REGULATION OF GENOMIC STRUCTURE AND TRANSCRIPTION IN DROSOPHILA

by

Christopher Randal Bauer

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF MOLECULAR AND CELLULAR BIOLOGY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2009

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by **Christopher Randal Bauer**

entitled **REGULATION OF GENOMIC STRUCTURE AND TRANSCRIPTION IN DROSOPHILA**

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

_____Date: 12/11/2009

Dr. Giovanni Bosco

_____Date: 12/11/2009

Dr. Roy Parker

_____Date: 12/11/2009

Dr. Ted Weinert

_____Date: 12/11/2009

Dr. Daniela Zarnescu

_____Date: 12/11/2009

Dr. Andrew Capaldi

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 12/11/2009

Dissertation Director: Dr. Giovanni Bosco

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Christopher Randal Bauer

ACKNOWLEDGEMENTS

I would first like to acknowledge my advisor, Dr. Giovanni Bosco. His enthusiasm for science has certainly fueled my own. I am grateful to him for his guidance, for allowing me to explore my own interests, for encouraging me to push myself, and for providing that push whenever I lacked it. Thank you, Gio.

I would also like to thank all of my colleagues in the Bosco lab for adding some fun and some food to my life.

I would finally like to thank the members of JMST. To you I owe my fondest memories of my time spent in Tucson. Whether discussing scientific endeavors or engaging in less lofty pursuits, I will remember these times for the rest of my life.

DEDICATION

I dedicate this to my parents, Mary Sue and Randy Bauer. I don't know how to thank you for all that you have done for me. All that I have and all that I am, I owe to you.

TABLE OF CONTENTS

ABSTRACT

Within the span of a single human lifetime, we have discovered that DNA is the basis of genetic inheritance, deciphered the genetic code, and determined the entire sequence of multiple human genomes. However, we still have only a basic understanding of many of the processes that regulate DNA structure, function, and dynamics. The work presented in this dissertation describes the roles of two sets of genes that regulate the expression of genetic information and its transmission from one generation to the next.

The condensin II complex has been implicated in the maintenance of genomic integrity during cell division and in transcriptional regulation during interphase. These roles stem from its ability to regulate chromosome structure though the mechanisms of this regulation are unclear. Evidence suggests that it is important for chromosome condensation and segregation during mitosis and meiosis. We have shown that this complex regulates the condensation of chromosomes during interphase. Its ability to reduce chromosome axial length provides a mechanism for the establishment of chromosome territories. We have also shown that condensin II differentially regulates interactions between homologous and heterologous DNA sequences. These findings contribute to our understanding of the overall structure of the nucleus, the regulation of chromosome structure, and the regulation of gene expression.

The function of the *Drosophila* gene, *sticky*, is poorly understood. It contributes to cytokinesis by phosphorylating myosin II, but it also has a role in the regulation of

chromatin structure. Mutations in *sticky* are associated with a wide range of developmental abnormalities. We provide evidence that this gene regulates the expression of numerous other genes which contribute to the phenotypes observed when *sticky* is mutated. We also show that *sticky* function overlaps with that of *dfmr1*, an ortholog of the gene associated with the most common form of human mental retardation. These findings contribute to our understanding of transcriptional regulation in chromatin and its implications in development and disease.

CHAPTER 1: INTRODUCTION

Explanation of the dissertation format

This dissertation is arranged into two chapters. Chapter 1 introduces the questions that are addressed in this work and provides a context for this work by summarizing the literature. Chapter 2 contains two studies. The first is an unpublished manuscript describing work that was performed to understand the role to the condensin II complex in interphase nuclear organization. The second is a published manuscript detailing a study of the regulation of gene expression and biological networks by the genes *sticky* and *dfmr1*.

Explanation of the problem and its context

It was long believed that chromosomes were only loosely organized in the interphase nucleus. During the past few decades, this paradigm has been broken by numerous studies that describe regulated and dynamic physical interactions of specific chromosomal loci with a variety of nuclear strucutres and other loci, in cis and in trans. Additionally, many cell types have been shown to have even higher order structural characteristics such as discrete chromosomal territories and interchromosomal compartments. These spatial characteristics are important for many of the critical cellular processes that occur within the nucleus including DNA replication, DNA repair, recombination, maintenance of genomic intergrity, and transcriptional regulation.

However, very little is known about the pathways that function to establish, maintain, and alter genomic spatial organization. This dissertation discusses work that has been conducted to understand how two sets of genes contribute to nuclear organization and the processes that are affected by altered nuclear organization.

Literature review

The basis of DNA structure

Watson and Crick's model for the structure of DNA is undoubtedly one of the greatest contributions to our understanding of molecular biology. They proposed that DNA is structured as two anti-parallel strands that form a right handed double helix with a periodicity of ten complimentary base pairs in each helical turn of 34Å (Watson and Crick, 1953). This model fits elegantly with the semiconservative replication of DNA (Messelson and Stahl, 1958) and many other enzymatic processes that are known to involve DNA. Thus, the double helix model enjoys virtually unanimous acceptance as the true structure of DNA.

Despite this acceptance, it is well known that this model does not represent the structure of most DNA in cells. It is based upon X-ray diffraction patterns from crystalized oligonucleotides, roughly 17 bp in length (Watson and Crick, 1953). Genomic DNA can consist of strands that are millions of base pairs in length and it is rarely crystalline within cells. The average length of a human chromosome would be 4.3

cm if it were structured as described above (Human Genome Project Build 33, 2003), yet we know that 46 human chromosomes are all contained within a nucleus with a diameter that is roughly 1000 times smaller than this (Alberts, 2002). Given this, claiming that the Watson and Crick model is the structure of DNA is similar to claiming that the alpha helix is the structure of protein. In fact, the helix described by Watson and Crick is now known as B-form DNA as other helix arrangements (A-form and Z-form) and non helical arrangements such as G-quartets have been observed in crystallized oligonucleotides (Milman,1967; Wang, 1979; Horvath, 2001). Taking a lesson from the field of protein structure, it is useful to consider DNA in the context of a structural hierarchy with DNA sequence representing primary structure, the B-form helix as one possible secondary structure, and additional levels of structural organization built upon these.

DNA tertiary structure: chromatin

Beyond the level of secondary structures, the conformation of DNA is intimately connected to DNA binding proteins. In the hierarchy of protein structures, interactions between separate molecules are considered to be quaternary structures; however, since proteins play such a critical role in regulating the structure of DNA, it is convenient to consider a core set of these proteins as part of tertiary structure of DNA. This amalgamation of protein and DNA is known as chromatin. Considerable progress has been made in understanding the most fundamental units of chromatin structure, but deciphering its full complexity remains far beyond our current capabilities. In most

eukaryotic model systems, the basis of chromatin structure is the nucleosome. X-ray diffraction of crystallized nucleosomes has allowed structural characterization at sufficient resolution to infer atomic coordinates. Based on this data, we know that the canonical nucleosome is an octamer composed of four histone proteins (H2A, H2B, H3, and H4) around which ~147 bp of DNA is coiled into ~1.7 gyres (Richmond and Davey, 2003).

Chromatin fibers

In most eukaryotic genomes, 70-90% of DNA is nucleosomal (Segal and Widom, 2009). Numerous images of purified histones bound to DNA have been successfully acquired by atomic force microscopy and cryo-electron microscopy. It is often reported that the structure of these complexes is a 10nm fiber of nucleosomes connected by linker DNA (Allen, 1993; Fritzsche, 1994). Structures similar to this are likely to exist, at least transiently, in vivo, near genes that are actively being transcribed or replicated. The presence of linker histone H1/H5 can result in the compaction of these nucleosomal arrays into fibers that are roughly 30nm in diameter (Fan and Roberts, 2006). Fibers of similar dimension are also commonly observed in fixed chromatin that is isolated from cells and are thus thought to be representative of a standard chromatin structure (Yang, 1994). Many models have been proposed for the structure of the 30nm fiber. Though they differ in their details, most models suggest that linker histones mediate contacts

between nucleosomes to arrange them into a helical structure with ~6 nucleosomes per 11nm of fiber length (van Holde, 2007).

The various proposed structures for the 30nm fiber suggest that it compacts the length DNA roughly 30 fold compared to the naked B-form double helix (van Holde and Zlatanova, 2007). This is still insufficient to account for the observed compaction of some genomic regions in many interphase cells and several orders of magnitude less than the compaction of most mitotic chromosomes. Light microscopy and electron micrographs of fixed and living cells have indeed revealed the presence of higher order chromatin fibers ranging in diameter from 60 to 300nm depending on cell type and cell cycle state (Belmont et al, 1987; Belmont and Bruce, 1994). Models for the structures of these fibers are purely speculative.

Euchromatin and heterochromatin

The characteristics of chromatin fibers can vary greatly within a chromosome and throughout the cell cycle. In the 1920's, Emil Heitz observed cell cycle dependent differences in the staining patterns along the lengths of chromosomes from several mosses, plants, and animals. He coined the terms euchromatin and heterochromatin to describe chromosomal loci that are visible only during mitosis or throughout the cell cycle, respectively. He was also able to relate his cytological studies of *D. melanogaster* chromosomes to gene density in the genetic maps of Thomas Hunt Morgan and concluded that euchromatin represented the genetically important part of chromosomes

while heterochromatin was genetically inert (Passarge, 1979). Remarkably, this dichotomous division of chromatin structure is still common in the scientific literature.

We now have a basic molecular understanding of the distinctions between euchromatin and heterochromatin. They are based primarily upon DNA sequence and post-translational histone modifications. The sequenced genomes of *D. melanogaster* confirms Heitz's hypothesis that most genes exist within the euchromatin. It is generally less compacted than heterochromatin during interphase and therefore more accessible to cellular machinery which can explain many of its properties such as high transcriptional activity and early replication. The histones in euchromatic regions are generally highly acetylated and marked by a number of specific covalent modifications.

Heterochromatin is characterized by a condensed structure of evenly spaced nucleosomes (Sun et al., 2001). It is usually constitutively associated with centromeres, telomeres, transposons, and other types of repetitive sequences (Kurenova et al., 1998). Some loci containing unique DNA sequences can exhibit characteristics of heterochromatin at times and characteristics of euchromatin at others. These regions are referred to as facultative heterochromatin (Trojer and Reinberg, 2007). This highly compacted structure tends to be repressive to DNA depedent, enzymatic processes such DNA replication, DNA repair, and transcription (Bucceri et al., 2006). The histones in heterochromatin are usually depleted of acetyl groups and contain specific methylation marks on histone H3. In some species, DNA methylation is also common in these regions (Kouzarides, 2007).

Position effect variegation

A number of chromosomal rearrangements in *D. melanogaster* have provided useful tools with which to study the effect of chromatin states on transcription. Translocation of euchromatic genes to regions near the centric heterochromatin can result in a phenomenon known as position effect variegation (PEV). These genes frequently exhibit a large variability in their expression levels as well as epigenetic inheritance of these levels. The most common example of PEV occurs in a chromosome called white mottled 4 (wm^4). This X chromosome has undergone an inversion which has repositioned the white gene (*wt*) near the centric heterochromatin. In flies that carry this chromosomal aberration, some clonal patches of cells in the eyes express the *wt* gene and appear red while others fail to express this gene and appear white (Reuter and Wolff, 1981). This variability in transcriptional activity suggests a similar variability in the boundaries between chromatin states and that the silencing influence of heterochromatin can spread in cis to nearby sequences.

Another notable chromosome that exhibits PEV is brown dominant (bw^D). This chromosome contains a translocation of heterochromatic repeat sequences into the brown locus, another eye pigment gene (Csink, 2002). This results in silencing of the modified *bw* locus but surprisingly, this chromosome has a dominant effect and also silences the allelic wild type locus in heterozygous flies. FISH studies have revealed the cause of this dominance. The heterochromatic insertion induces the bw^D locus to colocalize with other repeat sequences at the chromocenter. Since homologous loci are usually paired in

Drosophila, the wt *bw* locus is also directed to the chromocenter (Henikoff, 1995). This suggests that the silencing influence of heterochromatin can also spread in trans and may be exerted within the three dimensional space of the nucleus rather than just along the length of the chromosome.

More than 500 genes have been identified in screens for modifiers of PEV (Schotta et al, 2003). Most of these genes are uncharacterized, but of those that have been studied, many affect transcription by regulating chromatin structure. Suppressors of variegation include a number of histone modifying enzymes such as Su(var)3-9 which methylates histone H3, lysine 9 (Rea et al., 2000). Mutations in Su(var)3-9 are epistatic to almost all other modifiers of variegation and this protein is conserved from mammals to plants to fungi suggesting its central role in chromatin regulation (Schotta et al 2002; Aagaard et al,1999; Ekwall and Ruusala 1994). Heterochromatin Protein 1 (HP1), a structural components of heterochromatin, is also a strong suppressor of PEV. HP1 can bind to dimethyl-H3K9 and it localizes to centric microsatellite DNA, telomeres, and a number of bands at other locations in the genome where it generally serves to confer a heterochromatic state (Fanti et al., 2003).

Enhancers of variegation are even less well characterized than the suppressors, but the few that have been mapped and cloned also seem to be involved in chromatin. E(var)3-9 is a multiple zinc finger protein suggesting interactions with DNA (Weiler, 2007). E(var)3-93D has extensive homology to *broad* and *tramtrak*, two known chromatin interactors, and it is a strong enhancer of *Trithorax*, connecting it to a highly studied pathway of histone modification and chromatin remodeling (Dorn et al, 1993).

Histone Modifications

Given the large representation of histone modifiying enzymes and chromatin binding proteins among the known modifiers of variegation, it is clear that posttranslational modifications of histones are fundamental to the transcriptional activity of chromatin. Investigations into the variety of histone modifications have revealed that they are also incredibly complex. Mass spectrometry and antibodies have been used to detect histones that are phosphorylated, acetylated, methylated, ubiquitinated, sumoylated, ADP-ribosylated, deiminated, or proline iosomerized at more than sixty different residues. A number of histone protein variants have also been identified and implicated in processes ranging from centromere organization to DNA repair (Kouzarides 2007).

In attempts to deal with this complexity, the concept of a nucleosome code has been proposed. In this paradigm, each histone modification serves as a specific binding site for one or more effector proteins. The combination of chromatin binding proteins present at a locus leads to a distinct readout such as transcriptionally active euchromatin or silent heterochromatin (Jennuwein and Allis, 2001). If this hypothesis is true, then it should be possible to predict transcriptional activity across an entire genome by mapping the distributions of all histone modifications, identifying the proteins that bind them, and determining how these proteins, and combinations thereof, affect the behavior of chromatin.

Though it is clear that many chromatin factors do bind specific modifications and that many loci engaging in similar processes often share a variety of histone modificaitons, there are some problems with this model. For one, it underestimates the importance of sequence specific DNA binding proteins that can induce dramatically different transcriptional activities in two genes that are in the same chromatin environment. Though there is evidence that genes which are near each other on chromosomes tend to be similarly expressed, only about 20% of the genes in the *Drosophila* genome share expression patterns with their nearest neighbors (Spellman and Rubin, 2002). Since chromatin environments are likely to be similar in more than 20% of adjacent genes, there must be interactions between the regulatory regions of individual genes, their chromatin contexts, and trans-regulatory factors which combine to determine gene expression.

Several other lines of evidence also support a more context dependent view of chromatin. The *light* (*lt*) gene in D. melanogaster is one of a few hundred genes that reside within the constitutive heterochromatin near centromeres. Despite this, *lt* is highly expressed in many tissues (Chintapalli et al, 2007). Several X-ray induced chromosomal rearrangements (e.g. lt^{x13}) that moved the *lt* locus out of the heterochromatin were isolated and all resulted in variegated expression of this gene (Hearn, 1991). This suggests that *lt* transcription is dependent upon a heterochromatic environment that silences most other genes. If the nucleosome code were correct, it would be expected that previously identified modifiers of variegation in w^{m4} should have inverse effects in lt^{x13}. However, this is true for only a subset of genes (Weiler, 2002). Another line of

evidence comes from studies of HP1. Polytene chromosomes from the salivary glands of D. melanogaster show localization of HP1 throughout the heterochromatin and at about 100 euchromatic bands (Fanti et al., 2003). Surprisingly, it also localized to puffs, which are cytological regions of high transcriptional activity. Further investigation revealed that levels of Hsp70 expression, after heat shock, were positively correlated with HP1 dosage. This implicates HP1 as positive regulator of gene expression in this context (Piacentini et al, 2003).

Clearly, chromatin is a complex set of interacting components that has evolved to facilitate a wide variety of processes that occur in eukaryotic genomes. Similar to most aspects of biology, we can find many correlations between the molecular features of different chromatin environments and the behavior of those environments, but we can also find many exceptions. Much work is still needed to determine which aspects of the relationships between chromatin structure and function are general and which are highly dependent on the underlying DNA sequences or non-chromatin factors. PEV has been a useful tool for identifying regulators of chromatin function but with a limited number of variegating reporter genes, it is impossible to separate the general effects from the gene specific ones. With the recent availability of numerous fly strains containing fluorescent reporter constructs, it should be possible to isolate lines which show variegated expression in almost any gene. This would provide a means to observe and quantify the effects of modifiers of variegation at multiple loci, in multiple tissues, and in live cells. Thus, it will be possible to control for gene and cell type specific effects and to understand the dynamic properties of transcriptional regulation by chromatin structure.

Even without a detailed understanding of how different chromatin structures confer differential functions to specific regions of the genome, it is clear that these structures impinge upon the higher order conformation of chromosomes and that higher levels of DNA structure feedback to chromatin.

DNA quaternary structure: nuclear organization

The quaternary structure of DNA describes how numerous DNA elements interact with each other and with non-DNA factors to produce organizational patterns in large chromosomal domains, entire chromosomes, nuclear compartments, and the entire nucleus. This level of organization is the least well understood aspect of DNA structure though it was the first to be explored since it is accessible to light microscopy. The nucleolus was described as early as 1835 based upon its distinctive staining (Dundr and Misteli, 2002). We now know that this structure represents the site of ribosome biogenesis. In many organisms, rDNA loci and tRNA genes, which are scattered across the genome, all localize to the nucleolus. At this site, they are transcribed by RNA polymerases I and III to produce the majority of cellular RNA (Haeusler and Engelke, 2006). There is growing evidence that other nuclear processes are facilitated by concentrating specific genes and factors into subnuclear compartments to promote specific molecular interactions.

Homology dependent pairing

For decades, it has been known that some DNA sequences can interact with other sequences. These types of interactions can be divided into homology dependent interactions and homology independent interactions. The former is usually referred to as chromosome pairing. The roles of pairing are best understood during meiosis when homologous chromosomes find each other within the nucleus, become aligned so that homologous sequences are near each other in space, and become tightly associated (synapsed) via a proteinaceous glue called the synaptonemal complex. These processes are closely tied to homolgous recombination which serves to increase genetic diversity within a population. Physical attachments between homologous chromosomes also allows tension to be generated when the kinetochores of each homolog are oriented in a bipolar fassion. This provides a checkpoint to ensure fidelity in segregation such that the daughter cells of meiosis I each possess exactly one copy of each homologue (Pinsky and Biggins, 2005).

The establishment of homologue pairing is a mysterious process and it is clear that multiple mechanisms contribute to it. During meiosis, in S. pombe, chromosomes undergo a dramatic reorganization that results in a polarized arrangement with all telomeres clustered at one side of the nucleus. This organization reduces the search for homology from a three dimensional space to roughly two dimensions but it is not sufficient to produce pairing (Burgess, 2004). The polarization of chromosomes has also been described during meiosis in *Drosophila* and *C. elegans* though it is more subtle than

in *pombe* (Page and Hawley, 2003). The formation of double strand breaks can also contribute to pairing in yeast, mice, and plants. Mutations in spo11, the enzyme which produces the double stand breaks the can initiate recombination, substantially reduce meiotic pairing in these organisms (Peoples et al. 2002; Mahadevaiah et al. 2001; Pawlowski et al. 2004). This suggests that homologue pairing during meiosis likely evolved from the homology search required for the homolgous recombination pathway of DNA repair. However, double strand breaks are dispensable for meiotic homologue pairing in *Drosophila* and *C. elegans* (McKim and Hayashi-Hagihara, 1998; Dernburg et al, 1998) though pairing in worms does require the DNA damage checkpoint response gene chk-2 (MacQueen and Villeneuve, 2001).

In *Drosophila*, homolgous chromosomes pair during embryogenesis and remain paired in most interphase cells throughout the life of the fly. The mechanism responsible for this somatic pairing may fulfill the meiotic role of double strand breaks, though it is unclear what mediates this process. The pairing of the X and Y chromosomes in *Drosophila* male meiosis I depends upon a single repetitive sequence near the rDNA locus (McKee, 1996). In *C. elegans*, similar repetitive loci, termed homology recognition regions, can also serve as meiotic pairing sites, even in the absence of the synaptonemal complex (MacQueen et al, 2002). These findings may suggest that other pairing events depend upon interaction at specific sites; however, other lines of evidence suggest that somatic and meiotic pairing occur at many sites throughout the chromosomes of *Drosophila* (Fung et al, 1998; Vazquez et al, 2002). There is evidence that pairing in flies does not require any single gene product (Bateman and Wu, 2008). One interesting

possibility is that pairing occurs via a large number of weak interactions directly between homologous DNA sequences. A recent report has shown that as little as 1kb of dsDNA can interact with homologous dsDNA independent of protein or free DNA ends (Danilowicz, 2009). Though it is difficult to infer much about in-vivo chromosome pairing based on such studies, it is possible that pairing represents the lowest energy conformation for homologous dsDNA in solution and it may be the default state for homologous chromosomes. This model can explain the prevalence of pairing in *Drosophila*, why genetic disruption of pairing has not been observed, and why energy dependent mechanisms are required to separate paired homologous loci (Hartl et al., 2008).

Homology independent interactions

Even in the absence of homology, DNA sequences can interact with each other. These types of interactions are almost certainly mediated by proteins. The most notable example of homology independent DNA interactions occurs between enhancers and promoters. Traditionally, enhancers are thought to modulate the transcriptional activity of promoters within 50-100kb (Blackwood and Kadonga, 1998), so many of these interactions are probably mediated more by DNA tertiary structure than by quaternary structure. However, some enhancer-promoter interactions occur between loci that are located on different chromosomes (Lomvardas, 2006; Spilianakis, 2005). These types of

interactions are difficult to detect and thus are probably more prevalent than the current literature suggests.

Heterologous DNA sequences can also interact on a much larger scale. In most *Drosophila* cells, all of the centric heterochromatic regions colocalize to single site near the nuclear envelope known as a chromocenter (Zhang and Spradling, 1995). Similar phenomena are also common in Arabidopsis though a number of chromocenter are usually present (Fransz et al., 2002). Even in human cells, heterologous loci that share chromatin features or activities have a higher probability of physically interacting than would be expected by chance (Lieberman-Aiden et al, 2009), though these interactions may be transient.

Nuclear compartmentalization

Recently, a method combining chromosome conformation capture and high-throughput sequencing (Hi-C) has provided a way to measure the probability of interaction between all DNA sequences in an entire genome (Lieberman-Aiden et al, 2009). Limitations in DNA sequencing technology currently prevent the identification of specific enhancer-promoter interactions, but these limitations may be overcome in the near future. Regardless, application of Hi-C has identified numerous high probability contacts between some regions on different human chromosomes suggesting that inter-chromosomal contacts are common and nonrandom. This study also provided tremendous insight into the organization of chromosomes as well as the entire nucleus. It

was observed that loci on the same chromosome, even loci at opposite ends of a chromosome, are much more likely to exist in close proximity than loci on separate chromosomes. It was also found that the nucleus is divided into two separate spatial compartments, termed A and B. For 21 out of 23 chromosomes, the primary determinant for a given locus to exist in either compartment was not its position on the chromosome. Instead, compartment A was generally composed of loci with high gene density, transcriptional activity, and DNAseI sensitivity while compartment B was composed of loci that lack these features (Lieberman-Aiden et al, 2009). These findings highlight two general features of nuclear organization: 1) Chromosomes exist within discreet territories. 2) Active euchromatic loci and silent heterochromatic loci are spatially segregated.

Chromosome territories

Descriptions of chromosomes as discrete, organized structures first appeared over 100 years ago (Cremer and Cremer, 2006). These patterns were rediscovered many decades later when molecular techniques revealed that UV lasers could induce extensive DNA damage in some chromosomes while most other chromosomes in the same nucleus were unaffected (Cremer et al, 1982). Since this time, fluorescence in situ hybridization (FISH) has conclusively shown that chromosomal territories are common. Cytological studies in humans, mice, birds, reptiles, insects, worms, cnidarians, fungi, and plants have all revealed the existence of chromosome territories though the organizational patterns range from the polarized Rabl conformation to radial organization to globular (Kosak,

2009; Foster and Bridger, 2005). Despite this wealth of evidence and its acceptance within the field, many cell biology text books still depict the interphase nucleus as a disorganized mass of chromatin fibers. This is likely due to the phenomenological and correlative nature of most studies on nuclear organization.

There are currently very few models of how chromosome territories are formed, how they can be reorganized, or even what purposes they serve. Of those that have been proposed, practically all center on transcriptional regulation. The first hypothesis to get any traction suggested that chromosome territories divided the nucleus into two regions, the chromosome territories (CTs) and the inter chromatin domain (ICD or simply IC). Inherent to this model was the idea that the CT was a dense structure and relatively impenetrable to polymerases or other bulky machinery. Thus, active genes would reside on the surface of CTs and silent genes would be buried within them. This was supported by FISH and antibody staining which showed that certain transcripts, snRNPs, splicing factors, and highly expressed genes were enriched near the boundaries between CTs and the ICD (Zirble et al., 1993; Mahy et al, 2002). However, it is now evident that genes near the centers of CTs can also be actively transcribed (Abranches et al., 1998; Verschure et al., 1999) and the many loci can move free within large areas of their CTs (Chuang et al., 2006; Branco and Pombo, 2006). One territory that may follow the rules CT-ICD model is the Barr body. This inactivated X chromosome is associated with a high density of RNAs such as Xist and proteins that may function to physically sequester it away from the transcriptional machinery (Fackelmayer et al., 2005).

Some other models suggest that positions within territories are not as important as where the territories themselves are located within the nucleus. In human cells, which have some of the most well defined territories, it is becoming clear that positions of specific territories are nonrandom and several trends have emerged. The center of the nucleus tends to be enriched for chromosomes with high gene density and small size while the opposite features are common near the nuclear periphery (Kosak et al., 2002; Lieberman-Aiden et al, 2009). Transcriptional activity also shows correlations with radial positioning but the direction of causation is not clear if it exists at all. It may even be the case that CTs hinder gene regulation as much as they facilitate it. Thus, some loci must leave their territories to allow for specific, interallelic interactions or other regulatory processes (Murmann et al., 2005; Brown et al., 2006). There is certainly no consensus at the moment regarding the roles of CTs in gene regulation, but it will be interesting to see what emerges in the years to come.

The nuclear periphery

One of the initial observations suggesting a functional compartmentalization of the nucleus was that heterochromatin is usually found close to the nuclear periphery (Hochstrasser and Sedat, 1987). More recently, many loci have been found to associate with the nuclear envelope via interactions with components of the nuclear lamina. Chromatin immunoprecipitation experiments with lamin and emerin antibodies in Drosophila have revealed a large number of lamin associated domains. These regions

tend to have low transcriptional activity, similar to heterochromatin (Guelen et al., 2008). These findings have lead to the hypothesis that the nuclear envelope exerts a silencing effect. This is supported by studies in yeast which show that physically tethering a gene to the nuclear envelope can induce silencing and studies in live cells which show that activation of peripherally located genes can induce them to move toward the center of the nucleus (Andrulis, 1998; Chuang, 2006).

It is clear that transcriptional silencing is not associated with the entire nuclear periphery. In fact, some transcriptionally active genes are known to localize near nuclear pore complexes (Casolari et al. 2004). One stunning exception to silencing at the nuclear lamina has been found in the retinal cells of nocturnal mammals. In these cells, heterochromatin is packaged into a dense structure near the center of the nucleus and actively transcribed euchromatic sequences are found closer to the nuclear envelope. Since this type of organization is typical in retinal cells of nocturnal mammals and uncommon elsewhere, it was hypothesized that this arrangement allowed nuclei to function as lenses which condense light onto photoreceptors and thus enhance night visions. Computational models suggest that the electron dense heterochromatin could function as a lens if it were arranged into a smooth round shape within the less dense euchromatin (Solovei, 2009). This finding suggests that nuclear organization can evolve to produce a variety of patterns, but it supports the idea that the global structure of the nucleus contributes greatly to it functionality, whatever those functions may be.

Disruption of the nuclear lamina leads to defects in the structure of the nuclear envelope and chromatin (Clements, 2000). This can lead to a number of disease states in humans. One of the primary structural components of the nuclear lamina is lamin. Mutations in the human lamin A gene have been associated with at least 13 diseases including Emery-Dreifuss muscular dystrophy, cardiomyopathy, Werner syndrome, and Hutchinson-Gilford progeria (Online Mendelian Inheritance in Man). This wide range of lamin associated phenotypes suggests that lamin functions to regulate numerous processes, likely through its contribution to nuclear organization. Fibroblasts from children with progeria exhibit wrinkled nuclear envelopes and reduced staining for the heterochromatic histone modification, methyl-H3K9 (Shumaker, 2006). This undoubtedly has implications for the regulation of transcription and DNA replication. Since the symptoms of progeria resemble premature aging, it was hypothesized that nuclear architecture may contribute to normal aging. Remarkably, fibroblasts from elderly individuals, fibroblasts from young individuals with high passage number, and non-neuronal cells in C. elegans mutants with short life spans all exhibit abnormalities in nuclear envelope morphology similar to those seen in progeria patients (Scaffidi and Misteli, 2006; Cao et al, 2007; Haithcock et al, 2005).

