Title
Determining key features of X Chromosome Architecture in the context of the Inactive and Dampened States

Author
Miller, Jarrett

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Determining key features of X Chromosome Architecture in the context of the Inactive and Dampened States

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular Biology

by

Jarrett L. Miller

2020

ABSTRACT OF THE DISSERTATION

Determining key features of X Chromosome Architecture in the context of the Active, Inactive and Dampened gene expression states

by

Jarrett L. Miller

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2020

Professor Kathrin Plath, Chair

Present within the nuclei of female human cells are two X chromosomes of differing gene expression states, one active chromosome with a structurally open configuration and one inactive chromosome, harboring a distinct, spatially condensed structure, in which a majority of genes are transcriptionally silenced. A third gene expression state, the dampened X chromosome, is found early in human development in the pre-implantation blastocyst and displays neither the levels of transcription found in the active nor the inactive X chromosome but instead exhibits a reduction of transcriptional activity. Through a combination of current technologies like chromosome conformation capture techniques and microscopy, gene expression has been shown to correlate to chromosome structure. Although the dampened X chromosome is known to feature a unique gene expression profile, the structure of this X chromosome state remains unknown. To understand the impact of structural rearrangements of the human X chromosome throughout the transition period of X chromosome inactivation development we have sought to define the physical features of the active, inactive and dampened states of the X chromosome.

To determine the structure of the X chromosome, we chose the sequential fluorescent in

situ hybridization (FISH) technique and established the protocol in the lab. To this end, we optimized every step along the way including plating conditions for cells, fixation conditions, microscope relocation and alignment, and image analysis pipelines. We demonstrate that our sequential FISH approach allows for, at minimum, four rounds of hybridization, each containing multiple probes targeting various genomic locations for the purpose of defining chromosome structure in 3D, with imaging acquired on a conventional confocal microscope

Applying the optimized sequential FISH, we determined the variance of the size and higher-order organization of the X chromosome not only in different chromosome states of the same cell type but in the active X state across different cell types. We chose two different cell types to analyze: human fibroblast and naïve embryonic stem cells. The former representing a somatic state that carries an inactive (Xi) and active X chromosome (Xa) while the later illustrates a pluripotent, early in development cell type harboring an Xa and dampened X chromosome (Xd). In both cases, the lncRNA XIST, essential for X chromosome inactivation (XCI) which leads to a change in X chromosome gene transcription, is expressed from the non-active X chromosomes, Xi and Xd. This combination of cell types allows for the investigation of all three gene expression states but also provides an internal comparison between two different gene expression states in each cell. We found that Xa in fibroblast cells, while clearly larger in size and containing more intrachromosomal distance between genes, is absolutely smaller than the naïve Xa. In comparison, the Xd ranged from slightly smaller to practically the same size as the naïve Xa. Additionally, some Xd gene locations, such as XACT, were found to be incredibly distant to the other genes unlike in the naïve Xa.

We interrogated the structure of the Xd by comparing it to the structures of the Xa and Xi. We analyzed whole chromosome structure using genomic locations that span the length of the X chromosome as well as specific combinations of genes with relevance towards gene expression. We investigated the genes that become silenced upon XCI and found that the fibroblast Xi predictably contains small intrachromosomal distances between all probe

combinations. Conversely, both naïve and fibroblast Xa tended to be contain large intrachromosomal distances. The naïve Xd interestingly showed noticeably smaller intrachromosomal distances concerning measurements between genes in the latter half of the X chromosome. Genes that escape X chromosome dampening (XCD) showed similar arrangement between the naïve X chromosomes but not the fibroblast. Finally, we found that Xd shows intrachromosomal distances the same size or larger than naïve Xa when measuring between probes on either side of DXZ4.

This study ultimately characterizes distinct physical features of the active, inactive and dampened X chromosome states due to the analysis of the spatial locations of key genes regulated by these expression states. The chromosome organization data acquired allowed us to decipher specific features of the X chromosome with different overall activity states and, in the future, will aid in revealing the mechanisms underlying the transitions in chromosome organization during human development.

The dissertation of Jarrett Laine Miller is approved.

Douglas Black

Tracy Johnson

Stephen Smale

Kathrin Plath, Committee Chair

University of California, Los Angeles

2020

Dedication

I would like to dedicate this research to my mother.

"You the real MVP" – Kevin Durant

Table of Contents

List of Figures and Tables

Chapter 2

Chapter 3

Chapter 4

Acknowledgements

I would like to thank my mentor, Kathrin Plath for guiding me all these years through the highs and lows of graduate school. Thank you for putting up with me.

I would like to thank my thesis committee, Douglas Black, Tracy Johnson, and Stephen Smale for their attention and quality input concerning my project and eduction .

I would like to thank the members of the Plath lab. A special thanks to my former mentor Amy Pandya-Jones and the members of the Buff Bay: Justin Langerman, Iris Dror, and Amanda Collier.

I would like to acknowledge the many contributions of Yolanda Markaki and Iris Dror to the project.

I would like to thank the Molecular Biology Institute for funding and support.

VITA

June 15th, 1993	Born
2014	Bachelors of Science The University of Texas at Austin Biochemistry
2012-2014	Undergraduate Research Student Research Position, The University of Texas at Austin Dr. Hal Alper
2014-2015	Biochemistry and Molecular Biology Doctoral Program University of California, Los Angeles
2015-2020	MBIDP Graduate Program University of California, Los Angeles
2015	Teaching Assistant Chemistry and Biochemistry University of California, Los Angeles
2017-Present	Ph.D Candidate, Dr. Kathrin Plath Laboratory Molecular Biology Institute University of California, Los Angeles

Chapter 1

Introduction

X Chromosome Inactivation and X chromosome Dampening alter gene expression and chromosome architecture

Chromatin and Nuclear Organization

Chromatin manifests as the basic foundation of nuclear organization and center of information in the form of chromosomes (Zhu and Li 2016). Chromatin is assembled from deoxyribonucleic acid (DNA), composed of various combinations of four unique nucleotide bases in a double helix arrangement, wrapped around histones, octamers of protein that contain the capacity to be post-translationally modified at the N-terminus of these protein units (Venters and Pugh 2010; Zhu and Li 2016). The arrangement of the nucleotide bases within the DNA structure, also known as the sequence, carries linearly arranged information. DNA can be further organized into functional units for a multitude of cellular purposes that include protein coding genes and non-protein coding genes. Informational power of chromatin is derived from the transcription of ribonucleic acid (RNA), from the sequence of the DNA, which can function in its current state or alternatively be translated to functional protein (Hershey et al., 2012).

Modulation of gene transcription is also a relevant design of chromatin. Promoters, silencers, enhancers, and insulators all act as regulatory elements that facilitate or inhibit gene transcription (Bylino et al., 2020). Promoters signal to RNA Polymerases where to initiate transcription. Silencers seek to inhibit gene expression. Enhancers act as transcription boosters to increase the overall amount of gene transcription occurring (Calo and Wysocka 2014). Insulators act as barriers for either enhancer activity or general segregation of chromatin. But the information garnered from the DNA sequence alone does not represent the total power of chromatin. Epigenetic modifications to both DNA as well as to histones impart drastic changes to the makeup of the chromatin in addition to the effectiveness of transcription. Understanding chromatin organization and modifications and their link to gene expression is critical to the overall goals of this project in quantifying the structural differences of the X chromosome in its

three distinct global gene expression states.

Gene Expression and Chromatin Modifications

Gene expression is modulated by chromatin modifications (Dong and Weng 2013). Although many such modifications exist, a few of these modifications can grant a general glimpse at the many facets of this form of transcriptional regulation.

For example, methylation serves as a method to repress gene transcription. In the case of DNA methylation, a methyl group is usually added by a methyltransferase protein to a cytosine, potentially within a promoter region. Promoter methylation has been shown to vary significantly amongst different cell types (Suzuki and Bird 2008; Hodges et al., 2009; Varley et al., 2013). Histone methylation is another common epigenetic modification with gene expression consequences. Histone H3 lysine (K) 27 trimethylation and H3K9 di/trimethylation are often targeted in order to induce the inactivation of gene expression (Heard et al., 2001).

Histone ubiquitination serves as another important strategy for alteration of gene expression. Notably, ubiquitination of H2A and H2B, two units of the four core histones that compose the standard histone octamer, typically leads to the repression of gene expression (Bonnet et al., 2014). However, in some cases, H2B can in fact activate gene expression (Nickel et al., 1989; Davie and Murphy 1990; Henry et al., 2003).

Unlike the previously discussed histone modifications, acetylation of histone protein typically acts as a positive modulator of gene transcription (Grunstein 1997). When an acetyl group is added to a histone tail the chromatin undergoes a structure change and becomes less

condensed (Gorisch 2005).

Gene Silencing in Eutherian Mammals

Gene repression can be triggered in various ways but a common route taken involves the polycomb-group proteins. Polycomb repressive complex 2 (PRC2) is a conglomerate of four proteins: Suz12, Eed, Ezh1/2, and RbAp48, with the capacity to methylate lysine 27 of the N-terminus tail of histone H3 octamer (Cao et al., 2002). This activity will serve to downregulate the nearby genes by altering the region into a less accessible heterochromatin state. Following polycomb repressive complex 1 (PRC1), PRC2 will seek to better establish and maintain the silencing effect via the physical compaction of the chromatin (Zhao et al., 2008; Brockdorff 2017). Heterochromatinization due to silencing of gene expression effectively alters how a chromosome is spatially organized.

Chromosome Spatial Organization

In addition to the DNA sequence, RNA processing and epigenetic modifications of chromatin, the 3D spatial organization of the chromosome, such as the dynamic folding and looping of the chromstin, contributes to the regulation of gene expression (Kosak 2004; Misteli 2007; Dekker 2008). Genome folding produces regulatory segmentation for transcriptional control of the genome (Merkenschlager and Nora 2016; Pierro 2019). Chromatin composition and structure are inherently, intimately intertwined with gene regulation.

At the highest level of nuclear organization, chromosomes occupy separate spatial territories within the nucleus in a non-static manner with structural rearrangement occurring

allowing different regions of a given chromosome to become nearer or more distal (Cremer et al., 2001). In particular, regulatory elements can be separated on the linear DNA strand from their gene targets by multi-kilobase distances. Large regulatory domains bring regulatory elements and gene promoters together in three-dimensional (3D) space to provide specificity in the regulation of genes in a cell type-dependent manner (Symmons et al., 2014). Promoters and enhancers found within the same spatial domain are known to interact with each other to a higher degree (Merkenschlager and Nora 2016). Within this 3D genome space are distinct levels of organization that each contribute to gene regulation.

Hierarchical Levels of Genome Architecture

Chromatin can form loops to alter spatial distances via the proteins cohesion and CCCTC-binding factor (CTCF) (Holwerda et al., 2012). This action can insulate neighborhoods of genes or be used to force transcriptional activation of genes. These spatial arrangements can be defined using chromosome conformation capture techniques that crosslink DNA that is in close proximity to each other. Then, using sequencing, the identity and frequency of the interactors can be mapped on a single-cell or bulk population scale (Kalhor et al., 2011; van de Werken et al., 2012; Nagano et al., 2013).

Further organization of these folds includes chromatin compartments, labeled A or B, corresponding to multi kilobase to megabase regions of gene transcription or a lack thereof respectively. Additionally, compartment B is typically gene-poor as well (Merkenschlager and Nora 2016). The A and B compartments are composed of topologically associated domains (TADs) which are regions of the chromosome that are spatially adjacent to each other and share

regulatory elements such as enhancers. TADs inherently constitute preferential interactions involving the regulatory elements and genes within them (Dixon et al., 2012; Nora et al., 2012).

Hi-C is a technique that seeks to capture close range interactions using a process involving formaldehyde crosslinking in order to capture chromatin in close spatial proximity, followed by digestion and then subsequent ligation of these DNA fragments. Afterwards, these products can be sequenced in order to ascertain their identity as well as quantify their count, also known as their contact frequency (Lieberman-Aiden et al., 2009; Berknum et al., 2010; Belton et al., 2012). Based on Hi-C contact frequency data, interactions of loci within the same TAD are approximately two to three times that compared to interactions of the loci not found within the same TAD (Merkenschlager and Nora 2016). Based on Hi-C maps, TADs can encompass a wide range of sizes, anywhere from 40kb to 3Mb (Rao et al., 2014). This range in size of domains is achievable due to the folding of the chromosome allowing for locations which may be linearly distant to come closer together in order for these interactions to occur. Across Hi-C experiments of the same cell type and conditions, TAD positions remain largely consistent within populations of cells but single cell heterogeneity of chromatin folding does occur (Goetze et al., 2007). Data driven models and high-resolution imaging both point towards TADs and loop domains being dynamic (Fudenburg and Leonid 2012;Giorgetti et al., 2014;Nora et al., 2012). High resolution Hi-C experiments show that across cell types there is a portion of loops that remain unchanged. Current estimates suggest that 50% to 75% of peaks, regions of interaction which at a significant density threshold compared to neighboring interactions can be used to call loops, are shared across at least seven human cell lines (Rao et al., 2014; Wolff et al., 2020).

TADs and compartments are not static entities. Experimental recruitment of SUV39H1 in order to promote H3K9 trimethylation can produce a change in compartment occupancy for a

TAD, from the active A compartment to the repressive B compartment (Wijchers et al., 2016). Upon differentiation, it has been observed that within a single or series of TADs the A and B compartments have switched (Bickmore et al., 2013; Nora et al., 2013; Dixon et al., 2015). This indicates that TADs are dynamic with respect to gene expression.

Similar to the flexibility of TADs, enhancers work opportunistically with many gene targets. Using 5C, estimates show that an active gene is targeted on average by approximately four enhancers, with Hi-C revealing that a majority of promoters contact multiple enhancers (Pope et al., 2014;Phillips et al., 2009). For example, the β-globin promoter and enhancers have been experimentally shown to activate transcription at a distance of over 40kb (Deng et al., 2012). Enhancer activity does appear to correspond to regions organized as TADs (Symmons et al., 2014). Hi-C data in GM12878 cells, a human lymphoblastoid cell line, revealed that 30% of observed loops connect annotated promoters and enhancers with most interactions occurring within individual TADs (Rao et al., 2014). This is a noticeable increase when compared to the calculated 7% connection rate by chance.

Generally, genes that all fall within the same TAD tend to be more similarly regulated than genes belonging to separate TADs in the context of stem cell differentiation (Nora et al., 2012). Gene regulatory elements appear to engage in domain-wide contacts rather than preferentially engaging other regulatory elements. This gives credence to the idea that TADs, as spatial domains, play a role in organizing gene transcription. Reorganization of compartments is associated with stable changes in transcriptional state rather than rapid or transient changes (Jin et al., 2013;Le Dily et al., 2014;Dixon et al., 2015). The flexibility of genome architecture across multiple levels indicates a large degree of control concerning the correlated gene expression.

CTCF and Cohesin

Two main contributors of DNA loop formation and 3D genome organization are the transcription factor CTCF and the ring-like multiprotein complex cohesin. Cohesin participates in sister chromatid cohesion as well as other processes such as DNA double-strand-break repair and transcriptional control thanks to its ability to compact linear DNA (Nasmyth and Haering 2009; Kim et al 2019). On the other hand, CTCF acts as a DNA binding factor, being able to attach to chromatin and form loops. These chromatin loops are integral for forming long-range DNA interactions. To form a loop between two CTCF binding sites, even hundreds of kilobases apart, CTCF binding sites must be in proper, convergent orientation for communication with cohesin (Rao et al., 2014; Merkenschlager and Nora 2016). In the case of long-range loops, CTCF proteins are bound to DNA in the head-to-head configuration (Merkenschlager and Nora 2016). CTCF binding sites are enriched at TAD boundaries (Dixon et al., 2012). In comparison to long-range loops, the CTCF binding sites located at the proximity of TAD boundaries are in the divergent orientation rather than head-to-head. Inversion of the orientation of a CTCF binding site disrupts loop formation (de Wit et al., 2015;Guo et al., 2015). One such consequence of loop disruption due to binding site orientation inversion of clustered CTCF motifs is the access of previously insulated regions and alterations in chromatin folding (Guo et al., 2015). Cohesin, loaded onto DNA by the protein complex NIPBL/MAU2, utilizes two convergent CTCF sites as de facto boundaries to extrude the chromatin loops by moving both ends in opposite directions (Sanborn et al., 2015; Braccioli and Wit, 2019). CTCF is required to separate chromatin folding domains, while cohesin serves to shape said domains (Zuin et al., 2013). CTCF is essential for cellular function as well as embryonic development (Burcin et al., 1997; Fedoriw et al., 2004; Sleutels et al., 2012; Soshnikova et al., 2010;Wan et al., 2008). CTCF has

even been found to be DNA methylation-dependent in some cases (Beygo et al., 2013;Feldmann et al., 2013; Hadjur et al., 2009; Wang et al., 2012).

Disruption of the cohesin complex leads to a loss of local chromatin interactions (Zuin et al., 2013). However, depletion of CTCF reduced interactions on both local and long-range scales (Zuin et al., 2013). Critically linked to these interactions is gene expression. Loss of cohesion has shown to cause downregulation of genes (Rao et al., 2017). Importantly, this capacity to regulate gene expression and chromatin interactions has been shown to correlate during differentiation (Phillips-Cremins et al., 2013). CTCF has also been shown to play a role in the regulation of mRNA splicing on a genome-wide scale (Shukla et al., 2020).

CTCF and cohesin are known to co-occupy thousands of sites in the genome (Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wang et al., 2012). Incidentally, strong cohesin sites often agree with CTCF binding whereas weak cohesin sites are akin to locations of promoters and enhancers (Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wang et al., 2012; Faure et al., 2012; Kagey et al., 2010; Schmidt et al., 2010;Yan et al., 2013). CTCF is believed to affect cohesin distribution on chromosomes due to the finding that the enrichment of cohesin at CTCF binding sites is decreased upon CTCF depletion (Parelho et al., 2008; Wang et al., 2012). Conversely, the depletion of only cohesin subunits does not appear to disturb or significantly change the binding pattern of CTCF (Merkenschlager and Nora 2016). However, partial depletion of cohesin does appear to weaken interactions within TADs in Hi-C data (Zuin et al., 2013). Loss of both CTCF and cohesion binding has shown to promote the loss of chromatin contacts by Hi-C (Pugacheva et al., 2020).

CTCF and cohesion act as a duo that participate in genome organization on a scale that

ranges from local to long-range interactions. Chromatin looping is an essential part of the interplay between gene expression and structure.

Human Embryonic Stem Cells

Embryonic stem cells (ESCs) represent a powerful cellular model to understand the link between nuclear architecture and gene expression as they can easily be induced to differentiate and change their cell fate. ESCs have the capacity to grow indefinitely and differentiate into each of the three germ layers, allowing for the ability to serve as an incredible platform to model developmental states (Evans and Kaufman 1981; Martin 1981; Takahashi et al., 2006). Human embryonic stem cells (hESCs) can be harvested from pre-implantation blastocysts (Liu et al., 2019). Those that capture the state of epiblast cells in the pre-implantation blastocyst are referred to as naïve hESCs. HESCs in what is commonly known as the primed state correlates to a transcriptional profile that directly mirrors that of the post-implantation blastocyst cells (Pastor et al., 2016), which is induced through culture conditions during the derivation from the blastocyst. Additionally, induced pluripotent stem cells, which reflect the naïve or primed hESC state, can be reprogrammed from adult somatic cells via transcription factor-mediated reprogramming (Takahashi et al., 2006) combined with appropriate culture conditions. Morphologically, human primed and naïve pluripotent stem cells appear drastically different. Naïve and primed cells are known to use distinct enhancers and display many differences in gene expression. This would suggest different long-range chromatin interactions (Factor et al., 2014).

Long Noncoding RNA

Long non-coding RNA is a type of non-translated RNA that is longer than 200 nucleotides (Derrien et al., 2012; Yao et al., 2019). Composed of more than two-hundred nucleotides, long non-coding RNA does not adhere to the central dogma by not undergoing translation to produce functional protein. Long non-coding RNAs have been shown to be functional in various ways such as providing a scaffold for protein or DNA, recruitment of transcription factors, regulation of pluripotency, amongst other roles (Akhade et al., 2017). These RNAs are known to regulate chromatin structure and consequently regulate transcription by playing a role in altering DNA methylation as well as histone modifications (Akhade et al., 2017; Rinn and Chang 2012; Nagano and Fraser 2011; Kugel and Goodrich 2012; Mondal et al., 2010). Of particular note are the concepts that long non-coding RNAs may regulate chromatin in cis, found in the case of the ANRIL RNA, or by mediating looping, as seen in the function of HOTTIP RNA (Kugel and Goodrich 2012; Yap et al., 2010; Wang et al., 2011). Highly relevant to the gene expression states of the X chromosome as well as mammalian development, is XIST, one of the first long non-coding RNAs to be discovered (Brown 1991; Brockdorff et al., 1992; Heard and Loda 2019).

X Chromosome Inactivation in Mouse

X Chromosome inactivation (XCI) is a process which results in the silencing of most genes on the X chromosome to fulfill the gene dosage compensation issue of multiple X chromosomes in female cells. All X chromosomes except one will be inactivated (Lyon 1961; Belmont et al. 1986). In mouse, XCI occurs into rounds: directly after fertilization the paternal X

chromosome is inactivated via an imprinted XCI process and further into development, after a reactivation of the paternal X in the inner cell mass of the blastocyst, in the epiblast stage where a random inactivation of either the maternal or paternal X chromosome (Takagi 1978; Engel 2015). This process is dependent on the transcription of the lncRNA XIST, the master regulator of XCI (Nesterova et al., 2019).

XIST

Upregulation of the XIST gene is the essential step in initiation for X chromosome inactivation (Brown et al., 1991; Rastan et al., 1983). The long non-coding RNA XIST is conserved in all eutherian mammals and has been shown to have its function conserved in the corresponding marsupial gene RSX (Sprague et al., 2019). XIST, found on the X chromosome and responsible for controlling the inactivation of the X chromosome in cis, encodes a 17 kilobase transcript in humans containing six conserved nucleotide repeat regions of differing numbers of nucleotides but similar in unique features labeled alphabetically A through F (Brockdorff et al., 2002). The A-repeat interacts with SPEN, which is crucial for gene silencing by XIST. SPEN activates HDAC3 on chromatin, which in turn induces gene silencing and loss of transcriptional regulators (McHugh et al., 2015;). The C-repeat interacts with YY1 in tethering the XIST RNA to the X chromosome (Jeon et al., 2011). The B-repeat of the XIST transcript is crucial for maintaining PRC1 and PRC2 complexes and consequently XIST spreading (Colognori et al., 2019). Conversely, the E-repeat is necessary in late stage maintenance and helps to perpetuate gene silencing of the inactive X (Pandya-Jones et al., 2020)

X Chromosome Dampening

In addition to the inactive and active states of the X chromosome there is also a dampened state which occurs in the human pre-implantation embryo (Sahakyan et al., 2018). X chromosome dampening (XCD) is a process that has not been described in any other mammal (Sahakyan et al., 2018). Though an X dampening phenomenon has been observed in C. elegans and, in this organism, relies on a condensin protein assembly (Albritton et al., 2018). XCD is a state found in human pre-implantation embryos in which the X chromosome transcribes genes in a reduced manner compared to that of the active X chromosome but not functionally silenced like the inactive X chromosome. Essentially, dampening serves to lower the net expression of the chromosome. The dampened X chromosome is found in cells in the early stages of development within the pre-implantation blastocyst and can be recapitulated in naïve pluripotent hESCs in culture in 5iLAF or in t2iLAF + Gö conditions (Sahakyan et al., 2017; Vallot et al., 2017).

The standard media conditions for promoting naïve pluripotency are known as 5iLAF (Theunissen et al., 2014). This cocktail of inhibitors (PD0325901/IM12/SB590885/Y-27632/WH-4-023) supplemented with FGF and Activin A allows for generation of naïve pluripotent stem cells over a series of passages. PD0325901 acts as a MEK inhibitor and can moderately reduce the level of global methylation (Li et al., 2016). This hypomethylation may be critical for sustaining pluripotency of ESCs. IM-12 is an inhibitor of glycogen synthase kinase 3β (GSK-3β). GSK-3β plays a role in the Wnt signaling pathways which are known to influence embryonic development. SB590885, a RAF (Rapidly Accelerated Fibrosarcoma) inhibitor, also participates in the RAS-RAF-MEK-ERK signal transduction cascade. Y-27632 is a ROCK (Rho-associated protein kinase) inhibitor. Rho kinases play critical roles in self-renewal, lineage commitment, and apoptosis regarding stem cells (Wang et al., 2017). WH-4-023, a LCK/SRC

inhibitor, maintains expression of pluripotency factors (Gao et al., 2019). WH-4-023 also assists with proper cell morphology (Theunissen et al., 2014).

Similar to previously described inactive X chromosome, the dampened X chromosome expresses XIST (Sahakyan et al., 2018). XIST may in fact be the regulator for XCD by influencing the methylation of H3K27, leading to some degree of gene transcription repression found in the dampened state. The dampened state, in which two X chromosomes expressing the XIST are subject to a form of dosage compensation characterized by down regulation but not complete inactivation of X-linked gene expression, can be found in cultured naïve human pluripotent stem cells(hPSCs) (Petropoulos et al., 2016;Sahakyan et al., 2016).

While bi-allelic XIST expression can be captured in naïve hPSCs, in many cases XIST may be expressed from only a single X chromosome (Sahakyan et al., 2016). This contrasts with the finding that XIST accumulates on both active X chromosomes in early human pre-implantation embryos (Vallot et al., 2017). 5iLAF cultured cells, while initially facing substantial cell death, do effectively produce the characteristic dome-shaped colonies that exhibit upregulation of naïve pluripotency markers (Theunissen et al., 2014). Cells cultured in 5iLAF are also able to properly reactivate genes that would regularly be inactivated like *HUWE1, ATRX*, and *THOC2* by passage four (Sahakyan et al., 2016). The reactivation found in the cells cultured in the 5iLAF media is considered chromosome-wide based on the application of RNA-seq or Sanger sequencing of RT-PCR products utilizing single nucleotide polymorphisms (SNPs) in order to determine allelic origin (Sahakyan et al., 2016). The XIST cloud found in these naïve hPSCs is noticeably more diffuse and punctate than that of an XIST cloud found on a somatic inactive X (Sahakyan et al., 2016). This is encouraging towards validity of the naïve hPSC system as it mimics human pre-implantation blastocyst pattern (Okamoto et al., 2011). Naïve

hPSCs also feature a distinct lack of methylation, corresponding with the reactivation of X-linked genes, approaching that of near global hypo-methylation that is present in the human pre-implantation epiblast (Sahakyan et al., 2016;Smith et al., 2014;Pastor et al., 2016;Theunissen et al., 2016).

Dampening of gene expression is also detected in naïve hPSCs cultured in 5iLAF media. While neither of the X chromosomes in these cells are inactive, cells harboring XIST positive X chromosomes were found to express X-linked genes at a lower level than the XIST negative counterparts while autosomal genes remained unaffected (Sahakyan et al., 2016).

Enrichment of H3K27 methylation on the inactive X chromosome is a known and easily identifiable feature thanks to immunostaining (Plath et al., 2003). XIST positive chromosomes in naïve hPSC also display H3K27 methylation but critically lack a key facet of XCI, exclusion of RNA polymerase II (Sahakyan et al., 2016). Naïve hPSCs created via 5iLAF culture conditions can undergo de novo XCI (Sahakyan et al., 2016).

2iLAF+Gö media is an adaptation to the previous naïve hPSC conversion media. The difference being that the two inhibitors present in 2iLAF+Gö media are the previously detailed MEK inhibitor, PD0325901, and CHIR99021, another inhibitor of the GSK-3 enzyme. The noticeable addition being Gö6983, a protein kinase C inhibitor. Originally Gö6983 was tested in a titrated two inhibitor media in order to repress ES cell differentiation, which assists in the production of naïve cells (Takashima et al., 2014). Naïve cells produced from 2iLAF+Gö media appear transcriptionally identical to those of the 5iLAF cultured variety, recapitulating both the dampening effect as well as the XIST cloud.

XACT

Specific to primates is the novel long noncoding RNA, XACT (X active coating transcript). Like XIST, XACT is a repeat-rich lncRNA. In contrast to XIST, XACT is transcribed only from the active X chromosome. XACT may reduce the silencing capability of or act in an antagonistic manner towards XIST in early human embryonic development (Vallot et al., 2017). XACT is limited to expression in pluripotent cells. Therefore, XACT is only present in hESCs in development (Sahakyan et al., 2017). Like XIST, XACT is also known to cluster around the X chromosome in a cloud like formation (Vallot et al., 2017). In the case of naïve pluripotent stem cells, XIST and XACT are known to co-accumulate on the active X chromosome (Vallot et al., 2017). The antagonistic action of these two lncRNAs appear to be the crux of gene dosage compensation establishment in human (Vallot et al., 2017).

XACT expression seems to be tightly regulated in regards to development. At E4 stage of development of hESCs XACT is activated at significantly higher levels in females (Vallot et al., 2015). By E5 XACT levels in trophectoderm have already decreased (Petropoulos et al., 2016). This information points towards XACT playing a role in regulating gene expression of naïve pluripotent stem cells but the extent of which is still unknown.

XCD Escape

The dampened X state has a set of genes that escape the reduction of transcription across the rest of the chromosome. This group of genes is referred to as XCD escapers. XCD escapers have been identified due their relative higher gene expression when compared to the other genes

on the dampened X chromosome (Petropoulos et al., 2016). These genes may prove to have essential roles in navigating the pre-implantation landscape of development.

X Chromosome Inactivation in Human

Upon implantation of the blastocyst, XCI initiates and thereby causes random inactivation of either the maternal or paternal X chromosome during the transition from naïve to primed pluripotency (Minkovsky et al., 2012; Sahakyan et al., 2016). After XCI is initiated due to the cis accumulation of XIST RNA, silencing is brought about primarily via SPEN with contribution from PRC1 (van den Berg et al., 2009; Nesterova et al., 2019). The engagement of XIST in the inactivation process is obvious in the context of light microscopy. Numerous XIST transcripts are easily viewable as coating the X chromosome in what is regularly described as an XIST cloud (Jonkers et al., 2008; Brockdorff 2019). The concentration of the XIST transcripts in association with the RNA-binding proteins involved in XCI, PTBP1, MATR3, TDP43, and CELF1, are thought to lead to paraspeckle-like condensate formation that assists inactivation through phase separation (Cerase et al., 2019; Pandya-Jones et al., 2020). Another nuclear change that can be detected is the compaction of the inactive X chromosome into a heterochromatic structure (Plath et al., 2003; Rasmussen et al., 2001), also known as the Barr body, resulting in a 20-30% reduction in size. A bipartite structure of megadomains is formed around and separated by the noncoding region of DXZ4 (Froberg et al., 2018). The inactive X chromosome has a distinct superstructure to complement its stark change in gene expression profile (Darrow et al., 2016). The inactive X chromosome is often located at the periphery of the nucleus or neighboring the nucleolus (Bonora et al., 2018). This developmental process produces

a clear change in 3D structure of the X chromosome that has stark consequences to its gene expression.

During inactivation, transcription is instructed by the chromosome structure. Employing close spatial proximity, XIST is thought to spread to the neighboring TADs belonging to the A compartment (Engreitz et al., 2013). TAD construction of the inactive X degrades and aside from regions that contain CTCF binding events specific to the inactive X. Explicitly, this gives rise to entirety of the inactive X switching to the same compartment and developing a two super domain structure (Nora et al., 2012;Splinter et al., 2006). Deletion of the Xist gene after establishment of the inactive X does restore structural components of the chromosome such as loops, TADs, and compartments. It is important to note that transcriptional activity does not return to original levels (Minajigi et al., 2015; Splinter et al., 2011).

XCI enacts a change in overall chromosome structure and gene expression due to the lncRNA XIST and a cocktail of proteins that promote silencing. This produces what is essentially a classic example of facultative heterochromatin (Sidhu et al., 2008). However, there is order to the overall organization of the chromosome centered around the hinge region DXZ4.

DXZ4

The DXZ4 macrosatellite, a region of large repeat monomers featuring a high percentage of GC content (Giacalone et al., 1992), acts as an anchor for large chromatin loops in the inactive X chromosome, not only in humans but conserved across multiple mammalian species (Darrow et al., 2016). One such large chromatin loop that DXZ4 appears at the anchor of is none other

than the essential element for X chromosome inactivation, XIST (Darrow et al., 2016). In the case of the inactive X chromosome concerning both human and mouse the DXZ4/Dxz4 loci encode lncRNAs and bind CTCF and components of the cohesin complex (Bonora et al., 2018). Upon deletion of DXZ4, the bipartite, superdomain structure of the inactive X chromosome is completely lost. In addition to this large-scale change in structure caused by the loss of the DXZ4 topological boundary, modifications to compartmentalization patterns and distribution of chromatin marks also occur (Darrow et al., 2016). While compartment structure in the context of mouse is essentially lost in the inactive X chromosome, human inactive X chromosome still carries A and B compartment organization. The loss of DXZ4 alters the composition of these compartments, shifting some genes into the nontraditional compartment. In mouse, reorientation of the Dxz4 locus also leads to structural reorganization, further implicating it as a critical hub for the inactive X chromosome (Bonora et al., 2018). Although the DXZ4 domain acts as a conserved foundational piece for XCI, this does not mean the process of XCI itself is absolute.

XCI Escape

Not all genes are transcriptionally silenced on the inactive X chromosome. In humans, a population of about 15% of X-linked genes escape (Carrel et al. 1999; Yang et al. 2010) silencing and remain active. This particular group of genes are commonly referred to as XCI escapers. XCI escapers include a range of genes of various functions with their identities changing depending on cell type, tissues, and individuals (Andergassen et al., 2017). Additionally, these escapers are not all expressed at the same levels. Instead, XCI escapers must meet a minimum threshold identified by RNA-seq or may be determined based on comparative

mRNA expression levels of male and female samples (Yasukochi et al., 2010;Tukiainen et al., 2016;Talebizadeh et al., 2006). Because escapers are functionally active genes, they lack the repressive marks expected with the inactive X chromosome and instead are enriched in marks common to transcriptionally active genes (Berletch et al. 2011). This lack of methylation can be another indicator when determining escape genes (Cotton et al., 2011;Sharp et al., 2011).

Location of the escape genes is relevant to understanding the structure of the inactive X chromosome and how gene expression relates. Genes found within the pseudoautosomal regions (PARs) tend to escape X inactivation (Katsir and Linial 2019). In humans, escape genes are believed to located within clusters along the X chromosome (Tsuchiya et al. 2004; Carrel and Willard 2005).

A small group of escape genes contribute to male-female sex dimorphism (Snell and Turner 2018). Mutations in these genes has been linked to noncongenital diseases such as renal carcinoma and different forms of cancer (van Haaften et al. 2009; Dalgliesh et al. 2010). Another potent health concern regarding the misregulation of escape genes is breast cancer. Hypomethylation of the inactive X chromosome and reactivation can both appear in breast cancer cells (Chaligne et al. 2015;Sun et al. 2015).

Similar to XCD, XCI represents a gene expression state of the X chromosome where gene expression is suppressed on a large scale. Likewise, both gene expression states have genes that escape this suppression. XCI escape genes manage to avoid the silencing that a majority of genes on the Xi are affected by, and as expected, the active transcription of escape genes is reflected in their lack of repressive marks and location on the Xi.

Fibroblast Cells

Human fibroblast cells are a type of somatic cell line that is noticeably different than the previously discussed types of pluripotent stem cells. Fibroblast cells are mesenchymal-derived cells which play a role in epithelial differentiation as well as wound healing and have commonly been used as scaffolds for culturing of other cell types (Fernandes et al., 2014). Morphologically, fibroblast cells are characterized by their flat, stretched-out appearance. In relation to the X chromosome, the inactivation statuses as well as the transcriptional and methylation profiles are very clearly different.

The clear differences in these two cell types both morphologically and with regards to X chromosome gene expression state make these cell types as ideal candidates to study the differences in chromosome structure.

Confocal Microscopy

A common tool used to visualize cells at a sub nuclear level is a confocal microscope. Confocal microscopy is based on the principle of focusing a light source on a diffraction-limited spot via the objective lens of the microscope (Cox 2002). Commonly, a laser is employed as the goto light source for this purpose. To read the visual information, a detector receives light from the in-focus plane that is directed through a pinhole (Cox 2002). Confocal microscopy is not constrained to simply viewing samples as a flat image. By scanning multiple planes, a 3D reconstruction of the sample can be created.

3D image analysis has been used to identify chromosome territories, the organization of chromosomes into distinct regions within the nucleus (Cremer et al., 2006). Additionally,

quantitative 3D image analysis has been able to identify that gene-rich or actively transcribing chromatin is more spatially extensive. This has been identified in cells upon differentiation. Conversely, in the cases of inhibition of RNA Polymerase II or inhibition of histone deacetylation, the chromosomes became relatively smaller (Cremer et al., 2006). Image analysis of smaller, more specific regions of a chromosome is possible but must be met with a more specialized technique.

Fluorescence In Situ Hybridization

To specifically identify RNA transcripts or DNA regions of interest, fluorescence in situ hybridization (FISH) may be used in tandem with confocal microscopy. Fluorescence as a basic concept of science involves the idea that after absorption of high-energy photons via a light source, the subject will fluoresce at a longer wavelength. By utilizing the laser light source of a confocal microscope, a laser of a distinct wavelength can be targeted at the sample in order to trigger the specific excitation and subsequent fluorescence.

FISH probes are composed of complementary sequences to the RNA or DNA target of interest. The probe itself contains fluorescently labeled nucleotides. In order for these probes to be created, a template is necessary. One such source for a DNA template is a bacterial artificial chromosome (BAC). A BAC is a large plasmid housed in E. coli for purposes of transforming or cloning. BACs are particularly useful in their ability to hold ~200kb long sequences.

FISH has been previously utilized to view chromosome territories. With this technique, it has been shown that chromatin is organized on a nuclear scale with euchromatin often localizing towards in the inner regions of the nucleus whereas heterochromatin will more likely organize near the nuclear periphery (Bickmore et al., 2013; Cremer et al., 2006).

FISH provides a unique opportunity to probe specific regions of a chromosome in order to interrogate its actual structure within a cell. This technique serves as the cornerstone for the objectives of this project.

References

Akhade, V. S., Pal, D., & Kanduri, C. (2017). Long Noncoding RNA: Genome Organization and Mechanism of Action. *Advances in Experimental Medicine and Biology Long Non Coding RNA Biology*, 47–74. doi: 10.1007/978-981-10-5203-3_2

Albritton, S. E., & Ercan, S. (2018). Caenorhabditis elegans Dosage Compensation: Insights into Condensin-Mediated Gene Regulation. *Trends in Genetics, 34*(1), 41-53. doi:10.1016/j.tig.2017.09.010

Andergassen, D., Dotter, C. P., Wenzel, D., Sigl, V., Bammer, P. C., Muckenhuber, M., … Hudson, Q. J. (2017). Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression. *ELife*, *6*. doi: 10.7554/elife.25125

Augui, S., Nora, E.P., and Heard, E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. Nat. Rev. Genet. *12*, 429–442.

