A genomic analysis of meiosis in *Drosophila melanogaster*

By

Danny E. Miller

Submitted to the graduate degree program in Molecular and Integrative Physiology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

R. Scott Hawley, Ph.D. (Co-chair)

_____

David F. Albertini, Ph.D. (Co-chair)

_____

Casey M. Bergman, Ph.D.

_____

Justin P. Blumenstiel, Ph.D.

_____

Aron W. Fenton, Ph.D.

_____

Timothy Fields, M.D., Ph.D.

Date Defended: April 4, 2016

The Dissertation Committee for Danny E. Miller

certifies that this is the approved version of the following dissertation:

A genomic analysis of meiosis in *Drosophila melanogaster*

_____

R. Scott Hawley, Ph.D. (Co-chair)

_____

David F. Albertini, Ph.D. (Co-chair)

Date approved: April 4, 2016

Abstract

Meiosis is a specialized form of cell division in which a single diploid cell undergoes one round of genome duplication followed by two rounds of cell division to produce four haploid gametes. In most organisms, including *Drosophila melanogaster*, programmed double-strand breaks (DSBs) are created during meiosis that are typically repaired by one of two mechanisms: crossing over, which involves the exchange of flanking markers, or noncrossover gene conversion (NCO), which copies short segments of DNA from a homologous chromosome to repair the break. Crossing over is necessary for the proper segregation of homologous chromosomes at the first meiotic division, a process facilitated by the synaptonemal complex (SC), a large, multi-protein structure that holds homologs together during meiosis. Chromosomes that fail to crossover may not segregate properly, resulting in aneuploid gametes.

In many organisms, including humans, two forces primarily control the distribution of crossovers along the chromosome arm. The strongly polar centromere effect functions to reduce the frequency of centromere-proximal crossovers, while interference ensures that crossovers occurring on the same chromosome arm are widely spaced. It is unknown if these forces control the distribution of NCOs as well. In addition, while it is known that Drosophila mutants that fail to construct SC cannot repair DSBs by crossing over, it is unknown if these breaks can be repaired as NCOs. Finally, the forces that prevent crossing over are of interest as well. In Drosophila, multiply inverted balancer chromosomes are used either to suppress recombination or to prevent the recovery of recombinant chromosomes. While it is known that

inversion breakpoints themselves suppress nearby crossover events it is unclear over what distance they act.

In this work, I used whole-genome sequencing to investigate recombination in *D. melanogaster*. First, I precisely positioned CO and NCO events after a single round of meiosis in 196 individual wild-type males. While I found that CO distribution appears to be controlled, as expected, by the centromere effect and interference, NCOs surprisingly do not seem to respond to these same controls. In addition, I looked for evidence of NCOs in SC-deficient flies and recovered a single NCO event, suggesting that while rare, repair by NCO is possible in these mutants. These data also allowed me to identify novel meiotic events such as transposable element (TE)-mediated copy-number variations, which included evidence of recurrent CNV formation, which is known to contribute to disease in humans. Finally, I identified the precise genomic location of the majority of the inversion breakpoints of several of the most commonly used *X* and *3*rd chromosome balancers in Drosophila. This knowledge allows us to understand over what distance these breakpoints suppress crossing over. This analysis also allowed me to identify several instances of double crossovers, demonstrating that the mechanism by which balancers suppress exchange with their normal-sequence homologs is incomplete.

Dedication

For my parents, Robert and Karen Miller, the two most remarkable people I have ever known.

You taught me the value of curiosity and hard work without ever telling me you were.

Acknowledgements

I'd like to first thank my advisor, R. Scott Hawley who, for reasons that I am still not clear on, allowed me to intern in his lab as a post-baccalaureate student—a decision that certainly changed the course of my life for the better. Scott is the mentor that every student wants but few get. He is a kind, thoughtful, and appropriately critical advisor who maintains an open door, a keen interest, and a sharp wit.

The Stowers Institute has been a remarkable place to train at—a scientist's dream and a benefit for humanity. I would be remiss if I did not acknowledge the vision of Jim and Virginia Stowers and thank them for their gift. The world will be a better place because they were in it.

I would also like to thank all the members of the Hawley lab, too numerous to list, who I've worked closely with over the years. I am also grateful to other collaborators at the Stowers Institute including Sue Jaspersen and Eric Ross. The help of the Molecular Biology core at Stowers was critical for many of my experiments; specifically, I'd like to thank Anoja Perera, Kate Hall, and Kendra Walton for their help and patience over the years.

I have worked with many outside collaborators from the Drosophila community over the past four years, and all have demonstrated by their actions the open attitude and supportive nature that makes this community so great. They include Rob Kulathinal and Craig Stanley from Temple University; Kevin Cook from the Bloomington Drosophila Stock Center; and Terri Markow and Maxi Polihronakis Richmond of the University of California San Diego Drosophila Species Stock Center. I would also like to thank Dan Lindsley, who not only wrote the two books I used more than any other during my four years in the laboratory—Genetic Variations of

Drosophila melanogaster (1967) and The Genome of *Drosophila melanogaster* (1992)—but also gave me Georgina Zimm's personal copy of The Genome of *Drosophila melanogaster* after I told him I didn't have my own. This was a gift I will treasure, and hope to use, for the rest of my career.

I have had the pleasure of working with a number of people outside of the Drosophila community as well over the past four years. They include David Page and Jennifer Hughes from the Whitehead Institute at MIT; Erik Garrison at The University of Cambridge; and Michael Silberbach from Oregon Health and Sciences University. Several individuals at Children's Mercy Hospital in Kansas City allowed or encouraged my foray into human genetics, including Stephen Kingsmore, Neil Miller, and Michael Artman. Laura Salem and Jamie Dyer at Rockhurst University have not only been collaborators in an ongoing laboratory project but they also gave me an opportunity to teach a semester of Advanced Human Anatomy early in graduate school—an experience which affirmed my love for and desire to teach. James Calvet and Chad Slawson at the University of Kansas Medical Center have given me the yearly opportunity to tell medical students more than they ever wanted to know about next-generation sequencing and how it can be used clinically. Bill Pearson (University of Virginia) and Lisa Stubbs (University of Illinois at Urbana-Champaign) for allowing me to be a TA for Computational and Comparative Genomics at Cold Spring Harbor—not only did they take me as a student, they let me come back the next two years as a TA to see if I could actually learn any of the material. In addition, all of the Computational and Comparative Genomics students who not only trusted me enough to help them analyze their hard-earned data during all hours of the day or night, but pushed me to understand and analyze data that I had never seen or handled before—I couldn't have

imagined a broader, more challenging, or rewarding experience. Finally, Kelly Ranallo, for not only being one of the hardest working patient advocates I have ever met, but for her uncanny ability to make connections and get even the most difficult tasks done with what appears to be the greatest of ease.

The MD/PhD program at the University of Kansas Medical Center has been a great source of support. Specifically, I would like to thank Tim Fields, Brenda Rongish, and Janice Fletcher, all of whom are willing to drop whatever they are doing to help a student at a moments notice—their interest is truly the students, and it shows.

My committee has been extremely supportive over the years and all of my projects are markedly improved because of them. David Albertini gave me a clear human link to my work and provided many opportunities to expand my knowledge outside of Drosophila female meiosis; Casey Bergman for improving my computational biology skills and teaching me the value of reproducibility and the importance of open science; Justin Blumenstiel for patiently helping me work through the tedious details needed to complete more than one project; and Tim Fields for valuable guidance and mentorship since my first day of medical school and his unwavering commitment to the students in the MD/PhD program. I would also especially like to thank Aron Fenton, who gave me a job in his lab when my only real qualification was “an interest in research.”

My parents, in-laws, and other family members have been wonderful supporters throughout this process, for which I am grateful. Tyler, who reminds me often why I’ve chosen the path I have; Brooklyn and Hayden for convincingly pretending to tolerate my endless

scientific explanations; and my son Toby, whose easy smile and hearty laugh has captivated me in ways I never believed possible.

Finally, and most importantly, I'd like to thank my wife, Angie. I am remarkably fortunate that not only did she marry me—a questionable decision on it's own—but she helped edit this thesis, assisted with figures, and provided constructive and critical questions throughout, helping me clarify my thinking and improve my science. Without her I would most certainly be lost.

Table of Contents

List of Abbreviations

DSB	Double Strand Break
CNV	Copy-number Variation
CO	Crossover
DCO	Double Crossover
InDel	Insertion/Deletion Polymorphism
NAHR	Non-allelic Homologous Recombination
NCO	Non-crossover Gene Conversion
RNA-Seq	RNA Sequencing
SC	Synaptonemal Complex
SCO	Single Crossover
SNP	Single Nucleotide Polymorphism
TE	Transposable Element
WGS	Whole Genome Sequencing

List of Figures

List of Tables

Chapter 1: Introduction

The proper segregation of homologous chromosomes is essential for the formation of viable haploid gametes during meiosis. Recombination between homologs, which occurs during prophase of meiosis I, functions to ensure the proper segregation of homologous chromosomes at the first meiotic division. The exchange of alleles from one homolog to another also results in an increase in genetic diversity within a population. How the position, distribution, and number of recombination events is determined is of great interest to researchers and has a variety of clinical implications.

Drosophila melanogaster as a model system

One of the most thoroughly studied model organisms, *Drosophila melanogaster* has been used in laboratories throughout the world for over 100 years (Sturtevant 2001; Kenney and Borisy 2009). The advantages of *D. melanogaster* as a model system include, but are not limited to:

1. Rapid generation time: approximately 10 days from egg to hatched fly
2. Large number of offspring: a wild-type female can produce hundreds of progeny
3. Relatively small and well-annotated genome: the draft 140-Mb genome of *D. melanogaster* was completed in 2000 (Adams *et al.* 2000; Myers *et al.* 2000) and is regarded as one of the most complete genomes available today (Santos *et al.* 2014) **(Figure 1.1)**.
4. Numerous genetic tools available: an abundance of genetic tools are available for Drosophila (Ashburner *et al.* 2005; Mohr *et al.* 2014), including genome editing with

CRISPR/Cas9 (Gratz *et al.* 2013) or TALENs (Beumer *et al.* 2013) and RNAi knockdown lines (Perkins *et al.* 2015).

5. Large number of mutant lines available: mutants are available from a variety of sources, such as the Bloomington Drosophila Stock Center (http://www.flybase.org), the UCSD Species Stock Center (https://stockcenter.ucsd.edu/info/welcome.php), the Kyoto Stock Center (https://kyotofly.kit.jp/cgi-bin/stocks/index.cgi), and the Vienna Drosophila Resource Center (http://stockcenter.vdrc.at/control/main).
6. A long history of being an open and sharing community: researchers are willing to freely share stocks and regents, and are typically willing to discuss unpublished results. This openness has allowed the research community to thrive and is instilled in all new Drosophila researchers early in their careers.

Figure 1.1: The karyotype of Drosophila melanogaster.
(A) D. melanogaster has four pairs of homologous chromosomes. Males and females both carry two copies of the 2nd and 3rd chromosomes, which are relatively large autosomes, and two copies of the small 4th, or dot, chromosome. Females with a normal karyotype will carry two copies of the X chromosome, while males with a normal karyotype will carry one copy of the X chromosome and one copy of the Y chromosome. Image from Bridges (1916) (B) Detail of each of the chromosome arms in *Drosophila melanogaster*. Gray boxes indicate heterochromatin and white indicates euchromatin. Note that the 2nd and 3rd chromosomes have left and right arms.

Drosophila is a large genus in the family Drosophilidae that contains 553 individual described species (as of August 2015, NCBI Taxonomy site, there are likely more than 1,000 species, but not all are described on the NCBI site) covering approximately 40 million years of evolution (Lachaise *et al.* 1986; Schaeffer *et al.* 2008). The Drosophila genus can be further divided into three subgenera—Drosophila (256 species), Sophophora (151 species), and Hawaiian Drosophila (146 species) **(Figure 1.2)**. Drosophila melanogaster is a member of both the melanogaster group and subgroup within the Sophophora subgenus **(Figure 1.2)**.

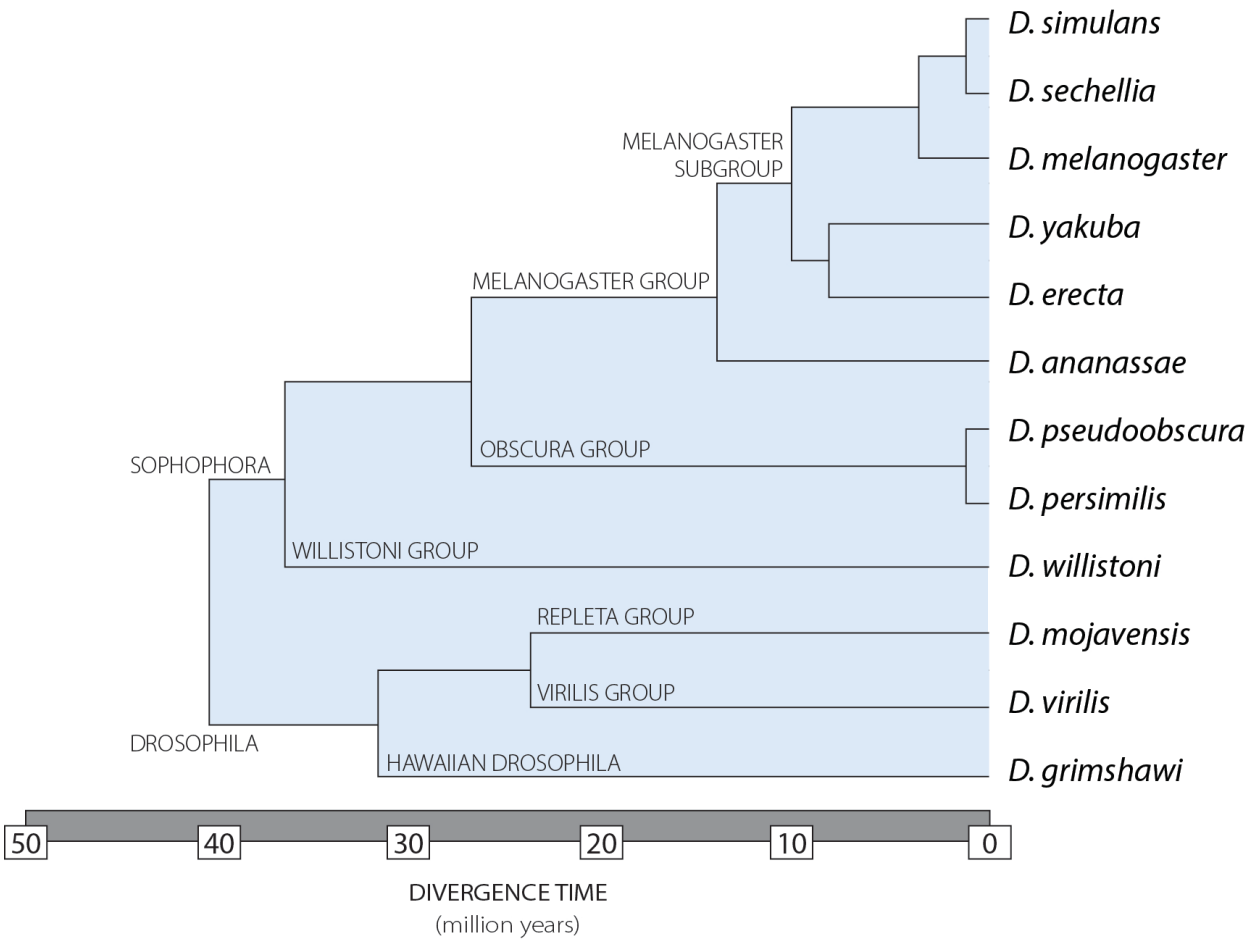

Figure 1.2: Phylogenetic tree of the 12 sequenced Drosophila species.

Approximately 40 million years separate all of the species within the Drosophila genus. While the majority of research has been done on *Drosophila melanogaster*, a growing body of work is looking at unique aspects of other species.

While the majority of research within the Drosophila genus has focused specifically on *Drosophila melanogaster*, a growing number of laboratories are describing novel phenotypes and behaviors in other closely related species. A search of PubMed for species within the Drosophila genus finds the majority of papers are from studies in *Drosophila melanogaster*, but a large (and growing) number of papers are being published on other species within the genus. Growth of research within the genus has been in part facilitated by the availability of high-quality genome assemblies for 11 additional species of Drosophila completed in the mid-2000s (Richards *et al.* 2005; "Evolution of genes and genomes on the Drosophila phylogeny." 2007; Schaeffer *et al.* 2008). Since the publication of the 12 genomes paper in 2007, a number of additional genome projects have been undertaken as of 2016**,** but none are as easily accessible as the original 12 genomes. With the recent advent of single-molecule long-read sequencing (PacBio or Oxford Nanopore, for example), it is expected that more high-quality genomes from other Drosophila species will become available in the coming years, facilitating further exploration into this diverse genus.

Meiosis in *Drosophila melanogaster*

Meiosis is a specialized form of cell division in which a diploid cell undergoes one round of genome duplication followed by two rounds of cell division to become a haploid gamete **(Figure 1.3)**. During Drosophila female meiosis, three of the four meiotic products are discarded as polar bodies, with only one product going on to form the egg, or oocyte. In male meiosis, all four haploid gametes will become small genetic torpedoes, known as sperm.

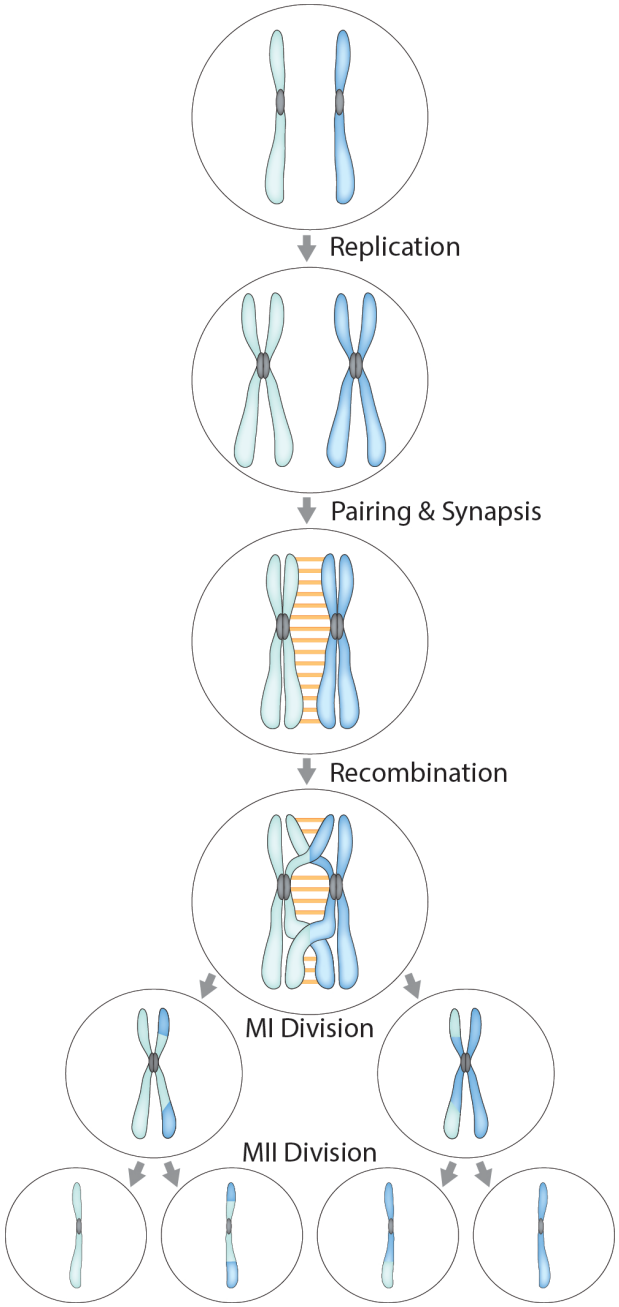

Figure 1.3: Meiosis I and meiosis II involve one round of genome duplication followed by two rounds of cell division.
After genome duplication homologs must find one another and pair and synapse along their lengths. A protenacious structure known as the synaptonemal complex (SC) is then built to hold these homologs together. Bivalents are homologs that are linked by crossover intermediates. At the first meiotic division homologs segregate from one another. At the second meiotic division sister chromatids segregate from one another. The end product is four haploid gametes.

Meiosis I can be further subdivided into four stages: prophase I, metaphase I, anaphase I, and telophase I. The steps that occur during prophase I will be considered in detail below, but briefly, crossing over occurs after homologous chromosomes pair along their lengths during prophase I. During metaphase I, bivalents (paired homologous chromosomes) migrate to a common location known as the metaphase plate and randomly align themselves with one of the two poles. In anaphase I, bioriented homologous chromosomes separate from one another, segregating to opposite poles. Finally, during telophase I the nuclear envelope will re-form around the separated chromosomes, creating two distinct nuclei, and cytokinesis, or the dividing of the cell itself, will occur. Importantly, at the end of meiosis I, sister chromatids are still attached to one another and will not separate until meiosis II.

Formation of SC and DSBs

Prophase I can be further subdivided into five stages which occur during oocyte development: leptotene, zygotene, pachytene, diplotene, and diakinesis. Oocyte development in Drosophila is divided into 16 stages, with the germarium, or the site of stem cell differentiation and DSB formation at stage 1, further divided into regions 1, 2A, 2B, and 3 **(Figure 1.4)**. Oocyte development starts when a germline stem cell divides to renew itself and create a cystoblast (Lake and Hawley 2012). This cystoblast then undergoes four incomplete mitotic divisions to produce a 16-cell cyst **(Figure 1.4)**. One of these 16 cells will go on to become the actual oocyte, with the other 15 becoming supporting cells known as nurse cells. Here, I am concerned with two specific events that occur during pachytene within the germarium: 1) the construction of

the synaptonemal complex (SC), a protenacious structure built between paired homologous chromosomes; and 2) the occurrence of programmed double-strand breaks (DSBs) **(Figure 1.4)**.

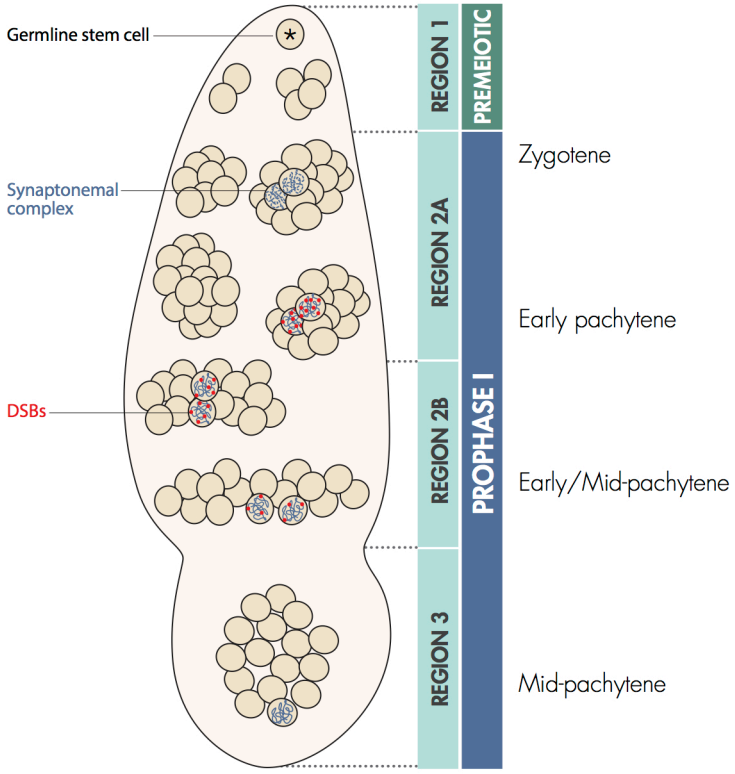

Figure 1.4: Prophase I occurs in Regions 2–3 of the female germarium.
The Drosophila melanogaster ovariale is divided into 16 stages, where stage 1 is also known as the germarium. The germarium can be further subdivided into regions 1–3. Prophase begins in region 2 with the construction of SC and the formation of DSBs. In Drosophila melanogaster DSBs are typically resolved by the end of stage 1 and a single cell, which will become the oocyte, will have SC present.

Oocyte selection begins in leptotene, classically described as a very short stage of prophase I in which chromosomes condense and homologous chromosomes begin to pair. However, recent evidence in Drosophila suggests that homologous chromosomes are paired even earlier—after the germline stem cell division but before the formation of the 16-cell cyst (Cahoon and Hawley

2013; Christophorou *et al.* 2013; Joyce *et al.* 2013). Next, during zygotene, chromosomes fully synapse and SC begins to become apparent along the arms of homologous chromosomes **(Figure 1.4)**.

The SC is a highly conserved, proteinaceous structure that forms between homologous chromosomes during prophase I. It consists of three distinct parts **(Figures 1.5)**: the lateral element, the central element, and the transverse filament. The transverse filament and central element proteins together make up the central region in the space between homologous chromosomes. In Drosophila the only known component of the lateral element is C(2)M, which attaches to the cohesion proteins SMC1 and SMC3. Acting as a homotetramer, the transverse filament protein C(3)G links the two lateral elements, and the central element protein Cona is thought to stabilize the central region. Another recently identified transverse-filament like protein, Corolla, appears by super-resolution microscopy to reside in the central region of the SC, but its exact position is yet to be determined (Collins *et al.* 2014).

Surprisingly, although the structure of the SC is highly conserved from yeast to mammals **(Figure 1.5),** very little sequence similarity is observed among these species and even within the Drosophila genus **(Figure 1.6)** (Fraune *et al.* 2012; Xiang *et al.* 2014). Indeed, within the Drosophila genus itself it is difficult to use protein homology to find SC components between species. For example, comparison of the protein sequence of the transverse filament protein C(3)G from the original 12 sequenced Drosophila species reveals very few conserved amino acids **(Figure 1.6B)**.

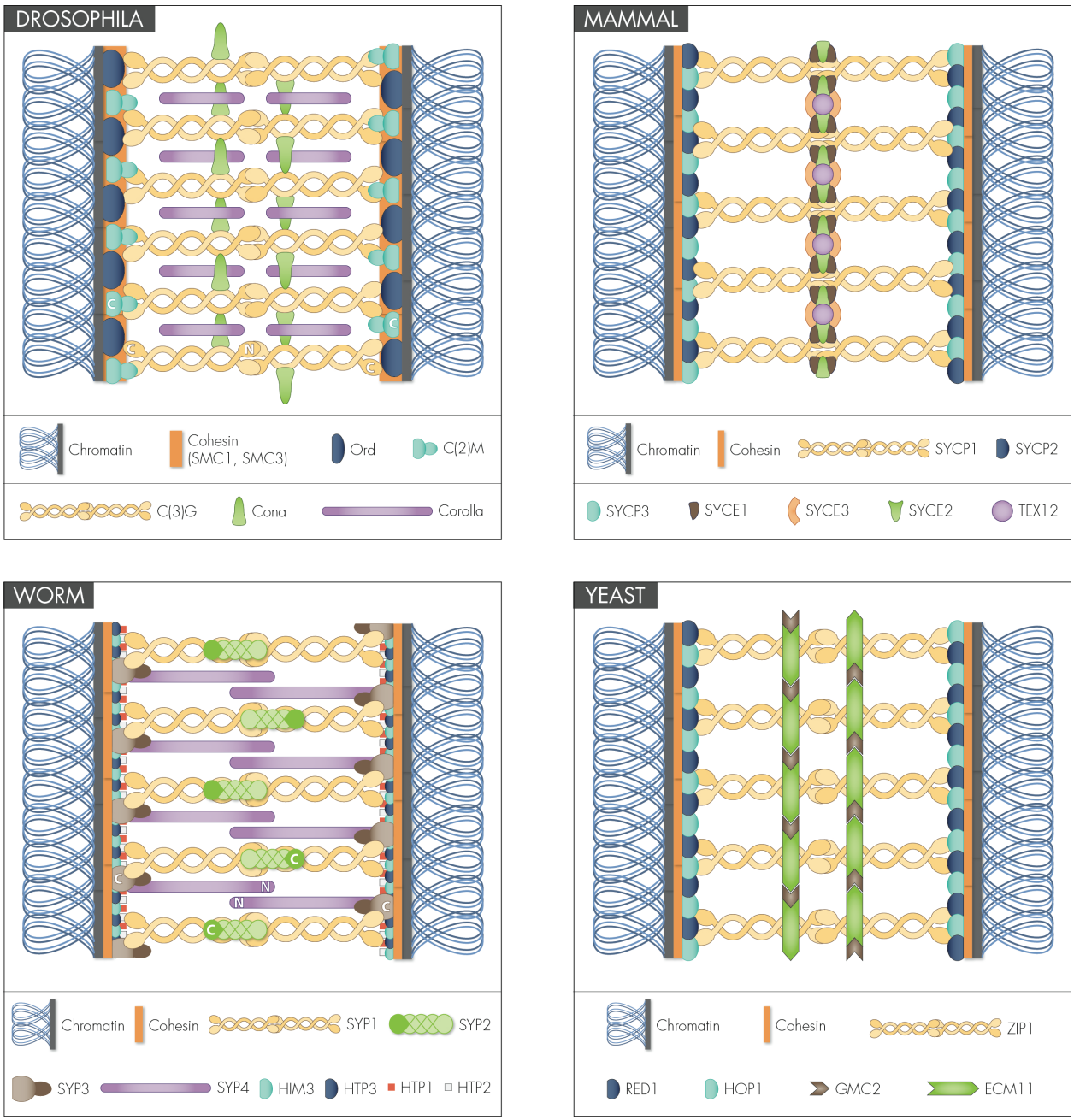

Figure 1.5: The structure of the synaptonemal complex is similar among many species.
The physical structure of the SC is conserved among a diverse set of organisms. The SC in all organisms has three basic structures: a lateral element, where SC components interact with cohesion components; transverse filament proteins, which act as linkers, holding homologs together; and a central element, which interact wit the transverse filament proteins. While the physical structure of the SC is conserved among organisms the components show a remarkable amount of sequence divergence, even among closely related species (Figure 1.6).

Figure 1.6 Representive alignments of transverse filament proteins.
Highlighted sequences indicate at least 50% similarity and are colored by amino acid type. Alignments were done using the entire transverse filament protein using ClustalO with default settings (Sievers *et al.* 2011) in Jalview (Waterhouse *et al.* 2009). (A) While the transverse filament protein is found in a wide variety of species there is very little sequence similarity among these species. (B) Within the Drosophila species group there is very little sequence conservation of the transverse filament protein C(3)G.

A second process that occurs during prophase I is the creation and repair of DSBs. Programmed breaks are created by the conserved topoisomerase-like protein Spo11 (Keeney *et al.* 1997), also known as Mei-W68 in Drosophila (McKim and Hayashi-Hagihara 1998). DSBs induced by Spo11 during meiosis are typically repaired by one of two pathways: crossing over (COs), or noncrossover-associated gene conversions (NCOs) **(Figure 1.7)**. Repair of DSBs by crossing over results in the exchange of flanking markers in which alleles segregate in a 2:2 ratio. Alternatively, repair of a DSB as an NCO involves copying short tracts of DNA from one chromatid onto another, resulting in a 3:1 segregation of alleles **(Figure 1.7)**.

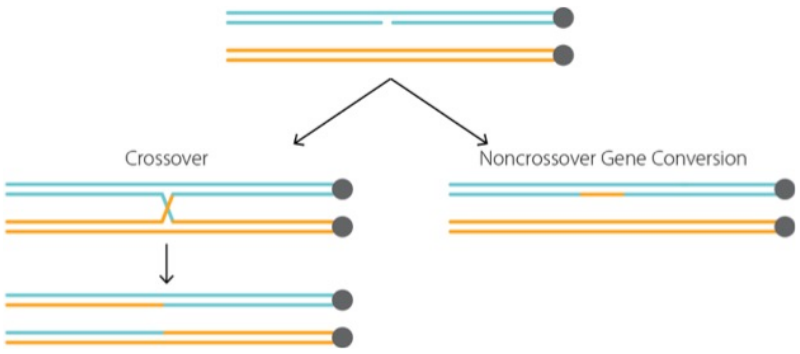

Figure 1.7: Non-crossover gene conversions result in a 3:1 segregation of alleles.
While crossing over involves the exchange of flanking markers, NCOs involve copying short stretches of DNA from a homolog in order to repair a DSB. This copying results in the 3:1 segregation of an allele, a violation of Mendel's laws.

DSBs undergoing repair may be visualized within the Drosophila germarium with an antibody to the phosphorylated histone H2A variant (γ-H2AV) (Lake *et al.* 2013). Using an antibody to γ-H2AV, it has been shown that 11–17 DSBs are formed during Drosophila female meiosis (Mehrotra and McKim 2006). Work by myself and others shows that approximately five of these breaks will be repaired as crossovers and eleven will be repaired as NCOs, accounting for nearly all the DSBs estimated to be produced. A more detailed analysis of how these breaks are repaired may be found in Chapter 2.

Drosophila females deficient in SC construction produce offspring with a higher rate of chromosome segregation errors and with no apparent crossing over (Page and Hawley 2001). These females also show a reduced number of DSBs—as low as 20% of the level observed in wild type (Mehrotra and McKim 2006). Whether the offspring of SC-deficient mothers carry NCOs has been an open question for some time. Although Carlson (1972) reported the recovery of no NCO events in such flies, the number of progeny screened is that study is unknown. In this

thesis, I report the recovery of a single NCO event defined by two polymorphisms after screening 95 individual male progeny by whole-genome sequencing (Chapter 3).

Numerous forces appear to control the distribution and number of recombination events

In Drosophila, a number of processes have been described that control the distribution of crossover events along the chromosome arm. Two of these processes, interference and the centromere effect are discussed in depth in Chapter 2. Briefly, **interference** acts to ensure that when crossovers occur on the same chromosome arm they are widely separated. It is a conserved process, observed from yeast to mammals, however there are some exceptions. For example, in *Drosophila mauritiana*, a species of Drosophila approximately 3 mya divergent from *D. melanogaster*, interference may be reduced or absent (True *et al.* 1995). While the mechanism of action of interference is poorly understood, work in *Caenorhabditis elegans* suggests that the SC may play a role in mediating interference (Sym and Roeder 1994). What role the SC has in mediating interference in Drosophila, if any, is an active area of research.

A second phenomenon, the **centromere effect,** acts to shift crossovers away from centromere-proximal regions of each chromosome arm. In Drosophila, the centromere effect results in the majority of crossover events falling in the distal 2/3 of any chromosome arm. Because a large amount of heterochromatin surrounds each centromere in Drosophila, there was some question about whether the centromere effect was mediated by proximity to the centromere or by proximity to heterochromatin. Work by Yamamoto and Miklos (1977; 1978) resolved this question by removing large blocks of pericentric heterochromatin and showing, as a result, that shifting euchromatin closer to the centromere **(Figure 1.1)** shifted crossovers even

more distally along the arm than seen in wild type. This demonstrates that proximity to the centromere itself, and not the presence of pericentric heterochromatin, reduces crossing over in nearby intervals.

Two additional forces, crossover assurance and the interchromosomal effect on crossing over, affect the distribution of crossover events in a number of organisms. Crossover assurance works to ensure that each chromosome arm has at least one crossover event and is best understood in the yeast *Saccharomyces cerevisiae* and the worm *C. elegans*. Indeed, crossover assurance is most clearly demonstrated in *C. elegans*, where exactly one crossover occurs per chromosome arm (Cline and Meyer 1996). The mechanism of crossover assurance is unclear, although recent work in *C. elegans* is beginning to illuminate some of the proteins involved. Rosu and colleagues (2013) identified that the protein DSB-2 localizes to chromatin in meiotic prophase at the same time DSB formation occurs, suggesting that it acts to promote the formation of DSBs. The authors propose that a mechanism exists to monitor the formation of a crossover event, resulting in the removal of DSB-2 from the chromatin of the chromosome arm the CO occurred on, effectively shutting down subsequent DSB formation.

In other organisms, such as Drosophila, it is less clear to what degree crossover assurance acts, if at all and how this mechanism is separated from interference. In Drosophila it is not uncommon to recover double- or triple-crossover chromatids, demonstrating that crossover assurance does not act to limit the number of crossovers per chromatid, as in *C. elegans*, but likely that interference is the major force in limiting the number of crossovers per chromatid. Similarly, the upper limit on crossovers per chromatid observed in *C. elegans* may also simply be a result of more complete interference in this species.

The interchromosomal effect on crossing over is simply the observation that when crossing over is reduced or suppressed on one chromosome arm it is increased on other chromosome arms (Lucchesi and Suzuki 1968). In all organisms, this suppression is most likely to occur when one homolog carries a large inversion or translocation that the other homolog does not. In Drosophila, this is most often observed in stocks carrying a balancer chromosome. Balancers are multiply inverted and rearranged chromosomes that prevent either the occurrence or recovery of crossovers and are discussed in detail in Chapters 4 and 5. In a stock with two balancer chromosomes, crossing over will be greatly increased on the third nonbalanced chromosome. Joyce and McKim (2011) demonstrated that the delay in formation of a crossover when an inversion heterozygote is present is due to activation of the Pch2-dependent pachytene checkpoint (Joyce and McKim 2009), delaying the exit from pachytene and allowing new DSBs to form on other chromosome arms and leading to an increase in crossovers on those chromosomes.

Hotspots of recombination do not appear to exist in Drosophila

Hotspots of recombination are fine-scale variations in the rate of recombination along a chromosome arm that are apparent in a number of organisms, including humans. In several species of mammals, the protein PRDM9 binds to the motif CCnCCnTnnCCnC and directs DSB formation to these regions (Baudat *et al.* 2010). Notably, in both humans and mice, different alleles of PRDM9 identify slightly different CCnCC-like motifs, explaining variation in recombination rate in different populations within those species (Baudat *et al.* 2010). However,

PRDM9 does not direct DSB formation in all mammals, such as dogs, which lack functional copies of PRDM9 (Muñoz-Fuentes *et al.* 2011).

Attempts to identify hotspots of recombination in Drosophila have thus far only identified fine-scale rate changes in small regions of the genome in several species. Fine-scale rate heterogeneity has been reported in *Drosophila pseudoobscura* (Cirulli *et al.* 2007), *Drosophila persimilis* (Stevison and Noor 2010), and *D. melanogaster* (Comeron *et al.* 2012; Singh *et al.* 2013), although the methods used to uncover rate heterogeneity in one study have been called into question by some (Gilliland 2015). Additional work has attempted to identify motifs associated with crossing over in Drosophila, with several groups reporting motifs in *D. melanogaster* (Miller *et al.* 2012; Comeron *et al.* 2012; Singh *et al.* 2013) (see Appendix A), and *D. pseudoobscura* (Cirulli *et al.* 2007). Yet, it appears that there is little support for any of these motifs to explain the apparent recombination rate variation observed in Drosophila (Heil and Noor 2012). Indeed, in follow-up work (Chapter 2), I find motifs that appear to be enriched around CO sites, but I am then able to find the same motifs when analyzing randomly selected sites of crossing over, suggesting that these motifs may not be real.

The apparent absence of a PRDM9-like protein in Drosophila does not itself rule out the possibility that recombination hotspots exist. A recent study in birds observed that while they lack PRDM9, hotspots of recombination remain at functional regions of the genome, suggesting the accessibility of these regions during meiosis facilitates the formation of crossovers (Singhal *et al.* 2015). A similar hypothesis for Drosophila has been formulated by Comeron and colleagues in which accessibility to regions transcribed in early ovarian development allows for the formation of DSBs (Adrian and Comeron 2013).

Summary

While much is known about meiosis and the distribution of recombination events in *Drosophila melanogaster*, the falling cost of 2nd generation sequencing allows us to ask new and exciting questions. For example, one may investigate classic genetic resources, such as balancer chromosomes, which have been used for decades but whose structure has yet to be elucidated because they were built in the pre-genomic era. In this thesis, I use 2nd generation sequencing to address a number of questions about female meiosis in the following chapters.

Chapter 2: Whole-genome analysis of individual meiotic events in *Drosophila melanogaster* reveals that noncrossover gene conversions are insensitive to interference and the centromere effect. In this chapter I analyze the distribution of CO and NCO events along the *X*, *2nd*, and *3rd* chromosomes in 196 individual males. I find that interference and the centromere effect are unique properties of each chromosome arm. I also observe that NCOs are insensitive to interference and the centromere effect, a previously unknown property of NCOs. In addition, I identify several cases of both inherited and *de novo* copy number variation (CNV), with every CNV appearing to be mediated by unequal crossing-over between transposable elements (TEs).

Chapter 3: A whole-genome analysis of offspring from mothers deficient in the transverse filament protein c(3)G. Here, I examine the genomes of 95 males from mothers who were deficient in SC formation. While, as expected, I observe no evidence of crossing over I do recover one NCO event along with several inherited and *de novo* CNV events. Remarkably I find

that one of the CNV events is identical to that observed in the 196 genomes studied in chapter 2—evidence for recurrent CNV formation, similar to that seen in humans. Finally, I describe three triploid progeny who phenotypically appeared to be male.

Chapter 4: Rare recombination events generate sequence diversity among balancer chromosomes in *Drosophila melanogaster*. Balancer chromosomes, or chromosomes that either suppress crossing over or the recovery of crossover products were synthesized in the pre-genomic era, so much about the molecular nature of these chromosomes is unknown. In this chapter I identify the breakpoints of the most commonly used Drosophila *X* chromosome balancer, *FM7*. I also describe several occurrences of rare double crossing over events within the central inversion that replaces a female sterile allele marking the balancer with a wild-type copy of the gene. Finally, I investigate the nature of the Bar duplication, identify the molecular breakpoints of the duplication, and characterize two revertants that arose due to unequal recombination events within the duplicated segment.

Chapter 5: Sequence analysis of *Drosophila melanogaster third* chromosome balancers reveals the loss of *p53* on the *TM3* balancer. This chapter expands on work done in chapter 4 and identifies the breakpoints of the third chromosome balancers *TM3*, *TM6*, and *TM6B*. Surprisingly, I find that one of the inversion breakpoints of *TM3* bisects the highly conserved tumor suppressor gene *p53*, meaning any stock balanced with *TM3* is a *p53* heterozygote. The *TM3* balancer leaves a large segment of the distal tip of *3L* unbalanced, which allows

examination of how close to the inversion breakpoint single crossover events can occur. In addition, I recover evidence of double crossover events onto both *TM3* and *TM6B*, extending our observations from the *X* chromosome balancer *FM7*. Finally, I identify the molecular nature of the majority of the alleles that these three balancer chromosomes carry.

Chapter 6: Conclusion. Summary of results and a discussion of future directions my research may take.

Chapter 2: Whole-genome analysis of individual meiotic events in *Drosophila melanogaster* reveals that noncrossover gene conversions are insensitive to interference and the centromere effect

This chapter is adapted from: Miller, D. E., C. B. Smith, N. Yeganeh Kazemi, A. J. Cockrell, A. V. Arvanitakis *et al.*, 2016 Whole-Genome Analysis of Individual Meiotic Events in *Drosophila melanogaster* Reveals that Noncrossover Gene Conversions are Insensitive to Interference and the Centromere Effect (Miller, Smith, et al. 2016).

ABSTRACT

A century of genetic analysis has revealed that multiple mechanisms control the distribution of meiotic crossover events. In *Drosophila melanogaster*, two significant positional controls are interference and the strongly polar centromere effect. Here, we assess the factors controlling the distribution of crossovers (COs) and noncrossover gene conversions (NCOs) along all five major chromosome arms in 196 single meiotic divisions in order to generate a more detailed understanding of these controls on a genome-wide scale. Analyzing the outcomes of single meiotic events allows us to distinguish among different classes of meiotic recombination. In so doing, we identified 291 NCOs spread uniformly among the five major chromosome arms and 541 COs (including 52 double crossovers and one triple crossover). We find that unlike COs, NCOs are insensitive to the centromere effect and do not demonstrate interference. Although the positions of COs appear to be determined predominately by the long-range influences of interference and the centromere effect, each chromosome may display a different pattern of sensitivity to interference, suggesting that interference may not be a uniform global property. In addition, unbiased sequencing of a large number of individuals allows us to describe the formation of *de novo* copy number variants, the majority of which appear to be mediated by unequal crossing over between transposable elements. This work has multiple implications for our understanding of how meiotic recombination is regulated to ensure proper chromosome segregation and maintain genome stability.

INTRODUCTION

The proper segregation of homologous chromosomes at the first meiotic division is essential for the production of viable haploid gametes. In most instances, proper homolog segregation is assured by the formation of crossovers (COs), reciprocal recombination events that link homologous chromosomes together. COs arise at a subset of programmed double-strand breaks (DSBs) that are induced during early prophase by Spo11. DSBs not repaired as COs must be repaired by another mechanism, such as noncrossover gene conversion (NCO) events, sister chromatid exchange events, or by nonhomologous end-joining (Do *et al.* 2013). In many organisms, COs and NCOs occur more frequently at specific regions of the genome, termed hotspots (Lichten and Goldman 1995; Hey 2004). The protein PRDM9 directs the formation of DSBs to these regions in some organisms (Baudat *et al.* 2010). Other organisms, such as *Saccharomyces cerevisiae*, lack a PRDM9-like protein but still have hotspots of recombination, and still other organisms lack PRDM9 and display a more even distribution of COs and NCOs (Auton *et al.* 2013; Singhal *et al.* 2015), suggesting that PRDM9-independent mechanisms may influence DSB formation. No equivalent of PRDM9 has been identified in *Drosophila melanogaster* or other species within the *Drosophila* genus (Heil and Noor 2012; Manzano-Winkler *et al.* 2013).

Drosophila oocytes experience approximately 11–17 DSBs per meiosis that are restricted to the euchromatin (Jang *et al.* 2003; Mehrotra and McKim 2006; Lake *et al.* 2013). How the position of these DSBs is determined and their fate (whether they become COs or NCOs) is poorly understood. Based on previous studies, the overall distribution of COs in *D. melanogaster* oocytes appears to be controlled by multiple mechanisms, most notably

crossover interference and the centromere effect (Dobzhansky 1930; Offermann and Muller 1932; Beadle 1932; Lindsley and Sandler 1977). The identification of these mechanisms began with the finding that the genetic distance between phenotypic markers examined was not consistent with the physical location of the genes on polytene maps for any of the five major chromosome arms (Dobzhansky 1930). This suggested that the frequency of crossing over was not proportional to physical distance. Indeed, as noted by Lindsley and Sandler (1977), the frequency of exchange is lowest in both the centromere-proximal euchromatin and telomeric regions and highest in the medial region of the chromosomes **(Figure 2.S1)**. Later studies showed that the reduced level of exchange in the proximal euchromatin reflects the activity of the centromere effect, which strongly reduces crossing over in a polar fashion in centromere-proximal regions of the genome (Offermann and Muller 1932; Beadle 1932; Sturtevant and Beadle 1936; Yamamoto and Miklos 1977). Recent work in *S. cerevisiae* has shown that the Ctf19 inner kinetochore subcomplex suppresses centromere-proximal COs by suppressing pericentric DSBs, the first demonstration of a specific protein or complex contributing to the centromere effect (Vincenten *et al.* 2015). Other studies suggest that the telomeres may also suppress exchange in a polar fashion, although the effect is substantially weaker than near the centromeres (reviewed in Hawley 1980).

The distribution of COs is also influenced by crossover interference, which can act over long distances. First described in Drosophila by Sturtevant and Muller (Sturtevant 1913; 1915; Muller 1916), interference prevents a second CO from forming near an existing CO, typically ensuring the wide spacing of double crossover (DCO) events. Although interference in other organisms appears to be mediated by modification of the synaptonemal complex (SC) in

response to COs (Sym and Roeder 1994; Libuda *et al.* 2013; Zhang *et al.* 2014), it remains unclear whether the SC also plays a role in mediating interference in *D. melanogaster* (Page and Hawley 2001). Finally, there is little information in Drosophila as to what degree, if any, interstitial sites or domains play in controlling the frequency of crossing over in specific euchromatic regions.

Several groups have employed whole-genome sequencing (WGS) to search for regions of increased crossing over, or recombination hotspots, in *D. melanogaster*. One method is to identify COs in pedigrees generated through controlled crossing schemes. Using this design, two recent studies failed to find strong evidence of hotspots in *D. melanogaster*, but did identify evidence for intervals of higher or lower rates of crossing over either within the region studied or at the whole-genome level (Comeron *et al.* 2012; Singh *et al.* 2013). These observations suggest that traditional recombination hotspots may not exist in Drosophila (Manzano-Winkler *et al.* 2013). A second approach, which infers recombination rates from population genetic data, also indicated that Drosophila likely does not have hotspots (Chan *et al.* 2012).

While it is known that interference and the centromere effect control crossover distribution, very little is known about the factors that control the distribution of NCOs in Drosophila. Although early genetic studies suggested that NCOs do not exert interference on COs or respond to interference from COs (Hilliker and Chovnick 1981), these studies looked at only a small number of loci, and only one of them (the *rosy* locus) in great detail. A recent study using WGS to analyze progeny that were allowed to freely recombine for one, two, five, or 10 generations has shown that, unlike COs, NCO sites appear to be evenly spaced throughout the

genome (Comeron *et al.* 2012). However, this study did not specifically investigate the joint distribution of COs and NCOs and their relationship to each other after a single round of meiosis. Thus, the effect, if any, of interference on NCOs has yet to be investigated on a genome-wide scale after a single round of meiosis in wild-type individuals.

In the present study, we determined the precise position of CO and NCO events on all five major *D. melanogaster* chromosome arms in 196 single meioses. We found a paucity of COs in the centromere-proximal one-half of most chromosome arms, consistent with the influence of the centromere effect on crossing over. Furthermore, our data suggests that the degree to which interference controls CO positioning may vary across the genome. However, proximity to the centromere does not seem to reduce the frequency of NCOs in this region, suggesting that NCOs are not sensitive to the centromere effect. We also observed NCOs near sites of crossing over and near other NCO events, supporting the hypothesis that NCOs are not sensitive to interference.

Unbiased sequencing of a large number of closely related individuals allows for the recovery of unexpected meiotic events. For example, we recovered several double crossovers (DCOs) much smaller than any previously observed in Drosophila. In addition, we observed three NCO events that appear to be the result of discontinuous repair, demonstrating the value of studying a large number of individual meiotic outcomes to elucidate novel or rare repair outcomes. Finally, analysis of all 196 individual male genomes revealed eight large copy number variants (CNVs) ranging in size from 17 kb to 855 kb, most of which appear to have been the result of unequal crossing over between transposable elements (TEs). This leads to a revision of the standard model of TE copy number control through ectopic recombination and suggests

that as in humans, TE-mediated copy number variation plays an important role in creating genetic heterogeneity in *D. melanogaster.*

RESULTS

We directly assessed the number and position of COs and NCOs in *D. melanogaster* using females obtained by crossing two divergent, isogenic stocks: w^{1118} and Canton-S, two wild-type laboratory lines commonly used in meiotic segregation assays (Page *et al.* 2008; Miller *et al.* 2012; Collins *et al.* 2014). Isogenized parental lines were found to be different at 486,549 single nucleotide polymorphisms (SNPs) by WGS. Specifically, we identified on average 1 SNP every 379 base pairs on the *X*, 1/192 bp on *2L*, 1/295 bp on *2R*, 1/241 bp on *3L*, and 1/302 bp on *3R*. Heterozygous F1 females were the crossed to either homozygous w^{1118} males or homozygous Canton-S males and 196 of the resulting F2 male offspring were individually whole-genome sequenced **(Figure 2.S2)**. The F2 males were sequenced to an average *X*-chromosome depth of 24x (min average: 8x, max average: 45x) and an average autosomal depth of 45x (min average: 14x, max average: 87x) **(Table 2.S3)**. By analyzing the euchromatic portion of the genome, 541 sites of crossing over and 291 NCOs were identified by changes in the haplotype origin along the maternally transmitted chromosomes (see Methods) **(Figure 2.1, Figure 2.S3, Figure 2.S4, and Tables S2.1–S2.2)**. We observed no significant difference in the number or distribution of recombination events recovered from heterozygous females that were the progeny of reciprocal crosses between w^{1118} and Canton-S **(Figure 2.S2)**.

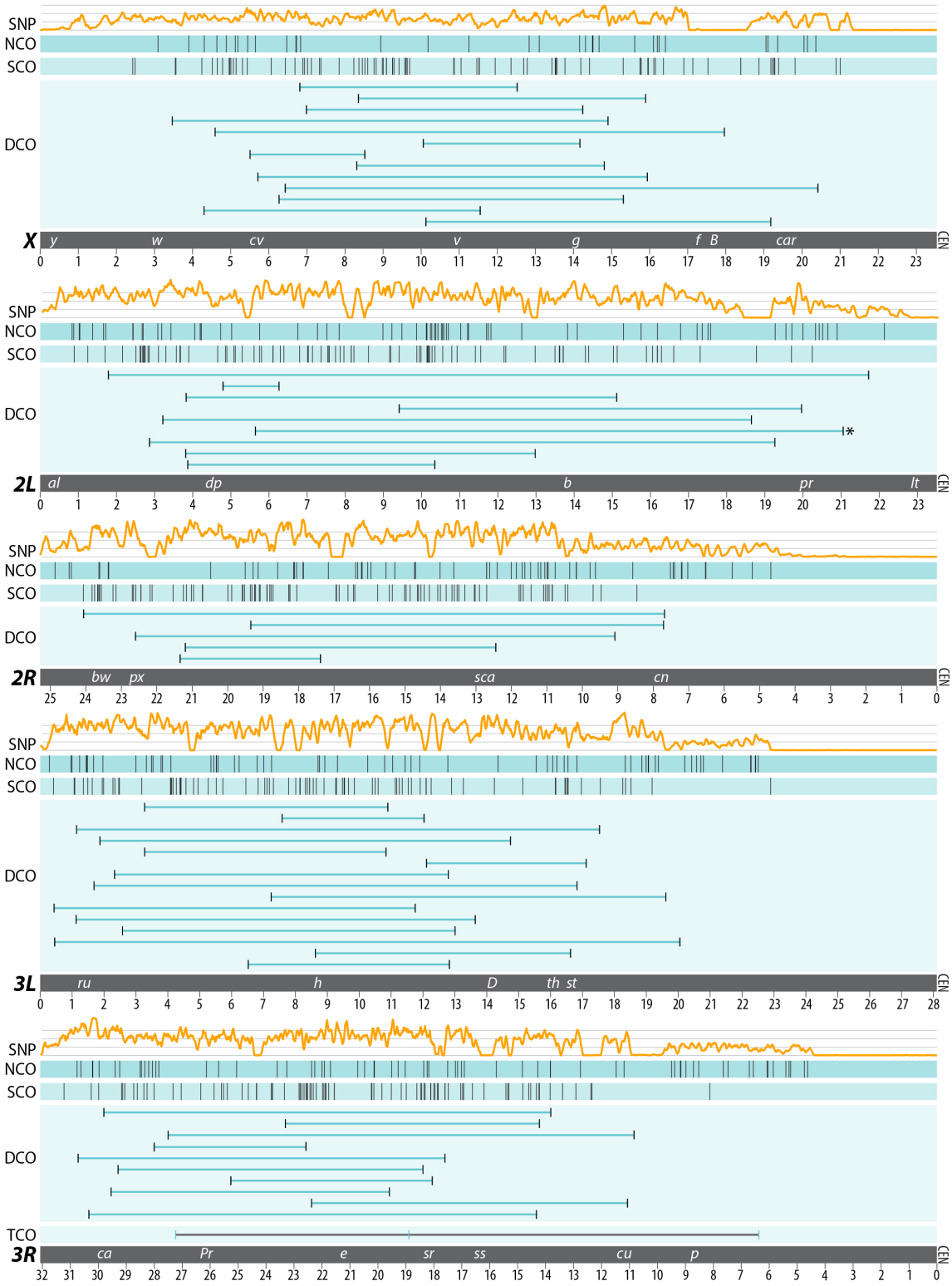

Figure 2.1. Distribution of 541 COs and 291 NCOs recovered in this study.
Each panel represents one of the five major *D. melanogaster* chromosome arms (the 1.4-Mb *4th* chromosome was not examined in this study). The centromere (CEN) resides on the right side of each panel. The top track in each

panel shows the SNP density observed when comparing the Canton-S and w^{1118} stocks. Note that SNP density drops to zero in the centromere-proximal regions of most chromosome arms, reflecting the recent addition of previously unmapped sequence to the latest *D. melanogaster* genome release. The NCO and SCO tracks show the locations of all NCOs and single COs recovered, respectively, the DCO tracks show the locations and spans of all double COs. One DCO on *2L* (denoted by *) was partly the result of unequal crossing over between two transposable elements. One triple CO (TCO) was recovered on *3R*. The centromere effect shifts crossovers distally on the autosomal arms; note that close to 80% of the SCOs on each autosomal arm occur in the distal one-half of the chromosome, but that frequency is only 60% in the distal one-half of the *X* chromosome. Commonly used visual markers are shown in the bottom track of each panel; descriptions of each can be found at FlyBase (http://www.flybase.org). Chromosome coordinates are in Mb along the *x*-axis.

To assess the number of false-positive NCOs recovered, we randomly selected 28 of 79 NCO events defined by a single SNP and 19 of 41 NCO events defined by two SNPs for sequence verification, because we considered those the most likely to be false-positive NCO events. We validated all of the 47 selected NCO events by PCR and Sanger sequencing, giving us high confidence that the remaining NCO events are, in fact, real **(Table 2.S4)**.

Distribution of single COs

The observed pattern of COs followed a distribution expected by traditional phenotypic marker analysis, with the four autosomal arms displaying a paucity of COs in the centromere-proximal euchromatic sequence due to the centromere effect, and a less pronounced telomere effect shifting COs away from the telomeric regions on all five major chromosome arms **(Figure 2.2)**. For example, for the four autosomal arms, 72%–83% of the SCOs were in the distal one-half of the chromosome arm **(Table 2.S1)**, demonstrating the ability of the centromere effect to alter the proximal distribution of single crossover (SCO) events.

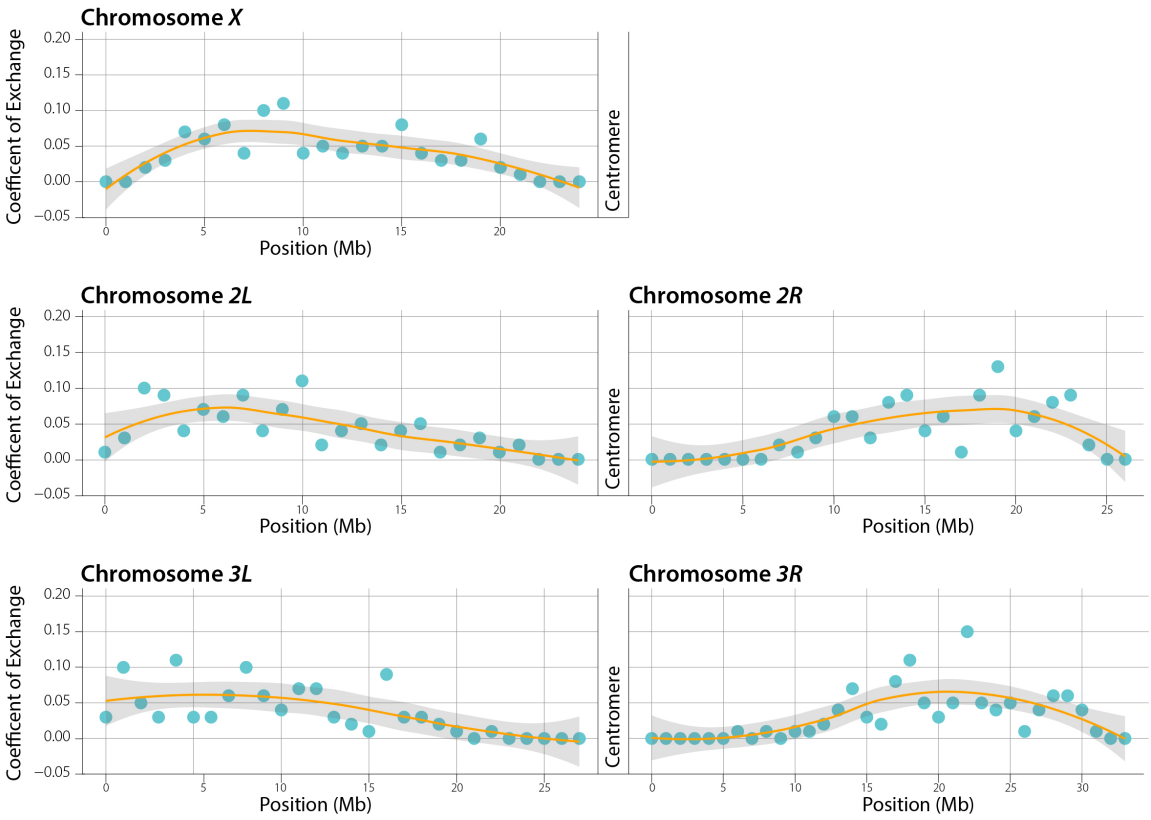

Figure 2.2. Coefficient of exchange.
COs are plotted in 1-Mb intervals for the five major chromosome arms. The orange line is a best-fit of the data and the gray shaded area indicates the 95% confidence interval of the best-fit line. The centromere effect is apparent along the four autosomal arms and the less pronounced telomere effect is apparent along all five arms.

