Homology and Heterology Effects in Drosophila: Cohesin and Condensin as Chromosome Choreographers

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Senaratne, Tharanga Niroshini. 2016. Homology and Heterology Effects in Drosophila: Cohesin and Condensin as Chromosome Choreographers. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493465">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493465</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Homology and Heterology Effects in Drosophila: 
Cohesin and Condensin as Chromosome Choreographers

A dissertation presented
by
Tharanga Niroshini Senaratne

to
The Division of Medical Sciences

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Genetics and Genomics

Harvard University
Cambridge, Massachusetts
March 2016
Homology and Heterology Effects in Drosophila:
Cohesin and Condensin as Chromosome Choreographers

Abstract

The organization of interphase nuclei has numerous consequences for gene expression and genome stability in eukaryotes. Here, we present studies of chromosome organization in Drosophila melanogaster over the course of the cell cycle. We report surprising observations suggesting that cohesin, a protein complex essential in mitosis for holding together the products of DNA replication known as sister chromatids, may not be required to keep sister chromatids in close proximity during interphase. These observations raise questions regarding the nature of cohesin-independent connections between chromatids, why the cell might have such connections in addition to those mediated by cohesin proteins, and how cohesin-independent mechanisms might contribute to other interchromosomal associations and nuclear organization throughout the cell cycle. A well-known feature of nuclear organization in Drosophila is the somatic pairing of maternal and paternal homologous chromosomes; we find that certain factors contributing to the pairing of homologs, specifically, condensin II and its regulators, also contribute to the organization of sister chromatids, findings which have interesting implications for the nature of pairing. Finally, we examine another type of inter-chromosomal interaction occurring in Drosophila nuclei, which is the nonhomologous clustering of centromeres, and identify candidate genes that regulate this organization. Overall, this work expands current knowledge on both homologous and heterologous inter-chromosomal interactions, and highlights the relationship between chromosome organization in interphase and in mitosis.
# Table of Contents

## Chapter 1: Introduction

- Overview ......................................................................................................................... 2
- Techniques for Studying Nuclear Organization .............................................................. 3
- Organization of Mitotic Chromosomes ............................................................................ 5
- Factors that Shape Mitotic Chromosomes .................................................................... 7
  - Topoisomerases ............................................................................................................. 7
  - SMC Complexes ........................................................................................................... 7
  - Kinetochore proteins and the organization of centromeres ........................................... 10
- The Different Layers of Interphase Nuclear Organization ............................................... 11
  - The nuclear lamina, nucleolus and other nuclear substructures .................................... 11
  - The organization of heterochromatin versus euchromatin .......................................... 13
  - The organization of individual chromosomes: CTs, TADs, and sub-TADs .................... 15
  - Heterologous interactions: gene regulation in *cis* and in *trans* ............................... 19
  - Heterologous interactions: centromere and telomere clustering ................................ 22
  - Homolog pairing, and homology effects in Drosophila .............................................. 23
  - Examples of homology effects outside Drosophila ..................................................... 25
- Factors that Organize the Interphase Nucleus ............................................................... 26
  - Architectural proteins: CTCF and other genome organizers ....................................... 26
  - Roles for cohesin in the interphase nucleus ................................................................ 27
  - Roles for condensin in the interphase nucleus ........................................................... 28
  - Long non-coding RNAs in nuclear organization .......................................................... 30
- Conclusions, and Aims of this Dissertation .................................................................... 31
  - Open questions on the roles of mitotic proteins in interphase nuclear organization .......... 31
  - Studying homologous and heterologous interactions by FISH ................................... 32
  - Drosophila as a model system for studying nuclear organization .................................. 32
- References for Chapter 1 ............................................................................................... 34
Chapter 2: The pairing of homologs and sister chromatids in interphase

nuclei of *Drosophila melanogaster* is cohesin-independent ................................................................. 66

Abstract ........................................................................................................................................ 67
Introduction ...................................................................................................................................... 67
Materials and Methods ................................................................................................................ 73
Results ........................................................................................................................................... 78
Discussion ....................................................................................................................................... 90
Acknowledgments ........................................................................................................................... 93
References for Chapter 2 ................................................................................................................. 93

Chapter 3: Condensin-regulated homolog pairing also contributes
to sister chromatid pairing in Drosophila ......................................................................................... 99

Abstract ........................................................................................................................................ 100
Introduction ...................................................................................................................................... 100
Materials and Methods ................................................................................................................ 106
Results ........................................................................................................................................... 108
Discussion ....................................................................................................................................... 118
Acknowledgments ........................................................................................................................... 120
References for Chapter 3 ................................................................................................................. 121

Chapter 4: Studies on the organization of pericentric heterochromatin
and the clustering of centromeres in *Drosophila* cells ..................................................................... 126

Abstract ........................................................................................................................................ 127
Introduction ...................................................................................................................................... 127
Materials and Methods ................................................................................................................ 133
Results ........................................................................................................................................... 137
Discussion ....................................................................................................................................... 158
Acknowledgments ........................................................................................................................... 161
References for Chapter 4 ................................................................................................................. 161
Glossary of Terms

Chr: Frequently used as an abbreviation for “chromosome”

IF: Immunofluorescence

FISH: Fluorescent in situ hybridization

Sisters: Sister chromatids, the products of DNA replication

Homologs: Maternal and paternal copies of chromosomes present in diploid cells

Pairing: Interactions between homologous chromosomes

Cohesion: The interaction by which sister chromatids are kept together until mitosis

Clustering: Interactions between heterologous chromosomes
Acknowledgements

First of all, I would like to thank my advisor, Ting Wu. From the day I met her, Ting has inspired me to “think big” and to be fearless in asking questions. She has constantly refreshed me with her excitement about science, and, I have learned much from her about conducting rigorous science, the art of writing, and the presentation of ideas. She also excels at developing a fun and creative research environment, from within the lab to the “Pairing Hoedown.” I consider myself lucky to have been in a lab where the PI genuinely cares about the people she mentors, and as I move forward, I will continue to remember and value all the things I have learned from her.

Secondly, I thank Eric Joyce, a post-doctoral fellow in the Wu lab. Eric was my mentor from my very first day in the lab as a rotation student, and in addition to training me and providing scientific guidance in all my projects, was my collaborator in the work presented in Chapter 4. Eric, I will miss all the conversations we had in our bay about science, lab, and life, and hope that at least some of them were as productive for you as they have been for me. As someone who has experienced and truly benefited from your mentorship over the years, I have no doubt that you will make a fantastic PI when you start your own lab in a couple months.

I would also like to thank Tiao Xie, a former fellow at the Image and Data Analysis Core at Harvard Medical School. Tiao’s help was invaluable in getting me started with MATLAB-based image analysis, and he was extremely generous in sharing code he had written for this purpose. Tiao’s help saved me countless hours that I would have spent counting FISH dots, and helped me develop image analysis skills that allowed me to participate in several collaborations both within the Wu lab and beyond, which truly enriched my graduate school experience. Special thanks go to the members of my dissertation advisory committee, Bob Kingston, Kami Ahmad, and Steve Elledge for all the scientific advice and guidance they have given me over the years. Thank you for the helpful meetings, and for caring.

Over the years I have spent in the Wu lab, I have also had the privilege of working with very talented and fun colleagues, including Brian, Caroline, Chamith, Fred, Guille, Guy, Hien, Jelena, Kwasi, Mohammed, Nick, Quentin, Ruth, and Sonny. I would especially like to thank Brian and Sonny for being such wonderful senior graduate students to look up to, Mohammed for all his smiles and for keeping the lab running, and Nick and
Chamith for becoming my close friends along the way. Wu lab members, thank you for all the times we shared in the break room, from morning lab meetings to afternoon coffee breaks, and especially for all the laughs.

Finally, I must acknowledge the HMS Department of Genetics and the Program in Genetics and Genomics for the support they have given me over the years, as well as the incredible community that is the PhD program in Biological and Biomedical Sciences. From Kate, Maria and Danny in the BBS office, to the program heads and faculty, the support we are given here as students is a real gift. My fellow students in BBS have provided an endless source of inspiration, support, and joy, and I am sure they will continue to do so in the years ahead. Special thanks to my classmates, Alex, Daniel, Danielle, Kaitlin, Kaloyan, Kostadin, Laura, Mary, Silvia, Stéphane, Stephen, Tyler, Wei Leong, and of course, Ben Vincent. I love you guys.
Dedication

This dissertation is dedicated to my family: my father, Chandana, who teaches me to be proud of my accomplishments (though I will never be as proud as he is), my mother, Vinitha, who inspires me to be strong, and my brother, Ruwan, who is a great source of “technical support” and reminds me that what seem like problems are actually “not a big cashew.”
Chapter 1

Introduction
Overview

The organization of eukaryotic genomes into chromosomes, and the changes chromosomes undergo over the life of a cell, have long fascinated scientists. In fact, it was observations about chromosome organization, their numbers, and their behavior during cell divisions that led Theodor Boveri and Walter Sutton to recognize beginning in 1902 that chromosomes were the units of heredity that Mendel had previously proposed [1–3]. Indeed, it is hard not to appreciate the organization of chromosomes during cell division, when compact chromosomes are segregated in a carefully orchestrated and rapidly executed cell cycle event. During mitotic divisions, segregation occurs between the replication products of each chromosome, known as sister chromatids, resulting in daughter cells with equal numbers of chromosomes as compared to the parent. During the meiotic divisions that form germ cells, it is pairs of homologous chromosomes that first associate and then segregate, ultimately forming daughter cells with half the number of chromosomes present in the parent. Yet, apart from stages of rapid growth such as in embryos, such divisions and the compaction associated with them are a small part of a chromosome’s life. During much of the cell cycle, known as interphase, chromosomes exist in relatively decompacted forms, allowing the information on them to be used in different ways.

Perhaps even more amazingly, these decondensed forms are themselves organized in the structure that is the interphase nucleus. While the existence of specific territories for interphase chromosomes was suggested as early as 1885 by Carl Rabl and in the early 1900s by Boveri and others, it was in the late twentieth century that evidence started to mount against the idea that interphase nuclei were unorganized “bowls of spaghetti” [reviewed by 4]. Studies by Thomas Cremer’s group starting in the 1970s showed that when nuclei were exposed to laser-UV-microirradiation, the resulting DNA damage occurred only on a small subset of chromosomes, supporting the hypothesis that chromosomes exist in distinct territories in interphase rather than being completely intermingled [5–8]. Subsequently, techniques were developed for in situ hybridization of chromatin and for the creation of chromosome-specific and region-specific probes, allowing further examination of chromosome territories. Recent years have seen a burst in the availability of higher resolution techniques for understanding genome organization, based both on visualization and other methods, and each day, our understanding of the interphase nucleus becomes more nuanced.
While we have learned much about interphase nuclear organization and its functional consequences, we still understand relatively little about the underlying mechanisms responsible, and this is the motivation for the work presented here. Our laboratory aims to identify genes and processes important for interphase nuclear organization and in doing so, we identify connections between the organization of chromosomes in interphase and in mitosis. In this introduction, I begin with a summary of the major techniques that have been used to study genome organization, to provide context for the results described. Next, I briefly discuss mitotic chromosomes and the proteins involved in their organization, particularly cohesin and condensin proteins. I then review the organization of interphase chromosomes, discussing different levels of nuclear organization, examples of the functional relevance of this organization, and known regulators of these processes, including interphase roles for cohesin and condensin. To conclude, I describe key questions in the field which motivated the work presented in this dissertation, and explain my choice of technique and model system: visual examination of chromosome organization in the cells of the fruit fly, Drosophila melanogaster.

Techniques for Studying Nuclear Organization

The methods used to study nuclear organization can be divided into two main categories, visual and molecular [reviewed by 9]. Early visual examination of the nucleus used light microscopy to identify structures such as the nucleolus [10], and cytological techniques that differentially stain different kinds of chromatin, which allowed the identification of chromosome banding patterns and initial investigations of chromosome organization in the interphase nucleus [for example, 11]. Modern-day techniques allow us to probe the interphase nucleus in a more detailed way using in situ hybridization (ISH), wherein nucleic acid probes marked with some sort of label are allowed to hybridize to the complementary sequences in the genome, allowing the visualization of those regions [reviewed by 12].

Early ISH experiments used probes prepared from subsets of genomic DNA or RNA and carrying a radioactive label [13–15]. A major advance was in the use of fluorescently labelled antibodies to identify the sites of probe binding, which was the beginning of fluorescent in situ hybridization or FISH [16]. Today, rather than labelling with antibodies, FISH probes can be directly conjugated with fluorescent molecules [17], and instead of
preparing FISH probes from genomic DNA clones, oligonucleotide probes can be designed to precisely target virtually any sequence, whether a 10 kilobase region or an entire chromosome, and synthesized in vitro [18–21]. FISH can be targeted to specific DNA loci, or to RNA to visualize transcription, and the position of these sequences can be examined relative to other nuclear substructures identified by DAPI staining of DNA or immunofluorescent labelling of proteins.

The other major class of technique used to examine nuclear organization involves molecular characterization of DNA sequences based on their physical proximity to nuclear features or to other DNA sequences. For example, a technique known as DamID fuses DNA adenine methyltransferase (Dam) from E. coli to nuclear proteins, resulting in the methylation of DNA sequences nearby, which can then be mapped using sequencing or other methods [22–25]. Other molecular methods to assay nuclear organization include chromosome conformation capture (3C) [26], and its many variants, including 4C, 5C and Hi-C technologies [reviewed by 9,27,28]. These techniques start with cross-linking of DNA by formaldehyde fixation, followed by restriction enzyme digestion, ligation of cross-linked sequences, and identification of the ligated sequences by PCR, microarray, or high-throughput sequencing [9,27]. Cross-linking frequency between different sequences is used as an assay of physical proximity. While 3C is used to assay proximity between two specific loci, the development of Hi-C technologies have allowed the assessment of interactions on a genome-wide scale [29–31].

Each of these types of method, visual and molecular, has its particular advantages and limitations, and a thorough examination of nuclear architecture requires both techniques [32]. With FISH, one is typically limited to the examination of a few regions at a time, depending on the number of colors that can be resolved. Additionally, the size of features that can be distinguished is limited by the resolution of light microscopy (around 200 nm), though recent technologies that combine FISH with super-resolution microscopy are pushing past this barrier [33]. Finally, FISH experiments involve fixing the cell and exposing it to high temperatures in order to denature the genomic DNA and allow probes to hybridize to their targets, and while experiments have shown that certain aspects of nuclear architecture are preserved [34], it is unclear how we might be damaging the nucleus through this process. Progress has been made towards live-imaging of DNA loci, taking advantage of sequence-specific DNA binding proteins such as LacI [35], TALEs [36], or Cas9 [37], but these approaches are limited in the types of loci that can
be assayed. Despite these drawbacks, FISH is an extremely powerful tool for examining nuclear organization. A major advantage of FISH is that it allows the examination of events taking place in single cells, and thus of rare or stage-specific events. Additionally, FISH is sensitive to the relative positioning of homologous sequences such as homologous chromosomes or sister chromatids, which are more difficult to resolve by molecular methods.

3C technologies generally provide higher resolution than FISH, since co-localization of two FISH signals means they are within hundreds of nanometers, while 3C-based methods are thought to be more sensitive to actual contact between DNA [38]. Additionally, 5C and Hi-C in particular are higher-throughput methods than FISH, allowing the entire genome to be assayed at once [38]. However, a drawback of 3C-based methods is that they provide an averaged view of the interactions occurring in an entire population of cells and relative contact frequencies, meaning that the frequency of individual events is hard to interpret and co-occurrence is difficult to demonstrate. While Hi-C has recently been developed for application to single cells, the resolution of this technique is limited [39,40]. Another limitation is that, since these methods detect physical interactions based on ligation between different sequences, interactions between homologous sequences are harder to detect. Recent studies have taken advantage of the single nucleotide polymorphisms (SNPs) that vary between homologs to map chromatin interactions in a haplotype specific way [41–43], but are still dependent on the positions and density of SNPs to identify inter-homolog interactions and cannot distinguish sister chromatids. For these reasons, insights from both visual and molecular techniques have shaped our understanding of chromosome organization in different ways, and will continue to do so going forward.

**Organization of Mitotic Chromosomes**

As mentioned above, mitotic chromosomes have been studied for over a hundred years, and while we know much about the genes that regulate their organization, the actual underlying structure of the chromosome is an area of ongoing research for which numerous models have been proposed [44–46]. Early studies examined the structure of mitotic chromosomes by electron microscopy following salt treatment to swell chromosomes, or nuclease treatment and removal of histone proteins to examine the shape of chromosomes independently of chromatin [47–50]. Based on these types of studies, researchers proposed the existence of some kind of scaffold or non-histone
protein core from which chromatin extrudes in “radial loops.” This scaffold was thought to contain proteins such as topoisomerase IIα and condensin, both of which localize to the axes of mitotic chromosomes [51–54]. Other models, also based on the results of imaging experiments, have proposed a model of hierarchical folding in which chromatin is folded into increasingly larger structures [55–57]. Additional models, and those that combine different aspects of the two models described here, have also been put forward [44–46].

A recent study applied chromosome conformation capture methods to examine the structure of mitotic chromosomes in more detail. Using 5C and Hi-C to analyze metaphase chromosomes, the authors identified a homogenous folding state that is most consistent with a model of loop extrusion [58]. These results also suggested that the mitotic chromosome is a series of consecutive loops, but that the loops are not always positioned at the same sequence and they are highly disordered, which would explain the microscopic observations that led to the model of hierarchical folding [46,58]. While these findings represent a major step in our understanding of mitotic chromosome organization, there are still open questions remaining, such as how disordered loop formation accounts for the observation that certain regions are more likely to be found at loop bases [59], how these loops relate to the relatively more ordered loops seen in interphase chromosomes [see 46, and below], and whether or not an axis or scaffold is actually part of the condensation process [60,61].

Setting aside the question of the substructure of mitotic chromosomes, the relative ease of visualizing chromosomes at this stage, and of screening for defects in their organization, means we know more about the genes required for mitotic chromosome organization than we do for their interphase counterparts. The proper organization of chromosomes in mitosis is crucial to ensure accurate segregation; during the preceding interphase, each chromosome is replicated to generate a set of sister chromatids, which must be segregated during mitosis. If sisters are not segregated, or chromosomes are not sufficiently condensed before cell division, the chromosomes can be damaged by cytokinesis, resulting in a cell untimely torn (“cut”) phenotype [62,63]. Studies of these and other phenotypes highlighted the role of topoisomerases and SMC family proteins in mitotic chromosome organization. In fact, a recent study showed that addition of topoisomerase II and the SMC complex condensin I, along with histones and histone chaperones, was sufficient to reconstitute mitotic chromosomes in vitro [64]. The major non-histone components of mitotic chromosomes are discussed in more detail below.
Factors that Shape Mitotic Chromosomes

Topoisomerases

Topoisomerases are a group of enzymes that are able to make breaks in DNA and pass strands through these breaks, which allows them to create or remove supercoils (over- or under-winding) in DNA or topological entanglements between strands [65]. Topoisomerases fall into two main families, type I topoisomerases, which cleave one DNA strand at a time, and type II topoisomerases, which make double-stranded DNA breaks [66]. These functions allow them to resolve supercoiling or catenations that form during DNA transcription, replication, or repair, and therefore, topoisomerases are essential for chromosome stability [65]. In particular, topo II in eukaryotes, which is a type IIA topoisomerase, is known to be important for mitotic chromosome organization [reviewed by 45].

Topo II is one of the most abundant proteins present in mitotic chromosomes apart from histones [67], and initial findings suggested that it might be a scaffold protein [51–54]. However, the association of topo II with chromatin is dynamic, suggesting that its role is not purely structural [68–70]. Numerous studies have demonstrated a role for topo II in the segregation of sister chromatids by removing catenations between them [71–76], in chromosome condensation [73,77–81], and in the proper organization of centromeres in mitosis [82]. These different functions may be related; it has been shown that mitotic chromosomes contain DNA entanglements, and topo II could function to remove entanglements between sisters chromatids but enhance entanglements within a chromosome for self-organization [83]. Additionally, mammalian cells have two isoforms of topo II, α and β, and the two proteins have different localizations on mitotic chromosomes [69,84,85], which may result in different functions.

SMC complexes

Another group of proteins initially found as part of a potential scaffold was eventually identified as condensin [86], which is part of the structural maintenance of chromosomes (SMC) family of complexes [reviewed by 87,88]. Members of this family were identified in early screens for genes involved in proper segregation and
condensation of mitotic chromosomes [89,90]. SMC proteins have two long coiled-coil domains with a hinge in the middle and globular domains at the N and C termini containing Walker A and Walker B DNA-binding motifs [reviewed by 88,91]. The proteins are folded at the hinge domain so that the coils associate in an anti-parallel manner and the two globular domains are brought together forming a functional ATPase unit [92–94]. Within an SMC complex, two SMC proteins associate at the hinge domain, with additional proteins associating with the N and C termini [95]. One of these non-SMC binding partners is always a member of another protein family known as the kleisins [96]. The three SMC complexes found in eukaryotes are cohesin, which maintains cohesion between sister chromatids up until mitosis, condensin, which is involved in chromosome condensation, and the Smc5/6 complex, which is involved in DNA repair [reviewed by 88]. These SMC complexes are highly conserved throughout eukaryotes. Notably, prokaryotic cells also have SMC-like complexes which are involved in similar functions [reviewed by 87,97,98], but below we focus on eukaryotic SMC proteins.

Cohesin

The SMC complex which is best understood is the cohesin complex [reviewed by 87,88,99,100]. Cohesin is required for sister chromatid cohesion, the process of keeping the products of DNA replication together until they are separated during cell division [101–103]. In addition to Smc1 and Smc3, cohesin includes a kleisin protein known as Rad21/Scc1 and another protein known as Stromalin/Scc3, and it associates with numerous other regulatory proteins [99]. The association of cohesin with chromatin is regulated in a cell-cycle-dependent manner, starting with the loading of cohesin during the G1/S transition in yeast [102,104] and even earlier in vertebrates [103,105,106]. The establishment of cohesion during S-phase is essential for its proper function in mitosis [107–109]. Once chromosomes have aligned at the metaphase plate, Rad21 is cleaved and cohesin dissociates from the chromatin, allowing sister chromatid separation [110–114]. Consistent with this, loss of cohesin leads to premature sister chromatid separation in mitosis [101–103,115–118]. Structurally, the cohesin complex forms a ring-shape [95,119,120], and artificially sealing the ring by cross-linking prevents cohesin dissociation from DNA [121,122]. Based on these and other data, several models for how cohesin holds sister chromatids together have been proposed [95,120–125]. For example, a single cohesin molecule could encircle two sister chromatids, or cohesin molecules could bind individual chromatids and then self-associate. Importantly, cohesin has also been associated with the
regulation of condensation in budding yeast [101,126–129] and has many interphase roles (discussed more below), and the mode of cohesin binding or function may be different in these different roles.

Condensin

Condensin proteins were first well-described in studies of Xenopus cell extracts, where they were found to be associated with mitotic chromosomes and required for condensation [130,131]. In addition to Smc2 and Smc4, condensin has three other subunits, which vary between different complexes [reviewed by 132]. While yeast only has one condensin complex, higher eukaryotes have two, known as condensin I and II [133,134]. Condensin I consists of a kleisin subunit known as Cap-H/Barren as well as other subunits Cap-G and Cap-D2. In condensin II, the kleisin subunit is known as Cap-H2, and the other subunits are named Cap-G2 and Cap-D3. Additionally, a third condensin complex has been identified in C. elegans, where it is involved in dosage compensation [135,136].

Both in yeast and in other organisms, condensin proteins are known to be involved in the process of compacting chromosomes prior to mitosis [reviewed by 132,137,138]. Additionally, when condensin proteins are depleted, cells are unable to completely segregate their DNA, from which a role in sister chromatid resolution has also been inferred [118,137,139–158]. This role might be explained by the fact that condensin proteins have also been implicated in the recruitment and activation of topoisomerase II [81,140,146,148,157,159,160]. In human cells, the two condensin complexes have different localization patterns; condensin II is present in the nucleus throughout interphase and in mitosis, while condensin I only enters following nuclear envelope breakdown in mitosis [158,161,162], so the two complexes are thought to be important at different stages of the condensation process. Along these lines, in mammalian, Xenopus, and chicken cells, the two complexes have been shown to contribute to the shape of mitotic chromosomes in different ways [133,155,163,164]. However, how condensin actually shapes chromosomes is still not fully understood. Condensin been proposed to mediate linkages in cis by interacting with chromosomes as well as with itself, which would allow compaction [137], and also been proposed to introduce supercoiling into DNA [138].
Smc5/6

The third SMC complex is known as Smc5/6 and contains a dimer of the Smc5 and Smc6 proteins, which have diverged significantly from Smc1 to Smc4 [reviewed by 88,165]. Additionally, this complex includes a kleisin protein, Nse4, and 3 other proteins, Nse1, Nse2 and Nse3. In the yeast Smc5/6 complex, there are two additional subunits, Nse5 and Nse6 [166]. Smc6 and Nse2 are also known as Rad18 and Mms21, respectively, and the complex was first identified as being involved in the response of cells to DNA damage [167–172]. In addition to roles in DNA recombination and repair [166,173–179], the complex is also known to be required for proper DNA replication [179–182], rDNA integrity [183], and telomere maintenance [184,185]. These functions have consequences for the mitotic organization of chromosomes; for example, if repair intermediates are not resolved, they may result in branched structures and other linkages between chromosomes that ultimately interfere with segregation [186,187]. Smc5/6 has been shown to contribute to the proper localization of Topo IIα and condensin [182] and to the SUMOylation of cohesin by Nse2/Mms21 [188], which is required for proper cohesion [189], and may also play a role in the regulation sister chromatid intertwinings [190]. These findings further implicate Smc5/6 in the organization of chromosomes in mitosis.

Kinetochore proteins and the organization of centromeres

In addition to forming compact and discrete structures prior to segregation, chromosomes must also attach to the mitotic spindle, a process that also involves specific organization of chromosomes at the region known as the centromere. Centromeres are defined as regions of chromosomes that give them stability during cell divisions, by providing a site for microtubule attachment [reviewed by 191,192]. Despite the conserved nature of centromere function, organisms vary widely in the size, positioning and composition of their centromeres [193,194]. For example, budding yeast centromeres are defined by a ~125 base pair core region [195–198], while in other eukaryotes the centromere may span several megabases [reviewed by 191,193], and in C. elegans, the chromosomes are described as holocentric, where the whole chromosome can function as a centromere [reviewed by 199]. Further, while in budding yeast, the centromere is defined by DNA sequence, in most other eukaryotes the centromere forms at regions rich in repetitive DNA and is defined epigenetically [200–202]. A major component of this epigenetic
definition and propagation is the presence of CENP-A (CID in Drosophila), a variant of histone H3, which is incorporated into nucleosomes specifically at centromeres [203–209].

During mitosis, centromeres are loaded with a group of proteins that form a specialized structure known as the kinetochore, which provides the interface between chromosomes and the microtubules [210–212]. The dynamics and fidelity of microtubule-kinetochore connections are essential for accurate cell division, and are highly regulated by cells [213–215]. Yeast kinetochores have over 65 different protein components [192,210] while vertebrate cells have over 100 [212], and the loading of many of these is dependent on CENP-A [216,217]. Kinetochore components can be divided into several different groups. Inner kinetochore proteins are closely associated with the DNA, including the constitutive centromere associated network (CCAN) that is associated with the centromere throughout the cell cycle, while outer kinetochore components are more peripheral and play a role in contacting microtubules, including the Ndc80 complex, which is assembled onto the centromeres during mitosis [192,212]. Additionally, the kinetochore includes a whole host of proteins that regulate spindle attachment, such as Aurora B kinase and its group of associated proteins known as the chromosomal passenger complex (CPC) [218–220]. Interestingly, these regulators may play additional roles; Aurora B has been shown to be required for proper condensin localization and chromosome compaction [161,221–227], and condensin has been implicated in the proper function of kinetochores [161,226,228,229], suggesting a link between these two aspects of mitotic chromosome organization.

**Different Layers of Interphase Nuclear Organization, and their Functional Relevance**

The nuclear lamina, nucleolus, and other structures that define nuclear organization

While the organization of chromosomes during cell divisions is crucial for their accurate transmission and genome stability, their organization in interphase has numerous consequences for gene expression, development and other processes. One basis for non-random organization of the interphase genome is the association of DNA with structural components of the nucleus. The nuclear lamina, which is a network consisting mainly of intermediate filament proteins known as lamins formed at the inner nuclear envelope [230], is associated with a transcriptionally silent compartment [231–234]. It has been shown that there are specific domains of DNA that associate with the
nuclear lamina known as LADs (lamin-associated domains), and that these generally show reduced levels of gene expression [235,236]. The A/T-richness of DNA sequences [237], histone modifications [238–240], as well as sequence motifs and transcription factors that bind them [241] may drive the association of specific domains with the lamina to form LADS. Interactions between nuclear envelope components and chromatin or chromatin-bound proteins may also drive LAD formation [242–245]. Certain LADs are cell-type specific [246–248], suggesting they may be involved in the establishment of specific expression patterns.

