a new technology called fluorescence *in situ* hybridization (FISH). The technique used fluorescently labeled probes complementary to the DNA sequences of specific chromosomes were used to interrogate the location of these structures in an intact chicken nucleus (Cremer and Cremer, 2001). In the field of microscopy, FISH became the workhorse when looking at genomic structures like chromosome territories. FISH together with high-resolution microscopy, it was discovered that these territories consist of multiple subdomains that occupied distinct regions in the nucleus, they can border each other closely, and they are not randomly arranged (Cremer and Cremer, 2001; Cremer et al., 2012; Langer-Safer et al., 1982).

Chromatin Conformation Capture transforms knowledge on genome organization

Since the Cremer and Cremer study in 2008, there was a ten-year lag in FISH and microscopy technologies to look at the inner workings of chromatin organization. This did not deter scientists, however, as multiple molecular techniques were developed to interrogate the structure of the genome. One such technique is chromatin conformation capture (3C). The 3C technology called HiC, has been instrumental in uncovering lower order structures of chromatin. Briefly, it works by cross linking chromatin to formaldehyde it is then digested into smaller fragments and relegated such that DNA fragments that are in close proximity to each other are covalently linked. These covalently linked DNA fragments are then sequenced, and subsequently those sequences can be aligned programmatically to the exact genomic coordinates from where they are located. In the end the readout is the locations of discrete points of the genome that are found in close proximity *in situ* (Dekker et al., 2002; Lieberman-Aiden et al., 2009). The use of HiC has corroborated the existence of chromosome territories, and it has also uncovered structures called compartments, and smaller domains called topologically associating domains (TADs) (Dixon et al., 2012). Discovery of Compartments happened through eigenvector chromatin interaction analysis that found there was higher level of compaction end specific regions of the genome. Compartments can be classified as having open chromatin states that are called A compartments, and structures having more closed or compact states are called B compartments. These distinct regions were confirmed by aligning these sequences too epigenetic markers. A compartments, for example, was found to have a high level of transcriptionally active genes that correlate with non-repressive markers such as H3K9ac and H3K27ac, high GC content, as well as hypersensitivity to DNAse I. Whereas the opposite was found in B compartment regions. B compartment regions contained a high level of silenced genes and had a high correlation with repressive histone marks such as H3K9me2, H3K9me3 (Lieberman-Aiden et al.,

2009). It was also shown that these compartments occupy non-random positions in the nucleus. A compartments had a tendency to localize in the center of the nucleus and B compartments preferentially localized to the periphery of the nucleus. Electron microscopy corroborate with these B compartment findings as it was found that heterochromatic regions of the genome localized to the nuclear lamin (Fawcett, 1966; Steensel and Belmont, 2017).

High resolution HiC and eigenvector analysis reveal novel sub-compartments

In a landmark study, these compartments were subdivided into sub-compartments. This was done in a high resolution (1 kb) HiC dataset on human diploid B-lymphoblasts, GM12878. What this group found was that each A/B compartment can be split into at least five sub-compartments defined by their long-range interaction patterns. A compartments are sub-divided into A1 and A2 sub-compartments. A1 and A2 sub-compartments are gene dense, they localize away from repressive regions of the nucleus such as the nuclear lamina and nucleolus associated domains (NADs), and they possess chromatin histone modification marks such as 3K36me3, H3K79me2, H3K27ac, and H3K4me1 associated with active gene expression. The B1 sub compartment is correlated with facultative heterochromatin where histone H3K27me3 is found and is depleted with histone H3K36me3. Subcompartments B2 and B3 do not possess the histone modification marks found on sub compartment B1. The majority (62%) of sub compartment B2 can be found in peri-centromeric regions, and it is enriched in nuclear lamin associated domains (LADs) and in NADs. Sub compartment B3 is enriched in LADs but depleted in NADs (Rao et al., 2014).

Topologically associating domains form functional neighborhoods for gene expression

Hi-C experiments revealed another level of compaction and chromatin organization that is a lower order to compartments, where it was observed that chromatin

was partitioned into self-interacting domains. These structures are known as topologically associating domains (TADs), and they occupy the genome in arbitrary size ranges from Mb to dozens of kb in genomic span (Dixon et al., 2012). TADs have been found among many mammalian genomes and thus are most likely highly conserved. TAD boundaries have been found to be enriched with CTCF genomic sites where interactions with the SMC-cohesin complex occur. In addition, it has been found that some types of genes are on the boundaries of TADs. These TAD boundaries are defined by cell type, by the loading and unloading of cohesin along the chromatin. Thus, the size of these TADs is dependent on the location of cohesin in respect to CTCF sites on chromatin, a highly dynamic process (Hughes et al., 2014; Nora et al., 2012). In contrast to this, other data suggest that the formation of TADs is random in cells that have been depleted of cohesin, this points to an alternative mechanism in which TADs can be generated through spontaneous self-interactions inherent in chromatin. It is possible that the CTCF-independent generation of TADs is connected to the location of gene transcription, as there is evidence that TAD boundaries occur near gene promotors—although only when cohesin protein is available (Lieberman-Aiden et al., 2009). TADs can be thought of as distinct neighborhoods in which the intricate process of gene expression regulation happens. Indeed, within TADs there is enrichment for enhancer and promoter interactions, as well as genes. Although the complete function of TADs is not fully understood, we can appreciate that they serve as fundamental meeting places where all the transcriptional players do gather.

Exploring the purpose of maintaining a hierarchy within the genome

Even though the genome is organized in a hierarchical manner, what would be the purpose of such organization? The answer is most apparent in the cases of diseases, or some sort of physiological dysfunction. If we assume that normal Physiology confers a consensus in chromatin organization at the level of compartments and at the level of TADs, it is possible to use the detection of these levels of nuclear organization as a

readout of disease or abnormal function. Certainly, when compared to wild type organization, the topology of a given locus may encounter events of deletion, duplication, inversions, impairment of double stranded break repair, as well as epigenetic changes, and novel insertions of viral genome (Boltsis et al., 2021; Misteli, 2010). Any one of these events would disrupt the normal pattern within the genome. In that case, it is not unconceivable that any one of these incidents could contribute to transcriptional alterations on a locus, or a genome wide fashion. Recent studies have shown that the cause of aberrant conditions or disease appears to be independent of genome conformation alteration mechanism. What this means is that genome architecture is closely intertwined with the way in which our genes are expressed and at what levels they are transcribed. One hypothesis on how genome organization is tied into abnormal transcription is how A/B compartments interact with each other (Bridger et al., 2000; Criscione et al., 2016; Xie et al., 2017). There is some evidence that A compartments and B compartments are isolated from one another in a membrane-less phase separate state. Any perturbation that happens with the any of the incidents mentioned would potentially disrupt the intricate and ordered phases of each compartment (Feric and Misteli, 2021; Feric et al., 2016). For example, some cancers show widespread epigenetic changes where A/B compartmental switching, A/B compartmental mixing, and TAD alterations are observed when compared to the compartment and TAD architecture of non-cancerous cells.

Although the recent discoveries of genome hierarchy have transformed the field of biology and how we think of the way the genome occupies space. This, with additional insight into the way it works, these findings are mostly connected through analysis of molecular techniques such as HiC. One significant limitation of these molecular techniques is that they are applied to large populations of nuclei to identify chromatin in close proximity. While this gives us reliable insight and to pooled population biology, it has

not been so reliable for features that may be found on a single cell level—this would include homolog resolved analysis in diploid cells. Therefore, it is important to have accessibility to a technology or technologies that would enable us to resolve features on a single cell level. Recently, the development of new fluorescent *in situ* hybridization (FISH) techniques as well as super-resolution microscopy has opened the doors for interrogating the structure of the genome.

Exploiting novel genome labeling and microscopy technologies to visualize DNA

The knowledge that we have gained from HiC over the past few years has been astounding. Scientists have gathered many insights into the organization of the genome through this molecular technique. HiC, however, has an inherent limitation where the observation of the genome is purely synthetic and computationally based. The Holy Grail would be the use of techniques that would enable us to observe the genome and it's architecture *in situ*. Surely, it cannot be argued that targeting specific regions of the genome and visualizing the genome has been a tall order for many decades. This is largely because there are several obstacles to overcome. First, if we were to target the genome the first set of obstacles would be physical in nature such as the cell membrane contents of the cytoplasm the nuclear membrane and heterogeneous milieu that we find in the nucleus. Second, we must probe specific sequences within the approximately 3 billion base pairs in the genome. Lastly, once your regions of interest are targeted, how does one visualize in a manner that suits your experimental questions? How large or how small are these regions of interest? For many years scientists have encountered and tried to resolve these problems. Fortunately, over the past decade there have been a few technologies that are beginning to address each one of these limitations. The first is advancement is the development of Oligopaints, and the second is the availability of super-resolution microscopy. At this period of time, we are at an inflection point where

these technologies can be combined to leverage our ability to target and observe the genome with high accuracy and high resolution.

Oligopaints: a flexible and precise way to probe DNA *in situ*

The recent implementation of the new FISH probe technology, Oligopaints, has transformed the field of genome detection in situ (Beliveau et al., 2012). Prior to Oligopaints, probe design was limited to longer sequences and less efficient single stranded DNA or RNA probes, and they did not always confer a high level of region specificity. The development of Oligopaints used a bioinformatics pipeline that simulated the thermodynamics of probe target hybridizations across the whole genome, and included added flexibility of testing for melting point temperature and GC content. The result of this simulation produced probes that were about 32-mer sequences, a length that is compatible with accurate and dense coverage of the genome. The human database for these probes is also cross-checked with multiple organisms such as organisms in the microbiome to make sure that there will be no overlaps in probe sequences. Thus, the result is a list of probes of short sequences that are designed to have complementarity to unique sequences of the genome with high specificity (Beliveau et al., 2012). Since the first iteration of Oligopaints, the technology has evolved to be even more specific targeting more regions of the genome—doing this with multiple human genome assemblies and other model organism genomes (Beliveau et al., 2018; Boettiger et al., 2016; Kishi et al., 2019; Nguyen et al., 2020; Nir et al., 2018). With the generation of probes that would have homology to the genome, there is also a generation of a database of probes/sequences that will not have any homology to the human genome. The list of these non-homology (nGHR) sequences added flexibility to Oligopaints. nGHR can be appended to the 5'-end or the 3'-end of the Oligopaints probe sequence. This gives an innate ability two bar code for potentially two different features on a probe-by-probe basis. to visualize these probes individually, fluorescent conjugates are typically used. These conjugates can either be

bound to the probe itself or they may be bound to other secondary probes with complementarity to 5' and 3' nGHR. At its highest potential, one can detect single bursts of fluorophore molecules to represent a single probe(Beliveau et al., 2012; Nir et al., 2018).

3D STORM: observing beyond the diffraction limit of light

What is the single best way to observe the genome in situ? What are the most important factors that one must consider? Especially in the case of interphase cells, the use of microscopy to visualize the genome has been elusive. This is because the signal-to-noise ratio/contrast has been low, our targeting technologies has been weak, and our resolution less than ideal. Historically the best method that was used to image the genome what is the use of electron microscopy (EM) (Dehghani et al., 2005; Golomb and Bahr, 1974). Although, EM had the main benefit of absolute resolution, it is limited and that I cannot select specific regions of interest to target and visualize. Recent work in the electron microscopy field, however, has made some headway and being able to isolate nuclear ultra-structures within their fields of view. The recent development of ChromEMT Is a good example of this (Ou et al., 2017). In addition, there has been some headway on the benchtop where fluorescent microscopy and electron microscopy are performed in tandem, opening exciting new ways to visualize biology. While it doesn't have the resolution of electron microscopy, fluorescence microscopy has evolved to have the capability to image at a resolution beyond the diffraction limit of light. The development of super-resolution microscopes has transformed the way we can visualize biology with high signal-to-noise and with high resolution (Diezmann et al., 2017; Rust et al., 2006; Vangindertael et al., 2018). In conjunction with microscope evolution, there has been great headway in the field of fluorophore chemistry, together this has made super-resolution possible (Dempsey et al., 2011; Kozma and Kele, 2018). Although there are many versions of super-resolution microscopy, the use of Stochastic Optical Reconstruction Microscopy (STORM) has gained traction in the genomic imaging field. 3D STORM is an

imaging platform that allows for the detection of the location of single molecules (SML) in three-dimensional space. Briefly, the way it works is by the detection of the activated state of fluorescence and photo-switchable molecules. This must be done in a manner that ensures sufficient number of photons to be omitted two and able a precise localization on that molecule before it becomes deactivated and goes into a dark state or is photobleached. Thus, there is a subset of fluorophore chemistries that will work in STORM. In addition, these activated fluorescent molecules must be in a distance that exceeds the diffraction limit of light (~250 nm) In order to have concurrent recording of many individual molecules. To overcome the diffraction limit of light, these fluorophore emitters have a unique quality that enables them to emit light in a spontaneous or stochastic manner. These bursts of emissions are recorded as point spread functions (PSF, the spread of light) of individual molecules with high precision. Next, the positions of the PSF molecular centroids are recorded in the x, y, z coordinates. The centroid position of each identified molecule is statistically fitted to a Gaussian function Where the precision is correlated to the number of detected photons from single molecules. Therefore, the precision increases with respect with the number of photons captured. The coordinate location of each molecule recorded in a table format which can be used to generate a reconstructed image from these coordinates (Bates et al., 2005; Bock et al., 2007; Dempsey et al., 2011; Heilemann et al., 2008; Hess et al., 2006; Lee et al., 2010; Rust et al., 2006; Vogelsang et al., 2009).

Oligopaints and 3D STORM: two forces combined to make *in situ* observation of the genome hierarchy achievable

Because of the added flexibility the barcoded nGHR provides on Oligopaints, several laboratories have exploited this unique feature so that multiplexing can be incorporated into imaging experiments. Typically, this involves the use of either a pneumatic or a peristaltic microfluidics system end which multiple reservoirs can be used

2 sequentially hybridize new probes, implement washes, all done by simple programming. For example, Nir, et.al. was able to trace an 8 megabase locus of human chromosome 19. the eight megabase region was split into 9 segments, where each individual segment can be targeted and imaged one at a time via a fluorophore-conjugated secondary oligo that has complementarity to and binds the 5'-end nGHR. This same segment can be further divided into several sub-regions where each individual sub-region is sequence barcoded differentially with the 3'-end nGHR. Thus, one can perform several rounds of hybridizations and microscopy acquisitions to target any region that is sequence barcoded on either 5'-end or 3'-end nGHR. By using this method Nir, et.al. was able to walk along the eight megabase region with an output that is in the nanoscale resolution. This landmark study was the first of its kind to display a contiguous trace at that scale (8 Mb) of two human chromosome homologs using sequential Oligopaints and 3D STORM (Beliveau et al., 2018; Nir et al., 2018).

Experimental rationale

The invention of Oligopaints and super-resolution microscopy has opened the door for exploration of the intricacies of genome organization. For the first time history, we can visualize and scrutinize the 3-dimensional makeup of the genome given the high resolution and coverage possible with these two technologies. Thus, one of the questions that we wanted to explore with the study detailed in this document was, how far can we push the technical limits of these methods? Specifically, with the direct goals of improving resolution, increasing total time in image acquisition, and targeting larger overall size of genomic loci. Prior to the study covered in this document, one of the most comprehensive publications in this field came from Nir, et. al. 2018, by using only OligoSTORM to peer into genome organization. In this study, they were able to interrogate several segments in ~8.0 Mb region of human chromosome 19 over the course of multiple rounds of automated microfluidics and OligoSTORM. This landmark study was a proof-of-concept

demonstration that components of genome organization can be isolated and visualized at a resolution greater than the diffraction limit of light—giving us insight into the structural features of lower-order portions of the genome, such as, topologically associating domains. To add to pushing the limits from a technical perspective, we must consider the infrastructure needed to undertake such data intensive procedures. The ultimate goal in the field is to image and record the 3-dimensional genome using super-resolution microscopy, and improvements are necessary in the aspects of data processing speeds, data analysis, and data storage capacity in order to make this "moonshot" possible.

Another open area for us to exploit is to observe key genome organization levels; from higher-order to lower-order, such as chromosome territories, functional compartments, sub-compartments, and topologically associating domains; imaged concurrently to assess their spatial relationship in respect to each other. We aimed to do this using super-resolution, as this is the only way in which we can observe all these features. Indeed, the question remains, what does the genome look like with the context of whole chromosome territory, compartment/sub-compartment, and topologically associating domain context.

Given the accessible resolution demonstrated by OligoSTORM, we sought to compare high resolution HiC of a cell type from a pooled analysis to that of single cell display from super-resolution microscopy—how reflective is pooled HiC data to single cell super-resolution microscopy?

Chapter 2: Materials and Methods

2.1 Manual curation of topologically associating domains from a HiC dataset

Genomic coordinates of GM12878 chromosome 9 (Chr 9) topologically associating domains (TAD) were manually curated using version 1.9.8 of Juicebox: Visualization software for HiC data. The final genomic coordinates recorded reflect the most resolute TADs using a balanced normalization in a viewing window of 1.0 kb, 10-15 kb or 25 kb. As such, recorded coordinates are regions that span the genomic length of a single topologically associating domain (TAD) at the highest resolution viewing window. In addition, all topologically-associating domains were assigned to sub-compartments (A1, A2, B1, B2, B3) as defined by Rao, et al (Rao et al., 2014; Robinson et al., 2018).

2.2 Mining of Oligopaints on human chromosome 9

OligoMiner was used to sort out a balanced selection of *in situ* hybridization oligonucleotide probes (Oligopaints) from a repeat-masked hg19 reference genome. These Oligopaints (GHR) of 35-42-mer of single-stranded DNA were qualified for their homology to genomic regions of curated TAD coordinates recorded from GM12878 Chr 9 HiC, at a coverage density of 2.2-10.3 probes $kb^{-1}TAD^{-1}$. Thus, an Oligopaints library was designed programmatically to extract the probes that intersect only our TAD genomic coordinate regions from manual curation. All oligopaints probes were then appended with non-genomic homology sequences that flanked the GHR on 5' (Mainstreet) and 3' (Backstreet) ends. The goal was to use Mainstreets and Backstreets as nucleotide barcodes to specifically target regions of interest, i.e. individual TADs or all TADs in a functional sub-compartment. These appended streets, together with a universal library forward priming index sequence (terminal 5'-end) and universal reverse priming index sequence (terminal 3'-end), were assembled with OligoLego where each probe is composed, from 5’ to 3’: universal forward priming index (FPR); 5’-non-genome-homology nucleotide barcode region (Mainstreet); GHR; 3’-non-genome-homology nucleotide barcode region (Backstreet); universal reverse priming index. Each Oligopaints probe also

included sites for a “toehold” site that facilitates multi-cycle and multi-directed imaging using automated microfluidics (Beliveau et al., 2018).

2.3 Design of an Oligopaints library targeting multiple regions of human chromosome 9

An Oligopaints library of 88,666 probes was synthesized by Twist Biosciences Oligo Pools service (San Francisco, CA) and were generated with the following specifications: genome homology regions (GHR) were designed to uniquely target and tile all regions of interest (ROI) in the library. Mainstreets are nucleotide barcoded for specific targeting of the pre-selected 80 individual TAD ROI. Backstreets confer specific targeting of the first TAD of segments 1-8 and the first TAD of segment 9 (BrT), these TADs are known as segment representatives. These segments, consisting of multiple sequential TADs each, were selected to isolate approximate genomic equidistant TAD clusters spanning the length, *p*-arm to *q*-arm, of chromosome 9. Minus the probes of segment representative TADs, all other Oligopaints probes were appended with Backstreet barcodes to identify their sub-compartment assignment on an individual sub-compartment basis. Given this, analysis of BrT Backstreet targets was performed using a multiple step qualification process via image co-localization and density-based clustering to confirm its sub-compartment assignment.

2.4 Oligopaints library synthesis and amplification to a working concentration

The Oligopaints library was synthesized by Twist Biosciences (San Francisco, CA) as a pool of a non-working stock solution. The amplification of our sub-libraries begins by validating and optimizing the concentrations of template and primers (Integrated DNA Technologies, Coralville, IA) using real-time polymerase chain reaction (PCR). After optimization, we performed large scale PCR to amplify our single-stranded template into a double-stranded product. Subsequent to PCR, product was purified using a DNA Clean and Concentrator-25 kit (Zymo Research, Irvine, CA). This purified product served as a

template to for *in vitro* transcription of our product using a T7 RNA polymerase reaction (targeting a T7 promotor sequence overhang from the reverse primer) (New England Biosciences E2040S). The single-stranded (ss) RNA product from this reaction is then reverse transcribed (ThermoFisher Scientific EP0752). Single-stranded is generated from the reverse transcription reaction and RNA contamination is digested by a mixture of (1:1) 0.25M ethylenediaminetetraacetic acid (EDTA) and 0.5M sodium hydroxide (NaOH) in 95°C high heat for 10.0 minutes. Finally, the ssDNA product is concentrated and purified using the DNA Clean and Concentrator-100 kit using a oligo binding solution for selective single-stranded clean-up. As a last validation step, we performed DNA gel electrophoresis to assess DNA band size to our expected probe length (Beliveau et al., 2012, 2018; Nir et al., 2018).

2.5 Cell culture of GM12878 B-lymphocytes

GM12878 cells were handled and maintained in accordance with the standard operating procedures formulated by the 4D Nucleome Consortium (National Institutes of Health). GM12878 cell cultures were grown in 5.0% CO_2 and atmospheric O_2 at 37.0°C to a confluency of ~2.0 million cells/mL in Roswell Park Memorial Institute 1640 (Corning 15-040-CM) media suspension conditioned with 2.0 mM L-glutamine (Gibco 25030081) and 15.0% fetal bovine serum (Seradigm 1500-500 lot 035B15). Cells were grown and expanded as needed in conditioned media.

2.6 Cell and cell substrate preparation

#1.5 40.0 mm german borosilicate glass coverslips (Bioptechs, Butler, PA) were cleaned and sterilized by ultrasonication (Branson 2800) in series with Branson solution, ddH_2O, and ethanol for 10.0 min each; and finally coated with 0.01% (w/v) poly-l-lysine (Millipore Sigma P8920). GM12878 cells were seeded on poly-l-lysine coated glass coverslips with conditioned media for 1.0 hr in 5.0% CO_2 and atmospheric O_2 at 37.0°C. Coverslips were then gently washed in 37.0°C phosphate buffer saline (PBS) and fixed in

4.0% paraformaldehyde (Electron Microscopy Sciences 15710) for 10.0 min at room temperature (RT). Formaldehyde polymerization reaction was quenched with 25.0 mM sodium borohydride (Millipore Sigma 452882) for 7.0 min at RT. The cells were then permeabilized in 0.5% (v/v) Triton-X100 (Thermo Scientific 85112) in PBS for 10.0 min and subsequently in 0.1 N HCl for 5.0 min both at RT (Nir et al., 2018).

2.7 3-dimensional DNA fluorescent *in situ* hybridization

Fixed and permeabilized samples were prepared for 3-dimensional DNA fluorescent *in situ* hybridization (3D DNA FISH) by washing in 0.1% (v/v) Tween-20 (Thermo Scientific J20605AP) 2X saline sodium citrate (Sigma Aldrich S6639) solution (2X SSC-T) at RT for 2.0 min two times. Unless stated otherwise, our sample of seeded GM12878 cells remained hydrated with buffer solution to aid in preserving the 3D integrity of chromatin. Next, we incubated in a 50.0% (v/v) Formamide (Millipore Sigma 344206) 2X SSC-T used to facilitate sample DNA denaturation for 5.0 min at RT followed by 20.0 min at 60.0°C (pre-conditioned). For each 40.0 mm coverslip, 35.0 µL of primary hybridization mixture was made by combining 50% formamide, 25% of 40% Dextran sulfate in 8X SSC-T, 1.4 µL RNase A (10.0 mg/mL, Thermo Scientific EN0531), 100.0 pmol Oligopaints probes (2.85 µM probe/1.5 Mb target), volume ddH_2O to 35.0 µL) is deposited on sample, sandwiched with a clean 22.0 mm X 30.0 mm glass coverslip, and edges sealed with rubber cement to prevent sample dehydration. The samples were immediately denatured for 3.0 min at 80.0°C and held at 47.0°C for ~48.0 hrs. After this primary hybridization incubation, samples were washed in pre-heated 2X SSC-T for 10.0 min at 60.0°C and repeated four times. Coverslips were finally washed and held in PBS before time of automated microfluidics and image acquisition preparation (Nir et al., 2018).

2.8 Preparation of microfluidics and OligoSTORM image acquisition

We used an integrated microfluidics (Bruker, Billerica, MA) and single molecule localization nanoscope (Bruker Vutara 352) to perform all image acquisitions. A 1:600

dilution of 90 nm gold nano-urchin beads (Millipore Sigma 797707):1X PBS was mixed and 150 µL of the mixture was deposited on each sample coverslip. The mixture was spread on the sample by sandwiching with another 40.0 mm glass coverslip. Each sample-containing coverslip was embedded with the gold nano-urchin beads by centrifugation at 600xg for 3.0 min. These gold nano-urchin beads now embedded onto the sample serve as fiducial markers for drift correction because of extended time in 3D STORM acquisition. After centrifugation, the sandwiched coverslips were separated, and the sample containing coverslips are stored in 1X PBS (Nir et al., 2018).

2.9 Diffraction limited image acquisition test of Oligopaints library

For validation of the TAD library, GM12878 cells were seeded on Ibidi µ-Slide VI 0.5 glass bottom slides (Ibidi 80607) and 3D DNA FISH was performed as described. Multi-stack fluorescence images were acquired using a Nikon Eclipse Ti2 Series inverted microscope. N=25 nuclei imaged for BrT backstreet target and N=22 nuclei imaged for TAD-1 target. Image processing and segmentation was performed using Imaris software suite (Oxford Instruments).

2.10 Multistep OligoSTORM for library validation

Multistep STORM was done using the 3D biplane Vutara 352 single molecule localization system from Bruker Nano equipped with a 60X silicon objective (Olympus Lifesciences) with a 1.3 numerical aperture (NA), and Bruker Microfluidics Accessory. This objective is not susceptible to spherical aberrations known to happen with other objectives of this type, making it suitable for interrogating the genome in 3D. For 3D STORM acquisition, our Oligopaints were detected with Alexa Fluor 647 tags and Alexa Fluor 405 for photoactivation. Detection of the fluorescent signal was done via a sCMOS camera (Hamamatsu Orca-flash 4.0). Regions of interest were capture using a 10.0 ms exposure time within a 40.0 µm x 40.0 µm x 11.0 µm (x,y,z) field of view (FOV). A total of 5 nuclei in 2 FOVs were imaged using 100 nm z-steps. Nuclei were selected based on their ploidy,

where only interphase diploid nuclei were chosen. For future reference, cells were numbered according to their x,y position in the FOV—the FOV was split into 4 quadrants where cells were identified numerically in the following manner: upper left, upper right, lower left, lower right.