Although the four autosomal arms displayed a relative paucity of centromere-proximal COs, a diminished centromere effect was seen on the *X* chromosome, with 59% of the SCOs occurring in the distal one-half of the chromosome arm **(Figure 2.1, Table 2.S1)**. This result parallels observations made in previous genetic studies ***(Baker and Hall 1976; Lindsley and Sandler 1977; Page et al. 2007)*** and was not unexpected because the *X* chromosome has a large block of heterochromatin residing between the centromere and the euchromatin that buffers the distance over which the centromere effect may act (Yamamoto and Miklos 1977; Lindsley and Zimm 1992). Indeed, studies have shown that the frequency of crossing over in the

centromere-proximal euchromatin of the *X* chromosome can be greatly reduced simply by deleting large blocks of proximal heterochromatin (Yamamoto and Miklos 1978).

Double crossovers and crossover interference

In *D. melanogaster*, researchers have traditionally used a limited number of variably spaced visible markers to measure recombination on each chromosome. (Commonly used markers and their approximate locations are shown in **Figure 2.1**.) Using this method of analysis, a CO event occurring distal to the most distally located visual marker on a chromosome would not be evident. Importantly, the distal CO may be part of a DCO that would thus be scored as an SCO instead. Similarly, a small DCO occurring between two adjacent visual markers would also be concealed using standard recombination assays.

As anticipated from studies dating back to Weinstein (1918), we recovered far more SCOs, fewer chromatids that did not experience a crossover event (or parental chromatids), fewer DCOs, and fewer triple crossovers (TCOs) per chromosome arm than expected by chance (**Table 2.S6**, chi-square $p < 0.0001$, based on 100,000 trials of randomly distributing 541 COs among 980 chromosome arms). Using WGS after a single round of meiosis further allowed us to precisely measure the distance between each crossover of a DCO and to identify closely spaced DCOs between visual markers, which could be missed by traditional recombination analysis. For example, the small 1.5-Mb DCO recovered between *dp* and *b* on *2L* **(Figure 2.1)** may not have been apparent using visual markers alone. Similarly, several of the SCO events, such as the distal-most SCO on *2R* **(Figure 2.1),** may also have been missed.

We recovered 52 chromatids with DCOs **(Figure 2.1, Figure 2.S3, Figure 2.S4, and Table 2.S5).** Specifically, we identified 13 DCOs on the *X* chromosome, nine on *2L*, five on *2R*, 15 on *3L*, and 10 on *3R*. The vast majority of all DCOs recovered were widely spaced, with an average distance between them of 10.5 Mb, significantly larger than expected by chance ($p < 0.0001$, expected average distance 8.1 Mb, binomial test based on 100,000 trials of randomly distributing 541 COs among the entire length of 980 chromosome arms). Of the 14 total DCOs recovered on the 2^{nd} chromosome, one was the largest DCO observed in this study (19.9 Mb in male cs8.6; see Methods for an explanation of naming conventions) and two were among the smallest recovered (1.5 Mb in male w4.8; and 4.0 Mb in male w12.2) **(Figure 2.1, Figure 2.S3, Figure 2.S4, and Table 2.S5)**. Recovery of DCOs both as small as 1.5 Mb and as large as 19.9 Mb was unexpected. To determine how often we would expect to recover DCO events of these sizes by chance, we randomly distributed 52 DCO events across each of the five major chromosome arms and recovered a 2.0-Mb or smaller DCO in only 0.1% of 100,000 trials and a 19-Mb or greater DCO in 0.2% of trials.

Interestingly, we also found that the strength of interference differed between chromosomes. Although the two arms of the 2^{nd} chromosome had a similar, albeit slightly greater, number of SCO events as the other chromosome arms (*X*: 86, *2L*: 88, *2R*: 90, *3L*: 84, *3R*: 86), we observed proportionally fewer DCOs on the 2^{nd} chromosome compared to the *X* and the 3^{rd} chromosomes ($p = 0.027$, Fisher's exact one-tailed test). If the strength of interference were equal across chromosome arms, we would expect all chromosomes to have a similar number of doubles. However, the number of doubles observed on the 2^{nd} was proportionally

about half that of the *X* and *3rd* chromosomes, driven, in part, by the paucity of doubles observed on *2R*.

These findings demonstrate the ability of crossover interference to influence the distribution of exchange events (Muller 1916; Lindsley and Sandler 1977; Berchowitz and Copenhaver 2010). They also may suggest that interference may not act equally among the five major chromosome arms. Indeed, the paucity of doubles observed on the *2nd* compared to the *X* and *3rd* suggest that interference may act differently on the *2nd* chromosome than it does on the other chromosome arms **(Figure 2.1)**.

To more quantitatively describe the distribution of chromosome arms that experienced no, one, or two CO events (denoted as E0, E1 and E2 bivalents, respectively), we employed the algebraic approach developed by Weinstein (1918) **(Table 2.S7)**. Our estimates of the frequencies of E1 and E2 bivalents are consistent with those obtained using data from much larger genetic studies of recombination on the *X, 2nd*, or *3rd* chromosomes (Baker and Carpenter 1972; Parry 1973; Page *et al.* 2007; Collins *et al.* 2014). Specifically, our E2 values of 27% for the *X*, 18% for *2L*, and 31% for *3L* are similar to previously published datasets with much larger *n* values, and our E2 value of 10% for *2R* is identical to a previously published study (Parry 1973) **(Table S7)**. Taken together, these observations provide additional evidence that mechanisms of crossover control may differ between chromosome arms.

NCOs fail to show interference and are insensitive to the centromere effect

Although much is known about the distribution of COs in *D. melanogaster*, less is known about the genome-wide distribution of NCOs. Because of the challenge of identifying the precise location of large numbers of NCOs after a single meiosis, it has been unclear whether NCO events follow the same rules pertaining to interference and the centromere effect as COs.

The 291 NCOs identified in our study contained an average of 5.0 SNPs per event (min 1, max 35; **Table 2.S2**). The average maximum conversion tract length, defined as the distance between unconverted polymorphisms, was 1,421 bp **(Figure 2.S5A, Table 2.S2)**; the average minimum conversion tract length, defined as the distance between the first and last converted polymorphism, was 290 bp **(Figure 2.S5B, Table 2.S2)**. The maximum likelihood estimate for the average tract length was found to be 440–442 bp (see Methods). This is consistent with previously reported estimates of conversion tract lengths, which range from 352 bp to 441 bp in studies using the *rosy* marker (Hilliker *et al.* 1994; Blanton *et al.* 2005) and from 476 bp to 518 bp in studies using WGS (Miller *et al.* 2012; Comeron *et al.* 2012). In addition, using maximum likelihood analysis, we find the conversion rate to be approximately 2.1×10^{-8} per base pair per meiosis, consistent with a rate of approximately 2.0×10^{-8} per base pair per meiosis reported for the *rosy* gene in two other studies (Hilliker *et al.* 1994; Blanton *et al.* 2005) and 1.8×10^{-8} per base pair per meiosis using WGS (Miller *et al.* 2012).

Determining the precise location of the observed NCOs on the genome sequence revealed 33 chromatids containing two or more conversion events, a number not significantly different than expected by chance ($p = 0.8$, based on 100,000 trials of randomly distributing 291 NCO events among 980 chromosome arms then counting the number of arms with two or more

NCOs present), with 11 of the 33 instances occurring within 4 Mb of each other (p = 0.5, based on 100,000 trials of randomly distributing 291 NCO events then counting those within 4 Mb of each other). Additionally, we found 128 instances where a conversion occurred on the same chromatid as either a single, double, or triple crossover event. Thirty-two of these conversions occurred within 4 Mb of a crossover, a number not significantly different from that expected by chance (p = 0.05, based on 100,000 trials of conservatively randomly distributing 541 CO and 291 NCO events then counting those within 4Mb of each other). Together, these data suggest that NCOs neither generate nor are subject to interference.

We next examined NCOs with respect to the centromere effect. On chromosome *2R*, although none (0/90) of the SCOs fell within the centromere-proximal one-third of the chromosome arm, 18% (11/61) of the NCOs recovered were in this region (p < 0.0001, Fisher's exact test) **(Figure 2.1)**. In addition, only 2% (3/170) of the SCOs on the *3rd* chromosome fell within the centromere-proximal one-third of either arm, whereas 27% (17/64) and 29% (19/66) of NCOs fell within that region on *3L* and *3R*, respectively (p < 0.0001, Fisher's exact test). For each chromosome arm, we modeled a random distribution of conversion events and found that the number of NCO events in the proximal 1/3 of each chromosome arm was not different from those placed by random chance (see Methods for individual chromosome arm values and 95% confidence intervals). The only exception was chromosome *2R*, in which we observed that significantly fewer NCOs occurred in the proximal 1/3 of the chromosome arm than expected by chance (p = 0.03). Note that although only 9% of the SNPs on chromosome *2R* are found in the centromere-proximal 1/3 of the arm, all tests are based on the conservative assumption that SNPs are equally distributed along the chromosome arm, suggesting that this deviation is,

in fact, explained by low SNP density in this region. These data therefore suggest that NCOs in *D. melanogaster* are insensitive to the centromere effect. Indeed, the paucity of SNPs in the most proximal region of each chromosome arm, including *2R*, prevents us from determining exactly how close to the euchromatic/heterochromatic boundary conversions may occur **(Figure 2.1)**, thus we are likely underestimating the frequency of proximal conversion events.

To estimate the number of NCOs we may have missed due to decreased SNP density, we used data from the 291 NCOs we recovered to estimate the genome-wide NCO rate to be 2.1×10^{-8} conversions per base pair per meiosis. Applying this rate to the entire 132.5 Mb haploid genome (excluding the *Y* and *4th* chromosomes and unmapped heterochromatic regions) yields 2.8 recoverable conversions per haploid meiosis (132.5 Mb x 2.1×10^{-8}). Thus, in the 196 individual flies examined in this study, we should have recovered approximately 549 NCO events had all events been equally detectable. Our observation that the NCOs we observe are insensitive to the centromere effect and to interference suggests that DSBs produced either near the centromere or in proximity to another DSB are preferentially repaired as NCOs. It is therefore likely that many of the 258 conversions we failed to detect (549 expected – 291 detected) occurred in regions of low SNP density, such as in the centromere-proximal euchromatic regions or in SNP deserts that occur randomly throughout the genome.

Recovery of complex NCO events

Unbiased recovery of NCO events on a genome-wide scale allows for the identification of unexpected meiotic repair products. We recovered three discontinuous NCOs on chromosome *2R* that appear to be the result of either a mitotic repair event or a complex meiotic repair

event. (When counting NCOs we considered these three discontinuous tracts as single NCO events unless otherwise noted.) All three discontinuous repair events appear as two short conversion tracts with a non-converted SNP between them **(Figure 2.3)**. These events appear remarkably similar to a complex conversion event at *rosy* recovered by Carpenter (1982) and analyzed at the molecular level by Curtis and Bender (1991). There are several processes, including bidirectional repair or template switching during repair, that may have given rise to these events (Merker *et al.* 2003; Whitby 2005). Recovery and identification of more complex repair events such as this, perhaps by methods designed to enrich for them, would certainly contribute to the mechanistic understanding of the repair processes at play during *D. melanogaster* female meiosis.

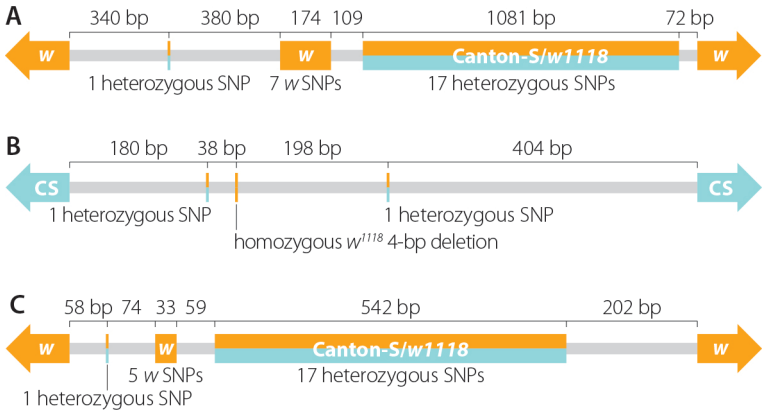

Figure 2.3. Recovery of complex NCO repair events.
We recovered three instances of complex NCO repair similar to an event recovered by Carpenter (Carpenter 1982) and described by Curtis and Bender (Curtis and Bender 1991). Exact coordinates for each NCO can be found in **Table S2** and are based *D. melanogaster* genome release 6 (dm6).

Transposable elements mediate copy number variation in Drosophila

In addition to the recovery of complex meiotic repair events, whole-genome sequencing of individual flies also allowed us to observe evidence of ectopic exchange events mediated by transposable elements (TEs). TEs are mobile DNA elements that can replicate within a genome by moving into or near genes, sometimes with deleterious effects to the host. TEs have been shown to be an important component of genome evolution and are thought to cause large deletions or duplications through ectopic exchange, or unequal crossing over either between homologs or sister chromatids **(Figure 2.4)** (Kaminker et al. 2002; Lee and Langley 2012). These copy-number variants (CNVs) may be visualized by plotting depth-of-coverage for an entire genome or region of interest **(Figure 2.4A)**. We recovered one DCO (male cs14.5) on chromosome *2L* in which the proximal of the two COs occurred at the same position as a TE present in the w^{1118} parental line but not in the Canton-S parental line **(Figure 2.1, Table 2.1)**. The position of this proximal CO also defined a change in read depth, with approximately 50% higher read depth on the distal side of the CO than on the proximal side **(Figure 2.5A)**. Plotting read depth for the entire chromosome arm revealed a 212-kb duplication precisely defined by two TEs, with the distal TE present only in the Canton-S parental line and the proximal TE present only in the w^{1118} parental line **(Table 2.1)**. We then created depth-of-coverage graphs for all 196 males sequenced in this study and identified three additional CNVs 10 kb or greater in size that were present only in individual male offspring **(Figure 2.5B–D, Table 2.1)**, as well as four CNVs shared among multiple male siblings **(Figure 2.5E–H, Table 2.1)**.

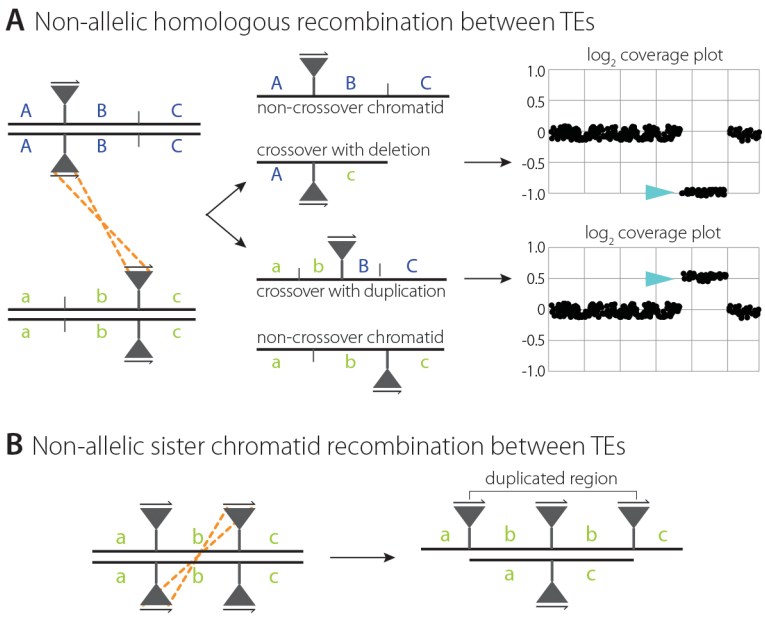

Figure 2.4. Model of unequal exchange between homologous chromosomes or sister chromatids.
(A) Non-allelic homologous recombination between identical TEs on homologous chromosomes creates a CO with one chromatid carrying a duplication and another carrying a deletion. Expected $\log_2$ depth-of-coverage graphs are shown for autosomal duplications and deletions. (B) Unequal sister chromatid exchange between identical TEs creates one sister chromatid carrying a duplication and one with a deletion. Note that in these models TEs are oriented in the same direction.

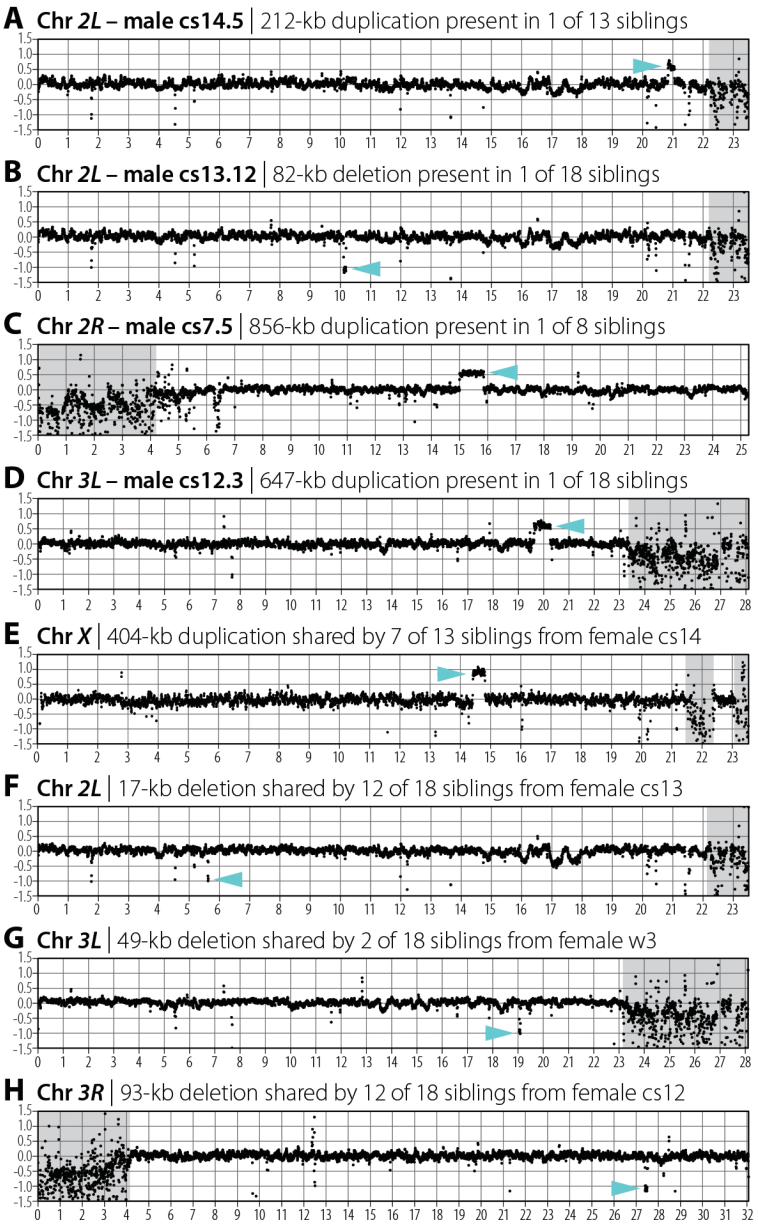

Figure 2.5. Large *de novo* and inherited CNVs.
Log_2 depth of coverage for each chromosome arm is shown. Alignment of reads to heterochromatic regions (shaded in gray) is poor. Siblings are the number of males sequenced from an individual female. Arrowheads (blue) indicate the position of a CNV. Note that there appear to be many small CNVs (represented by single dots) along each chromosome arm that are simply differences between each stock and the *D. melanogaster* reference genome. (A–D) Candidate *de novo* events that were observed in only one male. (E–H) Representative CNVs that were inherited by more than one male from either their heterozygous mother or homozygous father.

Figure	CNV Type	Chr	Proximal Coordinate	Distal Coordinate	Proximal Feature[a]	TE orientation[b]	Distal Feature	TE orientation
***De novo* CNVs**								
6A	Duplication	*2L*	20,845,594	21,057,582	w: - cs: Roo	- 3'-5'	w: Roo cs: -	unknown -
6B	Deletion	*2L*	10,113,178	10,194,791	w: - cs: McClintock	- unknown	w: - cs: -	- -
6C	Duplication	*2R*	15,005,073	15,860,851	w: hobo (DMHFL1) cs: -	3'-5' -	w: hobo (DMHFL1) cs: -	3'-5' -
6D	Duplication	*3L*	19,624,757	20,272,082	w: DMIS297 cs: DMIS297	unknown unknown	w: - cs: DMIS297	- unknown
Inherited CNVs								
6E	Duplication	*X*	14,413,980	14,817,705	w: Roo cs: -	unknown -	w: Roo cs: -	unknown -
6F	Deletion	*2L*	5,622,078	5,639,080	w: - cs: hobo (DMHFL1)	- 3'-5'	w: - cs: hobo (DMHFL1)	- 3'-5'
6G	Deletion	*3L*	19,053,516	19,102,247	w: - cs: DMIS297	- unknown	w: - cs: DMIS297	- unknown
6H	Deletion	*3R*	27,406,208	27,499,496	w: - cs: hobo (DMHFL1)	- 3'-5'	w: - cs: -	- -

Table 2.1. CNVs recovered in this study
Four CNVs (6A–6D) were recovered in only individual males and are thus likely to have arisen *de novo*. All four of these CNVs were defined by at least one TE in one of two parental genomes, and one CNV (6A) defined the proximal CO of a DCO event. Four CNVs (6E–6H) were shared among multiple individuals, and all four were defined by at least one TE present in one of the two parental genomes. The Figure column lists the panel in Figure 5 to which each event corresponds.

[a] Roo, McClintock, hobo, and DMIS297 are different TE families; w = w^{1118}; cs = Canton-S
[b] TE: transposable element

The presence of a CNV in one male on the maternal haplotype that is not present in his siblings would suggest that the CNV is a *de novo* event, thus it is likely that the CNVs in the four males represented in Figure 5A–D are indeed *de novo* CNVs. Three of these four CNVs had identical TEs present at both sides of the CNV in either one or both of the parents **(Table 2.1)**. The remaining *de novo* CNV was a deletion that contained a TE on only the distal side of the deletion in the Canton-S stock and no apparent parental TE or low-complexity sequence on the proximal side of the deletion **(Table 2.1).**

Because four of the CNVs observed were present in more than one sibling, it was presumed that these alleles were segregating in the parental germline. By analyzing TEs present in the w^{1118} or Canton-S stocks along with the flanking SNP profiles of the CNVs, we determined that three of the four inherited CNVs were likely sister chromatid recombination events mediated by identically oriented TEs **(Figure 2.4B, Figure 2.5E–H, Table 2.1)**. The remaining inherited CNV was a deletion defined on its proximal side by a TE present only in the Canton-S line but no apparent TE in either the w^{1118} or Canton-S stocks on its distal side. BLAST of read pairs from the distal side of the deletion revealed that all unmapped pairs matched a canonical hobo element (a TE family), suggesting that the deletion was mediated by the hobo element in the absence of an identical TE at the distal location.

Interestingly, four of the eight CNV events we observed lie in the proximal one-third of the chromosome arms, where crossing over is reduced. Previous studies have concluded that ectopic recombination is likely a major factor limiting the spread of TEs in natural populations (Charlesworth and Langley 1986). Our findings support this conclusion, however our data also show that ectopic recombination occurs at a significant frequency in regions of the genome with lower recombination. This is surprising because the accumulation of TEs in centromere-proximal genomic regions has historically been thought to be caused by a low rate of ectopic recombination in these regions (Charlesworth and Langley 1989; Lee and Langley 2010). Our data suggest, rather, that the reduced efficacy of selection against TE-mediated CNV formation in regions of reduced recombination may contribute to the accumulation in these regions.

DISCUSSION

Elucidating the properties controlling meiotic COs and NCOs has been of interest to Drosophila researchers for over a century. The present study helps to explain and clarify several observations that Drosophila researchers have made during that time. By examining 196 individual wild-type meiotic events, we are able to make accurate and precise observations about the number and position of COs and NCOs for the five major chromosome arms in *D. melanogaster*. We find, as expected, that COs are sensitive to the centromere effect and occur with less frequency in the centromere-proximal euchromatic regions of autosomal chromosome arms. NCOs, on the other hand, are not sensitive to the centromere effect and are often found in proximal euchromatic regions. NCOs also do not seem to be sensitive to or to generate interference and may occur close to crossover events, within double crossover events, and, surprisingly, even within close proximity to one another.

Crossover interference is evident in our dataset based on two observations. First, we recovered only 52 DCO events, significantly less than expected by random chance **(Table 2.S6)**, and only one TCO. Second, we find that DCOs are generally widely spaced, with an average distance of 10.5 Mb between the COs. Although interference is seen in many organisms from yeast to humans, the full mechanism remains a mystery (Berchowitz and Copenhaver 2010). *Drosophila mauritiana*, a species closely related to *D. melanogaster*, appears to have about twice as many COs—with more centromere-proximal COs—per chromosome arm as *D. melanogaster* (True *et al.* 1995). It would be interesting to perform an experiment in *D. mauritiana* similar to the one described here in order to obtain detailed insight into the distribution and distance between these crossovers. Because these two species are

approximately three million years divergent (Lachaise *et al.* 1986), it may be possible to identify genes or polymorphisms that play an important role in this process.

In some organisms there are two main pathways for repairing DSBs as COs. These are referred to as the ZMM-dependent, or Class I, pathway and the Mus81-dependent, or Class II, pathway (Whitby 2005). Class II crossovers, which appear to act as a "backup" system in some organisms, are infrequent and are insensitive to interference (Novak *et al.* 2001; Hollingsworth and Brill 2004). Recent work in tomato (*Solanum lycopersicum*) using high-resolution microscopy to visualize Class I and Class II crossover events (Anderson *et al.* 2014) reported fewer Class II events (18% of total COs) and found them much more often in the centromere-proximal euchromatin, suggesting that they may be less sensitive to the centromere effect than Class I events. Perhaps then, it is possible that the 1.4-Mb DCO recovered on chromosome *2L*, possibly the smallest DCO ever reported in *D. melanogaster*, is a product of the Class II, non-interfering pathway, similar to the 3.0-Mb DCO observed on the *X* and the 4.0-Mb DCO observed on *2R*. If these events are indeed Class II events, they may be the first demonstration of this pathway in *D. melanogaster*.

Several studies have identified motifs associated with sites of crossing over in both *D. melanogaster* (Miller *et al.* 2012; Comeron *et al.* 2012; Singh *et al.* 2013) and other species of Drosophila (Cirulli *et al.* 2007; Kulathinal *et al.* 2008; Stevison and Noor 2010). Analyzing our current dataset, we detected no motifs enriched over background using 201 of the SCO events defined by two SNPs 500 bp or less apart **(Figure 2.S6)**. This finding suggests that either several motifs may be associated with COs, as reported by Comeron and colleagues (2012), or that crossovers in *D. melanogaster* may be associated with open chromatin and transcription in

early meiosis (Adrian and Comeron 2013). Either one of these possibilities would make a true crossover-associated motif difficult to detect in a dataset of this size.

We recovered 541 total CO events and expected to recover 549 NCO events in our study (291 NCO events recovered, 258 NCO events not recovered). Early studies of NCO events at the *ry* locus recovered a significantly higher number of NCOs than COs (Chovnick *et al.* 1971), leading us to wonder if we were underestimating the total number of NCO events we expected to recover. The increased ratio of NCO:CO events at *ry* in previous studies makes sense in light of our finding that NCO events are evenly distributed along the chromosome arm while CO events are shifted to the distal 2/3 of the chromosome arm **(Figure 2.1)**. Indeed, only four of 41 SCO events (9%) on *3R* occurred between *ry* (*3R*:13,032,528–13,038,020, Cytological location 87D) and the centromere, demonstrating that few CO events are expected in this region. In addition, a smaller analysis of only the *X* chromosome from 30 individual males using the same genetic background recovered 15 COs, five NCOs, and calculated seven additional NCO events were not recovered (Miller *et al.* 2012)—similar to the NCO:CO ratio observed in the current study.

During prophase I of meiosis I, approximately 11–17 DSBs are produced per oocyte (Mehrotra and McKim 2006). Our likelihood analysis shows that 11 (2.8 per meiotic product times 4 haploid products of meiosis) of these DSBs will be repaired as NCO events. DSBs may also be repaired as COs and we recover an average of 2.8 COs per individual (541 COs/196 individuals) in this study. Because COs between homologs are apparent on only two chromatids, we estimate that 5.6 total COs (2.8 COs per individual * 2) are produced during a single meiosis, the same number reported in early estimates of 5.6 exchanges per meiosis

(Lindsley and Grell 1967; Carpenter 1982). Therefore the observed number of CO events plus our estimate of the total number of NCO events likely account for nearly all the DSBs formed during meiosis.

The reported number of DSBs in repair-deficient mutants, 19–23, is somewhat higher than that observed in wild-type (Jang *et al.* 2003; Mehrotra and McKim 2006; Lake *et al.* 2013). This difference may be because repair-deficient mutants may create more DSBs than wild-type, as resolution of breaks as either COs or NCOs may provide feedback that limits the number of DSBs produced (Thacker *et al.* 2014). If this is the case, the absence of this feedback may artificially inflate the number of DSBs expected in a wild-type background and 19–23 may not be an accurate number of DSBs in *D. melanogaster*. If the true number of DSBs created during meiosis is indeed slightly higher than the 11–17 reported in wild-type then our analysis may be missing evidence of other repair events. Examples of these events include mismatch repair of a NCO that may mask an identifying SNP, causing an underestimation of the number of NCO events (Radford *et al.* 2007), or nonhomologous end joining or sister chromatid repair may resolve a DSB in a way that is undetectable using SNP or InDel polymorphisms (McVey *et al.* 2004; Johnson-Schlitz *et al.* 2007; Goldfarb and Lichten 2010).

Unexpectedly, we identified four large *de novo* and four large inherited CNVs in the 196 individual male genomes that we studied. Previous studies of TE-mediated copy number variation in Drosophila focused on assaying unequal exchange between one family of TEs, *roo* elements, near the *white* locus (Davis *et al.* 1987; Montgomery *et al.* 1991). Separately, a screen for *de novo* mutations resulting in eye color changes recovered five large deletions that were presumed to be the result of unequal exchange between TEs at different genomic

positions (Watanabe *et al.* 2009). Using these five deletions, Watanabe and colleagues (2009) estimated the mutation rate for large deletions and duplications affecting multiple genes to be 1.7%, remarkably close to the 2% rate we observe for *de novo* CNVs in our study. The similar mutation rate observed in both studies supports the hypothesis that ectopic recombination is a common source of genetic variation in *D. melanogaster* and demonstrates the value of unbiased sequencing of individual meiotic products.

With one significant exception (Parry 1973), all previous genetic studies of recombination in Drosophila have focused on a single chromosome arm or studied offspring from recombinant inbred lines (Comeron *et al.* 2012). Ours is the first to characterize both NCOs and COs on all five major arms within a single meiosis. Our data show that the processes that position CO events are clearly distinct from those that position NCO events. It will be of considerable interest, as sequencing technology improves and declines in cost, to repeat this analysis in the presence of polar effect mutants or genotypes that elevate recombination. In addition, controlled crossing experiments such as ours should be repeated using recently isolated and characterized wild-type lines, which may carry polymorphisms affecting the distribution or number of meiotic repair events (Mackay *et al.* 2013; Lack *et al.* 2015), as it would be informative to identify lines in which these properties are significantly different from those described in this study or from each other. Much of this will involve repeating 20th century genetic assays with 21st century genomic approaches. But the goal will remain unchanged—to identify the mechanisms that ensure the proper number and position of exchanges and thus ensure the proper segregation of homologs at the first meiotic division.

METHODS

Fly Stocks and husbandry

Lab strains of w^{1118} and Canton-S were isogenized as described in (Miller *et al.* 2012). w^{1118} was isogenic for all four chromosomes while Canton-S was heterozygous for the fourth chromosome as well as for 15,718 SNPs along 3.9 Mb of *2R*, from 2R:21,413,827 to the telomere. All flies were kept on standard cornmeal-molasses and maintained at 25^{o}C.

DNA preparation, sequencing, alignment, and SNP calling

DNA for individual flies was prepared from single adult males using the Qiagen DNeasy Blood & Tissue Kit. DNA from parental lines was prepared from males and females. All flies were starved for 4 hr before freezing at -80^{o}C for at least 1 hr. One µg of DNA from each was fragmented to 250-bp fragments using a Covaris S220 sonicator (Covaris Inc.) by adjusting the treatment time to 85 sec. Libraries were prepared using a Nextera DNA Sample Prep Kit and Bioo Scientific NEXTflex™ DNA Barcodes. The resulting libraries were purified using Agencourt AMPure XP system (Beckman Coulter) then quantified using a Bioanalyzer (Agilent Technologies) and a Qubit Fluorometer (Life Technologies). For the first batch of 98 individual male flies libraries were pooled into four groups and were run in four lanes each of an Illumina HiSeq 2500 instrument on either a 150-bp or 100-bp paired-end flowcell **(Table 2.S3)**. For the second batch of 98 individual male flies libraries were pooled into four groups and run in two lanes each of an Illumina HiSeq 2500 instrument on a 125-bp paired-end flowcell **(Table 2.S3)**. For all runs HiSeq Control Software 2.0.12.0 and Real-Time Analysis (RTA) version 1.17.21.3 were used. Secondary

Analysis version CASAVA-1.8.2 was run to demultiplex reads and generate FASTQ files. Alignment to the *D. melanogaster* reference genome (dm6, UCSC) was performed using bwa version 0.7.7-r441 (Li and Durbin 2009). After alignment, Picard and GATK were used to mark duplicate reads and perform local realignment around InDels (McKenna *et al.* 2010). SNPs were identified using SAMtools version 0.1.19-44428cd and BCFtools version 0.1.19 (Li, Handsaker, *et al.* 2009).

Males were numbered based on if their father was homozygous w^{1118} or Canton-S and the number of their heterozygous mother. For example, male cs12.3 had a Canton-S father, its mother was female number 12, and it was the third male selected for DNA extraction. Sibling numbers may not be continuous, as males with low DNA concentrations after DNA extraction were not selected for sequencing.

Sequencing data from this project has been deposited at NCBI under the following project numbers: PRJNA285112 (isogenized w^{1118} and Canton-S parental stocks) PRJNA307070 (196 individual males).

Identification of sites of crossing over and gene conversion

Parental SNPs with quality scores greater than 220 and a read depth of at least 20 were used to identify CO and NCO events in offspring. Only locations with a SNP present in one parent and a reference allele in the other parent were considered in subsequent analysis of the offspring. For each offspring, SNPs with a quality score less than 200 and read depth below 10 were omitted from analysis. For the hemizygous *X* chromosome, instances where the parent of origin

switched from one stock to another were flagged as sites of a potential meiotic event. For the autosomes the same strategy was used except candidate events were flagged when the parent of origin switched from either a single parent to both parents or from both parents to a single parent of origin. Each putative CO and NCO event was then visually validated using IGV (Thorvaldsdottir *et al.* 2013). No CO events were excluded based on visual observation. Events flagged as potential NCOs that were due to local misalignment were excluded. While performing data analysis we found that lower quality thresholds for SNP calling in either parents or offspring resulted in a high number of false positive NCO events, with the overwhelming majority due to nearby InDel polymorphisms or low-complexity sequence. Scripts used to align genome data, call SNPs, and identify CO and NCO sites can be found at https://github.com/danrdanny/2016_CO_NCO_Paper.

Validation of NCO events by PCR

To verify the accuracy of NCO identification, 47 of 291 NCOs (16%) were validated by PCR and Sanger sequencing; Phusion polymerase (New England Biolabs) was used according to the manufacturer's instructions. All 47 NCOs validated as real. Primers and annealing temperatures are given in **Table 2.S4**.

Calculation of NCO tract length and conversion rate

To jointly estimate the rate of NCO events and NCO tract length, we used the maximum likelihood (ML) approach modified from (Miller *et al.* 2012). This method accounts for variable

spacing between SNPs by taking into account the likelihood that a DSB fated to become an NCO gene conversion occurred within the span of neighboring SNPs of arbitrary distance, but the conversion tract failed to extend far enough to allow conversion to be seen. Using the entire distribution of distances between unconverted SNPs and the positioning of converted SNPs, we jointly estimate the per-base rate of NCO-fated DSB formation and the tract length parameter, modeled as a geometric process. This allows us to estimate the genome-wide rate of DSB formation and tract length considering the fact that some NCO conversion events will be missed. Since estimation of NCO tract length is difficult when spans between SNPs are large, we first jointly estimated the ML NCO rate and tract length parameters using 225 of 291 conversion tracts in which the distance between the converted and unconverted SNPs on both the left and right side of the NCO tract was less than 1 kb. We then fixed the tract length parameter and determined the ML NCO rate parameter using 286 of the 291 NCO events defined by SNPs closer than 10 kb apart or that were not part of discontinuous repair tracts. Mathematica scripts used to estimate NCO rate and tract length can be found at https://github.com/danrdanny/2016_CO_NCO_Paper.

Motif searching with MEME

To test for the presence of a motif enriched in or around sites of COs, we used MEME version 3.9.0 (Bailey and Elkan 1994; Bailey *et al.* 2006) to search the sequence surrounding 201 SCOs defined by polymorphisms ≤500 bp apart **(Table 2.S1, Figure 2.S5C)**. To account for factors acting outside of the apparent CO interval, the search window was expanded to include 1 kb upstream and downstream of each CO interval. We searched for motifs 5–12 bp long. To create

a background distribution of motifs, we performed 100 trials where 201 COs were randomly placed along the five major chromosome arms with CO lengths randomly determined to be between 11 and 500 nucleotides long. Four motifs, [AT]GC[TA]GC[TA]GC[AT]GC[TA], ATAT[AG]TA[TC]ATAT, [TGC][TGC]TGGCCA[ACG][ACG], and AA[TA]T[GT][CA]A[AT]TT **(Figure 2.S6)** were found to be significantly enriched in the observed CO spans but were also found in 21 or more of our randomly sampled CO intervals, suggesting that they are unlikely to be real.

Statistical methods and modeling

The probability of recovering the observed number of SCO, DCO, and TCO or greater events was calculated by randomly distributing 541 CO events among 980 chromosome arms. Observed and expected values based on 100,000 trials can be found in **Table 2.S7**.

An expected distribution of distance between DCO events was created by conservatively assuming equal numbers of COs across the five chromosome arms of 196 individual flies and distributing 541 CO events randomly among 980 chromosome arms in 100,000 trials. The average distance between randomly placed DCO events was calculated to be 8.1 Mb.

An expected distribution of the distance between randomly distributed COs and NCOs was created by placing 541 COs and 291 NCO events randomly along the five major chromosome arms. The distance between events occurring on the same arm was then calculated. If one NCO and two COs occurred, then the distance from the NCO to both COs was calculated. If two NCOs and one CO occurred, then the distance from each NCO to the CO was calculated. The observed average distance between a CO and NCO was 8.4 Mb, the expected

average distance was 8.8 Mb based on 100,000 trials assuming uniform distribution of both NCO and CO events randomly placed on 980 chromosome arms. Similarly, the expected number of two or more NCO events per chromosome arm and the distance between multiple NCO events per chromosome arm was created by 100,000 trials of randomly placing 291 NCO events along the five major chromosome arms.

Similarly, to determine if NCOs are shifted with respect to the centromere, 291 NCO events were randomly distributed among the five major chromosome arms and the number of NCOs in the proximal 1/3 of each chromosome arm was calculated. The observed percentage of NCO events in the proximal 1/3 of each chromosome arm was: *X*: 26%, 95% CI: 22–46%; *2L*: 27%, 95% CI: 22–47%; *2R*: 18%, 95% CI: 21–45%; *3L*: 27%, 95% CI: 22–45%; *3R*: 29%, 95% CI: 21–45%; 100,000 trials were performed to calculate the confidence interval for each arm.

SUPPLEMENTAL FIGURES & TABLES

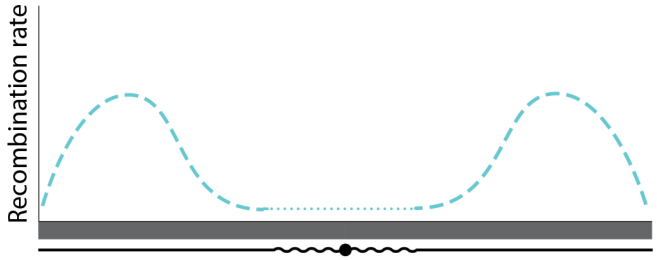

Figure 2.S1. Recombination rate is non-uniform in Drosophila.
Representation of the non-uniform rate of recombination for a metacentric chromosome, similar to what is observed for the five major *D. melanogaster* chromosome arms. The centromere effect shifts recombination away from the centromere and the milder telomere effect shifts recombination away from the telomere (per-chromosome population-wide estimates available in Mackay *et al.* 2013).

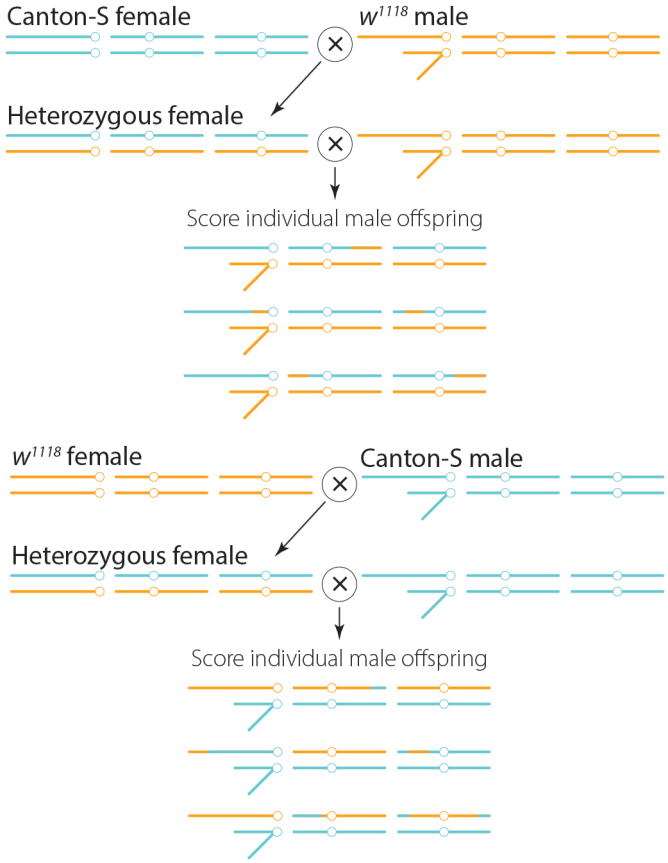

Figure 2.S2. Cross scheme.
Half of the 196 males sequenced in this study were generated by crossing isogenized Canton-S females to isogenized w^{1118} males, and half were generated by the reciprocal cross of isogenized w^{1118} females to isogenized Canton-S males. Individual heterozygous females were recovered, crossed to either individual w^{1118} or Canton-S males, and recombinant male offspring were analyzed.

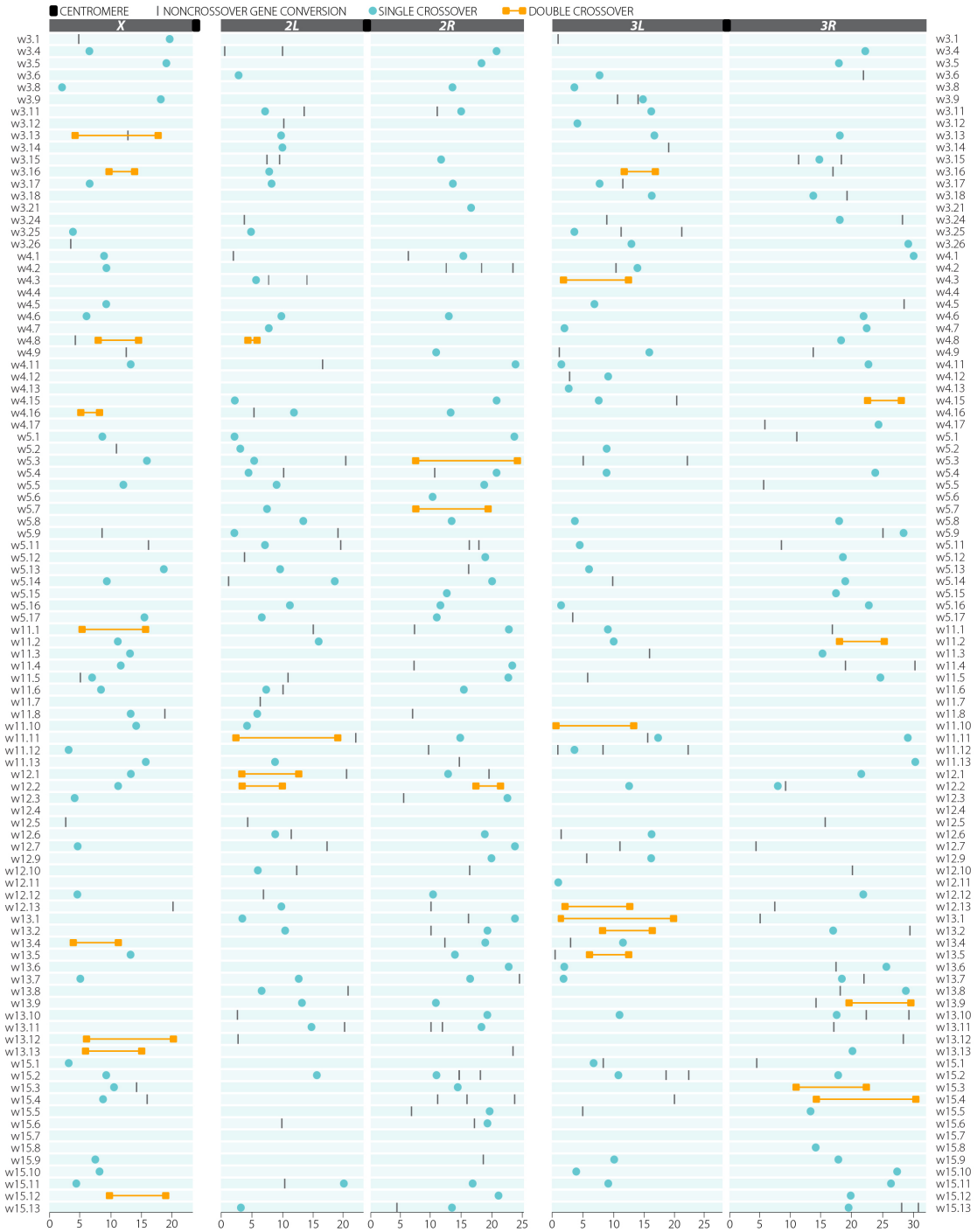

Figure 2.S3. Meiotic events recovered from 98 individual males from *w*^{1118} fathers.
Each row represents a single male analyzed and each column represents one of the five major *D. melanogaster* chromosome arms. Centromeres are located on the right side of the *X* chromosome, between *2L* and *2R*, and between *3L* and *3R*. The axis of each chromosome arm is shown in Mb. SCOs are represented by individual circles, DCOs by connected boxes, and NCOs by lines. Note that NCO events occur both near and within COs and within close proximity of one another, thus failing to demonstrate interference.

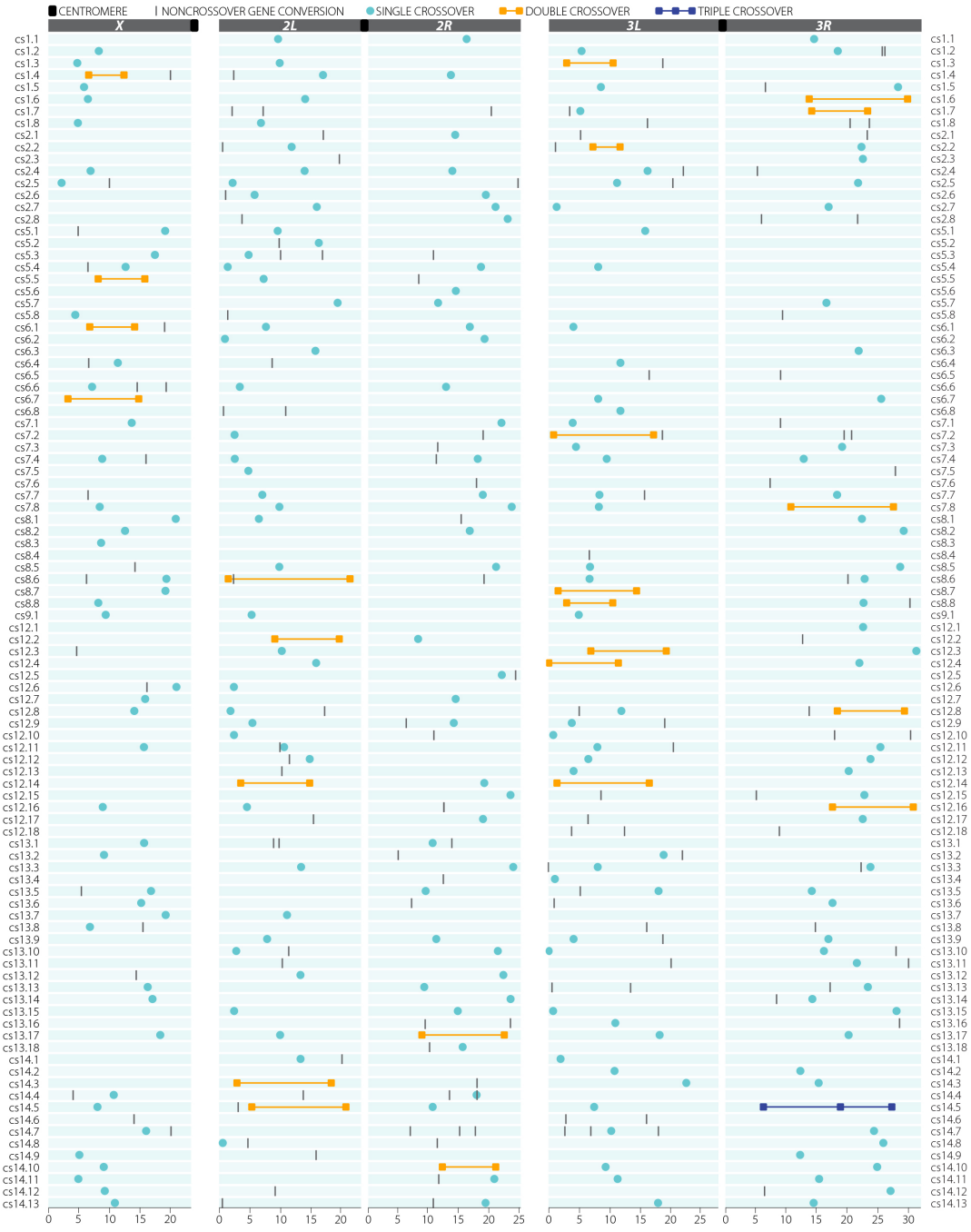

Figure 2.S4. Meiotic events recovered from 98 individual males from Canton-S fathers.
Each row represents a single male analyzed and each column represents one of the five major *D. melanogaster* chromosome arms. Centromeres are located on the right side of the *X* chromosome, between *2L* and *2R*, and between *3L* and *3R*. The axis of each chromosome arm is shown in Mb. SCOs are represented by individual circles, DCOs by connected boxes, and NCOs by lines. Three connected boxes represent the TCO recovered on *3R* in stock cs14.5. Note that NCO events occur both near and within COs and within close proximity of one another, thus failing to demonstrate interference.

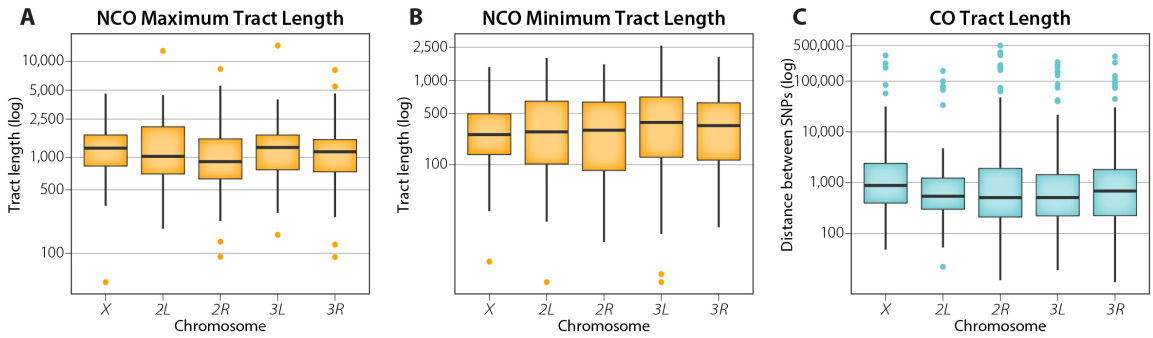

Figure 2.S5. Distribution of CO and NCO event sizes.
The middle bar of each box represents the mean, and the upper and lower boundaries of the box represent the 1st and 3rd quartiles, respectively. (A) Box-and-whisker plot of maximum conversion tract sizes for all 291 NCOs recovered. (B) Box-and-whisker plot of minimum conversion tract sizes for all 291 NCOs recovered. (C) Box-and-whisker plot for all 541 COs recovered in this study. COs show a much wider distribution than NCO events do, with some being defined by gaps of more than 100 kb.

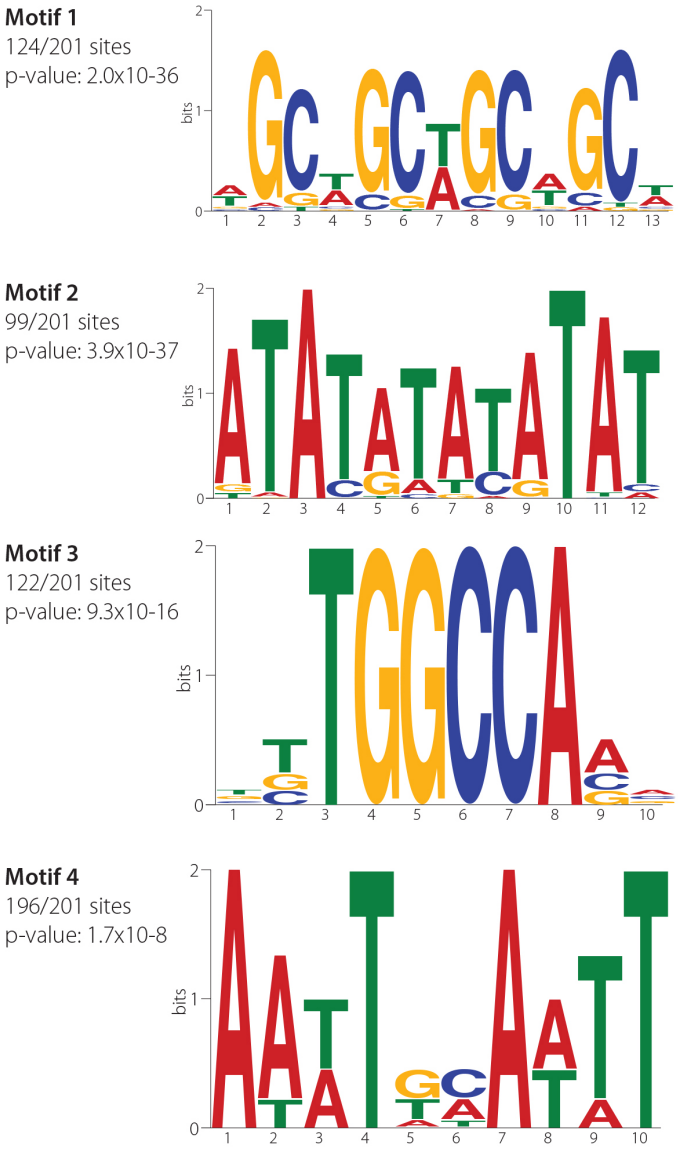

Figure 2.S6. Non-significant motifs recovered in this study.
Using MEME (Bailey and Elkan 1994), we identified four motifs significantly enriched around 201 SCOs defined by polymorphisms ≤500 bp apart **(Table S1)**. We performed 100 trials of randomly sampling 201 crossover events and found all four motifs significantly enriched in at least 21 of 100 trials, thus these motifs were all false-positive findings.

Table 2.S1. Detailed information on all 541 crossovers recovered in this study.