There is evidence to suggest that LAD organization plays a role in reducing gene expression. For example, loss of lamin caused derepression of testis-specific gene clusters that are associated with the lamina in Drosophila somatic cells [249]. This finding also highlights the role of tissue-specific LAD organization in gene expression during development. Other developmental processes that involve peripheral localization of silenced genes and repositioning of these genes when active include lymphocyte differentiation and VDJ recombination [250,251, see 252 for review], expression of β-globin during erythroid differentiation [253–255], differentiation of embryonic stem cells to neural precursor cells [246,256,257], development of neural progenitors [258], myogenesis [259], and monoallelic gene expression in astrocytes [260]. Peripheral localization also plays a role in gene expression processes that are more dynamic; in synchronized human embryonic stem cells, circadian genes are recruited to the lamina on an oscillating basis every 24 hours, which alternately represses and allows their transcription [261].

What is sometimes unclear is whether peripheral localization causes silencing or is instead a result of silencing – for example, some studies show that transcriptional changes may precede or be independent of nuclear repositioning [253,262], and further, targeting of activator domains to specific loci or induction of reporter genes can lead to their movement towards the center of the nucleus [263–265]. Studies in which genes are artificially tethered to the nuclear envelope to directly test the role of positioning have given mixed results, showing that peripheral localization is sufficient to cause silencing at some loci but not others [266–270]. A recent study examining three loci that are relocalized to the center of the nucleus in mouse embryonic stem cells following transcriptional activation showed that decompaction of chromatin (caused by targeting a synthetic peptide to the three loci) was also sufficient for relocalization [271]. These and other studies [272–274] suggest that chromatin condensation, in addition to transcription, can contribute to nuclear positioning.
While the nuclear periphery is generally a silent compartment, there are certain locations where active genes may be present, and these are the nuclear pores [234,275,276]. The nuclear pore complex (NPC) is a structure composed of nucleoporin proteins or Nups that is embedded in the nuclear membrane and allows transport between the nucleus and the cytoplasm, such as the export of RNAs and the import of proteins required for nuclear function [277,278]. Studies of certain loci have shown that they localize near the nuclear pores when expressed [279–281], consistent with early microscopy studies that found zones of euchromatin near the nuclear periphery [reviewed by 9,282,283]. While genome-wide studies have shown that Nups associate with active regions [284–286], Nups exist dynamically in the nucleoplasm as well as being part of the NPC [287], and many Nup-DNA interactions may be occurring within the nucleus rather than at the pore [288]. Therefore, further investigation is necessary to determine which subset of Nup associations with active genes also involve genome reorganization [276].

Another nuclear substructure that affects organization of DNA is the nucleolus [289], a compartment that forms around the rDNA genes where RNA polymerase I transcription and rRNA processing take place in order to produce ribosomes [290]. Nucleoli are also thought to have roles in cell cycle regulation, stress responses, and disease, as well as being genome organizers [290]. In addition to the rDNA, there are specific DNA sequences associated with the nucleolus known as nucleolus-associated domains or NADs [291,292]. The compartment surrounding the nucleolus is also generally a repressive domain [293,294], and NADs are gene-poor and AT-rich, similarly to LADs. In fact, there is a significant overlap of LADs with NADs, and the observation that some aspects of nuclear positioning are not maintained through mitosis but established in G1 suggests that the association of genomic regions with either the lamina or the nucleolus may reshuffle during cell division [240,295–297]. Localization to the nucleolus or its periphery is associated with specific cellular processes, including X chromosome inactivation in mammals [298], the regulation of imprinted loci [299], and centromere clustering in certain cell types (discussed in more detail below).

The organization of heterochromatin versus euchromatin within the nucleus

Since both the periphery of the nucleus and the nucleolus are regions that have relatively low transcriptional activity, it comes as no surprise that both of these regions are enriched for heterochromatin [294,300].
Heterochromatin was initially observed by Emil Heitz in 1928 as being cytologically distinct from euchromatin and forming a distinct compartment in eukaryotic cells [301–303]. Heterochromatin is more compacted and less “open” than euchromatin, enriched for histone modifications associated with gene silencing such as methylation, depleted for histone modifications associated with active transcription such as acetylation, has a unique signature of proteins which bind it, and often replicates late in S-phase compared to euchromatin [304–310]. Euchromatin, on the other hand, is associated with active transcription and a distinct set of histone modifications [305,311].

Heterochromatin can be further classified into two different types. Constitutive heterochromatin forms at the same regions in all cell types, such as pericentromeric regions and telomeres, and shows enrichment for trimethylation of histone 3 lysine 9 (H3K9me3) as well as proteins such as heterochromatin protein 1 (HP1) [reviewed by 312]. Facultative heterochromatin is similar to constitutive heterochromatin in that it is condensed and transcriptionally silent, but it can be dynamic in different cell types or over the course of the cell cycle and is often associated with distinct proteins and histone modifications, such as Polycomb group proteins and methylation of H3K27 [reviewed by 313]. Both types of heterochromatin are usually found colocalized together, and distributed between the nuclear periphery, the region surrounding the nucleolus, and regions where centromeres are clustered [294,308]. Furthermore, the formation of heterochromatin and the associated histone marks have been shown to be required for perinuclear localization as well as nucleolar integrity [239,240,314,315].

The spatial segregation of heterochromatin in the nucleus has several important roles. Firstly, heterochromatin establishes a compartment within the nucleus where transcription is generally repressed. Observations in Drosophila that chromosomal rearrangements placing genes near centromeric heterochromatin led to variable expression of those genes and variegating phenotypes, also known as position effect variegation or PEV, provided early evidence linking chromatin state to gene expression [316,reviewed by 317]. Such repression is crucial for genome stability, because heterochromatin contains elements such as transposons and endogenous retroviral elements whose replication would be deleterious to the cell [312]. The segregation of heterochromatin also restricts the access of recombination machinery to these sequences, which is important because many of the sequence elements especially in constitutive heterochromatin are highly repetitive, and their recombination can lead to duplication or loss of sequence, genomic rearrangements, and disease [318,319]. In fact, the repair of double-
strand breaks within heterochromatin in Drosophila and in mammalian cells requires special mechanisms by which the break is moved away from the heterochromatic compartment before being repaired [320–324].

In addition to playing a role in the regulation of repetitive elements, the silenced compartment(s) established by heterochromatin are also used by cells in gene regulation. Processes that involve the establishment of facultative heterochromatin and localization to a specific nuclear subcompartment include the inactivation of one copy of the X chromosome in mammalian females [325–329] as well as the silencing of developmental genes until they are needed by Polycomb group (PcG) proteins [306,330,331]. For example, PcG proteins are known to localize to specific foci within the nucleus known as PcG bodies [332–337] in mammalian and in Drosophila cells. Some genes that are regulated by PcG proteins have been shown to localize to PcG bodies when silenced but not when active [338,339]. These processes also often involve specific contacts between different chromosomal regions and sequence elements, which will be discussed further below.

Interestingly, the organization of heterochromatin, and peripheral versus central localization with the nucleus, may have functional impacts beyond gene expression and development. For example, rod photoreceptor nuclei in mice show an “inverted” organization, where euchromatin is at the periphery and pericentric heterochromatin is towards the center of the nucleus [340,341]. When this analysis was extended to other species, it was found that this inverted organization was found mainly in nocturnal species, while diurnal species had the traditional peripheral localization of heterochromatin and central localization of euchromatin [340]. Further analysis and computational modeling showed that nuclei with inverted as versus traditional localization had different optical properties, and since these nuclei are located in the retina, the inverted arrangement may have evolutionary advantages in nocturnal species [340]. This fascinating finding suggests that nuclear organization may have many diverse functions, and that studies of different cell types may reveal new and exciting types of nuclear organization.

The organization of individual chromosomes: CTs, TADs, and sub-TADs

Thus far we have mainly discussed positioning of chromosomes relative to different substructures of the nucleus, but of course, chromosomes also have different positions relative to one another, and are organized and
packaged in particular ways that have consequences for gene expression [reviewed by 342]. In this section, we discuss the way in which chromosomes are organized, both relative to one another and within each chromosome, followed in the next section by specific ways in which chromosomes contact each other, and their functional relevance.

**Chromosome territories (CTs)**

After the first microirradiation experiments suggested that chromosomes occupy distinct volumes within the nucleus [5–8], *in situ* hybridization techniques were developed that allowed the visualization of whole chromosomes. Some of the first experiments to do this examined mouse or hamster cells to which one or a few human chromosomes had been added artificially, followed by labelling with human-specific probes, and found that the human chromosomes occupied their own individual compartments [343,344]. Soon afterwards, the development of probes specific to individual chromosomes, known as chromosome paints, were used to visualize chromosome territories (CTs) in human cells [345–347].

Early studies revealed that gene-rich chromosomes such as human chromosome 19 are internally located, while gene-poor chromosomes such as chromosome 18 are closer to the nuclear periphery [348–351]. These findings were confirmed when visualization techniques advanced to the point of allowing all chromosomes to be examined simultaneously [352]. Further studies have shown that gene density [353–356], transcriptional profile [357], replication timing [358], GC content and the presence of repetitive elements [254] can be correlated with nuclear localization. These observations fit well with what we know about LADs, and suggest that silent domains at the nuclear lamina influence radial positioning within the whole nucleus. This radial arrangement of chromosome domains is evolutionarily conserved among primates, even though different primate species have undergone chromosomal rearrangements relative to each other [359,360], and radial arrangements have also been observed in chicken [361], mouse [355], reptiles and amphibians [362], and bovine cells [363].

As for the positioning of individual chromosomes relative to one another, the patterns are less clear. Chromosomes involved in common translocations in cancer cells are also found closer together in normal cells of
the same tissue [364,365], and studies in mice have shown tissue-specific arrangements of chromosomes [366]. However, while some cell-types may have such specific arrangements and even intermingling between CTs [367], many studies probing organization using ionizing radiation to induce translocations [368] and visualization [350,352,355] have found that the relative positioning of chromosomes, independent of radial positioning, is largely random. As chromosome paint technologies are further developed and used to examine additional cell types and organisms, such as model organisms with fewer chromosomes like Drosophila or C. elegans, it is possible that nonrandom patterns of relative chromosome positioning may be revealed. Additionally, as the ability to visually distinguish the maternal and paternal homolog of each chromosome present in diploid cells has only recently been developed [33], it is possible that adding this information to maps of mammalian chromosome organization may also reveal nonrandom preferences.

The organization of territories has a functional impact because of the relative positioning of genes within them. For example, transcriptionally active genes are often peripherally located within their territory, or even looped outside it [18,272,367,369–376]. Studies have shown that large protein complexes are not excluded from the interior of CTs, suggesting that the peripheral localization of active genes within a territory is not driven by access to transcription factors [377]. Further, localization to the periphery of a CT is not sufficient to cause transcriptional upregulation [378]. Therefore, we still have much to learn about what drives the organization of genes within territories; it is likely that processes such as chromatin condensation also play a role [274]. The organization of DNA within territories, like relative positioning between territories, also impacts the frequency of inter-chromosomal rearrangements, which is increased for regions that loop outside their territories [367,379].

Additional levels of genome organization have been revealed by the development of 3C technology and its high-resolution variants. These methods have confirmed that chromosomes territories exist in human [29], Drosophila [380], mouse [381], and C. elegans nuclei [382], by showing that most interactions occur within chromosomes rather than between them. These methods have also revealed various scales of organization within each chromosome [reviewed by 9]. Firstly, individual chromosome territories can be divided into two compartments, A and B compartments, on the basis of the interaction frequencies within them [29,42,381,383]. These compartments do not generally correspond with linear genome sequence, but regions within a compartment
interact more than regions in different compartments [29]. Compartments can be further divided based on the patterns of interactions within them and their chromatin landscape [42]. B compartments generally have higher interaction frequencies indicating that they are more condensed, while A compartments are more open and more transcriptionally active [29]. This suggests that each chromosome may segregate its active regions from its inactive ones.

*Topologically associating domains (TADs)*

Chromosomes can further be subdivided into units that interact more strongly with themselves than with neighboring regions [380,383–385]. Known as topologically associating domains, or TADs, these regions range from about 200 kilobases to one megabase in size, and at least for some TADs, visual inspection by FISH has confirmed that they form discrete structures [384]. A recent study showed that TADs in interphase cells correspond to cytological bands observed on Drosophila polytene chromosomes, demonstrating that TADs demarcate regions with distinct structures [386]. TADs also correlate well with the binding of specific proteins and/or histone modifications [42,380,383–385]. TAD boundaries, which are defined as being flanked by regions that do not interact with each other, are themselves special regions, being bound by numerous architectural proteins, containing sites of active transcription, and being enriched for housekeeping genes [42,380,383–385,387]. TADs may also be divided into smaller units known as sub-TADs or contact domains [42,387]. It is likely that as the resolution of Hi-C techniques increases we will learn about further layers of genome folding.

The organization of chromosomes into TADs has numerous functional consequences. Firstly, co-regulated genes are often found within TADs [384]. Secondly, long-range enhancer action is usually restricted to within TADs [388–390], and loops associated with gene regulation often have at least one end at a TAD boundary [42]. Thirdly, arrangement of TADs may be linked to specific gene expression programs. For example, in Drosophila cells, TADs reorganize after heat shock, when a specific transcriptional program is activated and the majority of genes are silenced [391]. Additionally, in *C. elegans* hermaphrodites and in mammalian females, where dosage compensation occurs, the X chromosome(s) being down-regulated have distinct TAD architectures [382,384], and the organization within TADs changes during embryonic stem cell differentiation [43]. Artificial disruption of TADs has been shown
to lead to novel enhancer-promoter connections and changes in gene expression [392]. Therefore TADs organization is likely somewhat related to transcription. However, the observations that TAD boundaries are conserved between cell types regardless of expression patterns [43,390], and that repositioning of sub-TADs does not necessarily cause transcriptional changes and vice versa [393], suggest that at least some aspects of TAD organization are independent of gene expression.

**Heterologous interactions: gene regulation in cis and in trans**

While chromosomes are folded into TADs and higher-order territories, individual loci can still be somewhat dynamic in their position [394,395]. Furthermore, there can be specific contacts between different sequences with consequences for gene expression [reviewed by 342,396,397]. These contacts can be in *cis*, when they occur within a given chromosome, including the folding that occurs to form TADs as well as regulatory contacts. Alternatively, contacts can form in *trans*, between different chromosomes; *trans* contacts can further be divided into heterologous contacts between unrelated chromosomes, or homologous contacts, which take place between the maternal and paternal homologs of each chromosome present in diploid cells. Here, we first discuss contacts between different sequence elements, such as contacts between enhancers and promoters, both in *cis* and in *trans* between heterologous chromosomes, followed by a discussion of homologous *trans* interactions and their consequences.

The formation of contacts between different chromosomal regions, particularly those that occur at long-range and between different chromosomes, has been linked to transcriptional regulation [398–402]. For example, studies of TADs and genome folding in Drosophila cells found not only that transcriptionally active and inactive domains form separate clusters, consistent with the formation of A and B compartments, but also, that domains of active transcription associate with active domains on different chromosomes [380]. This is consistent with the observation that when intermingling between or looping out from chromosome territories occurs, it generally involves sites of active transcription [272,367,369–376].
Such findings support the hypothesis that there are clusters of active transcription in the nucleus known as transcription factories, an idea originally based on the observation of discrete foci of nascent transcripts and RNA polymerase II in the nucleus [403–405], and supported by findings that transcription factors and active genes co-localize in small clusters distributed throughout the nucleus [406–410]. However, other studies, including one using super-resolution microscopy [411], have disputed the idea that transcription factories exist, and the idea is still under debate [see 412 for review]. For instance, while co-regulated erythroid genes have been suggested to cluster together in a transcription factory [409], others have suggested that, rather than clustering together, the five genes from four different chromosomes all co-localize with nuclear speckles [413]. Nuclear speckles are foci containing the spliceosome assembly factor SC35 [414,415], and clustering here results in a higher chance that genes will be found near each other, but does not mean co-localization is required for transcription [reviewed by 416].

In addition to the co-localization of active domains, there can be chromosomal contacts between specific sequences, and some of these are well-characterized. Here we discuss the co-localization of specific sequences in the regulation of gene expression. Gene expression by RNA polymerase II in eukaryotes usually involves two main types of regulatory sequences, gene proximal elements known as promoters, and more distal elements which may be enhancers, insulators, or other types of regulatory sequence [reviewed in 417]. Regulation of transcription depends on some kind of communication between these elements, and various models of looping, including those by which enhancers and promoters are brought into close physical proximity, have been proposed [reviewed by 396,417,418].

Looping is known to be involved in cis-activation at several loci [419], the first identified example being the mammalian β-globin locus. This locus contains several β-globin genes that are activated sequentially in development, and an upstream regulatory element known as the locus control region or LCR 40-60 kb away [reviewed by 342,420]. Initial studies of this locus using 3C and visualization of transcripts using a technique called RNA-TRAP showed that the LCR interacts with the active globin gene, while the intervening DNA, including inactive globin genes, is looped out [421,422]. Furthermore, the LCR changes its looping pattern during erythroid development, interacting with fetal globin genes in primitive erythroid cells and with adult globin genes in more differentiated cells [423]. This process is known to require multiple transcription factors [424,425] as well as a nuclear factor known as NL1/Ldb1 [426], and artificially targeting this Ldb1 protein to the β-globin promoter is
sufficient to induce looping and activation in cell types where the gene is not normally expressed [427,428]. These experiments demonstrate the importance of looping for transcription at this locus.

There are also examples where interactions between regulatory sequences occur in trans, that is between multiple chromosomes [reviewed by 429]. In olfactory sensory neurons in mammals, each neuron activates one out of around 2800 alleles to express a single olfactory receptor. The other alleles are silenced via localization to heterochromatin, and if nuclear organization is altered by overexpression of a lamin protein, multiple olfactory receptor genes start to be expressed [430]. The selection of the active allele may also be related to nuclear organization; there are at least 12 different enhancers that regulate these process, and multiple enhancers from different chromosomes co-localize at the active gene [431]. Trans interactions have also been observed between the H19 imprinted locus and imprinted regions on other chromosomes [432], between alternatively expressed genes such as interferon-γ and interleukin-4 in helper T-cells [399], as well as in the regulation of human anti-viral factors, estrogen-responsive genes, and developmentally-regulated genes [reviewed by 429, see also 433,434].

It is important to note that not all looping events result in gene activation. The Hox clusters are groups of developmental genes that are kept silenced by Polycomb group proteins so that they are expressed only in the appropriate tissues [435]. In Drosophila, the Hox genes are found in two clusters, Antennapedia and bithorax, each of which contains multiple Hox genes as well as Polycomb response elements (PREs), which are the regulatory elements involved in Polycomb-mediated repression. One such element is Fab-7, which has been shown to interact with Hox loci even when inserted as a transgene on different chromosomes, indicating that it can act in trans to facilitate silencing [436]. Additionally, it has been shown that the native bithorax locus contains looped structures, and these structures are changed when the genes are activated [338]. Finally, there is communication between the two clusters, Antennapedia and bithorax, which are separated by more than 10 megabases on the chromosome arm; genes from different clusters co-localize in cells in which they are both repressed, but rarely co-localize in cells in which one of the genes is activated [339]. The co-localization often occurs within a Polycomb body and depends on Polycomb group proteins [339]. This indicates that loops mediated by PcG proteins may play a functional role in repression.
Heterologous interactions: centromere and telomere clustering

Heterologous interactions between chromosomes may also be based on genomic features rather than being sequence-specific; for example, some cell types show clustering of centromeres and/or telomeres in interphase. This type of arrangement was first proposed in 1885 by Carl Rabl, based on studies of the interphase nuclei of *Salamandra* and *Proteus*. Rabl suggested that chromosomes keep the orientation established during anaphase, when centromeres cluster as they are pulled towards the spindle pole while telomeres lag behind and gather at the other end of the nucleus, thus establishing a polarity in the organization of the nucleus in the subsequent interphase [437]. Known as the Rabl configuration, such an arrangement of centromeres and telomeres has been noted in multiple cell organisms, including fission yeast [438,439], flies [440–448], and several plant species [449–453]. Importantly, centromeres and telomeres can be clustered without showing the polarized Rabl arrangement; mammalian cells often have multiple clusters distributed around the nucleus each containing 2-3 centromeres, known as chromocenters, and similar clusters of telomeres [454–463]. Centromere clusters are heterochromatic and may form at the nuclear periphery, or in the nuclear interior, associated with the nucleolus [354,464,461,465–469]. Interestingly, the extent of centromere clustering and its position within the nucleus can vary in different tissue types in the same organism, and over the course of the cell cycle [297,354,355,447,455–457,459,460,467,469–475].

The clustering of centromeres and telomeres has multiple functions. For example, telomeres cluster during meiotic prophase in several organisms, including yeast, insects, mice, zebrafish, and plants [reviewed by 476,477]. Known as “bouquet formation,” this clustering and subsequent chromosome movements facilitate homologous chromosome pairing, recombination, and the progression of meiosis. Centromere clustering has also been suggested to contribute to these processes [reviewed by 478,479]. In addition to these roles in establishing nuclear organization prior to divisions, since centromeres are enriched for heterochromatic sequences, the localization of genes near chromocenters is also associated with transcriptional silencing. Such patterns have been observed in mouse development [480–483], at the human β-globin locus [484], and in regulation of virulence genes in the eukaryotic parasite *Plasmodium falciparum* [485–487]. A direct connection between centromere clustering and heterochromatic silencing has been shown in Drosophila, where knocking down NLP, a protein required for centromere clustering at
the nucleolus, also caused derepression of transposable elements and DNA damage [468]. Therefore, centromere clustering plays a key role in the organization of heterochromatin, gene expression and genome integrity.

Homolog pairing and homology effects in Drosophila

The interactions described above take place between different sequence elements, whether on the same chromosome or on different chromosomes. However, homologous sequences can also interact with each other, and the interactions between maternal and paternal homologs of chromosomes are known as pairing. While the most robust pairing that occurs in many organisms is during meiosis, in preparation for segregation of homologs, pairing can also occur in somatic cells and in interphase, and these are the type of pairing interactions discussed here. Somatic pairing may take place between specific loci, or even between entire chromosomes.

In this respect, Drosophila and other dipteran insects have a remarkable form of interphase nuclear organization, where homologs are paired along their entire length [reviewed by 488,489]. Somatic pairing was first noted by Nettie Stevens and Charles Metz in the early 20th century, who found that in metaphase spreads, chromosomes of similar size were often found near each other, from which they inferred that these chromosomes were also paired in interphase [490,491]. Pairing is also observed in polytene chromosomes such as in the salivary glands of Drosophila third instar larvae, in which the genome replicates as many as ten times without cell division, resulting in more than a thousand copies of each chromosome which are held together in close alignment [492,493]. This alignment allows for the observation of bands and inter-bands, alternating regions of light and dark staining DNA, which, as mentioned above, have been shown by modern techniques to correspond to TADs and TAD boundaries, respectively [386]. The banding patterns of polytene chromosomes proved very useful for early cytogeneticists using Drosophila as a model system to study chromosome biology.

Development of in situ hybridization techniques allowed the demonstration that homologous regions of chromosomes are indeed paired in interphase nuclei in non-polytene cell types [494]. This pairing is established in early embryogenesis and occurs in most tissues of the fly [447,494–498], though notably not in germline progenitor cells, which suggests that somatic pairing is different from the meiotic pairing established in the germline [499].
Pairing of chromosomes in somatic cells has effects on gene expression, including transvection [reviewed by 500,501,502]. As first defined by Ed Lewis in 1954, the term transvection describes cases where gene expression is influenced by the close proximity of homologs [503]. This might occur through intra-allelic complementation; for example, when one allele is lacking a functional promoter, and another allele is lacking a functional enhancer, neither allele would result in transcription when homozygous. However, when both alleles are present in heterozygotes and positioned in close proximity, the enhancer of one allele can act on the promoter of the other allele, resulting in gene expression. This type of intra-allelic complementation occurs, for example, at the yellow locus in *Drosophila melanogaster* [504–507]. Studies of transvection in *Drosophila* were some of the first to demonstrate that enhancer action could occur in *trans* as well as in *cis*.

A prediction made by this model is that chromosomal rearrangements that disrupt pairing at a given locus should also disrupt transvection, whereas compensatory rearrangements that restore pairing should also restore expression [503]. Using this metric, several loci in *Drosophila* have been found to show transvection; in addition to *yellow*, and the Ultrabithorax locus that was studied by Lewis [503], transvection has also been observed at decapentaplegic [508], eyes absent [509], and Abdominal-B [510], among others [see 507 and references therein]. The combination of alleles required to observe classic transvection means that it has only been well-described at certain loci, however, recent observations suggest that transvection may occur widely in the *Drosophila* genome. For example, inserting complementary alleles of *yellow* at different sites in the *Drosophila* genome resulted in transvection at all eight positions tested [511], and studies of a reporter gene system showed that a single enhancer could act in *cis* as well as in *trans*, indicating that the occurrence of *cis*-regulation at a locus does not rule out transvection, though the two may be in competition [512]. Studies of transgenes inserted at common integration sites in *Drosophila* neuroblasts showed that transvection occurred between transgenes inserted at the same site, further demonstrating that transvection can occur at many different sites in the fly genome, involve diverse sequences, and occur in different tissue types [513]. Studies of malic enzyme expression in flies, which allow transvection at the malic enzyme locus to be quantitatively examined, show that the amount of expression, not just its occurrence, can depend on pairing [514 and references therein]. These studies highlight that there may be many new examples and types of transvection remaining to be discovered.
It is important to note that not all *trans* effects result in gene activation. A fascinating example is that of *brown* dominant in Drosophila [515–526]. Brown is a gene involved in the synthesis of eye pigment in the fly, and null alleles of *brown* give a recessive brown eye phenotype. Alleles of *brown* have been identified that show position effect variegation as a result of translocations or inversions that place *brown* near heterochromatic sequences. Interestingly, these alleles are dominant, because due to homolog pairing, the other allele will also be re-localized to heterochromatin and silenced [517,520,521]. Another example of pairing being involved in gene repression is pairing-sensitive silencing, or PSS [527,528]. This phenomenon has been described for certain regulatory elements in flies, especially PREs, which silence reporter constructs more efficiently in flies that are homozygous for the construct compared to those that are heterozygous, a phenotype that is sensitive to the position of these insertions relative to each other [reviewed by 529]. Therefore, it is possible that homologous pairing, like heterologous clustering of silenced genes, may facilitate Polycomb-mediated repression.

**Examples of homology effects outside Drosophila**

While the most extensive homolog pairing outside of meiosis, and most examples of homology effects, have been observed in Drosophila, there have been reports of transvection-like phenomena in other organisms [see 489,500,501 and the references therein]. Other phenomena that may involve pairing or at least some form of communication between homologous sequences include repeat-induced point mutation (RIP) in *Neurospora* [530–532] and paramutation in maize [533–535]. While mammalian cells do not show robust association between homologs [536], local pairing has been observed during several processes including the establishment of X-inactivation, when the two X chromosomes transiently pair, possibly facilitating the establishment of asymmetric expression patterns [537–540]. Pairing has also been observed at immunoglobulin loci in B-cell development, where only one allele undergoes recombination [252,541,542], and at imprinted loci [543–545], though one study in mice showed that while pairing does occur at imprinted regions, it is not necessarily related to imprinting or transcriptional status, and might be a more widespread feature of the mammalian genome [546].

Consistent with the idea of a broader function for pairing, pairing has also been observed at the Oct4 locus during embryonic development, where it is associated with the establishment of repressive histone modifications at
Oct4 and differentiation [547], and at sites of double-strand break repair in response to ionizing radiation [548,549]. Additionally, pairing has been associated with disease in the case of renal oncocytomas in humans, where, remarkably, the entire q-arm of Chromosome 19 is paired while the p-arm is unpaired, and there are changes in gene expression on Chromosome 19 [550]. It is likely that as nuclear organization is probed in greater detail in a growing number of cell types that we will find more examples of pairing, adding to the growing understanding that homology effects are not simply an unusual feature of Dipteran insects.

Factors that Organize the Interphase Nucleus

The many layers of organization described above require careful coordination, and in this section, we describe some of the factors involved. As alluded to above, many specific interactions involved in gene expression have been found to require transcription factors and other context-specific regulators [reviewed by 342]; similarly, lamins and other nuclear envelope proteins are required for the proper organization of LADs and domains near the nuclear pores [241,243,244,247,249,286,551,552]. Here we focus on regulators that shape the nucleus more broadly, by facilitating long-range interactions within and between chromosomes.

Architectural proteins: CTCF and other genome organizers

One class of proteins important for long-range chromatin interactions are architectural proteins, or insulators [553–555]. These proteins were initially identified as inhibiting enhancer-promoter interactions; binding of architectural proteins in the region in between a transgene and regulatory elements insulates the transgene from the effects of these elements, and from phenomena such as position effect variegation [556–558]. Well before the discovery of TADs, these proteins were thought to be important in dividing the genome into active and inactive domains. Now it is known that, in addition to having insulating activity by blocking enhancer action and preventing the spread of histone modifications [555,559], architectural proteins can also facilitate loop formation [387,553–555], and their binding sites, which are enriched at TAD boundaries [380,383–385,560], are often involved in loops. Therefore, architectural proteins play a context-dependent role in genome organization. These proteins have been particularly well-characterized in Drosophila, which have 11 different architectural proteins that bind specific DNA
binding sequences including CCCTC-binding factor (CTCF), Suppressor of Hairy-wing (Su(Hw)), transcription factor III-C (TFIIC), boundary element associated factor 32 (BEAF-32), and several others [554]. These often co-occur with other proteins that bind DNA in a non-sequence-specific manner, including cohesin and CP190, and these are also considered architectural proteins [554]. Fewer architectural proteins are known in mammalian cells, but the roles of CTCF and cohesin are conserved, and these have been particularly well studied.