2.11 Automated microfluidics and 3D STORM: setup and implementation

After primary hybridization and gold nano-urchin beads were embedded, the coverslip is mounted on the microfluidics chamber (Bioptechs FCS2) and microfluidics lines are connected to the Bruker Microfluidics Accessory. 2X SSC-Tween 20 (0.1%) buffer is pumped into the chamber displacing any air bubbles that may have formed. The small amount of detergent in our buffer also aids in preventing future bubbles from forming during the acquisition steps. Finally, the chamber with the sample is mounted on the microscope stage. The microfluidics and image acquisition cycles for OligoSTORM is programmed in Vutara SRX (Bruker's proprietary software) in the following fashion (flow rate=120.0 µL/min., and room temperature conditions, unless specified otherwise): (1) 550.0 µL of 2X SSC-T held for 3.0 min. (2) 700.0 µL hybridization mixture 1 (50.0 pmol backstreet bridge BrT, 50.0 pmol AlexaFluor647-conjugated secondary oligos, 30% formamide in 2X SSC-T) held for 45.0 min. (3) 550.0 µL of imaging washing buffer (40% formamide in 2X SSC-T) held for 3.0 min. and repeated once. (4) 800.0 µL of STORM/switching imaging buffer containing 10% (w/v) glucose, 2x SSC, 50 mM Tris, 1% (v/v) β-mercaptoethanol, and 2% (v/v) of a GLOX stock solution consisting of 75 mg/mL glucose oxidase (Sigma Aldrich G2133250KU), 7.5 mg/mL catalase (Sigma C40-500MG), 30 mM Tris, and 30 mM NaCl, flowed in the chamber twice and held for 5.0 min. so all fluid movement stops (5) FISH signal assessed using fluorescent widefield. (6) Fields of view (FOV) and their x, y, z-range coordinates are recorded, as well as a widefield 3D fluorescent image, stored as a reference for each FOV. (7) OligoSTORM single molecule localization (SML) begins (640 nm laser capturing 500 frames at 100 frames/sec. and 10.0

ms exposure time) with coordinates of FOV 1, capturing the whole range in the z-axis. After FOV 1 is captured to completion, the scope objective moves to FOV 2 and repeats the 3D STORM cycle. (8) Steps 1-7 are repeated for each additional target. SML of a total of ten targets were imaged: (1) backstreet bridge (BS) BrT (2) mainstreet bridge (MS) TAD-1 (3) MS TAD-59 (4) TAD-95 (5) TAD-131 (6) TAD-191 (7) TAD-249 (8) TAD-300 (9) TAD-363 (10) TAD-431.

2.12 Pre-analysis filtration and drift correction of SML data

The particle file for single molecule localization (SML) data from OligoSTORM acquisition was filtered where all particles fit ≤50.0 nm axial precision. Sample drift was corrected in x, y, z using at least 3 gold nano-urchin fiducials per field of view (FOV) using 15 time interval setting on Vutara SRX program. Fiducials for each round of microfluidics/imaging were qualified using low-tolerance density-based spatial clustering of applications with noise (DBSCAN) of a maximum particle distance of 30 nm and a minimum 500 particles to form a single cluster, where each cluster represents a single round of imaging.

2.13 Density-based analysis of single molecule localization data

We used Bruker's own SRX software to assess the spatial overlap and clustering of our single molecule localization (SML) data. First, we perform localization filtering where we disqualify all localizations under an axial precision of 50 nm. Drift correction is next performed using a center-of-mass function of fiducial clusters that are segmented using high stringency of density-based spatial clustering of applications with noise (DBSCAN). Qualification of fiducial localizations using high stringency of DBSCAN is applied so that the drift is corrected with very high accuracy. After drift correction, we applied DBSCAN again to filter only high-quality localization clusters. We next applied alpha surfaces to our clusters using an alpha radius of 100 nm, a global target distance setting (applied to all ROI) that is tethered to the maximum particle distance using in DBSCAN. Generation of

these alpha shape surfaces was essential to determine the spatial overlaps between more than one cluster. For example, we were able to determine quantitatively the overlap between DBSCAN qualified cluster(s) from round 1 of acquisition to round 2 of acquisition.

2.14 Sample preparation and strategy for OligoSTORM of 86 genomic targets

GM12878 cells underwent the same preparations in terms of culturing conditions, sample preparations, and 3D DNA FISH detailed in sections 2.10-2.13 of this document. Once sample preparations were complete, automated microfluidics protocol was programmed into SRX for sequential hybridizations of 87 steps/rounds to include flow of hybridization mixes, washing buffers, and STORM imaging buffer in a manner and sequence detailed in section 2.3.9. Fields of view (FOV) were selected based on the criteria of robust FISH signal from all cells, nuclei had to present no visible aneuploidy, maximum number of cells in super-resolution FOV (40 μm x 40 μm). The first step of the sequential hybridization targeted the universal priming index regions with probes tagged to Alexa Fluor 568, and the BrT Backstreet barcode with probes tagged to Alexa Fluor 647 (AF647). The diffraction limited FISH signal for BrT was used to assess the quality of the cells on the sample coverslip, and to select all the FOVs. Next the z-range was determined by the Alexa Fluor 568 FISH signal for the priming index sequences. This step was significant because we wanted to make certain the whole contents of our Oligopaints library had the potential to be imaged in OligoSTORM—capturing the whole potential z-range of chromosome 9. Both diffraction limited BrT and priming index sequence FISH signals had the addition benefit of serving as region references, later used in ROI selection and confirmation steps. In all, six FOVs were selected, each containing 5-6 nuclei, utilizing the same microscope and microfluidics system settings detailed in section 2.3.9. Our 3D STORM system utilized single molecule detection of AF647 in all rounds of imaging, where AF647 is photoactivated in stoichiometric 1:1 ratio by a 405 nm-oligo-dye. The sequential OligoSTORM rounds included: round 1 targeted BrT Backstreet barcode, rounds 2-6

targeted a single sub-compartment Backstreet barcode in each round (B1>A1>B2>A2>B3). Image acquisition was terminated at round 23, where rounds 7-23 targeted 3 individual TADs at a time, each belonging to a unique sub-compartment using their Mainstreet barcode. A maximum of all five sub-compartments and 52 of 80 TADs were acquired per chromosome 9 territory in each nucleus.

2.15 Particle localization filtration and drift correction of 52 TAD multiplexed dataset

After 23 rounds of 3D STORM image acquisition, we encountered a large amount of noise in our localization data. For this reason, we needed to establish a novel protocol to isolate our nuclei of interest and to segment our signal with the least background possible. As a first step, we filtered all localizations that fit an axial precision of <50 nm. Doing this aided in cutting out about 1/3 of total localizations. We next located all nuclei that were imaged and numbered each nucleus according to their quadrant position. We then selected nuclei that were near three or more fiducial gold nanourchin beads. The reason for doing this is because it was observed that the bead drift pattern was not all the same, even in the same field of view. The opposite observation was made of beads that were in close proximity to each other and to their respective nuclei. A possible cause of this may be related to the position of nuclei in the FOV contributing to altered fluidics flow patterns. We cropped each nucleus that met the qualification for drift correction, so that it was less taxing on the workstation gpu and cpu. A drift correction setting of 15 intervals from the volumetric centroids of the fiducial beads where the tract of 23 beads is clearly visible. This process is repeated for each nucleus that fits in the fiducial bead location and visibility guidelines.

2.16 Decoding of regions of interest using density-based spatial clustering of applications

The decoding of our regions of interest (ROI) involved multiple layers of criterion steps. The first step was to apply a density-based spatial clustering of application

(DBSCAN) analysis to rule in particles that were no further than 150 nm and contained at least 100 localizations to form a cluster. We also applied alpha shape surfaces on clusters using a 150 nm setting, reflecting DBSCAN analysis parameters. Next, we selected the most high-quality (clusters with high particle counts) super-resolution clusters that formed an overlay with the widefield reference BrT signal. This overlay method proved to be a reliable way to call the 9 TAD representatives of each segment per homolog. Together with the BrT clusters and the widefield reference of the universal priming index signal, we were able to spatially populate rounds that targeted sub-compartment TADs. Building off the known positions of BrT and clusters for TADs in each sub-compartment, we were able to use those as reference points to qualify individual TADs that were imaged as a multiplex of three TADs per round. This manual process of decoding was performed on 12 nuclei. Each nucleus underwent this multistep process of cluster selection three times as to limit manual selection bias.

2.17 Analysis of 3D structural TAD features

Each individual TAD was examined for possible 3D physical feature correlations. These physical features included volume, surface area, compaction factor, sphericity, radius of gyration. To do this we applied a student's T-test to the comparing the differences in means of each physical feature in a homolog specific manner. Variances between cell to cell was also assessed. For sub-compartment statistics, a multi-variable one-way ANOVA was used on all collective data for all cells analyzed.

2.18 Analysis of human chromosome 9 pairing

All called TADs underwent manual inspection on an individual basis and on a collective basis. This gave us insight and context into the shapes and spatial distributions of matched TADs across homologs. First, we determined how close both homologs of chromosome 9 were to each other by visual inspection, and application of cluster co-localization on SRX to assess how truly close observed TADs were to each

other. TADs were ruled to be paired based on a 500 nm threshold. Paired TADs were recorded and plotted with frequency of their pairing occurrences.

Chapter 3: Results

3.1 Manual curation of a high resolution HiC dataset yields 431 domains spanning whole human chromosome 9

To date several studies have shown with super-resolution microscopy the interaction of chromatin on the level of compartment domains and consecutive domains (TADs). Very little is known, however, of shortrange interactions like we observe in TADs, with other characterized long range hierarchical structures such as sub-compartments and whole chromosome territories—and how they appear and interact with each other in 3-dimensional (3D) space in nano-scale resolution using microscopy. Our goal in this study was to explore the possibility of assessing the conformation status of the genome at the levels of TADs, sub-compartment, and chromosome territories at an unprecedented scale in situ. To do this we employed the use of two landmark technologies, Oligopaints and 3D stochastic optical reconstruction microscopy (STORM). Together, these technologies allowed for a flexible experimental design and execution technique to target and visualize the genome that would be representative of the genome's 3D scale in x, y, z planes.

3.2 Selection of domains and sub-compartments of human chromosome 9

In this work, we combined a stepwise and multiplexed approach using 3D DNA FISH, microfluidics, and STORM (collectively known as OligoSTORM) to target sequential TADs and groups of TADs spanning from *p*-arm to *q*-arm of human chromosome 9 (Chr 9). Due to the spatial resolution necessary to employ OligoSTORM it was essential that an Oligopaints library design be at the highest resolution possible. To do this, we used the *in situ* HiC dataset from Roa, et al. 2014, which gave us a maximum contact frequency resolution of 1.0-kb of human B-lymphoblast cell type, GM12878. Since there is no standardized way of calling TADs using a HiC dataset of this type, we elected to use the JavaScript program Juicebox.js to manually curate the genomic coordinates of TADs along the whole span of human Chr 9. Chromosome 9 was selected for based on its relatively middling size, and for not being enriched or depleted of genes. The size range

of TADs vary greatly, and there is little consensus from previous studies to point to exact size. Since we worked with a high-resolution dataset, we called TADs based on and outward-in method where the genomic coordinates are recorded starting from what can be observed from a 10- to 15-kb resolution on Juicebox, which yields TADs in Mb size range (**Figure 1a**). This locus is re-assessed at the 5-kb level to yield TADs in the 100- to 500-kb size range (**Figure 1b**). A last assessment was done at 1-kb resolution to call TADs at the highest level when observed, to yield TAD size of 20- to 80-kb (**Figure 1c**). Called TADs were aligned with genomic coordinates of sub-compartments according to the same study to reflect their epigenetic status. Altogether, 431 TADs with size ranges of their sub-compartmental addresses were recorded spanning the length of chr 9 (Rao et al., 2014; Robinson et al., 2018).

3.3 Mining for Oligopaints probes to interrogate the hierarchical organization of human chromosome 9

Considering the technical and financial constraints of sequentially imaging 431 TADs in super-resolution, we decided to subset the total Oligopaints library into a more feasible 80 TADs. We programmatically selected Oligopaints probes through an OligoMiner hg19 (same genome as Rao, et.al. study) probe database and appended these probes with unique barcodes using OligoLego. TADs qualified in this library were between 40-kb to 1.2-Mb in size, with sub-compartment distribution of 12 TADs in B1, 8 TADs in B2, 6 TADs in B3, 9 TADs in A1, and 32 TADs in A2 (**Table 1**). As previously stated, this 80-TAD Oligopaints library contained a pool of 88,666 unique probes that are appended with unique nucleotide barcodes on the 5'-end (Mainstreet) and the 3'-end (Backstreet) to reflect specific sub-regions that can be targeted on an individual-TAD manner (**Figure 2a**).

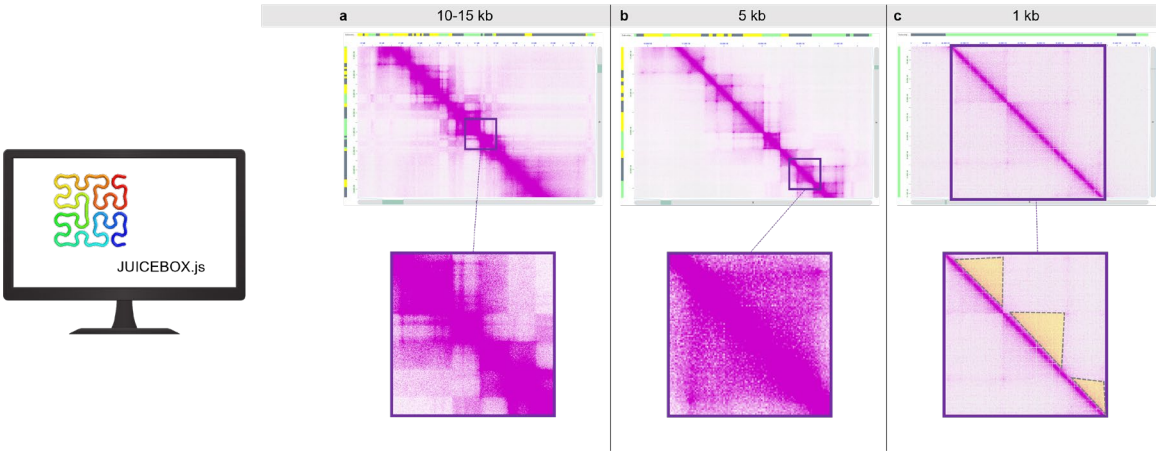


Figure 1|Manual curation of topologically associating domains (TAD) using HiC analysis tool Juicebox.

The java script software analysis tool Juicebox was used to visually call TADs using the *in situ* HiC dataset of human B-lymphoblast GM12878 from Rao, et al. 2014. TADs were called at the highest resolution possible. Regions of enrichment where chromatin self-interactions appear on three levels of resolution, aligned to sub-compartment placement. **a,** 10-15 kb resolution is where domains can be observed in the megabase-scale (insert), however, toggling to the next highest resolution. **b,** 5 kb resolution is where domains are observed on the 100s kb-scale (insert). In this locus the genomic coordinates of three individual TADs. **c,** (shaded yellow) all belonging to sub-compartment A2 were called and recorded. A total of 431 TADs from whole chromosome 9 were called using this method.

TAD number	Genomic size (kb)	Number of probes	Density (Probes/kb)	Segment number	Mainstreet (MS) unique barcode	Backstreet (BS) unique barcode
TAD1	71	443	6	1	MS-1	BS-BrT
TAD2	198	1663	8		MS-2	BS-B1
TAD3	181	1352	7		MS-3	BS-B1
TAD4	79	493	6		MS-4	BS-B1
TAD5B1	105	1024	10		MS-5	BS-B1
TAD6B1	428	2993	7		MS-6	BS-B1
TAD7A2	158	791	5		MS-7	BS-A2
TAD8B1	63	336	5		MS-8	BS-B1
TAD9A2	169	968	6		MS-9	BS-A2
TAD10A2	65	389	6		MS-10	BS-A2
TAD11A2	67	525	8		MS-11	BS-A2
TAD12A2	70	670	10		MS-12	BS-A2
TAD59B2	940	7389	8	2	MS-13	BS-BrT
TAD60B2	343	801	2		MS-14	BS-B2
TAD61B3	271	1730	6		MS-15	BS-B3
TAD62B2	402	2746	7		MS-16	BS-B2
TAD63B2	773	5304	7		MS-17	BS-B2
TAD64A2	126	483	4		MS-18	BS-A2
TAD65A2	40	124	3		MS-19	BS-A2
TAD66A2	139	336	2		MS-20	BS-A2
TAD67A2	81	440	5		MS-21	BS-A2
TAD68B3	453	2825	6		MS-22	BS-B3
TAD69B3	423	2874	7		MS-23	BS-B3
TAD70A2	259	1546	6		MS-24	BS-A2
TAD95B3	1220	3297	3	3	MS-25	BS-BrT
TAD96B3	682	2412	4		MS-26	BS-B3
TAD97B3	517	1884	4		MS-27	BS-B3
TAD98A2	199	1248	6		MS-28	BS-A2
TAD99B2	376	2184	6		MS-29	BS-B2
TAD100A2	75	409	5		MS-30	BS-A2
TAD101A1	185	1430	8		MS-31	BS-A1
TAD102A1	107	623	6		MS-32	BS-A1
TAD103A1	142	800	6		MS-33	BS-A1
TAD104B1	105	229	2		MS-34	BS-B1
TAD105A1	97	261	3		MS-35	BS-A1
TAD141A2	104	575	6	4	MS-36	BS-BrT
TAD142A2	68	468	7		MS-37	BS-A2
TAD143A2	76	583	8		MS-38	BS-A2
TAD144A2	66	461	7		MS-39	BS-A2
TAD145A2	142	866	6		MS-40	BS-A2
TAD146B2	210	1372	7		MS-41	BS-B2
TAD147B2	132	1073	8		MS-42	BS-B2
TAD148B2	135	853	6		MS-43	BS-B2
TAD149B3	259	1131	4		MS-44	BS-B3
TAD150B2	147	1194	8		MS-45	BS-B2
TAD191B2	312	2016	6	5	MS-46	BS-BrT
TAD192B2	211	1670	8		MS-47	BS-B2
TAD193A2	90	505	6		MS-48	BS-A2
TAD194A2	105	250	2		MS-49	BS-A2
TAD195A2	78	71	1		MS-50	BS-A2
TAD196A2	50	219	4		MS-51	BS-A2
TAD197A2	273	1306	5		MS-52	BS-A2
TAD198A2	67	674	10		MS-53	BS-A2
TAD199A2	59	280	5		MS-54	BS-A2
TAD249B1	336	1638	5	6	MS-55	BS-BrT
TAD250A2	164	491	3		MS-56	BS-A2
TAD251A2	65	196	3		MS-57	BS-A2
TAD252B1	84	10	0		MS-58	BS-B1
TAD253B1	273	550	2		MS-59	BS-B1
TAD254A1	102	840	8		MS-60	BS-A1
TAD255A1	202	1538	8		MS-61	BS-A1
TAD256A1	66	447	7		MS-62	BS-A1
TAD300A2	69	377	5	7	MS-63	BS-BrT
TAD301B1	46	275	6		MS-64	BS-B1
TAD302B1	113	555	5		MS-65	BS-B1
TAD303A2	75	374	5		MS-66	BS-A2
TAD304A2	666	3226	5		MS-67	BS-A2
TAD305A2	163	767	5		MS-68	BS-A2
TAD306A2	144	593	4		MS-69	BS-A2
TAD363A2	110	1129	10	8	MS-70	BS-BrT
TAD364A1	109	826	8		MS-71	BS-A1
TAD365A1	78	496	6		MS-72	BS-A1
TAD366A1	97	550	6		MS-73	BS-A1
TAD367A1	52	304	6		MS-74	BS-A1
TAD368A1	82	475	6		MS-75	BS-A1
TAD427A1	148	962	7	9	MS-76	BS-A1
TAD428A1	122	476	4		MS-77	BS-A1
TAD429A1	129	597	5		MS-78	BS-A1
TAD430A1	162	919	6		MS-79	BS-A1
TAD431B1	360	1540	4		MS-80	BS-BrT

Table 1|80 TADs are selected for an Oligopaints library.

Individual TADs are colored according to their sub-compartment address; sub-compartment A1, blue; A2, aqua; B1, red; B2, dark red, B3, brown. Mainstreet (MS) unique barcodes are assigned to each individual TAD, green; and backstreet (BS) unique barcodes are assigned to the first TAD of segments 1-8 and the last TAD of segment 9, purple; all remaining TADs are uniquely barcoded on BS to reflect the sub-compartment they belong, color-coded same as above.

In addition, groups of sequential TADs are segmented into 9 parts. Thus, individual TADs can be viewed individually through the Mainstreet (MS) barcode, and each TAD sub-compartment assignment or the first TAD in segments 1-8 and last TAD in segment 9 can be targeted by the Backstreet (BS) barcode (**Figure 2b**). The library pool of oligos goes through a rigorous process of in vitro amplification and purification steps that have been previously described.

3.4 Oligopaints library validation by diffraction-limited fluorescence microscopy

As a first step, it was necessary to validate the Oligopaints library by 3D FISH. Resources were used to culture GM12878 human diploid B-lymphocytes. These naturally spherical-shaped cells were seeded onto cover glass to take on a more convex morphology—including the nucleus. The cover glass with seeded GM12878 underwent a 3-dimensional fluorescent *in situ* hybridization (3D FISH) procedure so not to alter the spatial integrity of the chromatin structure. Using fluorescent widefield microscopy, we first targeted BS for BrT. We were able to visualize a robust signal-to-noise ratio (SNR) counting a median of 9 foci per cell (n=25) using post-acquisition segmentation, spanning most of the spatial breath of the nucleus (**Figure 3a, 3c**). TAD-1 in sub-compartment B1 (SC-B1) was also selected as a target where we were able to detect a median of 2 foci per cell (n=22) after segmentation (**Figure 3b, 3c**). Based on these observations we can firstly concluded that our parameters and SNR for 3D FISH are optimal to proceed to 3D STORM; second, diffraction-limited microscopy is not sufficient to detect all 9 segments of our TAD Oligopaints library. For 2 homologs, we expected a total of 18 foci per cell to be visualized. Third, a single TAD (TAD-1), belonging to the smallest third in genomic size and average Oligopaints probe density, can be individually observed. This indicated that image acquisition of a single TAD of this composition in 3D STORM is plausible, giving us confidence to pursue sequential acquisition of TADs (**Figure 3a-c**).

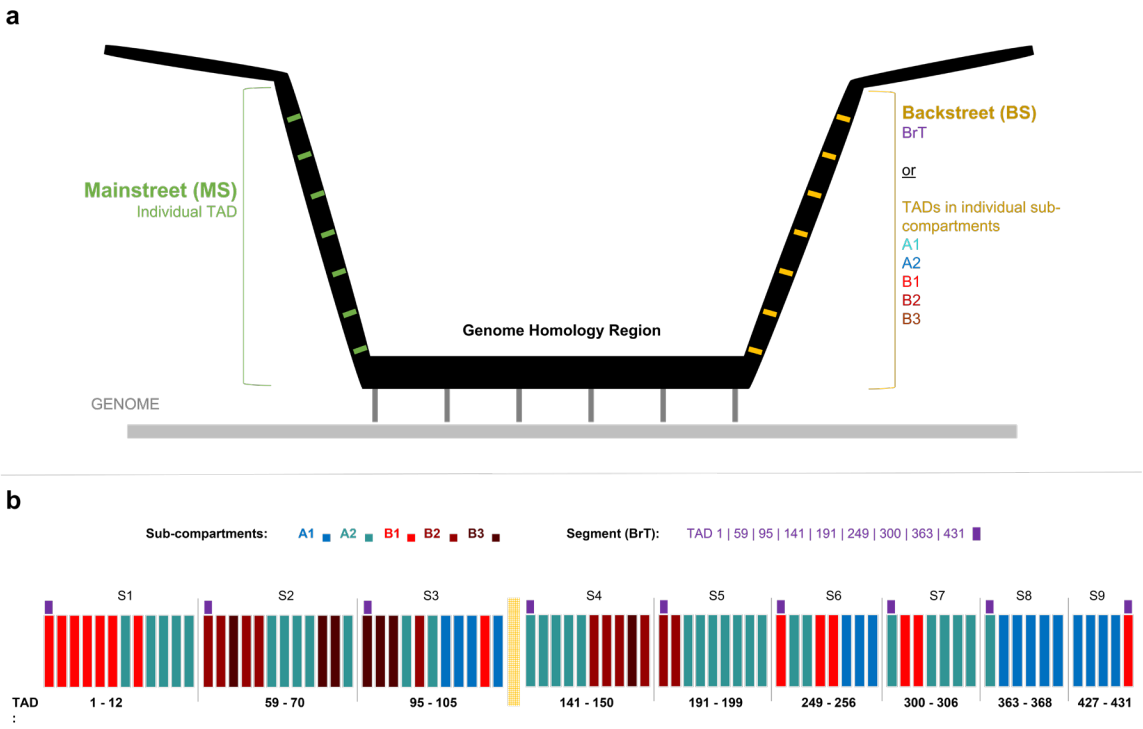


Figure 2|Oligopaints probe anatomy and distribution of 80 TADs spaning the length of human chromosome 9.

a, Each Oligopaints probe consists of five distinct areas. The Genome Homology Region (GHR) complements specific regions of the genome (in this case, TAD loci on human chr 9) with high affinity. The GHR is flanked on both sides by nucleotide barcode sequences that do not complement the human genome. On the 5' end, Mainstreet (green) is barcoded for each individual TAD (any TAD can be imaged individually). On the 3' end, Backstreet is barcoded to target all together the first TAD of segments 1-8 and the last TAD of segment 9 (BrT). Backstreet is also barcode-assigned to all TADs that belong to their respective sub-compartment (A1, A2, B1, B2, B3). **b,** The TAD distribution across human chromosome 9, each colored bar represents a single TAD and its designated sub-compartment. Purple bars indicate TADs belonging to BrT. Each segment (S1-S9) is shown with their sequential TADs assignment. Yellow bar in between S3 and S4 is the location of chr 9 centromere, separating *p*-arm from *q*-arm.

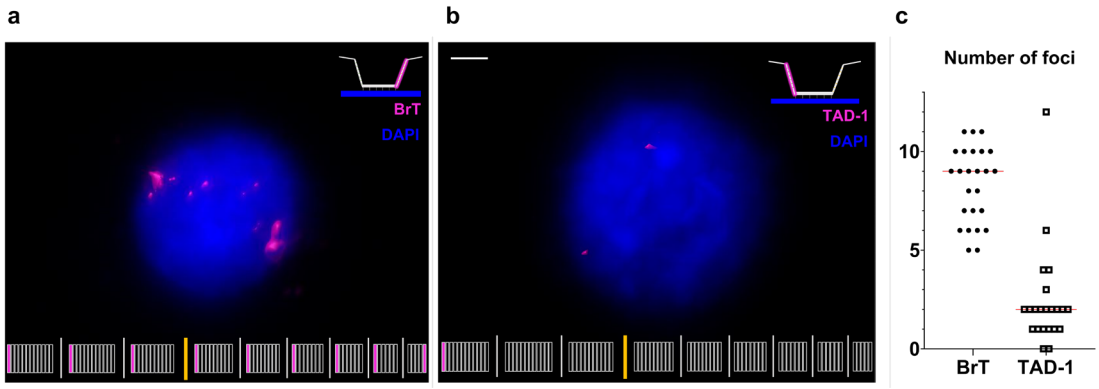


Figure 3|Widefield fluorescence microscopy displays relative location of 9 TAD segments and a single TAD of human chromosome 9.