Chr	Stock	Class	5' Parent	5' SNP	3' Parent	3'SNP
chrX	cs2.5	SCO	Canton-S	2,432,574	w^{1118}	2,432,632
chrX	w3.8	SCO	Canton-S	2,495,015	w^{1118}	2,495,709
chrX	cs6.7	DCO	Canton-S	3,468,155	w^{1118}	3,472,332
chrX	w11.12	SCO	w^{1118}	3,548,268	Canton-S	3,550,070
chrX	w15.1	SCO	w^{1118}	3,565,970	Canton-S	3,566,269
chrX	w3.25	SCO	Canton-S	4,246,446	w^{1118}	4,246,782
chrX	w13.4	DCO	w^{1118}	4,303,012	Canton-S	4,304,169
chrX	w12.3	SCO	w^{1118}	4,523,250	Canton-S	4,523,840
chrX	w3.13	DCO	Canton-S	4,594,909	w^{1118}	4,595,412
chrX	cs5.8	SCO	Canton-S	4,646,302	w^{1118}	4,651,444
chrX	w15.11	SCO	w^{1118}	4,789,599	Canton-S	4,792,203
chrX	w12.12	SCO	Canton-S	4,962,689	w^{1118}	4,963,090
chrX	cs1.3	SCO	Canton-S	4,981,299	w^{1118}	4,981,624
chrX	w12.7	SCO	Canton-S	5,006,355	w^{1118}	5,011,442
chrX	cs1.8	SCO	w^{1118}	5,073,831	Canton-S	5,074,709
chrX	cs14.11	SCO	Canton-S	5,145,127	w^{1118}	5,146,150
chrX	cs14.9	SCO	w^{1118}	5,313,510	Canton-S	5,314,395
chrX	w13.7	SCO	Canton-S	5,437,853	w^{1118}	5,437,963
chrX	w4.16	DCO	w^{1118}	5,511,783	Canton-S	5,512,607
chrX	w11.1	DCO	w^{1118}	5,717,433	Canton-S	5,717,848
chrX	cs1.5	SCO	Canton-S	6,070,261	w^{1118}	6,071,438
chrX	w13.13	DCO	Canton-S	6,274,810	w^{1118}	6,275,098
chrX	w13.12	DCO	Canton-S	6,435,748	w^{1118}	6,436,145
chrX	w4.6	SCO	Canton-S	6,435,748	w^{1118}	6,436,145
chrX	cs1.6	SCO	Canton-S	6,689,024	w^{1118}	6,695,857
chrX	cs1.4	DCO	w^{1118}	6,817,266	Canton-S	6,822,515
chrX	w3.4	SCO	w^{1118}	6,898,008	Canton-S	6,898,474
chrX	w3.17	SCO	w^{1118}	6,940,724	Canton-S	6,943,061
chrX	cs6.1	DCO	Canton-S	6,992,967	w^{1118}	6,993,084
chrX	cs13.8	SCO	Canton-S	7,021,738	w^{1118}	7,023,293
chrX	cs2.4	SCO	Canton-S	7,122,322	w^{1118}	7,123,914
chrX	w11.5	SCO	w^{1118}	7,335,724	Canton-S	7,339,991
chrX	cs6.6	SCO	w^{1118}	7,369,774	Canton-S	7,378,406
chrX	w15.9	SCO	Canton-S	7,850,857	w^{1118}	7,851,840
chrX	cs14.5	SCO	Canton-S	8,238,244	w^{1118}	8,238,465
chrX	w4.8	DCO	w^{1118}	8,312,878	Canton-S	8,314,975
chrX	cs5.5	DCO	Canton-S	8,357,722	w^{1118}	8,357,809
chrX	cs8.8	SCO	Canton-S	8,370,462	w^{1118}	8,372,265
chrX	cs1.2	SCO	Canton-S	8,459,530	w^{1118}	8,459,719
chrX	w4.16	DCO	Canton-S	8,525,007	w^{1118}	8,525,160
chrX	w15.10	SCO	w^{1118}	8,528,676	Canton-S	8,529,455
chrX	cs7.8	SCO	Canton-S	8,600,013	w^{1118}	8,600,304
chrX	w11.6	SCO	Canton-S	8,766,307	w^{1118}	8,766,911
chrX	cs8.3	SCO	Canton-S	8,823,415	w^{1118}	8,824,454
chrX	w5.1	SCO	Canton-S	8,983,530	w^{1118}	8,983,698
chrX	cs7.4	SCO	Canton-S	9,007,434	w^{1118}	9,008,395
chrX	w15.4	SCO	Canton-S	9,092,074	w^{1118}	9,092,777
chrX	cs12.16	SCO	w^{1118}	9,092,804	Canton-S	9,093,755
chrX	w4.1	SCO	w^{1118}	9,250,434	Canton-S	9,254,613

chrX	cs14.10	SCO	Canton-S	9,256,424	w^{1118}	9,258,212
chrX	cs13.2	SCO	Canton-S	9,295,221	w^{1118}	9,295,377
chrX	cs14.12	SCO	Canton-S	9,426,297	w^{1118}	9,426,345
chrX	cs9.1	SCO	Canton-S	9,573,699	w^{1118}	9,574,478
chrX	w4.5	SCO	Canton-S	9,606,848	w^{1118}	9,607,481
chrX	w15.2	SCO	Canton-S	9,611,375	w^{1118}	9,611,645
chrX	w4.2	SCO	w^{1118}	9,636,599	Canton-S	9,639,194
chrX	w5.14	SCO	w^{1118}	9,702,274	Canton-S	9,702,954
chrX	w3.16	DCO	w^{1118}	10,060,100	Canton-S	10,061,373
chrX	w15.12	DCO	w^{1118}	10,129,849	Canton-S	10,130,727
chrX	w15.3	SCO	w^{1118}	10,857,721	Canton-S	10,864,131
chrX	cs14.4	SCO	w^{1118}	10,871,561	Canton-S	10,873,767
chrX	cs14.13	SCO	Canton-S	11,050,148	w^{1118}	11,056,207
chrX	w11.2	SCO	Canton-S	11,466,164	w^{1118}	11,467,685
chrX	w12.2	SCO	Canton-S	11,522,707	w^{1118}	11,524,650
chrX	cs6.4	SCO	Canton-S	11,530,767	w^{1118}	11,533,558
chrX	w13.4	DCO	Canton-S	11,553,529	w^{1118}	11,556,662
chrX	w11.4	SCO	Canton-S	11,948,652	w^{1118}	11,949,164
chrX	w5.5	SCO	Canton-S	12,362,667	w^{1118}	12,363,123
chrX	cs1.4	DCO	Canton-S	12,523,696	w^{1118}	12,525,465
chrX	cs8.2	SCO	Canton-S	12,696,534	w^{1118}	12,697,845
chrX	cs5.4	SCO	Canton-S	12,788,323	w^{1118}	12,789,261
chrX	w11.3	SCO	Canton-S	13,438,843	w^{1118}	13,439,270
chrX	w13.5	SCO	w^{1118}	13,524,955	Canton-S	13,525,668
chrX	w11.8	SCO	w^{1118}	13,537,245	Canton-S	13,537,508
chrX	w4.11	SCO	Canton-S	13,544,655	w^{1118}	13,545,120
chrX	w12.1	SCO	w^{1118}	13,572,695	Canton-S	13,573,285
chrX	cs7.1	SCO	w^{1118}	13,778,227	Canton-S	13,779,455
chrX	w3.16	DCO	Canton-S	14,173,476	w^{1118}	14,173,806
chrX	cs12.8	SCO	Canton-S	14,197,058	w^{1118}	14,197,321
chrX	cs6.1	DCO	w^{1118}	14,246,281	Canton-S	14,246,644
chrX	w11.10	SCO	w^{1118}	14,425,176	Canton-S	14,425,590
chrX	w4.8	DCO	Canton-S	14,811,317	w^{1118}	14,811,543
chrX	cs6.7	DCO	w^{1118}	14,909,071	Canton-S	14,912,134
chrX	cs13.6	SCO	Canton-S	15,309,269	w^{1118}	15,310,584
chrX	w13.13	DCO	w^{1118}	15,313,277	Canton-S	15,313,694
chrX	w5.17	SCO	w^{1118}	15,749,882	Canton-S	15,753,494
chrX	cs12.11	SCO	w^{1118}	15,755,802	Canton-S	15,756,445
chrX	cs13.1	SCO	w^{1118}	15,782,279	Canton-S	15,783,774
chrX	cs5.5	DCO	w^{1118}	15,895,424	Canton-S	15,898,581
chrX	w11.1	DCO	Canton-S	15,938,988	w^{1118}	15,946,126
chrX	cs12.7	SCO	Canton-S	15,954,423	w^{1118}	15,957,370
chrX	w11.13	SCO	Canton-S	15,968,836	w^{1118}	15,972,102
chrX	cs14.7	SCO	Canton-S	16,113,850	w^{1118}	16,114,048
chrX	w5.3	SCO	Canton-S	16,154,599	w^{1118}	16,156,725
chrX	cs13.13	SCO	w^{1118}	16,364,539	Canton-S	16,366,418
chrX	cs13.5	SCO	Canton-S	16,901,928	w^{1118}	16,902,213
chrX	cs13.14	SCO	w^{1118}	17,108,581	Canton-S	17,166,185
chrX	cs5.3	SCO	w^{1118}	17,421,825	Canton-S	17,644,950
chrX	w3.13	DCO	w^{1118}	17,803,352	Canton-S	18,125,787
chrX	cs13.17	SCO	w^{1118}	18,301,070	Canton-S	18,485,722
chrX	w3.9	SCO	Canton-S	18,301,070	w^{1118}	18,485,722
chrX	w5.13	SCO	Canton-S	18,871,915	w^{1118}	18,876,967

chrX	w15.12	DCO	Canton-S	19,184,197	w^{1118}	19,184,687
chrX	cs5.1	SCO	Canton-S	19,185,697	w^{1118}	19,186,054
chrX	cs8.7	SCO	w^{1118}	19,228,780	Canton-S	19,228,832
chrX	cs13.7	SCO	Canton-S	19,267,635	w^{1118}	19,267,683
chrX	w3.5	SCO	w^{1118}	19,287,728	Canton-S	19,288,676
chrX	cs8.6	SCO	Canton-S	19,390,405	w^{1118}	19,390,959
chrX	w3.1	SCO	w^{1118}	19,822,035	Canton-S	19,822,731
chrX	w13.12	DCO	w^{1118}	20,420,859	Canton-S	20,420,926
chrX	cs8.1	SCO	Canton-S	20,880,939	w^{1118}	20,912,505
chrX	cs12.6	SCO	w^{1118}	20,969,339	Canton-S	21,053,189
chr2L	cs14.8	SCO	Canton-S	897,205	both	898,277
chr2L	cs6.2	SCO	Canton-S	1,251,649	both	1,251,737
chr2L	cs5.4	SCO	Canton-S	1,701,448	both	1,701,755
chr2L	cs8.6	DCO	both	1,791,196	Canton-S	1,791,580
chr2L	cs12.8	SCO	Canton-S	2,160,402	both	2,160,961
chr2L	cs2.5	SCO	Canton-S	2,510,625	both	2,511,171
chr2L	w5.9	SCO	both	2,619,282	w^{1118}	2,619,501
chr2L	w5.1	SCO	both	2,641,945	w^{1118}	2,642,804
chr2L	w4.15	SCO	w^{1118}	2,689,171	both	2,689,539
chr2L	cs12.6	SCO	Canton-S	2,713,155	both	2,713,208
chr2L	cs12.10	SCO	Canton-S	2,733,036	both	2,733,389
chr2L	cs13.15	SCO	both	2,755,517	Canton-S	2,755,836
chr2L	cs7.2	SCO	Canton-S	2,829,548	both	2,830,262
chr2L	cs7.4	SCO	both	2,862,432	Canton-S	2,862,583
chr2L	w11.11	DCO	both	2,867,730	w^{1118}	2,869,931
chr2L	cs13.10	SCO	both	3,103,251	Canton-S	3,103,273
chr2L	cs14.3	DCO	Canton-S	3,217,726	both	3,219,174
chr2L	w3.6	SCO	both	3,297,352	w^{1118}	3,297,584
chr2L	w5.2	SCO	both	3,573,367	w^{1118}	3,574,085
chr2L	w15.13	SCO	w^{1118}	3,667,432	both	3,667,493
chr2L	cs6.6	SCO	Canton-S	3,680,561	both	3,680,687
chr2L	w12.1	DCO	w^{1118}	3,818,381	both	3,818,636
chr2L	cs12.14	DCO	both	3,830,340	Canton-S	3,830,741
chr2L	w12.2	DCO	both	3,871,360	w^{1118}	3,871,457
chr2L	w13.1	SCO	both	3,900,004	w^{1118}	3,901,241
chr2L	w11.10	SCO	both	4,655,040	w^{1118}	4,658,075
chr2L	w4.8	DCO	both	4,797,280	w^{1118}	4,797,641
chr2L	cs12.16	SCO	both	4,858,604	Canton-S	4,858,919
chr2L	w5.4	SCO	w^{1118}	4,897,172	both	4,897,839
chr2L	cs7.5	SCO	Canton-S	5,083,880	both	5,084,338
chr2L	cs5.3	SCO	both	5,119,789	Canton-S	5,120,459
chr2L	w3.25	SCO	w^{1118}	5,301,826	both	5,301,925
chr2L	cs9.1	SCO	Canton-S	5,610,309	both	5,610,814
chr2L	cs14.5	DCO	Canton-S	5,645,388	both	5,645,794
chr2L	cs12.9	SCO	both	5,751,032	Canton-S	5,751,624
chr2L	w5.3	SCO	both	5,804,919	w^{1118}	5,805,805
chr2L	cs2.6	SCO	Canton-S	6,106,536	both	6,107,160
chr2L	w4.3	SCO	w^{1118}	6,109,352	both	6,109,629
chr2L	w4.8	DCO	w^{1118}	6,263,560	both	6,268,311
chr2L	w11.8	SCO	w^{1118}	6,301,379	both	6,301,693
chr2L	w12.10	SCO	w^{1118}	6,391,857	both	6,392,050
chr2L	cs8.1	SCO	both	6,803,667	Canton-S	6,803,897
chr2L	w13.8	SCO	w^{1118}	7,014,162	both	7,015,299

chr2L	w5.17	SCO	both	7,034,334	w^{1118}	7,035,282
chr2L	cs1.8	SCO	Canton-S	7,142,976	both	7,143,745
chr2L	cs7.7	SCO	Canton-S	7,367,735	both	7,368,360
chr2L	w5.11	SCO	both	7,543,148	w^{1118}	7,544,445
chr2L	w3.11	SCO	w^{1118}	7,571,595	both	7,571,762
chr2L	cs5.5	SCO	both	7,580,688	Canton-S	7,584,439
chr2L	w11.6	SCO	w^{1118}	7,744,680	both	7,747,457
chr2L	w5.7	SCO	both	7,872,224	w^{1118}	7,872,325
chr2L	cs6.1	SCO	Canton-S	7,966,770	both	7,966,945
chr2L	cs13.9	SCO	Canton-S	8,120,931	both	8,199,459
chr2L	w4.7	SCO	w^{1118}	8,120,931	both	8,199,459
chr2L	w3.16	SCO	both	8,229,596	w^{1118}	8,230,053
chr2L	w3.17	SCO	both	8,608,924	w^{1118}	8,612,126
chr2L	w11.13	SCO	w^{1118}	9,162,578	both	9,164,368
chr2L	w12.6	SCO	w^{1118}	9,196,725	both	9,197,023
chr2L	cs12.2	DCO	both	9,415,071	Canton-S	9,415,530
chr2L	w5.5	SCO	both	9,417,440	w^{1118}	9,417,501
chr2L	cs5.1	SCO	Canton-S	9,878,469	both	9,879,092
chr2L	cs1.1	SCO	both	9,947,253	Canton-S	9,947,731
chr2L	w5.13	SCO	w^{1118}	9,988,204	both	9,988,393
chr2L	w3.13	SCO	w^{1118}	10,143,717	both	10,144,942
chr2L	cs8.5	SCO	both	10,147,327	Canton-S	10,148,333
chr2L	cs7.8	SCO	Canton-S	10,165,958	both	10,166,124
chr2L	w12.13	SCO	both	10,186,535	w^{1118}	10,186,910
chr2L	w4.6	SCO	w^{1118}	10,188,137	both	10,188,751
chr2L	cs1.3	SCO	Canton-S	10,213,373	both	10,213,725
chr2L	cs13.17	SCO	both	10,277,169	Canton-S	10,279,874
chr2L	w12.2	DCO	w^{1118}	10,349,447	both	10,349,955
chr2L	w3.14	SCO	both	10,352,995	w^{1118}	10,353,848
chr2L	cs12.3	SCO	Canton-S	10,556,920	both	10,557,400
chr2L	w13.2	SCO	both	10,794,673	w^{1118}	10,795,052
chr2L	cs12.11	SCO	both	10,933,745	Canton-S	10,937,320
chr2L	cs13.7	SCO	Canton-S	11,412,495	both	11,412,710
chr2L	w5.16	SCO	both	11,554,876	w^{1118}	11,555,278
chr2L	cs2.2	SCO	Canton-S	12,149,014	both	12,182,830
chr2L	w4.16	SCO	both	12,208,983	w^{1118}	12,210,743
chr2L	w13.7	SCO	w^{1118}	12,948,525	both	13,016,246
chr2L	w12.1	DCO	both	12,948,525	w^{1118}	13,016,246
chr2L	w13.9	SCO	w^{1118}	13,500,584	both	13,501,816
chr2L	cs13.12	SCO	both	13,615,010	Canton-S	13,615,545
chr2L	cs14.1	SCO	Canton-S	13,626,276	both	13,627,186
chr2L	cs13.3	SCO	Canton-S	13,713,299	both	13,714,906
chr2L	w5.8	SCO	w^{1118}	13,717,577	both	13,717,990
chr2L	cs2.4	SCO	both	14,298,837	Canton-S	14,298,963
chr2L	cs1.6	SCO	both	14,383,788	Canton-S	14,384,833
chr2L	w13.11	SCO	both	15,036,640	w^{1118}	15,036,886
chr2L	cs12.14	DCO	Canton-S	15,119,244	both	15,121,276
chr2L	cs12.12	SCO	both	15,131,505	Canton-S	15,131,609
chr2L	w15.2	SCO	w^{1118}	15,896,222	both	15,898,950
chr2L	cs6.3	SCO	both	16,061,390	Canton-S	16,062,501
chr2L	cs12.4	SCO	Canton-S	16,182,911	both	16,183,365
chr2L	w11.2	SCO	w^{1118}	16,187,855	both	16,188,361
chr2L	cs2.7	SCO	Canton-S	16,290,811	both	16,293,513

chr2L	cs5.2	SCO	Canton-S	16,614,591	both	16,615,169
chr2L	cs1.4	SCO	Canton-S	17,308,219	both	17,308,592
chr2L	cs14.3	DCO	both	18,573,073	Canton-S	18,732,227
chr2L	w5.14	SCO	both	18,732,812	w^{1118}	18,831,775
chr2L	w11.11	DCO	w^{1118}	19,267,232	both	19,268,010
chr2L	cs5.7	SCO	Canton-S	19,701,889	both	19,702,056
chr2L	cs12.2	DCO	Canton-S	19,966,121	both	19,966,373
chr2L	w15.11	SCO	w^{1118}	20,245,642	both	20,248,327
chr2L	cs14.5	DCO	both	21,057,413	Canton-S	21,058,356
chr2L	cs8.6	DCO	Canton-S	21,724,631	both	21,725,983
chr2R	w5.3	DCO	w^{1118}	7,682,478	both	7,683,941
chr2R	w5.7	DCO	w^{1118}	7,709,714	both	7,710,375
chr2R	cs12.2	SCO	both	8,462,796	Canton-S	8,462,835
chr2R	cs13.17	DCO	Canton-S	9,081,566	both	9,084,198
chr2R	cs13.13	SCO	both	9,467,118	Canton-S	9,469,060
chr2R	cs13.5	SCO	both	9,696,555	Canton-S	9,697,431
chr2R	w5.6	SCO	w^{1118}	10,375,307	both	10,449,157
chr2R	w12.12	SCO	both	10,454,005	w^{1118}	10,517,050
chr2R	cs13.1	SCO	both	10,843,717	Canton-S	10,844,325
chr2R	cs14.5	SCO	Canton-S	10,850,511	both	10,851,167
chr2R	w13.9	SCO	both	10,937,732	w^{1118}	10,937,911
chr2R	w4.9	SCO	w^{1118}	10,982,785	both	10,983,096
chr2R	w15.2	SCO	both	11,035,216	w^{1118}	11,035,311
chr2R	w5.17	SCO	w^{1118}	11,090,327	both	11,090,430
chr2R	cs13.9	SCO	both	11,443,859	Canton-S	11,443,962
chr2R	w5.16	SCO	w^{1118}	11,668,140	both	11,668,812
chr2R	cs5.7	SCO	both	11,737,410	Canton-S	11,739,035
chr2R	w3.15	SCO	both	11,783,126	w^{1118}	11,783,145
chr2R	cs14.10	DCO	Canton-S	12,439,516	both	12,439,961
chr2R	w5.15	SCO	both	12,696,098	w^{1118}	12,696,249
chr2R	w12.1	SCO	w^{1118}	12,901,039	both	12,901,501
chr2R	w4.6	SCO	both	13,023,142	w^{1118}	13,023,262
chr2R	cs6.6	SCO	both	13,045,482	Canton-S	13,045,751
chr2R	w4.16	SCO	w^{1118}	13,313,445	both	13,313,878
chr2R	w5.8	SCO	both	13,479,714	w^{1118}	13,480,591
chr2R	w15.13	SCO	both	13,530,614	w^{1118}	13,530,956
chr2R	w3.8	SCO	w^{1118}	13,609,820	both	13,611,742
chr2R	w3.17	SCO	w^{1118}	13,676,466	both	13,677,087
chr2R	cs1.4	SCO	both	13,840,809	Canton-S	13,840,821
chr2R	w13.5	SCO	w^{1118}	14,004,010	both	14,004,254
chr2R	cs2.4	SCO	Canton-S	14,078,525	both	14,078,709
chr2R	cs12.9	SCO	Canton-S	14,227,542	both	14,402,326
chr2R	w15.3	SCO	w^{1118}	14,460,987	both	14,461,222
chr2R	cs2.1	SCO	Canton-S	14,551,523	both	14,551,975
chr2R	cs12.7	SCO	both	14,619,895	Canton-S	14,619,973
chr2R	cs5.6	SCO	Canton-S	14,650,558	both	14,650,929
chr2R	w11.11	SCO	both	14,858,952	w^{1118}	14,858,982
chr2R	cs13.15	SCO	Canton-S	14,969,237	both	14,969,646
chr2R	w3.11	SCO	both	15,012,571	w^{1118}	15,013,548
chr2R	w4.1	SCO	both	15,359,565	w^{1118}	15,359,812
chr2R	w11.6	SCO	both	15,443,679	w^{1118}	15,447,586
chr2R	cs13.18	SCO	Canton-S	15,772,149	both	15,774,135
chr2R	cs1.1	SCO	Canton-S	16,416,548	both	16,416,752

chr2R	w13.7	SCO	both	16,453,721	w^{1118}	16,454,287
chr2R	w3.21	SCO	both	16,609,368	w^{1118}	16,609,430
chr2R	cs8.2	SCO	both	16,747,424	Canton-S	17,117,240
chr2R	cs6.1	SCO	both	16,775,235	Canton-S	17,117,322
chr2R	w15.11	SCO	both	16,775,238	w^{1118}	16,939,919
chr2R	w12.2	DCO	both	17,384,447	w^{1118}	17,384,973
chr2R	cs14.4	SCO	both	18,054,834	Canton-S	18,055,102
chr2R	cs7.4	SCO	Canton-S	18,243,747	both	18,243,927
chr2R	w13.11	SCO	w^{1118}	18,272,724	both	18,273,082
chr2R	w3.5	SCO	both	18,278,758	w^{1118}	18,278,971
chr2R	w5.5	SCO	w^{1118}	18,699,449	both	18,747,336
chr2R	cs5.4	SCO	both	18,778,807	Canton-S	18,778,894
chr2R	w12.6	SCO	both	18,804,868	w^{1118}	18,825,933
chr2R	w5.12	SCO	w^{1118}	18,887,275	both	18,887,965
chr2R	w13.4	SCO	both	18,892,612	w^{1118}	18,893,584
chr2R	cs7.7	SCO	both	19,084,445	Canton-S	19,084,780
chr2R	cs12.17	SCO	Canton-S	19,110,141	both	19,110,782
chr2R	w13.10	SCO	w^{1118}	19,217,382	both	19,217,461
chr2R	w13.2	SCO	w^{1118}	19,238,121	both	19,238,187
chr2R	w15.6	SCO	w^{1118}	19,238,444	both	19,238,478
chr2R	cs12.14	SCO	both	19,317,131	Canton-S	19,318,772
chr2R	w5.7	DCO	both	19,348,019	w^{1118}	19,351,094
chr2R	cs6.2	SCO	both	19,362,595	Canton-S	19,362,859
chr2R	cs14.13	SCO	Canton-S	19,524,433	both	19,525,433
chr2R	cs2.6	SCO	both	19,573,783	Canton-S	19,577,755
chr2R	w15.5	SCO	w^{1118}	19,589,537	both	19,589,722
chr2R	w12.9	SCO	w^{1118}	19,887,445	both	19,888,503
chr2R	w5.14	SCO	w^{1118}	19,990,332	both	19,993,781
chr2R	w5.4	SCO	both	20,704,780	w^{1118}	20,706,108
chr2R	w3.4	SCO	both	20,709,219	w^{1118}	20,710,393
chr2R	w4.15	SCO	both	20,713,862	w^{1118}	20,714,601
chr2R	cs14.11	SCO	both	20,963,695	Canton-S	20,963,833
chr2R	w15.12	SCO	both	21,025,968	w^{1118}	21,026,185
chr2R	cs2.7	SCO	both	21,165,143	Canton-S	21,167,046
chr2R	cs14.10	DCO	both	21,195,525	Canton-S	21,195,769
chr2R	cs8.5	SCO	Canton-S	21,251,092	both	21,251,147
chr2R	w12.2	DCO	w^{1118}	21,342,738	both	21,342,767
chr2R	cs13.10	SCO	Canton-S	21,534,128	both	21,546,790
chr2R	cs7.1	SCO	both	21,971,056	Canton-S	22,295,778
chr2R	cs12.5	SCO	both	22,089,883	Canton-S	22,295,778
chr2R	cs13.12	SCO	Canton-S	22,356,858	both	22,543,209
chr2R	w12.3	SCO	w^{1118}	22,356,858	both	22,543,178
chr2R	w11.5	SCO	both	22,597,335	w^{1118}	22,597,681
chr2R	cs13.17	DCO	both	22,599,825	Canton-S	22,600,966
chr2R	w13.6	SCO	both	22,660,461	w^{1118}	22,660,891
chr2R	w11.1	SCO	both	22,690,815	w^{1118}	22,691,310
chr2R	cs2.8	SCO	Canton-S	23,152,681	both	23,153,202
chr2R	w11.4	SCO	both	23,235,667	w^{1118}	23,235,762
chr2R	w5.1	SCO	w^{1118}	23,564,105	both	23,566,121
chr2R	cs12.15	SCO	Canton-S	23,611,564	both	23,611,748
chr2R	cs13.14	SCO	both	23,634,275	Canton-S	23,634,562
chr2R	w12.7	SCO	both	23,663,142	w^{1118}	23,663,467
chr2R	w13.1	SCO	w^{1118}	23,667,315	both	23,668,487

chr2R	w4.11	SCO	both	23,754,523	w^{1118}	23,754,853
chr2R	w5.3	DCO	both	23,813,940	w^{1118}	24,314,292
chr2R	cs7.8	SCO	both	23,824,108	Canton-S	23,827,040
chr2R	cs13.3	SCO	both	23,978,611	Canton-S	24,161,424
chr3L	cs13.10	SCO	both	423,397	Canton-S	423,450
chr3L	cs12.4	DCO	both	438,695	Canton-S	439,457
chr3L	w13.1	DCO	both	455,837	w^{1118}	459,300
chr3L	cs13.15	SCO	Canton-S	1,066,558	both	1,067,046
chr3L	cs12.10	SCO	Canton-S	1,087,159	both	1,087,429
chr3L	w11.10	DCO	both	1,128,058	w^{1118}	1,128,835
chr3L	cs7.2	DCO	both	1,142,358	Canton-S	1,143,990
chr3L	cs13.4	SCO	Canton-S	1,355,369	both	1,355,489
chr3L	w12.11	SCO	both	1,496,091	w^{1118}	1,497,420
chr3L	cs2.7	SCO	Canton-S	1,634,046	both	1,634,248
chr3L	cs12.14	DCO	both	1,689,000	Canton-S	1,689,549
chr3L	cs8.7	DCO	both	1,876,134	Canton-S	1,876,267
chr3L	w5.16	SCO	w^{1118}	1,947,908	both	1,949,195
chr3L	w4.11	SCO	w^{1118}	1,989,821	both	1,990,848
chr3L	cs14.1	SCO	both	2,283,293	Canton-S	2,283,725
chr3L	w4.3	DCO	w^{1118}	2,335,040	both	2,335,059
chr3L	w13.7	SCO	both	2,344,773	w^{1118}	2,344,842
chr3L	w13.6	SCO	w^{1118}	2,464,653	both	2,464,807
chr3L	w4.7	SCO	both	2,486,804	w^{1118}	2,487,161
chr3L	w12.13	DCO	both	2,582,609	w^{1118}	2,582,815
chr3L	w4.13	SCO	w^{1118}	3,183,452	both	3,183,836
chr3L	cs1.3	DCO	Canton-S	3,272,605	both	3,284,655
chr3L	cs8.8	DCO	Canton-S	3,272,605	both	3,281,471
chr3L	w3.8	SCO	w^{1118}	4,089,884	both	4,090,193
chr3L	w3.25	SCO	both	4,090,571	w^{1118}	4,090,900
chr3L	w11.12	SCO	both	4,101,283	w^{1118}	4,101,505
chr3L	cs12.9	SCO	both	4,123,267	Canton-S	4,123,984
chr3L	w5.8	SCO	both	4,169,248	w^{1118}	4,169,471
chr3L	cs7.1	SCO	both	4,273,153	Canton-S	4,273,203
chr3L	cs6.1	SCO	both	4,387,251	Canton-S	4,387,430
chr3L	w15.10	SCO	both	4,390,883	w^{1118}	4,391,731
chr3L	cs12.13	SCO	Canton-S	4,413,897	both	4,413,994
chr3L	cs13.9	SCO	Canton-S	4,416,390	both	4,418,296
chr3L	w3.12	SCO	w^{1118}	4,572,340	both	4,573,727
chr3L	cs7.3	SCO	Canton-S	4,693,470	both	4,929,882
chr3L	w5.11	SCO	both	4,953,060	w^{1118}	4,955,985
chr3L	cs9.1	SCO	Canton-S	5,271,329	both	5,271,631
chr3L	cs1.7	SCO	both	5,521,314	Canton-S	5,521,829
chr3L	cs1.2	SCO	both	5,730,451	Canton-S	5,732,435
chr3L	w5.13	SCO	w^{1118}	6,449,599	both	6,450,567
chr3L	w13.5	DCO	both	6,524,877	w^{1118}	6,525,102
chr3L	cs12.12	SCO	both	6,833,285	Canton-S	6,834,431
chr3L	cs8.6	SCO	both	7,041,672	Canton-S	7,042,495
chr3L	cs8.5	SCO	Canton-S	7,112,051	both	7,113,416
chr3L	w15.1	SCO	both	7,187,626	w^{1118}	7,187,833
chr3L	cs12.3	DCO	both	7,244,355	Canton-S	7,244,397
chr3L	w4.5	SCO	w^{1118}	7,313,147	both	7,314,715
chr3L	cs2.2	DCO	both	7,543,981	Canton-S	7,630,516
chr3L	cs14.5	SCO	both	7,791,858	Canton-S	7,791,950

chr3L	w4.15	SCO	w^{1118}	8,013,724	both	8,014,626
chr3L	w3.6	SCO	w^{1118}	8,039,841	both	8,245,622
chr3L	w3.17	SCO	both	8,060,764	w^{1118}	8,245,649
chr3L	cs12.11	SCO	both	8,318,766	Canton-S	8,319,182
chr3L	cs13.3	SCO	both	8,376,029	Canton-S	8,376,427
chr3L	cs5.4	SCO	Canton-S	8,447,572	both	8,447,697
chr3L	cs6.7	SCO	Canton-S	8,449,014	both	8,449,443
chr3L	cs7.8	SCO	Canton-S	8,561,751	both	8,562,770
chr3L	w13.2	DCO	both	8,628,955	w^{1118}	8,629,074
chr3L	cs7.7	SCO	Canton-S	8,658,541	both	8,658,607
chr3L	cs1.5	SCO	both	8,899,683	Canton-S	8,900,643
chr3L	w5.4	SCO	both	9,268,112	w^{1118}	9,268,362
chr3L	w5.2	SCO	both	9,288,299	w^{1118}	9,288,597
chr3L	w11.1	SCO	both	9,478,291	w^{1118}	9,478,392
chr3L	w4.12	SCO	both	9,528,457	w^{1118}	9,528,908
chr3L	w15.11	SCO	w^{1118}	9,544,026	both	9,544,193
chr3L	cs14.10	SCO	both	9,670,648	Canton-S	9,670,968
chr3L	cs7.4	SCO	both	9,848,327	Canton-S	9,849,891
chr3L	w11.2	SCO	both	10,417,522	w^{1118}	10,417,851
chr3L	w15.9	SCO	w^{1118}	10,498,397	both	10,503,460
chr3L	cs14.7	SCO	both	10,595,428	Canton-S	10,595,688
chr3L	cs8.8	DCO	both	10,843,455	Canton-S	10,843,969
chr3L	cs1.3	DCO	both	10,879,474	Canton-S	10,921,660
chr3L	cs14.2	SCO	both	11,135,967	Canton-S	11,136,386
chr3L	w15.2	SCO	both	11,174,736	w^{1118}	11,175,067
chr3L	cs13.16	SCO	Canton-S	11,261,093	both	11,261,294
chr3L	w13.10	SCO	w^{1118}	11,355,413	both	11,355,902
chr3L	cs2.5	SCO	Canton-S	11,545,654	both	11,546,210
chr3L	cs14.11	SCO	both	11,639,279	Canton-S	11,640,756
chr3L	cs12.4	DCO	Canton-S	11,759,765	both	11,760,423
chr3L	w13.4	SCO	both	11,913,828	w^{1118}	11,914,101
chr3L	cs2.2	DCO	Canton-S	12,025,593	both	12,047,362
chr3L	cs6.8	SCO	Canton-S	12,025,593	both	12,178,406
chr3L	cs6.4	SCO	Canton-S	12,025,634	both	12,177,420
chr3L	w3.16	DCO	both	12,047,362	w^{1118}	12,177,420
chr3L	cs12.8	SCO	Canton-S	12,259,571	both	12,260,147
chr3L	w4.3	DCO	both	12,796,346	w^{1118}	12,796,408
chr3L	w13.5	DCO	w^{1118}	12,830,149	both	12,830,465
chr3L	w12.2	SCO	both	12,900,942	w^{1118}	12,901,116
chr3L	w12.13	DCO	w^{1118}	13,008,017	both	13,010,830
chr3L	w3.26	SCO	both	13,266,887	w^{1118}	13,267,062
chr3L	w11.10	DCO	w^{1118}	13,641,405	both	13,642,438
chr3L	w4.2	SCO	w^{1118}	14,239,524	both	14,239,824
chr3L	cs8.7	DCO	Canton-S	14,745,396	both	14,745,912
chr3L	w3.9	SCO	w^{1118}	15,131,162	both	15,134,335
chr3L	w4.9	SCO	both	16,151,423	w^{1118}	16,151,531
chr3L	cs5.1	SCO	both	16,166,552	Canton-S	16,166,740
chr3L	w12.9	SCO	w^{1118}	16,452,664	both	16,453,034
chr3L	w3.11	SCO	w^{1118}	16,463,550	both	16,464,055
chr3L	w12.6	SCO	both	16,516,501	w^{1118}	16,517,031
chr3L	w3.18	SCO	both	16,547,854	w^{1118}	16,547,967
chr3L	cs2.4	SCO	Canton-S	16,549,471	both	16,557,598
chr3L	w13.2	DCO	w^{1118}	16,624,503	both	16,631,771

chr3L	cs12.14	DCO	Canton-S	16,832,576	Canton-S	16,832,971
chr3L	w3.13	SCO	both	16,969,066	w^{1118}	16,970,047
chr3L	w3.16	DCO	w^{1118}	17,119,829	both	17,121,623
chr3L	cs7.2	DCO	Canton-S	17,536,276	both	17,536,873
chr3L	w11.11	SCO	both	17,561,685	w^{1118}	17,565,523
chr3L	cs14.13	SCO	both	18,267,251	Canton-S	18,268,705
chr3L	cs13.5	SCO	Canton-S	18,353,165	both	18,353,367
chr3L	cs13.17	SCO	Canton-S	18,526,247	both	18,526,314
chr3L	cs13.2	SCO	Canton-S	19,185,874	both	19,186,402
chr3L	cs12.3	DCO	Canton-S	19,579,736	both	19,654,059
chr3L	w13.1	DCO	w^{1118}	20,056,244	both	20,057,583
chr3L	cs14.3	SCO	both	22,881,747	Canton-S	22,921,874
chr3R	cs14.5	TCO	Canton-S	6,466,846	both	6,474,296
chr3R	w12.2	SCO	w^{1118}	8,225,505	both	8,225,752
chr3R	cs7.8	DCO	both	10,922,991	Canton-S	10,948,412
chr3R	w15.3	DCO	w^{1118}	11,170,623	both	11,170,745
chr3R	cs14.9	SCO	Canton-S	12,285,585	both	12,591,515
chr3R	cs14.2	SCO	Canton-S	12,359,258	both	12,591,515
chr3R	cs7.4	SCO	Canton-S	13,009,923	both	13,023,020
chr3R	w15.5	SCO	both	13,430,619	w^{1118}	13,557,154
chr3R	cs1.6	DCO	Canton-S	13,915,057	both	13,915,359
chr3R	w3.18	SCO	w^{1118}	13,938,444	both	13,938,913
chr3R	w15.8	SCO	both	14,320,305	w^{1118}	14,322,111
chr3R	cs1.7	DCO	Canton-S	14,320,337	both	14,322,111
chr3R	cs13.5	SCO	both	14,337,545	Canton-S	14,337,894
chr3R	cs13.14	SCO	Canton-S	14,427,031	both	14,427,452
chr3R	w15.4	DCO	both	14,427,890	w^{1118}	14,428,085
chr3R	cs14.13	SCO	Canton-S	14,631,515	both	14,631,633
chr3R	cs1.1	SCO	both	14,707,596	Canton-S	14,707,970
chr3R	w3.15	SCO	both	14,910,185	w^{1118}	14,910,498
chr3R	w11.3	SCO	w^{1118}	15,415,096	both	15,415,183
chr3R	cs14.3	SCO	Canton-S	15,452,160	both	15,453,571
chr3R	cs14.11	SCO	Canton-S	15,519,717	both	15,522,659
chr3R	cs13.10	SCO	Canton-S	16,255,612	both	16,350,494
chr3R	cs5.7	SCO	Canton-S	16,715,920	both	16,715,983
chr3R	cs13.9	SCO	both	17,031,197	Canton-S	17,031,813
chr3R	cs2.7	SCO	both	17,068,663	Canton-S	17,068,947
chr3R	w13.2	SCO	both	17,134,799	w^{1118}	17,137,033
chr3R	w5.15	SCO	both	17,585,414	w^{1118}	17,585,507
chr3R	cs12.16	DCO	both	17,660,114	Canton-S	17,739,708
chr3R	cs13.6	SCO	both	17,660,114	Canton-S	17,739,708
chr3R	w13.10	SCO	both	17,660,114	w^{1118}	17,739,708
chr3R	w15.2	SCO	w^{1118}	17,922,486	both	17,952,899
chr3R	w15.9	SCO	both	17,953,115	w^{1118}	17,983,264
chr3R	w3.5	SCO	both	18,055,084	w^{1118}	18,055,961
chr3R	w5.8	SCO	w^{1118}	18,077,706	both	18,122,288
chr3R	w11.2	DCO	w^{1118}	18,154,471	both	18,154,512
chr3R	w3.24	SCO	both	18,206,484	w^{1118}	18,206,614
chr3R	w3.13	SCO	w^{1118}	18,211,655	both	18,212,792
chr3R	w4.8	SCO	w^{1118}	18,408,970	both	18,412,933
chr3R	cs7.7	SCO	both	18,445,604	Canton-S	18,447,407
chr3R	cs12.8	DCO	both	18,486,435	Canton-S	18,487,137
chr3R	w13.7	SCO	w^{1118}	18,530,556	both	18,530,717

chr3R	cs1.2	SCO	Canton-S	18,560,152	both	18,563,276
chr3R	w5.12	SCO	both	18,704,427	w^{1118}	18,705,183
chr3R	cs14.5	TCO	both	18,989,341	Canton-S	18,990,501
chr3R	w5.14	SCO	w^{1118}	19,083,606	both	19,083,617
chr3R	cs7.3	SCO	both	19,262,960	Canton-S	19,264,006
chr3R	w15.13	SCO	w^{1118}	19,619,933	both	19,620,156
chr3R	w13.9	DCO	w^{1118}	19,687,449	both	19,687,549
chr3R	w15.12	SCO	both	19,956,611	w^{1118}	19,957,029
chr3R	w13.13	SCO	both	20,244,350	w^{1118}	20,244,561
chr3R	cs13.17	SCO	both	20,315,976	Canton-S	20,316,527
chr3R	cs12.13	SCO	both	20,333,086	Canton-S	20,333,710
chr3R	cs13.11	SCO	both	21,655,944	Canton-S	21,656,021
chr3R	w12.1	SCO	w^{1118}	21,656,419	both	21,656,872
chr3R	cs2.5	SCO	both	21,859,772	Canton-S	21,859,909
chr3R	cs6.3	SCO	Canton-S	21,952,516	both	21,953,183
chr3R	w12.12	SCO	w^{1118}	21,980,754	both	21,981,468
chr3R	w4.6	SCO	both	22,031,094	w^{1118}	22,031,217
chr3R	cs12.4	SCO	both	22,074,118	Canton-S	22,074,969
chr3R	w3.4	SCO	w^{1118}	22,303,990	both	22,306,458
chr3R	cs2.2	SCO	both	22,411,846	Canton-S	22,411,946
chr3R	w15.3	DCO	both	22,471,827	w^{1118}	22,472,312
chr3R	cs8.1	SCO	both	22,484,505	Canton-S	22,485,337
chr3R	w4.7	SCO	w^{1118}	22,527,516	both	22,529,157
chr3R	cs12.17	SCO	both	22,617,386	Canton-S	22,617,557
chr3R	w4.15	DCO	both	22,635,522	w^{1118}	22,710,817
chr3R	cs2.3	SCO	both	22,636,058	Canton-S	22,636,169
chr3R	cs12.1	SCO	both	22,637,407	Canton-S	22,710,798
chr3R	cs8.8	SCO	Canton-S	22,740,926	both	22,741,106
chr3R	w4.11	SCO	both	22,813,875	w^{1118}	22,814,230
chr3R	cs12.15	SCO	both	22,873,525	Canton-S	22,874,011
chr3R	w5.16	SCO	both	22,878,004	w^{1118}	22,878,648
chr3R	cs8.6	SCO	Canton-S	22,917,883	both	22,919,287
chr3R	cs1.7	DCO	both	23,406,581	Canton-S	23,409,147
chr3R	cs13.13	SCO	both	23,444,465	Canton-S	23,444,484
chr3R	cs13.3	SCO	Canton-S	23,866,960	both	23,867,985
chr3R	cs12.12	SCO	Canton-S	23,877,828	both	23,877,907
chr3R	w5.4	SCO	w^{1118}	23,911,117	both	23,911,320
chr3R	cs14.7	SCO	Canton-S	24,390,134	both	24,499,438
chr3R	w4.17	SCO	w^{1118}	24,390,134	both	24,478,803
chr3R	w11.5	SCO	both	24,734,613	w^{1118}	24,735,097
chr3R	cs14.10	SCO	Canton-S	24,959,795	both	24,960,553
chr3R	w11.2	DCO	both	25,359,203	w^{1118}	25,359,914
chr3R	cs12.11	SCO	Canton-S	25,488,789	both	25,489,017
chr3R	cs6.7	SCO	both	25,621,000	Canton-S	25,621,176
chr3R	w13.6	SCO	both	25,697,258	w^{1118}	25,699,653
chr3R	cs14.8	SCO	Canton-S	25,962,221	both	25,963,030
chr3R	w15.11	SCO	both	26,438,132	w^{1118}	26,438,877
chr3R	cs14.12	SCO	Canton-S	27,134,978	both	27,135,381
chr3R	cs14.5	TCO	Canton-S	27,338,451	both	27,339,070
chr3R	w15.10	SCO	w^{1118}	27,429,512	both	27,429,860
chr3R	cs7.8	DCO	Canton-S	27,603,887	both	27,604,729
chr3R	w4.15	DCO	w^{1118}	28,100,433	both	28,102,063
chr3R	cs13.15	SCO	both	28,106,068	Canton-S	28,107,852

chr3R	cs1.5	SCO	Canton-S	28,356,421	both	28,358,277
chr3R	w5.9	SCO	w^{1118}	28,470,969	both	28,471,718
chr3R	cs8.5	SCO	both	28,701,193	Canton-S	28,706,896
chr3R	w13.8	SCO	w^{1118}	28,834,450	both	28,834,579
chr3R	w11.11	SCO	w^{1118}	29,154,345	both	29,156,251
chr3R	w3.26	SCO	w^{1118}	29,226,053	both	29,226,402
chr3R	cs8.2	SCO	both	29,267,970	Canton-S	29,268,436
chr3R	cs12.8	DCO	Canton-S	29,393,790	both	29,393,966
chr3R	w13.9	DCO	both	29,643,787	w^{1118}	29,644,877
chr3R	cs1.6	DCO	both	29,901,731	Canton-S	29,902,513
chr3R	w4.1	SCO	w^{1118}	30,089,042	both	30,089,122
chr3R	w11.13	SCO	both	30,355,503	w^{1118}	30,357,322
chr3R	w15.4	DCO	w^{1118}	30,439,429	both	30,440,821
chr3R	cs12.16	DCO	Canton-S	30,822,917	both	30,823,149
chr3R	cs12.3	SCO	both	31,326,496	Canton-S	31,326,989

Table 2.S2. Detailed information about all 294 NCO events recovered in this study.

Stock	Chr	Orig	5' SNP	NCO	5'NCO	3'NCO	SNP	3' SNP	Min Tract Len	Max Tract Len
cs1.2	3R	both	25,797,087	cs	25,797,702	25,798,877	18	25,799,116	1,175	2,029
cs1.2	3R	both	26,229,609	cs	26,230,138	26,230,187	3	26,230,505	49	896
cs1.3	3L	cs	19,063,124	both	19,063,363	19,063,488	3	19,063,768	125	644
cs1.4	2L	cs	2,699,765	both	2,700,086	2,701,390	17	2,701,940	1,304	2,175
cs1.4	X	w	20,055,079	cs	20,056,910	20,057,317	5	20,058,234	407	3,155
cs1.5	3R	cs	6,811,320	both	6,811,568	6,812,266	8	6,813,683	698	2,363
cs1.7	2L	cs	2,438,094	both	2,440,378	2,440,638	13	2,440,980	260	2,886
cs1.7	2L	cs	7,522,845	both	7,523,012	7,523,137	6	7,523,201	125	356
cs1.7	2R	cs	20,480,585	both	20,480,791	20,481,018	4	20,482,763	227	2,178
cs1.7	3L	both	3,792,394	cs	3,792,602	3,792,602	1	3,792,738	0	344
cs1.8	3L	cs	16,547,180	both	16,547,185	16,547,382	5	16,547,719	197	539
cs1.8	3R	cs	20,573,427	both	20,574,254	20,574,493	2	20,575,410	239	1,983
cs1.8	3R	cs	23,703,931	both	23,704,355	23,704,355	1	23,704,840	0	909
cs12.10	2R	both	11,059,163	cs	11,059,289	11,059,770	7	11,059,821	481	658
cs12.10	3R	both	18,334,037	cs	18,334,415	18,335,044	3	18,335,256	629	1,219
cs12.10	3R	both	30,724,432	cs	30,724,573	30,725,251	16	30,725,361	678	929
cs12.11	2L	both	10,270,245	cs	10,270,799	10,270,931	4	10,271,039	132	794
cs12.11	3L	cs	20,793,503	both	20,793,608	20,793,730	5	20,794,540	122	1,037
cs12.12	2L	both	11,821,923	cs	11,822,463	11,822,463	1	11,822,907	0	984
cs12.13	2L	cs	10,583,358	both	10,584,641	10,584,641	1	10,585,653	0	2,295
cs12.15	3L	both	8,923,967	cs	8,924,062	8,926,679	15	8,927,522	2,617	3,555
cs12.15	3R	both	5,331,450	cs	5,332,327	5,332,701	3	5,333,950	374	2,500
cs12.16	2R	cs	12,698,057	both	12,698,237	12,698,237	1	12,698,274	0	217
cs12.16	2R	cs	12,698,278	both	12,698,476	12,698,476	1	12,698,880	0	602
cs12.17	2L	cs	15,758,136	both	15,758,371	15,758,491	5	15,759,094	120	958
cs12.17	3L	both	6,802,801	cs	6,803,210	6,804,260	20	6,804,422	1,050	1,621
cs12.18	3L	cs	4,099,573	both	4,099,983	4,100,010	3	4,100,258	27	685
cs12.18	3L	cs	12,782,223	both	12,782,527	12,782,722	5	12,783,012	195	789
cs12.18	3R	cs	9,077,837	both	9,077,898	9,079,821	19	9,080,080	1,923	2,243
cs12.2	3R	both	12,850,547	cs	12,851,117	12,851,375	2	12,851,774	258	1,227
cs12.3	X	w	4,887,025	cs	4,889,571	4,889,622	4	4,889,833	51	2,808
cs12.5	2R	cs	24,471,658	both	24,471,762	24,472,024	4	24,472,130	262	472
cs12.6	X	w	16,251,375	cs	16,251,556	16,251,780	5	16,252,417	224	1,042
cs12.8	2L	both	17,581,980	cs	17,582,420	17,582,555	2	17,582,753	135	773
cs12.8	3L	cs	5,354,120	both	5,355,010	5,355,010	1	5,355,966	0	1,846
cs12.8	3R	both	13,925,399	cs	13,925,856	13,925,958	2	13,926,063	102	664
cs12.9	2R	cs	6,529,172	both	6,529,429	6,529,429	1	6,529,675	0	503
cs12.9	3L	cs	19,379,625	both	19,380,222	19,380,222	1	19,381,207	0	1,582
cs13.1	2L	both	9,224,679	cs	9,224,793	9,224,880	5	9,225,050	87	371
cs13.1	2L	both	10,126,562	cs	10,127,024	10,127,024	1	10,127,847	0	1,285
cs13.1	2R	cs	14,011,507	both	14,011,636	14,011,650	2	14,011,760	14	253
cs13.10	2L	cs	11,706,965	both	11,707,045	11,707,080	2	11,707,145	35	180
cs13.10	3R	both	28,040,865	cs	28,041,281	28,041,281	1	28,041,558	0	693
cs13.11	2L	both	10,653,770	cs	10,655,037	10,655,630	6	10,657,515	593	3,745
cs13.11	3L	both	20,425,449	cs	20,425,526	20,425,526	1	20,425,737	0	288
cs13.11	3R	cs	30,071,213	both	30,071,394	30,071,571	7	30,071,810	177	597
cs13.12	X	w	14,516,292	cs	14,516,533	14,517,549	13	14,517,837	1,016	1,545

cs13.13	3L	both	969,147	cs	969,293	969,472	3	970,119	179	972
cs13.13	3L	both	15,548,633	cs	15,561,281	15,561,281	1	15,563,193	0	14,560
cs13.13	3R	both	17,326,933	cs	17,327,356	17,327,541	7	17,327,936	185	1,003
cs13.14	3R	cs	8,621,616	both	8,622,114	8,622,114	1	8,623,082	0	1,466
cs13.16	2R	cs	9,630,382	both	9,631,478	9,631,478	1	9,632,523	0	2,141
cs13.16	2R	cs	23,610,020	both	23,610,443	23,610,782	6	23,611,162	339	1,142
cs13.16	3R	both	28,602,735	cs	28,602,875	28,603,104	3	28,603,287	229	552
cs13.18	2R	cs	10,357,250	both	10,357,447	10,357,476	3	10,357,885	29	635
cs13.2	2R	both	5,210,313	cs	5,211,005	5,211,005	1	5,211,111	0	798
cs13.2	3L	both	22,272,997	cs	22,273,942	22,273,942	1	22,274,164	0	1,167
cs13.3	3L	both	293,913	cs	295,659	295,695	2	295,887	36	1,974
cs13.3	3R	cs	22,356,879	both	22,357,756	22,357,918	3	22,358,012	162	1,133
cs13.4	2R	both	12,612,483	cs	12,613,088	12,613,088	1	12,613,345	0	862
cs13.5	3L	cs	5,525,947	both	5,526,883	5,527,279	7	5,527,823	396	1,876
cs13.5	X	cs	5,650,268	w	5,650,323	5,651,353	12	5,652,112	1,030	1,844
cs13.6	2R	both	7,398,693	cs	7,399,211	7,400,262	5	7,400,532	1,051	1,839
cs13.6	3L	both	1,238,394	cs	1,239,996	1,239,996	1	1,241,184	0	2,790
cs13.8	3L	cs	16,429,266	both	16,430,717	16,431,316	6	16,432,022	599	2,756
cs13.8	3R	cs	14,929,297	both	14,929,818	14,929,818	1	14,930,385	0	1,088
cs13.8	X	w	15,610,345	cs	15,610,898	15,611,416	3	15,611,597	518	1,252
cs13.9	3L	both	19,073,541	cs	19,075,242	19,075,422	5	19,076,768	180	3,227
cs14.1	2L	both	20,428,935	cs	20,429,217	20,429,217	1	20,429,349	0	414
cs14.11	2R	both	11,860,899	cs	11,861,117	11,861,903	2	11,862,045	786	1,146
cs14.12	2L	cs	9,486,011	both	9,486,165	9,486,825	15	9,487,350	660	1,339
cs14.12	3R	cs	6,667,690	both	6,667,708	6,668,552	7	6,669,414	844	1,724
cs14.13	2L	cs	839,425	both	839,664	839,972	4	840,497	308	1,072
cs14.13	2R	cs	10,966,447	both	10,967,228	10,967,363	5	10,968,035	135	1,588
cs14.3	2R	cs	18,143,882	both	18,143,939	18,144,002	2	18,144,680	63	798
cs14.4	2L	both	14,084,565	cs	14,086,188	14,086,188	1	14,086,743	0	2,178
cs14.4	2R	both	13,622,406	cs	13,622,451	13,623,415	11	13,623,556	964	1,150
cs14.4	2R	both	18,143,882	both	18,143,939	18,143,939	1	18,144,680	0	798
cs14.4	X	cs	4,310,527	w	4,310,750	4,310,757	2	4,310,854	7	327
cs14.5	2L	cs	3,430,497	both	3,430,560	3,431,013	10	3,431,064	453	567
cs14.6	3L	cs	3,538,035	both	3,538,503	3,538,503	1	3,539,504	0	1,469
cs14.6	3L	cs	18,523,747	both	18,524,612	18,524,612	1	18,524,639	0	892
cs14.6	X	cs	14,164,480	w	14,164,669	14,165,068	5	14,165,963	399	1,483
cs14.7	2R	cs	7,189,811	both	7,190,060	7,190,207	3	7,190,395	147	584
cs14.7	2R	cs	15,288,609	both	15,288,696	15,288,709	3	15,289,189	13	580
cs14.7	2R	cs	17,864,516	both	17,864,621	17,864,707	2	17,865,150	86	634
cs14.7	3L	both	3,001,547	cs	3,001,863	3,001,931	3	3,002,307	68	760
cs14.7	3L	both	7,255,544	cs	7,255,665	7,255,963	7	7,256,319	298	775
cs14.7	3L	cs	18,341,119	both	18,341,349	18,341,349	1	18,341,382	0	263
cs14.7	X	cs	20,145,715	cs	20,145,729	20,145,729	1	20,145,765	0	50
cs14.8	2L	both	5,022,553	cs	5,022,584	5,022,584	1	5,022,864	0	311
cs14.8	2R	both	11,620,787	cs	11,621,086	11,621,355	7	11,621,376	269	589
cs14.9	2L	cs	16,184,698	both	16,184,800	16,185,259	10	16,185,700	459	1,002
cs2.1	2L	cs	17,364,999	both	17,365,149	17,365,205	2	17,365,889	56	890
cs2.1	3L	cs	5,575,800	both	5,579,178	5,579,550	6	5,579,801	372	4,001
cs2.1	3R	cs	23,355,950	both	23,356,022	23,356,314	14	23,356,378	292	428
cs2.2	2L	cs	881,627	both	881,702	882,129	9	882,611	427	984
cs2.2	3L	both	1,462,457	cs	1,462,825	1,462,830	2	1,463,469	5	1,012
cs2.3	2L	both	19,998,787	cs	19,999,807	20,000,444	11	20,000,997	637	2,210

cs2.4	3L	both	22,430,490	cs	22,431,594	22,431,594	1	22,432,070	0	1,580
cs2.4	3R	both	5,498,924	cs	5,499,307	5,499,458	3	5,507,023	151	8,099
cs2.5	2R	both	24,865,265	cs	24,865,481	24,865,481	1	24,873,586	0	8,321
cs2.5	3L	both	20,709,299	cs	20,709,624	20,711,074	8	20,711,474	1,450	2,175
cs2.5	X	w	10,187,078	cs	10,187,375	10,187,639	3	10,187,754	264	676
cs2.6	2L	cs	1,373,024	both	1,374,199	1,374,199	1	1,374,863	0	1,839
cs2.8	2L	cs	4,054,814	both	4,054,986	4,055,191	2	4,055,661	205	847
cs2.8	3R	both	6,173,812	cs	6,174,195	6,174,488	4	6,174,961	293	1,149
cs2.8	3R	both	21,795,289	cs	21,796,205	21,796,205	1	21,796,897	0	1,608
cs5.1	X	cs	5,125,196	w	5,126,286	5,126,439	4	5,126,587	153	1,391
cs5.2	2L	cs	10,143,717	both	10,144,942	10,144,963	2	10,145,962	21	2,245
cs5.3	2L	cs	10,371,819	both	10,372,302	10,372,548	5	10,372,877	246	1,058
cs5.3	2L	cs	17,227,951	both	17,228,153	17,228,519	3	17,229,702	366	1,751
cs5.3	2R	cs	10,990,374	both	10,991,196	10,991,260	2	10,991,518	64	1,144
cs5.4	X	cs	6,715,621	w	6,715,858	6,716,055	4	6,716,178	197	557
cs5.5	2R	cs	8,579,734	both	8,579,986	8,580,076	3	8,580,833	90	1,099
cs5.8	2L	both	1,719,369	cs	1,719,552	1,719,600	3	1,719,701	48	332
cs5.8	3R	both	9,592,539	cs	9,592,935	9,593,018	3	9,593,703	83	1,164
cs6.1	X	cs	19,106,881	w	19,107,331	19,107,559	2	19,108,873	228	1,992
cs6.4	2L	both	8,990,233	cs	8,990,244	8,990,244	1	8,991,034	0	801
cs6.4	X	cs	6,838,375	w	6,839,453	6,839,453	1	6,840,094	0	1,719
cs6.5	3L	both	16,826,331	cs	16,826,402	16,827,717	14	16,827,924	1,315	1,593
cs6.5	3R	both	9,276,279	cs	9,276,599	9,277,526	5	9,278,400	927	2,121
cs6.6	X	cs	14,674,612	w	14,675,085	14,675,471	2	14,676,516	386	1,904
cs6.6	X	cs	19,361,996	w	19,362,525	19,363,980	9	19,364,463	1,455	2,467
cs6.8	2L	both	1,023,563	cs	1,024,690	1,024,690	1	1,025,592	0	2,029
cs6.8	2L	both	11,208,602	cs	11,208,745	11,210,550	21	11,211,407	1,805	2,805
cs7.1	3R	cs	9,257,136	both	9,258,167	9,258,711	3	9,262,600	544	5,464
cs7.2	2R	both	19,142,890	cs	19,142,914	19,142,952	2	19,143,418	38	528
cs7.2	3L	both	18,994,046	cs	18,994,326	18,994,851	2	18,996,234	525	2,188
cs7.2	3R	both	19,607,313	cs	19,607,717	19,607,744	2	19,608,202	27	889
cs7.2	3R	both	20,817,149	cs	20,817,171	20,817,189	2	20,817,240	18	91
cs7.3	2R	cs	11,695,205	both	11,695,739	11,695,848	5	11,696,000	109	795
cs7.4	2R	cs	11,459,111	both	11,459,496	11,459,496	1	11,461,937	0	2,826
cs7.4	X	w	16,102,283	cs	16,102,614	16,102,959	3	16,103,521	345	1,238
cs7.5	3R	cs	27,935,021	both	27,936,304	27,936,346	2	27,936,570	42	1,549
cs7.6	2R	cs	18,057,857	both	18,058,040	18,058,040	1	18,058,080	0	223
cs7.6	3R	both	7,570,334	cs	7,570,491	7,570,491	1	7,570,858	0	524
cs7.7	3L	both	16,070,455	cs	16,070,603	16,071,100	9	16,071,427	497	972
cs7.7	X	w	6,730,605	cs	6,730,664	6,730,664	1	6,731,407	0	802
cs8.1	2R	cs	15,536,934	both	15,537,679	15,538,807	17	15,539,111	1,128	2,177
cs8.4	3L	cs	7,007,956	both	7,008,374	7,009,028	6	7,009,287	654	1,331
cs8.5	X	w	14,319,556	cs	14,320,505	14,320,726	3	14,321,223	221	1,667
cs8.6	2L	cs	2,675,337	both	2,675,843	2,675,987	3	2,676,388	144	1,051
cs8.6	2R	both	19,285,945	cs	19,287,112	19,287,112	1	19,288,656	0	2,711
cs8.6	3R	cs	20,221,867	both	20,222,747	20,222,747	1	20,223,967	0	2,100
cs8.6	X	cs	6,475,005	w	6,475,735	6,475,735	1	6,479,615	0	4,610
cs8.8	3R	both	30,297,496	cs	30,297,671	30,298,113	3	30,298,691	442	1,195
w11.1	2L	both	15,292,965	w	15,293,161	15,293,161	1	15,295,425	0	2,460
w11.1	2R	both	7,513,996	w	7,514,177	7,514,534	4	7,515,996	357	2,000
w11.1	3R	w	17,007,874	both	17,008,296	17,008,455	5	17,008,544	159	670
w11.11	2L	both	22,136,812	w	22,137,741	22,137,741	1	22,137,883	0	1,071
w11.11	3L	both	15,903,139	w	15,903,492	15,903,906	16	15,904,153	414	1,014