CTCF is highly conserved protein with 11 zinc finger DNA binding domains which binds a wide range of sequences containing an 11-15 base pair consensus sequence [reviewed by 9,561]. Additionally, CTCF proteins oligomerize [reviewed by 562], offering a model for how they facilitate looping. CTCF is also important for the organization of TADs and overall genome structure [563,564], as well as enhancer-promoter looping and other aspects of gene regulation at specific loci [reviewed by 9,553–555,561]. Recent assays of CTCF binding by chromatin immunoprecipitation experiments estimate that the mammalian genome has 55,000-66,800 CTCF binding sites [565,566], and while these are enriched at TAD boundaries, the majority of them are actually found within TADs [reviewed by 561]. It is possible that CTCF plays different roles at TAD boundaries compared to within TADs; in Drosophila, CTCF sites at TAD boundaries are co-occupied by many other architectural proteins, while within TADs architectural protein binding sites are found at low density [560]. It is thought that CTCF sites at boundaries isolate neighboring domains while sites within TADs facilitate intra-TAD gene regulatory interactions [554,555]. Consistent with this idea, in human cells, genes that are flanked by CTCF sites are more likely to be co-regulated than genes that are separated by CTCF sites [567,568].

Roles for cohesin in the interphase nucleus

Interestingly, in mammals, CTCF often is found co-localized with the SMC complex cohesin [569–572]. Cohesin was first suggested to have roles beyond the cohesion of sister chromatids when it was found that the cohesin loading factor Nipped-B facilitates distal activation of homeobox genes in Drosophila [573]. Nipped-B is conserved in humans (NIPBL) and mutations in this gene and other cohesin subunits are associated with a disease known as Cornelia de Lange syndrome [574–577], which is thought to be caused by transcriptional misregulation rather than cohesion defects [578]. Cohesin is now known to play a role in gene regulation at many loci in
Drosophila (where it does not co-occur with CTCF [579]) as well as in mammals, potentially by facilitating enhancer-promoter looping as well as by regulating RNA polymerase II pausing and interacting with Polycomb Group proteins [580–582].

In mammalian cells cohesin may play different roles at sites that are co-occupied by CTCF as versus sites where it binds alone. Cohesin binding sites without CTCF are enriched at transcription start sites and may be involved in tissue-specific enhancer-promoter looping, while cohesin binding sites with CTCF are enriched at TAD boundaries and may mediate constitutive long-range interactions [reviewed by 583]. Evidence for the different impacts of cohesin and CTCF on genome architecture comes from the fact that cohesin depletion decreases intra-TAD interactions but does not alter TAD boundaries [563,584,585], while CTCF depletion also increases inter-TAD interactions, changing their relative organization [563]. Another factor that interacts with cohesin and CTCF in different combinations is the transcriptional regulator Mediator [387,586], which also participates in looping and genome organization.

In addition to roles in gene regulation, cohesin also plays a role in repair of DNA double-strand breaks [587,588, reviewed by 589]. Of course, in G2, these functions in gene expression and repair must be balanced with the important role of keeping sister chromatids together, and it has been hypothesized that different pools of cohesin may be involved in these different functions [88,590,591]. For example, live-imaging in human cells has shown that cohesin is more stably bound to chromatin in G2 while being more dynamic in G1 [590], and decreasing the amount of cohesin in yeast cells affects condensation, rDNA stability and repair before cohesion is disrupted [591]. It will be of great interest going forward to determine how these different functions of cohesin are regulated and how the stable connection required for cohesion is balanced with the potentially more dynamic role of cohesin in other aspects of chromosome organization.

Roles for condensin in the interphase nucleus

Like cohesin, condensin plays roles outside the organization of mitotic chromosomes. As mentioned above, *C. elegans* have a third condensin complex that is similar to condensin I but contains a different protein, DPY-27, in
the place of Smc4 [135,136]. This complex is known as the dosage compensation complex (DCC) and plays a role in down-regulating each of the two X chromosomes in *C. elegans* hermaphrodites to achieve an equal level of expression with males, which have one X chromosome [reviewed by 592]. Observations consistent with a role for condensin complexes in transcriptional silencing have also been made in organisms that do not have a condensin-related DCC. For example, condensin I in Drosophila has been observed to contribute to Polycomb-mediated silencing [593], condensin II in mice T-cells maintains gene silencing and a quiescent state [594], and yeast condensin mutants have defects in silencing of mating type loci [140] and telomeres, while silencing at the rDNA is enhanced [595]. These gene silencing roles could reflect a novel function of condensin; alternatively, they could be related to the role of condensin in compaction, for example, compact chromatin may be less accessible to transcription factors.

Consistent with an architectural role for condensin proteins outside mitosis, they have been shown to have multiple roles in interphase nuclear organization. In budding yeast, tDNA genes from the 16 different chromosomes cluster with RNA polymerase III at the nucleolus, where tRNA transcription occurs, and condensin is required for this clustering [596–598]. In mammalian cells, it has been shown that condensin II begins the process of resolving sister chromatids as soon as they are replicated [145], which could have functional consequences for nuclear organization in G2. A specific role for condensin proteins in nuclear organization has been particularly well-described in Drosophila. Early studies showed that condensin components regulate PEV, silencing that is associated with proximity to heterochromatin [150,599], which could reflect a role in chromosome organization as well as in transcriptional silencing. Subsequently, it was shown that condensin II antagonizes the pairing of homologous chromosomes in Drosophila, both in polytene chromosomes and in cycling cells as assayed by FISH [600,601]. Consistent with this role, condensin II over-expression disrupts transvection [600], and several condensin regulators have been identified which also regulate pairing levels [601–605]. Condensin II also promotes the formation of chromosome territories [606]. Therefore, condensin II plays a key role in nuclear organization in Drosophila, and given that condensin is so well-conserved, potentially could do so in other organisms as well.
Long non-coding RNAs in nuclear organization

In addition to proteins, long noncoding RNAs (lncRNAs) have also been shown to contribute to higher-order genome organization. Several lncRNAs have been proposed to play organizational roles by providing scaffolding for nuclear bodies, including nuclear speckles and paraspeckles, which are regions within the nucleus where specific proteins co-localize [reviewed by 607]. In these cases, the lncRNAs localize at the sites of their transcription and recruit proteins to them, allowing various functions such as RNA processing or sequestration [607].

There are also lncRNAs that regulate long-range inter-chromosomal interactions or the organization of whole chromosomes. For example, the lncRNA Xist, which spreads across one of the two X chromosomes in mammalian females in an essential step of X-inactivation, has been shown to alter the topological organization of the X chromosome [608–610]. As for inter-chromosomal interactions, a recently identified long intergenetic noncoding RNA (lincRNA) known as the functional intergenic repeating RNA element, or Firre, localizes at a 5 Mb domain around its transcription site on the X chromosome but also recruits at least five other loci on different chromosomes [611]. Firre has also been shown to cause the inactive X to localize near the nucleolus [612]. These results demonstrate that lncRNAs can play a role in nuclear positioning and inter-chromosomal interactions.

Interestingly, RNAs have also been proposed to play a role in homolog pairing. How homologous sequences recognize each other in the nucleus, whether in Drosophila somatic pairing, in the more localized examples of pairing seen in mammals, or in the search for repair templates during homologous recombination, is not fully understood. Since RNAs are reporters of DNA sequence, it is possible that they might be the molecules through which homology is recognized. In fact, RNA may mediate pairing in S. pombe meiosis, where, as described above, telomere clustering gives way to homologous pairing. One study found that transcripts from a specific locus were retained on the chromosome and facilitated pairing at that locus, while translocation of this locus to other sites caused ectopic pairing [613]. It will be of great interest to determine if chromosomally-retained RNAs play a role in pairing in other organisms.
Conclusions, and Aims of this Dissertation

In this chapter, I have presented a review describing techniques for studying nuclear organization, the organization of chromosomes both in mitosis and in interphase, the functional impacts of this organization on gene expression and other processes, and several factors known to be involved in determining these different types of organization. To conclude, I give a brief overview of the questions that motivated the work described in this dissertation, and explain our choice of experimental system.

Open questions on the roles of mitotic proteins in interphase nuclear organization

As this review has highlighted, there are several proteins that play key roles in the organization of chromosomes both in interphase and in mitosis, including cohesin and condensin. For example, cohesin must maintain a unique inter-chromosomal connection between sister chromatids, but is also involved in enhancer-promoter looping and other types of intra-chromosomal contacts in interphase. Cohesion between sister chromatids has mostly been studied in the context of mitosis, but is also a prominent feature of interphase nuclear organization in G2, when sister chromatids are kept in close alignment. In Drosophila, the situation is even more complicated, because in addition to the interactions between sister chromatids, homologous chromosomes are paired. Here I investigate the role of cohesin prior to mitosis in Drosophila, examining the organization of sister chromatids in interphase. These studies are presented in Chapter 2.

I also explore the relationship of sister chromatid cohesion to the pairing of homologs, both of which are interactions occurring between chromosomes that share homology. As described above, condensin proteins are known to regulate both somatic homolog pairing in Drosophila as well as sister chromatid segregation in many different organisms. Therefore, cohesin and condensin have functional interactions in shaping chromosomes. I have examined these interactions in Drosophila cells both in interphase and in mitosis. In particular, I have investigated whether mechanisms of homolog pairing contribute to the organization of sister chromatids in interphase, and these studies are the focus of Chapter 3.
My studies of cohesin and condensin reveal how mitotic proteins may have roles throughout the cell cycle. This is not the only way in which mitotic regulators may establish interphase nuclear organization. For example, the Rabl configuration seen in interphase nuclei is thought to be a consequence of the way in which chromosomes are pulled during segregation in anaphase and telophase. Interestingly, the extent of centromere clustering can vary between different organisms, in different cell types of the same organism, and even over the course of the cell cycle, suggesting that pathways exist to regulate clustering beyond establishment in mitosis. While a few factors required for centromere clustering are known, we sought to identify additional regulatory factors, expanding our knowledge of how clustering is both established and maintained. These efforts are the focus of Chapter 4.

**Studying homologous and heterologous interactions by FISH**

To answer these questions, we have used visual techniques to examine nuclear organization. While Hi-C based methods have developed to the point of distinguishing homologs based on SNP differences, sister chromatids are identical by sequence and their relative positioning cannot be examined. FISH does not necessarily distinguish sisters or homologs (depending on the FISH technology used), but the number of signals and their relative positioning can be used to examine the interactions between these different chromosomes. Additionally, FISH allows us to study the co-occurrence of multiple events in the same cell, such as the clustering of multiple nonhomologous centromeres. Rather than average interaction frequencies obtained over a population of cells, we can examine combinations of events taking place in the same cell. Finally, FISH is amenable to high-throughput screening, which facilitates the identification of new genes involved in nuclear organization.

**Drosophila as a model system for studying nuclear organization**

In the work presented here, we focus on the fruit fly *Drosophila melanogaster* as a model system for studying nuclear organization. Historically, the fruit fly has been an important and widely-used system in this field for several reasons, including the ability to examine chromosome structure in polytene tissues and the existence of visible phenotypes related to chromosome organization, such as position effect variegation and transvection. Additionally, Drosophila have fewer chromosomes than mammalian cells, four diploid pairs in *Drosophila*
melanogaster, providing a relatively simple system for defining inter-chromosomal associations. Finally, interesting aspects of Drosophila nuclear organization, including both homolog pairing [601,614] and centromere clustering [468], can be observed in cell culture, facilitating the study of genes involved in these processes in addition to allowing screening for new ones. A screen conducted in our laboratory for factors involved in pairing identified genes that are conserved from Drosophila to mammals, suggesting that while Drosophila are extreme in the extent to which they pair their chromosomes, the underlying mechanisms may be conserved in other organisms, where they perhaps act in a more localized or temporally restricted fashion [601]. Indeed, many features of nuclear organization, including Polycomb-mediated repression and the role of cohesin in gene expression, were initially studied in Drosophila but are conserved in mammals. Therefore, we view the Drosophila system as an ideal one for continued studies of the mechanisms driving nuclear organization.
References

3. Boveri T. Ergebnisse über die Konstitution der chromatischen Substanz des Zelkerns. Fish Jena. 1904;


51


59


Chapter 2

The pairing of homologs and sister chromatids in interphase nuclei of *Drosophila melanogaster* is cohesin-independent

This chapter is based on:


Author contributions: TNS performed experiments and carried out data analysis. TNS and EFJ provided reagents. TNS, EFJ and C-TW conceived and designed the experiments. TNS and C-TW wrote the paper.
Abstract

Following DNA replication, sister chromatids must stay connected for the remainder of the cell cycle in order to ensure accurate segregation in the subsequent cell division. This important function involves an evolutionarily conserved protein complex known as cohesin; any loss of cohesin causes premature sister chromatid separation in mitosis. We examined the role of cohesin prior to mitosis, using fluorescence *in situ* hybridization (FISH) to assay the association of sister chromatids in interphase Drosophila cells. Because maternal and paternal homologs are paired in Drosophila cells, these assays allowed us to also investigate the role of cohesin in homolog pairing. Surprisingly, we found that cohesin is not required for homolog pairing in interphase nuclei, and further, that sister chromatids remain closely aligned in G2 with little to no cohesin. We present data showing that cohesin-independent pairing is widespread in Drosophila, unlike in other systems where it has only been observed at specific chromosomal regions, and further, that the cohesin-independent pairing of sisters observed in interphase is not dependent on the presence of a homolog. These results suggest that there may be mechanisms contributing to the organization of sister chromatids in interphase independently of cohesin proteins.

Introduction

Nuclear organization is a fundamental aspect of genome regulation and function [1–4]. In addition to cell-type specific interactions between and within chromosomes [5,6], this organization is shaped by chromosome-wide changes in structure that are inherent to the process of nuclear division. For instance, in addition to condensing their chromosomes into the compacted forms found in metaphase, mitotically dividing cells double their DNA content and thus their chromosome number during S-phase. Diploid cells therefore transition from a G1 phase with two copies of each chromosome, called the maternal and paternal homologs, to a G2 phase with four copies of each chromosome, each homolog having been replicated to form a set of sister chromatids. Importantly, sister chromatids are held together by physical connections that are critical for ensuring that the two chromatids ultimately segregate between daughter cells in mitosis [7–9]. Remarkably, these connections, known as cohesion, exist in interphase amidst a variety of other inter- and intra-chromosomal interactions, and yet are uniquely maintained between sisters.
Cohesion between sister chromatids is known to require the highly conserved SMC complex known as cohesin [10–12]. Cohesin consists of four subunits known as Rad21, Smc1, Smc3 and SA, which come together to form a ring-shaped protein complex [13]. Cohesin is loaded onto chromatin before or during S-phase (depending on the organism), and the presence of cohesin during replication is essential for proper sister cohesion at later stages in the cell cycle, suggesting that S-phase is the stage when establishment of cohesion occurs [12,14–16]. Ultimately, the Rad21 subunit of cohesin is cleaved at anaphase causing cohesin dissociation, a step that is both necessary and sufficient for sister chromatid separation during mitosis [17–22]. These and other results have led to multiple models of how cohesin functions, including the “embrace” model, wherein a single cohesin protein encircles two sister chromatids from the time of replication until mitosis (Figure 2.1), or “snap” models, in which cohesin molecules bind individual chromatids and then self-associate [13,22–26]. Regardless of how cohesin associates with DNA, the temporal pattern of its association, and the requirement for cohesion establishment in S-phase, suggest that cohesin is involved in maintaining cohesion from the time of DNA replication, through G2, and until segregation in mitosis. Consistent with this idea, observations in many different organisms demonstrate that loss of cohesin results in premature sister chromatid separation in metaphase [10–12,27–30]. While most studies of cohesin function in sister chromatid organization have focused on mitosis, studies in human and in chicken cells that have examined cohesion in S/G2 using fluorescent in situ hybridization (FISH) have found that loss of cohesin results in phenotypes consistent with sister chromatid separation in interphase, including an increase in the number of FISH signals, increased distance between signals, or abnormally shaped signals (Figure 2.2) [27,31–35]. These results further highlight the importance of cohesin at multiple stages in the cell cycle.
Figure 2.1 Model showing how the yeast cohesin complex may function in sister chromatid cohesion. Figure reproduced with permission from Haering et al., 2002 [23]. (A) Cohesin components form a ring shape. A 10 nm chromatin fiber and a DNA double helix are shown on the right for scale comparison. (B) Representation of the “embrace” model.
Figure 2.2 Examples of FISH studies that observed disruption of cohesion in interphase. Figures reproduced with permission from: (A) Sonoda et al., 2001 [27], (B) Schmitz et al., 2007 [32], (C) Nishiyama et al., 2010 [33], (D) Toyoda and Yanagida, 2006 [31], (E) Manning et al., 2014 [34], (F) Dupont et al., 2014 (content available under a Creative Commons Attribution (CC-BY) license) [35]. These studies found an increase in the number of FISH signals, increased distance between signals, or abnormally shaped signals in G2 or S-phase when cohesin or cohesin regulators were knocked down or mutated. Sccl is Rad21; Sororin (Sor) and pRB regulate the association of cohesin with DNA; Roberts Syndrome is caused by mutations in ESCO2, a factor required for the establishment of cohesion.
The nuclei of Drosophila melanogaster and other Dipteran insects present a unique situation, where sister chromatid cohesion is not the only association between chromosomes that are homologous to each other. In these insects, maternal and paternal homologs show robust pairing in somatic cells, occurring in different stages of the cell cycle and in many different tissues [36–41]. Additionally, this pairing impacts gene expression through mechanisms known as transvection [42–50]. Therefore, homolog pairing is a prominent and functional feature of Dipteran nuclear organization. Since cohesion and somatic homolog pairing are both interactions between chromosomes sharing sequence homology, it is reasonable to ask whether mechanisms of cohesion between sisters play a role in somatic homolog pairing in Drosophila. For example, cohesin molecules may encircle or otherwise spatially restrict homologs as well as sisters, or the presence of cohesion between sister chromatids may facilitate alignment and/or recombination of homologs, as occurs during meiosis in many organisms [51,52]. The role of cohesins in pairing homologs is supported by observations that meiosis I, which involves segregation of homologs but not sisters, requires separase, an enzyme that cleaves cohesin proteins [53–56]. While Drosophila is somewhat divergent from other organisms in the proteins that provide cohesion in meiosis, cohesion is still required for meiotic homolog pairing [57–62]. Therefore, cohesion might also be expected to play a role in somatic homolog pairing.

In an effort to identify mechanisms responsible for somatic homolog pairing and other forms of interphase nuclear organization, our laboratory previously conducted a genome-wide RNAi screen in Drosophila cells, assaying pairing at two distinct heterochromatic loci using high-throughput fluorescent in situ hybridization (Hi-FISH) [63]. In this study, a single FISH signal for a given target indicated that all copies of that locus were in close proximity to each other, and RNAi knockdowns leading to more or fewer FISH signals than observed in control cells identified candidate genes for promoting or antagonizing pairing respectively [63]. As such, we expected to identify cohesin proteins in this screen; loss of cohesion was expected to cause separation of sister chromatids in G2, thus increasing the number of FISH signals observed (Figure 2.3). Surprisingly, cohesin proteins were not among the 105 genes identified. This was our first indication that Drosophila cells might have cohesin-independent pairing of homologous chromosomes as well as of sister chromatids in interphase, though there were many other reasons why cohesin proteins might not have been identified in the screen.
In this chapter, I present characterization of this phenomenon of cohesin-independent pairing, demonstrating that our results cannot be explained by inefficient cohesin knockdown, cell cycle arrest, or other complicating factors. Additionally, I show that cohesin-independent pairing occurs genome-wide in Drosophila interphase cells, including pericentric repetitive as well as single-copy euchromatic regions. Finally, I present evidence suggesting that cohesin-independent pairing does not require the presence of a homolog, or of pre-established homolog pairing, suggesting that sister chromatids themselves may have some amount of cohesin-independent pairing between them, which allows them to remain in close proximity in G2.

Figure 2.3 Cohesin knockdown was expected to cause sister chromatid separation in G2 and in metaphase. (A) When a specific DNA sequence is targeted by fluorescent in situ hybridization (FISH) in Drosophila cells, most cells show a single FISH signal (red spot) in interphase because of homolog pairing and sister chromatid cohesion. (B) Knockdown of cohesin was expected to cause sister chromatid separation and thus an increase in the number of FISH signals both in G2 and in metaphase. Note that only one homolog is shown for clarity; in reality, multiple homologs may be present, and an increase in the number of FISH signals would be expected if there was homolog unpairing, loss of sister chromatid cohesion, or both.
Materials and Methods

Cell culture and RNAi

*Kc*<sub>167</sub>, S2R+ and Clone 8 cells were cultured according to standard protocols (see www.flyrnai.org). RNAi treatments were started in each case one day after the cells had been split as part of their regular passaging. RNAi treatments lasted for four days unless otherwise specified. For *Kc*<sub>167</sub> and S2R+, cells were seeded at 0.5-0.8 million cells/mL with 15 ug of RNA per well in a 6-well plate or 5 ug of RNA per well in a 24-well plate. For Clone 8, cells were transfected with dsRNA using Effectene transfection reagent from Qiagen, with a GFP-expressing plasmid as a co-transfection marker. When using Effectene, the amount of dsRNA was reduced to 1.2 ug per well in a 24-well plate. dsRNA primers were designed using the SnapDragon tool for primer design (http://www.flyrnai.org/cgi-bin/RNAi_find_primers.pl) and synthesized by PCR amplification from genomic DNA followed by an *in vitro* transcription reaction using a MEGAscript® T7 Transcription Kit (Thermo Fisher Scientific).

Quantitative PCR

Quantitative PCR was used to assay efficiency of RNAi knockdowns according to standard techniques. Briefly, total RNA was isolated from cells using a Qiagen RNeasy Plus kit and then converted to cDNA using the Invitrogen SuperScript® III First-Strand Synthesis System for RT-PCR. Primers for qPCR were designed using the Primer-BLAST website (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Reactions were set up using KAPA SYBR® FAST qPCR kits and run on an Applied Biosystems 7300 Real-Time PCR System.

Western blot

Cells were collected after four days of RNAi and their protein levels were analyzed by western blot according to standard protocols. Blots were probed using a rabbit anti-Rad21 antibody (generous gift from Dr. Stefan Heidmann; used at 1:3000) to assay cohesin knockdown and a mouse anti-α-tubulin antibody (Sigma-Aldrich; 1:5000) to assay loading, followed by secondary antibodies conjugated to HRP (GE Healthcare Life
Sciences), anti-rabbit (1:5000) and anti-mouse (1:10000). Blots were then stained using Pierce ECL Western Blotting Substrate (ThermoFisher Scientific).

**Fluorescence Activated Cell Sorting (FACS)**

Following RNAi treatment, cells were harvested, resuspended in ice cold 100% ethanol, allowed to warm up to 37°C and then stained using a PI/RNase Staining Buffer (BD Pharmingen). Cell populations were assayed based on DNA content to determine their cell cycle profile using a BD LSR II Analyzer.

**Immunofluorescence (IF)**

Cells were plated onto slides at concentrations of 1-5 million cells/mL and allowed to adhere for 1-2 hours. Slides were washed in PBS and fixed in 4% paraformaldehyde for 5-10 minutes. The slides were then washed in PBS and used immediately or stored in PBS at 4°C. IF slides were washed in PBS-T (PBS with 0.1% Tween-20), blocked in 1% BSA/PBS-T for 30 minutes at room temperature, and incubated with primary antibody at 4°C overnight followed by three more PBS-T washes and incubation with secondary antibody either for 2 hours at room temperature or overnight at 4°C. Slides were then washed in PBS-T and mounted using Slowfade with DAPI (Thermo Fisher Scientific), followed by imaging. Primary antibodies used: rabbit α-Rad21 (gift of Dr. Stefan Heidmann; 1:200), mouse α-cyclin B (Developmental Studies Hybridoma Bank; 1:100). Secondary antibodies used: Cy3-conjugated anti-rabbit (1:165), 488-conjugated anti-mouse (1:100), Cy5-conjugated anti-mouse (1:20) (all from Jackson ImmunoResearch Laboratories).

**Fluorescence in situ hybridization (FISH)**

Our FISH protocol has been previously published [63,64] and was adapted from standard protocols [65–67]. In brief, cells were fixed as above and then washed in PBS, 2x SSCT (0.3M sodium chloride, 0.03M sodium citrate, 0.1% Tween-20), and 50% formamide/2x SSCT. Slides were either used for FISH immediately or stored in 50% formamide/2x SSCT at 4°C. FISH slides were pre-denatured in 50% formamide/2x SSCT at 92°C for 2.5
minutes and then at 60°C for 20 minutes. FISH probes were added in a hybridization solution of 10% dextran sulphate/2x SSCT/50% formamide containing 10-20 pmol of probe per hybridization. The slides were then denatured by placing them on a heat block at 92°C for 2.5 minutes and allowed to hybridize overnight at room temperature for heterochromatic probes and at 37-42°C for euchromatic probes. Following hybridization, slides were washed in 2x SSCT at 60°C for 15 minutes, 2x SSCT at room temperature for 10 minutes, and 0.2x SSC at room temperature for 10 minutes before being mounted using Slowfade with DAPI (Thermo Fisher Scientific) and imaged. In cases where both IF and FISH were carried out, generally the two protocols were carried out in succession and the slides imaged afterwards. For some more sensitive antibodies, such as that identifying Cyclin B, the cells were imaged following IF, then washed, used for FISH, and re-imaged, using software-assisted stage navigation to relocate the same fields. We applied this method to Clone 8 cells to examine the localization of FISH probes in cells where we had previously determined the presence or absence of various proteins by IF (Figure 2.6).

Most euchromatic FISH probes used in this study were designed and generated using our published Oligopaints protocol [64], including 5A, 16E, 24D, 28B, 69C, 89B, 89E, and 100B, as well as the chromosome paints on 2R (41E-44C, 50D-53C, 58D-60E). One experiment at the 28B locus used a probe synthesized from a P1 plasmid (Berkeley Drosophila Genome Project) containing cloned Drosophila genomic DNA corresponding to chromosomal regions 28B1-28B2 (DS01529) and then labeled by nick translation/direct labeling (Vysis). Heterochromatic repeat regions were assayed using previously described FISH probe sequences [65,66] synthesized by Integrated DNA Technologies (IDT).

Metaphase spreads

Metaphase spreads were prepared using protocols adapted from published methods [68,69]. Cells were obtained from actively growing cultures without the use of any drugs to increase mitotic index unless otherwise specified. In the case where microtubule inhibitors were used, colchicine was added to the media of growing cells at a concentration of 30 μM for 2 hours prior to spread preparation. Cells from 5 mL of culture were spun down, washed once in PBS, and then gradually resuspended in 10 mL 1% sodium citrate. Sodium citrate was added in a dropwise manner, interspersed with frequent mixing to ensure even resuspension of the cells. The cells were then
incubated at room temperature for 30 minutes. We then added 1 mL of cold fixative (3:1 methanol: glacial acetic acid solution), spun down the cells, and washed three more times in 10 mL of the same fixative. Addition of fixative was also performed slowly. Finally, cells were resuspended in 100-500 μL fixative and dropped onto a glass slide under humidified conditions. The slide was allowed to dry in a humidified chamber and then washed in 70%, 90% and 100% ethanol successively, before being dried and imaged. For metaphase FISH, these slides were then denatured in 70% formamide/2X SSCT at 70°C for 90 seconds followed by washes in cold 70%, 90% and 100% ethanol. FISH probes were added and the cells were allowed to hybridize without any additional denaturing, followed by our standard FISH washes.

**Image acquisition and automated analysis using MATLAB**

All images were obtained using an Olympus IX83 epifluorescence microscope with a 60x oil objective and the CellSens acquisition software. Raw TIFF files were analyzed using custom-written MATLAB scripts (described in [63] and subsequently adapted) for measuring different properties such as the number of FISH dots per nucleus, their area, and the intensity of IF signals. First, nuclei were identified in the DAPI image, starting with correction for background followed by image segmentation based on a Ridler-Calvard thresholding algorithm [70]. Identified nuclei were filtered on the basis of shape and size, to eliminate clusters of nuclei or dead cells. Next, the areas corresponding to the identified nuclei were mapped in the images corresponding to the different FISH signals, and a threshold based on local background was used to identify all signals within each nucleus, which were then filtered on the basis of size. All uniquely identifiable foci of fluorescent signal (above background) were counted as FISH signals, regardless of the distance between them. When identifying FISH signals corresponding to large chromosomal targets, we found that the intensity of the FISH signals varied depending on the compaction status of the chromosome, so different thresholds were used based on the maximum intensity within that nucleus. Counts of FISH signals per nucleus along with other statistics were recorded, along with segmented output images, allowing the examination of results on a per-cell or per-image basis. Counts obtained from MATLAB algorithms were compared to manual counts of the same images to confirm that the results were comparable.
Statistical analyses

When assaying the number of FISH signals in a nucleus, a whole population of cells was scored and each nucleus classified as either having one signal (homolog pairing as well as sister chromatid cohesion intact) or more than one signal (homologs and/or sisters have become unpaired). The relative numbers of cells having one signal or more than one signal were then compared between different conditions using a two-tailed Fisher’s Exact Test. Similar analyses were used when comparing the percentages of metaphase spreads with intact versus perturbed sister chromatid cohesion. When examining the areas of FISH signals representing larger chromosomal regions, a Student’s t-test was used to compare the distribution of areas obtained. Finally, when examining the number of FISH signals in metaphase spreads, a Mann-Whitney U test was used to compare the different conditions.
Results

Pairing of homologs and sister chromatids at heterochromatic loci in Drosophila interphase cells can be maintained with little to no cohesin protein.