Nuclear Dynamics

Though nuclei are highly structured, they require the ability to alter their structure in order to perform a variety of basic cellular functions and respond to stimuli. At the

most basic level, the double helical structure of DNA can be distorted by interactions with proteins. Interaction with TATA binding protein causes DNA to unwind and bend sharply (Chasman et al, 1993). Unique DNA structures such as these can serve as binding sites for additional proteins resulting in multistep regulatory processes with high specificity (Williams, 2008). This is probably also a thermodynamic necessity as many processes that occur on DNA require dramatic changes in the underlying structure that are typically unstable. Both DNA and RNA polymerases require unwinding of the DNA helix which generates superhelical tension and melting of the base pairs to produce stretches of single stranded DNA. These distortions are produced by ATP dependent helicase activity and are stabilized by topoisomerases and single stranded DNA binding proteins (Collins, 2001). Since most DNA in eukaryotic genomes is associated with histones, enzymatic processes that occur on DNA must also alter DNA-histone interactions.

Chromatin remodeling

Though recent mapping of preferred nucleosome positions suggests that many functional transcription factor binding sites are depleted of nucleosomes, this is insufficient to explain how all transcription factors gain access to their binding sites (Sekinger et al, 2005). A number of chromatin remodeling complexes can actively disrupt DNA-histone interactions. In vitro assays for the ISWI complex activity show that it can hydrolyze ATP to produce regular spacing between nucleosomes on plasmid

DNA (Ito et al, 1997). Yeast lacking *Sth1*, the catalytic subunit of the RSC chromatin remodeling complex, show changes in both nucleosome density and spacing at a variety of genes (Parnell et al. 2008). These findings show that histones can be actively moved around or displaced from DNA to expose binding sites for a variety of proteins.

Without the activity of chromatin remodeling complexes, nucleosomes appear to be quite stable. SWI/SNF can actively displace nucleosomes from preferred positions in vitro but, when ATP is depleted, nucleosomes remain displaced for several hours (Ulyanova and Schnitzler, 2005). FRAP experiments in live cells expressing GFP tagged histones shows that H3 and H4 were incorporated into DNA after replication and about 80% remained permanently bound, even through S phase and mitosis. H2B exhibited slightly higher exchange rates. 3% of GFP-H2B showed transcription dependent exchange with $t_{1/2}$=6 minutes, 43% exchanged with $t_{1/2}$=130 minutes, and the remaining population did not exchange over a 16 hour period (Kimura and Cook, 2001). Unfortunately, FRAP only shows that most histone proteins are not freely diffusible and does not indicate anything about their mobility on DNA.

Though nucleosomes may be quite stable unless forced to move, proteins may be able to access nucleosomal DNA without displacing histones. The photolyase enzyme can only efficiently repair pryrimidine dimers in naked DNA. However, when it is present in high concentrations, it can repair 50% of UV-induced dimers in the yeast genome in 5 seconds. Many of these repaired sites are loci with high densities of histones (Bucceri et al, 2006). The rate of this repair seems too fast to be dependent upon

active chromatin remodeling. Thus, it seems that many histone-DNA interactions may be loose enough to allow rapid DNA unwrapping, looping, or sliding without displacement of the histones from the DNA.

Chromatin fiber dynamics

Chromatin fibers also exhibit constrained motion within the nucleus. Tandem arrays of the lac operator have been integrated into many genomic locations in many organisms. Expression of lacI-GFP fusion proteins allows the lacO sites to be visualized in live cells and followed in real time. This approach has been used to observe the spatial dynamics of chromosomes in yeast, flies, and mammals. Most of these studies conclude that interphase chromosomal loci exhibit Brownian motion that is confined within a small fraction of the nuclear volume (Heun et al, 2001; Marshall et al, 1997; Chubb et al, 2002). The typical interpretation of these findings is that diffusion governs motions within the nucleus and that loci are either tethered via chromosomal associations with the nuclear envelope or that the crowded nuclear environment prevents long range motions. Another investigation into chromosome dynamics over the course of several hours suggests a modification to this model. *Drosophila* spermatocyte chromosomes exhibit rapid, short range motion consistent with constrained diffusion on the time scale of seconds but show slower long range motions over the course of several minutes. These data suggest that a given locus can explore an entire chromosome territory in about one

hour. Furthermore, the long range motion can be seen in early G2 but are gradually reduced and cease completely prior to the first meiotic division (Vazquez et al, 2001).

Many reports suggest that chromosomes can exhibit very rapid movements over large distances during meiosis I. In a wide variety of eukaryotes, telomeres cluster on the nuclear envelope near the spindle pole body (SPB) with their centromeres oriented toward the opposite pole. This is known as the bouquet arrangement. In fission yeast, this arrangement is followed by a change in the structure of the entire nucleus from a sphere to a highly elongated body called a horsetail nucleus (Scherthan, 2001). The extenion of microtubules, primarily in one direction, causes the SPB to move toward the opposite end of the cell, dragging the telomeres along with it. Once the nucleus reaches one end of the cell, microtubules become stabilized in the opposite direction, pushing the telomeres back to the other side of the cell. The telomeres continue to lead the nucleus back and forth across the cell body every 5-10 minutes for a few hours (Ding et al, 1998). The purpose of this process is not completely clear but it is thought that chromosome movements induced by the whipping of the nucleus allow homologous chromosomes to find each other and pair (Hiraoka, 1998). Nuclear movements have also been observed in rat spermatocytes and maize meiocytes during prophase I. These involve oscillating rotations of the entire chromatin mass within the nucleus rather than translational movement of the entire nucleus. Synapsis in rat spermatocytes coincides with these rotational movements (Parvinen and Soderstrom, 1976). Chromatin rotational oscillations in maize depend upon polymerization of both microtubules and actin. Blocking either of these processes causes defects in homologue pairing (Sheehan and

Pawlowski, 2009). Though there are few other observations of such chromosome movements, the facts that these three observations come from three different phyla and that they seem to share functional and mechanistic features suggests that similar phenomena may be universal and simply await observation in other eukaryotes.

Mitosis and meiosis

In addition to prophase I chromosome movements, many levels of DNA structure exhibit their greatest dynamic range during events associated with mitosis and meiosis. Prior to and during cell division, the genome experiences histone modification, chromatin condensation, restructuring of chromosome territories, and long range directed movements. Mitotic and meiotic chromosomes are characterized by extensive histone phosphorylation, particularly on histone H3. During late G2 and M phases, several serine and threonine residues are phosphorylated in all organisms that have been studied (Ito, 2007). However, substantial variability is observed between organisms, between mitosis and meiosis, between meiosis I and II, between different regions of the chromosomes, and during mitotic progression (Houben et al, 2007; Wang et al, 2006). How these modifications contribute to the dramatic structural changes of chromosomes is far from clear. Chromosome condensation, cohesion, and segregation have been shown to depend on specific phosphorylated residues, but again, these dependencies are often organism specific (Houben et al, 2007; Nowak and Corces, 2004).

Chromosome Condensation

During prophase, chromosomes condense into discrete entities. This process involves axial compaction of chromatin fibers and resolution of entanglements between chromosomes. Chromosome condensation represents yet another aspect of the nucleus that has been well known for over a century but still eludes our understanding. Andrew Belmont describes the structure of mitotic chromosomes as "a riddle, wrapped in a mystery, inside an enigma" (Belmont, 2006). Many models of chromosome condensation involve the formation of chromatin loops that are folded and attached at intervals to a proteinaceous filament that forms the core of the mitotic chromosome. These models are supported by a number of observations. Isolation of mitotic chromosomes followed by nuclease digestion revealed that they possessed a filamentous core of proteins. Most prominent among these were topoisomerase II and a number of mysterious chromosome associated proteins (CAPs) (Laemmli et al, 1978). Years later, a five subunit complex of CAP proteins was identified as a factor necessary for the condensation of sperm DNA in Xenopus oocyte extracts (Hirano et al, 1997). This complex was named condensin and found to be highly conserved throughout eukaryotes. Based purely upon this evidence, condensin and topoisomerase II seemed to provide the functions necessary to condense chromosomes, physically support mitotic chromosome structure, and resolve entanglements between sister chromatids.

Many experiments have since argued against this model of chromatin loop connected to a core of topoisomerase II and condensin. In a wide array of organisms and

cell culture systems, mutation or depletion of topoisomerase II and/or condensin subunits does not prevent chromosome condensation. These genetic perturbations frequently lead to segregation defects such as anaphase bridges, altered mitotic chromosome morphology, or mitotic delay; however, chromosomes are almost always observed to achieve fairly normal levels axial shortening (Savvidou et al., 2005; Dej et al., 2004; Hagstrom et al., 2002). If condensin was the main enzyme responsible for condensation or a structural component of a chromosomal scaffold, it is difficult to imagine how mitotic chromosomes could form or be maintained in the absence of condensin. FRAP experiments have also determined that topoisomerase II exchanges very rapidly from mitotic chromosomes in vivo making it a poor candidate for a structural scaffold protein (Christensen et al, 2002).

The identification of a second condensin complex in higher eukaryotes offered additional candidates for scaffold proteins. However, condensin I and II share the same ATPase subunits, SMC2 and SMC4, and differ in their other three subunits (Ono et al, 2003). Thus, previous experiments that involved mutation or knockdown of the SMC subunits should have inhibited both condensin complexes and prevented condensation. Later analyses revealed that RNAi knockdown of the unique subunits in both complexes was also insufficient to prevent chromosome condensation though it was significantly delayed (Hirota, 2004).

Some experimental evidence suggests that the structure of mitotic chromosomes does not depend on a protein scaffold at all. First, chromosome micromanipulation

reveals that mitotic chromosomes are elastic and can be reversibly stretched to five times their typical length. Digestion with restriction enzymes eliminates this elasticity and eventually cause the chromosome to become severed (Poirier et al, 2002). Another line of evidence comes from a Xenopus egg extract system that supports DNA replication, chromosome condensation, and spindle assembly in vitro. In this assay the nucleosome linker, histone H1, was shown to be necessary for condensation of sperm chromosomes that had gone through DNA replication. Depletion of histone H1 by antibody pull-down resulted in mitotic chromosomes that were thin and 50% longer than chromosomes in control extracts. The H1 depleted chromosomes and control chromosomes contained similar levels of several CAPs including condensins and topoisomerase II (Maresca, 2005). Histone H1 is known to facilitate interactions between nucleosomes to produce of 30nm fibers and possibly higher order fibers. Thus, if the mitotic chromosome was structured as chromatin loops that emanate from an axial protein scaffold, it would be expected that mitotic chromosomes lacking H1 would simply have longer loops and appear wider or fuzzier than normal chromosomes. The opposite finding suggests that the compaction of chromatin fibers contributes directly to the axial shortening of mitotic chromosomes. This is consistent with a model in which contacts between DNA form the basis of mitotic chromosome structure.

Functions of the condensin complex

If mitotic chromosomes do not possess a protein scaffold that organizes their structure then we are left to determine how chromatin fibers are linked in order to condense chromosomes. Clearly the condensin complexes play some role in this process. Though loss of condensin function does not eliminate condensation, it does delay condensation, reduce the hypercondensation observed during mitotic arrest, alter the morphology of condensed chromosomes, and destabilize mitotic chromosomes (Savvidou et al, 2005**;** Hirota et al, 2004). The most likely explanation is that multiple, or partially redundant, pathways contribute to condensation. There may exist an unidentified factor that possess the capacity to compact chromatin. Alternatively, proteins that are already known to be part of chromatin may interact in ways we do not yet appreciate to accomplish this task. Condensin may provide some condensation on its own but its activity probably contributes to a more fundamental aspect of DNA structure that facilitates or stabilizes condensation and contributes to other processes as well.

This idea is supported by the finding that bacteria possess homologues of the condensin SMC subunits, even though they do not condense their chromosomes in a way that is directly analogous to eukaryotic chromatin condensation. E. coli lacking the SMC homologue, mukB, have disordered nucleoid structures and produce a high proportion of anucleate cells due to chromosome segregation defects (Case et al, 2004). In addition to the phenotypic similarities between mukB and SMC protein mutants, the proteins themselves have many conserved structural and functional characteristics. Both proteins

have a partial ATPase domain connected to a dimerization domain by a coiled coil. The ATPase domains are capable of binding ATP but they cannot hydrolyze it as monomers. When dimerization occurs, the two ATPase domains with bound ATP can interact with each other to form two complete ATPases. ATP hydrolysis results in separation of the ATPase domains (Woo et al., 2009).

SMC proteins also resemble mukB in their ability to bind and alter the topology of DNA. Purified Xenopus condensin is capable to inducing and stabilizing positive supercoils in plasmid DNA. This activity requires the presence of topoisomerase I and ATP (Kimura and Hirano, 1997). Additionally, this complex can produced right handed knots in plasmid DNA in the presence of topoisomerase II (Kimura et al, 2000). mukB possesses very similar activities in vitro. It can produce right handed plasmid knots in the presence of topoisomerase II and it can induce and stabilize supercoils in the presence of topoisomerase I. The only differences seem to be that mukB induces negative rather than positive supercoils and its topological effects are independent of ATP hydrolysis.

It is tempting to hypothesize that mukB alters the topology of chromosomes much like it does in plasmids. In theory, the circular chromosome of *E. coli* could be compacted by local supercoiling induced by mukB and compensatory writhe in other regions of the chromosome. However, it is unlikely that the behavior of the 4.6Mb *E. coli* chromosome resembles that of a 3kb plasmid. Furthermore, numerous other enzymes such as helicases and polymerases can also induce supercoils into DNA and *E. coli* possesses a gyrase enzyme, which naturally relieves superhelical tension (Reyes-

Domínguez, 2003). It is unclear how the effects of mukB would manifest in the context of these other enzymatic processes.

The molecular functions of condensins on linear chromosomal DNA probably also involves some form of topological modulation. There is currently no evidence to suggest that it can induce the formation of knots in this context, though, such a mechanism could provide a rather intuitive model for condensation in which condensin links regions of DNA by simply tying them into a knot. The interaction of condensin with topoisomerase II suggests that this is a possibility and these knots could be resolved by topo II during decondensation. It is difficult to imagine what role supercoiling could play in a linear eukaryotic chromosome. Given the enormous size of eukaryotic chromosomes and the topologically constrained nature of chromatin, it seems likely that superhelical tension would remain in a localized region once induced, unless relieved by a topoisomerase. Superhelical tension in a linear DNA molecule would also be difficult to detect and could explain why the function of condensin has remained so elusive. Some intercalating molecules preferentially interact with negatively supercoiled DNA and these molecules have been used to visualize superhelicity in chromosomes (Matsumoto and Hirose, 2004). Methods such as this may help to elucidate a role for superhelical tension in chromosome condensation. Recently, the plectonemic model of chromosome condensation has been proposed (Banfalvi, 2008). Though this model was proposed to explain the structure of some less than convincing images of individual chromatin fibers in decondensing chromosomes, it predicts the formation of coiled hairpin structures

composed of 10nm fibers which might be expected to form in the presence of superhelical tension.

Perhaps the most appealing feature of model in which condensin regulates the superhelical topology of DNA, besides the direct biochemical evidence for it, is its generality. Many of the models that have been proposed for the molecular function of condensin rest upon speculations about its ability to form oligomers, rosette structures or rings that topologically embrace DNA loops (Hagstrom and Meyer, 2003; Bazett-Jones et al, 2002; Hirano, 2006). There is little evidence that condensin is able to form these structures and these models usually suggest a direct role in chromosome condensation, which it may not even have. Condensin has also been implicated in a number of other mitotic and meiotic roles such as promoting sister chromatid segregation and promoting rDNA segregation as well as interphase roles such as inhibiting chromosome pairing, promoting the colocalization of tRNA genes, enhancement of PEV, suppression of PEV, and promoting DNA repair (Hartl, 2008; Hirano, 2006; Cobbe, 2006; Haeusler and Engelke, 2006). In the following chapter, we provide evidence that *Drosophila* condensin II is necessary for a 10 fold axial compaction of interphase chromosomes and we propose that this activity is directly responsible for establishing chromosome territories. We also show that the inhibition of chromosome pairing by condensin II is restricted to the euchromatin. Finally, we show that condensin prevents heterologous heterochromatic repeats from aggregating into a chromocenter. In order for a single complex to play a role in so many diverse functions related to DNA, it most likely functions to regulate fundamental properties of DNA structure such as superhelicity.

CHAPTER II: PRESENT STUDY

The methods, results, and discussion of this dissertation are included as two manuscripts. Some additional background information, specific to each study, is provided as well. The first manuscript, **Drosophila Condensin II Induces Interphase Genome Spatial Reorganization via Chromosome Axial Shortening**, will be submitted in the near future. It describes an investigation into the functions of the Drosophila condensin II complex in the regulation of chromosome structure and global nuclear organization during interphase. The second manuscript, **Genetic and systems level analysis of *Drosophila sticky/citron kinase* and *dFmr1* mutants reveals common regulation of genetic networks**, was published in BMC Systems Biology on the 25th of November, 2008. Presented within are the findings and interpretations of a study which involved the integration of microarray gene expression data, molecular interaction networks, genetic manipulations, and phenotypic characterization to elucidate the mode of interaction between two pleiotropic mutants in the genes *sticky* and *dFmr1* and to understand how these two genes regulate such diverse biological processes.

REFERENCES

Aagaard L, Laible G, Selenko P, Schmid M, Dorn R, Schotta G, Kuhfittig S, Wolf A, Lebersorger A, Singh PB, Reuter G, Jenuwein T. (1999) Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J. 18(7):1923-38.

Abranches R et al., (1998) Transcription sites are not correlated with chromosome territories in wheat nuclei. J. Cell Biol. 143 , pp. 5–12

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson J D. Molecular Biology of the Cell. Fourth Edition. New York, Garland Science. 2002

Allen MJ, Dong XF, O'Neill TE, Yau P, Kowalczykowski SC, Gatewood J, Balhorn R, Bradbury EM. (1993) Atomic force microscope measurements of nucleosome cores assembled along defined DNA sequences. Biochemistry. 32(33):8390-6.

Banfalvi G. (2008) Chromatin fiber structure and plectonemic model of chromosome condensation in Drosophila cells. DNA Cell Biol. 27(2):65-70

Bazett-Jones DP, Kimura K, Hirano T. (2002) Efficient supercoiling of DNA by a single condensin complex as revealed by electron spectroscopic imaging. Mol Cell. 9(6):1183-90

Belmont AS, Bruce K. (1994) Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. J Cell Biol. 127(2):287-302.

Belmont AS, Sedat JW, Agard DA. (1987) A three-dimensional approach to mitotic chromosome structure: evidence for a complex hierarchical organization. J Cell Biol. 105(1):77-92.

Belmont AS. (2006) Mitotic chromosome structure and condensation. Curr Opin Cell Biol. 18(6):632-8

Branco MR, Pombo A. (2006) Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. 4(5):e138. Epub 2006 Apr 25.

Brown JM, J. Leach, J.E. Reittie, A. Atzberger, J. Lee-Prudhoe, W.G. Wood, D.R. Higgs, F.J. Iborra and V.J. Buckle, (2006) Coregulated human globin genes are frequently in spatial proximity when active, J Cell Biol 172, pp. 177–187

Bucceri A, Kapitza K, Thoma F. (2006) Rapid accessibility of nucleosomal DNA in yeast on a second time scale. EMBO J. 25(13):3123-32

Burgess SM. (2004) Homolog pairing in S. pombe: the ends are the means. Mol Cell.13(6):766-8

Cao K, Capell BC, Erdos MR, Djabali K, Collins FS. (2007) A lamin A protein isoform overexpressed in Hutchinson-Gilford progeria syndrome interferes with mitosis in progeria and normal cells. Proc Natl Acad Sci U S A. 104(12):4949-54

Case RB, Chang YP, Smith SB, Gore J, Cozzarelli NR, Bustamante C. (2004) The bacterial condensin MukBEF compacts DNA into a repetitive, stable structure. Science. 305(5681):222-7

Chasman DI, Flaherty KM, Sharp PA, Kornberg RD. (1993) Crystal structure of yeast TATA-binding protein and model for interaction with DNA. Proc Natl Acad Sci U S A. 90(17):8174-8

Chintapalli VR, Wang J, Dow JA. (2007) Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat Genet. 39(6):715-20.

Christensen MO, Larsen MK, Barthelmes HU, Hock R, Andersen CL, Kjeldsen E, Knudsen BR, Westergaard O, Boege F, Mielke C. (2002) Dynamics of human DNA topoisomerases IIalpha and IIbeta in living cells. J Cell Biol. 157(1):31-44

Chubb JR, Boyle S, Perry P, Bickmore WA. (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol. 12(6):439-45

Clements L, Manilal S, Love DR, Morris GE. (2000) Direct interaction between emerin and lamin A. Biochem Biophys Res Commun. 267(3):709-14

Cremer T, Cremer C, Baumann H, Luedtke EK, Sperling K, Teuber V, Zorn C. (1982) Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. Hum Genet. 60(1):46-56

Dej KJ, Spradling AC. (1999) The endocycle controls nurse cell polytene chromosome structure during Drosophila oogenesis. Development. 126(2):293-303.

Danilowicz C, Lee CH, Kim K, Hatch K, Coljee VW, Kleckner N, Prentiss M. (2009) Single molecule detection of direct, homologous, DNA/DNA pairing. Proc Natl Acad Sci U S A. 106(47):19824-9

Dernburg AF, McDonald K, Moulder G, Barstead R, Dresser M, Villeneuve AM. (1998) Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell. 94(3):387-98

Ding DQ, Chikashige Y, Haraguchi T, Hiraoka Y. (1998) Oscillatory nuclear movement in fission yeast meiotic prophase is driven by astral microtubules, as revealed by continuous observation of chromosomes and microtubules in living cells. J Cell Sci.111:701-12

Dorn R, Krauss V, Reuter G, Saumweber H. (1993) The enhancer of position-effect variegation of Drosophila, E(var)3-93D, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. Proc Natl Acad Sci U S A. 90(23):11376-80.

Dundr M, Misteli T. (2002) Nucleolomics: an inventory of the nucleolus. Mol Cell. 9(1):5-7.

Ekwall K, Ruusala T. (1994) Mutations in rik1, clr2, clr3 and clr4 genes asymmetrically derepress the silent mating-type loci in fission yeast. Genetics. 36(1):53-64.

Fackelmayer FO, (2005) A stable proteinaceous structure in the territory of inactive X chromosomes, J Biol Chem 280, pp. 1720–1723

Fanti L, Berloco M, Piacentini L, Pimpinelli S. (2003) Chromosomal distribution of heterochromatin protein 1 (HP1) in Drosophila: a cytological map of euchromatic HP1 binding sites. Genetica. 117(2-3):135-47

Foster HA, Bridger JM. (2005) The genome and the nucleus: a marriage made by evolution. Genome organisation and nuclear architecture. Chromosoma. 114(4):212-29

Fritzsche W, Schaper A, Jovin TM. (1994) Probing chromatin with the scanning force microscope. Chromosoma. 103(4):231-6.

Fung JC, Marshall WF, Dernburg A, Agard DA, Sedat JW. (1998) Homologous chromosome pairing in Drosophila melanogaster proceeds through multiple independent initiations. J Cell Biol. 1998 141(1):5-20

Haeusler RA, Engelke DR. (2006) Spatial organization of transcription by RNA polymerase III. Nucleic Acids Res. 34(17):4826-36

Hagstrom KA, Meyer BJ. (2003) Condensin and cohesin: more than chromosome compactor and glue. Nat Rev Genet. 4(7):520-34

Haithcock E, Dayani Y, Neufeld E, Zahand AJ, Feinstein N, Mattout A, Gruenbaum Y, Liu J. (2005) Age-related changes of nuclear architecture in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 102(46):16690-5

Heun P, Laroche T, Shimada K, Furrer P, Gasser SM. (2001) Chromosome dynamics in the yeast interphase nucleus. Science. 294(5549):2181-6

Hirano T, Kobayashi R, Hirano M. (1997) Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a Xenopus homolog of the Drosophila Barren protein. Cell. 89(4):511-21

Hirano T. (2006) At the heart of the chromosome: SMC proteins in action. Nat Rev Mol Cell Biol. 7(5):311-22.

Hiraoka Y. Meiotic telomeres: a matchmaker for homologous chromosomes. Genes Cells. 3(7):405-13

Hirota T, Gerlich D, Koch B, Ellenberg J, Peters JM. (2004) Distinct functions of condensin I and II in mitotic chromosome assembly. J Cell Sci. 117(Pt 26):6435-45

Hochstrasser M, Sedat JW. (1987) Three-dimensional organization of Drosophila melanogaster interphase nuclei. II. Chromosome spatial organization and gene regulation. J Cell Biol. 104(6):1471-83

Houben A, Demidov D, Caperta AD, Karimi R, Agueci F, Vlasenko L. (2007) Phosphorylation of histone H3 in plants--a dynamic affair. Biochim Biophys Acta. 1769(5-6):308-15

Human Genome Project, Build 33 (2003) Chromosome FAQs: http://www.ornl.gov/sci/techresources/HumanGenome/posters/chromosome/faqs.shtml

Ito T, Bulger M, Pazin MJ, Kobayashi R, Kadonaga JT. (1997) ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell. 90(1):145-55

Ito T. (2007) Role of histone modification in chromatin dynamics. J Biochem.141(5):609-14

Jenuwein T, Allis CD. (2001)Translating the histone code. Science. 293(5532):1074-80

Kimura H, Cook PR. (2001) Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J Cell Biol. 153(7):1341-53

Kimura K, Hirano T. (1997) ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation. Cell. 90(4):625-34

Kimura K, Hirano T. (2000) Dual roles of the 11S regulatory subcomplex in condensin functions. Proc Natl Acad Sci U S A. 97(22):11972-7

Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, Singh H. (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science. Apr 5;296(5565):158-62.

Kouzarides T. (2007) Chromatin modifications and their function.Cell. 128(4):693-705.

Laemmli UK, Cheng SM, Adolph KW, Paulson JR, Brown JA, Baumbach WR. (1978) Metaphase chromosome structure: the role of nonhistone proteins. Cold Spring Harb Symp Quant Biol. 42 Pt 1:351-60

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. (2006) Interchromosomal interactions and olfactory receptor choice. Cell. 126(2):403-1

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 326(5950):289-93.

MacQueen AJ, Colaiácovo MP, McDonald K, Villeneuve AM. (2002) Synapsis-dependent and -independent mechanisms stabilize homolog pairing during meiotic prophase in C. elegans. Genes Dev. 16(18):2428-42

MacQueen AJ, Villeneuve AM. (2001) Nuclear reorganization and homologous chromosome pairing during meiotic prophase require C. elegans chk-2. Genes Dev. 15(13):1674-87

Mahadevaiah SK, Turner JM, Baudat F, Rogakou EP, de Boer P, Blanco-Rodríguez J, Jasin M, Keeney S, Bonner WM, Burgoyne PS. (2001) Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet. 27(3):271-6

Mahy NLet al., (2002) Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. J. Cell Biol. 157, pp. 579–589

Maresca TJ, Freedman BS, Heald R. (2005) Histone H1 is essential for mitotic chromosome architecture and segregation in Xenopus laevis egg extracts. J Cell Biol. 169(6):859-69

Marshall WF, Straight A, Marko JF, Swedlow J, Dernburg A, Belmont A, Murray AW, Agard DA, Sedat JW. (1997) Interphase chromosomes undergo constrained diffusional motion in living cells. Curr Biol. 7(12):930-9

Matsumoto K, Hirose S. (2004) Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of Drosophila. J Cell Sci. 117(Pt 17):3797-805

McKee BD. (1996) The license to pair: identification of meiotic pairing sites in Drosophila. Chromosoma. 105(3):135-41

McKim KS, Hayashi-Hagihara A. (1998) mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes Dev. 12(18):2932-4

Meselson M, Stahl FW (1958) Proc. Natl. Acad. Sci. USA 44:671–682.

Murmann AE, J. Gao, M. Encinosa, M. Gautier, M.E. Peter, R. Eils, P. Lichter and J.D. Rowley, (2005) Local gene density predicts the spatial position of genetic loci in the interphase nucleus, Exp Cell Res 311 (, pp. 14–26

Nowak SJ, Corces VG. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet. 20(4):214-20

Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), {date of download}. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/

Ono T, Losada A, Hirano M, Myers MP, Neuwald AF, Hirano T. (2003) Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell.115(1):109-21

Parvinen M, Söderström KO. (1976) Effects of ESH and testosterone on the RNA synthesis in different stages of rat spermatogenesis. J Steroid Biochem. 7(11-12):1021-3

Passarge E. (1979) Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. Am J Hum Genet. 31(2): 106–115.

Page SL, Hawley RS. (2003) Chromosome choreography: the meiotic ballet. Science. 301(5634):785-9

Parnell TJ, Huff JT, Cairns BR. (2008) RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. EMBO J. 27(1):100-10

Peoples TL, Dean E, Gonzalez O, Lambourne L, Burgess SM. (2002) Close, stable homolog juxtaposition during meiosis in budding yeast is dependent on meiotic recombination, occurs independently of synapsis, and is distinct from DSB-independent pairing contacts. Genes Dev. 16(13):1682-9

Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S. (2003) Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. J Cell Biol. 161(4):707-14

Pinsky BA, Biggins S. (2005) The spindle checkpoint: tension versus attachment. Trends Cell Biol. 15(9):486-93

Pawlowski WP, Golubovskaya IN, Timofejeva L, Meeley RB, Sheridan WF, Cande WZ. (2004) Coordination of meiotic recombination, pairing, and synapsis by PHS1. Science. 303(5654):49-50.