w11.12	2R	w	9,780,801	both	9,781,245	9,781,559	4	9,782,010	314	1,209
w11.12	3L	both	1,443,623	w	1,444,015	1,444,130	6	1,444,499	115	876
w11.12	3L	w	8,711,353	both	8,711,806	8,711,876	2	8,712,007	70	654
w11.12	3L	both	22,419,779	w	22,420,117	22,420,683	9	22,421,184	566	1,405
w11.13	2R	both	14,732,633	w	14,732,778	14,732,778	1	14,733,772	0	1,139
w11.3	3L	w	16,200,915	both	16,201,345	16,201,492	4	16,202,103	147	1,188
w11.4	2R	both	7,434,398	w	7,439,495	7,439,771	3	7,439,968	276	5,570
w11.4	3R	w	19,119,523	both	19,120,657	19,121,013	6	19,122,263	356	2,740
w11.4	3R	w	30,310,949	both	30,311,038	30,311,708	9	30,312,056	670	1,107
w11.5	2L	both	11,233,317	w	11,233,639	11,233,643	2	11,234,223	4	906
w11.5	3L	both	6,239,818	w	6,239,930	6,240,293	2	6,240,496	363	678
w11.5	X	w	5,451,026	cs	5,451,197	5,451,346	5	5,452,036	149	1,010
w11.6	2L	w	10,430,586	both	10,431,000	10,431,000	1	10,431,462	0	876
w11.7	2L	both	6,755,412	w	6,755,529	6,755,529	1	6,756,911	0	1,499
w11.8	2R	both	7,205,276	w	7,205,831	7,205,831	1	7,207,040	0	1,764
w11.8	X	cs	19,057,431	w	19,057,536	19,057,564	4	19,057,743	28	312
w12.1	2L	w	20,647,383	both	20,650,022	20,650,080	3	20,660,246	58	12,863
w12.1	2R	both	19,507,783	w	19,507,922	19,508,175	9	19,509,105	253	1,322
w12.10	2L	both	12,629,786	w	12,629,801	12,629,801	1	12,629,974	0	188
w12.10	2R	both	16,391,946	w	16,392,135	16,392,135	1	16,393,000	0	1,054
w12.10	3R	both	20,238,920	w	20,238,934	20,239,515	10	20,239,756	581	836
w12.12	2L	both	7,273,166	w	7,273,992	7,273,992	1	7,274,497	0	1,331
w12.13	3R	both	7,743,368	w	7,743,534	7,743,715	5	7,743,963	181	595
w12.13	X	w	20,367,477	cs	20,367,515	20,367,822	4	20,367,981	307	504
w12.2	3R	both	9,483,096	w	9,483,445	9,483,558	2	9,484,473	113	1,377
w12.3	2R	w	5,770,073	both	5,770,645	5,770,949	2	5,772,000	304	1,927
w12.5	2L	w	4,735,516	both	4,737,743	4,737,873	4	4,738,122	130	2,606
w12.5	3R	both	15,861,811	w	15,862,255	15,862,255	1	15,862,732	0	921
w12.5	X	cs	3,096,697	w	3,096,864	3,096,981	4	3,097,142	117	445
w12.6	2L	both	11,749,165	w	11,749,523	11,749,741	6	11,751,342	218	2,177
w12.6	3L	both	1,969,142	w	1,969,167	1,969,472	10	1,969,698	305	556
w12.7	2L	both	17,511,467	w	17,511,647	17,512,095	9	17,512,545	448	1,078
w12.7	3L	w	11,441,258	both	11,441,592	11,441,888	9	11,442,773	296	1,515
w12.7	3R	w	4,724,519	both	4,725,092	4,725,092	1	4,725,162	0	643
w12.9	3L	w	6,092,918	both	6,092,945	6,092,960	2	6,093,074	15	156
w13.1	2R	w	16,209,421	both	16,209,728	16,209,812	2	16,209,965	84	544
w13.1	3R	both	5,389,489	w	5,390,167	5,390,167	1	5,390,348	0	859
w13.10	2L	w	3,086,353	both	3,086,822	3,086,849	2	3,087,035	27	682
w13.10	3R	w	22,468,466	both	22,468,785	22,469,341	11	22,469,546	556	1,080
w13.10	3R	w	29,336,093	both	29,336,412	29,336,434	2	29,336,883	22	790
w13.11	2L	w	20,333,724	both	20,333,904	20,334,029	4	20,334,221	125	497
w13.11	2R	w	12,006,441	both	12,006,887	12,007,016	4	12,007,169	129	728
w13.11	3R	both	17,247,743	w	17,248,227	17,248,227	1	17,248,693	0	950
w13.12	2L	w	3,188,550	both	3,188,693	3,188,693	1	3,188,991	0	441
w13.12	3R	both	28,437,439	w	28,437,716	28,438,088	5	28,438,211	372	772
w13.13	2R	w	23,368,576	both	23,369,217	23,369,284	4	23,369,353	67	777
w13.2	3R	w	29,501,854	both	29,502,203	29,502,506	3	29,503,321	303	1,467
w13.4	2R	both	12,399,554	w	12,400,593	12,400,748	2	12,401,054	155	1,500
w13.4	3L	both	3,493,539	w	3,493,758	3,494,459	9	3,494,629	701	1,090
w13.5	3L	w	999,224	both	999,241	1,000,086	13	1,000,597	845	1,373
w13.6	3R	both	17,602,447	w	17,602,811	17,602,811	1	17,603,193	0	746
w13.7	2R	w	24,410,367	both	24,410,425	24,411,028	9	24,411,260	603	893

w13.7	3R	both	22,097,788	w	22,097,806	22,097,806	1	22,097,911	0	123
w13.8	2L	both	20,899,413	w	20,899,531	20,899,565	2	20,899,804	34	391
w13.8	3R	w	18,280,325	both	18,280,445	18,280,814	5	18,280,891	369	566
w13.9	3R	w	14,380,139	both	14,380,494	14,380,494	1	14,381,450	0	1,311
w15.1	3L	w	8,756,837	both	8,756,944	8,756,944	1	8,757,116	0	279
w15.1	3R	w	4,840,339	both	4,841,378	4,841,437	2	4,841,646	59	1,307
w15.11	2L	w	10,702,467	both	10,702,533	10,702,861	7	10,703,489	328	1,022
w15.13	2R	both	4,684,864	w	4,685,028	4,685,040	2	4,685,109	12	245
w15.13	3R	both	28,166,394	w	28,166,523	28,166,542	5	28,166,765	19	371
w15.13	3R	both	30,861,946	w	30,862,087	30,862,087	1	30,862,184	0	238
w15.2	2R	w	14,707,329	both	14,707,387	14,707,387	1	14,707,461	0	132
w15.2	2R	w	14,707,494	both	14,707,553	14,708,095	17	14,708,297	542	803
w15.2	2R	w	18,117,429	both	18,117,693	18,118,235	8	18,118,947	542	1,518
w15.2	3L	w	18,862,293	both	18,862,465	18,862,469	2	18,862,962	4	669
w15.2	3L	w	22,512,156	both	22,512,563	22,513,303	10	22,513,961	740	1,805
w15.3	X	cs	14,500,840	w	14,501,131	14,501,202	11	14,501,490	71	650
w15.4	2R	both	11,243,206	w	11,243,438	11,243,486	3	11,243,711	48	505
w15.4	2R	both	15,962,802	w	15,963,212	15,964,780	5	15,964,915	1,568	2,113
w15.4	2R	both	23,632,920	w	23,634,275	23,634,275	1	23,634,562	0	1,642
w15.4	3L	both	20,210,881	w	20,210,981	20,210,981	1	20,211,178	0	297
w15.4	X	w	16,203,974	cs	16,204,044	16,204,676	22	16,205,059	632	1,085
w15.5	2R	w	7,025,757	both	7,026,202	7,026,202	1	7,026,593	0	836
w15.5	3L	both	5,442,768	w	5,443,158	5,443,175	3	5,443,259	17	491
w15.6	2L	w	10,235,657	both	10,235,809	10,235,809	1	10,236,527	0	870
w15.6	2R	w	17,162,333	both	17,162,485	17,162,680	8	17,162,734	195	401
w15.9	2R	both	18,590,172	w	18,590,475	18,590,512	3	18,590,544	37	372
w3.1	3L	w	1,482,240	both	1,482,481	1,482,758	5	1,483,378	277	1,138
w3.1	X	w	5,187,973	cs	5,190,961	5,190,961	1	5,191,569	0	3,596
w3.11	2L	both	13,829,140	w	13,829,423	13,830,211	5	13,831,334	788	2,194
w3.11	2R	both	11,173,177	w	11,173,410	11,173,863	5	11,174,456	453	1,279
w3.12	2L	both	10,554,176	w	10,554,747	10,554,931	4	10,555,222	184	1,046
w3.13	X	w	13,104,981	cs	13,105,118	13,105,459	3	13,106,504	341	1,523
w3.14	3L	both	19,283,024	w	19,285,314	19,285,749	7	19,285,929	435	2,905
w3.15	2L	both	7,842,242	w	7,842,418	7,843,622	28	7,843,678	1,204	1,436
w3.15	2L	both	9,873,348	w	9,873,483	9,874,768	35	9,875,477	1,285	2,129
w3.15	3R	both	11,575,861	w	11,576,815	11,576,815	1	11,577,090	0	1,229
w3.15	3R	w	18,459,776	both	18,460,755	18,460,755	1	18,461,025	0	1,249
w3.16	3R	both	17,110,837	w	17,111,279	17,112,002	4	17,112,273	723	1,436
w3.17	3L	w	11,901,481	both	11,902,997	11,903,104	6	11,903,449	107	1,968
w3.18	3R	both	19,371,681	w	19,373,001	19,373,221	7	19,374,304	220	2,623
w3.24	2L	both	4,196,679	w	4,196,989	4,197,610	4	4,197,683	621	1,004
w3.24	3L	both	9,322,510	w	9,323,195	9,323,506	6	9,323,872	311	1,362
w3.24	3R	w	28,296,711	both	28,297,291	28,297,838	7	28,298,349	547	1,638
w3.25	3L	w	11,627,016	both	11,627,355	11,628,329	18	11,628,571	974	1,555
w3.25	3L	w	21,391,017	both	21,391,705	21,392,176	4	21,392,664	471	1,647
w3.26	X	w	3,905,310	cs	3,905,585	3,905,784	5	3,906,151	199	841
w3.4	2L	both	1,041,871	w	1,042,323	1,042,323	1	1,042,450	0	579
w3.4	2L	both	10,346,414	w	10,346,642	10,346,642	1	10,346,912	0	498
w3.6	3R	both	22,014,728	w	22,014,875	22,015,241	2	22,015,289	366	561
w3.9	3L	w	11,049,075	both	11,049,906	11,050,639	14	11,051,418	733	2,343
w3.9	3L	w	14,358,057	both	14,358,674	14,359,406	13	14,359,437	732	1,380
w4.1	2L	both	2,429,537	w	2,429,897	2,431,771	11	2,433,980	1,874	4,443
w4.1	2R	both	6,515,204	w	6,515,733	6,515,860	4	6,516,104	127	900

w4.11	2L	both	16,796,507	w	16,797,044	16,797,261	3	16,797,376	217	869
w4.12	3L	both	3,329,144	w	3,330,222	3,330,412	4	3,330,659	190	1,515
w4.15	3L	both	20,575,839	w	20,576,146	20,577,099	7	20,577,511	953	1,672
w4.16	2L	both	5,752,541	w	5,753,073	5,753,782	16	5,754,748	709	2,207
w4.17	3R	w	6,145,376	both	6,145,615	6,145,953	4	6,146,708	338	1,332
w4.2	2R	w	16,054,201	both	16,054,377	16,055,057	12	16,056,541	680	2,340
w4.2	2R	w	23,357,682	both	23,358,022	23,358,022	1	23,358,402	0	720
w4.2	2R	w	23,358,581	both	23,358,690	23,359,771	17	23,359,843	1,081	1,262
w4.2	3L	w	10,800,267	both	10,800,377	10,801,257	17	10,801,501	880	1,234
w4.3	2L	both	11,027,059	w	11,027,195	11,027,195	1	11,027,615	0	556
w4.3	2L	both	19,556,624	w	19,557,175	19,557,236	3	19,557,558	61	934
w4.5	3R	both	28,556,876	w	28,557,836	28,557,836	1	28,561,498	0	4,622
w4.8	X	w	4,640,912	cs	4,641,638	4,641,725	2	4,642,152	87	1,240
w4.9	3L	both	1,673,428	w	1,673,619	1,673,673	3	1,673,955	54	527
w4.9	3R	w	13,929,028	both	13,929,132	13,929,246	5	13,929,489	114	461
w4.9	X	cs	12,838,609	w	12,838,748	12,838,850	3	12,839,444	102	835
w5.1	3R	w	11,290,918	both	11,291,153	11,291,173	2	11,291,312	20	394
w5.11	2L	w	19,713,343	both	19,713,512	19,714,061	9	19,714,415	549	1,072
w5.11	2R	w	16,335,888	both	16,335,973	16,337,415	9	16,337,591	1,442	1,703
w5.11	2R	w	17,878,078	both	17,878,444	17,879,339	14	17,879,381	895	1,303
w5.11	3R	w	8,809,908	both	8,811,078	8,811,412	5	8,812,129	334	2,221
w5.11	X	cs	16,416,836	w	16,417,040	16,417,592	5	16,417,903	552	1,067
w5.12	2L	w	4,224,713	both	4,225,022	4,225,046	2	4,225,327	24	614
w5.13	2R	both	16,229,089	w	16,229,122	16,229,122	1	16,229,181	0	92
w5.14	2L	both	1,663,265	w	1,663,300	1,663,793	7	1,664,893	493	1,628
w5.14	3L	w	10,266,680	both	10,267,000	10,267,094	19	10,267,243	94	563
w5.17	3L	both	3,834,385	w	3,834,723	3,835,049	3	3,835,300	326	915
w5.2	X	w	11,259,002	cs	11,259,951	11,259,951	1	11,260,251	0	1,249
w5.3	2L	w	20,522,579	both	20,522,674	20,522,760	3	20,522,900	86	321
w5.3	3L	both	5,525,261	w	5,526,883	5,527,279	8	5,527,823	396	2,562
w5.3	3L	both	22,288,565	w	22,288,980	22,288,980	1	22,289,862	0	1,297
w5.4	2L	both	10,526,928	w	10,526,973	10,527,466	16	10,527,586	493	658
w5.4	2R	both	10,766,166	w	10,766,576	10,767,179	5	10,767,222	603	1,056
w5.5	3R	w	5,952,322	both	5,952,608	5,952,712	3	5,953,484	104	1,162
w5.9	2L	w	19,274,310	both	19,275,502	19,276,149	18	19,277,129	647	2,819
w5.9	3R	w	25,153,037	both	25,153,700	25,153,700	1	25,154,434	0	1,397
w5.9	X	w	8,943,852	cs	8,944,235	8,944,235	1	8,945,205	0	1,353

Table 2.S3. Summary sequencing statistics for all 196 individuals and two parental lines used in this study.

				Depth of coverage					
Stock	Sequencing Batch	Read Length (bp)	Average Insert Size (bp)	Chr *X*	Chr *2L*	Chr *2R*	Chr *3L*	Chr *3R*	Chr *4*
cs1.1	1	151	244	34x	66x	64x	63x	65x	75x
cs1.2	1	151	231	33x	64x	61x	60x	62x	58x
cs1.3	1	151	233	32x	61x	59x	58x	60x	66x
cs1.4	1	151	227	37x	73x	69x	69x	71x	83x
cs1.5	1	151	244	40x	80x	76x	76x	78x	62x
cs1.6	1	151	251	39x	78x	74x	73x	76x	62x
cs1.7	1	151	232	35x	68x	65x	65x	67x	70x
cs1.8	1	151	234	40x	79x	75x	75x	77x	62x
cs12.1	2	126	245	18x	37x	35x	35x	36x	76x
cs12.2	2	126	249	16x	31x	30x	30x	30x	54x
cs12.3	2	126	252	16x	33x	31x	32x	32x	70x
cs12.4	2	126	249	20x	41x	39x	39x	40x	72x
cs12.5	2	126	245	18x	36x	34x	34x	35x	65x
cs12.6	2	126	246	20x	41x	39x	39x	40x	56x
cs12.7	2	126	248	18x	37x	35x	35x	36x	73x
cs12.8	2	126	252	18x	37x	35x	35x	36x	81x
cs12.9	2	126	246	17x	34x	32x	33x	33x	62x
cs12.10	2	126	250	18x	37x	35x	36x	36x	70x
cs12.11	2	126	245	18x	37x	35x	35x	35x	57x
cs12.12	2	126	251	17x	35x	33x	33x	34x	58x
cs12.13	2	126	250	21x	42x	40x	40x	41x	60x
cs12.14	2	126	248	19x	38x	35x	36x	36x	57x
cs12.15	2	126	243	18x	37x	35x	35x	36x	61x
cs12.16	2	126	248	16x	33x	32x	32x	32x	69x
cs12.17	2	126	251	20x	42x	39x	39x	40x	52x
cs12.18	2	126	242	16x	32x	31x	31x	31x	63x
cs13.1	2	126	256	17x	33x	31x	31x	32x	78x
cs13.2	2	126	259	13x	26x	24x	24x	25x	60x
cs13.3	2	126	236	20x	40x	38x	38x	39x	60x
cs13.4	2	126	255	16x	33x	31x	31x	32x	79x
cs13.5	2	126	250	19x	37x	35x	35x	36x	59x
cs13.6	2	126	259	21x	42x	40x	40x	41x	52x
cs13.7	2	126	248	20x	39x	37x	37x	38x	71x
cs13.8	2	126	187	15x	29x	28x	28x	28x	64x
cs13.9	2	126	252	14x	28x	27x	27x	27x	50x
cs13.10	2	126	242	18x	36x	34x	34x	35x	69x
cs13.11	2	126	251	16x	33x	31x	31x	32x	67x
cs13.12	2	126	256	16x	33x	31x	31x	32x	63x
cs13.13	2	126	256	20x	40x	38x	38x	39x	74x
cs13.14	2	126	252	18x	35x	34x	34x	34x	70x
cs13.15	2	126	258	19x	37x	36x	35x	36x	46x
cs13.16	2	126	250	19x	39x	37x	37x	38x	74x
cs13.17	2	126	254	20x	40x	38x	38x	39x	71x
cs13.18	2	126	258	15x	31x	30x	30x	30x	67x
cs14.1	2	126	252	17x	35x	33x	33x	34x	61x
cs14.2	2	126	248	16x	33x	31x	32x	32x	38x
cs14.3	2	126	244	17x	35x	33x	33x	34x	38x

cs14.4	2	126	241	19x	38x	36x	36x	37x	38x
cs14.5	2	126	248	20x	40x	38x	38x	39x	48x
cs14.6	2	126	252	20x	42x	39x	40x	40x	43x
cs14.7	2	126	252	18x	37x	35x	35x	36x	47x
cs14.8	2	126	253	17x	34x	32x	32x	33x	45x
cs14.9	2	126	244	15x	31x	29x	29x	30x	46x
cs14.10	2	126	240	19x	37x	35x	35x	36x	52x
cs14.11	2	126	250	20x	41x	39x	39x	40x	55x
cs14.12	2	126	245	20x	40x	37x	37x	38x	75x
cs14.13	2	126	251	21x	43x	41x	41x	42x	74x
cs2.1	1	151	238	41x	80x	76x	76x	78x	57x
cs2.2	1	151	247	25x	49x	47x	47x	48x	49x
cs2.3	1	151	242	40x	78x	75x	74x	77x	36x
cs2.4	1	151	231	39x	77x	73x	73x	75x	45x
cs2.5	1	151	236	39x	77x	73x	73x	75x	34x
cs2.6	1	151	259	35x	69x	66x	65x	67x	35x
cs2.7	1	151	249	38x	74x	71x	70x	73x	47x
cs2.8	1	151	258	43x	84x	81x	80x	83x	49x
cs5.1	1	151	255	28x	55x	53x	53x	54x	50x
cs5.2	1	151	242	36x	71x	68x	68x	70x	41x
cs5.3	1	151	223	32x	63x	60x	60x	61x	41x
cs5.4	1	151	242	35x	69x	66x	65x	67x	40x
cs5.5	1	151	232	33x	64x	61x	61x	63x	39x
cs5.6	1	151	242	27x	53x	50x	50x	51x	49x
cs5.7	1	151	244	30x	59x	56x	56x	58x	43x
cs5.8	1	151	243	33x	64x	61x	61x	63x	39x
cs6.1	1	151	251	36x	69x	66x	66x	68x	39x
cs6.2	1	151	245	34x	65x	62x	62x	64x	42x
cs6.3	1	151	247	30x	58x	56x	56x	57x	36x
cs6.4	1	151	240	28x	56x	53x	53x	54x	39x
cs6.5	1	151	250	40x	77x	73x	73x	75x	33x
cs6.6	1	151	243	34x	68x	65x	64x	66x	42x
cs6.7	1	151	248	40x	78x	75x	74x	76x	34x
cs6.8	1	151	245	32x	63x	60x	60x	62x	39x
cs7.1	1	151	234	32x	61x	58x	58x	59x	35x
cs7.2	1	151	254	39x	75x	72x	72x	74x	36x
cs7.3	1	151	250	34x	65x	62x	61x	63x	41x
cs7.4	1	151	252	41x	79x	76x	75x	77x	44x
cs7.5	1	151	256	42x	82x	80x	78x	80x	45x
cs7.6	1	151	264	36x	70x	67x	67x	69x	35x
cs7.7	1	151	252	32x	63x	60x	60x	61x	41x
cs7.8	1	151	231	45x	87x	84x	83x	86x	47x
cs8.1	1	151	246	43x	84x	80x	80x	82x	41x
cs8.2	1	151	231	33x	64x	61x	61x	63x	30x
cs8.3	1	151	229	41x	81x	78x	77x	79x	53x
cs8.4	1	151	235	38x	74x	71x	70x	72x	36x
cs8.5	1	151	259	34x	67x	64x	64x	66x	45x
cs8.6	1	151	229	38x	74x	71x	71x	73x	46x
cs8.7	1	151	254	35x	68x	65x	65x	67x	44x
cs8.8	1	151	231	36x	70x	67x	66x	68x	46x
cs9.1	1	151	238	28x	54x	52x	52x	53x	42x
w11.1	2	126	247	17x	34x	32x	32x	33x	34x
w11.2	2	126	258	18x	37x	36x	35x	36x	41x

w11.3	2	126	256	17x	34x	32x	32x	33x	31x
w11.4	2	126	255	19x	37x	35x	35x	36x	39x
w11.5	2	126	255	21x	41x	39x	39x	40x	43x
w11.6	2	126	250	17x	35x	33x	33x	34x	39x
w11.7	2	126	258	15x	30x	29x	29x	29x	36x
w11.8	2	126	254	18x	36x	34x	34x	35x	30x
w11.10	2	126	258	16x	32x	31x	31x	32x	38x
w11.11	2	126	255	18x	36x	34x	34x	35x	33x
w11.12	2	126	259	17x	34x	32x	32x	33x	37x
w11.13	2	126	255	16x	32x	30x	30x	31x	28x
w12.1	2	126	251	21x	43x	40x	41x	41x	41x
w12.2	2	126	252	19x	37x	35x	35x	36x	37x
w12.3	2	126	261	19x	39x	37x	37x	38x	33x
w12.4	2	126	268	8x	15x	15x	14x	15x	37x
w12.5	2	126	261	18x	36x	35x	35x	36x	36x
w12.6	2	126	257	20x	39x	37x	38x	38x	35x
w12.7	2	126	254	17x	33x	32x	32x	32x	33x
w12.9	2	126	256	19x	39x	37x	37x	38x	32x
w12.10	2	126	256	18x	37x	35x	35x	35x	38x
w12.11	2	126	256	20x	40x	38x	38x	38x	34x
w12.12	2	126	255	24x	48x	46x	45x	47x	35x
w12.13	2	126	261	22x	44x	42x	42x	43x	35x
w13.1	2	126	258	17x	34x	32x	32x	33x	40x
w13.2	2	126	263	19x	38x	35x	36x	36x	39x
w13.4	2	126	256	17x	35x	33x	33x	34x	35x
w13.5	2	126	253	18x	37x	35x	35x	36x	32x
w13.6	2	126	255	22x	44x	42x	42x	43x	36x
w13.7	2	126	256	18x	35x	33x	33x	34x	33x
w13.8	2	126	245	16x	33x	31x	31x	32x	36x
w13.9	2	126	248	16x	31x	30x	30x	31x	35x
w13.10	2	126	254	18x	35x	34x	34x	34x	33x
w13.11	2	126	255	15x	31x	29x	29x	30x	37x
w13.12	2	126	253	21x	43x	40x	40x	41x	32x
w13.13	2	126	256	19x	39x	37x	37x	38x	38x
w15.1	2	126	248	20x	39x	37x	37x	38x	41x
w15.2	2	126	249	18x	37x	35x	35x	36x	35x
w15.3	2	126	251	20x	40x	38x	37x	39x	34x
w15.4	2	126	247	20x	40x	38x	38x	39x	38x
w15.5	2	126	271	20x	40x	38x	38x	39x	35x
w15.6	2	126	244	19x	38x	36x	36x	37x	35x
w15.7	2	126	242	17x	35x	33x	33x	34x	32x
w15.8	2	126	257	17x	34x	33x	33x	33x	38x
w15.9	2	126	259	19x	38x	36x	36x	37x	36x
w15.10	2	126	244	19x	39x	37x	37x	38x	37x
w15.11	2	126	246	18x	37x	35x	35x	36x	32x
w15.12	2	126	225	23x	46x	43x	44x	45x	36x
w15.13	2	126	291	11x	23x	22x	22x	22x	30x
w3.1	1	101	250	19x	36x	34x	34x	35x	32x
w3.4	1	101	244	23x	45x	42x	42x	44x	36x
w3.5	1	101	249	21x	42x	40x	40x	41x	39x
w3.6	1	101	242	21x	40x	38x	38x	39x	36x
w3.8	1	101	250	18x	36x	34x	34x	35x	34x
w3.9	1	101	243	26x	51x	49x	48x	50x	25x

w3.11	1	101	243	19x	39x	37x	36x	38x	38x
w3.12	1	101	240	30x	59x	56x	56x	58x	15x
w3.13	1	101	242	22x	44x	42x	42x	43x	32x
w3.14	1	101	248	23x	46x	44x	44x	45x	42x
w3.15	1	101	241	25x	48x	46x	46x	47x	45x
w3.16	1	101	241	26x	51x	48x	48x	50x	23x
w3.17	1	101	238	22x	43x	41x	41x	42x	36x
w3.18	1	101	238	16x	32x	30x	30x	31x	32x
w3.21	1	101	227	24x	48x	46x	46x	47x	30x
w3.24	1	101	236	22x	44x	42x	42x	43x	42x
w3.25	1	101	242	22x	44x	41x	41x	43x	40x
w3.26	1	101	242	18x	35x	33x	33x	34x	34x
w4.1	1	101	244	19x	37x	35x	35x	36x	46x
w4.2	1	101	242	19x	38x	36x	36x	37x	34x
w4.3	1	101	216	20x	40x	38x	38x	39x	41x
w4.4	1	101	239	20x	40x	38x	38x	39x	37x
w4.5	1	101	234	22x	42x	40x	40x	42x	39x
w4.6	1	101	243	23x	44x	42x	42x	43x	37x
w4.7	1	101	235	22x	43x	41x	40x	42x	40x
w4.8	1	101	242	24x	48x	45x	45x	47x	34x
w4.9	1	101	234	26x	53x	50x	50x	51x	39x
w4.11	1	101	226	23x	44x	42x	42x	43x	39x
w4.12	1	101	240	21x	41x	39x	39x	40x	40x
w4.13	1	101	223	24x	47x	44x	44x	46x	38x
w4.15	1	101	233	22x	42x	40x	40x	41x	38x
w4.16	1	101	228	25x	48x	46x	45x	47x	41x
w4.17	1	101	226	26x	51x	48x	48x	49x	28x
w5.1	1	101	224	25x	50x	47x	47x	49x	37x
w5.2	1	101	233	25x	49x	46x	46x	48x	38x
w5.3	1	101	223	21x	40x	38x	38x	39x	29x
w5.4	1	101	217	22x	42x	40x	39x	41x	40x
w5.5	1	101	221	21x	42x	39x	39x	41x	36x
w5.6	1	101	236	26x	51x	48x	48x	50x	32x
w5.7	1	101	231	19x	38x	36x	36x	38x	39x
w5.8	1	101	221	24x	47x	45x	44x	46x	31x
w5.9	1	101	212	28x	57x	54x	54x	55x	36x
w5.11	1	101	233	23x	45x	42x	42x	44x	34x
w5.12	1	101	243	24x	47x	45x	45x	46x	31x
w5.13	1	101	240	22x	44x	42x	42x	43x	30x
w5.14	1	101	238	25x	51x	48x	48x	49x	36x
w5.15	1	101	240	25x	49x	46x	46x	48x	35x
w5.16	1	101	245	21x	41x	39x	39x	40x	31x
w5.17	1	101	227	24x	48x	46x	46x	47x	36x

Table 2.S4. PCR primers used to validate selected NCO events.

Stock	Chr	Forward Primer	Reverse Primer
cs2.8	*2L*	AATCGCCATTATCCCAAAGA	AATGCAGTGGGAACTCAAAAA
w4.3	*2L*	ACTGTGGAAGGAGGGGTTTT	ATTGATATTGCCTCGGTGGA
w5.12	*2L*	AGCGGGCTGCAGATTAGATA	GGAGAGCTCGACCAAAGACT
cs2.8	*2L*	CACCCAAATTATACCCGGATT	TCAACAGAAGAAGGGAATCCA
w4.3	*2L*	CGAAACGCATCAGTCAGTGT	CATCATAATTTGCCGCCCTA
cs5.8	*2L*	CGAGTAGCTTGTCCCATTCC	CTGCAAACATGTGCTCATACAA
w4.11	*2L*	GCAATTGACCAGTGTGTTTGA	GTCGGTCTGATGGGTCCTT
cs6.4	*2L*	TCAATTTCTGGTTTTATGGAATTTT	TTTGCTGATAGGAAGTGTTGGA
cs6.4	*2L*	TCACGTGCAGTCACTGAAAA	CAATGGCACGGTCAACAA
cs2.2	*2L*	TCCCCTCGACGTCAGATACT	GGAGTTCGAGGAGGAAGTGA
w3.4	*2L*	TCGATACTTGTGGCGATGTG	TGCTAAATCCCCTTGTGGAC
w5.12	*2L*	TTGCGTCTTCTAATGCTAATGC	GCCGCAGCTATTCATCAACT
cs2.1	*2L*	TTTTTGCCACACACACACAT	CATTCCAAACCACTCCATCC
w5.13	*2R*	ACAAGGATGCCAAGTTCGAG	GTCGCCCTGGTAGTGAAGAG
cs5.5	*2R*	ACTAGCGAACACGCCACCT	TTTGTTATAGACCGACCAATGC
w4.2	*2R*	AGCGCTTGCACATAACTCCT	TGAGCTGGAGACTTATAGCAACC
w3.12	*2R*	ATGAACGGCGGTCACACT	TTAGATGTTGATAATTGTGGTATGC
cs7.6	*2R*	CACTGGTGCTGATACTGAAGAA	CCGTTACTCTTTCCCAACCA
cs7.6	*2R*	CGGATTATTCAAACTTGACATTT	CCCAACCACACACACTTCAA
cs5.3	*2R*	CTGCAGTTGCGTCATTGTG	ACGAGTTGCGAAATGAGTCC
cs5.3	*2R*	GGCGGGTGTGGAGAATATGT	GCCTCCATTAGGCAAAGTCA
cs7.3	*2R*	TAACCATAGACCGCATCCA	CTGTCCGAGCCACAGACC
cs7.4	*2R*	TCAGCCGCAGGCTATTACTT	GCTGGGGATTCTGATTTTGA
cs7.2	*2R*	TTTGTCTGCACCCAAAACCT	CTCGCTCTTTCCTTGTGCTC
cs8.6	*2R*	TTTTGTAACAGCGATGATTTTGA	AGCATTCGCATTCAATTAACA
w5.3	*3L*	AACTAGAGCGGAACGACGTG	TGGTATGCGTACCGTGTGAC
cs2.2	*3L*	AGGAAGCCGAAAGCTACTCC	CACTGCAAGTGGCCAAAAAT
cs8.4	*3L*	CAACAGTTGGTGTTGGCTTG	CTCCCGGCTGTTTGTTTAGA
cs7.2	*3L*	CGTTCTTGGCACAATTAGCA	TTGATGTCATTTCCGTCGAA
cs1.8	*3L*	GGACTTTAGATGGGCAGCA	ACCAATTTATTGGGCCTGAG
cs2.4	*3L*	GTCCACTGTGCAAATGGTGA	CAGACGACGAGCACAAACTC
cs7.6	*3R*	AAAACACGTCAGCCATAATTTTT	AAACAATTGCCCAAAATGGT
cs1.8	*3R*	AAGTGTTTTCTGGCCAATGC	CTCGCTTCTCGGTTCAGTTT
w4.17	*3R*	CACGATAGTAGAAATTTGCACACA	ACCCACTTTCACATCCGAAG
cs7.2	*3R*	CCACCTGCTCGTTCTACATTG	CGCTTTCGTGGAGAAACACT
cs7.5	*3R*	GCGATGATGAATCCTCCTTC	CACGGAGGTGGATCTGAAAT
w5.9	*3R*	GTGGACGACGCAAAGATGTA	GACTAGATGGCTTGCCTTGC
cs7.5	*3R*	TACGGGATCTGGGTCATAGC	CACTGAAGAGCCGAAAGACC
w3.15	*3R*	TGAAGCGAATCAACTCTAGGC	CAGAATGGTGGCTGGATCTT
cs7.6	*3R*	TGCAAGTTGTTTTGCTGCAC	TGGGAAAATTAAGCAATGGAA
cs8.6	*3R*	TTCTGCTGGCAAGCAACTAA	TTGCAGCTAGTCTCGGGTTT
w3.1	*X*	AAAACTTGAGAGCCTTTCTTGG	AACTTTTTCTGATGGTATACACAAATG
cs8.6	*X*	AATGCCCCATCCTCCATATT	GGGGGAACTCTCTCTCTCGT
cs13.8	*X*	AGGCGGCTGTGATAATTTGA	AGACTCCATGCGGAATTAGG
w5.11	*X*	ATAATATCTGGTCGTACAGGACACT	GAATTGGCACCAATGACAC
cs1.4	*X*	CATTGCACTGCTCTCGAAAC	TTTCGGCCAAGATTCAGACT
cs6.4	*X*	CCAATACGGAAATTTGCATTC	GTTTGACCTACCGACCGAAA
w3.26	*X*	CGGGAAGCGATAGATGTGG	AGCAGTACGCTGATGACACC
w3.21	*X*	CGGTGGCTCTGCCTCTTC	GAAGCACTTATGGGTGAACGA

w3.1	*X*	GAGGACATGCCTGCTTCTTC	TGTTGGTGTACAAGGGGTGA
cs8.5	*X*	GCAAAGGATGGAAGGATGAA	TCCGGTGTGGACTCTATTGG
cs6.6	*X*	GCATGTGTGCGTGAATGAAT	TAATTTCCAATCGCCTGACG
w5.9	*X*	GCGTCGAGTCGAGTTGAGTT	TTCGGACGATTTAATCAAAAA
w3.13	*X*	GGAACAAAAGCCATTTCCAA	CATTCCCACATTGACCAACA
cs7.7	*X*	GGCTCGGTTCTTAGCTTGTG	GGTTTCGGCCAGGATTTTAG
w4.8	*X*	GGCTTCTCCGTGATCGAGT	CATGCCCAGCTCCCTGAC
cs5.4	*X*	TACGGAATGCAATCCCCTAT	TCTCCATGGTGGAGGAGTTC
cs6.1	*X*	TAGAAGTGACTGCGCCACAC	GATGCAACATGTCGATGCTC
cs7.7	*X*	TGGACAATGCGTTCCAAGTA	ATTTGCAGCGAGCCATAAAG
cs6.1	*X*	TGGCTACACTTGGAGAAATGC	ACAGGTGGATGCAGAAGGAG
cs5.1	*X*	TGTCAGCTACGGTTTTCACG	TGGCCAGAGTAGAACCAAGTG
w4.8	*X*	TGTCCTTTGGCTTGTCCTTC	GAGCTACCGCGTCGAATAAC
cs2.5	*X*	TTTTAGAGTGCCCGAGCCTA	GGGCTACTGTCATTCGAGGA

Table 2.S5. Detailed information on all 52 DCO events and one TCO event.

Stock	Chromosome	Class	Father	5' SNP ID	3' SNP ID	Gap (bp)
cs6.7	chrX	DCO	Canton-S	3,470,244	14,910,603	11,440,359
w13.4	chrX	DCO	w^{1118}	4,303,591	11,555,096	7,251,505
w3.13	chrX	DCO	w^{1118}	4,595,161	17,964,570	13,369,409
w4.16	chrX	DCO	w^{1118}	5,512,195	8,525,084	3,012,889
w11.1	chrX	DCO	w^{1118}	5,717,641	15,942,557	10,224,917
w13.13	chrX	DCO	w^{1118}	6,274,954	15,313,486	9,038,532
w13.12	chrX	DCO	w^{1118}	6,435,947	20,420,893	13,984,946
cs1.4	chrX	DCO	Canton-S	6,819,891	12,524,581	5,704,690
cs6.1	chrX	DCO	Canton-S	6,993,026	14,246,463	7,253,437
w4.8	chrX	DCO	w^{1118}	8,313,927	14,811,430	6,497,504
cs5.5	chrX	DCO	Canton-S	8,357,766	15,897,003	7,539,237
w3.16	chrX	DCO	w^{1118}	10,060,737	14,173,641	4,112,905
w15.12	chrX	DCO	w^{1118}	10,130,288	19,184,442	9,054,154
cs8.6	chr2L	DCO	Canton-S	1,791,388	21,725,307	19,933,919
w11.11	chr2L	DCO	w^{1118}	2,868,831	19,267,621	16,398,791
cs14.3	chr2L	DCO	Canton-S	3,218,450	18,652,650	15,434,200
w12.1	chr2L	DCO	w^{1118}	3,818,509	12,982,386	9,163,877
cs12.14	chr2L	DCO	Canton-S	3,830,541	15,120,260	11,289,720
w12.2	chr2L	DCO	w^{1118}	3,871,409	10,349,701	6,478,293
w4.8	chr2L	DCO	w^{1118}	4,797,461	6,265,936	1,468,475
cs14.5	chr2L	DCO	Canton-S	5,645,591	21,057,885	15,412,294
cs12.2	chr2L	DCO	Canton-S	9,415,301	19,966,247	10,550,947
w5.3	chr2R	DCO	w^{1118}	7,683,210	24,064,116	16,380,907
w5.7	chr2R	DCO	w^{1118}	7,710,045	19,349,557	11,639,512
cs13.17	chr2R	DCO	Canton-S	9,082,882	22,600,396	13,517,514
cs14.10	chr2R	DCO	Canton-S	12,439,739	21,195,647	8,755,909
w12.2	chr2R	DCO	w^{1118}	17,384,710	21,342,753	3,958,043
cs12.4	chr3L	DCO	Canton-S	439,076	11,760,094	11,321,018
w13.1	chr3L	DCO	w^{1118}	457,569	20,056,914	19,599,345
w11.10	chr3L	DCO	w^{1118}	1,128,447	13,641,922	12,513,475
cs7.2	chr3L	DCO	Canton-S	1,143,174	17,536,575	16,393,401
cs12.14	chr3L	DCO	Canton-S	1,689,275	16,832,774	15,143,499
cs8.7	chr3L	DCO	Canton-S	1,876,201	14,745,654	12,869,454
w4.3	chr3L	DCO	w^{1118}	2,335,050	12,796,377	10,461,328
w12.13	chr3L	DCO	w^{1118}	2,582,712	13,009,424	10,426,712
cs8.8	chr3L	DCO	Canton-S	3,277,038	10,843,712	7,566,674
cs1.3	chr3L	DCO	Canton-S	3,278,630	10,900,567	7,621,937
w13.5	chr3L	DCO	w^{1118}	6,524,990	12,830,307	6,305,318
cs12.3	chr3L	DCO	Canton-S	7,244,376	19,616,898	12,372,522
cs2.2	chr3L	DCO	Canton-S	7,587,249	12,036,478	4,449,229
w13.2	chr3L	DCO	w^{1118}	8,629,015	16,628,137	7,999,123
w3.16	chr3L	DCO	w^{1118}	12,112,391	17,120,726	5,008,335
cs14.5	chr3R	TCO	Canton-S	6,470,571	18,989,921	12,519,350
cs7.8	chr3R	DCO	Canton-S	10,935,702	27,604,308	16,668,607
w15.3	chr3R	DCO	w^{1118}	11,170,684	22,472,070	11,301,386
cs1.6	chr3R	DCO	Canton-S	13,915,208	29,902,122	15,986,914
cs1.7	chr3R	DCO	Canton-S	14,321,224	23,407,864	9,086,640
w15.4	chr3R	DCO	w^{1118}	14,427,988	30,440,125	16,012,138
cs12.16	chr3R	DCO	Canton-S	17,699,911	30,823,033	13,123,122

w11.2	chr3R	DCO	w^{1118}	18,154,492	25,359,559	7,205,067
cs12.8	chr3R	DCO	Canton-S	18,486,786	29,393,878	10,907,092
cs14.5	chr3R	TCO	Canton-S	18,989,921	27,338,761	8,348,840
w13.9	chr3R	DCO	w^{1118}	19,687,499	29,644,332	9,956,833
w4.15	chr3R	DCO	w^{1118}	22,673,170	28,101,248	5,428,079

Table 2.S6. Observed and expected numbers of noncrossover chromatids, SCOs, DCOs, TCOs, and greater.

Class	Observed	Expected
Noncrossover chromatid	493	564
SCO	434	312
DCO	52	86
TCO & greater	1	18

SCO: single crossover chromatid, DCO: double crossover chromatid, TCO: triple crossover chromatid.

Table 2.S7. E-values from this study and previously published studies.

This Study					
Chromosome	***X***	***2L***	***2R***	***3L***	***3R***
E0	12%	10%	8%	14%	12%
E1	61%	71%	82%	55%	67%
E2	27%	18%	10%	31%	20%
n	196	196	196	196	196
Published Datasets					
Chromosome	***X***	***2L***	***2R***	***3L***	***3R***
E0	7%	15%	16%	5%	–
E1	49%	76%	76%	71%	–
E2	39%	9%	10%	24%	–
n	2,505	11,495	2,921	1,269	–

*E-values calculated using Weinstein tetrad analysis (Weinstein 1918; but see Hawley and Walker 2003) for this study and three other studies using larger sample sizes (*2L* data: X data: Page et al. 2007; Baker and Carpenter 1972; 2R data: Parry 1973; *3L* data: Collins *et al.* 2012).

Chapter 3: Whole-genome sequencing identifies triploid offspring and a rare noncrossover gene conversion in flies mutant for the synaptonemal complex protein C(3)G

INTRODUCTION

In many organisms, including mice, functional synaptonemal complex (SC) is required both for double-strand break (DSB) formation during meiotic prophase I and for proper DSB repair (Bolcun-Filas *et al.* 2009). Drosophila is unique in that it still forms DSBs at about 20% the level of wild type even in the absence of functional SC (Mehrotra and McKim 2006). The SC in Drosophila is, however, required to repair DSBs into crossovers (COs) (Page and Hawley 2001), which ensure the proper segregation of homologous chromosomes during meiosis I. SC mutants that fail to form COs create progeny with a high rate of chromosome nondisjunction (Page and Hawley 2001).

The Drosophila SC protein C(3)G is functionally homologous to the transverse filament proteins SYCP-1 in mammals and ZIP1 in budding yeast (Page and Hawley 2001), and Drosophila females homozygous for a loss-of-function *c(3)G* allele do not build SC and are thus unable to properly repair DSBs by meiotic crossing over (Page and Hawley 2001). Whether DSBs can be repaired as noncrossover gene conversion events (NCOs) in Drosophila females lacking functional SC remains an open question. Although a single study did report no NCO events in *c(3)G* mutants, it examined conversion at only a single locus and did not report the number of progeny studied (Carlson 1972). Separately, we previously reported an unexpectedly high amount of transposable element (TE)-mediated copy number variation (CNV) in wild-type Drosophila offspring (Miller, Smith, *et al.* 2016). It is unknown whether this rate would be affected—or if TE-mediated CNVs would even form—in a mutant with defective homolog synapsis and a reduced number of DSBs.

Here, we used whole-genome sequencing (WGS) to investigate the occurrence of CNV and NCO events along the Drosophila *X* and *2nd* chromosome arms in 95 individual male offspring from females homozygous for a loss-of-function *c(3)G* allele **(Figure 3.1)**. We recovered a single presumed NCO event, suggesting that although extremely rare, repair of DSBs via NCO may be possible in females lacking SC. Additionally, we identified three triploid males, 25 *X-null* males lacking a *Y* chromosome, and six instances of *4th* chromosome gain or loss. The recovery of these individuals is consistent with the observation that SC mutants exhibit high levels of chromosome missegregation. Finally, we identified several large-scale TE-mediated CNVs, one of which, remarkably, was identical to a CNV observed in a different male from a separate study (Miller, Smith, *et al.* 2016), suggesting that recurrent CNVs may occur in Drosophila as they do in humans (Itsara *et al.* 2009).

Figure 3.1: C(3)G cross scheme.
Isogenic Canton-S females carrying the loss-of-function mutant *c(3)G68* were crossed to isogenic *w1118* males carrying the loss-of-function mutant *c(3)G68 e,ca*. Individual females hemizygous for both mutant alleles were collected and crossed to individual isogenic *w1118* males. Individual male offspring were collected on days 10–14 and prepared for WGS.

RESULTS

The SC mutant *c(3)G* may allow rare noncrossover gene conversion events

To assay for the presence of NCO events in *c(3)G* mutants, we whole-genome sequenced 98 individual male offspring from females who were heterozygous for w^{1118} and Canton-S SNPs on the *X* and 2^{nd} chromosomes and homozygous for the loss-of-function mutation $c(3)G^{68}$ **(Figure 3.1)**. (Note: 95 of 98 offspring were analyzed for NCO events as the three triploid offspring discussed below were excluded from this analysis.)

Consistent with previous studies (Page and Hawley 2001), we recovered no COs from *c(3)G* homozygous females, however we did recover and validate a single NCO event on chromosome *2R* from male c3g6.4. This potential NCO was minimally defined by a 4-bp deletion on the 5' end (*2R*:23,350,969, release 6 coordinates) and a single polymorphism on the 3' end (*2R*:23,351,148) **(Figure 3.2)**. The average depth of coverage within the 1kb interval surrounding the two polymorphisms was 54x, similar to the average depth of coverage for chromosome *2R*, demonstrating that this NCO event was not due to a deletion or duplication of this interval. The minimum and maximum possible widths of the gene conversion were 180 bp and 2,507 bp, respectively, falling within the range observed in wild type (Chovnick *et al.* 1971; Miller *et al.* 2012; Miller, Smith, *et al.* 2016). The presence of two polymorphisms, validated by PCR and Sanger sequencing (see Methods), suggests that this NCO is unlikely to be a false positive caused by *de novo* mutation.

Figure 3.2: Structure of the single NCO event recovered in this study.
This NCO, validated by PCR and Sanger Sequencing was defined by a 4 bp deletion on one side and a SNP on the other, both from the w^{1118} line. The NCO has a maximum possible size of 2,507 bp and a minimum size of 180 bp.

Identification of triploid and nondisjunctional progeny by WGS

We identified three males that appeared to be heterozygous for all *X*-chromosome SNPs. Specifically, plots of allele frequency uncovered three males (c3g2.8, c3g2.9, and c3g10.7, see methods for a description of naming conventions) with 50% w^{1118}/Canton-S *X*-chromosome allele frequency as opposed to the 100% w^{1118} or 100% Canton-S *X*-chromosome allele frequency expected for hemizygous males **(Figure 3.3)**. Subsequent depth-of-coverage analysis of these stocks revealed 65% *X*-chromosome depth of coverage compared to each of the other four major autosomal arms **(Figure 3.4, Table 3.S1)**. Phenotypically, these three flies appeared to be male (sex combs were not noted before males were collected and frozen for DNA isolation). Two of the three males came from the same mother (c3g2.8 and c3g2.9), and based on depth-of-coverage analysis, two of the individuals (c3g2.8 and c3g10.7) carried a *Y* chromosome while the third male (c3g2.9) did not **(Table 3.S1)**. The SNP profile of each *X* chromosome suggested the presence of one w^{1118} chromosome and one Canton-S chromosome, supporting the hypothesis that these flies carried two different *X* chromosomes. Using allele-frequency along the 2^{nd} chromosome, we observed that SNPs from the w^{1118} genotype were present 66% of the time while SNPs from the Canton-S genotype were present 33% of the time in all three males, strongly suggesting the presence of three autosomes and

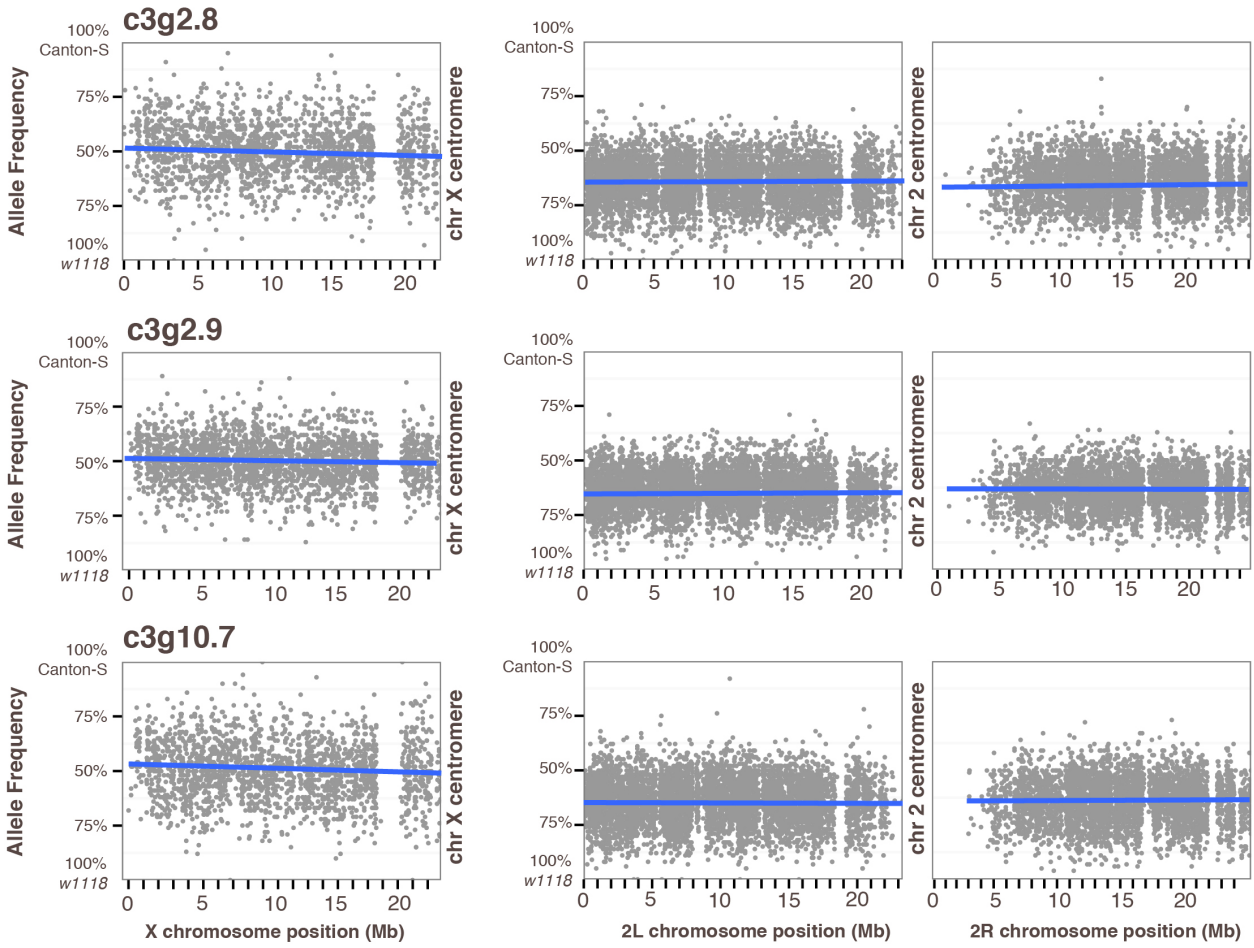

Figure 3.3: Three males were identified as triploid based on their autosomal allele frequency.
A heterozygous male should have a 50%/50% w^{1118}/Canton-S allele frequency for the 2^{nd} chromosome, and because they are hemizygous for the *X* chromosome, a 100% allele frequency for the *X* chromosome. These three males carry a 50% w^{1118}/Canton-S allele frequency for the *X*, suggesting that they carry two distinct *X* chromosomes. They also carry a 67%/33% w^{1118}/Canton-S allele frequency for both arms of the 2^{nd} chromosome, with 67% of the SNPs from the w^{1118} stock, and 33% of the SNPs from the Canton-S genome—which is evidence for the presence of three 2^{nd} chromosomes.

suggesting that the heterozygous female created diploid heterozygous oocytes. These three individuals were classified as triploids and excluded from any NCO or CNV analysis.

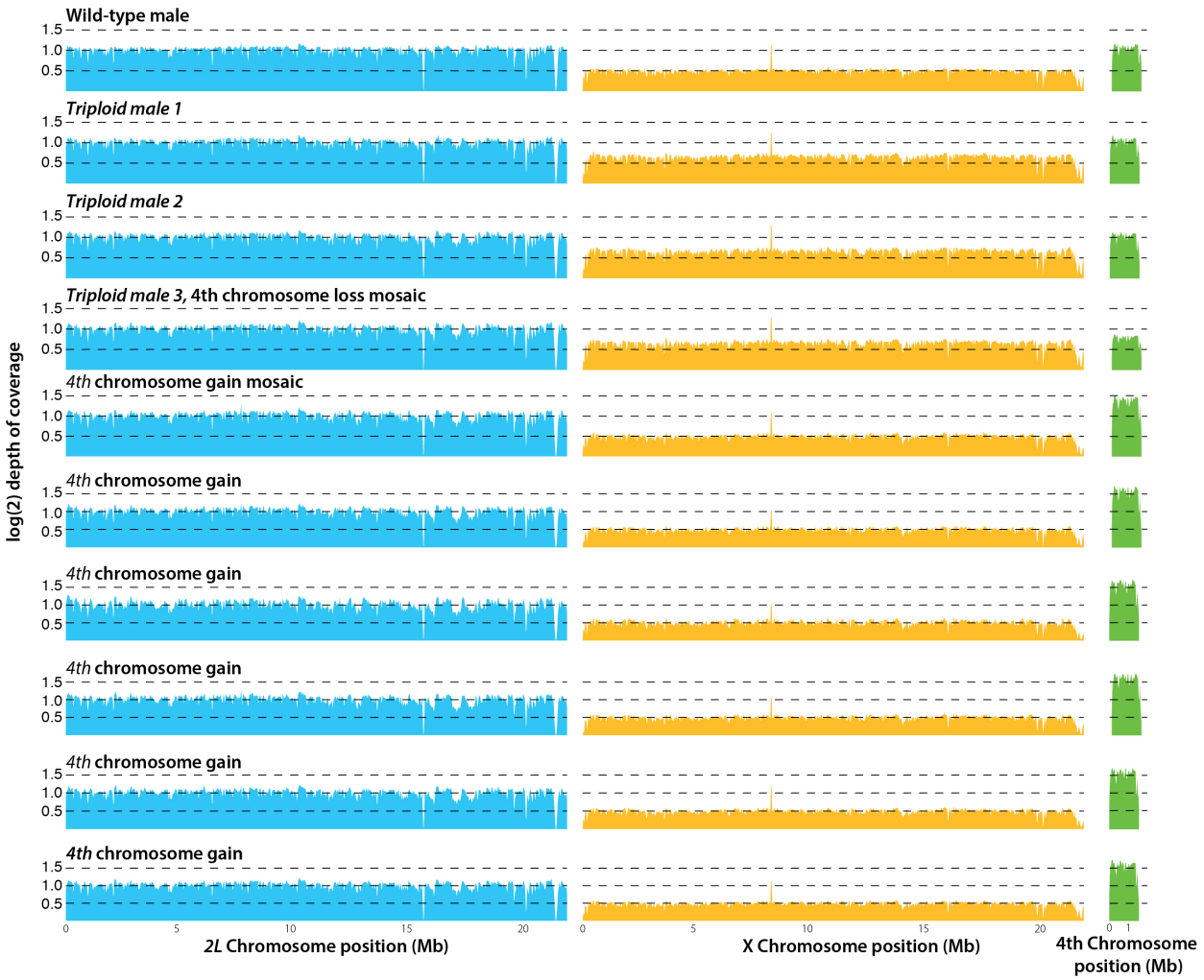

Figure 3.4: Log_2 depth-of-coverage analysis for the *X*, *2nd*, and *4th* chromosomes.
Analysis uncovered a wild-type male, three triploid males, and six individuals with an extra copy of chromosome *4*. Note that one of the triploid males is also a *4th* chromosome mosaic. The log_2 differences for the *X* and *4th* chromosomes use chromosome *2L* as the basis of their log_2 ratio calculation.

Depth-of-coverage analysis also identified one triploid male that was *44/444* mosaic, one diploid male that was *44/444* mosaic, and five diploid males carrying three copies of the *4th* chromosome **(Figure 3.4)**. In addition, we identified 25 *X0* males carrying no *Y* chromosome—evidence of female *X*-chromosome nondisjunction **(Table 3.S1)**. These data give us a rate of female *X*-chromosome nondisjunction of 43% and *4th* chromosome nondisjunction of 10% **(Table 3.1)**. A high degree of chromosome nondisjunction such as this is characteristic of SC

mutants in Drosophila (Page and Hawley 2001). Indeed, previous studies with two different alleles of *c(3)G* [$c(3)G^{17}$ and $c(3)G^{68}$] reported *X* chromosome nondisjunction rates of 32.4% and 39.2%, respectively, and 4^{th} chromosome nondisjunction rates of 17.9% and 26.8%, respectively (Hall 1972). Given our small sample size, it is notable that we recovered a similar *X*-chromosome nondisjunction rate as that reported by Hall (1972). We did recover a significantly lower percentage of 4^{th} chromosome nondisjunction, however we intentionally did not select minute males (those carrying only one 4^{th} chromosome) for sequencing, thus this result is not surprising.

	Counts	**% Adj**
WT	62	
X NDJ	24	43%
4 NDJ	5	10%
X & 4 NDJ	1	
AdJ Total	117	

Table 3.1: Nondisjunction data for 92 non-triploid individuals recovered in this study.
Note that 4^{th} chromosome NDJ is affected by the poor viability of nullo-*4* individuals, and that they were purposely excluded from collection in this study.

Copy-number variation is apparent in males from SC-deficient mothers

Because we had previously identified cases of both inherited and *de novo* CNV formation in individual wild-type males (Miller, Smith, *et al.* 2016), we wondered if females deficient in SC assembly would produce offspring with large CNVs as well. We identified one TE-mediated CNV (a deletion) inherited by multiple progeny and three novel CNV events **(Figure 3.5, Table 3.2)** on the *X* and 2^{nd} chromosome that were each present in only one of the 95 non-triploid individuals studied.