In our previously published screen for genes involved in homolog pairing [63], we applied Hi-FISH to tetraploid Kc167 cells in 384-well plates seeded with a whole-genome RNAi library and assayed pairing at two pericentric heterochromatic loci, one consisting of the 359 satellite repeats [71] on the X chromosome and the other consisting of the dodeca satellite repeats [72,73] on chromosome 3. In this study, the extent of homolog pairing was defined operationally as the percentage of nuclei in a population with one FISH signal per locus, and because a single signal would also require the close proximity of sister chromatids, our assay of homolog pairing also reflected sister chromatid pairing. Surprisingly, no cohesin subunits or associated proteins were identified in the screen, suggesting that pairing of homologs as well as sisters can occur independently of cohesin protein at the two loci assayed.

Given the unexpected nature of these findings, we began our studies by determining whether our failure to identify cohesin in the screen was due to an artifact of culturing cells in 384-well plates and using Hi-FISH as part of the screening protocol and/or due to incomplete knockdown by RNAi. To do so, we performed studies using conventional slide-based FISH in both Kc167 cells and another Drosophila cell line, S2R+. Under these conditions, we found that Rad21 was depleted below detectable levels by western blot or immunofluorescence after four days of RNAi, which was the duration of RNAi knockdown used in the screen (Figure 2.4, A&B). Furthermore, the mRNA levels of Rad21, Smc1, Smc3, and Stromalin (SA) were reduced by 80-90% as compared to levels in control cells when they were individually knocked down and assayed by qPCR (Supplementary Figure A1.1). Importantly, no significant reduction in the percentage of nuclei with only a single FISH signal was observed in cohesin RNAi-treated cells as compared to controls when we targeted FISH to 359 (81.8% versus 78.9%, P=0.4526) and dodeca (40.9% versus 36.0%, P=0.3909), as well as the AACAC pericentric heterochromatic repeat locus [74] on chromosome 2 (55.0% versus 58.0%, P=0.6214) (Figure 2.4C; values are for S2R+ cells, but lack of a significant reduction was also observed with Kc167 cells, see Figure 2.5 for more details). We also tested simultaneous
knockdowns of multiple cohesin proteins (Supplementary Figure A1.2A) and longer RNAi treatments (Supplementary Figure A1.2B) and found no consistently significant effects on the percentage of nuclei with single FISH signals.

Efficient knockdown of cohesin was also confirmed by the observation of premature sister chromatid separation in metaphase following cohesin RNAi (Figure 2.4D and Supplementary Figure A1.3). That is, after knockdown of Rad21, Smc1, or Smc3, sister chromatids appeared as single chromosomes rather than the pairs of connected chromatids that are normally observed in metaphase spreads (Supplementary Figure A1.3). We then extended this analysis to quantify the amount of separation caused by cohesin RNAi, focusing on Rad21 because the extent of unpairing observed following Rad21 knockdown was more severe than that observed following knockdown of either Smc1 or Smc3 (Supplementary Figure A1.3).
Figure 2.4 Cohesin knockdown in Drosophila cells disrupts sister chromatid cohesion in mitosis, but not the pairing of sisters or homologs in interphase. (A) Western blot showing that, after four days of RNAi in S2R+ cells, Rad21 knockdown is efficient. (B) Immunofluorescence for Rad21 confirms knockdown at the level of individual cells. (C) FISH targeting pericentric heterochromatin of the X chromosome and chromosomes 2 and 3 shows that Rad21 knockdown does not increase the number of FISH signals observed in interphase (dotted line, perimeter of DAPI signal). Graph shows percentages of nuclei with single FISH signals in Rad21 RNAi-treated cells compared to controls (controls include untreated cells and those treated with LacZ dsRNA; result is mean of 8-10
Figure 2.4 (Continued) independent trials; error bars = SD; n>190 nuclei per knockdown per trial; differences between controls and Rad21 RNAi were not significant by Fisher’s exact test (calculated for each trial) and by Student’s t-test for pooled averages from multiple trials (P=0.3909, P=0.6214, P=0.4526 for 359, AACAC and dodeca, respectively)). (D) Metaphase spreads with FISH show that Rad21 knockdown causes sister chromatid separation in mitosis, increasing the number of FISH signals per locus. Boxplot shows results from a single trial (n≥32 mitotic nuclei per knockdown; differences between controls and Rad21 RNAi treated cells were significant by Mann-Whitney U test, P<0.0001 for AACAC and dodeca, P=0.0126 for 359). For another independent trial, see Supplementary Figure A1.6. All scale bars represent 5 μm.

Here, we first determined the copy number of each chromosome in control cells by performing FISH on metaphase spreads, confirming that our S2R+ cell line has an irregular but stable karyotype, with two copies of the X, three copies of Chr. 2, and four copies of Chr. 3 (Figure 2.4D). We then applied FISH after Rad21 knockdown and observed an increase in the number of FISH signals at AACAC and dodeca (P<0.0001 for both); the median numbers of FISH signals increased from 3.0 and 4.0 signals in control cells at AACAC and dodeca, respectively, to 6.0 and 7.5 in Rad21 RNAi-treated cells (Figure 2.4D). While we cannot rule out contributions from aneuploidy, the approximate doubling of the median along with the observation of unpaired chromatids is strongly indicative of sister chromatid separation after cohesin knockdown. Importantly, this loss of cohesion in mitotic cells was unlikely to have had a significant impact on the overall percentage of nuclei with single FISH signals, since mitotic nuclei represent only a small percentage of a cycling population, around 7% following Rad21 knockdown in our assays (consistent with published results [28]).

Note that, while cohesin knockdown approximately doubled the median numbers of FISH signals in mitotic cells at AACAC and dodeca, it was not sufficient to completely disrupt cohesion at 359; while there was sometimes a significant increase in the number of FISH signals at 359 (P=0.0126), the median was unchanged, remaining at 2.0 and indicating that in many cells, sister chromatids remained connected (Figure 2.4D; note that significance was not always achieved for 359, see Supplementary Figure A1.6). It is possible that this could be explained by the presence of residual cohesin following RNAi treatment that is specifically enriched at this locus. Alternatively, it is possible
that 359 may retain a cohesin-independent connection between chromatids in mitosis, similar to that observed in interphase.

To better understand the progression of cohesin depletion, we performed a timecourse of Rad21 knockdown in Kc167 cells and observed premature sister chromatid separation in mitotic cells as early as the third day following RNAi treatment (Supplementary Figure A1.4C). As Kc167 cells complete the cell cycle in 24-30 hours [75], this observation argues that populations of cells that have been treated with RNAi for four days should have experienced cohesin depletion for the duration of at least one cell cycle. These studies also enabled us to address whether our inability to observe an effect of cohesin knockdown in interphase cells resulted from inadvertent disruption of the cell cycle; for example, arrest in G1, prior to S phase, would necessarily preclude sister chromatid separation. Evidence against this explanation was the fact that, while Rad21 knockdown caused an increased mitotic index, cells continued cycling, albeit with a delay as compared to control cells (Supplementary Figure A1.4A), consistent with published results [28,76]. In addition, both FACS analysis (Supplementary Figure A1.4B) and immunofluorescence for cyclin B, a protein that is expressed from S-phase through G2/M-phase [77–80] (Supplementary Figure A1.8B), confirmed that at least two-thirds of the cell population is in G2 following cohesin knockdown. These observations argue that the apparent maintenance of sister chromatid pairing following cohesin knockdown cannot be explained by a paucity of G2 nuclei. Thus, in conjunction with the findings described above, our studies indicate that the 384-well FISH format cannot explain why cohesin was not identified as a candidate gene in our pairing screen [63], and further, that neither inefficient knockdown nor a paucity of G2 nuclei can explain why cohesin RNAi treatment does not disrupt pairing in interphase cells. As such, our studies suggest the existence of a cohesin-independent mechanism maintaining the pairing of homologs throughout interphase as well as of sister chromatids in G2.

Metaphase separation of sister chromatids in the absence of cohesin is not dependent on spindle assembly.

Given that we observed sister chromatid separation in mitosis but not in interphase following Rad21 knockdown, we considered the possibility that this phenotype requires not only cohesin depletion but also the physical forces that separate sister chromatids in mitosis. As such, in G2 cells, perhaps sister chromatid pairing is
maintained not because sisters are held together in the absence of cohesin, but because they are not actively pulled apart. To address this possibility, we sought to determine whether the sister chromatid separation observed in metaphase following Rad21 knockdown is dependent on the formation of a mitotic spindle, by preparing metaphase spreads both in the absence and presence of the microtubule polymerization inhibitor colchicine.

Addition of colchicine increased the number of metaphase spreads that were obtained from 1.25% to 3.88% in control cells, and from 1.86% to 4.84% in Rad21 RNAi-treated cells (Table 2.1), consistent with its role in inhibiting the passage of cells through mitosis by blocking spindle assembly. However, we found that cohesion in metaphase following Rad21 knockdown was not significantly altered by the addition of colchicine (Table 2.1); the percentage of metaphase spreads with intact sister chromatid cohesion after Rad21 RNAi was 25.0% without colchicine and 33.6% with colchicine (P=0.1570), both of which were significantly less than the 77.1% observed in control cells treated with colchicine (P<0.0001) demonstrating that loss of cohesion in mitosis does not depend on spindle assembly. This suggests that the failure of Rad21 knockdown to cause sister chromatid separation in G2 is due to some other difference between interphase and mitosis, perhaps the existence of pairing mechanisms in interphase.

Table 2.1 Separation of sister chromatids following Rad21 RNAi is not dependent on microtubule polymerization.

<table>
<thead>
<tr>
<th>RNAi</th>
<th>Colchicine</th>
<th>Percentage of Metaphase Spreads with Intact Sister Chromatid Cohesion</th>
<th>Percentage of Metaphase Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>-</td>
<td>70.9% (n = 103)</td>
<td>1.25% (n = 1844)</td>
</tr>
<tr>
<td>Untreated</td>
<td>+</td>
<td>77.1% (n = 118)</td>
<td>3.88% (n = 1700)</td>
</tr>
<tr>
<td>Rad21</td>
<td>-</td>
<td>25.0% (n = 120)</td>
<td>1.86% (n = 1021)</td>
</tr>
<tr>
<td>Rad21</td>
<td>+</td>
<td>33.6% (n = 119)</td>
<td>4.84% (n = 826)</td>
</tr>
</tbody>
</table>
Cohesin-independent pairing of sisters and homologs is widespread in the Drosophila genome during interphase.

To determine whether cohesin-independent pairing of chromosomes is unique to centromeric regions and/or pericentric heterochromatin, we used FISH to target eleven euchromatic regions in a variety of genomic locations (Figure 2.5). Applying Oligopaint [64] FISH probes to control and Rad21 RNAi-treated cells, we targeted eight euchromatic loci ranging in size from tens to hundreds of kilobases and representing all major Drosophila chromosomes: 5A (X chromosome, target size 672.0 kb), 16E (X, 700.0 kb), 24D (2L, 490.6 kb), 28B (2L, 680.0 kb), 69C (3L, 674.0 kb), 89B (3R, 49.7 kb), 89E (3R, 49.7 kb) and 100B (3R, 462.3 kb) (Figure 2.5, A&C). Strikingly, we did not find the percentage of nuclei with a single FISH signal to be consistently and significantly reduced at any locus following Rad21 knockdown in either Kc_{167} or S2R+ cells (Figure 2.5C). These data suggest that cohesin-independent pairing is widespread in the Drosophila genome and pertains to single-copy euchromatic as well as to pericentric repetitive regions.
Figure 2.5 Pairing of homologs and sister chromatids in interphase remains unaffected genome-wide after knockdown of Rad21. (A) FISH targeting the 69C euchromatic locus (chromosome 3L, 674 kb) shows that the number of FISH signals per nucleus did not increase following Rad21 RNAi (dotted line, DAPI perimeter; scale bar = 5 μm). (B) FISH targeting the 41E, 50D and 58D loci (chromosome 2R, total of 8.4 Mb) shows that following Rad21 RNAi the size of the FISH signals did not increase (dotted line = DAPI perimeter; scale bar = 10 μm). (C) Schematic showing the three major chromosomes of Drosophila and summarizing the effects of Rad21 knockdown,
Figure 2.5 (Continued) with heterochromatic FISH targets highlighted in grey, smaller euchromatic FISH targets (tens to hundreds of kilobases) highlighted in yellow, and larger euchromatic FISH targets (several megabases) highlighted in pink. The sizes of these targets are as follows: 359 (estimated size ~11 Mb [65]), AACAC (size unknown), dodeca (size unknown), 5A (672.0 kb), 16E (700.0 kb), 24D (490.6 kb), 28B (680.0 kb), 69C (674.0 kb), 89B (49.7 kb), 89E (49.7 kb), 100B (462.3 kb), 41E-44C (3.1 Mb), 50D-53C (2.7 Mb), 58D-60E (2.6 Mb). For the grey and yellow targets, percentages represent the proportion of nuclei in a cycling cell population with a single FISH signal, while, for the pink targets, percentages represent the proportion of the nuclear area occupied by the FISH signals. Results for Kc167 are shown above each chromosome while those for S2R+ cells are shown below. (Data are from single trials except for 359, AACAC, dodeca, 16E, and 100B, where the percentages are the mean results from 2 or more trials; for all loci, n≥150 nuclei per trial. Significance was assessed using Fisher’s exact test (calculated separately for each trial), except for the pink targets, when a t-test was used to compare FISH signal areas (asterisks, P<0.05 for difference between control and Rad21 knockdowns in each trial)). A simple separation of sister chromatids would be expected to cause a decrease in the percentage of nuclei with a single FISH signal, or an increase in the area of FISH signals; we observed neither (see main text for details).

Considering the possibility that cohesin might be required to maintain pairing on a more global scale in ways not obvious from the analysis of short chromosomal regions, we also tested cohesin knockdowns with Oligopaints targeting three large regions on the right arm of chromosome 2 (3.1 Mb, 2.7 Mb and 2.6 Mb) (Figure 2.5B). Examining a large region minimized the chances of visualizing only late-replicating regions where, in early G2, sister chromatids may not yet have formed. Additionally, large FISH targets allowed a greater dynamic range in the size of the FISH signals, permitting us to more easily measure the area of the FISH signals in maximum-Z projections, in addition to counting the number of signals. We reasoned that an assay of signal size might be more sensitive to local unpairing of sister chromatids and/or homologs occurring anywhere along the chromosome arm even if complete separation had not occurred.

Following knockdown of Rad21 and FISH targeting these regions on chromosome 2R, neither the number of FISH signals nor the area of the image covered by these signals showed a significant increase, contrary to what
might have been expected if sister chromatids had simply separated (Figure 2.5C). The combined area of the FISH signals was 23.8% and 17.5% of the nuclear area in control and Rad21 RNAi-treated cells, respectively (P<0.0001 indicating a significant decrease). The decrease in signal areas we observed following Rad21 knockdown was unexpected, and could indicate an interesting role for cohesin in antagonizing compaction of chromatin, though further experiments are necessary to confirm this trend. Along these lines, it is interesting to note the significant increase in the percentage of nuclei with a single FISH signal at some loci examined with smaller FISH probe sets, specifically 89B in S2R+ cells and AACAC in Kc167 cells (see Figure 2.5C). Regardless, our data suggest that sister chromatids as well as homologs can remain paired in interphase across all major chromosome arms in the absence of Rad21. Therefore, cohesin-independent pairing may be a genome-wide phenomenon.

**Cohesin-independent pairing of sister chromatids in interphase does not require the presence of a homologous chromosome.**

Because homolog pairing does not appear to require cohesin proteins, we wondered if cohesin-independent mechanisms that regulate the pairing of homologs also regulate the pairing of sisters. For example, the mechanisms that pair homologs might also act directly between sister chromatids, keeping them in close proximity in interphase even in the absence of cohesin. Alternatively, it is possible that, because homologs are paired in G1, the replication products of these chromosomes can remain closely associated in G2 without mechanisms acting directly to hold sisters together (Figure 2.6A). To test this second possibility, we investigated whether cohesin-independent pairing of chromatids requires the presence of a homolog, by studying a chromosome that does not have a homolog, that is, the single X chromosome in a diploid XY male cell line. We reasoned that when we knock down cohesin, the alignment of sister chromatids in interphase will depend on cohesin-independent mechanisms, and if these mechanisms require the presence of a homolog, the X chromosome in a male cell line should display disrupted cohesion while the autosomes, which are present in two copies, should not.
Figure 2.6 Cohesin-independent pairing between sister chromatids does not require the presence of a homolog. (A) Cartoon showing theoretical interactions between chromosomes. Homologs and sister chromatids can be held together by a combination of homolog-homolog and sister-sister interactions (left) or just homolog-homolog interactions that indirectly hold sister chromatids together (middle, right). (B) Karyotype of Clone 8 cells, which have two copies of the autosomes (long arrows) and a single X chromosome (arrowhead). Scale bar = 5 μm. (C) Clone 8 cells following Rad21 knockdown (top row, immunofluorescence; bottom row, FISH). GFP serves as a marker for cells transfected with dsRNA, while Cyclin B (CycB) indicates cells in the G2 stage of the cell cycle. The cell marked by the arrow is in G2 and depleted for cohesin, but FISH indicates that sister chromatid pairing is unperturbed, both for an autosome (dodeca, 3rd Chr.) and for the X chromosome (16E). Scale bar = 5 μm. (D)
Figure 2.6 (Continued) Quantification of results for experiment illustrated in (C) showing percentage of nuclei with a single FISH signal at dodeca and 16E. (Means represent three independent trials; error bars = SD; n≥40 per trial per knockdown). Differences between control and Rad21 knockdown cells were not significant by Fisher’s exact test (P=0.0654 and P=0.7992 for dodeca and 16E, respectively).

For these studies we selected Drosophila Clone 8 (Cl.8+) cells, which we confirmed by karyotyping to be stably diploid and XY (Figure 2.6B). Given the low efficiency of RNAi in Clone 8, as versus Kc167 or S2R+ cells, we used GFP as a co-transfection marker for dsRNA. We carried out immunofluorescence for GFP, cyclin B (a G2 marker), and Rad21 to identify the cells of interest (positive for GFP and cyclin B and negative for Rad21) followed by FISH targeting 16E on the X chromosome and dodeca on chromosome 3 (Figure 2.6C). This allowed the identification of G2 cells in which cohesin was efficiently depleted, and the analysis of FISH results specifically from these cells. Consistent with our results in other cell lines, the percentage of nuclei with a single FISH signal at dodeca was not significantly different between control G2 cells and those treated with Rad21 RNAi (78.9% and 87.8%, respectively, P=0.0654) (Figure 2.6D). Remarkably, sister chromatid pairing at 16E on the X chromosome was also unaffected by cohesin knockdown, with 95.2% and 95.9% of control and Rad21 RNAi-treated G2 cells, respectively, having a single FISH signal (P=0.7992) (Figure 2.6D). These observations argue that, barring intrinsic features that may be specific to the X chromosome, cohesin-independent pairing of sister chromatids requires neither the presence of a homolog in G2 nor the pairing of homologs in the preceding G1. Thus, our data suggest that, whatever mechanism might be compensating for the loss of cohesin and keeping sister chromatids in close proximity in interphase, it initiates in G2 and acts directly between sister chromatids.

Note that 16E is located within the euchromatic arm of the X and >5 Mb away from the rDNA locus near the centromere. Thus, while the rDNA loci of the X and Y chromosomes support local pairing [81], we consider it unlikely that rDNA pairing accounts for cohesin-independent pairing of sister chromatids at 16E. That being said, it remains possible that inter-chromosomal associations occurring near the centromere might influence larger-scale organization. For example, pairing of X and Y near the centromere might lead to nonhomologous “pairing” between their arms, which could influence sister chromatid pairing as well.
Discussion

Our results show that Drosophila interphase cells lacking cohesin display levels of pairing comparable to control cells, as assayed by FISH at resolutions allowed by light-microscopy. These data suggest that pairing between maternal and paternal homologs as well as between sister chromatids can be maintained with little to no cohesin in interphase. While our studies cannot rule out that very low levels of cohesin persisted after RNAi treatment and/or that cohesin knockdown only affected a small proportion of cells, it remains surprising that the majority of cells maintain interphase chromosome organization with such low levels of cohesin (Figure 2.4, A&B). Indeed, non-mitotic functions of cohesin are thought to be more sensitive to cohesin depletion than mitotic ones based on studies in yeast [82], and mouse models of Cornelia de Lange syndrome, a developmental disorder caused by mutations in cohesin and other associated proteins, show that these mutations disrupt transcription but not chromosome cohesion in metaphase [83]. These observations suggest that the phenotypes we have observed in Drosophila cells may involve additional mechanisms that contribute to the organization of sister chromatids in interphase. We cannot rule out that such cohesin-independent pairing is induced in response to cohesin knockdown. Even so, our results demonstrate that cohesin is largely dispensable for the pairing of homologs throughout interphase and of sister chromatids in G2, suggesting that there may be cohesin-independent mechanisms contributing to this organization.

An alternative interpretation of these results is that, since sister chromatids are related by replication, they are necessarily in close proximity and remain that way in G2 without the need for a connection between them. Findings that argue against this interpretation are FISH results from other organisms following knockdowns of cohesin proteins, which demonstrate phenotypes consistent with a separation of sister chromatids in G2 (Figure 2.2) [27,31–35]. Additionally, our results obtained from metaphase cells (Table 2.1) indicate that spindle assembly and the resulting physical forces pulling sister chromatids apart in mitosis are not necessary to observe sister chromatid separation following loss of Rad21. Further, our results clearly demonstrate that cohesin is not required for the pairing of homologs, implying the existence of cohesin-independent mechanisms that can hold together chromosomes that share homology but are not related by replication. Importantly, homolog pairing can be disrupted
in interphase by RNAi knockdowns of genes other than cohesin, to an extent that is detectable by FISH [eg. 63,84,85].

While all the experiments presented here were done in Drosophila tissue culture cells, there is also evidence to suggest that cohesin is not required for the pairing of sisters and homologs in vivo. Studies in flies that have a TEV-protease cleavage site inserted into the Rad21 protein have been used to examine the effects of inducing premature dissociation of cohesin from chromatin in a tissue- and stage-specific manner [86]. Interestingly, in the salivary glands of Drosophila third instar larvae, Rad21 cleavage did not disrupt the alignment of polytene chromosomes [86]. This result is particularly striking because polytene chromosomes have been endoreduplicated, resulting in over 1000 sister chromatids held together in close alignment [87,88]. It will be of great interest to determine if this lack of a phenotype seen in polytene chromosomes will also be seen in other, actively dividing fly tissues. Another example of Rad21-independent cohesion is seen in Drosophila meiotic cells, which have meiosis-specific cohesion proteins such as Solo or Ord [57–61]. Our experiments in which multiple cohesin subunits, including Rad21 as well as Smc1 and Smc3, were knocked down (Supplementary Figure A1.2) argue against the idea that the pairing of sister chromatids in these cells is mediated by other proteins taking the place of Rad21 in the cohesin complex. However, it is possible that there are multiple novel cohesion proteins that have not yet been identified in Drosophila somatic cells.

A very interesting aspect of the biology of these cells is that they continue to cycle even in the absence of cohesin. We have shown that even following Rad21 depletion, cell populations continue to grow (Supplementary Figure A1.4), consistent with published results [28], and furthermore, that at least two-thirds of the population is in G2, the stage when sister chromatids are present (Supplementary Figures A1.4 & A1.8). This means that the loss of cohesion observed in metaphase chromosomes does not prevent passage through mitosis, even though this phenotype would be expected to activate the spindle assembly checkpoint (SAC) [89]. Interestingly, work done in Drosophila cells has shown that TEV-induced Rad21 cleavage only results in weak SAC activation and a short mitotic delay, the average length of mitosis being increased from 12 minutes to 38.1±13 minutes, with the frequency of unattached kinetochores being very low [76]. These results support our findings regarding cell cycle progression
following Rad21 knockdown, and hint at the possibility of cohesin-independent mechanisms that contribute to sister chromatid segregation in mitosis.

Interestingly, there is evidence from other organisms suggesting that mechanisms other than those mediated by cohesin proteins can contribute to cohesion [90–92]. For example, sister chromatid separation at the rDNA locus during mitosis in *S. cerevisiae* requires the activity of other proteins in addition to cohesin cleavage, suggesting that this locus has connections between sisters that are independent of cohesin [93–97]. Intriguingly, the rDNA locus in Drosophila is proximal to the 359 locus [65], where we saw that cohesin knockdown does not cause complete sister chromatid separation even in mitosis. In yeast, however, cohesin-independent pairing of sisters is region-specific, as other loci show sister chromatid separation soon after or even before the completion of DNA replication if cohesion is not established [11,90,98]. Similarly, the extent of premature sister chromatid separation observed in metaphase in the absence of cohesin varies across chromosomes in Xenopus [30,99], chicken [27], and human cells [100]. What we have observed in Drosophila is somewhat unique in that, at least in interphase, cohesin-independent pairing is observed throughout the genome, including both repetitive heterochromatic sequences near the centromere as well as single-copy euchromatic regions at different locations along chromosome arms, ranging in size from 50 kb to multiple megabases (Figure 2.5). It may well be no coincidence that Drosophila also supports extensive pairing of homologs.

Our work in Clone 8 cells demonstrates that cohesin-independent cohesion does not require the presence of a homolog, nor pre-established homolog pairing (Figure 2.6). One implication of this result is that whatever mechanisms maintaining the pairing of sisters in interphase can act directly between sisters, rather than relying on an indirect connection mediated through paired homologs. However, it is still possible that cohesin-independent pairing of sisters relies on homolog pairing in another way, by sharing a common regulatory mechanism. This possible relationship between sister chromatids and homologs, and the interactions taking place between them, is discussed in Chapter 3.
Acknowledgments

The work described in this chapter was supported by a grant and an NIH Director’s Pioneer Award to Chao-ting Wu from NIH/NIGMS (RO1GM085169 and 5DP1GM106412) and a Ruth L. Kirschstein National Research Service Award to Eric Joyce from NIH/NCI (F32CA157188). Additionally, I was supported by the Harvard Medical School Department of Genetics and by a training grant to the Harvard Medical School Program in Genetics and Genomics from NIH/NIGMS (T32GM096911). I am also grateful to the following: Stefan Heidmann of the University of Bayreuth for generously sharing the anti-Rad21 antibody, Suzan Lazo and other members of the Flow Cytometry facility at the Dana-Farber Cancer Institute for their assistance, Tiao Xie for computational expertise, and members of the Wu laboratory including Nicholas Apostolopoulos, Brian Beliveau, Mohammed Hannan, Hien Hoang and Son Nguyen for providing technical advice and assistance.

References


Chapter 3

Condensin-regulated homolog pairing also contributes to sister chromatid pairing in Drosophila

This chapter is partly based on:


Author contributions: TNS performed experiments and carried out data analysis. TNS and EFJ provided reagents. TNS, EFJ and C-TW conceived and designed the experiments. TNS and C-TW wrote the paper.
Abstract

Given that sister chromatids are able to maintain relative positioning in interphase with little to no cohesin protein, we wondered if mechanisms that pair maternal and paternal homologous chromosomes in Drosophila also contribute to sister pairing. We found that condensin II and the SCF Ubiquitin ligase component Slmb, in addition to regulating pairing of homologs, also regulate the pairing of sister chromatids when cohesin is absent. These results support a model wherein cohesin and condensin II oppose each other’s functions in the alignment of sister chromatids, and suggests that homologs and sisters may be paired by common, cohesin-independent, mechanisms. Interestingly, not all genes identified as regulators of homolog pairing also regulate sister pairing, suggesting that these other factors may play roles in different steps of pairing regulation or at different stages in the cell cycle. We also show that the role of condensin II in unpairing sisters in interphase is at least partially redundant with the role of condensin I in sister chromatid resolution in mitosis, suggesting that the mechanism by which condensin II regulates homolog pairing in Drosophila may be related to that by which condensin I regulates cohesion in mitosis, a function that is conserved in many species.

Introduction

The surprising finding that Drosophila cells are able to maintain pairing of sister chromatids in interphase with little to no cohesin protein, as described in Chapter 2, led us to question if other mechanisms might be keeping sister chromatids in close alignment, and what those might be. Such a mechanism, in addition to playing an important cohesin-independent role in the organization of the interphase nucleus, might also facilitate cohesin’s other roles in the regulation of gene expression [1–4] and DNA repair [5,6]. Furthermore, given our current understanding of how cohesin mediates cohesion from S-phase onwards [7–9], the observation of cohesin-independent pairing in G2 suggests that there may be more to be understood about the function of cohesin in G2, at least in Drosophila. In other organisms, there have been examples of “cohesin-independent cohesion” observed in metaphase chromosomes [reviewed by 10] but these reports have been limited to specific regions. The results described in Chapter 2 suggest that cohesin-independent pairing of sisters appears to occur chromosome-wide in Drosophila interphase cells. Since Drosophila are also somewhat unique in another aspect of nuclear organization,
which is the pairing of maternal and paternal homologs (Figure 3.1) [11–15], an attractive model is that the same mechanisms that pair homologs also act between sisters, thus maintaining the close alignment of sister chromatids in interphase even when cohesin proteins have been depleted to very low levels. If true, this idea would have several interesting implications for the nature of homolog pairing, which is still not well understood.