Poirier MG, Marko JF. Micromechanical studies of mitotic chromosomes. J Muscle Res Cell Motil. 23(5-6):409-31

Ragoczy T, Telling A, Sawado T, Groudine M, Kosak ST. (2003) A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. Chromosome Res.;11(5):513-25.

Rajapakse I, Perlman MD, Scalzo D, Kooperberg C, Groudine M, Kosak ST. (2009) The emergence of lineage-specific chromosomal topologies from coordinate gene regulation. Proc Natl Acad Sci U S A. 106(16):6679-84

Ringrose L, Ehret H, Paro R. (2004) Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. Mol Cell.16(4):641-53.

Savvidou E, Cobbe N, Steffensen S, Cotterill S, Heck MM. (2005) Drosophila CAP-D2 is required for condensin complex stability and resolution of sister chromatids. J Cell Sci.118(Pt 11):2529-43

Scaffidi P, Misteli T. (2006) Good news in the nuclear envelope: loss of lamin A might be a gain. J Clin Invest. 116(3):632-4

Scherthan H. A bouquet makes ends meet. Nat Rev Mol Cell Biol. 2001 2(8):621-7

Schotta G, Ebert A, Dorn R, Reuter G. (2003) Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. Semin Cell Dev Biol. 14(1):67-75

Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, Reuter G. (2002) Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. 21(5):1121-31

Sekinger EA, Moqtaderi Z, Struhl K. (2005) Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol Cell. 18(6):735-48

Sheehan MJ, Pawlowski WP. (2009) Live imaging of rapid chromosome movements in meiotic prophase I in maize. Proc Natl Acad Sci U S A. [Epub ahead of print]

Shumaker DK, Dechat T, Kohlmaier A, Adam SA, Bozovsky MR, Erdos MR, Eriksson M, Goldman AE, Khuon S, Collins FS, Jenuwein T, Goldman RD. (2006) Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. Proc Natl Acad Sci U S A. 103(23):8703-8

Segal E, Widom J (2009) What controls nucleosome positions? Trends. Genet. 25(8):335-43

Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. (2005) Interchromosomal associations between alternatively expressed loci. Nature. 435(7042):637-45

Ulyanova NP, Schnitzler GR. (2005) Human SWI/SNF generates abundant, structurally altered dinucleosomes on polynucleosomal templates. Mol Cell Biol. 25(24):11156-70

van Holde K, Zlatanova J. (2007) Chromatin fiber structure: Where is the problem now? Semin Cell Dev Biol. 5:651-8

Vazquez J, Belmont AS, Sedat JW. (2001) Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. Curr Biol. 11(16):1227-39

Vazquez J, Belmont AS, Sedat JW. (2002) The dynamics of homologous chromosome pairing during male Drosophila meiosis. Curr Biol. 12(17):1473-8

Verschure, P.J. et al. (1999) Spatial relationship between transcription sites and chromosome territories. J. Cell Biol. 147, 13–24

Wang Y, Ng TY. (2006) Phosphatase 2A negatively regulates mitotic exit in Saccharomyces cerevisiae. Mol Biol Cell. 17(1):80-9

Watson JD, Crick FH (1953) A Structure for Deoxyribose Nucleic Acid Nature. 171:964–967

Weiler KS. (2007) E(var)3-9 of Drosophila melanogaster encodes a zinc finger protein. Genetics.

Weiler KS, Wakimoto BT. (2002) Suppression of heterochromatic gene variegation can be used to distinguish and characterize E(var) genes potentially important for chromosome structure in Drosophila melanogaster. Mol Genet Genomics. 266(6):922-3

Yang G, Leuba SH, Bustamante C, Zlatanova J, van Holde K. Role of linker histones in extended chromatin fibre structure. Nat Struct Biol. (11):761-3.

Zirbel RM et al., (1993) Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries. Chromosome Res. 1, pp. 93–106

APPENDIX A

DROSOPHILA CONDENSIN II INDUCES INTERPHASE GENOME SPATIAL REORGANIZATION VIA CHROMOSOME AXIAL SHORTENING

Christopher R. Bauer, Thomas A. Hartl, and Giovanni Bosco

Introduction

The nucleus is a highly organized and dynamic organelle. Within the nucleus, chromosomes are also highly organized (Takizawa, 2008). The spatial arrangement of DNA elements and other factors within the three dimensional confines of the nucleus is crucial for proper gene expression (Kosak, 2004)(Takizawa, 2008), DNA repair (Misteli, 2005), DNA replication (Muck and Zink, 2009), chromosome segregation (Haeusler and Engelke, 2006), cellular aging (Scaffidi, 2005), and many other processes that are central to basic cellular function (Misteli, 2007). Defects in nuclear organization have been implicated in cancer, aging, cardiomyopathy, and evolution (Lever and Sheer, 2009; Shumaker, 2006; Roix et al., 2003). However, the details of the regulatory mechanisms that establish and maintain the organizational patterns of chromosomes remain enigmatic.

Specific chromosomal loci colocalize with structures such as the nuclear lamina, nuclear pores, the nucleolus, and other chromosomal loci (Branco and Pombo, 2006; Lamond and Spector, 2003; Cremer and Cremer, 2001). It is also known that environmental cues can induce changes in these localization patterns (Chaung 2006). Similar processes are known to affect entire chromosomes as well. Specific chromosomal territories exist in many contexts and the spatial relationships between these territories are regulated (Branco and Pombo 2006, Kosak 2009). Dosage compensation mechanisms can serve to enhance or reduce transcription of thousands of genes on a chromosome simultaneously by changing chromosomal architecture (Augui 2007). Perhaps the most remarkable structural reorganization of chromosomes occurs

during mitosis when they are condensed, resolved of catenations, spatially segregated, and decondensed, all while maintaining the fidelity of the underlying DNA sequences and epigenetic marks.

Though many patterns of chromosome organization have been well characterized, the molecular processes responsible for these patterns are not well understood. This is especially true during interphase. It remains to be determined how chromosome territories are established and maintained, how interactions between distant chromosomal loci are mediated, and which proteins function to generated the forces necessary to organize and reorganize the interphase genome.

One of the biggest technical challenges in understanding the organization of chromosomes is their scale. Chromosomes are the largest biological molecules known to exist. They are also too dynamic to be structurally characterized by X-ray diffraction or other techniques that are useful for studying protein or small nucleic acid structures. Despite their size, the fundamental units of chromosome organization are far too small to be studied by light microscopy. The giant polytene chromosomes in the larval salivary glands of Drosophila melanogaster and other dipteran insects provide one solution to this problem. Many researchers have utilized this system to investigate chromosome structural characteristics (Zhimulev, 1996; Zhimulev, 2004; Mathog, 1984; Mathog, 1985; Mathog, 1989; Urata, 1995).

Many tissues in larval flies and some adult tissues contain polyploidy nuclei where homologous chromosomes and sister chromatids are aligned and tightly

associated. Salivary gland nuclei have exceptionally high ploidy (up to 1024 or 2048 C) and thus display, in stunning detail, a reproducible pattern of dense bands and diffuse interbands that are thought to represent different states of chromatin organization (Zhimulev, 2004). Detailed drawings of these banding patterns were produced by Bridges in the 1930's and have since provided a valuable tool to identify the cyotogenetic locations of numerous genes, histone modifications, and DNA binding proteins on squashed chromosomes (Ringrose et al., 2004). Banding patterns have also been used to determine how chromosomes are folded in three dimensions within intact nuclei (Agard and Sedat, 1983; Mathog et al., 1985).

One feature that is lacking in the salivary gland polytene chromosome system is the ability to perturb nuclear organization and assess the effects. Only through an approach such as this is it possible to gain an understanding of how nuclear structure relates to function. The nurse cells in the Drosophila ovary provide a unique alternative to salivary glands (Spradling, 1993 #1262). These germ line cells develop huge nuclei that are similar in ploidy to salivary gland nuclei (Lilly, 1996; Koryakov 1997; Dej, 1999). The Drosophila ovaries also contain egg chambers from all the developmental stages of oogenesis. This makes it is easy to observe developmental changes in nurse cell nuclear morphology and chromosome structure within a single preparation. Furthermore, nurse cell chromosomes undergo a developmentally regulated shift in chromosome structure from an aligned, polytene arrangement to a dispersed structure (Dej and Spradling, 1999). The progression of nurse cell chromosomes from polytene to non-polytene structure is perturbed in a variety of RNA binding protein mutants such as *otu*,

cup, *sqd, and glo* (Koryakovet al., 2004; kalifa et al., 2009). Unfortunately, these mutants have provided little insight into the mechanisms that control chromosome structure. However, one group of genes, known to directly interact with chromatin, may provide some clues.

Reorganization of the nurse cell nucleus has recently been shown to be dependent upon condensin II. Mutantions in several subunits of this complex cause persistent polytene chromosome structure resulting in nurse cell nuclei that resemble those of salivary glands (Hartl et al., 2008). Conversely, expression of one condensin II subunit, Cap-H2, in salivary glands is capable of inducing dispersal of polytene chromosomes to produce nuclei that resemble those of wild type nurse cells (Hartl et al., 2008). These two cell types thus provide complimentary systems in which condensin function can be modulated to investigate genomic organization, developmentally regulated reorganization, and the biological consequences of different organizational patterns.

The *Drosophila* condensin II complex consists of a heterodimer of SMC2 and SMC4 (*gluon*) as well as the kleisin subunit, Cap-H2, and the HEAT repeat protein, Cap-D3. The SMC subunits also contribute to the condensin I complex along with Cap-H (barren), Cap-D2, and Cap-G (Hirano, 2005; Hirano, 2006; Nasmyth, 2005). The precise *in vivo* functions of the condensin complexes have remained elusive despite many efforts, though some general features are understood. It is clear that condensins play crucial roles in chromosome structural organization, maintenance of genomic integrity, and mitotic

and meiotic chromosome segregation (Losada and Hirano, 2005). In C. elegans condensins have also been shown to regulate dosage compensation (Tsai et al., 2008).

The condensin complex was originally identified as a factor necessary for chromosome condensation in Xenopus egg extracts (Hirano, 1997). The *Drosophila* gene *barren,* a member of this complex, had previously been shown to have defects in chromosome segregation and to interact with topoisomerase II (Bhat, 1996). A similar phenotype was also demonstrated for YCS4, a yeast condensin subunit (Bhalla, 2002). Interestingly, the Drosophila and yeast condensin mutants did not have apparent chromosome condensation defects. Condensin knockout studies in chicken DT40 cells similarly revealed that chromosome condensation did occur, despite some delays (Hudson, 2003). Only recently has a report directly implicated condensin as a factor required for chromosome condensation in vivo. Several condensin mutants fail to sufficiently shorten mitotic X chromosomes in worms, yet wild type chromosomes are only 20-38% shorter than mutant chromosomes (Mets, 2009). If condensin is truly a major contributor to chromosome condensation *in vivo*, as its name would suggest, it must be responsible for reducing chromosome lengths by more than 38% given that mitotic chromosome condensation results in 4-50 fold chromatin compact relative to interphase (Belmont, 2006).

Here, we have utilized the giant nuclei of the *Drosophila* ovarian nurse cells and salivary glands to allow detailed analysis and quantitation of nuclear structure in 3 dimensions. By genetically modulating condensin II function, we have assessed

numerous, dramatic changes in genomic organization that are regulated by this complex. We report that *Drosophila* condensin II functions to reduce interphase chromosome length by approximately 10 fold. Condensin II also induces the dispersal of chromatids from the euchromatic arms of polytene chromosomes yet not from the heterochromatic or telomeric regions. Finally, condensin II functions to promote chromosome territory formation, while inhibiting chromocenter formation. These effects are remarkably similar in salivary gland and nurse cell nuclei suggesting that they are general features of condensin function. Though our studies were conducted in interphase cells, it is likely that condensin II contributes to similar processes during mitosis. These findings also provide insights into the underlying structure of polytene chromosomes and many general features of nuclear architecture.

Methods

Fly Strains

The following fly strains were used in this study: y[1] w[1118] (wild type control for all nurse cell studies), y[1] w[67c23]; P(w[+mC] y[+mDint2]=EPgy2)Cap-D3[EY00456]/Df (2L)Exel7023 (Cap-D3 mutant), y[1] w[67c23]; P(w[+mC]=lacW)glu[k08819]/+; Cap-H2$^{Z3\text{-}0019}$/+ (SMC4, Cap-H2 double heterozygote), Cap-H2^{0019}/ Df (R3)Exel6159 (Cap-H2 mutant), Oregon-R-S (wild type control for all salivary gland studies), w[*]; P(w[+mC]=GAL4-nos.NGT)40 P(w[+mC]=lacO.256x)43 P(lacO.256x)50F P(lacO.256x)57A P(lacO.256x)60AB/CyO; P(w[+mC]=UAS-GFP.lacI)1.2/TM3, Sb[1] (lacO array 50F insertion line), w[*]; P(w[+mC]=lacO.256x)60F (lacO array 60F insertion line)

Flies were maintained at 25 degrees C on standard cornmeal molasses media.

Labeling of Probes

BAC clones were selected to span the X chromosome at regular intervals. The following BACs were ordered from CHORI BACPAC Resources: BACR25p24, BACR18c23, BACR23m08, BACR32h11, BACR33k15, BACR32l12. These map approximately to 3.9, 8.3, 12.9, 18.5, 20.5, and 22.1 Mb, respectively. To mark the heterochromatic region of the X chromosome we synthesized PCR products to the 359 bp repeat sequence using the primers 5'-CGGTCATCAAATAATCATTTATTTTGC-3' and 5'-CGAAATTTGGAAAAACAGACTCTGC-3'. We also used the following

oligonucleotides to label the 2L, 2R, 3R, and 4 heterochromatic regions, respectively: $(AATAG)_5$, $(AACAC)_5$, 5'-CCCGTACTGGTCCCGTACTGGTCCCGTACTCGGTCCCGTACTCGGT-3', and $(AATAT)_5$. Telomeric probes were made by pcr amplification of plasmids containing a 2 kb ApaI–ApaI fragment of the 3′ UTR of HeT-A and the 23Zn-1 fragment containing ORF1+ORF2 of HeT-A

Midipreps were performed on bacterial cultures to isolate BAC DNA. In some cases, probes were made by whole genome amplification (Sigma WGA kit) of previously isolated DNA. 20 ug of BAC DNA or amplified DNA was digested overnight with the restriction enzymes AluI, HaeIII, MseI, MspI, and RsaI . DNA fragments were purified using a Qiagen PCR cleanup column, ethanol precipitated, and end labeled with aminoallyl dUTP (Invitrogen ARES Alexa Fluor DNA Labeling Kit) using terminal deoxytransferase as described in (Hartl, 2008 #3429). DNA fragments were purified again on a PCR cleanup column (Qiagen) and ethanol precipitated. These labeled DNA fragments were then conjugated to Alexa488, Alexa546, or Alexa647 dyes as described in the ARES DNA labeling kits. Labeled probes were purified on a PCR cleanup column, ethanol precipitated, and resuspended in 20 uL hybridization buffer.

Fluorescence in-situ hybridizations/Imaging.

FISH protocol was adapted primarily from (Dernburg, 2000). 0-2 day old female flies were fattened on yeast with males at 25 degrees C for 2 days. 5 ovary pairs were dissected in Grace's medium and fixed within 15 minutes. Fixation was performed for 4

minutes in 100 mM Sodium cacodylate,100 mM sucrose, 40 mM Sodium acetate, 10mM EGTA with 3.7% formaldehyde. Ovaries were washed twice for 5 minutes each in 2X SSCT and individual ovarioles were separated in 2X SSCT. Ovarioles were washed for 10 minutes each in 2X SSCT with 20%, 40%, and 50% formamide at room temperature and once more for 2 hours in 2X SSCT with 50% formamide at 30 degrees C (for AT rich heterochromatic probes) or 37 degrees C. 200-500 ng of probe was mixed into a total of 40 uL hybridization buffer, denatured at 95C, and snap frozen in liquid nitrogen. Formamide mixture was aspirated from ovarioles and replaced with probe/hybridization buffer mixture. This suspension was mixed gently and incubated at 30 or 37C for 10 minutes prior to a 2 min denaturation at 92oC. Tubes were immediately returned to hybridization oven and left overnight at the specified temperature.

The next day, ovarioles were washed 4 times for 30 minutes each in 2X SSCT with 50% formamide at the hybridization temperature. Ovarioles were then washed 10 minutes each in 2X SSCT with 40% and 20% formamide. Three 5 minute washes with 2X SSCT were performed to remove to remaining formamide. To visualize the nuclear envelope, ovarioles were incubated for 20 minutes in 2X SSCT containing __mg/ml wheat germ agglutinin conjugated to Alexa488. The ovarioles were then stained with 10ng/mL DAPI in 2X SSCT for 10 minutes and washed twice with 2X SSCT for 10 minutes each. Ovarioles were mounted in vectashield. Pieces of number 1½ coverslips were used as spacers between the actual coverslip and slide to prevent flattening of the egg chambers. Slides were imaged on a Zeiss 510 meta confocal microscope using a 1.4 NA, 63X objective.

Nurse cell Image Analysis

Individual channels were separated from 3D image stacks and analyzed using the 3D Object Counter plugin for imageJ. The spatial coordinate and intensity measurements were then analyzed in the R programming environment (R scripts available in supplements). Intensity thresholds were set as follows: 25 for DAPI, 35 for Alexa488 and Alexa647, 55 for Alexa546. Objects smaller than 10 voxels were ignored. Spot counts for each probe are simply the number of objects detected. For each probe in each nucleus, the centroid was defined as the center of total fluorescence intensity. Principal component analysis was performed using point weights equal to the total fluorescence intensity of each object. Total dispersion is defined as the mean of the three principal components. The center of fluorescence intensity in the DAPI channel was set as the nuclear center and radial distances for each object were measured in relation to this point.

Salivary Gland Imaging and Measurements

Transgenic lines carrying 256-repeat array of the Lac-O site at chromosomal position 50F were crossed to lines carrying the same Lac-O site inserted at position 60F and transgenes Hsp70>Gal4 and UAS>Cap-H2, as previously described (Hartl, 2008 #3429). These lines also had the ability to express a GFP-LacI fusion protein that binds to the LacO arrays and marks the chromosomal insertion site of the LacO array. Expression of GFP-LacI and Cap-H2 was controlled with heatshock at 32oC, as detailed in the text and figure legends. GFP-LacI in salivary glands was imaged as previously described (Hartl, 2008 #3429). Distances between the two LacO arrays was measured in

3D as described for nurse cell imgages. The nuclear envelope was stained with wheat germ agglutinin conjugated to Alexa488.

Results

Condensin II induces the dispersal of euchromatic loci

Previous work has shown that Drosophila Cap-H2 is necessary for the dispersal of nurse cell polytene chromosomes and can induce dispersal of salivary gland chromosomes (Hartl et al., 2008). We first sought to characterize the effect of condensin on global chromosome architecture in order to better understand the mechanism by which condensin acts as well as the underlying chromosomal structure upon which it acts. To address this, we synthesized fluorescently labeled DNA probes to six euchromatic loci on the X chromosome as well as probes specific to the telomeric repeats and the pericentric heterochromatin of all 4 chromosomes. We used these probes to perform FISH in the egg chambers of wild type flies and of three allelic combinations of mutated condensin subunit with a range of severities in phenotype (Figure 1). Three dimensional image stacks of all developmental stages were collected by confocal microscopy. These images were then analyzed using the 3D object counter plugin for imageJ to determine the number of distinct fluorescent foci (spots) as well as the volume, intensity, and spatial coordinates of each spot.

Consistent with previous findings (Dej et al., 1999)(Hartl et al., 2008), our analysis revealed that nurse cell polytene chromosomes undergo a rapid transition from aligned chromatids to a more dispersed structure between stages 5 and 6. In the stage 5 egg chambers of wild type and condensin II mutants, all euchromatic FISH probes mark a single spot within each nucleus (data not shown) showing that all homologous loci are

alligned. However, in wild type stage 6 egg chambers, all probes targeting the Xchromosome mark a number of distinct foci due to the dispersal of the chromatids. The number of foci observed varied greatly depending on the chromosomal position (Figure 2A). Unexpectedly, we found that the number of spots was always greatest in the region approximately 8 Mb from the telomere. Toward the telomere and centric heterochromatin, chromatids remained more tightly associated and aligned. This pattern was generally preserved in later stage egg chambers though the number of spots visible at almost all loci tended to increase gradually as the nuclei increased in size and ploidy. A region between 18 and 20 Mb did not show much dispersal in stage 6 egg chambers but was highly dispersed by stage 8 (Figure 2). The maximum number of spots that we observed for any probe, in any nucleus was 31, consistant with a previous report of 32 based on squashed chromosome preparations (Dej, 1999). We believe that 32 is the true number of chromatid bundles as it can be derived from a dilploid nucleus by 5 rounds DNA synthesis. However, we only observed 31 spots in a single stage 10 nucleus with a probe to the locus at 8.3 Mb. Maximum dispersal is never reached at earlier developmental stages and the chromatid bundles at most chromosomal loci never completely separate.

The chromatid separation at all euchromatic loci was highly dependent upon condensin II. In all of the mutant allelic combinations we tested, dispersal was significantly reduced from wild type levels (Figure 2). In flies homozygous for a mutation in Cap-H2, chromatids were almost always tightly associated (Figure 1A and Figure 2). Occasionally, we observeded two spots with similar size and intensity. This

was probably due to asynapsis of homologues and was slightly more common near the centric heterochromatin as has been observed in salivary gland polytene chromosomes (Evgen'ev and Polianskaya, 1976). Chromosomes in Cap-D3 mutants were similiar, though this genotype did exhibit slightly more dispersal near the centromere (Figure 2). In SMC4, Cap-H2 double heterozygous flies, we observed more dispersal than in the other mutants. The pattern of dispersal in the double heterozygous nurse cells was generally similar to that of wild type flies; however, the number of spots visible at any locus was always far less than that observed in the wild type (compare Figure 1C to 1D; Figure 2). This lack of dispersal in the condensin mutants cannot be explained by massive under-replication of sequences along the X chromosome, as our previous studies showed that condensin mutants have no defects in nurse cell DNA replication or ploidy (Hartl et al., 2008). From these data we conclude that the repetive sequences at the ends of the X-chromosome are not susceptible to the condensin II dependent processes that induce chromatid dispersal at euchromatic loci.

Condensin does not disperse centric or telomeric heterochromatin

The dispersal pattern of the X chromosome centric heterochromatin differed substantially from that of the euchromatin. While nearly all euchromatic regions exhibited condensin II dependent dispersal only after stage 5, the dispersal of the heterochromatin began earlier and was not strongly correlated with condensin activity (Figure 2). We thus investigated the behavior of the 2^{nd}, 3^{rd}, and 4^{th} chromosome

heterochromatin to determine if this was a general feature of heterochromatic DNA. FISH revealed that heterochromatic regions, particularly on the X, had already begun to disperse by stage 5 in some nuclei (data not shown). By stage 6, nearly all nuclei displayed more than one spot for all heterochromatic probes (Figure 4). In wild type nuclei, the heterochromatin was generally less dispersed than euchromatic loci. This could be due to under-replicaiton as previous reports show that heterochromatic sequences are under-replicated in nurse cells (Lilly, 1996)(Dej, 1999)(Royzman, 2002). Similar to the X, the presence of condensin II activity had little bearing on the degree of heterochromatic dispersal in the 2nd and 3rd chromosomes. The number of spots was very similar for the 2nd and 3rd chromosomes, but the X always exhibited more dispersal. (Figures 3 and 4). The heterochromatin of the 4th chromosome was an exception. This was the only heterochromatic region to show a clear dependence upon condensin activity (Figure 4). Interestingly, it did not show a large amount of dispersal in stages 6 and 8 but reached levels comparable to that of many euchromatic loci by stage 10 (Figures 2 and 4).

Similar to the centric heterochromatin, telomeric regions were much less dispersed than the euchromain in wild type nurse cell nuclei (Figure 4). In the polytene nurse cell chromosomes of Cap-H2 mutants, as well as in squashed wild type salivary gland polytene chromosomes, probes to the telomeres usually marked 5 separate loci (one on the X and two on each major autosome). Thus, five spots corresponds to the absence of dispersal. Nurse cell nuclei in stage 6 and stage 8 egg chambers almost always had between 4 and 8 spots regardless of condensin II activity suggesting minimal amounts of

telomere dispersal (Figure 4). Telomeric probes did mark 17.7 spots, on average, in wild type stage 10 nuclei. However, this would equate to only 3 or 4 spots at each telomere which is far less than euchromatic loci. It is possible that this dispersal was due to condensin activity, though the Cap-D3 mutant nurse cells also had a similar number of telomeric spots at this stage (Figure 4) and these mutants had minimal dispersal of euchromatin loci suggesting very little condensin activity. Perhaps a version of the condensin II complex that lacks Cap-D3 can function at the telomeres. Regardless, the telomeres of nurse cell chromosomes are much more tightly associated than loci within the euchromatin at all stages of development in wild type nurse cell nuclei.

Chromatid Dispersal is Topologically Restricted

Once we had determined the effect of chromosome position chromatid dispersal for the X chromosome, we sought to analyze the spatial relationships between homologous chromatid bundles as they separate from each other in wild type nurse cells. To this end, we determined the x, y, and z coordinates of all FISH spots for any given locus. For all X chromosome probes that displayed at least 5 spots in a given nucleus, we performed principal component analysis, weighting each spot by its total fluorescence intensity. The three principal components are three mutually orthogonal vectors which describe the direction and magnitude of spatial dispersion with respect to the center of fluorescence intensity. The first principal component (PC1) provides the axis of maximal dispersion as well as its magnitude, the second (PC2) provides the magnitude and

direction of maximal dispersion in a plane orthogonal to PC1, and the third (PC3) provides the magnitude of dispersion in the remaining dimension. Alternatively, the magnitudes of the principal components correspond to the variances of a 3 dimensional Gaussian distribution. Thus, these quantities describe the total amount of spatial dispersion as well as the shape of this dispersion.

If a number of molecules begin to diffuse freely from a single point in space, these molecules will tend to disperse evenly in all directions producing a spherical ellipsoid. Principal component analysis of such dispersal pattern would return three vectors with similar magnitudes (Figure 5D black). Alternatively, a number of molecules that have dispersed from a single point due to some active, directed process, or by diffusion with topological constraints, will exhibit significant differences in the magnitudes of the principal components. Specifically, if the ratio of the magnitudes of the first and second principal components (PC1/PC2) is large, this implies that the shape of dispersal is a prolate ellipsoid (i.e. hotdog shaped, Figure 5D blue). If the ratio of the magnitudes of the second and third principal components (PC2/PC3) is large, this implies that the shape of dispersal is an oblate ellipsoid (i.e. hamburger shaped, Figure 5D red).

Thus, for each set of spots observed with a given probe in a given nucleus, we can describe the spatial character of the dispersion by plotting the ratios PC1/PC2 vs. PC2/PC3 (Figure 5). Points that lie near the point (1,1) represent loci that have dispersed equally in three dimensions. Points that lie near the y-axis represent loci that have dispersed primarily in two dimension and points near the x-axis represent spots that have

dispersed primarily in one dimension. This analysis revealed several interesting trends. In stage 6 egg chambers, where little time had passed since the initial chromosome dispersal, we find that dispersion of most euchromatic loci is confined to two dimensions. The most likely explanation for this effect is that individual chromatids remain roughly aligned during the initial phase of dispersal and move radially outward from the central chromosome axis. This pattern is similar to what was often seen in stage 10 egg chambers with reduced condensin function (Figure 1C). This observation is also consistant with previous structural analyses of salivary gland polytenes revelaing a toroidal configuration of chromatids (Urata, 1995). Later in development, euchromatic loci tend to lose their axial alignment and exhibit a more spherical shape (Figure 5A).

This spatial arrangement of dispersing heterochromatic loci on the X was quite different from that observed for the euchromatic regions. There was no obvious spatial pattern to the dispersal of heterochromatin loci in the nuclei of stage 6 and stage 8 egg chambers. However, by stage 10, many heterochromatic loci were arranged in a highly linear pattern (Figure 5). This increase in order suggests that heterochromatic sequences are either being actively directed into a linear arrangement or that they diffuse randomly but become anchored to some nuclear scaffold that possesses a linear organization.

Condensin is Necessary for Axial Compaction of Nurse Cell Chromosomes

In addition to the chromosome dispersal activity of Condensin II, we also observed a reduction in the axial lengths of chromosomes in response to Condensin II

activity. A comparison of the mean distances between several pairs of probes at different developmental stages clearly shows that condensin II is required for a dramatic compaction of the entire X chromosome (Figure 6). Without exception, the mean distances between loci on the X chromosome were greater in Cap-H2 mutants than in the wild type. In nearly all cases, the SMC4, Cap-H2 double heterozygote had chromosomes of intermediate length, between the wild type and strong Cap-H2 mutant (Figure 6). It may appear that the Cap-D3 mutant has similar chromosome lengths to the Cap-H2 mutant, but we believe the chromosomes are actually much shorter in the Cap-D3 mutants as well. A few considerations must be taken into account in order to properly interpret this data.