3-dimensional fluorescent *in situ* hybridization was performed on GM12878 cells using the amplified Oligopaints library. FISH signal (magenta) is shown for, **a,** the diffraction-limited location of the 9 TAD segments and for, **b,** the first most upstream TAD in the oligopaints library, TAD-1. **c,** total number of FISH signal foci were counted for each FISH condition where number of cells, n=25 for BrT and n=22 for TAD-1. Pictographs of an Oligopaints probe and chr 9 scale indicates a Mainstreet or Backstreet target and specific TAD targeted, respectively.

3.5 Super-resolution 3D STORM is sufficient to spatially resolve single TADs

To address the questions of the genomic spatial limits of view TADs in super-resolution, and the reproducibility of multi-round STORM using fluidics, we devised an experimental plan to do so. We proceeded with a general workflow that has been tested for multi-round OligoSTORM (**Figure 4a-d**). The imaging sequence included, in the first round of imaging targeting the TADs of the BrT BS barcode (TADs 1, 59, 95, 131, 191, 249, 300, 363, 431). This allowed us to visualize a TAD representative for each segment all together. Subsequent rounds of 3D STORM acquisition targeted TADs belonging to BrT on an individual basis using their specific TAD MS barcode (**Figure 4e**). This ultimately allowed us to overlay and decode each individual TAD based on the BrT BS barcode round of imaging (round 1). Interphase nuclei of GM12878 cells (n=5) in 2 fields of view (FOV) went through 10 rounds of microfluidics/image acquisition. We filtered raw single molecule localizations to fit a 50 nm axial precision to give us an average particle localization precision of 23.14 ± 2.91 nm and 21.22 ± 1.32 nm in the x and y axes, respectively. To select for the highest quality localizations, we applied global selection parameters of density-based spatial clustering of applications with noise (DBSCAN) of BrT (round 1). Completion of this analysis qualified all clusters that co-localizes with a fluorescent widefield reference overlay image of BrT taken before 3D STORM acquisition (**Figure 4f**). DBSCAN parameters were toggled to reveal all 9 TADs of BrT in each of two homologs of Chr 9 and each cluster was spatially distinguishable one from another (**Figure 4g**). The same DBSCAN parameters were subsequently applied to the single molecule localizations from rounds 2-10. A high-quality cluster from each round of image acquisition co-localized with clusters from BrT in round 1 of imaging with a high percentage of volumetric overlap (**Figure 4h**). These data suggest that single molecules detection from our Oligopaints library are done in nano-scale resolution. Additionally, TADs can be visualized with a relatively high degree of confidence and reproducibility using multiple

and mixed use of microfluidics, and OligoSTORM targeting MS and BS barcodes—validating our Oligopaints library and experimental design process (**Figure 4a-h**).

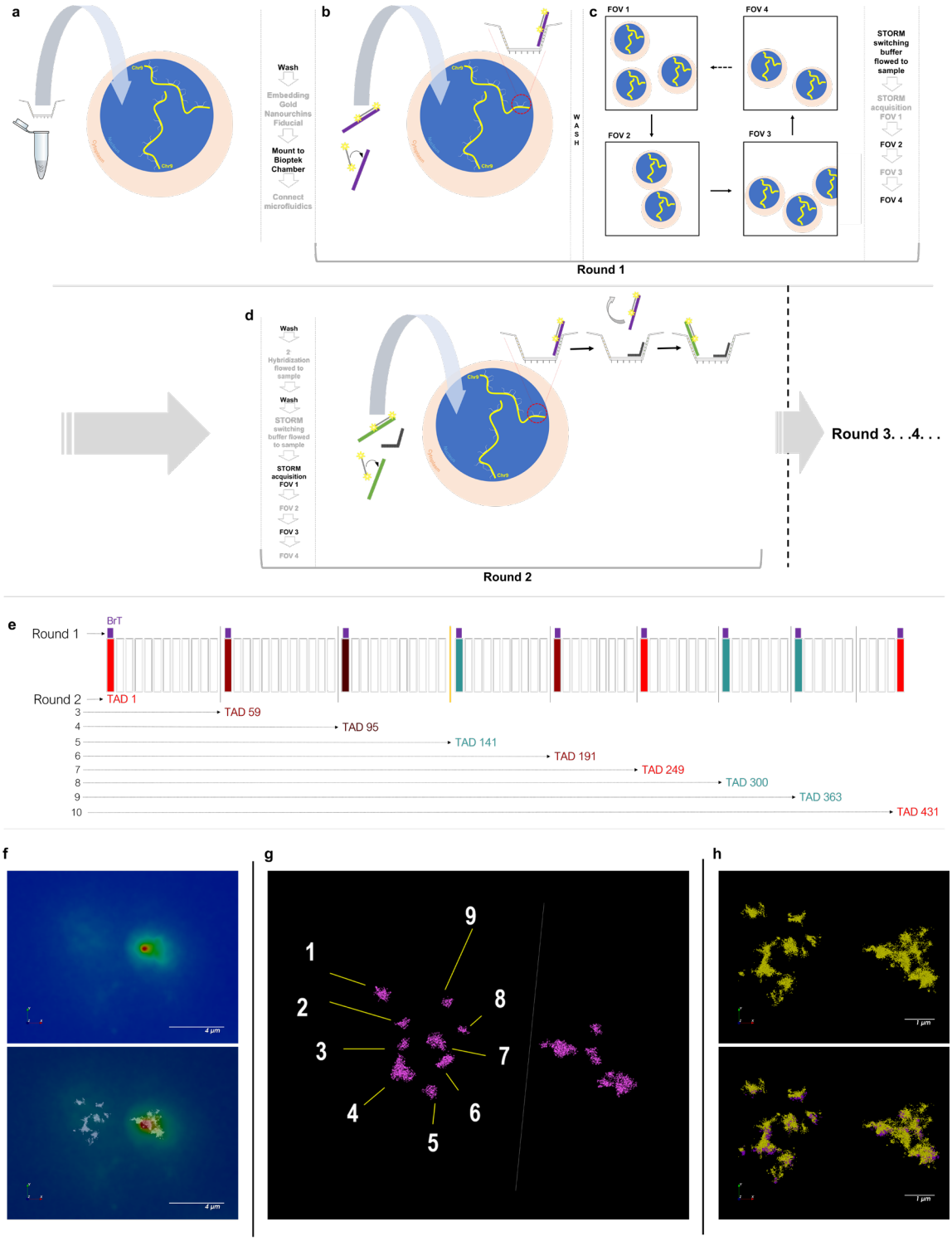


Figure 4|General workflow for multistep OligoSTORM and Oligopaints library validation.

a, The Oligopaints library is included in a primary hybridization that targets all 80 TADs on both homologs of chromosome 9 in a 3D-FISH procedure. After hybridization, the sample is embedded

with gold nano-urchins to serve as fiducials to correct for drift after STORM acquisition. The cover glass containing the hybridized sample is installed and mounted to a microfluidics chamber, and finally to the microfluidics instrument. **b,** A secondary hybridization step included hybridization of oligos called a Bridge (purple) that has complementarity to the Backstreet barcode. Additionally, another secondary oligo with AlexaFluor 647 (AF647) conjugates on both ends complements and binds one end of the Bridge—the AF647 signal, thus, serves as a tether for the Oligopaints probe itself. After a washing step to decrease background levels, **c,** widefield fluorescent microscopy is used to confirm FISH signal and to designate the x, y, and z-range coordinates for all fields of view (FOV). These exact coordinates serve as reference points for all rounds of fluidics and STORM acquisition. For each FOV, a widefield reference image is acquired. Next, the microfluidics system is programmed to flow in STORM switching buffer into the sample chamber and STORM image acquisition begins with FOV 1 and the microscope continuously acquires until the last FOV is imaged to completion. Round 1 includes the first fluidics and first STORM imaging cycle, **b** and **c.** The fluidics and imaging cycle is repeated for Round 2, **d,** with the inclusion of an additional oligo meant to remove the previous round's Bridge with its attached fluorescent secondary oligo using a Toehold (black/white dots). This removal of the previous round's Bridge ensures non-conflicting signal in subsequent rounds. Another Bridge (green) with its secondary fluorescent oligo is also flowed into the sample chamber to bind its specific Oliogopaints probe Mainstreet barcode. **e,** Oligopaints library validation was performed in 10 rounds. **f-g**, results of 10 rounds of imaging. **f,** widefield fluorescent reference image of BrT, top; overlay of single molecule localizations (SML) from STORM after round 1, bottom. **g,** SML displaying 9 segments placed along the length of chr 9, dashed line demarks chr 9 homologs. **h,** SML of BrT (yellow), top; SML overlay of TADs 1, 59, 95,141,191, 249, 300, 363, 431 after 10 rounds (purple). Superimposition (yellow and purple) of matched TADs belonging to BrT aligned to an average of 73% ± 6% (n=6) based on centers of mass.

3.6 Grouped TADs between sub-compartments are spatially distinct

In addition to validating the spatial occupancy of the 9 segments spanning Chr 9 in super-resolution, we also questioned the placement of sub-compartments relative to each other. We, therefore, performed OligoSTORM to target the BS barcode for the groups of TADs that correspond to their sub-compartment address. Six rounds of imaging were performed, one round for each sub-compartment. Single molecule precision and DBSCAN parameters were pre-defined on previous metrics from the library validation steps. All five sub-compartments were detected with high precision and formed cluster numbers consistent with the number of TADs that constitute their appropriate sub-compartment (**Figure 5a-f**). Finally, sub-compartment localization clusters of higher particle count aligned with BrT fluorescent widefield and super-resolution clusters. Except for sub-compartment B3, we were able to distinguish both territories/homologs of Chr 9 in a sub-compartment manner. Sub-compartment B3 occupied significantly more space relative to other sub-compartments imaged (**Figure 5e**). We next analyzed inter-sub-compartment volumetric centroid distances and spatial overlap and found that TADs in each sub-compartment are spatially isolated from each other. However, no pairwise median distance relationship was uncovered between individual TADs between sub-compartments. Together, these observations suggest that sub-compartments mostly occupy distinct spaces with little overlap. Also, since these sub-compartments can be singled out one from another, it can serve as another qualifying variable to aide in a multiplexed imaging approach, similar to an approach previously done that utilized BrT and reference images.

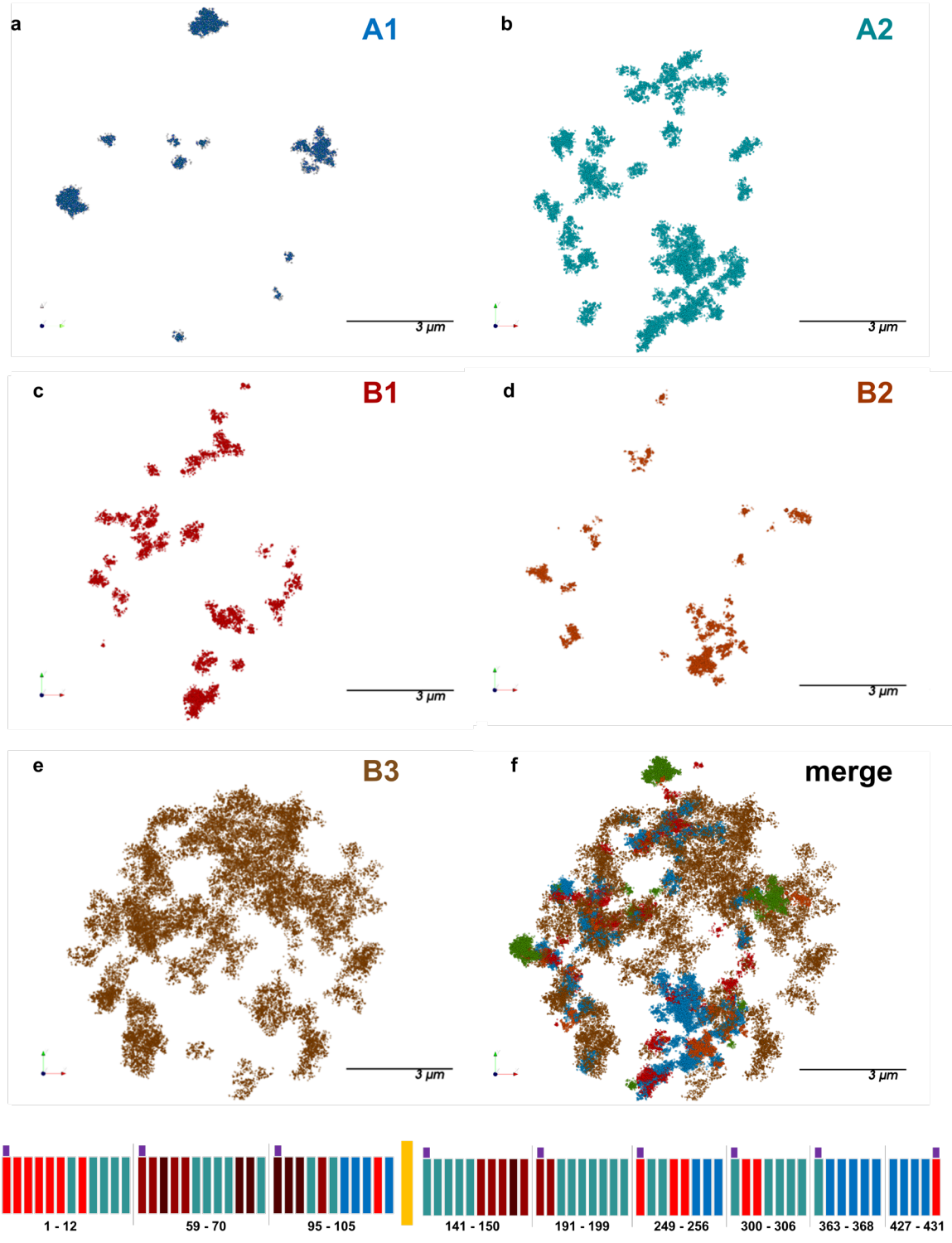


Figure 5|Observing human chromosome 9 sub-compartments in nanoscale.

a-f, automated microfluidics and OligoSTORM was performed in five rounds, one round per sub-compartment, to reveal distribution of sub-compartments as single molecule localization point (20

nm) cloud views. Shown is sub-compartment A1, **a**; A2, **b**; B1, **c**; B2, **d**; B3, **e**. **f**, displays a view of sub-compartments merged.

3.7 A multiplexed microfluidics and imaging approach captures 52 TADs located end to end on human chromosome 9

Extensive consideration was taken when deciding a strategy for image acquisition of all 80 TADs in our Oligopaints library. The fact that multiple GM12878 nuclei in a single FOV, together, occupies an imaging z-range of average of 9.6 ± 1.6 μm, the sample time in acquisition, time in automated microfluidics, and half-life of STORM switching buffer, was a determining factor for the precision of single molecule localizations. Thus, we proposed a logical workflow based on previous observations that we made from target BrT backstreet (**Figure 4g**), sequential imaging of BrT TADs on mainstreet (**Figure 4h**), and sub-compartment spatial distribution (**Figure 5**). An added benefit of our imaging system is the integration of the use of reference images. Considering all of these variables, our workflow followed this sequence using the same general workflow in **figure 4a-d**: (1) initial 2° hybridization of AF488-cojugated oligo targeting the 5' and 3' primer index sequence (2) initial 2° hybridization also included BrT BS Bridge (AF647-conjugate) (3) FISH signal-to-noise ratio assessed for primer index and BrT (4) selection and loading of x,y,z-range coordinates for 6 FOV containing non-aneuploid nuclei (5) 3D STORM to complete round 1 of imaging of BrT (7-12) all sub-compartments are individually imaged completing rounds 2-6 (13) three TADs belonging to a mix of sub-compartment assignments are targeted and imaged concurrently, i.e. TAD-5/B1; TAD-150/B2; TAD-363/A2 to complete round 7 (14) another three TADs imaged using same qualifications as previous step, i.e. TAD-69/B3; TAD-364/A1; TAD-198/A2 (15) format repeated for the next multiplex of three TADs (**Table 2, figure 6a-e**).

Round	Target	Genomic start (hg19)	Genomic end (hg19)	Size (kb)
1	BrT	variable	variable	3510
2	SubC-B1	variable	variable	2070
3	SubC-A1	variable	variable	979
4	SubC-B2	variable	variable	3500
5	SubC-A2	variable	variable	4070
6	SubC-B3	variable	variable	3550
7	TAD-5 B1	897001	1002000	105
	TAD-150 B2	72584001	72731000	147
	TAD-363 A2	127064001	127174000	110
8	TAD-69 B3	19927001	20350000	423
	TAD-364 A1	127187001	127296000	109
	TAD-198 A2	86879001	86936000	57
9	TAD-2 B1	277001	475000	198
	TAD-192 B2	85951001	86162000	211
	TAD-12 A2	2016001	2086000	70
10	TAD-68 B3	19470001	19923000	453
	TAD-427 A1	139986001	140134000	148
	TAD-11 A2	1947001	2014000	67
11	TAD-1 B1	202001	273000	71
	TAD-59 B2	16160001	17100000	940
	TAD-7 A2	1452001	1610000	158
12	TAD-95 B3	29795001	31015000	1220
	TAD-365 A1	127297001	127375000	78
	TAD-9 A2	1702001	1871000	169
13	TAD-3 B1	614001	795000	181
	TAD-60 B2	17122001	17465000	343
	TAD-10 A2	1875001	1940000	65
14	TAD-96 B3	31123001	31805000	682
	TAD-366 A1	127376001	127473000	97
	TAD-64 A2	19051001	19177000	126
15	TAD-4 B1	811001	890000	79
	TAD-62 B2	17808001	18210000	402
	TAD-65 A2	19188001	19228000	40
16	TAD-97 B3	31827001	32344000	517
	TAD-367 A1	127476001	127528000	52
	TAD-66 A2	19235001	19374000	139
17	TAD-6 B1	1023001	1451000	428
	TAD-63 B2	18257001	19030000	773
	TAD-67 A2	19375001	19456000	81
18	TAD-149 B3	72321001	72580000	259
	TAD-368 A1	127537001	127619000	82
	TAD-70 A2	20361001	20620000	259
19	TAD-8 B1	1625001	1688000	63
	TAD-99 B2	32590001	32966000	376
	TAD-98 A2	32380001	32579000	199
20	TAD-104 B1	33625001	33730000	105
	TAD-428 A1	140182001	140304000	122
	TAD-100 A2	33001001	33076000	75
21	TAD-249 B1	99090001	99426000	336
	TAD-191 B2	85602001	85914000	312
	TAD-141 A2	71197001	71301000	104
22	TAD-301 B1	113101001	113147000	46
	TAD-429 A1	140305001	140434000	129
	TAD-142 A2	71326001	71394000	68

23	TAD-302 B1	113191001	113304000	113
	TAD-430 A1	140508001	140670000	162
	TAD-143 A2	71464001	71540000	76

Table 2|Sequential OligoSTORM on multiple TAD targets on human chromosome 9.

List of all targets in chromosome 9 that underwent sequential OligoSTORM. The first six rounds of automated fluidics and 3D STORM targeted one Backstreet barcode in each round. Rounds 7-23 targeted a multi-plex of three TADs at a time with their Mainstreet barcode. Each TAD in the multi-plex belonged to unique sub-compartments.

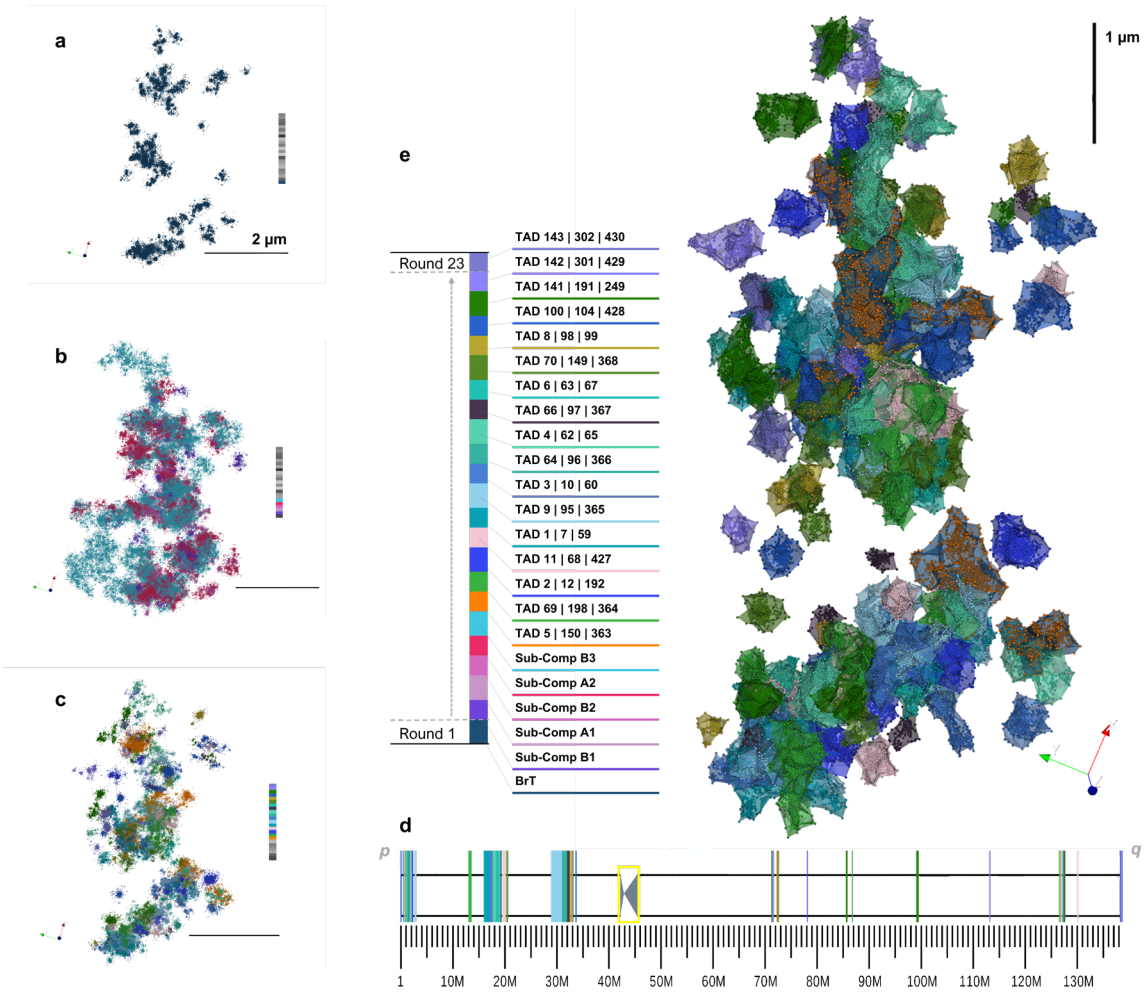


Figure 6|3D STORM composite display of human chromosome 9.

a-e, displays composite SML point cloud and alpha hull cluster reconstructions of 23 rounds of automated fluidics and 3D STORM. **a**, round 1 point cloud clusters (sphere point=20 nm diameter) of target BrT Backstreet barcode. **b**, rounds 2-6 composite point cloud clusters (1 point=20 nm diameter) targeting TADs of each sub-compartment on their Backstreet barcode. **c**, rounds 7-23 composite point cloud clusters (1 point=20 nm diameter) targeting individual TADs on their mainstreet barcode. **d**, ideogram of human chromosome 9 showing all TADs acquired after 23 rounds of automated fluidics and 3D STORM with their relative genome location and genomic size. **e**, display of TADs of Chr 9 as cluster alpha hulls

3.8|Walking along the TADs of human chromosome 9 from *p*-arm to *q*-arm

The TAD walk along chromosome 9 was terminated on round 23 after 16 days of continuous automated microfluidics and 3D STORM acquisition for six fields of view. This included two hours for automated fluidics to flow in secondary hybridization/STORM switching buffer and subsequent single molecule localization of 11 hours for each individual round. The size range in the Z-axis of Chr 9 selected measured 10.3 ± 1.4 μm is a compatible depth range of the single molecule localization microscope used for these data (Bruker Nanosystems Vutara 352). This dataset employed significant time and storage resources where raw particle counts exceeded 18-billion and occupied about 32 terabytes of storage targeting 58 total sub-regions of Chr 9. Thus, a defined strategy was necessary to filter the raw data for into sub-parts was vital to facilitate real-time and reasonable single molecule reconstruction image manipulation (through Graphics Processing Unit) and STORM data table processing (through Central Processing Unit). First, we aimed to partition the data by selecting nuclei <0.5 μm of at least two 90 nm drift correction-purposed gold nano-urchin beads. This was done because we observed that the total number of fiducial beads varied by FOV, and the drift pattern in x,y,z of each bead showed significant variability even in the same FOV. Selecting beads in close proximity to each nucleus and bead-to-bead, was essential to have homogeneity, reproducibility, and accuracy in x,y,z drift correction. Nuclei that did not meet this drift correction criterion was discarded and not used for any further analysis. After successful particle drift correction, we used DBSCAN to assess the general density clusters that fit in our regions of interest (ROI) starting with round 1, BrT. After DBSCAN, the resultant super-resolution clusters for BrT were confirmed by its widefield reference overlay, the widefield overlay of the universal reference (signal coverage of all Oligopaints probes), and comparison of particle counts from previous runs of STORM (**Figure 6a**). The collective location of each segment using the BrT Backstreet barcode in round 1 of automated fluidics and 3D STORM (**Figure**

6a) was selected by particle number, particle density, and a widefield reference overlay/co-localization. The result can then be applied to select sub-compartment TADs (**Figure 6b**) in rounds 2-6 with the additional help of the widefield reference overlay of the universal labeled probes that targets the entirety of the Oligopaints library (all TADs at once). Individual TADs that were imaged using Mainstreet barcodes were selected using these same parameters. By doing this, we were able to observe 110 (52 TADs in each homolog) single clusters with an approximate 3D tract that mimics that of the BrT and sub-compartment distribution of clusters (**Figure 6c**). To visualize more clearly the borders of TADs we applied a cluster hull around singular localization clusters using a hull alpha measure of 50 nm (**Figure 6e**). This allowed us the opportunity to visualize clear distinctions between TAD entanglements, spatially distant TADs, TADs in *p* versus *q*-arm of Chr 9, and the relative boundaries separating Chr 9 homologs.

All subsequent decoding of sub-compartments (**Figure 6b**) and by the 3-TAD-multiplex (**Figure 6c, 6e**) took on a similar workflow of confirmatory steps where a sub-compartment ROI is qualified by BrT cluster co-localization (TADs with matching sub-compartment) and universal reference overlay (overall diffraction limited Oligopaints probes signal tract). Individual TAD (mainstreet) molecule localization clusters after DBSCAN were decoded and called based on a few qualifying steps; for example, a DBSCAN candidate cluster for TAD-1 (imaged MS in round 11) was super-imposed or entangled with a cluster from BrT (imaged BS in round 1), near flanking TADs i.e., TAD-2 spatially occupied between TAD-1 and TAD-3. These general steps were repeated for each individual TAD of the 52 TADs that were potentially imaged in each cell—in addition, the process from drift correction, DBSCAN cluster analysis, to individual TAD selection was repeated independently three times to refine the TAD calls. Using these high levels of stringency resulted in a final TAD call in 12 nuclei in four separate fields of view. Overall,

all TADs in sub-compartments A1, B2, and B3 were potentially called (homolog dependent), while 11/12 B1 and 17/32 A2 TADs were called.