**Figure 3.1 Examples of paired Drosophila chromosomes in mitosis, in interphase, and in polytene nuclei.** (Left) A mitotic spread from Drosophila cells, with homologous chromosomes highlighted by chromosome-specific FISH probes, shows that homologs are associated even in metaphase. (Middle) Image from Drosophila cell culture, showing nuclei in interphase with chromosome-specific FISH and DAPI staining. Note that there are four copies of this chromosome in this aneuploid cell line, but they are all kept in close proximity. (Right) Image of polytene chromosomes from Drosophila third instar larvae, showing that pairing provides close alignment of homologous sequences. Polytenes image courtesy of Ting Wu.

Studies in other organisms have revealed several factors that may regulate cohesion independently of cohesin proteins [10]. For example, in budding yeast, components of the origin recognition complex (ORC) have been implicated in mediating cohesion [16,17]. ORC proteins bind at eukaryotic replication origins where they recruit other replication proteins [18]. Mutants in ORC components in yeast have premature loss of cohesion, an effect which is additive in double mutants with cohesin proteins, and overexpression of ORC components can rescue the cohesion loss observed in cohesin mutants [16,17]. Moreover, the cohesion defects observed following ORC depletion can be rescued if ORC production is restored in G2 [17], which is not the case for cohesin-mediated cohesion [19–21]. Therefore, it is likely that ORC contributes to cohesion independently of cohesin. Conversely,
there are also proteins that antagonize cohesion independently of cohesin removal, implying the existence of cohesin-independent connections between sister chromatids. For example, topoisomerase II is an enzyme that removes catenations formed during replication and other processes, and proper regulation of catenations is important for sister chromatid segregation [22–35]. Recent work has shown not only that catenations contribute to cohesion but also that cohesin may play a role in maintaining catenations [36,37]. It is unclear from these studies if catenations can also contribute to cohesion independently of cohesin.

Another group of proteins known to antagonize cohesion are condensin complexes [38,39]. Like cohesin, condensin has two SMC components, Smc2 and Smc4, in addition to non-SMC proteins [40]. Budding yeast has only one condensin complex, while most higher eukaryotes, including Drosophila, have two, both of which contribute to chromosome condensation [38,39]. Condensins also participate in other aspects of mitotic chromosome organization; for example, condensin mutants have errors in segregation such as anaphase bridges, which could be caused by defects in sister chromatid separation, in addition to condensation defects (for example, see Figure 3.2) [39,41–52]. Beyond these anaphase phenotypes, there is evidence from multiple organisms that condensin plays a direct role in antagonizing cohesion. For example, at the rDNA locus in budding yeast, condensin is required for the separation of sister chromatids even after cohesin cleavage, suggesting that there are cohesin-independent connections at this locus that are antagonized by condensin [53–57]. Additionally, in human cells, studies using premature chromosome condensation assays have found that condensin II begins the process of sister chromatid resolution beginning in late S-phase, prior to cohesin removal [58]. These results hint at an interesting interaction between cohesin and condensin, or at least, between their functions.
Figure 3.2 Examples of mitotic phenotypes caused by loss of the condensin I protein Barren in Drosophila.

Figures reproduced with permission from: (A) Bhat et al., 1996 [41] and (B) Oliveira et al., 2005 [50]. In (A), embryos have been stained for DNA (red) and α-tubulin (green), and white arrows highlight mitotically dividing cells. In (B), panels represent images from a time-lapse movie of mitotically dividing cells having fluorescently-labelled histones. Both barren mutants and Barren RNAi cells have defects in segregation, including anaphase bridges, which could be caused by failure to fully resolve sister chromatids.

Intriguingly, condensin II is also known to regulate homolog pairing in Drosophila. Overexpression of condensin II causes unpairing of polytene chromosomes and disrupts transvection in vivo (Figure 3.3), suggesting a role for condensin II in antagonizing pairing [59]. Condensin II components were also found to be important for resolving chromosomes in anaphase I of Drosophila meiosis [60]. Additionally, a high-throughput FISH screen for factors that regulate homolog pairing in cell culture identified multiple components of condensin II as anti-pairers, while a negative regulator of condensin II, the SCF ubiquitin ligase component Slmb, was found to be a pairing-promoter [61]. Further work has found condensin II to be important in the formation of chromosome territories [62] and identified additional regulators of condensin II that are required for proper nuclear organization [63–66].
example, Slmb is required for ubiquitination of the Cap-H2 subunit of condensin II, which results in down-regulation of the complex [63]. Therefore, the effect of Slmb on homolog pairing is dependent on Cap-H2, and consistent with this, double knockdowns of the two genes have the same effect as knocking down Cap-H2 alone, which is to enhance pairing [61].

The fact that condensin proteins have been implicated in the regulation of homolog pairing in Drosophila, as well as in the regulation of cohesion in organisms where homolog pairing is not prevalent, raises the intriguing idea that homolog pairing in flies might be mechanistically related to cohesin-independent cohesion in other species [58]. Here, we present data from Drosophila cells showing that knockdown of pairing regulators such as Slmb and Cap-H2 modify the phenotypes observed following cohesin loss in interphase. These results are consistent with what one might expect if mechanisms of homolog pairing are also contributing to the pairing of sister chromatids in interphase. Additionally, our data show that the organization of sister chromatids reflects an interplay between the activity of cohesin and condensin proteins, supporting the idea that these important SMC complexes contribute to chromosome organization in multiple ways.
Figure 3.3 **Condensin II overexpression causes polytene chromosome dispersal and homolog unpairing in Drosophila, while cohesin cleavage does not.** Figures reproduced with permission from: (A) Hartl, Smith, and Bosco, 2008 [59] and (B) Pauli et al., 2008 (content available under a Creative Commons Attribution (CC-BY) license) [67]. (A) Salivary gland polytene chromosomes showing that overexpression of Cap-H2, a condensin II subunit, causes polytene chromosome disassembly as well as homolog unpairing. (B) Salivary gland polytene chromosomes from flies expressing TEV-cleavable Rad21, compared to flies without a cleavage site in Rad21 (left column). Induction of TEV-protease by heat shock causes Rad21 cleavage and dissociation from DNA (images labelled “myc”), but polytene chromosome alignment, assayed by DAPI staining, is unperturbed.
Another potential player in this process in the interplay between sisters, homologs, and the inter-chromosomal interactions between them is the condensin I complex. Unlike condensin II, condensin I has not been implicated in the pairing of homologs in Drosophila interphase, but this is consistent with its known localization during this stage. Work done in mammalian cells has shown that the two condensin complexes have distinct subcellular localizations; condensin II is present in the nucleus throughout interphase and in mitosis, while condensin I is restricted to the cytoplasm until nuclear envelope breakdown in mitosis [52,68,69]. Consistent with this, condensin I proteins play many roles in mitotic cells, including in sister chromatid resolution, as described above [39,41–52]. To determine if the role played by condensin II in antagonizing sister chromatid pairing in interphase is similar to that played by condensin I in antagonizing cohesion in metaphase, we also present studies investigating the contributions of the two condensin complexes to sister chromatid cohesion in mitosis.

**Materials and Methods**

For a description of the methods used for cell culture, RNAi, qPCR, IF, FISH, metaphase spread preparation, image processing and statistical analyses, see the Materials and Methods of Chapter 2. Modifications specific to the work presented in this chapter are described below.

**RNAi**

The amount of dsRNA used for individual RNAi knockdowns in Kc167 or S2R+ cells was 15 ug of RNA per well in a 6-well plate or 5 ug of RNA per well in a 24-well plate. When knocking down multiple genes by RNAi, these amounts were scaled accordingly, with the amount of RNA for each target being kept the same as in the single knockdowns, resulting in a doubling of the total amount of dsRNA when simultaneously knocking down two genes, etc. We confirmed by qPCR that the increase in the amount of dsRNA did not affect the efficiency of RNAi knockdown (Supplementary Figure A1.5).
Antibodies

Primary antibodies used: mouse α-cyclin B (Developmental Studies Hybridoma Bank; 1:100), rabbit α-pH3 S10 (Epitomics, 1:100). Secondary antibodies used (Jackson ImmunoResearch Laboratories): Cy3-conjugated anti-rabbit (1:165), 488-conjugated anti-mouse (1:100), Cy5-conjugated anti-mouse (1:20).

Statistical Analysis

When analyzing knockdowns that increased the number of FISH signals, nuclei were sorted based on the relevant criteria (for example, one FISH signal versus more than FISH signal when considering chromosome pairing, or fewer than four versus four or more FISH signals when considering possible sister chromatid separation at AACAC, etc.) followed by a two-tailed Fisher’s Exact Test to compare the effect of different dsRNAs on the distribution of nuclei between these different categories. Additionally, when multiple trials of certain conditions were being compared, a two-tailed Student’s t-test was used to compare the percentages of nuclei having a certain number of FISH signals obtained under the different conditions.
Results

Slmb, a gene regulating the pairing of homologs, also regulates the pairing of sister chromatids in interphase.

To determine if mechanisms contributing to homolog pairing also contribute to the pairing of sister chromatids in interphase, we combined cohesin knockdown with knockdown of Slmb, a gene which is required for homolog pairing. Slmb is a negative regulator of condensin II, and Slmb knockdown leads to an increased number of FISH signals [61]. If Slmb is also required for the pairing of sister chromatids, we might expect simultaneous knockdown of both Slmb and cohesin to disrupt the pairing of sisters as well as of homologs, leading to even more FISH signals than when Slmb alone is knocked down.

We performed double knockdowns of Slmb and Rad21 in S2R+ cells and, having confirmed by qPCR that the knockdown of each gene was efficient (Supplementary Figure A1.5), assayed the number of FISH signals observed at pericentric heterochromatin (Figure 3.4A). Knockdowns of Slmb, whether alone or in combination with Rad21, reduced the percentages of nuclei with a single FISH signal at AACAC and dodeca from, respectively, 52.5% and 38.0% in control cells to 18.8% and 10.0% after Slmb knockdown and 26.0% and 11.9% following knockdown of both Slmb and Rad21 (Figure 3.4B). Therefore, pairing levels were similarly reduced whether we knocked down only Slmb or both Rad21 and Slmb; however, when unpairing did occur, we often observed more FISH signals when both Rad21 and Slmb were knocked down (Figure 3.4A). In particular, the double knockdown of Rad21 and Slmb produced nuclei with four to six or more FISH signals at AACAC (Chr 2), or five to eight or more FISH signals at dodeca (Chr 3), which is noteworthy because our S2R+ cells typically carry only three copies of chromosome 2 and four copies of chromosome 3 (see Chapter 2, Figure 2.4D). We reasoned that the "extra" FISH signals likely represented the separation of sister chromatids and applied this approach in subsequent analyses. That is, we considered the presence of more than three AACAC FISH signals or four dodeca FISH signals in a nucleus as indicative of sister chromatid separation.

Using this metric for identifying instances of sister chromatid separation, we observed that there is little sister chromatid separation following Slmb knockdown; the percentages of nuclei with more than three FISH signals
at AACAC or more than four FISH signals at dodeca were 4.7% and 11.2%, respectively, differing little from those of control cells (Figure 3.4C). This is consistent with findings for Slmb-depleted cells in mitosis, which have paired sister chromatids implying that cohesion is intact at this stage (Supplementary Figure A1.6). In contrast, the percentages of nuclei with more than three AACAC signals or more than four dodeca signals were 15.2% and 23.5% when both Rad21 and Slmb were knocked down, both values representing significant increases as compared to the outcome of knocking down Slmb alone (P<0.0001 and P=0.0035 for AACAC and dodeca, respectively) (Figure 3.4C). These findings suggest that, unlike knockdown of either Slmb or Rad21 alone, the double knockdown of Slmb and Rad21 results in sister chromatid separation as well as homolog unpairing.
Figure 3.4 Slmb and condensin II regulate sister chromatid pairing in interphase when cohesin is absent. (A) Nuclei from different RNAi knockdowns in S2R+ cells with FISH targeting AACAC (Chr 2) and dodeca (Chr 3). Knockdown of both Rad21 and Slmb produced more FISH signals as compared to knockdown of Slmb alone. The extra FISH signals seen in the double knockdown of Rad21 and Slmb were suppressed in a triple knockdown with Cap-H2 (dotted line, DAPI perimeter; scale bar = 5 μm). (B) Quantification of results for experiments illustrated in
Figure 3.4 (Continued) (A) showing percentages of nuclei with a single FISH signal, indicating nuclei where all sisters and homologs are paired. (C) Quantification of results for experiments illustrated in (A) showing percentages of nuclei with more than 3 AACAC signals or more than 4 dodeca signals, indicating nuclei with possible sister chromatid separation in addition to homolog unpairing. (In B & C, means represent 4-6 independent trials; error bars = SD; n≥100 nuclei per knockdown per trial. Significance was calculated for each trial using Fisher’s exact test and across several trials using a Student’s t-test; results from pooled tests are shown (*, P<0.05; **, P<0.0001).)

As the extra FISH signals could be explained by aneuploidy, we analyzed metaphase spreads following knockdowns, but did not find evidence for increased aneuploidy after double knockdown of Rad21 and Slmb as compared to knockdown of Slmb alone (Supplementary Figure A1.6). The extra FISH signals were also unlikely to reflect decompaction or fragmentation of heterochromatin, as double knockdowns of Rad21 and Slmb increased the number of signals at three out of five euchromatic loci studied (Supplementary Figure A1.7). The relatively modest effects observed at euchromatic as versus heterochromatic loci may stem from the overall higher levels of homolog pairing at euchromatin [61,70,71]. We also considered the possibility that the increase of nuclei with extra FISH signals represented the arrest of cells in mitosis, when sister chromatid cohesion is lost following cohesin knockdown. Here, we combined FISH with immunofluorescence to phosphorylated histone H3 (pH3) to identify mitotic cells [72] and, while the mitotic index is increased following knockdowns of Rad21 and Slmb, that increase cannot account for the overall increase in nuclei with extra FISH signals. In this experiment, after knockdown of both Rad21 and Slmb, the percentage of nuclei with more than three FISH signals at AACAC was 14.9%, and exclusion of pH3-stained nuclei dropped this number only slightly to 13.6%, which is still significantly higher than the percentage of interphase nuclei with more than three FISH signals after knockdown of Slmb alone (5.16%, P=0.0005). Furthermore, immunofluorescence for cyclin B confirmed that the increase in the number of FISH signals in the double knockdown was not caused by an enrichment of G2 cells (Supplementary Figure A1.8). In fact, knockdowns decreased the proportion of G2 cells from 66.9% in control cells to 46.2% after Slmb knockdown and 41.9% after knockdown of both Rad21 and Slmb. This decrease limits the number of nuclei where sister chromatid separation is possible, and may explain why the percentage of nuclei with extra FISH signals was never more than 30%. Therefore, while we cannot rule out the contributions of aneuploidy, disorganization of heterochromatin, or
cell cycle arrest, we favor the hypothesis in which the extra FISH signals in the double knockdowns of Rad21 and Slmb are caused by sister chromatid separation in interphase. This interpretation suggests that Slmb regulates sister chromatid pairing independently of cohesin.

The effect of Slmb on sister chromatid pairing is dependent on condensin II.

In addition to Slmb being identified as a pairing gene in the screen, several components of the condensin II complex were identified as causing increased pairing when knocked down (anti-pairers). Additionally, double knockdown of Slmb with condensin II component Cap-H2 suppressed the effect of Slmb, indicating that the effect of Slmb on pairing is dependent on condensin II [61,63]. Slmb is a negative regulator of condensin II, causing ubiquitination of the Cap-H2 subunit and destabilization of the complex [63]. Therefore, proper regulation of Cap-H2 by Slmb is required to maintain homolog pairing at wild-type levels. To determine if the extra FISH signals we observed in the double knockdown of Rad21 and Slmb was also dependent on condensin II, we carried out a triple knockdown of Rad21, Slmb, and Cap-H2. Remarkably, the number of nuclei with extra FISH signals was suppressed to levels comparable to that observed for the knockdown of Rad21 or Slmb alone (Figure 3.4C). In particular, while the percentages of nuclei with more than three FISH signals at AACAC or more than four FISH signals at dodeca were, respectively, 15.2% and 23.5% for the double knockdown of Rad21 and Slmb, the triple knockdown of Rad21, Slmb and Cap-H2 gave significantly lower percentages, with only 5.0% and 4.8% of nuclei having extra FISH signals (P=0.0008 and P=0.0016 for AACAC and dodeca, respectively) (Figure 3.4C). Importantly, we confirmed that this effect was not due to a reduction in the percentage of G2 cells in the triple knockdown (Supplementary Figure A1.8). The capacity of the triple knockdown to suppress the extra FISH signals argues strongly that the extra FISH signals reflect a genetic interaction between Rad21 and Slmb and not an artifact of the double knockdown. These results suggest that condensin II activity is required for the extra FISH signals. We propose that condensin II regulates sister chromatid pairing in interphase; specifically, condensin II may separate sister chromatids independently of cohesin removal.
Not all genes involved in homolog pairing also regulate sister pairing.

To determine if the interaction we observed between Rad21 and Slmb is a general consequence of unpairing homologs, rather than the result of a specific effect of Slmb and condensin II on sister chromatids, we tested if other candidate pairing genes showed similar genetic interactions with Rad21. Two other genes identified in the pairing screen [61] were Meph1, a component of the microtubule organizing center (MTOC) [73], and Pavarotti (Pav), a protein involved in cytokinesis [74]. Both of these genes were found to reduce levels of pairing when knocked down, putting them in the category of pairing-promoting genes, like Slmb [61]. We combined knockdown of Meph1 and Pav with that of Rad21, predicting that if Meph1 and Pav are required for sister chromatid pairing as well as for homolog pairing, the double knockdowns with Rad21 should show an increase in the number of FISH signals, similar to that observed in the double knockdown of Rad21 and Slmb.

Examining pairing at AACAC and dodeca, we found that, compared to the single knockdowns of Meph1 and Pav, double knockdowns with Rad21 did not significantly increase the number of FISH signals (Figure 3.5). Comparing knockdowns of Meph1 to those of both Rad21 and Meph1, there was a slight increase in the percentages of nuclei with extra FISH signals observed following double knockdowns (2.2% versus 4.8% for AACAC and 3.8% versus 7.8% for dodeca) but these increases were not quite significant/not consistently significant between trials (Figure 3.5A). Knockdowns of Pavarotti, in addition to generating multinucleate cells, gave rise to individual nuclei having extra FISH signals because of aneuploidy, and the proportion of these nuclei in the double knockdown with Rad21 was not further increased (20.2% versus 15.7% for AACAC, and 22.3% versus 20.8% for dodeca) (Figure 3.5A). Considering the possibility that a high level of aneuploidy makes it difficult to choose an appropriate threshold for the number of FISH signals, we also examined the total number of FISH signals observed and compared the distributions obtained after different knockdowns (Figure 3.5B). The distribution of nuclei was not observably changed in Rad21+Pav knockdowns compared to Pav knockdown, in contrast to the comparison between Rad21+Slmb knockdowns and Slmb knockdown (Figure 3.5B). These results raise two interesting points. Firstly, they suggest that the effect of Slmb/condensin II on sister chromatid pairing is not a general feature of genes that regulate homolog pairing. Secondly, because Pavarotti knockdown also leads to aneuploidy, the results of the Rad21+Pav double knockdown suggest that depleting cohesin from cells that are aneuploid does not necessarily lead
to an increased number of FISH signals. This is consistent with our hypothesis that the extra FISH signals observed in the Rad21+Slmb double knockdown are caused not by aneuploidy but rather by sister chromatid separation.

Figure 3.5 Two genes that regulate homolog pairing, Mcph1 and Pavarotti, do not regulate the pairing of sister chromatids in interphase. (A) Graphs representing the proportion of nuclei having more than three FISH signals at AACAC and more than four FISH signals at dodeca following knockdowns of Mcph1, Rad21+Mcp1, Pav, and Rad21+Pav (shown are means from two independent trials, error bars = SD, with n>200 nuclei per knockdown per trial). The differences between knockdown of Mcph1 and Rad21+Mcp1 were nonsignificant for AACAC (P_{trial1}=0.0868 and P_{trial2}=0.2642) and not significant borderline significant for dodeca (P_{trial1}=0.0570 and P_{trial2}=0.048). The differences between knockdown of Pav and Rad21+Pav were nonsignificant at both loci (P_{trial1}=0.5325 and P_{trial2}=0.0772 for AACAC, and P_{trial1}=0.9055 and P_{trial2}=0.4189 for dodeca). (B) Distribution of the number of FISH signals per nucleus obtained after knockdown of Slmb or Pav, compared to double knockdowns with Rad21 (data in top histogram is the same as that shown in Figure 3.4C, right, while that in the bottom histogram is the same as that shown in Figure 3.5A, right). The distribution of nuclei after knockdown of Rad21+Slmb is shifted to the right compared to the distribution seen after Slmb knockdown, consistent with an
The role of condensin II in regulating sister pairing in interphase is redundant with the role of condensin I in mitosis.

While the role of condensin II in antagonizing homolog pairing in interphase may be unique to Drosophila, Drosophila do have a second condensin complex, known as condensin I, and this complex plays multiple roles in mitosis including in sister chromatid resolution that appear to be conserved between Drosophila and other species [75]. For example, mutants or knockdowns of the Drosophila condensin I components Barren [41,44,50], Cap-G [47,48], and Cap-D2 [49] have segregation defects including anaphase bridges, which can result from failure to fully segregate sister chromatids in mitosis. We considered two possibilities for the role of condensin II in pairing. Firstly, condensin II might separate homologs and sisters in interphase by a completely different mechanism than that used by condensin I to segregate sisters in mitosis; alternatively, the two complexes might act by similar mechanisms and remove similar types of inter-chromosomal connections, with condensin II beginning the process in interphase and condensin I completing it in mitosis. If the latter is true, we would expect the contributions of condensin II in interphase to be redundant with those provided by condensin I in mitosis.

Therefore, we assayed the roles of both condensin complexes in regulating cohesion in mitosis. At this stage, in addition to increasing the number of FISH signals, Rad21 knockdown results in abnormal nuclear morphologies (Figure 3.6, A&B). In particular, the shape of mitotic chromatin, as revealed by staining for pH3, was highly irregular in cohesin knockdowns, often showing multiple pH3-staining bodies rather than one compact structure, which may reflect unpaired chromatids that fail to congress at the metaphase plate. In support of previous results [43,76], we found that double knockdowns of Rad21 and Smc2 (condensin I and II), and double knockdowns of Rad21 and Barren (condensin I only) rescued the defects in morphology of mitotic nuclei observed following Rad21 depletion (Figure 3.6, C&D). In particular, the percentage of cells with multiple pH3-staining bodies
decreased from 72.5% in a Rad21 knockdown to 21.8% in the double knockdown of Rad21 and Smc2, and to 20.4% in the double knockdown of Rad21 and Barren (for both double knockdowns, P<0.0001 compared to Rad21 knockdown). There was also a decrease in mitotic index, indicating reduced mitotic arrest.

Figure 3.6 Knockdowns of condensin I proteins modify the mitotic phenotypes of cohesin knockdown, while condensin II does not play an additional role in regulating cohesion at this stage. (A) Rad21 knockdown leads to an increase in the number of FISH signals in mitotic nuclei, identified by staining for phosphorylated histone H3.
Figure 3.6 (Continued) (pH3). (B) Rad21 knockdown also leads to abnormal morphologies of mitotic nuclei, including multi-lobed structures with breaks in the mass of chromosomes formed at mitosis. The dotted line indicates the DAPI perimeter. (C) Depletion of condensin protein Smc2 partially rescues defects observed after Rad21 knockdown; nuclear morphology resembles that of control cells. (D) Quantification of results for experiments in (B) and (C). Mitotic nuclei were classified as either having (Discontinuous) or not having (Continuous) discontinuities in structure, as revealed by pH3 staining. Results are shown for knockdowns of Rad21 alone and in combination with that of Smc2 (condensin I and II), Barren (condensin I only), and Cap-H2 (condensin II only). Significant rescues were observed in double knockdowns with condensin I proteins, while condensin II co-depletion did not consistently have a significant effect (Means represent three independent trials in S2R+ cells; error bars = SD; n≥53 mitotic nuclei per knockdown per trial; significance calculated for each trial using Fisher’s exact test (*, P<0.0001 in both trials; ~, P=0.1914 in one trial and P=0.0111 in another)). Mitotic index for each knockdown, determined as the percentage of pH3-positive nuclei, is shown beneath the graph (n≥850 per genotype).

The fact that knocking down Smc2 (both condensins) did not significantly enhance cohesion compared to knocking down Barren (condensin I only) suggests that the function played by condensin II in interphase to separate sisters may be redundant with that of condensin I in mitosis. Consistent with this idea, the double knockdown of Rad21 and Cap-H2 (condensin II only), did not significantly rescue the mitotic phenotypes of Rad21 depletion; the percentage of mitotic cells with multiple pH3-staining bodies was 52.9% in the Rad21+Cap-H2 knockdown, which was not significantly different from the 72.5% observed following Rad21 knockdown (see Figure 3.6 legend). There is some amount of rescue, however, so we cannot completely rule out that condensin II has additional roles at this stage. Notably, the rescue of the mitotic phenotypes by knockdown of either condensin complex is not complete; the double knockdown of Rad21 and Smc2 still exhibits the increased number of FISH signals seen after Rad21 knockdown (Figure 3.6C). Preliminary evidence suggests that this observation is true both at heterochromatic and euchromatic loci, and is not suppressed by the addition of colchicine to the cells to prevent spindle assembly (data not shown). This indicates that while reduced condensin activity improves the ability of cohesin-depleted cells to maintain cohesion in mitosis, cohesin-independent connections are not enough to maintain full alignment of sister chromatids at this stage like they seem to be in interphase.
Discussion

Here, we report a genetic interaction between Rad21 and Slmb, a gene required for homolog pairing [61,63]. Specifically, we observed that simultaneous knockdown of Rad21 and Slmb resulted in more FISH signals than when Slmb alone was knocked down (Figure 3.4), suggesting that in the double knockdown, there is separation both of homologs and of sister chromatids. We cannot completely rule out other explanations for the increase in the number of FISH signals such as aneuploidy, cell cycle arrest, or disorganization of heterochromatic repeats. However, we also note that, even if Slmb does not regulate sister pairing as we have proposed, any of the alternative explanations outlined above would still indicate an interesting relationship between Rad21 and Slmb and, therefore, between cohesion and homolog pairing. For example, if the double knockdown of Rad21 and Slmb causes more aneuploidy or mitotic arrest than knockdowns of each individual gene, it would indicate that the combined loss of homolog pairing and cohesion causes significantly more issues in mitosis than the loss of just pairing or just cohesion. Nonetheless, the results of control experiments presented in this chapter suggest that aneuploidy or cell cycle arrest are unlikely, leading us to favor the hypothesis that the double knockdown of Rad21 and Slmb leads to separation of homologs as well as sisters in interphase. This suggests that Slmb, through regulation of Cap-H2, contributes to cohesin-independent pairing of sister chromatids, as well as to homolog pairing.

Interestingly, this effect on sister pairing is not a general result of all genes that unpair homologs. Specifically, the factors Mcph1 and Pav, both of which increase the number of FISH signals when knocked down, likely reflecting unpaired homologs [61], do not show additional FISH signals when Rad21 is also knocked down, suggesting that the pairing of sister chromatids is intact in these cells (Figure 3.5). These findings highlight that the Slmb/condensin II pathway may be somewhat unique in its regulation of inter-sister as well as inter-homolog pairing. Perhaps this indicates that different pathways contribute to homolog pairing at different stages in the cell cycle. For example, since Mcph1 and Pav are both proteins that have their primary functions during mitosis, they may be involved in establishing patterns of chromosome organization, including homolog pairing, during this stage. The possible roles of mitotic factors in the establishment of interphase chromosome positioning will be discussed further in Chapter 4. Condensin II, on the other hand, may regulate pairing directly in interphase, including in G2, when sisters are present. In future studies, the presence or absence of interactions with Rad21 could be assayed for
additional candidate pairing genes and used as a metric to sort them into different pathways, which will help further our understanding of homolog pairing.

Along these lines, the big question remaining is how condensin II is acting to antagonize sister as well as homolog pairing. Regarding the role of condensin I in sister chromatid segregation, various models have proposed that condensin may remove cohesin protein from chromosomes, or that it may recruit topoisomerases to remove catenations between chromosomes [75]. Our results might also be explained by condensin complexes having a role in cohesin removal; in this case, depletion of either condensin may stabilize a small population of cohesin protein remaining after RNAi, explaining the rescue of mitotic phenotypes in a double knockdown of cohesin and condensin (Figure 3.6). In interphase, Slmb knockdown may lead to more FISH signals by enhancing condensin II activity, which removes residual cohesin following Rad21 knockdown and resulting in sister chromatid separation. Two lines of evidence argue against this interpretation. Firstly, the efficiency of our Rad21 knockdown (Chapter 2, Figure 2.4, A&B) argues strongly in favor of a model wherein condensin regulates sister pairing in interphase and sister cohesion in metaphase independently of cohesin protein. Secondly, the knockdown of Slmb alone does not disrupt sister pairing, as demonstrated by the fact that sisters arrive in metaphase with intact cohesion (Supplementary Figure A1.6). Further investigations into the stability of Rad21 on interphase chromatin following Slmb knockdown will help answer this question.