Belmont, A., Bignone, F., & Ts'o, P. (1986). The relative intranuclear positions of barr bodies in XXX non-transformed human fibroblasts. *Experimental Cell Research, 165*(1), 165-179. doi:10.1016/0014-4827(86)90541-0

Belton, J., Mccord, R. P., Gibcus, J. H., Naumova, N., Zhan, Y., & Dekker, J. (2012). Hi–C: A comprehensive technique to capture the conformation of genomes. *Methods, 58*(3), 268-276. doi:10.1016/j.ymeth.2012.05.001

Berkum, N. L., Lieberman-Aiden, E., Williams, L., Imakaev, M., Gnirke, A., Mirny, L. A., . . . Lander, E. S. (2010). Hi-C: A Method to Study the Three-dimensional Architecture of Genomes. *Journal of Visualized Experiments,* (39). doi:10.3791/1869

Berletch, J. B., Yang, F., Xu, J., Carrel, L., & Disteche, C. M. (2011). Genes that escape from X inactivation. *Human Genetics, 130*(2), 237-245. doi:10.1007/s00439-011-1011-z

Beygo, J., Citro, V., Sparago, A., Crescenzo, A. D., Cerrato, F., Heitmann, M., . . . Riccio, A. (2012). The molecular function and clinical phenotype of partial deletions of the IGF2/H19 imprinting control region depends on the spatial arrangement of the remaining CTCF-binding sites. *Human Molecular Genetics, 22*(3), 544-557. doi:10.1093/hmg/dds465

Bickmore, W., & Van Steensel, B. (2013). Genome Architecture: Domain Organization of Interphase Chromosomes. *Cell, 152*(6), 1270-1284. doi:10.1016/j.cell.2013.02.001

Bonnet, J., Devys, D., & Tora, L. (2014). Histone H2B ubiquitination: Signaling not scrapping. *Drug Discovery Today: Technologies, 12*. doi:10.1016/j.ddtec.2012.09.002

Bonora, G., Deng, X., Fang, H., Ramani, V., Qiu, R., Berletch, J. B., . . . Disteche, C. M. (2018). Orientation-dependent Dxz4 contacts shape the 3D structure of the inactive X chromosome. *Nature Communications, 9*(1). doi:10.1038/s41467-018-03694-y

Braccioli, L., & Wit, E. D. (2019). CTCF: A Swiss-army knife for genome organization and transcription regulation. *Essays in Biochemistry, 63*(1), 157-165. doi:10.1042/ebc20180069

Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S., & Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell, 71(3), 515–526. doi:10.1016/0092-8674(92)90519-i

Brockdorff, N. (2002). X-chromosome inactivation: closing in on proteins that bind Xist RNA. *Trends in Genetics*, *18*(7), 352–358. doi: 10.1016/s0168-9525(02)02717-8

Brockdorff, N. (2017). Polycomb complexes in X chromosome inactivation. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1733), 20170021. doi:10.1098/rstb.2017.0021

Brockdorff, N. (2019). Localized accumulation of Xist RNA in X chromosome inactivation. Open Biology, 9(12), 190213. doi:10.1098/rsob.190213

Brown, S. D. M. (1991). XIST and the mapping of the X chromosome inactivation centre. BioEssays, 13(11), 607–612. doi:10.1002/bies.950131112

Brown, C., Ballabio, A., and Rupert, J. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature.

Burcin, M., Arnold, R., Lutz, M., Kaiser, B., Runge, D., Lottspeich, F., . . . Renkawitz, R. (1997). Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. *Molecular and Cellular Biology, 17*(3), 1281-1288. doi:10.1128/mcb.17.3.1281

Bylino, O. V., Ibragimov, A. N., & Shidlovskii, Y. V. (2020). Evolution of Regulated Transcription. *Cells*, *9*(7), 1675. MDPI AG. doi:10.3390/cells9071675

Calo, E., & Wysocka, J. (2013). Modification of Enhancer Chromatin: What, How, and Why? *Molecular Cell, 49*(5), 825-837. doi:10.1016/j.molcel.2013.01.038

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039–1043.

Carrel, L., Cottle, A. A., Goglin, K. C., & Willard, H. F. (1999). A first-generation X-inactivation profile of the human X chromosome. *Proceedings of the National Academy of Sciences, 96*(25), 14440-14444. doi:10.1073/pnas.96.25.14440

Carrel, L., & Willard, H. F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature, 434*(7031), 400-404. doi:10.1038/nature03479

Cerase, A., Armaos, A., Neumayer, C., Avner, P., Guttman, M., & Tartaglia, G. G. (2019). Phase separation drives X-chromosome inactivation: a hypothesis. *Nature Structural & Molecular*

Biology, *26*(5), 331–334. doi: 10.1038/s41594-019-0223-0

Chaligné, R., Popova, T., Mendoza-Parra, M., Saleem, M. M., Gentien, D., Ban, K., . . . Heard, E. (2015). The inactive X chromosome is epigenetically unstable and transcriptionally labile in breast cancer. *Genome Research, 25*(4), 488-503. doi:10.1101/gr.185926.114

Clemson, C.M., McNeil, J. a, Willard, H.F., and Lawrence, J.B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J. Cell Biol. *132*, 259–275.

Colognori, D., Sunwoo, H., Kriz, A. J., Wang, C.-Y., & Lee, J. T. (2019). Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X. *Molecular Cell*, *74*(1). doi: 10.1016/j.molcel.2019.01.015

Cotton, A. M., Lam, L., Affleck, J. G., Wilson, I. M., Peñaherrera, M. S., Mcfadden, D. E., . . . Brown, C. J. (2011). Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. *Human Genetics, 130*(2), 187-201. doi:10.1007/s00439-011-1007-8

Cox, G. (2002). Biological confocal microscopy. *Materials Today, 5*(3), 34-41. doi:10.1016/s1369-7021(02)05329-4

Cremer, T., & Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Reviews Genetics*, *2*(4), 292–301. doi: 10.1038/35066075

Cremer, T., Cremer, M., Dietzel, S., Müller, S., Solovei, I., & Fakan, S. (2006). Chromosome territories – a functional nuclear landscape. *Current Opinion in Cell Biology, 18*(3), 307-316. doi:10.1016/j.ceb.2006.04.007

Darrow, E. M., Huntley, M. H., Dudchenko, O., Stamenova, E. K., Durand, N. C., Sun, Z., . . . Aiden, E. L. (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. *Proceedings of the National Academy of Sciences, 113*(31). doi:10.1073/pnas.1609643113

Dalgliesh, G. L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A., . . . Futreal, P. A. (2010). Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature, 463*(7279), 360-363. doi:10.1038/nature08672

Davie, J. R., & Murphy, L. C. (1990). Level of ubiquitinated histone H2B in chromatin is coupled to ongoing transcription. *Biochemistry, 29*(20), 4752-4757. doi:10.1021/bi00472a002

de Wit, E., Vos, E., Holwerda, S., Valdes-Quezada, C., Verstegen, M., Teunissen, H., . . . De Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell, 60*(4), 676-684. doi:10.1016/j.molcel.2015.09.023

Dekker, J. (2008). Gene Regulation in the Third Dimension. *Science, 319*(5871), 1793-1794. doi:10.1126/science.1152850

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P., . . . Blobel, G. (2012). Controlling

Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. *Cell, 149*(6), 1233-1244. doi:10.1016/j.cell.2012.03.051

Deng, X., Berletch, J. B., Nguyen, D. K., & Disteche, C. M. (2014). X chromosome regulation: Diverse patterns in development, tissues and disease. *Nature Reviews Genetics, 15*(6), 367-378. doi:10.1038/nrg3687

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D. G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J. B., Lipovich, L., Gonzalez, J. M., … Guigo, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Research, 22(9), 1775–1789. doi:10.1101/gr.132159.111

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., . . . Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature, 485*(7398), 376-380. doi:10.1038/nature11082

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., . . . Ren, B. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature, 518*(7539), 331-336. doi:10.1038/nature14222

Dong, X., & Weng, Z. (2013). The correlation between histone modifications and gene expression. *Epigenomics, 5*(2), 113-116. doi:10.2217/epi.13.13

Engel, N. (2015). Imprinted X chromosome inactivation offers up a double dose of epigenetics. Proceedings of the National Academy of Sciences, 112(47), 14408–14409. doi:10.1073/pnas.1520097112

Engreitz, J. M., Pandya-Jones, A., Mcdonel, P., Shishkin, A., Sirokman, K., Surka, C., . . . Guttman, M. (2013). The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome. *Science, 341*(6147), 1237973. doi:10.1126/science.1237973

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature, 292(5819), 154–156. doi:10.1038/292154a0

Factor, D., Corradin, O., Zentner, G., Saiakhova, A., Song, L., Chenoweth, J., . . . Tesar, P. (2014). Epigenomic Comparison Reveals Activation of "Seed" Enhancers during Transition from Naive to Primed Pluripotency. *Cell Stem Cell, 14*(6), 854-863. doi:10.1016/j.stem.2014.05.005

Faure, A. J., Schmidt, D., Watt, S., Schwalie, P. C., Wilson, M. D., Xu, H., . . . Flicek, P. (2012). Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Research, 22*(11), 2163-2175. doi:10.1101/gr.136507.111

Fedoriw, A. M. (2004). Transgenic RNAi Reveals Essential Function for CTCF in H19 Gene Imprinting. *Science, 303*(5655), 238-240. doi:10.1126/science.1090934

Feldmann, A., Ivanek, R., Murr, R., Gaidatzis, D., Burger, L., & Schübeler, D. (2013). Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions. *PLoS Genetics, 9*(12). doi:10.1371/journal.pgen.1003994

Fernandes, I. R., Russo, F. B., Pignatari, G. C., Evangelinellis, M. M., Tavolari, S., Muotri, A. R., & Beltrão-Braga, P. C. (2014). Fibroblast sources: Where can we get them? *Cytotechnology, 68*(2), 223-228. doi:10.1007/s10616-014-9771-7

Froberg, J. E., Pinter, S. F., Kriz, A. J., Jégu, T., & Lee, J. T. (2018). Megadomains and superloops form dynamically but are dispensable for X-chromosome inactivation and gene escape. *Nature Communications*, *9*(1). doi: 10.1038/s41467-018-07446-w

Fudenberg, G., & Mirny, L. A. (2012). Higher-order chromatin structure: Bridging physics and biology. *Current Opinion in Genetics & Development, 22*(2), 115-124. doi:10.1016/j.gde.2012.01.006

Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., . . . Liu, P. (2019). Establishment of porcine and human expanded potential stem cells. *Nature Cell Biology, 21*(6), 687-699. doi:10.1038/s41556-019-0333-2

Giacalone, J., Friedes, J., & Francke, U. (1992). A novel GC–rich human macrosatellite VNTR in Xq24 is differentially methylated on active and inactive X chromosomes. *Nature Genetics, 1*(2), 137-143. doi:10.1038/ng0592-137

Giorgetti, L., Galupa, R., Nora, E., Piolot, T., Lam, F., Dekker, J., . . . Heard, E. (2014). Predictive Polymer Modeling Reveals Coupled Fluctuations in Chromosome Conformation and Transcription. *Cell, 157*(4), 950-963. doi:10.1016/j.cell.2014.03.025

Goetze, S., Mateos-Langerak, J., Gierman, H. J., Leeuw, W. D., Giromus, O., Indemans, M. H., . . . Driel, R. V. (2007). The Three-Dimensional Structure of Human Interphase Chromosomes Is Related to the Transcriptome Map. *Molecular and Cellular Biology, 27*(12), 4475-4487. doi:10.1128/mcb.00208-07

Gorisch, S. M. (2005). Histone acetylation increases chromatin accessibility. Journal of Cell Science, 118(24), 5825–5834. doi:10.1242/jcs.02689

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature, 389*(6649), 349-352. doi:10.1038/38664

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D., . . . Wu, Q. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell, 162*(4), 900-910. doi:10.1016/j.cell.2015.07.038

Hadjur, S., Williams, L. M., Ryan, N. K., Cobb, B. S., Sexton, T., Fraser, P., . . . Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature, 460*(7253), 410-413. doi:10.1038/nature08079

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C., & Spector, D. L. (2001). Methylation of Histone H3 at Lys-9 Is an Early Mark on the X Chromosome during X Inactivation. *Cell, 107*(6), 727-738. doi:10.1016/s0092-8674(01)00598-0

Henry, K. W. (2003). Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes & Development, 17*(21), 2648-2663. doi:10.1101/gad.1144003

Hershey, J. W. B., Sonenberg, N., & Mathews, M. B. (2012). Principles of Translational Control: An Overview. Cold Spring Harbor Perspectives in Biology, 4(12), a011528–a011528. https://doi.org/10.1101/cshperspect.a01152

Hodges, E., Smith, A. D., Kendall, J., Xuan, Z., Ravi, K., Rooks, M., . . . Hicks, J. B. (2009). High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Research, 19*(9), 1593-1605. doi:10.1101/gr.095190.109

Holwerda, S., & Laat, W. D. (2012). Chromatin loops, gene positioning, and gene expression. *Frontiers in Genetics*, *3*. doi: 10.3389/fgene.2012.00217

Jeon, Y., & Lee, J. T. (2011). YY1 Tethers Xist RNA to the Inactive X Nucleation Center. *Cell*, *146*(1), 119–133. doi: 10.1016/j.cell.2011.06.026

Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., . . . Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature, 503*(7475), 290-294. doi:10.1038/nature12644

Jonkers, I., Monkhorst, K., Rentmeester, E., Grootegoed, J. A., Grosveld, F., & Gribnau, J. (2008). Xist RNA Is Confined to the Nuclear Territory of the Silenced X Chromosome throughout the Cell Cycle. Molecular and Cellular Biology, 28(18), 5583–5594. doi:10.1128/mcb.02269-07

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., Berkum, N. L., . . . Young, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature, 467*(7314), 430-435. doi:10.1038/nature09380

Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., & Chen, L. (2011). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nature Biotechnology, 30*(1), 90-98. doi:10.1038/nbt.2057

Katsir, K., & Linial, M. (2019). Human genes escaping X-inactivation revealed by single cell expression data. BMC Genomics, 20(1). doi:10.1186/s12864-019-5507-6

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J., & Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. Science, 366(6471), 1345–1349. https://doi.org/10.1126/science.aaz4475

Kosak, S. T. (2004). Form follows function: The genomic organization of cellular differentiation. *Genes & Development, 18*(12), 1371-1384. doi:10.1101/gad.1209304

Kugel, J. F., & Goodrich, J. A. (2012). Non-coding RNAs: Key regulators of mammalian transcription. *Trends in Biochemical Sciences, 37*(4), 144-151. doi:10.1016/j.tibs.2011.12.003

Le Dily, F., Baù, D., Pohl, A., Vicent, G. P., Serra, F., Soronellas, D., . . . Beato, M. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes & Development, 28*(19), 2151-2162. doi:10.1101/gad.241422.114

Li, C., Liu, B., Zhong, S., & Wang, H. (2016). MEK inhibitor PD0325901 and vitamin C synergistically induce hypomethylation of mouse embryonic stem cells. *Oncotarget, 7*(26), 39730-39739. doi:10.18632/oncotarget.9452

Lieberman-Aiden, E., Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., . . . Dekker, J. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science, 326*(5950), 289-293. doi:10.1126/science.1181369

Liu, G., David, B. T., Trawczynski, M., & Fessler, R. G. (2019). Advances in Pluripotent Stem Cells: History, Mechanisms, Technologies, and Applications. *Stem Cell Reviews and Reports, 16*(1), 3-32. doi:10.1007/s12015-019-09935-x

Loda, A., & Heard, E. (2019). Xist RNA in action: Past, present, and future. PLOS Genetics, 15(9), e1008333. doi:10.1371/journal.pgen.1008333

Lyon, M. F. (1961). Gene Action in the X-chromosome of the Mouse (Mus musculus L.). *Nature, 190*(4773), 372-373. doi:10.1038/190372a0

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proceedings of the National Academy of Sciences, 78(12), 7634–7638. doi:10.1073/pnas.78.12.7634

McHugh, C. A., Chen, C.-K., Chow, A., Surka, C. F., Tran, C., Mcdonel, P., … Guttman, M. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, *521*(7551), 232–236. doi: 10.1038/nature14443

Merkenschlager, M., & Nora, E. P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annual Review of Genomics and Human Genetics, 17*(1), 17-43. doi:10.1146/annurev-genom-083115-022339

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., . . . Lee, J. T. (2015). A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science, 349*(6245). doi:10.1126/science.aab2276

Minkovsky, A., Patel, S., & Plath, K. (2011). Concise Review: Pluripotency and the Transcriptional Inactivation of the Female Mammalian X Chromosome. STEM CELLS, 30(1), 48–54. doi:10.1002/stem.755

Misteli, T. (2007). Beyond the Sequence: Cellular Organization of Genome Function. *Cell, 128*(4), 787-800. doi:10.1016/j.cell.2007.01.028

Mondal, T., Rasmussen, M., Pandey, G. K., Isaksson, A., & Kanduri, C. (2010). Characterization of the RNA content of chromatin. *Genome Research, 20*(7), 899-907. doi:10.1101/gr.103473.109

Nagano, T., & Fraser, P. (2011). No-Nonsense Functions for Long Noncoding RNAs. *Cell, 145*(2), 178-181. doi:10.1016/j.cell.2011.03.014

Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., . . . Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature, 502*(7469), 59-64. doi:10.1038/nature12593

Nasmyth, K., & Haering, C. H. (2009). Cohesin: Its Roles and Mechanisms. *Annual Review of Genetics, 43*(1), 525-558. doi:10.1146/annurev-genet-102108-134233

Nesterova, T. B., Wei, G., Coker, H., Pintacuda, G., Bowness, J. S., Zhang, T., Almeida, M., Bloechl, B., Moindrot, B., Carter, E. J., Alvarez Rodrigo, I., Pan, Q., Bi, Y., Song, C.-X., & Brockdorff, N. (2019). Systematic allelic analysis defines the interplay of key pathways in X chromosome inactivation. Nature Communications, 10(1). doi:10.1038/s41467-019-11171-3

Nickel, B. E., Allis, C. D., & Davie, J. R. (1989). Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochemistry, 28*(3), 958-963. doi:10.1021/bi00429a006

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., . . . Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature, 485*(7398), 381-385. doi:10.1038/nature11049

Nora, E. P., Dekker, J., & Heard, E. (2013). Segmental folding of chromosomes: A basis for structural and regulatory chromosomal neighborhoods? *BioEssays, 35*(9), 818-828. doi:10.1002/bies.201300040

Pandya-Jones, A., Markaki, Y., Serizay, J., Chitiashvilli, T., Mancia, W., Damianov, A., … Plath, K. (2020). An Xist-dependent protein assembly mediates Xist localization and gene silencing. doi: 10.1101/2020.03.09.979369

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., . . . Merkenschlager, M. (2008). Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell, 132*(3), 422-433. doi:10.1016/j.cell.2008.01.011

Pastor, W., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., . . . Clark, A. (2016). Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell, 18*(3), 323-329. doi:10.1016/j.stem.2016.01.019

Penny, G., Kay, G., and Sheardown, S. (1996). Requirement for Xist in X chromosome inactivation. Nature.

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S., Codeluppi, S., . . . Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell, 165*(4), 1012-1026. doi:10.1016/j.cell.2016.03.023

Phillips, J. E., & Corces, V. G. (2009). CTCF: Master Weaver of the Genome. *Cell, 137*(7), 1194-1211. doi:10.1016/j.cell.2009.06.001

Phillips-Cremins, J., Sauria, M., Sanyal, A., Gerasimova, T., Lajoie, B., Bell, J., . . . Corces, V. (2013). Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. *Cell, 153*(6), 1281-1295. doi:10.1016/j.cell.2013.04.053

Pierro, M. D. (2019). Inner workings of gene folding. *Proceedings of the National Academy of Sciences, 116*(11), 4774-4775. doi:10.1073/pnas.1900875116

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K. a, Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. Science (80-.). *300*, 131–135.

Pope, B. D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., . . . Gilbert, D. M. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature, 515*(7527), 402-405. doi:10.1038/nature13986

Pugacheva, E. M., Kubo, N., Loukinov, D., Tajmul, M., Kang, S., Kovalchuk, A. L., Strunnikov, A. V., Zentner, G. E., Ren, B., & Lobanenkov, V. V. (2020). CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. Proceedings of the National Academy of Sciences, 117(4), 2020–2031. https://doi.org/10.1073/pnas.1911708117

Rao, S., Huntley, M., Durand, N., Stamenova, E., Bochkov, I., Robinson, J., . . . Aiden, E. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell, 159*(7), 1665-1680. doi:10.1016/j.cell.2014.11.021

Rao, S. S. P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D., Bochkov, I. D., Huang, X., Shamim, M. S., Shin, J., Turner, D., Ye, Z., Omer, A. D., Robinson, J. T., Schlick, T., Bernstein, B. E., … Aiden, E. L. (2017). Cohesin Loss Eliminates All Loop Domains. Cell, 171(2), 305-320.e24. doi:10.1016/j.cell.2017.09.026

Rasmussen, T.P., Wutz, a P., Pehrson, J.R., and Jaenisch, R.R. (2001). Expression of Xist RNA is sufficient to initiate macrochromatin body formation. Chromosoma *110*, 411–420.

Rastan, S. (1983). Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. J. Embryol. Exp. Morphol. *78*, 1–22.

Rinn, J. L., & Chang, H. Y. (2012). Genome Regulation by Long Noncoding RNAs. *Annual Review of Biochemistry, 81*(1), 145-166. doi:10.1146/annurev-biochem-051410-092902

Rubio, E. D., Reiss, D. J., Welcsh, P. L., Disteche, C. M., Filippova, G. N., Baliga, N. S., . . .

Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proceedings of the National Academy of Sciences, 105*(24), 8309-8314. doi:10.1073/pnas.0801273105

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T. W., . . . Plath, K. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell, 20*(1), 87-101. doi:10.1016/j.stem.2016.10.006

Sahakyan, A., Yang, Y., & Plath, K. (2018). The Role of Xist in X-Chromosome Dosage Compensation. *Trends in Cell Biology*, *28*(12), 999–1013. doi: 10.1016/j.tcb.2018.05.005

Sanborn, A. L., Rao, S. S., Huang, S., Durand, N. C., Huntley, M. H., Jewett, A. I., . . . Aiden, E. L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences, 112*(47). doi:10.1073/pnas.1518552112

Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., Hurtado, A., Brown, G. D., Carroll, J. S., . . . Odom, D. T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Research, 20*(5), 578-588. doi:10.1101/gr.100479.109

Sharp, A. J., Stathaki, E., Migliavacca, E., Brahmachary, M., Montgomery, S. B., Dupre, Y., & Antonarakis, S. E. (2011). DNA methylation profiles of human active and inactive X chromosomes. *Genome Research, 21*(10), 1592-1600. doi:10.1101/gr.112680.110

Sidhu, S. K., Minks, J., Chang, S. C., Cotton, A. M., & Brown, C. J. (2008). X chromosome inactivation: heterogeneity of heterochromatin. Biochemistry and Cell Biology, 86(5), 370–379. doi:10.1139/o08-100

Sleutels, F., Soochit, W., Bartkuhn, M., Heath, H., Dienstbach, S., Bergmaier, P., . . . Galjart, N. (2012). The male germ cell gene regulator CTCFL is functionally different from CTCF and binds CTCF-like consensus sites in a nucleosome composition-dependent manner. *Epigenetics & Chromatin, 5*(1). doi:10.1186/1756-8935-5-8

Smith, Z. D., Chan, M. M., Humm, K. C., Karnik, R., Mekhoubad, S., Regev, A., . . . Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. *Nature, 511*(7511), 611-615. doi:10.1038/nature13581

Snell, D. M., & Turner, J. M. A. (2018). Sex Chromosome Effects on Male–Female Differences in Mammals. Current Biology, 28(22), R1313–R1324. doi:10.1016/j.cub.2018.09.018

Soshnikova, N., Montavon, T., Leleu, M., Galjart, N., & Duboule, D. (2010). Functional Analysis of CTCF During Mammalian Limb Development. *Developmental Cell, 19*(6), 819-830. doi:10.1016/j.devcel.2010.11.009

Splinter, E. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes & Development, 20*(17), 2349-2354. doi:10.1101/gad.399506

Splinter, E., Wit, E. D., Nora, E. P., Klous, P., H. J. G. Van De Werken, Zhu, Y., . . . Laat, W. D. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is

dependent on Xist RNA. *Genes & Development, 25*(13), 1371-1383. doi:10.1101/gad.633311

Sprague, D., Waters, S. A., Kirk, J. M., Wang, J. R., Samollow, P. B., Waters, P. D., & Calabrese, J. M. (2019). Non-linear sequence similarity between the Xist and Rsx long noncoding RNAs suggests shared functions of tandem repeat domains. doi: 10.1101/630475

Stedman, W., Kang, H., Lin, S., Kissil, J. L., Bartolomei, M. S., & Lieberman, P. M. (2008). Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *The EMBO Journal, 27*(4), 654-666. doi:10.1038/emboj.2008.1

Sun, Z., Prodduturi, N., Sun, S. Y., Thompson, E. A., & Kocher, J. A. (2015). Chromosome X genomic and epigenomic aberrations and clinical implications in breast cancer by base resolution profiling. *Epigenomics, 7*(7), 1099-1110. doi:10.2217/epi.15.43

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. *Nature Reviews Genetics, 9*(6), 465-476. doi:10.1038/nrg2341

Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., . . . Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research, 24*(3), 390-400. doi:10.1101/gr.163519.113

Takagi, N. (1978). Preferential Inactivation of the Paternally Derived X Chromosome in Mice. In Genetic Mosaics and Chimeras in Mammals (pp. 341–360). Springer US. doi:10.1007/978-1-4684-3390-6_24

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell, 126*(4), 663-676. doi:10.1016/j.cell.2006.07.024

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., . . . Smith, A. (2014). Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell, 158*(6), 1254-1269. doi:10.1016/j.cell.2014.08.029

Talebizadeh, Z., Simon, S. D., & Butler, M. G. (2006). X chromosome gene expression in human tissues: Male and female comparisons. *Genomics, 88*(6), 675-681. doi:10.1016/j.ygeno.2006.07.016

Theunissen, T., Powell, B., Wang, H., Mitalipova, M., Faddah, D., Reddy, J., . . . Jaenisch, R. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell, 15*(4), 471-487. doi:10.1016/j.stem.2014.07.002

Theunissen, T., Friedli, M., He, Y., Planet, E., O'Neil, R., Markoulaki, S., . . . Jaenisch, R. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell, 19*(4), 502-515. doi:10.1016/j.stem.2016.06.011

Tsuchiya, K. D. (2004). Comparative Sequence and X-Inactivation Analyses of a Domain of Escape in Human Xp11.2 and the Conserved Segment in Mouse. *Genome Research, 14*(7),

1275-1284. doi:10.1101/gr.2575904

Tukiainen, T., Villani, A., Yen, A., Rivas, M. A., Marshall, J. L., Satija, R., . . . Macarthur, D. G. (2016). Landscape of X chromosome inactivation across human tissues. doi:10.1101/073957

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Ali, T. M. L., … Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell*, *20*(1), 102–111. doi: 10.1016/j.stem.2016.10.014

van de Werken, H. J. G., Landan, G., Holwerda, S. J. B., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Öz, Y., Bouwman, B. A. M., Verstegen, M. J. A. M., de Wit, E., Tanay, A., & de Laat, W. (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nature Methods, 9(10), 969–972. doi:10.1038/nmeth.2173

van den Berg, I. M., Laven, J. S. E., Stevens, M., Jonkers, I., Galjaard, R.-J., Gribnau, J., & Hikke van Doorninck, J. (2009). X Chromosome Inactivation Is Initiated in Human Preimplantation Embryos. The American Journal of Human Genetics, 84(6), 771–779. doi:10.1016/j.ajhg.2009.05.003

van Haaften, G., Dalgliesh, G. L., Davies, H., Chen, L., Bignell, G., Greenman, C., . . . Futreal, P. A. (2009). Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nature Genetics, 41*(5), 521-523. doi:10.1038/ng.349

Varley, K. E., Gertz, J., Bowling, K. M., Parker, S. L., Reddy, T. E., Pauli-Behn, F., . . . Myers, R. M. (2013). Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Research, 23*(3), 555-567. doi:10.1101/gr.147942.112

Venters, B. J., & Pugh, B. F. (2009). How eukaryotic genes are transcribed. *Critical Reviews in Biochemistry and Molecular Biology, 44*(2-3), 117-141. doi:10.1080/10409230902858785

Wan, L., Pan, H., Hannenhalli, S., Cheng, Y., Ma, J., Fedoriw, A., . . . Bartolomei, M. S. (2008). Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development, 135*(16), 2729-2738. doi:10.1242/dev.024539

Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., . . . Chang, H. Y. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature, 472*(7341), 120-124. doi:10.1038/nature09819

Wang, H., Maurano, M. T., Qu, H., Varley, K. E., Gertz, J., Pauli, F., . . . Stamatoyannopoulos, J. A. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Research, 22*(9), 1680-1688. doi:10.1101/gr.136101.111

Wang, T., Kang, W., Du, L., & Ge, S. (2017). Rho-kinase inhibitor Y-27632 facilitates the proliferation, migration and pluripotency of human periodontal ligament stem cells. *Journal of Cellular and Molecular Medicine, 21*(11), 3100-3112. doi:10.1111/jcmm.13222

Wijchers, P. J., Krijger, P. H., Geeven, G., Zhu, Y., Denker, A., Verstegen, M. J., . . . Laat, W. D. (2016). Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments. *Molecular Cell, 61*(3), 461-473. doi:10.1016/j.molcel.2016.01.001

Wolff, J., Backofen, R., & Grüning, B. (2020). Loop detection using Hi-C data with HiCExplorer. doi:10.1101/2020.03.05.979096

Yan, J., Enge, M., Whitington, T., Dave, K., Liu, J., Sur, I., . . . Taipale, J. (2013). Transcription Factor Binding in Human Cells Occurs in Dense Clusters Formed around Cohesin Anchor Sites. *Cell, 154*(4), 801-813. doi:10.1016/j.cell.2013.07.034

Yang, F., Babak, T., Shendure, J., & Disteche, C. M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Research, 20*(5), 614-622. doi:10.1101/gr.103200.109

Yao, R.-W., Wang, Y., & Chen, L.-L. (2019). Cellular functions of long noncoding RNAs. Nature Cell Biology, 21(5), 542–551. doi:10.1038/s41556-019-0311-8

Yap, K. L., Li, S., Muñoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., . . . Zhou, M. (2010). Molecular Interplay of the Noncoding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a. *Molecular Cell, 38*(5), 662-674. doi:10.1016/j.molcel.2010.03.021

Yasukochi, Y., Maruyama, O., Mahajan, M. C., Padden, C., Euskirchen, G. M., Schulz, V., . . . Weissman, S. M. (2010). X chromosome-wide analyses of genomic DNA methylation states and gene expression in male and female neutrophils. *Proceedings of the National Academy of Sciences, 107*(8), 3704-3709. doi:10.1073/pnas.0914812107

Zhao, J., Sun, B. K., Erwin, J. A., Song, J.-J., & Lee, J. T. (2008). Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. Science, 322(5902), 750–756. doi:10.1126/science.1163045

Zhu, P., & Li, G. (2016). Structural insights of nucleosome and the 30-nm chromatin fiber. *Current Opinion in Structural Biology, 36*, 106-115. doi:10.1016/j.sbi.2016.01.013

Zuin, J., Dixon, J. R., M. I. J. A. Van Der Reijden, Ye, Z., Kolovos, P., Brouwer, R. W., . . . Wendt, K. S. (2013). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences, 111*(3), 996-1001. doi:10.1073/pnas.1317788111

Chapter 2

Sequential FISH Method Establishment

Abstract

The X chromosome is known to change in structure as it transitions through different activation states predicated on vast gene expression changes. In order to study these alterations in chromatin folds accompanying each activation state, we developed a method that allows for the imaging of multiple, informative regions along the chromosome that can be extracted for use in structure analysis, all within the working limitations of an ordinary biology-focused laboratory, permitting access to a quality confocal microscope. We found that the simplest and most effective protocol involved the utilization of sequential rounds of DNA FISH probe hybridization combined with a treatment to strip the probes from the sample. We found that a relocation system employing a tissue culture dish equipped with a high-quality imaging surface rather than a conventional coverslip approach in addition to an alignment method exploiting multi-channel fluorescent microspheres as landmarks is effective for relocating the X chromosome from round to round. This protocol serves as the foundation for the interrogation of the three-dimensional structure of the Xa, Xi and Xd.

Introduction

Chromosome structure and gene expression are intimately linked (Bickmore et al., 2013; Dixon et al., 2015). The quintessential example of this phenomenon is the case of the X chromosome in the context of XCI (Disteche and Berletch 2015). In this example, the inactive X shrinks in size by 30% and adopts a new folded structure based on a dual super domain configuration originating around the DXZ4 microsatellite (Bonora et al., 2018). The established inactive X houses chromosome wide repressive marks and a vast number of genes are silenced with a small number of XCI escapers still showing any level of gene expression (Carrel et al.

1999; Yang et al. 2010). XCD is another state of X chromosome activation where, unlike the nearly complete silencing the inactive X undergoes, the dampened X chromosome displays a reduction of gene expression aside from XCD escapers (Sahakyan et al., 2017; Vallot et al., 2017). However, the structure of the dampened X, either by itself or in relation to the inactive or active X, has yet to be defined. In order to understand this novel state of the X chromosome we sought to develop a technique to interrogate its structure.

While it has been possible to view chromosomes in totality due to chromosome paint techniques, the overall image lacks resolution when attempting to focus on particular regions or chromosome folds (Cremer and Cremer 2001). As a method, FISH has previously been applied to highlight specific genes within chromosomes with high specificity. However, a limitation arises when considering viewing multiple regions simultaneously. Detection of unique colors of light is limited to a few colors at a time. This limitation can be circumvented employing combinatorial labeling spectral barcoding or sequential FISH.

Spectral barcoding is based on the principle that different combinations of fluorescent probes can be resolved and subsequently linked to the specific regions they are targeted at (Lubek and Cai 2012). This technique can identify multiple regions of a chromosome at the same time but does have limitations. In particular, the number of probe hybridizations that must occur in order to acquire informative data increases at a considerable rate the more locations one is interested in. For example, identifying nine distinct locations along the chromosome would require four different, resolvable fluorophores with fourteen separate hybridization events. While this is a minor microscopy issue the more pervasive problem is ensuring all hybridization events occur, as FISH, while a powerful technique, is sometimes fickle in its execution.

Sequential FISH is a technique that circumvents the dependency on a high number of

simultaneous hybridization events by probing targets over multiple rounds (Lubeck et al., 2014; Wang et al., 2016). One hybridization event per DNA target, using the same fluorophores from round to round serves to simplify the procedure and the analysis. This can either be accomplished as a direct binding event or as a fluorescent secondary probe hybridizing to a primary probe which serves as a targeting scaffold (Wang et al., 2016). However, the sequential FISH technique is not without its pitfalls. Firstly, there is the issue concerning how to remove the fluorescent signals of the previous round of probes. Secondly, there is the issue of aligning the imaging of the subject from round to round of hybridization.

Photobleaching of fluorescent signals is a common method of removal. In this method, a high-powered laser light is directed at the probes which causes the end of their fluorescent lifetime, effectively photobleaching them. For example, each field of view (FOV), the region of the sample within focus of the microscope, will be illuminated with 640-nm light for 12 seconds (Wang et al., 2016). This has previously been shown to work in strategies utilizing the method of hybridizing primary and secondary probes. Alternatively, there exist protocols that remove hybridized probes using a formamide solution. In this case the probes are no longer bound to the DNA and merely wash off.

Image alignment has previously been solved by avoiding the problem completely using a fixed stage which locks the sample in for the entirety of the protocol, including the microscopy as well as the washes and probe hybridization (Wang et al., 2016). While this system is effective, it is not particularly practical for an ordinary biology focused laboratory. Here, we developed and optimized a sequential FISH method, based on a multi-step hybridization, relocation, and image alignment approach to determine the structure of the X, that targets several genes per round in order to acquire spatial data at a single chromosome level.

Results

Glass bottom chamber slides allow for Probe Hybridization and Relocation

Selection of a proper viewing apparatus is critical for the retrieval of quality microscopy data. For acquiring super-resolution or close to super-resolution microscopy data, the slide must meet standards such as oil compatibility, a proper thickness compatible for such microscopes, and exhibiting low levels of autofluorescence so as not to reduce the quality of the image.

Glass coverslips provide an assortment of issues that must be overcome. Removing and replacing the coverslip onto the glass slide is problematic for relocation of the cells that were previously imaged as well as for potentially damaging the structural integrity of the cells.

The chamber slides that were chosen are equipped with natural borders that can serve as landmarks that can assist in relocation of the previously imaged cells (Fig 2-1a). We found that using the edges of the chambers can provide a point of reference when searching for the same FOV of cells. Normally this limitation could be ignored with the use of an automated stage that can move to previously saved locations representing FOVs. In our case, the confocal microscope is part of a shared microscopy core where the stage position can be reset or altered over time. By mapping out the location of the FOVs of the sample the automated stage can still be of use but is not a necessity (Fig 2-1b).

FISH Probe Signals are Separate and Resolvable

Using BACs to generate fluorescently labeled DNA probes, we have been able to clearly define three targets on the X chromosome at the same time. By selecting the Atto 488, Texas Red or Cy3, and Cy5 fluorophores we were able to detect unique, resolvable signals in both chromosomes (Fig 2-2a). Although some overlap may appear in the signals due to close

proximity of genes location along the chromosome, the wavelengths of emitted fluorescence display a wide range and therefore the signals are independent.

We also tested used XIST RNA FISH to mark the non-active X chromosome in the cells (Fig 2-2b). This hybridization takes place before the DNA FISH. The XIST cloud will only be present over inactive and dampened X chromosomes, assisting in downstream analysis.

Probe Stripping and Rehybridization are Effective for Sequential FISH

After imaging a round of hybridized DNA probes, it becomes imperative to remove their fluorescence from remaining in the next round. Using a solution of 70% formamide in 2xSodium-saline citrate (SSC) buffer (2xSSC + 70% formamide) and placing the glass bottom chamber slide on a 70C heat block for 5 minutes, DNA probes release from the chromosome. Formamide is a denaturant that is found in many hybridization protocols as a way to modulate stringency (Yilmaz et al., 2012; He et al., 2007; Manz et al., 1992). Based on the lack of signal originating from the locations found in the previous round we have determined that the stripping and rehybridization of DNA probes is effective within the context of our sequential FISH protocol(Fig 2-3)

Multi-channel Fluorescent Microspheres Provide a Visual and Physical Landmark for Image Alignment

Multi-channel fluorescent microspheres (TetraSpeck) appear in each channel of the microscope during imaging. This property allows for the microspheres to be utilized as physical landmarks for alignment purposes. 100nm microspheres have previously been shown to withstand 20 rounds of mock hybridization involving 3 minutes on an 80°C heat block in a

2xSSC + 50% formamide solution for an initial hybridization with subsequent hybridizations involving 30 minute room temperature incubations in 2xSSC + 30% formamide and washes of 2xSSC + 40% formamide, with both while maintaining an average localization precision of 50 nm (Wang et al., 2016). The microspheres are added to the bottom of the wells before the addition of Geltrex hESC-Qualified, Ready-To-Use, Reduced Growth Factor Basement Membrane Matrix (Geltrex). The second application of microspheres is during the fixation of cells. This ensures that the microspheres will be present both above and below the cells. We find that using multi-channel fluorescent microspheres and in combination with the MultiStackReg function of Fiji, beads align across three rounds of probe hybridization with an average of <200nm of movement (Fig 2-4). This indicates that the microspheres are useful for image alignment over multiple rounds without extra preparation at the microscope. This also points towards the stability of the fixed microspheres, as the matrix of microspheres retains its overall structure in each FOV (Fig 2-4).

Multi-channel Fluorescent Microspheres signals in multiple channels define Chromatic Aberration in Images

TetraSpeck multi-channel fluorescent microspheres serve another purpose within the sample. We found that the multi-channel properties of the microspheres are excellent tools for calibrating images for adjusting for chromatic aberration. Chromatic aberration is a common microscopy issue where the lens fails to focus all colors to a singular point. In our case, the multi-channel microsphere is physically the same point and therefore all colors should align at its location. However, in images captured via microscope it is possible to see a ghosting or halo like effect of light from other channels when in reality the light source originates from the same point

(Fig 2-5). By extracting the location of the microspheres using their light intensity in each channel, we have calculated the chromatic shift between the channels as a way to obtain more precise location information.

Image Alignment over Multiple Rounds of Hybridization Ensures Accurate Spatial Coordinates

Proper alignment for multiple rounds of images taken over the course of days with many reapplications of the chamber well to the microscope stage must correct for disparities in all three dimensions. Initially, stacks of images should be trimmed down so that the nuclei start to appear at roughly the same slice number of each stack of images that are part of the same FOV. From here, the MultiStackReg function of Fiji is applied to both stacks, using the first round of DNA probe hybridization as the reference stack. This process rotates each image of a stack to best match the image of the reference stack. However, this initial processing only aligns in the X and Y axes. To align the stacks of images in the Z axis, the stacks of images must be resliced, essentially to viewing the FOV from a side-to-side arrangement rather than a top-to-bottom arrangement. By utilizing the MultiStackReg function a second time the alignment is now present in all three axes. We find that after aligning images with reslicing we are able to transform all subsequent rounds of hybridizations into proper configuration (Fig 2-6).