A representative virtual walk along genomically-sequential TADs was created to observe a cumulative view of human chromosome 9, where each TAD sub-compartment assignment is also shown (**Figure 7**). After TAD calls for each cell, we were pleased to observe chromosome 9 occupying two spatially distinct territories with no entanglements. Number of total TADs called differed on a homolog basis with minimum 36 and maximum of 52. The reason for this is likely due to the high amount of noise present in the raw data, making it not suitable to qualify TADs with confidence. No correlations were found between TAD genomic size, Oligopaints probe density, of sub-compartment assignment, with singular levels of noise contribution. There is more enrichment of TADs captured on the *p*-arm compared to the *q*-arm, which is reflective of the Oligopaints library design—although the tract of each Chr 9 territory is still well defined.

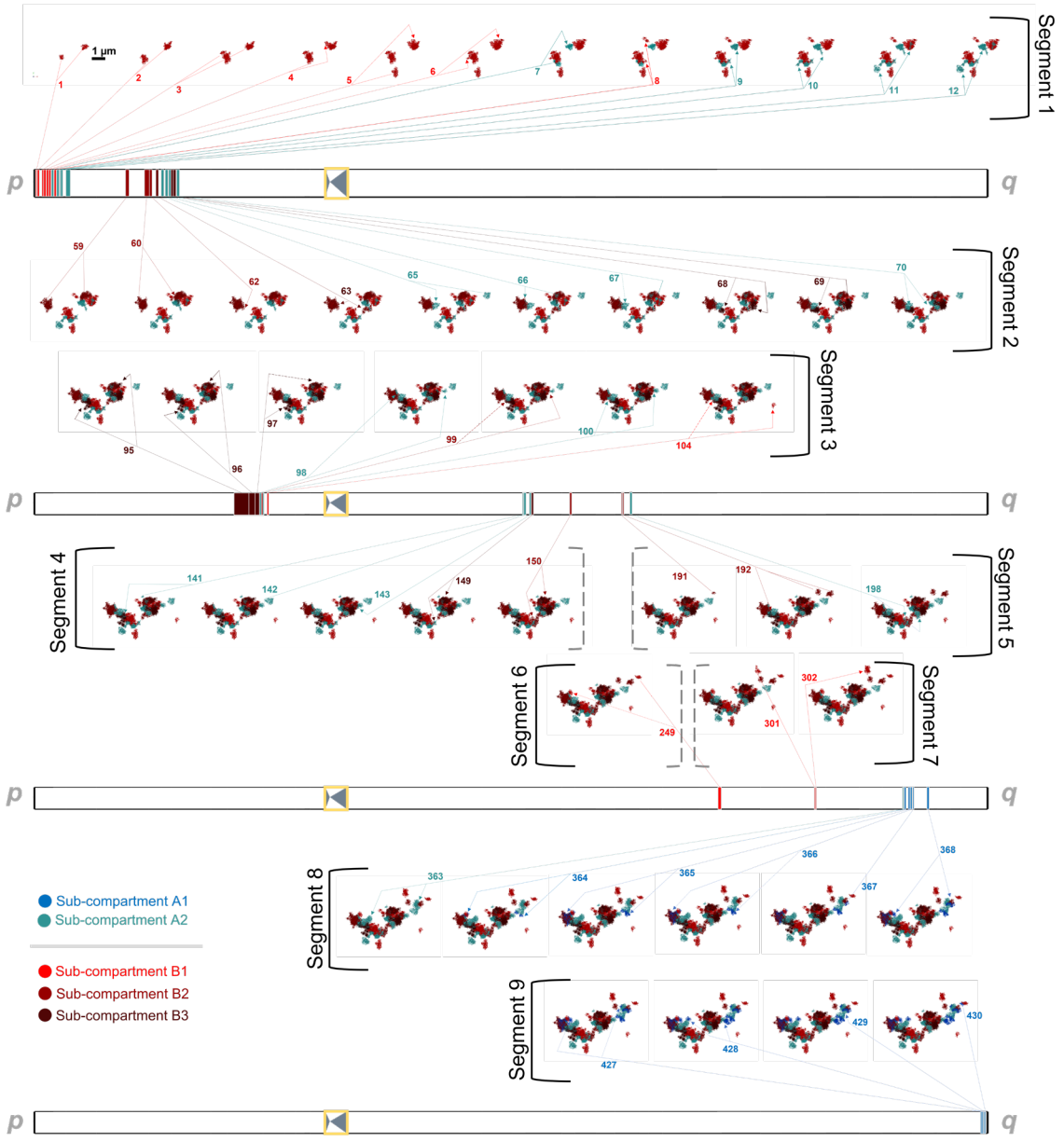


Figure 7|A TAD walk from *p*-arm to *q*-arm along human chromosome 9

A virtual TAD-walk displays an accumulative assembly of human chromosome 9. From the top left, the walk starts with a SML point cloud (spheres, 20 nm) cluster hull view of both copies of TAD-1 (number 1) of sub-compartment B1 (red). TAD-1 is traced to its relative position and relative genomic size on the most upstream location on Chr 9, *p*-arm. TAD-1 is also the first TAD in segment 1 (bracket). All 9 segments are bracketed appropriately to contain their assigned TADs and genomic locations. All other TADs imaged and decoded are added onto the whole Chr 9 view in

numerical order, from *p*-arm to *q*-arm. Solid lines with arrows are directed to the exact location of a single TAD in view. Dotted lines with arrows are directed to TADs that are behind or masked by the current view. The centromeric region is indicated by yellow box and all TAD sub-compartment assignments are color-coded as shown.

3.9 Multiple TADs display variability in 3D STORM measured structural features

Subsequent to TAD decoding with DBSCAN, we were able to assess the spatial distribution of all TADs and correlations we can extract from their sub-compartment assignment, their position in respect to each other in the same homolog and from their opposing homolog, and how they distribute on the whole (**Figure 8a**). We have observed from a chromosome territory perspective a few things: in 5 out of the 12 nuclei that made it through analysis, we observed the occurrence of homolog pairing (pairing qualified as an inter-distance of <500 nm); wide variance in sub-compartment entanglement, *p*-arm orientation in the 5' direction of homologs is typically mirrored (**Figure 8a**). Because the amount of total data we had to analyze was insurmountable, to make more key observations looking at chromosome 9, we were prompted to analyze on a single TAD level to parse out any significant difference that may be contained in our data. To leverage the strengths of 3D STORM and since a super-resolution view of a whole chromosome containing groups of sequential TADS on this scale was not yet performed using single molecule localization data, we were interested in an application-based assessment of volume (μm^3), surface area (μm^2), compaction factor (nucleotides/μm^3) which measures TAD chromatin density, sphericity which measures how closely TADs reflect the shape of a sphere, and radius of gyration (nm) which measures the average radius distance of a 3D shape given its known center of mass. Based on alpha shape cluster hulls generated with a 100 nm alpha limit, the volumetric space occupied, TAD to TAD, varied between 1.2 $nm^3 \cdot kb^{-1}$ - 1.4 $\mu m^3 \cdot kb^{-1}$. Even though this is 1000-fold difference in range of volumes (**Figure 8b-d**, **Table 3, Supplementary data S1**), there were few significant differences of volumes between matched TADs of the same cell (between homologs), whereas we observed more volumetric differences between matched TADs between cell to cell (**Table 3**)**.** Interestingly, measures of surface area of TADs imaged showed largely the same trends as volumetric measurements in terms of absolute surface area measurements, and

in terms of the TADs that show significant differences in measurements when compared to the mean (**Figure 8e, Table 3, Supplementary data S2**). Among all the physical measurements that were assessed in this study, we were intrigued by the comparison of the sphericity between multiple TADs, TADs between homologs, and TADs between cells. In general, the sphericity of TADs measured were not significantly different one from another or much difference was found between homologs or matched TADs between cells. Only 3 TADs appeared to be different between homologs and four different TADs showed some significant differences on a cell-to-cell basis (**Figure 8g, Table 3, Supplementary data S4**). The radius of gyration of matched TADs were, in general, about the same between homologs and not much difference was found between nuclei (**Figure 8h, Table 3, Supplementary data S5**). TAD compaction factors or chromatin density, however, displayed significant differences between matched TADs for most of the ROI acquired via 3D STORM but not much difference between matched TADs in the same cell (**Figure 8f, Table 3, Supplementary data S3**). Together these data suggest that TAD volumes, surface areas, radiuses of gyration and compaction factors show more variation between different cells than between homologs.

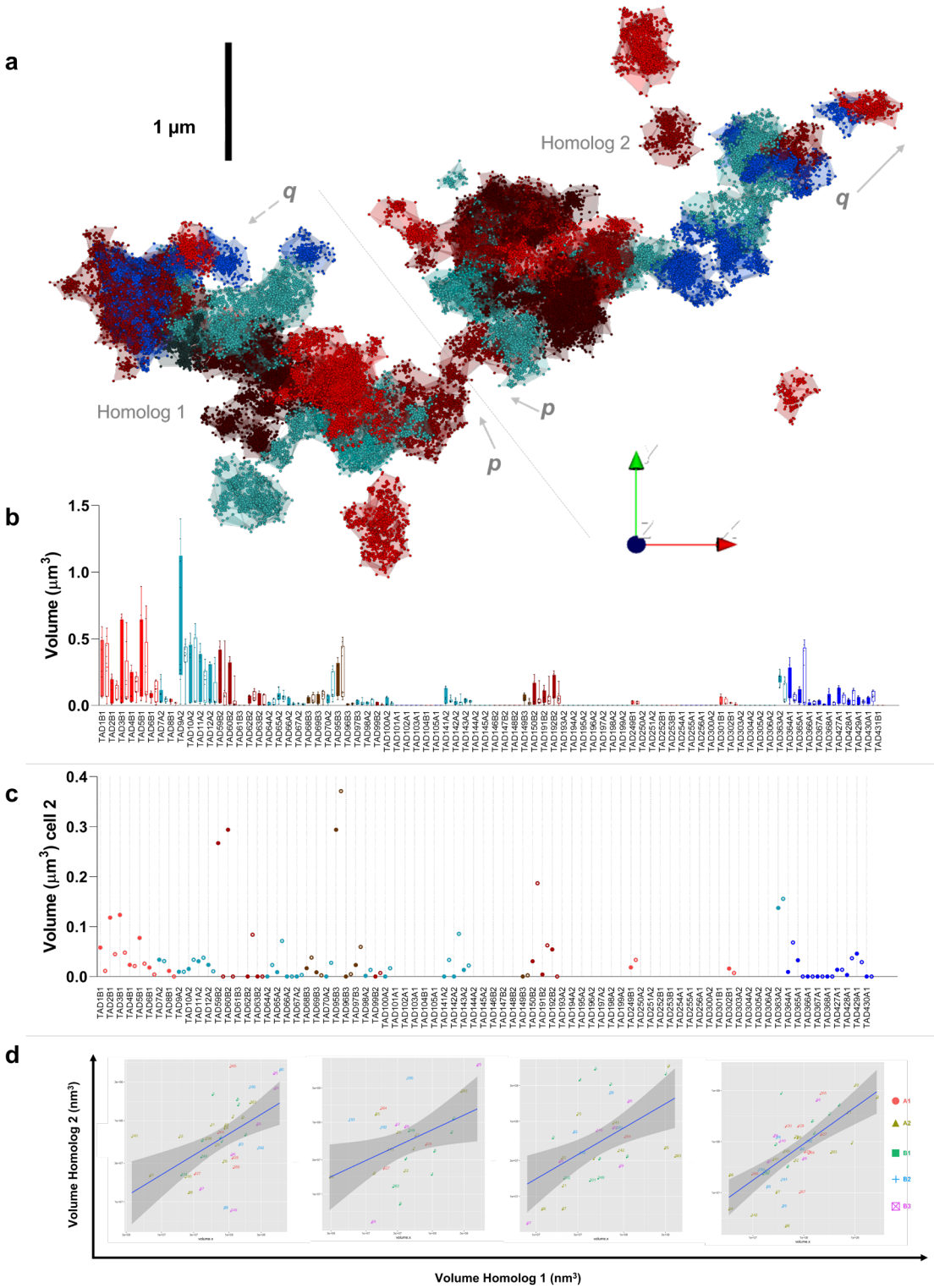
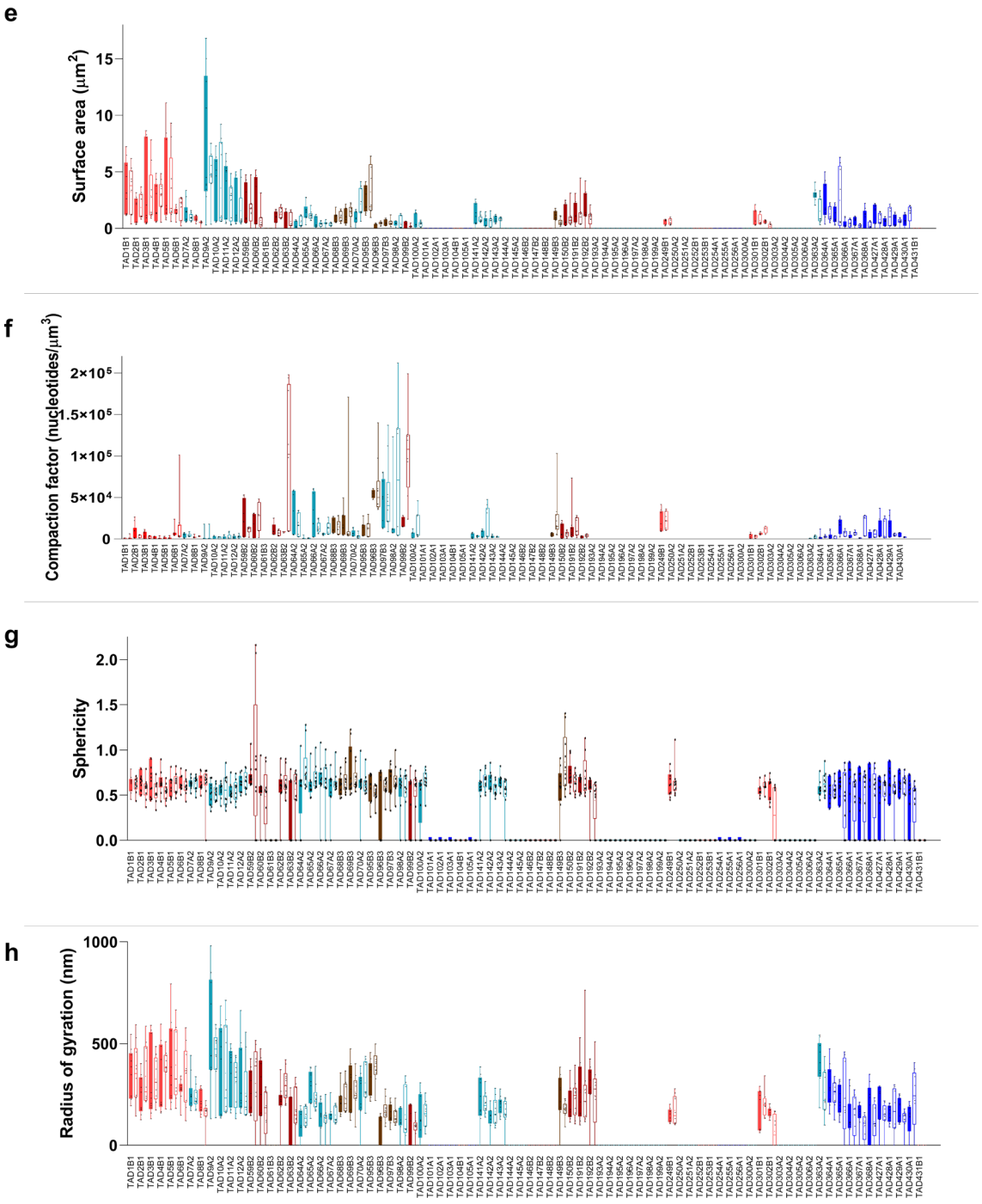


Figure 8|Structural composition of TADs in human chromosome 9



Extended figure 8|Structural composition of TADs in human chromosome 9

a-h, view of whole chromosome 9 and 3D STORM-based structural features measured. **a**, cumulative TAD view of two homologs of chromosome 9, each TAD is represented as alpha shape hulls of single molecule localization point (colored spheres=20 nm) clouds. **b-h**, boxplots displaying

the range and median of measured features for each individual TAD (displayed as black dots) in each homolog (solid bars indicate homolog 1, open bars indicate homolog 2). Boxplot color is an indicator of sub-compartment assignment. Individual absolute TAD volumes, **b**, are measured in μm^3 by homolog. **c**, displays TAD volumes in a single cell and homolog level. **d**, display TAD volumes (nm^3) of homologs plotted against each other, colored by sub-compartment designation; **e**, surface area measured in μm^2; **f**, compaction factor in nucleotides/ μm^3; **g**, sphericity; and **h**, radius of gyration (nm).

	Vol. homologs	S.A. homologs	Sph. homologs	R.G. homologs	C.F. homologs	Vol. nuclei	S.A. nuclei	Sph. nuclei	R.G. nuclei	C.F. nuclei
TAD1B1										***
TAD2B1										****
TAD3B1										**
TAD4B1			*							
TAD5B1										
TAD6B1						**	*		**	****
TAD7A2						****	**		*	
TAD8B1	**	**			*	****	**		*	**
TAD9A2						***	**			
TAD10A2										
TAD11A2										
TAD12A2										
TAD59B2				*				****	*	**
TAD60B2										
TAD62B2										*
TAD63B2	**	**			**	**				***
TAD64A2										
TAD65A2										***
TAD66A2						***	**			**
TAD67A2	**	**			**					****
TAD68B3										
TAD69B3								*		***
TAD70A2	*	*			*	***	**			
TAD95B3										
TAD96B3			**			*			*	**
TAD97B3										
TAD98A2						*	**		**	
TAD99B2		**		**	**	*				**
TAD100A2						**	**			**
TAD141A2						**	**			**
TAD142A2			*							***
TAD143A2						*	*			*
TAD149B3	**	**		*					*	****
TAD150B2										**
TAD191B2								**		
TAD192B2							*			
TAD249B1									*	
TAD301B1										
TAD302B1		*						*		**
TAD363A2	**	**		**	*		*			****
TAD364A1						****	*			*
TAD365A1						***	**			
TAD366A1										*
TAD367A1	**	**		**		***	**			
TAD368A1	**	**		**	*	*				****
TAD427A1				*						
TAD428A1						**	*		*	
TAD429A1						**	**		*	****
TAD430A1	**	**		**	*				*	**

Table 3|Distribution of TADs with significant differences in structural measures of volume, surface area, sphericity, radius of gyration, and compaction factor.

TADs are colored by significance comparing TAD-TAD by homolog or TAD-TAD by cell.

3.10 TADs display shape heterogeneity between homologs and between cells

We next sought to assess the shape of TADs in 3D. To do this, we visually inspected each individual TAD called through our DBSCAN analysis. We observed these TADs as cluster hulls which allowed us to view structures where its borders are visible, as point (spheres=20 nm) clouds which allowed us to visualize each individual particle (single molecule localization), and point splatting which allows for a inferred view of signal density and overlap on any given viewing angle (**Figure 9a-d**). We have observed that matched TADs from opposing homologs show single particle distribution variation, in terms of local area density (**Figure 9a**). This was the case for the same homolog matched TADs in all other cells (**Figure 9b**). The multi-pockets of particle density were visible in almost every case when manually observing each TAD for all homologs for all cells. Because we had the opportunity to scrutinize the shape of each TAD on an individual basis, two remarkable features struck us: B-compartment TADs (n=69) displayed extensions from a central mass taking on a *drumstick* appearance (**Figure 9c**), and A-compartment TADs (n=53) displayed a shape that consisting of two polar sides of higher density of particles taking on a *dumbbell* appearance (**Figure 9d**). More strikingly, is the observation that these shapes were homolog locked. This means that all *drumsticks* and all *dumbbells* in any given cell was found only in one of the two homologs of Chr 9, where the homolog location for *drumsticks* or *dumbbells* appeared to be independent of each other.

3.11 Structural differences of TADs can be isolated by functional sub-compartments

After analyzing 3D STORM-based features of structural differences on an individual TAD level, we next wanted to assess if significant TAD structural variance can be deciphered at the level of functional compartments and sub-compartments (**Figure 10a-c**). Our analysis shows that the volumes (**Figure 10a**) and surface areas (**Figure 10b**) between compartment A and B TADs are not different, however, significant TAD differences can be isolated at the level of sub-compartments. Interestingly, this was not

the case when looking at sphericity where TADs between A and B compartments, as well as between sub-compartments vary significantly (**Figure 10c**). Together, this means TADs can be grouped based on these physical feature measurements on the basis of functional sub-compartments.

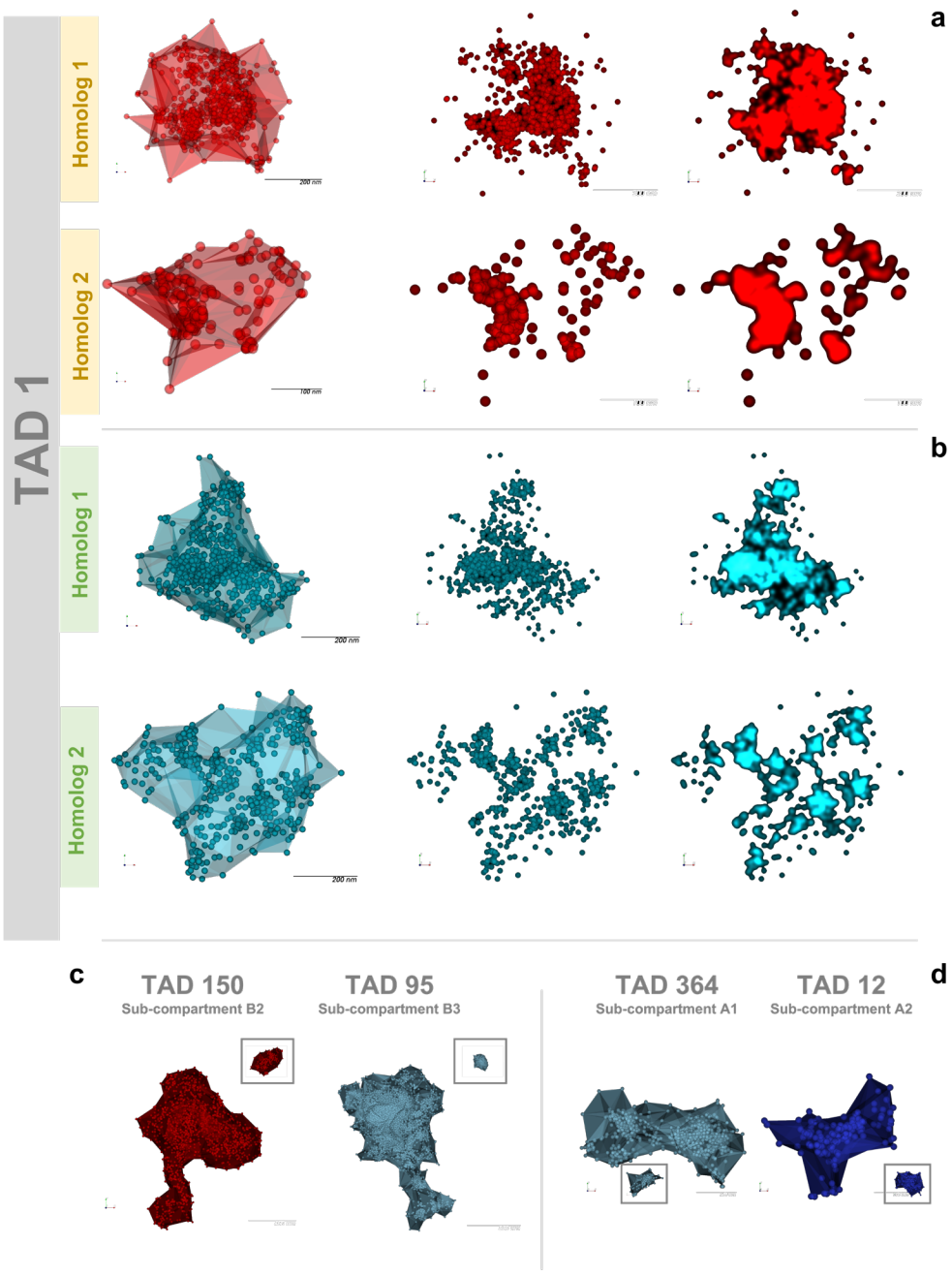


Figure 9|TAD structural differences are observed by single cells and single homologs

Singular and close-up views of individual TADs and its match in the opposing homolog. **a-b**, alpha shape hull view (left), point cloud (colored spheres=20 nm) (center), and point splatting displays (right) of TAD-1 in two different nuclei (red/blue). B compartment *drumsticks*, **c.** A compartment *dumbbells*, **d.**

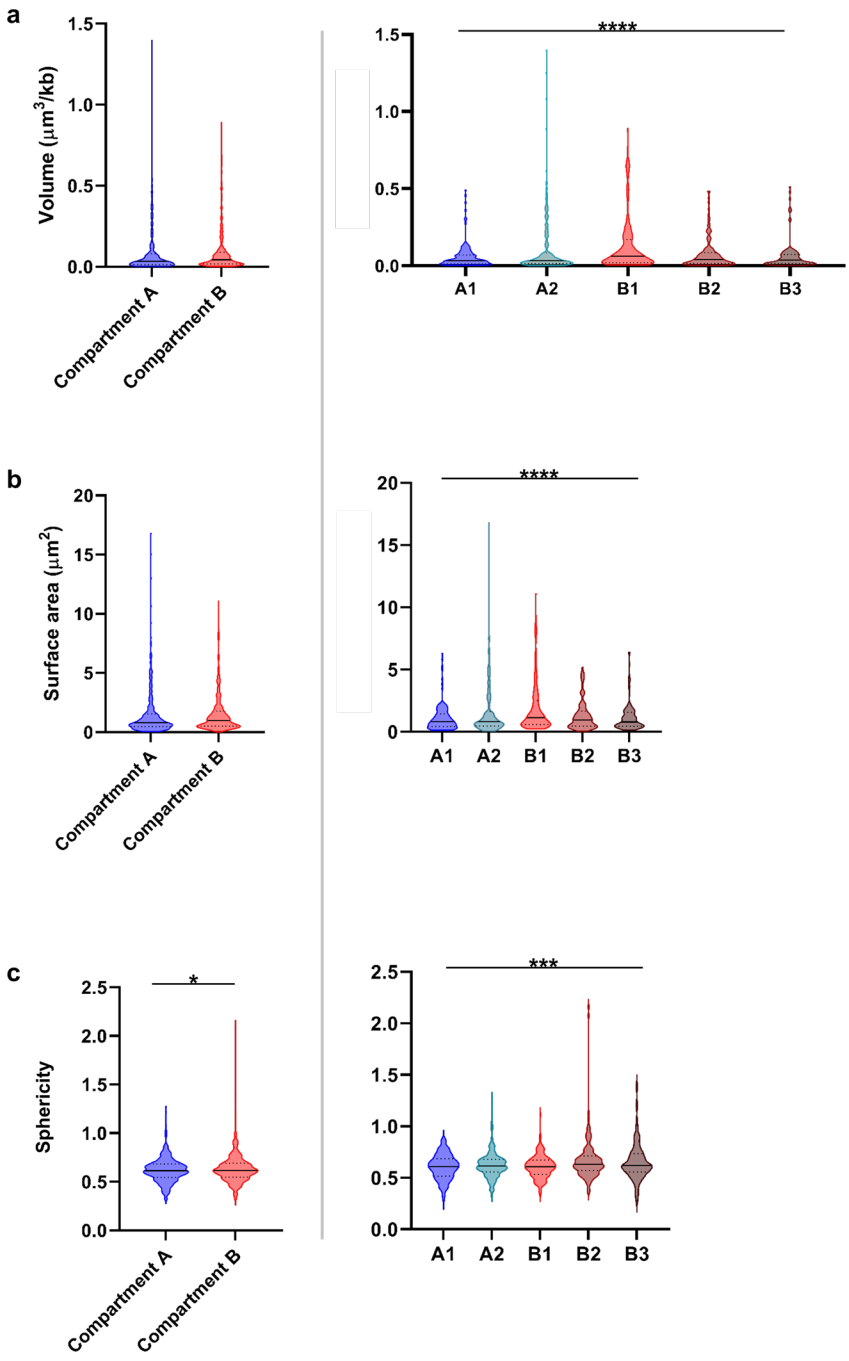


Figure 10|TAD structural signatures by functional compartment and sub-compartment

Violin plots display the distribution, range, mean, and significance of mean differences of TADs plotted in a compartment A or B, or sub-compartment approach. TAD volume, normalized to genomic size (µm³/kb) is shown in **a**; surface area in **b**; and sphericity in **c**.