An alternate possibility is that condensin activity removes catenations that form between chromosomes during DNA replication. In fact, one of the earliest models for cohesion, proposed before cohesin proteins were known, posited that sister chromatids could be held together by DNA catenations [77–79]. Condensin proteins may resolve catenations by recruiting topoisomerase II to chromosomes, or in other ways [see 75 for review]. In addition to the roles played by mitotic condensin in sister chromatid segregation in many organisms [39,41–52], condensin II in human cells has been shown to begin separating sister chromatids beginning in late S-phase, when cohesin is still present on the chromosomes [58]. Further experiments investigating the exact timing of removal of pairing by condensin II, and determining whether or not topoisomerase II is involved in this regulation, will reveal more about condensin’s function in antagonizing sister chromatid pairing.
If catenations are the basis of cohesin-independent sister pairing, it would imply that catenations contribute to interphase nuclear organization, perhaps being the basis of pairing between homologs as well. In this regard, our work suggests that the pairing of homologs and its regulation by condensin II may not be a weird phenomenon specific to flies, but one that shares a mechanistic basis with condensin I-mediated sister chromatid resolution, a function that is conserved in many different species. Our results also add to the growing list of examples of how different SMC complexes, including both cohesin and condensin, have interacting functions. These ideas will be discussed further in Chapter 5.

Acknowledgments

The work described in this chapter was supported by a grant and an NIH Director’s Pioneer Award to Chao-ting Wu from NIH/NIGMS (RO1GM085169 and 5DP1GM106412) and a Ruth L. Kirschstein National Research Service Award to Eric Joyce from NIH/NCI (F32CA157188). Additionally, I was supported by the Harvard Medical School Department of Genetics and by a training grant to the Harvard Medical School Program in Genetics and Genomics from NIH/NIGMS (T32GM096911). I am also grateful to Tiao Xie for computational expertise, and members of the Wu laboratory for their technical advice and assistance.
References


124


Chapter 4

Studies on the organization of pericentric heterochromatin and the clustering of centromeres in Drosophila cells

This chapter presents the results of an unpublished collaboration between myself and Eric Joyce. Both Eric and I conceived and designed experiments, while Eric performed FISH experiments and carried out the imaging, and I developed image processing algorithms and carried out image and data analysis. I wrote all parts of this chapter.

In the second part of this chapter, the images used to screen for candidate clustering genes were initially taken as part of a high-throughput FISH screen for genes that regulate homolog pairing, conducted by Eric and others, which is described in:

Abstract

Drosophila cells show robust clustering of their centromeres and pericentric heterochromatin in interphase, but only a few factors that regulate this important aspect of nuclear organization are currently known. Centromere clustering is also observed in Drosophila cultured cells, and we took advantage of this to further characterize this clustering and to screen for additional factors that regulate it. We present data from FISH experiments in which each centromere is uniquely identified showing that centromeres do not have preferences for clustering with other specific chromosomes, beyond their preference for homologous associations. We also present reanalysis of images taken during a previous high-throughput FISH screen for genes involved in somatic homolog pairing; specifically, we examined distances between nonhomologous FISH targets to identify genes involved in the clustering of pericentric heterochromatin. From a primary screen of 352 genes, we identified 54 candidate pro-clustering factors and 23 candidate anti-clustering factors, including several previously known regulators. Candidates were enriched for mitotic genes, highlighting the importance of the mitotic stage of the cell cycle in the establishment of centromere positioning in interphase. Finally, we also found that increased clustering was correlated with greater distances from the nuclear periphery, demonstrating that nuclear positioning is also an important aspect of clustering and suggesting that the clustering of centromeres and/or pericentric heterochromatin may influence nuclear organization on a broader scale.

Introduction

The clustering of centromeres in interphase is a common feature of nuclear organization that has been observed in different organisms. Since the regions surrounding centromeres are enriched for heterochromatic sequences, centromere clustering also involves heterochromatic associations [1,2], and these two features have multiple roles in the cell. For example, during interphase in fission yeast, centromeres are clustered at the nuclear envelope at a site associated with the spindle pole body (SPB) in the cytoplasm [3], and this clustering is important for nuclear envelope stability and the association of microtubules in the cytoplasm with the nucleus [4]. Centromere clustering is also tightly regulated; for example, during fission yeast meiosis, centromere clustering is disrupted
while telomeres form a bundle at the SPB, an exchange which is important for meiotic homolog pairing [5–7]. This role of clustering a certain part of the chromosome in establishing inter-chromosomal interactions elsewhere along the chromosome arm highlights that, in addition to playing structural roles in the communication between the nucleus and the cytoskeleton, centromere clustering also affects the structure of chromosomes within the nucleus, and does so beyond the organization of heterochromatin. Mathematical modeling of chromosome movements observed in vivo in budding yeast (which, like fission yeast, have both centromere clustering at the SPB as well as telomere clustering [8–11]) indicates that tethering of chromosomes at centromeres and telomeres determines the fluctuations of a chromosome arm and the formation of chromosome territories [12]. Therefore, centromere clustering may play multiple roles in chromosome and nuclear structure.

Centromere clustering and pericentric heterochromatin also play roles in gene regulation. For example, in mammalian cells, rather than all centromeres forming one cluster, multiple small clusters containing more than one centromere are formed [13–18]. Known as chromocenters, these clusters establish heterochromatic compartments distributed around the nucleus, the extent and positioning of clustering varying by cell type and by organism [19–23]. Genes that are not pericentromeric may be recruited to chromocenters for silencing [24]. In mouse primary B lymphocytes, silenced genes are associated with centromeres, a localization that is thought to involve the transcription factor Ikaros [25–29], and a similar stage-specific repositioning of genes is observed in the development of helper T cells, though these effects may be locus-specific [30]. Furthermore, in addition to playing a role in the regulation of specific genes, centromere clustering may be important at a broader scale in the establishment of a heterochromatic compartment. In Drosophila cells, where heterochromatin is spatially segregated from euchromatin (Figure 4.1A), knockdowns of genes required for centromere clustering also cause disorganization of heterochromatin and as a result, mis-regulation of transposable elements and genome instability (Figure 4.1B) [31]. These findings highlight the importance of centromere clustering for developmental processes as well as for functions essential in all cells.
Figure 4.1 Euchromatin and heterochromatin are spatially segregated in Drosophila nuclei, and the proper clustering of centromeres is required for this organization. (A) (Right) Example nucleus stained for histone 3 lysine 4 dimethylation (H3K4me2) and Heterochromatin Protein 1α (HP1α) show that euchromatin and heterochromatin, respectively, form separate compartments. (Left) Map of Drosophila genome, showing location of major heterochromatic blocks. Figure courtesy of Eric Joyce. (B) Knockdown of NLP, a gene required for centromere clustering also disrupts organization of heterochromatin. Figures reproduced with permission from...
As the studies described above highlight, a few factors required for centromere clustering are already known. In fission yeast, the clustering of centromeres at the SPB requires centrosome components [32], kinetochore proteins [33,34], nuclear proteins as well as nuclear envelope components [35], though the role of some nuclear envelope proteins is debated [4,36]. In Drosophila cells, centromere clustering occurs at the nucleolus and requires the protein NLP, or nucleoplasmin [31]. NLP is a Drosophila homolog of the nucleophosmin/nucleoplasmin family of nuclear chaperones [37]; nucleophosmin in humans is known to localize to the nucleolus [38,39], to associate with the nucleolar protein nucleolin (known as Modulo in Drosophila) [40], and to interact with centromeric histone H3 [41]. These observations in human cells are consistent with the idea that NLP/nucleophosmin might tether centromeres to the nucleolus [42], and work by Patrick Heun’s group showed that this is indeed the case in Drosophila, where NLP, along with Modulo and another protein known to interact with nucleophosmin, CTCF [43], are essential for both centromere clustering and tethering to the nucleolus (Figure 4.1B) [31]. Artificial targeting of NLP to the nuclear periphery also caused over 40% of centromere clusters to be localized at the periphery instead of at the nucleolus, suggesting that it might be sufficient to recruit centromeres [31], though in this case since the mechanisms required for native centromeres to cluster at the nucleolus are still intact, it is difficult to say whether NLP is also sufficient for clustering. Additionally, it is unclear when in the cell cycle this tethering takes place; for example, it is possible that repositioning of centromere clusters can take place during interphase, or alternatively, that passage through mitosis is required for centromere clusters to be recruited to either the periphery, in the case where NLP has been artificially targeted, or in the case of normal cells, to the nucleolus.

The idea that features of interphase nuclear organization might be established during mitosis was first suggested by Carl Rabl in 1885, who proposed that the clustering of centromeres that occurs as they are pulled towards the spindle pole during anaphase is maintained in the subsequent interphase [44]. Consistent with this idea,
nuclei in many organisms, including in certain cell types in Drosophila, show clustering of centromeres at one end of the nucleus with clustering of telomeres at the other, an organization known as the “Rabl configuration” [3,35,45–58]. Interestingly, the knockdown of NLP does not disrupt telomere clustering in Drosophila, so its effect is specific to centromeres [31]. Consistent with the idea that centromere clustering can be regulated independently of telomere organization and/or mitosis, centromere clustering is seen in cells that do not show the Rabl configuration, and the extent of centromere clustering can vary between cell types, during development, and even over the course of the cell cycle [14–16,22,23,52,59–69]. These findings suggest that there may be dynamic regulation of centromere clustering outside mitosis, and in cell types where clustering is observed, mechanisms to maintain this organization in interphase. Additionally, there have been studies showing that passage through mitosis is not necessarily required for establishing centromere clustering, further highlighting the possibility of interphase mechanisms involved in centromere clustering [9,70]. In Drosophila cells, the amount of clustering may vary between G1 and G2 (Figure 4.2). Therefore, it is of great interest to identify factors that act to establish or regulate the clustering of centromeres in interphase as well as in mitosis, and to study the interactions between them.

![Figure 4.2](image.png)

**Figure 4.2 Centromere clustering in Drosophila interphase cells observed by CID staining.** Immunofluorescence for CID (the centromeric histone H3 variant in Drosophila) and Fibrillarin identifies centromeres and the nucleolus, respectively, while cyclin B (CycB) identifies a cell in G2. Note that the amount of centromere clustering varies between the G1 nucleus (left) and the G2 nucleus (right) shown here. Figure courtesy of Eric Joyce.
Drosophila cells provide an ideal system for studying the different pathways that contribute to centromere clustering and positioning. The study examining the role of NLP in centromere clustering was done using immunofluorescence for centromeric proteins in Drosophila S2 cells, which show robust centromere clustering [31]. Given the ease of RNAi knockdown in Drosophila tissue culture cells and of high-throughput FISH assays, Drosophila cells can easily be screened for factors involved in nuclear organization. Our laboratory previously published a screen for genes involved in somatic homolog pairing, which examined the number of FISH signals observed at two different pericentric heterochromatic loci on different chromosomes [71]. Due to the genomic proximity of these sequences to centromeres, we realized that by considering the distance between the two loci, we could also use the data from this screen to identify potential factors involved in centromere clustering. In this chapter, I first describe results from five-color FISH experiments in which all centromeres are uniquely identified, showing that FISH to pericentric loci can be used to assay nonhomologous clustering. These results also demonstrate that there are no pairwise association preferences between nonhomologous centromeres, and that clustering is largely stochastic. I then describe our results from reanalysis of data from a previous screen, to identify genes that regulate the clustering of pericentric heterochromatin, and potentially, of centromeres.

The work presented in this chapter was done in collaboration with Eric Joyce, a post-doctoral fellow in the Wu laboratory. Eric carried out the FISH and imaging experiments, and therefore contributed data to all figures and tables in this chapter. Since my contribution to the project was largely in the form of image analysis and data processing, in this chapter I focus on image analysis methods I developed for this project and the results obtained from these.
Materials and Methods

For details of our FISH protocol, see the Materials and Methods of Chapter 2. The modifications presented here are specific to re-hybridization (experiments in which multiple rounds of FISH were carried out) and image registration (aligning images of the same nuclei taken on separate occasions). Additionally, the identification of candidate clustering factors described in the second part of this chapter was based on the analysis of images originally taken in a previous screen, using a high-throughput FISH protocol in 384-well plates that has been published [71]. Here describe how we identified candidate clustering genes from these images.

Re-hybridization

Our re-hybridization experiments used S2R+ and Clone 8 cells, with FISH probes targeting pericentric heterochromatic repeat loci. For multiple rounds of FISH, the first FISH experiment was carried out as previously described (Materials and Methods, Chapter 2). The FISH probes used in the first experiment and their conjugated fluorophores were dodeca-Alexa488 (Chr. 3), AACAC-Cy3 (Chr. 2), and 359-Cy5 (X Chr.). As in our other FISH experiments, probes were added following the denaturation step, in a hybridization mixture of 10% dextran sulphate/2x SSCT/50% formamide, with each probe being used at a 1:100 dilution. Probes were then allowed to hybridize at 37°C overnight, followed by washing, mounting the slide in DAPI, and microscopy. When imaging, the stage navigation tools in the Olympus CellSens software were used to save the position of each image acquisition.

Following imaging, the coverslip of the slide was carefully removed and the slide was washed in 50% formamide/2x SSCT (2x SSCT = 0.3M sodium chloride, 0.03M sodium citrate, 0.1% Tween-20) at 60°C for 10 minutes, followed by washing in 50% formamide/2x SSCT at room temperature for another 10 minutes. The slides were then allowed to dry and the second set of FISH probes was added (using the same hybridization mixture and probe concentration as before). The FISH probes used in the second experiment were dodeca-Alexa488 (Chr. 3), AATAC-Cy3 (Chr. Y), and AATAT-Cy5 (Chr. 4). Additionally, probes that were not conjugated with any fluorophores were added to the mixture, these probes being specifically designed to displace the AACAC and 359 FISH probes used in the first experiment. Known as “toehold exchange” probes, these probes shift the
thermodynamic equilibrium in favor of strand displacement, by forming base-paired structures that are thermodynamically more stable than the original FISH probe duplexed with genomic DNA [72]. Toehold probes for AACAC and 359 were added at concentrations of 25 pmol. Some slides were imaged without the addition of the second set of labelled FISH probes, confirming that the first set was efficiently removed (data not shown). After the addition of probes, the FISH slides were placed on a heat block at 70°C for 5 minutes, and then allowed to hybridize overnight at 37°C, followed by the usual FISH washes, mounting in DAPI, and imaging. The stage navigation tools allowed approximate relocation of the same fields of cells acquired during the first round of imaging.

**Automated Image Registration**

When aligning images taken from different FISH experiments, I first analyzed the images corresponding to dodeca (the 488 channel) from each hybridization. Signals were identified in maximum-Z projections of each image in the same way as described previously (Materials and Methods, Chapter 2), and the centroid of each FISH signal was found. The two images were overlaid, with the FISH signals from separate images in different colors. Then, for each dodeca signal from the first image, the surrounding area was searched for the closest FISH centroid from the second image. For this step, I used nuclei with only one dodeca signal, to simplify identification of the corresponding signal from the second image. The pixel search area was varied by image and depended on the density of nuclei and the distance by which the two images were offset. When a corresponding FISH signal was identified, the distances between the signals in the X- and Y-planes were measured. After measuring distances for each nucleus, the average offset in XY was calculated for the image. Nuclei that were shifted by more than a standard deviation away from the mean were discarded, and the average offset was recalculated.

Then, the dodeca image from the second hybridization was shifted in XY by the average offset calculated across the whole image, and the two dodeca images were again overlaid. The amount of pixel overlap between dodeca signals from different hybridizations was measured as a percentage of the area of the signal for each nucleus. Nuclei in which the overlap was less than 40% were noted, to be discarded from further analysis. This threshold was chosen based on visual analysis of aligned images, to allow the maximum number of nuclei for analysis while maintaining a good alignment (Supplementary Figure A2.1). Then, the additional three channels from the second
hybridization (DAPI, Cy3, Cy5) were all shifted in XY by the same offset calculated for the 488 (dodeca) channel, and all eight channels were overlaid.

These registered images were then analyzed as usual, with DAPI segmentation followed by analysis of the six FISH channels, but only in nuclei that had been determined to have sufficient overlap between dodeca signals. The number of FISH signals in each channel was counted, and touching between signals (defined as any pixel overlap between different colors) was assayed in a pairwise fashion for each FISH probe. Pairwise frequencies of nonhomologous clustering were defined as the percentage of nuclei showing touching between the two FISH probes representing each pair of chromosomes. Frequency of homologous clustering was defined as the percentage of nuclei showing fewer FISH signals than expected for a given chromosome based on its copy number in that cell line (for example, if the cell line being examined had three copies of a given chromosome, nuclei with one or two FISH signals for that chromosome would be counted as having homologous touching). Finally, all six FISH signals were treated as a signal color, and the total number of centromere clusters was counted for each nucleus.

Modification of high-throughput FISH screen analysis algorithm for the identification of candidate clustering genes

The previously conducted FISH screen was analyzed by a custom-written MATLAB script that identified nuclei within an image and counted the number of FISH signals within them [71]. For identification of candidate clustering genes, I modified this script in several ways. Firstly, I added features to automatically find input images (rather than providing MATLAB with a list of images to analyze) to facilitate analysis of images from 384-well plates. Secondly, I added measures of distance between nonhomologous signals (that is, between dodeca and 359) to identify factors that change levels of clustering when knocked down. Thirdly, I identified the nuclear periphery as the perimeter of the DAPI signal, and calculated the closest distance between each FISH signal and the nuclear periphery. Measurements such as the mean, minimum and maximum distances between signals or between signals and the periphery were recorded for each cell, regardless of the number of FISH signals in each cell. Distances between FISH signals were then normalized to the approximate diameter of each nucleus. This diameter was calculated based on the area of DAPI signal for each nucleus, assuming a circular shape. Distances between FISH
signals and the nuclear periphery were normalized to nuclear radius, which was calculated based on DAPI area similarly to diameter. Subsequent analysis showed that these distance measures were not correlated with the number of FISH signals in a given nucleus (Supplementary Figure A2.2), indicating that they could be used to assay clustering and positioning regardless of homologous pairing status. The average and standard deviation of these distances were reported across the whole image and across each well of the 384-well plate.

Finally, to facilitate identification of screen hits, I wrote a script to sort control wells from experimental wells on each plate. In total we analyzed four 384-well plates, two plates with different dsRNAs tested in duplicate. Each plate had 157 control wells (a combination of no dsRNA and LacZ dsRNA wells, which were confirmed to not be significantly different from each other), and wells with experimental dsRNAs were compared to the average of control wells on the same plate. A z-score was calculated for each well in each different parameter measured, by taking the difference between the value in that well and the mean of control wells, divided by the standard deviation among control wells. The wells in each plate were then sorted by z-score to identify pro- and anti-clustering candidates. A gene was considered a candidate if its knockdown resulted in a z-score of either more than 1.5 or less than 1.5 in both duplicate wells.
Results

MATLAB-based image registration for alignment of FISH images from multiple rounds of hybridization.

Most FISH experiments are limited in the number of loci that can be simultaneously examined due to the number of fluorophores that can be accurately distinguished. For example, most widefield microscopes have filters that allow them to recognize the emission spectra of DAPI, a DNA dye, as well as several other fluorophores, including Alexa488/GFP, Cy3, and Cy5. This means that in addition to using DAPI to recognize nuclei, one has three additional colors to use for FISH probes. When examining the centromeres of all Drosophila chromosomes simultaneously, we needed four colors, or even five colors when examining male cell lines, which have both the X and Y chromosomes in addition to the three sets of autosomes. Therefore, we decided to use multiple rounds of hybridization, using three sets of FISH probes in the first round, imaging the slide, and washing off the FISH probes, before performing a second round of FISH and re-imaging the same nuclei.

While software-assisted stage navigation is available to assist in finding the same field of cells in subsequent rounds of microscopy, we found that the alignment between the first and second image was never perfect, with differences of up to 3 µm in the position of FISH signals. These differences are quite significant, considering the fact that Drosophila nuclei themselves are 5 to 10 µm in diameter. For analysis of the relative positioning of FISH probes within a nucleus between multiple rounds of hybridization, a much closer alignment was needed. We believe the differences between successive images to be caused by small shifts in the placement of the slide on the microscope stage each time it was imaged, so to circumvent this issue, I wrote a MATLAB script for automated registration of two or more images of the same field of nuclei. To obtain the most accurate alignment, we decided to use a common FISH probe rather than the nuclear outline as the basis for aligning the two images. Therefore, the first hybridization used dodeca (Chr. 3), AACAC (Chr. 2) and 359 (X Chr.), while the second hybridization again used dodeca, but alongside AATAC (Y Chr.) and AATAT (Chr. 4) (Figure 4.3). Using these five probes allowed each chromosome to be uniquely identified, and using dodeca twice provided a common signal between the two images to use for their alignment.
Figure 4.3 Map of the Drosophila genome, showing FISH targets used in this study. Figure reproduced with permission from Dernburg and Sedat, 1998 [73], showing all Drosophila chromosomes and highlighting repeat loci that can be used as FISH targets. Colored boxes have been added to highlight FISH probes used in this study to uniquely identify each chromosome.

Following image segmentation and thresholding to identify nuclei and the dodeca signals within them (see Materials and Methods of Chapter 2), the dodeca signals from the first image were compared to the dodeca signals from the corresponding second image (see Materials and Methods, this Chapter). In cases where the same dodeca signal was confidently identified in both images, the offset between them was measured, and the second image was shifted by the same distance to obtain overlap of the dodeca signals (Figure 4.4). Note that I assumed all nuclei within a given field had shifted by the same amount between images; however, there were cases where individual nuclei had shifted relative to those around them (for example, Figure 4.4C). This might be the result of perturbations caused when the coverslip was removed from the slide in between hybridizations. These nuclei were dropped from the analysis.
The nuclei we did score had at least 40% overlap of pixel area between the dodeca signals from different hybridizations; this threshold was chosen based on visual inspection of the quality of the alignment (Supplementary Figure A2.1), and allowed us to analyze >70% of all nuclei (Supplementary Table A2.1). Once the optimum shift for realignment of the two images had been determined based on analysis of dodeca signals, the images from other channels (AATAC and AATAT) were shifted by the same amount. Quality of the alignment was also assessed by measuring clustering frequencies between each signal and dodeca from the first hybridization, and comparing to the clustering frequencies between each signal and dodeca from the second hybridization (Table 4.1). If the alignment was perfect, these two frequencies should be the same. In our data, there were small variations in these frequencies (in Table 4.1, the average discrepancy between two channels was 3.9%, and the highest discrepancy between any two channels was 6.3%), and this margin of error must be kept in mind when interpreting the results described below. However, the fact that these differences are quite small indicated that the quality of the alignment is fairly high, and that these images can be used to estimate pairwise clustering frequencies between different chromosomes.

**Figure 4.4** Example images of the same nuclei acquired in different FISH experiments, and the use of dodeca signals to align them. The first two columns show images from the first and second hybridization, respectively.
Figure 4.4 (Continued) each of which has multiple FISH signals. The third column shows only the dodeca signals from each hybridization, those from Hyb. #1 colored in green with those from Hyb. #2 colored in purple. The offset between them is the result of small shifts in the position of the slide on the microscope stage each time it is imaged. The fourth column shows the alignment of dodeca signals following the use of our image registration algorithm, while the fifth column highlights the nuclei chosen for future analysis. (A) An image from Clone 8 cells, in which most nuclei had a single dodeca FISH signal, indicating that all the copies of Chromosome 3 were paired in those cells. (B) An image from S2R+ cells, in which most nuclei had more than one dodeca FISH signal, but the correct corresponding signals were still identified for realignment. (C) Another image from S2R+ cells, but in this case some nuclei had shifted relative to those around them and could not be properly aligned. As the image in the fifth column highlights, only the nuclei showing a good amount of overlap following registration were scored in further analyses. All scale bars = 10 µm.

Table 4.1 Assessment of image registration based on clustering frequencies. Data shown is for Clone 8 cells (n=255 nuclei). For each FISH signal, its clustering frequency with dodeca from the first hybridization was compared to its clustering frequency with dodeca from the second hybridization. Clustering frequency is defined as the percentage of nuclei in which two different FISH signals are touching in overlaid images. Clustering between dodeca signals from the two hybridizations was always 100%. While there are small discrepancies in each row (mean difference = 3.9%) and possible reasons for this are discussed below, the clustering frequencies calculated for the two dodeca images were fairly similar, suggesting accurate alignment.

<table>
<thead>
<tr>
<th>Locus (Chr)</th>
<th>% clustered with dodeca (Hyb #1)</th>
<th>% clustered with dodeca (Hyb #2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>359 (X)</td>
<td>49.8%</td>
<td>50.2%</td>
</tr>
<tr>
<td>AATAC (Y)</td>
<td>37.2%</td>
<td>40.0%</td>
</tr>
<tr>
<td>AACAC (2nd)</td>
<td>33.3%</td>
<td>39.6%</td>
</tr>
<tr>
<td>AATAT (4th)</td>
<td>78.4%</td>
<td>84.3%</td>
</tr>
</tbody>
</table>
FISH targeting pericentric regions can be used to assay clustering, revealing that clustering of nonhomologous centromeres in Drosophila is stochastic.

After realignment of FISH images from two rounds of hybridization, we were able to obtain images in which each the pericentric heterochromatin of each chromosome was uniquely identified (Figure 4.5). This analysis was done in two different Drosophila cell lines, Clone 8 and S2R+, which were originally derived from different tissues and have different karyotypes as well as different levels of homolog pairing (see Chapter 2). Specifically, S2R+ cells are embryonic in origin, while Clone 8 cells were derived from wing discs of third instar larvae. Our studies have shown that Clone 8 cells have higher percentages of nuclei showing a single FISH signal, indicating higher levels of homolog pairing (for example, see Chapter 2). Additionally, by performing FISH on metaphase spreads, I have confirmed that our S2R+ cells are stably aneuploid, having two copies of the X, three copies of chromosome 2, four copies of chromosome 3, and no Y chromosome (Figure 4.6, left). Our Clone 8 cells are diploid and male, having two copies of each autosome as well as an X and a Y (Figure 4.6, right). Chromosome 4 in Drosophila, which is quite small (~5 Mb), is often lost in metaphase spreads and therefore difficult to reliably karyotype, so a diploid copy number of 2 was assumed for both cell lines.

![Figure 4.5 Example of an interphase Clone 8 nucleus in which each chromosome has been uniquely identified.](image)

Overlay highlights that pericentric heterochromatin of each chromosome is in close proximity, indicative of centromere clustering. This particular nucleus was scored as having two clusters (see final panel). Blue line = DAPI perimeter; scale bar = 5 µm.
Figure 4.6 Metaphase spreads of Drosophila S2R+ and Clone 8 cells. Note that these images are from a single FISH experiment (not multiple hybridizations) so only three different chromosomes can be simultaneously labelled. S2R+ cells have two chromosomes labelled with 359 (X), three chromosomes labelled with AACAC (2nd), and four chromosomes labelled with dodeca (3rd). Clone 8 cells have one chromosome labelled with 359 (X) and two chromosomes each labelled with AACAC and dodeca (2nd and 3rd). The unlabeled chromosome in Clone 8 cells has been confirmed in other experiments to be the Y chromosome. The copy number of chromosome 4 is difficult to estimate from spreads (the 4th is the small dot chromosome present near the pair of 2nd chromosomes in the Clone 8 spread). Scale bar = 5 um.

To compare how well FISH targets adjacent to the centromere reflected centromere positioning and clustering, I ignored the color of each signal, treating them all as centromere markers, and counted the number of clusters that were observed in each nucleus (Figure 4.5, last panel). These results show that the FISH signals do cluster, giving an average of 1.6 clusters per nucleus (n=255 nuclei) in Clone 8 cells, in which 5 different FISH targets were assayed (Chr. 2, 3, 4, X and Y), and 2.2 clusters per nucleus (n=103 nuclei) in S2R+ cells, in which 4 different FISH targets were assayed (same as Clone 8 except for the Y). The small number of clusters, compared to the number of FISH targets assayed, indicates that pericentric regions do show nonhomologous clustering in these cells. The number of clusters observed by FISH are slightly lower than those observed by immunofluorescence targeting centromeric proteins such as CID (the Drosophila centromeric histone H3 variant), which allows direct visualization of the centromere. CID staining in these cells revealed an average of 2.2 signals per nucleus in Clone 8.
cells (n=372 nuclei) and 3.4 signals per nucleus in S2R+ cells (n=303 nuclei). The differences might be explained by the fact that some of our FISH targets were quite large (the size of the 359 repeat is estimated to be ~11 Mb [73]) and the large signals produced might be more likely to overlap with other FISH signals, compared to the relatively small signals obtained from immunofluorescence for CID (compare the CID signals seen in Figure 4.2 to the FISH signals in Figure 4.5). Therefore, FISH targeting pericentric heterochromatic regions provides a slight over-estimate of centromere clustering, but is still sensitive to relative differences between cell lines. Therefore, we are confident in using FISH results as a proxy for centromere positioning.

Given these findings, we then took the identity of each FISH probe into consideration and examined pairwise associations between them, with the aim of determining whether, beyond the pairing of homologs, centromeres show any preferential clustering. For example, the centromere of chromosome 2 might associate with the centromere of chromosome 3 more often than it does with that of the X chromosome. For nonhomologous associations, we scored nuclei where there was any overlap between the FISH signals of the two chromosomes being examined (see Materials and Methods). For homologous associations, we scored nuclei where there were fewer FISH signals than might be expected given the copy number of that chromosome (for example, S2R+ cells have three copies of Chr. 2/AACAC, so any nuclei with 1 or 2 AACAC FISH signals would be scored as having homologous associations of chromosome 2). For chromosome 4, any nuclei with 1 AATAT FISH signal were scored as having homologous associations. Note that if the copy number of chromosome 4 is actually higher than 2, this analysis would provide an under-estimate of the amount of homologous pairing of the 4th chromosome occurring in these cells. Finally, all of the FISH probes used in these studies were extremely efficient (Supplementary Table A2.2), meaning that we can be confident the decreases in the numbers of signals represented homologous chromosome pairing, rather than inefficient labeling.