Discussion

By attempting to find a method that will allow us to properly interrogate the structure of the dampened X chromosome, we have ultimately developed a technique that is flexible for chromosome structure investigation, scalable, and accessible to an ordinary biology focused laboratory. So long as there is access to a high-resolution microscope and a few materials, such as multi-channel microspheres and glass bottom chamber wells, this protocol can acquire data of 3D locations of specific regions of a chromosome.

A complication of the alignment procedure is that much if it currently depends on manual inputs from the user rather than a completely automated system. Matching starting images from stack to stack is problematic as the edges of a nucleus may not totally align based on how large the z sections are. Logically, the more z sections of a sample one takes during the imaging session would mitigate this issue but unfortunately will drastically increase the imaging time per sample, exposing the sample to the laser for a significantly larger amount of time which can lead to the unintentional photobleaching of fluorescent probes as well as damage to the sample itself.

When extracting probe locations using their high intensity a lack of automation is also present. 3D Objects Counter using a sliding threshold in order to identify high intensity regions. While it theoretically should be possible to set a baseline or a standard for the function to adhere to there are scenarios where this will not yield quality results. For example, debris found within the sample may appear as false positives or distort the actual fluorescent probes. Of course, more thorough washes can help to alleviate this issue but it must be a consideration nonetheless.

Multi-channel microspheres applied to the sample provide a physical landmark for alignment in addition to providing a source to calculate chromatic aberration. While the microspheres applied to the base of the well at the face of the glass bottom slide stay affixed for

the entirety of the multiple hybridizations, the microspheres introduced to the sample during the cell fixation step have the capacity to wash off during the multiple rounds of hybridization.

We hybridized four rounds of DNA probes and found that alignment was still within the 200nm precision limit. We did observe that the quality of the DAPI staining with regards to the nuclei appeared to become less effective from round to round indicating a difficulty in re-staining or degradation of the nuclei.

We believe that this method is practical for imaging multiple rounds of DNA FISH hybridizations due to the ability to resolve multiple signals, efficiency of probe stripping, and the reliability of image alignment.

Materials & Methods

Probe Design and Synthesis

We generated the fluorescently labeled probes from a BAC library containing genes found across the human X chromosome. The nick translation mixture is composed of 1 µg of the BAC of choice, 5 µL of NT Buffer, 5 µL of 0.1M β-Me, 5 µL of dNTP Mix 1mM, 3 µL of fluorescently labeled dUTP (Atto 488 (Sigma-Aldrich), Cy3 (Sigma-Aldrich), Texas Red (Thermo Fisher Scientific), Cy5 (VWR)), 2 µL of DNase (2000U/mL) at 1:100 dilution in cold water, 1 µL of Pol I (10U/µL), and the remainder of the 50 µL volume as water. Using a thermocycler, incubate the mixture at 15°C for 12 hours. Then stop the reaction by heating the mixture to 65°C for 10 minutes.

Purification of labelled DNA can be accomplished using AMPure beads. After adding 120 µL of DNA binding AMPure beads to the reaction mixture, pipette to gently resuspend the new mixture. Incubate at room temperature for 10 minutes under foil or a light-blocking container. Transfer the mixture to a magnetic tube rack and again incubate at room temperature for 10 minutes under foil or a light-blocking container. While the tube containing the mixture remains on the magnetic rack, remove the supernatant and add 400 µL of freshly prepared 80% ethanol (EtOH). Proceed to incubate at room temperature for 10 minutes under foil or a light-blocking container. Again remove the supernatant while the tubs are still attached to the magnetic rack and repeat the 80% EtOH wash. Add 100 µL of water to just the AMPure beads, resuspend, and incubate the tube separate from the magnetic rack for 10 minutes. Replace the tube onto the magnetic rack and once again incubate for 10 minutes. Retrieve the purified product.

Combine the purified product with 10 µL of COT-1 DNA, 10 µL of Salmon Sperm

DNA, 340 µL of 100% EtOH, and 15 µL of 3M Sodium Acetate pH 5.2. Precipitate the mixture at -80°C overnight. Centrifuge the product for 30 minutes at 4°C for14,000g. Aspirate the ethanol. Add 1 mL of 70% EtOH. Then centrifuge for 15 minutes at 4°C for14,000g. Again, remove the EtOH. Add 1 mL of 100% EtOH. Remove EtOH. Air dry for 20 minutes. Add 50 µL of de-ionized formamide onto the pellet. Resuspend at 37°C mixing at 10,000 RPM overnight or for at least 6 hours. Add 50 µL 20% Dextran Sulphate + 4xSSC. Then incubate for 10 minutes at 37°C mixing at 10,000 RPM. The mixture can be used right away or stored at -20°C.

Cell Culturing

Naïve hPSCs were cultured in t2iLGö media (Takashima et al., 2014;Watanabe et al., 2019) consisting of a 1:1 mixture of DMEM/F12 and Neurobasal, 0.5 % N2 supplement, 1 % B27 supplement, 1x nonessential amino acids, 2 mM L-Glutamine, 0.5x Penicillin/Streptomycin (all media components from ThermoFisher Scientific), 0.1 mM β-mercaptoethanol (Sigma-Aldrich) (N2B27 basal media) with 1 µM PD0325901 (Cell Guidance Systems), 1 µM CHIR99021 (Cell Guidance Systems), 20 ng/ml human LIF (Millipore) and 2 µM Gö6983 (Tocris) on irradiated MEFs. Passaging of naïve hPSCs occurred every 4 days using Accutase (ThermoFisher Scientific) at 37° C for 5 minutes. All cell culturing took place in 5% O2, 5% CO2 at 37°C with daily media changes.

Human fibroblast cells are cultured in hESC medium containing DMEM/F12 (Invitrogen) with the addition of 20% KSR (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM βmercaptoethanol (Invitrogen), and 10 ng/ml FGF2 (Peprotech)(Gu et al., 2016).

Sample preparation

After vigorous vortexing, 0.1 µm TetraSpeck™ microspheres are first resuspend in water at a ratio of 1:20. This dilution is then mixed into 100% EtOH at a ratio of 1:18. 6 µL of the microsphere-EtOH mixture is added to per well of the µ-Slide 8 Well Glass Bottom (Ibidi) chamber slide. Using the tip of a pipette, the mixture is gently spread across the surface of the well. Incubate at room temperature for 10 minutes to let the mixture dry. Geltrex LDEV-Free Reduced Growth Factor Basement Membrane Matrix is then added to cover the bottom of the well. Incubate at room temperature for 30 minutes. Aspirate the Geltrex off the chamber slide. Next, plate ~15,000 human fibroblasts or naive hPSCs per well for a sparse coating. Incubate at 37°C in 5% CO2 for 6 hours. Aspirate media and wash twice with Phosphate-Buffered Saline (PBS). Add 3 mL of microsphere-EtOH per well followed by 4% Paraformaldehyde (PFA) in PBS. Incubate for 10 minutes at room temperature to achieve cell fixation. Aspirate 4%PFA and wash PBS twice. Proceed to probe hybridization steps or store at 4°C.

Initial Hybridization

Begin by washing the sample PBS 3 times. Wash with PBS + 0.05% Tween (PBST). To permeabilize cells add PBS + 0.5% Triton X-100. Incubate with rocking for 20 minutes. Wash with PBS. Add PBS + 20% glycerol. Incubate sample for 1 hour at room temperature. Wash with PBST. Add 0.1 N HCl. Incubate for 5 minutes at room temperature. Wash twice with PBS. Wash 3 times with 2xSSC with 5 minutes of rocking. Add pre-warmed 2xSSC + 50% Formamide pH7. Prepare probes by denaturing in a 95°C heat block for 5 minutes. Spin down probes and place on ice. Denature for 3 minutes by placing the chamber slide on a 70°C heat block. Remove 2xSSC + 50% Formamide. Add probe. Hybridize for 48 hours at 37°C in a light-

blocking container with a water-soaked paper towel.

Washes and Imaging Preparation

For all following washes, cover the sample with aluminum foil or keep in a light-blocking container. Wash sample with 2xSSC, rocking for 15 minutes. Wash 3 times with pre-warmed 0.1xSSC at 60°C, rocking for 5 minutes. Wash 3 times with pre-warmed 4xSSC + 0.05% Tween at 42°C, rocking for 5 minutes. Wash once with 2xSSC. Wash twice with PBST. Stain with DAPI (0.5 mg\mL) in PBS at a 1:200 ratio, rocking for 6 minutes. Wash 3 times with 2xSSC, rocking for 10 minutes.

Microscopy

All imaging takes place using a motorized stage equipped Zeiss LSM 880 microscope. Zen Blue and Zen Black software was used for all image acquisition while Fiji and Adobe Photoshop were used for any image editing.

Probe Stripping and Rehybridization

Incubate sample in 2xSSC + 70% formamide for 30 minutes at 37°C. Denature probes at 95°C for 5 minutes. Spin down probes and place on ice. Heat chamber slide for 5 minutes on a 70°C heat block. Add probe. Hybridize for 48 hours at 37°C in a light-blocking container with a water-soaked paper towel. Repeat steps detailed in the “Washes and Imaging Preparation” section.

Relocation

Begin imaging by focusing on a corner of a well. Tile a 20x20 image in this location using the Plan-Apochromat 100X/1.40 Oil DIC (100X) objective with only the channel detecting DAPI staining. Move the focus towards the center of the well 2-3 more times, imaging and saving more 20x20 tiled sections. Proceed to image the FOVs of the round of sequential FISH hybridization. Upon returning to image a subsequent hybridization round, start from the same corner and consult the tiled sections to aid in returning to the FOVs.

Alignment

Consolidate all images of the same FOV. Using the first round of DNA FISH hybridization as a baseline, match the images to the first appearance of a nucleus. Using the baseline stack, find the slice where the last nuclei disappears from view. By subtracting the initial slice by the final slice the number of total slices per stack is calculated. Remove slices from all images starting from the first appearance of nuclei to the total number of slices calculated. This should make all stacks start and end roughly at the same points.

To align using the multi-channel microspheres, select only the channel with the detected light shortest wavelength, excluding the DAPI channel. Using the MultiStackReg function of Fiji, adjust the round of hybridization to the baseline stack and save the transformation file. Re-slice both the baseline and the round of hybridization. Again, use the MultiStackReg function and save the transformation file. Apply both transformation files to the other channels of the round of hybridization stack in MultiStackReg.

Chromatic Shift Calculation and Application

Using the baseline stack of images, split the channels. Extract the coordinates of the microspheres in each channel using the 3D Objects Counter. Subtract each channel's microsphere from the microsphere in the channel with the detected light shortest wavelength. This yields a chromatic shift from each channel in relation to that of the channel with the detected light shortest wavelength. This chromatic shift can then be applied to probe coordinates from probe hybridization rounds specifically based on the channel the probe signal originates from.

Microsphere Realignment

After a stack of images of a round of hybridization has been aligned to the baseline the microsphere locations must be obtained. Using the channel with the detected light shortest wavelength, excluding the DAPI channel, extract the location of the microspheres using 3D Objects Counter. Once the microsphere locations have been found, subtract the round of hybridization microsphere coordinates from the baseline microsphere coordinates, yielding the shift of the image.

Figures

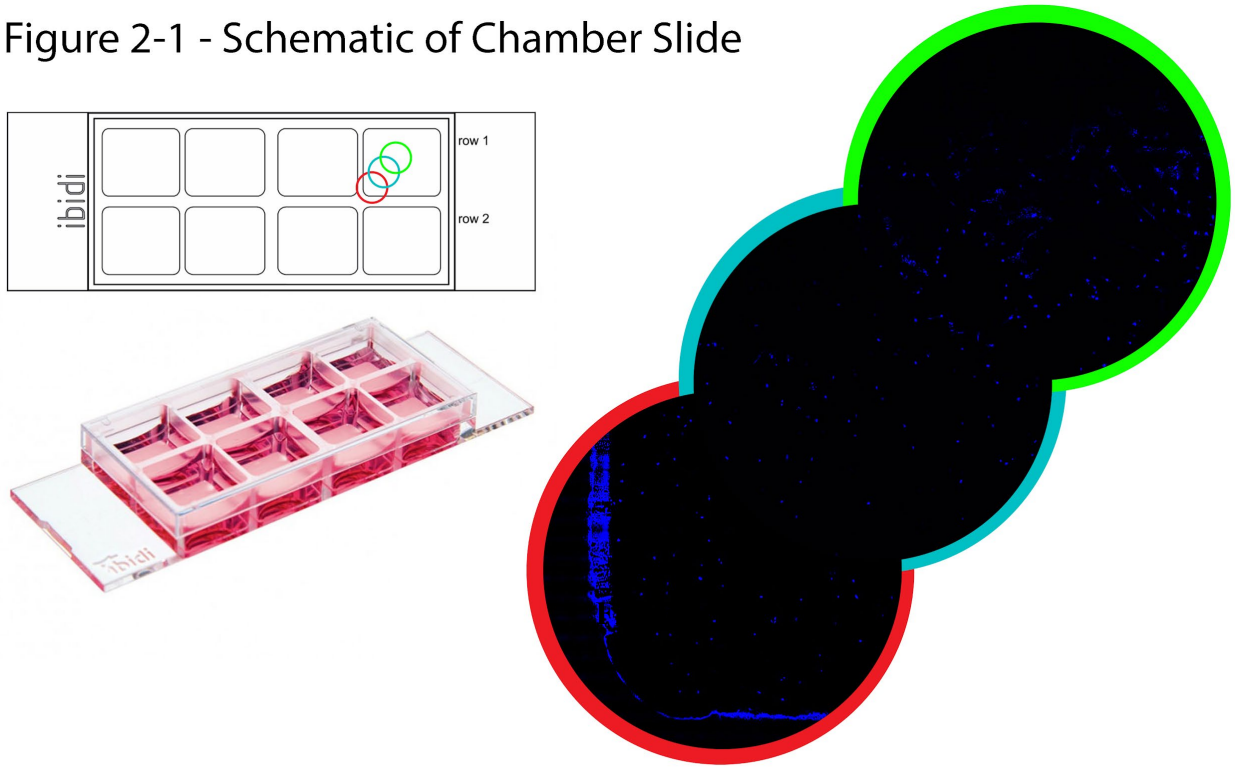


Figure 2-1 - Schematic of Chamber Slide

(A) µ-Slide 8 Well Glass Bottom (Ibidi) showcases multiple 1.0 cm² wells that feature physical walls. Colored circles indicate subsequent tiled images. (B) DAPI stained nuclei are visible through the glass bottom slide and their positions can be located relative to the walls of the wells. By overlapping three subsequent images a path to the center of the well is visualized.

Figure 2-2 - FISH Illustrates Separate Chromosomal Regions and Non-Active X Chromosome

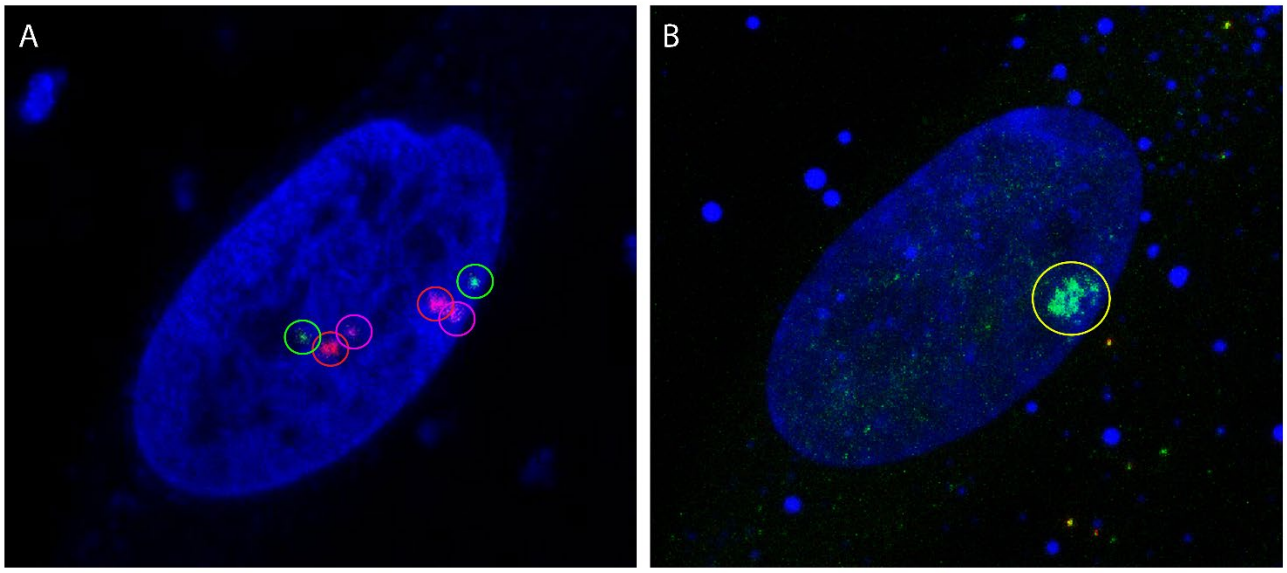


Figure 2-2 - FISH Illustrates Separate Chromosomal Regions and Non-Active X Chromosome

(A) TSIX, HUWE1, and THOC2 are targeted by DNA FISH probes containing Atto 488, Cy3, and Cy5 fluorophores respectively. All regions clearly separate within the DAPI stained nuclei for both X chromosomes of the fibroblast cell. (B) XIST RNA FISH using Atto 488 shows a cloud over the non-active X chromosome

Figure 2-3 - Probe Stripping and Rehybridization Replaces DNA FISH Signals

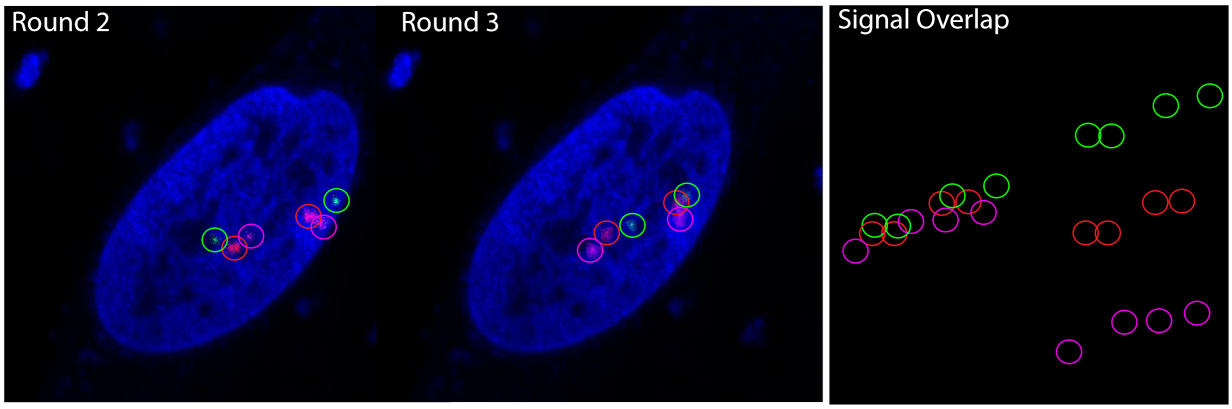
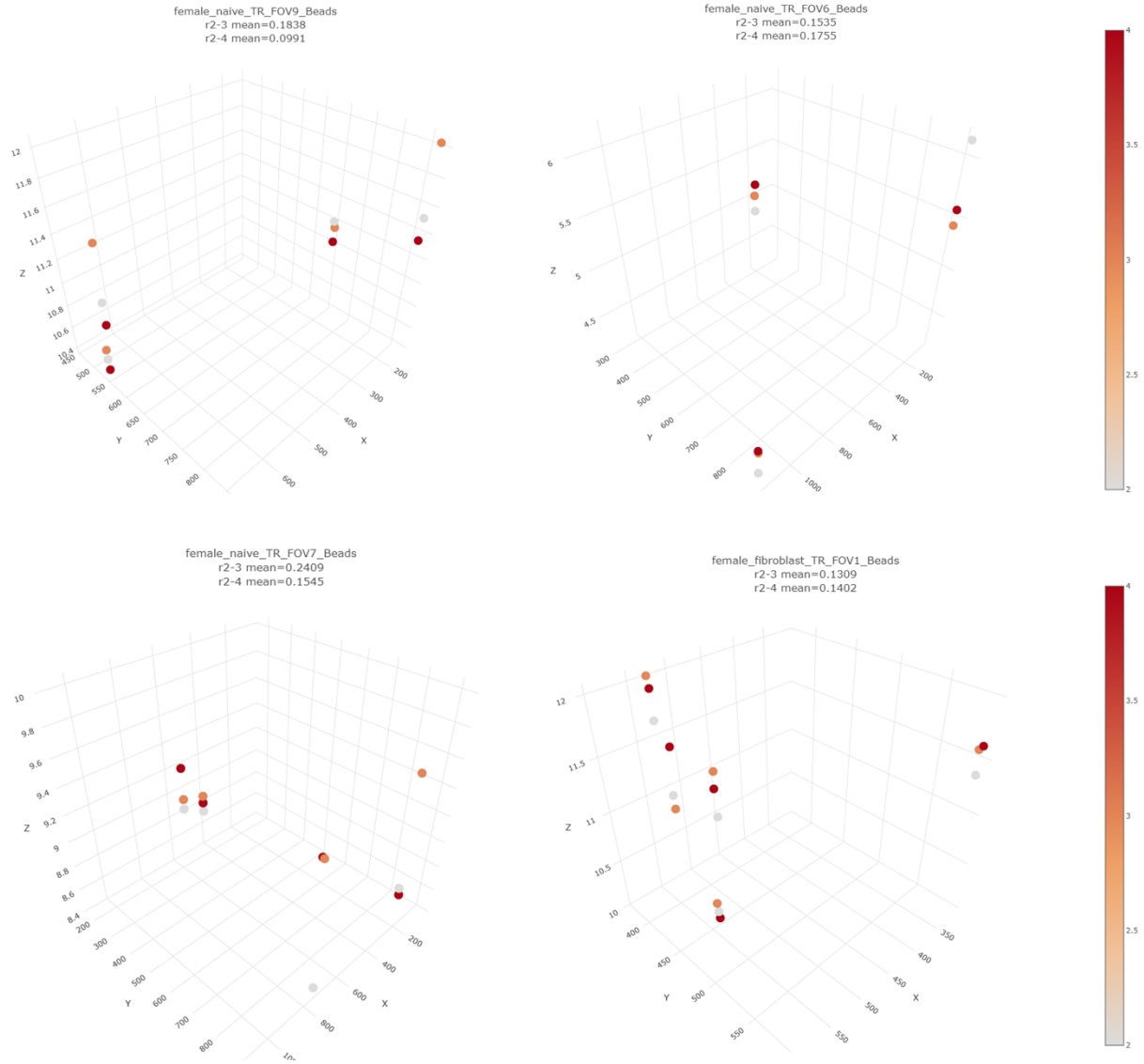


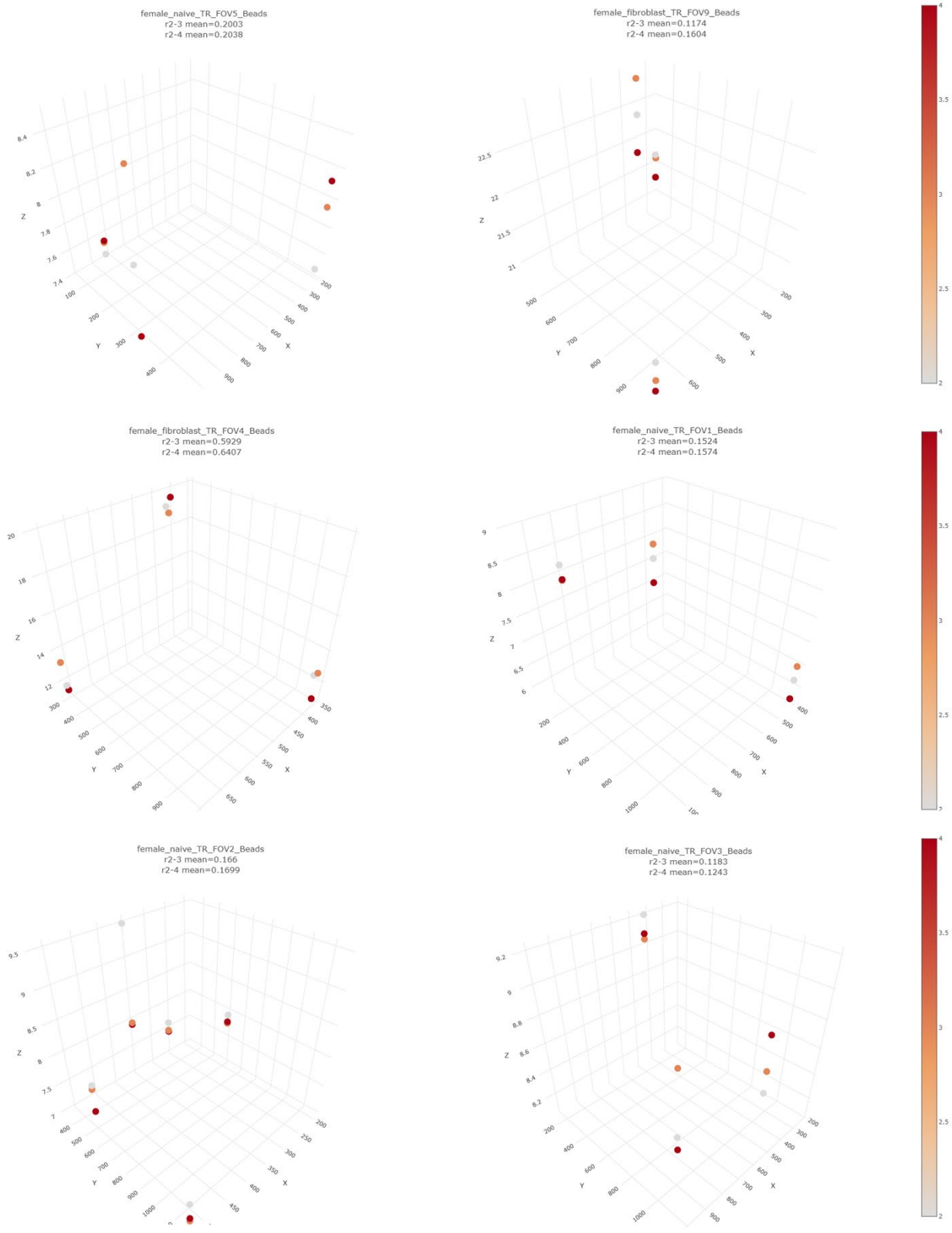
Figure 2-3 - Probe Stripping and Rehybridization Replaces DNA FISH Signals

TSIX, HUWE1, and THOC2 are targeted by DNA FISH probes in Round 2 of hybridization. XACT, GPC3, and UTX are targeted by DNA FISH probes in Round 3 of hybridization. Overlap of Rounds 2 and 3 of hybridization showing successful stripping of probes between hybridizations.

Figure 2-4 - Multi-channel Microspheres Aid in Image Alignment

Bead Distribution per FOV after Reslicing Alignment




female_naive_TR_FOV5_Beads
r2-3 mean=0.2003
r2-4 mean=0.2038
female_fibroblast_TR_FOV9_Beads
r2-3 mean=0.1174
r2-4 mean=0.1604
female_fibroblast_TR_FOV4_Beads
r2-3 mean=0.5929
r2-4 mean=0.6407
female_naive_TR_FOV1_Beads
r2-3 mean=0.1524
r2-4 mean=0.1574
female_naive_TR_FOV2_Beads
r2-3 mean=0.166
r2-4 mean=0.1699
female_naive_TR_FOV3_Beads
r2-3 mean=0.1183
r2-4 mean=0.1243

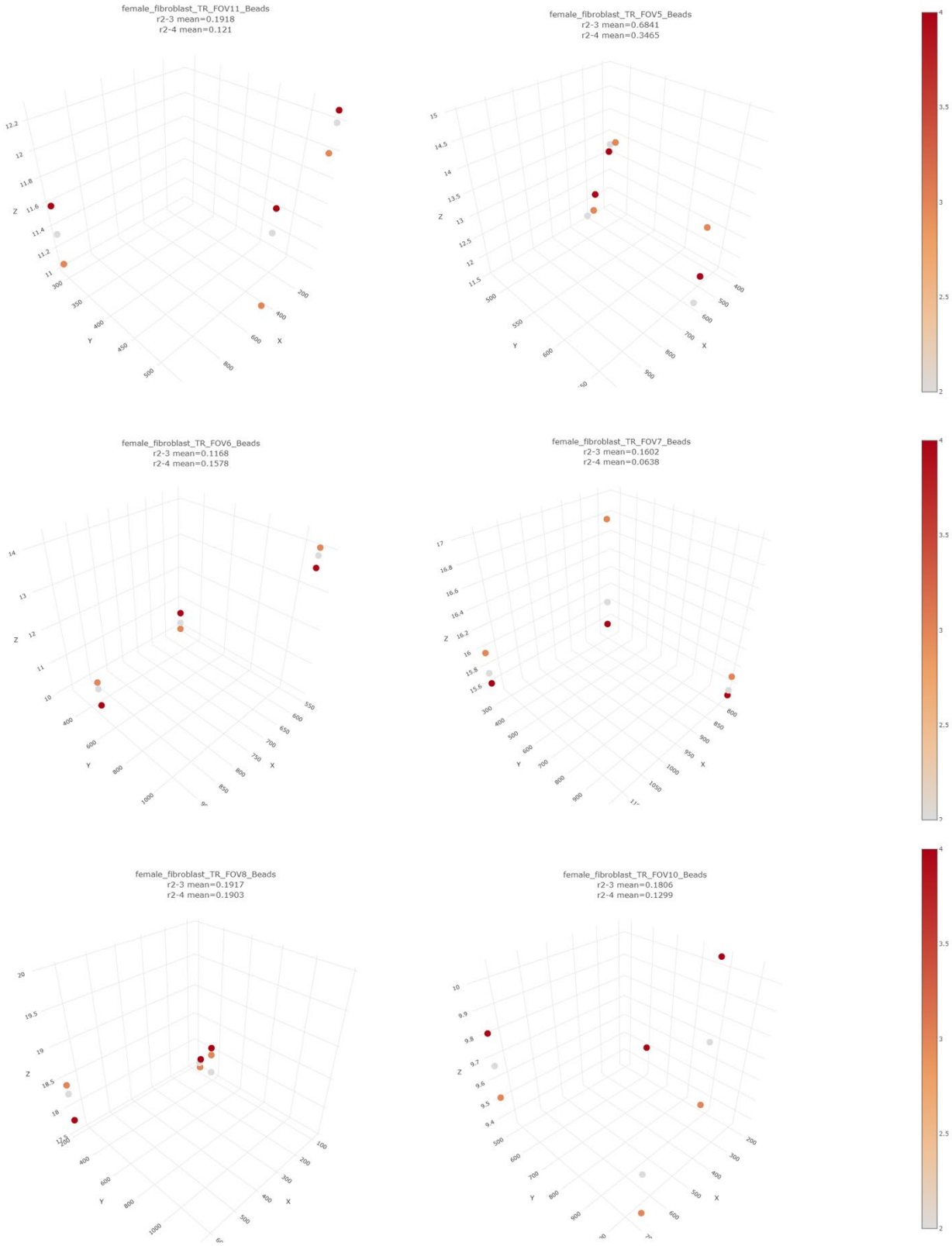

female_fibroblast_TR_FOV11_Beads
r2-3 mean=0.1918
r2-4 mean=0.121
female_fibroblast_TR_FOV5_Beads
r2-3 mean=0.6841
r2-4 mean=0.3465
female_fibroblast_TR_FOV6_Beads
r2-3 mean=0.1168
r2-4 mean=0.1578
female_fibroblast_TR_FOV7_Beads
r2-3 mean=0.1602
r2-4 mean=0.0638
female_fibroblast_TR_FOV8_Beads
r2-3 mean=0.1917
r2-4 mean=0.1903
female_fibroblast_TR_FOV10_Beads
r2-3 mean=0.1806
r2-4 mean=0.1299

Figure 2-4 - Multi-channel Microspheres Aid in Image Alignment

Individual FOVs from both naïve and fibroblast cells illustrating the average shift in multi-channel microshperes by each round when compared to the initial round of DNA FISH hybridization. Colors represent the round of hybridization the microspheres are imaged from.

Figure 2-5 - Multi-Channel Microspheres Adjust for Chromatic Aberration

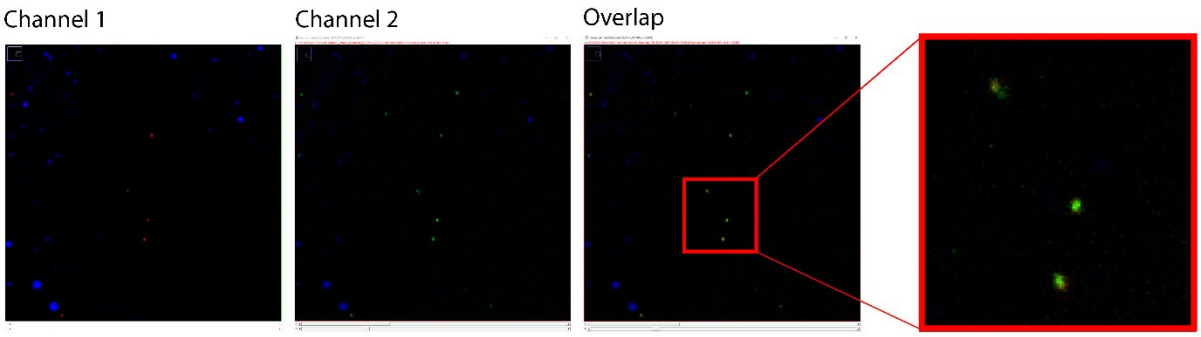


Figure 2-5 - Multi-channel Microspheres Adjust for Chromatic Aberration

TetraSpeck multi-channel microspheres (Invitrogen) appear in each channel of an image. The microspheres appear as red dots in Channel 1 and green dots in Channel 2. Overlap of images with a zoomed image illustrating the other edges of the red channel extending beyond the green channel indicating a small case of misalignment in light detection.

Figure 2-6 - Fiji Reslicing Enables Image Alignment

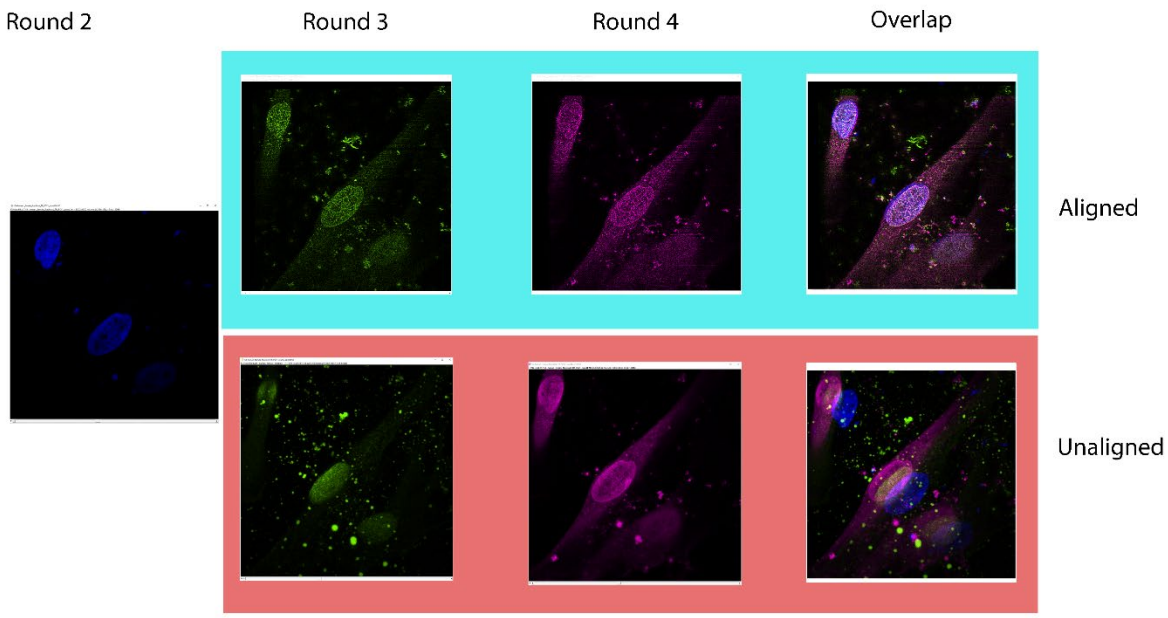


Figure 2-6 - Fiji Reslicing Boosts Image Alignment

Differently colored DAPI channels of rounds 2-4 of a FOV of sequential FISH experiment in fibroblast cells. The overlap of the aligned image demonstrates the alignment ability using the multi-channel microspheres versus the poor overlap of the unaligned DAPI channels.

References

Bickmore, W., & Van Steensel, B. (2013). Genome Architecture: Domain Organization of Interphase Chromosomes. *Cell, 152*(6), 1270-1284. doi:10.1016/j.cell.2013.02.001

Bonora, G., Deng, X., Fang, H., Ramani, V., Qiu, R., Berletch, J. B., . . . Disteche, C. M. (2018). Orientation-dependent Dxz4 contacts shape the 3D structure of the inactive X chromosome. *Nature Communications, 9*(1). doi:10.1038/s41467-018-03694-y

Carrel, L., Cottle, A. A., Goglin, K. C., & Willard, H. F. (1999). A first-generation X-inactivation profile of the human X chromosome. *Proceedings of the National Academy of Sciences, 96*(25), 14440-14444. doi:10.1073/pnas.96.25.14440

Cremer, T., & Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Reviews Genetics*, *2*(4), 292–301. doi: 10.1038/35066075

Disteche, C. M., & Berletch, J. B. (2015). X-chromosome inactivation and escape. *Journal of Genetics, 94*(4), 591-599. doi:10.1007/s12041-015-0574-1

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., . . . Ren, B. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature, 518*(7539), 331-336. doi:10.1038/nature14222

Fluorescence in situ hybridization. (2005). *Nature Methods, 2*(3), 237-238. doi:10.1038/nmeth0305-237

Gu, W., Gaeta, X., Sahakyan, A., Chan, A., Hong, C., Kim, R., . . . Christofk, H. (2016). Glycolytic Metabolism Plays a Functional Role in Regulating Human Pluripotent Stem Cell State. *Cell Stem Cell, 19*(4), 476-490. doi:10.1016/j.stem.2016.08.008

He, Z., Gentry, T. J., Schadt, C. W., Wu, L., Liebich, J., Chong, S. C., . . . Zhou, J. (2007). GeoChip: A comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *The ISME Journal, 1*(1), 67-77. doi:10.1038/ismej.2007.2

Lubeck, E., & Cai, L. (2012). Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nature Methods, 9*(7), 743-748. doi:10.1038/nmeth.2069

Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M., & Cai, L. (2014). Single-cell in situ RNA profiling by sequential hybridization. *Nature Methods, 11*(4), 360-361. doi:10.1038/nmeth.2892

Manz, W., Amann, R., Ludwig, W., Wagner, M., & Schleifer, K. (1992). Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. *Systematic and Applied Microbiology, 15*(4), 593-600. doi:10.1016/s0723-2020(11)80121-9

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T. W., . . . Plath, K. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell, 20*(1), 87-101. doi:10.1016/j.stem.2016.10.006

Theunissen, T., Powell, B., Wang, H., Mitalipova, M., Faddah, D., Reddy, J., . . . Jaenisch, R. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell, 15*(4), 471-487. doi:10.1016/j.stem.2014.07.002

Theunissen, T., Friedli, M., He, Y., Planet, E., O'Neil, R., Markoulaki, S., . . . Jaenisch, R. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell, 19*(4), 502-515. doi:10.1016/j.stem.2016.06.011

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Ali, T. M. L., … Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell*, *20*(1), 102–111. doi: 10.1016/j.stem.2016.10.014

Wang, S., Su, J., Beliveau, B. J., Bintu, B., Moffitt, J. R., Wu, C., & Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. *Science, 353*(6299), 598-602. doi:10.1126/science.aaf8084

Watanabe, M., Haney, J. R., Vishlaghi, N., Turcios, F., Buth, J. E., Gu, W., . . . Novitch, B. G. (2019). TGFβ superfamily signaling regulates the state of human stem cell pluripotency and competency to create telencephalic organoids. doi:10.1101/2019.12.13.875773

Yang, F., Babak, T., Shendure, J., & Disteche, C. M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Research, 20*(5), 614-622. doi:10.1101/gr.103200.109

Yilmaz, L. S., Loy, A., Wright, E. S., Wagner, M., & Noguera, D. R. (2012). Modeling Formamide Denaturation of Probe-Target Hybrids for Improved Microarray Probe Design in Microbial Diagnostics. *PLoS ONE, 7*(8). doi:10.1371/journal.pone.0043862

Chapter 3

Sequential FISH describes X Chromosome Structure at a Single Cell Level

Abstract

By utilizing the sequential FISH protocol we developed, we extracted the 3D spatial coordinates of nine genomic regions across the X chromosome in two different cell types, human fibroblast and naive hPSCs. We have created a probe scheme specifically for the human X chromosome that spans across the length of the chromosome and selects genes that are inactivated during XCI and escape dampening in XCD in order to both capture the overall structure of the chromosome as well as understand specific changes in structure related to different gene expression states. XIST RNA FISH highlights XIST clouds coating inactive or dampened chromosomes in fibroblasts and naïve hPSCs, respectively. By having knowledge of the identity of each X chromosome, analysis concerning the gene expression state can be factored into the understanding of the acquired chromosome structure. From the 3D spatial coordinates of our nice different genomic locations, we calculated probe to probe distances in an intra-chromosomal context in order to better understand the relationship of the probe locations within the overall structure. We also sought to validate that magnitude of intrachromosomal distances acquired by this method are within logical reason. To do this, we measured the maximum distance across each cell's XIST cloud to serve as a baseline for the intrachromosomal distances measured from the sequential FISH method. Finally, we sought to determine if any hybridization round biases, wherein intrachromosomal distances from one hybridization round gravitate towards a single, inaccurately obtained value, are present within the data. Here, we collect 3D spatial data of X chromosomes relating to gene expression state and attempt to validate the legitimacy of the data.