First, the average distance between two random points within a sphere is equal to 36/35r, where r is the radius of the sphere. Thus, we should never observe a mean distance between any two loci that is much greater than the nuclear radius. Secondly, polytene chromosomes exhibit a large amount of curvature. Thus, the average distance measured between any two loci, separated by at least 1 Mb, on a polytene chromosome will be an underestimate of the actual distance that would be measured if we could trace along the path of the chromosome. These two considerations are exemplified nicely by the distance measurements in the Cap-H2 mutants, which exhibit strong polyteny (Figure 5, red bars). As the cytogenetic distances between loci increase, the 3-dimensional, linear distances between those loci quickly plateau to a distance equal to the nuclear radius. Given this, maximum distances are likely to be better estimates of chromosome length than mean distances for curvaceous polytene chromosomes. Furthermore, it seems that

condensin activity induces a reduction of the curvature in chromosomes (compare Figure 1A with 1C; Figure 7C with 7F). Thus, while condensin II activity reduces the distances between loci along the path of the chromosome, it also reduces the likelihood of underestimating the linear distances between loci. Therefore, the best estimates of the lengths of chromosomal regions come from probes that are near each other in cytogenetic space, which are less likely to have extensive curvature between them. In this case, that is the probes that are 2Mb apart. Finally, chromosome length and the regulation of this length by condensin II are best assessed in nuclei from egg chambers at late stages of development. One simple reason for this is that they are larger and distances measurements can be calculated with more accuracy. Large nuclei also allow loci to reach a greater maximal distance so underestimation of chromosome length is less likely. Additionally, it is possible that condensin II is acting throughout development and that it does not induce maximal compaction until stage 10.

Consideration of all of these factors suggests that we should focus on the distance measurements of the 2 Mb chromosomal segments in stage 10 egg chambers and place more weight on maximum distances rather than mean distances for chromosomes with substantial polytene structure. Cap-H2 mutants displayed a completely penetrant nurse cell polytene phenotype and thus probably have negligible condensin II activity. In this genotype, the two loci that are 2Mb apart (cytogenetic locations 18.5 Mb and 20.5 Mb) were separated by a maximum of 13 um within the 3-dimensional space of the nucleus. In a few cases, it was possible to trace the path of the polytene chromosome from one probe to the next and confirm that minimal chromosome curvature was present. Thus, we

are highly confident that 13 um is the true length of this chromosomal segment. Extrapolation from this figure leads to an estimate of 145 um for the entire euchromatic region of the X chromosome in Cap-H2 mutants. This fits very well with a previous estimate of 140 um for the length of the polytene X chromosome in intact salivary glands, which have comparable ploidy to nurse cells (Hochstrasser and Sedat, 1987).

The Cap-D3 mutants seem to have a small amount of condensin II activity as they still maintain some nurse cell polytene structure, evident by some visible bands and minimal dispersal of chromatids. In this genotype, the maximum distance observed between the pair of probes spaced 2Mb (same as above), was 7.7um. The SMC4, Cap-H2 double heterozygotes, must possess a moderate amount of condensin activity since they have at least one wild type copy of the genes for all condensin II subunits. They also lack any visible bands in nurse cell chromosomes. In this genotype the maximum distance was 4.7 um. Since these chromosomes do not appear to have a large amount of curvature, the mean distance (3.5 + 0.56 um) may be a better estimate if chromosome length is very sensitive to condensin II dosage.

In wild type nuclei we find that the maximum length of the same 2Mb region is 1.6 um. This suggests that, in nurse cells, condensin II is responsible for an 8 fold reduction in chromosome axial length. Though the high degree of chromosome dispersal present in stage 10 wild type nurse cell nuclei may make it seem unintuitive to measure the distances between centroids, at least two lines of evidence indicate that our estimate is reasonable at the least, and it may even be too conservative. One indication is that the 3-

dimensional distances between loci continue to increase as the cytogenetic distance increases (Figure 6C, black bars). Even when separated by 18.5 Mb, more than 80% of the length of the entire euchromatic arm, the distance between loci is never greater than 12.2 um. This leads to an estimate of more than 9 fold compaction. The amount of curvature present in the dispersed chromatids cannot be determined based on our data. The trend suggested by the other genotypes may suggest that there is very little and that the mean distance between probes (0.8 ± 0.14 um) may better reflect the true length of this chromosome segment. If this is true, then the chromosomes could be compacted as much as 16 fold in the presence of condensin II. Data from salivary gland chromosomes provide additional insight.

Condensin Induces Axial Compaction of Salivary Gland Chromosomes

It was possible that the effects of chromosome axial shortening we observed were due to some unique feature of nurse cell chromosome dynamics. Therefore, we overexpressed Cap-H2 in salivary gland cells to see how a change in condensin activity would affect the structure of these chromosomes. This was performed in a fly strain that was hemizygous for two different insertions of tandem LacO repeats approximately 10.8 Mb apart on the right arm of the 2nd chromosome. We found that salivary gland chromosomes, in response to expression of Cap-H2, behaved quite similar to those in wild type nurse cells (Figure 7). Cap-H2 expression was driven by a heat shock promoter so it was possible to vary the amount of expression by changing heat shock duration or

temperature. With a short heat shock or no heat shock at all, we were able to produce phenotypes with shortened, but still highly aligned chromosomes, similar to those seen in SMC4/+; Cap-H2/+ and Cap-D3/Df nurse cell nuclei, respectively (Figure 7 and figure 1). Long heat shock times resulted in completely dispersed chromatids that closely resembled wild nurse cell nuclei (Figure 7 and Figure 1).

In order to measure the distances between the two lacO arrays, we performed an 8 hour heat shock to induce a maximal amount of chromosome shortening without extensive dispersal. For nuclei that possessed two distinct GFP bands, the mean 3-dimensional distance between the LacO arrays was 6.0 + 0.40 um. This suggests that the entire right arm of the 2nd chromosome was reduced to approximately 12 um in these cells. Since each chromosome arm was distinguishable in many of the nuclei, we are confident that this is an accurate estimate of the chromosome length. This figure also matches our conservative estimate for the length of the similarly sized X chromosome in wild type nurse cell nuclei. Based on this and a previous measurement of the 2R chromosome arm in the same cell type (Hochstrasser and Sedat, 1987), condensin II can induces chromosome axial compaction of 12 fold. Since longer heat shock times lead to a more severe dispersal phenotype, additional compaction is quite possible.

We also observed that the LacO array inserted at band 50F, near the center of the chromosome arm, was always wider than the band at 60F, near the telomere. The centrally located region also began to disperse prior to the subtelomeric region (Figure 7). This matches the dispersal pattern observed on the X chromosome in nurse cells where

probes near the telomere show minimal dispersal relative to probes closer to the middle of the chromosome (Figure 2).

Condensin Inhibits Aggregation of Heterologous Heterochromatic Repeats

Measuring distances between loci on the same chromosome conclusively demonstrated that condensin II is required for compaction of DNA sequences in cis, but this analysis does not provide any information about the effects of condensin on the spatial relationships between loci on different chromosomes. This is a particularly interesting question for centromeric heterochromatic sequences since heterologous centromeres often aggregate into subnuclear compartments, also know as chromocenters. Previous reports indicate that nurse cells do not maintain a large chromocenter where all centromeres are clustered (Dej, 1999)(Royzman, 2002). However, salivary gland nuclei exhibit prominent, well organized chromocenters (Zhang and Spradling, 1995). Since the nurse cell chromosomes in Cap-H2 mutants had very similar morphology to wild type salivary chromosomes, we wondered if chromocenters might also be observed in nurse cells when condensin activity is absent.

To address this, we measured the distances between heterochromatic repeats on all 4 chromosomes. Consistent with the previous reports, we found that heterologous centromeric repeats had little tendency to associate with one another in wild type nurse cells (Figures 3 and 8). By contrast, the centromeric repeats on different chromosomes were strongly associated with one another in all condensin mutants compared to the wild

type (Figure 9). This was true for almost all pairs of centromeres and this effect was persistent throughout development (Figure 8). Similar to salivary chromocenters, the clustered centromers in the condensin mutant nurse cells were localized to the nuclear peripherary (Mathog, 1984 #3532). Interestingly, the distance between two heterochromatic probes on the left and right arms of the 2nd chromosome did not vary in response to condensin activity. This suggests that the heterochromatin is not susceptible to the condensin dependent axial compaction as observed in euchromatic regions (Figure 8). Perhaps heterochromatic regions are already maximally compacted. It is also possible that very subtle compaction of centric and pericentric regions occurs but is not detectable within our limits of resolution. Taken together, these data suggest that although condensin II does not have a critical role in dispersal of allelic heterochromatic sequences from sister-chromatids or homologs, it does play an important function in resolving inter-chromosomal associations between heterologous repeats.

Centromeres and Telomeres are Associated with the Nuclear Envelope in Nurse Cell Nuclei

Several lines of evidence suggest that gene expression can be affected by the radial position of a gene within a nucleus (Kosak, 2004)(Takizawa, 2008)(Misteli, 2007). Therefore, we hypothesized that a non-random radial organization of euchromatic and heterochromatic sequences may be present in nurse cell nuclei. To test this, we determined the geometric center of DAPI fluorescence in nurse cell nuclei and then

measured the distances of all loci on the X chromosome from this point. These distances were then normalized to an estimate of the nuclear radius that was based on the volume of DAPI staining (Figure 9). In stage 6 egg chambers, there was a tendency for centric and telomeric loci to localize near the nuclear periphery in both the wild type and Cap-H2 mutant nurse cells. Most euchromatic loci tended to be located more centrally within the nucleus (Figure 9A). This effect became increasingly exaggerated as development proceeded in Cap-H2 mutants. We were initially surprised to find that in wild type nurse cells nuclei all X chromosome loci move toward the periphery by stage 10 (Figure 9). This suggested that while centromeres and telomeres remain at the periphery throughout development, euchromatic regions are progressively moved toward the nulcear periphery in the presence of condensin II activity. When we consider that condensin induces substantial compaction in the euchromatin, the movement of euchromatic loci toward the periphery makes sense. If the X chromosome is only 12um long (Figure 6 and associate text), and both ends are located near the nuclear envelope, then the rest of the chromosome must be nearby as well given that the nuclear diameter is 20-30 um in stage 10 egg chambers.

When all of the findings presented above are considered together, an interesting model emerges. The 5 major chromosome arms are each reduced to about 12 um in length in stage 10 egg chambers, when wild type levels of condensin activity are present. The ends of all chromosome arms consist of either a centromere or a telomere. It is also clear that centromeres and telomeres maintain some associations between allelic loci and that both of these regions localize to the nuclear envelope. Since the nuclear diameter is

approximately 40 um in stage 10 egg chambers, this suggests that each chromosome would be confined to a compartment that represents only a small fraction of the nucleus. This is precisely the definition of a chromosome territory. A previous report suggests that chromosome territory formation in *Drosophila* primary spermatocytes is disrupted by mutations in Cap-H2 (Hartl et al., 2008). Thus, it seems likely that condensin II functions to promote chromosome territory formation and that this function is a direct consequence of euchromatic axial.

Discussion

Condensin proteins were originally identified by their ability to dramatically alter higher order chromatin structure in vitro (Hirano, 1997). This activity has long been thought to be one of the major forces responsible for physically compacting chromosomes along their axes during mitosis and meiosis. However, it has yet to be persuasively demonstrated that condensin function plays a large role in reducing the lengths of chromosomes in vivo. This is probably due to some redundancy in the functions of the condensin I, condensin II, as well as other factors. It is imposible to isolate mutants which completely disrupt mitotic chromosome condensation since they will prevent cell division. Here, we have utilized the highly polyploid tissues of *Drosophila* to provide a resolution at which chromosome structure and dynamics can be readily observed and quantified. We show that condensin II as a potent inducer of chromosome axial compaction and allelic loci separation in the euchromatin. These processes mirror each other in the larval salivary glands and in the germ line nurse cells suggesting they are general features condensin II funciton. These effects also suggest an intuitive model for the condensin II dependent formation of chromosome territories.

Characterization of chromatid dispersal along the length of the X chromosome in nurse cells provided several insights into the function of condensin II and the structure of polytene chromosomes. By stage 6, close to the time that Cap-H2 protein is first detectable within nurse cell nuclei (Hartl, 2008), the chromatids that make up polytene chromosomes have already begun to separate from each other. Initially, the DNA fibers

remain reasonably well aligned and disperse primarily in the plane orthogonal to the major chromosome axis. As time progresses, they become increasingly distant from each other and alignment along their lengths diminishes. This may suggests that some residual homology dependent associations are still present in stage 6 nurse cell nuclei which hold some chromatids together at various positions along their lengths and thus topologically restrict diffusion. Alternatively, it may be that persistent associations of telomeric and centromeric loci with the nuclear envelope are sufficient to limit diffusion to the plane orthogonal to the chromosome axis. This model seems most plausible if the chromosomes do not exhibit a large amount of curvature. It does seem that one of the results of condensin II activity is a reduction in chromosome curvature.

Though all euchromatic regions of the genome seem to be susceptible to the anti-pairing activity of condensin, and they all begin to disperse simultaneously, the degree of dispersal, as measured by the number of discrete fluorescence foci, varies significantly between different chromosomal loci. Specifically, dispersal is greatest near the center of chromosome arms and diminishes toward telomeric and centromeric regions. This result was unexpected. Since DNA replication results in catenations between sister chromatids (Downes et al., 1994; D'Ambrosio et al., 2008), we suspected that the sister chromatids in polytene chromosomes were also probably entangled, similar to the fibers in a piece of rope. Thus, as a rope begins to fray only at the ends, we anticipated that polytene chromosomes would first disassemble near the free DNA ends at the telomeres and the centromere of the X chromosome. This result seems to suggest, however, that sets of chromatids are arranged in parallel units that are free of catenations prior to dispersal.

It is possible that entanglements are resolved during the dispersal process via the enzymatic activity of a type II topoisomerase. Although condensin I is known to facilitated sister chromatin resolution by localizing topoisomerase II to chromosomes during mitosis in flies (Coelho et al., 2003), there are some problems with such a model. A previous study which analyzed nurse cell polytene dispersal in squashed nuclei concluded that nurse cell chromatid dispersal does not begin to occur until nuclei reach 64 C ploidy (Dej and Spradling, 1999). In order for 64 entangled chromatids to disperse into 32 pairs, topoisomerases would have to make two double strand breaks and pass a pair of double helices through both breaks. This would have to occur many times on each chromatid pair in short period of time which makes this seem unlikely. Furthermore, if another S-phase was initiated prior to the resolution of all entanglements, four double helices would have to be passed through another set of four double helices. Unlike the previous study, we find that dispersal continues throughout nurse cell development while multiple S-phases take place. Such a process seems increasingly implausible with each subsequent S-phase.

The disassembly of salivary gland chromosomes suggests that these polytene chromosomes are also composed of units that are free of entanglements. Even when Cap-H2 expression is driven by only 6 hours of heat shock, significant disassembly of polytene chromosomes is observed (Hartl, 2008). During these six hours, the media must reach temperature, the heat shock response must be initiated, the Cap-H2 transgene must be transcriped, the transcript must be spliced and processed, the mRNA must be exported from the nucleus, Cap-H2 protein must be translated, the protein must be imported back

into the nucleus, sufficient quantities of protein must accumulate, and only then can the chromatid dispersal process begin. Thus, it is likely that Cap-H2 is actually active in the nucleus for much less than six hours. Since feeding 3rd instar larva already have salivary gland ploidy that is close to the maximum of 1024 or 2048C (Rudkin, 1972), it would seem that dispersal of these polytene chromosomes would involve a huge number of topoisomerase II mediated breakages if the individual chromatids were entangled.

We have observed that salivary glands, in which Cap-H2 expression has been driven by heat shock for 24 hrs, can contain as many as 52 spots (data not shown). Interestingly, a detailed analysis of salivary gland polytene chromosome 3D structure suggested that they are composed of 25 to 53 units that are 200-400 nm in diameter (Urata and Parmelee et al, 1995). These units are interpreted to be bundles of 20-40 chromatids. It was also observed that polytene substructure was consistent between neighboring bands, even those separated by interbands. This could be taken as circumstantial evidence of a parallel structure in chromatid bundles.

Given that the number of spots we observed in nurse cells and salivary glands was always orders of magnitude less than the total number of chromatids, it must be the case that polytene chromosome structure is maintained by at least two levels of organization. One of these levels of organization can be disrupted by condensin and the other is condensin resistant. We believe the following model best fits the collective observations of previous studies and our own. During the first 4 or 5 endocycles, sister chromatids are resolved to produce 32 or 64 uncatenated double helices that remain tightly associated. Each of these chromatids then serves as a template to produce a chromatid bundle as

observed by Urata and Parmelee et al. In most tissues, these bundles remain tightly associated with each other resulting in polytene chromosomes. The associations between bundles can be disrupted by condensin II activity. However, the chromatid associations within the bundles are resistant to condensin. Catenations may be present within the bundles but the precise natures of both types of associations remain enigmatic.

Dej and Spradling also concluded that all euchromatic FISH probes marked exactly 32 spots in every nurse cell nucleus at stage 6 and all later developmental stages (Dej and Spradling, 1997). We confirm that 32 is indeed the maximum number of discrete spots. In our study, more than 300 nuclei were marked with FISH probes at 11 unique loci, and not a single nucleus was ever observed to have more than 32 spots. However, we usually found significantly fewer than this. Generally, the number of spots increased with the age of the egg chamber and with the distance of a locus from centromeres and telomeres. To reconcile the differences between our observations and those of Dej and Spradling, it is important to consider the different approaches and their respective limitations. The latter involved a very harsh treatment in which cells were exposed to acetic acid and then squashed flat on slides before FISH was performed. This method does not preserve information about the spatial relationships between loci in the nucleus and could break weak associations between chromatid bundles by denaturing necessary proteins or shearing DNA. Our analysis was conducted in whole mount, formaldehyde fixed tissues, which maintain much of the structure found in vivo. However, our method is unable to distinguish between chromatid bundles that co-localize

due to physical interactions, and those that are not physically linked but are not spatially separated enough to be distinguished by light microscopy.

We believe that the reduced dispersal detected at euchromatic loci near heterochromatic and telomeric regions is probably a result of these loci being physically tethered. Since we find that telomeres exhibit minimal dispersal and heterochromatic repeats exhibit limited dispersal as well, it is intuitive that loci near these regions could not disperse much more than their adjacent loci. This effect reduces the distance between spots and increases the probability that two spots will touch each other and be counted as a single spot. Squashing nurse cell nuclei is certainly capable of disrupting the condensin resistant associations between telomeres or heterochromatic sequences given that it is known to break under replicated regions of polytene chromosomes.

There are several possible explanations for the different dispersal patterns observed in telomeric, euchromatic, and heterochromatic loci in our whole mount preparations. One possibility is that the function of condensin at repetitive sequences is distinct from that of the euchromatin. It may play no role in the dispersal of these regions as we see a similar amount of heterochromatic dispersal in wild type and condensin II mutant nurse cells. We also see that condensin II does not induce measureable axial shortening of heterochromatic sequences. A second possibility is that associations with the nuclear envelope restrict the dispersal of telomeres and centromeres. If the tendency for these regions to localize to the nuclear periphery is due to some physical attachment, then these attachments would certainly inhibit dispersal. Even if these repeats could

move freely across the nuclear envelope, they would be confined to a two dimensional surface and thus would disperse less than loci free to move in three dimensions. A third possibility is that underreplicaiton of heterochromatin and telomeres holds these regions together. None of these models necessarily excludes the others.

The behavior of the chromosome 4 heterochromatin is a striking exception to the observations made at other heterochromatic loci. In wild type nurse cell nuclei it disperses to a similar extent as euchromatic regions, though this dispersal is delayed until after stage 8. This may be simple consequence of the small size of chromosome 4. If polytene chromosomes are held together by weak interactions along there lengths, then the forces that hold the 4^{th} chromosome together would be weaker than that of the other chromosomes. Alternatively, smaller molecules diffuse at a greater rates than large molecules and it is possible that, given enough time, we would see all heterochromatic regions behave in a manner similar to that of chromosome 4.

Although condensin does not play a substantial role in the dispersal of homologous heterochromatic sequences, we found that condensin does seem to limit or prevent the association of non-homologous heterochromatic repeats. It is consistently observed that salivary gland polytene chromosomes exist as five arms that all extend from a single chromocenter where all the centric heterochromatin colocalizes (Zhang and Spradling, 1995). It has previously been stated that the wild type nurse cell nuclei of Drosophila do not possess a chromocenter (Dej and Spradling, 1999). Our findings support this claim. However, when condensin function is reduced, the heterochromatic

regions of all chromosomes have a much greater tendency to co-localize in a single region, much like the salivary gland chromocenter. Again, the heterochromatin of the 4th chromosome is an exception to this in stage 10 nuclei since it completely disperses.

Perhaps the most revealing finding of this study is the dramatic axial compaction that condensin induces in chromosomes. This effect is strong in both the nurse cells and salivary glands suggesting that this is a general function of condensin and not a tissue specific effect. This observation is not necessarily a surprising one as the axial compaction we observed is reminiscent of the mitotic condensation activity that the condensin complex was originally named for. Yet, many previously studies have failed to prove that condensin is necessary for chromosome condensation in vivo. Mutants classically exhibit defects in chromosome segregation and some defects in mitotic chromosome morphology, but condensation appears reasonably normal along the chromosome axis (Hirano, 2005). A recent report has shown that mitotic chromosomes in condensin mutants are longer than in wild type cells but the length differences were modest (Meyer, 2009).

We show here that Drosophila condensin II can induce, at least, an 8 fold reduction in nurse cell chromosome length, and possibly more. In salivary gland chromosomes we are comfortable claiming that chromatin is compacted 12 fold relative to interphase. This is highly significant in terms of the total amount of axial compaction that is known to occur during mitosis. The naked relaxed B-DNA of the Drosophila X chromosome is roughly 7.5 mm long. The addition of histones would reduce this figure

to about 1.34 mm based on the predicted nucleosome occupancy of Drosophila (Widom and Segal 2009). Numerous models of the 30 nm fiber suggest that this structure would reduce the length of a nucleosomal array by approximately six fold resulting in a chromosome that is 224 um long (van Holde and Zlatanova, 2007). This is similar to the length of the salivary gland polytene X chromosome as described by Bridges', though analysis of whole mount nuclei suggests a figure of about 140 um (Hochstrasser and Sedat, 1987). Our estimate of polytene chromosome length in stage 10 nurse cells is practically identical to the latter suggesting that interphase chromatin is about 50% more compact than the 30nm fiber.

We find that condensin II can induce the formation of fibers that are 8-12 times more compact than those present in interphase, resulting in chromosome arms that are 12-17 um long. This chromatin structure would require additional reduction of roughly 3 fold to produce a chromosome with a length that is typical of a mitotic chromosome (Ashburner). It may be that we have underestimated the effect of condensin since the 12 fold figure comes from salivary glands that were heat shocked for 8 hours. A 24 hour heat shock produced a much more dramatic dispersal phenotype and may have produce levels chromatin compaction very similar to those in mitotic chromosomes. It is also notable that these condensed chromosomes are not individualized chromatids, like mitotic chromosome, but bundles comprising 16 to 64 sister chromatids. The lengths that these chromosomes can attain is even more remarkable when this is considered.

Though the chromosomes we have studied are not mitotic, the nurse cells are believed to undergo a mitosis like state after the fifth endocycle (Dej and Spradling, 1999) which is the same point at which condensin induces chromatid dispersal. Thus, it is reasonable to hypothesize that the function of condensin in this modified cell cycle state is similar to its role in typical cell cycles. There certainly must be some redundancy during normal mitoses since Cap-H2 mutants do not show any severe mitotic defects. The fact that expression of cap-H2 by itself in salivary glands is sufficient to induce axial compaction is somewhat puzzling. These cells are not thought to exhibit any mitotic characteristics. The SMC proteins are most likely the active part of the condensin II complex since they are ATPases. Interestingly, mRNAs for both of these genes are present at low to moderated levels in the larval salivary glands (Chintapalli et al., 2007). Perhaps they have a role during interphase as part of condensin I or possibly just as a heterodimer.

One important question that remains is whether the axial shortening of chromosomes and the dispersal of chromatid bundles are two results of a single molecular function of condensin or two separate processes. This is difficult to address since we are currently unable to separate these events. Mutations in different subunits have differences in phenotype severity, but the axial compaction and dispersal phenotypes in the nurse cells are correlated. One model that our data argue against is that condensin II forms an axial scaffold along the core of the chromosomes. If this were the case, then it would be unlikely that we would observe such a wide phenotypic range. How could leaky expression from a hsp70 promoter produce a small amount of Cap-H2

protein in the salivary glands and induce very minor compaction across the entire chromosome by the formation of a small amount of scaffold (see figure 7)? We do observe shortening of chromosomes prior to dispersal, but this may be due to the scale of the distances we are trying to measure. Two allelic loci could not be observed as individual spots until the distance between them have reached a few hundred nanometers. A major obstacle in addressing the issue of mechanism(s) is that no studies have provided real insight into the molecular nature of the changes in chromatin structure that condensin II can induce in vivo.

So far, the best insights have been gained by in-vitro studies. These have shown that condensin can induce and/or stabilize positive supercoils in plasmid DNA or tie plasmids into knots when topoisomerases are present (Bazett-Jones et al., 2002). Presumably, if positive supercoils were induced or trapped in chromosomal DNA, this could result in compensatory negative writhe that would reduce chromosome length. However, it is not at all clear that the supercoiling effects of condensin/topoisomerase, acting on small circular DNA molecules, bear any resemblance to the effects that this complex has, in the presence of numerous other factors, on giant linear chromosomes that are already substantially condensed and coiled around histones. If the activity of topoisomerase II is necessary for some aspect of condensin II function, then feeding etoposide to flies could have interesting consequences. It is easy to imagine how this could have impacts of chromatid dispersal. If axial compaction was also affected, then this would suggest that the mechanism of condensin II dependent chromatin condensation involved the generation of DNA knots or some other process that involved creating

catenations within a chromosome. If etoposide had differential affects on dispersal and condensation, then it is clear that multiple mechanisms are in play. If condensin II is acting through multiple mechanisms, then there is also a chance that these mechanisms are genetically separable. Perhaps additional condensin II subunit alleles with be isolated in the future.

Another interesting question that emerges from this study is related to the role that condensin II plays in the formation of chromosome territories. These subnuclear compartments have roused considerable interest in recent years but there is little evidence as to how they are formed or why they exist. If a chromosome is substantially compacted along it major axis while the chromosome ends remain associated with the nuclear envelope, it is fairly obvious that this will induce all loci on the chromosome to move closer to each other and occupy a fraction of the nuclear volume (see Figure 11 and Figure 1D). By definition, this is a chromosome territory. To our knowledge, no other gene has been previously been implicated in the regulation of chromosome territory formation. This may simply be an issue of semantics, however. In a way, all mitoses and meioses involve the formation of discrete chromosome territories. The presence of these organizational patterns in interphase nuclei may just represent a holdover from the previous mitosis. Alternatively, chromosome territories may be regulated and facilitate important nuclear processes or other levels of organization that occur within the nucleus. Regardless, we now have a system with which was can test these models. It will be quite interesting to compare cells which express condensin II, and have territories, to those that do not and see what differences we are able to detect in processes such as transcription,

DNA replication, DNA repair, chromosome dynamics, or nuclear envelope structure. Of course, it will be difficult to control for local effects that condensin II may exert at specific genomic regions. Along those lines, it will also be interesting to see if other models systems, such as human cell lines which display very prominent chromosome territories, require condensin II to establish or main them.

References

Aagaard L, Laible G, Selenko P, Schmid M, Dorn R, Schotta G, Kuhfittig S, Wolf A, Lebersorger A, Singh PB, Reuter G, Jenuwein T. (1999) Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J. 18(7):1923-38.

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson J D. Molecular Biology of the Cell. Fourth Edition. New York, Garland Science. 2002

Allen MJ, Dong XF, O'Neill TE, Yau P, Kowalczykowski SC, Gatewood J, Balhorn R, Bradbury EM. (1993) Atomic force microscope measurements of nucleosome cores assembled along defined DNA sequences. Biochemistry. 32(33):8390-6.

Banfalvi G. (2008) Chromatin fiber structure and plectonemic model of chromosome condensation in Drosophila cells. DNA Cell Biol. 27(2):65-70

Bazett-Jones DP, Kimura K, Hirano T. (2002) Efficient supercoiling of DNA by a single condensin complex as revealed by electron spectroscopic imaging. Mol Cell. 9(6):1183-90

Belmont AS, Bruce K. (1994) Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. J Cell Biol. 127(2):287-302.

Belmont AS, Sedat JW, Agard DA. (1987) A three-dimensional approach to mitotic chromosome structure: evidence for a complex hierarchical organization. J Cell Biol. 105(1):77-92.

Belmont AS. (2006) Mitotic chromosome structure and condensation. Curr Opin Cell Biol. 18(6):632-8

Branco MR, Pombo A. (2006) Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. 4(5):e138. Epub 2006 Apr 25.