3.12 Single molecule localization microscopy reveals inter-TAD distance anomalous patterns

A TAD distance transformation matrix was created for each homolog and for matched TADs in each cell (**Figure 11a-f**) using the volumetric centroids of each TAD. We have observed distinct distance patterns when comparing maps of TAD distances between homologs of the same cells (**Figure 11a, b**). Using our example nucleus, distances between TAD-1 to TAD-150 region are, in general, closer in proximity when compared to all other TADs distributed in the rest of the *q*-arm—downstream of TAD-150. The general inter-TAD distances, however, is not consistent with genomic distances as we did notice that we are not able to visualize the centromeric region in this pair of chromosome 9 territories' distance heat map, which lies between TAD-105 (chr9: 33,818,001- 33915000) and TAD-141 (chr9: 71197001- 71301000) (**Figure 11a, b**). Other differences include pockets of TAD clusters found in one homolog, and not in the other. Possibly indicating possible folding in this region. Due to this finding we sought to interrogate what the distances are between TADs of opposing homologs of the same cell plotted TADs of one homolog against the other (**Figure 11c**). This analysis revealed that the close proximity of TADs is not limited to constitutive TADs of the same homolog. We observe through this analysis that homolog pairing of TADs do occur in interphase nuclei. To get a strengthened idea of the placement of these TADs, made ball-and-stick models (**Figure 11d, f**) to have another visualization representation of the TAD distance transformations. Ball-and-stick models corroborate what we observe in heat maps. In addition, we compared what we observed in heat maps and with ball-and-stick models to the 3D composite of Chr 9 of the same cell, where it also corroborates with these homolog-paired TADs (**Figure 11e**).

3.13 Evidence of chromosome homolog pairing in mammalian diploid interphase cells

By doing further analysis of inter-TAD distance transformations, we uncovered 6 of 12 nuclei that were imaged showed some form of homolog pairing. We counted a total of 94 pairing events within 6 nuclei (**Figure 12, extended figure 12a**). A pairing event was qualified based on TADs that were closer than 500 nm, edge-to-edge. These paired borders were as close as 50-60 nm, which is a resolution well within our imaging resolution (**Extended figure 12b**). Decoding the identity of the paired TADs indicated that there is enrichment of pairing events in 4 segments of Chr 9. These regions include TADs of segment 1, segment 4, segment 8, and segment 9 (**Figure 12**). TADs that displayed a higher frequency of pairing typically paired with its matched counterpart (**Extended figure 12c**) on the opposing homolog, and these TADs were isolated to the extreme telomeric ends of the *p*-arm and *q*-arm (**Figure 12, extended figure 12c-d**). Together these data points to the capabilities of a super-resolution system like OligoSTORM to target the genome in order to extract fine details not possible with other methodologies.

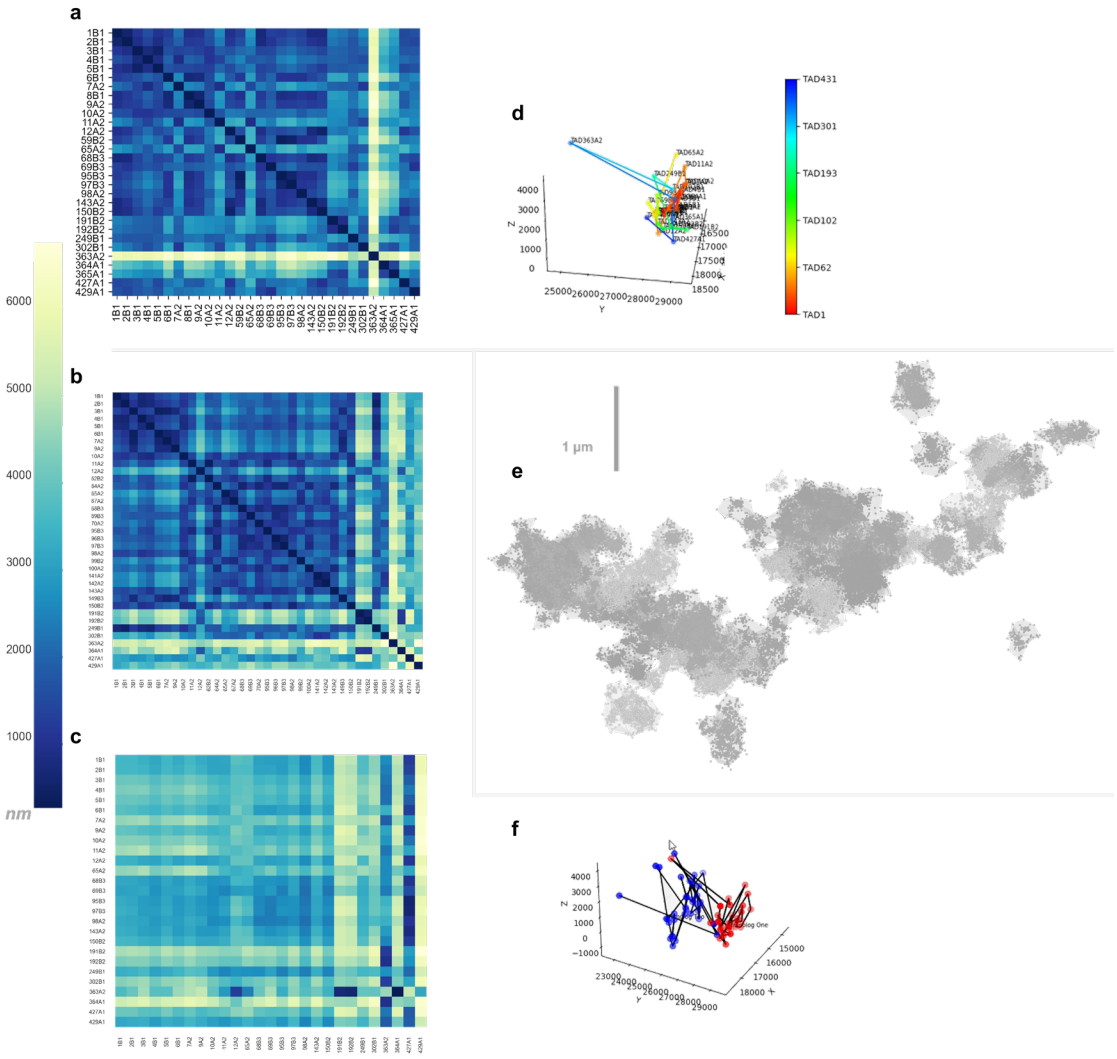


Figure 11|Inter-TAD and inter-homolog distance matrices coincide with ball-and-stick representations

Distances between the center of masses of TADs of a single cell, per homolog, are measured and represented as a distance matrix, **a** and **b**. Distance matrix of matched TADs in a single cell were plotted to assess how TADs interact on a homolog basis in **c**. A ball-and-stick representation, **d**, of heatmap of figure **a** is displayed. **e**, is a whole chromosome 9 view, colors point to the location highlighted in the heatmaps. A ball-and-stick view in **f**, displays a representation that mirror the heatmap in figure **c** of TADs in homolog 1 and homolog 2.

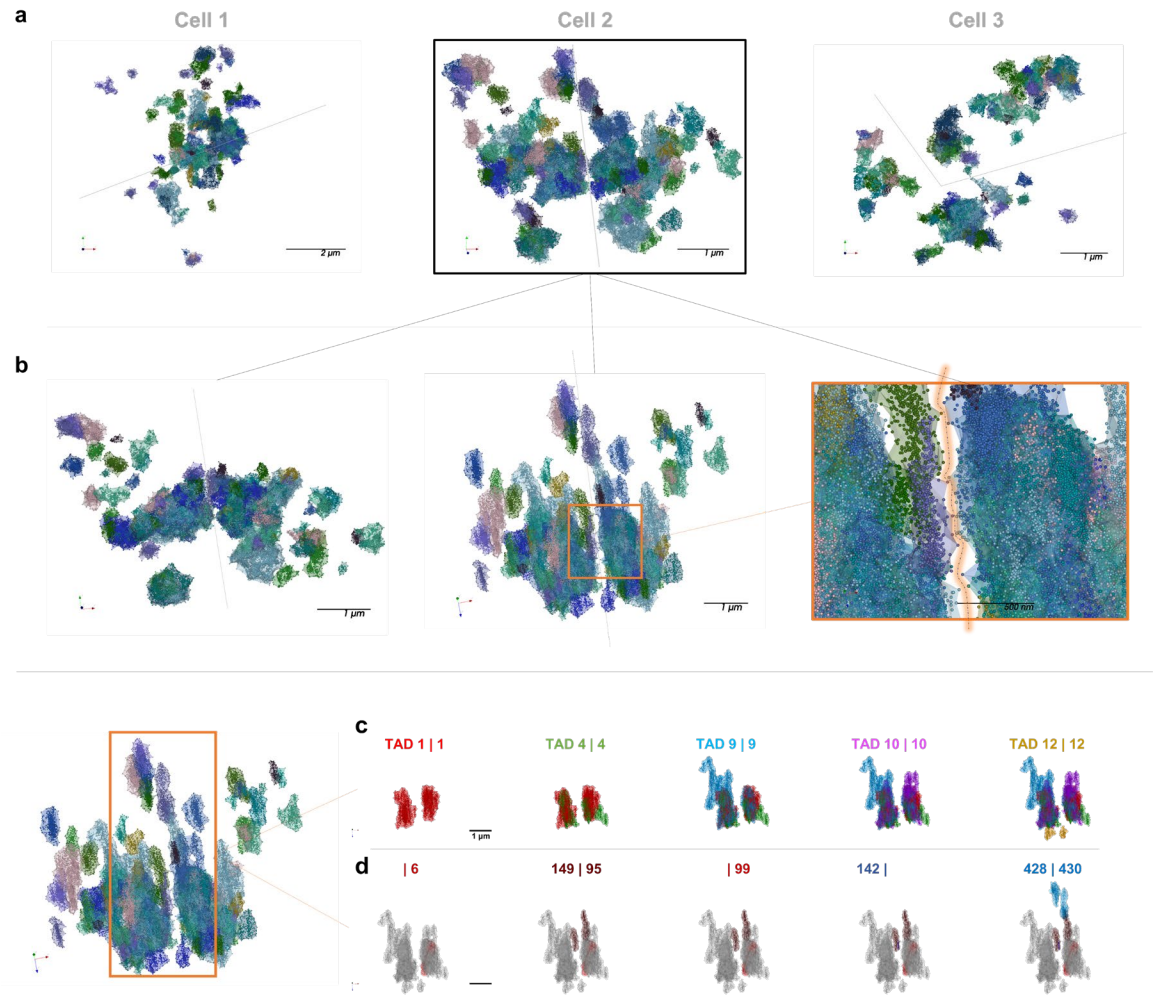
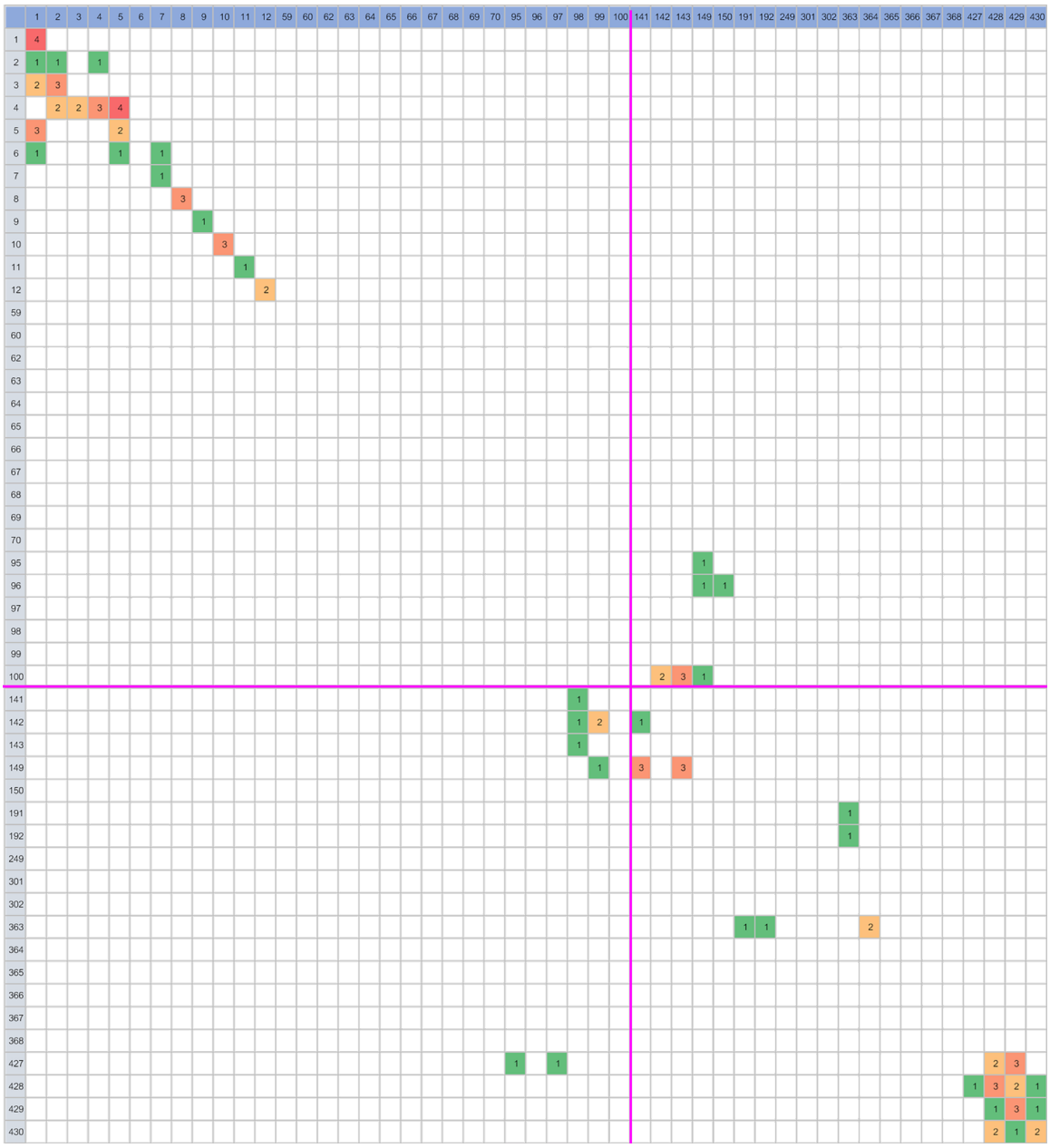


Figure 12|Close proximity of TADs between paired homologs

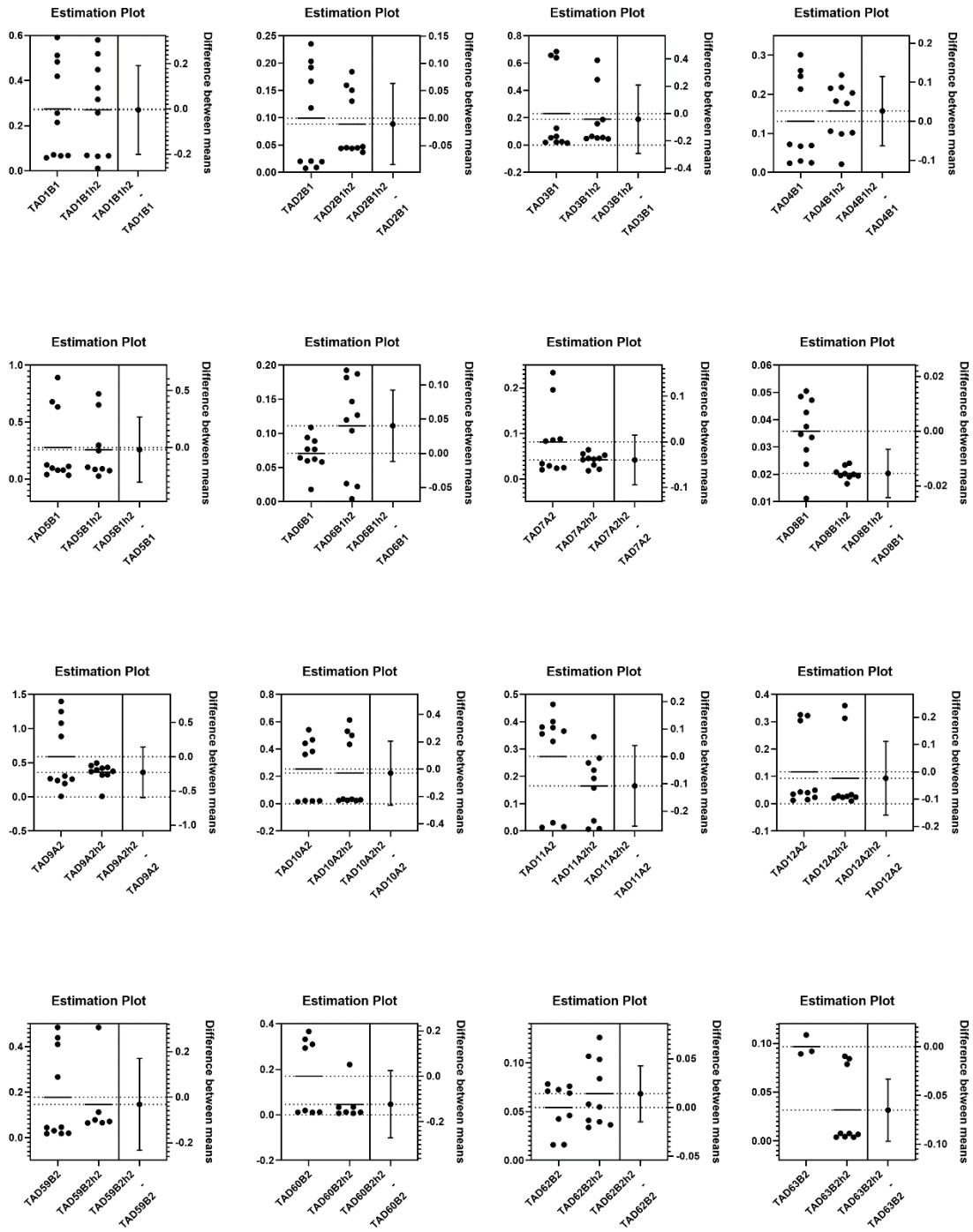
Chromosome 9 homolog pairing was observed in multiple nuclei, **a**. A distinct case in one Chr 9 pairing event, **b**. A cumulative view of matched TADs from figure **b** are shown in **c**. A cumulative view of non-matched paired TADs from the same nucleus, **d**, TADs from figure **c** are grayed out.



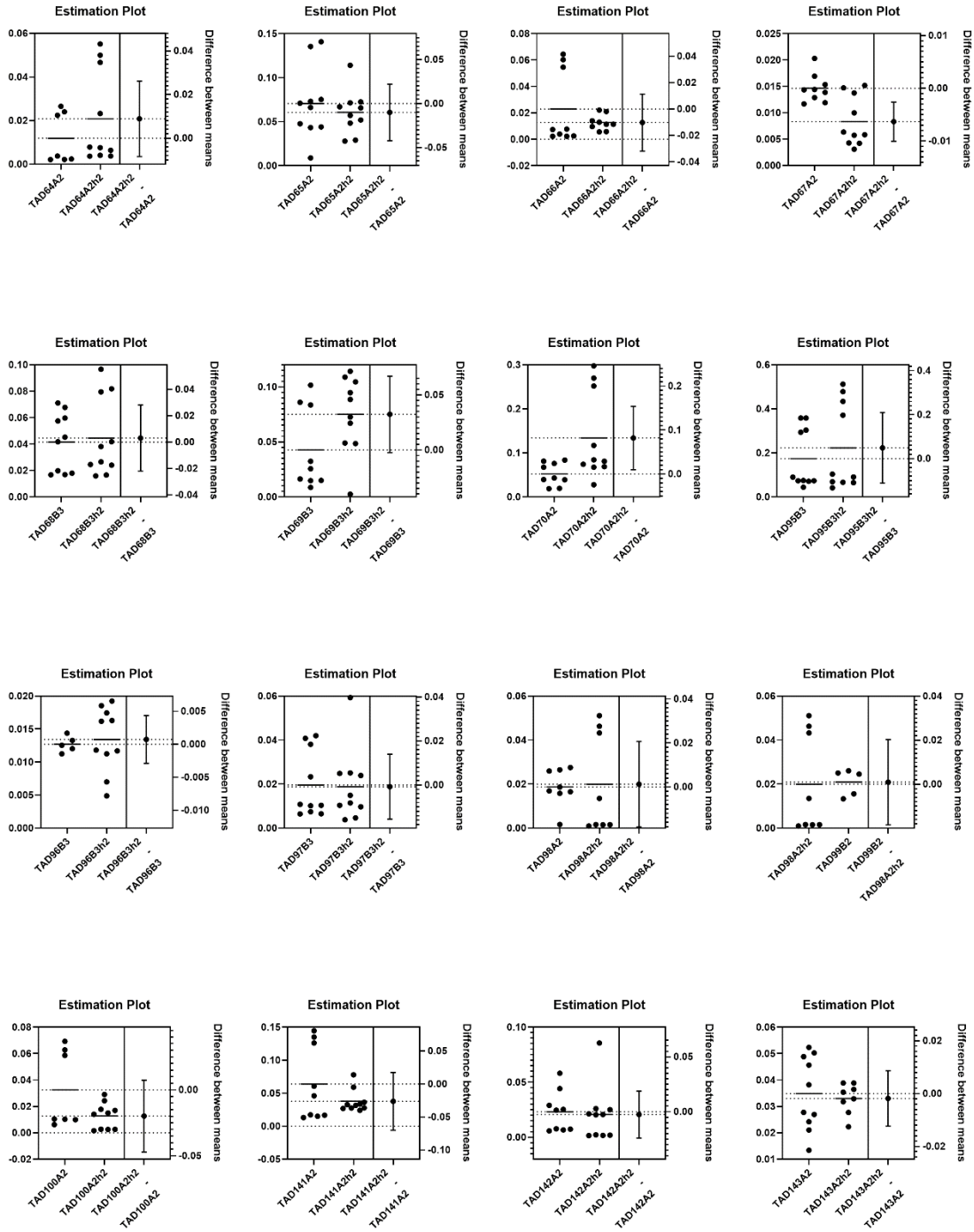
Extended figure 12|Heatmap of interphase homolog pairing events.

Observed in human B-lymphoblast cells (GM12878) showing pairwise TADs of homolog 1 on vertical (light blue) and homolog 2 on horizontal. Approximate location of centromere is indicated by magenta crosslines.