Introduction

The three X chromosome activation states, inactive, active, and dampened, all display different gene expression profiles (Jin et al., 2013). Because gene expression is intertwined with chromosome structure, it is logical to presume that the dampened X chromosome, which displays a reduction in gene expression but not silencing, will have a novel structure that is unsimilar to the structures of the inactive or active X chromosomes (Galupa and Heard 2018; Petropoulos et al., 2016; Sahakyan et al., 2017). To determine the structure of the dampened X as well as to ascertain how the structure of the dampened X relates, be it by overall size or unique features, to that of the inactive and active X structures, we developed a sequential FISH protocol that allows for the acquisition of 3D spatial locations of regions of the chromosome on a single cell resolution.

Choosing informative locations on the chromosome is a critical component of designing a sequential FISH experiment. One must consider genes of for detection, spacing of probes across the chromosome, and how many probes and therefore how many rounds of hybridizations must be performed. The experimental design must then be compatible with the limitations of any protocol needed to acquire such data. For example, the sequential FISH technique must be free of any round to round biases in order to faithfully recapitulate the chromosome structure.

When seeking to interrogate the whole chromosomal structure of the X chromosome, it becomes a necessity to have probes target the across the entire chromosome body in fairly even spacing for the highest efficiency assay. This ensures that the overall shape of the chromosome is captured and may also provide near max range of intrachromosomal distances which can prove useful to understanding relative spatial organization based on probe distance.

Overall chromosome structure, while critical to understanding the dampened X

chromosome in comparison to the inactive and active X chromosomes, is not the end-all concerning interrogation of the structure of the X chromosome. To find key details, specific genes of consequence must also be a component of the sequential FISH probe scheme. When highlighting genes of importance within the X chromosome, features relevant to gene expression. Critical to X chromosome gene expression is inactivation and dampening. In XCI, most genes become silenced except for a small group of XCI escapers (Carrel and Brown 2017; Disteche and Berletch et al., 2015; Patrat et al., 2020). However, gene expression is merely reduced, aside from the XCD escapers, in the case of XCD (Petropoulos et al., 2016; Sahakyan et al., 2017). By targeting genes that play a part in XCI or XCD, there is a greater opportunity to learn how each of these chromosomes of differing gene expression states are spatially organized.

Results

A total of 16 fibroblast cells and 20 naive cells have been acquired using the sequential FISH method we have developed (Fig 3-1). In sum, 72 individual X chromosomes are found in this dataset. This dataset allows for comparison across cell types as well as comparison of the complementary X's within each cell.

The sequential FISH probe schematic encompasses nine genes: CDKL5, UTX, PAGE1, HUWE1, TSIX, XACT, LAMP2, THOC2, and GPC3 (Fig 3-2). The collection of genes gives an average spacing of 14,301kb and from head to tail covers approximately 73% of the X chromosome. CDKL5 is translated to form a kinase and is believed to play a role in neuronal morphology (Kilstrup-Nielsen et al., 2012). The UTX gene, also known as KDM6A, deals with the demethylation of H3K27, essential for embryonic development, differentiation, and proper reprogramming (Lee et al., 2007; Meulen et al., 2014). PAGE1 is known to be a prostate

associated protein coding gene (Brinkmann et al., 1998). Another protein coding gene, HUWE1, produces an enzyme involved in ubiquitination that regulates homeostasis of cellular development (Kao et al., 2018). TSIX transcribes a lncRNA antisense to the Xist gene, and is known to suppress XIST function in regards to modulation of X chromosome gene expression state (Sado et al., 2005; Bemmel et al., 2019). Similarly, XACT also produces noncoding RNA that regulates XIST in the context of development, notably in the naive cell type (Vallot et al., 2013; Vallot et al., 2017). LAMP2 is known to regulate autophagy (Jalali and Parvaz 2020). Relevant to gene regulation, THOC2 plays a role in mRNA export (Chi et al., 2012; Kumar et al., 2020). GPC3 plays a role in Wnt signaling, a pathway integral to embryonic development (Stigliano et al., 2008). This collection of genes displays a wide berth of functions, many of which are important for gene regulation and development.

Genes that are targeted for silencing during XCI include: CDKL5, HUWE1, LAMP2, THOC2, and GPC3 (Fig 3-2). The XCI genes that undergo inactivation have an average spacing of 28,602 kb and from head to tail covers approximately 73% of the X chromosome.

UTX, HUWE1, XACT, and THOC2 are all genes that escape XCD (Fig 3-2). 50% of the X chromosome is covered with an average spacing of 25,983 kb. Of note, HUWE1 and THOC2, are shared between the XCD escapers and the XCI inactive genes present in the probe scheme.

By extracting the probe intensities of each individual cell, it is possible to view both complementary chromosomes (either the Xi and Xa or the Xd and Xa) in 3D space (Fig 3-3). From a basic glance, there is an obvious trend that the inactive X chromosomes appear to be, for the most part, smaller than the complementary active X chromosomes in the fibroblast cells. However, there is noticeable cell to cell variation, which is only detectable by the single cell technique. By overlapping the signals from each round into one z-projection of the cell it is

possible to view how the probes spread the breadth of each chromosome but also how the density of the fluorescent signals would pose a challenge to resolve if all hybridized simultaneously (Fig 3-3). In viewing the extracted signals compared to the z-projection, one can appreciate just how much space the chromosomes occupy and how probe placement already appears different in the complementary chromosomes.

In both the naive and fibroblast cells the presence of XIST clearly marks the dampened or inactive X chromosome respectively (Fig 3-4). While the XIST cloud is easy to spot in both cell types, the naïve culture may not display an XIST cloud in every cell. The naïve cell FOV is representative of the total culture, not displaying an XIST cloud in all cells (Fig 3-4). In these cases, only the cells in which an XIST cloud is present is counted. This ensures that the dampened X chromosome gene expression state it accurately captured rather than a scenario where there are two active X chromosomes.

To understand the visual, spatial structure of the chromosomes we recreated their image in a single cell context (Fig 3-5). In this way we could view the complementary X chromosomes as well as asses if there are any instances of interest in structure in either same FOV or cell type groups. Present in most of the fibroblast cells is a probe or two, at minimum, found in the active X that separates from the main body of probes. Obvious examples of this include FOV1 Cell3 with UTX and FOV5 Cell1 with THOC2 separating from the main body (Fig 3-5). This phenomenon is for the most part not found in the inactive X chromosomes with totality of probes being found in very close spatial proximity to one another. In naïve, this is even more apparent in the active X chromosomes, with the overall structure appearing both large and plastic. The inactive X chromosomes found in fibroblasts were smaller than their complementary active X chromosomes. In contrast, the dampened X chromosomes in naïve appear to be a similar size to

and mirror the openness of the complementary active X chromosomes such as in FOV5 and FOV6 (Fig 3-5). Additionally, the XACT probe appears much further away in comparison to the main body of probes in some cases of the dampened X chromosome, examples include: FOV7 Cell3 and FOV5 Cell1 (Fig 3-5).

In order to assess each cell in a more quantitative fashion, we calculated the 3D distance between each of the probes in a chromosome. Coloring by magnitude of these measurements, these distances can then be compiled into a spatial distance map showing a comprehensive view of all available probe to probe connections within a chromosome (Fig 3-6). While the magnitude of the probe distances appears more muted in the fibroblast spatial distance maps due to the larger size of the naïve chromosomes (all distances are on the same scale), the fibroblast cells nonetheless appear to show larger distances in a more linear chromosome fashion. For example, in nearly all the distances from the first three genes: CDKL5, UTX, and PAGE1 to the last three genes: LAMP2, THOC2 and GPC3 are in the 2-3 μm range in active X chromosomes (Fig 3-6). Whereas the inactive X chromosomes only have sporadic larger distances scattered throughout the chromosome. In naïve cells, the active X seems to harbor larger intra-chromosomal distances in many cases across the chromosome but does retain the same feature of the fibroblast active X in regards to the long-range interactions. We believe that the large degree of openness found in the Xa of naïve cells is caused due to the stage of development, wherein the pluripotent nature of the cell gives rise to a different organization of compartments as well as local chromatin interactions (up to 200 kb regions) in comparison to the XA of fibroblast cells (Schmitt et al., 2016). As seen in the 3D spatial data in naïve cells, the dampened X chromosomes range from slightly smaller to comparable in size and openness to their complementary active X chromosomes, with a small population of dampened X chromosomes displaying the XACT

probe more distal to the main body of probes than even the complementary active X chromosomes (Fig 3-6).

Using the measurement tool in Fiji the longest vectors across the XIST cloud were measured in order to produce a baseline for intra-chromosomal distances (Fig 3-7a). The concept being that the XIST cloud should cover roughly the same area as the actual chromosome. We found that the sizes of the XIST clouds of the dampened X chromosomes of the naive cells clearly appeared to be larger in comparison to the XIST cloud sizes of the inactive X chromosomes of the fibroblast cells (Fig 3-7b). This is not surprising considering the overall sizes of the naïve X chromosomes but is a respectable sanity check (Fig 3-6). Another measurement taken in Fiji is nuclear volume. We sought to understand how the nucleus size relates to the size of the accompanying X chromosomes. By dividing the longest intrachromosomal distances by the nuclear volume, we can generate a ratio to explain the relative amount of space the chromosome occupies (Fig 3-7c). Across the board, fibroblast X chromosomes, while smaller relative to the naïve cell X chromosomes, take up a proportionally larger space within the nucleus (Fig 3-7c). This may be due, in part, to the clear difference in 3D topology of the cells, with the fibroblast cells being noticeably flatter than the naïve cells.

Now that we have acquired quantitative data concerning the location of the probes, we find it to be imperative to inspect the data for any distance biases due to hybridization round irregularities. By clearing this hurdle, we can further legitimize the sequential FISH method as well as have confidence in the structural observations we have made. By accumulating all probe to probe distances and separating them by hybridization round, we have determined that hybridization round based biases are not present within the dataset (Fig 3-8a). If we instead look only at the mean distances of the probe distances separated by hybridization round, we find clear

separation of values between each X species. Potentially the only area of concern may be the final round of hybridization in naïve with 2 out of 3 distances being on average quite close (Fig 3-8b). However, this may just be a shared feature of the X chromosomes (the Xd and Xa) in naïve cells. If we instead compare all probe distances that occur only within the same hybridization rounds, we find that there are a range of probe distances rather than a concentration all at one specific distance (3-8c).

To further investigate this potential issue, we created violin plots for each individual probe distance as well as each probe difference split by cell type and found no obvious distance biases (Fig 3-9a-b).

Discussion

The probe scheme we have designed manages to capture positions spanning a majority of the X chromosome while still limiting the number of rounds of probe hybridization to only four, including a round of RNA FISH solely for certainty in calling the gene expression state of each of the chromosomes. Within the probe scheme are sets of probes that are targeted for inactivation during XCI and well as genes that escape during XCD. This probe scheme has provided data from the sequential FISH technique that not only represents the 3D spatial coordinates of nine genes across the chromosome but can further be broken down into informative groups pertaining to structural rearrangements of the X chromosome.

An initial hybridization of XIST RNA probe has shown to be effective in identifying the X chromosome expressing XIST, either the inactive or dampened X chromosome in fibroblast or naive cell type respectively. This in turn identifies the complementary active X chromosome. This simple but effective step ensures the proper calling of chromosomes for downstream analysis. The XIST cloud itself is a clear marker, spanning multiple microns at its largest length. This measurement then acts as a sort of guideline for how large intrachromosomal distances are expected to be since the XIST cloud thoroughly covers the inactive X chromosome.

While the XIST RNA probe hybridization effectively assesses the identity of each X chromosome in each cell, the 3D spatial representation of the chromosomes can aid in assessment of the overall structure of the chromosome. An obvious visual observation is the difference in size of the two chromosomes, more apparent in the fibroblast cells. We also observed many cell-to-cell variations in individual chromosome structure. Spatial distances from probe to probe give a direct quantification of the structure of the chromosome. We found certain probe distances to appear more regularly, such as long-range genes showcasing longer distances

in the active X chromosomes and XACT appearing to be more spatial distant in some of the dampened X chromosomes.

We arranged all probe distances by round of hybridization in order to obtain a view of any biases or irregular trends present within the data. Based on the lack of a high density of distances being in a small range of each other, we can confidently say that the technique does not cause a uniform distancing effect. This is critical to ensuring that the sequential FISH technique we have developed can faithfully recapitulate the shape of the X chromosome. Based on this data, we can say that for the first time the structure of the Xd was uncovered.

Materials & Methods

Cell Culturing

Naïve hPSCs were cultured in t2iLGö media (Takashima et al., 2014;Watanabe et al., 2019) consisting of a 1:1 mixture of DMEM/F12 and Neurobasal, 0.5 % N2 supplement, 1 % B27 supplement, 1x nonessential amino acids, 2 mM L-Glutamine, 0.5x Penicillin/Streptomycin (all media components from ThermoFisher Scientific), 0.1 mM β-mercaptoethanol (Sigma-Aldrich) (N2B27 basal media) with 1 μM PD0325901 (Cell Guidance Systems), 1 μM CHIR99021 (Cell Guidance Systems), 20 ng/ml human LIF (Millipore) and 2 μM Gö6983 (Tocris) on irradiated MEFs. Passaging of naïve hPSCs occurred every 4 days using Accutase (ThermoFisher Scientific) at 37° C for 5 minutes. All cell culturing took place in 5% O2, 5% CO2 at 37°C with daily media changes.

Human fibroblast cells are cultured in hESC medium containing DMEM/F12 (Invitrogen) with the addition of 20% KSR (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM βmercaptoethanol (Invitrogen), and 10 ng/ml FGF2 (Peprotech)(Gu et al., 2016).

Sample preparation

After vigorous vortexing, 0.1 μm TetraSpeck™ microspheres are first resuspend in water at a ratio of 1:20. This dilution is then mixed into 100% EtOH at a ratio of 1:18. 6 μL of the microsphere-EtOH mixture is added to per well of the μ-Slide 8 Well Glass Bottom (Ibidi) chamber slide. Using the tip of a pipette, the mixture is gently spread across the surface of the well. Incubate at room temperature for 10 minutes to let the mixture dry. Geltrex LDEV-Free Reduced Growth Factor Basement Membrane Matrix is then added to cover the bottom of the

well. Incubate at room temperature for 30 minutes. Aspirate the Geltrex off the chamber slide. Next, plate ~15,000 human fibroblasts or naive hPSCs per well for a sparse coating. Incubate at 37°C in 5% CO2 for 6 hours. Aspirate media and wash twice with Phosphate-Buffered Saline (PBS). Add 3 mL of microsphere-EtOH per well followed by 4% Paraformaldehyde (PFA) in PBS. Incubate for 10 minutes at room temperature to achieve cell fixation. Aspirate 4%PFA and wash PBS twice. Proceed to probe hybridization steps or store at 4°C.

Probe Design and Synthesis

We generated the fluorescently labeled probes from a BAC library containing genes found across the human X chromosome. The nick translation mixture is composed of 1 µg of the BAC of choice, 5 µL of NT Buffer, 5 µL of 0.1M β-Me, 5 µL of dNTP Mix 1mM, 3 µL of fluorescently labeled dUTP, 2 µL of DNase (2000U/mL) at 1:100 dilution in cold water, 1 µL of Pol I (10U/µL), and the remainder of the 50 µL volume as water. Using a thermocycler, incubate the mixture at 15°C for 12 hours. Then stop the reaction by heating the mixture to 65°C for 10 minutes.

Purification of labelled DNA can be accomplished using AMPure beads. After adding 120 µL of DNA binding AMPure beads to the reaction mixture, pipette to gently resuspend the new mixture. Incubate at room temperature for 10 minutes under foil or a light-blocking container. Transfer the mixture to a magnetic tube rack and again incubate at room temperature for 10 minutes under foil or a light-blocking container. While the tube containing the mixture remains on the magnetic rack, remove the supernatant and add 400 µL of freshly prepared 80% ethanol (EtOH). Proceed to incubate at room temperature for 10 minutes under foil or a light-blocking container. Again, remove the supernatant while the tubs are still attached to the

magnetic rack and repeat the 80% EtOH wash. Add 100 µL of water to just the AMPure beads, resuspend, and incubate the tube separate from the magnetic rack for 10 minutes. Replace the tube onto the magnetic rack and once again incubate for 10 minutes. Retrieve the purified product.

Combine the purified product with 10 µL of COT-1 DNA, 10 µL of Salmon Sperm DNA, 340 µL of 100% EtOH, and 15 µL of 3M Sodium Acetate pH 5.2. Precipitate the mixture at -80°C overnight. Centrifuge the product for 30 minutes at 4°C for14,000g. Aspirate the ethanol. Add 1 mL of 70% EtOH. Then centrifuge for 15 minutes at 4°C for14,000g. Again, remove the EtOH. Add 1 mL of 100% EtOH. Remove EtOH. Air dry for 20 minutes. Add 50 µL of de-ionized formamide onto the pellet. Resuspend at 37°C mixing at 10,000 RPM overnight or for at least 6 hours. Add 50 µL 20% Dextran Sulphate + 4xSSC. Then incubate for 10 minutes at 37°C mixing at 10,000 RPM. The mixture can be used right away or stored at -20°C.

XIST RNA FISH

XIST RNA FISH probes were made using the same FISH protocol except with exclusively Atto 488 fluorophores (Sigma-Aldrich). The XIST FISH takes place before any of the DNA FISH and before any RNAse A or RNAse H is added to the sample.

Initial Hybridization

Begin by washing the sample PBS 3 times. Wash with PBS + 0.05% Tween (PBST). To permeabilize cells add PBS + 0.5% Triton X-100. Incubate with rocking for 20 minutes. Wash with PBS. Add PBS + 20% glycerol. Incubate sample for 1 hour at room temperature. Wash with PBST. Add 0.1 N HCl. Incubate for 5 minutes at room temperature. Wash twice with PBS.

Wash 3 times with 2xSSC with 5 minutes of rocking. Add pre-warmed 2xSSC + 50% Formamide pH7. Prepare probes by denaturing in a 95°C heat block for 5 minutes. Spin down probes and place on ice. Denature for 3 minutes by placing the chamber slide on a 70°C heat block. Remove 2xSSC + 50% Formamide. Add probe. Hybridize for 48 hours at 37°C in a light-blocking container with a water-soaked paper towel.

Washes and Imaging Preparation

For all following washes, cover the sample with aluminum foil or keep in a light-blocking container. Wash sample with 2xSSC, rocking for 15 minutes. Wash 3 times with pre-warmed 0.1xSSC at 60°C, rocking for 5 minutes. Wash 3 times with pre-warmed 4xSSC + 0.05% Tween at 42°C, rocking for 5 minutes. Wash once with 2xSSC. Wash twice with PBST. Stain with DAPI (0.5 mg\mL) in PBS at a 1:200 ratio, rocking for 6 minutes. Wash 3 times with 2xSSC, rocking for 10 minutes.

Microscopy

All imaging takes place using a motorized stage equipped Zeiss LSM 880 microscope. Zen Blue and Zen Black software was used for all image acquisition while Fiji and Adobe Photoshop were used for any image editing.

Probe Stripping and Rehybridization

Incubate sample in 2xSSC + 70% formamide for 30 minutes at 37°C. Denature probes at

95°C for 5 minutes. Spin down probes and place on ice. Heat chamber slide for 5 minutes on a 70°C heat block. Add probe. Hybridize for 48 hours at 37°C in a light-blocking container with a water-soaked paper towel. Repeat steps detailed in the "Washes and Imaging Preparation" section.

Relocation

Begin imaging by focusing on a corner of a well. Tile a 20x20 image in this location using the Plan-Apochromat 100X/1.40 Oil DIC (100X) objective with only the channel detecting DAPI staining. Move the focus towards the center of the well 2-3 more times, imaging and saving more 20x20 tiled sections. Proceed to image the FOVs of the round of sequential FISH hybridization. Upon returning to image a subsequent hybridization round, start from the same corner and consult the tiled sections to aid in returning to the FOVs.

Alignment

Consolidate all images of the same FOV. Using the first round of DNA FISH hybridization as a baseline, match the images to the first appearance of a nucleus. Using the baseline stack, find the slice where the last nuclei disappears from view. By subtracting the initial slice by the final slice the number of total slices per stack is calculated. Remove slices from all images starting from the first appearance of nuclei to the total number of slices calculated. This should make all stacks start and end roughly at the same points.

To align using the multi-channel microspheres, select only the channel with the detected light shortest wavelength, excluding the DAPI channel. Using the MultiStackReg function of Fiji, adjust the round of hybridization to the baseline stack and save the transformation file. Re-

slice both the baseline and the round of hybridization. Again, use the MultiStackReg function and save the transformation file. Apply both transformation files to the other channels of the round of hybridization stack in MultiStackReg.

Chromatic Shift Calculation and Application

Using the baseline stack of images, split the channels. Extract the coordinates of the microspheres in each channel using the 3D Objects Counter. Subtract each channel's microsphere from the microsphere in the channel with the detected light shortest wavelength. This yields a chromatic shift from each channel in relation to that of the channel with the detected light shortest wavelength. This chromatic shift can then be applied to probe coordinates from probe hybridization rounds specifically based on the channel the probe signal originates from.

Microsphere Realignment

After a stack of images of a round of hybridization has been aligned to the baseline the microsphere locations must be obtained. Using the channel with the detected light shortest wavelength, excluding the DAPI channel, extract the location of the microspheres using 3D Objects Counter. Once the microsphere locations have been found, subtract the round of hybridization microsphere coordinates from the baseline microsphere coordinates, yielding the shift of the image.

Figures

Figure 3-1 - Cell and Chromosome Data Table

Cell type	Cell number	Xa	Xd	Xi
Fibroblast	16	16	0	16
Naïve	20	20	20	0

Figure 3-1 – Cell and Chromosome Data Table

Table displaying number and species of X chromosomes acquired via sequential FISH.

Figure 3-2 - Gene Table

A

Gene	Start	Stop	XCI inactive status	XCD escape status	Hybridization round	Fluorophore
CDKL5	18443725	18671749	Inactive		4	Cy3
UTX	44732423	44971845	Escape	Escape	3	Cy5
PAGE1	49452054	49460596			4	488
HUWE1	53559057	53713697	Inactive	Escape	2	Cy3
TSIX	73012040	73049066	Escape		2	488
XACT	112859587	113181506		Escape	3	488
LAMP2	119560003	119603204	Inactive		4	Cy5
THOC2	122734412	122866904	Inactive	Escape	2	Cy5
GPC3	132884115	133043964	Inactive		3	TR

B

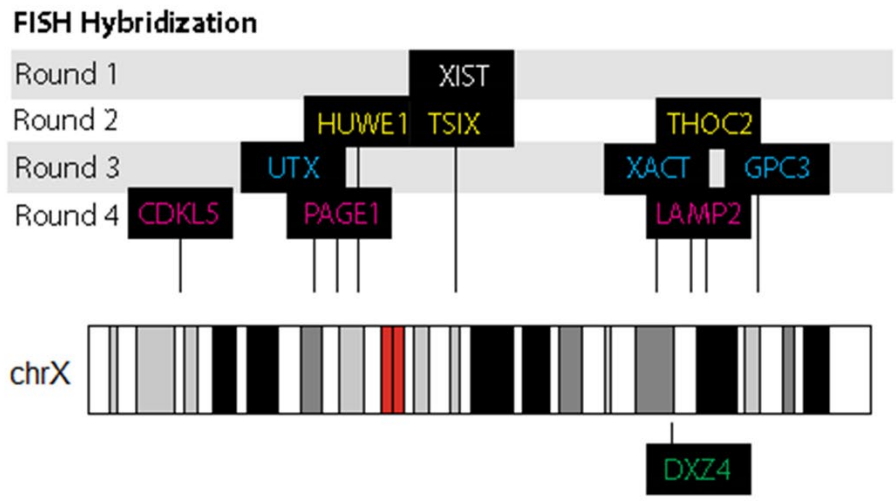


Figure 3-2 – Gene Table

(A) Table displaying location, gene expression status, hybridization round, and fluorophore used for each gene from the sequential FISH experiment. (B) Sequential FISH probe schematic indicating which genes are targeted during each hybridization as well as their relative locations along the X chromosome. DXZ4 is not targeted in the sequential FISH experiment but is indicated on the X chromosome map due to its structural significance.

Figure 3-3 - Individual Cell

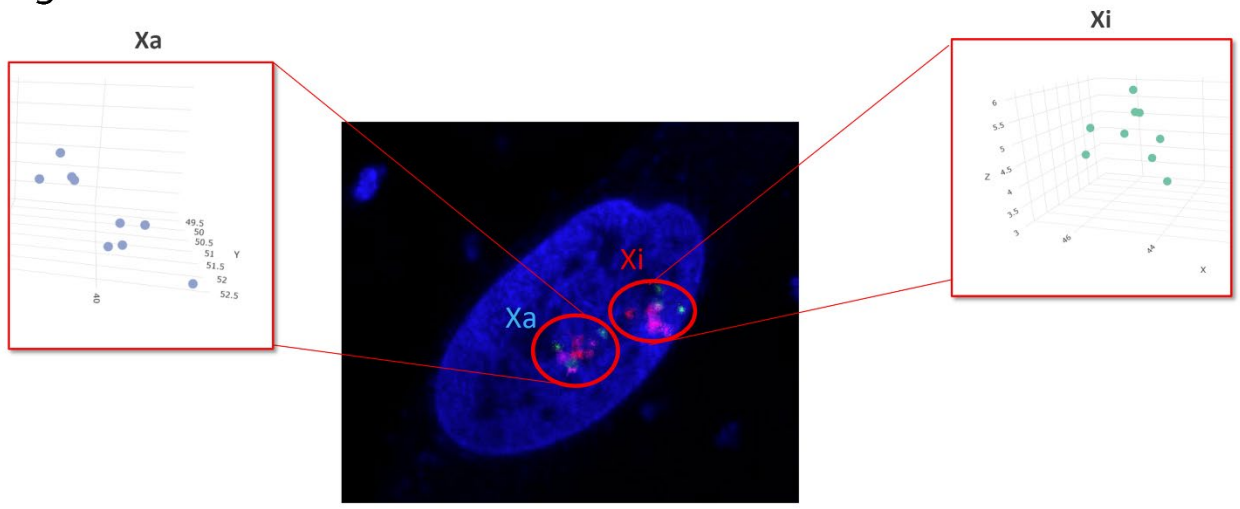


Figure 3-3 – Individual Cell Representation

Example fibroblast cell containing overlap of all 9 DNA FISH probes with X chromosome activation status marked from XIST RNA FISH. Fluorescent intensity signals extracted via Fiji and recapitulated in 3D form using the Plotly function via R.

Figure 3-4 - XIST RNA Probe Identifies Non-Active X Chromosome

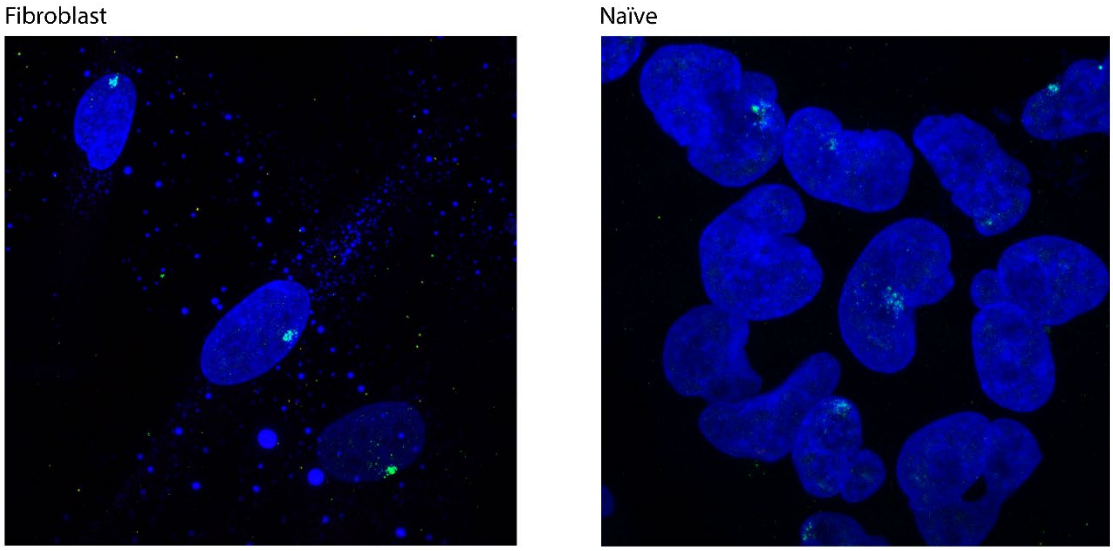
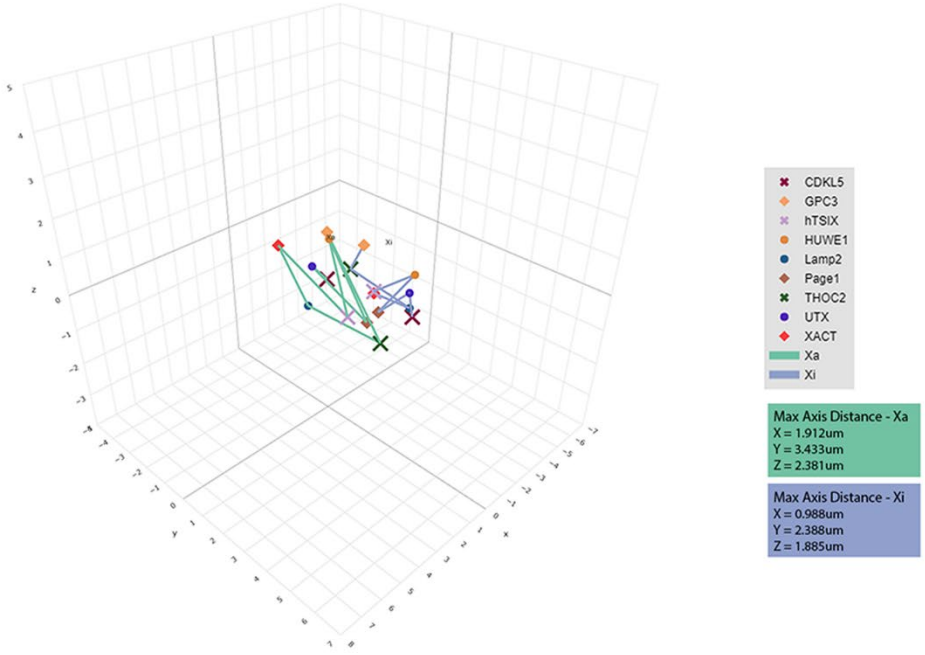


Figure 3-4 - XIST RNA Probe Identifies Non-Active X Chromosome

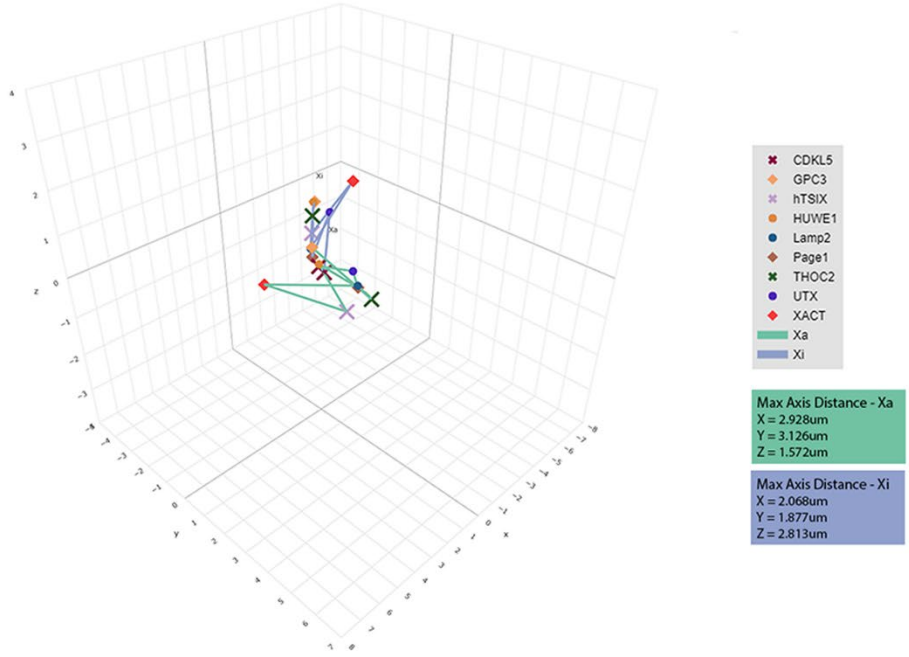
FOVs from both fibroblast and naïve cells demonstrating the clarity of XIST clouds via RNA FISH. Probes found within these regions can be accurately described as inactive or dampened X chromosomes in fibroblast and naïve respectively.