Having analyzed the pairwise associations between centromeres in this manner, we came up with estimates of the clustering frequencies of all the different chromosomes in Clone 8 cells (Table 4.2). As these results show, the frequencies of homologous touching (78 – 92%) are generally higher than those of nonhomologous touching (31 – 52% for all chromosomes except Chr. 4), indicating that while all centromeres do cluster, there is still a preference for homologous pairing. Notably, this observation does not hold for chromosome 4, which shows nonhomologous
association frequencies as high as its frequency of homologous pairing. These results might be explained by the fact that chromosome 4 is almost entirely heterochromatic. For example, given the high proportion of heterochromatin, chromosome 4 may be segregated to the heterochromatic compartment more efficiently and cluster more strongly; alternatively, the presence of euchromatic arms that are paired may be necessary for the homologous preferences observed at the centromeres of other chromosomes.

Table 4.2 Pairwise clustering frequencies in Clone 8 cells. Numbers along the diagonal represent homologous clustering, while those off-diagonal represent nonhomologous clustering (n=255 nuclei, percentages for dodeca are those from the first hybridization).

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>-</td>
<td>52%</td>
<td>31%</td>
<td>50%</td>
<td>88%</td>
</tr>
<tr>
<td>Y</td>
<td>52%</td>
<td>-</td>
<td>31%</td>
<td>37%</td>
<td>89%</td>
</tr>
<tr>
<td>2nd</td>
<td>31%</td>
<td>31%</td>
<td>78%</td>
<td>33%</td>
<td>41%</td>
</tr>
<tr>
<td>3rd</td>
<td>50%</td>
<td>37%</td>
<td>33%</td>
<td>92%</td>
<td>78%</td>
</tr>
<tr>
<td>4th</td>
<td>88%</td>
<td>89%</td>
<td>41%</td>
<td>78%</td>
<td>82%</td>
</tr>
</tbody>
</table>

Considering the nonhomologous associations of other chromosomes, there seems to be no preference for specific pairwise interactions beyond the preference for homologous pairing. For example, the X chromosome clusters with chromosome 3 in 50% of Clone 8 nuclei, comparable to its clustering frequency with the Y chromosome, which is 52%. These results indicate that nonhomologous centromere clustering is somewhat stochastic, with centromeres being equally likely to cluster with any of the other chromosomes. Notably, chromosome 2 is somewhat of an exception, with lower interaction frequencies across the board; the nonhomologous clustering frequencies of AACAC with other chromosomes range from 31 – 41% in Clone 8 cells, while for dodeca the range is from 33 – 78%. This could indicate that there is something distinct about the organization of the pericentromere of chromosome 2. Interestingly, the AACAC probe on chromosome 2 is somewhat more distally located from its centromere compared to some of the other FISH probes used (see Figure
which could also explain why it appears to interact less with the pericentric heterochromatin of other chromosomes.

Notably, S2R+ cells show some of the same trends as just described for Clone 8, including higher homologous interaction frequencies than nonhomologous, and nonhomologous interaction frequencies that are generally comparable between different pairs of chromosomes (Table 4.3). Chromosome 4 shows a lower homologous association frequency in S2R+ than Clone 8, but as mentioned above, this may be caused by inaccurate karyotyping of chromosome 4 in this aneuploid cell line. Interestingly, some of the chromosome-specific trends for nonhomologous associations observed in Clone 8 cells are not as strong or absent in S2R+ cells. For example, chromosome 2 only has slightly lower nonhomologous interaction frequencies compared to other chromosomes. Also, the nonhomologous interaction frequencies for chromosome 4 are not markedly higher than those for other chromosomes, except for the interaction frequency between chromosome 4 and the X (89%), which is the most frequent nonhomologous interaction observed. These results might could indicate cell-type specific differences in the organization of pericentric heterochromatin and/or centromeres.

**Table 4.3** *Pairwise clustering frequencies in S2R+ cells.* Numbers along the diagonal represent homologous clustering, while those off-diagonal represent nonhomologous clustering (n=103 nuclei, percentages for dodeca are those from the first hybridization).

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt;</th>
<th>3&lt;sup&gt;rd&lt;/sup&gt;</th>
<th>4&lt;sup&gt;th&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>89%</td>
<td>51%</td>
<td>61%</td>
<td>89%</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>51%</td>
<td>93%</td>
<td>62%</td>
<td>52%</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>61%</td>
<td>62%</td>
<td>98%</td>
<td>59%</td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>89%</td>
<td>52%</td>
<td>59%</td>
<td>59%</td>
</tr>
</tbody>
</table>
Reanalysis of data from high-throughput FISH screen done in Drosophila cells identifies genes involved in clustering and positioning of pericentric regions.

Having determined that FISH can be used to study centromere positioning, we realized that the results of a previously published FISH-based screen conducted by our laboratory could be used to identify genes that either promote or antagonize the nonhomologous clustering of pericentric heterochromatin as well as of centromeres. The original screen aimed to identify genes that promoted or antagonized somatic homolog pairing [71]. Briefly, Drosophila Kc167 cells were grown in 384-well plates that had been treated with a dsRNA library covering the entire Drosophila genome, and following four days of RNAi, prepared for high-throughput FISH targeting dodeca and 359 (see Figure 4.7, and ref. [71]). The resulting images were analyzed by a custom-written MATLAB script for the number of FISH signals per locus and the distances between them. The original analysis did consider the possibility that nonhomologous clustering might influence the frequencies of homolog pairing observed after RNAi knockdowns of different genes, and therefore, used two methods to estimate the amount of clustering. Firstly, the average pairwise distances between all signals, dodeca and 359, were calculated for each nucleus. By this assay, 25% of anti-pairing genes also decreased the average distance between signals when they were knocked down by RNAi, indicating that they might also be anti-clustering genes [71]. Secondly, nuclei having one signal for both dodeca and 359 were assayed for the frequency of touching between the two signals. Surprisingly, by this assay, none of the anti-pairing genes affected the frequency of nonhomologous touching [71]. These results indicated that the candidate anti-pairers identified by the screen did indeed have effects on homologous pairing independent of any effect on nonhomologous clustering.

While sufficient to rule out that the results of the screen were driven by the effects of nonhomologous clustering, the analysis described above was insufficient to identify genes that regulate clustering. Firstly, the average pairwise distance between all signals includes homologous distances, and therefore, a decrease in the average could be driven by increased pairing (at least, a decreased distance between homologs) in addition to increased clustering. Secondly, the analysis of touching between dodeca and 359 signals was only done in nuclei having one signal for both dodeca and 359, indicating that all copies of chromosome 3 and all copies of the X chromosome were homologously paired. Limiting the analysis in this way simplified the possible outcomes, because
when there are multiple FISH signals at either locus, it is possible that none, some, or all of these FISH signals could show nonhomologous clustering, and these different outcomes need to be weighted accordingly.

Figure 4.7 Experimental design used in high-throughput FISH screen to identify candidate pairing genes. Figure courtesy of Eric Joyce, based on ref. [71].

Keeping these different possibilities in mind, we decided to reanalyze the images taken during the first screen with a new MATLAB script. Specifically, I added criteria to examine nonhomologous distances between FISH signals, and to determine the minimum, maximum, and average nonhomologous distances in each nucleus. These measures allow the effect of a gene on clustering to be examined regardless of its effect on pairing. Additionally, I adapted the script to allow examination of the position of centromeres within the nucleus. As described in the introduction, in Drosophila cells, centromeres are clustered at the nucleolus [31], as opposed to in other organisms where centromere clustering takes place at the nuclear periphery [3,8–10]. It is possible that genes disrupting clustering also disrupt nucleolar localization, as has been shown for NLP [31], or that there are genes that affect centromere localization without influencing clustering. To parse these different contributions, I determined the shortest distance between each FISH signal and the nuclear periphery. While we have no way to directly assay positioning relative to the nucleolus in these images, one would expect that nucleolar tethering would restrict their nuclear positioning, while if centromeres are not tethered to the nucleolus, the distance to the periphery will vary more in a population of cells. Therefore, measuring distance to the periphery allows us to both identify genes that
influence the positioning of pericentric regions, and to make inferences regarding the relationship between clustering and positioning.

Identification of candidate clustering genes.

After analysis of several different parameters, we decided to focus on the average distance between nonhomologous FISH signals as the main criteria for the identification of genes regulating the clustering of pericentric heterochromatin. We are also confident that this measure reflects the amount of centromere clustering present in a given nucleus, because using this measure to sort the screen data identified several genes previously identified as being required for centromere clustering in flies. These included genes encoding Cal1, the functional ortholog of HJURP in humans that loads CID at centromeres and is required for centromere clustering in Drosophila meiosis [74], and NLP, which is required for clustering in Drosophila S2 cells [31]. Therefore, the hits we describe as “clustering genes” that regulate the clustering of pericentric heterochromatin can also be considered candidates for the regulation of centromere clustering.

Additional groups of genes that were found to regulate clustering are described below. In the analysis presented here, we focus on genes that were identified in the primary screen for pairing factors (not all of which were ultimately validated as candidate pairing promoters or anti-pairers), our reasoning being that genes that influence centromere clustering will most likely alter the number of FISH signals observed at individual loci by affecting clustering of homologous centromeres. This analysis can easily be extended genome-wide in the future, to determine if there are additional genes that only affect nonhomologous clustering. The candidates presented here are the results from screening dsRNAs targeting 352 different genes, tested in duplicate with FISH targeting the dodeca and 359 pericentric heterochromatic repeat loci on the 3rd and the X chromosomes, respectively. The average distance between dodeca and 359 signals, normalized to the approximate nuclear diameter, was calculated for each cell and then for an entire well, and a z-score was assigned for this distance relative to the control wells present on the same plate (Figure 4.8; see Materials and Methods for more details). Ribosomal genes and histone proteins were dropped from the list, as their effects may be indirect. Using these cut-offs, we identified 54 dsRNAs that consistently increased the distance between nonhomologous signals (z-score ≥ 1.5), representing genes that cause
disruption of clustering when knocked down. These are candidate pro-clustering genes, which may be involved in establishment or maintenance of clustering. We also identified 23 dsRNAs that consistently decreased the distance between nonhomologous signals (z-score ≤ -1.5), representing genes that cause increased clustering when knocked down. These are candidate anti-clustering genes, which may be involved in the dispersal of clustering or some other aspect of its negative regulation.

![Figure 4.8 Distribution of nonhomologous distances obtained in the screen.](image)

Data shown for 352 different dsRNAs tested in duplicate and 628 control wells. The control range was defined as the average distance among all controls, ±1.5 standard deviations. Pro-clustering candidates increased the distance between dodeca and 359 when knocked down, while anti-clustering candidates decreased the distance between dodeca and 359 when knocked down.

Some examples of genes in both the pro- and anti-clustering categories are shown in Figure 4.9, including sample FISH images and bar graphs showing the distributions of nonhomologous distances (359 to dodeca distance) and distances to the nuclear periphery (359 to periphery distance) for control cells and for knockdowns of different genes. We provide further descriptions of the candidate genes below. The putative functions of candidates were assigned based on gene ontology terms found on Flybase (www.flybase.org), and where possible, these terms were used to sort the candidates into different functional groups (Figure 4.10). For candidates with multiple roles, the
putative function most relevant to chromosome organization and/or the cell cycle was used to assign candidates to different groups.

<table>
<thead>
<tr>
<th>Candidate</th>
<th>DAPI 359 dodeca</th>
<th>FISH only</th>
<th>359 to dodeca distance</th>
<th>359 to periphery distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td><img src="controls.png" alt="Image" /></td>
<td><img src="controls.png" alt="Image" /></td>
<td>Mean = 0.295</td>
<td>Mean = 0.332</td>
</tr>
<tr>
<td>(n=1301)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1mb RNAi</td>
<td><img src="s1mb.png" alt="Image" /></td>
<td><img src="s1mb.png" alt="Image" /></td>
<td>Mean = 0.412**</td>
<td>Mean = 0.281**</td>
</tr>
<tr>
<td>(n=1218)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nlp RNAi</td>
<td><img src="nlp.png" alt="Image" /></td>
<td><img src="nlp.png" alt="Image" /></td>
<td>Mean = 0.336**</td>
<td>Mean = 0.298**</td>
</tr>
<tr>
<td>(n=1333)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cal1 RNAi</td>
<td><img src="cal1.png" alt="Image" /></td>
<td><img src="cal1.png" alt="Image" /></td>
<td>Mean = 0.326**</td>
<td>Mean = 0.348</td>
</tr>
<tr>
<td>(n=1096)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cap-H2 RNAi</td>
<td><img src="cap-h2.png" alt="Image" /></td>
<td><img src="cap-h2.png" alt="Image" /></td>
<td>Mean = 0.257**</td>
<td>Mean = 0.361*</td>
</tr>
<tr>
<td>(n=794)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup358 RNAi</td>
<td><img src="nup358.png" alt="Image" /></td>
<td><img src="nup358.png" alt="Image" /></td>
<td>Mean = 0.268*</td>
<td>Mean = 0.404**</td>
</tr>
<tr>
<td>(n=885)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.9** Candidate pro- and anti-clustering genes alter the distances between nonhomologous signals. Example nuclei are shown for control cells and for knockdowns of candidate pro- and anti-clustering genes. The
Figure 4.9 (Continued) images shown were taken during the original FISH screen for genes involved in homolog pairing [71]; the panel on the left shows DAPI signals as well as FISH signals, while those on the right show just FISH signals and the outline of each nucleus (scale bars = 5 μm). The bar graphs show distributions of distances obtained in a population of nuclei, with each nucleus being an individual data point; the x-axis shows normalized distance, while the y-axis shows the proportion of nuclei in which that distance was observed. Graphs represent pooled results from two replicate wells. The bar graphs on the left show the average nonhomologous distance between 359 and dodeca, normalized to nuclear diameter, while those on the right show the minimum distance between 359 and the nuclear periphery, normalized to nuclear radius. Bar graphs for control cells are in blue, while those for candidate pro-clustering genes are in red and those for candidate anti-clustering genes are in green. The mean of the distribution is shown on each bar graph; significant changes from control cells are marked with asterisks. Those marked with two asterisks represent a z-score $> +2$ or $< -2$ in each replicate, while those marked with a single asterisk represent a z-score $> +1.5$ or $< -1.5$ in each replicate.
Figure 4.10 Summary of candidate genes identified in our screen that regulate clustering of pericentric heterochromatin and/or centromeres. Candidate pro-clustering genes are highlighted in red while candidate anti-clustering genes are highlighted in green.
**Candidate pro-clustering genes.**

Here, we describe pro-clustering factors, which caused increased nonhomologous distances/decreased clustering when knocked down by RNAi. In this category we identified 54 candidate genes, including centromere and kinetochore proteins, nucleolar proteins, and known regulators of the cell cycle, including many mitotic genes.

*The kinetochore, and associated proteins.* In addition to genes that were known to affect clustering in Drosophila, we also identified genes that might have been expected to be important for centromere clustering given findings in other organisms. For example, in fission yeast, kinetochore proteins are important for centromere clustering [33,34]. In our screen, we identified several kinetochore components as being necessary for centromere clustering in Drosophila, including the inner kinetochore protein Cenp-C. Cenp-C is part of the constitutive centromere-associated network (CCAN) which is found at centromeres throughout the cell cycle, and is essential for localization of other centromere-associated proteins [75]. We also identified members of the KNL1/Mis12/Ndc80 (KMN) outer kinetochore network, including Ndc80, Spc25, and Kmn2 [76], which form the part of the kinetochore where spindle attachment occurs during mitosis [77,78]. Another group of hits associated with the kinetochore was the chromosomal passenger complex or CPC, which includes Aurora-B kinase, Borealin, Deterin (also known as Survivin), and Incenp [79–81]. Three out of four CPC subunits (Aurora-B, Borealin, Deterin) were identified as consistently causing disruptions in clustering when knocked down. The CPC and its catalytic subunit Aurora-B have a well-described role in monitoring the connection of kinetochores with microtubules in mitosis and eliminating incorrect attachment [79–81]. These results highlight the importance of centromere components, and the attachment of centromeres to microtubules, for the clustering of pericentric heterochromatin.

*Microtubules, actin filaments, and proteins involved in cytoskeleton regulation.* Consistent with the idea that microtubule attachment is important for clustering, we identified genes encoding both α- and β-tubulin proteins, genes required for proper spindle assembly at the microtubule organizing center (MTOC) including dynein (Dhc64C) [82] and Abnormal spindle (asp) [83–85], as well as the microtubule-binding protein Mars [86–89] and its associated proteins Protein-phosphatase 1 (Pp1-87b) [87,90] and tousled-like kinase (Tlk) [91]. Other factors directly or indirectly associated with spindle organization that were identified as pro-clustering factors include three
rows (thr) and eIF3ga. Other cytoskeletal proteins that were identified include Actin-related protein 1 (Arp87C), and proteins that bind and/or regulate the actin cytoskeleton including Rho1 (a Rho GTPase) and twinstar (tsr).

**Nucleolar proteins.** Since centromeres in Drosophila cluster at the nucleolus and require NLP to do so, we might expect additional nucleolar components to also be required for clustering. Consistent with this, we identified Non1 (novel nucleolar protein 1) as well as Ns1 (nucleostemin), proteins associated with the nucleolus involved in spindle assembly as well as other functions [92,93].

**Cell cycle regulators.** We also identified several different types of cell cycle regulators that are required for wild-type levels of centromere clustering. We identified several components of the anaphase promoting complex or APC, including Cdc16, Cdc27, shattered (shtd), and imaginal discs arrested (ida). Additional regulators of the metaphase to anaphase transition were also identified as clustering genes, including vihar (vih) and Spindly [94]. We also identified known regulators of cytokinesis including pavarotti (pav) [95], polo kinase, which also has other functions in cell cycle regulation [96,97], and a kinase, CG7236, that was identified in a screen for genes involved in cytokinesis [98]. Additional cell cycle regulators important for clustering include regulators of the G2-M transition Cyclin B and Not1.

**Negative regulators of condensin II, condensin I.** Four known negative regulators of condensin II, including three components of the SCF-ubiquitin ligase complex, Slmb [71,99], Lin19, and SkpA, as well as Casein Kinase 1 alpha [100], were identified as being required for clustering. This is consistent with the fact that condensin II specific proteins were identified as being negative regulators of clustering both in this screen (see further discussion below) and in previously published work [71,101]. Interestingly, a condensin I specific protein, barren, and a protein present in both condensin I and II, Sme2, were both identified as candidate pro-clustering factors. This could indicate that condensin I and II play opposite roles in the organization of centromeres.

**Other factors.** Additional factors found to be required for clustering were the chromatin remodeling protein E(bx), the heat shock protein Hsc70Cb, and several other factors whose functions are unknown or for which the connection to chromatin or chromosome organization is unclear. For a full list, see Figure 4.10.
Candidate anti-clustering genes.

Here, we describe anti-clustering factors, which caused decreased nonhomologous distances/increased clustering when knocked down by RNAi. We obtained fewer candidate genes in this category (Figure 4.10), suggesting that there are relatively few pathways that disrupt as opposed to those that establish or maintain the clustering of centromeres and pericentric regions. Additionally, many of the candidates were associated with transcription, metabolism, and proteosomal degradation, suggesting their effects may be indirect. Candidate genes that might play a more direct role in clustering are described in more detail below.

Condensin II. Condensin II plays a well-characterized role in antagonizing homolog pairing [71,102] and in promoting the formation of chromosome territories while causing the dispersal of heterochromatin [101]. Consistent with this, we identified Cap-H2, a component of the condensin II complex, as a hit in this screen. Therefore, it is likely that condensin II plays a role in the dispersal of centromere clusters, perhaps at a specific stage in the cell cycle. As described above, negative regulators of condensin II have the opposite phenotype, promoting clustering, highlighting condensin II as a key player in the regulation of centromere clustering.

Cell cycle regulators. We identified the cell cycle regulator CycA, which regulates S-phase as well as the G2-M transition, MBD-R2, a protein implicated in the G2 DNA damage checkpoint [103], and Hel25E, an RNA helicase that regulates splicing and is also implicated in mitotic spindle organization [104]. These findings highlight that, like in the establishment of clustering, the cell cycle may play an important role in the dispersal of clustering.

Nuclear pore complexes. Interestingly, we identified the nuclear pore protein Nup358 as well as a predicted nuclear pore component, CG4673, as anti-clustering factors. Since the nuclear periphery is a region where centromeres cluster in many other organisms, it is possible that dispersal of centromere clustering in Drosophila involves release of nucleolar tethering and localization near the nuclear periphery and/or nuclear pores instead. This possibility is discussed more below.
Possible link to telomere clustering. One candidate anti-clustering gene, effete (also known as UbcD1), is an E2 ubiquitin ligase involved in several different processes in fly development. Interestingly, this gene has also been implicated in the proper organization of telomeres [105]. Drosophila UbcD1 mutants often show telomere-telomere fusions, which are thought to form when telomeres are closely juxtaposed in interphase [105]. This result suggests that some genes involved in antagonizing centromere clustering also act to antagonize telomere interactions, or that the two aspects of chromosome organization are related in some other way.

Relationship between centromere clustering and nuclear positioning.

As mentioned above, in Drosophila cells, centromeres are generally clustered at the periphery of the nucleolus, and knockdown of NLP, in addition to disrupting clustering, was also found to disrupt this nucleolar tethering [31]. We were curious to examine among our candidate genes whether all knockdowns that disrupt clustering also disrupt positioning, or if there might be specific genes that affect one phenotype but not the other. First we examined control wells from the screen, to look at the relationship between clustering and positioning in wild-type cells. Specifically, we looked at the relationship between average nonhomologous distances and the minimum distance to the nuclear periphery for either dodeca or 359. We found a modest negative correlation between these two parameters ($R^2=0.4618$ for dodeca and $R^2=0.4330$ for 359) indicating that in control cells, when dodeca and 359 are further apart from each other, reflecting centromeres that are not clustered, they tend to be closer to the nuclear periphery (Supplementary Figure A2.3). This finding supports the hypothesis that clustering status may be related to nuclear positioning. The relatively modest nature of the correlations may be because we cannot control for nucleolar positioning in these data; it is possible that centromeres could be clustered at the nucleolus and still be close to the nuclear periphery. However, the fact that a trend is still visible suggests that perinucleolar clustering tends to occur more centrally within the nucleus.

We then examined hits which had been identified as candidate genes that regulate centromere clustering. Interestingly, knockdowns of either pro- or anti-clustering genes disrupted the correlation between nonhomologous distances and distance to the periphery, resulting in much lower $R^2$ values than those found in control cells (Supplementary Figure A2.3). This may be explained by the fact that when clustering levels are altered by RNAi,
nuclear positioning of centromeres is shifted away from the range observed in control cells. Importantly, the direction of this shift was different for candidate pro- and anti-clustering genes. Cells in which pro-clustering genes were knocked down showed FISH signals that were on average closer to the nuclear periphery than those observed in control cells (see Slmb, Nlp and Cal1 in Figure 4.9). However, the total distribution of distances in these knockdowns was not significantly shifted towards the nuclear periphery, indicating that there were still many FISH signals located in the center of the nucleus. These results could reflect a randomization of centromere positioning once clustering is lost, including an ability to be closer to the periphery than in control cells. Consistent with this idea, in knockdowns of the pro-clustering genes (all of which had a z-score of $\geq +1.5$ for nonhomologous distances), the average z-score for distances to the nuclear periphery was +0.44 for dodeca and -0.18 for 359. This indicates that knockdowns that disrupt clustering do not significantly influence distance from the nuclear periphery in one direction or the other.

A different situation was observed among the 23 candidate anti-clustering genes, which cause increased clustering when knocked down. In these knockdowns, the distances to the periphery tended to be larger, indicating that FISH signals were located non-randomly far from the nuclear periphery (see Cap-H2 and Nup358 in Figure 4.9). Furthermore, this shift towards the interior was significantly different compared to control cells. In knockdowns of anti-clustering genes (all of which had a z-score of $\leq -1.5$ for nonhomologous distances), the average z-score for distances to the periphery was +2.95 for dodeca and +2.83 for 359. These results indicate that when candidate anti-clustering genes are knocked down and clustering is increased, the distance between FISH signals and the nuclear periphery is almost 3 standard deviations higher than the mean distance observed in control cells. Consistent with this finding, 18 out of the 23 anti-clustering candidates were also hits for increasing the distance of either dodeca or 359 from the nuclear periphery (hits for nuclear positioning were determined in the same way as for clustering hits, see Materials and Methods). Combined with the results obtained for candidate pro-clustering genes, these data indicate that an increase in clustering is associated with a shift away from the nuclear periphery, whereas a decrease in clustering results in distances from the nuclear periphery that are more variable.
Discussion

In this chapter, we present studies of centromere clustering in Drosophila cells. Using five-color FISH experiments, we have demonstrated that FISH targeting pericentric heterochromatin can be used as a proxy for centromere positioning. We believe that the image registration algorithm described here will also prove useful in future experiments, either in place of or as a supplement to software-assisted stage navigation. Using a common FISH signal as the basis for image registration both allows a more accurate degree of alignment and provides some amount of quality control, demonstrating that nuclear architecture was not perturbed between the two FISH experiments, at least in the region surrounding the locus being examined.

We note that there is room for improvement in the alignment between images (Table 4.1), and there are two strategies to address this issue. Firstly, our requirement that two signals have pixel overlap in order to be scored as clustered may be too strict, given the small variabilities there may be in the quality of the alignment from nucleus to nucleus, so it may be better to use a distance-based cutoff instead. Secondly, it is possible (though more computationally intensive) to align each nucleus individually rather than the whole image at once. Our data suggest that within a certain image, most nuclei have shifted by the same amount, consistent with the shift being caused by changes in the position of the slide on the microscope stage, but there may be variations between nuclei of 1 – 2 pixels, which corresponds to approximately 108 – 216 nm, and these variations could explain the differences in clustering frequencies estimated using dodeca signals from the two different images. Even with this caveat in mind, however, we can make certain observations about the patterns of centromere clustering in these cells.

By uniquely identifying each chromosome, we demonstrate that there is no pairwise preference in centromere clustering beyond that for homologous association. There were two outliers in this data, namely, chromosome 2, which appeared to interact less frequently with all other chromosomes than with itself, and chromosome 4, which interacts as frequently with other chromosomes as with itself, at least in Clone 8 cells. Regarding chromosome 2, since AACAC is located closer to euchromatin than any of the other FISH probes used (see Figure 4.3), it may be less reliable as a proxy for the position of the centromere. In future, it would be interesting to repeat the experiment with a FISH probe targeting the Responder locus (Rsp), another heterochromatic
repetitive region on chromosome 2 that is located closer to the centromere. Regarding chromosome 4, it could be that the high nonhomologous interaction frequencies could be explained by the fact that this chromosome is mostly heterochromatic. Notably, the high proportion of heterochromatin is also true of the Y chromosome, which did not show such high nonhomologous interaction frequencies, so perhaps the small size or some other unique feature of chromosome 4 drives its interactions with other chromosomes. It would be interesting to determine in future experiments if chromosome 4 plays any special role in the nucleation of centromere clustering, perhaps by analyzing cell lines with different copy numbers for chromosome 4, which has frequently been gained or lost in Drosophila cultured cells [106].

We also present results from reanalysis of images taken during a previous high-throughput FISH screen [71], to identify candidate genes that regulate the clustering of pericentric heterochromatin and of centromeres. By modifying the image analysis algorithm initially used in the screen to specifically examine distances between nonhomologous FISH signals, we analyzed images taken in a primary screen of 352 genes and identified 54 candidate pro-clustering factors and 23 candidate anti-clustering factors. These included genes that had previously been implicated in the regulation of centromere clustering in Drosophila [31,71,74,101] as well as novel candidates. Over half of the candidate pro-clustering genes have functions in mitosis, highlighting the important role of mitosis in establishing interphase chromosome positioning and supporting early hypotheses put forward by Rabl [44]. Additionally, there may be pathways that maintain or disrupt centromere clustering during interphase, such as condensin II and its regulators [71,99–101]. Future work to determine when in the cell cycle the different candidate genes identified have their effects on clustering will help sort out these different pathways.

Investigation into the timing of their effects is just one of many follow-up experiments that will be done with the candidate genes identified here. Since these candidates were identified using a FISH-based screen, confirmation of their effect on centromere clustering will be done using immunofluorescence for the centromeric protein CID. Additionally, it will be interesting to determine if any of these candidates also affect telomere clustering. Telomeres can be identified in Drosophila cells by immunofluorescence for telomeric proteins such as HipHop or HOAP [107–109]. Since the Rabl configuration of centromeres and telomeres is thought to be established in anaphase, genes that affect both types of clustering may be involved in their establishment, while
genes that affect centromeres but not telomeres may indicate stage-specific regulators, or those specific to different parts of chromosomes.

Additionally, in future experiments, we can combine examination of centromere positioning with examination of nucleolar positioning, by staining for nucleolus specific proteins such as fibrillarin. Our results indicate that increased levels of clustering are correlated with localization away from the nuclear periphery. Studies in which the position of the nucleolus is also known will help determine whether this increased clustering occurs at the nucleolus or whether there are other locations where centromeres cluster following RNAi, as well as to clarify any distinctions between position relative to the nuclear periphery and position relative to the nucleolus in cells where clustering has been lost. Furthermore, our study could be expanded in future to identify genes that affect positioning independently of clustering, potentially identifying pathways that affect radial positioning not just of centromeres but of the whole nucleus.