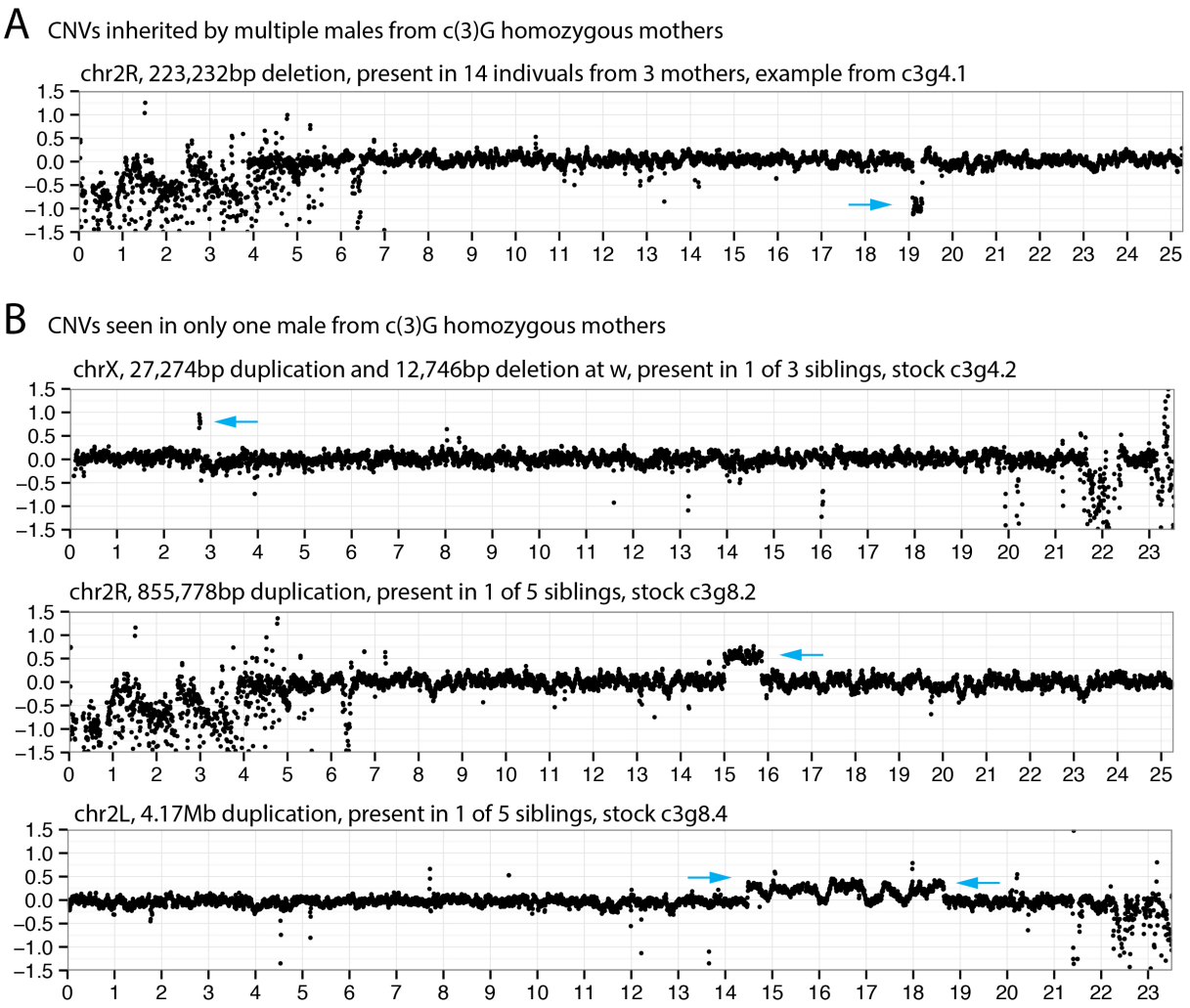

Figure 3.5: Copy-number variants are seen in males from mothers homozygous for *c(3)G* loss-of-function mutations.

Details of each CNV can be found in Table 3.2. (A) One deletion shared among 14 males from 3 different mothers was recovered in this study. (B) Three unique CNVs were identified in this study. Two of these appeared to occur during meiosis, while the third, based on a lower log(2) ratio, likely happened during the first mitotic division.

Type	Chr	Start	End	Size (bp)	Genes Affected	Males Observed In	Event Between	5' Feature	TE Orientation	3' Feature	TE Orientation
Deletion	2R	19,092,386	19,315,617	223,231	27	14 siblings from 3 females	Sisters	cs: DMHFL1 w: -	5'-3' -	cs: DMHFL1 w: -	5'-3' -
Duplication	X	2,753,999	2,781,273	27,274	1	1/3	Sisters	w: Roo cs: -	unknown -	w: Roo cs: -	unknown -
Deletion	X	2,781,273	2,794,019	12,746	3	1/3	Sisters	w: Roo cs: -	unknown -	w: Roo cs: -	unknown -
Duplication	2R	15,005,073	15,860,851	855,778	107	1/5	Sisters	w: hobo (DMHFL1) cs: -	3'-5' -	w: hobo (DMH cs: -	3'-5' -
Duplication	2L	14,482,976	18,655,730	4,172,754	513	1/5	Sisters	w: - cs: -	- -	w: - cs: -	- -

Table 3.2. Detail of TE-mediated copy-number variants recovered in this study.
Visualization of CNVs can be seen in Figure 3.5.

We recovered one apparent *de novo* CNV that occurred at the *w* locus on chromosome *2R.* This was a complex event involving both a deletion and a duplication **(Figure 3.5B, Table 3.2)** and was mediated by unequal crossing over between *Roo* elements. Interestingly, previous work describing ectopic recombination in *D. melanogaster* focused on unequal exchange between *Roo* elements at the *w* locus, similar to the event recovered in this study (Goldberg *et al.* 1983). A second apparent *de novo* CNV recovered in this study was a large duplication that included 107 genes mediated by unequal crossing over between two hobo elements. A CNV with the exact same breakpoints was identified in a previous study and is considered in the discussion. Both *de novo* events were validated using the haplotypes of the siblings—specifically that other siblings had inherited the same chromosome from their moms that did not carry the CNV observed in each individual. Furthermore, these are likely not variants segregating at low frequency as these were single female/male matings.

The remaining novel CNV on chromosome *2L* created a 4.2-Mb duplication not defined by either a TE or low-complexity sequence **(Figure 3.5B, Table 3.2)**. Read pairs show it to be a tandem duplication, with reads on the proximal end of the duplication linked to reads mapping

to the distal end of the duplication. The $\log_2$ depth-of-coverage ratio for this interval is 0.25 **(Figure 3.5B)**—half of that expected for an autosomal duplication that occurs before the first mitotic division. Therefore, it is likely that this duplication occurred during the first mitotic division, possibly as a consequence of a re-replication event that was then repaired by recombination between the duplicated segments (Green *et al.* 2010). It is notable that the fly was able to tolerate this large duplication (which involved 513 genes) present in half of the cells. Although we are unable to fully exclude the possibility that there was selection against cells carrying the large duplication, a $\log_2$ depth-of-coverage ratio of 0.25 does strongly suggest there was no selection against duplication-carrying cells—if there was, the $\log_2$ ratio would fall below 0.25 and perhaps become undetectable.

The recovery of TE-mediated CNVs in females unable to construct SC demonstrates that these CNVs can occur independently of normal meiotic synapsis and DSB formation. As would be expected in a mutant defective in homologous chromosome pairing, all three TE-mediated CNV events recovered appear, based on allele frequency and TE positioning, to be events between sister chromatids and not between homologous chromosomes.

DISCUSSION

The analysis of offspring from females deficient in SC formation allows us to look for evidence of NCO events from females that are unable to repair DSBs as COs during meiosis. Although previous work in Drosophila had reported no conversions at the *rosy* locus in females homozygous for a loss-of-function allele of the transverse filament protein *c(3)G*, it did not report the number of offspring scored (Carlson 1972). In the present study, we recovered one presumed NCO event from 95 *c(3)G* offspring, or 285 chromosome arms scored (NCOs were not assayed from the three triploid individuals identified in this study).

We considered several alternatives to explain the recovery of this candidate NCO event **(Figure 3.2)**. This NCO event occurred on a chromosome with the Canton-S haplotype that, because heterozygous w^{1118}/Canton-S females were crossed to homozygous w^{1118} males, excludes the possibility that it was contributed from the isogenic w^{1118} father. Another possibility is that this event could have been the result of somatic mutation; however, because it was defined by two closely located polymorphisms (a 4-bp deletion and a SNP) that created changes identical to the other haplotype used in this study, this alternative is highly unlikely. A third possibility is that because homologous chromosomes are paired in somatic tissue in Drosophila, we may have recovered an example of a mitotic repair event in a cell fated to become a germline stem cell (Joyce *et al.* 2012; Bosco 2012), although there are no reliable estimates of the rate at which this occurs.

On the other hand, it is also possible that we failed to identify additional NCO events that may have occurred in SNP-poor regions of the genome. Indeed, in a recent study, we identified only 291 of the 549 NCO events that were expected to have occurred, demonstrating

the difficulty in recovering all NCOs from a single meiosis (Miller, Smith, *et al.* 2016). If the apparent NCO recovered was indeed a real meiotic NCO, and if it was indeed the only NCO event that occurred, then we can estimate the NCO rate in a *c(3)G* mutant to be approximately $1x10^{-10}$ per bp per meiosis. This is much lower than the ~$1.9x10^{-8}$ rate reported in wild type (Hilliker *et al.* 1994; Blanton *et al.* 2005; Miller *et al.* 2012; Miller, Smith, *et al.* 2016), which is not surprising given that females homozygous for *c(3)G* mutations produce DSBs at a much lower level (approximately 20%) than wild-type females (Mehrotra and McKim 2006).

Taking into account the reduced number of breaks, if NCOs had occurred at a wild-type rate in a *c(3)G* mutant, we would have expected a total of 27.6 NCO events from the *X* and *2nd* chromosomes in the 92 individuals studied (see Methods). Based on randomly distributing 28 NCO events along the *X* and *2nd* chromosome we expect to recover 11–21 NCOs, as some will be missed due to low SNP density. The recovery of only a single NCO event suggests either that repair by NCO is extremely rare in an SC-deficient mutant or that this single event was due to an alternate repair mechanism discussed above and perhaps did not occur during meiosis.

Not only do *c(3)G* mutants make fewer breaks, they also fail to form crossovers, which leads to increased levels of chromosome nondisjunction. Accordingly, we were able to identify three triploid individuals and several flies with an extra *4th* chromosome **(Figure 3.4).** Aneuploidy for the *4th* chromosome is well tolerated in Drosophila (Ashburner *et al.* 2005), and the occurrence of aneuploidy in a *c(3)G* mutant is not itself surprising, so we were not surprised to recovering several triploid flies in this experiment. Indeed, previous studies have noted an increase in the number of triploid individuals recovered from c(3)G mutants (Lindsley and Zimm 1992).

Finally, we wondered how a lack of synapsis would affect the occurrence of TE-mediated CNVs. Analysis of individual genomes by WGS allowed the identification of one inherited and two apparent *de novo* TE-mediated CNV events. Thus, these events occur at approximately a wild-type rate, even in a background of reduced DSB formation, suggesting that they may be DSB-independent events **(Figure 3.5)**. Remarkably, the 855-kb duplication on *2R* that arose on the w^{1118} haplotype seen in male c3g8.2 exhibited identical duplication breakpoints as a male (cs7.5) recovered in a separate study (Miller, Smith, *et al.* 2016) **(Figure 3.5B, Table 3.2)**. In that work, analysis of the siblings of cs7.5 suggested that the duplication arose *de novo* in that cross, as it was present in only one of five males that inherited the w^{1118} chromosome from their heterozygous mother (their father contributed a Canton-S chromosome). That, combined with the fact that the duplication is seen in only one individual (c3g8.2) in this study as well, strongly suggests that it was also a *de novo* event here.

We also observed a large CNV duplicating over 500 genes that does not appear to have been mediated by recombination between TEs; it likely occurred during the first mitotic division in about half the cells in that individual. Elucidating the rate and mechanisms that contribute to the formation of these events, as well as how the organism tolerates such large duplications will be of great interest as similar duplications and deletions occur in humans, typically with clinically-relevant consequences. Understanding which regions of the genome are amenable to large-scale duplication or deletion may help us predict how duplication and deletion of genes with homologs in humans may be tolerated.

Taken together, this study demonstrates that unbiased sequencing of individual genomes and individual meiotic products even from well-studied mutants such as *c(3)G* may

reveal novel insights. Here we have demonstrated that repair of DSBs in the absence of pairing and synapsis may be possible, that TE-mediated CNV occurs at a rate close to wild-type in a mutant with a reduced number of DSBs, and that flies may be unexpectedly tolerant of large genome duplications created during the mitotic divisions. This study provides a framework for similar experiments in other meiotic mutants in which individual products are studied—all of which are likely to produce unexpected findings.

METHODS

Fly Stocks and husbandry

The loss-of-function alleles *c(3)G^{68} e or c(3)G^{68} e,ca* (Page and Hawley 2001) were placed into *w^{1118}* and Canton-S isogenized stocks (Miller *et al.* 2012)**.** Females heterozygous for *w^{1118}* and Canton-S *X* and *2nd* chromosomes and homozygous for *c(3)G^{68}* were then crossed to isogenic *w^{1118}* males **(Figure 3.1)**. Females were removed after 7 days of egg laying. Male offspring used for sequencing were collected between days 12 and 15. All flies were kept on standard cornmeal-molasses and maintained at 25^{o}C.

DNA preparation and sequencing

DNA was prepared from single adult males using the Qiagen DNeasy Blood & Tissue Kit. All flies were starved for 4 hr before freezing at -80^{o}C for at least 1 hr. One µg of DNA from each was fragmented to 250-bp fragments by adjusting the treatment time to 85 sec using a Covaris S220 sonicator (Covaris Inc.). Libraries were prepared using a Nextera DNA Sample Prep Kit and Bioo Scientific NEXTflex™ DNA Barcodes. The resulting libraries were purified using Agencourt AMPure XP system (Beckman Coulter) then quantified using a Bioanalyzer (Agilent Technologies) and a Qubit Fluorometer (Life Technologies). All samples were run either 150-bp paired-end or 125-bp paired-end on an Illumina HiSeq 2500 run in Rapid Mode using HiSeq Control Software 2.0.12.0 and Real-Time Analysis (RTA) version 1.17.21.3. Secondary Analysis version CASAVA-1.8.2 was run to demultiplex reads and generate FASTQ files. See **Table 3.S1** for summary alignment statistics.

Alignment of DNA sequences, SNP calling, and identification of NCO events

Alignment to the Drosophila reference genome (dm6) was preformed using bwa version 0.7.7-r441 (Li and Durbin 2009). SNPs were identified using SAMtools version 0.1.19-44428cd (Li, Handsaker, *et al.* 2009). Candidate NCO events were identified as described in (Miller, Smith, *et al.* 2016). No candidate COs were identified.

Validation of NCOs by PCR

Nine candidate NCO events were identified in 92 males and examined by PCR; Phusion polymerase (NEB) was used according to the manufacturer's instructions. Only one of the nine putative conversion events validated as real in male c3g6.4 using primers c3g6_4_F1 (5'-GCCACTCCATGTTCTCTTCG-3') and c3g6_4_R1 (5'-CACCCGTGATCAGATTGTCC-3'). Primers used to validate the putative conversion events as false positives are available upon request.

Stock	Sequence
c3g7_1_F1	AAAGCGGGCAATACGAAAA
c3g7_1_R1	GGAGTGCCCAGATTCTCAAG
c3g5_3_F1	CCCAAGAAGCGACACATTG
c3g5_3_R1	CCGATTGTCACAAACTCTGC
c3g6_7_F1	CCATTGTGAGATTTCATTCAGC
c3g6_7_R1	CGAGTGGAGCAAAGCAAAA
c3g7_6_F2	GCGGTGACGAACCAAAAATA
c3g7_6_R2	TGTTGTTTCTGTTGCCTTGG
c3g7_13_F1	TGGTTTCAATTGCCATCACT
c3g7_13_R1	TGCATGCCTGACTAATTGCT
c3g9_5_F1	AAAACATGCGGACGAACAC
c3g9_5_R1	AAGCGCTTATCGAATCAAAC
c3g9_6_F1	AACCCAATGCCACATCAGTT
c3g9_6_R1	TGAACTTATTGGCTTTCAATGG
c3g10_15_F1	TTGCTTGGTACGAATGTTGC
c3g10_15_R1	AAGTGGCCACAAGTGTGCT

Table 3.3: Primers used to check gene conversions in males from C(3)G homozygous mothers

Calculation of expected NCO events

The number of expected NCO events was calculated by performing 100,000 trials of randomly distributing 28 NCO events between chromosomes *X*, *2L*, and *2R*. Events were randomly assigned a size between 50 and 2,000 bp and were counted as observable if they covered a region that included at least one SNP. The average number of recoverable NCOs per trial was 17. Perl code used to run this model is available at https://github/danrdanny/thesis/c3g/.

NCO events are expected to occur at ~1.9×10^{-8} conversions per base pair per meiosis (Chovnick *et al.* 1970; Miller, Smith, *et al.* 2016). Applying this rate to the *X* and *2nd* chromosome only we expect to recover 1.5 NCO events per haploid product (2.1×10^{-8} * (23.5Mb + 23.5Mb + 25.2Mb)). Therefore, with a wild-type number of DSBs, we would have expected to recover 138 NCO events from the *X* and *2nd* chromosomes of 92 individuals (1.5 NCOs * 92 individuals). Using the most basic assumption that NCOs are reduced in concert with the number of DSBs, or 20%, we can calculate the number of NCOs we would expect to recover in *c(3)G* mutants as 27.6 (138 NCOs * 20%). To determine how many of these 27.6 NCOs we should expect to recover when analyzing 92 individual offspring we created a model in which NCO events are randomly placed along the three chromosome arms (*X*, *2L*, and *2R*) in order to create a confidence interval for the number of NCOs we expect to recover. We performed 100,000 trials of randomly distributing 28 NCO events with tract lengths ranging from 50 bp to 2,000 bp along the *X* and *2nd* chromosomes and counted the times an NCO event fell within an interval containing a SNP that differentiates w^{1118} from Canton-S. These NCOs are the only ones that would have been apparent to us. Using this method, we can generate a 95% confidence interval of 11–21 recoverable NCO events from 92 progeny in a background in which DSBs

occur at 20% of wild-type. The recovery of only a single NCO event when 11–21 NCO events were expected to be recovered suggests that either repair by NCO is extremely rare in a SC-deficient mutant, or that this single event was due to an alternate repair mechanism and perhaps did not occur during meiosis.

Depth-of-coverage calculations

Depth of coverage for each chromosome arm was calculated by summing the total read depth for each base position then dividing by the length of the entire chromosome arm. Because of the repetitive nature of the *Y* chromosome, the region studied was limited to *chrY*:332,000–510,000.

Identification of CNV events

CNV events were identified as described in (Miller, Smith, *et al.* 2016).

SUPPLEMENTAL TABLES

Table 3.S1: Summary statistics for stocks sequenced in this study.

		Depth						
Stock	**Read Len**	**X**	**2L**	**2R**	**3L**	**3R**	**4th**	**Y**
c3g1.1	150bp PE	20.0	40.6	38.9	38.9	39.6	41.5	6.3
c3g1.2	150bp PE	26.3	53.1	51.1	51.0	52.0	53.8	8.2
c3g10.1	125bp PE	23.2	45.0	43.3	43.1	44.0	45.2	5.3
c3g10.10	125bp PE	19.2	38.8	37.1	37.2	37.9	39.3	7.0
c3g10.11	125bp PE	19.0	37.5	36.0	36.0	36.7	37.3	4.6
c3g10.12	125bp PE	16.4	33.0	31.6	31.6	32.3	33.3	0.1
c3g10.13	125bp PE	14.6	29.5	28.3	28.2	28.9	28.9	4.3
c3g10.14	125bp PE	15.1	30.1	28.8	28.9	29.5	30.3	4.8
c3g10.15	125bp PE	14.8	30.3	28.9	28.9	29.6	30.3	0.1
c3g10.16	125bp PE	18.4	36.9	35.3	35.2	36.1	36.7	0.1
c3g10.17	125bp PE	21.8	42.6	40.9	40.8	41.6	42.3	4.9
c3g10.18	125bp PE	19.9	39.7	38.0	38.1	38.8	40.3	0.2
c3g10.19	125bp PE	14.8	29.7	28.5	28.6	29.1	29.3	4.3
c3g10.2	125bp PE	16.4	32.8	31.3	31.4	32.0	33.1	5.1
c3g10.20	125bp PE	14.8	30.0	28.9	28.8	29.5	28.4	4.3
c3g10.21	125bp PE	23.1	44.8	43.1	42.9	43.9	43.9	0.1
c3g10.3	125bp PE	18.4	36.8	35.1	35.4	35.9	37.6	5.8
c3g10.4	125bp PE	21.3	42.2	40.5	40.5	41.2	43.1	5.8
c3g10.5	125bp PE	17.1	34.2	32.7	32.7	33.4	34.7	4.9
c3g10.6	125bp PE	15.4	31.4	30.1	30.1	30.7	30.8	4.0
c3g10.7	125bp PE	22.6	34.1	32.6	32.8	33.5	33.8	4.7
c3g10.8	125bp PE	21.5	43.0	41.1	41.1	42.1	43.6	0.1
c3g10.9	125bp PE	21.1	41.7	40.0	40.0	41.0	42.2	0.2
c3g2.11	125bp PE	16.4	32.8	31.6	31.5	32.2	32.5	4.8
c3g2.12	125bp PE	20.4	41.1	39.3	39.5	40.1	61.0	6.9
c3g2.3	150bp PE	24.1	48.1	46.1	46.1	47.0	48.8	7.4
c3g2.6	125bp PE	15.1	30.1	28.9	29.0	29.5	30.4	5.2
c3g2.7	125bp PE	15.9	31.8	30.2	30.5	31.1	32.0	6.7
c3g2.8	125bp PE	22.9	34.5	33.1	33.2	33.9	33.8	4.2
c3g2.9	125bp PE	29.0	43.8	41.9	42.1	42.9	32.5	0.2
c3g21.10	125bp PE	20.0	39.5	38.0	37.9	38.6	38.9	4.6
c3g3.1	125bp PE	19.5	39.4	37.6	37.8	38.5	38.6	6.6
c3g3.2	125bp PE	18.8	37.3	35.9	36.0	36.5	37.8	5.7
c3g3.3	125bp PE	18.6	37.5	36.0	36.0	36.7	36.5	4.9
c3g3.4	125bp PE	20.0	39.8	38.0	38.2	38.9	52.0	0.2
c3g3.5	125bp PE	19.1	38.0	36.3	36.5	37.2	37.8	0.2
c3g4.1	125bp PE	20.3	40.6	38.6	38.9	39.7	40.8	0.2
c3g4.2	125bp PE	15.9	31.9	30.4	30.6	31.2	31.0	5.1
c3g4.3	125bp PE	17.0	34.1	32.9	32.8	33.4	34.5	5.1
c3g5.1	125bp PE	18.3	36.0	34.3	34.4	35.2	35.3	0.1
c3g5.1	125bp PE	18.3	36.0	34.3	34.4	35.2	35.3	0.1
c3g5.10	125bp PE	19.4	39.4	37.6	37.7	38.5	38.9	5.8
c3g5.11	125bp PE	18.6	37.9	36.3	36.3	37.1	38.0	5.5
c3g5.2	125bp PE	19.8	39.6	37.8	37.9	38.7	57.7	5.1
c3g5.3	125bp PE	18.8	37.6	35.9	36.2	36.8	37.8	5.7
c3g5.4	125bp PE	19.0	38.1	36.4	36.6	37.3	37.2	5.2
c3g5.5	125bp PE	19.5	38.7	37.0	37.1	37.8	39.3	0.2

c3g5.6	125bp PE	20.2	39.8	37.9	38.1	38.8	58.5	4.7
c3g5.8	125bp PE	18.2	36.3	34.7	34.8	35.5	35.6	0.1
c3g5.9	125bp PE	18.2	36.8	34.8	35.2	35.9	37.2	0.1
c3g6.1	125bp PE	17.7	35.2	33.8	33.9	34.5	35.5	5.8
c3g6.10	125bp PE	20.5	39.4	37.6	37.7	38.6	38.8	4.0
c3g6.11	125bp PE	21.3	42.5	40.6	40.7	41.6	42.4	0.1
c3g6.12	125bp PE	19.1	38.3	36.5	36.8	37.3	38.6	6.1
c3g6.13	125bp PE	18.2	36.7	35.0	35.1	35.9	36.9	5.3
c3g6.14	125bp PE	17.1	34.8	33.3	33.3	34.0	34.4	0.1
c3g6.15	125bp PE	16.7	33.8	32.1	32.3	33.0	32.7	5.1
c3g6.16	125bp PE	13.4	27.0	25.5	25.8	26.4	27.2	0.1
c3g6.17	125bp PE	18.7	37.4	35.7	35.8	36.5	37.9	5.1
c3g6.17	125bp PE	18.7	37.4	35.7	35.8	36.5	37.9	5.1
c3g6.2	125bp PE	18.0	36.3	34.7	34.8	35.5	36.2	5.1
c3g6.3	125bp PE	18.0	35.7	34.2	34.3	34.9	35.5	4.3
c3g6.4	125bp PE	21.0	42.1	40.1	40.4	41.1	42.3	5.4
c3g6.5	125bp PE	17.8	35.9	34.3	34.4	35.0	35.3	4.9
c3g6.6	125bp PE	17.5	34.7	33.2	33.4	33.9	36.4	5.6
c3g6.7	125bp PE	21.0	41.1	39.4	39.4	40.1	41.1	4.9
c3g6.8	125bp PE	18.0	36.8	34.8	35.2	35.8	37.2	5.8
c3g6.9	125bp PE	18.7	37.4	35.8	35.8	36.6	37.6	0.1
c3g7.1	125bp PE	19.3	38.6	36.9	36.9	37.8	39.0	0.1
c3g7.10	125bp PE	18.3	36.7	35.3	35.2	36.0	37.0	4.8
c3g7.11	125bp PE	18.3	36.5	35.0	35.1	35.7	35.8	5.8
c3g7.12	125bp PE	17.3	34.0	32.9	32.8	33.4	34.0	4.4
c3g7.13	125bp PE	17.7	35.4	33.9	33.9	34.6	35.6	4.8
c3g7.14	125bp PE	20.4	41.0	39.3	39.3	40.1	40.8	5.4
c3g7.15	125bp PE	19.1	38.6	36.9	36.9	37.7	38.1	5.6
c3g7.16	125bp PE	20.1	39.5	38.0	38.0	38.5	39.9	4.1
c3g7.17	125bp PE	18.5	36.4	35.0	35.0	35.7	35.6	4.3
c3g7.2	125bp PE	18.8	37.6	35.9	35.9	36.8	37.4	0.1
c3g7.3	125bp PE	19.5	39.3	37.5	37.6	38.3	39.3	6.8
c3g7.4	125bp PE	17.8	35.9	34.4	34.5	35.1	35.8	5.4
c3g7.5	125bp PE	21.4	42.6	40.7	41.0	41.7	43.5	0.2
c3g7.6	125bp PE	17.9	35.7	34.0	34.3	34.9	36.0	0.2
c3g7.7	125bp PE	17.6	35.6	33.9	34.1	34.7	35.8	5.1
c3g7.8	125bp PE	18.7	37.0	35.4	35.5	36.1	36.7	4.3
c3g7.9	125bp PE	18.8	36.8	35.5	35.4	36.0	36.5	4.4
c3g8.10	125bp PE	17.5	35.4	33.9	33.9	34.6	35.6	5.2
c3g8.2	125bp PE	13.2	26.7	26.1	25.6	26.1	27.3	3.5
c3g8.3	125bp PE	20.4	41.3	39.6	39.6	40.4	60.9	4.9
c3g8.4	125bp PE	15.1	31.9	29.0	29.2	29.7	30.0	4.4
c3g8.5	125bp PE	15.0	30.3	29.0	28.9	29.7	28.9	5.0
c3g9.1	125bp PE	14.7	29.3	28.1	28.2	28.6	29.9	4.1
c3g9.2	125bp PE	15.1	30.3	29.1	29.0	29.7	30.5	3.6
c3g9.3	125bp PE	14.1	27.9	26.7	26.8	27.2	28.5	3.8
c3g9.4	125bp PE	15.7	31.7	30.3	30.3	31.0	31.2	4.4
c3g9.5	125bp PE	16.3	33.2	31.6	31.7	32.4	33.1	0.2
c3g9.6	125bp PE	17.2	34.4	32.9	32.9	33.6	49.8	4.5
c3g9.7	125bp PE	17.2	35.2	33.5	33.6	34.4	34.5	0.2
c3g9.8	125bp PE	21.5	43.4	41.6	41.6	42.6	42.9	0.1

Chapter 4: Rare recombination events generate sequence diversity among balancer chromosomes in *Drosophila melanogaster*

This chapter is adapted from: Miller, D. E., K. R. Cook, N. Yeganeh Kazemi, C. B. Smith, A. J. Cockrell et al., 2016 Rare recombination events generate sequence diversity among balancer chromosomes in Drosophila melanogaster. Proceedings of the National Academy of Sciences 201601232.

ABSTRACT

Multiply inverted balancer chromosomes that suppress exchange with their homologs are an essential part of the genetic toolkit in *Drosophila melanogaster*. Despite their widespread use, the organization of balancer chromosomes has not been characterized at the molecular level, and the degree of sequence variation among copies of any given balancer chromosome is unknown. To map inversion breakpoints and study potential diversity in the descendants of a structurally identical balancer chromosome, we sequenced a panel of laboratory stocks containing the most widely used *X*-chromosome balancer, *First Multiple 7* (*FM7)*. We mapped the locations of *FM7* breakpoints to precise euchromatic coordinates and identified the flanking sequence of breakpoints in heterochromatic regions. Analysis of SNP variation revealed megabase-scale blocks of sequence divergence among currently used *FM7* stocks. We present evidence that this divergence arose by rare double crossover events that replaced a female-sterile allele of the *singed* gene (sn^{X2}) on *FM7c* with wild type sequence from balanced chromosomes. We propose that, although double crossover events are rare in individual crosses, many *FM7c* chromosomes in the Bloomington *Drosophila* Stock Center have lost sn^{X2} by this mechanism on a historical timescale. Finally, we characterize the original allele of the *Bar* gene (B^1) that is carried on *FM7* and validate the hypothesis that the origin and subsequent reversion of the B^1 duplication is mediated by unequal exchange. Our results reject a simple non-recombining, clonal mode for the laboratory evolution of balancer chromosomes and have implications for how balancer chromosomes should be used in the design and interpretation of genetic experiments in *Drosophila*.

INTRODUCTION

Balancer chromosomes are genetically engineered chromosomes that suppress crossing over with their homologs and are used for many purposes in genetics, including construction of complex genotypes, maintenance of stocks, and estimation of mutation rates. Balancers typically carry multiple inversions that suppress genetic exchange or result in the formation of abnormal meiotic products if crossing over does occur **(Figure 4.1A)**. For example, single crossovers inside the inverted segment create acentric or dicentric chromosomes that will fail to segregate properly during meiosis or large deletions or duplications that likely result in inviable gametes (Beadle and Sturtevant 1935; Novitski and Braver 1954). Balancers also often carry recessive lethal or sterile mutations to prevent their propagation as homozygotes as well as dominant markers for easy identification. First developed for use in *Drosophila melanogaster*, balancer chromosomes remain one of the most powerful tools for genetic analysis in this species (Ashburner *et al.* 2005).

Despite their widespread use, very little is known about the organization of *Drosophila* balancer chromosomes at the molecular level. Since their original syntheses decades ago, balancers have undergone many manipulations including the addition or removal of genetic markers. Additionally, rare recombination events can cause spontaneous loss of deleterious alleles on chromosomes kept over balancers in stock, as well as loss of marker alleles on balancer chromosomes themselves (Ashburner *et al.* 2005). Likewise, recent evidence has shown that sequence variants can be exchanged between balancer chromosomes and their wild type homologs *via* gene conversion during stock construction or maintenance (Cooper *et al.* 2008;

Blumenstiel *et al.* 2009). Thus, substantial variation may exist among structurally identical balancer chromosomes due to various types of sequence exchange.

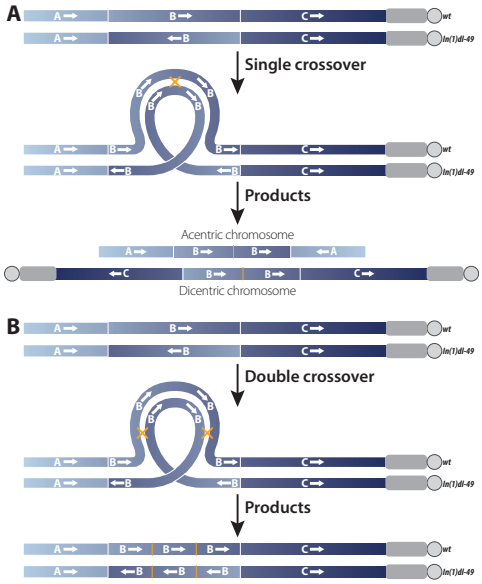

Figure 4.1. Consequences of a single or double crossover between a wild-type X- chromosome (wt) and an X-chromosome carrying a single inversion [*In(1)dl-49*].
Euchromatin is shown in blue, heterochromatin is shown in grey, and centromeres are depicted as circles. Thin white lines mark locations of inversion breakpoints, and yellow crosses/thin lines mark locations of crossover events. A) A single crossover event within the inverted segment results in the formation of chromosomes with deletions and zero (acentric) centromeres or duplications and two (dicentric) centromeres, neither of which will segregate properly during meiosis. B) A double crossover within an inverted segment results in intact chromosomes with one centromere that will segregate properly during meiosis.

To gain insight into the structure and evolution of balancer chromosomes, we have undertaken a genomic analysis of the most commonly used *X*-chromosome balancer in *D. melanogaster, First Multiple 7* (*FM7*). We have focused on *FM7* because this *X*-chromosome balancer series lacks lethal mutations and can therefore easily be sequenced in a hemizygous or homozygous state. In addition, the *FM7* chromosome has been shown to pair normally along most of its axis with a standard *X*-chromosome, providing a structural basis for possible exchange events (Gong

et al. 2005). Moreover, although details of how early balancers in *D. melanogaster* were created are not fully recorded, the synthesis and cytology of the *FM7* series is reasonably well documented (Ashburner *et al.* 2005).

The earliest chromosome in the *FM7* series, *FM7a*, was constructed using two progenitor *X*-chromosome balancers, *FM1* and *FM6*, to create a chromosome carrying three inversions – *In(1)sc*^{8}, *In(1)dl-49*, and *In(1)FM6* – relative to the wild type configuration (Merriam 1968; 1969) (**Figure 4.2A**). Subsequently, a female-sterile allele of *singed* (*sn*^{X2}) was introduced onto *FM7a* to create *FM7c*, which prevents the loss of balanced chromosomes carrying recessive lethal or female-sterile mutations (Merriam and Duffy 1972). More recently, versions of *FM7a* and *FM7c* have been generated that carry transgene insertions that allow balancer genotypes to be determined in embryonic or pupal stages (Casso *et al.* 2000; Le *et al.* 2006; Abreu-Blanco *et al.* 2011; Lattao *et al.* 2011; Pina and Pignoni 2012).

To identify the inversion breakpoints in *FM7* balancers and to study patterns of sequence variation that may have arisen since the origin of the *FM7* series, we sequenced genomes of eight *D. melanogaster* stocks carrying the *FM7* chromosome (four *FM7a* and four *FM7c*). We discovered several megabase-scale regions where *FM7c* chromosomes differ from one another, which presumably arose *via* double crossover (DCO) events from balanced chromosomes **(Figure 4.1B)**. These DCOs eliminate the female-sterile *sn*^{X2} allele in the centrally located *In(1)dl-49* inversion and are expected to confer a fitness advantage to *sn*$^{+}$ chromosomes, either by allowing propagation of *sn*$^{+}$ *FM7* as homozygotes in females or by *sn*$^{+}$ *FM7* males out-competing *sn*^{X2} *FM7* males in culture. We show that loss of the *sn*^{X2} allele is common in *FM7c* chromosomes by screening other *FM7c*-carrying stocks at the Bloomington

Drosophila Stock Center. We also identified the breakpoints of the B^1 duplication carried on *FM7* and provide direct molecular evidence for the role of unequal exchange in the origin and reversion of the B^1 allele (Sturtevant and Morgan 1923; Sturtevant 1925; Muller 1936; Peterson and Laughnan 1963; Gabay and Laughnan 1973). Our results provide clear evidence that the common assumption that balancers are fully non-recombining chromosomes is incorrect on a historical timescale and that substantial sequence variation exists among balancer chromosomes in circulation today.

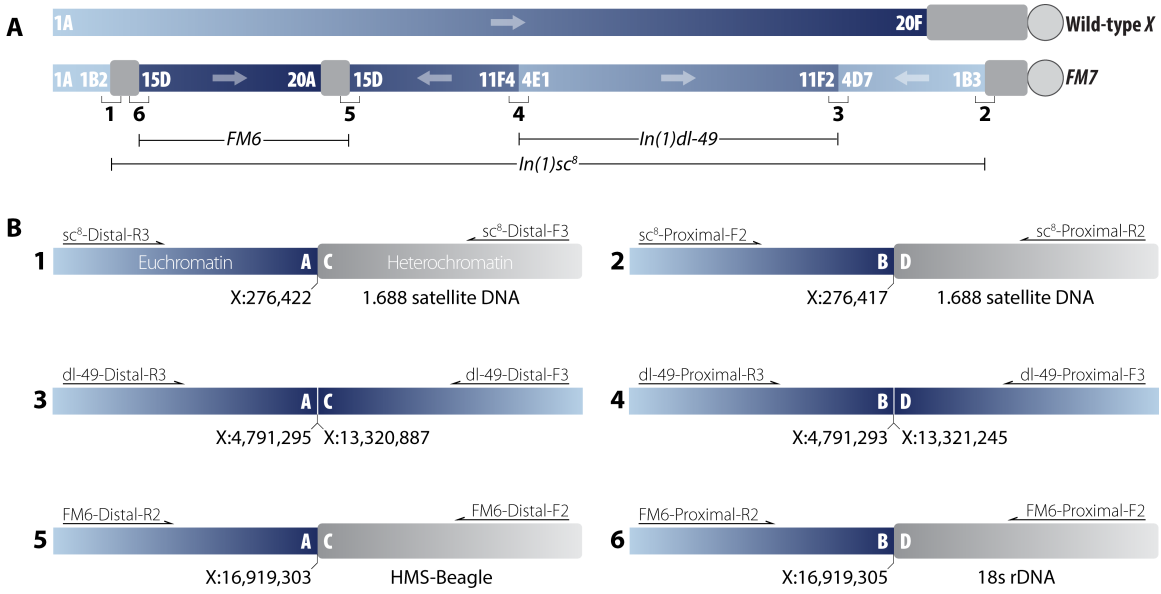

Figure 4.2. Structure of the *FM7* balancer chromosome.
Euchromatin is shown in blue and heterochromatin is shown in grey. **A)** Schematic view of the organization of wild type and *FM7* X-chromosomes. *FM7* contains three inversions (*In(1)sc*^{8}, *In(1)dl-49*, and *In(1)FM6*) relative to wild type. The six breakpoint junctions for the three inversions are numbered 1-6 and are shown in detail in panel B. **B)** Location and organization of inversion breakpoints in *FM7*. Each inversion has two breakpoints that can be represented as A/B and C/D in the standard wild type arrangement and A/C and B/D in the inverted *FM7* arrangement, where A, B, C and D represent the sequences on either side of the breakpoints. Locations of euchromatic breakpoints are on Release 5 genome coordinates, and the identity of the best BLAST match in FlyBase is shown for heterochromatic sequences. Primers used for PCR amplification are shown above each breakpoint (see methods, **Table 4.S2,** and **Table 4.S3** for details). Forward and reverse primers are named with respect to the orientation of the assembled breakpoint contigs, not the orientation of the wild type or *FM7 X*-chromosome.

RESULTS

Identification of *FM7* inversion breakpoints

The inversions carried by *FM7* that confer the ability to suppress recombination were generated by *X*-ray mutagenesis and characterized using genetic and cytogenetic data in the pre-genomic era, and thus the precise locations and molecular nature of their breakpoints remain unknown. To better understand the genomic organization of *FM7* chromosomes, we used whole-genome sequencing to identify breakpoints for the three inversions present on *FM7*: *In(1)sc*^{8}, *In(1)dl-49*, and *In(1)FM6* (**Figure 4.2A**). Based on cytological data, it is known that both breakpoints of *In(1)dl-49* lie in euchromatic regions (Painter 1934; Hoover 1938; Lindsley and Zimm 1992). However, for both *In(1)sc*^{8} and *In(1)FM6*, one breakpoint is euchromatic and the other lies in centric heterochromatin (Sidorov 1931; Patterson 1933; Muller and Prokofyeva 1934; Patterson and Stone 1935; Grell and Lewis 1956; Lindsley and Zimm 1992).

Our general strategy to identify breakpoint regions is as follows. We sequenced eight *FM7*-carrying stocks to approximately 50-fold coverage with paired-end Illumina data and mapped reads to the *D. melanogaster* reference genome (see summary statistics in **Table 4.S1**). We identified clusters of split or discordantly mapped reads from all stocks in the vicinity of expected breakpoint locations based on previous cytological data, then performed *de novo* assembly of split/discordant reads and their mate-pairs (reads from the other end of the same paired-end sequenced fragments). Breakpoint contigs identified by sequence analysis were then used to design PCR amplicons that span breakpoints, and resulting PCR amplicons were Sanger sequenced to verify breakpoint assemblies. Using this approach, we were able to map

euchromatic breakpoints of all three inversions on the *FM7* chromosome to reference genome coordinates, as well as characterize the sequence composition of the heterochromatic breakpoints for both *In(1)sc*^{8} and *In(1)FM6* (**Figure 4.2B**).

The distal breakpoint of the X-ray-induced *In(1)sc*^{8} inversion has been localized near bands 1B2-3 between the *achaete (ac)* and *scute (sc)* genes (Sidorov 1931; Patterson 1933; Muller and Prokofyeva 1934; Patterson and Stone 1935; Campuzano *et al.* 1985; Lindsley and Zimm 1992). We identified a cluster of split/discordant reads in *FM7* stocks around X:276,500 (predicted band 1A7) of the type expected in the vicinity of an inversion breakpoint. Split/discordant reads from ±1.5 kb around the putative *In(1)sc*^{8} inversion breakpoint (which map to the A and B regions) and their mate-pairs (which map to the C and D regions) were extracted from all *FM7* strains, pooled together and assembled to identify candidate A/C and B/D breakpoint sequences. BLAST analysis of the resulting assembly revealed two contigs of 506 bp and 551 bp. The euchromatic components of these contigs mapped to nucleotides X:276,417–276,422 in the Release 5 genome sequence between *ac* and *sc*, within an intron of *CG32816*. The heterochromatic components of these contigs contained copies of the 1.688 satellite DNA repeat (Hsieh and Brutlag 1979) that covers approximately half of the *X*-chromosome centric heterochromatin (Lohe *et al.* 1993). The locations and sequences of candidate breakpoints for *In(1)sc*^{8} were used to design PCR primers that yielded amplicons in all stocks carrying *In(1)sc*^{8} but not in stocks lacking this inversion (**Table 4.S2**). Sanger sequencing of PCR amplicons spanning breakpoint regions confirmed the sequence of A/C and B/D *de novo* assemblies. Comparison of A/C and B/D fragments revealed a 6-bp sequence (TTTCGT) from the *ac–sc* region that is present at both breakpoint junctions, suggesting the X-ray-induced

inversion event created a small, staggered break at the euchromatic end. Our candidate A/C and B/D breakpoint regions also had strong BLAST hits to an *In(1)sc*^{8} A/C junction from the *Dp(1;f)1187* mini-chromosome and the corresponding wild type A/B junction identified in a previous study (Glaser and Spradling 1994). Both our A/C fragment and that obtained by Glaser & Spradling (1994) map the euchromatic part of the distal *In(1)sc*^{8} breakpoint to the same location in the *D. melanogaster* euchromatin and contain 1.688 satellite DNA in their heterochromatic part **(Table 4.S2)**.

In(1)dl-49 is an X-ray-induced inversion (Muller 1926) with both distal and proximal breakpoints in euchromatic regions at bands 4D7–E1 and 11F2–4, respectively (Painter 1934; Hoover 1938; Lindsley and Zimm 1992). We identified clusters of split/discordant reads for the distal breakpoint near X:4,791,300 (predicted band 4D5) and for the proximal breakpoint from approximately X:13,321,200–13,321,900 (predicted band 11F6). These candidate breakpoint intervals were also identified using Breakdancer (Chen *et al.* 2009), an independent method which is able to predict inversions that have two euchromatic breaks. We extracted split/discordant reads within ±1.5 kb of each of the putative *In(1)dl-49* breakpoint intervals plus their mate-pairs, pooled reads from both breakpoints, then performed *de novo* assembly followed by PCR and Sanger sequencing **(Table 4.S1)**. As expected, PCR amplification was successful in stocks carrying *In(1)dl-49* but failed in stocks lacking *In(1)dl-49* (**Table 4.S2**). Sanger sequencing verified the sequence of the A/C and B/D breakpoint assemblies. Both the proximal and distal breakpoints were found in unique genomic regions, with the distal break occurring between X:4,791,293–4,791,295 in an intron of *CG42594* and the proximal break occurring from X:13,320,887–13,321,245 in an intergenic region between *SET domain containing 2 (Set2)*

and *Neuropilin and tolloid-like (Neto)* **(Figure 4.2B)**. The breakpoint in the A/C fragment contained a small 3-bp duplication that is not present in the reference genome, suggesting repair of a small staggered break during the inversion process. A 358-bp deletion was found in the B/D fragment, possibly due to resection during the repair event, which explains why the split/discordant reads for the proximal breakpoint mapped to an interval in the reference genome rather than to a single point.

The distal euchromatic breakpoint of the X-ray-induced *In(1)FM6* was reported to be near bands 15D–E (Grell and Lewis 1956; Lindsley and Zimm 1992). We identified a cluster of split/discordant reads near X:16,919,300 (predicted band 15D3) in all *FM7* stocks and used these reads and the corresponding reads from the other end of the same paired-end sequenced fragments for *de novo* assembly. PCR using primers based on the two resulting putative A/C and B/D contigs validated that this breakpoint was present in all *FM7* stocks but not in stocks that lack the *In(1)FM6* inversion (**Table 4.2**), and Sanger sequencing of amplicons verified the predicted breakpoint sequences **(Table 4.S3)**. Euchromatic components of the A/C and B/D fragments map to the same location within an intron of *CG45002* and reveal that the inversion introduced a 1-bp deletion (X:16,919,304) **(Figure 4.2B)**. The heterochromatic part of the A/C fragment contains sequence from the transposable element *HMS-Beagle* (Snyder *et al.* 1982), and the heterochromatic part of the B/D fragment contains 18S rDNA sequence, consistent with the proximal breakpoint being in *X*-chromosome centric heterochromatin (Tartof and Dawid 1976). The fact that the heterochromatic regions in the A/C and B/D fragments are not the same sequence suggests either a complex breakage/repair event following irradiation or post-inversion rearrangement of sequences at either the A/C or B/D breakpoint. Nevertheless,

the structure of the euchromatic junctions for the *In(1)sc*^{8}, *In(1)dl-49*, and *In(1)FM6* inversions carried on *FM7* show that X-ray-induced mutagenesis can often generate rearrangements with relatively precise breakpoints.

Recombination generates sequence variation among *FM7* chromosomes

It is widely believed that balancers seldom undergo recombination (Theurkauf and Hawley 1992; Hughes *et al.* 2009), giving rise to the idea that they should diverge from each other clonally and thus accumulate deleterious mutations under Muller's Ratchet (Araye and Sawamura 2013). However, previous studies have shown that sequence exchange can occur, albeit rarely, both into and out of balancer chromosomes (Cooper *et al.* 2008; Blumenstiel *et al.* 2009), although the frequency and genomic scale of such events is unknown. To test if ongoing sequence exchange between balancers and homologous chromosomes has occurred since the original synthesis of the first *FM7* chromosome, we identified variants present on only one of the eight *FM7* chromosomes in our sample. Unique variants that differentiate one *FM7* from all others in our sample can arise by either by *de novo* mutation or recombination events that donate sequence from homologous chromosomes to balancers (by either gene conversion or crossing over). However, crossing over is the only mechanism that can explain the large contiguous tracts of sequence variation that are unique to individual *FM7* chromosomes.

As shown in **Figure 4.3B**, we observe megabase-scale tracts of unique variation on three of the eight *FM7* chromosomes (*FM7c-5193*, *FM7c-36337*, *FM7a-23229*), superimposed on a relatively even distribution of unique variants along the remainder of the chromosome. Notably, all of these tracts of unique variation are contained within the *In(1)dl-49* inversion,

span the *sn* locus, and are found only in sn^{+} stocks. These tracts of variants were not caused by placement of the sn^{X2} allele onto *FM7a* to create *FM7c* (Merriam and Duffy 1972), since *FM7c*'s marked with sn^{X2} (*FM7c-616*, *FM7c-3378*) do not differ substantially in their SNP profile from *FM7a*'s in the *sn* region (**Figure 4.5B**). In fact, similarity between *FM7a* and the original *FM7c* is expected in the *sn* region since a *In(1)dl-49* chromosome was a progenitor of *FM7a* (Merriam 1968; 1969), the sn^{X2} allele arose on a *In(1)dl-49* chromosome (Bender 1960), and a sn^{X2} marked *In(1)dl-49* was used as the donor to move sn^{X2} onto *FM7a* to create *FM7c* (Merriam and Duffy 1972). The nature of the sn^{X2} allele was not determined in earlier studies (Paterson and O'Hare 1991); however, we identified a cluster of split/discordant reads at X:7,878,402–7,878,413 that arises from the insertion of an *F*-element in the 2nd coding exon of *sn* that is present only in the sn^{-} stocks *FM7c-616* and *FM7c-3378*. We propose that this *F*-element insertion is the lesion that causes the sn^{X2} allele. Additionally, if the tracts of variants in *FM7c-5193*, *FM7c-36337*, *FM7a-23229* arose from movement of sn^{X2} onto *FM7c*, they would not be unique. Rather, they would form a haplotype shared by all other *FM7c* chromosomes, as is observed in the region surrounding the *g* locus (**Figure 4.5B**). The *FM7c g* haplotype on *FM7a-23229* is unexpected, and suggests that this balancer is actually an *FM7c* that has been mislabeled as *FM7a* because of its sn^{+} phenotype. Together, these results indicate that all chromosomes with large tracts of unique SNPs are *FM7c*'s that lack the sn^{X2} allele.

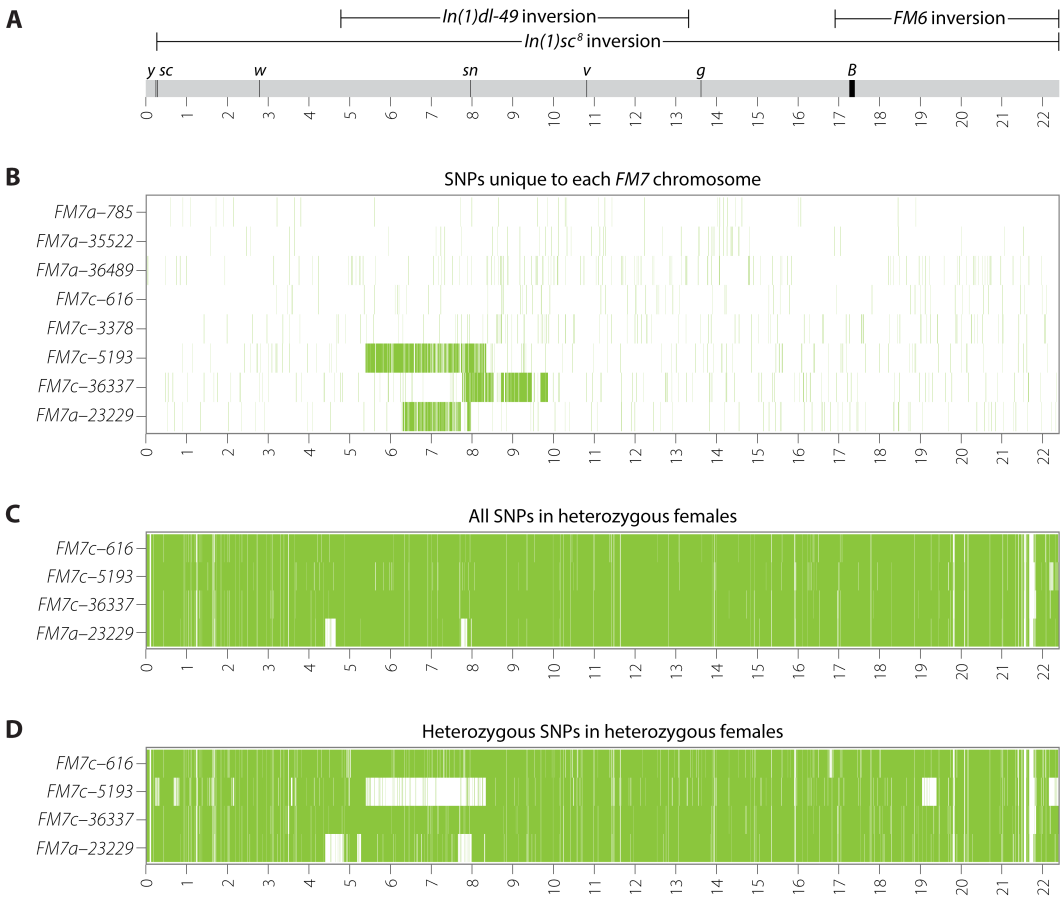

Figure 4.3. Recombination generates sequence diversity among *FM7* balancer chromosomes.
A) Schematic of the wild type *X*-chromosome showing the locations of inversions (oriented with respect to the reference genome, not *FM7*), visible genetic markers, and Release 5 genome coordinates (in Mb). **B)** Heatmap of unique SNPs found in only one *FM7* chromosome in our sample. The density of unique SNPs is plotted in 5 kb windows with a 5 kb offset. The three large tracts of unique SNPs on *FM7c-5193*, *FM7c-36337*, and *FM7a-23229* all are contained fully within *In(1)dl-49* and replace the sn^{X2} allele with wild type sequence. The *FM7a-23229* chromosome is a mislabeled *FM7c* (see **Figure 4.5B**). **C)** Heatmap of all SNPs found in heterozygous female samples carrying *FM7* balancers over different balanced *X*-chromosomes. Genotypes of balanced *X*-chromosomes can be found in **Table 4.1**. Small tracts where few SNPs are present in *FM7a-23229* arise because of common ancestry among the *X*-chromosomes in *FM7*, the balanced chromosome, and the *ISO-1* reference genome (see **Figure 4.5C**). **D)** Heatmap of heterozygous SNPs found in heterozygous female samples carrying *FM7* balancers over different balanced X-chromosomes. Loss of heterozygosity (LOH) is observed for a large tract in *FM7c-5193* that corresponds to the large tract of unique variants for this chromosome shown in panel B. LOH is also observed in *FM7c-5193* for two deletions in the balanced chromosome (*Df(1)JA27* and an uncharacterized deletion on the *Df(1)JA27* chromosome), and for tracts in *FM7a-23229* that share ancestry with y^{1}-ncd^{D} and *ISO-1* (see **Figure 4.5C**)

The number of unique single nucleotide variants expected on each *FM7* chromosome if they evolved clonally and independently under *de novo* mutation alone since their origin in 1968 (Merriam 1968; 1969) to the time our lines were sequenced is approximately 150 (45 years * 26

generations/year * $22x10^{6}$ bp * $5.8x10^{-9}$ mutations/bp/generation (Haag-Liautard *et al.* 2007)). Shared ancestry among chromosomes in our sample, such as for the *FM7c* chromosomes that were generated several years later (Merriam and Duffy 1972), would lower the number of unique variants observed from this expectation. The number of unique variants observed for five out of eight *FM7* chromosomes (56–152 unique SNPs) is less than or nearly equal to the expected value under independent clonal evolution with *de novo* mutation alone. However, the number of unique variants observed for *FM7c-5193*, *FM7c-36337*, *FM7a-23229* (between 541–3,564 unique SNPs) is more than three times higher than expected under clonal evolution with mutation alone, suggesting that the action of additional processes such as gene conversion or crossing over is required to explain these observations. The large tracts of unique variation on *FM7c-5193*, *FM7c-36337*, *FM7a-23229* range between 1.7–3.0 Mb in length and encompass 195–356 genes. Since the average tract length of gene conversion in *D. melanogaster* is approximately 350–450 bp (Hilliker *et al.* 1994; Miller *et al.* 2012), we propose that the large tracts of unique variants on *FM7c-5193*, *FM7c-36337,* and *FM7a-23229* arose by independent DCOs from unrelated chromosomes onto different *FM7* balancer chromosome lineages that replaced sn^{X2} with sn^{+}.

The most obvious donor for sequence exchange onto a balancer chromosome is the chromosome with which it is kept in stock. To test whether the large tracts of unique sequence variation we observe on *FM7* chromosomes are the result of recombination with their homolog in stock, we sequenced heterozygous females from the three stocks with putative DCO events (*FM7c-5193*, *FM7c-36337,* and *FM7a-23229*) and from one negative control with no putative DCO event (*FM7c-616*). If a recent exchange event occurred between the balanced

chromosome and its homolog, we would expect to see a loss of heterozygosity (LOH) in the region where the two chromosomes underwent recombination. As shown in **Figure 4.3C**, the distribution of all SNPs (both homozygous and heterozygous variants) in heterozygous samples is high and relatively constant across the entire *X*-chromosome for three of the four stocks, with two small regions in *FM7a-23229* yielding a paucity of SNPs because of shared ancestry between all *FM7*'s and the y^1 chromosomes in both *ISO-1* and the balanced chromosome (see **Figure 4.5C**). Analysis of heterozygous SNPs in heterozygous females (**Figure 4.3D**) shows a relatively uniform distribution of heterozygous SNPs across the *X*-chromosome, with clear LOH in the exact region of the predicted exchange event for *FM7c-5193*, but not for *FM7c-36337* or *FM7a-23229*. These results indicate that recent exchange between *FM7c-5193* and its balanced homolog can explain the large tract of unique variants on this chromosome. However, the predicted exchange events for *FM7c-36337* or *FM7a-23229* must have occurred sometime in the past with different chromosomes other than those with which they are currently kept in stock.

Intriguingly, all three putative DCOs are contained within the central *In(1)dl-49* inversion, occur on *FM7c* chromosomes, and replace the female-sterile sn^{X2} allele that was present on the original *FM7c* (Merriam and Duffy 1972) with a wild type allele. Although DCOs fully within the *In(1)dl-49* regions are rare (Sturtevant and Beadle 1936; Novitski and Braver 1954), such events would lead to viable *FM7*-bearing gametes. Furthermore, replacement of the female-sterile sn^{X2} allele with sn^{+} is expected to generate *FM7* chromosomes with a fitness advantage relative to the ancestral *FM7c* and thus these rare recombinant chromosomes could quickly increase in frequency in stock. Loss of sn^{X2} could lead to a fitness advantage by allowing

propagation of *sn*$^{+}$ *FM7* as homozygotes in females, although this would lead to loss of balanced mutations in culture, which occurs rarely. Alternatively, *sn*$^{+}$ *FM7c* males may have a fitness advantage in crowded cultures relative to *sn*^{X2} *FM7c* males who have bristle and mechanosensory defects (Lindsley and Zimm 1992; Cant *et al.* 1994). We favor the advantage of *sn*$^{+}$ *FM7c* males in culture as the predominant mechanism by which *sn*$^{+}$ *FM7c* chromosomes replace *sn*^{X2} *FM7c* chromosomes because *FM7c*'s have likely accumulated other female sterile mutations over time, which would reduce the fitness of homozygous *sn*$^{+}$ *FM7c* females even in the absence of *sn*^{X2}.