Figure 3-5 - Single Cell Chromosome Structures

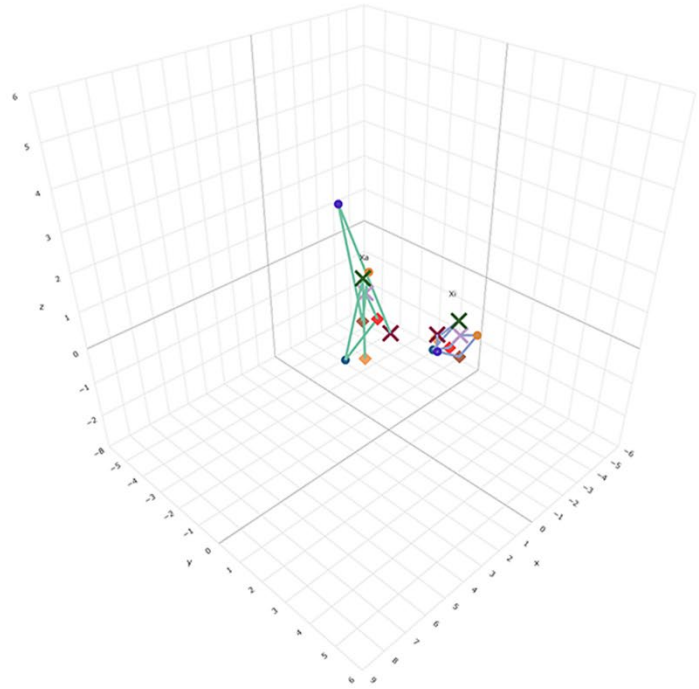
Fibroblast FOV1 Cell1



Fibroblast FOV1 Cell2



Fibroblast FOV1 Cell3



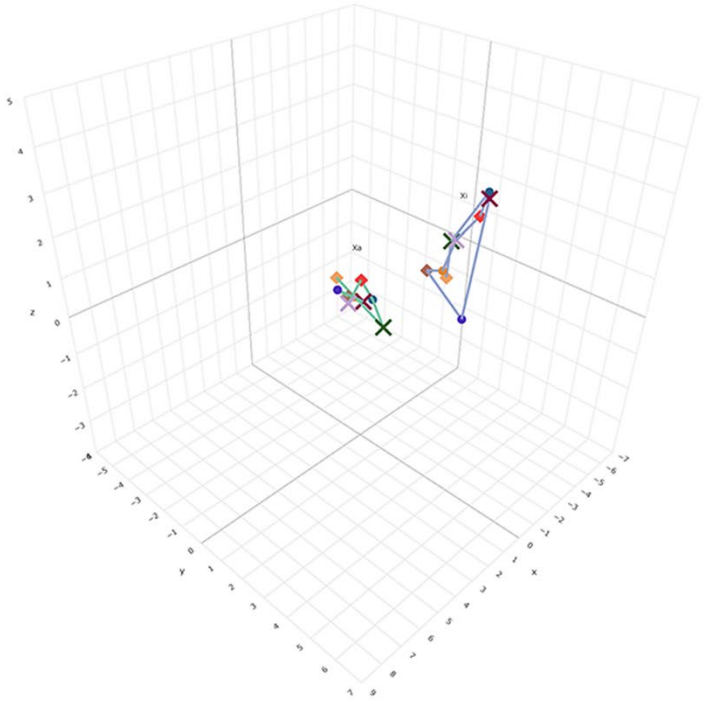
Bead Mean Distributions
r2-3 = 0.1309um
r2-4 = 0.1402um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 2.724um
Y = 2.314um
Z = 4.013um

Max Axis Distance - Xi
X = 2.146um
Y = 0.852um
Z = 0.901um

Fibroblast FOV4 Cell1

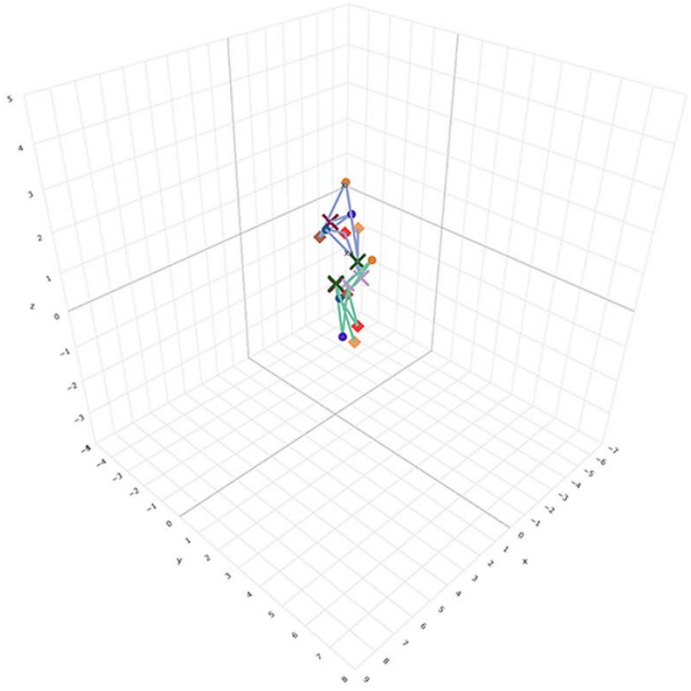


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 2.343um
Y = 3.072um
Z = 0.925um

Max Axis Distance - Xi
X = 1.057um
Y = 2.264um
Z = 2.834um

Fibroblast FOV4 Cell2



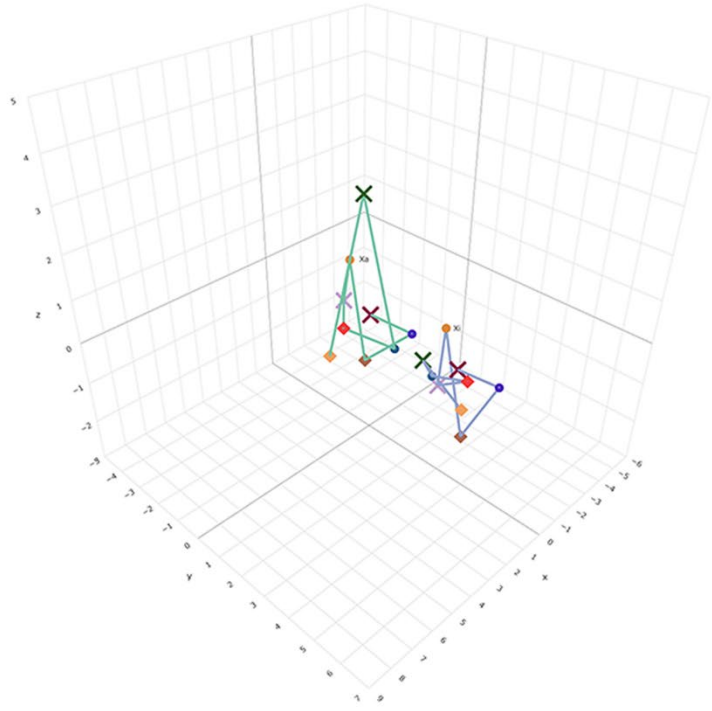
Bead Mean Distributions
r2-3 = 0.1183um
r2-4 = 0.1243um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 1.754um
Y = 1.995um
Z = 1.954um

Max Axis Distance - Xi
X = 1.733um
Y = 2.576um
Z = 1.906um

Fibroblast FOV5 Cell1

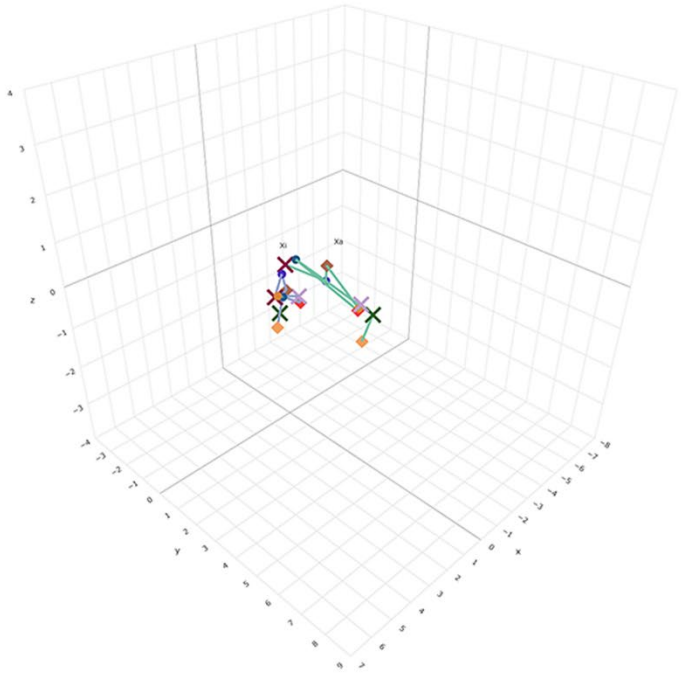


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 2.999um
Y = 2.747um
Z = 4.123um

Max Axis Distance - Xi
X = 0.715um
Y = 2.619um
Z = 2.162um

Fibroblast FOV5 Cell2



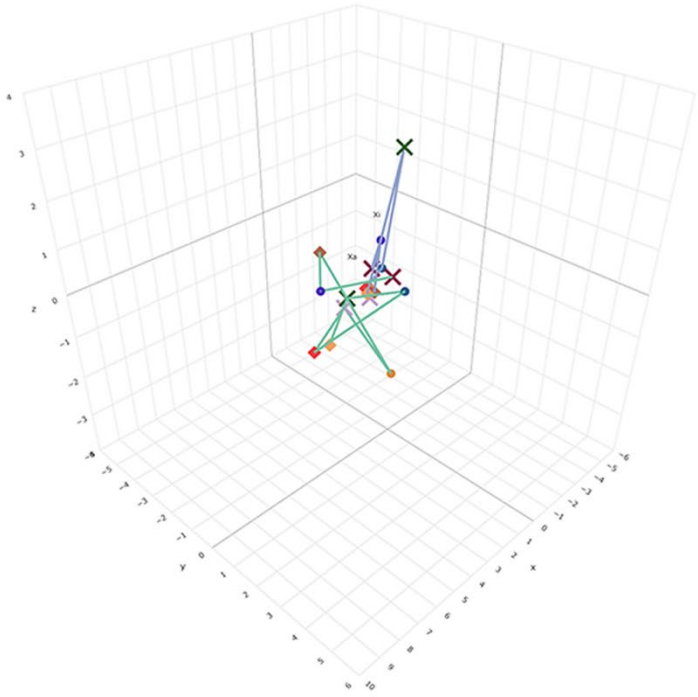
Bead Mean Distributions
r2-3 = 0.0713um
r2-4 = 0.1651um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 1.267um
Y = 3.252um
Z = 1.491um

Max Axis Distance - Xi
X = 1.569um
Y = 0.777um
Z = 1.005um

Fibroblast FOV6 Cell1

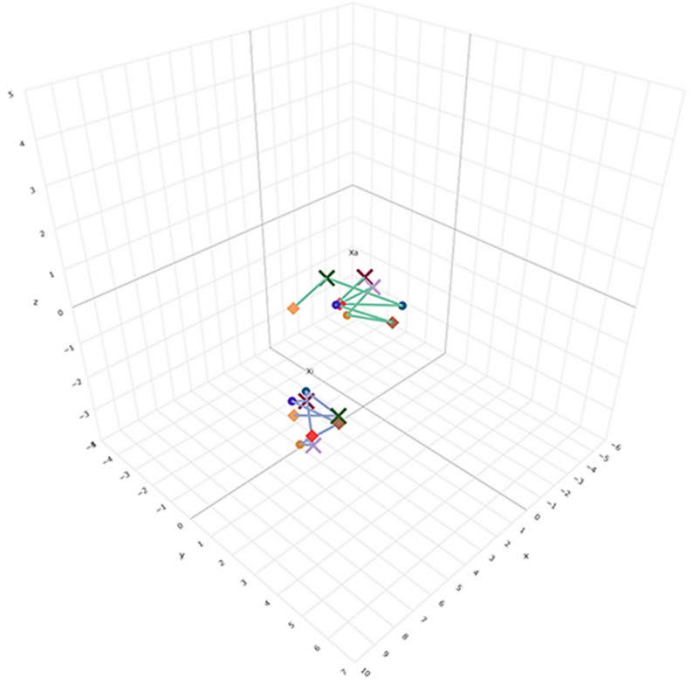


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 3.939um
Y = 2.700um
Z = 2.564um

Max Axis Distance - Xi
X = 1.818um
Y = 0.774um
Z = 3.504um

Fibroblast FOV6 Cell2



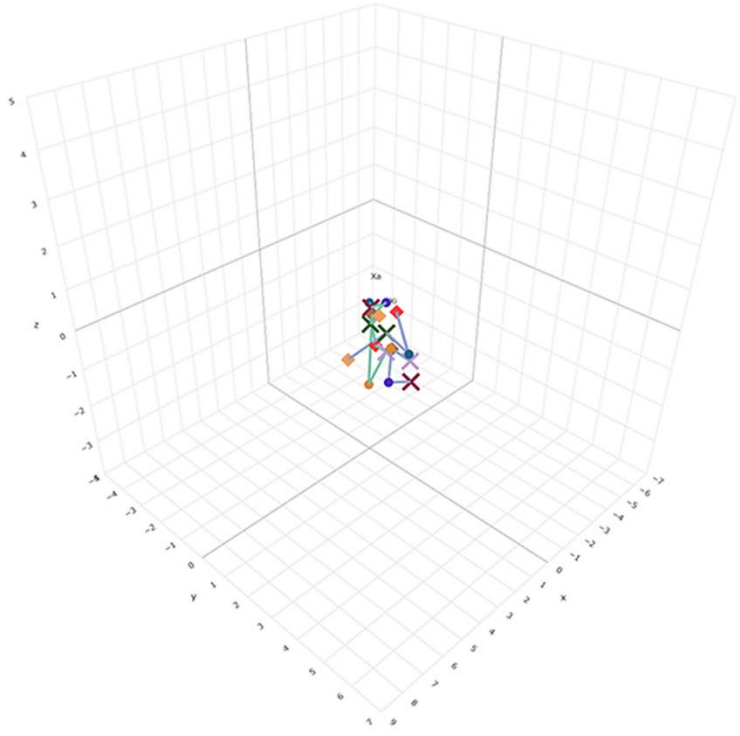
Bead Mean Distributions
r2-3 = 0.1535um
r2-4 = 0.1755um

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xi

Max Axis Distance - Xa
X = 4.440um
Y = 2.300um
Z = 1.772um

Max Axis Distance - Xi
X = 1.074um
Y = 1.355um
Z = 1.377um

Fibroblast FOV7 Cell1

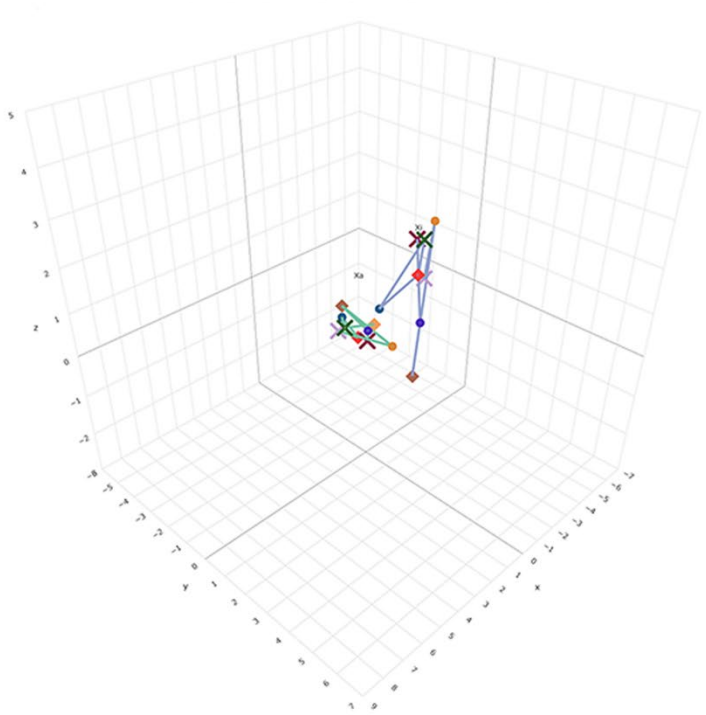


- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xi

Max Axis Distance - Xa
X = 2.083um
Y = 1.946um
Z = 1.494um

Max Axis Distance - Xi
X = 1.975um
Y = 3.442um
Z = 2.332um

Fibroblast FOV7 Cell2



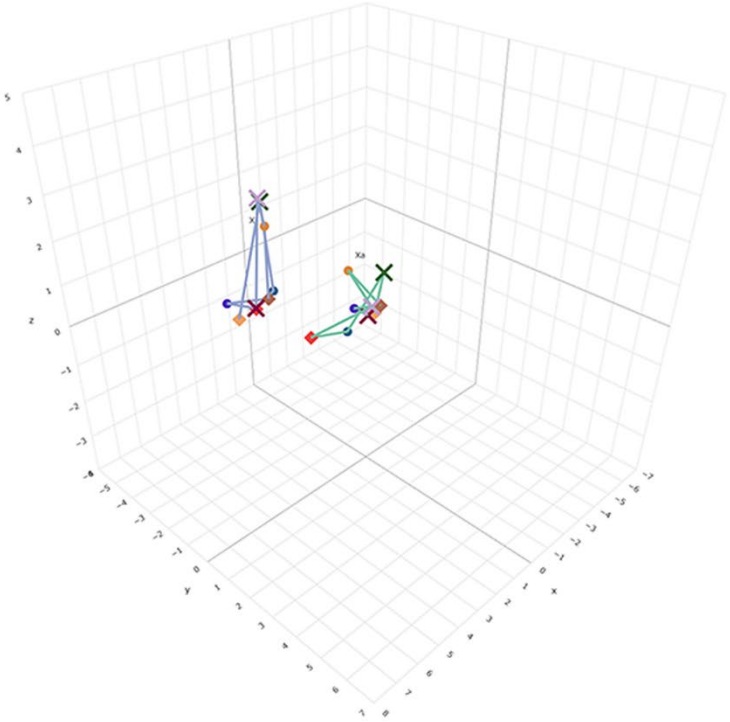
Bead Mean Distributions
r2-3 = 0.1602um
r2-4 = 0.0638um

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xi

Max Axis Distance - Xa
X = 1.709um
Y = 2.082um
Z = 0.800um

Max Axis Distance - Xi
X = 1.976um
Y = 1.794um
Z = 3.142um

Fibroblast FOV8 Cell1

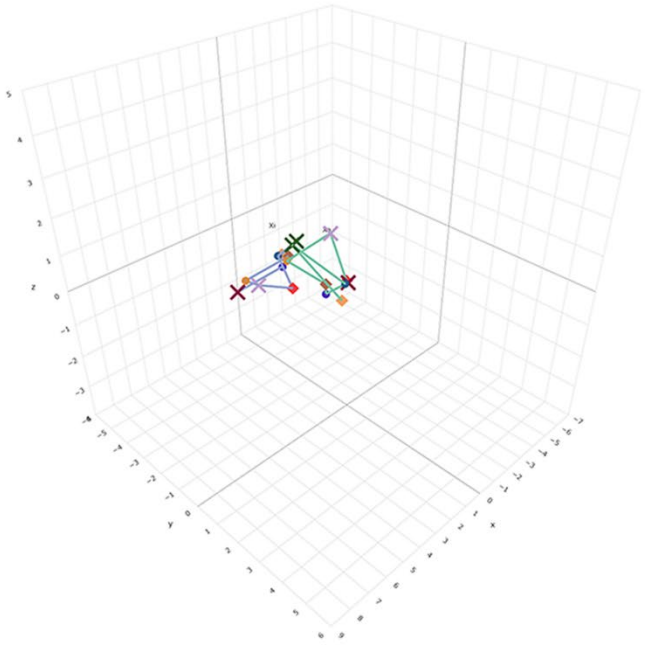


- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xi

Max Axis Distance - Xa
X = 2.458um
Y = 2.028um
Z = 2.181um

Max Axis Distance - Xi
X = 1.932um
Y = 1.524um
Z = 2.760um

Fibroblast FOV8 Cell2



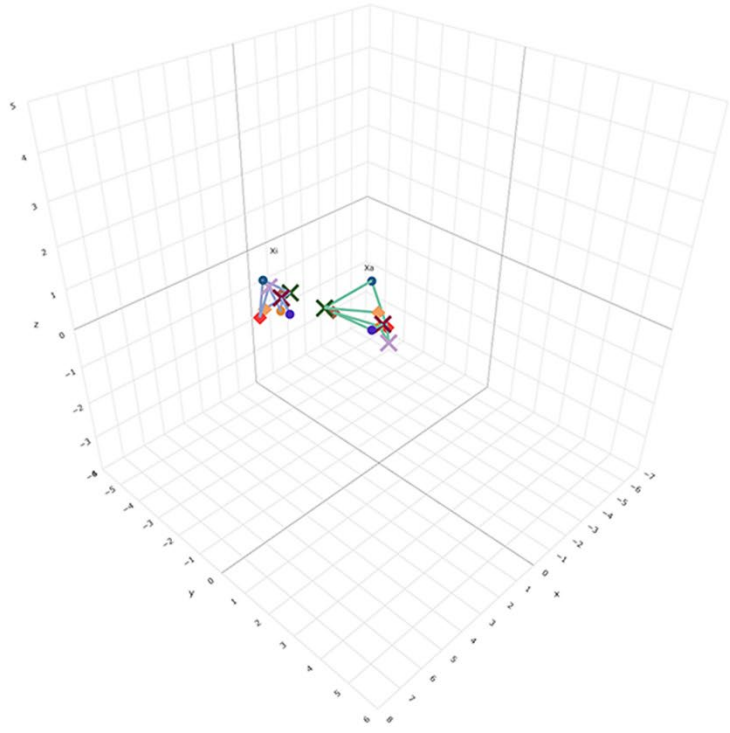
Bead Mean Distributions
r2-3 = 0.1917um
r2-4 = 0.1903um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 1.213um
Y = 2.276um
Z = 2.129um

Max Axis Distance - Xi
X = 3.612um
Y = 1.609um
Z = 1.496um

Fibroblast FOV9 Cell2



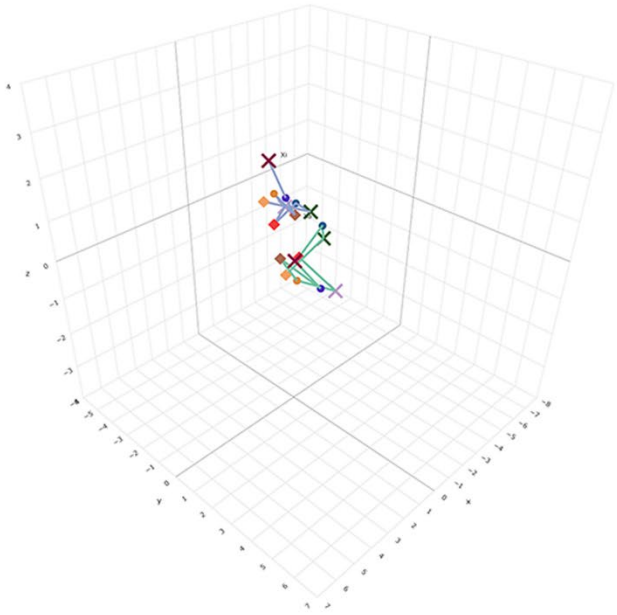
Bead Mean Distributions
r2-3 = 0.1838um
r2-4 = 0.0991um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xi

Max Axis Distance - Xa
X = 1.824um
Y = 2.631um
Z = 0.991um

Max Axis Distance - Xi
X = 1.542um
Y = 1.534um
Z = 0.801um

Fibroblast FOV10 Cell1



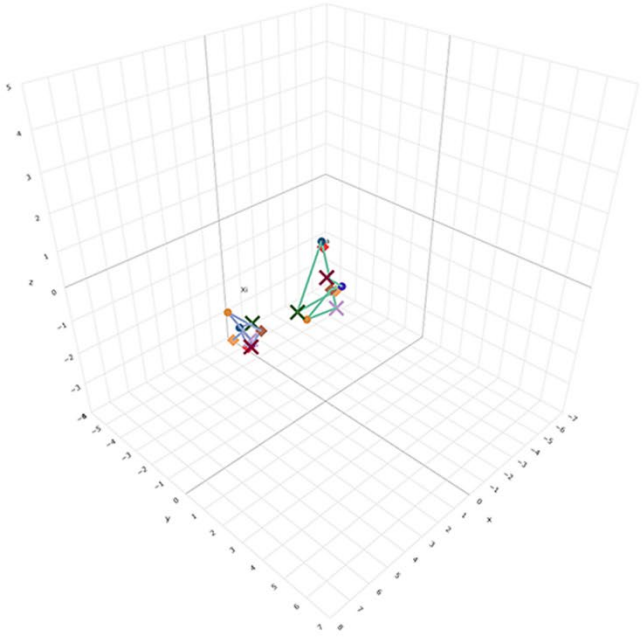
Bead Mean Distributions
r2-3 = 0.1806um
r2-4 = 0.1299um

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xi

Max Axis Distance - Xa
X = 2.492um
Y = 1.214um
Z = 1.333um

Max Axis Distance - Xi
X = 2.116um
Y = 0.978um
Z = 1.530um

Fibroblast FOV11 Cell1



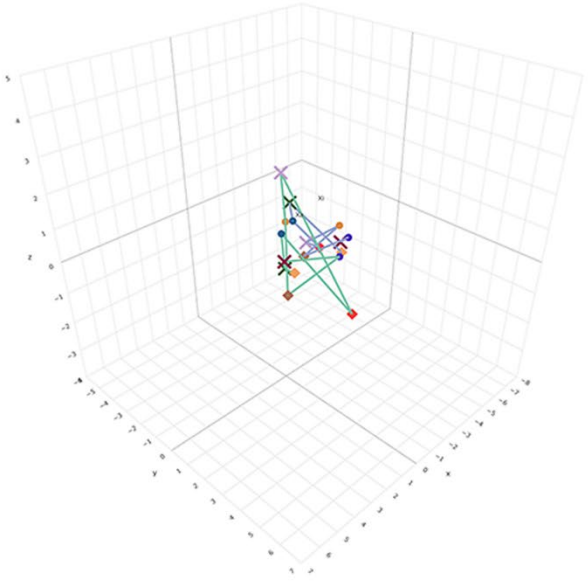
Bead Mean Distributions
r2-3 = 0.1524um
r2-4 = 0.1574um

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xi

Max Axis Distance - Xa
X = 3.324um
Y = 1.632um
Z = 1.415um

Max Axis Distance - Xi
X = 0.959um
Y = 0.988um
Z = 1.049um

Naïve FOV1 Cell1

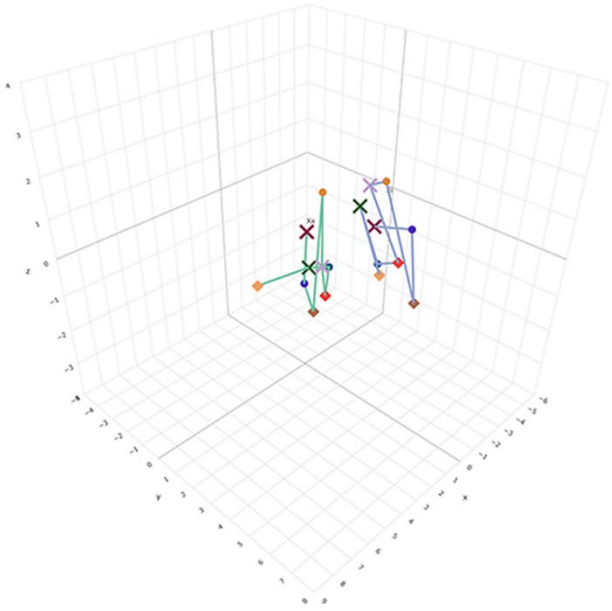


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.725um
Y = 3.263um
Z = 4.009um

Max Axis Distance - Xd
X = 2.782um
Y = 2.265um
Z = 1.465um

Naïve FOV1 Cell2

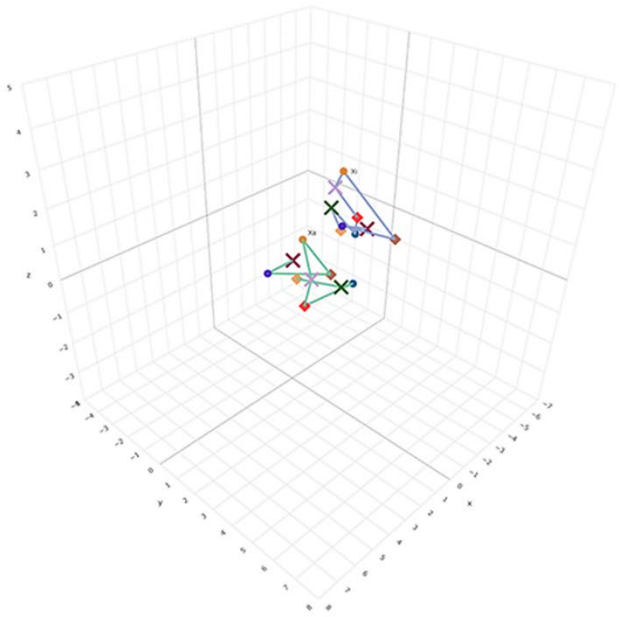


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.400um
Y = 2.065um
Z = 3.698um

Max Axis Distance - Xd
X = 2.733um
Y = 1.817um
Z = 3.456um

Naïve FOV1 Cell3



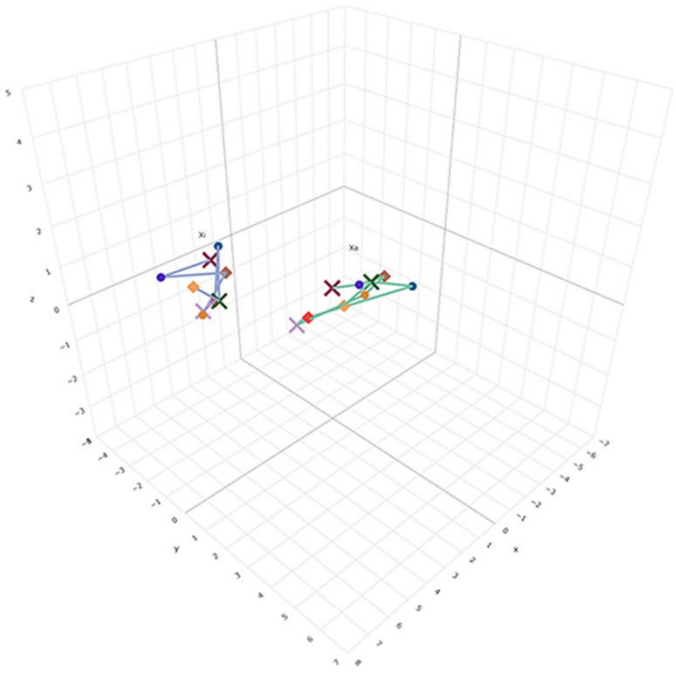
Bead Mean Distributions
r2-3 = 0.1918um
r2-4 = 0.1210um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.225um
Y = 3.015um
Z = 1.484um

Max Axis Distance - Xd
X = 1.727um
Y = 2.153um
Z = 1.845um

Naïve FOV2 Cell1



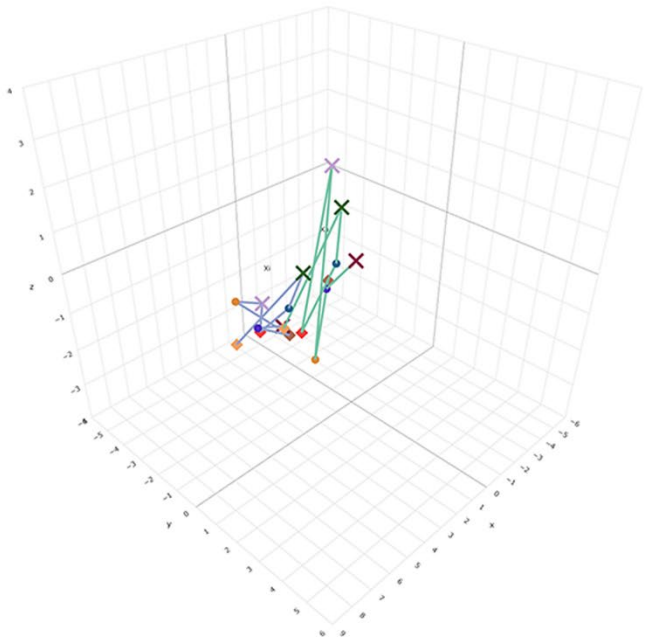
Bead Mean Distributions
r2-3 = 0.1660um
r2-4 = 0.1699um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 5.604um
Y = 2.094um
Z = 1.252um

Max Axis Distance - Xd
X = 4.860um
Y = 3.223um
Z = 1.811um

Naïve FOV2 Cell2

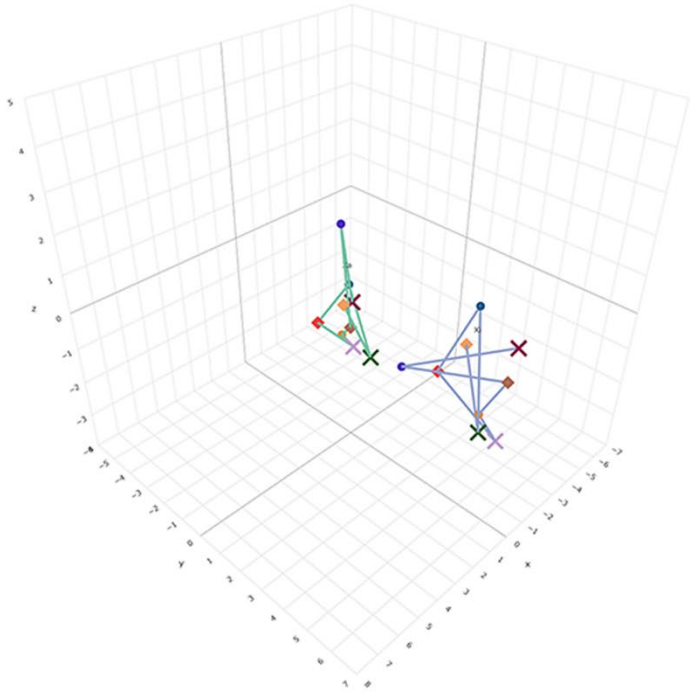


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 2.802um
Y = 1.487um
Z = 5.251um

Max Axis Distance - Xd
X = 2.827um
Y = 1.943um
Z = 2.311um

Naïve FOV3 Cell1

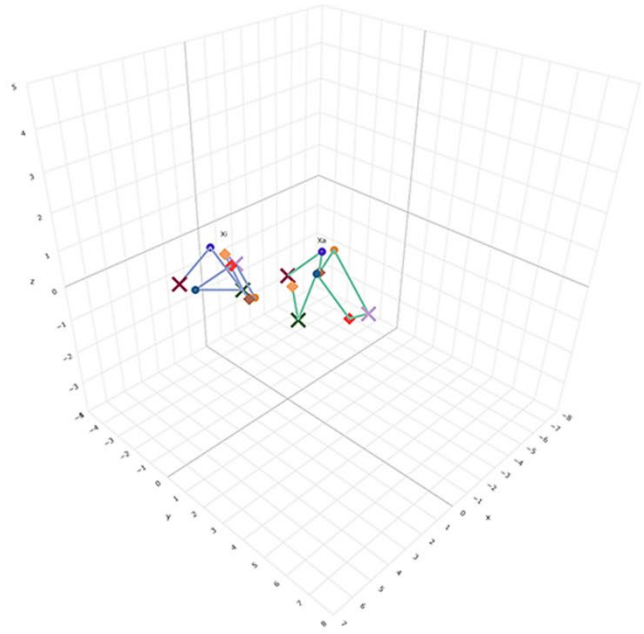


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 1.939um
Y = 1.691um
Z = 3.326um

Max Axis Distance - Xd
X = 3.699um
Y = 3.226um
Z = 2.707um

Naïve FOV3 Cell2

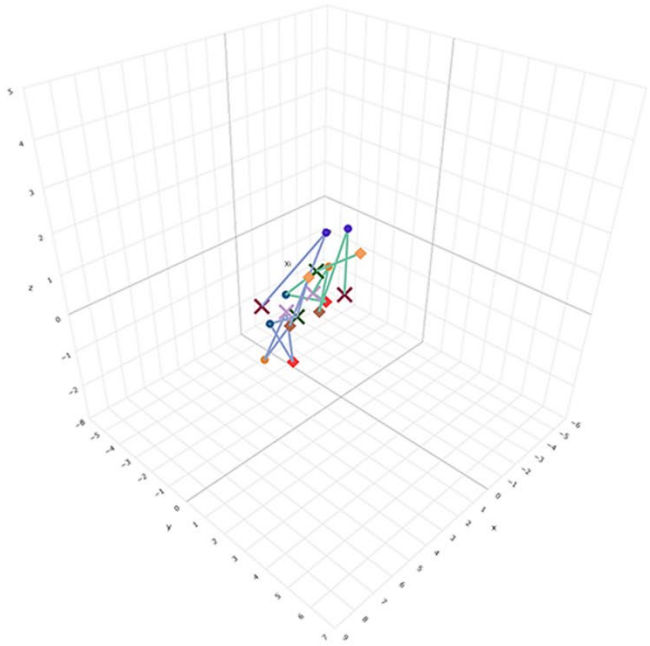


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 2.883um
Y = 3.209um
Z = 2.251um

Max Axis Distance - Xd
X = 4.317um
Y = 2.464um
Z = 1.737um

Naïve FOV3 Cell3



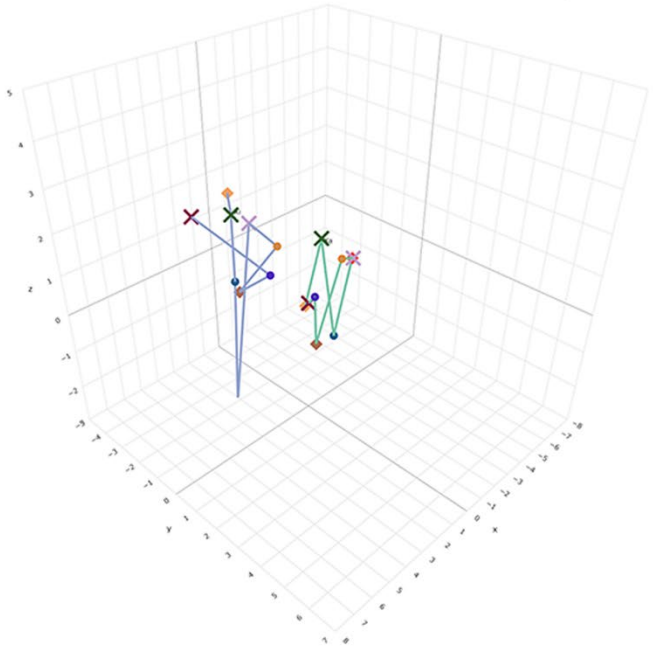
Bead Mean Distributions
r2-3 = 0.1183um
r2-4 = 0.1243um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 1.567um
Y = 3.194um
Z = 2.718um

Max Axis Distance - Xd
X = 2.218um
Y = 1.711um
Z = 3.188um

Naïve FOV5 Cell1

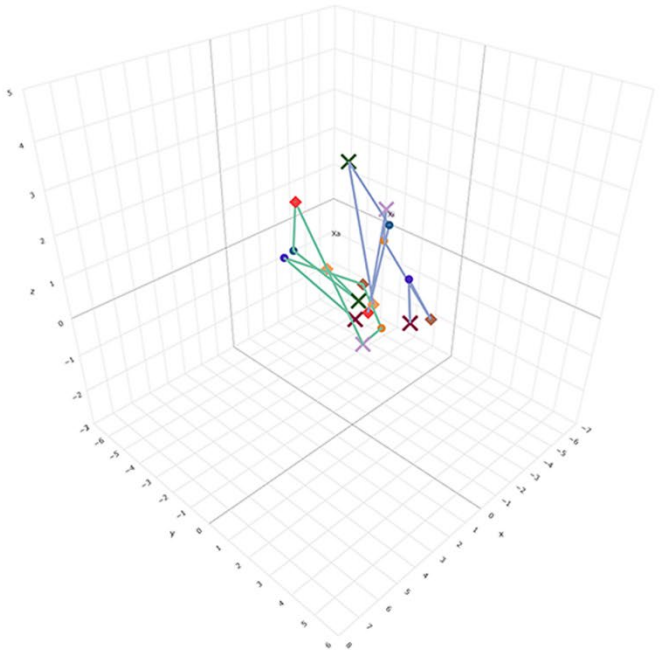


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.178um
Y = 1.690um
Z = 2.973um

Max Axis Distance - Xd
X = 1.788um
Y = 3.483um
Z = 5.699um

Naïve FOV5 Cell2



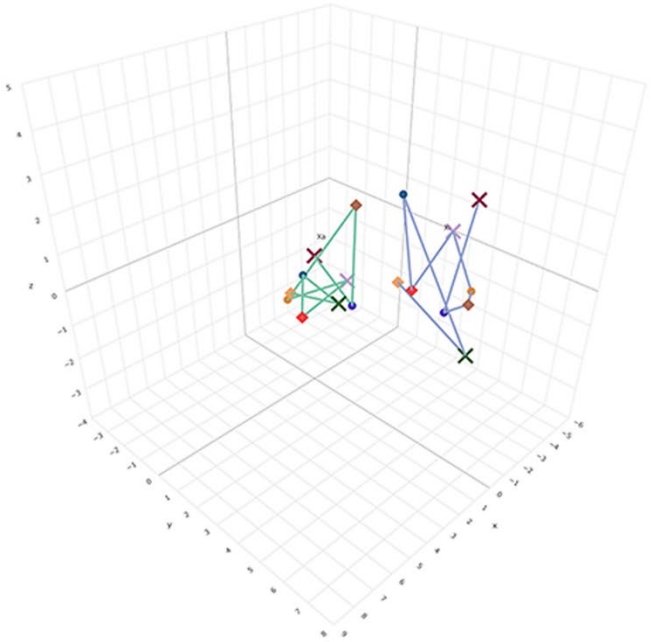
Bead Mean Distributions
r2-3 = 0.2003um
r2-4 = 0.2008um

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.393um
Y = 3.858um
Z = 2.541um

Max Axis Distance - Xd
X = 1.667um
Y = 3.547um
Z = 3.582um

Naïve FOV6 Cell1

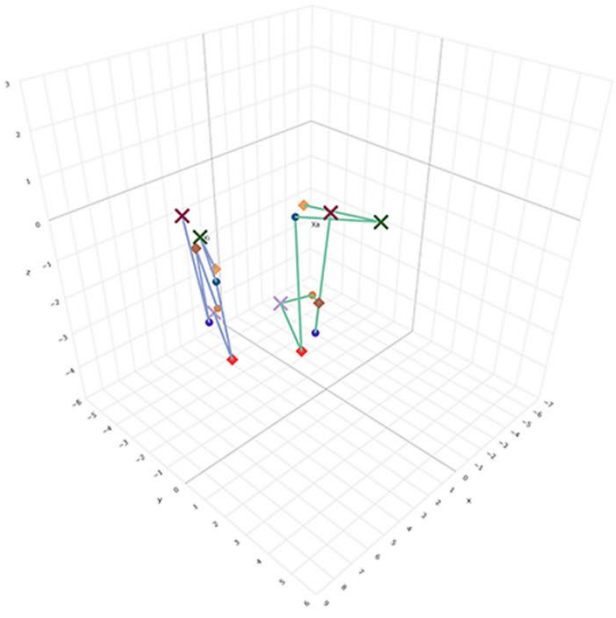


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 4.365um
Y = 2.695um
Z = 2.879um

Max Axis Distance - Xd
X = 2.356um
Y = 3.456um
Z = 3.532um

Naïve FOV6 Cell2



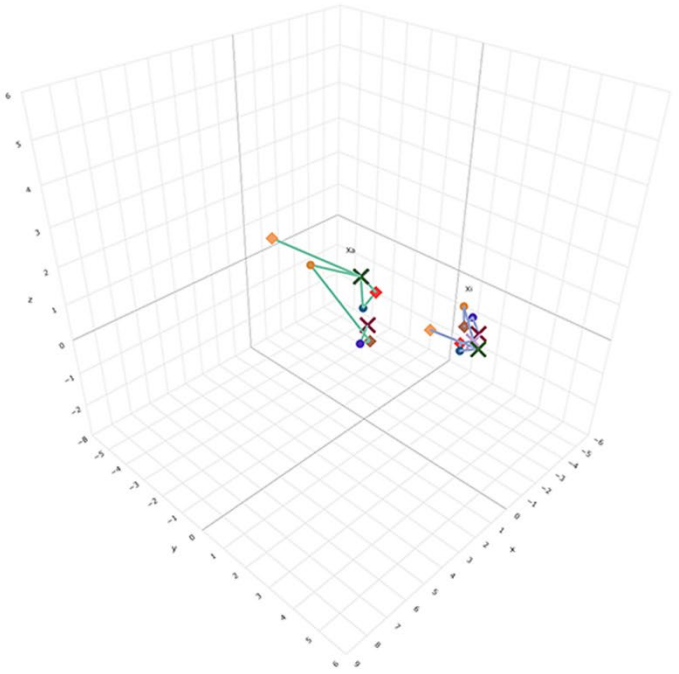
Bead Mean Distributions
r2-3 = 0.1168
r2-4 = 0.1578

CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.637um
Y = 3.150um
Z = 3.140um

Max Axis Distance - Xd
X = 2.017um
Y = 2.283um
Z = 3.140um

Naïve FOV7 Cell1

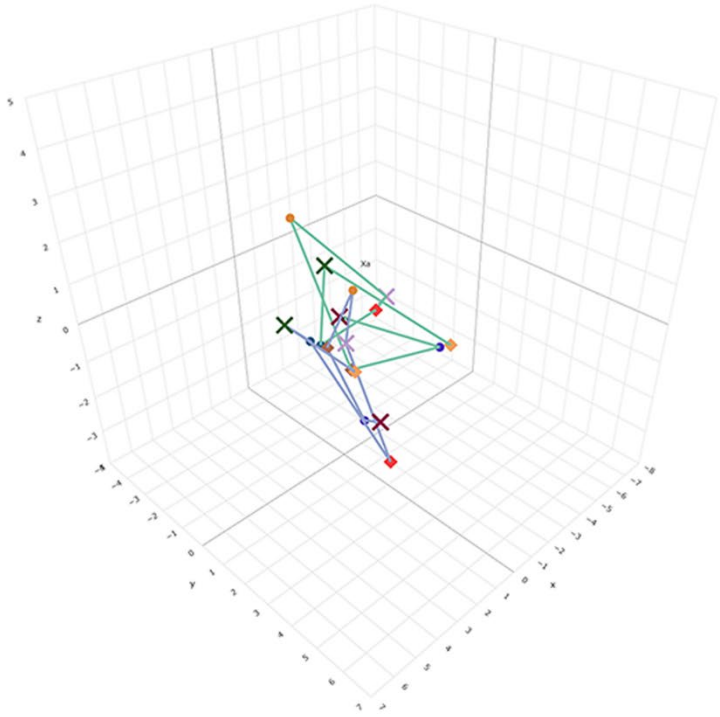


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 2.961um
Y = 2.384um
Z = 2.743um

Max Axis Distance - Xd
X = 2.597um
Y = 1.393um
Z = 0.654um

Naïve FOV7 Cell2

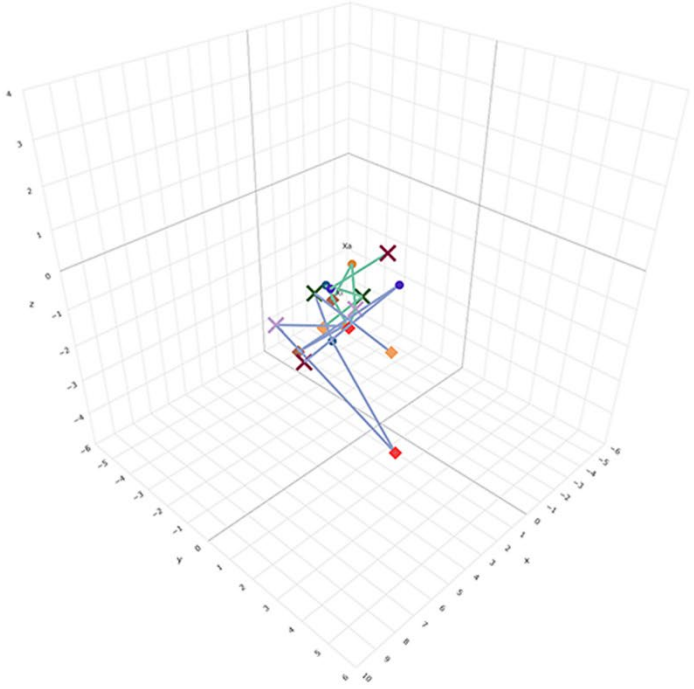


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 5.061um
Y = 3.368um
Z = 3.558um

Max Axis Distance - Xd
X = 1.418um
Y = 3.304um
Z = 3.513um

Naïve FOV7 Cell3



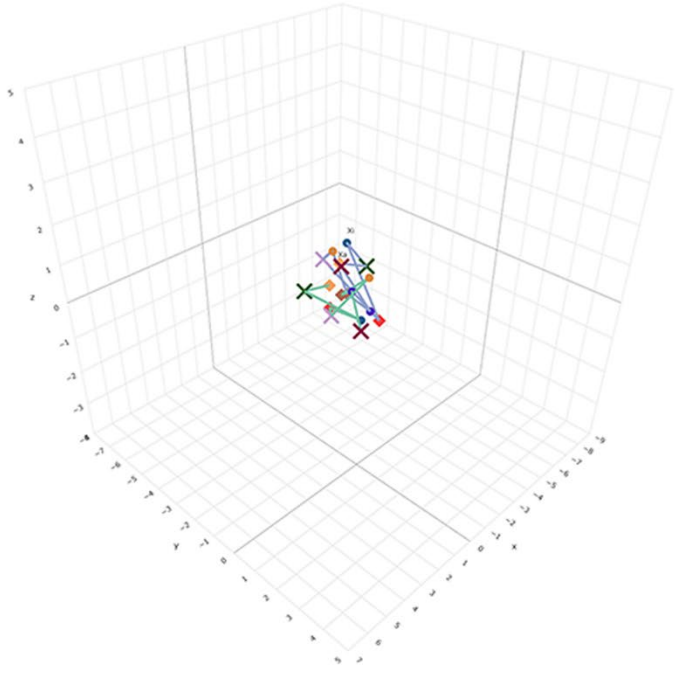
Bead Mean Distributions
r2-3 = 0.1602um
r2-4 = 0.0638um

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xd

Max Axis Distance - Xa
X = 3.459um
Y = 2.200um
Z = 1.368um

Max Axis Distance - Xd
X = 3.964um
Y = 3.918um
Z = 3.427um

Naïve FOV8 Cell1

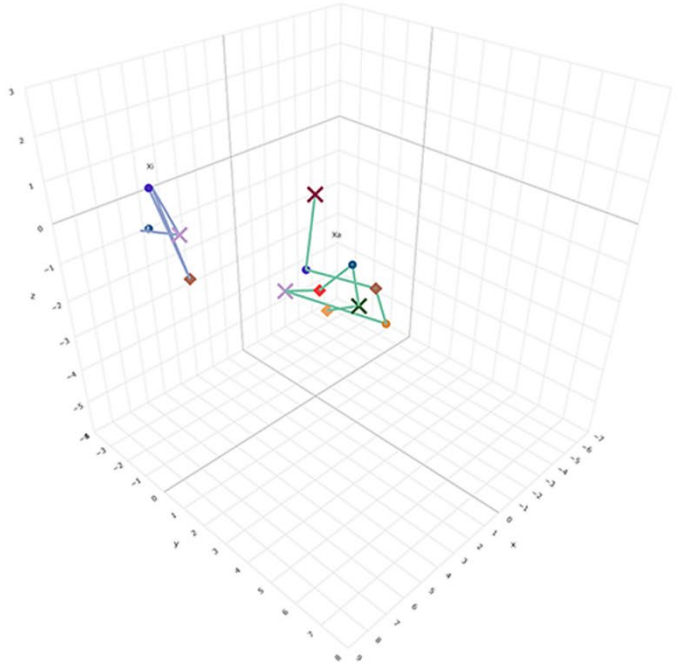


- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xd

Max Axis Distance - Xa
X = 2.024um
Y = 3.349um
Z = 1.084um

Max Axis Distance - Xd
X = 2.384um
Y = 1.566um
Z = 2.392um

Naïve FOV8 Cell2

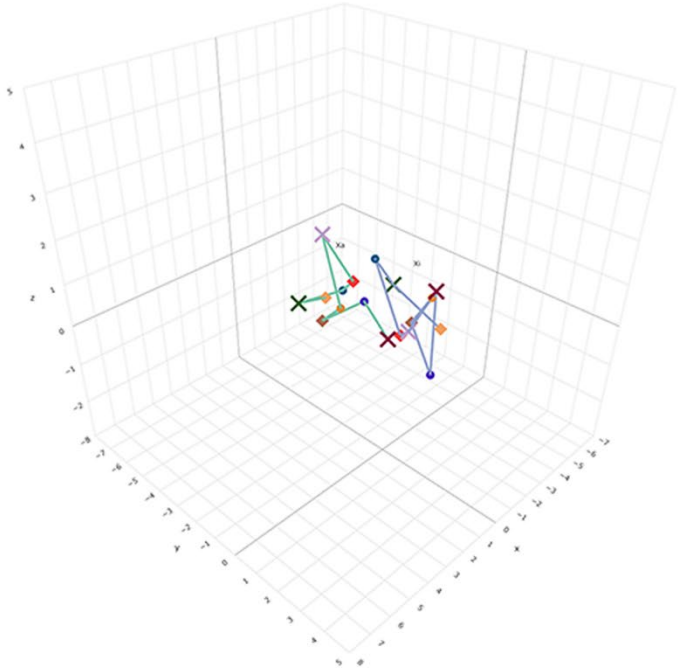


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 2.376um
Y = 2.780um
Z = 2.848um

Max Axis Distance - Xd
X = 1.359um
Y = 2.143um
Z = 2.021um

Naïve FOV8 Cell3

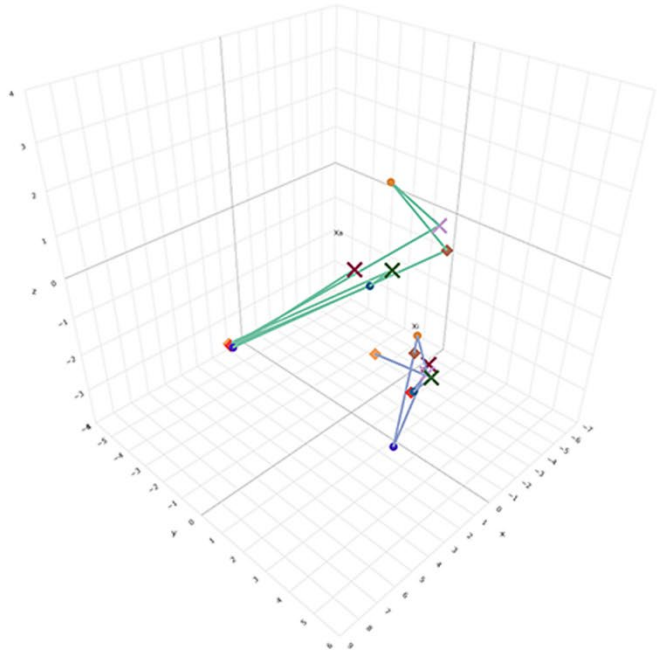


CDKL5
GPC3
hTSIX
HUWE1
Lamp2
Page1
THOC2
UTX
XACT
Xa
Xd

Max Axis Distance - Xa
X = 3.200um
Y = 3.193um
Z = 1.820um

Max Axis Distance - Xd
X = 1.231um
Y = 2.761um
Z = 2.195um

Naïve FOV8 Cell4



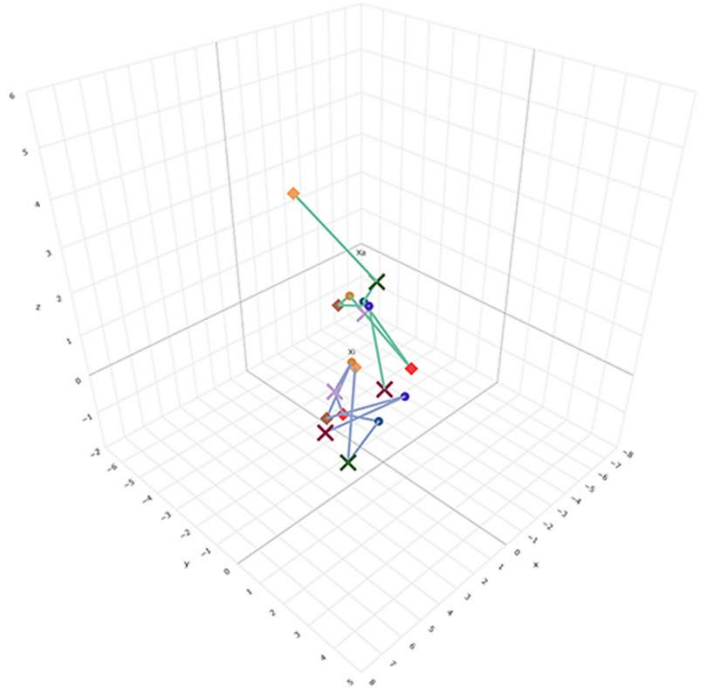
Bead Mean Distributions
r2-3 = 0.1917
r2-4 = 0.1903

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xd

Max Axis Distance - Xa
X = 6.350um
Y = 5.677um
Z = 4.503um

Max Axis Distance - Xd
X = 2.165um
Y = 1.919um
Z = 3.294um

Naïve FOV9 Cell1



Bead Mean Distributions
r2-3 = 0.1838um
r2-4 = 0.0991um

- CDKL5
- GPC3
- hTSIX
- HUWE1
- Lamp2
- Page1
- THOC2
- UTX
- XACT
- Xa
- Xd

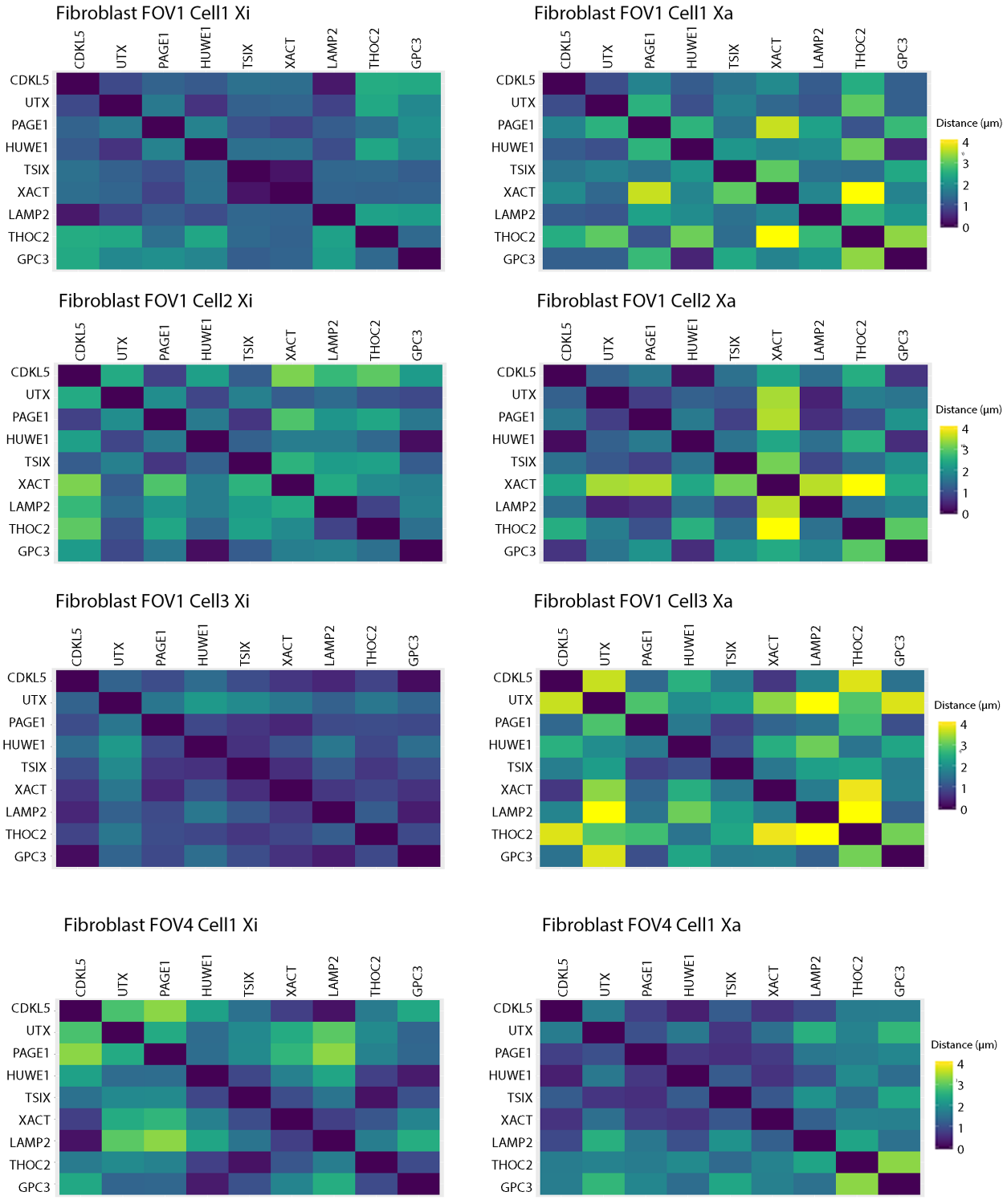
Max Axis Distance - Xa
X = 1.304um
Y = 4.954um
Z = 3.386um

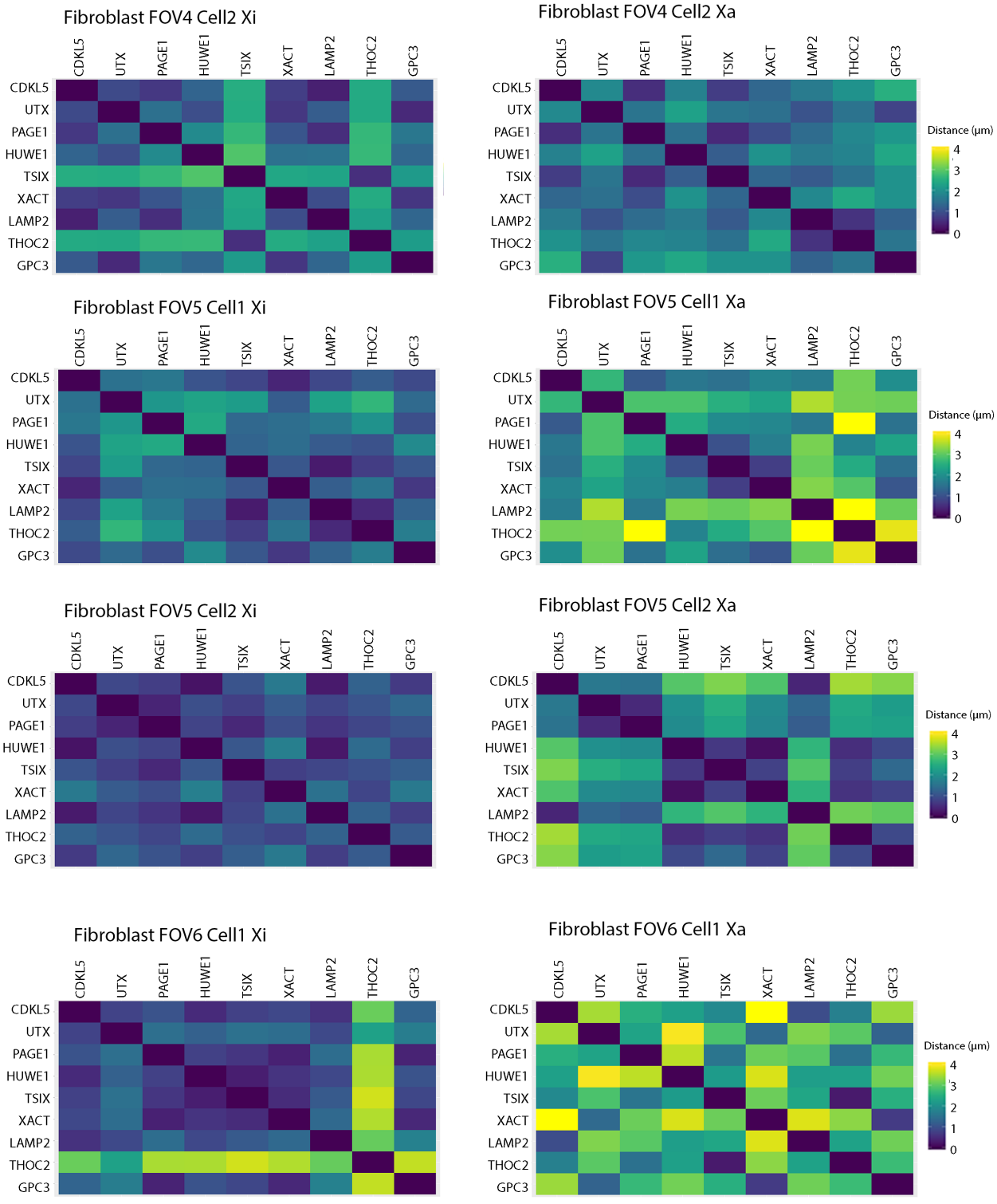
Max Axis Distance - Xd
X = 2.780um
Y = 1.949um
Z = 0.861um

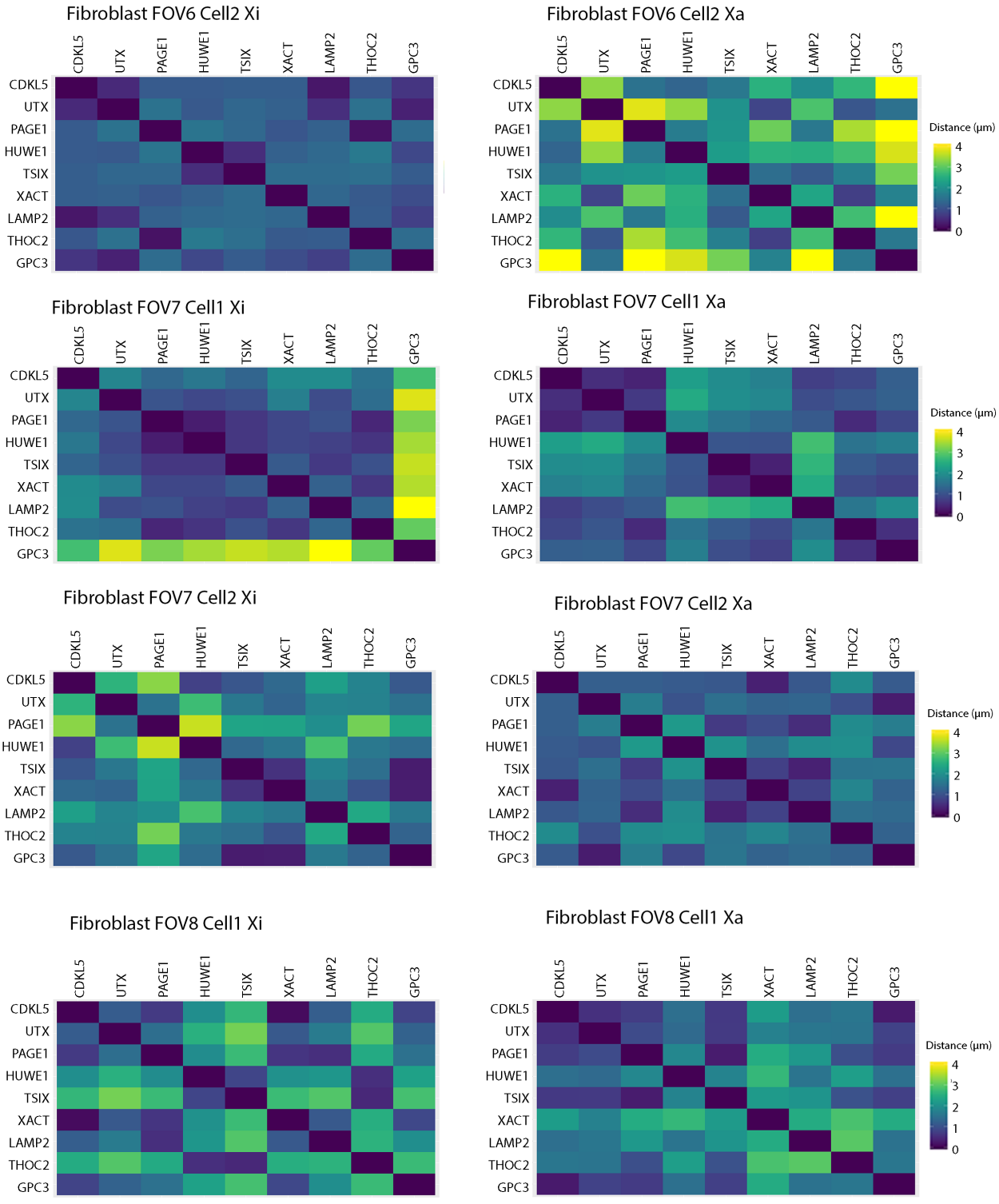
Figure 3-5 – Single Cell Chromosome Structures

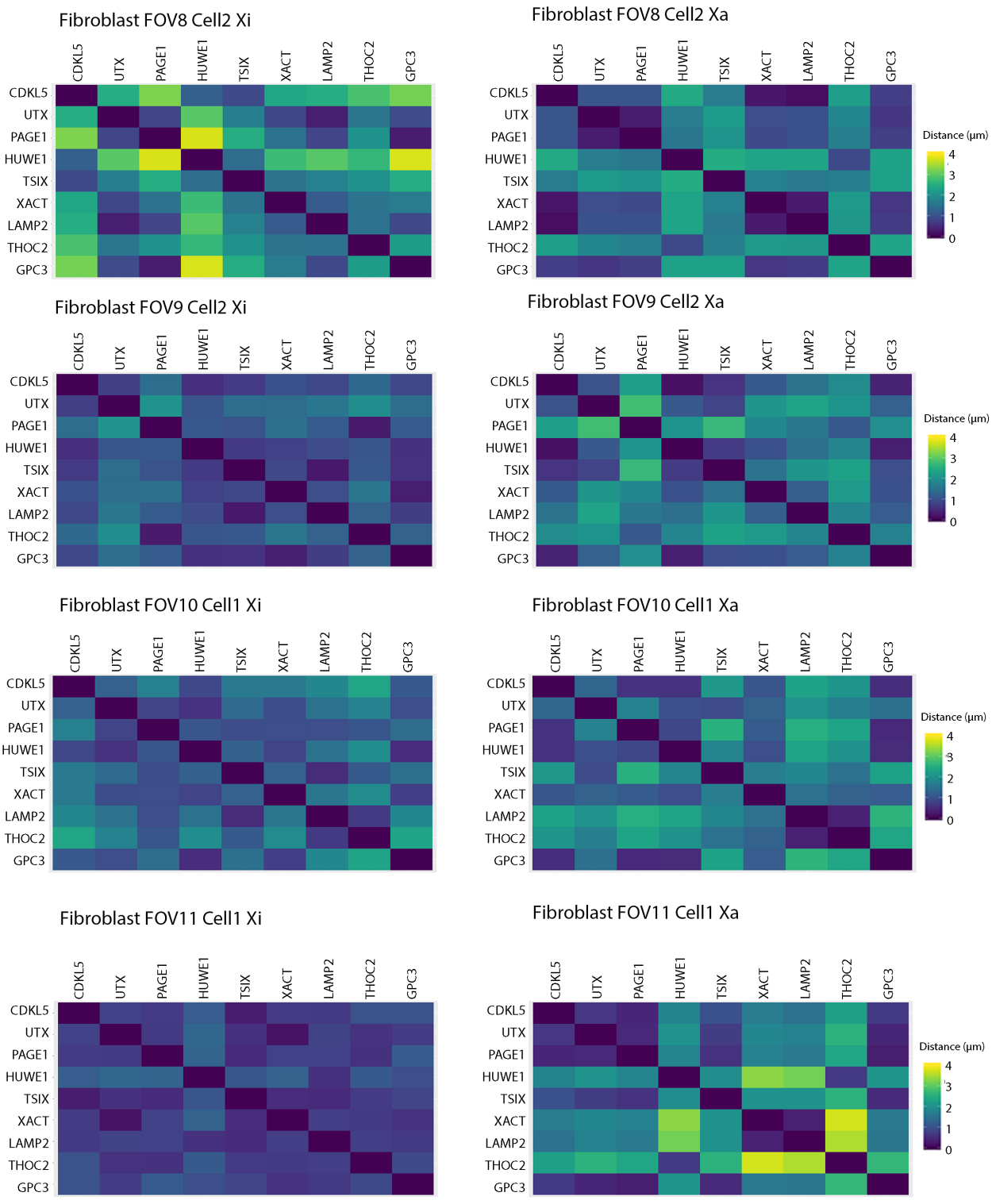
Individual cells separated by cell type and categorized by FOV. Plotly function within R assembled the 3D coordinates, each with a unique marker designated by gene, with a connecting line, colored by gene expression state, based on a head to tail organization of gene locations found on the X chromosome. Max Axis Distance indicates the largest distances of each chromosome in the X, Y, and Z axes. Bead Mean Distributions are listed per FOV, indicating accuracy of alignment based on the distance the microspheres shift per hybridization round.

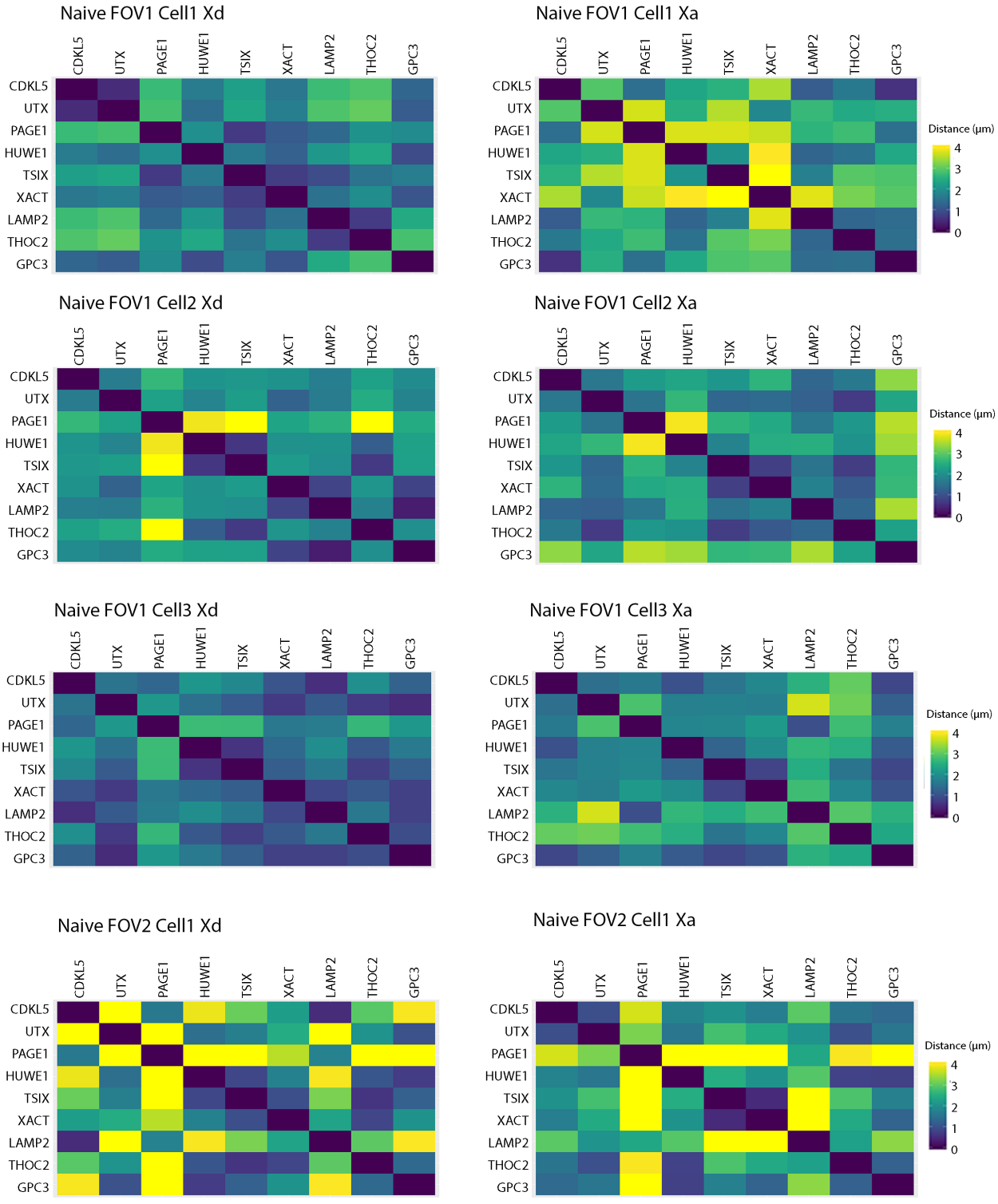
Figure 3-6 - X Chromosome Spatial Distance Maps

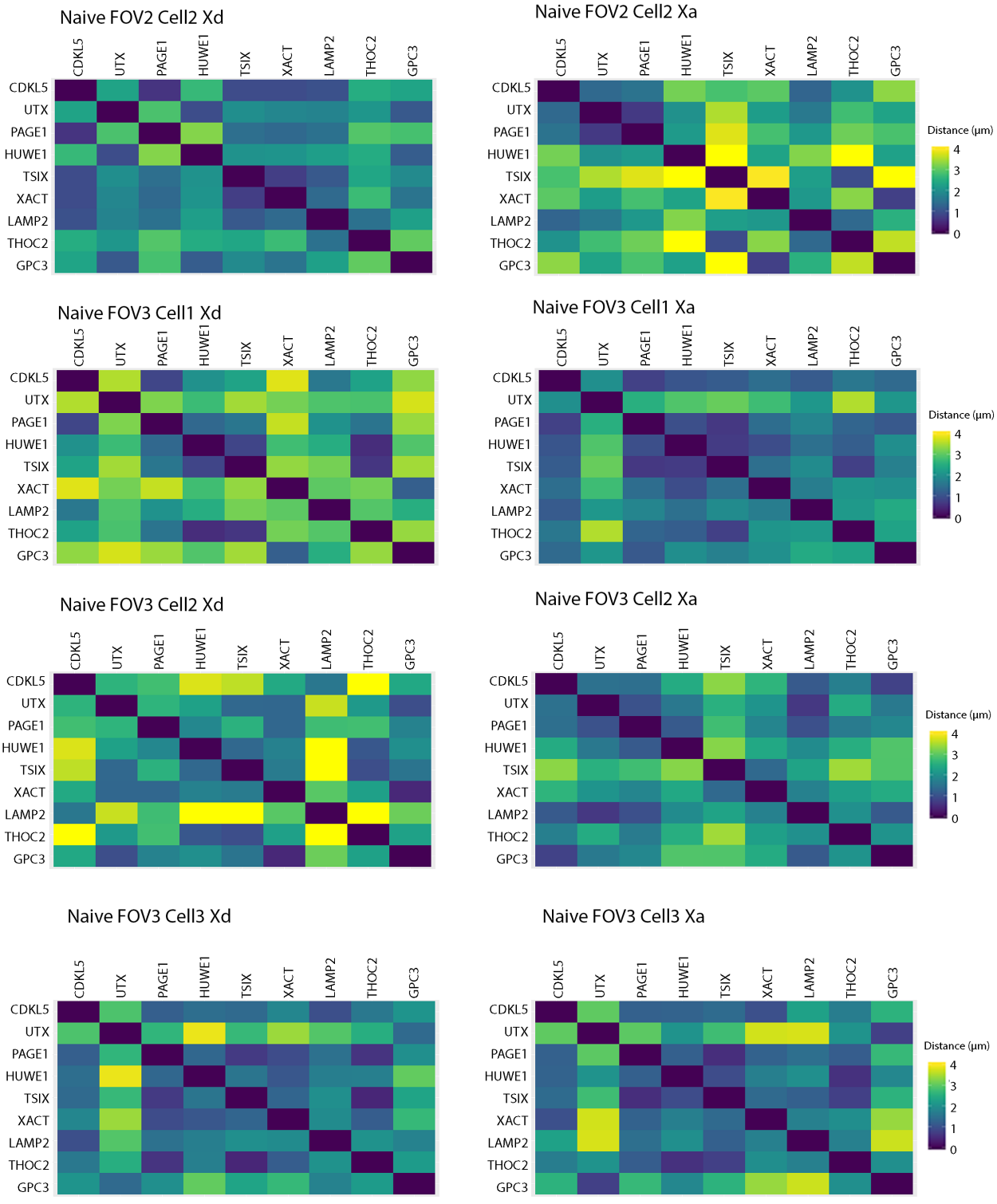


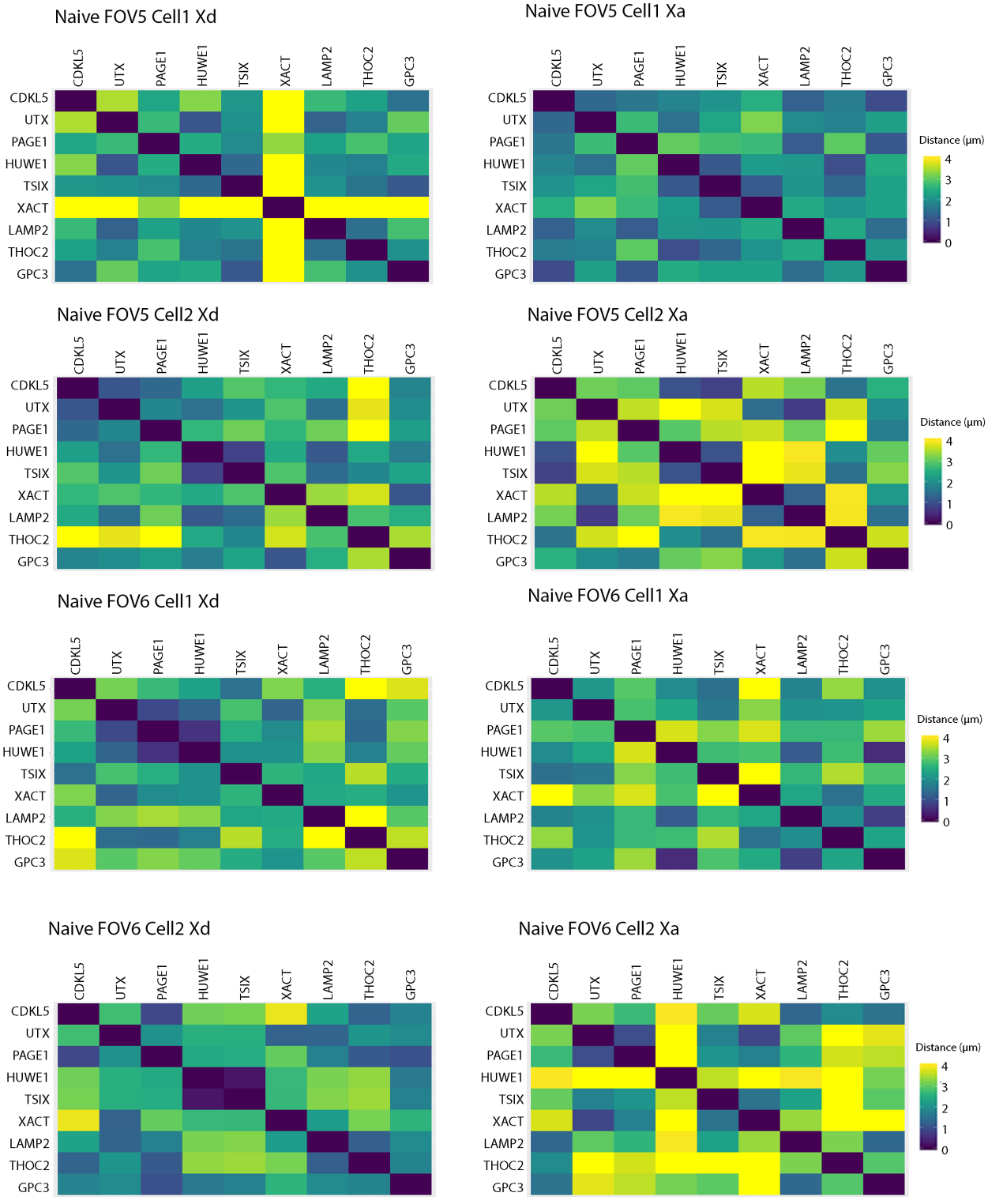

Fibroblast FOV4 Cell2 Xi
Fibroblast FOV4 Cell2 Xa
Fibroblast FOV5 Cell1 Xi
Fibroblast FOV5 Cell1 Xa
Fibroblast FOV5 Cell2 Xi
Fibroblast FOV5 Cell2 Xa
Fibroblast FOV6 Cell1 Xi
Fibroblast FOV6 Cell1 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0