An interesting result from the screen was that condensin I and II have opposite effects on centromere clustering; knockdown of the condensin I component Barren disrupted clustering, while knockdown of the condensin II component Cap-H2 enhanced clustering. Smc2, a component of both condensin complexes, had a phenotype similar to Barren, possibly indicating that condensin I is more sensitive to Smc2 depletion than condensin II, or that the role of condensin I in centromere clustering is epistatic to that of condensin II. While these hits need to be confirmed before we can draw any conclusions, the possibility that the two complexes contribute to centromere clustering in different ways is especially interesting in light of their localizations over the course of the cell cycle. In many organisms, condensin II is present in the nucleus throughout interphase, while condensin I is present mainly during mitosis [110–112]. Therefore, the different roles of condensin I and II might reflect the differences in centromere organization in interphase and in mitosis. These ideas will be discussed more in the next chapter.

In summary, this chapter presents studies using FISH to assay the clustering of pericentric heterochromatin and of centromeres in Drosophila cells, and to identify potential candidate genes that regulate clustering. Additional experiments to verify and further characterize these candidate genes are in progress. Implications of our findings, and their overall relevance to studies of nuclear organization, will be further discussed in the next chapter.
Acknowledgments

The work described in this chapter was supported by a grant and an NIH Director’s Pioneer Award to Chao-ting Wu from NIH/NIGMS (RO1GM085169 and 5DP1GM106412) and a Ruth L. Kirschstein National Research Service Award to Eric Joyce from NIH/NCI (F32CA157188). Additionally, I was supported by the Harvard Medical School Department of Genetics and by a training grant to the Harvard Medical School Program in Genetics and Genomics from NIH/NIGMS (T32GM096911). I am grateful to my collaborator in this work, Eric Joyce, for carrying out FISH experiments and for sharing data with me, to Tiao Xie for his computational expertise and advice, and to the curators of Flybase (www.flybase.org), as this resource greatly facilitated the description of candidate genes identified in our screen.

References


Chapter 5

Discussion

Parts of this chapter are based on:

Overview

The work presented in this dissertation was motivated by an interest in interphase nuclear organization, the mechanisms that drive it, and the connections between chromosome organization in interphase and in mitosis. I have described three major findings in these areas. Firstly, cohesin, a protein essential for sister chromatid cohesion in mitosis, does not appear to be necessary for keeping sister chromatids in close proximity in G2 in Drosophila. This is particularly surprising given what is known about cohesin’s cell cycle association with chromatin and data from other organisms, and implies the existence of cohesin-independent connections that are able to keep chromosomes in close alignment during interphase. Secondly, I show that these cohesin-independent connections between sisters may be the same type of connection that form between homologs in Drosophila; at least, the two processes share regulatory mechanisms. This finding has several interesting implications for our understanding of pairing, in particular, demonstrating that the pairing of homologs, which has only been widely-observed in Dipteran insects, may be related to cohesin-independent sister chromatid pairing, which has been reported in other organisms as well. Thirdly, I describe results showing that the clustering of pericentric heterochromatin and of centromeres in Drosophila interphase cells is a dynamically-regulated process, with cells having mechanisms that both promote and antagonize clustering. These findings set the stage for future investigations into the regulation of clustering over the course of the cell cycle, and the establishment of cell type-specific clustering patterns. In this chapter, I discuss further implications of these results and speculate regarding the answers to several open questions, including the nature and purpose of cohesin-independent connections between chromosomes, how condensin might be regulating these, and how much of interphase chromosome positioning is established in mitosis.

Is there cohesin-independent pairing in vivo in flies?

The results described in this dissertation are based on experiments done in Drosophila cell culture; interestingly, there is evidence to suggest that cohesin-independent pairing of chromosomes, including both sisters and homologs, also occurs in vivo. Studies that induced cohesin cleavage via a TEV protease in Drosophila larvae and found that cohesin cleavage does not noticeably disrupt polytene chromosome alignment [1]. In contrast, overexpression of the condensin II component Cap-H2 disrupted polytene alignment in the same cell type [2],
supporting the idea that condensin II antagonizes cohesin-independent connections between sisters and homologs. Of course, polytene chromosomes represent a special case, because these cells are not proceeding through a normal cell cycle with replication followed by division, so experiments in additional cell types in the fly are needed. However, these results from polytene chromosomes are promising and support our model. Studies in the male Drosophila germline have shown that in normal cells, the chromosome arms of sister chromatids begin to separate in G2, suggesting little to no cohesin-independent pairing at these regions [3]; however, since the Drosophila germline has also been shown to be different from somatic cells in the onset of homolog pairing [4,5], perhaps this is not too surprising.

**Observations regarding the nature of homolog pairing.**

Our observations suggest that the same cohesin-independent mechanisms that act between sister chromatids may also act between homologs, suggesting that sister pairing and homolog pairing are one and the same. This model is consistent with the idea that recognition of a pairing partner is based on DNA sequence or chromatin structure [reviewed by 6, see also 7,8], and that pairing can accommodate more than two copies of a chromosome [9–13]. Our work also pertains to the question of whether or not Drosophila cells distinguish sisters versus homologs [see 14]. If they do not distinguish sisters from homologs at least in terms of pairing, then it is possible that sister chromatids may influence gene expression beyond increasing chromosome copy number. For example, as homolog pairing can influence the communication between regulatory elements and promoters in cis as well as in trans [15–23], sister chromatids may be able to join and influence this dialogue [11]. Indeed, just as transvection can occur between paired homologs, so might it occur between sister chromatids, making it possible for these two forms of transvection to be synergistic or mutually inhibitory during G2 (Figure 5.1). Importantly, inter-homolog communication and the contribution of sister chromatids to that process could vary by cell type, depending on the levels of cohesin-independent pairing between sisters and homologs.
Our data also address the long-standing question of when in the cell cycle pairing can be established. While some studies have shown that levels of pairing are higher in G1 than in G2, suggesting that S-phase is a stage when pairing is more dynamic and possibly disrupted [12,24], other work shows that pairing levels are similar in G1 and G2 [11], possibly reflecting variability between cell types. Our work suggests that S-phase/G2 is a stage when cohesin-independent sister pairing and, perhaps also, cohesin-independent homolog pairing can be established. In particular, our experiments in the male diploid Clone 8 cell line presented in Chapter 2 suggest that cohesin-independent pairing of sisters in G2 is not dependent on the presence of a homolog, which would indicate that cohesin-independent pairing can be established \textit{de novo} in G2. As such, perhaps the pairing of homologs can also be established at this stage. Of course, S-phase/G2 may not be the only stage of the cell cycle when homolog pairing is established; G1 homolog pairing could represent either an additional establishment event following the disruption of pairing in anaphase [25] or, theoretically, the maintenance of homologous connections from the previous cell cycle through mitosis [11,12].
What are the cohesin-independent connections between chromosomes?

A major remaining question concerns the nature of the cohesin-independent connection—between chromosomes. There are several possible answers, including the contribution of additional factors, such as proteins or RNA, that function similarly to cohesin in bringing chromosomes together, but act specifically in interphase. For example, it is possible that cohesion in somatic cells involves multiple novel cohesin complexes, as is known to be the case in Drosophila meiosis [26–30]. Alternatively, cohesin-independent cohesion might involve direct connections between chromosomes themselves without any need for bridging factors, perhaps involving some kind of nontraditional base pairing or DNA catenations resulting from replication. Catenations were one of the earliest models put forward for cohesion [31–33], and topoisomerase II, an enzyme that regulates catenations, also regulates cohesion [34–47]. Cohesin may help to maintain catenations [48,49]. Conversely, a complex related to cohesin and condensin known as Smc5/6 is thought to bind chromosomes in response to sister chromatid intertwinnings or other forms of topological stress and facilitate their resolution in yeast [50,51]. Our work suggests that catenations may exist independently of cohesin protein [see 41,47,52] and are sufficient to maintain the close proximity of sister chromatids in interphase in Drosophila cells. In future, it would be of great interest to test the role of topoisomerase II and Smc5/6 in this interphase pairing of sisters, to determine if they are acting with condensin II in this process, or if condensin II has taken over these functions in Drosophila cells.

Given the mechanistic relatedness we have observed between homolog pairing and sister pairing, it is possible that homolog pairing is also mediated at least in part by DNA catenations or entanglements [see 3,11,12,19,53]. While catenations are generally thought to result from DNA replication and therefore more likely to form between sisters than between homologs, catenations can be formed through multiple processes [54]. It is possible that homologs could become catenated when they are replicated in close proximity, perhaps via replication fork collapse and repair, or when there is DNA recombination between homologs, especially at the repetitive sequences of pericentric heterochromatin [55]. Interestingly, inhibition of topoisomerase II, can also influence levels of homolog pairing in Drosophila cells, though it is unclear if the latter is through similar roles in sister as versus homolog pairing [11; see discussion within].
If catenations are the basis of cohesin-independent connections between chromosomes, how might they be antagonized by condensin complexes? There is evidence that condensin complexes recruit topoisomerase II to chromosomes [56–61], or facilitate topoisomerase II activity in other ways [62], but there are also effects of condensin on sister chromatid resolution that may be independent of topoisomerase II [63–70]. Condensin is also important in restoring the structure of DNA following transcription [71], and it has been proposed to remove connections between sister chromatids that are formed during transcription [72,73]. Alternatively, the action of condensin to compact a chromosome, forming more interactions in cis, might limit its ability to interact with other chromosomes in trans [2,12,74,75].

**Why are Drosophila so different from other organisms in the amount of pairing?**

Given our observations in mitotic cells (Chapter 3), it seems likely that condensin I and II remove similar types of connections between chromosomes (possibly DNA catenations), but do so at different stages in the cell cycle. This is consistent both with the known localization of condensin I and II over the course of the cell cycle [76–78] and with findings in mammalian cells that condensin II begins the process of resolving sister chromatids in S-phase [79]. Perhaps in Drosophila, condensin II starts to remove DNA catenations in interphase but does so incompletely, such that many cohesin-independent connections between sisters and between homologs remain until mitosis when they are removed by condensin I as required for segregation, possibly in response to spindle formation [62] or other mitosis-specific factors. (Figure 5.2). This would explain why knockdown of cohesin did not disrupt G2 pairing of sister chromatids (Chapter 2), why knockdown of condensin I had a greater effect on mitotic cohesion than did that of condensin II (Chapter 3), as well as why the pairing of homologs in somatic cells is so much more widespread in Drosophila than in other organisms [6]. Interestingly, in flies, the condensin I component Barren has been shown to enter the nucleus earlier than in other organisms, binding to chromosomes in early prophase [80], which suggests that the balance between the activities of the two complexes might be different in Drosophila. This might be connected to the timing with which Drosophila cells remove catenations between chromosomes. Of course, this does not mean that in other organisms all catenations are resolved prior to mitosis or that they do not contribute to nuclear organization in G2. The identification of ultrafine anaphase bridges in normal human cells demonstrates
that catenations can remain until anaphase at certain regions, specifically, at centromeres, and that other regions have resolved catenations by this point under normal conditions [43,44,46,49,81].

The possible relationship between the timing of removal of catenations and the degree of cohesin-independent pairing suggests that the length of time a cell spends in G2 may be correlated with its requirements for cohesin-independent pairing mechanisms. Thus, in cells that divide soon after DNA replication, such as in yeast, the short length of G2 may minimize the need for cohesin-independent pairing of sisters. In contrast, cells with a long G2, such as the Drosophila cells used in our studies, may rely more heavily on cohesin-independent mechanisms to maintain interphase genome organization. Alternatively, it is possible that, rather than being a requirement of cells with a prolonged G2, extensive cohesin-independent pairing lengthens G2 by requiring more time to be resolved. Perhaps cells even use the regulation of pairing as a way to regulate the amount of time spent in G2. Further examinations into the exact timing of removal of catenations/homolog pairing in Drosophila cells will shed light on these models.
Figure 5.2 Model demonstrating how Drosophila might differ from other organisms in the extent of cohesin-independent pairing between sisters and homologs. Representation of the hypothetical cohesin-independent connections that form between chromosomes and the timing of their removal. Vertical blue lines represent sections of chromosomes, purple rings represent cohesin, orange rings represent condensin II, pink rings represent condensin I, and dashed blue lines indicate cohesin-independent connections (possibly DNA catenations). In Drosophila, the same cohesin-independent connections that form between sisters also form between homologs, but extensive inter-homolog interactions have not been observed in other organisms (homolog not pictured). As shown, cohesin-independent connections may be resolved later in the cell cycle in Drosophila than in other organisms; alternatively, it is possible that cohesin-independent connections are formed more efficiently in Drosophila, so that they are more widespread. Condensin I and II may antagonize cohesin-independent connections by compacting chromosomes in cis, disrupting trans interactions [2,12,74,75]. Note that in the top panel representing Drosophila, all cohesin-independent linkages are resolved by metaphase, but in theory, some connections between homologs might remain in metaphase and in anaphase. Similarly, in the lower panel representing other organisms, all cohesin-independent...
Figure 5.2 (Continued) connections between sisters have been resolved by G2, but it is possible that they remain in certain regions.

Why have cohesin-independent pairing of sisters?

If cohesin-independent connections exist between sister chromatids, why evolve another mechanism of cohesion in the form of the highly conserved and essential cohesin proteins? One explanation may lie with the importance of having connections that are unique to sister chromatids, thus ensuring their segregation into different daughter cells. Specifically, the unique connections can be provided by cohesin, whose establishment is coupled to replication [82–84], while, as we have shown, cohesin-independent mechanisms may contribute to genome organization in other ways. Secondly, cohesin-independent connections may allow sister chromatid alignment to be maintained at chromosomal regions where cohesin protein is not always bound at a high density. This would enable cohesin binding to be spatially and temporally dynamic [85,86] and permit additional roles of cohesin in interphase, such as in the regulation of transcription and DNA repair [87,88].

Another possible reason for having multiple mechanisms holding together sister chromatids is that they allow for a more layered regulation of cohesion removal as cells enter mitosis [see discussion of ref. 89]. In fact, in higher eukaryotes, cohesin proteins themselves are removed from different parts of the chromosome by distinct pathways at different times; while a small population of cohesin is retained at the centromeres and cleaved at anaphase [90–92], the bulk of cohesin on the chromosome arms is removed during prophase by Wapl and Pds5 [90,93–98]. Recent work suggests that telomeric cohesion involves yet additional regulation [99–101]. Cohesin-independent pairing provides a further layer to be removed in the segregation of sister chromatids and thus provides another potential point of regulation, which may be useful in determining the order of segregation [52]. Since all these processes must be coordinated with the condensation of chromosomes prior to mitosis, perhaps it is not surprising that condensin proteins also regulate cohesion [52,56–58,61,67–69,78,79,102–113]. Furthermore, it is possible that cohesion itself controls condensation in some way; in fact, there is much evidence in budding yeast that cohesin and cohesin regulators play a role in condensation [114–118]. Having a diversity of cohesion mechanisms
may allow for a more nuanced regulation of condensation, which may vary from one region of a chromosome to
another or occur most efficiently when it is gradual.

**What is the role of mitosis in establishing interphase chromosome positioning?**

Our work on cohesin and condensin complexes highlights that proteins with important roles in shaping
mitotic chromosomes also contribute to the organization of chromosomes in interphase. This also appears to be the
case for the clustering of pericentric heterochromatin and potentially of centromeres, as many of the candidate genes
identified, particularly in the pro-clustering category, are genes with well-characterized roles in mitosis. It may be
that some of these proteins have additional roles in interphase, or, that they help to establish patterns of organization
in mitosis that are maintained through the rest of the cell cycle, as Rabl first proposed for the organization of
centromeres and telomeres over a century ago [119]. Our findings suggest centromere clustering is a balance
between establishment in mitosis and regulation during interphase, when the non-mitotic factors we identified may
participate.

A potentially interesting finding from the screen is that condensin I and II may regulate clustering in
different ways, with condensin I being required for the establishment of clustering while condensin II antagonizes
clustering. While additional validation is necessary to confirm these results, the idea that the two complexes play
different roles is not unprecedented. For example, there is evidence that the loss of either complex results in
different phenotypes in mitosis, suggesting that the two complexes play different roles in determining the shape of
mitotic chromosomes in human cells [120], mice [121], chicken cells [122,123] and in Xenopus egg extracts
[111,120]. Alternatively, the different effects of condensin I and II could be related to the cell cycle timing of their
activities. Condensin II is known to be present in the nucleus throughout interphase, while condensin I enters during
mitosis [76–78,80]. Again, this finding could highlight the existence of mechanisms to establish centromere
clustering in mitosis, while there are mechanisms to antagonize or otherwise regulate centromere clustering in
interphase.
How do chromosomes balance homologous with heterologous interactions?

The fact that Drosophila nuclei show both homolog pairing as well as centromere clustering highlights the fact that different parts of a chromosome can participate in different types of interactions; the arms of Drosophila chromosomes are homologously paired while their centromeres associate with other, nonhomologous chromosomes. It is possible that these two types of interactions may influence each other. For example, the clustering of centromeres and/or telomeres which occurs in meiotic prophase of several organisms has been proposed to facilitate homolog pairing, by restricting the volume in which chromosomes have to search for homologous sequences [124–127]. Homolog pairing may also influence centromere clustering, as we have shown in Chapter 4, where we found that centromeres have a tendency to cluster with homologous centromeres over heterologous ones. Interestingly, this effect was not true for chromosome 4, which is mostly heterochromatic, and maybe is not subjected to the same pairing mechanisms that act at the euchromatic arms of other centromeres. Along these lines, there is evidence to suggest that pairing may be different at heterochromatin versus euchromatin; certainly, levels of pairing are lower at heterochromatic sequences compared to euchromatic ones [4,11,12].

The fact that condensin II plays a role in antagonizing homolog pairing as well as in antagonizing nonhomologous centromere associations [2,12,75,128] is further evidence that these two aspects of nuclear organization may be related. Additional characterizations of candidate genes identified in our clustering screen may reveal further connections between the two processes. Many of our candidate clustering genes were also described as pairing genes in the original screen (29 out of 54 pro-clustering genes and 15 out of 23 anti-clustering genes) but given that this candidate screen that was enriched for pairing regulators, this finding may not be too surprising. Extending the analysis of potential clustering genes genome-wide will give us a better sense of the overlap between clustering and pairing regulators. The fact that there were many genes we found to identify clustering but not pairing, particularly in the pro-clustering category, suggests that genome-wide studies will reveal further interesting candidates.
Concluding Remarks

In sum, our work highlights that chromosomes are shaped by several processes and factors over the course of the cell cycle, and understanding the balance between these factors is key to understanding nuclear organization. For example, the processes of chromosome condensation and sister chromatid resolution are intimately tied, as are the relative positioning of centromeres in mitosis and their clustering in interphase. In further studies of potential regulators of nuclear organization, it will be important to determine when in the cell cycle these factors have their effects, and to consider the vastly different states that the nucleus is in, for example, between G1 and G2. It will also be of great interest to determine what other features of interphase nuclear organization are affected by the cell having been through a mitotic division, or indeed, its preparation for the next one.
References


Appendix 1

Supplementary information related to chapters 2 & 3

The figures in this appendix are adapted from:


Author contributions: TNS performed experiments and carried out data analysis. TNS and EFJ provided reagents. TNS, EFJ and C-TW conceived and designed the experiments. TNS and C-TW wrote the paper.
Supplementary Figure A1.1 Quantitative PCR confirmed efficient knockdown of cohesin subunits by RNAi.

Results shown are for S2R+ cells following four days of RNAi. Relative mRNA levels were normalized to levels of rp49, a ribosomal gene, in each sample, and each sample was then normalized to levels in LacZ dsRNA-treated cells. Rad21 knockdown was confirmed to be more than 80% effective across multiple trials (left). Knockdowns of Smc1, Smc3 and SA were also found to be more than 80% effective in a single trial (right).
Supplementary Figure A1.2 Knockdowns of multiple cohesin subunits, and longer RNAi treatments, did not disrupt interphase pairing. (A) Graphs show the percentage of nuclei with a single FISH signal at several different loci following knockdowns of different cohesin proteins Rad21, Smc1, Smc3 and SA in various combinations, in both S2R+ and Kc_{167} cells. (B) Graphs showing the percentage of nuclei with a single FISH signal after Rad21 was knocked down for periods longer than our standard 4 day RNAi treatment, in both S2R+ and Kc_{167} cells. (For all graphs, shown are percentages from single trials, n≥290 nuclei per knockdown.)
Supplementary Figure A1.3 Knockdowns of cohesin subunits cause premature loss of cohesion in mitosis.
Metaphase spreads obtained from Kc167 cells (tetraploid) after (A) no dsRNA treatment, (B) Rad21 RNAi, (C) Smc1 RNAi and (D) Smc3 RNAi. Spreads were prepared following four days of RNAi without use of colchicine or any other drugs to increase mitotic index. Rad21 knockdown caused a more severe loss-of-cohesion phenotype as compared to knockdowns of Smc1 and Smc3.
Supplementary Figure A1.4 Cells continue to cycle following Rad21 knockdown while exhibiting metaphase cohesion defects. (A) Growth curves of Kc167 cells with no dsRNA (dark blue) and treated with Rad21 dsRNA (light blue). dsRNA was added at day zero and cell count was assayed every day for 6 days. Rad21 knockdown
Supplementary Figure A1.4 (Continued) caused a slight cell cycle delay compared to untreated cells. (B) FACS profiles of S2R+ cells after 5 days of Rad21 knockdown and stained with propodium iodide to assay DNA content. Rad21 RNAi caused a slight enrichment for G2 cells compared to cells treated with LacZ RNAi. (C) Timecourse showing gradual onset of the premature loss of cohesion phenotype in response to Rad21 RNAi in Kc167 cells. dsRNA was added at day zero and metaphase spreads were prepared each day for 6 days from untreated and Rad21 RNAi-treated cells.

Supplementary Figure A1.5 Knockdowns of genes in double and triple RNAi treatments are as efficient as single knockdowns. Quantitative PCR results are shown for S2R+ cells (processed in the same way as for Supplementary Figure A1.1). For single knockdowns, 5 µg of dsRNA was used while for multiple knockdowns, 5 µg of the dsRNA of each species was used (i.e. 10 µg total in double knockdowns, 15 µg in triple knockdowns) and the controls used were untreated cells.
Supplementary Figure A1.6 Knockdown of Rad21 and Slmb did not cause more aneuploidy than did either gene alone. Metaphase spreads were prepared from S2R+ cells following four days of RNAi (results are from a single trial; n=30 mitotic nuclei per knockdown). Knockdowns involving Rad21 double the number of FISH signals observed at AACAC and dodeca because of sister chromatid separation (for discussion of 359, see main text; P-values from Mann-Whitney U-test comparing LacZ and Rad21 RNAi are <0.0001 for AACAC and dodeca, P=0.0568 for 359). In the double knockdown of Rad21 and Slmb, this doubling in the number of FISH signals was also observed, but the overall number of chromatid pairs was not significantly increased compared to knockdowns of either Rad21 or Slmb alone (P-values from Mann-Whitney U-test comparing Rad21 RNAi and Rad21+Slmb RNAi are 0.7864, 0.5132, and 0.1652 for 359, AACAC, and dodeca, respectively).
Supplementary Figure A1.7 Knockdown of Rad21 and Slmb increases the number of FISH signals at certain euchromatic loci. Results shown are for S2R+ cells following four days of RNAi. The number of FISH signals described as “extra” depends on the copy number of the chromosome being examined. For FISH targets on the X chromosome (16E), chromosome 2 (24D and 28B), and chromosome 3 (69C and 100B), respectively, nuclei with greater than or equal to 3, 4 or 5 signals were classified as having extra FISH signals (see main text for more details). Shown are the percentages from single trials (n≥300 per knockdown). While the increases observed were modest compared to those seen at heterochromatic loci, significant increases were seen at 3 out of 5 loci examined (significance calculated by Fisher’s exact test, *, P<0.05 for difference between Slmb knockdown and Rad21+Slmb double knockdown; for 24D, 100B, 16E, 69C and 28B, respectively, P=0.0150, P=0.0447, P=0.0026, P=0.1382, P=0.0991).
Supplementary Figure A1.8 Slmb knockdown, alone or in combination with Rad21 and/or Cap-H2, leads to fewer G2 cells. (A) Representative control cells stained with antibodies against cyclin B. Cells that would be identified as “G2” are outlined with a dashed line (scale bar = 10 μm). Mitotic cells also express cyclin B, but were excluded on the basis of DAPI morphology. (B) Quantification of the percentage of G2 cells observed after four days of RNAi knockdowns in S2R+ cells (n≥100 cells per knockdown).
Appendix 2

Supplementary information related to chapter 4

The information presented here is based on an unpublished collaboration between myself and Eric Joyce. Both Eric and I conceived and designed experiments, while Eric performed FISH experiments and imaging, and I developed image processing algorithms and carried out image and data analysis.

Supplementary Figures A2.2 and A2.3 were made based on analysis of images taken during a previously published FISH screen:

Supplementary Figure A2.1 Example results from image registration algorithm using different thresholds for the amount of signal overlap required. All three images are from the same field of nuclei. Green and purple represent dodeca signals from the first and second hybridizations, respectively, while white indicates overlap between them following image registration. This particular field was chosen to have nuclei with varying qualities of overlap, to demonstrate how our thresholding distinguished between them. The white outlines highlight the nuclei meeting the requirement outlined above the panel (i.e. 25% overlap between dodeca signals, 40% overlap between dodeca signals, etc.). Scale bars = 10 um.
Supplementary Figure A2.2 Distances between FISH signals and between FISH and the nuclear periphery were not correlated with the number of FISH signals in a given nucleus. The data on each graph is taken from a single 384-well plate, pooling all wells (controls and experimental, total n=176774 nuclei) and dropping nuclei with zero FISH signals, which were out of the plane of focus. In each case, the number of FISH signals is shown on the x-axis (counting both dodeca and 359 in the panel on the left). The y-axis shows normalized distances. Mean distance between dodeca and 359 for each nucleus (left) was normalized to the nuclear diameter, while mean distances between either FISH signal and the nuclear periphery for each nucleus (middle, right) were normalized to the nuclear radius within that cell. Values above 1 are obtained because the values of nuclear radius and diameter are estimates based on nuclear area and assuming a circular shape (see main text of Chapter 4). The fact that there is little to no correlation between distance and the number of FISH signals in these graphs indicates that the distances can be used as measures of nonhomologous clustering and nuclear positioning independently of homolog pairing.
Supplementary Figure A2.3 Control cells have a modest correlation between nonhomologous clustering distance and distance to the nuclear periphery, which is lost following knockdown of either candidate pro-clustering or candidate anti-clustering genes. The data on each graph is taken from all four 384-well plates analyzed in our screen, pooling control wells (652 wells, top row), all wells that increased the distance between nonhomologous signals when knocked down (120 wells, middle row, candidate pro-clustering genes), and all wells
Supplementary Figure A2.3 (Continued) that decreased the distance between nonhomologous signals when knocked down (50 wells, bottom row, candidate anti-clustering genes). Each point on the graph indicates a single well, with the x-axis representing average distance between 359 and dodeca in that well, and the y-axis representing the shortest distance to the periphery for 359 (left column) and dodeca (right column) in each nucleus, averaged over the whole well. In control cells, there was a modest negative correlation between nonhomologous distances and distance to the periphery (R²=0.433 for 359 and R²=0.4618 for dodeca) indicating that when dodeca and 359 are closer to each other (reflecting possible centromere clustering), they tend to be further away from the nuclear periphery. This association is lost in knockdowns of candidate clustering genes of either category.

Supplementary Table A2.1 The percentage of successfully registered nuclei obtained in analysis of FISH images from successive rounds of hybridization. In each case n represents the number of nuclei. The percentage of registered nuclei is a measure of both the preservation of nuclei through both rounds of FISH, and of the success of our image registration.

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>n Hyb #1</th>
<th>n Hyb #2</th>
<th>n Registered</th>
<th>% Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl8</td>
<td>432</td>
<td>365</td>
<td>311</td>
<td>85.2%</td>
</tr>
<tr>
<td>S2R+</td>
<td>161</td>
<td>173</td>
<td>125</td>
<td>72.3%</td>
</tr>
</tbody>
</table>

Supplementary Table A2.2 FISH probes used in our re-hybridization experiments. Target locations and approximate sizes are based on Dernburg and Sedat, 1998. Copy numbers are based on my own karyotyping of metaphase spreads. Note that the copy numbers of Chromosome 4 are an estimate (see main text), and that our S2R+ cell line has lost the Y chromosome. FISH efficiencies are based on the Clone 8 data presented in this chapter (n=296 nuclei, efficiency is defined as the percentage of nuclei showing any FISH signals).

<table>
<thead>
<tr>
<th>Target</th>
<th>359</th>
<th>AACAC</th>
<th>dodeca</th>
<th>AATAT</th>
<th>AATAC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X Chr</td>
<td>Chr 2</td>
<td>Chr 3</td>
<td>Chr 4</td>
<td>X Chr</td>
</tr>
<tr>
<td>Approx. Size</td>
<td>~11 Mb</td>
<td>Unknown</td>
<td>Unknown</td>
<td>~3.5 Mb</td>
<td>~3.5 Mb</td>
</tr>
<tr>
<td>Copy # (Cl8)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2*</td>
<td>1</td>
</tr>
<tr>
<td>Copy # (S2R+)</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2*</td>
<td>-</td>
</tr>
<tr>
<td>FISH efficiency</td>
<td>98.6%</td>
<td>99.3%</td>
<td>100%</td>
<td>99.3%</td>
<td>94.2%</td>
</tr>
</tbody>
</table>