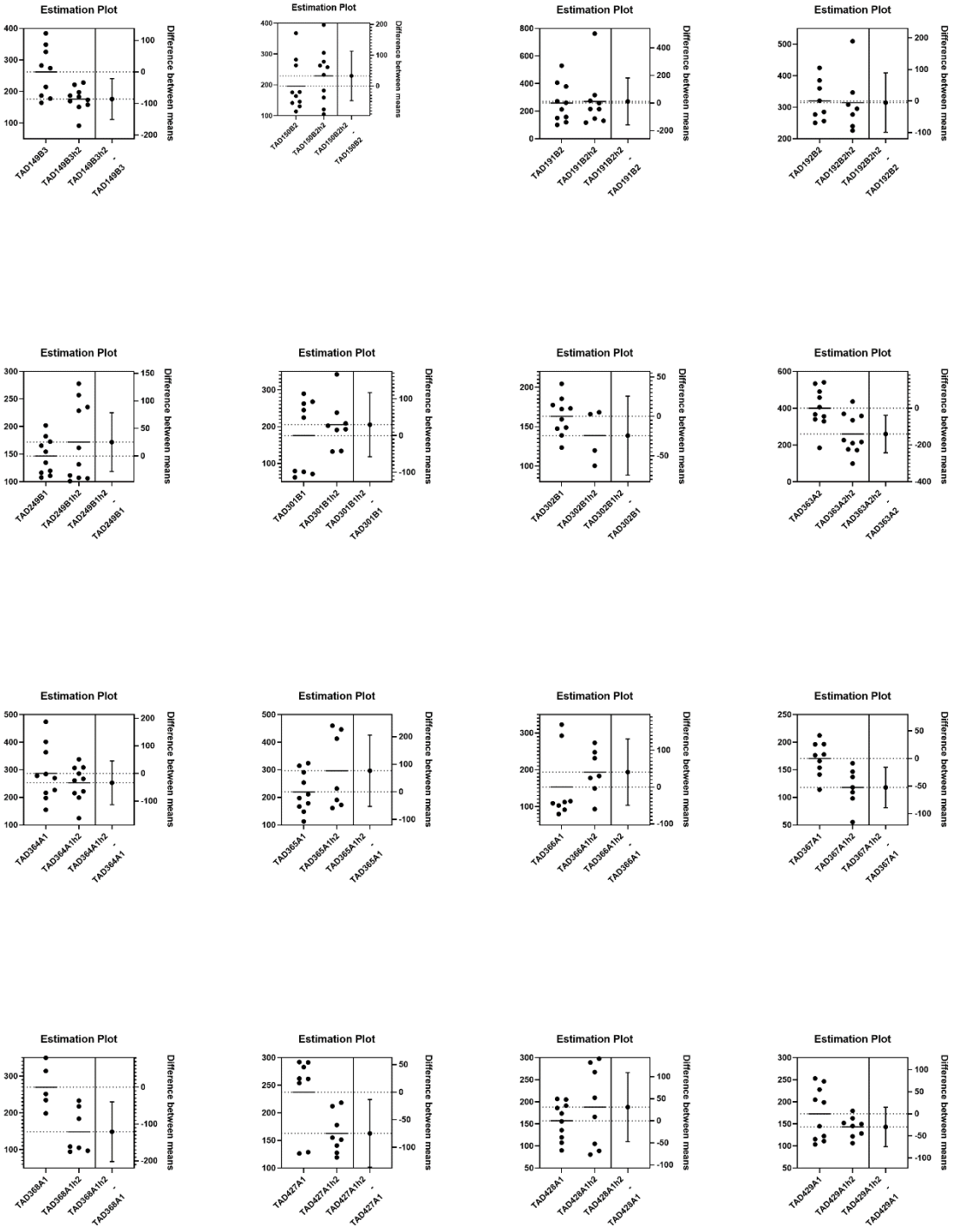
Supplementary data



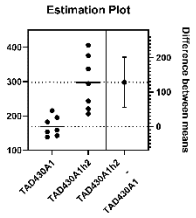
Supplementary data S1: Volume TADs 1-63



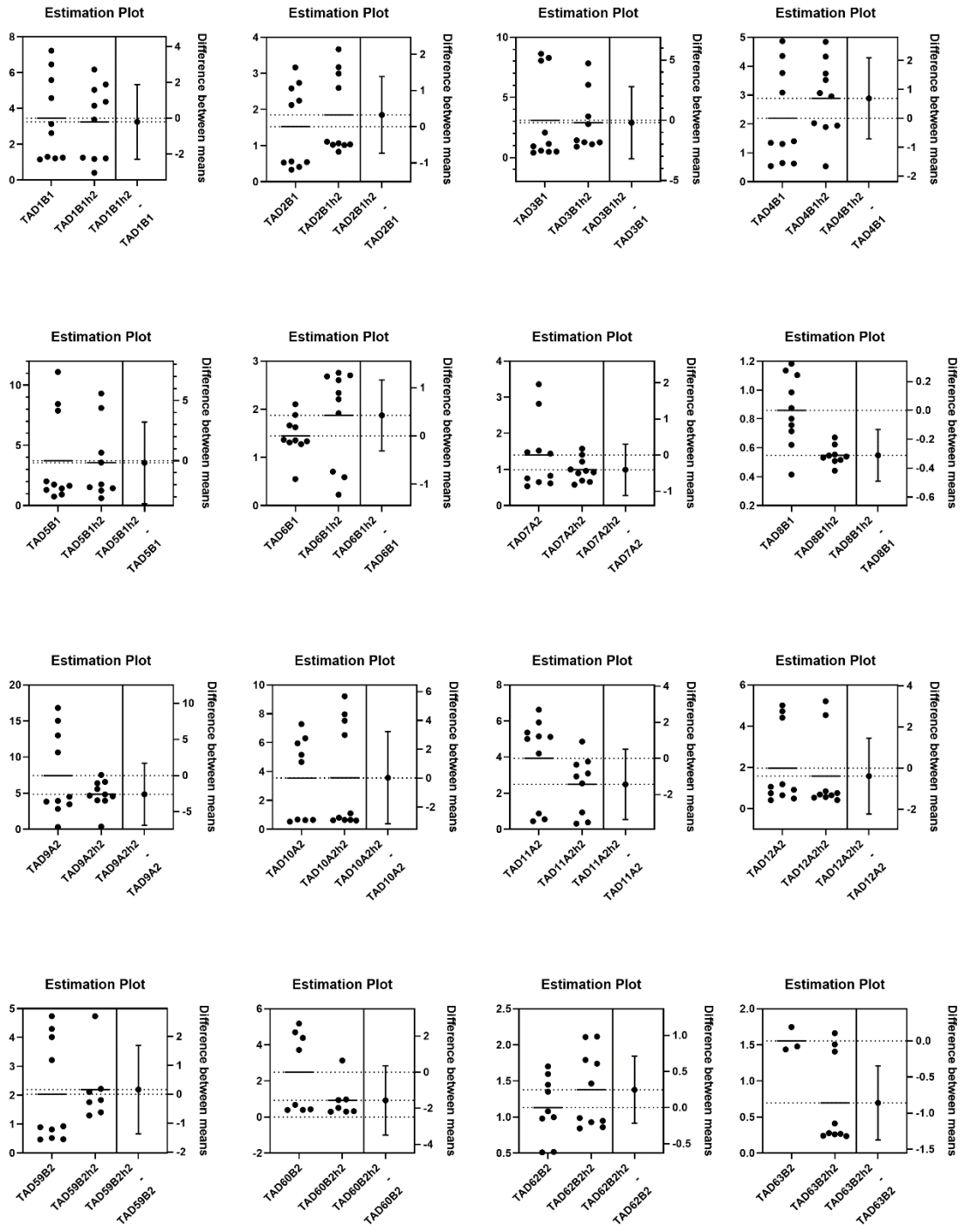
Supplementary data S1: Volume TADs 64-143



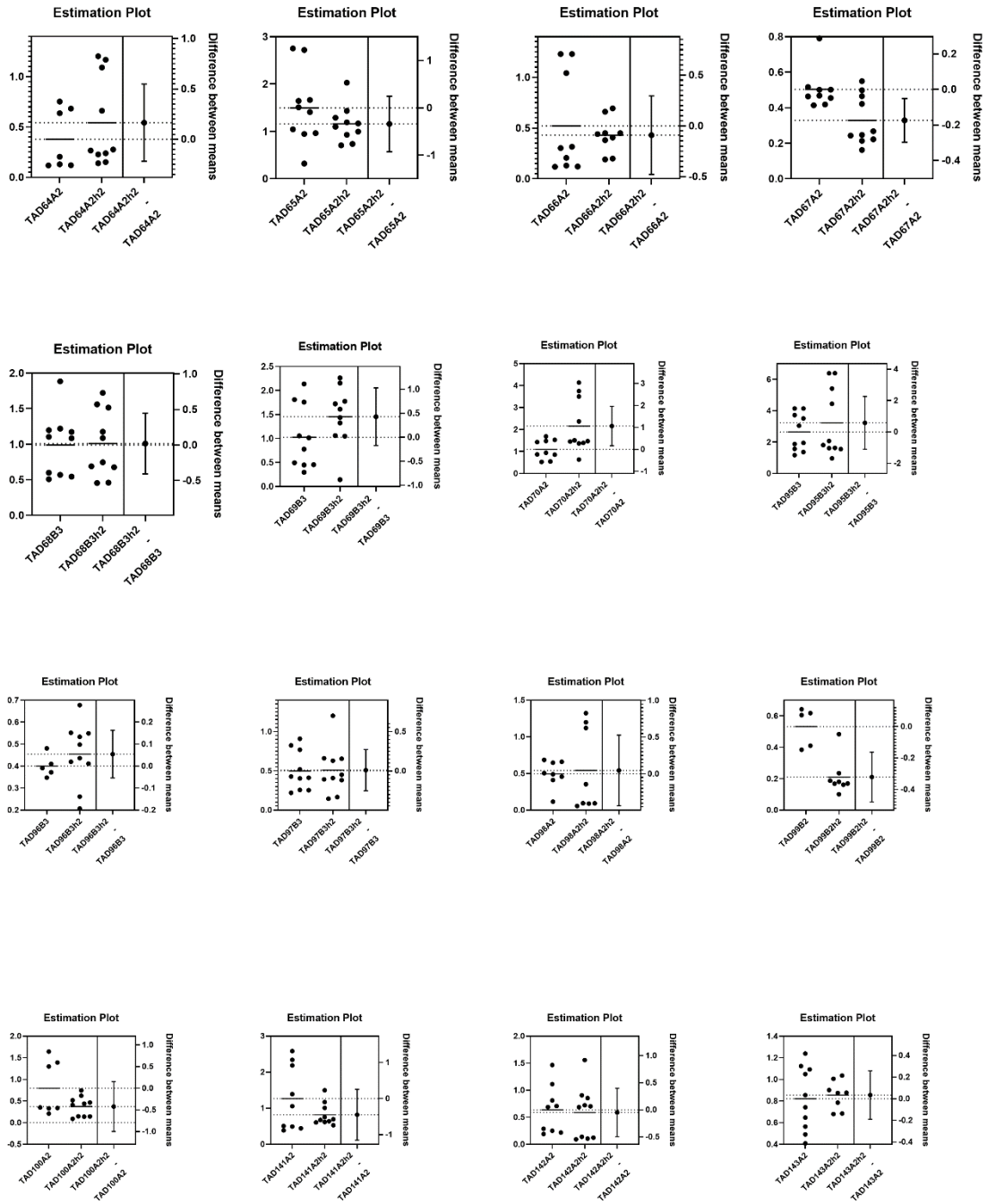
Supplementary data S1: Volume TADs 149-429



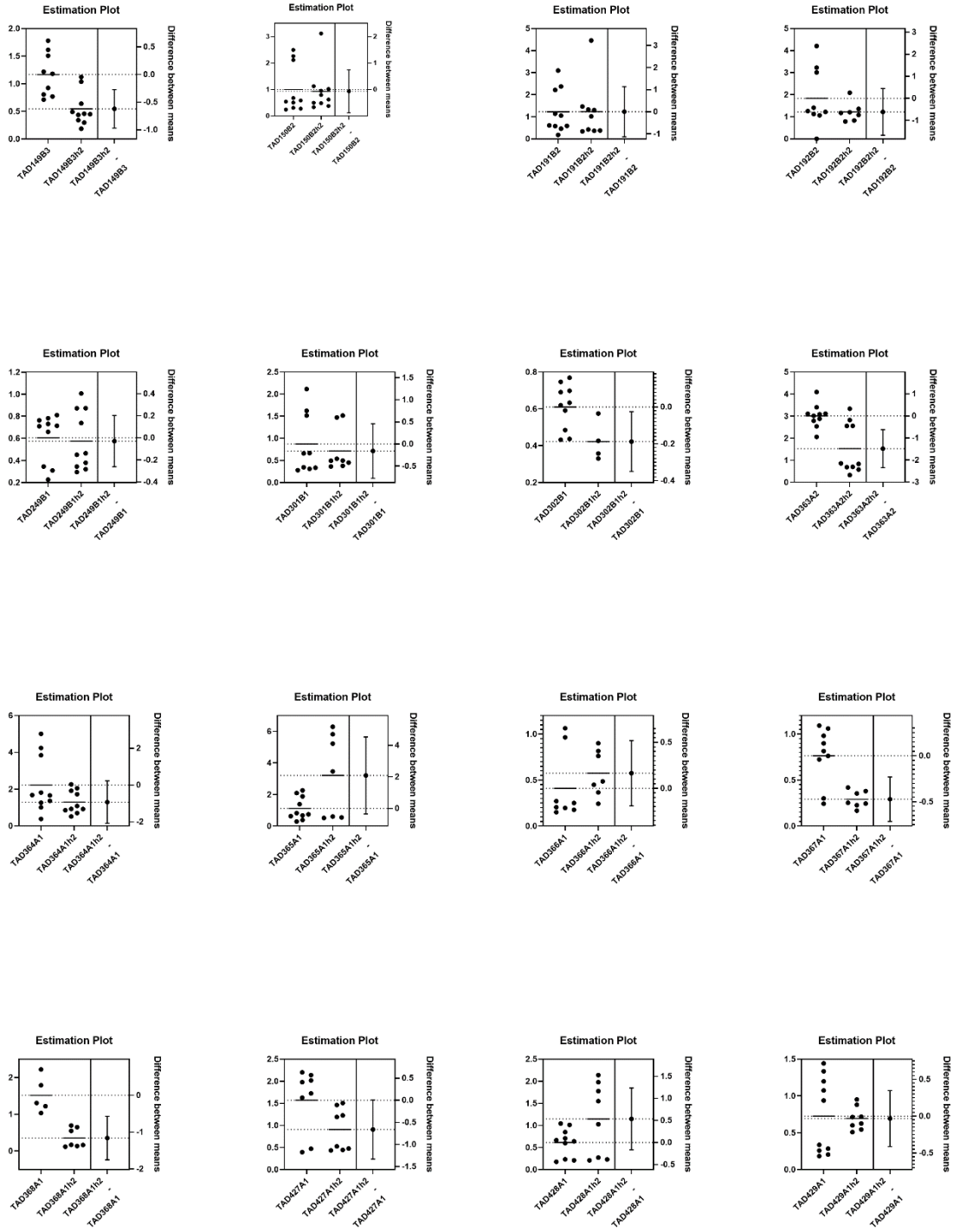
Supplementary data S1: Volume TAD 430



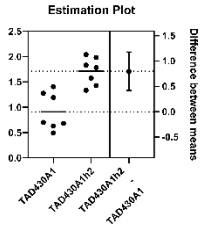
Supplementary data S2: Surface area TADs 1-63



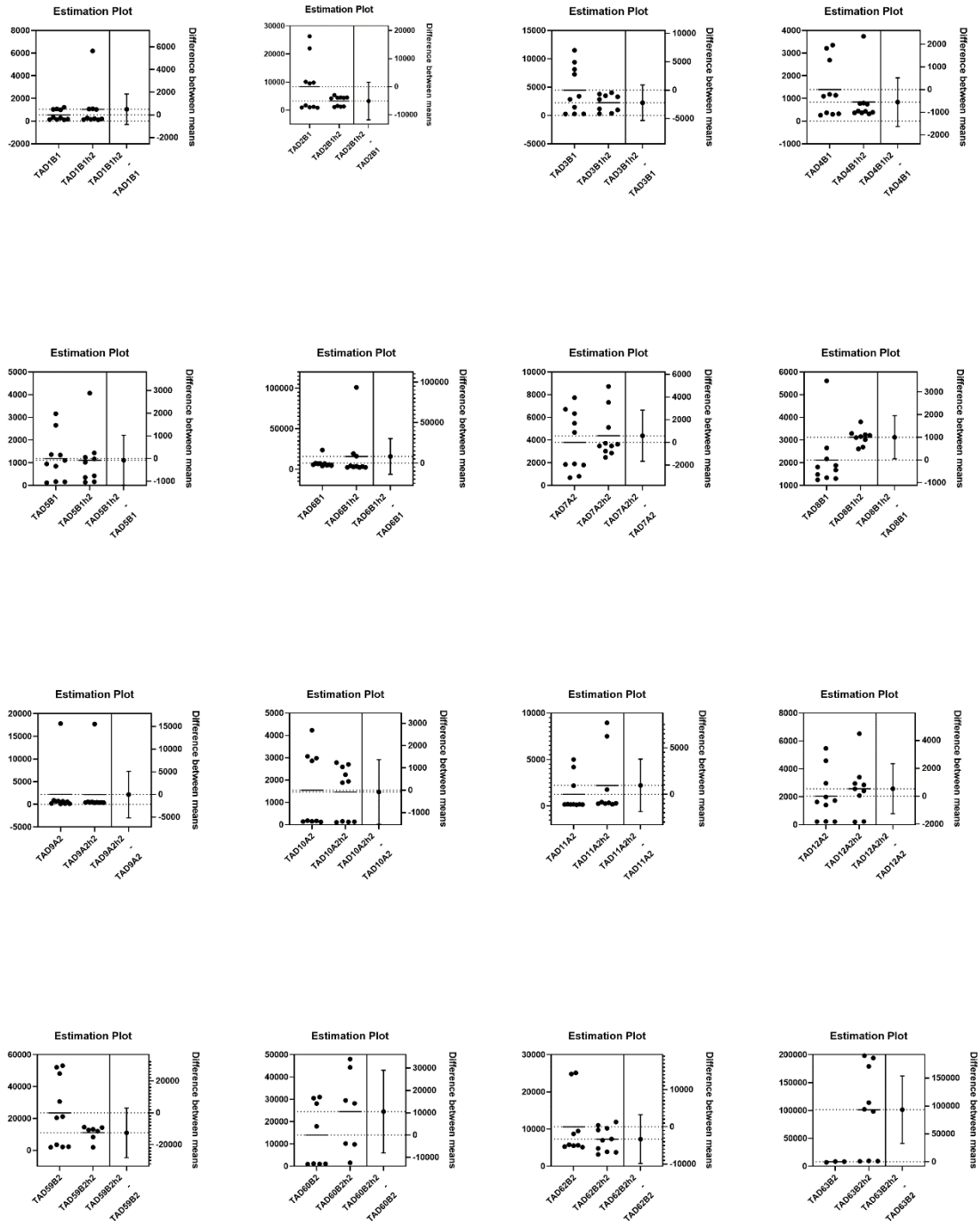
Supplementary data S2: Surface area TADs 64-143



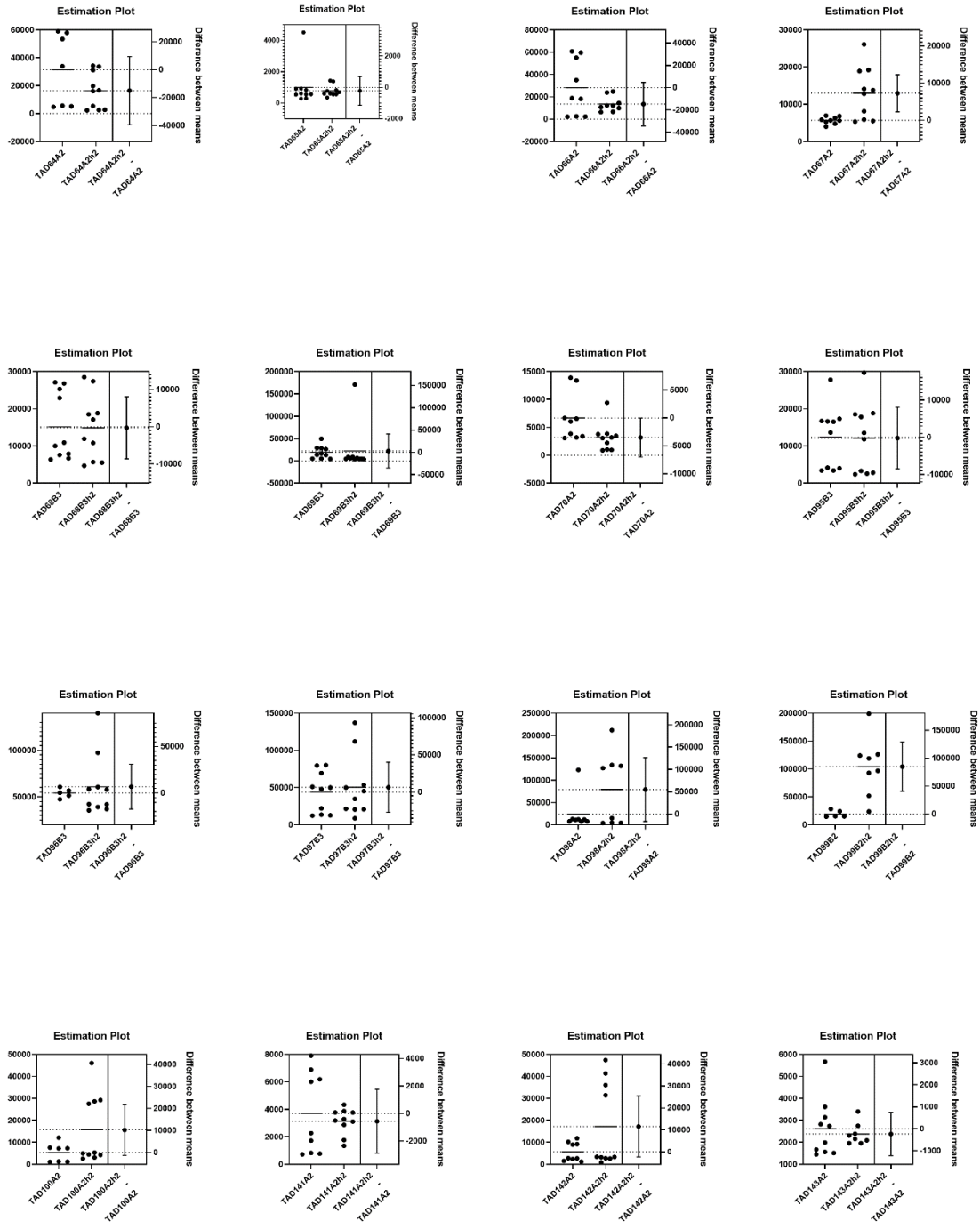
Supplementary data S2: Surface area TADs 149-429



Supplementary data S2: Surface area TAD 430

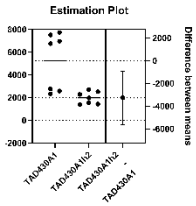


Supplementary data S3: Compaction factor TAD 1-63

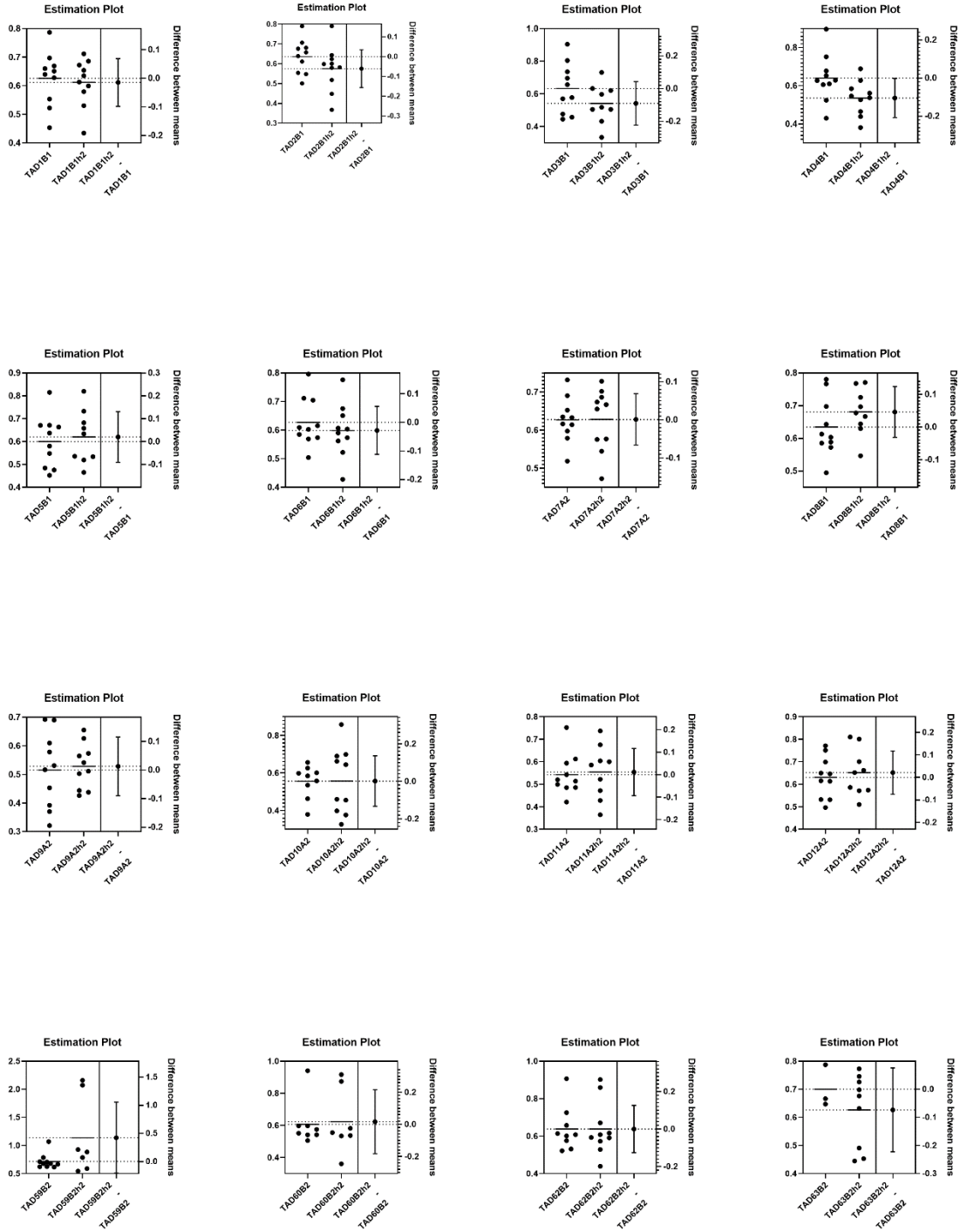


Supplementary data S3: Compaction factor TAD 64-143

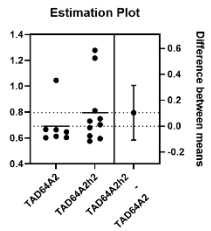
Supplementary data S3: Compaction factor TAD 149-429

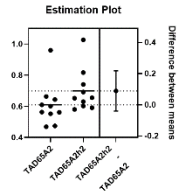


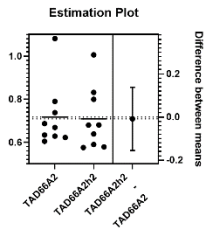
Supplementary data S3: Compaction factor TAD 430

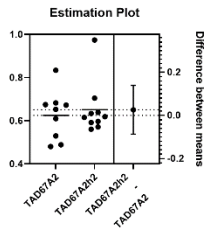


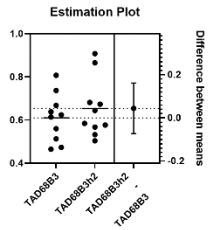
Supplementary data S4: Sphericity TAD 1-63