Bucceri A, Kapitza K, Thoma F. (2006) Rapid accessibility of nucleosomal DNA in yeast on a second time scale. EMBO J. 25(13):3123-32

Burgess SM. (2004) Homolog pairing in S. pombe: the ends are the means. Mol Cell.13(6):766-8

Cao K, Capell BC, Erdos MR, Djabali K, Collins FS. (2007) A lamin A protein isoform overexpressed in Hutchinson-Gilford progeria syndrome interferes with mitosis in progeria and normal cells. Proc Natl Acad Sci U S A. 104(12):4949-54

Case RB, Chang YP, Smith SB, Gore J, Cozzarelli NR, Bustamante C. (2004) The bacterial condensin MukBEF compacts DNA into a repetitive, stable structure. Science. 305(5681):222-7

Chasman DI, Flaherty KM, Sharp PA, Kornberg RD. (1993) Crystal structure of yeast TATA-binding protein and model for interaction with DNA. Proc Natl Acad Sci U S A. 90(17):8174-8

Chintapalli VR, Wang J, Dow JA. (2007) Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat Genet. 39(6):715-20.

Christensen MO, Larsen MK, Barthelmes HU, Hock R, Andersen CL, Kjeldsen E, Knudsen BR, Westergaard O, Boege F, Mielke C. (2002) Dynamics of human DNA topoisomerases IIalpha and IIbeta in living cells. J Cell Biol. 157(1):31-44

Chubb JR, Boyle S, Perry P, Bickmore WA. (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol. 12(6):439-45

Clements L, Manilal S, Love DR, Morris GE. (2000) Direct interaction between emerin and lamin A. Biochem Biophys Res Commun. 267(3):709-14

Cremer T, Cremer C, Baumann H, Luedtke EK, Sperling K, Teuber V, Zorn C. (1982) Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. Hum Genet. 60(1):46-56

Dej KJ, Spradling AC. (1999) The endocycle controls nurse cell polytene chromosome structure during Drosophila oogenesis. Development. 126(2):293-303.

Danilowicz C, Lee CH, Kim K, Hatch K, Coljee VW, Kleckner N, Prentiss M. (2009) Single molecule detection of direct, homologous, DNA/DNA pairing. Proc Natl Acad Sci U S A. 106(47):19824-9

Dernburg AF, McDonald K, Moulder G, Barstead R, Dresser M, Villeneuve AM. (1998) Meiotic recombination in C. elegans initiates by a conserved mechanism

and is dispensable for homologous chromosome synapsis. Cell. 94(3):387-98

Ding DQ, Chikashige Y, Haraguchi T, Hiraoka Y. (1998) Oscillatory nuclear movement in fission yeast meiotic prophase is driven by astral microtubules, as revealed by continuous observation of chromosomes and microtubules in living cells. J Cell Sci.111:701-12

Dorn R, Krauss V, Reuter G, Saumweber H. (1993) The enhancer of position-effect variegation of Drosophila, E(var)3-93D, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. Proc Natl Acad Sci U S A. 90(23):11376-80.

Dundr M, Misteli T. (2002) Nucleolomics: an inventory of the nucleolus. Mol Cell. 9(1):5-7.

Ekwall K, Ruusala T. (1994) Mutations in rik1, clr2, clr3 and clr4 genes asymmetrically derepress the silent mating-type loci in fission yeast. Genetics. 36(1):53-64.

Fanti L, Berloco M, Piacentini L, Pimpinelli S. (2003) Chromosomal distribution of heterochromatin protein 1 (HP1) in Drosophila: a cytological map of euchromatic HP1 binding sites. Genetica. 117(2-3):135-47

Foster HA, Bridger JM. (2005) The genome and the nucleus: a marriage made by evolution. Genome organisation and nuclear architecture. Chromosoma. 114(4):212-29

Fritzsche W, Schaper A, Jovin TM. (1994) Probing chromatin with the scanning force microscope. Chromosoma. 103(4):231-6.

Fung JC, Marshall WF, Dernburg A, Agard DA, Sedat JW. (1998) Homologous chromosome pairing in Drosophila melanogaster proceeds through multiple independent initiations. J Cell Biol. 1998 141(1):5-20

Haeusler RA, Engelke DR. (2006) Spatial organization of transcription by RNA polymerase III. Nucleic Acids Res. 34(17):4826-36

Hagstrom KA, Meyer BJ. (2003) Condensin and cohesin: more than chromosome compactor and glue. Nat Rev Genet. 4(7):520-34

Haithcock E, Dayani Y, Neufeld E, Zahand AJ, Feinstein N, Mattout A, Gruenbaum Y, Liu J. (2005) Age-related changes of nuclear architecture in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 102(46):16690-5

Heun P, Laroche T, Shimada K, Furrer P, Gasser SM. (2001) Chromosome dynamics in the yeast interphase nucleus. Science. 294(5549):2181-6

Hirano T, Kobayashi R, Hirano M. (1997) Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a Xenopus homolog of the Drosophila Barren protein. Cell. 89(4):511-21

Hirano T. (2006) At the heart of the chromosome: SMC proteins in action. Nat Rev Mol Cell Biol. 7(5):311-22.

Hiraoka Y. Meiotic telomeres: a matchmaker for homologous chromosomes. Genes Cells. 3(7):405-13

Hirota T, Gerlich D, Koch B, Ellenberg J, Peters JM. (2004) Distinct functions of condensin I and II in mitotic chromosome assembly. J Cell Sci. 117(Pt 26):6435-45

Hochstrasser M, Sedat JW. (1987) Three-dimensional organization of Drosophila melanogaster interphase nuclei. II. Chromosome spatial organization and gene regulation. J Cell Biol. 104(6):1471-83

Houben A, Demidov D, Caperta AD, Karimi R, Agueci F, Vlasenko L. (2007) Phosphorylation of histone H3 in plants--a dynamic affair. Biochim Biophys Acta. 1769(5-6):308-15

Human Genome Project, Build 33 (2003) Chromosome FAQs: http://www.ornl.gov/sci/techresources/HumanGenome/posters/chromosome/faqs.shtml

Ito T, Bulger M, Pazin MJ, Kobayashi R, Kadonaga JT. (1997) ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell. 90(1):145-55

Ito T. (2007) Role of histone modification in chromatin dynamics. J Biochem.141(5):609-14

Jenuwein T, Allis CD. (2001)Translating the histone code. Science. 293(5532):1074-80

Kimura H, Cook PR. (2001) Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J Cell Biol. 153(7):1341-53

Kimura K, Hirano T. (1997) ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation. Cell. 90(4):625-34

Kimura K, Hirano T. (2000) Dual roles of the 11S regulatory subcomplex in condensin functions. Proc Natl Acad Sci U S A. 97(22):11972-7

Kouzarides T. (2007) Chromatin modifications and their function.Cell. 128(4):693-705.

Laemmli UK, Cheng SM, Adolph KW, Paulson JR, Brown JA, Baumbach WR. (1978) Metaphase chromosome structure: the role of nonhistone proteins. Cold Spring Harb Symp Quant Biol. 42 Pt 1:351-60

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. (2006) Interchromosomal interactions and olfactory receptor choice. Cell. 126(2):403-1

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 326(5950):289-93.

MacQueen AJ, Colaiácovo MP, McDonald K, Villeneuve AM. (2002) Synapsis-dependent and -independent mechanisms stabilize homolog pairing during meiotic prophase in C. elegans. Genes Dev. 16(18):2428-42

MacQueen AJ, Villeneuve AM. (2001) Nuclear reorganization and homologous chromosome pairing during meiotic prophase require C. elegans chk-2. Genes Dev. 15(13):1674-87

Mahadevaiah SK, Turner JM, Baudat F, Rogakou EP, de Boer P, Blanco-Rodríguez J, Jasin M, Keeney S, Bonner WM, Burgoyne PS. (2001) Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet. 27(3):271-6

Maresca TJ, Freedman BS, Heald R. (2005) Histone H1 is essential for mitotic chromosome architecture and segregation in Xenopus laevis egg extracts. J Cell Biol. 169(6):859-69

Marshall WF, Straight A, Marko JF, Swedlow J, Dernburg A, Belmont A, Murray AW, Agard DA, Sedat JW. (1997) Interphase chromosomes undergo constrained diffusional motion in living cells. Curr Biol. 7(12):930-9

Matsumoto K, Hirose S. (2004) Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of Drosophila. J Cell Sci. 117(Pt 17):3797-805

McKee BD. (1996) The license to pair: identification of meiotic pairing sites in Drosophila. Chromosoma. 105(3):135-41

McKim KS, Hayashi-Hagihara A. (1998) mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes Dev. 12(18):2932-4

Meselson M, Stahl FW (1958) Proc. Natl. Acad. Sci. USA 44:671–682.

Nowak SJ, Corces VG. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet. 20(4):214-20

Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), {date of download}. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/

Ono T, Losada A, Hirano M, Myers MP, Neuwald AF, Hirano T. (2003) Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell.115(1):109-21

Parvinen M, Söderström KO. (1976) Effects of ESH and testosterone on the RNA synthesis in different stages of rat spermatogenesis. J Steroid Biochem. 7(11-12):1021-3

Passarge E. (1979) Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. Am J Hum Genet. 31(2): 106–115.

Page SL, Hawley RS. (2003) Chromosome choreography: the meiotic ballet. Science. 301(5634):785-9

Parnell TJ, Huff JT, Cairns BR. (2008) RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. EMBO J. 27(1):100-10

Peoples TL, Dean E, Gonzalez O, Lambourne L, Burgess SM. (2002) Close, stable homolog juxtaposition during meiosis in budding yeast is dependent on

meiotic recombination, occurs independently of synapsis, and is distinct from DSB-independent pairing contacts. Genes Dev. 16(13):1682-9

Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S. (2003) Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin. J Cell Biol. 161(4):707-14

Pinsky BA, Biggins S. (2005) The spindle checkpoint: tension versus attachment. Trends Cell Biol. 15(9):486-93

Pawlowski WP, Golubovskaya IN, Timofejeva L, Meeley RB, Sheridan WF, Cande WZ. (2004) Coordination of meiotic recombination, pairing, and synapsis by PHS1. Science. 303(5654):49-50.

Poirier MG, Marko JF. Micromechanical studies of mitotic chromosomes. J Muscle Res Cell Motil. 23(5-6):409-31

Rajapakse I, Perlman MD, Scalzo D, Kooperberg C, Groudine M, Kosak ST. (2009) The emergence of lineage-specific chromosomal topologies from coordinate gene regulation. Proc Natl Acad Sci U S A. 106(16):6679-84

Ringrose L, Ehret H, Paro R. (2004) Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. Mol Cell.16(4):641-53.

Savvidou E, Cobbe N, Steffensen S, Cotterill S, Heck MM. (2005) Drosophila CAP-D2 is required for condensin complex stability and resolution of sister chromatids. J Cell Sci.118(Pt 11):2529-43

Scaffidi P, Misteli T. (2006) Good news in the nuclear envelope: loss of lamin A might be a gain. J Clin Invest. 116(3):632-4

Scherthan H. A bouquet makes ends meet. Nat Rev Mol Cell Biol. 2001 2(8):621-7

Schotta G, Ebert A, Dorn R, Reuter G. (2003) Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. Semin Cell Dev Biol. 14(1):67-75

Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, Reuter G. (2002) Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. 21(5):1121-31

Sekinger EA, Moqtaderi Z, Struhl K. (2005) Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol Cell. 18(6):735-48

Sheehan MJ, Pawlowski WP. (2009) Live imaging of rapid chromosome movements in meiotic prophase I in maize. Proc Natl Acad Sci U S A. [Epub ahead of print]

Shumaker DK, Dechat T, Kohlmaier A, Adam SA, Bozovsky MR, Erdos MR, Eriksson M, Goldman AE, Khuon S, Collins FS, Jenuwein T, Goldman RD. (2006) Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. Proc Natl Acad Sci U S A. 103(23):8703-8

Segal E, Widom J (2009) What controls nucleosome positions? Trends. Genet. 25(8):335-43

Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. (2005) Interchromosomal associations between alternatively expressed loci. Nature. 435(7042):637-45

Ulyanova NP, Schnitzler GR. (2005) Human SWI/SNF generates abundant, structurally altered dinucleosomes on polynucleosomal templates. Mol Cell Biol. 25(24):11156-70

van Holde K, Zlatanova J. (2007) Chromatin fiber structure: Where is the problem now? Semin Cell Dev Biol. 5:651-8

Vazquez J, Belmont AS, Sedat JW. (2001) Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. Curr Biol. 11(16):1227-39

Vazquez J, Belmont AS, Sedat JW. (2002) The dynamics of homologous chromosome pairing during male Drosophila meiosis. Curr Biol. 12(17):1473-8

Wang Y, Ng TY. (2006) Phosphatase 2A negatively regulates mitotic exit in Saccharomyces cerevisiae. Mol Biol Cell. 17(1):80-9

Watson JD, Crick FH (1953) A Structure for Deoxyribose Nucleic Acid Nature. 171:964–967

Weiler KS. (2007) E(var)3-9 of Drosophila melanogaster encodes a zinc finger protein. Genetics. 177(1):167-78.

Weiler KS, Wakimoto BT. (2002) Suppression of heterochromatic gene variegation can be used to distinguish and characterize E(var) genes potentially important for chromosome structure in Drosophila melanogaster. Mol Genet Genomics. 266(6):922-3

Yang G, Leuba SH, Bustamante C, Zlatanova J, van Holde K. Role of linker histones in extended chromatin fibre structure. Nat Struct Biol. (11):761-3.

Figure Legends

Figure 1 – Condensin II Globally Inhibits Chromosome Alignment in the Euchromatin

2D projections of stage 10 nurse cell nuclei stained with 3 FISH probes to the X chromosome 12.9Mb-Red, 18.5Mb-White, 20.5Mb-Green, DNA-blue. A) Cap-H2/Df , B) Cap-D3/Df , C) SMC4/+;Cap-H2/+, D) yw (wt) Scale bar is equal to 10 um for all nuclei. X chromosome is diagrammatically depicted below images. The telomere is at the left end and the circle represents the centromere.

Figure 2 – Chromatid Dispersal Varies Between Loci

Locus specific X Chromosome dispersal as measured by mean number of spots in nurse cell nuclei. wt is in black, SMC4/+;Cap-H2/+ is in green, Cap-D3/Df is in blue, and Cap-H2/Df is in red. Data from stage 6 (top left), stage 8 (top right), and stage 10 (bottom left) are shown. Cytogenetic location 24Mb corresponds to the pericentric heterochromatin. Dispersal of X chromosome telomere (cytogenetic location 0 Mb) was inferred by dividing the total number of telomeric spots by 5 (the mean number of spots seen in polytene salivary gland polytene chromosome squashes). Error bars correspond to standard errors.

Figure 3 – Dispersal of Heterochromatic Loci is Condensin II Independent

2D Projections of stage 10 nurse cell nuclei stained with 3 FISH probes to heterochromatic loci. X is in green, 3R is in red, and 2L is in white, DNA is in blue. A)

Cap-H2/Df, B) Cap-D3/Df, C) SMC4/+;Cap-H2/+, D) yw (wt) Scale is equal to 10 um for all nuclei.

Figure 4 – Dispersal of Heterochromatin On Each Chromosomes

Dispersal of chromosome specific heterochromatic loci and telomeric loci as measured by mean number of spots in stage 10 nurse cell nuclei. yw (wt) is in black, SMC4/+;Cap-H2/+ is in green, Cap-D3/Df is in blue, and Cap-H2/Df is in red. Data from stage 6 (top left), stage 8 (top right), and stage 10 (bottom left) are shown. Here the total number of telomeric spots (from all chromosomes) are shown. Salivary gland polytene chromosomes most commonly display 5 separate telomeric foci. Error bars correspond to standard errors.

Figure 5 – Chromatid Dispersal is Topologically Restricted

Plots of principal component ratios. For all graphs: The X-axis represents the ratio of the first principal component to the second principal component, the Y-axis represents the ratio of the second principal component to the third principal component. Large X values indicate a highly linear pattern of chromatid dispersal. Large Y values indicate a highly planar pattern of chromatid dispersal. The point (1,1) would represent a perfectly spherical pattern of chromatid dispersal. The 3-D Diagram on the bottom right displays three idealized dispersal patterns. The Z-axis is for 3-dimensional perspective only. The 2-D graphs on the top row and left side display actual dispersal data. Each point represents a set of spots from a single probe in a single nucleus. Nuclei with less than 5 spots were not considered in this analysis. Red points correspond to all euchromatic loci

on the X chromosome. Blue points correspond to the X chromosome heterochromatin. Data from wild type egg chambers at: stage 6 (top left), stage 8 (top right), and stage 10 (bottom left).

Figure 6 – Condensin II is Necessary for Chromosome Axial Compaction in Nurse Cells

The X-axis displays the cytogenetic distance between each pair of loci on the X chromosome in megabases. Y-axis displays the mean distances between each pairs of probes in the nucleus in microns. When multiple spots were detected for a probe, the coordinates were considered to be those of the fluorescence intensity weighted centroid. yw (wt) is in black, SMC4/+;Cap-H2/+ is in green, Cap-D3/Df is in blue, and Cap-H2/Df is in red. Error bars correspond to standard errors.

Figure 7 – Condensin Induces Axial Compaction and Dispersal of Salivary Gland Chromatids

The effects of induced Cap-H2 expression on salivary gland chromosome structure. A) DAPI stained unspread polytene chromosome squash from wild type larva at 25C. B) DAPI stained unspread polytene squash from hsp70>Gal4/UAS>Cap-H2 raised at 25C. C,D,E) Confocal slice from DAPI stained whole mount salivary gland nucleus from UAS>lacI-GFP/+; 50F-LacO/60F-LacO with 8 hr, 32C heat shock. F,G,H) Confocal slice from DAPI stained whole mount salivary gland nucleus from UAS>lacI-GFP/+; 50F-LacO/60F-LacO;hsp70>Gal4/UAS>Cap-H2 with 8 hr, 32C heat shock. I) Confocal slice from DAPI stained whole mount salivary gland nucleus UAS>lacI-GFP/+; 60F-lacO/+; hsp70>GAL4/+ with 24, 32C hour heat shock. J) Confocal slice from DAPI

stained whole mount salivary gland nucleus UAS>lacI-GFP/+; 60F-lacO/+; hsp70>GAL4/UAS-Cap-H2 with 8 hour, 32C heat shock. K) Confocal slice from DAPI stained whole mount salivary gland nucleus UAS>lacI-GFP/+; 60F-lacO/+; hsp70>GAL4/UAS-Cap-H2 with 24 hour, 32C heat shock.

Figure 8 – Condensin is Inhibits Chromocenter Formation

Distances measured between pairs of heterologous heterochromatic loci. Pairs of loci are listed on the X-axis. Y-axis corresponds to mean distances between probes within the nucleus. When multiple spots were detected for a probe, the coordinates were considered to be those of the fluorescence intensity weighted centroid. yw (wt) is in black, SMC4/+;Cap-H2/+ is in green, Cap-D3/Df is in blue, and Cap-H2/Df is in red. Data from stage 6 (top left), stage 8 (top right), and stage 10 (bottom left) are shown. Error bars correspond to standard errors.

Figure 9 – Telomeric and Centromeric Loci Tend to Localize Near the Nuclear Periphery

The X-axis corresponds to the cytogenetic location of the loci in megabases. The Y-axis corresponds to the mean radial distance of a given probe. Radial distances are calculated as the distance between a probe and the geometric center of DAPI fluorescence and normalized to $V^{(1/3)}$, where V=volume of DAPI fluorescence. yw (wt) is in black and Cap-H2/Df is shown in red. Telomere data is based on signal from all telomeres. Data are shown from stage 6 (top left), stage 8 (top right), and stage 10 (bottom left) nurse cell nuclei. Error bars correspond to standard errors.

Figure 10 – Polytene Chromosome Dynamics in Response to Condensin II

Interpretation of the transition from polytene chromosome structure to dispersed structure. This figure is meant for illustration of basic principles only. It does not capture any subtle aspects of chromosome structure, it is not to scale, and it does not necessarily depict chromosomes in a way that in true to reality. Top) The absence of condensin II activity results in polytene chromosome structure: chromatids are elongated and tightly associated with sister chromatids and homologous chromosomes, alignment is precise and banding is clearly visible. Chromosomes exhibit extensive curvature. Middle) In the presence of moderate amounts of condensin II activity (or initially after condensin II begins to act on chromosomes), chromosome length is reduced and associations between sister chromatids and homologous chromosomes are at least partially resolved. Euchromatic regions exhibit more dispersal than telomeric or centromeric regions. Banding patterns are difficult to resolve though chromatids remain roughly aligned. Chromosomes appear to stiffen and straighten out. Bottom) With high levels of condensin activity and sufficient time, chromosomes are further reduced in length and chromatid dispersal is complete. Telomeric and centromeric still do not disperse completely. No banding is detectable.

Figure 11 – Condensin Regulated Chromosome Structure in the Nurse Cell Nucleus

Interpretation of the transiation from polytene to dispersed structure within the nucleus. This figure is meant for illustration of basic principles only. It does not capture any subtle aspects of chromosome structure, it is not to scale, and it does not necessarily

depict chromosomes in a way that in true to reality. Large, dark blue spheres represent nuclear volumes and small bright spheres represent hypothetical FISH probes: black=centromere, blue=telomere, red/white/green=euchromatic loci. Individual chromatids are depicted as gray lines. A) In Cap-H2/Df nurse cell nuclei or wild type salivary gland nuclei, chromosomes are completely polytene (FISH signals are almost always single spots) though centromeric regions often exhibit some dispersal. B) In Cap-D3/Df nurse cell nuclei or with minimal expression of Cap-H2 in salivary glands nuclei, chromosomes are shorter and exhibit some dispersal at euchromatic loci as well as near the centromere. C) In SMC4/+; Cap-H2/+ nurse cell nuclei or with moderate expression of Cap-H2 in salivary gland nuclei, chromosomes are much shorter than polytenes and dispersal is extensive throughout the euchromatin though allelic loci maintain some alignment. Centromeric regions are still only moderately dispersed and telomeres show little dispersal. D) In wild type nurse cell nuclei or with strong expression of Cap-H2 in salivary glands nuclei, chromosomes are shorter than the nuclear diameter. All euchromatic loci are completely dispersed though centromeres exhibit only moderate dispersal and telomeres are still loosely associated. Euchromatic loci are pulled toward the nuclear periphery, and spread out completely within the chromosome territory.

Figure 1

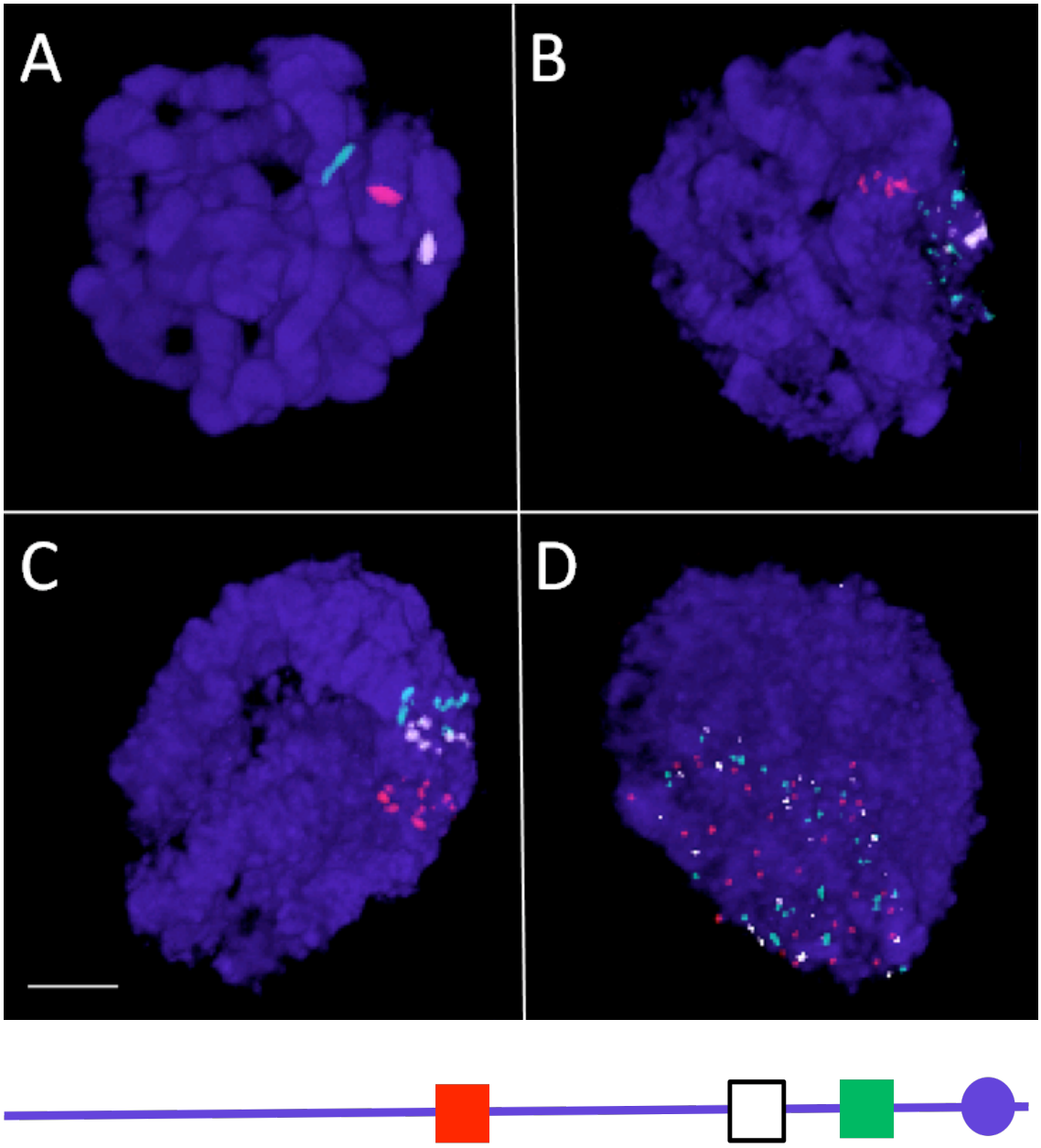


Figure 2

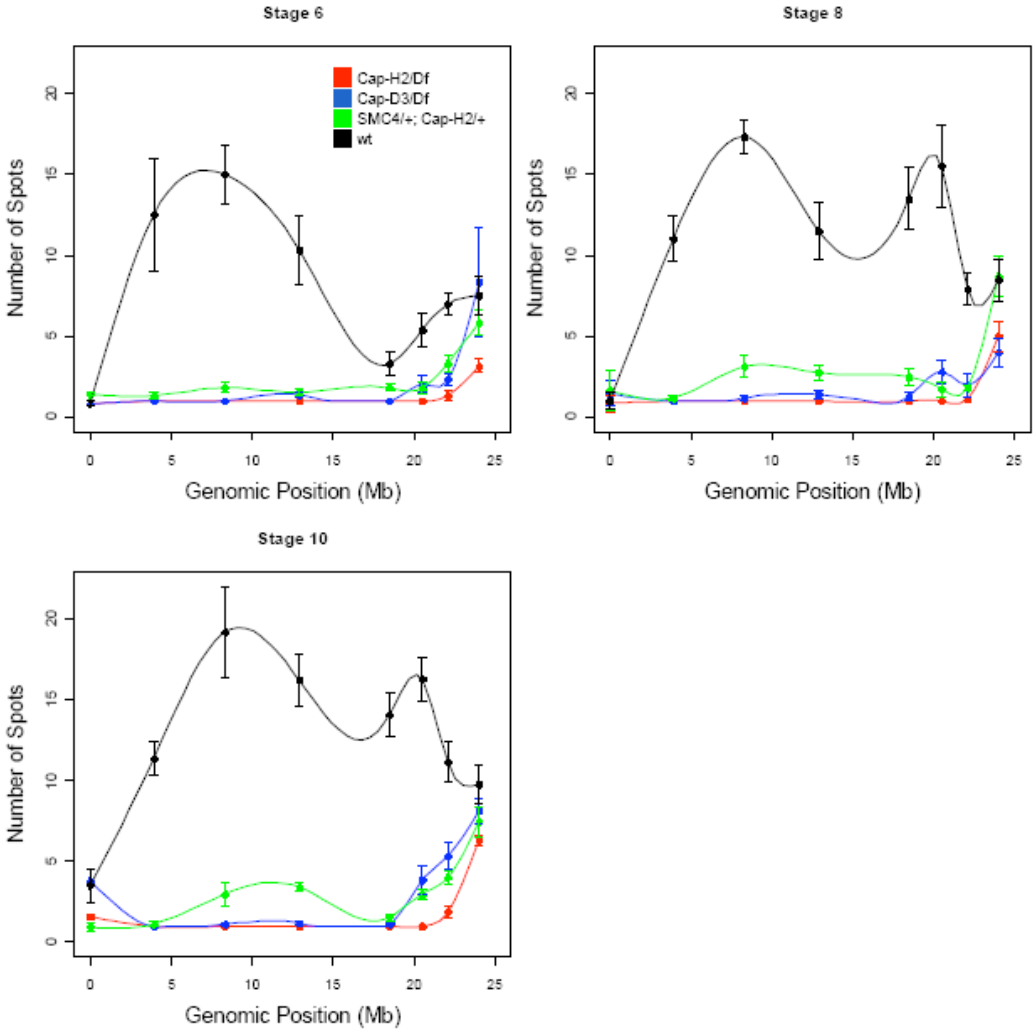

Stage 6
Cap-H2/Df
Cap-D3/Df
SMC4/+; Cap-H2/+
wt
Number of Spots
Genomic Position (Mb)
Stage 8
Number of Spots
Genomic Position (Mb)
Stage 10
Number of Spots
Genomic Position (Mb)