To address how often loss of *sn*^{X2} occurs in *FM7c* chromosomes, we screened and classified the *sn* phenotype in males from 630 stocks carrying a *FM7c* chromosome in the Bloomington *Drosophila* Stock Center (**Table 4.S4**). Of 630 stocks labeled as carrying *FM7c*, we found 82 (13%) had the revertant *sn*$^{+}$ phenotype in *B*-eyed males, consistent with loss of the female-sterile *sn*^{X2} allele on *FM7c* chromosomes by DCO with a balanced homolog inside the *In(1)dl-49* inversion while maintained in stock. Of these 82 stocks, only 16 (20%) had any prior evidence of *sn*^{X2} reversion in their genotype or description, underscoring how commonly the *sn*^{X2} reversion may occur without notice. The genotypes of these *sn*$^{+}$ stocks have now been updated in the Bloomington *Drosophila* Stock Center database.

Since at least one of the *FM7a* stocks we sequenced (*FM7a-23229*) was in reality a *FM7c* stock mislabeled as a *FM7a* stock, the lack of *sn*^{X2} on *FM7* chromosomes could simply reflect that these chromosomes are actually *FM7a*'s that are mislabeled as *FM7c*'s, rather than true loss of *sn*^{X2} by a DCO inside *In(1)dl-49* on *FM7c*. To resolve these alternatives, we took advantage of the fact that all *bona fide FM7c*'s are expected to carry the same allele at the

garnet locus (g^4), whereas all *FM7a*'s should lack this marker. Within the mutant *g* gene on all *FM7c* (and *FM7a-23229*) chromosomes (**Figure 4.5B**), we found a diagnostic 24-bp deletion that spans an intron-exon junction and results in a frame-shift in the RB and RD transcripts (FBtr0331709 and FBtr0073842), and also ablates the ATG start codon of the RF transcript (FBtr0331710). We tested 76 of the 82 revertant sn^+ stocks labeled as *FM7c* in Bloomington for the presence or absence of this putative g^4-causing deletion by PCR and Sanger sequencing. We found that 71/76 (93%) of the sn^+ stocks screened by PCR and sequencing carried the g^4 allele present on all *FM7c* chromosomes (**Table 4.S4**), indicating that the majority of these are *bona fide FM7c*'s and thus are truly revertants. Because *g* lies outside the *In(1)dl-49* inversion and *sn* resides inside it, it is highly unlikely that one DCO event could have replaced both sn^{X2} and g^4 in any of the five putative *FM7c* sn^+ stocks that lack the g^4 deletion. We therefore conclude that these five stocks have been mislabeled as *FM7c* when, in fact, they are actually *FM7a's*. Thus, the vast majority of sn^+ stocks labeled as *FM7c* in the Bloomington *Drosophila* Stock Center are indeed *FM7c*'s, but mislabeling of *FM7* subtypes (a versus c) occurs in about of 7% of stocks. Overall, these results support the conclusion that the DCOs within the *In(1)dl-49* interval occur at an appreciable frequency, endangering mutations in homologous chromosomes kept in stock over balancer chromosomes, and leading to sequence diversity among *FM7c* balancers in circulation today.

Origin and reversion of the B^1 allele

X-chromosome balancers including *FM7* carry the B^1 allele, a dominant mutation affecting eye morphology, discovered more than 100 years ago (Tice 1914). B^1 is an unusual allele that

reverts to wild type at a high frequency in females (May 1917; Zeleny 1921) through either inter-chromosomal or intra-chromosomal unequal exchange (Sturtevant and Morgan 1923; Sturtevant 1925; Peterson and Laughnan 1963; Gabay and Laughnan 1973). B^1 is known to revert on *FM7* (Ashburner *et al.* 2005) and previous work suggests that B^1 reversion rates may be higher in inverted *X*-chromosomes (Sturtevant and Beadle 1936; Gabay and Laughnan 1973). B^1 has been shown to be associated with a tandem duplication of a large segment containing cytological bands 16A1–7, and B^1 revertants lack this duplicated segment (Muller *et al.* 1936; Bridges 1936). Muller (1936) argued that B^1 arose by unequal exchange between two sister chromatids or homologous chromosomes, rather than a duplicative insertion event as suggested by Bridges (1936). Muller's model for the origin of B^1 was supported by the work of Tsubota *et al.* (1989) who used a *P*-element-induced revertant of B^1 to clone the putative breakpoint of the B^1 duplication. These authors found a *roo* transposable element located exactly at the breakpoint between the two duplicated segments, and proposed that the B^1 allele originated by unequal exchange between *roo* elements located at 16A1 and 16A7, respectively, on two different homologous chromosomes (Tsubota *et al.* 1989) **(Figure 4.4A)**. However, the exact nature of the B^1 rearrangement remains to be clarified, since the 16A7 breakpoint of B^1 identified by Tsubota *et al.* (1989) contained a short segment of DNA not found in wild type flies. Moreover, neither the genomic extent nor gene content of the B^1 duplication has been investigated in the context of modern genomic data.

Figure 4.4. Genomic evidence for the role of unequal exchange at the *Bar* locus.
A) Model for the origin of the B^1 allele by unequal exchange (Muller 1936) between two different *roo* transposable elements (Tsubota *et al.* 1989). The distal and proximal segments of the B^1 duplication are shown in blue and orange, respectively, and *roo* elements are shown in green. **B)** Genome annotation and depth of coverage for *X*-chromosome balancers carrying B^1 (*FM7a-785*) and wild type revertants (*Binsc-107-614* and *Binscy-107-624*). Note the twofold increase in depth that starts downstream of *B-H2* and ends upstream of *CG12432* in the *FM7a-785* chromosome carrying B^1 that is lacking in *Binsc-107-614* and *Binscy-107-624* revertants. **C)** Model for the reversion of the B^1 allele to wild type by unequal exchange between the two duplicated regions. The model shows an inter-chromosomal exchange event (Sturtevant and Morgan 1923; Sturtevant 1925) however intra-chromosomal exchange events are also possible (Peterson and Laughnan 1963; Gabay and Laughnan 1973). **D)** Schematic of sequence variants in B^1 chromosomes (*FM7a-785*) and wild type revertants (*Binsc-107-614* and *Binscy-107-624*). Sequences from the distal and proximal duplicated regions in B^1 chromosomes map to the same coordinates in the reference genome, resulting in apparent "heterozygosity". The two revertant chromosomes are characterized by different haplotypes of homozygous SNPs. Sequences shared by both revertants at their 5' and 3' ends can be used to define the boundaries of unequal exchange events and partially phase the distal and proximal haplotypes, respectively. Diagnostic SNPs from fragments that span the junctions of putative unequal exchange events can then be used to phase haplotypes on both sides of both exchange junctions in B^1 chromosomes (dotted arcs), which together with the sequence of the revertants, can be used assign the location of each exchange event to the appropriate revertant stock.

We identified the precise genomic limits of the B^1 duplication on the basis of a contiguous 203,476-bp region between X:17,228,526–17,432,002 with two-fold higher sequencing depth in all eight *FM7* stocks sequenced **(Figure 4.4B).** Sequences flanking the duplicated interval correspond exactly to the B^1 breakpoints identified by Tsubota *et al.* (1989). We found that previous uncertainty in the wild type configuration of the 16A7 B^1 breakpoint region reported by Tsubota *et al.* (1989) is due to inclusion of phage DNA in their sequence. The B^1 duplicated interval contains the *BarH1 (B-H1)* homeodomain gene that has been shown to be involved in the *Bar* eye phenotype (Kojima *et al.* 1991; Higashijima, Kojima, *et al.* 1992), plus seven other predicted protein-coding genes and a putative ncRNA gene (*CR43491*) that likely corresponds to the *T1*/*T2* or *BarA* transcript identified previously (Higashijima, Kojima, *et al.* 1992; Norris *et al.* 1992). As predicted by Higashijima and colleagues (1992), the B^1 breakpoint lies in an intergenic region upstream of *B-H1* and downstream of *BarH2 (B-H2)* **(Figure 4.4B)**, a related homeodomain gene that is also involved in eye morphogenesis (Higashijima, Kojima, *et al.* 1992). Thus, the B^1 duplication on *FM7* chromosomes carries an intact *B-H2–B-H1* locus, plus an additional copy of *B-H1* fused downstream of *CG12432* **(Figure 4.4B)**.

Tsubota *et al.* (1989) proposed that unequal exchange between two *roo* insertions at different positions on homologous chromosomes caused the B^1 duplication **(Figure 4.4A)**. To provide an independent assessment of this hypothesis, we extracted split/discordant reads and their mate-pairs in the ±1.5-kb intervals at either end of the duplicated segment, then performed *de novo* assembly as above for the *FM7* inversions and recovered two contigs spanning the 16A1 and 16A7 sides of the B^1 breakpoint. Both of these contigs contained *roo* sequences that began after the exact point at which alignment to the reference genome ended.

We used long-range PCR to amplify an approximately 8-kb fragment spanning the breakpoint from the end of 16A7 to the beginning of 16A1 in *FM7*-carrying but not in wild type stocks **(Table 4.S2)**. Sanger sequencing of the 5' and 3' ends of this breakpoint-spanning fragment revealed a *roo* element in the expected location and orientation **(Table 4.S2)**. Together, these results confirm the work of Tsubota *et al.* (1989), showing that the B^1 breakpoint contains a *roo* element in the 5' to 3' orientation located precisely at the junction between the duplicated segments.

Our genomic data also allows us to investigate sequence variation directly within the B^1 duplication, which provides new insights into the origin and reversion of the B^1 allele. Analysis of sequence variation in the region duplicated in B^1 revealed a large number of "heterozygous" SNPs in each hemizygous or homozygous *FM7* stock (min: 1242, max: 1250). "Heterozygous" SNPs in hemizygous or homozygous stocks can arise from calling variants in duplicated regions that are mapped to the same single-copy interval of the reference genome (Remnant *et al.* 2013). This apparent heterozygosity in the B^1 interval implies substantial sequence divergence existed between the two ancestral haplotypes that underwent unequal exchange to form the original B^1 allele, providing independent support for the origin of B^1 by unequal exchange between two homologous chromosomes rather than two sister chromatids (Tsubota *et al.* 1989). Additionally, the "heterozygous" SNP profile was nearly identical among all eight *FM7* stocks, supporting a single origin for the B^1 allele, consistent with the historical record (Tice 1914).

These "heterozygous" variants also give us a rich set of molecular markers that, together with depth of coverage in the *B* region, can be used to investigate the mechanism of B^1

reversion. If reversion of the B^1 allele is due to either inter-chromosomal or intra- chromosomal unequal exchange (Sturtevant and Morgan 1923; Sturtevant 1925; Peterson and Laughnan 1963; Gabay and Laughnan 1973), we expect a twofold reduction in the depth of coverage to be associated with loss of "heterozygosity" across the entire B^1 duplicated region in revertant chromosomes **(Figure 4.4C)**. To test this hypothesis, we identified two *X*-chromosome balancer stocks carrying reversions of B^1 (*Binsc-107-614* and *Binscy-107-624*) and sequenced their genomes. As expected, depth of coverage in both B^1 revertants was at wild type levels across the B^1 interval X:17,228,526–17,432,002 **(Figure 4.4B)**. Additionally, no high quality heterozygous SNPs or split/discordant reads were observed in the B^1 interval in either revertant. These results demonstrate that the duplicated segment is strictly associated with the *B* phenotype, as shown previously at the cytological level (Muller *et al.* 1936; Bridges 1936).

Comparison of the single-copy haplotypes in the two revertants revealed likely sites of unequal exchange **(Figure 4.4D)**. *Binsc-107-614* and *Binscy-107-624* haplotypes in the B^1 interval contained the same SNPs from X:17,228,526–17,283,005 and again from X:17,388,394–17,432,002, but differed from each other in the central X:17,283,375–17,388,155 interval. This result indicated that unequal exchange must have occurred in a 370-bp window between X:17,283,005 and X:17,283,375 in one stock, and in a 239-bp window between positions X:17,388,155 and X:17,388,394 in the other stock. This result also implied that the haplotype from the beginning of B^1 to 17,283,005 is from the 5' duplicated segment, and the haplotype from X:17,388,394 to the end of B^1 is from the 3' duplicated segment. Because the SNPs defining the sites of unequal exchange were close to one another, we were able to phase haplotypes from the distal and proximal duplicates using read-pair data in non-recombinant

FM7 "heterozygotes". Knowing the phase and location of both non-recombinant haplotypes in the B^1 duplication allowed us to infer that unequal exchange occurred between X:17,283,005 and X:17,283,375 in *Binsc-107-614*, and independently between X:17,388,155 and X:17,388,394 in *Binscy-107-624*. Together, these data provide definitive genomic evidence that B^1 reversion is associated with unequal exchange among duplicated segments directly within the B^1 interval.

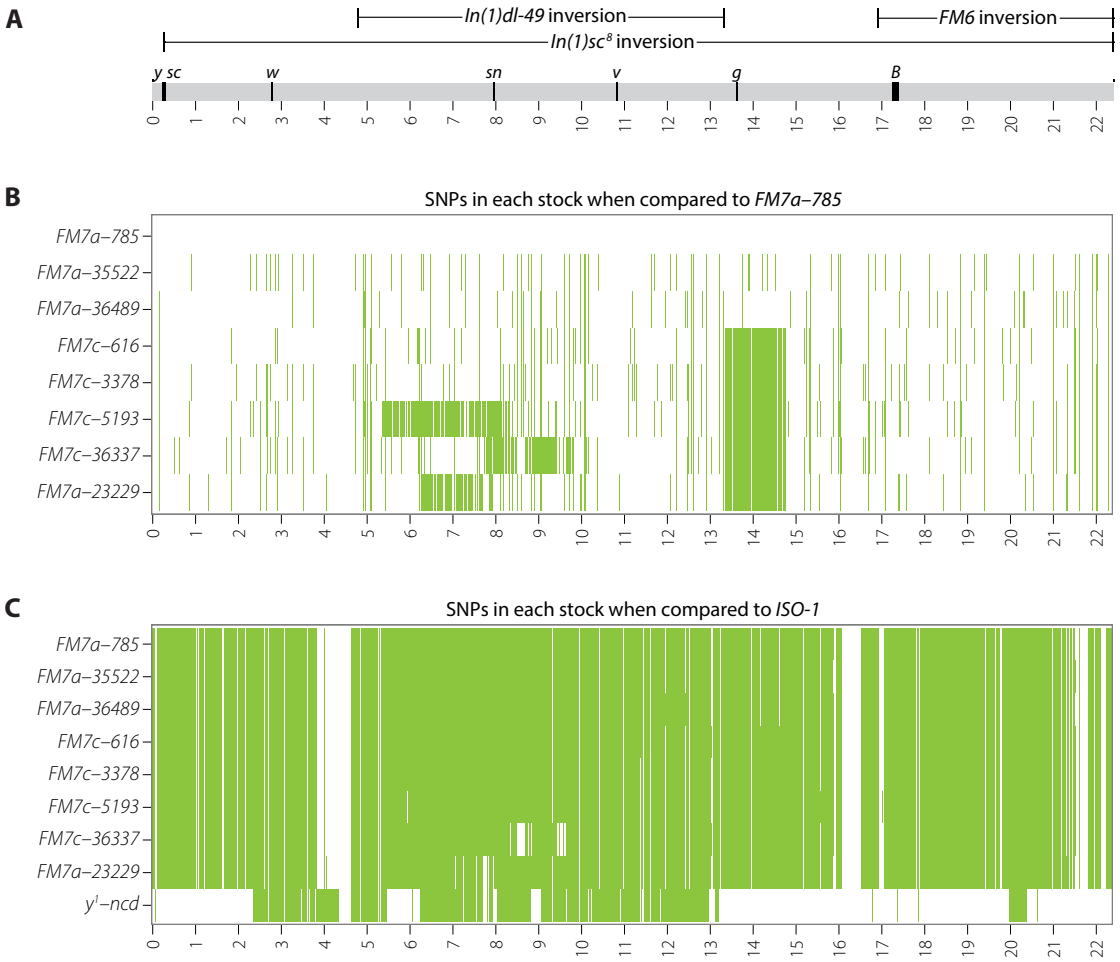

Figure 4.5. Polymorphisms are evident both within *FM7* stocks and when comparing *FM7* stocks to the *ISO-1* reference genome.

A) Schematic of the wild type *X*-chromosome, showing the locations of inversions (oriented with respect to the reference genome, not *FM7*), marker genes, and Release 5 genome coordinates (in Mb). **B)** Heatmap of SNPs detected in the eight *FM7* stocks used in this study when using *FM7a-785* as a genome reference. Increased SNP density covering the *sn* region in stocks *FM7c-5193*, *FM7c-36337*, and *FM7a-23229* indicates the region replaced by a DCO event. Note the increased SNP density between 13,369,185 Mb and 14,812,237 Mb present in all *FM7c* stocks (and the mislabeled *FM7a-23229*) that defines the haplotype containing g^4 present on *FM7c*. **C)** Heatmap of SNPs detected among all *FM7* stocks and y^1-ncd^D compared to the *ISO-1* reference genome. Sequence diversity among the eight *FM7* stocks is apparent at this scale as differing levels of SNP density surrounding the *sn* locus. Blocks of similarity between all *FM7*'s and *ISO-1* suggest a common ancestor for these regions. Blocks of diminished SNP density (in white) shared between *FM7a-23229* and y^1-ncd^D are apparent in Figure 4.3D as an apparent absence of SNPs.

DISCUSSION

Our work provides detailed insight into the structure and diversity of the most commonly used *X*-chromosome balancer in *D. melanogaster, FM7*. We mapped and characterized breakpoints of the three large inversions present on *FM7* and identified major sequence differences in the vicinity of *g* between the two subtypes of *FM7* (*FM7a* and *FM7c*). Surprisingly, we identified megabase-scale tracts of unique sequence in different *FM7c*'s that likely arose from DCOs removing the female-sterile sn^{X2} allele within the *In(1)dl-49* inversion. We further show that loss of the sn^{X2} allele affects a substantial proportion of *FM7c* chromosomes at the Bloomington *Drosophila* Stock Center. Finally, we clarified the molecular organization of the B^1 allele carried on *FM7*, and provide definitive genomic evidence for origin and reversion of B^1 by unequal exchange. In contrast to the prevailing notion of balancers as clonal non-recombining chromosomes, our results provide evidence that rare recombination events have led to large-scale sequence differences among balancers currently used by *Drosophila* researchers.

Our work has a number of implications for the design and interpretation of experiments that use *X*-chromosome balancers in *D. melanogaster*. Knowing the precise molecular location of inversion breakpoints on *FM7* reveals regions of the *X*-chromosome that are more or less susceptible to exchange events. Furthermore, the fact that many *FM7c's* carry megabase-scale tracts of unique variation, and that a substantial proportion of *FM7* chromosomes are mislabeled, should motivate researchers to characterize which *FM7* subtype their stocks actually carry. Characterization of an *FM7* subtype may be carried out by PCR and Sanger sequencing of *g*, or by simply crossing the *FM7* chromosome in question to a stock carrying a loss-of-function allele of *g* and scoring the eye phenotype of heterozygous females. The

genomic scale of sequence differences between *FM7* subtypes is sufficiently large such that, without controlling properly for *FM7* subtype, effects attributed to balanced chromosomes in heterozygotes could arise from differences in the *FM7* genetic background. Our finding that reversion of the female sterile sn^{X2} allele by DCO in the *In(1)dl-49* interval is common suggests researchers should be cautious when using *FM7c* for long-term stock maintenance of mutations in this region. We advise that replicate copies of such stocks be maintained and periodically checked for sn^{+}, B^{1} males that could indicate breakdown of the balanced chromosome by a DCO event. Alternatively, such mutations could be maintained using attached-*X* stocks instead of balancer chromosomes (Ashburner *et al.* 2005). Unavoidable DCOs within the *In(1)dl-49* region that remove the sn^{X2} allele on *FM7c* may motivate synthesis of a new generation of female-sterile *X*-chromosome balancers, perhaps by introducing additional inversions inside the *In(1)dl-49* interval on *FM7c*. While our work documents that DCOs do occur within *FM7* on a historical timescale, we emphasize that such events are not common enough to impair the utility of *FM7* as balancer chromosomes in routine genetic analysis.

Our study also demonstrates the value of sequencing classical stocks of *D. melanogaster* to uncover the molecular basis of uncharacterized mutations and better understand the genetic background of mutant stocks. Despite the availability of a nearly complete, richly annotated genome sequence, over 1,000 existing classical mutations in *D. melanogaster* have not been associated with gene models or linked to genomic sequences. Here we identified the causal molecular basis of three classical inversions (*In(1)sc*^{8}, *In(1)dl-49*, and *In(1)FM6*), mapped the locations of the B^{1} duplication and *Df(1)JA27* deletion (356 kbp deletion from X:19,043,642-19,399,862), proposed candidates for the lesions that causes the g^{4} and sn^{X2} alleles, and

identified an uncharacterized deletion in the *Df(1)JA27* chromosome (X:22,164,372–heterochromatin). Further analysis of our genomic data should lead to insights about the molecular basis of additional mutations carried by these strains, including the sites of transgene insertions that mark some *FM7* balancer chromosomes (Casso *et al.* 2000; Le *et al.* 2006; Abreu-Blanco *et al.* 2011; Lattao *et al.* 2011; Pina and Pignoni 2012). Sequencing classical lab stocks can also lead to the identification of mislabelled strains (e.g. that *FM7a-23229* is in fact a *FM7c* chromosome) and unreported genotypes (e.g. sn^+ in *FM7a-23229*), and thereby reduce sources of unwanted experimental variation. Systematic sequencing of stocks in the Bloomington *Drosophila* Stock Center could therefore improve the precision of *Drosophila* genetics and, in conjunction with extensive phenotypic information in FlyBase, provide a powerful model to develop workflows to identify rare disease variants in humans.

Future work on second and third chromosome balancers is needed to generalize the findings reported here, although such studies would be more challenging because genomic analysis would need to be performed in heterozygotes. Sequencing larger samples of *FM7* chromosomes could also provide deeper insight into the mechanisms of exchange in highly inverted chromosomes (Sturtevant and Beadle 1936; Novitski and Braver 1954). We identified 71 *FM7c* sn^+ stocks here that are all *bona fide FM7c*'s likely to have undergone DCO with a balanced stock, which should provide a rich sample to study how DCOs are distributed relative to the locations of breakpoints in inversion heterozygotes. Likewise, sequencing of additional B^1 revertants can now be used as a model system to study unequal exchange at the molecular level, especially given our finding that the two duplicated regions in B^1 differ by many variants. By generating a large sample of B^1 revertants in heterozygotes that differ from *FM7* outside the

B^{1} interval, it will be possible to precisely measure the relative contribution of inter- and intra-chromosomal unequal exchange events, and to understand how unequal exchange events are distributed across the duplicated region. More in-depth analysis of sequence variation among *FM7* chromosomes could also lead to insights about gene conversion between balancers and balanced chromosomes (Cooper *et al.* 2008; Blumenstiel *et al.* 2009), as well as whether the predicted accumulation of deleterious mutations on balancers is observed at the molecular level (Araye and Sawamura 2013). Finally, sequencing a larger panel of *FM7* chromosomes and more primitive *X*-chromosome balancers could shed light on the ancestral state of *FM7* at the time of its origin, and how inversions were integrated within inversions to create the founders of the *FM* series (Lewis and Mislove 1953).

METHODS

Fly stocks used

X-chromosome balancer stocks used in this experiment were obtained from either the Bloomington *Drosophila* Stock Center or from the *Drosophila* Genetic Resource Center (see **Table 4.1** for stock identifiers). The y^{1}-ncd^{D} stock that was used as a parental *X*-chromosome in the construction of the *ISO-1* reference genome strain (Brizuela *et al.* 1994) was obtained from Jim Kennison. Full genotypes of stocks as labeled at the outset of this project are listed in **Table 4.1** and are referred to in the text by their abbreviated genotype followed by their stock number (where available). All flies were kept on standard cornmeal-molasses and maintained at 25°C.

DNA preparation and whole-genome sequencing

DNA was prepared from 10 adult hemizygous *FM7*-carrying *Bar* eyed males for stocks *FM7a-785, FM7a-23229, FM7a-35522, FM7a-36489, FM7c-616, FM7c-3378*, *Binsc-107-614*, and *Binscy-107-624.* Because of the poor viability of *FM7*-carrying hemizygous males in *FM7c-5193* and *FM7c-36337*, DNA was prepared from a mixture of 10 adult hemizygous *FM7* male and homozygous *FM7* females for these two samples. Ten heterozygous adult females were used for the *FM7c-616*, *FM7c-5193*, *FM7c-36337*, and *FM7a-23229* heterozygous samples. Ten adult hemizygous yellow males were used for y^{1}-ncd^{D} sample. All DNA samples were extracted using the Qiagen DNeasy Blood & Tissue Kit (catalog number 69504). Flies were starved for 4 hr before freezing at –80°C for at least 1 hr prior to DNA extraction. 600- to 800-bp fragments of

DNA were selected after shearing and libraries were prepared using a Nextera DNA Sample Prep Kit (catalog number FC-121-1031) from Illumina following the manufacturer's directions. Hemizygous males and homozygous females from stocks *FM7a-785, FM7a-23229, FM7a-35522, FM7a-36489, FM7c-616, FM7c-5193, FM7c-3378*, and *FM7c-36337* were sequenced as 100-bp paired-end samples on an Illumina HiSeq 2500. Heterozygous females from stocks *FM7c-616*, *FM7c-5193*, *FM7a-23229*, and *FM7c-36337* were sequenced as 150-bp paired-end samples on an Illumina HiSeq 2500. Hemizygous males from stocks y^{1}-ncd^{D}, *Binsc-107-614*, and *Binscy-107-624* were sequenced as 150-bp paired-end samples on an Illumina NextSeq.

Genome alignment and SNP calling

Alignment to the UCSC Genome Bioinformatics dm3 version of the Release 5 *D. melanogaster* reference genome sequence was performed using bwa (version 0.7.7-r441) (Li and Durbin 2009). Variants were called using SAMtools and BCFtools (version 0.1.19-44428cd) (Li, Handsaker, *et al.* 2009). Indels and low quality SNPs (qual<200) were filtered out of single-sample Variant Call Format (VCF) files. Unique SNPs were identified by additionally filtering out heterozygous SNPs from single sample VCF files and merging samples to identify SNPs present in only one sample using VCFtools (version 0.1.12b) and visualized as heatmaps using R (version 3.1.3).

Identification, assembly, and validation of rearrangement breakpoints

Rearrangement breakpoints were identified using three strategies. For the *In(1)sc*^{8}, *In(1)dl-49*, *In(1)FM6*, and *B* breakpoints, split/discordant *X*-chromosome read pairs were identified using samblaster (Faust and Hall 2014) and visualized using the UCSC genome browser (Rosenbloom *et al.* 2015). Clusters of split/discordant reads corresponding to putative breakpoints were identified in the approximate locations where rearrangements were expected based on classical work. Original fastq sequences of split/discordant reads and their mate-pairs from ±1.5kb around putative breakpoints from the same rearrangement were then merged from all eight *FM7* stocks into a single per-rearrangement file. SOAPdenovo2 (version 2.04) was then used to perform *de novo* assemblies for both breakpoints of each rearrangement at the same time using a kmer size of 41 or 51 for the *In(1)sc*^{8}, *In(1)dl-49*, and *In(1)FM6* inversions and a kmer size of 73 for the *B*^{1} duplication breakpoint (Li, Yu, *et al.* 2009). To identify the *In(1)dl-49* inversion, we also ran Breakdancer (version 1.4.4) (Chen *et al.* 2009) using default options with the exception that only the *X*-chromosome was analyzed (-o X) and any event with fewer than 10 supporting reads was ignored (-r 10). For the *B*^{1} duplication, we also identified an interval with the expected two-fold higher read-depth coverage in the location where the duplication was expected to be found (**Figure 4.4B**) (Lindsley and Zimm 1992).

Contigs spanning candidate breakpoints were used to design PCR primers on either side of each candidate breakpoint region using Primer3 (Rozen and Skaletsky 2000). PCR was performed using Phusion DNA polymerase (NEB, catalog #M0530L) using a 62^{o}C annealing temperature and 45 second extension time. PCR products were purified from a gel using a QIAquick PCR Purification Kit (Qiagen, catalog #28106) and Sanger sequenced. Long-range PCR

of the junction between the two duplicated B^1 regions was performed using a Qiagen LongRange PCR Kit (catalog #206402) using 250 ng of genomic DNA, 59^{o}C annealing temperature, and nine minute extension time.

Screen for *sn* reversion in *FM7* stocks at the Bloomington *Drosophila* Stock Center.

We visually screened 630 stocks from the Bloomington *Drosophila* Stock Center that were labeled as carrying *FM7c* for the presence or absence of the *sn* phenotype in *B* males. Eighty-two stocks yielded *B*, sn^+ males and were classified as putative *FM7c* revertants. To determine if putative *FM7c* revertants were in fact mislabeled *FM7a*'s, 76 of these putative *FM7c* revertants were screened for the presence of a diagnostic 24-bp deletion associated with the g^4 allele that is present on all *bona fide FM7c*'s. Primers used to amplify a fragment spanning the g^4 deletion were garnet_F2 (5'-ACACCCGCATCGTATTGATT-3') and garnet_R2 (5'-CCAGTTGGCTGAAACTGAAA-3'). DNA was prepared by placing single *B*, sn^+ males in a standard fly squish buffer (50 µL of 1M Tris pH 8.0, 0.5M EDTA, 5M NaCl) plus 1 µL of 10 mg/ml Proteinase K. Extracts were then placed in a thermocycler at 37^{o}C for 30 minutes, 95^{o}C for 2 minutes followed by a 4^{o}C hold. PCR was performed using 4 µL of fly squish product in a total volume of 50 µL. Fragments were amplified using Phusion polymerase (NEB catalog number M0530L) reaction conditions were per manufactures instructions except for a 64^{o}C annealing temperature, and 45 second extension time. PCR amplicons were Sanger sequenced and resulting sequences were aligned to the reference genome to determine the presence or absence of the 24-bp deletion.

SUPPLEMENTAL TABLES

Table 4.S1: Stocks sequenced in this study.
Stocks sequenced, along with depth of coverage for each chromosome arm.

			Depth of Coverage					
Stock Name	**ID**	**Full Genotype**	**X**	**2L**	**2R**	**3L**	**3R**	**4**
Binscy-hemi	107-624	In(1)sc[S1L]sc[8R]+dl-49 y[c4] sc[S1L] sc[8R] B[1]	46	96	97	96	99	92
y1-ncd-hemi	na	y[1]; ncd[D]/TM6, Tb; pol	60	123	124	122	126	117
FM7c-36337-hemi	36337	FM7c, P{w[+mC]=2xTb[1]-RFP}FM7c, sn[+]/oc[otd-XC86]	96	125	126	126	129	118
FM7c-5193-hemi	5193	Df(1)JA27/FM7c, P{w[+mC]=GAL4-Kr.C}DC1, P{w[+mC]=UAS-GFP.S65T}DC5, sn[+]	73	128	129	128	133	118
FM7c-36337-het	36337	FM7c, P{w[+mC]=2xTb[1]-RFP}FM7c, sn[+]/oc[otd-XC86]	54	54	55	54	56	50
Binsc-hemi	107-614	In(1)sc[S1L]sc[8R]+dl-49 sc[S1L] sc[8R] B[1]	53	110	111	109	113	101
FM7a-23229-het	23229	w[*] baz[4] P{w[+mW.hs]=FRT(w[hs])}9-2/FM7a, P{Dfd-GMR-nvYFP}1	51	51	53	52	54	47
FM7c-3378-hemi	3378	y[1] arm[1]/FM7c	73	147	149	147	152	141
FM7c-616-het	616	y[1] arm[4] w[*]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	57	58	59	59	60	54
FM7a-785-hemi	785	FM7a	65	132	132	131	136	123
FM7a-35522-hemi	35522	FM7a, P{w[+mC]=sChFP}1	64	128	130	128	132	122
FM7a-23229-hemi	23229	w[*] baz[4] P{w[+mW.hs]=FRT(w[hs])}9-2/FM7a, P{Dfd-GMR-nvYFP}1	51	101	102	101	105	96
FM7c-5193-het	5193	Df(1)JA27/FM7c, P{w[+mC]=GAL4-Kr.C}DC1, P{w[+mC]=UAS-GFP.S65T}DC5, sn[+]	54	55	56	56	57	49
FM7c-616-hemi	616	y[1] arm[4] w[*]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	57	113	114	114	117	108
FM7a-36489-hemi	36489	FM7a, P{w[+mC]=Tb[1]}FM7-A	66	134	136	134	139	124

Table 4.S2: Primers used to verify FM7 inversion breakpoints.

Inversion Fragment	**sc8 A/C**	**sc8 B/D**
Forward Primer	sc8-Distal-F3	sc8-Proximal-F2
Stowers ID	482	486
Forward Sequence	5'-AACAGACTCTGCAAAAATGTTGA-3'	5'-TTTGCGGAATTCATAATCCA-3'
Primer cordinates (dm3)	multi-hit (1.688 satellite)	X:276,895-276,877
Reverse Primer	sc8-Distal-R3	sc8-Proximal-R2
Stowers ID	483	487
Reverse Sequence	5'-TCAGACCACCAAGACACCAC-3'	5'-TCAAACTGTGTTCAAAAATGGAA-3'
Primer cordinates (dm3)	X:276,152-276,171	multi-hit (1.688 satellite)
Annealing Temp (^{o}C)	62	62
Extension Time (sec)	45	45
Polymerase	Phusion	Phusion
Stock		
ISO-1-2057	not amplified, not expected	not amplified, not expected
FM7a-785	amplified, expected	amplified, expected
FM7a-23229	amplified, expected	amplified, expected
FM7c-616	amplified, expected	amplified, expected
FM7a-35522	amplified, expected	amplified, expected
FM7c-3378	amplified, expected	amplified, expected
FM7c-36337	amplified, expected	amplified, expected
FM7a-36489	amplified, expected	amplified, expected
FM7c-5193	amplified, expected	amplified, expected
Inversion Fragment	**dl-49 A/C**	**dl-49 B/D**
Forward Primer	dl-49-Distal-F3	dl-49-Proximal-F3
Stowers ID	468	474
Forward Sequence	5'-GCATAAAGATCTGCGTCCAA-3'	5'-CGGCCAGAGATAAAATGAGG-3'
Primer cordinates (dm3)	X:13,321,576-13,321,557	X:13,320,799-13,320,818
Reverse Primer	dl-49-Distal-R3	dl-49-Proximal-R3
Stowers ID	471	477
Reverse Sequence	5'-TTGATTGCGATGGAAAATCC-3'	5'-GCGAAAAAGTTGTCCCTTGT-3'
Primer cordinates (dm3)	X:4,791,489-4,791,470	X:4,791,152-4,791,171
Annealing Temp (^{o}C)	62	62
Extension Time (sec)	45	45
Polymerase	Phusion	Phusion
Stock		
ISO-1-2057	not amplified, not expected	not amplified, not expected
FM7a-785	amplified, expected	amplified, expected
FM7a-23229	amplified, expected	amplified, expected
FM7c-616	amplified, expected	amplified, expected
FM7a-35522	amplified, expected	amplified, expected
FM7c-3378	amplified, expected	amplified, expected
FM7c-36337	amplified, expected	amplified, expected
FM7a-36489	amplified, expected	amplified, expected
FM7c-5193	amplified, expected	amplified, expected
Inversion Fragment	**FM6 A/C**	**FM6 B/D**
Forward Primer	FM6-Distal-F2	FM6-Proximal-F2
Stowers ID	498	492
Forward Sequence	5'-TCGAGGACAAGCGTCAATTA-3'	5'-AATTTGGAAGAATTATCATGTGC-3'
Primer cordinates (dm3)	multi-hit (rDNA)	multi-hit (rDNA)

Reverse Primer	FM6-Distal-R2	FM6-Proximal-R2
Stowers ID	499	493
Reverse Sequence	5'-TGTTTGCCTTGCAAATGTGT-3'	5'-AGCGAATAAGGCGACAAAAC-3'
Primer cordinates (dm3)	X:16,919,027-16,919,046	X:16,919,497-16,919,478
Annealing Temp (^{o}C)	62	62
Extension Time (sec)	45	45
Polymerase	Phusion	Phusion
Stock		
ISO-1-2057	not amplified, not expected	not amplified, not expected
FM7a-785	amplified, expected	amplified, expected
FM7a-23229	amplified, expected	amplified, expected
FM7c-616	amplified, expected	amplified, expected
FM7a-35522	amplified, expected	amplified, expected
FM7c-3378	amplified, expected	amplified, expected
FM7c-36337	amplified, expected	amplified, expected
FM7a-36489	amplified, expected	amplified, expected
FM7c-5193	amplified, expected	amplified, expected
Inversion Fragment	**Bar Duplication**	
Forward Primer	BarDuplication_F1	
Stowers ID		
Forward Sequence	5'-GAGCAGCACCAACTGCAAC-3'	
Primer cordinates (dm3)	X:17,431,706-17,431,724	
Reverse Primer	BarDuplication_R2	
Stowers ID		
Reverse Sequence	5'-CCCAGACAGCCAGAGGATG-3'	
Primer cordinates (dm3)	X:17,228,751-17,228,733	
Annealing Temp (^{o}C)	59	
Extension Time (sec)	9 min	
Polymerase	NEB Long-Range	
Stock		
ISO-1-2057	not amplified, not expected	
FM7a-785	amplified, expected	
FM7a-23229	N.A.	
FM7c-616	amplified, expected	
FM7a-35522	N.A.	
FM7c-3378	N.A.	
FM7c-36337	N.A.	
FM7a-36489	N.A.	
FM7c-5193	N.A.	

Table 4.S3: Inversion breakpoint sequences from the *X*-Chromosome balancer *FM7*

Sequence of all six *FM7* inversion breakpoints and sequence of the *Bar* duplication breakpoints based on Sanger sequencing results. A *roo* element sits between the duplicated *Bar* regions, 16A7-*roo*-16A1, so the 3’ end of one breakpoint and the 5’ end of the other breakpoint contains the 5' and 3' ends, respectively, of the *roo* sequence.

>In(1)sc[8]_AC_fragment
ACCAAGACACCACCACCAGACGCATAAAACCGCCTTGGCTTACGTCTTCAAATCACCTGGACGATTATGACAAATGCATCTTCT
AGGCATGGTTCGGTCCACAAACAGGGCCAACAATCAAAAACTGTATGGAGTTTAATCAACTGAATGAAGTATTCTAATCTAAAT
TTAAATACTTAAATTTATATGTACATAATGTTTAATGTTTTATTTTAGACTGCCTTTCGTCATAAAGACAACGTAGAACTTTGC
TTTTGATTTCGTAATCACTGAGCTCGTAATAAAATTTCCAATCAAACTGTGTTCAAAAATGGAAATTAAATTTTTTGGCCATAT
TTTGCAAATTTTGATGACCCCCCTCCTTACAAAAAATGCGAAAATTGATCCAAAAATTAATTTCCCTAAATCCTTCAAAAAGTA
ATAGGGATCGTAAGCACTGGTAATTAGCTGCTCAAAACAGTTATTCTTACATCCATGTGACCATTTTAGCCAAGTTATAACGAA
AGTTTCGTTTGTAAATATCAACATTTTTGCAGAAGTCTGTT

>In(1)sc[8]_BD_fragment
AATCAAACTGTGTTCAAAAATGGAAGTTAAATTTTTCGTTTTTCTGTTAGTGCATGAACATTTGAAATTTATTTGGTATTTTTT
TTTACATTTTTATGCTTAGTATCAGTCTTGTACTTCCCCCATCGTCAACTTTCTTTCCAACAGGAATCCACTGATTAACCCCCT
ATCTTCCTATTAAGCTTATGAAACTGCTTCTGTTCCGAATTCAGATTCAAATTACTAATCGGCGCAAGCTGCAAAAATCGTTGC
ACATGTCGGGGCGTTTTCGTCTACTTGATTTTGATTCAGTTCGCCTGATCCCGCTGATATATAGTTCCTGCTCATTCGGATACG
ACTCTTATGTCACTGGGTTAAGGGTTCGTATTTATTTTGCGTGCGTTCAATGCTGCAGATGCATTTTGATTTTGTCTGTCGCAC
CCATTCCTTTACACTTAAAAAATAAATTGTCAAGTGTAAAGATCTTTAAAACATACCTACAAATTATGAATTATGGATTATGAA
TTCCGCAAA

>In(1)dl-49_AC_fragment
CGAAAAAGTTGTCCCTTGTTAGGAGGAAATCATCAAAATGACACAAACAATTATTTAACTTTCGCTAGGCCGTGCATGAGACCG
GTGACGTCACATTACACACACGCACACACACATGACAGATGACTTTTACCCCAACCCGCTTGCCACTGGCCAGTAAGTCCGCTG
TCAGCGCGGCAAGATAATCAGCTCCGAACTCGACGTGGACTTTCCTCATTTTATCTCTGGCCG

>In(1)dl-49_BD_fragment
TTGATTGCGATGGAAAATCCGACATATTCGACTTTATTGAGCGGGAAATGCCAGCGTATCGGTAGTTATATATTTGTCAAAGAA
ATTGTTTATAGTTCCCGTCATTGGGCTCAAAAATGTGTAAATGGCTTATTGTTCCAACTGCAGAGGCTTCAATTTAAATGGGCT
TGTAACTTTGGCTTGAGCAAAAATGTGCGTGTTATGTGTTATGTGTGTCTCAAAACAGAGGGTTAATCTGAAAACACTCAACTT
ATTATTGAATTAAAGTGTACTTAAATACAGTATAAATCTTTATAATTATATCTTTATATCTTTATCAACTTGCATCCCATTTCT
TGCAGTGCTCGCAGGGCGCTAATAGTCCAAGCATCCGAAGGAAACCCCCAAACCCCGCCGGAATAATTCAAGGTGGCCGTAATG
GCTGAATGCCCAACTTGGTAGACACCATCGTCGTCGTGTCCTTTCCAACCTGCCTATCCTCCTGCTCCGAATTCCCGGACCTCG
GTGACCTCGTCCTTGGACGCAGATCTTTATGCA

>In(1)FM6_AC_fragment
TGTGTGTTTGCCTTGCAAATGTGTGCGTGGGCGGGCATGTGTGTGTGAACATGTGGGTTTTCTTATTTGCCAACCGCTTGTCTC
TCGACTTGCCCCCAAAATGGCTCAAAGCAATTTTCAAATTGTTTTCACAATCACAACGAAATCGATTTGCAAATATACACAACG
CACAGCCCTCTTGGCAACCTACAACAAGTGCCACACCCACTTGGGCTCACACAATGTGGGTATGAGTATGTTCTGCATTTGATG
TCCTTTTTCTAGCTACAATTATGTCCGGAGTGCTATATAAAAATGGCCGTATTCGAATGGATTTATTTTTATAAATATATTTAA
AATTTTTACCCAAAGGCAAAATATTGAATTACATTCAATAATATAAAAAAATGGAATTATATAAGTTAATAATTACAGTTATCG
ATTTGATTTTGAAGATCGCAAGCGACCGTTTATTGCAATTTATCATTTGAAACTAAATCTAGCGTACAAAATGTTTCCCTAAGT
CCCTAGCAATCAAGTGAAGTCGTCGGCAGCGGCGCAGCAGGCGTCGGCCGCGGCGCAGCGCAGAAGTGTCGATGTCGCGCTTAA
CCGTTCGTTGGCGTTGATGGCAGCGGAGACTATGTGGAACCACAAGATGTTAGAGAATCAATTGCAGGGCAATAACTCCTCCCC
TCTTAATTGACGCTTG

>In(1)FM6_BD_fragment
ATTTGGAAGAATTATCATGTGCGCTCGGTTTTATGTTATATATTACCAGAGAGTTATATGAAAAGAGATAAATTTTAAATTTAT
CATCAAGATGCAAATGATTTAACTTATATTTGGTTAAACAAAAATTGTACAAGTGTGGATACAAAATTTATGTATGTTGGAAAT
AAAATGATATTTTAGAATGAAATATATGTATATATAAAGACAAAATTATAGAAAATATATTACAATAATTGTATGATCTTCTTG
TTATATTGGTAAAACAAGTAGAATTTAAAAATGGAAATACGTAGTTAAAATGCCCATGTGTTTATGGCCCTGGCCAAGGATACG
CAAGATTTGCTCCTCGGATTTCCGTCGATTTATTTCCGCTGATAAAGTGCGATTTGAAAAGCGATTTACACAAACAAATCTCTC
TTTTGTCCGCCCAGCAACTGATTGCTTTTATGTCCATGTCAATTGTTTTGTCGCCTTATTC

>Bar_16A7_roo #Sequence from euchromatic sequence into 5' of the roo element

CAGGCTGGCTGGCCCCTGCAAAAAAGTGTTGCATGTTCTTCTATCATATTTGCAATTTATGTGTCGCCAATTCACTCAGACTCC

CCCTGCCGCGCCCGCAGCTCCTGCTGCCCCCCCGAGAGTCGTCCTCCCGATTGTGCAGCGGATTCCCCCGTCCGTTTATGGCAT
GTTGATGCCCGAGTGCGAACAAGTAATTAGAAAACATATTTCGAAATTGTTTCGCCATCCGTTTGTGCAACTTTGCGGCTCGCA
GCCGTGCTGAACTCAAGGGGTGTTCACACATGAACACGAATATATTTAAAGACTTACAATTTTGGGCTCCGTTCATATCTTATG
TAAATGAATCGAGAGCGATAAATTATATTTAGGATTTTGTTATCTAAGGCGACATGGGTGCATTGCTCAAAAACATGTAATTTA
AGTGCACACTACATGAGTCAGTCACTTGAGATCGTTCCCCGCCTCCTAAAATAGTCCCTTAGTGGGAGACCACAGATAAGGTCC
TCGCCGCTCAAGATAGGCAGATGTGCCCGAGCGTGGGACCTCGATAAGGCGGGGACTATTTACTTAGGCCTCTGCGTAGGCCAT
TTACTTTAAGATGCGATTCTCATGTCACCTATTTAAACCGAAGATATTTCCAAATAAAATCAGTTTCTTACAAAAACTCAACGA
GTAAAGTCTTCTTATTTGGGATTTTACATTTGGTCAATCGAGCCTTTAATCGACTCTGCAGTTTCCCCCTACCAAAGGTAAGGA
ACTCAGAGAAAGGCCAGCTCCTTTAAGCATCTTACAGCTAAAGGTAGCAAAAATAAGTGACTCTTGTTTCCCCC

>Bar_16A1_roo #Sequence from 3' end of the roo element into euchromatic sequence
AAAAAAGCAAAATGTTTAAAATAAGTTAATTGAGTACAAATTGTTGAATTAAAAATAAATATAAACCATAATTGTAATCCAATA
AAATTAAAAGCCCAGAAAAACTAGGGCCCATTGAAATCTTAGTTGCAAAATAAATGAACATATATCAAATAAATACAGTCCACT
ACTGTTATAAATGCAACTAATATACTAATGTACATCTCAGCTTGCTGGCCCTTTGGCAGAATGTTCACACATGAACACAAATAT
ATTTAAAGACTTACAATTTTGGGCTCCGTTCATATCTTATGTAAATGAATCGAGAGCGATAAATTATATTTAGGATTTTGTTAT
CTAAGGCGACATGGGTGCATTGCTCAAAAACATGTAATTTAAGTGCACACTACATGAGTCAGTCACTTGAGATCGTTCCCCGCC
TCCTAAAATAGTCCCTTAGTGGGAGACCACAGATAATGTCCTCGCCGCTCAAGATAGGCAGATGTGCCCGAGCGTGGGACCTCG
ATAAGGCGGGGACTATTTACTTAGGCCTCTGCGTAGGCCATTTACTTTAAGATGCGATTCTCATGTCACCTATTTAAACCGAAG
ATATTTCCAAATAAAATCAGTGTTTCTTACAAAAACTCAACGAGTAAAGTCTTCTTATTTGGGACATTACAATGATTATTTCCC
AACTACTCCCCCATTTTTCCCCAACATTAAGTGAAAGTCTCATAGGAGTCTGGATAATCTTAAAATTGTTTTAAGCTGCGTCAT
CTGAAGGGCTTAACCCTTAAACCCAACCGGAAGTAACGCCCGCCCTCTGGATGGAAAACGGGAAAGAGACGGGG

Table 4.S4: Results of singed (sn) screen at the Bloomington Drosophila Stock Center.
Results of the *sn* reversion screen for *FM7* stocks in the Bloomington *Drosophila* Stock Center. Columns include data on the Bloomington *Drosophila* Stock Center stock identifier, the phenotype at the *sn* locus in *Bar*-eyed males (sn^+ or sn^-), and the presence or absence of a 24-bp deletion associated with the g^4 allele that marks *bona fide FM7c* chromosomes. Presence or absence of the g^4 allele at the molecular level was only determined for 76 of the 82 sn^+ stocks, and the g^4 status for rest of the stocks is not determined (n.d.).

bdsc_id	genotype	singed_status	garnet_status
616	y[1] arm[4] w[*]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	sn[-]	n.d.
731	Df(1)N-264-105/FM7c	sn[-]	n.d.
816	w[1] ovo[svb-2]/FM7c	sn[+]	g[-]
935	Df(1)JC19/FM7c	sn[-]	n.d.
941	Df(1)HC244/FM7c	sn[-]	n.d.
943	Df(1)RC40/FM7c, sn[+]	sn[+]	g[-]
944	Df(1)JC70/FM7c, sn[+]	sn[+]	g[-]
950	Df(1)RA2/FM7c	sn[-]	n.d.
951	Df(1)KA14/FM7c	sn[-]	n.d.
955	Df(1)HC133, Imp[HC133]/FM7c	sn[-]	n.d.
959	Df(1)HA85/FM7c	sn[-]	n.d.
960	Df(1)KA6/FM7c	sn[-]	n.d.
961	Df(1)RA47/FM7c	sn[+]	g[-]
963	Df(1)KA10, sn[3] m[1]/FM7c	sn[-]	n.d.
964	Df(1)JA26/FM7c	sn[-]	n.d.
965	Df(1)HF368, cac[HF368]/FM7c	sn[-]	n.d.
968	Df(1)HA92/FM7c	sn[-]	n.d.
969	Df(1)KA9/FM7c	sn[-]	n.d.
971	Df(1)JA27/FM7c	sn[+]	g[+]
972	Df(1)HF396/FM7c	sn[-]	n.d.
978	Df(1)JC4/FM7c	sn[-]	n.d.
1098	Df(1)AD11, w[*]/FM7c	sn[-]	n.d.
1144	Df(1)AC7, w[*]/FM7c	sn[-]	n.d.
1399	In(1)JA9, l(1)7Aa[2]/FM7c	sn[-]	n.d.
1411	y[1] ph-d[503] w[1] f[36a]/FM7c, sn[+]/Dp(1;2;Y)w[+]	sn[-]	n.d.
1483	y[1] mew[M6] f[36a] P{ry[+t7.2]=neoFRT}18A/FM7c	sn[-]	n.d.
1494	Df(1)cho2, y[1] w[a]/FM7c	sn[-]	n.d.
1877	Df(1)GA102/FM7c	sn[-]	n.d.
2100	y[1] fog[S4]/FM7c	sn[-]	n.d.
2176	g[2] if[B2] f[36a]/FM7c	sn[-]	n.d.
2177	g[2] if[k27e] f[36a]/FM7c	sn[-]	n.d.
2181	y[1] sdt[XN]/FM7c	sn[+]	g[-]
2187	w[*] tsg[2]/FM7c	sn[-]	n.d.
2208	y[1] phm[E7]/FM7c	sn[-]	n.d.
2497	y[1] sog[S6]/FM7c, sn[+]	sn[+]	g[-]
3070	Df(1)E128/FM7c, w[+] <P>	sn[-]	n.d.
3146	otu[7]/FM7c	sn[-]	n.d.
3200	Df(1)bi-DL1, y[59b] z[1] w[i] ct[6] f[1]/FM7c, sn[+]	sn[+]	n.d.
3203	Df(1)bi-D2, w[*]/FM7c	sn[-]	n.d.
3204	Df(1)bi-D3, w[*]/FM7c	sn[-]	n.d.
3206	T(1;2)bi[D2], w[*] bi[D2]/FM7c	sn[+]	n.d.
3208	In(1)rb[D1], w[*] rb[D1]/FM7c	sn[+]	g[-]
3241	y[1] peb[hnt-E8]/FM7c	sn[-]	n.d.
3292	y[1] qs[8]/FM7c	sn[-]	n.d.
3293	y[1] ec[1] cv[1] ct[1] v[1] exd[1]/FM7c	sn[-]	n.d.

3294	y[1] btd[1]/FM7c	sn[-]	n.d.
3297	w[*] rtv[11]/FM7c	sn[-]	n.d.
3347	Df(1)sd72b/FM7c	sn[-]	n.d.
3372	Df(1)Sp(rb), y[1]/FM7c	sn[-]	n.d.
3378	y[1] arm[1]/FM7c	sn[-]	n.d.
3572	Df(1)bi-DL2, y[59b] z[1] w[i] ct[6] f[1]/FM7c	sn[+]	g[+]
3651	Df(1)lz-90b24, y[2] w[a]/FM7c	sn[-]	n.d.
3689	Df(1)18.1.15, y[1]/FM7c	sn[-]	n.d.
3694	Df(1)9a4-5, y[1] cv[1] v[1] f[1] car[1]/FM7c	sn[-]	n.d.
3728	w[*] P{w[+mC]=lacW}nej[P]/FM7c	sn[-]	n.d.
3729	w[*] nej[3]/FM7c	sn[-]	n.d.
4166	w[*] Chc[1]/FM7c/Dp(1;Y)shi[+]1, y[+] B[S]	sn[+]	g[-]
4167	w[*] Chc[4]/FM7c	sn[-]	n.d.
4168	mus109[lS]/FM7c	sn[-]	n.d.
4307	fs(1)A456[1] v[24]/FM7c	sn[-]	n.d.
4492	cv[1] ct[1] v[1] os[upd-1] mal[1]/FM7c	sn[-]	n.d.
4557	Df(1)JB254, P{w[+mC]=snf[+],dhd[+]}SL2, w[*]/FM7c	sn[-]	n.d.
4593	y[1] Sxl[f2]/FM7c	sn[-]	n.d.
4610	y[1] cv[1] ptg[13-342] lz[50e]/FM7c	sn[-]	n.d.
4618	y[1] ovo[D1rv20] v[24]/FM7c	sn[-]	n.d.
4649	Ns3[VE795]/FM7c	sn[-]	n.d.
4707	y[1] l(1)6PP7[1] f[1]/ FM7c <P>	sn[-]	n.d.
4721	w[1118] xmas-1[692-58] f[1] P{w[+mW.hs]=FRT(w[hs])}9-2/FM7c	sn[-]	n.d.
4852	olfF[x27]/FM7c	sn[+]	g[-]
4928	y[1] cv[1] v[1] l(1)15Db[815-14] f[1]/FM7c/Dp(1;Y)y[+]; sv[spa-pol]	sn[-]	n.d.
4929	y[1] cv[1] v[1] l(1)15De[692-47] f[1]/FM7c; sv[spa-pol]	sn[-]	n.d.
4953	Df(1)BK10, r[*] f[1]/FM7c	sn[-]	n.d.
5268	w[1118] E(Pc)17B[147]/FM7c	sn[-]	n.d.
5275	C(1;YL)1, y[1] N[TA17V] l(1)7Ad[TA17V]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5276	C(1;YL)1, y[1] WC1[WC1] Oce[WC1]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5277	C(1;YL)1, y[1] cv[1] v[1] g[2] exd[S136]/FM7c	sn[-]	n.d.
5278	C(1;YL)1, y[1] fs(1)h[rnc]/FM7c	sn[-]	n.d.
5284	cv[1] fs(1)h[1] v[1]/FM7c	sn[-]	n.d.
5285	fs(1)h[18]/FM7c	sn[-]	n.d.
5288	stout[UA104] v[1] f[1]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5289	w[1118] P{w[+mC]=EP}txl[Ab2]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5290	w[1118] P{w[+mC]=EP}txl[As1]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5293	y[1] qed[1] v[1] f[1]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5296	y[1] w[1] cv[1] v[1] UC119[UC119] f[1]/FM7c/Dp(1;Y)y[+]	sn[-]	n.d.
5371	Df(1)RC29, w[*]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	sn[-]	n.d.
5380	l(1)10Aj[1]/FM7c	sn[-]	n.d.
5383	l(1)10Bk[1]/FM7c	sn[-]	n.d.
5384	l(1)10Bo[3]/FM7c	sn[-]	n.d.
5401	y[1] mei-9[a] mei-41[D5] l(1)ESHS46[2]/FM7c	sn[-]	n.d.
5444	ph-d[401] ph-p[602] w[1]/FM7c	sn[+]	g[-]
5595	Nrg[l7]/FM7c	sn[+]	g[-]
5633	l(1)1CDa[28:26A2]/FM7c	sn[-]	n.d.
5634	l(1)1CDb[28:76A]/FM7c	sn[-]	n.d.
5635	l(1)ESHS10[1]/FM7c	sn[-]	n.d.
5636	l(1)ESHS12[1]/FM7c	sn[-]	n.d.
5637	l(1)ESHS15[1]/FM7c	sn[-]	n.d.

5638	l(1)ESHS30[1]/FM7c	sn[-]	n.d.
5639	l(1)ESHS47[1]/FM7c	sn[-]	n.d.
5640	l(1)ESHS32[3]/FM7c	sn[-]	n.d.
5641	l(1)ESHS33[2]/FM7c	sn[-]	n.d.
5644	l(1)ESHS45[1]/FM7c	sn[-]	n.d.
5645	l(1)ESHS49[1]/FM7c	sn[-]	n.d.
5646	l(1)ESHS50[1]/FM7c	sn[-]	n.d.
5648	pch[26]/FM7c	sn[-]	n.d.
5649	sbr[12]/FM7c	sn[-]	n.d.
5650	y[1] l(1)2Fd[19] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5651	y[1] l(1)ESHS11[1] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5653	y[1] l(1)ESHS19[1] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5654	y[1] l(1)ESHS1[1] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5655	y[1] l(1)ESHS3[3] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5656	y[1] l(1)ESHS4[1] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5657	y[1] l(1)ESHS6[1] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5660	y[1] mei-9[a] l(1)ESHS26[1] mei-41[D5]/FM7c	sn[-]	n.d.
5661	y[1] mei-9[a] l(1)ESHS27[1] mei-41[D5]/FM7c	sn[-]	n.d.
5662	y[1] mei-9[a] l(1)ESHS29[1] mei-41[D5]/FM7c	sn[-]	n.d.
5663	y[1] mei-9[a] l(1)ESHS35[1] mei-41[D5]/FM7c	sn[-]	n.d.
5664	y[1] mei-9[a] l(1)ESHS36[1] mei-41[D5]/FM7c	sn[-]	n.d.
5665	y[1] mei-9[a] l(1)ESHS37[1] mei-41[D5]/FM7c	sn[-]	n.d.
5666	y[1] mei-9[a] l(1)ESHS38[1] mei-41[D5]/FM7c	sn[-]	n.d.
5667	y[1] mei-9[a] l(1)ESHS41[1] mei-41[D5]/FM7c	sn[-]	n.d.
5668	y[1] mei-9[a] l(1)ESHS43[1] mei-41[D5]/FM7c	sn[-]	n.d.
5669	y[1] mei-9[a] mei-41[D5] l(1)19Ed[17]/FM7c	sn[-]	n.d.
5672	y[1] mei-9[a] rtv[10] mei-41[D5]/FM7c	sn[-]	n.d.
5673	y[1] mul[3] mei-9[a] mei-41[D5]/FM7c	sn[-]	n.d.
5706	Df(1)M38-C5/FM7c	sn[-]	n.d.
5707	Df(1)v[N124B]/FM7c	sn[-]	n.d.
5708	Nrg[l4]/FM7c	sn[-]	n.d.
5709	ccw[1] sma[1] up[1] mal[F1]/FM7c	sn[-]	n.d.
5710	drw[2]/FM7c	sn[-]	n.d.
5713	l(1)1Bm[8-12-2] mei-9[a]/FM7c/Dp(1;Y)y[2]67g19.1	sn[-]	n.d.
5714	l(1)2Bu[1]/FM7c	sn[-]	n.d.
5715	l(1)7Ab[23]/FM7c	sn[-]	n.d.
5716	l(1)7Ci[7-87]/FM7c	sn[-]	n.d.
5717	l(1)7Df[1]/FM7c	sn[-]	n.d.
5732	pch[12] y[2] w[i] ct[6] f[1]/FM7c	sn[-]	n.d.
5734	ras[1] v[1] m[1] l(1)10Fd[1]/FM7c	sn[-]	n.d.
5735	rud[1] v[54k] tc[1] sl[2] smd[1]/FM7c, sn[*]	sn[+]	g[-]
5738	v[1] ty[2]/FM7c	sn[-]	n.d.
5741	wapl[2]/FM7c	sn[-]	n.d.
5981	Df(1)cho10, y[1] w[1]/FM7c	sn[-]	n.d.
5982	Df(1)cho24, y[1] w[1]/FM7c	sn[-]	n.d.
5983	Df(1)cho25, y[1] w[1]/FM7c	sn[-]	n.d.
5986	Df(1)EA113/FM7c	sn[-]	n.d.
5988	Df(1)GA37/FM7c	sn[-]	n.d.
5989	Df(1)GA104/FM7c/Dp(1;Y)y[+]mal[106]	sn[-]	n.d.
5990	Df(1)HC194/FM7c	sn[-]	n.d.
5992	Df(1)HF359/FM7c	sn[-]	n.d.
5994	Df(1)JA21/FM7c	sn[-]	n.d.
5995	Df(1)v-JA22/FM7c	sn[-]	n.d.