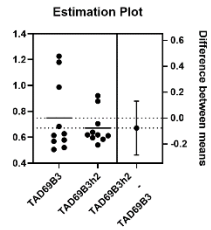


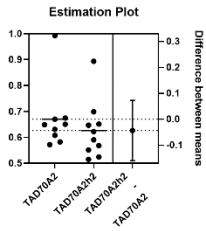


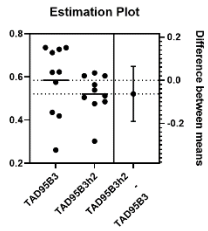


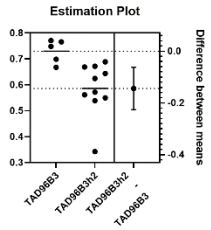


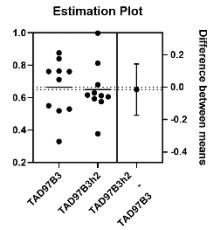


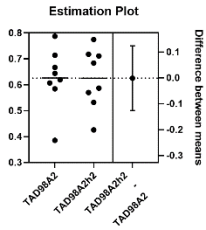


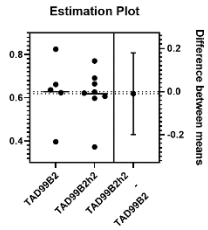


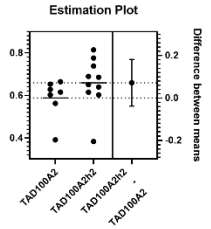


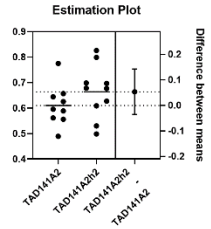


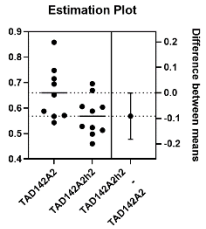


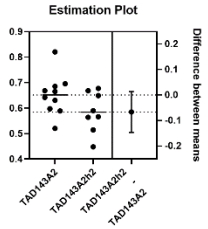




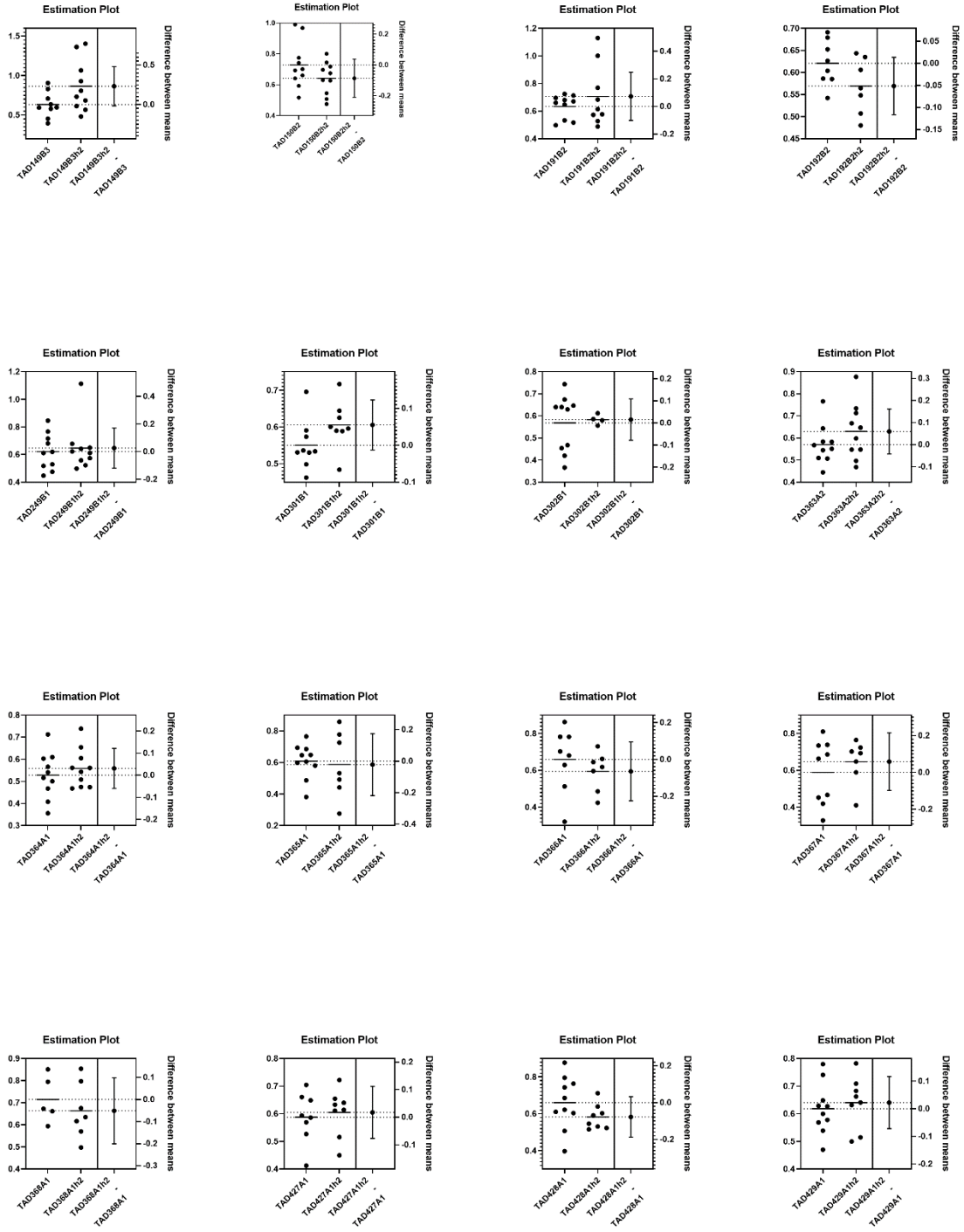




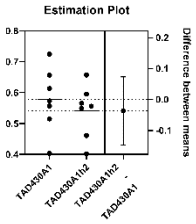




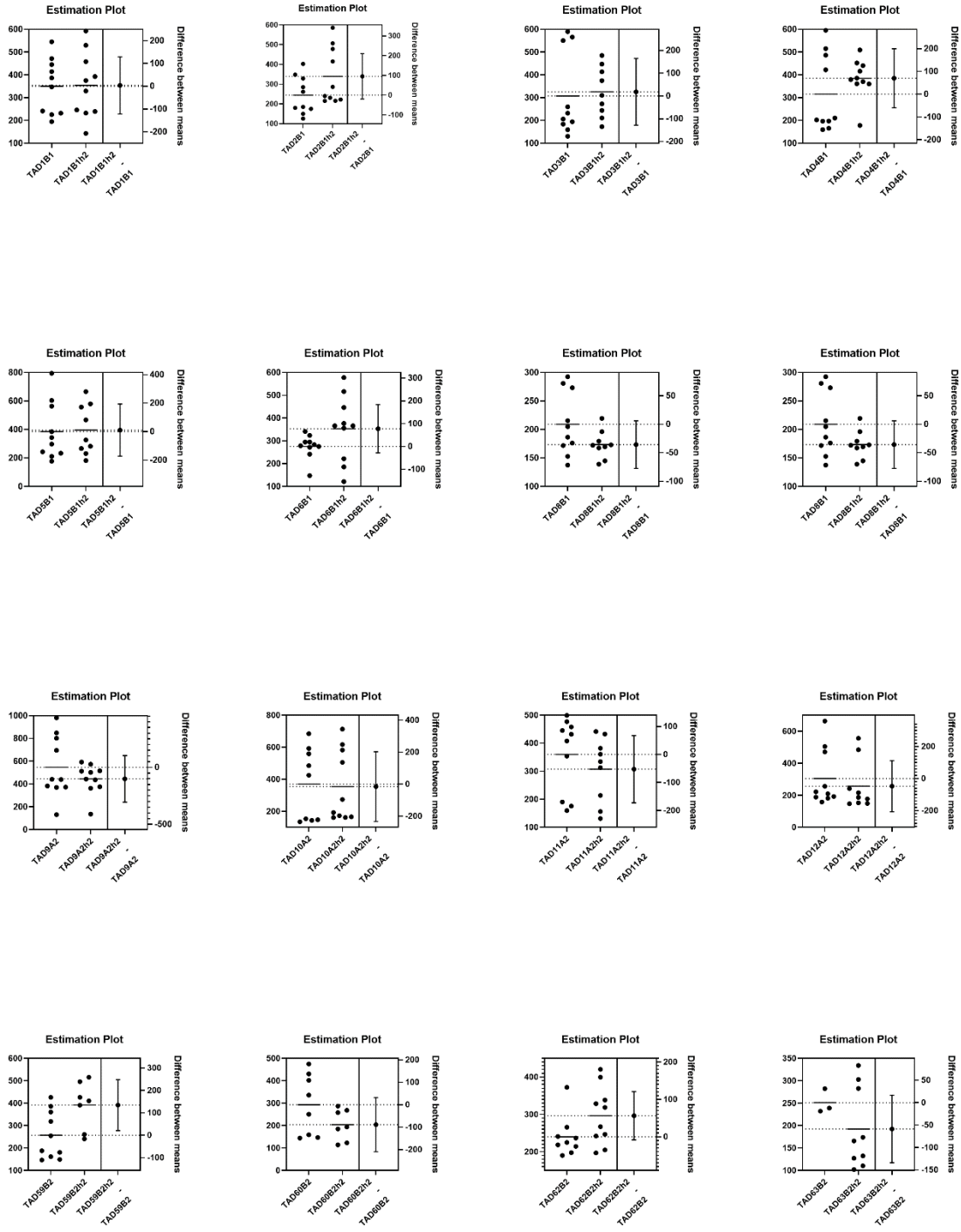
Supplementary data S4: Sphericity TAD 64-143



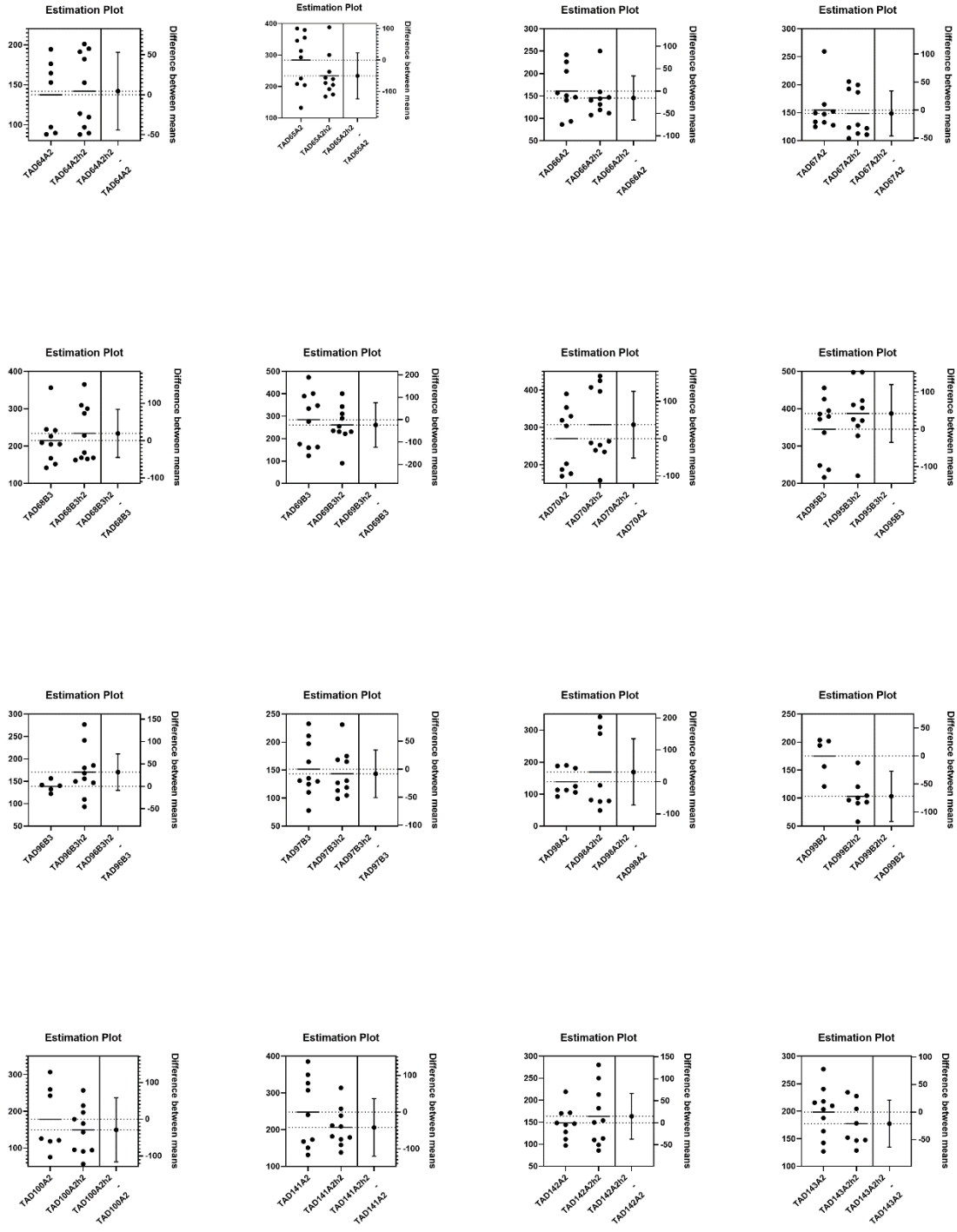
Supplementary data S4: Sphericity TAD 149-429



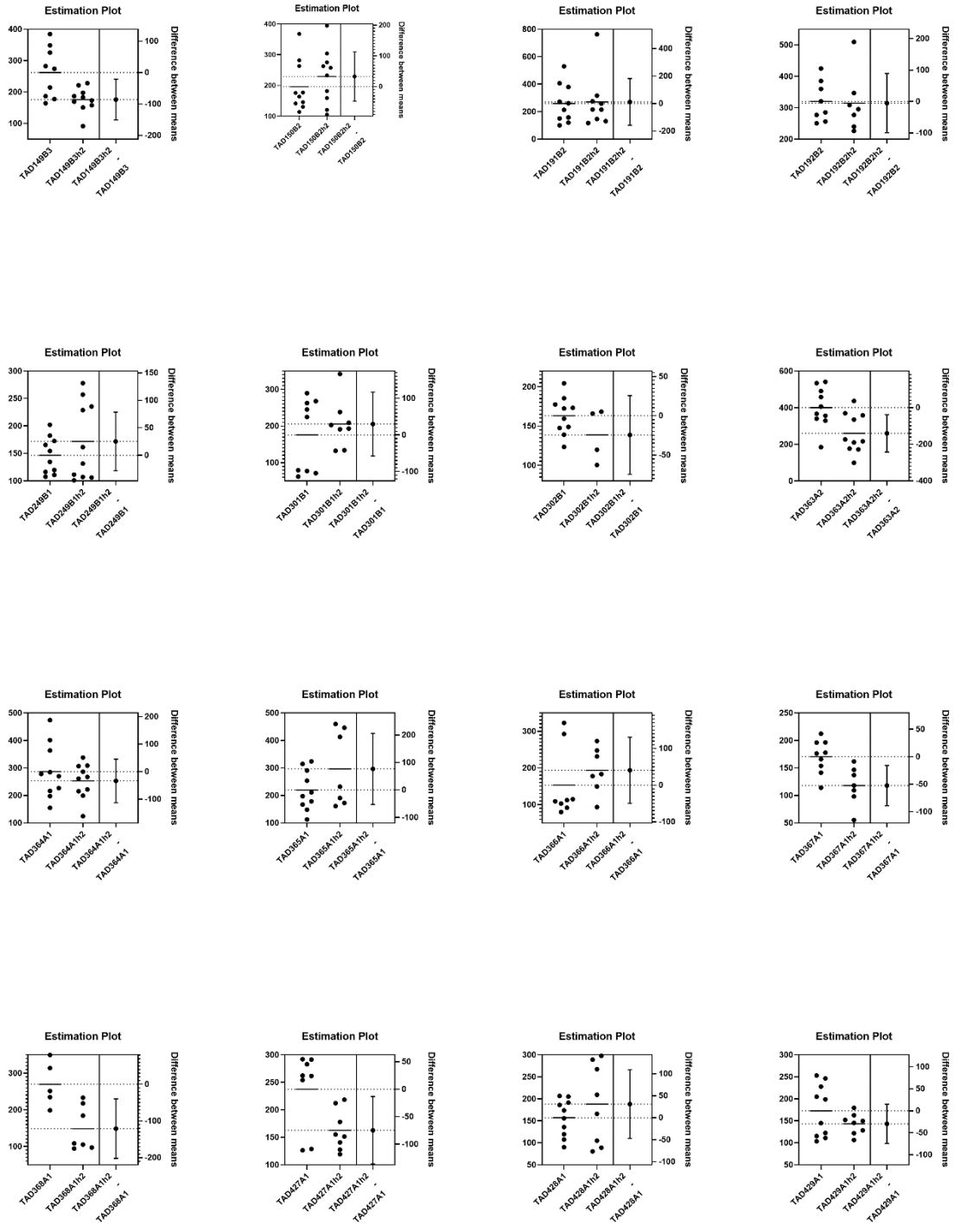
Supplementary data S4: Sphericity TAD 430



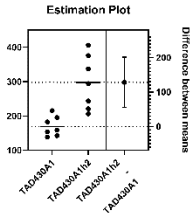
Supplementary data S5: Radius of Gyration TAD 1-63



Supplementary data S5: Radius of Gyration TAD 64-143



Supplementary data S5: Radius of Gyration TAD 149-429



Supplementary data S5: Radius of Gyration TAD 430

CHAPTER 4: DISCUSSION

4.1 Making use of novel technologies to benefit and drive genomic imaging

In this study we were able to leverage two technologies of 3D STORM and Oligopaints to interrogate the genome in unprecedented detail. In addition to this, we were able to evaluate genome organization at the level of topologically associating domains, sub-compartments, and the chromosome territory of human chromosome 9. To our knowledge, assessment on these different levels of chromosome organization and structure concurrently has not been previously done at this scale. through this process we were able to work out some of the drawbacks and some of the benefits of performing such a study. Our desire in doing this study was to assess what was already known about human chromosome organization through our knowledge of HiC, and to test the limits of OligoSTORM in terms of time in acquisition, data storage capabilities, storm analysis pipelines, data and graphics processing of single molecule localization data sets. All of these matters are significant to consider, since there is an ongoing movement to capture larger portions of the genome, i.e. imaging the whole genome using super-resolution microscopy (Gagnon, 2020).

4.2 Curation of HiC targets for the use in Oligopaints library comes with limitations

To date, we have used the molecular method of chromosome confirmation capture method, HiC, to get a glimpse of how the genome is organized. This is the rationale why we wanted to look at TADs, as it serves it's the smallest functional unit of chromosome interaction. In addition, it is at this level of genome organization where transcriptional players interact. The generation up our Oligopaints library came with a few challenges. The first challenge was the classic view of TADs changed when we apply different settings in the Jukebox application. Even though we carried out the manual curation of TADs of chromosome 9 six different times, we recognize that there were inconsistencies based on what display monitor was used, what calibrations were used for the monitor, what resolution we were viewing within the jukebox application, as well as the native TAD calling

feature on the application itself. Because of this we were not able to call some TAD positions Due to disagreements and the different iterations of our calling procedure. Indeed, there are multiple TAD colors that are available now, but still the native TAD calling functions are not consistent between these applications (Durand et al., 2016; Robinson et al., 2018). Another drawback in terms of our Oligopaints library, is the exclusion of some key repetitive element domains or regions. These would include telomeric regions and centromeric regions (satellite DNA) at the very least. We recognize that this may have been an easily missed opportunity to assess how these large regions of repetitive sequences may interface with the genome hierarchy. Certainly, repetitive regions represent an area of chromosome architecture that is overlooked, as it is a difficult to target in HiC and in FISH/microscopy. We also encountered some discrepancies with calling established sub-compartments based on the Rao, et.al. study. In Juicebox, some of the sub-compartment calls were mis-aligned or would be placed in a fraction of a single TAD or multiple sub-compartments what occupy the same TAD. In these instances, these TADs were also eliminated from our final call as they were confounding and did not amount to a significant proportion of the total library.

4.3 Selection and handling of a model system used for imaging is significant

Based on our study goals, we selected the HiC data set from Rao, et.al. that used the human diploid B-lymphoblast cell type GM12878 two gather a remarkable 1 kb resolution. This is a well-established and studied cell type, they are easy to culture and are easy to handle. There are limitations, however, that we should consider of the cell type. First, these cells were transformed with human Epstein Barr Virus (EBV) as a means to alter their proliferative capabilities. As such, these cells have indefinite proliferative potential. The drawback is that there are several studies that suggest EBV may produce aberrant chromatin structure changes at the level of the epigenome (Avolio-Hunter and Frappier, 2003; Day et al., 2007; Lieberman, 2008). We are not fully aware, however, of

how this may or may not transiently contribute to chromosome organization. Another factor that may be worth discussing, is the way GM 12878 cells are prepared in order to go through OligoSTORM. The cells are seeded onto a layer of poly L lysine for a time period in which these cells, with their cell bodies and nuclei, take on a convex shape rather than their native spherical shape. Even though this is a necessary step, we may consider what this dramatic physical transformation may do to our overall 3-dimensional assessment of the spatial occupancy of different regions of interest in the genome. If we were able to image GM 12878 cells in their native state, how would this reflect in 3D STORM analysis? An interesting future proposition could be to image at super-resolution nuclei of different cell types that possess dramatically differences in their gross appearance, such as B-lymphocytes compared to multi-lobular neutrophils. Another avenue that is worth some thought is the manner in which cells are fixed to preserve the three-dimensional structure of the genome. It is worth noting that there are multiple ways in which cells can be crosslinked prior to a 3D FISH procedure. In our case, we elected to fix in formaldehyde, as this has been at work-horse chemical for the use of FISH procedures, however, there is evidence that time and fixation it's critical two per serving cellular compartment integrity. In addition to this, and taking into account the extended auto-microfluidics and OligoSTORM session detailed in this study, an important consideration was the structural integrity of our samples. Indeed, we did observe a slow depletion of number of particle count based on sample time in STORM imaging/switching buffer (FOV1->FOV6 in the 52 TADs imaging sequence) and degradation obtained over consecutive days of imaging. The question here, remains, at what point is the imaging session deleterious or at what point can we rely on or not relay on the accuracy of our data? Is it plausible that a standard operating procedure can be made where pulses of sample fixation can aid in a multi-day imaging session such that structural integrity is held in place—and what fixative will work best in this scenario? There is room for improvement when these matters are concerned.

4.4 Multi-day 3D STORM requires a robust infrastructure in computational processing and data storage

The use of single molecule localization (SML) microscopy has transformed the way we view biology. To clearly visualize parts of a cell beyond the diffraction limit of light has truly been an exhilarating experience. In this study, we have pushed the limits of SML acquisition in terms of total time in imaging and the total particle count acquired. We were encouraged that the Bruker Vutara 352 SML microscope used in this study performed optimally throughout the 16-day course of imaging. But we did not fully expect the amount of data that would be generated through this imaging session (imaging of 5 sub-compartments and 52 TADs). All told we generated dozens of billions of single molecule localizations, and we had the daunting task of filtering each one of those particles. This quantity of data took a toll the computational power that we had on hand—from a CPU and GPU standpoint. Computational tasks that normally take one or two minutes to perform, in our case took hours or even days to complete. Because of this, we were extremely limited and how the data was processed post imaging. We came up with creative means to manipulate the data by separating it out into data subsets. Even so manipulation and analysis of data was extremely taxing. Taking into consideration that an overarching goal is to image the whole human genome using super resolution microscopy, serious consideration for data manipulation and analysis must be at the forefront of discussion and implementation.

4.5 Ways to improve on 3D STORM data analysis and infrastructure for genomic imaging

The first point to consider is data storage infrastructure. We have imaged in this study a small portion of the genome, yet it occupied over 30 terabytes of storage. Serious attention should be taken in terms of total capacity of storage as well as lossless compression or compartmentalization of data such that we can minimize total storage

capacity needed. The second point to consider is real time handling of SML data. A possible field to explore is that of automation through machine learning or artificial intelligence. There is an open window of opportunity for machine learning to be used within the workflow of single molecule localization filtration and analysis. For example, for a large data set as we have dealt with, our first obstacle was to find the gold nano urchin fiducials (beads) within the three-dimensional field of dense particles. Given that these beads have a unique signature of particle density and particle number, the task of drift correction through these signatures can be a simple and defined task to train machine learning algorithms. Because of the slow image acquisition nature of STORM imaging, however, it would be seemingly difficult to gather a number of data points sufficient for training. A possible way in which we can overcome this is to implement mock, multi-day experiments where fiducial beads are embedded unto coverslips without any biological samples at high density. This method may possibly give us a way to generate the number of data points necessary for machine learning training application for the purpose of drift correction automation. Another way in which automation can be helpful in the analysis of SML data, is through selection of clustering algorithms in order to filter through which localizations belong to our regions of interest and those localizations that do not. Using density based spatial clustering of applications (DBSCAN), the clustering algorithm that we implemented in this study, there are defined and discrete tasks within that analysis workflow that can be automated. For example, the determination of maximum particle distance and the number of minimum particles to form a cluster can be predefined such that, in the end of automation, there would be multiple iterations of clustering parameters available. In other words, the output could be multiple data tables of clusters, which can visualize one at a time one data table output versus others. This type of process would be extremely helpful because it gives multiple options for the researcher to select the best fit, the best representation of data based on clustering of SML particles.

4.6 Making the case for using 3D STORM-based super-resolution microscopy over HiC

A goal in our study was to attempt to answer open questions of genome organization with the use of microscopy. We planned to connect how well HiC data would recapitulate with imaging data. The major differences between interrogating the structural genome through microscopy is that it is a single cell process, whereas the most reliable HiC datasets are performed on pooled data. Another benefit of using microscopy is that we are able to observe striking differences between the structures of chromosome territories or homologs. Our decision to use 3D STORM was because this is a super-resolution technology that allows for an appropriate image acquisition compatibility with the depth of field we required to image the whole span of chromosomes on the z-axis. Unlike the latest innovations in point spread function centroid based super-resolution extrapolation, we were interested in the space-filling qualities of 3D STORM. For this reason, we focused on structural features such a volume, surface area, sphericity, compaction factor, and radius of gyration.

4.7 TAD volumes and surface area follow the same measured trends

By doing analysis of TAD absolute volumes showed us substantial variation and very little correlation to genomic size of individual TADs. For this reason, we used a normalized measurement of volume against the genomic size of each TAD. Because we did this, it represented a better size comparison between TADs as volume/kb (genomic span) (**Figure 8)**. This remained true when we normalized volumes to genomic span of these TADs. This observation does not support previous studies that showed a correlation between volume of TADs and genomic size (Bintu et al., 2018; Boettiger et al., 2016; Nir et al., 2018). This could be because of selection bias, as very few unique TADs were imaged in these studies. Another difference between our study and others is that we have assessed only using 3D STORM a more diverse set of regions of interest along parts of a

whole chromosome—allowing for the capture for up to 52 unique TADs. This gave us the unique position to compare multiple TADs from different regions of a whole chromosome. A question here is, do TADs in the same chromosome region have significant differences in volume or do they follow a predictable pattern? An interesting finding based on 3D STORM-based physical feature analysis is that TAD surface area followed the same quantitative trends in measurements as our volumetric measurements. What this could mean is that the surface area to volume ratios of TADs is close to 1. In future analysis of this type of data, a surface area:volume measurement could serve as a metric or a readout to assess structural differences between genomic features of physiologically altered cells when compared to normal cells.

4.8 Genomic feature sphericity displays measures that are tightly packed between TADs

An unexpected finding that was observed in this study was the measurements of TAD sphericity (**Extended figure 8**). What we found was a TAD sphericity that oscillated groups of consecutive TADs with a predictable amplitude. This is truly a striking finding which suggests that how spherical TADs are *in situ* is tied to their genomic position. To our knowledge, this is the first study to show this feature when looking at multiple sequential domains. This finding may support the concept of DNA taking on a similar chemical behavior consistent with a complex polymer folded into a confined space. In general, the hypothesis states that the topology of DNA can be predicted based position due to the intermolecular forces that are involved between complex structure of charged macromolecules (Dalal et al., 2005; Li and Miramontes, 2006). It is plausible that there might be some interplay between this observed spherical periodicity and the location of different groups of sequences, repetitive or otherwise, or epigenetic differences between consecutive TADs. What this implies is that we may have serendipitously found, in TADs, the scale of DNA topology at which these oscillations occur.

The measures of radius of gyration and compaction factor did not yield any remarkable results in terms of correlating these physical measurements to genomic size of TADs, or to their compartment and sub-compartment assignments. Bases on measured compaction factor, we did observe the same oscillation pattern present in sphericity, but with the difference of relative higher amplitude in measure. One other observation that we made was that the compaction factor/density of our TADs followed an unexpected trend when compared with genomic scale of these TADs. Larger TADs had a tendency to be less dense than smaller TADs. A possible explanation for this might be that larger TADs have more pockets of compact DNA versus smaller TADs. These two observations of radius of gyration have not be quantified, and it is unclear whether this is a result because of an artifact of STORM specific imaging or if this is a true feature at the level of the genome. To answer this question of compaction more studies with a higher number of data points will have to be performed to test this.