Figure 3

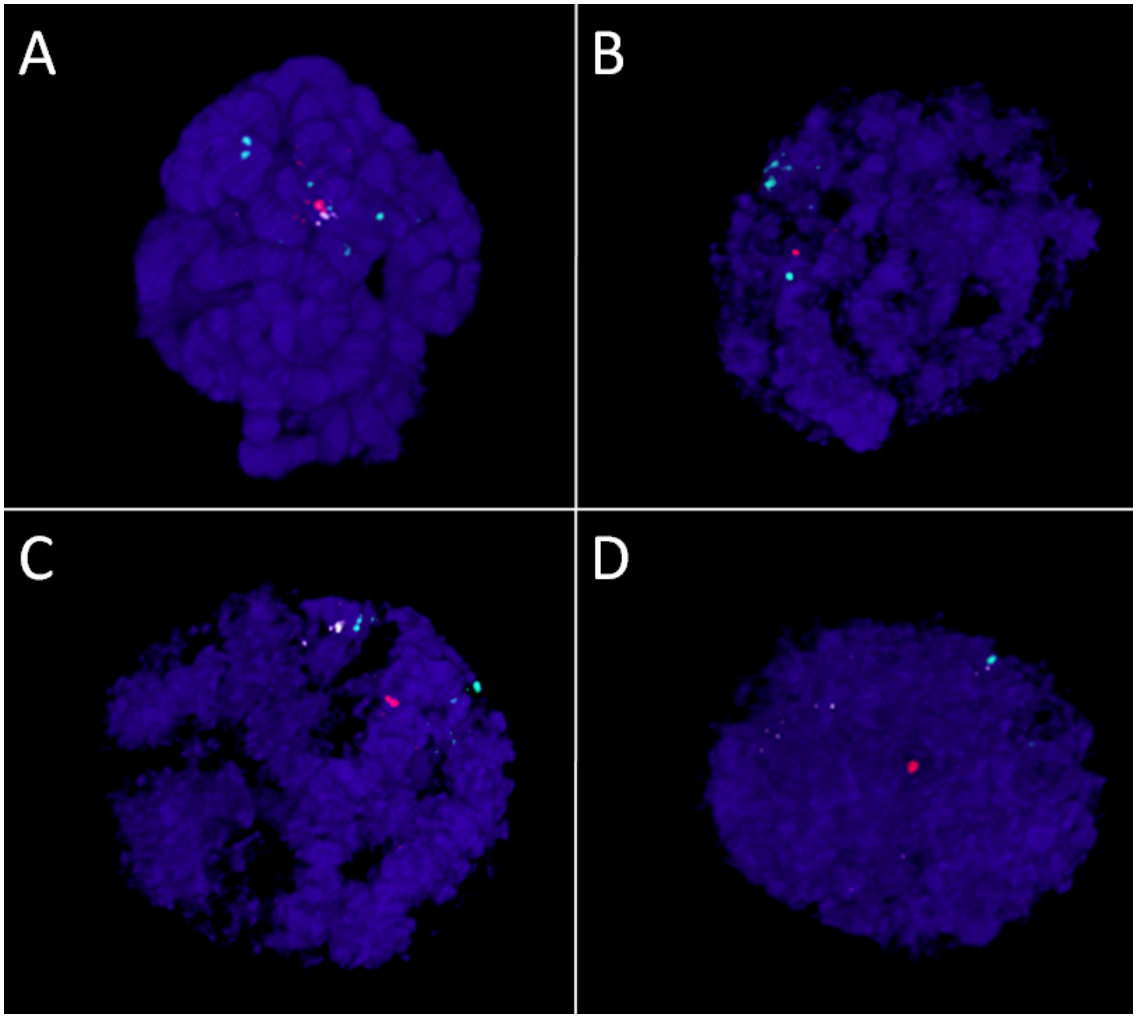


Figure 4

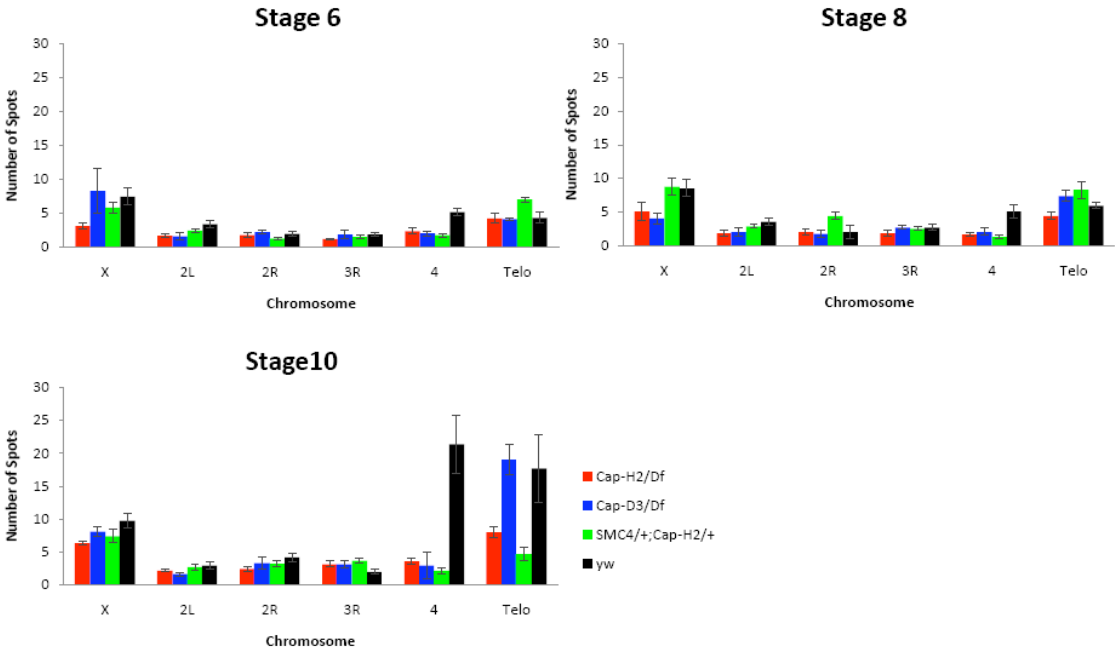

Stage 6
Stage 8
Stage10
Number of Spots
Chromosome
X
2L
2R
3R
4
Telo
0
5
10
15
20
25
30
Cap-H2/Df
Cap-D3/Df
SMC4/+;Cap-H2/+
yw

Figure 5

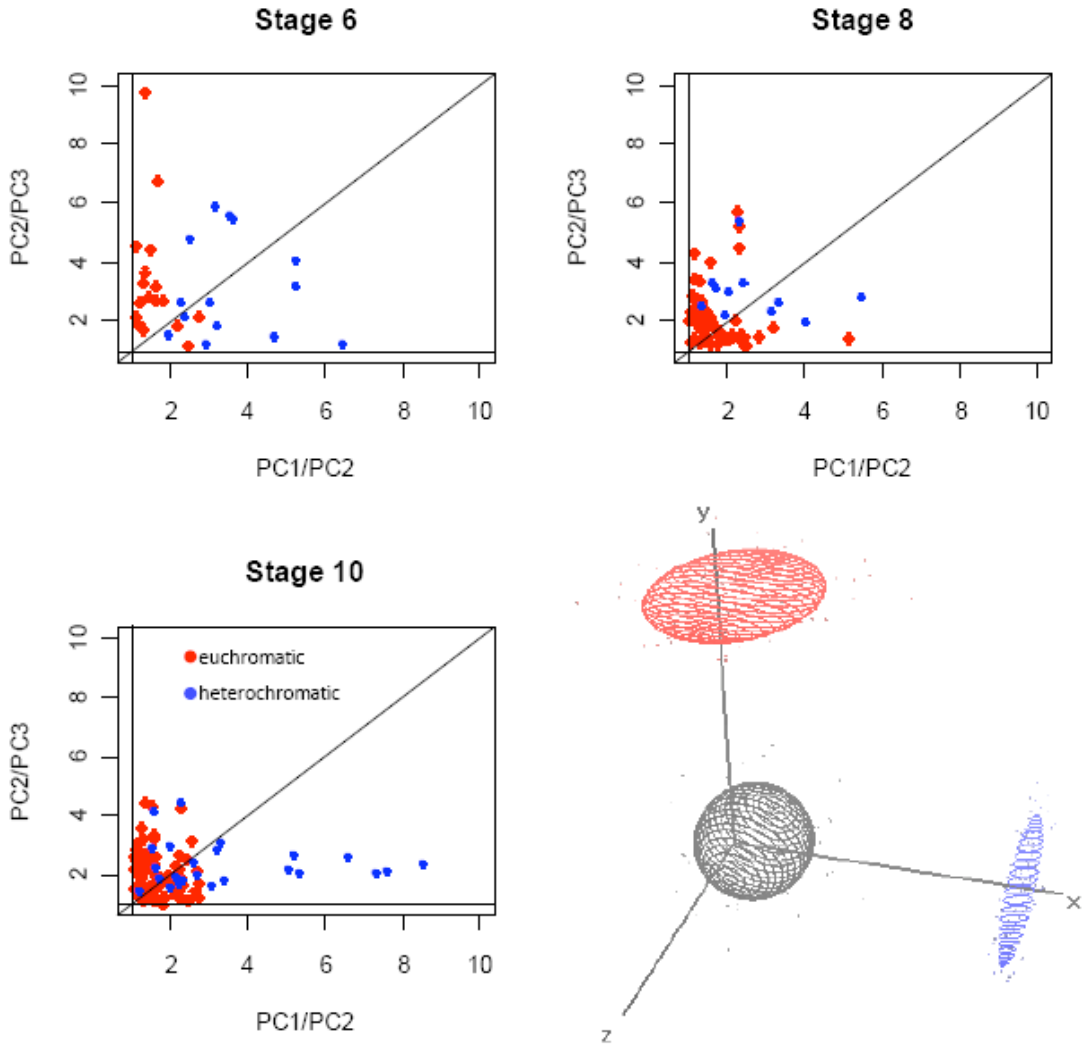

Stage 6
PC2/PC3
PC1/PC2
Stage 8
PC2/PC3
PC1/PC2
Stage 10
euchromatic
heterochromatic
PC2/PC3
PC1/PC2
y
x
z

Figure 6

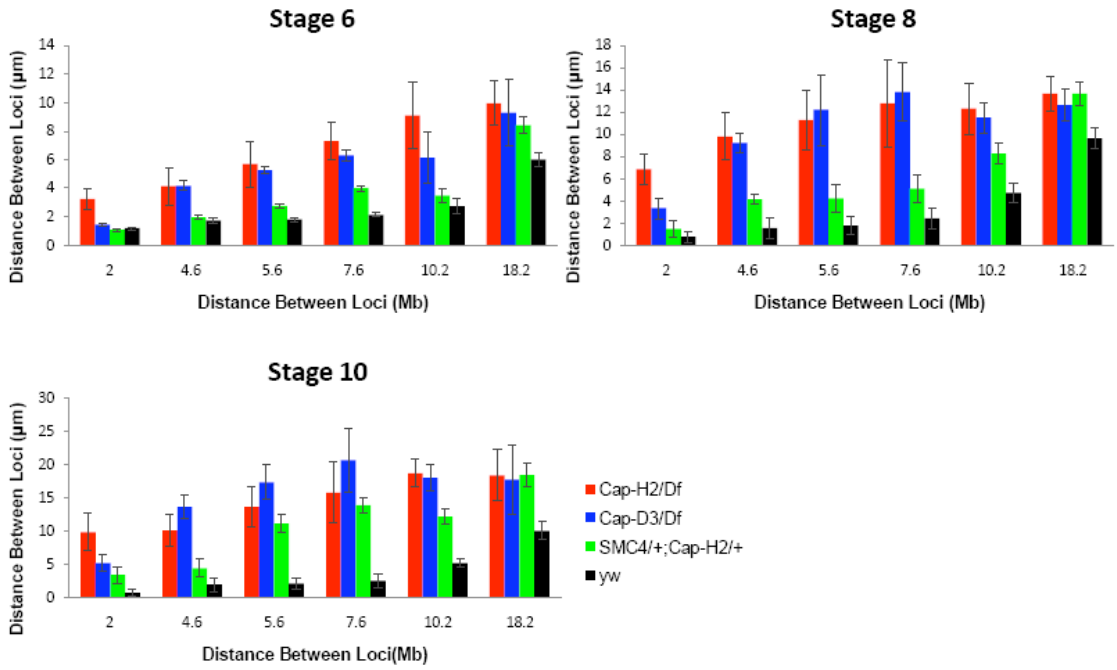

Stage 6
Distance Between Loci (µm)
Distance Between Loci (Mb)
Stage 8
Distance Between Loci (µm)
Distance Between Loci (Mb)
Stage 10
Distance Between Loci (µm)
Distance Between Loci(Mb)
2
4.6
5.6
7.6
10.2
18.2
Cap-H2/Df
Cap-D3/Df
SMC4/+;Cap-H2/+
yw

Figure 7

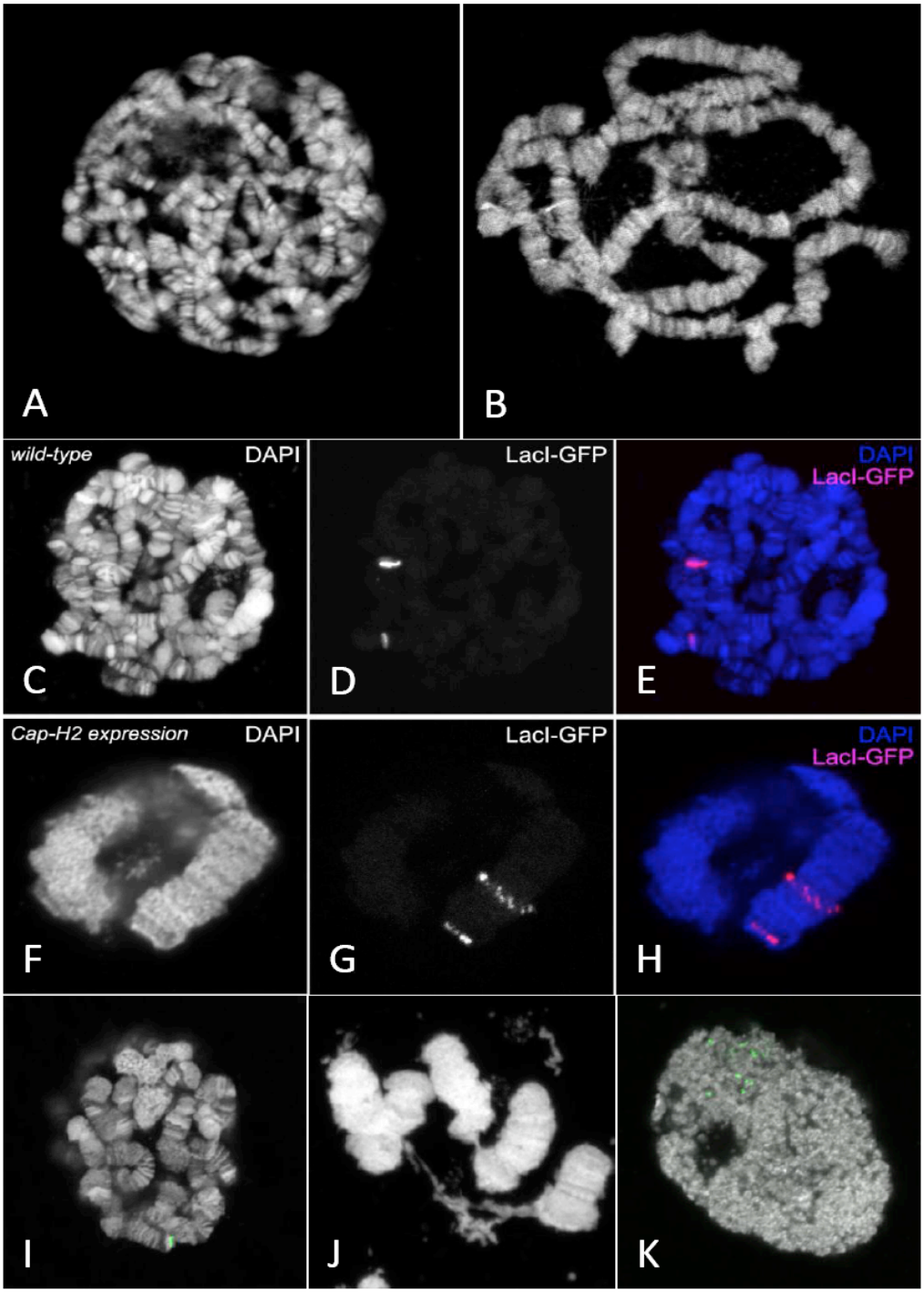

A
B
wild-type
DAPI
C
LacI-GFP
D
DAPI
LacI-GFP
E
Cap-H2 expression
DAPI
F
LacI-GFP
G
DAPI
LacI-GFP
H
I
J
K

Figure 8

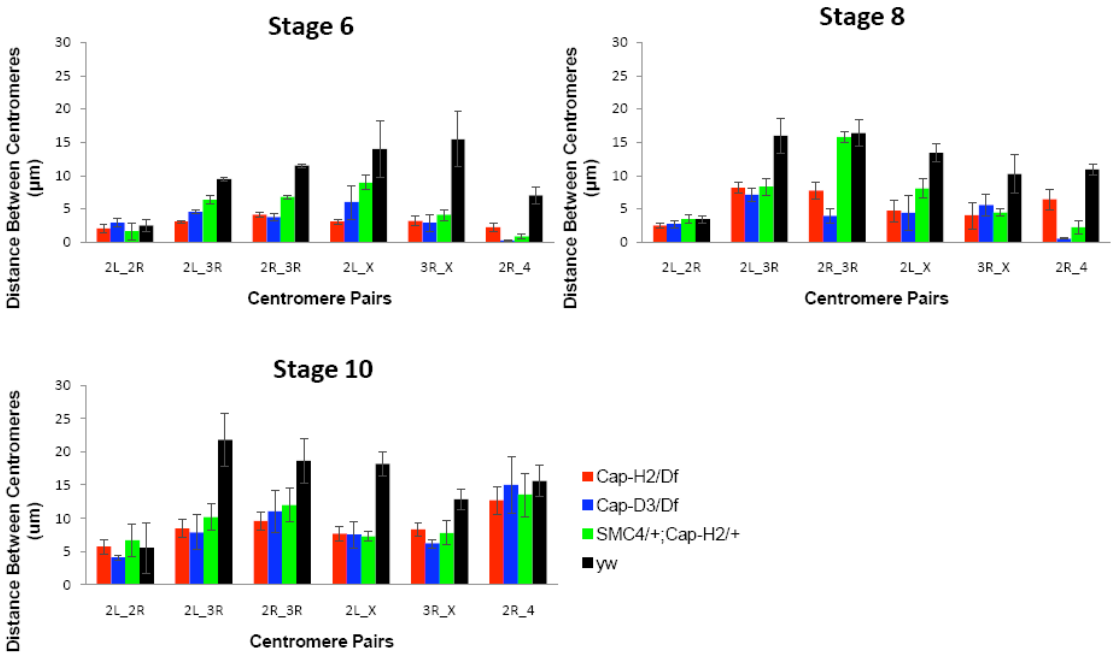

Stage 6
Distance Between Centromeres (µm)
30
25
20
15
10
5
0
2L_2R
2L_3R
2R_3R
2L_X
3R_X
2R_4
Centromere Pairs
Stage 8
Distance Between Centromeres (µm)
Centromere Pairs
Stage 10
Distance Between Centromeres (um)
Centromere Pairs
Cap-H2/Df
Cap-D3/Df
SMC4/+;Cap-H2/+
yw

Figure 9

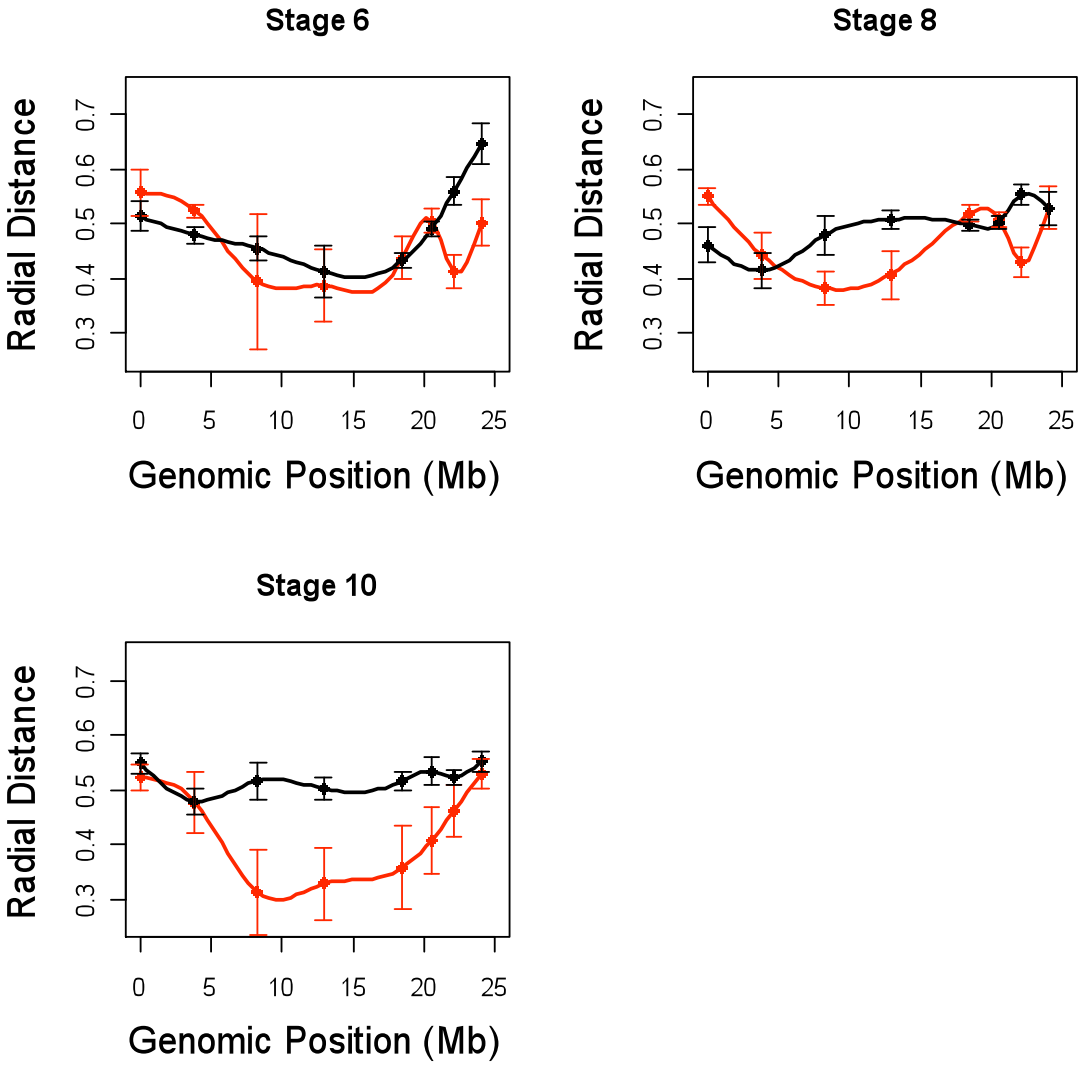

Stage 6
Radial Distance
0.3
0.4
0.5
0.6
0.7
0
5
10
15
20
25
Genomic Position (Mb)
Stage 8
Radial Distance
0.3
0.4
0.5
0.6
0.7
0
5
10
15
20
25
Genomic Position (Mb)
Stage 10
Radial Distance
0.3
0.4
0.5
0.6
0.7
0
5
10
15
20
25
Genomic Position (Mb)

Figure 10

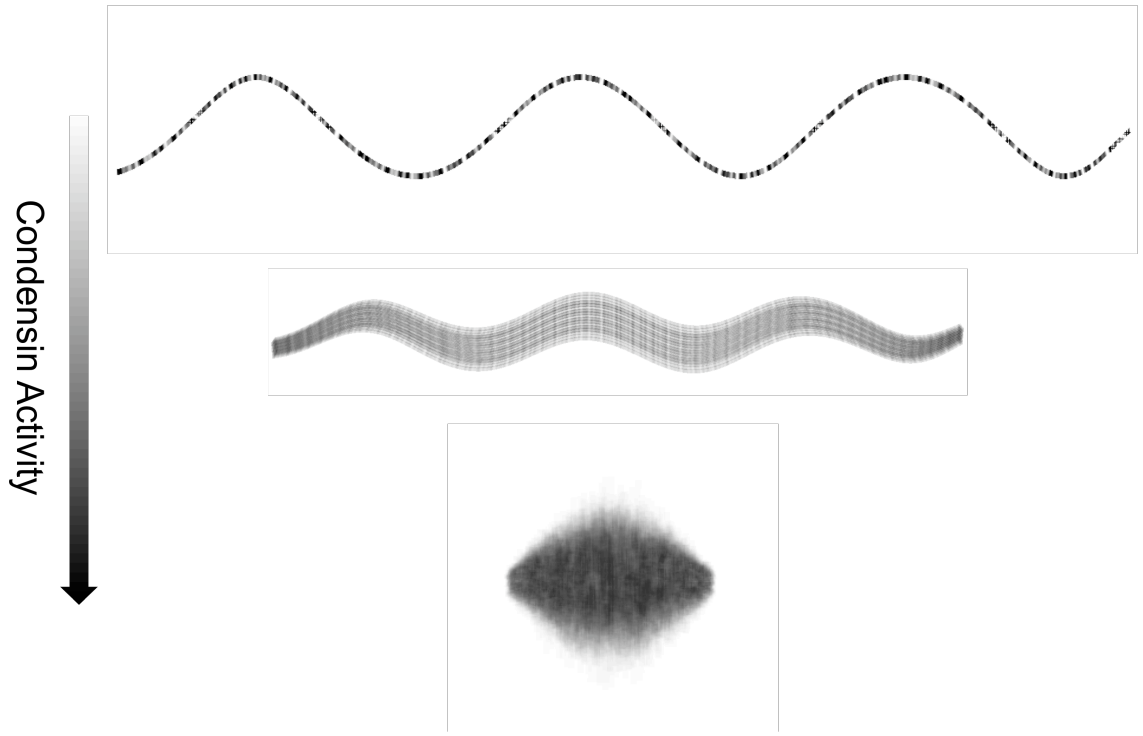
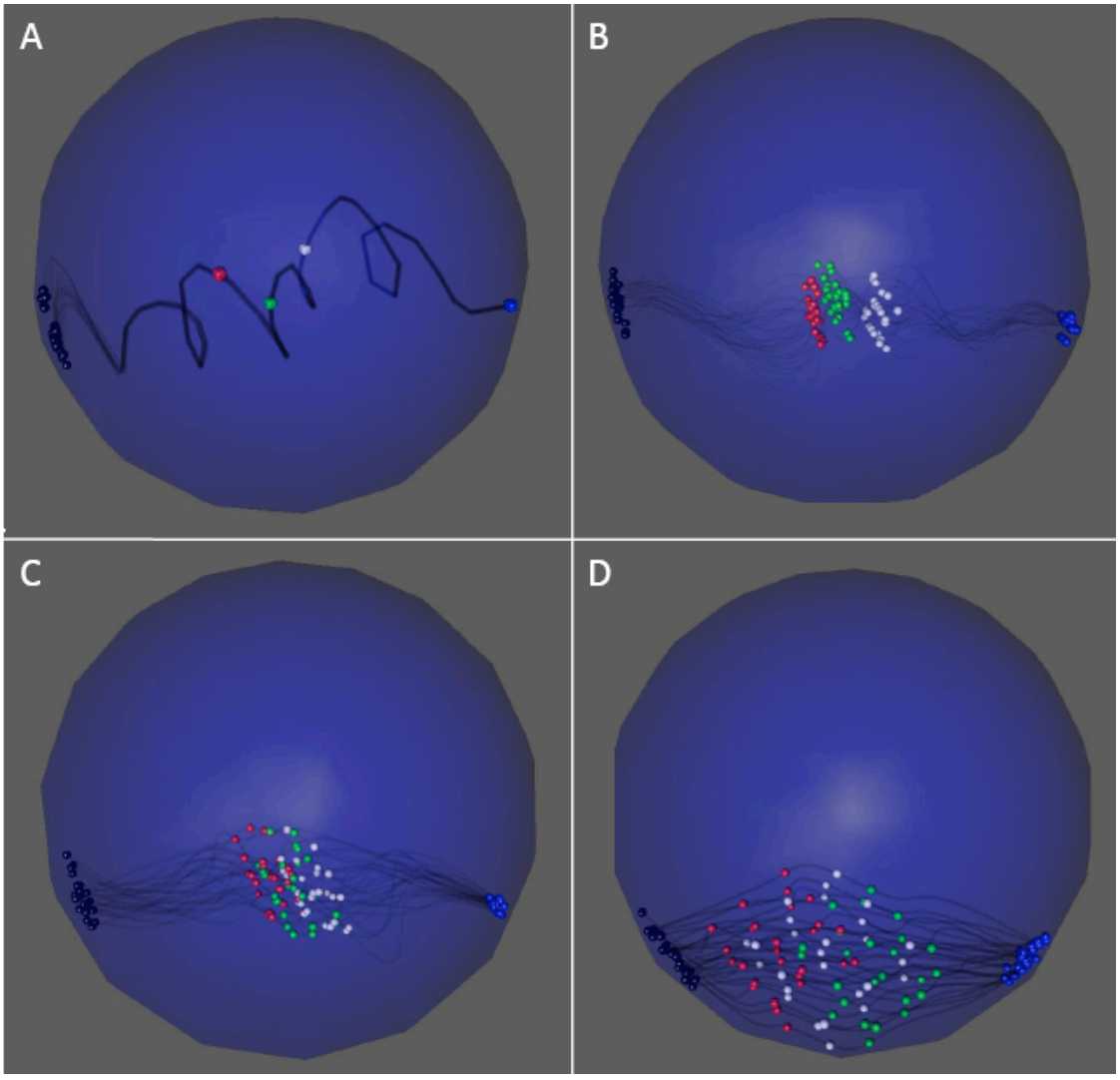


Figure 11



APPENDIX B

GENETIC AND SYSTEMS LEVEL ANALYSIS OF DROSOPHILA STICKY/CITRON KINASE AND DFMR1 MUTANTS REVEALS COMMON REGULATION OF GENETIC NETWORKS

BMC Systems Biology 2008, **2**:101doi:10.1186/1752-0509-2-101

Christopher R. Bauer[1], Andrew M. Epstein[1], Sarah J. Sweeney[1], Daniela C. Zarnescu[1] and Giovanni Bosco*[1]

[1] Department of Molecular and Cellular Biology, University of Arizona, Tucson, USA

*Corresponding author

Email addresses:

CRB: chrisb@email.arizona.edu

AME: epstein1@email.arizona.edu

SJS: ssweeney@email.arizona.edu

DCZ: zarnescu@email.arizona.edu

GB: gbosco@email.arizona.edu

Abstract

Background

In *Drosophila*, the genes *sticky* and *dFmr1* have both been shown to regulate cytoskeletal dynamics and chromatin structure. These genes also genetically interact with Argonaute family microRNA regulators. Furthermore, in mammalian systems, both genes have been implicated in neuronal development. Given these genetic and functional similarities, we tested *Drosophila sticky* and *dFmr1* for a genetic interaction and measured whole genome expression in both mutants to assess similarities in gene regulation.