5996	Df(1)JC12/FM7c	sn[-]	n.d.
5998	Df(1)JC77/FM7c/Dp(1;Y)y[+]mal[106]	sn[+]	g[-]
6006	Df(1)R8A/FM7c	sn[+]	g[+]
6009	Df(1)RR62, w[1118]/FM7c	sn[-]	n.d.
6011	Df(1)rb14, y[1] w[1]/FM7c	sn[+]	g[-]
6012	Df(1)rb23, y[1] w[1]/FM7c	sn[-]	n.d.
6013	Df(1)rb42, y[1] w[1]/FM7c	sn[-]	n.d.
6014	Df(1)RF19, In(1)RF19/FM7c	sn[-]	n.d.
6018	Df(1)w67k30, lz[1] ras[1] v[1]/FM7c	sn[-]	n.d.
6030	In(1LR)pn2a, Vinc[1]/FM7c/Dp(1;Y)B[S]	sn[+]	g[-]
6039	y[2] sc[1] z[1] w[*] N[spl-1] sn[3] v[1] g[2] f[1]/FM7c	sn[-]	n.d.
6278	Df(1)ct4b1, Dp(1;1)sn[S93]/FM7c	sn[-]	n.d.
6307	Df(1)R29, y[1]/FM7c	sn[-]	n.d.
6328	y[1] ac[Hw-1] v[1] RpII215[Ubl]/FM7c	sn[-]	n.d.
6666	y[1] w[1118] rok[2] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
6761	w[*] hep[r75]/FM7c	sn[-]	n.d.
6873	y[1] w[*] N[1]/FM7c, P{w[+mC]=GAL4-twi.G}108.4, P{UAS-2xEGFP}AX	sn[-]	n.d.
6888	tsg[4] g[1]/FM7c	sn[-]	n.d.
6890	T(1;2)JC68, l(1)8Aa[5]/FM7c	sn[-]	n.d.
6892	Df(1)TEM7, y[2] w[i] ct[6] f[1]/FM7c	sn[-]	n.d.
6895	l(1)7Eb[7]/FM7c	sn[-]	n.d.
6896	l(1)air7[28]/Dp(1;Y)ct[+]y[+]/FM7c	sn[-]	n.d.
7059	Exp6[8-36-2]/Dp(1;Y)y[2]67g19.1/FM7c	sn[-]	n.d.
7152	Df(1)w-194A/FM7c	sn[-]	n.d.
7202	Df(1)fu-Z4, y[1] w[1] sn[3] f[1]/FM7c	sn[-]	n.d.
7339	In(1)AC2[L]AB[R], y[1] w[1]/FM7c	sn[-]	n.d.
7435	Rbf[14] w[1118]/FM7c	sn[-]	n.d.
7699	Df(1)Exel6221, P{w[+mC]=XP-U}Exel6221 w[1118]/FM7c	sn[-]	n.d.
7700	Df(1)Exel6223, P{w[+mC]=XP-U}Exel6223 w[1118]/FM7c	sn[+]	g[-]
7702	Df(1)Exel6225, P{w[+mC]=XP-U}Exel6225 w[1118]/FM7c	sn[-]	n.d.
7703	Df(1)Exel6226, P{w[+mC]=XP-U}Exel6226 w[1118]/FM7c	sn[-]	n.d.
7704	Df(1)Exel6227, P{w[+mC]=XP-U}Exel6227 w[1118]/FM7c	sn[-]	n.d.
7705	Df(1)Exel6230, P{w[+mC]=XP-U}Exel6230 w[1118]/FM7c	sn[-]	n.d.
7706	Df(1)Exel6231, P{w[+mC]=XP-U}Exel6231 w[1118]/FM7c	sn[-]	n.d.
7707	Df(1)Exel6233, w[1118] P{w[+mC]=XP-U}Exel6233/FM7c	sn[-]	n.d.
7708	Df(1)Exel6234, w[1118] P{w[+mC]=XP-U}Exel6234/FM7c	sn[-]	n.d.
7709	Df(1)Exel6235, w[1118] P{w[+mC]=XP-U}Exel6235/FM7c	sn[+]	g[-]
7710	Df(1)Exel6236, w[1118] P{w[+mC]=XP-U}Exel6236/FM7c	sn[-]	n.d.
7711	Df(1)Exel6237, w[1118] P{w[+mC]=XP-U}Exel6237/FM7c	sn[-]	n.d.
7712	Df(1)Exel6238, w[1118] P{w[+mC]=XP-U}Exel6238/FM7c	sn[-]	n.d.
7713	Df(1)Exel6239, w[1118] P{w[+mC]=XP-U}Exel6239/FM7c	sn[-]	n.d.
7714	Df(1)Exel6240, w[1118] P{w[+mC]=XP-U}Exel6240/FM7c	sn[-]	n.d.
7715	Df(1)Exel6241, w[1118] P{w[+mC]=XP-U}Exel6241/FM7c	sn[-]	n.d.
7716	Df(1)Exel6242, w[1118] P{w[+mC]=XP-U}Exel6242/FM7c	sn[-]	n.d.
7717	Df(1)Exel6244, w[1118] P{w[+mC]=XP-U}Exel6244/FM7c	sn[-]	n.d.
7718	Df(1)Exel6245, w[1118] P{w[+mC]=XP-U}Exel6245/FM7c	sn[-]	n.d.
7719	Df(1)Exel6248, w[1118] P{w[+mC]=XP-U}Exel6248/FM7c	sn[-]	n.d.
7720	Df(1)Exel6251, w[1118] P{w[+mC]=XP-U}Exel6251/FM7c	sn[-]	n.d.
7721	Df(1)Exel6253, w[1118] P{w[+mC]=XP-U}Exel6253/FM7c	sn[-]	n.d.
7722	Df(1)Exel6254, w[1118] P{w[+mC]=XP-U}Exel6254/FM7c	sn[-]	n.d.
7723	Df(1)Exel6255, w[1118] P{w[+mC]=XP-U}Exel6255/FM7c	sn[-]	n.d.
7756	P{w[+mC]=XP}XPG-L, w[1118]/FM7c	sn[-]	n.d.

8247	P{w[+mC]=UAS-HLH106.P450}1, y[1] w[1118]/FM7c	sn[-]	n.d.
8342	Rbf[cas-21]/FM7c	sn[-]	n.d.
8343	Rbf[sls-15] w[1118] P{ry[+t7.2]=neoFRT}18A/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	sn[-]	n.d.
8413	P{w[+mC]=dpp-lacZ.Exel.2}1, y[1] w[1118]/FM7c	sn[-]	n.d.
8492	y[1] v[1] hop[Tum]/FM7c	sn[-]	n.d.
8493	y[1] hop[27]/FM7c	sn[-]	n.d.
8494	y[1] hop[25]/FM7c	sn[-]	n.d.
8495	y[1] w[*] hop[3]/FM7c	sn[-]	n.d.
8675	w[*] oc[otd-YH13] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
8676	y[1] sn[3] oc[2] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
9103	w[*] Mer[3] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
9165	y[1] cv[1] rux[4]/FM7c	sn[-]	n.d.
9167	y[1] cv[1] shtd[1]/FM7c	sn[-]	n.d.
9242	y[1] cv[1] rux[8]/FM7c	sn[-]	n.d.
10070	w[67c23] P{w[+mC]=lacW}l(1)G0004[G0004]/FM7c	sn[-]	n.d.
10071	w[67c23] P{w[+mC]=lacW}kdn[G0033]/FM7c	sn[-]	n.d.
10079	w[67c23] P{w[+mC]=lacW}Sas10[G0106]/FM7c	sn[-]	n.d.
10092	P{w[+mC]=lacW}sgg[G0183] w[67c23]/FM7c	sn[-]	n.d.
10096	w[67c23] P{w[+mC]=lacW}Mipp2[G0303]/FM7c	sn[-]	n.d.
10111	w[67c23] P{w[+mC]=lacW}l(1)G0455[G0455]/FM7c	sn[-]	n.d.
10112	P{w[+mC]=lacW}l(1)G0458[G0458] w[67c23]/FM7c	sn[-]	n.d.
10125	w[67c23] P{w[+mC]=lacW}l(1)G0473[G0473]/FM7c	sn[-]	n.d.
10127	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0476]/FM7c	sn[-]	n.d.
10132	w[67c23] P{w[+mC]=lacW}sog[G0479]/FM7c	sn[-]	n.d.
10134	w[67c23] P{w[+mC]=lacW}Trxr-1[G0481]/FM7c	sn[-]	n.d.
10138	w[67c23] P{w[+mC]=lacW}ras[G0482]/FM7c	sn[-]	n.d.
10148	w[67c23] P{w[+mC]=lacW}l(1)G0141[G0141]/FM7c	sn[-]	n.d.
10151	P{w[+mC]=lacW}G0145a w[67c23] P{lacW}Tango5[G0145b] P{lacW}G0145c, l(1)G0145[G0145]/FM7c	sn[-]	n.d.
10255	w[67c23] P{w[+mC]=lacW}l(1)G0277[G0277]/FM7c	sn[-]	n.d.
10263	w[67c23] P{w[+mC]=lacW}dome[G0282]/FM7c	sn[-]	n.d.
10270	w[67c23] P{w[+mC]=lacW}e(y)3[G0283]/FM7c	sn[-]	n.d.
10282	w[67c23] P{w[+mC]=lacW}l(1)G0255[G0290]/FM7c	sn[-]	n.d.
10317	w[67c23] P{w[+mC]=lacW}l(1)G0298[G0298]/FM7c	sn[-]	n.d.
10597	P{w[+mC]=lacW}deltaCOP[G0301] w[67c23]/FM7c	sn[+]	g[-]
10600	w[67c23] P{w[+mC]=lacW}sd[G0309]/FM7c	sn[-]	n.d.
10760	w[67c23] P{w[+mC]=lacW}Nat1[G0311b]/FM7c	sn[-]	n.d.
10785	y[1] w[1118] PBac{w[+mC]=PB}eIF2B-beta[c02002]/FM7c	sn[-]	n.d.
10805	w[67c23] P{w[+mC]=lacW}l(1)G0045[G0312]/FM7c	sn[-]	n.d.
10812	w[67c23] P{w[+mC]=lacW}sd[G0315b]/FM7c	sn[-]	n.d.
10897	w[67c23] P{w[+mC]=lacW}Moe[G0323]/FM7c, B[+]	sn[-]	n.d.
10958	w[67c23] P{w[+mC]=lacW}l(1)G0324[G0324]/FM7c	sn[-]	n.d.
10987	y[1] w[67c23] P{y[+t7.7]=Mae-UAS.6.11}bou[GG01077]/FM7c	sn[-]	n.d.
11171	P{w[+mC]=lacW}G0329a w[67c23] P{lacW}spri[G0329b], l(1)G0329[G0329]/FM7c	sn[-]	n.d.
11466	w[67c23] P{w[+mC]=lacW}Rip11[G0003]/FM7c	sn[+]	g[-]
11467	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0007]/FM7c	sn[-]	n.d.
11470	w[67c23] P{w[+mC]=lacW}Act5C[G0009] dome[G0009]/FM7c	sn[-]	n.d.
11471	y[1] P{w[+mC]=lacW}pck[G0012] w[67c23]/FM7c	sn[-]	n.d.

11475	w[67c23] P{w[+mC]=lacW}Tcp-1zeta[G0022]/FM7c	sn[-]	n.d.
11477	P{w[+mC]=lacW}trol[G0023] w[67c23]/FM7c	sn[-]	n.d.
11496	P{w[+mC]=lacW}l(1)G0024[G0024] w[67c23]/FM7c	sn[-]	n.d.
11499	w[67c23] P{w[+mC]=lacW}Act5C[G0025]/FM7c	sn[-]	n.d.
11503	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0028]/FM7c	sn[-]	n.d.
11516	w[67c23] P{w[+mC]=lacW}kdn[G0030]/FM7c	sn[-]	n.d.
11517	w[67c23] P{w[+mC]=lacW}lnx2[G0035]/FM7c	sn[-]	n.d.
11519	w[67c23] P{w[+mC]=lacW}lnx2[G0036]/FM7c	sn[-]	n.d.
11523	P{w[+mC]=lacW}skpA[G0037] w[67c23]/FM7c	sn[-]	n.d.
11546	w[67c23] P{w[+mC]=lacW}Trf2[G0039]/FM7c	sn[-]	n.d.
11547	w[67c23] P{w[+mC]=lacW}RpII215[G0040]/FM7c	sn[-]	n.d.
11550	w[67c23] P{w[+mC]=lacW}lnx2[G0043]/FM7c	sn[-]	n.d.
11559	y[1] P{w[+mC]=lacW}pck[G0044] w[67c23]/FM7c	sn[-]	n.d.
11573	w[67c23] P{w[+mC]=lacW}l(1)G0045[G0045]/FM7c	sn[-]	n.d.
11576	P{w[+mC]=lacW}G0046a w[67c23] P{lacW}G0046b, l(1)G0046[G0046]/FM7c	sn[-]	n.d.
11580	w[67c23] P{w[+mC]=lacW}Mipp2[G0050]/FM7c	sn[-]	n.d.
11588	P{w[+mC]=lacW}deltaCOP[G0051] w[67c23]/FM7c	sn[-]	n.d.
11593	w[67c23] P{w[+mC]=lacW}tlk[G0054]/FM7c	sn[+]	g[-]
11596	P{w[+mC]=lacW}sgg[G0055] w[67c23]/FM7c	sn[-]	n.d.
11640	w[67c23] P{w[+mC]=lacW}Tcp-1zeta[G0057]/FM7c	sn[-]	n.d.
11649	P{w[+mC]=lacW}skpA[G0058] w[67c23]/FM7c	sn[-]	n.d.
11665	w[67c23] P{w[+mC]=lacW}schlank[G0061] schlank[G0061]/FM7c	sn[-]	n.d.
11673	w[67c23] P{w[+mC]=lacW}l(1)G0062[G0062]/FM7c	sn[-]	n.d.
11680	w[67c23] P{w[+mC]=lacW}Spt6[G0063]/FM7c	sn[-]	n.d.
11691	y[1] w[67c23] P{y[+t7.7]=Mae-UAS.6.11}CG43689[GG01071]/FM7c	sn[-]	n.d.
11696	w[67c23] P{w[+mC]=lacW}Dlic[G0065]/FM7c	sn[-]	n.d.
11701	P{w[+mC]=lacW}Unc-76[G0066] w[67c23]/FM7c	sn[-]	n.d.
11708	w[67c23] P{w[+mC]=lacW}Moe[G0067]/FM7c	sn[-]	n.d.
11727	w[67c23] P{w[+mC]=lacW}Rip11[G0070]/FM7c	sn[-]	n.d.
11739	w[67c23] P{w[+mC]=lacW}Trf2[G0071]/FM7c	sn[-]	n.d.
11798	w[67c23] P{w[+mC]=lacW}Imp[G0072]/FM7c	sn[-]	n.d.
11799	w[67c23] P{w[+mC]=lacW}beta-Spec[G0074]/FM7c	sn[-]	n.d.
11800	w[67c23] P{w[+mC]=lacW}Ran[G0075]/FM7c	sn[-]	n.d.
11801	w[67c23] P{w[+mC]=lacW}l(1)G0076[G0076]/FM7c	sn[-]	n.d.
11802	P{w[+mC]=lacW}Actn[G0077] w[67c23]/FM7c	sn[-]	n.d.
11803	w[67c23] P{w[+mC]=lacW}mew[G0078a] P{lacW}G0078b, l(1)G0078[G0078]/FM7c	sn[-]	n.d.
11805	w[67c23] P{w[+mC]=lacW}Fas2[G0081]/FM7c	sn[-]	n.d.
11807	w[67c23] P{w[+mC]=lacW}e(y)3[G0084]/FM7c	sn[-]	n.d.
11808	w[67c23] P{w[+mC]=lacW}Ntf-2[G0086]/FM7c	sn[-]	n.d.
11809	w[67c23] P{w[+mC]=lacW}lnx2[G0090a] P{lacW}G0090b, l(1)G0090[G0090]/FM7c	sn[-]	n.d.
11811	w[67c23] P{w[+mC]=lacW}ctp[G0094a] P{lacW}ras[G0094b], l(1)G0094[G0094]/FM7c	sn[-]	n.d.
11812	w[67c23] P{w[+mC]=lacW}IntS4[G0095]/FM7c	sn[-]	n.d.
11814	w[67c23] P{w[+mC]=lacW}Nrg[G0099]/FM7c	sn[-]	n.d.
11815	w[67c23] P{w[+mC]=lacW}Hsc70-3[G0102]/FM7c	sn[-]	n.d.
11819	w[67c23] P{w[+mC]=lacW}beta-Spec[G0108]/FM7c	sn[-]	n.d.
11820	w[67c23] P{w[+mC]=lacW}Hsc70-3[G0111]/FM7c	sn[-]	n.d.
11822	w[67c23] P{lacW}tlk[G0113a] P{lacW}G0113b	sn[-]	n.d.

	P{lacW}wisp[G0113c], l(1)G0113[G0113]/FM7c		
11823	P{w[+mC]=lacW}l(1)G0115[G0115] w[67c23]/FM7c	sn[-]	n.d.
11824	w[67c23] P{w[+mC]=lacW}l(1)G0116[G0116]/FM7c	sn[-]	n.d.
11826	w[67c23] P{w[+mC]=lacW}lnx2[G0118]/FM7c	sn[-]	n.d.
11827	w[67c23] P{w[+mC]=lacW}nej[G0119a] P{lacW}G0119b, l(1)G0119[G0119]/FM7c	sn[-]	n.d.
11828	w[67c23] P{w[+mC]=lacW}l(1)G0120[G0120]/FM7c	sn[-]	n.d.
11830	w[67c23] P{w[+mC]=lacW}Ant2[G0126] sesB[G0126]/FM7c	sn[-]	n.d.
11832	w[67c23] P{w[+mC]=lacW}l(1)G0128[G0128]/FM7c	sn[-]	n.d.
11833	P{w[+mC]=lacW}l(1)G0129[G0129] w[67c23]/FM7c	sn[-]	n.d.
11834	P{w[+mC]=lacW}sta[G0130] w[67c23]/FM7c	sn[-]	n.d.
11835	P{w[+mC]=lacW}l(1)G0132[G0132] w[67c23]/FM7c	sn[-]	n.d.
11836	w[67c23] P{w[+mC]=lacW}l(1)G0137[G0137]/FM7c	sn[-]	n.d.
11837	w[67c23] P{w[+mC]=lacW}Vav[G0147]/FM7c	sn[-]	n.d.
11838	P{w[+mC]=lacW}cin[G0142] CG42376[G0142] w[67c23]/FM7c	sn[-]	n.d.
11839	P{w[+mC]=lacW}l(1)G0259[G0259] w[67c23] /FM7c	sn[+]	g[-]
11840	P{w[+mC]=lacW}l(1)G0296[G0296] w[67c23]/FM7c	sn[-]	n.d.
11841	P{w[+mC]=lacW}br[G0284a] P{lacW}G0284b w[67c23], l(1)G0284[G0284]/FM7c	sn[-]	n.d.
11842	P{w[+mC]=lacW}arm[G0268a] P{lacW}G0268b w[67c23], l(1)G0268[G0268]/FM7c	sn[-]	n.d.
11843	P{w[+mC]=lacW}csw[G0170] w[67c23]/FM7c	sn[-]	n.d.
11844	P{w[+mC]=lacW}l(1)G0144[G0144] w[67c23]/FM7c	sn[-]	n.d.
11845	P{w[+mC]=lacW}sgg[G0263] w[67c23]/FM7c	sn[-]	n.d.
11846	P{w[+mC]=lacW}l(1)G0211[G0211] w[67c23]/FM7c	sn[+]	g[-]
11847	w[67c23] P{w[+mC]=lacW}ctp[G0244a]/FM7c	sn[-]	n.d.
11848	P{w[+mC]=lacW}trol[G0271] w[67c23]/FM7c	sn[-]	n.d.
11849	P{w[+mC]=lacW}wds[G0251] w[67c23]/FM7c	sn[-]	n.d.
11851	w[67c23] P{w[+mC]=lacW}ctp[G0153]/FM7c	sn[-]	n.d.
11853	w[67c23] P{w[+mC]=lacW}Act5C[G0245]/FM7c	sn[+]	g[-]
11855	w[67c23] P{w[+mC]=lacW}l(1)G0254[G0254]/FM7c	sn[-]	n.d.
11856	w[67c23] P{w[+mC]=lacW}lnx2[G0059]/FM7c	sn[-]	n.d.
11857	w[67c23] P{w[+mC]=lacW}lnx2[G0317]/FM7c	sn[-]	n.d.
11858	w[67c23] P{w[+mC]=lacW}Ykt6[G0155]/FM7c	sn[-]	n.d.
11859	w[67c23] P{w[+mC]=lacW}Tom40[G0216]/FM7c	sn[-]	n.d.
11860	w[67c23] P{w[+mC]=lacW}CG42593[G0307a] P{lacW}G0307b, l(1)G0307[G0307]/FM7c	sn[-]	n.d.
11862	w[67c23] P{w[+mC]=lacW}mys[G0233]/FM7c	sn[-]	n.d.
11863	w[67c23] P{w[+mC]=lacW}cyr[G0199] P{lacW}dome[G0199b], l(1)G0199[G0199]/FM7c	sn[-]	n.d.
11864	w[67c23] P{w[+mC]=lacW}Trxr-1[G0154]/FM7c	sn[-]	n.d.
11865	w[67c23] P{w[+mC]=lacW}l(1)G0228[G0228]/FM7c	sn[-]	n.d.
11866	w[67c23] P{w[+mC]=lacW}l(1)G0203[G0203]/FM7c	sn[-]	n.d.
11867	w[67c23] P{w[+mC]=lacW}l(1)G0219[G0219]/FM7c	sn[-]	n.d.
11868	w[67c23] P{w[+mC]=lacW}l(1)G0178[G0178]/FM7c	sn[-]	n.d.
11869	w[67c23] P{w[+mC]=lacW}l(1)G0249[G0249]/FM7c	sn[-]	n.d.
11870	w[67c23] P{w[+mC]=lacW}l(1)G0200[G0200]/FM7c	sn[-]	n.d.
11871	w[67c23] P{w[+mC]=lacW}l(1)G0286[G0286]/FM7c	sn[-]	n.d.
11872	w[67c23] P{w[+mC]=lacW}l(1)G0196[G0196]/FM7c	sn[-]	n.d.
11873	w[67c23] P{w[+mC]=lacW}l(1)G0270[G0270]/FM7c	sn[-]	n.d.
11875	w[67c23] P{w[+mC]=lacW}dsh[G0267]/FM7c	sn[-]	n.d.
11876	w[67c23] P{w[+mC]=lacW}dlg1[G0276]/FM7c	sn[-]	n.d.

11877	w[67c23] P{w[+mC]=lacW}Kmn1[G0237]/FM7c	sn[-]	n.d.
11879	w[67c23] P{w[+mC]=lacW}RPA3[G0241]/FM7c	sn[-]	n.d.
11880	w[67c23] P{w[+mC]=lacW}lic[G0252]/FM7c	sn[-]	n.d.
11881	w[67c23] P{w[+mC]=lacW}l(1)G0191[G0191]/FM7c	sn[-]	n.d.
11882	w[67c23] P{w[+mC]=lacW}l(1)G0186[G0186]/FM7c	sn[-]	n.d.
11883	w[67c23] P{w[+mC]=lacW}Top1[G0229]/FM7c	sn[-]	n.d.
11886	w[67c23] P{w[+mC]=lacW}l(1)G0136[G0136]/FM7c	sn[-]	n.d.
11887	w[67c23] P{w[+mC]=lacW}sd[G0239]/FM7c	sn[-]	n.d.
11888	w[67c23] P{w[+mC]=lacW}sd[G0262]/FM7c	sn[-]	n.d.
11890	w[67c23] P{w[+mC]=lacW}l(1)G0212[G0212]/FM7c	sn[-]	n.d.
11891	w[67c23] P{w[+mC]=lacW}G0221a P{lacW}mei-41[G0221b], l(1)G0221[G0221]/FM7c	sn[-]	n.d.
11892	w[67c23] P{w[+mC]=lacW}l(1)G0272[G0272]/FM7c	sn[-]	n.d.
11894	w[67c23] P{w[+mC]=lacW}l(1)G0013[G0013]/FM7c	sn[-]	n.d.
11895	w[67c23] P{w[+mC]=lacW}l(1)G0041[G0041]/FM7c	sn[-]	n.d.
11896	w[67c23] P{w[+mC]=lacW}l(1)G0156[G0156]/FM7c	sn[-]	n.d.
11897	w[67c23] P{w[+mC]=lacW}e(y)3[G0266]/FM7c	sn[-]	n.d.
11900	w[67c23] P{w[+mC]=lacW}l(1)G0223[G0223]/FM7c	sn[-]	n.d.
11902	w[67c23] P{w[+mC]=lacW}l(1)G0179[G0179]/FM7c	sn[-]	n.d.
11903	P{w[+mC]=lacW}l(1)G0355[G0355] w[67c23]/FM7c	sn[-]	n.d.
11904	P{w[+mC]=lacW}Unc-76[G0360] w[67c23]/FM7c	sn[-]	n.d.
11905	w[67c23] P{w[+mC]=lacW}CG7884[G0363a] P{lacW}G0363b, l(1)G0363[G0363]/FM7c	sn[-]	n.d.
11906	w[67c23] P{w[+mC]=lacW}l(1)G0384[G0384]/FM7c	sn[-]	n.d.
11907	w[67c23] P{w[+mC]=lacW}l(1)G0394[G0394]/FM7c	sn[-]	n.d.
11908	w[67c23] P{w[+mC]=lacW}sog[G0395]/FM7c	sn[-]	n.d.
11910	w[67c23] P{w[+mC]=lacW}e(y)3[G0409]/FM7c	sn[-]	n.d.
11911	P{w[+mC]=lacW}Unc-76[G0423a] w[67c23] P{lacW}G0423b, l(1)G0423[G0423]/FM7c	sn[-]	n.d.
11912	w[67c23] P{w[+mC]=lacW}vfl[G0427]/FM7c	sn[-]	n.d.
11913	w[67c23] P{w[+mC]=lacW}Ntf-2[G0428]/FM7c	sn[+]	g[-]
11914	w[67c23] P{w[+mC]=lacW}Tcp-1zeta[G0027]/FM7c	sn[-]	n.d.
11915	Df(1)G0124, w[67c23] P{w[+mC]=lacW}G0124 Tis11[G0124] Smr[G0124]/FM7c	sn[-]	n.d.
11916	w[67c23] P{w[+mC]=lacW}CG43658[G0143b] P{lacW}CG12991[G0143a]/FM7c	sn[-]	n.d.
11917	w[67c23] P{w[+mC]=lacW}G0161a P{lacW}Trf2[G0161b], l(1)G0161[G0161]/FM7c	sn[-]	n.d.
11918	w[67c23] P{w[+mC]=lacW}l(1)G0164[G0164]/FM7c	sn[-]	n.d.
11919	w[67c23] P{w[+mC]=lacW}Aats-his[G0165]/FM7c	sn[-]	n.d.
11921	w[67c23] P{w[+mC]=lacW}Act5C[G0177]/FM7c	sn[-]	n.d.
11922	w[67c23] P{w[+mC]=lacW}beta-Spec[G0198]/FM7c	sn[-]	n.d.
11923	w[67c23] P{w[+mC]=lacW}Top1[G0201]/FM7c	sn[-]	n.d.
11924	P{w[+mC]=lacW}l(1)G0220[G0220] w[67c23]/FM7c	sn[-]	n.d.
11925	w[67c23] P{w[+mC]=lacW}l(1)G0250[G0250]/FM7c	sn[-]	n.d.
11926	w[67c23] P{w[+mC]=lacW}Act5C[G0330]/FM7c	sn[+]	g[-]
11927	w[67c23] P{w[+mC]=lacW}l(1)G0332[G0332]/FM7c	sn[-]	n.d.
11928	w[67c23] P{w[+mC]=lacW}Rpt4[G0345]/FM7c	sn[-]	n.d.
11929	w[67c23] P{w[+mC]=lacW}l(1)G0346[G0346]/FM7c	sn[-]	n.d.
11930	w[67c23] P{w[+mC]=lacW}l(1)G0006[G0006]/FM7c	sn[-]	n.d.
11931	y[1] w[67c23] P{w[+mC]=lacW}l(1)G0011[G0011]/FM7c	sn[-]	n.d.
11932	w[67c23] P{w[+mC]=lacW}ctp[G0445b]/FM7c	sn[-]	n.d.
11933	P{w[+mC]=lacW}br[G0042] w[67c23]/FM7c	sn[-]	n.d.

11934	w[67c23] P{w[+mC]=lacW}ras[G0098]/FM7c	sn[-]	n.d.
11935	w[67c23] P{w[+mC]=lacW}dm[G0139]/FM7c	sn[-]	n.d.
11936	w[67c23] P{w[+mC]=lacW}kdn[G0140a] P{lacW}G0140b, l(1)G0140[G0140]/FM7c	sn[-]	n.d.
11937	w[67c23] P{w[+mC]=lacW}l(1)G0148[G0148]/FM7c	sn[-]	n.d.
11938	w[67c23] P{w[+mC]=lacW}l(1)G0148[G0149]/FM7c	sn[-]	n.d.
11939	P{w[+mC]=lacW}G0151a P{lacW}G0151b w[67c23] P{lacW}sog[G0151c], l(1)G0151[G0151]/FM7c	sn[-]	n.d.
11940	w[67c23] P{w[+mC]=lacW}l(1)G0152[G0152]/FM7c	sn[-]	n.d.
11941	w[67c23] P{w[+mC]=lacW}lnx2[G0157]/FM7c	sn[-]	n.d.
11942	P{w[+mC]=lacW}Unc-76[G0158] w[67c23]/FM7c	sn[-]	n.d.
11943	w[67c23] P{w[+mC]=lacW}kdn[G0159]/FM7c	sn[-]	n.d.
11946	w[67c23] P{w[+mC]=lacW}l(1)G0167[G0167]/FM7c	sn[-]	n.d.
11947	w[67c23] P{w[+mC]=lacW}lnx2[G0173a] P{lacW}G0173b, l(1)G0173[G0173]/FM7c	sn[-]	n.d.
11948	w[67c23] P{w[+mC]=lacW}l(1)G0175[G0175]/FM7c	sn[-]	n.d.
11949	P{w[+mC]=lacW}l(1)G0181[G0181] w[67c23]/FM7c	sn[-]	n.d.
11951	w[67c23] P{w[+mC]=lacW}Dlic[G0190]/FM7c	sn[-]	n.d.
11952	w[67c23] P{w[+mC]=lacW}G0213a P{lacW}G0213b, l(1)G0213[G0213]/FM7c	sn[-]	n.d.
11953	w[67c23] P{w[+mC]=lacW}dome[G0218]/FM7c	sn[-]	n.d.
11955	w[67c23] P{w[+mC]=lacW}l(1)G0230[G0230]/FM7c	sn[-]	n.d.
11956	w[67c23] P{w[+mC]=lacW}l(1)G0232[G0232]/FM7c	sn[-]	n.d.
11957	w[67c23] P{w[+mC]=lacW}ras[G0238]/FM7c	sn[-]	n.d.
11959	w[67c23] P{w[+mC]=lacW}Ant2[G0247] sesB[G0247]/FM7c	sn[-]	n.d.
11960	w[67c23] P{w[+mC]=lacW}ctp[G0248]/FM7c	sn[-]	n.d.
11961	w[67c23] P{w[+mC]=lacW}l(1)G0255[G0255]/FM7c	sn[-]	n.d.
11962	w[67c23] P{w[+mC]=lacW}dome[G0264]/FM7c	sn[-]	n.d.
11964	w[67c23] P{w[+mC]=lacW}mys[G0281]/FM7c	sn[-]	n.d.
11965	w[67c23] P{w[+mC]=lacW}l(1)G0285[G0285]/FM7c	sn[-]	n.d.
11966	w[67c23] P{w[+mC]=lacW}l(1)G0289[G0289]/FM7c	sn[-]	n.d.
11967	w[67c23] P{w[+mC]=lacW}Hsc70-3[G0292]/FM7c	sn[+]	n.d.
11968	w[67c23] P{w[+mC]=lacW}l(1)G0299[G0299]/FM7c	sn[-]	n.d.
11969	w[67c23] P{w[+mC]=lacW}l(1)G0222[G0316]/FM7c	sn[-]	n.d.
11970	w[67c23] P{w[+mC]=lacW}l(1)G0320[G0320]/FM7c	sn[-]	n.d.
11971	w[67c23] P{w[+mC]=lacW}dome[G0321]/FM7c	sn[-]	n.d.
11973	P{w[+mC]=lacW}sgg[G0335] w[67c23]/FM7c	sn[-]	n.d.
11974	w[67c23] P{w[+mC]=lacW}Fas2[G0336]/FM7c	sn[-]	n.d.
11976	w[67c23] P{w[+mC]=lacW}dlg1[G0342]a P{lacW}dlg1[G0342]b dlg1[G0342]/FM7c	sn[-]	n.d.
11977	w[67c23] P{w[+mC]=lacW}l(1)G0343[G0343]/FM7c	sn[+]	g[-]
11978	w[67c23] P{w[+mC]=lacW}nej[G0350]/FM7c	sn[-]	n.d.
11979	w[67c23] P{w[+mC]=lacW}ras[G0351]/FM7c	sn[-]	n.d.
11980	w[67c23] P{w[+mC]=lacW}vfl[G0353]/FM7c	sn[-]	n.d.
11983	w[67c23] P{w[+mC]=lacW}Aats-his[G0358]/FM7c	sn[-]	n.d.
11986	w[67c23] P{w[+mC]=lacW}dome[G0367]/FM7c	sn[-]	n.d.
11987	w[67c23] P{w[+mC]=lacW}l(1)G0369[G0369]/FM7c	sn[-]	n.d.
11989	w[67c23] P{w[+mC]=lacW}Rala[G0373]/FM7c	sn[-]	n.d.
11991	w[67c23] P{w[+mC]=lacW}l(1)G0375[G0375]/FM7c	sn[+]	g[-]
11992	P{w[+mC]=lacW}l(1)G0377[G0377] w[67c23]/FM7c	sn[-]	n.d.
11994	w[67c23] P{w[+mC]=lacW}Trxr-1[G0379]/FM7c	sn[-]	n.d.
11995	w[67c23] P{w[+mC]=lacW}ras[G0380b]/FM7c	sn[-]	n.d.
11996	w[67c23] P{w[+mC]=lacW}e(y)3[G0381]/FM7c	sn[-]	n.d.

11998	P{w[+mC]=lacW}Pgd[G0385] w[67c23]/FM7c	sn[-]	n.d.
11999	w[67c23] P{w[+mC]=lacW}ctp[G0387]/FM7c	sn[-]	n.d.
12001	P{w[+mC]=lacW}skpA[G0389] w[67c23]/FM7c	sn[-]	n.d.
12002	w[67c23] P{w[+mC]=lacW}ras[G0391]/FM7c	sn[-]	n.d.
12003	w[67c23] P{w[+mC]=lacW}l(1)G0392[G0392]/FM7c	sn[-]	n.d.
12004	w[67c23] P{w[+mC]=lacW}su(f)[G0393]/FM7c	sn[-]	n.d.
12007	P{w[+mC]=lacW}br[G0401] w[67c23]/FM7c	sn[-]	n.d.
12008	w[67c23] P{w[+mC]=lacW}Moe[G0404]/FM7c	sn[-]	n.d.
12009	w[67c23] P{w[+mC]=lacW}dome[G0405]/FM7c	sn[-]	n.d.
12010	P{w[+mC]=lacW}l(1)G0406[G0406] w[67c23]/FM7c	sn[-]	n.d.
12011	w[67c23] P{w[+mC]=lacW}Hsc70-3[G0407]/FM7c	sn[-]	n.d.
12012	w[67c23] P{w[+mC]=lacW}l(1)G0411[G0411]/FM7c	sn[-]	n.d.
12014	w[67c23] P{w[+mC]=lacW}l(1)G0414[G0414]/FM7c	sn[-]	n.d.
12015	w[67c23] P{w[+mC]=lacW}Moe[G0415]/FM7c	sn[-]	n.d.
12016	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0416]/FM7c	sn[-]	n.d.
12017	w[67c23] P{w[+mC]=lacW}Moe[G0417a] P{lacW}G0417b, l(1)G0417[G0417]/FM7c	sn[-]	n.d.
12018	w[67c23] P{w[+mC]=lacW}l(1)G0419[G0419]/FM7c	sn[-]	n.d.
12019	w[67c23] P{w[+mC]=lacW}Act5C[G0420]/FM7c	sn[-]	n.d.
12021	w[67c23] P{w[+mC]=lacW}mew[G0429]/FM7c	sn[-]	n.d.
12022	w[67c23] P{w[+mC]=lacW}l(1)G0430[G0430]/FM7c	sn[-]	n.d.
12023	P{w[+mC]=lacW}Ns3[G0431] w[67c23]/FM7c	sn[-]	n.d.
12024	w[67c23] P{w[+mC]=lacW}ctp[G0432]/FM7c	sn[-]	n.d.
12025	w[67c23] P{w[+mC]=lacW}l(1)G0433[G0433]/FM7c	sn[+]	g[-]
12027	w[67c23] P{w[+mC]=lacW}ras[G0436]/FM7c	sn[-]	n.d.
12030	w[67c23] P{w[+mC]=lacW}dome[G0441]/FM7c	sn[-]	n.d.
12054	w[67c23] P{w[+mC]=lacW}ras[G0002]/FM7c	sn[-]	n.d.
12098	P{w[+mC]=lacW}elav[G0031] arg[G0031] w[67c23]/FM7c	sn[-]	n.d.
12099	w[67c23] P{w[+mC]=lacW}Fas2[G0032]/FM7c	sn[-]	n.d.
12113	w[67c23] P{w[+mC]=lacW}l(1)G0087[G0087]/FM7c	sn[-]	n.d.
12120	w[67c23] P{w[+mC]=lacW}G0091a P{lacW}mys[G0091b]/FM7c	sn[-]	n.d.
12125	w[67c23] P{w[+mC]=lacW}fs(1)h[G0093]/FM7c	sn[-]	n.d.
12131	w[67c23] P{w[+mC]=lacW}Rpt4[G0114]/FM7c	sn[-]	n.d.
12139	w[67c23] P{w[+mC]=lacW}Grip91[G0122]/FM7c	sn[-]	n.d.
12141	w[67c23] P{w[+mC]=lacW}Top1[G0134]/FM7c	sn[-]	n.d.
12147	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0176]/FM7c	sn[-]	n.d.
12149	P{w[+mC]=lacW}G0184a w[67c23] P{lacW}Rph[G0184b], l(1)G0184[G0184]/FM7c	sn[-]	n.d.
12153	w[67c23] P{w[+mC]=lacW}l(1)G0185[G0185]/FM7c	sn[-]	n.d.
12228	w[67c23] P{w[+mC]=lacW}ctp[G0204]/FM7c	sn[-]	n.d.
12230	w[67c23] P{w[+mC]=lacW}l(1)G0214[G0214]/FM7c	sn[-]	n.d.
12231	w[67c23] P{w[+mC]=lacW}vfl[G0225]/FM7c, B[+]	sn[-]	n.d.
12232	P{w[+mC]=lacW}G0226a w[67c23] P{lacW}G0226b, l(1)G0226[G0226]/FM7c	sn[-]	n.d.
12233	w[67c23] P{w[+mC]=lacW}Vap-33A[G0231]/FM7c	sn[-]	n.d.
12235	w[67c23] P{w[+mC]=lacW}l(1)G0279[G0279]/FM7c	sn[-]	n.d.
12236	w[67c23] P{w[+mC]=lacW}Trf2[G0295]/FM7c	sn[-]	n.d.
12237	w[67c23] P{w[+mC]=lacW}Rip11[G0297]/FM7c	sn[-]	n.d.
12238	w[67c23] P{w[+mC]=lacW}p115[G0306]/FM7c	sn[-]	n.d.
12239	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0308]/FM7c	sn[-]	n.d.
12240	P{w[+mC]=lacW}elav[G0319] arg[G0319] w[67c23]/FM7c	sn[-]	n.d.
12241	w[67c23] P{w[+mC]=lacW}rap[G0326]/FM7c	sn[-]	n.d.

12244	w[67c23] P{w[+mC]=lacW}l(1)G0344[G0344]/FM7c	sn[-]	n.d.
12246	w[67c23] P{w[+mC]=lacW}Smox[G0348]/FM7c	sn[-]	n.d.
12247	w[67c23] P{w[+mC]=lacW}dm[G0359]/FM7c	sn[-]	n.d.
12249	w[67c23] P{w[+mC]=lacW}l(1)G0366[G0366]/FM7c	sn[-]	n.d.
12250	w[67c23] P{w[+mC]=lacW}l(1)G0376[G0376]/FM7c	sn[-]	n.d.
12251	w[67c23] P{w[+mC]=lacW}Ant2[G0386] sesB[G0386]/FM7c	sn[-]	n.d.
12252	w[67c23] P{w[+mC]=lacW}Nrg[G0413]/FM7c	sn[-]	n.d.
12253	P{w[+mC]=lacW}l(1)G0422[G0422] w[67c23]/FM7c	sn[-]	n.d.
12254	w[67c23] P{w[+mC]=lacW}Trf2[G0425]/FM7c	sn[-]	n.d.
12255	w[67c23] P{w[+mC]=lacW}l(1)G0442[G0442]/FM7c	sn[-]	n.d.
12256	w[67c23] P{w[+mC]=lacW}mew[G0443]/FM7c	sn[-]	n.d.
12259	P{w[+mC]=lacW}l(1)G0451[G0451] w[67c23]/FM7c	sn[-]	n.d.
12260	w[67c23] P{w[+mC]=lacW}l(1)G0459[G0459]/FM7c	sn[-]	n.d.
12261	w[67c23] P{w[+mC]=lacW}l(1)G0148[G0461]/FM7c	sn[-]	n.d.
12262	w[67c23] P{w[+mC]=lacW}l(1)G0464[G0464]/FM7c	sn[-]	n.d.
12265	w[67c23] P{w[+mC]=lacW}nej[G0470]/FM7c	sn[-]	n.d.
12268	P{w[+mC]=lacW}phl[G0475] w[67c23]/FM7c	sn[-]	n.d.
12269	w[67c23] P{w[+mC]=lacW}Trxr-1[G0477]/FM7c	sn[-]	n.d.
12270	w[67c23] P{w[+mC]=lacW}G0478a P{lacW}Clic[G0478b], l(1)G0478[G0478]/FM7c	sn[-]	n.d.
12271	w[67c23] P{w[+mC]=lacW}sd[G0483]/FM7c	sn[-]	n.d.
12272	w[67c23] P{w[+mC]=lacW}baz[G0484]/FM7c	sn[-]	n.d.
12275	w[67c23] P{w[+mC]=lacW}Nrg[G0488b]/FM7c	sn[-]	n.d.
12276	w[67c23] P{w[+mC]=lacW}schlank[G0489]/FM7c	sn[-]	n.d.
12277	w[67c23] P{w[+mC]=lacW}l(1)G0490[G0490]/FM7c	sn[-]	n.d.
12278	w[67c23] P{w[+mC]=lacW}l(1)G0007[G0491]/FM7c	sn[-]	n.d.
12279	w[67c23] P{w[+mC]=lacW}l(1)G0493[G0493]/FM7c	sn[-]	n.d.
12280	w[67c23] P{w[+mC]=lacW}fs(1)h[G0495]/FM7c	sn[-]	n.d.
12282	P{w[+mC]=lacW}east[G0500] w[67c23]/FM7c	sn[-]	n.d.
12283	w[67c23] P{w[+mC]=lacW}Rala[G0501]/FM7c	sn[-]	n.d.
12287	w[67c23] P{w[+mC]=lacW}flw[G0172]/FM7c	sn[-]	n.d.
12288	w[67c23] P{w[+mC]=lacW}Rpt4[G0227]/FM7c	sn[-]	n.d.
12289	w[67c23] P{w[+mC]=lacW}l(1)G0236[G0236]/FM7c	sn[-]	n.d.
12290	w[67c23] P{w[+mC]=lacW}Dd[G0269]/FM7c	sn[-]	n.d.
12291	P{w[+mC]=lacW}G0280a w[67c23] P{lacW}G0280b, l(1)G0280[G0280]/FM7c	sn[-]	n.d.
12292	P{w[+mC]=lacW}l(1)G0193[G0327] w[67c23]/FM7c	sn[-]	n.d.
12295	w[67c23] P{w[+mC]=lacW}l(1)G0372[G0372]/FM7c	sn[-]	n.d.
12296	P{w[+mC]=lacW}l(1)G0399[G0399] w[67c23]/FM7c	sn[-]	n.d.
12298	w[67c23] P{w[+mC]=lacW}l(1)G0435[G0435]/FM7c	sn[-]	n.d.
12299	P{w[+mC]=lacW}sta[G0448] w[67c23]/FM7c	sn[-]	n.d.
12300	w[67c23] P{w[+mC]=lacW}l(1)G0453[G0453]/FM7c	sn[-]	n.d.
12301	w[67c23] P{w[+mC]=lacW}dlg1[G0456]/FM7c	sn[-]	n.d.
12302	w[67c23] P{w[+mC]=lacW}l(1)G0462[G0462]/FM7c	sn[-]	n.d.
12303	w[67c23] P{w[+mC]=lacW}CkIalpha[G0492]/FM7c	sn[-]	n.d.
12817	w[1118] P{w[+mGT]=GT1}CG1789[BG02603]/FM7c	sn[+]	g[-]
12834	w[1118] P{w[+mGT]=GT1}Inx2[BG02429]/FM7c	sn[+]	g[-]
12839	w[1118] P{w[+mGT]=GT1}mRpL33[BG01040]/FM7c	sn[+]	g[-]
13310	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}beta-Spec[KG02312]/FM7c	sn[+]	g[-]
13350	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}Ntf-2[KG03852]/FM7c	sn[+]	g[-]
13477	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-	sn[+]	g[-]

	P}CG3560[KG02424]/FM7c		
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}sesB[KG04431]		
13597	Ant2[KG04431]/FM7c	sn[+]	g[+]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
13742	P}Rip11[KG02485]/FM7c	sn[+]	g[-]
13899	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}sw[KG05547]/FM7c	sn[+]	g[-]
14192	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}KG03828/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14291	P}ctp[KG06321]/FM7c	sn[+]	g[-]
	y[1] w[67c23] P{y[+t7.7]=Mae-UAS.6.11}VhaAC39-		
14394	1[GG01465]/FM7c	sn[-]	n.d.
14395	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}KG01741/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14424	P}CG5599[KG02236]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14436	P}kdn[KG04873]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14460	P}Kap3[KG05921]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14505	P}Usp7[KG06814]/FM7c	sn[+]	g[-]
14592	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}KG04566/FM7c	sn[+]	g[-]
14605	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}N[KG06588]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14611	P}mod(r)[KG07005]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14904	P}l(1)G0148[KG03467]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14936	P}sno[KG08094]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14953	P}brk[KG08470]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
14966	P}Pgd[KG08676]/FM7c	sn[+]	g[-]
	y[1] w[67c23] P{w[+mC]		
15037	y[+mDint2]=EPgy2}l(1)G0255[EY00709]/FM7c	sn[+]	g[-]
15049	y[1] w[67c23] P{w[+mC] y[+mDint2]=EPgy2}EY00929/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
15098	P}CG43736[KG02072]/FM7c	sn[+]	g[-]
15103	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-P}KG05904/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
15159	P}HDAC4[KG09091]/FM7c	sn[+]	g[-]
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
15203	P}RpL22[KG09650]/FM7c	sn[+]	g[-]
15425	y[1] P{w[+mC] y[+mDint2]=EPgy2}EY02603 w[67c23]/FM7c	sn[+]	g[-]
15713	y[1] P{w[+mC] y[+mDint2]=EPgy2}EY04066 w[67c23]/FM7c	sn[-]	n.d.
	y[1] w[67c23] P{w[+mC]		
15725	y[+mDint2]=EPgy2}Clic[EY04209]/FM7c	sn[+]	g[-]
	y[1] P{w[+mC] y[+mDint2]=EPgy2}east[EY05235]		
15791	w[67c23]/FM7c	sn[+]	g[-]
	y[1] P{w[+mC] y[+mDint2]=EPgy2}Atf3[EY02562]		
15872	w[67c23]/FM7c	sn[+]	g[-]
	y[1] w[67c23] P{w[+mC]		
15877	y[+mDint2]=EPgy2}sicily[EY02706]/FM7c	sn[+]	g[-]
15899	y[1] w[67c23] P{w[+mC]	sn[+]	g[-]

	y[+mDint2]=EPgy2}shf[EY03173]/FM7c		
	y[1] P{y[+mDint2] w[BR.E.BR]=SUPor-		
16469	P}Gbeta13F[KG08410]/FM7c	sn[+]	g[-]
18269	w[1118] PBac{w[+mC]=RB}CHOp24[e04526]/FM7c	sn[-]	n.d.
18270	w[1118] PBac{w[+mC]=RB}CG3527[e04544]/FM7c	sn[-]	n.d.
18272	w[1118] PBac{w[+mC]=RB}Aats-lys[e04554]/FM7c	sn[-]	n.d.
	w[1118] PBac{w[+mC]=RB}pot[e04564]		
18273	PBac{RB}fw[e04564]/FM7c	sn[-]	n.d.
18275	w[1118] PBac{w[+mC]=RB}l(1)10Bb[e04588]/FM7c	sn[-]	n.d.
18276	PBac{w[+mC]=RB}arm[e04595] w[1118]/FM7c	sn[-]	n.d.
18728	w[1118] PBac{w[+mC]=WH}Arp2[f04069]/FM7c	sn[+]	g[-]
18729	w[1118] PBac{w[+mC]=WH}Lim1[f04087]/FM7c	sn[-]	n.d.
	w[1118] PBac{w[+mC]=WH}CG9940[f04110]		
18730	NnaD[f04110]/FM7c	sn[-]	n.d.
18731	w[1118] PBac{w[+mC]=WH}xmas-2[f04114]/FM7c	sn[-]	n.d.
	PBac{w[+mC]=WH}cin[f05298] CG42376[f05298]		
18854	w[1118]/FM7c	sn[-]	n.d.
19116	w[1118]/FM7c; PBac{w[+mC]=WH}CG32112[f07936]	sn[-]	n.d.
19132	w[1118]/FM7c; PBac{w[+mC]=WH}Eip93F[f08111]	sn[-]	n.d.
19184	w[1118] P{w[+mC]=XP}Dok[d02937]/FM7c	sn[-]	n.d.
	y[1] w[67c23] P{y[+t7.7]=Mae-		
19989	UAS.6.11}lic[GG01785]/FM7c	sn[-]	n.d.
	y[1] w[67c23] P{w[+mC]		
21111	y[+mDint2]=EPgy2}tlk[EY14954]/FM7c	sn[+]	g[-]
	y[1] w[67c23] P{w[+mC] y[+mDint2]=EPgy2}Tcp-		
21199	1zeta[EY16253]/FM7c	sn[+]	g[-]
21429	y[1] w[67c23] P{w[+mC] y[+mDint2]=EPgy2}EY14160/FM7c	sn[+]	g[-]
21875	w[1118] P{w[+mC]=EPg}fu[HP10439]/FM7c	sn[-]	n.d.
	y[1] w[67c23] P{w[+mC] y[+mDint2]=EPgy2}rdgB[EY20869]		
22434	CG32625[EY20869]/FM7c	sn[+]	g[+]
23295	Df(1)FDD-0024486, w[1118]/FM7c	sn[-]	n.d.
23296	Df(1)FDD-0230908, w[1118]/FM7c	sn[-]	n.d.
	ph-d[504] ph-p[504] w[*]		
24162	P{w[+mW.hs]=FRT(w[hs])}101/FM7c	sn[-]	n.d.
25095	Rala[EE1]/FM7c	sn[-]	n.d.
25712	w[*] sqh[AX3] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
26815	y[1] w[*] btd[XA]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	sn[-]	n.d.
27408	Df(1)FDD-0230186, w[1118]/FM7c	sn[-]	n.d.
27409	Df(1)FDD-0089480, w[1118]/FM7c	sn[-]	n.d.
27411	Df(1)FDD-0369033, w[1118]/FM7c	sn[-]	n.d.
27415	Df(1)FDD-0369024, w[1118]/FM7c	sn[+]	g[-]
27906	In(1)drd[x1], drd[x1]/FM7c	sn[-]	n.d.
28813	N[55e11] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
28870	w[*] P+PBac{w[+mC]=XP5.WH5}RhoGAPp190[1]/FM7c	sn[-]	n.d.
32007	sgg[M11] w[*] f[36a]/FM7c, P{ftz/lacC}YH1, sn[+]	sn[+]	n.d.
32183	w[1] mxc[G43]/FM7c	sn[-]	n.d.
33828	Df(1)BK18/FM7c/Dp(1;Y)y[+]; Dp(1;4)rK20, f[+]/sv[spa-pol]	sn[-]	n.d.
	Df(1)815-6, y[1] cv[1] v[1] f[1]/FM7c/Dp(1;Y)y[+];		
33829	Dp(1;4)rK20, f[+]/+	sn[-]	n.d.
33833	sno[EF531]/FM7c	sn[-]	n.d.
34040	w[*] nej[0.3]/FM7c	sn[-]	n.d.
	w[*] ct[C145] P{ry[+t7.2]=neoFRT}19A/FM7c;		
36496	P{w[+mC]=UAS-ct.P}2, P{w[+mC]=UAS-	sn[-]	n.d.

	mCD8::GFP.L}LL5/CyO		
	Lim1[E9] P{ry[+t7.2]=neoFRT}19A/FM7c;		
	P{w[+mW.hs]=GawB}GH146 P{w[+mC]=UAS-		
36499	mCD8::GFP.L}LL5/CyO	sn[-]	n.d.
	Df(1)f08066-f07791, w[1118] PBac{w[+mC]=WHr}f08066-		
39614	f07791/FM7c	sn[-]	n.d.
39615	Df(1)e00904-d02459, w[1118]/FM7c	sn[-]	n.d.
41112	w[*] wuho[56]/FM7c	sn[+]	g[-]
41460	w[*] wuho[7]/FM7c	sn[+]	g[-]
41811	y[1] w[*] VhaAC39-1[FY38] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
41812	y[1] w[*] VhaAC39-1[FZ29] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
	usp[5] w[*] P{w[+mC]=UAS-mCD8::GFP.L}Ptp4E[LL4]		
44383	P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
	y[1] w[*] P{w[+mC]=UAS-mCD8::GFP.L}Ptp4E[LL4]		
44384	Smox[MB388] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
	Df(1)YO17, y[1] w[67c23] mthl1[YO17]/FM7c,		
51320	P{w[+mC]=GAL4-twi.G}108.4, P{UAS-2xEGFP}AX	sn[-]	n.d.
51323	y[1] M{vas-Cas9}ZH-2A w[1118]/FM7c	sn[-]	n.d.
51334	y[1] w[*] Rbcn-3A[FV10] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51335	y[1] w[*] Rbcn-3A[FE6] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51336	y[1] Rbcn-3B[FK39] w[*] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51337	y[1] Rbcn-3B[GA20] w[*] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51338	y[1] w[*] flw[FP41] g[2] f[1] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51339	y[1] w[*] PI4KIIIalpha[FQ88] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51340	y[1] w[*] PI4KIIIalpha[GS27] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51341	y[1] w[*] Crag[CJ101] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51342	y[1] w[*] AP-1-2beta[GH73] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
51343	y[1] w[*] shi[FL54] P{ry[+t7.2]=neoFRT}19A/FM7c	sn[-]	n.d.
	Df(1)l11, y[1] w[1] v[1] f[1] mal[1] su(f)[1]/FM7c,		
	flam[FM7c]/Dp(1;Y)y[+]mal[126]; P{w[+mC]=gypsy-		
53723	lacZ.p12}3	sn[+]	g[-]
55125	w[1] CycD[1]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	sn[+]	n.d.
55126	w[1] CycD[2]/FM7c, P{ry[+t7.2]=ftz/lacC}YH1	sn[+]	n.d.

Chapter 5: *Third* chromosome balancer inversions disrupt protein-coding genes and influence distal recombination events in *Drosophila melanogaster*

ABSTRACT

Balancer chromosomes are multiply inverted chromosomes that suppress meiotic crossing over and prevent the recovery of crossover products. Balancers are commonly used in *Drosophila melanogaster* to maintain deleterious alleles and in stock construction. They exist for all three major chromosomes, yet the molecular location of the breakpoints and the exact nature of many of the mutations carried by the 2^{nd} and 3^{rd} chromosome balancers has not been available. Here, we precisely locate eight of 10 of the breakpoints on the 3^{rd} chromosome balancer *TM3*, six of eight on *TM6*, and nine of 11 breakpoints on *TM6B*. We find that one of the inversion breakpoints on *TM3* bisects the highly conserved tumor suppressor gene *p53*, a finding that may have important consequences for a wide range of studies in Drosophila. We also identify evidence of single and double crossovers between several *TM3* and *TM6B* balancers and their normal-sequence homologs that have created genetic diversity among these chromosomes. Furthermore, knowledge of the precise location of the most distal *TM3* breakpoint allows us to investigate how close to the inversion breakpoint exchange may occur, providing insight into the distance over which inversions suppress exchange. Overall, this work demonstrates the practical importance of precisely identifying the position of inversion breakpoints of balancer chromosomes and characterizing the mutant alleles carried by them.

INTRODUCTION

Balancer chromosomes are multiply rearranged chromosomes that are extensively used in *Drosophila melanogaster* for tasks such as stock construction and the maintenance of recessive deleterious alleles in populations (Ashburner *et al.* 2005). Balancers work by suppressing meiotic recombination, by creating recombinant chromatids that will not segregate properly during the first meiotic division (Novitski and Braver 1954) or, in the case of pericentric inversions, by creating recombinants that carry duplications or deficiencies large enough to result in zygotic lethality. While all balancer chromosomes carry easily scored dominant marker alleles that allow for visual identification of flies carrying the balancer, most balancers also carry recessive lethal mutations that prevent the balancer from becoming homozygous in stock (Lindsley and Zimm 1992; Ashburner *et al.* 2005).