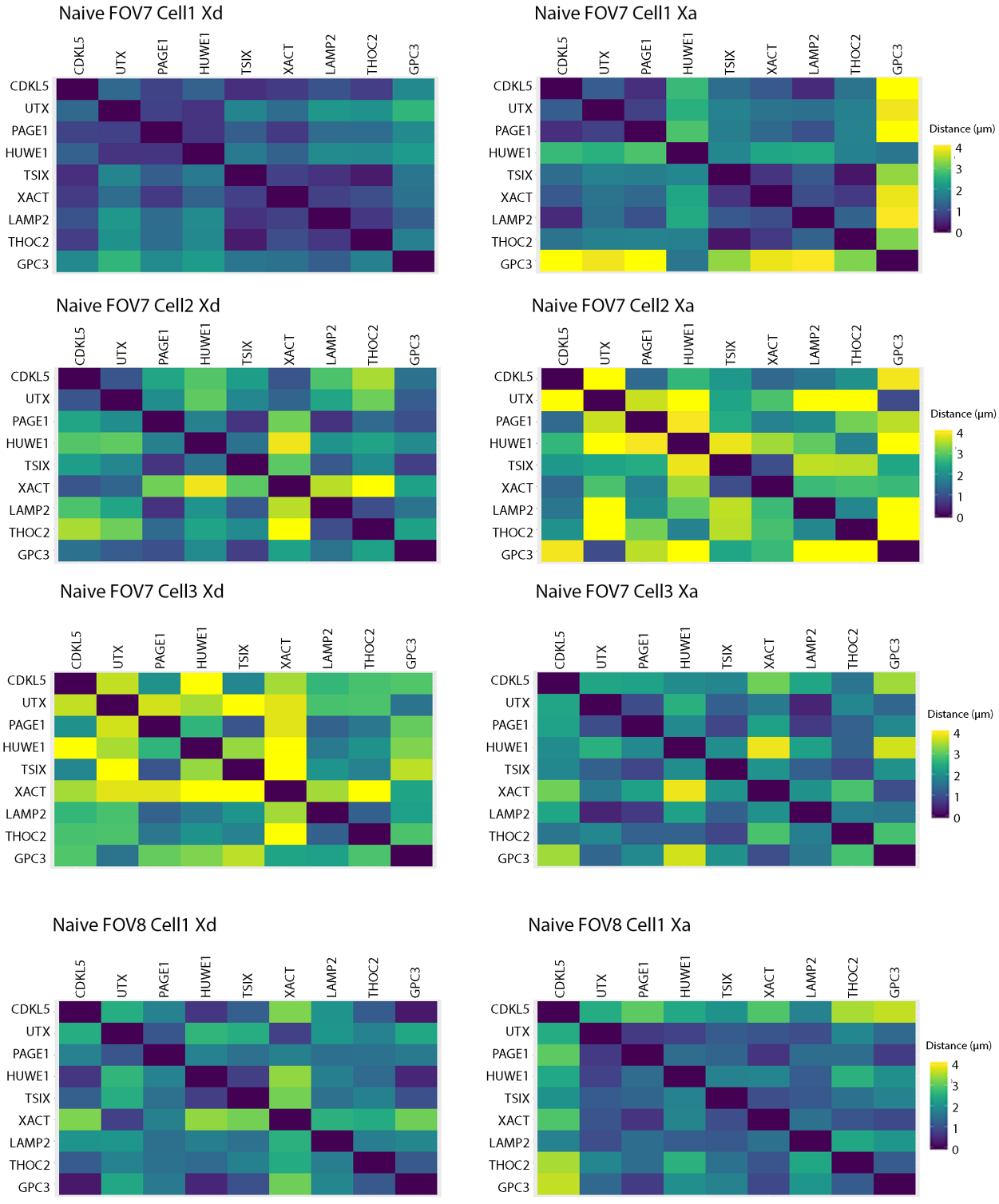

Fibroblast FOV6 Cell2 Xi
Fibroblast FOV6 Cell2 Xa
Fibroblast FOV7 Cell1 Xi
Fibroblast FOV7 Cell1 Xa
Fibroblast FOV7 Cell2 Xi
Fibroblast FOV7 Cell2 Xa
Fibroblast FOV8 Cell1 Xi
Fibroblast FOV8 Cell1 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0


Fibroblast FOV8 Cell2 Xi
Fibroblast FOV8 Cell2 Xa
Fibroblast FOV9 Cell2 Xi
Fibroblast FOV9 Cell2 Xa
Fibroblast FOV10 Cell1 Xi
Fibroblast FOV10 Cell1 Xa
Fibroblast FOV11 Cell1 Xi
Fibroblast FOV11 Cell1 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0


Naive FOV1 Cell1 Xd
Naive FOV1 Cell1 Xa
Naive FOV1 Cell2 Xd
Naive FOV1 Cell2 Xa
Naive FOV1 Cell3 Xd
Naive FOV1 Cell3 Xa
Naive FOV2 Cell1 Xd
Naive FOV2 Cell1 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0


Naive FOV2 Cell2 Xd
Naive FOV2 Cell2 Xa
Naive FOV3 Cell1 Xd
Naive FOV3 Cell1 Xa
Naive FOV3 Cell2 Xd
Naive FOV3 Cell2 Xa
Naive FOV3 Cell3 Xd
Naive FOV3 Cell3 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0


Naive FOV5 Cell1 Xd
Naive FOV5 Cell1 Xa
Naive FOV5 Cell2 Xd
Naive FOV5 Cell2 Xa
Naive FOV6 Cell1 Xd
Naive FOV6 Cell1 Xa
Naive FOV6 Cell2 Xd
Naive FOV6 Cell2 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0


Naive FOV7 Cell1 Xd
Naive FOV7 Cell1 Xa
Naive FOV7 Cell2 Xd
Naive FOV7 Cell2 Xa
Naive FOV7 Cell3 Xd
Naive FOV7 Cell3 Xa
Naive FOV8 Cell1 Xd
Naive FOV8 Cell1 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0

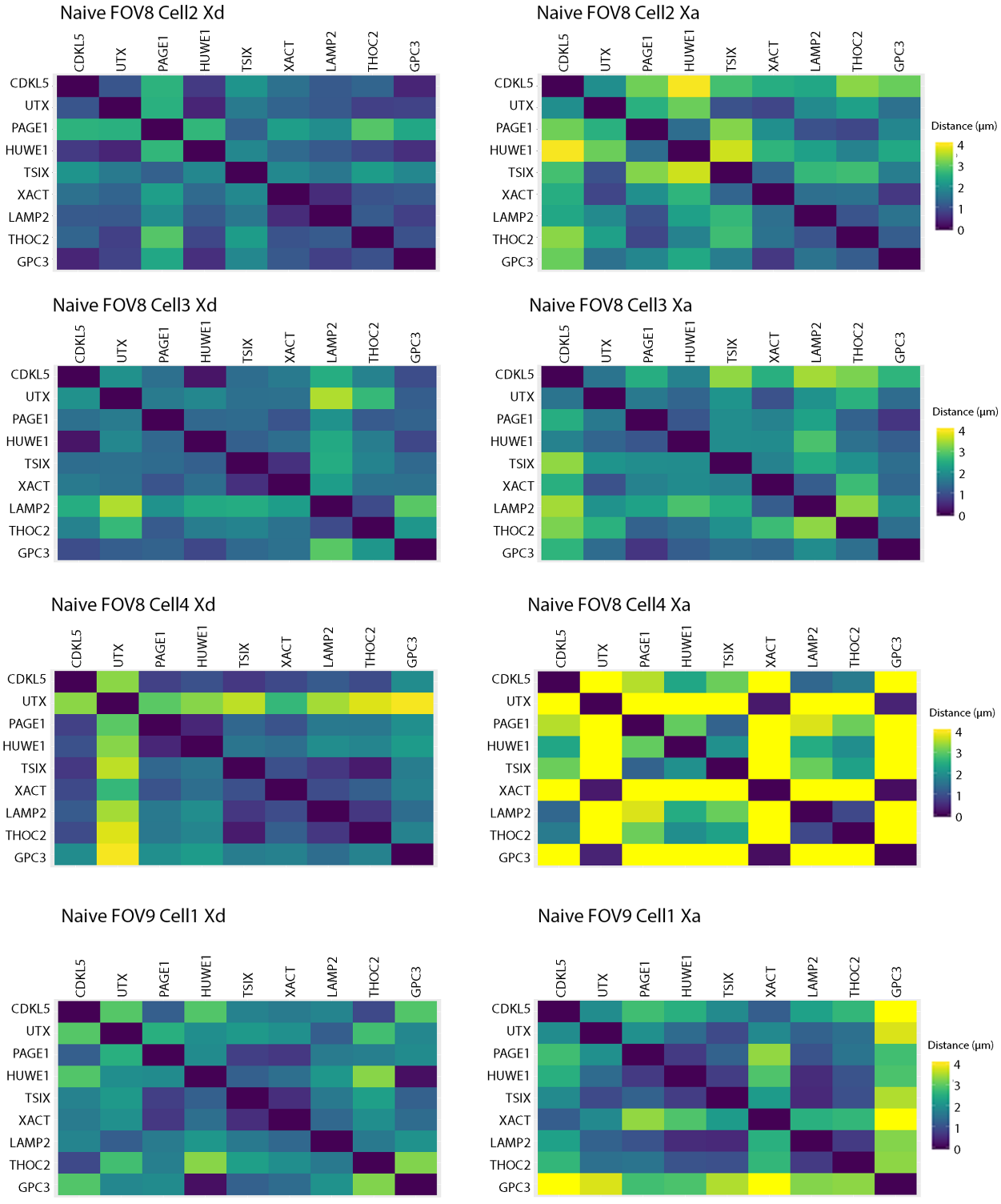

Naive FOV8 Cell2 Xd
Naive FOV8 Cell2 Xa
Naive FOV8 Cell3 Xd
Naive FOV8 Cell3 Xa
Naive FOV8 Cell4 Xd
Naive FOV8 Cell4 Xa
Naive FOV9 Cell1 Xd
Naive FOV9 Cell1 Xa
CDKL5
UTX
PAGE1
HUWE1
TSIX
XACT
LAMP2
THOC2
GPC3
Distance (µm)
4
3
2
1
0

Figure 3-6 – X Chromosome Spatial Distance Maps

Single X intra-chromosomal distances from probe to probe are separated by cell type and displayed using the ggplot2 function of R. Probes are found in each row and column and distances between the two probes can be found by locating the intersection of the row and column of the desired pair of probes. Distances are colored by magnitude and are arranged in a matrix based on head to tail organization of gene locations found on the X chromosome. Each row indicates a single cell with the active X chromosome in the right column and the non-active in the left column.

Figure 3-7 - Longest XIST cloud distance indicates that Naïve Xd skews larger than Fibroblast Xi

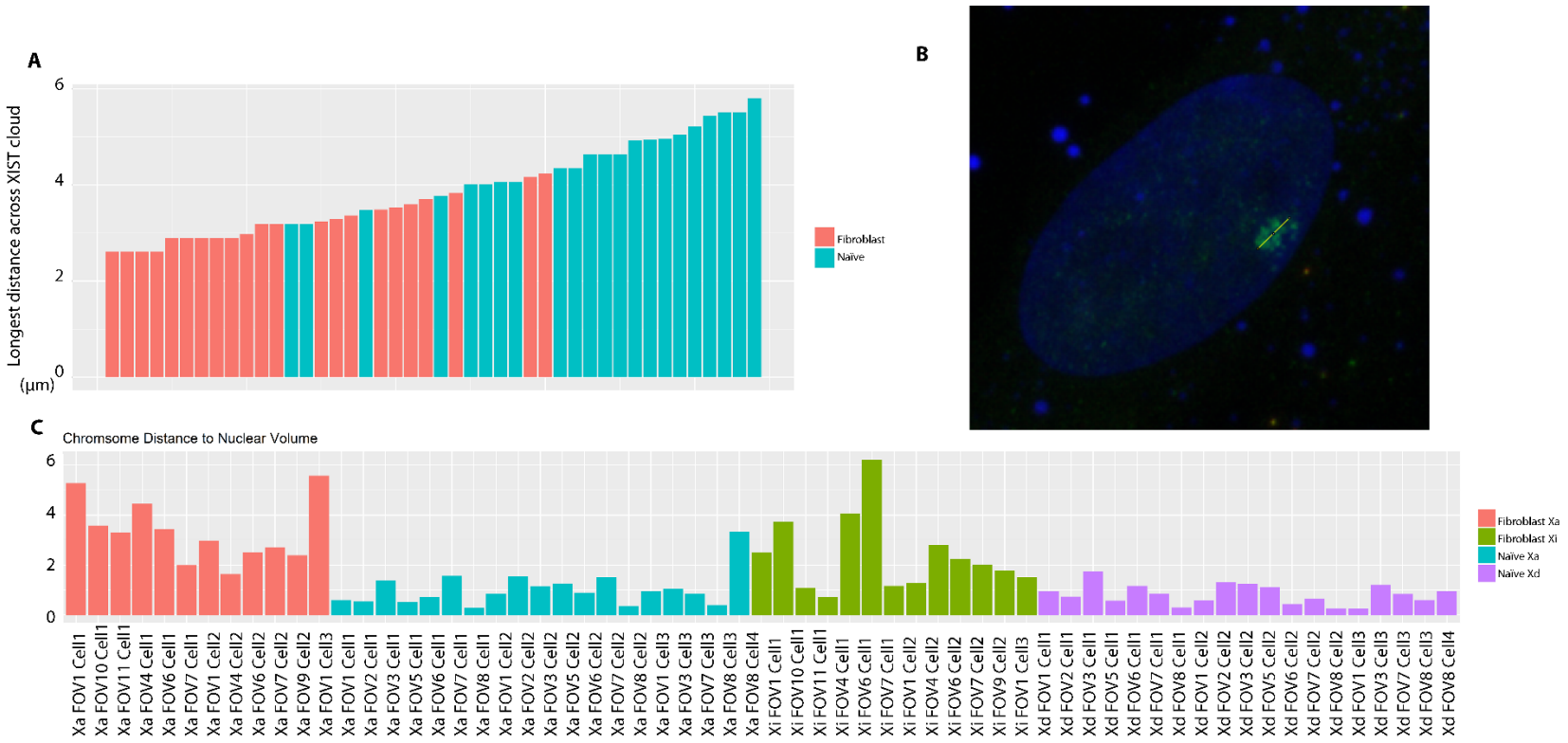


Figure 3-7 – Longest XIST cloud Distance Indicates that Naïve Xd skews Larger than Fibroblast Xi

(A) Using the measuring tool in Fiji, the longest XIST cloud vectors were recorded and then organized in a bar graph colored by cell type. (B) Largest probe distance divided by nuclei volume, measured in Fiji by masking DAPI signal, is colored by X chromosome species. (C) Longest XIST cloud vectors divided by the corresponding nuclear volume, measured by creating a mask in Fiji. Relative distances are colored by X chromosome gene expression state.

Figure 3-8 - Sequential FISH Lacks Hybridization Round Distance Biases

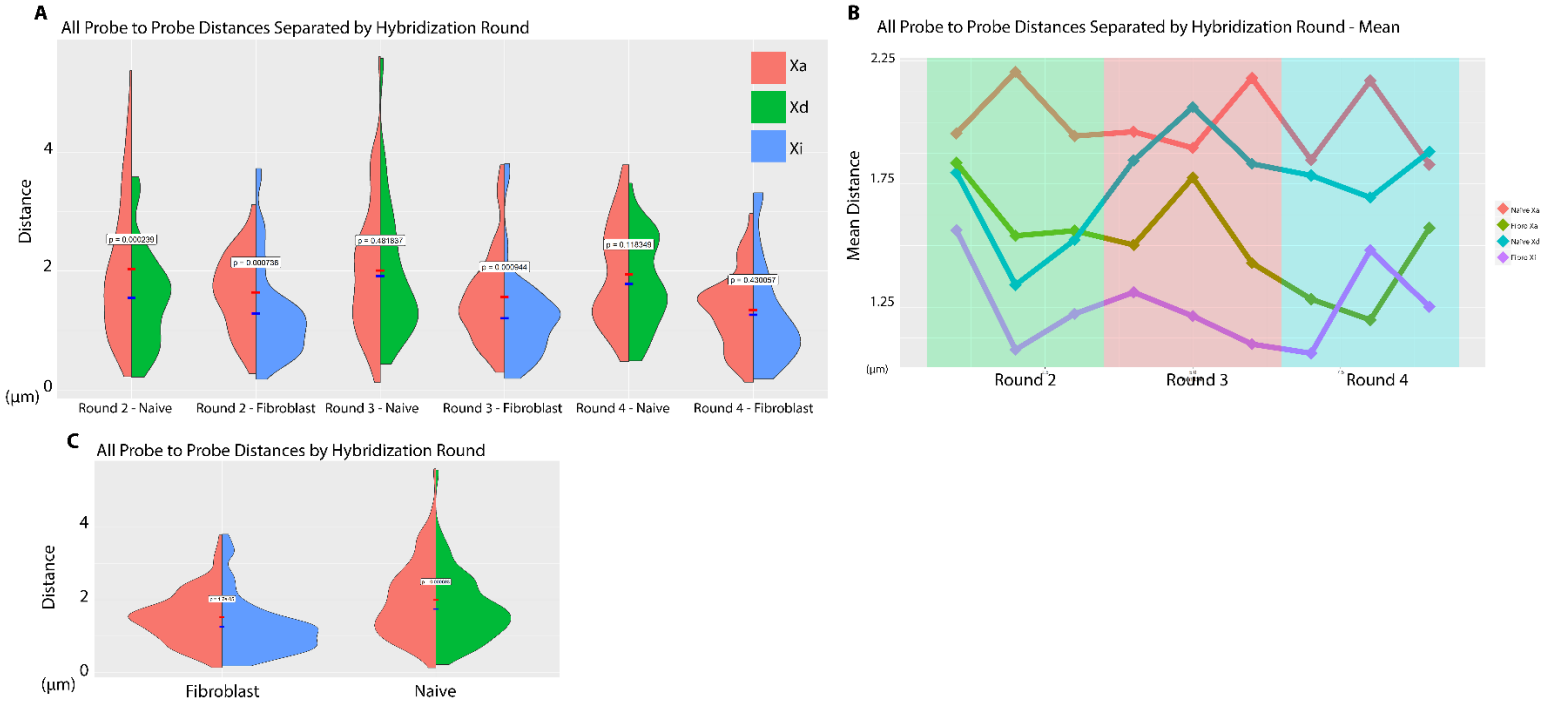


Figure 3-8 – Sequential FISH Lacks Hybridization Round Distance Biases

(A) Split violin plots showing all probe distances organized by hybridization round and cell type and colored by X chromosome gene expression state. (B) Mean probe distances, limited to the probes present within each round of hybridization, colored by X chromosome species. Round 2: THOC2, TSIX, HUWE1. Round 3: UTX, XACT, GPC3. Round4: CDKL5, PAGE1, LAMP2. (C) Split violin plots showing the probe distances contained to each hybridization round. No cross hybridization round distances are present.

Figure 3-9 - All Probe Distances by Cell Type

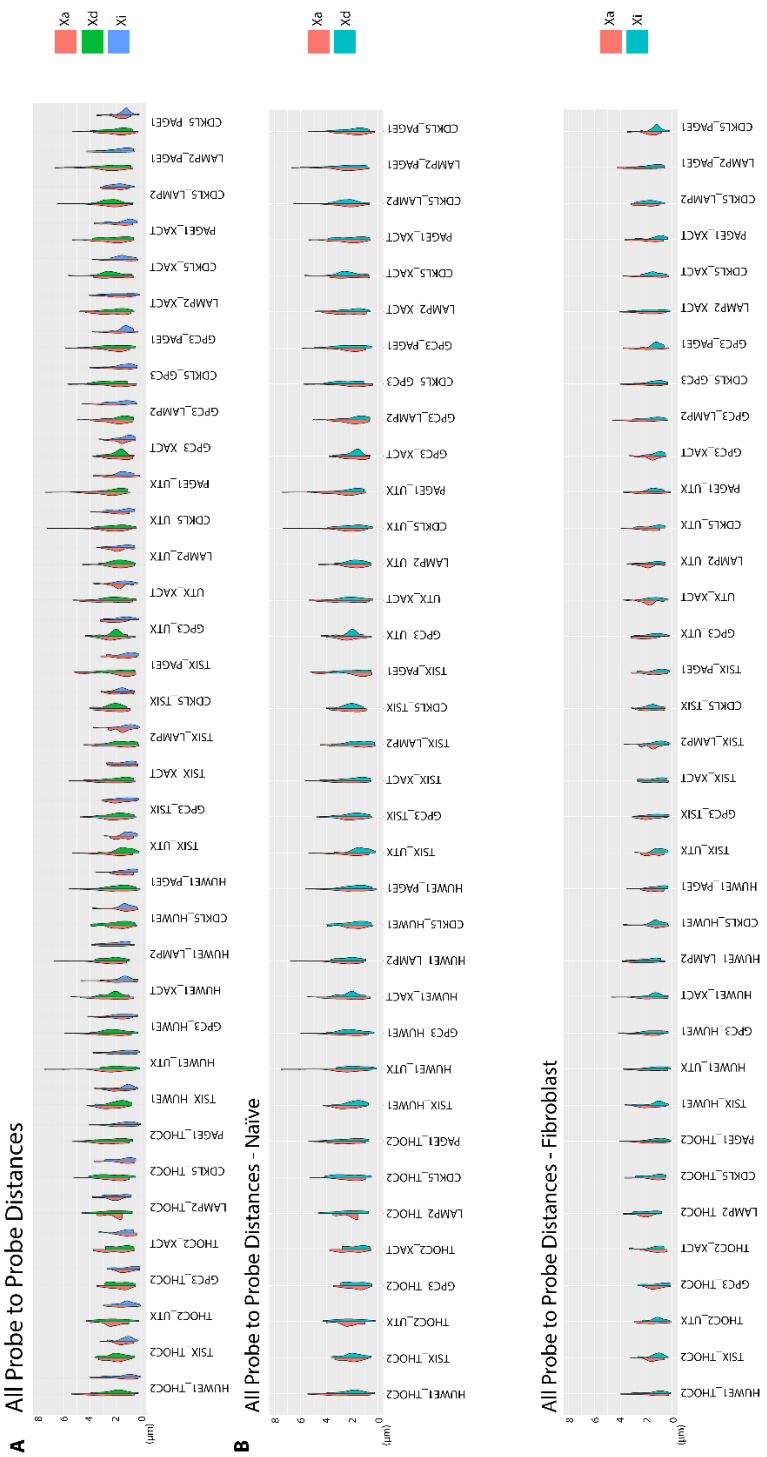


Figure 3-9 – All Probe Distances by Cell Type

(A) Split violin plots showing each probe distance, organized by cell type and colored by X chromosome gene expression state. (B) Split violin plots showing each probe distance, colored by X chromosome gene expression state. Naïve only above and fibroblast only below.

References

Bemmel, J. G., Galupa, R., Gard, C., Servant, N., Picard, C., Davies, J., . . . Heard, E. (2019). The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nature Genetics, 51*(6), 1024-1034. doi:10.1038/s41588-019-0412-0

Brinkmann, U., Vasmatzis, G., Lee, B., Yerushalmi, N., Essand, M., & Pastan, I. (1998). PAGE-1, an X chromosome-linked GAGE-like gene that is expressed in normal and neoplastic prostate, testis, and uterus. *Proceedings of the National Academy of Sciences, 95*(18), 10757-10762. doi:10.1073/pnas.95.18.10757

Carrel, L., & Brown, C. J. (2017). When the Lyon(ized chromosome) roars: Ongoing expression from an inactive X chromosome. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1733), 20160355. doi:10.1098/rstb.2016.0355

Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., . . . Cheng, H. (2012). Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Research, 41*(2), 1294-1306. doi:10.1093/nar/gks1188

Disteche, C. M., & Berletch, J. B. (2015). X-chromosome inactivation and escape. *Journal of Genetics, 94*(4), 591-599. doi:10.1007/s12041-015-0574-1

Galupa, R., & Heard, E. (2018). X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation. *Annual Review of Genetics, 52*(1), 535-566. doi:10.1146/annurev-genet-120116-024611

Jalali, Z., & Parvaz, N. (2020). Molecular evolution of autophagy rate-limiting factor LAMP2 in placental mammals. *Gene, 727*, 144231. doi:10.1016/j.gene.2019.144231

Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., . . . Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature, 503*(7475), 290-294. doi:10.1038/nature12644

Kao, S., Wu, H., & Wu, K. (2018). Ubiquitination by HUWE1 in tumorigenesis and beyond. *Journal of Biomedical Science, 25*(1). doi:10.1186/s12929-018-0470-0

Kilstrup-Nielsen, C., Rusconi, L., Montanara, P. L., Ciceri, D., Bergo, A., Bedogni, F., & Landsberger, N. (2012). What We Know and Would Like to Know about CDKL5 and Its Involvement in Epileptic Encephalopathy. *Neural Plasticity, 2012*, 1-11. doi:10.1155/2012/728267

Kumar, R., Palmer, E., Gardner, A. E., Carroll, R., Banka, S., Abdelhadi, O., . . . Gecz, J. (2020). Expanding Clinical Presentations Due to Variations in THOC2 mRNA Nuclear Export Factor. *Frontiers in Molecular Neuroscience, 13*. doi:10.3389/fnmol.2020.00012

Lee, M. G., Villa, R., Trojer, P., Norman, J., Yan, K., Reinberg, D., . . . Shiekhattar, R. (2007). Demethylation of H3K27 Regulates Polycomb Recruitment and H2A Ubiquitination. *Science, 318*(5849), 447-450. doi:10.1126/science.1149042

Meulen, J. V., Speleman, F., & Vlierberghe, P. V. (2014). The H3K27me3 demethylase UTX in normal development and disease. *Epigenetics, 9*(5), 658-668. doi:10.4161/epi.28298

Patrat, C., Ouimette, J., & Rougeulle, C. (2020). X chromosome inactivation in human development. *Development, 147*(1). doi:10.1242/dev.183095

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S., Codeluppi, S., . . . Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell, 165*(4), 1012-1026. doi:10.1016/j.cell.2016.03.023

Sado, T., Hoki, Y., & Sasaki, H. (2005). Tsix Silences Xist through Modification of Chromatin Structure. *Developmental Cell, 9*(1), 159-165. doi:10.1016/j.devcel.2005.05.015

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T. W., . . . Plath, K. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell, 20*(1), 87-101. doi:10.1016/j.stem.2016.10.006

Schmitt, A. D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C. L., Li, Y., Lin, S., Lin, Y., Barr, C. L., & Ren, B. (2016). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. Cell Reports, 17(8), 2042–2059. doi:10.1016/j.celrep.2016.10.061

Stigliano, I., Puricelli, L., Filmus, J., Sogayar, M. C., Joffé, E. B., & Peters, M. G. (2008). Glypican-3 regulates migration, adhesion and actin cytoskeleton organization in mammary tumor cells through Wnt signaling modulation. *Breast Cancer Research and Treatment, 114*(2), 251-262. doi:10.1007/s10549-008-0009-2

Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur-Griscelli, A., . . . Rougeulle, C. (2013). XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. *Nature Genetics, 45*(3), 239-241. doi:10.1038/ng.2530

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Ali, T. M., . . . Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell, 20*(1), 102-111. doi:10.1016/j.stem.2016.10.014

Chapter 4

X Chromosome Structure trends present in Gene Expression States

Abstract

By comparing 3D spatial data we collected using the developed sequential FISH technique, we can evaluate population of cells from both human naive and fibroblast cells to understand trends better and determine not just interesting instances in chromosomes but potentially gene expression state identity defining traits. We found that the active X chromosomes in both cell types tend to be predictably larger than their counterpart. Additionally, the dampened X chromosome showed that probe distances of regions downstream of DXZ4 and genes that escape XCD are of comparable or larger sizes when compared to the naive active X. We also find that the dampened X chromosome adopts its own unique structural features and correlative to its gene expression, is more open than an inactive X but less open than an active X.

Introduction

Single cell analysis of cells holds powerful information, such as being able to identify individual quirks or uncommon characteristics on a cell by cell basis. This style of data acquisition has become more popular in recent times with the rise of technologies like highly parallel microfluidics based single-cell sequencing methods (Macosko et al., 2015). However, when attempting to parse out more widespread trends, applying population focused analysis to single cell data can produce quality information. By utilizing the single cell sequential FISH data in a population-based approach, we hope to find trends amongst the different gene expression states of the X chromosome that help to define their structural identities.

The general structure of the inactive X chromosome, conserved across multiple eutherian mammals including both human and mouse, is a condensed, dumbbell-like shape (Darrow et. al, 2016; Deng et al., 2015; Giorgetti et al., 2016; Minajigi et al., 2015; Rao et al., 2014). This

structure is apparent in context of high-resolution DNA FISH (Giorgetti et al., 2016; Wang et al., 2016). In comparison the active X chromosome is known to have a larger, more open configuration (Jégu et al., 2017).

XCI, which silences a majority of the genes on the X chromosome, initiates in cis via the expression of XIST from the X chromosome initiation center. The prevailing thought being that XIST spreads to nearby regions in spatial proximity (Pandya-Jones and Plath 2016). The spreading of XIST across the chromosome may then play an integral role in influencing how the inactive X chromosome is formed and the final structure of the chromosome. Genes that escape XCI are believed to congregate towards the outer edges, away from the highly heterochromatic core of the inactive X chromosome (Chow et al., 2010; Clemson et al., 2006). Therefore, genes that undergo inactivation, in our case: CDKL5, HUWE1, LAMP2, THOC2, and GPC3, will then most likely be directed towards the center of the X chromosome, away from the more accessible exterior regions.

Unlike XCI, the process by which XCD occurs is less understood and therefore the confidence in predicting the locations of dampened or escape genes is lacking. In both cases, XIST RNA is present and plays a role in overall gene expression of the chromosome (Sahakyan et al., 2017; Vallot et al., 2017). While assumptions can be made about the structure if considering the connection between gene expression and the spatial organization of chromosomes, these predictions are nonetheless conjecture. A feature unique to human naïve cells is the greater enrichment of XIST concentrated in the region of the X chromosome downstream of DXZ4 relative to that of fibroblast cells. This oddity may prove to be a significant contributor to the structure of the dampened X chromosome due to the known association of XIST and the condensed structure of the inactive X chromosome (Jégu et al.,

2017).

Whilst analyzing the spatial data of the X chromosomes and finding general trends in chromosome structure, there will undoubtedly become apparent that some samples may show deviations from the norm or may appear to be outliers. In order to determine the validity of the samples, there may be key, identifiable features that regardless of the variance of chromosome to chromosome may still shine through, even in a blind call scenario. This investigation may in return provide more X chromosome gene expression state specific features.

Results

To first ascertain a better understanding of the average X chromosome of each gene expression state, we took the mean and median of every probe to probe distance in order to create mean spatial distance maps (Fig 4-1). The mean and median maps essentially appear the same, so going forward we will focus on the mean representations. For the most part, the observations form the single X chromosome spatial maps still ring true here. Probably the most obvious trend is just how uniformly small the probe distances are in the fibroblast inactive X chromosome (Fig 4-1). The mean naïve dampened X and fibroblast active X chromosomes do have spikes of larger probe distances at the near and far end of the chromosome respectively (Fig 4-1). In comparison, the naïve active X chromosome seems to lack these hotspots and happens to have larger probe distances throughout the chromosome with little patterning apparent (Fig 4-1).

We sought to more accurately compare the mean spatial distance maps of each gene expression state to one another. To accomplish this, we subtracted all variants of the mean probe distances (Fig 4-2). Of note is the “Xa Fibrobast – Xi Fibroblast” difference map resembles that of the “Xa Naïve – Xd Naïve” difference map, with the non-active X having generally smaller

probe distances aside from distances involving genes at the start of the chromosome (Fig 4-2). This indicates a similar yet unique relationship of the complementary X chromosomes in different cell types. While the "Xa Naïve – Xi Fibroblast" difference map is nearly solid blue, indicating that the active X is larger in every distance, the "Xd Naïve – Xa Fibroblast" difference map is a mostly pale to white color, indicating very similar values in many of the probe distances (Fig 4-2). To further study the mean probe distances, we decided to take advantage of the probe scheme and target genes involved in specific changes in gene transcription.

By removing genes except for those that are targeted for inactivation during XCI, we can assess the structural dynamics of these regions of the chromosome in their fundamental inactive state as well as the other gene expression states (Fig 4-3). The dampened X chromosome stands out by harboring long-range gene connections that have long probe distances but relatively smaller probe distances in the context of the last three targeted genes: LMAP2, THOC2, GPC3 (Fig 4-3a). This suggests a more elongated structure with the end portion of the chromosome more compact than either of the two active chromosomes. When comparing the XCI inactive probe distances in a hybridization round dependent manner, it can seen that the dampened X has the most gentle of distribution curves, lacking a distance predisposition when compared to the other gene expression states (Fig 4-3b).

We chose to investigate the arrangement of XCD escape genes within the X chromosome in order to ascertain whether this group behaved in an interesting manner in the context of the dampened X chromosome in comparison to the other gene expression states (Fig 4-4). We found that while a clear difference in distances can be found in fibroblast, the distances in naïve are striking similar (Fig 4-4a). Even the distribution looks nearly mirrored in the naïve violin plot (Fig 4-4b). The XCD escapers essentially share chromosome structure in both the dampened and

active X chromosomes.

We found another surprising trend in the spatial distances of the dampened X chromosome. When taking into account the probe distances specific to genes within the region of the chromosome downstream of DXZ4, we found that the distances tended to be not similar but larger than that of the complementary naïve active X chromosome in half of the cells (Fig 4-5b). Based on data from another lab member, Iris Dror, we know that XIST RNA enriches in the region downstream of DXZ4. Based on the previous observations in the XCI and XCD gene groups as well as this single cell comparison specifically looking at the DXZ4 region, it is likely that XIST RNA manages to have a condensing structural effect on the region downstream of DXZ4 due to XIST's expression in the dampened X chromosome. However, the data also seems to imply that either because of the lack of relative XIST coverage on the rest of the chromosome or some intrinsic structural trait of DXZ4, this hinge region is not akin to a closed door in the inactive X but open being able to be even larger than the naïve active X chromosome.

We considered that some of the single cell variations among the probe distances do not agree with the mean distance data. In order to certify whether the cell to cell variations of distance are still within the bounds of the greater collection of chromosomes of the same gene expression state we attempted a variety of methods to blindly call X chromosome identity, knowing only the cell type and probe distance of the two X chromosomes, to match the known identities due to the XIST RNA FISH (Fig 4-6). Based on Binary Call (percentage of probe distances that are larger than the complementary X chromosomes), Standard Deviation Call (matching an individual probe distance to each of the gene expression state's mean probe distances with the lowest standard deviations), and TOP (the mean highest value probe distances), we managed to calculate a composite X Chromosome Call Confidence score (Fig 4-

6a). 100% of fibroblast cells and 96% of naïve cells are able to be called correctly using the blind call method. By just looking at Binary or Standard Deviation Call most fibroblast cells can be properly called (Fig 4-6b). It is likely that the dampened X and active X chromosome share a good portion of their structure and need the extra inputs, such as TOP distances, likely to help call the exceedingly large active X chromosome, to identify the chromosomes correctly (Fig 4-6b).

Based on the data we have collected we are proposing the general structures of the X chromosome expression states (Fig 4-7). The well documented inactive X chromosome is relatively small and condensed. Both active X chromosomes are much larger and more open with the naïve active X chromosome being specifically plastic and housing regions that extend far from the rest of the chromosome. The dampened X chromosome, similar in size to the naïve active X chromosome harbors a slightly condensed region within the region downstream of DXZ4 and a widely fairly open, and quite distal region preceding DXZ4 (Fig 4-7).

Discussion

In both human fibroblast and naïve cells, we found that the active X chromosome skewed towards larger intrachromosomal distances as well as a larger overall structure. As expected, the inactive X in fibroblast cells shows the smallest overall intrachromosomal distances. The active X in naïve cells appeared to not only be the largest of X chromosomes overall but had the largest individual probe distances, solidifying the structure as the most open of the X chromosome gene expression states in the assay. We found that comparing the mean dampened X chromosome structure showed many similarities to the active X chromosome of fibroblast cells. This combination evidence speaks to the more open nature of the dampened X chromosome due to its active gene expression but also indicates that the size of the X chromosomes fluctuates to an extent depending on cell type. To further illustrate further divergence of the dampened X chromosome structure from that of the active or inactive, one only needs to look at the specific gene groups based on expression state.

We found that the genes that are inactivated during XCI in the dampened X chromosome show a wide distribution of probe distances with a smaller, more condensed distances appearing towards the end of the chromosome. This is in direct contrast to the XCD escape genes, where in the dampened x chromosome these probe distances mirror those of the spatially open, complementary active X chromosome. Additionally, we created a method to confidently call the identity of chromosomes within a cell. Based on this data and the single cell comparison of the probes downstream of DXZ4, we believe we have a more accurate picture of the structure of the dampened X chromosome.

Materials & Methods

X Chromosome Confidence Call

Binary Distance Probe Score

Compares all unique probe distances (36) and assigns a “1” or “0” based on if the distance is larger or smaller than the other X. Sum these points to produce the Binary Distance Probe Score.

Standard Deviation Score

Take a set of the probe distances with the lowest Standard Deviations. Then compare each X to the mean values of these distances. Score each X’s distance as “Xa” or “Xi/Xd”. A Double positive indicates that the Xa is scored as “Xa” and the Xi/Xd is scored as “Xi/Xd”. A Single positive indicates that one of the X’s is scored correctly and the other X’s identity can be derived based on the positive X’s higher score. A Single positive is scored as a “2/3” (based on the average Binary Distance Probe Score). A Double positive is scored as a “1”.

TOP Score

Take a small set of probe distances containing the largest and smallest on average distances (the far ends of the distance spectrum). Assign a “1” or “0” based on which X is closer the mean distance

Figures and Legends

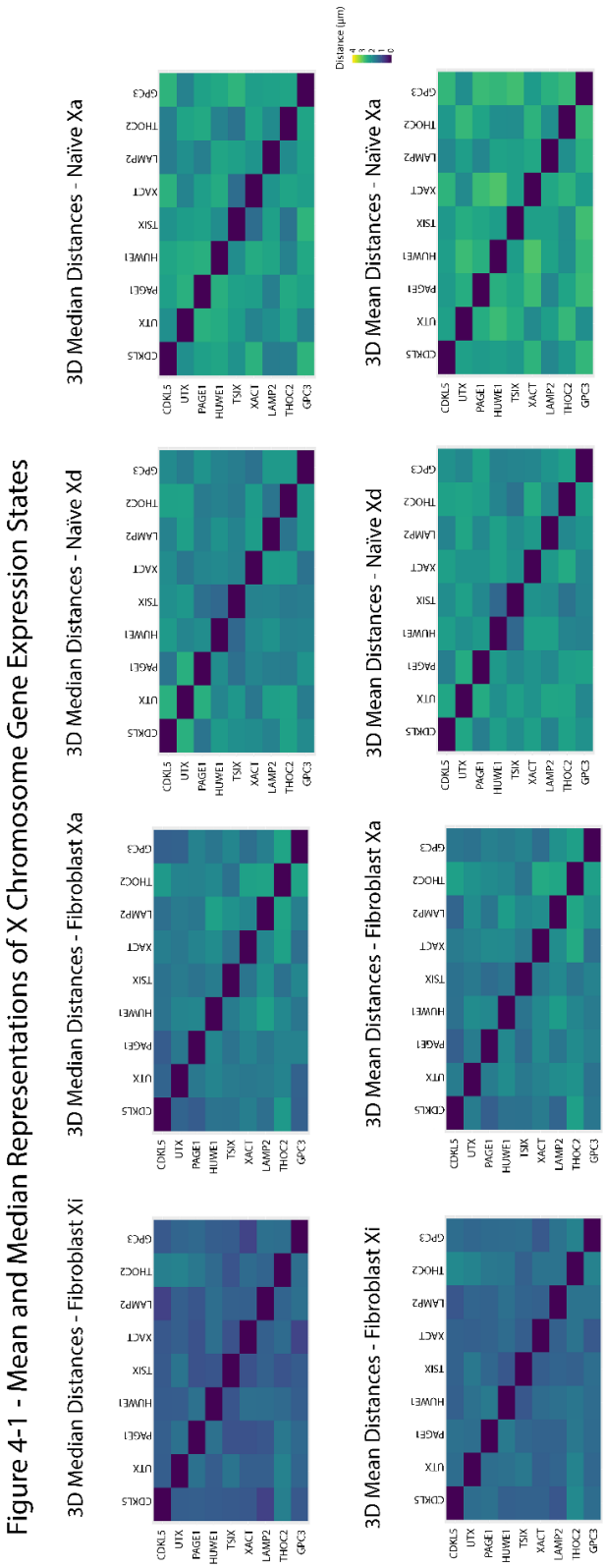


Figure 4-1 - Mean and Median Representations of X Chromosome Gene Expression States

Figure 4-1 - Mean and Median Representations of X Chromosome Gene Expression States

Mean and median mean spatial distance maps of each X Chromosome species, displayed using the ggplot2 function of R. Probes are found in each row and column and distances between the two probes can be found by locating the intersection of the row and column of the desired pair of probes. Distances are colored by magnitude and are arranged in a matrix based on head to tail organization of gene locations found on the X chromosome. Each row indicates a single cell with the active X chromosome in the right column and the non-active in the left column.