Results

We found that *sticky* mutations can dominantly suppress a *dFmr1* gain-of-function phenotype in the developing eye, while phenotypes produced by RNAi knock-down of *sticky* were enhanced by *dFmr1* RNAi and a *dFmr1* loss-of-function mutation. We also identified a large number of transcripts that were misexpressed in both mutants suggesting that *sticky* and *dFmr1* gene products similarly regulate gene expression. By integrating gene expression data with a protein-protein interaction network, we found that mutations in *sticky* and *dFmr1* resulted in misexpression of common gene networks, and consequently predicted additional specific phenotypes previously not known to be associated with either gene. Further phenotypic analyses validated these predictions.

Conclusions

These findings establish a functional link between two previously unrelated genes. Microarray analysis indicates that *sticky* and *dFmr1* are both required for regulation of many developmental genes in a variety of cell types. The diversity of transcripts regulated by these two genes suggests a clear cause of the pleiotropy that *sticky* and *dFmr1* mutants display and provides many novel, testable hypotheses about the functions of these genes. As both of these genes are implicated in the development and function of the mammalian brain, these results have relevance to human health as well as to understanding more general biological processes.

Background

In multicellular organisms, developmental processes must coordinate cytoskeletal dynamics and morphogenesis with cell proliferation. For example, microtubule mediated trafficking and cell motility are inhibited during mitosis since centrosome and mitotic spindle mediated chromosome segregation requires dramatic reorganization of microtubules. In addition, actin/myosin mediated cell migration ceases during cell division when cortical actin must be reorganized in order to specify where the cleavage furrow is activated during cytokinesis. Consequently, gene expression patterns that control differentiation, morphogenesis and cell division must respond to a myriad of developmental cues and integrate them into a concerted cellular response in order to achieve proper tissue size and shape (for review see [1]).

It is also thought that changes in chromatin structure accompany developmental changes in order to establish and/or maintain tissue specific gene expression states [2, 3]. How changes in chromatin are coordinated with cell division and cell differentiation remains poorly understood. However, it is clear that these processes must be linked in order to ensure accurate propagation of epigenetic states and maintenance of cell fates [2-4]. As these are fundamental processes ubiquitous to all metazoans, it is of great interest to uncover the factors that link cell cycle progression, differentiation and morphogenesis to developmental changes in chromatin.

Drosophila development, and oogenesis in particular, has proven to be an excellent model for understanding how developmental cues coordinate differentiation with cell cycle progression [1, 5, 6]. The *Drosophila* model has recently been used to reveal novel epigenetic functions of *sticky* and the *Fragile-X mental retardation-1* (*dFmr1*) gene [7, 8]. Furthermore, both genes have been shown to regulate actin/myosin cytoskeletal organization and mutants in both genes exhibit pleiotropy. These findings suggest that *sticky* and *dFmr1* may possess similar biological functions and may represent regulatory hubs that coordinate diverse cellular and developmental processes. If this is true, it would be expected that both of these genes would have many interactors as it has been suggested that pleiotropy may result from disrupting hubs in protein-protein interaction networks as well as in gene regulatory networks [9, 10].

sticky is the *Drosophila* homologue of mammalian *citron kinase*. It is related to protein kinase B, protein kinase C, Rho-Kinase (ROK), myotonic dystrophy protein

kinase (DMPK) and myotonin-related Cdc42-binding kinase (MRCK) [11, 12] (for review see [13]). The only known substrate for citron kinase is myosin II, the primary motor protein responsible for cleavage furrow ingression during cytokinesis [14]). Although it is clear that *sticky/citron kinase* plays a critical function in cytokinesis, this kinase has other functions as it likely has more than one substrate. In addition, it may possess activities not dependent upon its kinase domain. *Citron kinase* deficient mice exhibit cytokinesis failure and increased apoptosis in the central nervous system resulting in severe ataxia and epilepsy. However, some non-neuronal cells develop normally [15-18]. Normal development of non-neuronal cells suggests that, in mice, *citron kinase* has a neuron specific function. In a Down syndrome mouse model *citron kinase* is responsible for inhibiting neurite extension [19]. Interestingly, this *citron kinase* mediated neurite inhibition is through a direct interaction with tetratricopeptide repeat protein TTC3, the ortholog of *Drosophila dTPR2*, which suppresses polyglutamine toxicity in a fly Huntington's disease model [19, 20]. Further support for a neuro-specific function comes from the observation that a splice variant of citron kinase protein, which lacks the kinase domain, directly interacts with post-synaptic density protein 95 (PSD95) and localizes to synapses [21-24].

The Fragile-X mental retardation-1 protein (FMRP) is a microRNA regulator, and like citron kinase protein, FMRP is thought to have critical functions in neurodevelopment [25]. Inactivation of the human *FMR1* ortholog leads to Fragile-X syndrome, which is the most common form of inherited mental retardation and is associated with severe neurodevelopmental and behavioral abnormalities (for review see

[26]). FMRP is an RNA binding protein, and its molecular function is thought to be in regulating mRNA translation and mRNA trafficking in neurons. Interestingly, in *Drosophila, dFmr1* has also been shown to regulate heterochromatin mediated gene silencing [7]. Therefore, it is clear that Fmrp is capable of affecting the expression of many genes, some of which produce mRNA molecules that are direct binding targets of the Fmrp protein while others may be targets of Fmrp chromatin mediated epigenetic gene regulation [7, 25]. Both *sticky* and *dFmr1* have been shown to genetically interact with the microRNA regulator *Ago1*, further supporting a connection between these genes in translational and epigenetic control of gene expression [8, 27].

Fmrp also regulates actin/myosin cytoskeletal dynamics through the Rac1 GTPase [7, 28-31]. Similarly, in both *Drosophila* and mouse, *sticky* (*citron kinase*) is regulated by Rho and Rac GTPase activities [12, 32]. This raises the interesting possibility that both citron kinase protein and Fmrp could regulate actin/myosin cytoskeletal dynamics *via* a common signaling pathway. Given the phenotypic and genetic similarities between the *Drosophila sticky* and *dFmr1* genes, we first investigated whether mutations in these two genes exhibited a genetic interaction. Secondly, we assessed whether both mutants had a common suite of genes that were misregulated that could account for the similarities between *sticky* and *dFmr1* mutant phenotypes.

In this report, we show that the *Drosophila citron kinase* ortholog, *sticky*, exhibits a genetic interaction with *dFmr1*, and we show that *sticky* and *dFmr1* regulate a common set of genes. These findings have important implications for the functions of both citron

kinase protein and FMRP in human neurodevelopment as well as human pathologies including Down syndrome and Fragile-X syndrome.

Results and Discussion

***sticky* mutants dominantly suppress *dFmr1* gain of function rough eye phenotype.**

We have previously shown that an *Argonaute1* (*Ago1*) gain of function phenotype can be dominantly suppressed by loss-of-function mutations in *sticky* [8]. This observation suggested that wild-type Ago1 protein function was dependent on proper *sticky* dosage (Figure 1F). Fmrp, like Ago1, is a regulator of microRNA mediated gene silencing [25]. In addition, *Drosophila dFmr1* mutants have been shown to have heterochromatin structure and gene silencing defects similar to those reported for *sticky* mutants [7, 8]. Therefore, we tested whether *sticky* genetically interacts with *dFmr1*.

Two different *sticky* alleles, sti^{3} and $sti^{Z3\text{-}5829}$, were crossed to a *sev:dFmr1* strain where *dFmr1* is being overexpressed specifically in the eye under the *sevenless* gene enhancer [33]. As previously reported the *sev:dFmr1* eyes were rough (Fig. 1A). At higher magnification the *sev:dFmr1* eye showed that 100% (n=10 eyes) had >50 ommatidia with crater-like necrosis (see arrows in Fig. 1D). Both sti^{3} and $sti^{Z3\text{-}5829}$ mutant alleles were able to dominantly suppress the *dFmr1* gain-of-function rough eye phenotypes yielding a less severe rough eye (Fig. 1B and 1C). Although the degree of suppression was variable, it occurred in 100% of progeny (n>100). Furthermore, the

crater-like necrosis of the individual ommatidia was almost completely suppressed as all (n=10) eyes we observed at high magnification had <5 of these defects per eye (see arrows, Fig. 1D and 1E). We conclude that Fmrp function is sensitive to *sticky* gene dosage. This also suggests that, in part, Fmrp effector functions are *sticky* dependent, and therefore we propose that *sticky* may function upstream of Fmrp (Fig. 1F). However, since the Rac-GTPase may be positively or negatively regulated by Fmrp [28, 29], it is also possible that Fmrp activity modulates *sticky* through Rac activation or Rac mRNA translational repression (Fig. 1F).

Loss of *dFmr1* function enhances *sticky* knockdown phenotypes in the eye.

The ability of *sticky* mutations to rescue *dFmr1* overexpression suggests positive regulation between these genes. In a model such as this (Fig. 1F), it would be predicted that a loss of function in both genes would result in phenotype enhancement. It has previously been shown that RNAi knockdown of *sticky*, under control of the eyeless promoter, results in a rough eye phenotype and a reduction in eye size [12]. We have used this system to demonstrate that loss of function in *dFmr1* can enhance the phenotypes caused by *sticky* RNAi knockdown in the fly eye (Fig. 2).

Flies containing an *eyeless*-*GAL4* driver were crossed to a line containing a UAS-*sticky*-RNAi construct. This resulted in rough and reduced eyes in many of the progeny. We also observed a strong effect on the bristles that border the eye as well as a variety of low frequency phenotypes including complete loss of the eye, reduced or missing maxillary palps, and necrotic ommatidia. We crossed the *ey*-*GAL4*; *UAS*-*sticky*-RNAi

lines to *UAS-dFmr1*-RNAi lines and to $dFmr1^{3}$ mutants. In both cases, reducing the function of *dFmr1* enhanced the phenotypes caused by *sticky* RNAi knockdown. In the cases of rough eyes, reduced eyes, vibrissae, and postorbital bristles, the phenotypes caused by sticky knockdown were significantly enhanced by both *dFmr1* knockdown and a $dFmr1^{3}$ loss of function allele ($p<.05$, Fisher's Exact test). When the *ey-GAL4* driver was used to express the *dFmr1* RNAi construct alone, no phenotypes were observed (n=62).

sticky and _dFmr1_ regulate a common set of genes.

sticky and *dFmr1* genetically interact and share a common set of genetic interactors such as *Ago1*, *Rac*-GTPase and possibly others (Fig. 1F). Since mutations in the *sticky* and *dFmr1* genes have been demonstrated to affect chromatin structure, we also hypothesized that regulation of gene expression by sticky kinase and/or Fmrp could occur at the level of transcription and therefore mRNA levels would be altered in each of these mutants. To assess which specific transcripts were being regulated by these genes, we used NimbleGen *Drosophila* whole genome arrays containing ~385,000 features (12 unique features per gene, synthesized in duplicate on each array) to measure the levels of 15,634 specific transcripts in Cy3 labeled cDNA made from whole female fly total RNA. Due to the number of measurements per gene, we were able to identify a large set of misexpressed genes with high significance.

Expression array data obtained from three biological replicates of each mutant were compared to a pooled set of two biological replicates for each of two independent

wild-type controls (Fig. 3A). One wild-type control was *Oregon-R*. The second wild-type control was from a stock made in a w^{1118} background, homozygous for $dFmr1^{3}$, and also carried a fully functional wild-type *dFmr1* transgene that rescued the $dFmr1^{3}$ mutant phenotypes. Therefore, this second control stock had a genetic background identical to that of the $dFmr1^{3}$ mutant. This served to reduce the chance of identifying genes which showed different transcript abundance due to genetic background differences. The use of two independent wild type controls also served to reduce the chance of detecting transcripts which may have large natural variances in abundance but do not dramatically impact phenotype. Adult female flies from control stocks were grown and aged under identical conditions as the mutants.

We observed that $sti^{3}/sti^{Z3\text{-}5829}$ females had 249 genes and 771 genes misexpressed with a false discovery rate of $q<0.001$ and $q<0.01$, respectively (Fig. 3 and Additional File 1). Similarly, $dFmr1^{3}/dFmr1^{3}$ females had 1,044 and 2,894 genes misexpressed with $q<0.001$ and $q<0.01$, respectively (Fig. 3 and Additional File 2). Strikingly, a comparison of these two sets revealed a highly significant overlap (Fig. 3). With the more stringent q-value threshold, 40 genes were found to be differentially expressed on both lists. This is significantly greater than the number one would expect to find by chance ($p=1.91\times10^{-7}$, see methods). Each of these 40 genes showed a change in the same direction in both mutants (Fig. 4). With the more relaxed threshold of $q<0.01$, 246 genes are common to both lists ($p=3.23\times10^{-20}$) with 227 of these genes (~92%) commonly up-regulated in the mutants (Additional File 3).

To assess the validity of these findings we used semi-quantitative RT-PCR to measure the abundance of 8 different transcripts that were enriched in both mutants. We confirmed that 8 of 8 genes are indeed being overexpressed in *sticky* mutants and 7 of 8 genes were upregulated in *dFmr1* mutants (Fig. 5). Though the *Cbl*-L transcript signal appeared to be higher in the *dFmr1* mutant than in the wild-type control, we could not conclude that this difference was statistically significant. This may be because *Cbl*-L is expressed at very low levels in all genotypes and our assay was simply not sensitive enough to detect such a small difference. Regardless, this RT-PCR validation provides general evidence that *sticky* and *dFmr1* both serve to normally repress a common set of genes, and these data suggest that a large number of similar processes may be affected by both *sticky* and *dFmr1*.

sticky kinase and Fmrp regulate a common set of biological processes.

Though the analyses described above demonstrate that many genes are commonly regulated by sticky kinase and Fmrp with a high degree of statistical significance, the biological significance of this regulation was not evident from the gene lists alone. To overcome this problem we associated our expression data with a network of *Drosophila* protein-protein interactions. Transcriptional regulation is only one of many levels of regulation found in biology. Though microarrays do not directly provide data on other types of regulation, the effects of transcriptional changes on other levels of biological regulation can be inferred based on additional information. For example, if a gene does not show a change in transcript abundance but nevertheless interacts with a number of

genes that show significant expression changes, it is likely that its functionality will be altered. Thus, we used protein-protein interaction data to expand the scope of our microarray analysis. After associating our microarray gene expression data set with a protein-protein interaction network (Fig. 6), we used jActiveModules to identify and score sub-networks of interacting genes that collectively showed significant changes in expression [34].

For each mutant expression data set the highest scoring sub-networks revealed several interesting features (Fig. 7 and 8). First, we observed that for both mutants, the *Mer*, *ci* and *aPKC* genes emerge as common hubs for the two sub-networks even though none of these genes are being misexpressed in either mutant (gray nodes in Fig. 7 and 8). As noted above, this is due to many of the same genes that are being misexpressed in both mutants. Second, we found that although both *sticky* and *dFmr1* mutants generally tend to overexpress genes (Fig. 4), the *dFmr1* sub-network shows many more underexpressed genes than the *sticky* sub-network (Fig. 7 and 8, green nodes). Third, although the *dFmr1* sub-network shows many more underexpressed genes than the *sticky* sub-network these are not necessarily genes that are misregulated only in the *dFmr1* mutant. For example, the DNA repair genes, *Irbp* and *Pms2*, are overexpressed in *sticky* mutants whereas both are underexpressed in *dFmr1* mutants (Fig. 7 and 8). Similarly, the Autophagy-specific gene 1 kinase, *Atg1,* gene is overexpressed in *sticky* mutants whereas it is underexpressed in *dFmr1* mutants (Fig. 7 and 8). Lastly, although DNA damage response and repair genes are common in both sub-networks, we noted that different subsets of these genes are misexpressed in each of the two mutants. For example, the

Lig4 and *mus309* genes are present in the *sticky* sub-network but not the *dFmr1* sub-network, whereas the *RecQ*, *mus209* and *mus210* genes emerge from the *dFmr1* sub-network and not observed in the *sticky* sub-network. This suggests that although both mutants may impinge on similar biological processes by altering expression of similar sets of genes, the mechanism by which *sticky* and *dFmr1* effect these same processes may be different. In the case of these DNA repair genes, it will be interesting to see whether these differences in the two sub-networks reflect physiologically different types of DNA damage and/or repair processes.

As a quantitative measurement, we used the genes within these sub-networks to identify significantly overrepresented gene ontology terms (GO-terms). Several common GO-terms were overrepresented in the highest scoring sub-networks for both *sticky* and *dFmr1* mutant gene expression patterns. These GO-terms included many related to embryonic and larval development, nervous system development, oogenesis, cytoskeletal organization, axis specification, cell cycle regulation, and DNA damage responses (Fig. 9 and 10). Many of these processes fit well with our current understanding of the *in vivo* functions of these two genes yet some suggest novel functions. For example, according to our analysis "cytoskeleton organization and biogenesis" is altered in both *sticky* and *dFmr1* mutants. This is consistent with previous reports in the literature. Intriguingly, our GO-term analysis predicts disruptions in oocyte and embryonic axis specification and DNA repair, two processes that have not been reported to be affected in either *sticky* or *dFmr1* mutants.

As a control for this analysis, we compared gene expression in our two different wild-type control strains. We observed 191 and 365 genes ($q<.001$ and $q<.01$, respectively) to be differentially expressed between these two strains. These expression changes were integrated with our protein-protein interaction network and used to identify significantly altered genetic modules. The highest scoring sub-network did not possess any informative GO-terms which were significantly over-represented (data not shown).

As a further test of the significance of this GO-term analysis, we randomized the microarray data sets by randomly reassigning gene names to the microarray gene expression p-values (see methods), and we asked whether these would yield any specifically enriched GO-terms that were also found in the non-randomized data set. We found that the *sticky* mutant randomized expression data set produced no significantly enriched GO-terms (data not shown). By contrast, the *dFmr1* mutant randomized data set did yield a small number of significantly enriched GO-terms, such as "RNA-pol II transcription" and "transcriptional regulation" (data not shown). This may represent bias that is inherent to our protein-protein interaction network.

sticky and _dFmr1_ regulate genes with similar tissue specificity.

The GO-terms found to be over-represented in the identified sub-networks (Fig. 9 and 10) suggested some misexpression may have been tissue specific. However, since our RNA was prepared from whole adult flies, it was not possible to directly determine which tissues were misexpressing which genes. In order to address this issue we compared our data to the FlyAtlas dataset [35]. This data set compares gene expression

profiles between individual *Drosophila* tissues. However, the criteria used to define tissue specific genes in Chintapalli et al. were as strict as possible (no detectable signal in more than one tissue) and resulted in small lists of tissue specific genes. We wished to determine larger sets of "tissue enriched" genes from FlyAtlas that could be more robustly compared to our microarray data. Therefore, we defined our own sets of "tissue enriched" genes (see methods) based on the ratio of gene expression in a given tissue to that of the whole fly, as reported in [35]. We were then able to calculate the probability of observing a given amount of overlap between sets.

By comparing our misexpression lists with tissue enrichment lists derived from FlyAtlas, we observed both significant over- and under-representation of tissue enriched genes. We found that genes enriched in the head, midgut, malphigian tubules, larval tubules, and larval fatbody were significantly over-represented in the lists of misexpressed genes for both *sticky* and *dFmr1* mutants ($q<.001$ and $q<.01$) (Fig. 11). Hindgut enriched genes were also highly represented on both lists though this was just shy of significance for *dFmr1* misexpressed genes. We also found that genes enriched in ovaries and testes were under-represented in both the *sticky* and *dFmr1* misexpression sets though the significance of this varied based on which q-value cutoff was used (Fig. 11). The under-representation of testes enriched genes was expected since we used female total RNA. However, under-representation of ovary specific genes was surprising since nearly half the RNA of mature adult females comes from ovaries, and this could have biased our analysis toward ovary enriched genes. Alternatively, *sticky* and *dFmr1*

could regulate genes that are not particularly abundant in ovaries but nevertheless have critical ovary specific functions, and therefore give rise to tissue specific phenotypes.

sticky and _dFmr1_ mutants exhibit predicted oocyte polarity defect.

In order to validate some GO-term analysis predictions we turned again to the *sticky* and *dFmr1* mutants. GO-term analysis predicted that the *sticky* and *dFmr1* mutants may exhibit an oocyte axis polarity defect (Fig. 7 and 8). This hypothesis was further supported by the common misexpression of the *Cbl* gene (Table 1 and Figure 5), which encodes for a receptor tyrosine kinase-associated, E3 ubiquitin ligase [36]. The Cbl protein attenuates EGFR signaling, which is a major receptor tyrosine kinase signaling pathway in the developing ovary [36, 37]. Furthermore, Fmrp has been shown to directly bind the *Cbl* mRNA, possibly regulating its stability and translation, thus adding an additional level of regulation to Cbl gene expression [33]. Disruption of EGFR signaling during oogenesis can lead to oocyte polarity defects [38], and thus overexpression of *Cbl* could potentially lead to polarity defects by inappropriately dampening EGFR signaling [36]. Alternatively, *sticky* is known to phosphorylate the *Drosophila* myosin II light chain, *spaghetti squash* (*sqh*), in order to promote cytokinesis [39]. Although genetic mosaic studies have demonstrated that loss of *sqh* function leads to oocyte dorsal/ventral axis polarity defects [40, 41], it is not known whether sticky kinase is the effector kinase that regulates myosin II in a non-cytokinesis context such as oocyte axis determination.

We examined stage 14 oocytes in $sti^{3}/sti^{Z3\text{-}5829}$ females and observed that 35% of oocytes were indeed ventralized, whereas the $sti^{3}/+$ sibling heterozygous controls had

normal oocytes (Fig. 12 A,C,E). The variability in severity of oocyte ventralization is likely due to the fact that the *sti*$^{Z3\text{-}5829}$ mutation is not a complete loss-of-function [8]. Interestingly, the severity of this phenotype was temperature sensitive (data not shown), also consistent with our previous report demonstrating that the temperature sensitivity of the *sti*$^{Z3\text{-}5829}$ mutant allele [8]. Similarly, stage 14 oocytes from *dFmr1*^{3}/*dFmr1*^{3} mutant females were also severely ventralized, whereas a sibling heterozygous *dFmr1*^{3}/+ had normal oocytes (Fig. 12 B,D,F). This oocyte ventralization defect suggests that localization of axis determining factors requires sticky kinase and Fmrp function and/or that *Cbl* overexpression in the ovary can lead to polarity defects. In either case, we conclude that the polarity defects predicted by the GO-term analysis are valid, and this points to clear and testable mechanistic models of *sticky* and *dFmr1* function in *Drosophila* oocyte axis formation.

Conclusions

Mutations in *Drosophila sticky* and *dFmr1* both result in a wide variety of phenotypes associated with many cell types and cellular processes. We believe many of these diverse phenotypes are caused by effects on the populations of a wide range of transcripts, either through transcriptional regulation or mRNA stability. Genes that function to regulate such a diverse array of processes can be resistant to characterization by traditional genetic methods. This is because pleiotropic effects make it difficult to infer function from phenotypes. Thus, we have used a combination of genomic and

protein-protein interaction network analyses in order to compare the effects of mutations in *sticky* and *dFmr1*. We have found that these two genes function to negatively regulate a large number of common transcripts and that their targets are involved in a number of similar biological processes. It is noteworthy that this interpretation is also supported by previous reports that both sticky kinase and Fmrp are regulators of chromatin mediated epigenetic gene silencing. It has been shown that Fmrp is required for centric heterochromatin assembly during embryogenesis [7]. It has also been shown that *sticky* wild-type gene function is required for proper silencing of genes by heterochromatin [8]. Therefore, the finding that both mutants display overexpression of many genes is consistent with genetic evidence demonstrating their general gene silencing functions.

Based on our analyses, we predict that *sticky* and *dFmr1* both function to control many aspects of development in imaginal disc derived appendages, oocytes, neurons, and possibly other tissues. Surprisingly, other biological processes that have not previously been associated with either gene are also predicted to be perturbed in *sticky* and *dFmr1* mutants. For example, a number of genes functioning in DNA damage repair such as *DNA pol delta, Irbp, Ligase 4, mei-41, mre11, Mlh1, mus209, mus210, mus309, Pms2* and *tefu,* among others, are revealed in the high scoring sub-networks of either one or both mutants (Fig. 9 and 10, Additional Files 1 and 2). We have validated at least one GO-analysis prediction by showing that *sticky* and *dFmr1* mutants indeed exhibit an oocyte axis polarity defect (Fig. 12). Interestingly, *Drosophila* mutants in DNA repair functions, including *mei-41*, disrupt receptor tyrosine kinase (EGFR) signaling and results in oocyte polarity defects [42, 43]. Therefore, it is possible that the *sticky* and

dFmr1 mutant oocyte polarity defects could be due to failure to repair meiotic or germline chromosomal breaks.

Lastly, it is also noteworthy that the *Drosophila Merlin* (*Mer*) gene, the ortholog of the human tumor suppressor neurofibromatosis-2, emerged as a major hub for both *sticky* and *dFmr1* active sub-networks (Fig. 7 and 8). This is important because the Mer protein serves to anchor actin to transmembrane proteins and is critical for establishing oocyte axis polarity [44]. It remains to be tested whether DNA damage repair, Cbl mediated attenuation of EGFR signaling, or actin/myosin defects is the primary cause of this axis polarity mutant phenotype. In addition, it remains to be tested whether other predictions from these analyses are correct.

Our claims of overlapping functions for the *sticky* and *dFmr1* genes are supported by the demonstration of a genetic interaction and phenotypic similarities. The phenotypes that we observe in these mutants also serve to validate our computational predictions about the functions of these genes. Genes that regulate many biological processes also tend to exhibit pleiotropic effects, leading to multiple and seemingly unrelated phenotypes. As is the case with many other proteins that function as hubs within biological networks, understanding the molecular mechanisms through which sticky kinase and Fmrp affect any cellular process is complicated by the very fact that they impinge on many other regulatory molecules. The idea that mutations with pleiotropic effects are common to genes whose products have many protein-protein interactions is supported by a recent study of yeast pleiotropic genes [10]. Though few

direct targets of Fmrp are known, and sticky kinase has only one known phosphorylation target, our expression data suggest these genes may nevertheless be hubs in the gene expression network (for reviews see, [13, 45]). In addition, the great number of genes that are transcriptionally misregulated in both *dFmr1* and *sticky* loss-of-function mutations could also explain pleiotropic effects [9]. Our systems level analysis of gene expression patterns in *sticky* and *dFmr1* mutants has revealed gene networks, and potential direct downstream targets, that are regulated by both sticky kinase and Fmrp. It is our hope the future studies that seek to identify the molecular mechanisms causing the myriad phenotypes presented in these mutants may be facilitated by the gene networks reported in this study and the multiple testable hypotheses that arise from them, for example the predicted function of sticky kinase and Fmrp in cell polarity and DNA damage repair.

Methods

Drosophila strains

The *P* w^{+} *[sev:dFmr1]*, CyO and w^{1118}*; +/+;* $dFmr1^{3TJ}/dFmr1^{3TJ}$ lines were previously described in [25]. The *dFmr1* wild-type control used in the expression arrays has the genotype w^{1118}*; P[w+,* $dFmr1^{+}$*]/+;* $dFmr1^{3TJ}/dFmr1^{3TJ}$ and was previously described [46, 47]. This wild-type control, although homozygous mutant for the $dFmr1^{3TJ}$, carries a fully functional transgene rescuing wild-type *dFmr1* gene inserted into the *Drosophila* genome by P[*w*+] transposable element. Therefore, it is nearly

genetically identical to the w^{1118}; +/+; $dFmr1^{3TJ}/dFmr1^{3TJ}$ mutant. The *sti* ^{3}/*TM6B Hu* and sti $^{Z3\text{-}5829}$/*TM6B Hu* lines were previously described in [8].