4.9 TADs are more different between sub-compartments rather than A/B compartments

When analyzing these TAD physical features in the context of compartments and sub-compartments we were able to extract significant differences between these features. Looking at all data points of TAD volume separated into their respective A or B compartment, there is no significant difference between these. However, when these same data points are compared by their sub-compartment assignments, we do see significant differences between these groups for volume and for surface area. However, in terms of sphericity, TADs can be distinguished at the compartment and the sub-compartment level. In our study, the measure of sphericity was the most consistent and predictable for the comparison of compartment functional status. The relationship between epigenetic condition and to what extent the structure of genomic domains, such as TADs, resembles a sphere is understudied. Understandably, our work here may be the first to

report on this particular sub-compartment-TAD correlation with functional compartment assignment. This may be because sub-compartment assignments of TADs were done at the level of the primary Oligopaints library, thus, contributing to inherent bias originating from the HiC dataset used to design the library. Even so, this would still tie back to the way sub-compartments were defined. According to Rao, et.al. the sub-compartments were in large part defined by epigenetic status, each with their own histone modification patterns. What this may indicate is that TAD volumes can be distinguished on the level of sub-compartments and thus the level to which these TADs are transcriptionally active or not, could be prudent to correlate TAD assignments to their position in the nucleus—that is not an analysis that was not performed in this study. We know from previous work from multiple groups that nuclear position of chromatin is determinant of histone modifications status (Steensel and Belmont, 2017; Vilarrasa-Blasi et al., 2021). This would be an intriguing motivator to corroborate a TAD-epigenitic-nuclear position enrichment study. To do this, a cell type agnostic coordinate system of nuclear position would have to be developed. A possible way to do this is to treat each nucleus to be contained in a perfect sphere, where its radius is equal to the distance from its volumetric centroid to the furthest distance point on the nuclear surface. This sphere can then be segmented into 100 nm grids in x, y, and z axes; thus, creating a conserved cubed-sphere geometry that can be transformed arbitrarily to reflect the context of what nuclear structures are being studied.

4.10 TAD-to-TAD observations reveal striking structural differences between cells and homologs

Another unique opportunity that we had in doing the study was to do comparison of TADs on an individual TAD-to-TAD basis. We were able to evaluate the shapes of TADs in the context of homologs, sub-compartment assignments, and between cells. The rationale for doing this was because even though the same a matched TAD would share very similar volumes and their respective homolog; we did observe striking differences in

their overall alpha hull shapes. In addition to these three-dimensional shapes, we also observed that the chromatin density patterns are also very different between homolog matched TADs of the same cell. Although a quantitative analysis was not done, we can speculate that the differences in chromatin density may be reflective of the overall compaction difference is we observed in homologs of the same cell. Another shape feature that was very striking what's the appearance of *drumstick* and *dumbbell* shaped TADs (**Figure 9**). the *drumstick* shaped TADs were only reserved to TADs of the B compartment. The *dumbbell* shaped TADs were reserved to TADs of the A compartment. The most peculiar fact about these unique shapes was the way they were distributed. In the case where *dumbbell* or *drumstick* shaped TADs were observed, they were found and distributed on one but not both homologs of chromosome 9. To which homolog the *drumsticks* and *dumbbells* were distributed was independent of each other. The exact reason for the appearance of these dramatic shapes is unclear, we can contemplate that it could be connected to transcriptional status, or to a loop extrusion event—which case, the two may be connected. Much like the "stripes" that other groups have observed in HiC contact heat maps, it could be that these structures are in a transitional state reflective of CTCF and cohesin position (Yoon and Vahedi, 2021). Therefore, what we may be observing are very large loop extrusions within TADs. A possible way in which we can test for this it's to co-stain with immune markers/primary antibodies for cohesin such as RAD 21 on the current Oligopaints TAD library to evaluate the positioning of these components in respect to each other. If cohesin is involved, then we can expect it to be localized along the "pinch points" of our TAD masses. Because of the compartment B connection to only *drumsticks* and the compartment A connection to only *dumbbells*, the loop extrusion hypothesis may not hold up, as we would expect to observe a more expanded repertoire of shapes connected to the dynamic process of loop extrusion (**Figure 9**).

4.11 Chromosome 9 homolog pairing was only possible through 3D STORM

Indeed, one of the breakthrough findings of this human chromosome 9 study is the observation of homolog pairing in interphase cells. The overall design of the Oligopaints library is one of the reasons that made this possible. This is because Oligopaints probes incorporated targets that would enable us to view each significant domain and compartment that compose a more complete view of chromosome structural hierarchy. In addition to this, we can make the case that 3D STORM and the combined OligoSTORM exercise was instrumental in uncovering these TAD pairing events. To date, the combined technology of specific genome targeting and microscopy resolution had not intersected to make this type of observation possible. This finding of TAD pairing is significant because, to our knowledge, this is the first time that homolog pairing has been detected in interphase mammalian cells *in situ*. Homolog pairing within interphase has been observed in drosophila somatic cells where pairing events were characterized by the amount of minimum distance to which homologs came together known as *loose* and *tight* pairing, respectively (Abed et al., 2019; Erceg et al., 2019). In our work we surveyed pairing on a whole chromosome territories context, with the added assessment of TAD positions within the span of the chromosome. Therefore, we had the framework to visualize multiple areas of enrichment where pairing occurred. What we have observed are chromosomal regions of enrichment in which pairing events happened (**Figure 11, figure 12, extended figure 12**). These hotspot regions are the telomeric ends of both *p*-arm (~200kb from the terminal end) and *q*-arm (~50-100 kb from terminal end), as well as regions in proximity to the centromere; that are gene poor, and TAD-depleted in the middle of chromosome 9. Because of where on chromosome 9 these paired TADs are found; it is convenient to postulate that their proximity to large regions of repetitive sequences may point to a mechanism in which pairing happens. Our knowledge to date has only characterized and observed pairing of homologs during prophase of meiosis in mammals. Although the

pairing mechanism is a largely understudied area in biology, there has been strong correlations of structural mechanistic players of telomeres, centromeres, histone modifications, and small non-coding RNAs being involved in the pairing process in meiosis—can it be similar in somatic cells? While we did not observe a complete alignment of interphase homologs in our study, we did observe pairing events in regions of chromosome 9 that are, at least, suggestive of these mechanistic players being involved due to their proximity to hotspot TADs. One of the main reasons homolog pairings is understudied is because they are difficult to find in interphase nuclei. With the recent explosion of and implementation of several genome targeting and microscopy techniques, it would be just a matter of time before a consensus is made in terms of human homolog chromosomes pairing in interphase cells. A possible proposal to do this would involve whole genome imaging by Oligopaints, targeting by chromosome territory. Doing this process would have to be performed to answer the first question of whether there is enrichment of any territories that pair to one another. This can be achieved using diffraction-limited methods like spinning disk confocal microscopy or conventional widefield fluorescence microscopy—to influence precision and speed with the goal of imaging thousands of loci in an efficient manner. To obtain qualitative data of pairing events, this case would require OligoSTORM or another single molecule localization-base method. Just as we have done in our TAD study, techniques such as STORM has the inherent ability to provide the "space filling" quality that is needed to make more defined assessments of pinpointing the finite regions that are in close proximity; at a resolution not possible with other fluorescent-based methodologies.

4.12 Future directions to explore close proximity chromosome homologs in human cells

Indeed, further exploration into past studies into pairing phenomena are needed to pursue further the prospect of understanding a mechanism by which these pairing events

happen in somatic cells in mammals. Our observations here can be characterized in two ways. Homolog pairing of matched TADs and homolog pairing of non-matched but adjacent TADs. To our knowledge in the case of pairing events that occur in Drosophila, homologous chromosomes align themselves where the locus from one homolog is in line with its counterpart in the other homolog. Our observations do not perfectly fit this mechanism of homolog positioning where we do observe locus mismatches across one homolog from another. Based on this, an alternative mechanism should be explored to describe our observed homolog "kissing" events, rather than pairing.

The reasons for why human chromosome 9 is involved in these homolog kissing events in somatic cells is perplexing. There are multiple factors to consider such as, are these events chromosome specific, are they cell type dependent, what are the biological precursors needed to promote homolog kissing? We first should explore the hotspots for the kissing events. A next step would be to explore the chromosome territory position enrichment of all human chromosomes. This can be performed at a lower resolution as the assessment would only involve gross positioning in the nucleus. This assessment will be needed to be performed in different cell types to disclose what tissue differences there may be. If there are enrichments of chromosome territory position tied together with homolog kissing events, a last step can be done using low coverage Oligopaints and conventional microscopy to do a comparative study of cells of physiologically normal cells to aberrant cell types. Even though this type of assessment would be significant and informative, to the point that it could be used for medical diagnostics, it would not shed light on the mechanism of how these kissing events occur. As a future direction of exploration on this topic and given the homolog kissing distribution in the terminal ends of chromosome 9 arms, we can hypothesize that these kissing events are connected to regions of large repetitive sequences of the genome (**Figure 13**). To test for this, super-resolution microscopy is necessary as these events should be characterized at the highest

resolution possible. The biological areas of exploration are almost limitless because it is still unclear the prerequisites necessary for these events to happen.

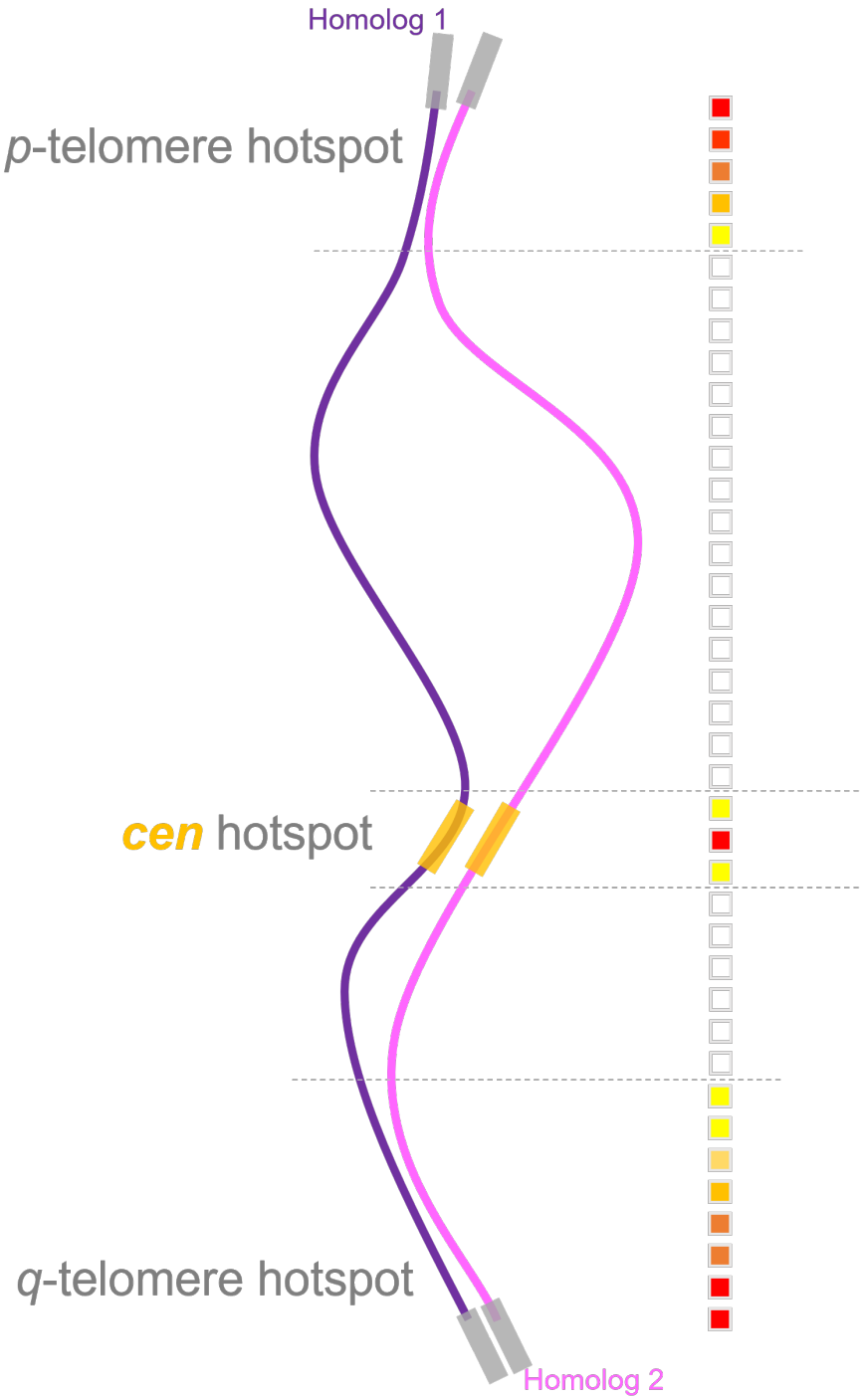


Figure 13|Human chromosome 9 homolog kissing

Proposal of chromosome homolog kissing. Mechanism involves the alignment based on large regions of repetitive sequence. In this case, telomeric regions (grey) and centromeric regions (yellow).

Bibliography

Abed, J.A., Erceg, J., Goloborodko, A., Nguyen, S.C., McCole, R.B., Saylor, W., Fudenberg, G., Lajoie, B.R., Dekker, J., Mirny, L.A., et al. (2019). Highly structured homolog pairing reflects functional organization of the Drosophila genome. Nat Commun *10*, 4485. https://doi.org/10.1038/s41467-019-12208-3.

Avolio-Hunter, T.M., and Frappier, L. (2003). EBNA1 efficiently assembles on chromatin containing the Epstein–Barr virus latent origin of replication. Virology *315*, 398–408. https://doi.org/10.1016/s0042-6822(03)00561-0.

Bates, M., Blosser, T.R., and Zhuang, X. (2005). Short-Range Spectroscopic Ruler Based on a Single-Molecule Optical Switch. Phys Rev Lett *94*, 108101. https://doi.org/10.1103/physrevlett.94.108101.

Beliveau, B.J., Joyce, E.F., Apostolopoulos, N., Yilmaz, F., Fonseka, C.Y., McCole, R.B., Chang, Y., Li, J.B., Senaratne, T.N., Williams, B.R., et al. (2012). Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. Proc National Acad Sci *109*, 21301–21306. https://doi.org/10.1073/pnas.1213818110.

Beliveau, B.J., Kishi, J.Y., Nir, G., Sasaki, H.M., Saka, S.K., Nguyen, S.C., Wu, C., and Yin, P. (2018). OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. Proc National Acad Sci *115*, E2183–E2192. https://doi.org/10.1073/pnas.1714530115.

Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N., and Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science *362*, eaau1783. https://doi.org/10.1126/science.aau1783.

Bock, H., Geisler, C., Wurm, C.A., Middendorff, C. von, Jakobs, S., Schönle, A., Egner, A., Hell, S.W., and Eggeling, C. (2007). Two-color far-field fluorescence nanoscopy based on photoswitchable emitters. Appl Phys B *88*, 161–165. https://doi.org/10.1007/s00340-007-2729-0.

Boettiger, A.N., Bintu, B., Moffitt, J.R., Wang, S., Beliveau, B.J., Fudenberg, G., Imakaev, M., Mirny, L.A., Wu, C., and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature *529*, 418–422. https://doi.org/10.1038/nature16496.

Boltsis, I., Grosveld, F., Giraud, G., and Kolovos, P. (2021). Chromatin Conformation in Development and Disease. Frontiers Cell Dev Biology *9*, 723859. https://doi.org/10.3389/fcell.2021.723859.

Brack, C., and Griffith, J.D. (2008). DNA Electron Microscopy - Part. Critical Rev Biochem *10*, 113–141. https://doi.org/10.3109/10409238109114551.

Bridger, J.M., Boyle, S., Kill, I.R., and Bickmore, W.A. (2000). Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. Curr Biol *10*, 149–152. https://doi.org/10.1016/s0960-9822(00)00312-2.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292–301. https://doi.org/10.1038/35066075.

Cremer, M., Grasser, F., Lanctôt, C., Müller, S., Neusser, M., Zinner, R., Solovei, I., and Cremer, T. (2012). The Nucleus, Volume 1: Nuclei and Subnuclear Components. Methods Mol Biology *463*, 205–239. https://doi.org/10.1007/978-1-59745-406-3_15.

CRICK, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563. https://doi.org/10.1038/227561a0.

Criscione, S.W., Cecco, M.D., Siranosian, B., Zhang, Y., Kreiling, J.A., Sedivy, J.M., and Neretti, N. (2016). Reorganization of chromosome architecture in replicative cellular senescence. Sci Adv *2*, e1500882. https://doi.org/10.1126/sciadv.1500882.

Dahm, R. (2008). Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. Hum Genet *122*, 565–581. https://doi.org/10.1007/s00439-007-0433-0.

Dalal, Y., Fleury, T.J., Cioffi, A., and Stein, A. (2005). Long-range oscillation in a periodic DNA sequence motif may influence nucleosome array formation. Nucleic Acids Res *33*, 934–945. https://doi.org/10.1093/nar/gki224.

Dame, R.T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol Microbiol *56*, 858–870. https://doi.org/10.1111/j.1365-2958.2005.04598.x.

Day, L., Chau, C.M., Nebozhyn, M., Rennekamp, A.J., Showe, M., and Lieberman, P.M. (2007). Chromatin Profiling of Epstein-Barr Virus Latency Control Region. J Virol *81*, 6389–6401. https://doi.org/10.1128/jvi.02172-06.

Dehghani, H., Dellaire, G., and Bazett-Jones, D.P. (2005). Organization of chromatin in the interphase mammalian cell. Micron *36*, 95–108. https://doi.org/10.1016/j.micron.2004.10.003.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science *295*, 1306–1311. https://doi.org/10.1126/science.1067799.

Dempsey, G.T., Vaughan, J.C., Chen, K.H., Bates, M., and Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. Nat Methods *8*, 1027–1036. https://doi.org/10.1038/nmeth.1768.

Diezmann, A. von, Shechtman, Y., and Moerner, W.E. (2017). Three-Dimensional Localization of Single Molecules for Super-Resolution Imaging and Single-Particle Tracking. Chem Rev *117*, 7244–7275. https://doi.org/10.1021/acs.chemrev.6b00629.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380. https://doi.org/10.1038/nature11082.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst *3*, 99–101. https://doi.org/10.1016/j.cels.2015.07.012.

Erceg, J., Abed, J.A., Goloborodko, A., Lajoie, B.R., Fudenberg, G., Abdennur, N., Imakaev, M., McCole, R.B., Nguyen, S.C., Saylor, W., et al. (2019). The genome-wide multi-layered architecture of chromosome pairing in early Drosophila embryos. Nat Commun *10*, 4486. https://doi.org/10.1038/s41467-019-12211-8.

Fawcett, D.W. (1966). On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. Am J Anat *119*, 129–145. https://doi.org/10.1002/aja.1001190108.

Feric, M., and Misteli, T. (2021). Phase Separation in Genome Organization across Evolution. Trends Cell Biol *31*, 671–685. https://doi.org/10.1016/j.tcb.2021.03.001.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686–1697. https://doi.org/10.1016/j.cell.2016.04.047.

Franklin, R.E., and Gosling, R.G. (1953). The structure of sodium thymonucleate fibres. I. The influence of water content. Acta Crystallogr *6*, 673–677. https://doi.org/10.1107/s0365110x53001939.

Franks, A., Airoldi, E., and Slavov, N. (2017). Post-transcriptional regulation across human tissues. Plos Comput Biol *13*, e1005535. https://doi.org/10.1371/journal.pcbi.1005535.

Gagnon, L. (2020). Imaging the Genome in 3D at Super Resolution. Microsc Today *28*, 18–26. https://doi.org/10.1017/s1551929520001546.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin Architecture of the Human Genome Gene-Rich Domains Are Enriched in Open Chromatin Fibers. Cell *118*, 555–566. https://doi.org/10.1016/j.cell.2004.08.011.

Golomb, H.M., and Bahr, G.F. (1974). Electron microscopy of human interphase nuclei. Chromosoma *46*, 233–245. https://doi.org/10.1007/bf00284879.

Gregory, S.G., Barlow, K.F., McLay, K.E., Kaul, R., Swarbreck, D., Dunham, A., Scott, C.E., Howe, K.L., Woodfine, K., Spencer, C.C.A., et al. (2006). The DNA sequence and biological annotation of human chromosome 1. Nature *441*, 315–321. https://doi.org/10.1038/nature04727.

Hansen, J.C. (2002). CONFORMATIONAL DYNAMICS OF THE CHROMATIN FIBER IN SOLUTION: Determinants, Mechanisms, and Functions. Annu Rev Bioph Biom *31*, 361–392. https://doi.org/10.1146/annurev.biophys.31.101101.140858.

Hansen, J.C. (2012). Human mitotic chromosome structure: what happened to the 30-nm fibre? Embo J *31*, 1621–1623. https://doi.org/10.1038/emboj.2012.66.

Heilemann, M., van de Linde, S., Schüttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., and Sauer, M. (2008). Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes. Angewandte Chemie Int Ed *47*, 6172–6176. https://doi.org/10.1002/anie.200802376.

Hess, S.T., Girirajan, T.P.K., and Mason, M.D. (2006). Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. Biophys J *91*, 4258–4272. https://doi.org/10.1529/biophysj.106.091116.

Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., Gobbi, M.D., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat Genet *46*, 205–212. https://doi.org/10.1038/ng.2871.

Ito, K., Masuda, M., Fujiwara, K., Hayashi, H., and Sato, H. (1994). Metaphase and anaphase in the artificially induced monopolar spindle. Proc National Acad Sci *91*, 3921–3925. https://doi.org/10.1073/pnas.91.9.3921.

Kishi, J.Y., Lapan, S.W., Beliveau, B.J., West, E.R., Zhu, A., Sasaki, H.M., Saka, S.K., Wang, Y., Cepko, C.L., and Yin, P. (2019). SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. Nat Methods *16*, 533–544. https://doi.org/10.1038/s41592-019-0404-0.

Koonin, E.V., Gorbalenya, A.E., and Chumakov, K.M. (1989). Tentative identification of RNA-dependent RNA polymerases of dsRNA viruses and their relationship to positive strand RNA viral polymerases. Febs Lett *252*, 42–46. https://doi.org/10.1016/0014-5793(89)80886-5.

Kozma, E., and Kele, P. (2018). Fluorogenic probes for super-resolution microscopy. Org Biomol Chem *17*, 215–233. https://doi.org/10.1039/c8ob02711k.

Langer-Safer, P.R., Levine, M., and Ward, D.C. (1982). Immunological method for mapping genes on Drosophila polytene chromosomes. Proc National Acad Sci *79*, 4381–4385. https://doi.org/10.1073/pnas.79.14.4381.

Lee, H.D., Lord, S.J., Iwanaga, S., Zhan, K., Xie, H., Williams, J.C., Wang, H., Bowman, G.R., Goley, E.D., Shapiro, L., et al. (2010). Superresolution Imaging of Targeted Proteins in Fixed and Living Cells Using Photoactivatable Organic Fluorophores. J Am Chem Soc *132*, 15099–15101. https://doi.org/10.1021/ja1044192.

Li, W., and Miramontes, P. (2006). Large-scale oscillation of structure-related DNA sequence features in human chromosome 21. Phys Rev E *74*, 021912. https://doi.org/10.1103/physreve.74.021912.

Lieberman, P.M. (2008). Chromatin organization and virus gene expression. J Cell Physiol *216*, 295–302. https://doi.org/10.1002/jcp.21421.

Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293. https://doi.org/10.1126/science.1181369.

Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? Curr Opin Cell Biol *22*, 291–297. https://doi.org/10.1016/j.ceb.2010.03.001.

Mandelkern, M., Elias, J.G., Eden, D., and Crothers, D.M. (1981). The dimensions of DNA in solution. J Mol Biol *152*, 153–161. https://doi.org/10.1016/0022-2836(81)90099-1.

McClintock, B. (1953). INDUCTION OF INSTABILITY AT SELECTED LOCI IN MAIZE. Genetics *38*, 579–599. https://doi.org/10.1093/genetics/38.6.579.

Misteli, T. (2010). Higher-order Genome Organization in Human Disease. Csh Perspect Biol *2*, a000794. https://doi.org/10.1101/cshperspect.a000794.

Misteli, T., and Finn, E.H. (2021). Chromatin architecture is a flexible foundation for gene expression. Nat Genet *53*, 426–427. https://doi.org/10.1038/s41588-021-00813-2.

Nguyen, H.Q., Chattoraj, S., Castillo, D., Nguyen, S.C., Nir, G., Lioutas, A., Hershberg, E.A., Martins, N.M.C., Reginato, P.L., Hannan, M., et al. (2020). 3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing. Nat Methods *17*, 822–832. https://doi.org/10.1038/s41592-020-0890-0.

Nicodemi, M., and Pombo, A. (2014). Models of chromosome structure. Curr Opin Cell Biol *28*, 90–95. https://doi.org/10.1016/j.ceb.2014.04.004.

Nir, G., Farabella, I., Estrada, C.P., Ebeling, C.G., Beliveau, B.J., Sasaki, H.M., Lee, S.D., Nguyen, S.C., McCole, R.B., Chattoraj, S., et al. (2018). Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. Plos Genet *14*, e1007872. https://doi.org/10.1371/journal.pgen.1007872.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Berkum, N.L. van, Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385. https://doi.org/10.1038/nature11049.

Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., and O'Shea, C.C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science *357*. https://doi.org/10.1126/science.aag0025.

Paweletz, N. (2001). Walther Flemming: pioneer of mitosis research. Nat Rev Mol Cell Bio *2*, 72–75. https://doi.org/10.1038/35048077.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021.

Robinson, J.T., Turner, D., Durand, N.C., Thorvaldsdóttir, H., Mesirov, J.P., and Aiden, E.L. (2018). Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. Cell Syst *6*, 256-258.e1. https://doi.org/10.1016/j.cels.2018.01.001.

Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods *3*, 793–796. https://doi.org/10.1038/nmeth929.

Steensel, B. van, and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780–791. https://doi.org/10.1016/j.cell.2017.04.022.

Strachan, T., and Read, and A.P. (2019). Human Molecular Genetics.

Struhl, K. (1999). Fundamentally Different Logic of Gene Regulation in Eukaryotes and Prokaryotes. Cell *98*, 1–4. https://doi.org/10.1016/s0092-8674(00)80599-1.

Vangindertael, J., Camacho, R., Sempels, W., Mizuno, H., Dedecker, P., and Janssen, K.P.F. (2018). An introduction to optical super-resolution microscopy for the adventurous biologist. Methods Appl Fluores *6*, 022003. https://doi.org/10.1088/2050-6120/aaae0c.

Vilarrasa-Blasi, R., Soler-Vila, P., Verdaguer-Dot, N., Russiñol, N., Stefano, M.D., Chapaprieta, V., Clot, G., Farabella, I., Cuscó, P., Kulis, M., et al. (2021). Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. Nat Commun *12*, 651. https://doi.org/10.1038/s41467-020-20849-y.

Vogelsang, J., Cordes, T., Forthmann, C., Steinhauer, C., and Tinnefeld, P. (2009). Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. Proc National Acad Sci *106*, 8107–8112. https://doi.org/10.1073/pnas.0811875106.

Xie, W.J., Meng, L., Liu, S., Zhang, L., Cai, X., and Gao, Y.Q. (2017). Structural Modeling of Chromatin Integrates Genome Features and Reveals Chromosome Folding Principle. Sci Rep-Uk *7*, 2818. https://doi.org/10.1038/s41598-017-02923-6.

Yoon, S., and Vahedi, G. (2021). Stripenn detects architectural stripes from chromatin conformation data using computer vision. Biorxiv 2021.04.16.440239. https://doi.org/10.1101/2021.04.16.440239.

(1970). IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Abbreviations and symbols for nucleic acids, polynucleotides and their constituents. Recommendations 1970. Biochem J *120*, 449–454. https://doi.org/10.1042/bj1200449.