Figure 4-2 - Mean Spatial Distance Comparison

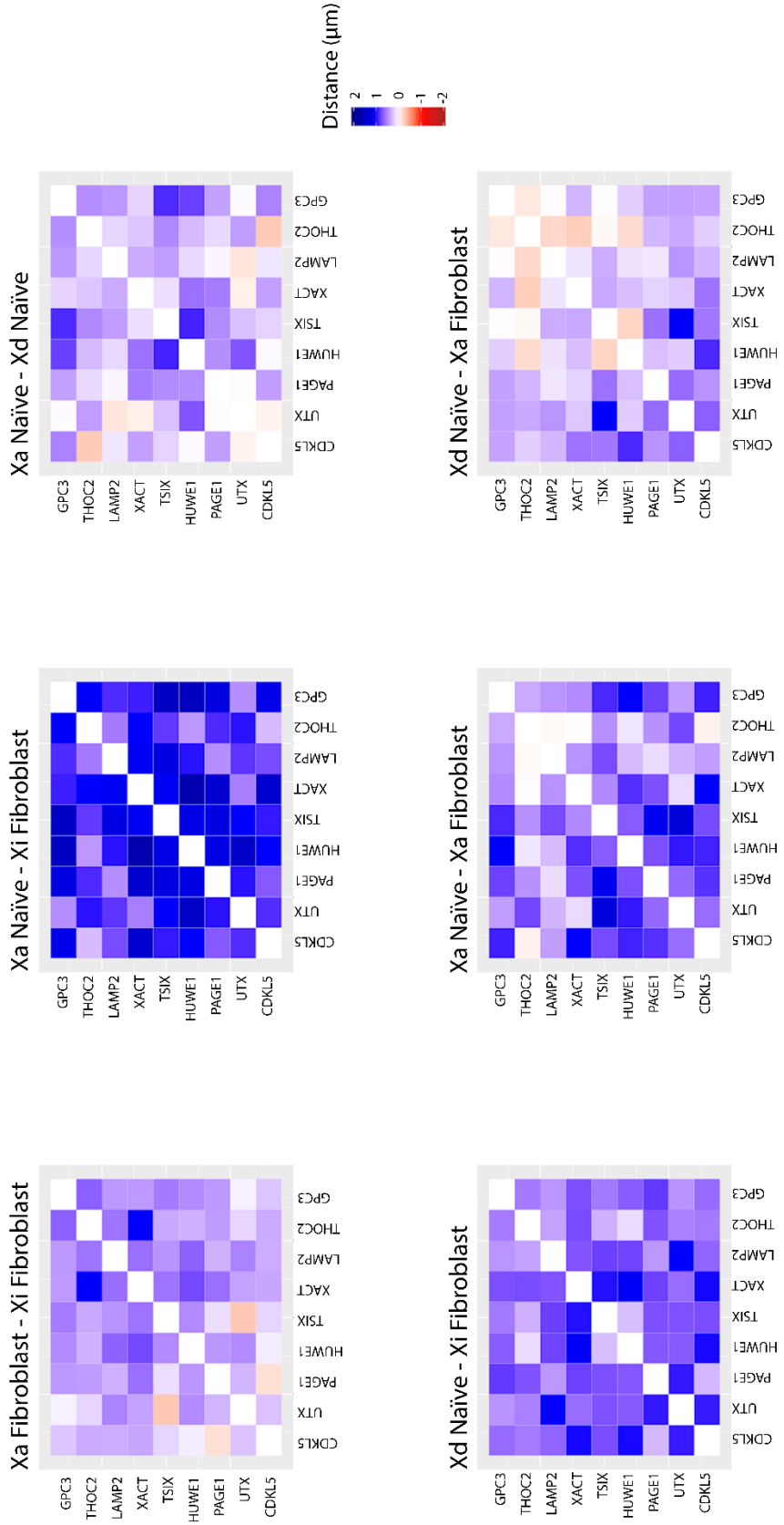


Figure 4-2 – Mean Spatial Distance Comparison

Difference of mean spatial maps. Color of tiles displays the average difference of probe distance between the two X chromosome species. The bluer the tile, the larger the relative mean probe distance of the first listed X chromosome gene expression state. The redder the tile, the larger the relative mean probe distance of the second listed X chromosome gene expression state.

Figure 4-3 - XCI Inactive Genes

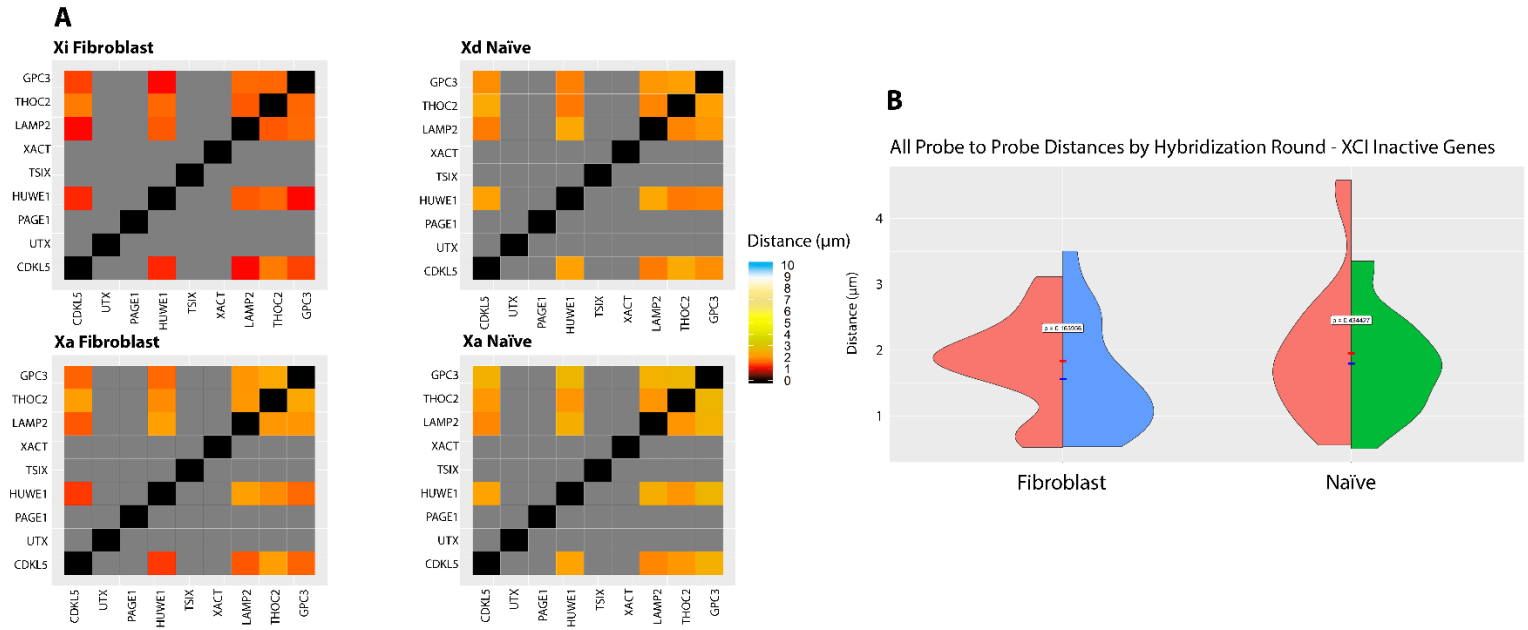


Figure 4-3 - XCI Inactive Genes

(A) Mean spatial distance map containing only XCI inactive genes in each cell type. Grey indicates that both probes are not XCI inactive genes. Magnitude of probe distance is indicated by color. (B) Split violins showcasing probe distances of XCI inactive genes that fall within the same hybridization round, separated by cell type and colored by X chromosome gene expression state.

Figure 4-4 - XCD Escape Genes

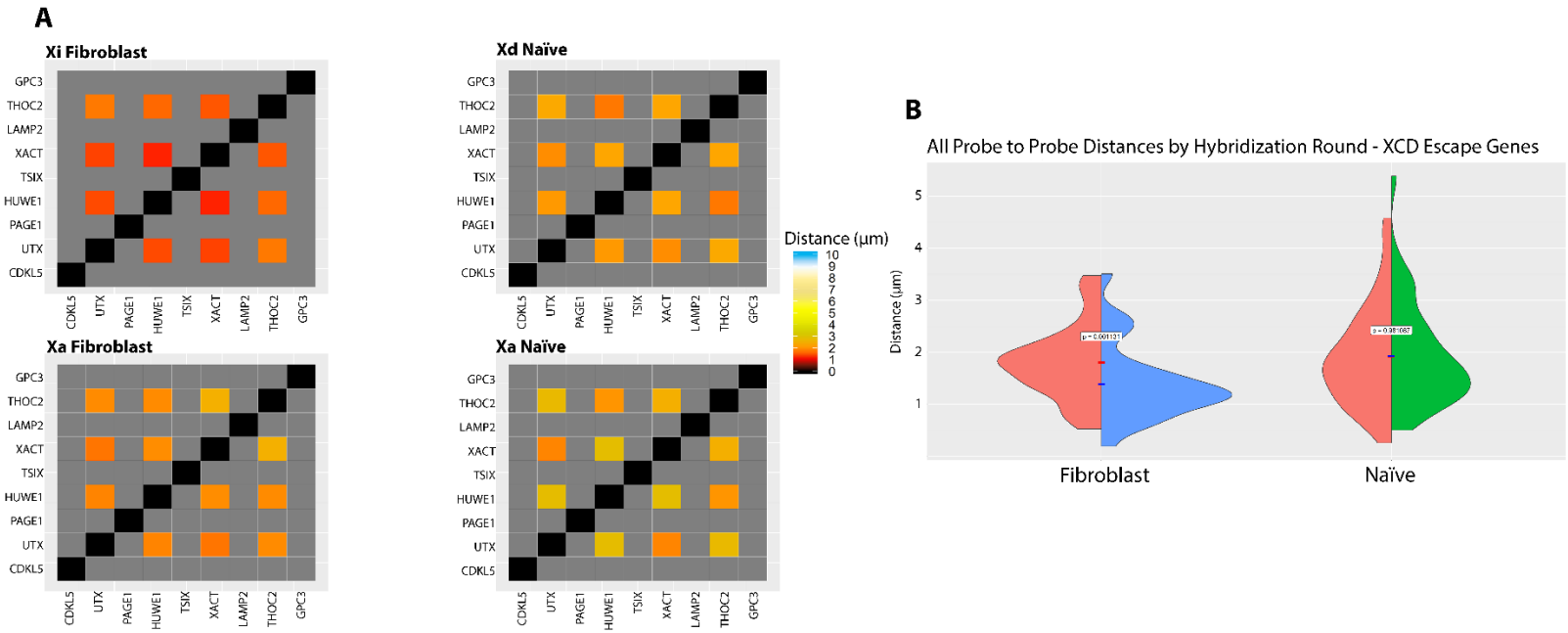
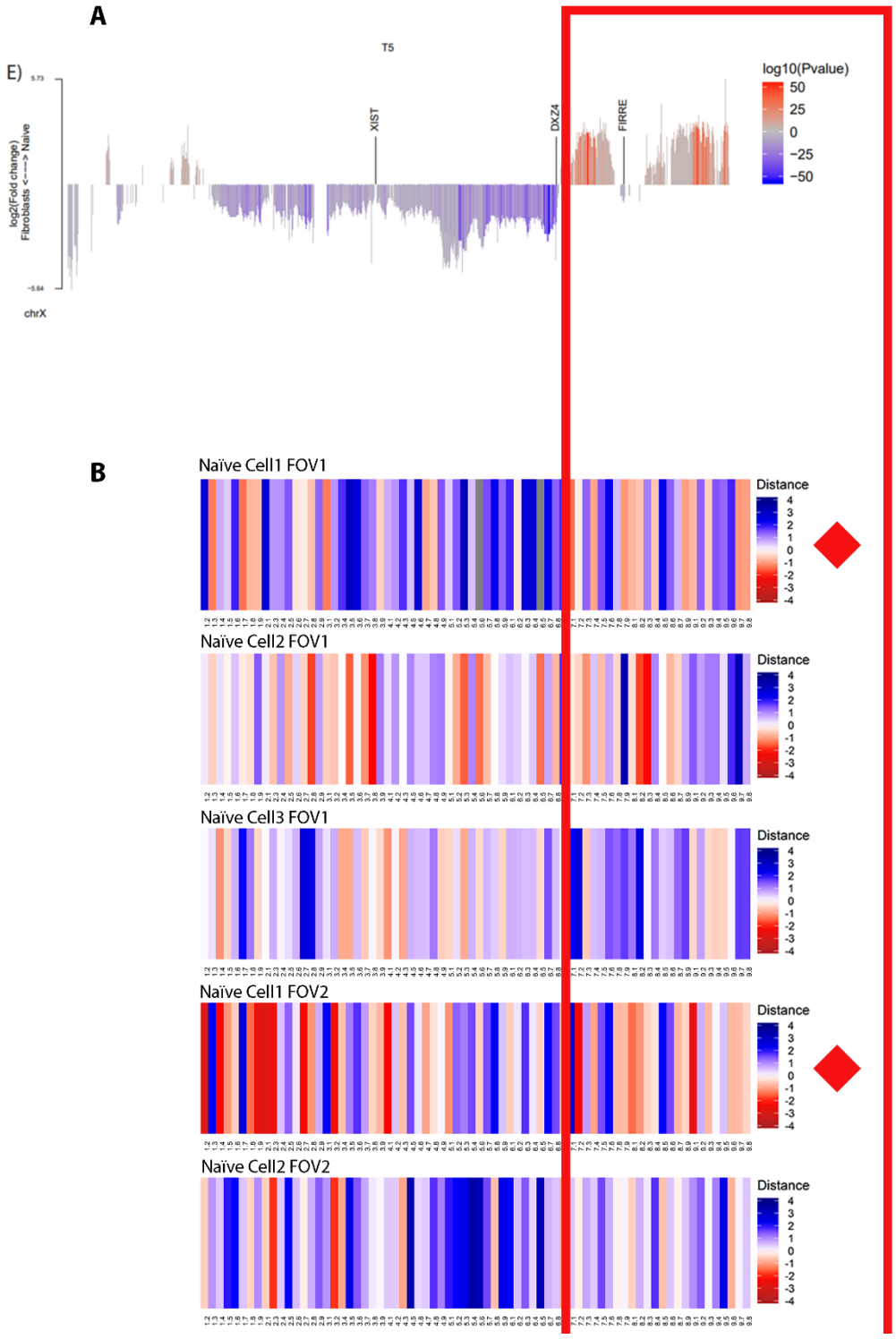
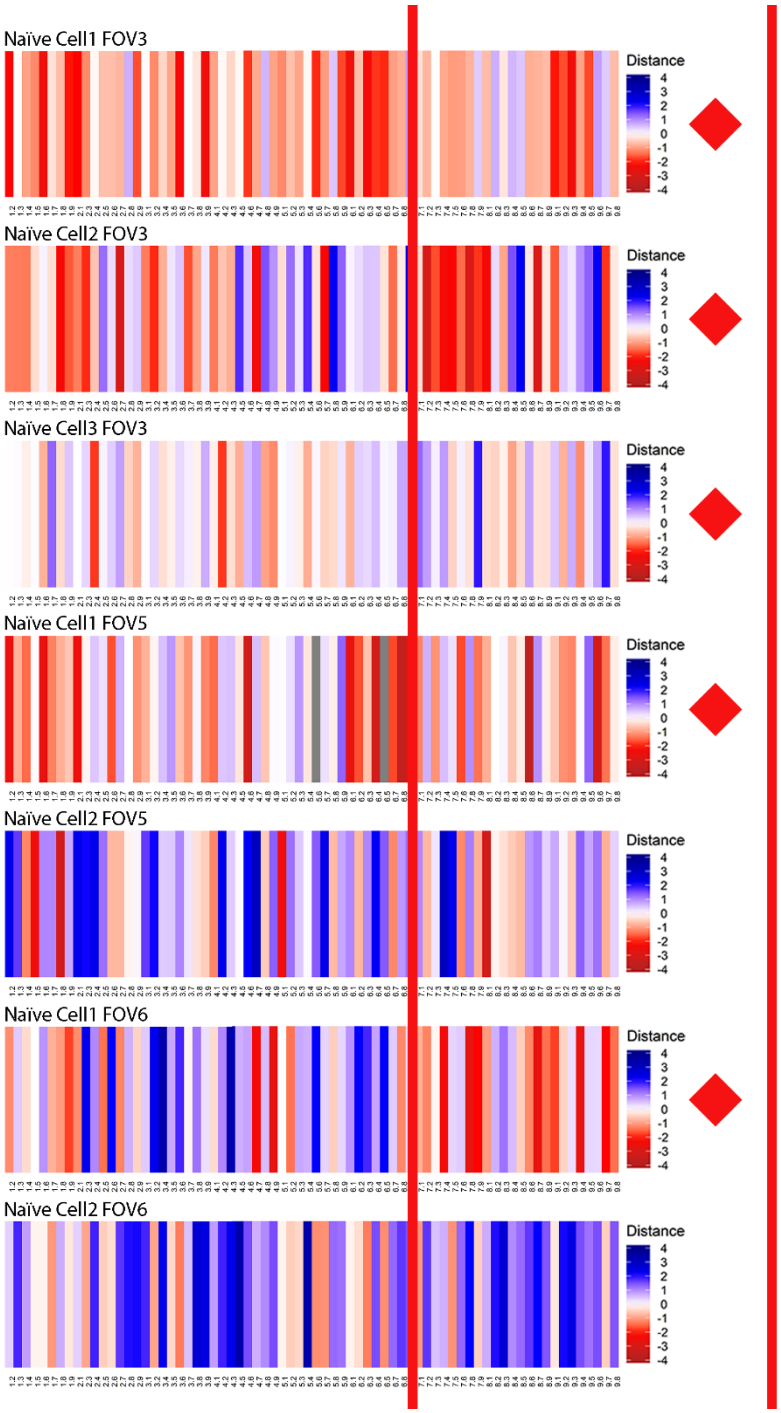


Figure 4-4 - XCD Escape Genes

(A) Mean spatial distance map containing only XCD escape genes in each cell type. Grey indicates that both probes are not XCD escape genes. Magnitude of probe distance is indicated by color. (B) Split violins showcasing probe distances of XCD escape genes that fall within the same hybridization round, separated by cell type and colored by X chromosome gene expression state.

Figure 4-5 - Genes Downstream of DXZ4




Naïve Cell1 FOV3
Distance
4
3
2
1
0
-1
-2
-3
-4
Naïve Cell2 FOV3
Distance
4
3
2
1
0
-1
-2
-3
-4
Naïve Cell3 FOV3
Distance
4
3
2
1
0
-1
-2
-3
-4
Naïve Cell1 FOV5
Distance
4
3
2
1
0
-1
-2
-3
-4
Naïve Cell2 FOV5
Distance
4
3
2
1
0
-1
-2
-3
-4
Naïve Cell1 FOV6
Distance
4
3
2
1
0
-1
-2
-3
-4
Naïve Cell2 FOV6
Distance
4
3
2
1
0
-1
-2
-3
-4

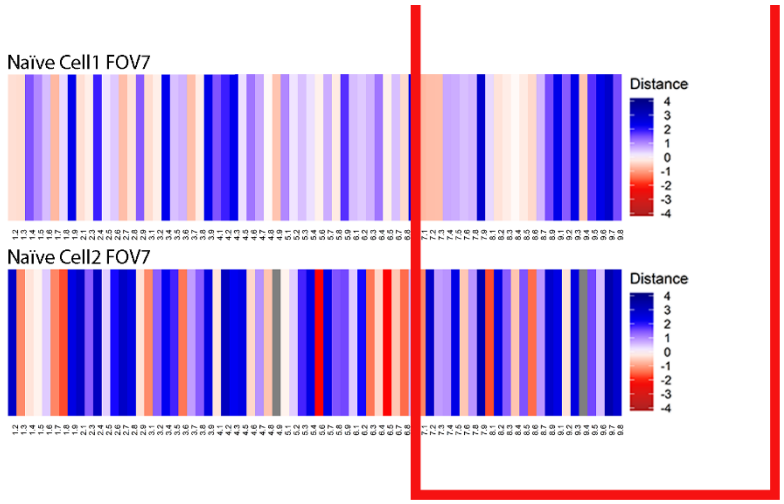


Figure 4-5 - Genes Downstream of DXZ4

(A) XIST RNA enrichment map of the X Chromosome in different cell types provided by Iris Dror. (B) Linearized spatial distance maps of individual naïve cells demonstrating probe distance difference in Xa vs Xd. Probes distances are labeled as such: "probe number on linear chromosome[.] probe number on linear chromosome (See Fig 3-2)" (Ex. "1.2" is CDKL5 to UTX). Probe distances are organized in ascending order within the probe and across the X chromosome (Ex. "1.2". "1.3", "1.4" … "2.1". "2.3", "2.4" … "3.1". "3.2", "3.4"). Diamonds indicate that in the region downstream of DXZ4, probe distances are found to be longer in Xd. Genes within the region downstream of DXZ4 is highlighted by the red box.

Figure 4-6 - Blind Chromosome Call Identifies Chromosome Gene Expression State on Population Characteristics

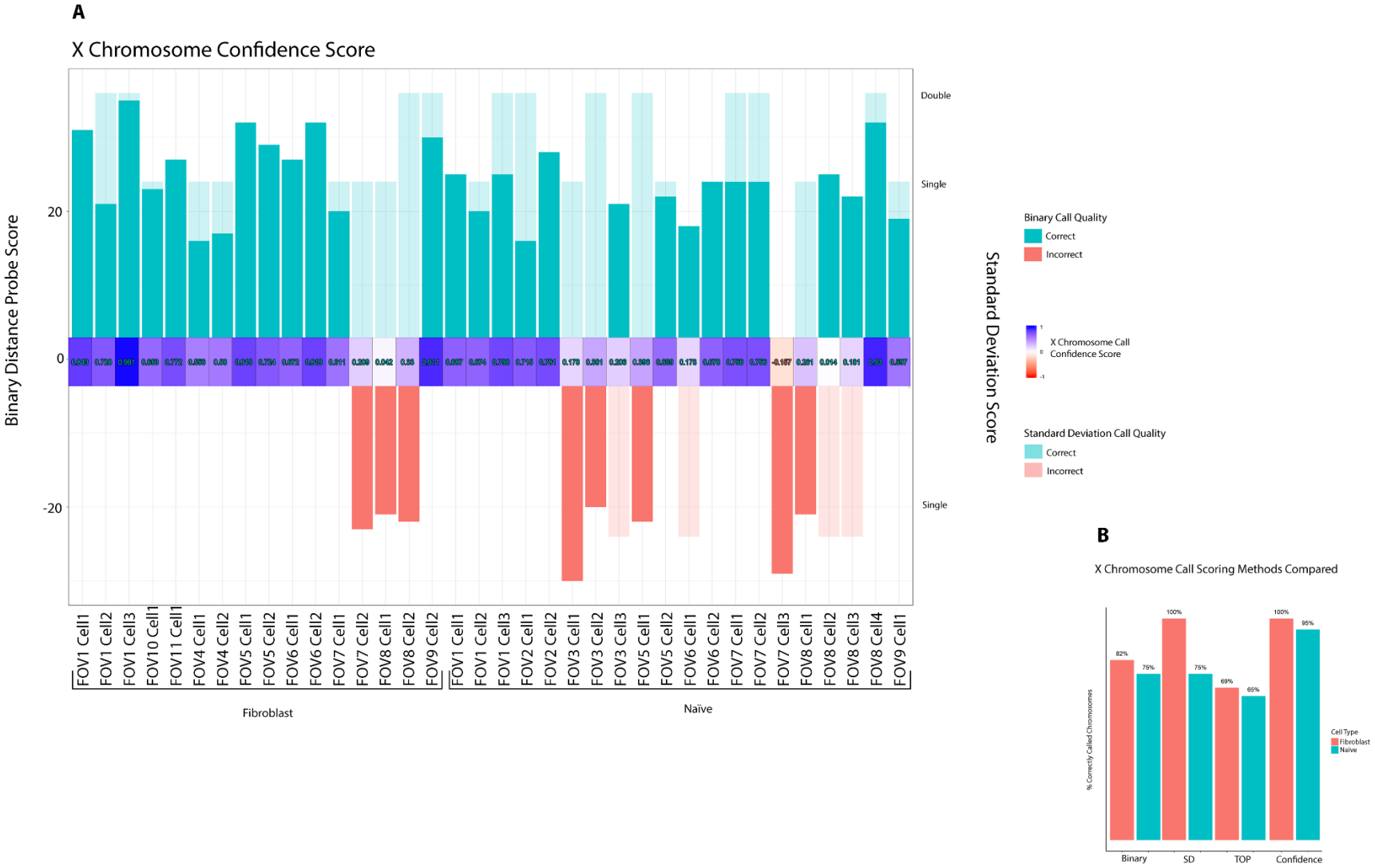


Figure 4-6 - Blind Chromosome Call Identifies Chromosome Gene Expression State on Population Characteristics

(A) Mixed bar plot indicating ability to call chromosome identity correctly based on the Binary Call, Standard Deviation Call, and the combined Confidence Score. Solid bars indicate the ratio of the larger distances per cell (Binary Call). The color of the solid bars indicates whether or not the majority of the larger distances belongs to the active X (Correct) or the non-active X chromosome (Incorrect). Transparent bars indicate if a one (Single) or both (Double) X chromosomes from the same cell can be identified based on comparing their distances to the mean distances with the lowest amount of deviation (Standard Deviation Call). The X Chromosome Call Confidence Score is represented in a numeric fashion with the color of the number indicating a correct or incorrect call (teal or red). The number represents the confidence of the call of the identity of the chromosomes (the further from zero the higher the confidence). The color of the tiles reflects the combination of the degree of confidence in addition to the accuracy of the call (the bluer indicating higher confidence and higher accuracy, the redder indicating high confidence and low accuracy) based on a combination of Binary, Standard Deviation and TOP Calls. (B) Bar plot indicating the effectiveness of each of the various methods used to call chromosome identity as well as the combination of the three, split and colored by cell type.

Figure 4-7 - X Chromosome Structure by Gene Activation State Model

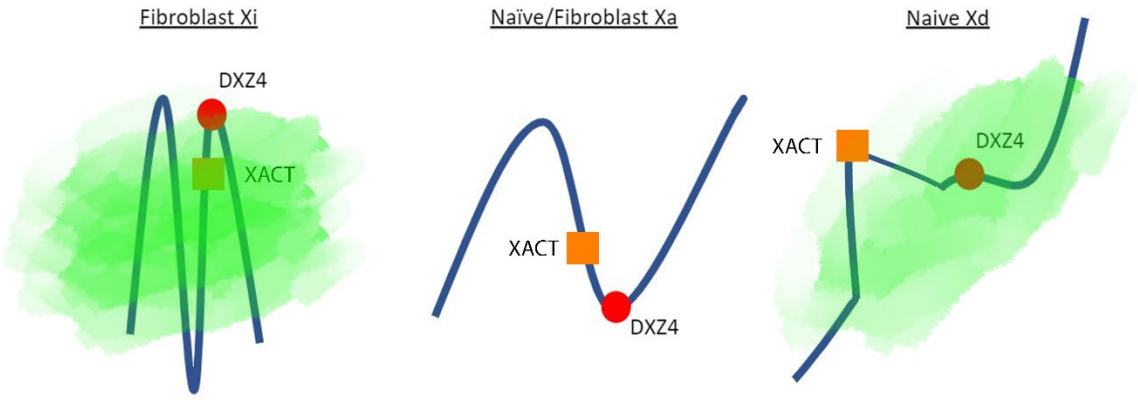


Figure 4-7 – X Chromosome Structure by Gene Activation State Model

Artistic interpretations of the inactive, active, and dampened X chromosomes. DXZ4 is marked by a red dot, XACT gene location marked by an orange diamond, and the XIST RNA denoted by the green cloud. Naïve Xd shows long distances between region downstream of DXZ4 and across cell, small distances withing region downstream of DXZ4, and XACT gene locus separate from the rest of the chromosome body.

References

Chow, J. C., Ciaudo, C., Fazzari, M. J., Mise, N., Servant, N., Glass, J. L., . . . Heard, E. (2010). LINE-1 Activity in Facultative Heterochromatin Formation during X Chromosome Inactivation. *Cell, 141*(6), 956-969. doi:10.1016/j.cell.2010.04.042

Clemson, C. M., Hall, L. L., Byron, M., Mcneil, J., & Lawrence, J. B. (2006). The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proceedings of the National Academy of Sciences, 103*(20), 7688-7693. doi:10.1073/pnas.0601069103

Darrow, E. M., Huntley, M. H., Dudchenko, O., Stamenova, E. K., Durand, N. C., Sun, Z., . . . Aiden, E. L. (2016). Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. *Proceedings of the National Academy of Sciences, 113*(31). doi:10.1073/pnas.1609643113

Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., . . . Disteche, C. M. (2015). Bipartite structure of the inactive mouse X chromosome. *Genome Biology, 16*(1). doi:10.1186/s13059-015-0728-8

Giorgetti, L., Lajoie, B. R., Carter, A. C., Attia, M., Zhan, Y., Xu, J., . . . Dekker, J. (2016). Structural organization of the inactive X chromosome in the mouse. *Nature, 535*(7613), 575-579. doi:10.1038/nature18589

Jégu, T., Aeby, E., & Lee, J. T. (2017). The X chromosome in space. *Nature Reviews Genetics, 18*(6), 377-389. doi:10.1038/nrg.2017.17

Macosko, E., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., . . . Mccarroll, S. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell, 161*(5), 1202-1214. doi:10.1016/j.cell.2015.05.002

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., . . . Lee, J. T. (2015). A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science, 349*(6245). doi:10.1126/science.aab2276

Pandya-Jones, A., & Plath, K. (2016). The "lnc" between 3D chromatin structure and X chromosome inactivation. *Seminars in Cell & Developmental Biology, 56*, 35-47. doi:10.1016/j.semcdb.2016.04.002

Rao, S., Huntley, M., Durand, N., Stamenova, E., Bochkov, I., Robinson, J., . . . Aiden, E. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell, 159*(7), 1665-1680. doi:10.1016/j.cell.2014.11.021

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T. W., . . . Plath, K. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell, 20*(1), 87-101. doi:10.1016/j.stem.2016.10.006

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Ali, T. M. L., … Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell*, *20*(1), 102–111. doi: 10.1016/j.stem.2016.10.014

Wang, S., Su, J., Beliveau, B. J., Bintu, B., Moffitt, J. R., Wu, C., & Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. *Science, 353*(6299), 598-602. doi:10.1126/science.aaf8084

Chapter 5

Conclusion

Concluding Remarks

The X chromosome fulfills a unique niche in not only being a sex chromosome but one that undergoes a distinct structural change related to gene expression over the natural course of development (Lyon 1961; Dixon et al., 2015). In the case of most eutherian mammals there are two known distinct gene expression states, the active and inactive X chromosomes (Patrat et al., 2002). The active and inactive X chromosomes have visually obvious and well documented discrete chromosome structures matching their overall active and inactive gene expression profile, respectively (Disteche and Berletch 2015; Carrel and Brown 2017; Galupa and Heard 2018). The active X with its prolific gene expression profile has a notably open configuration with a larger structure overall. Conversely, the inactive X chromosome is known for its smaller, more condensed structure harboring mostly silenced genes. Importantly, this difference in gene expression states has been captured using next generation sequencing approaches as well as microscopy (Jin et al., 2013; Cremer and Cremer 2001). While this general X chromosome paradigm in eutherian mammals is known, specific to human is another gene expression state with a previously uncharacterized chromosome structure, the dampened X chromosome (Sahakyan et al., 2017; Vallot et al., 2017). Data concerning the gene expression of this X chromosome state describes a pattern in which many genes have a reduction or dampening gene expression with some genes escaping this phenomenon (Petropoulos et al., 2016). We expected that due to the unique gene expression state, a corresponding novel chromosome structure will also exist.

To test this idea, we sought to develop a technique that would allow for the visualization of the X chromosomes on a single cell level that is accessible to an ordinary biology focused laboratory. Based on previously published data in which chromosome structure was interrogated using a sequential FISH hybridization technique, we developed our own variant to fit our

specific needs: minimal hybridization events, practical setup, and informative, modular design of probe targets (Wang et al., 2016). By combining image tiling of the sample well plus the addition of multi-channel microspheres, we were able to create a method that utilized image relocation and alignment between rounds in order to compensate for lacking a fixed microscope and fluidics device. Our current protocol allows for up to three fluorescent signals per round for a total of four rounds of probe hybridization with the capacity for further hybridization rounds.

Using this technique, we acquired X chromosome structure information from 36 individual cells: 20 naïve and 16 fibroblast. We then recreated the 3D structures based on the coordinates of the gene targets. We saw that the there was a notable difference in the size of the X chromosome in fibroblasts (the Xi and Xa), whereas the X chromosomes in naïve hESCs (the Xd and Xa) appeared much closer in size, although both are larger than the fibroblast X chromosomes. Additionally, we calculated the intrachromosomal distances of the probes covering the X chromosome length in order to understand the finer details of the chromosome organization in a way which can be compared from chromosome to chromosome and across cell types. We identified probe distances that consistently stood out in distance in individual samples, like XACT, which produces a lncRNA that coats the active X chromosome in humans, showing greater separation from the other probes in multiple dampened X chromosomes (Vallot et al., 2013).

By comparing the mean spatial intrachromosomal distances of each of the X chromosome gene expression states, we discerned significant structural features and relationships between the different X chromosome species. We also specifically looked at how genes specific to changes in gene expression related to chromosome structure. For this purpose, we honed in on XCI and XCD escaping genes. We found a general resemblance of the dampened X chromosome to both

active X chromosomes. XCD escape genes (on the Xd) were as open as on the active X chromosome, while the XCI escape genes on the dampened X chromosome saw smaller distances in the region of the X chromosome downstream of DXZ4.

DXZ4 is a region of key interest with regards to the X chromosome. In the context of XCI, DXZ4 acts as a hinge region that separates the two super domains of the inactive X chromosome structure (Bonora et al., 2018). For naïve hESCs, based on data provided by another member of the Plath lab, the region downstream of DXZ4 contains a higher enrichment of XIST RNA relative to that of fibroblasts.

Looking forward, it would be pertinent to determine what mediates the structural changes of the Xd. One potential method may be a candidate screen for XIST or XACT interactors. Already, many protein interactors of XIST have been identified (Chu et al., 2015; Mira-Bontenbal et al., 2016). By screening these proteins via depletion or deletion in conjunction with DNA FISH in key areas, for example two probes downstream of DXZ4 and at least one upstream or XACT, we can detect structural changes in regions of Xd that are known to be fairly condensed and quite distal, respectively. Alternatively, a more direct approach targeting the key lnRNAs, such as a deletion of XACT or XIST, could also shed light on the mechanisms of controlling the Xd structure.

Our original goal was to determine the structure of the dampened X chromosome with respects to the inactive and active X chromosomes. We interrogated multiple single chromosome representations of the dampened X chromosome to gain a greater knowledge of the overall shape as well as hone in on specific features relevant to gene expression. In broad strokes, we have determined that the dampened X chromosome in naïve hESCs, which on average tends to be slightly smaller in overall stature than the complementary active X, appears to have structural

similarities concerning genes that escape XCD. Additionally, within the region downstream of DXZ4 the dampened X chromosome appears to have the capacity to be larger in intrachromosomal distance than the active X chromosome when measuring distances across the chromosome but smaller when looking specifically at distances measured within the downstream DXZ4 region. This may point to a dampened X chromosome structure that is more condensed than an active X, more open than an inactive X, and having a large portion of the chromosome physically separated from the rest of the chromosome, potentially granting a structure larger than its complementary active X chromosome. This complex, unique structure may all be due to the enrichment of XIST RNA within the downstream DXZ4 region. This projects a scenario where XIST and DXZ4 consistently plays a critical role in X chromosome structure in all gene expression states, with the idea of the hinge region being applied to the dampened X chromosome as not a closed or open door but one curiously cracked open.

References

Bonora, G., Deng, X., Fang, H., Ramani, V., Qiu, R., Berletch, J. B., . . . Disteche, C. M. (2018). Orientation-dependent Dxz4 contacts shape the 3D structure of the inactive X chromosome. *Nature Communications, 9*(1). doi:10.1038/s41467-018-03694-y

Carrel, L., & Brown, C. J. (2017). When the Lyon(ized chromosome) roars: Ongoing expression from an inactive X chromosome. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1733), 20160355. doi:10.1098/rstb.2016.0355

Chu, C., Zhang, Q. C., da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., Magnuson, T., Heard, E., & Chang, H. Y. (2015). Systematic Discovery of Xist RNA Binding Proteins. Cell, 161(2), 404–416. doi:10.1016/j.cell.2015.03.025

Cremer, T., & Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Reviews Genetics*, *2*(4), 292–301. doi: 10.1038/35066075

Disteche, C. M., & Berletch, J. B. (2015). X-chromosome inactivation and escape. *Journal of Genetics, 94*(4), 591-599. doi:10.1007/s12041-015-0574-1

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., . . . Ren, B. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature, 518*(7539), 331-336. doi:10.1038/nature14222

Galupa, R., & Heard, E. (2018). X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation. *Annual Review of Genetics, 52*(1), 535-566. doi:10.1146/annurev-genet-120116-024611

Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., . . . Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature, 503*(7475), 290-294. doi:10.1038/nature12644

Lyon, M. F. (1961). Gene Action in the X-chromosome of the Mouse (Mus musculus L.). *Nature, 190*(4773), 372-373. doi:10.1038/190372a0

Mira-Bontenbal, H., & Gribnau, J. (2016). New Xist -Interacting Proteins in X-Chromosome Inactivation. Current Biology, 26(8), R338–R342. doi:10.1016/j.cub.2016.03.022

Patrat, C., Ouimette, J., & Rougeulle, C. (2020). X chromosome inactivation in human development. *Development, 147*(1). doi:10.1242/dev.183095

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S., Codeluppi, S., . . . Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell, 165*(4), 1012-1026. doi:10.1016/j.cell.2016.03.023

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T. W., . . . Plath, K. (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. *Cell Stem Cell, 20*(1), 87-101. doi:10.1016/j.stem.2016.10.006

Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur-Griscelli, A., . . . Rougeulle, C. (2013). XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. *Nature Genetics, 45*(3), 239-241. doi:10.1038/ng.2530

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Ali, T. M. L., … Rougeulle, C. (2017). XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell*, *20*(1), 102–111. doi: 10.1016/j.stem.2016.10.014

Wang, S., Su, J., Beliveau, B. J., Bintu, B., Moffitt, J. R., Wu, C., & Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. *Science, 353*(6299), 598-602. doi:10.1126/science.aaf8084