Genetic interaction tests

Female *sti* ^{3}/*TM6B Hu*$^{-}$ and *sti* $^{Z3\text{-}5829}$/*TM6B Hu*$^{-}$ flies were crossed to males w/Y; *P* w^{+} *[sev:dFmr1]*, CyO; +/+ carrying a *dFmr1* eye-specific, overexpression transgene (*sev:dFmr1*). Adult progeny from these two crosses that were *P* w^{+} *[sev:dFmr1]*, CyO; *sti*$^{-}$*/Hu*+ (mutant *sticky*) and *P* w^{+} *[sev:dFmr1]* , *CyO; sti*$^{+}$*/TM6B Hu*$^{-}$ (wild-type *sticky*) were scored for the rough eye phenotype caused by the Fmrp overexpression, as previously described [25]. Eyes were first scored under a dissecting light microscope, and adult flies were placed in 70% ethanol and stored up to one week. Samples were washed 5 times with 100% ethanol and dried under vacuum. Dried whole flies were coated with gold and imaged with an Electroscan E3 scanning electron microscope at 150-200x magnification. Images were captured directly as digital images.

For the *eyeless*-GAL4, double RNAi crosses, female *ey*-GAL4; UAS-*sticky*-RNAi/SM-TM6B flies were crossed to male UAS-*dFmr1*-RNAi/UAS-*dFmr1*-RNAi;+/+ flies. Females from the individual RNAi lines were also crossed to the *ey*-GAL4 stock males. Progeny of the appropriate genotypes were scored under a dissecting light microscope. For *ey*-GAL4/+; UAS-*sticky*-RNAi/+, n=162. For *ey*-GAL4/UAS-*dFmr1*-RNAi; UAS-*sticky*-RNAi/+, n=36. For *ey*-GAL4/+; UAS-*sticky*-RNAi/$dFmr1^{3}$, n=32. For *ey*-GAL4/UAS-*dFmr1*-RNAi, n=62.

RNA preparation and microarray statistics

We used NimbleGen *Drosophila* whole genome arrays containing ~385,000 unique features (12 features per gene, in duplicate) to measure the levels of 15,634 specific transcripts in Cy3 labeled cDNA made from the RNA of whole female flies fed on live yeast. RNA was purified using mirVana protocols from Ambion. Ten arrays were used to probe four different genotypes: Three arrays for w^{1118}; +/+; $dFmr1^{3}/dFmr1^{3}$. Three arrays for +/+; +/+; $sti^{3}/sti^{Z3\text{-}5829}$. Two arrays for Oregon R-S wild-type control. Two arrays for w^{1118}; P[w+, $dFmr1^{+}$]/+; $dFmr1^{3}/dFmr1^{3}$ transgene rescue strain control. Hybridizations, image analysis and quantile normalization were performed by NimbleGen.

We used the limma package for bioconductor to fit linear models to the log transformed, normalized fluorescence intensities and calculated p-values for differential expression using empirical Bayesian methods [48]. These p-values were then used to calculate q-values with the Q-Value package for R [49]. All lists of differentially expressed genes were obtained by q-value thresholds (<.01, <.001). We calculated p-values for the chance of observing a given intersection of two gene lists, a and b, where b is the larger list, by summing over a hypergeometric distribution density function in R. This was performed using the function sum(dhyper(x,m,n,k)) where x represents a vector {i,i+1,i+2,…,j-1,j} (i being the number of genes in the intersection of a and b, and j being the number of genes in a), m is the number of gene in b, n=15,634-m (i.e. the number of genes on the arrays but not in b), and k is the number of gene in a.

In order to determine the random occurrence of enriched GO-terms, we used our experimental expression array data set and randomized the gene names associated with expression values using the "sample" function in R, and used in the following context: sample(x, length(x), replace=FALSE, prob=NULL) where x is a vector of gene IDs. The GO-term analysis was performed on this randomized set as described below.

RT-PCR quantitation

We designed primers specific for 8 genes that were misexpressed in both *sticky* and *dFmr1* mutants (Additional File 4). RNA was prepared as above from females derived from different crosses than those used for the expression arrays. First strand cDNA synthesis was done using oligo-dT and Omniscript reverse transcriptase from Qiagen. Serial dilutions of cDNA were PCR amplified for w^{1118}; +/+; $dFmr1^3/dFmr1^3$ females and +/+; +/+; $sti^3/sti^{Z3\text{-}5829}$ and Oregon R-S females. PCR products were electrophoresed on agarose gels and visualized with ethidium bromide. PCR products in the linear range were imaged and fluorescence was quantitated using ImageQuant v5.2 software. Local average background fluorescence was subtracted and the signal for each gene was normalized to a housekeeping gene (*gpdh*) that showed consistent expression between genotypes on the arrays. Up to ten replicates were done for each gene and p-values were calculated by Student's T-test.

Genetic network analysis

Approximately 46,000 interactions (edges) between 9,196 genes (nodes) were downloaded from the Biomolecular Interaction Database (BIND) and the Database of

interacting proteins (DIP) and integrated into Cytoscape using BioNetBuilder [50, 51] to construct a protein-protein interaction network for *Drosophila melanogaster*. We were able to associate transcript fold changes and p-values with 8846 (>96%) of these genes. The jActiveModules plug-in [34] was used to find and score sub-networks based on the significance of their aggregate changes in expression. We searched for a single path adjusting score for size and using regional scoring with a search depth of 1 and max depth from start nodes of 2. We then identified biological process gene ontology terms that were significantly over represented (Benjamini Hochberg corrected $p < .001$) within the highest scoring sub-networks using the BiNGO plug-in [52].

FlyAtlas

We downloaded the entire FlyAtlas annotated dataset from www.flyatlas.org/annotator.cgi [35]. We then parsed out only the gene names and enrichment factors for each tissue (i.e. ratios of single tissue expression levels to whole fly expression). To construct lists with similar numbers of tissue enriched genes we had to determine an appropriate enrichment factor cutoff for each tissue. This was necessary because different tissues contribute different amounts of gene expression to the total observed in a whole fly. For each tissue, we plotted enrichment factors versus the total number of genes represented on the arrays that had an enrichment factors at least as great. We then fit power distributions ($y=ax^b$) to these data and calculated the enrichment factor cutoffs predicted to provide the top 500 most enriched genes. These cutoffs had to be manually adjusted for ovaries and testes. We compared each tissue enriched gene list to

each of our mutant misexpression lists (*sti*:q<.001, *sti*:q<.01, *dFmr1*:q<.001, and *dFmr1*:q<.01). We were then able to calculate p-values based on the number of genes present in both lists. This was performed by summing over the density of a hypergeometric distribution in R.

Authors' contributions

C.R.B. performed the RNA extractions and computational analyses, designed experiments, and participated in drafting the manuscript. A.M.E. collected oocyte polarity defect data. S.J.S performed the genetic interaction tests. D.C.Z. designed experiments and participated in drafting the manuscript. G.B. collected oocyte polarity defect data, designed experiments, and participated in drafting the manuscript.

Acknowledgements

We are grateful to Jay Konieczka for critical reading of the manuscript and extensive help with expression data analysis and advice on statistical methods. We thank Bruce Walsh for advice on statistical methods, and David Bentley for assistance with SEM imaging. We also thank JMST for useful discussions. S.J.S. was supported in part by the University of Arizona Undergraduate Biology Research Program funded by the Howard Hughes Medical Institute (#71196-521304). C.R.B. was supported in part by an NSF IGERT fellowship in Genomics. This work was supported by a grant to G.B. from the National Institutes of Health (RO1 GM069462) and University of Arizona start-up funds to D.C.Z.

References

1. Lee LA, Orr-Weaver TL: **Regulation of cell cycles in Drosophila development: intrinsic and extrinsic cues**. *Annu Rev Genet* 2003, **37**:545-78.

2. Maurange C, Lee N, Paro R: **Signaling meets chromatin during tissue regeneration in Drosophila**. *Curr Opin Genet Dev* 2006, **16**:485-9.

3. McClure KD, Schubiger G: **Transdetermination: Drosophila imaginal disc cells exhibit stem cell-like potency**. *Int J Biochem Cell Biol* 2007, **39**:1105-18.

4. Baker NE: **Patterning signals and proliferation in Drosophila imaginal discs**. *Curr Opin Genet Dev* 2007, **17**:287-93.

5. Spradling AC: **Developmental genetics of oogenesis**. In: *The development of Drosophila melanogaster* Edited by M Bate, A Martinez Arias, vol. 1. pp. 1-70. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1993: 1-70.

6. Bosco G, Orr-Weaver TL: **Regulation of the cell cycle during oogenesis and early embryogenesis in Drosophila**. In: *Regulation of gene expression at the beginning of animal development* Edited by M DePamphilis. pp. pp.107-154. Amsterdam: Elsevier; 2002: pp.107-154.

7. Deshpande G, Calhoun G, Schedl P: **The drosophila fragile X protein dFMR1 is required during early embryogenesis for pole cell formation and rapid nuclear division cycles**. *Genetics* 2006, **174**:1287-98.

8. Sweeney SJ, Campbell P, Bosco G: **Drosophila sticky/citron kinase is a regulator of cell cycle progression, genetically interacts with Argonaute 1 and modulates epigenetic gene silencing**. *Genetics* 2008, **178**:1311-1325.

9. Featherstone DE, Broadie K: **Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network**. *Bioessays* 2002, **24**:267-74.

10. He X, Zhang J: **Toward a molecular understanding of pleiotropy**. *Genetics* 2006, **173**:1885-91.

11. Naim V, Imarisio S, Di Cunto F, Gatti M, Bonaccorsi S: **Drosophila citron kinase is required for the final steps of cytokinesis**. *Mol Biol Cell* 2004, **15**:5053-63.

12. D'Avino PP, Savoian MS, Glover DM: **Mutations in sticky lead to defective organization of the contractile ring during cytokinesis and are enhanced by Rho and suppressed by Rac**. *J Cell Biol* 2004, **166**:61-71.

13. Zhao ZS, Manser E: **PAK and other Rho-associated kinases--effectors with surprisingly diverse mechanisms of regulation**. *Biochem J* 2005, **386**:201-14.

14. Matsumura F: **Regulation of myosin II during cytokinesis in higher eukaryotes**. *Trends Cell Biol* 2005, **15**:371-7.

15. Di Cunto F, Imarisio S, Hirsch E, Broccoli V, Bulfone A, Migheli A, Atzori C, Turco E, Triolo R, Dotto GP, et al: **Defective neurogenesis in citron kinase knockout mice by altered cytokinesis and massive apoptosis**. *Neuron* 2000, **28**:115-27.

16. Ackman JB, Ramos RL, Sarkisian MR, Loturco JJ: **Citron kinase is required for postnatal neurogenesis in the hippocampus**. *Dev Neurosci* 2007, **29**:113-23.

17. Di Cunto F, Ferrara L, Curtetti R, Imarisio S, Guazzone S, Broccoli V, Bulfone A, Altruda F, Vercelli A, Silengo L: **Role of citron kinase in dendritic morphogenesis of cortical neurons**. *Brain Res Bull* 2003, **60**:319-27.

18. LoTurco JJ, Sarkisian MR, Cosker L, Bai J: **Citron kinase is a regulator of mitosis and neurogenic cytokinesis in the neocortical ventricular zone**. *Cereb Cortex* 2003, **13**:588-91.

19. Berto G, Camera P, Fusco C, Imarisio S, Ambrogio C, Chiarle R, Silengo L, Di Cunto F: **The Down syndrome critical region protein TTC3 inhibits neuronal differentiation via RhoA and Citron kinase**. *J Cell Sci* 2007, **120**:1859-67.

20. Kazemi-Esfarjani P, Benzer S: **Genetic suppression of polyglutamine toxicity in Drosophila**. *Science* 2000, **287**:1837-40.

21. Furuyashiki T, Fujisawa K, Fujita A, Madaule P, Uchino S, Mishina M, Bito H, Narumiya S: **Citron, a Rho-target, interacts with PSD-95/SAP-90 at glutamatergic synapses in the thalamus**. *J Neurosci* 1999, **19**:109-18.

22. Zhang W, Vazquez L, Apperson M, Kennedy MB: **Citron binds to PSD-95 at glutamatergic synapses on inhibitory neurons in the hippocampus**. *J Neurosci* 1999, **19**:96-108.

23. Madaule P, Furuyashiki T, Eda M, Bito H, Ishizaki T, Narumiya S: **Citron, a Rho target that affects contractility during cytokinesis**. *Microsc Res Tech* 2000, **49**:123-6.

24. Zhang W, Benson DL: **Targeting and clustering citron to synapses**. *Mol Cell Neurosci* 2006, **31**:26-36.

25. Jin P, Zarnescu DC, Ceman S, Nakamoto M, Mowrey J, Jongens TA, Nelson DL, Moses K, Warren ST: **Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway**. *Nat Neurosci* 2004, **7**:113-7.

26. Penagarikano O, Mulle JG, Warren ST: **The pathophysiology of fragile x syndrome**. *Annu Rev Genomics Hum Genet* 2007, **8**:109-29.

27. Yang L, Duan R, Chen D, Wang J, Jin P: **Fragile X mental retardation protein modulates the fate of germline stem cells in Drosophila**. *Hum Mol Genet* 2007, **16**:1814-20.

28. Lee A, Li W, Xu K, Bogert BA, Su K, Gao FB: **Control of dendritic development by the Drosophila fragile X-related gene involves the small GTPase Rac1**. *Development* 2003, **130**:5543-52.

29. Schenck A, Bardoni B, Langmann C, Harden N, Mandel JL, Giangrande A: **CYFIP/Sra-1 controls neuronal connectivity in Drosophila and links the Rac1 GTPase pathway to the fragile X protein**. *Neuron* 2003, **38**:887-98.

30. Castets M, Schaeffer C, Bechara E, Schenck A, Khandjian EW, Luche S, Moine H, Rabilloud T, Mandel JL, Bardoni B: **FMRP interferes with the Rac1 pathway and controls actin cytoskeleton dynamics in murine fibroblasts**. *Hum Mol Genet* 2005, **14**:835-44.

31. Reeve SP, Bassetto L, Genova GK, Kleyner Y, Leyssen M, Jackson FR, Hassan BA: **The Drosophila fragile X mental retardation protein controls actin dynamics by directly regulating profilin in the brain**. *Curr Biol* 2005, **15**:1156-63.

32. Di Cunto F, Calautti E, Hsiao J, Ong L, Topley G, Turco E, Dotto GP: **Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the Rho-Rac-binding protein Citron**. *J Biol Chem* 1998, **273**:29706-11.

33. Zarnescu DC, Jin P, Betschinger J, Nakamoto M, Wang Y, Dockendorff TC, Feng Y, Jongens TA, Sisson JC, Knoblich JA, et al: **Fragile X protein functions with lgl and the par complex in flies and mice**. *Dev Cell* 2005, **8**:43-52.

34. Ideker T, Ozier O, Schwikowski B, Siegel AF: **Discovering regulatory and signalling circuits in molecular interaction networks**. *Bioinformatics* 2002, **18 Suppl 1**:S233-40.

35. Chintapalli VR, Wang J, Dow JA: **Using FlyAtlas to identify better Drosophila melanogaster models of human disease**. *Nat Genet* 2007, **39**:715-20.

36. Pai LM, Wang PY, Chen SR, Barcelo G, Chang WL, Nilson L, Schupbach T: **Differential effects of Cbl isoforms on Egfr signaling in Drosophila**. *Mech Dev* 2006, **123**:450-62.

37. Guy GR, Wong ES, Yusoff P, Chandramouli S, Lo TL, Lim J, Fong CW: **Sprouty: how does the branch manager work?** *J Cell Sci* 2003, **116**:3061-8.

38. Nilson LA, Schupbach T: **EGF receptor signaling in Drosophila oogenesis**. *Curr Top Dev Biol* 1999, **44**:203-43.

39. Karess RE, Chang XJ, Edwards KA, Kulkarni S, Aguilera I, Kiehart DP: **The regulatory light chain of nonmuscle myosin is encoded by spaghetti-squash, a gene required for cytokinesis in Drosophila**. *Cell* 1991, **65**:1177-89.

40. Wheatley S, Kulkarni S, Karess R: **Drosophila nonmuscle myosin II is required for rapid cytoplasmic transport during oogenesis and for axial nuclear migration in early embryos**. *Development* 1995, **121**:1937-46.

41. Jordan P, Karess R: **Myosin light chain-activating phosphorylation sites are required for oogenesis in Drosophila**. *J Cell Biol* 1997, **139**:1805-19.

42. Ghabrial A, Ray RP, Schupbach T: **okra and spindle-B encode components of the RAD52 DNA repair pathway and affect meiosis and patterning in Drosophila oogenesis**. *Genes and Dev.* 1998, **12**:2711-2723.

43. Klattenhoff C, Bratu DP, McGinnis-Schultz N, Koppetsch BS, Cook HA, Theurkauf WE: **Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response**. *Dev Cell* 2007, **12**:45-55.

44. MacDougall N, Lad Y, Wilkie GS, Francis-Lang H, Sullivan W, Davis I: **Merlin, the Drosophila homologue of neurofibromatosis-2, is specifically required in posterior follicle cells for axis formation in the oocyte**. *Development* 2001, **128**:665-73.

45. Zarnescu DC, Shan G, Warren ST, Jin P: **Come FLY with us: toward understanding fragile X syndrome**. *Genes Brain Behav* 2005, **4**:385-92.

46. Dockendorff TC, Su HS, McBride SM, Yang Z, Choi CH, Siwicki KK, Sehgal A, Jongens TA: **Drosophila lacking dfmr1 activity show defects in circadian output and fail to maintain courtship interest**. *Neuron* 2002, **34**:973-84.

47. Costa A, Wang Y, Dockendorff TC, Erdjument-Bromage H, Tempst P, Schedl P, Jongens TA: **The Drosophila fragile X protein functions as a negative regulator in the orb autoregulatory pathway**. *Dev Cell* 2005, **8**:331-42.

48. Smyth GK: **Limma: linear models for microarray data.** In: *Bioinformatics and Computational Biology Solutions using R and Bioconductor* Edited by VC R. Gentleman, S. Dudoit, R. Irizarry, W. Huber. pp. 397-420. New York: Springer; 2005: 397-420.

49. Storey JD, Tibshirani R: **Statistical significance for genomewide studies**. *Proc Natl Acad Sci U S A* 2003, **100**:9440-5.

50. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome Res* 2003, **13**:2498-504.

51. Avila-Campillo I, Drew K, Lin J, Reiss DJ, Bonneau R: **BioNetBuilder: automatic integration of biological networks**. *Bioinformatics* 2007, **23**:392-3.

52. Maere S, Heymans K, Kuiper M: **BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks**. *Bioinformatics* 2005, **21**:3448-9.

Figure Legends

Figure 1 - *sticky* mutations rescue *dFmr1* overexpression phenotypes in the *Drosophila* eye

(A) Scanning electron microscope visualization of eye developmental defects due to overexpression of *dFmr1* by the *sevenless* promoter in a *sticky* wild-type background (*sev:dFmr1, sti*$^{+/+}$). (B-C) One mutant copy of the *sticky*, *sti* $^{3/+}$ or *sti* $^{Z3\text{-}5829/+}$ dominantly suppresses the rough eye phenotype caused by *sev:dFmr1* overexpression. (D-E) The crater-like necrotic adult eye tissue caused by the *sev:dFmr1* overexpression (see arrows) is almost completely suppressed by one mutant copy of *sti* $^{Z3\text{-}5829/+}$. (F) Rho is known to directly activate *sticky* (solid arrow) and Rac is known to inhibit (solid bar) *sticky* activity. Other, unknown (?) developmental and cellular signals may also regulate *sticky* activity. We propose that both Ago1 and Fmrp activities are downstream of *sticky*, and that some wild-type or gain-of-function Ago1 and Fmrp dependant processes are sensitive to *sticky*. Dashed arrows indicate possible indirect interactions. This model is supported by the genetic interaction between *sticky* and *dFmr1* shown above and that previously reported for *sticky* and *Ago1* [8]. Both Ago1 and Fmrp are known to directly control microRNA mediated gene regulation (solid arrows). Fmrp also directly targets Rac1 mRNA for repression [28]. *sticky* and Fmrp proteins also control gene expression through heterochromatin gene silencing (bars and dashed arrows). It is not known whether these effects on heterochromatin are direct or indirect.

Figure 2 – dFmr1 loss of function enhances sticky RNAi knockdown phenotypes

Plot of phenotype penetrance for the genotypes: *ey-GAL4*/+;*UAS-sticky* RNAi/+ (n=162), *ey-GAL4/UAS-dFmr1* RNAi; *UAS-sticky* RNAi/+ (n=36), and *ey-GAL4*/+;*UAS-sticky* RNAi/$dFmr1^3$ (n=32). No phenotypes were observed in flies of the genotype *ey-GAL4/UAS-dFmr* RNAi (data not shown). Phenotypes included are Rough Eye (RE), Reduced Eye (Red), No Eye (NE), Vibrissae (Vib), Frontorbital bristles (FO), Orbital and Vertical bristles (OV), Postorbital bristles (PO), Frontal bristles (F), Maxillary Palps (MP), and Necrosis (N). For RE, Red, Vib and PO we found statistically significant enhancement in the double RNAi lines when compared to the *sticky*-RNAi line alone, $p<0.05$ in a Fisher's Exact test. This was also true for *sticky*-RNAi line in a $dFmr1^3$ heterozygous compared to the *sticky*-RNAi line alone.

Figure 3 - A common set of genes is misexpressed in *sticky* and *dFmr1* mutants

(A) Three biological replicates of *sticky* total RNA and three biological replicates of *dFmr1* mutants were analyzed by NimbleGen gene expression arrays. These data were compared to expression data from two biological replicates of *OrR* and two biological replicates of w^{1118} wild-type control RNA. Expression data from the two controls were combined in order to better account for differences in genetic background. (B-C) Venn diagrams showing the number of genes with significantly altered expression levels in *sticky* and *dFmr1* mutants. (B) With a significance threshold of q<.001, we found 249 misexpressed genes and 1,044 misexpressed genes in *sticky* and *dFmr1* mutants, respectively. There were 40 genes differentially expressed in both mutant strains

(p=2.37e-6) and 14,322 genes with no significant change in expression. (C) With a threshold of q<.01, we found 771 misexpressed genes and 2,894 misexpressed genes in *sticky* and *dFmr1* mutants, respectively. There were 246 genes differentially expressed in both mutants (p=4.27e-14) and 11,999 genes with no significant change in expression. (D) Of the genes that were misexpressed in either mutant the vast majority of them were overexpressed.

Figure 4 - Commonly misexpressed genes tend to be overexpressed

Plot of microarray log fold changes observed in *sticky* versus *dFmr1* for the 40 genes found to be differentially expressed in both mutants. R^2=.581 and a two-tailed P value is <0.0001 for 39 degrees of freedom.

Figure 5 - RT-PCR quantitation of mRNA levels for candidate genes

(A) Mean gene expression levels with standard errors as measured by RT-PCR. For each gene, band intensity was normalized to *gpdh* expression from the same cDNA template. (B) Mean gene expression levels with standard errors as measure by microarray.

Figure 6 - Expression Network

Protein-protein interaction network overlaid with *sticky* mutant expression data. Red indicates genes that are up-regulated in *sticky* mutant compared to wild-type and green indicates down-regulation. Yellow nodes represent genes with little to no change in expression and gray nodes represent genes that could not be assigned expression data.

Node size indicates the significance (i.e. p-value) of the change in expression, where larger nodes are more significant and smaller nodes are less significant.

Figure 7 - *sticky* mutant active sub-network

Highest scoring sub-network found by jActiveModules using *sticky* mutant expression data. Red indicates genes that are up-regulated in the *sticky* mutant compared to wild type and green indicates down-regulation. Yellow nodes represent genes with little to no change in expression and gray nodes represent genes that could not be assigned expression data. Node size indicates the significance (i.e. p-value) of the change in expression, where larger nodes are more significant and smaller nodes are less significant.

Figure 8 - *dFmr1* mutant active sub-network

Highest scoring sub-network found by jActiveModules using *dFmr1* expression data. Red indicates genes that are up-regulated in *dFmr1* compared to wild type and green indicates down-regulation. Yellow nodes represent genes with little to no change in expression and gray nodes represent genes that could not be assigned expression data. Node size indicates the significance (i.e. p-value) of the change in expression, where larger nodes are more significant and smaller nodes are less significant.

Figure 9 - *sticky* BiNGO

Directed acyclic graph of over-represented GO-terms in the highest scoring *sticky* active sub-network. Node size indicates the number of genes associated with each GO-term

(i.e. larger nodes represent GO-terms associated with many genes in the active sub-network). Node color indicates significance of over-representation based on Benjamini-Hochberg corrected p-value: white = p>.0001, yellow = p<.0001, red = p<<.0001.

Figure 10 - *dFmr1* BiNGO

Directed acyclic graph of over-represented GO-terms in the highest scoring *dFmr1* active sub-network. Node size indicates the number of genes associated with each GO-term (i.e. larger nodes represent GO-terms associated with many genes in the active sub-network). Node color indicates significance of over-representation based on Benjamini-Hochberg corrected p-value: white = p>.0001, yellow = p<.0001, red = p<<.0001.

Figure 11 - Tissue specificity of misexpressed genes

For each tissue (head, brain, crop, midgut, hindgut, Malpighian tubule, ovary, testes, accessory gland, larval Malpighian tubule, larval fat body, thoracico-abdominal ganglia, and thoracic and abdominal carcass), a list of genes enriched in that tissue was compiled based on the FlyAtlas data set (see methods). These lists were then compared to the lists of misexpressed genes in *sticky* and *dFmr1* mutants at q<.01 and q<.001. Each column represents the percent of all misexpressed genes in a given list that were enriched in a given tissue. The black columns represent the percent of all genes in the fly genome that were defined as enriched in a particular tissue. These can be interpreted as the percentage of tissue enriched gene that would be expected in a random list of genes.

Figure 12 - *sticky* and *dFmr1* mutant oocytes exhibit predicted axis polarity defects

(A-B) Phase contrast images of *sticky* $^{+/-}$ wild-type and *dFmr1*$^{+/-}$ wild-type, stage 14 oocytes showing normal dorsal/ventral axis formation as well as normal dorsal appendages. (C-E) Stage 14 *sticky* $^{Z3\text{-}5829}$/*sticky* ^{3} (*sticky* $^{-/-}$) and *dFmr1*$^{3/3}$ (*dFmr1*$^{-/-}$) mutant oocytes are completely ventralized with missing or severely malformed dorsal appendages. Arrowheads point to regions of the oocytes where dorsal appendages would normally emanate.

Table 1 Genes similarly regulated by *sticky* and *dFmr1*

Gene Name	*Sticky* Log Fold Change	*Sticky* p-value	*dFmr1* Log Fold Change	*dFmr1* p-value
Adh	0.65	1.62E-05	0.68	8.66E-06
Ag5r2	0.85	3.84E-08	0.71	7.76E-07
antdh	0.89	3.97E-05	0.90	3.34E-05
Cbl	0.59	4.90E-06	0.56	1.09E-05
CG10659	0.70	4.79E-06	0.64	1.74E-05
CG10908	0.48	3.98E-05	0.56	5.33E-06
CG12656	0.65	4.04E-05	0.61	9.80E-05
CG13324	1.08	2.56E-08	0.93	3.17E-07
CG13325	1.17	1.42E-08	0.76	1.40E-05
CG13420	0.70	4.44E-05	0.78	1.02E-05
CG13482	1.06	9.09E-09	1.07	7.54E-09
CG15194	0.88	1.05E-07	0.66	8.93E-06
CG16904	0.48	2.70E-05	0.54	6.41E-06
CG16996	0.59	4.88E-06	0.77	7.63E-08
CG17012	0.78	4.42E-05	1.11	2.52E-07
CG17327	0.69	6.19E-07	0.52	3.42E-05
CG17352	1.24	7.63E-11	0.59	2.49E-05
CG31681	1.36	4.07E-09	0.89	4.61E-06
CG32642	-0.70	8.48E-06	-0.62	4.01E-05
CG4461	0.68	1.34E-05	0.84	5.85E-07
CG4653	1.21	2.86E-11	1.38	1.92E-12
CG4734	1.62	9.69E-10	1.13	5.17E-07
CG4847	1.15	3.32E-05	1.64	1.82E-07

CG5506	1.13	1.66E-09	0.68	7.26E-06
CG6432	0.66	1.66E-05	0.80	1.08E-06
CG8562	0.68	4.98E-06	0.63	1.35E-05
CG8774	0.67	2.55E-05	0.66	2.96E-05
CG9080	0.62	3.00E-05	0.82	4.75E-07
CG9396	1.17	4.96E-09	0.73	1.13E-05
CG9672	0.72	1.30E-05	1.01	6.99E-08
lambdaTry	0.75	1.75E-06	0.68	6.62E-06
lectin-37Db	0.77	5.70E-09	0.44	3.34E-05
mex1	0.87	5.54E-08	0.52	9.37E-05
PGRP-SB1	1.09	1.77E-08	0.84	1.24E-06
PGRP-SC1a	0.71	1.16E-05	0.67	2.63E-05
Reg-3	0.67	2.46E-05	0.66	3.41E-05
Tsf1	1.58	1.46E-06	1.27	3.19E-05