A variety of balancers are available for the *X*, *2*nd, and *3*rd chromosomes in Drosophila, and they have become increasingly effective as the number of inversions has increased and as visible markers and recessive lethal or sterile alleles have been added. For example, *First multiple one* (*FM1*), an *X* chromosome balancer, improved upon earlier single-inversion balancers such as *In(1)dl-49*, *In(1)sc*, and *ClB* by combining the *In(1)dl-49* and *In(1)sc* inversions into one chromosome (Lindsley and Zimm 1992; Ashburner *et al.* 2005). Further improvements generated a series of FM balancers, and similar series exist for the *2*nd (Second multiple; SM) and *3*rd (Third multiple; TM) chromosomes (Lindsley and Zimm 1992). The current study will focus on the *3*rd chromosome balancers *TM3*, *TM6*, and *TM6B*.

TM3 was created in the late 1950s by repeated X-raying of a chromosome marked with *kni*$^{ri-1}$, *p*^{p}, *sep*^{1}, *Ubx*$^{bx-34e}$ and *e*^{1} and carrying two inversions, *In(3LR)sep* (65D2-3;85F2-4) and *In(3R)C* (92D1-E1;100F2-3). The irradiation superimposed three additional inversions on this chromosome, creating a balancer with 10 total inversions (Lewis 1960) **(Figure 5.1)**. Tinderholt (1960) introduced the dominant markers *Serrate* (*Ser*) and *Stubble* (*Sb*) into inverted regions of this chromosome by double crossing over, relying on the increased recombination created by the so-called interchromosomal effect (Schultz and Redfield 1951; Ramel 1966) to obtain these double crossovers (DCOs). Specifically, he performed this synthesis in a female heterozygous for three balancers to increase the likelihood of recombination within the desired inversions (Tinderholt 1960). In doing so, he demonstrated that segments could be swapped into *TM3*—even if such events were uncommon.

The progenitor chromosome that was X-rayed to produce *TM3* also carried *Dp(1;3)sc*$^{260-20}$, an aberration that replaced the tip of chromosome *3L* with the tip of an *X* chromosome carrying a wild-type allele of the *yellow* (*y*) gene (Sutton 1943). However, this *y*$^{+}$ marker was frequently lost by a single crossover event between *TM3* and normal-sequence chromosomes in the region distal to the 65D inversion breakpoint; consequently, most *TM3* chromosomes now carry a normal *3L* tip (Shearn 1980). This is one of several observations indicating that the relatively large uninverted region distal to 65D undergoes frequent exchange events—even though recombination is largely suppressed in regions proximal to 65D.

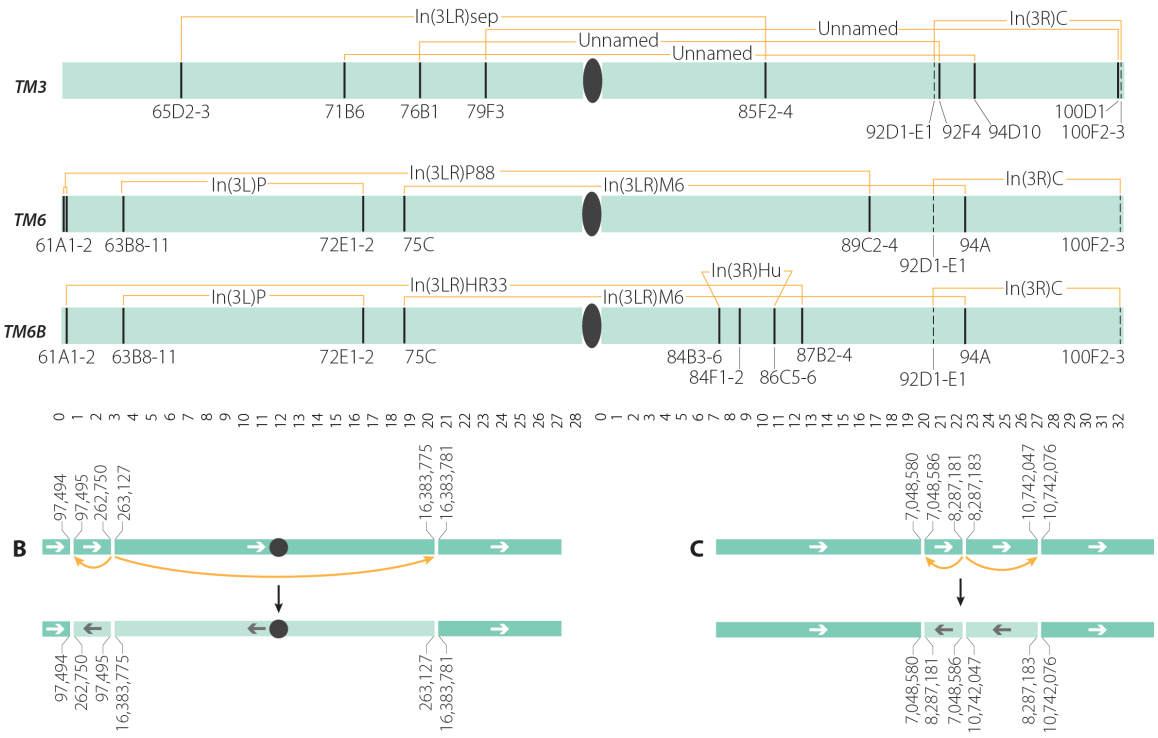

Figure 5.1: *TM3*, *TM6*, and *TM6B* inversion breakpoints.
Black circles indicate centromeres and left-facing arrows indicate an inverted segment. **(A)** The inversions carried by the 3^{rd} chromosome balancers *TM3*, *TM6*, and *TM6B.* Breakpoints that have been molecularly identified are shown as solid lines; those that are estimates are shown as dashed lines; numbers are cytological bands of breakpoints given in Lindsley and Zimm (1992). **(B)** The *In(3LR)P88* (61A1-2;89C2-4) rearrangement on *TM6* is a previously unreported three-breakpoint rearrangement with a breakpoint at *3L*:263,127–263,132 that bisects the gene *Tudor-SN*, while the breakpoint at *3R*:16,383,781 bisects *spineless*, an allele previously reported to be carried by this chromosome (Duncan *et al.* 1998), and the breakpoint at *3L*:97,494 is intergenic. **(C)** In the *In(3R)Hu* (84B1;84F4;86C7-8) three-breakpoint rearrangement on *TM6B*, the breakpoint at *3R*:8,287,181 bisects the non-coding RNA *CR44318* while the *3R*:10,742,076 breakpoint bisects *TkR86C*. We hypothesize that the breakpoint at 3R:7,048,580 causes the $Antp^{Hu}$ phenotype.

TM6 was created by X-ray mutagenesis of a chromosome marked with $Ubx^{bx\text{-}34e}$ and e^{1} and carrying three preexisting inversions: *In(3L)P* (63C;72E1-2) lying inside *In(3LR)P88* (61A;89CD) with *In(3R)C* (92D1-E1;100F2-3) to the right **(Figure 5.1)**. Irradiation resulted in an additional inversion, *In(3LR)M6*, between bands 75C and 94A (Lindsley and Zimm 1992). *TM6B* was built from *TM6* by replacing the left breakpoint of *In(3LR)P88* with the left end of *In(3LR)HR33* (61A1-2;87B) (Ashburner 1972) by a single crossover **(Figure 5.1)**. The three-breakpoint rearrangement *In(3R)Hu* (84B1;84F4;86C7-8) (Hazelrigg and Kaufman 1983) was

carried onto the recombinant chromosome, along with the left end of *In(3LR)HR33* from a double-aberration progenitor. An internal segment spanning the right breakpoint of *In(3LR)P88* was then replaced with a segment spanning the right breakpoint of *In(3LR)HR33* by double crossing over. Finally, the dominant *Tubby* (Tb^1) marker was added by a DCO event within an inverted segment near the right end of the newly created *TM6B* (Craymer 1981; 1984; Lindsley and Zimm 1992).

Because balancers are widely used in Drosophila experiments, sometimes as heterozygous controls, it is informative for the community to determine the exact position of their breakpoints and the nature of the alleles carried by them. A recent study reported rare DCO events between the *X* chromosome balancer *FM7* and its normal sequence homologs that were selected for because they conferred an advantage to flies carrying the recombinant chromosome (Miller, Cook, *et al.* 2016). A similar whole-genome analysis of commonly used autosomal balancers has not yet been conducted.

Here, we use whole genome sequencing to identify all but one of the inversions on the *TM3*, *TM6*, and *TM6B* balancer chromosomes **(Figure 5.1)**. Surprisingly, we find that the breakpoint at 94D on *TM3* splits the highly conserved tumor suppressor gene *p53* in half, demonstrating that any stock balanced with *TM3* is heterozygous for a *p53* loss-of-function allele. We also find evidence of single and double crossover events on more than half of the *TM3* chromosomes sampled and on one *TM6B* chromosome and are able to estimate the distance over which inversions suppress exchange by examining single crossover events that occur in an unbalanced region of the *TM3* chromosome. These findings demonstrate that,

similar to the *X* chromosome balancer *FM7*, sequence diversity exists among *3*rd chromosome balancers and suggests that this variation may influence experimental outcomes.

RESULTS

Using whole-genome sequencing, we precisely identified eight of the 10 breakpoints on *TM3*, six of eight breakpoints on *TM6*, and nine of 11 breakpoints on *TM6B* **(Figure 5.1A, Table 5.1)**. The three balancers share an inversion, *In(3R)C*, between cytological bands 92D1-E1 and 100F2-3 (Sturtevant 1913; Muller 1918) that we were unable to accurately position because its location near the telomere suggests that it most likely involves highly repetitive sequences. Note that throughout the manuscript we refer to breakpoints by the names of the inversions that created them and the historical cytological bands reported in Lindsley and Zimm (1992) and not the estimated cytological bands that are available on FlyBase or the UCSC genome browser.

Because autosomal balancers carry recessive lethal mutations, the recovery of homozygous progeny for sequencing is not feasible. To circumvent this problem, we crossed males from each *TM3* and *TM6B* balancer stock to females from the ISO-1 stock, which was used to construct the Drosophila reference genome, and recovered heterozygous individuals for sequencing (see Methods). We confirmed breakpoints by two methods: first, we whole-genome sequenced large-insert (2–12 kb) library preparations for two *TM3* and one *TM6B* stocks (see Methods); and second, we PCR and Sanger sequenced all identified breakpoints on *TM3* and *TM6*, and selected breakpoints on *TM6B* **(Table 5.S1)**.

***Third* chromosome balancer breakpoints disrupt protein-coding genes**

After identifying the exact position of each inversion breakpoint, we found that the breakpoints on *TM3* altered six characterized (*Glut4EF*, *FucTA*, *p53*, *ms(3)76Ba*, *Lrrk*, and *kek6*) and two uncharacterized (*CG32206* and *CG14459*) protein-coding genes **(Table 5.1)**. Perhaps most surprisingly, we observed that the 94D inversion breakpoint on *TM3* at *3R*:23,050,763–23,050,764 bisects the fifth intron of the highly-conserved tumor suppressor *p53* (Jin *et al.* 2000) and affects all reported *p53* isoforms. We also confirmed that the allele $Glut4EF^{TM3}$ is caused by the inversion at 85F2 on *TM3*, as reported by Yazdani and colleagues (2008) **(Table 5.2)**. Finally, we found that the y^{+} *X* chromosome fragment originally present on *TM3* (Lewis 1960; Shearn 1980) was the result of a break of the *X* chromosome at *X*:416,997 and subsequent attachment to the 3^{rd} chromosome at *3L*:149,709, in agreement with its original isolation as a reciprocal translocation affecting the *X*-linked *scute* gene (Sutton 1943). This rearrangement deletes or disrupts 10 protein-coding and eight noncoding RNA genes from the 3^{rd} chromosome in the distal 150-kb interval of *TM3*.

Balancer	Inversion	Chr	Reported bands[1]	5' Break	3' Break	Duplication (+) /Deletion (-)	Affected gene/region
TM3	*In(3LR)sep*	*3L*	65D2-3	6,925,034	6,926,125	- 1,090	intergenic
		3R	85F2-4	9,943,831	9,944,040	- 208	*Glut4EF*
TM3	Unnamed	*3L*	71B6	15,150,269	15,150,272	- 2	*FucTa*
		3R	94D10	23,050,763	23,050,764	0	*p53*
TM3	Unnamed	*3L*	76B1	19,386,273	19,388,151	- 1,877	*CG32206, ms(3)76Ba*
		3R	92F4	20,637,930	20,637,930	+ 1	*Lrrk*
TM3	Unnamed	*3L*	79F3	22,637,876	22,637,952	- 75	*CG14459*
		3R	100D1	31,653,695	31,653,707	- 11	*kek6*
TM3, TM6,	*In(3R)C*	*3R*	92D1-E1	unknown	unknown	-	unknown
TM6B		*3R*	100F2-3	unknown	unknown	-	unknown
TM6	*In(3LR)P88*	*3L*	61A1-2	97,494	97,495	0	intergenic
		3L	61A1-2	263,127	263,132	- 4	*Tudor-SN*
		3R	89C2-4	16,383,781	16,383,775	+ 7	*ss*
TM6, TM6B	*In(3LR)M6*	*3L*	75C	18,693,657	18,693,663	- 5	*CR43987*
		3R	94A	22,393,827	22,393,828	0	*CG13857*
TM6, TM6B	*In(3L)P*	*3L*	63B8-11	3,173,046	3,173,053	- 6	*CG14964*
		3L	72E1-2	16,308,841	16,308,845	- 3	intergenic
TM6B	*In(3LR)HR33*	*3L*	61A1-2	233,562	233,565	- 2	intergenic
		3R	87B2-4	12,227,473	12,227,471	+ 3	intergenic
TM6B	*In(3R)Hu*	*3R*	86C5-6	10,742,047	10,742,076	-28	*TkR86C*
		3R	84F1-2	8,287,181	8,287,183	-1	*CR44318*
		3R	84B3-6	7,048,580	7,048,586	-5	intergenic

Table 5.1. Molecular details of the *TM3*, *TM6*, and *TM6B* inversion breakpoints

[1] Reported bands are those found in Lindsley and Zimm (1992) and are not based on estimated genomic position.

Gene	Allele	Balancer(s)	Observed aberration	Previous reports
ebony	e^{1}	*TM3*, *TM6*, *TM6B*	TE (family: 412) at *3R*:21,231,832–21,231,838, 6 nt into the 2nd exon	—
Ultrabithorax	$Ubx^{bx\text{-}34e}$	*TM3*, *TM6*	TE (family: DMIS176) insertion in the first intron of *Ubx* at approximately *3R*:16,731,980	Gypsy insertion (Bender *et al.* 1983)
knirps	$kni^{ri\text{-}1}$	*TM3*	252-bp deletion at *3L*:20,707,101-20,707,352.	(Lunde 2003)
pink	p^{p}	*TM3*	1-bp deletion at *3R*:6,661,619 resulting in a frameshift	1-bp deletion at *3R*:6,661,624 (Syrzycka *et al.* 2007)
lethal (3) 89Aa	$l(3)89aA^{1}$	*TM3*	Unknown	Mapped to 89A2-89A5
ventral veins lacking	vvl^{sep}	*TM3*	Unknown	—
Stubble	Sb^{1}	*TM3*	TE (family: 412) insertion in 4th exon of *Sb* at *3R*:16,141,939-16,141,942.	TE insertion (Hammonds and Fristrom 2006)
Serrate	Ser^{1}	*TM3*	TE (family: TIRANT) insertion at *3R*:27,172,910-27,172,913 in the 3' UTR of *Ser*	TE insertion (Fleming *et al.* 1990)
Ultrabithorax	Ubx^{P15}	*TM6*	Unknown	—
Henna	Hn^{P}	*TM6*	Multiple deletions within the first intron and a G->A mutation at splice acceptor site (AG becomes AA) in the third intron of the gene.	—
spineless	ss^{aP88}	*TM6*	Gene is split by the *In(3LR)P88* (61A1-2;89C2-4) rearrangement.	Break in the transcription unit (Duncan *et al.* 1998)
Antennapedia	$Antp^{Hu}$	*TM6B*	Unknown. Phenotype may be a result of the *In(3R)HR33* triple rearrangement (Figure 3).	—
Tubby	Tb^{1}	*TM6B*	An in-frame 15-nt deletion in the 2^{nd} exon from *3R*:26,656,728-26,656,742; a 69-nt in-frame deletion of 23 amino acids from *3R*:26,657,089-26,657,157; and a T->G mutation (Ser->Ala) at *3R*:26,657,334.	—

Table 5.2. Genomic aberrations of marker and recessive lethal alleles carried by *TM3, TM6*, and *TM6B*

The breakpoints on *TM6* affected four protein-coding genes (*Tudor-SN*, *ss*, *CG13857*, *CG14964*) and one noncoding RNA (*CR43987*) **(Table 5.1)**. Using whole-genome data, we confirmed that the previously reported *spineless* allele (ss^{aP88}) on *TM6*, reported as a break in the transcription unit (Duncan *et al.* 1998), is indeed caused by the inversion at 89C4 **(Table 5.2)**. We also observed that the *In(3LR)P88* (61A;89CD) inversion on *TM6*, which had been reported to be a simple inversion of 61A to 89C, is actually a three-breakpoint rearrangement that creates a previously unknown 165-kb inversion **(Figure 5.1B, Table 5.1)**.

Finally, the *TM6B* breakpoints affect three protein-coding genes (*CG13857*, *CG14964*, *TkR86c*) and two noncoding RNAs (*CR43987*, *CR44318*) **(Table 5.1)**. We also characterized the three-breakpoint *In(3R)Hu* (84B1;84F4;86C7-8) rearrangement on *TM6B* and found that it consists of 1.2-Mb and 2.5-Mb inverted segments **(Figure 5.1C, Table 5.1)**. Based on the position of these breakpoints, we propose that the gain-of-function mutation $Antennapedia^{Hu}$ ($Antp^{Hu}$) is a regulatory mutation caused by the 84B1 inversion breakpoint that lies approximately 50 kb away from *Antp* (Thom Kauffman, personal communication).

In addition to the mutations caused by inversion breakpoints, balancer chromosomes carry a number of presumably innocuous mutations that provide visible markers for easy identification as well as recessive lethal alleles that prevent balancers from becoming homozygous in stock. Some of these markers are shared by more than one balancer—such as *ebony* (e^1), present on *TM3*, *TM6*, and *TM6B*—while others are present on only one balancer—such as *Tubby* (Tb^1), present only on *TM6B* (Table 2). The general nature of many of these alleles has been previously described (such as that a transposable element (TE) insertion in *Ultrabithorax* gives rise to the $Ubx^{bx\text{-}34e}$ allele carried by *TM3* and *TM6* (Bender *et al.* 1983), or

that a TE insertion is responsible for *Ser*^{1} on *TM3* (Fleming *et al.* 1990)), but the specific lesions that convey their respective phenotypes are unknown for most alleles. Using our whole-genome sequencing data, we were able to identify the precise nature of nine of 13 visible or recessive alleles carried by the three balancers analyzed in this study. These data are summarized in Table 2.

The *TM3* balancer allows single crossover events distal to 65D

Inversion breakpoints are known to suppress exchange in nearby regions, but the mechanism by which they do this and over what distance they act is unknown (Sturtevant and Beadle 1936; Novitski and Braver 1954). Previous work has shown that balancer chromosomes pair along their lengths with their normal sequence homologs (Gong *et al.* 2005) and that both crossover and noncrossover gene conversion events occur between balancers and their normal-sequence homologs (Blumenstiel *et al.* 2009; Miller, Cook, *et al.* 2016). Because the distal-most inversion breakpoint on the left arm of *TM3* is 6.9 Mb from the telomere (estimated cytological band 65D3), we hypothesized that single crossover events would be common in this region **(Figure 5.1A)**. Evidence of recombination within this interval would manifest as tracts of unique SNPs among *TM3* chromosomes, thus we sequenced a panel of seventeen stocks from the Bloomington Drosophila Stock Center and one laboratory stock carrying the *TM3* chromosome **(Table 5.S2)** to identify how close to the inversion breakpoint these crossovers occurred.

We saw evidence of crossing over between the telomere and the most distal *3L* inversion breakpoint in 11 of 18 *TM3* stocks **(Figure 5.2)**. Crossovers in stocks *TM3-560* and *TM3-1614* are observed as close as approximately 2 Mb from the inversion breakpoint, the first

evidence of the distance over which an inversion breakpoint may suppress exchange. We also observed that several of the crossover tracts recovered were shared among multiple stocks, highlighting the relatedness of these chromosomes. For example, stocks *TM3-560* and *TM3-1614* share identical SNPs in the 1- to 5-Mb interval of *3L* and stocks *TM3-500* and *TM3-9013* are nearly identical over 2.5–3.5 Mb in this same region **(Figure 5.2A)**.

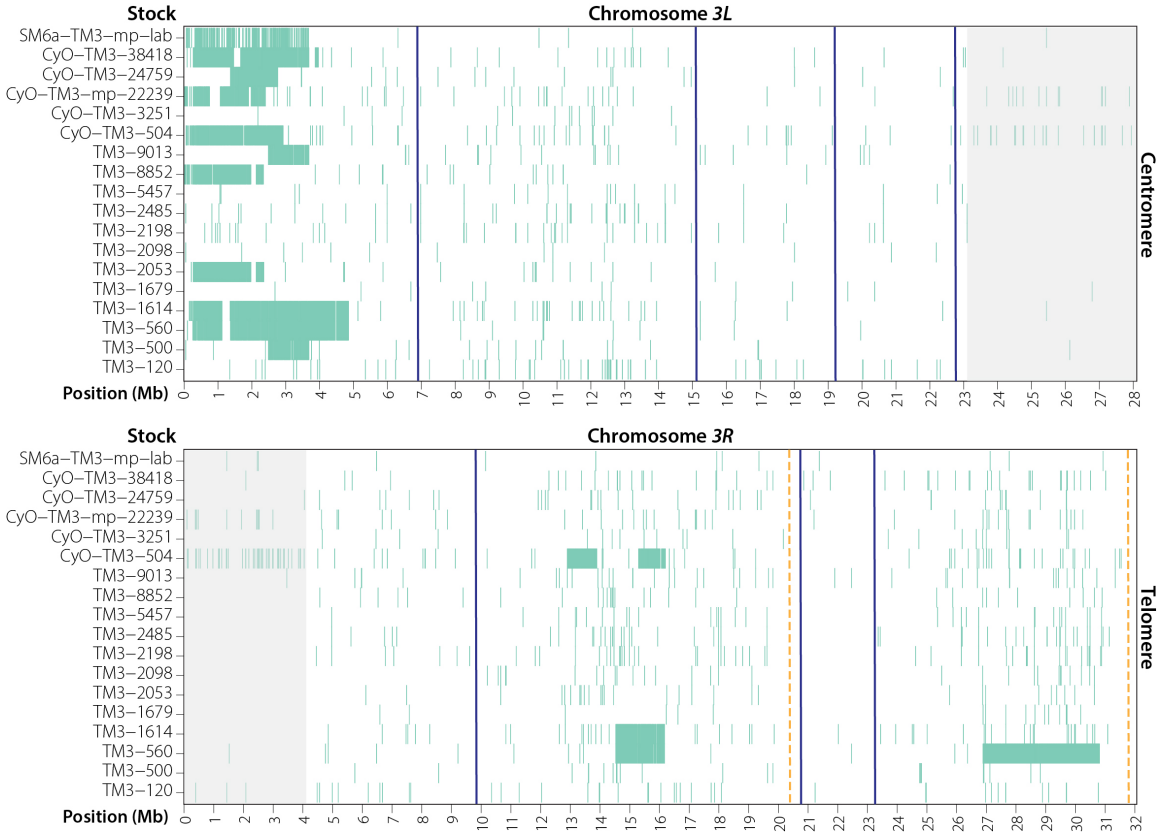

Figure 5.2. Visualizing SNPs present in five or fewer *TM3* chromosomes reveals numerous single crossover events on *3L* and several DCO events on *3R*.
Blue lines indicate the positions of inversion breakpoints whose precise location is known, orange dashed lines show the approximate position of the unidentified *In(3R)C* (92D1-E1;100F2-3) inversion breakpoints. **(A)** Single crossovers are common in the region distal to the 65D inversion breakpoint at position *3L*:6,925,034 and occur within about 2 Mb of the breakpoint. **(B)** Several DCOs are apparent on *3R*. Stocks *TM3-560* and *TM3-1614* may be versions of *TM3* before Ser^1 was added to a *TM3,* Sb^+ Ser^+ chromosome (*TM3-560*) and before Sb^1 was added to a *TM3,* Sb^+ Ser^1 chromosome (*TM3-1614*).

Double crossover events can occur on *TM3* and *TM6B*

We were also able to identify DCOs that had occurred within inverted segments on both the *TM3* and *TM6B* balancers. We found three DCO events that replaced a mutant copy of *Stubble* (Sb^1) with a wild-type copy in stocks that are phenotypically Sb^+ **(Figure 5.2B)**. Based on shared SNPs, the 1.7-Mb segment between 14.5 Mb and 16.2 Mb in *TM3-1614* and *TM3-560* appears to have originated from a single DCO in a common progenitor, while the 900-kb segment in *CyO-TM3-500* likely arose by an independent DCO event. In addition, we also found a 3.9-Mb DCO that replaced a mutant copy of *Serrate* (*Ser*) with a wild-type copy **(Figure 5.2B)**. While difficult to confirm, *TM3-560* may be similar to the original isolate of *TM3* before Ser^1 and Sb^1 were added by Tinderholt (1960) and *TM3-1614* may be the Sb^+ Ser^1 version of the chromosome after Ser^1 was added and before Sb^1 was added (Tinderholt 1960).

The *TM3* chromosome carried by the *CyO-TM3-504* stock carries a second 1-Mb DCO event near the DCO that replaced Sb^1 with Sb^+ **(Figure 5.2B)**. Analysis of this region using SnpEFF (Cingolani *et al.* 2012) finds no obvious deleterious mutations in this interval on any other *TM3* chromosome. We do, however, find a 10-kb tandem duplication within this DCO that fully duplicates *CG31157* and *CG7966*, two uncharacterized genes highly expressed in a variety of tissues, which may confer a competitive advantage to flies carrying the duplication. Interestingly, *CG7966*, which encodes a selenium-binding protein, is conserved from Drosophila to humans (SELENBP1), which makes this duplication a provocative candidate for further study.

The two presumed DCO events on *CyO-TM3-504* are also interesting because of their sizes. At 900 kb and 1 Mb, these are likely the smallest DCO events yet reported in Drosophila—even smaller than the 1.5-Mb DCO observed in a recent study (Miller, Smith, *et al.* 2016). It is

unlikely these two DCOs are the result of a single larger DCO at coordinates 12.9–16.2 Mb followed by a second DCO at coordinates 13.9–15.3 Mb, because the second DCO would have had to occur with a homologous *TM3* or *TM3* progenitor chromosome. A simpler explanation is that these were two independent DCOs.

Finally, we identified a single 1.4-Mb DCO at *3L*:9,216,999–10,625,261 on *TM6B-587* **(Figure 5.3)**. It replaces three separate frameshifting deletions in the uncharacterized genes *CG46121*, *CG16711*, and *CG32055* with wild-type copies—a potential advantage for flies carrying this chromosome. Overall, our findings provide molecular evidence that, while rare, DCO events do occur between *TM3* or *TM6B* balancers and their normal-sequence homologs.

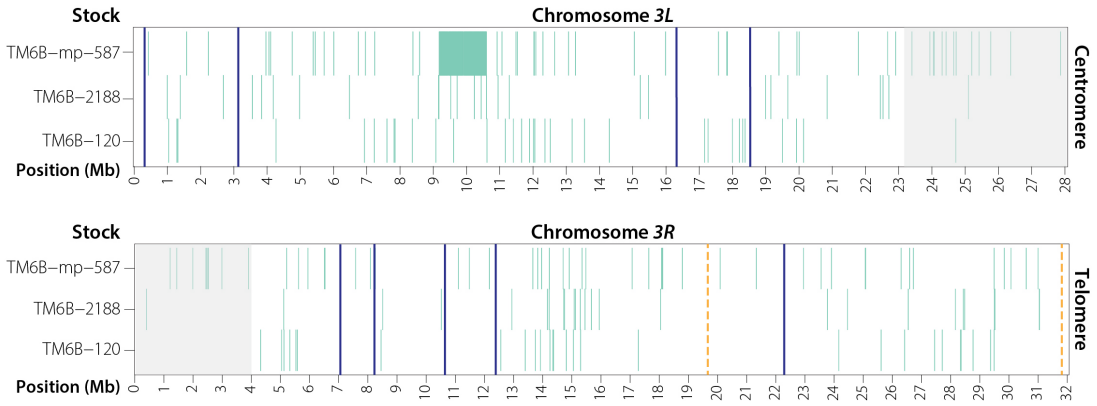

Figure 5.3. Unique SNPs present among the three *TM6B* chromosomes sequenced in this study.
Blue lines indicate the positions of inversion breakpoints, orange dashed lines indicate the approximate position of the unidentified *In(3R)C* (92D1-E1;100F2-3) inversion breakpoints. A single DCO event was recovered in stock *TM6B-587*.

DISCUSSION

We have identified the precise locations of all inversion breakpoints from the Drosophila *3rd* chromosome balancers *TM3*, *TM6*, and *TM6B* except for the *In(3R)C* (92D1-E1;100F2-3) inversion shared by all three chromosomes. Surprisingly, we find that one of the *TM3* inversion breakpoints bisects all transcripts of the tumor suppressor *p53*, with implications for a wide range of studies in Drosophila. As hypothesized, we identified evidence of single crossover events in the 6.9-Mb interval between the telomere and the most distal inversion breakpoint on *TM3* in nearly two-thirds of the stocks we sequenced. These single crossover events provide the first evidence for the distance over which inversion breakpoints can suppress meiotic exchange.

Eleven of 18 *TM3* stocks carry evidence of a recombination event between the 65D breakpoint and the telomere, with the closest exchange event occurring approximately 2 Mb from the 65D breakpoint. Do all inversion breakpoints suppress exchange in a similar way and over a similar distance? Perhaps the most instructive case is that of the *X* chromosome inversion *In(1)dl-49*. The distal-most breakpoint of the inversion lies approximately 4.9 Mb from the telomere (2 Mb closer to the telomere than the 65D breakpoint on *TM3*). Recombination in a single generation was previously measured between the distal-most breakpoint of *In(1)dl-49* and the telomere using *yellow*, a marker near the telomere, and *echinus*, a marker approximately 1 Mb from the most distal *In(1)dl-49* breakpoint, and was found to be approximately 10% of what it would be in the absence of the inversion (Stone and Thomas 1935; Sturtevant and Beadle 1936). Although we did recover a substantial number of *TM3*

chromosomes that had undergone distal exchanges, it must be remembered that these could have occurred at any point in the history of each *TM3* balancer. While not examined here, it would be interesting to see if recombination is reduced between 65D and the telomere on *TM3* within a single generation; we would indeed predict such a reduction. Alternatively, future studies using methods similar to ours could determine exactly how close to other inversions, such as *In(1)dl-49*, recombination can occur. Either way, the consequence for balanced chromosomes remains the same—crossing over is possible within this region. One feasible explanation for the high diversity in the region distal to 65D observed among the panel of *TM3* chromosomes we sampled is that exchange events may confer a competitive advantage in this region and can propagate throughout a stock, although the exact advantage of a recombinant *TM3* chromosome remains unclear.

An appreciation that single crossovers can occur distal to the 65D inversion on *TM3* also has practical purposes for long-term maintenance of deleterious alleles in stocks. At least 550 stocks at the Bloomington Stock Center have a mutation, transgene insertion, or chromosomal deletion distal to 65D that could be lost by recombination with the *TM3* present. Although, this number assumes that recombination can occur anywhere from the tip of *3L* to the 65D breakpoint, our data suggests an approximately 2-Mb buffer over which recombination may be suppressed, potentially reducing the number of vulnerable alleles. Yet the practical implication remains that genetic components thought to be present on all non-balancer chromosomes in a population may be present in only a subset of individuals in the population, or may have been moved to the balancer chromosome itself. Therefore, it may be prudent for researchers to check for the presence of the desired genetic element distal to 65D in any *TM3* stock before

undertaking experiments. Furthermore, poorly balanced regions exist at the ends of other popular balancers—including *CyO*, *In(2LR)Gla*, and *TM1*—and these balancers should generally be avoided in constructing stocks with distally located genetic components **(Figure 5.4)**.

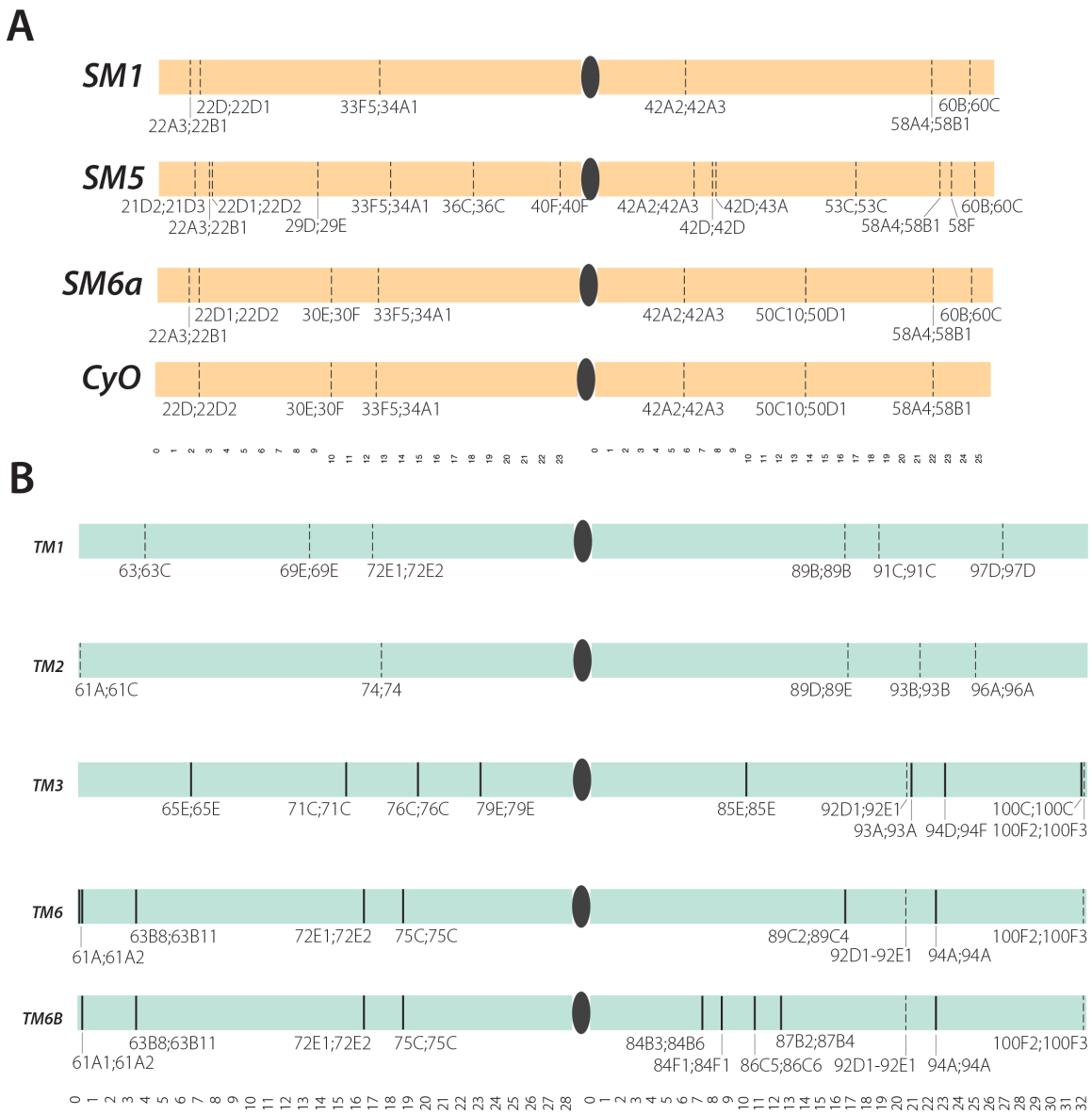

Figure 5.4. Inversion breakpoints for commonly used *2nd* and *3rd* chromosome balancers.
Breakpoints that have been molecularly identified are shown as solid lines; those that are estimates are shown as dashed lines; centromeres are represented by black dots. **(A)** Inversion breakpoints of four commonly used *2nd* chromosome balancers. **(B)** Inversion breakpoints of five commonly used *3rd* chromosome balancers, including the three balances sequenced in this study.

We also recovered evidence of double crossing over between *TM3* and *TM6B* and their normal sequence homologs. Two of the stocks with DCO events, *TM3-560* and *TM3-1614*, are unique in that they appear to be examples of the *TM3* balancer before Sb^1 and Ser^1 (*TM3-1614*) or before Sb^1 (*TM3-560*) were added to *TM3* through double crossing over in a triple-balanced female (Tinderholt 1960). Recovery of DCO events on these balancer chromosomes was not surprising, as similar exchanges were recently shown to occur within the inverted *In(1)dl-49* segment of the *X* chromosome balancer *FM7c*. DCO events on *FM7c* always replaced the female sterile *singed* (sn^{X2}) allele with a wild-type copy of the gene, resulting in sn^+ progeny with reproductive advantages (Miller, Cook, *et al.* 2016). Similarly, the DCO events recovered in the *TM3-504* and *TM6B-587* stocks created a small duplication and the elimination of three frameshifting deletions, respectively, each of which may confer selective advantages.

The precise identification of inversion breakpoints and the knowledge that rare DCO events are possible within inverted segments should encourage researchers to carefully consider the proper balancer to use when keeping any allele over a balancer for a long period of time. We suggest using a balancer with an inversion breakpoint as close to the allele of interest as possible to prevent loss through double crossing over **(Figure 5.4)**. In cases when this is not feasible, then keeping multiple copies of a stock along with periodic validation of the allele is likely in order.

Drosophila has a rich history. It has been over 100 years since Muller realized the power of this tiny fly as a potent tool for scientific inquiry (Sturtevant 2001). The success and rapid progress of experimentation in Drosophila today relies on genetic tools that have been built over the past century. Balancers have been especially important to the development of

Drosophila as a genetic model organism. Molecular characterization of balancers helps explain how they work, how they vary, and what their inherent limitations are. This study endeavors to help Drosophila geneticists make better use of these invaluable tools.

METHODS

Stocks used for breakpoint identification and validation

Stocks used in this study along with their associated Bloomington ID and genotype are listed in **Table 5.2**. Laboratory strains of *wg*$^{Sp\text{-}1}$/*SM6a duox*^{Cy}; *Pr*^{1}/*TM3 Sb*^{1} *Ser*^{1}, and +/*TM6* that were used in this study are available upon request. The ISO-1 (*y*^{1}; *Gr22b*$^{iso\text{-}1}$ *Gr22d*$^{iso\text{-}1}$ *cn*^{1} *CG33964*$^{iso\text{-}1}$ *bw*^{1} *sp*^{1}; *LysC*$^{iso\text{-}1}$ *MstProx*$^{iso\text{-}1}$ *GstD5*$^{iso\text{-}1}$ *Rh6*^{1}) stock used to create heterozygous *TM3* and *TM6B* flies for sequencing was obtained from Sue Celniker. The single *TM6* chromosome used in this study was not sequenced as an ISO-1/*TM6* heterozygote, but as a +/*TM6* heterozygote. All flies were kept on standard cornmeal-molasses and maintained at 25^{o}C.

DNA preparation and genome alignment

DNA for sequencing was prepared from either heterozygous males or a combination of heterozygous males and females using the Qiagen DNeasy Blood & Tissue Kit. All flies were starved for 1 hr before freezing at -80^{o}C for at least 1 hr. Mate pair DNA libraries for stocks *CyO-TM3-22239*, *SM6a-TM3-lab*, and *TM6B-587* were generated from 1 µg of high-quality genomic DNA. Following the manufacturer's directions, libraries were generated using the gel-free method of the Illumina Nextera Mate Pair Library Preparation kit with 10 cycles of PCR amplification. Resulting libraries were checked for quality and quantity using a Bioanalyzer 2100 (Agilent) and Qubit Fluorometer (Life Technologies). All libraries were pooled, requantified and sequenced as 150-bp paired end on the Illumina NextSeq 500 instrument. Following sequencing, Illumina Real Time Analysis version 2.4.6 was run to demultiplex reads and

generate FASTQ files. 100 ng of sample *TM6-Ubx-lab* was sheared using the Covaris s220 instrument to 300bp and prepared using the KAPA HTP Library Prep Kit for Illumina and Bioo Scientific NEXTflex DNA barcodes. The resulting library was quantified using a LabChip GXII (Perkin Elmer) and a Qubit Fluorometer (Life Technologies). This library was pooled with others, requantified and sequenced as 150-bp paired end on an Illumina HiSeq 2500 in rapid mode. Following sequencing, Illumina Real Time Analysis version 1.17.21.3 and CASAVA version 1.8.2 were run to demultiplex reads and generate FASTQ files. For the remainder of samples used in this study, 500 ng of DNA from each was fragmented to 600-bp fragments using a Covaris S220 sonicator by adjusting the treatment time to 30 seconds, except for sample *CyO-TM3-504,* which was sonicated using 89 ng of DNA and was not size selected. Libraries were prepared using the KAPA HTP Library Prep Kit for Illumina and Bioo Scientific NEXTflex DNA barcodes. The resulting libraries were quantified using a Bioanalyzer (Agilent Technologies) and a Qubit Fluorometer (Life Technologies). All libraries were pooled, requantified and sequenced as 150-bp paired end on the Illumina NextSeq 500 instrument. Following sequencing, Illumina Real Time Analysis version 2.4.6 was run to demultiplex reads and generate FASTQ files. Alignment to the *D. melanogaster* reference genome (dm6) was performed using bwa version 0.7.7-r441 (Li and Durbin 2009).

Identification and validation of inversion breakpoints

Breakpoints were identified as reported in Miller et al. (2016). Briefly, split or discordant read pairs were isolated using Samblaster (Faust and Hall 2014) and known regions of repetitive or low-complexity sequence were masked with repeatmasker (Chen 2004). Separately, we used

BreakDancer (Chen *et al.* 2009) to identify candidate inversion breakpoints. Regions where BreakDancer identified large inversion polymorphisms and where rearrangements were previously reported to be present (Lindsley and Zimm 1992) were analyzed in 1-kb windows for regions that contained more than 10 split or discordant read pairs. Breakpoints were visually validated using Integrative Genomics Viewer (Thorvaldsdottir *et al.* 2013) and the UCSC Genome Browser (Rosenbloom *et al.* 2015). Original fastq reads from each breakpoint were collected and *de novo* assembled using SOAPdenovo2 with a kmer size of 41 (Luo *et al.* 2012). Primers for PCR validation were designed using Primer3 (Rozen and Skaletsky 2000). PCR was done with Phusion polymerase, and Sanger sequencing confirmed each breakpoint. PCR primers used to validate inversion breakpoints are listed in **Table 5.S1**.

SUPPLEMENTAL TABLES

Breakpoint	Balancer	F Primer Seq	R Primer Seq	Annealing Temp
65D-85F	*TM3*	CGATGGACAGAAGCAACAGA	ATCACCAGGACTACGGCAAC	60
65D-85F	*TM3*	TAATTCCTGTGAGCGACGTG	TAATGGGCATCAAGCATACG	60
71B-94D	*TM3*	TGGACGAACAAGCTAAACGA	TCTAAAATGCCCATCCAACC	60
71B-94D	*TM3*	AACAGCTCTTGAGGCGAGAC	TACACGAGTTTTGGCAGACG	60
76B-92F	*TM3*	GTAAGGGTTCCCTGGATGGT	GGCGATCAAACAACCAAAGT	60
76B-92F	*TM3*	TCAGGTGATGTGCTGGAATC	AGGAAGATCCCGCAATAGGT	60
79F-100D	*TM3*	CCTCCGAAACGCATTGTATT	TGCAGTTGGATAGGTTCGTG	60
79F-100D	*TM3*	ATTTGGATCCATTCCGTTGA	AACAGGGCGGCTACTTGTTA	60
61B-87B	*TM6B*	TCACTTTAGCAGGTTCCATCG	TTGAACCCGAAATGGCTTTA	60
61B-87B	*TM6B*	TTGACAGGGTGGTCCAATTA	AATTTGCTTCGCAATGAAGG	60
75D-94A	*TM6, TM6B*	AAATTGCCGATCAAAAGGTG	ATTAATTGGCCCAGGACCTC	60
75D-94A	*TM6, TM6B*	AACCCACGAGTCCCCTAACT	ACCCCGAAGTGTGCAGTATC	60
61A-89C	*TM6*	GCACACTCCGCACACTTG	CGGGTAAGAGCATGACCAAT	63
61A-89C	*TM6*	TTTGAGCTGCACTCTTGCAC	TTCCCCATCAACTCCTCTCA	61
61A-89C	*TM6*	ACCCCATTTTTCACTTGCTG	ACGTGTGGGGTCACTAGAGG	61

Table 5.S1. PCR primers used to validate selected inversion breakpoints

Bloomington ID	FlyBase Genotype
TM3 Stocks	
N/A	*wg^{Sp-1}/SM6a, Pr^{1} Dr^{1}/TM3, Sb^{1} Ser^{1}*
120	*TM3, ry^{RK} Sb^{1} Ser^{1}/TM6B, Tb^{1}*
500	*eg^{1}/TM3, Sb^{1} Ser^{1}*
504	*amd^{21} Bl^{1}/CyO; DCTN1-p150^{1}/TM3, Sb^{1} Ser^{1}*
560	*Pri^{1} Dr^{1}/TM3*
1614	*y^{1} w*; TM3 y^{+} Ser^{1}/ Sb^{1}*
1679	*Dp(1;Y)B^{S}; dsx^{1} p^{p}/TM3, Sb^{1}*
2053	*twr^{1} red^{1} e^{1}/TM3, Sb^{1} Ser^{1}*
2098	*sas^{15} p^{p} cu^{1}/TM3, Sb^{1} Ser^{1}*
2198	*Df(3R)ro80b, st^{1} e^{1}/TM3, Sb^{1}*
2485	*ru^{1} h^{1} Diap1^{1} st^{1} cu^{1} srp^{3} sr^{1} e^{s} ca^{1}/TM3, Sb^{1} Ser^{1}*
3251	*CyO, l(2)DTS513^{1}/l(2)**; BicF^{1}/TM3, Sb^{1} Ser^{1}*
5457	*hkb^{2} p^{p}/TM3, p^{+} Sb^{1}*
8852	*mwh^{2} ru^{1} kni^{ri-1}/TM3, mwh^{2} ru^{1} Sb^{1}*
9013	*l(3)SG43^{1} red^{1}/TM3, Sb^{1} Ser^{1}*
22239	*y^{1} w*; P{y[+t7.7]=Mae-UAS.6.11}Dpit47^{LA00491}/CyO; l(3)**/TM3, Sb^{1} Ser^{1}*
24759	*w*; sna^{Sco}/CyO; P{w[+mC]=ninaD-GAL4.W}3/TM3, Sb^{1}*
38418	*w*; P{w[+mW.hs]=GawB}nubbin-AC-62/CyO; P{w[+mC]=UAS-Nslmb-vhhGFP4}3/TM3, Sb^{1} Ser^{1}*
TM6 Stocks	
N/A	*+/TM6, Ubx*
TM6B Stocks	
120	*TM3, ry^{RK} Sb^{1} Ser^{1}/TM6B, Tb^{1}*
587	*Sb^{Spi}/TM6B, Tb^{1}*
2188	*w^{1118}; Scr^{4} red^{1} e^{1}/TM6B, Tb^{1}*

Table 5.S2. Stocks sequenced in this study

Chapter 6: Conclusions and future directions

In this thesis I have used whole-genome sequencing to answer a number of questions related to the distribution of crossovers (COs), non-crossover gene conversions (NCOs), and to understand how inversion breakpoints suppress exchange in multiply inverted chromosomes. For researchers that study the distribution of COs and NCOs this work will be of interest because it suggests that the formation of DSBs in Drosophila is not controlled by phenomenon such as the centromere effect or interference, it is only the repair of those events that is controlled. Importantly, my work looked at individual meiotic events in both wild-type and a mutant background—a strategy that has not been undertaken on the genome-wide scale in Drosophila. By looking at individual meiotic events I was able to also identify evidence for TE-mediated CNV formation in Drosophila—a phenomenon which has been previously proposed in Drosophila, but never conclusively proven. Finally, I provided several observations about how recombination occurs in multiply inverted, or balancer, chromosomes, which will have a broad impact throughout the Drosophila community. Specifically, I have shown the distance over which inversion breakpoints appear to suppress exchange and demonstrated that reversion of duplicated segments in

The distribution of crossover and non-crossover gene conversions in Drosophila

Here I have shown that while CO distribution is primarily controlled by the centromere effect and interference, which was previously known, NCOs do not respond to either of those forces. Surprisingly, I identified NCOs near COs, near other NCOs, and in the centromere proximal 1/3

of each chromosome arm. This finding suggests that DSBs in Drosophila do not respond to either the centromere effect or interference, but does not support the idea that DSBs are randomly distributed throughout the genome—there may still be a level of control, such as targeting to open chromatin—that was not readily apparent in my data.

I also analyzed a mutant that failed to build SC to look for NCO events. While it has been known that SC mutants fail to repair breaks by crossing over, it is unknown if they are able to repair breaks via NCO. I recovered evidence for one NCO event, defined by two polymorphisms (a SNP and a small deletion) and validated them by PCR and Sanger sequencing. That DSBs may be repaired by NCO in a SC mutant that pairs poorly and fails to synapse was surprising, so I cannot rule out that this conversion event occurred at a pre-meiotic stage in a precursor cell.

It would be interesting to look at CO and NCO distribution in so-called polar-effect mutants, or those that shift crossover distribution into the centromere proximal regions, such as *Mei-218* (McKim *et al.* 1996). One question that could be addressed in these mutants is if more distal NCO events are observed, suggesting that breaks that would normally be repaired as COs would instead be repaired as NCOs. Second, it would also be informative to look at CO and NCO distribution in either wild-caught Drosophila melanogaster lines or in the background of other commonly used laboratory strains. Although, as pointed out in Chapter 2, it will be critical to understand what factors, if any contribute to an altered CO or NCO distribution or rate as these lines have not been as extensively studied in the laboratory as others.

Unequal crossing over between transposable elements is a common source of genetic variability in Drosophila

Whole-genome sequencing of individual meiotic products allowed me to identify and describe several inherited and *de novo* CNVs in offspring from both wild-type and SC-mutant mothers. In both classes of offspring we observed CNVs in 1–2% of offspring, in-line with a previous estimate of the genome-wide CNV rate in Drosophila. Strikingly, we recovered an identical 855 kb *de novo* CNV in two males from different moms, demonstrating that CNV formation in Drosophila may be recurrent, as is observed in humans. These observations open the door to using Drosophila as a model to study CNV formation in humans, bringing a difficult problem to a tractable model system.

Future studies of CNV formation should be performed in a variety of meiotic mutants, such as those unable to make DSBs (a *mei-W68* mutant, for example), as whether or not these CNVs are DSB-dependent is an open question. In addition, broader studies would be able to determine if CNVs are more frequently mediated by certain types of TEs. While we observed several examples of CNVs mediated by both *roo* and *foldback* elements, it is unclear if these are the most frequent mediators of CNV formation.

Studying inversion heterozygotes helps us understand how DSBs are repaired

In this thesis I also reported the precise location of all inversion breakpoints for the *X* chromosome balancer *FM7* and for the majority of breakpoints present on the *3rd* chromosome balancers *TM3*, *TM6*, and *TM6B*. Knowledge of the precise location of balancer chromosome breakpoints has important consequences in both study design and understanding phenotypes

that may be associated with balancer chromosomes. For example, I observed that the *3rd* chromosome balancer *TM3* contains an inversion that bisects the highly conserved tumor suppressor gene p53 in half, with potential consequences for studies of DNA repair, among others. I also found that a rearrangement on the TM6B chromosome may be causing the previously uncharacterized *Antp^{Hu}* phenotype as it lies in close proximity to *Antp*.

I was also able to recover evidence of DCO events between both the *X* chromosome balancer *FM7* and the *3rd* chromosome balancers and their normal-sequence homologs, demonstrating that recombination is not completely suppressed on balancer chromosomes. The observation that DCO is possible on balancers explains the loss of balanced alleles in many cases, and helps Drosophila workers using balancers to decide exactly which balancer should be used to balance particular alleles.

Finally, inversion breakpoints have long been expected to suppress exchange, yet is has not been shown over what distance they act. Using a panel of *TM3* chromosomes which can freely recombine over the 6.9 Mb region between the telomere and the first inversion breakpoint of the *3rd* chromosome I observed that single crossover events occurred up to approximately 2 Mb from the inversion breakpoint, the first direct observation of the suppression distance.

An obvious extension of this work will be to look at the *2nd* chromosome balancers *SM1*, *SM5*, *SM6a*, and *CyO*. *CyO* maintains large unbalanced regions between both telomeres and the distalmost inversion breakpoints, similar to the unbalanced region on *3L* of *TM3*, meaning single crossovers should be abundant among a panel of *CyO* chromosomes. Additionally, few of the inversion breakpoints have been precisely identified on the *2nd* chromosome balancers,

work which may begin using the data generated here for the *3rd* chromosome balancers, as five of the stocks whole-genome sequenced carry a *2nd* chromosome balancer as well.

A final question that may be addressed using the data generated for the *X* and *3rd* chromosome balancers here is the frequency and distribution of NCO events onto these chromosomes. It is apparent that there are short tracts of SNPs unique to each of these balancers that may have been NCO events between the balancer and a normal-sequence homolog. Whether NCO events occur closer to inversion breakpoints than CO events do would be an interesting question to address that may help researchers further understand how these breakpoints suppress exchange.

Drosophila as a model research organism: another 100 years

The success and rapid progress of experimentation in Drosophila over the past 100 years is built upon tools and observations made by others. As demonstrated in this thesis, there is great value in careful analysis of the tools and understanding of processes we use each day with the shared assumption that they are as they have always been. As new technologies such as whole genome sequencing become available it is important to take pause and consider how we may use that technology to validate and update our existing assumptions and knowledge. Only with clarity and understanding can we guarantee another 100 years of exciting and unexpected findings in Drosophila.

References

Abreu-Blanco, M. T., J. M. Verboon, and S. M. Parkhurst, 2011 Cell wound repair in Drosophila occurs through three distinct phases of membrane and cytoskeletal remodeling. The Journal of Cell Biology 193: 455–464.

Adams, M. D., S. E. Celniker, R. A. Holt, and C. A. Evans, 2000 The Genome Sequence of Drosophila melanogaster. Science.

Adrian, A. B., and J. M. Comeron, 2013 The Drosophila early ovarian transcriptome provides insight to the molecular causes of recombination rate variation across genomes. BMC Genomics 14: 794.

Anderson, L. K., L. D. Lohmiller, X. Tang, D. B. Hammond, L. Javernick *et al.*, 2014 Combined fluorescent and electron microscopic imaging unveils the specific properties of two classes of meiotic crossovers. Proceedings of the National Academy of Sciences 111: 13415–13420.

Araye, Q., and K. Sawamura, 2013 Genetic decay of balancer chromosomes in *Drosophila melanogaster*. Fly (Austin) 7: 184–186.

Ashburner, M., 1972 New Mutants Report. DIS 49: 34.

Ashburner, M., K. Golic, and R. S. Hawley, 2005 *Drosophila—A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Auton, A., Y. Rui Li, J. Kidd, K. Oliveira, J. Nadel *et al.*, 2013 Genetic recombination is targeted towards gene promoter regions in dogs. PLoS Genet 9: e1003984.

Bailey, T. L., and C. Elkan, 1994 Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2: 28–36.

Bailey, T. L., N. Williams, C. Misleh, and W. W. Li, 2006 MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Research 34: W369–73.

Baker, B. S., and A. T. Carpenter, 1972 Genetic analysis of sex chromosomal meiotic mutants in *Drosophila melanogaster*. Genetics 71: 255–286.

Baker, B. S., and J. C. Hall, 1972 Meiotic mutants: genetic control of meiotic recombination and chromosome segregation.

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober *et al.*, 2010 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327: 836–840.

Beadle, G. W., 1932 A possible influence of the spindle fibre on crossing-over in Drosophila. Proc. Natl. Acad. Sci. U.S.A. 18: 160–165.

Beadle, G. W., and A. H. Sturtevant, 1935 *X* chromosome inversions and meiosis in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 21: 384–390.

Bender, H. A., 1960 Studies on the expression of various *singed* alleles in *Drosophila melanogaster*. Genetics 45: 867–883.

Bender, W., M. Akam, F. Karch, P. A. Beachy, M. Peifer *et al.*, 1983 Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. Science 221: 23–29.

Berchowitz, L. E., and G. P. Copenhaver, 2010 Genetic interference: don't stand so close to me. Curr. Genomics 11: 91–102.

Beumer, K. J., J. K. Trautman, M. Christian, T. J. Dahlem, C. M. Lake *et al.*, 2013 Comparing zinc finger nucleases and transcription activator-like effector nucleases for gene targeting in Drosophila. G3 (Bethesda) 3: 1717–1725.

Blanton, H. L., S. J. Radford, S. McMahan, H. M. Kearney, J. G. Ibrahim *et al.*, 2005 REC, Drosophila MCM8, drives formation of meiotic crossovers. PLoS Genet 1: e40.

Blumenstiel, J. P., A. C. Noll, J. A. Griffiths, A. G. Perera, K. N. Walton *et al.*, 2009 Identification of EMS-induced mutations in *Drosophila melanogaster* by whole-genome sequencing. Genetics 182: 25–32.

Bolcun-Filas, E., E. Hall, R. Speed, M. Taggart, C. Grey *et al.*, 2009 Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between synaptonemal complex structural components and DNA repair. PLoS Genet 5: e1000393.

Bosco, G., 2012 Chromosome Pairing: A Hidden Treasure No More (G. P. Copenhaver, Ed.). PLoS Genet 8: e1002737.

Bridges, C. B., 1916 Non-Disjunction as Proof of the Chromosome Theory of Heredity. Genetics 1: 1–53.

Bridges, C. B., 1936 The Bar “gene” a duplication. Science 83: 210–211.

Brizuela, B. J., L. Elfring, J. Ballard, J. W. Tamkun, and J. A. Kennison, 1994 Genetic analysis of the *brahma* gene of *Drosophila melanogaster* and polytene chromosome subdivisions 72AB. Genetics 137: 803–813.

Cahoon, C. K., and R. S. Hawley, 2013 Flies Get a Head Start on Meiosis (G. P. Copenhaver, Ed.). PLoS Genet 9: e1004051–3.

Campuzano, S., L. Carramolino, C. V. Cabrera, M. Ruíz-Gómez, R. Villares *et al.*, 1985 Molecular genetics of the *achaete-scute* gene complex of *D. melanogaster*. Cell 40: 327–338.

Cant, K., B. A. Knowles, M. S. Mooseker, and L. Cooley, 1994 Drosophila singed, a fascin

homolog, is required for actin bundle formation during oogenesis and bristle extension. The Journal of Cell Biology 125: 369–380.

Carlson, P. S., 1972 The effects of inversions and the C(3)G mutation on intragenic recombination in Drosophila. Genet. Res. 19: 129–132.

Carpenter, A. T., 1982 Mismatch repair, gene conversion, and crossing-over in two recombination-defective mutants of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 79: 5961–5965.

Casso, D., F. Ramírez-Weber, and T. B. Kornberg, 2000 GFP-tagged balancer chromosomes for *Drosophila melanogaster*. Mech. Dev. 91: 451–454.

Chan, A. H., P. A. Jenkins, and Y. S. Song, 2012 Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. PLoS Genet 8: e1003090.

Charlesworth, B., and C. H. Langley, 1986 The evolution of self-regulated transposition of transposable elements. Genetics 112: 359–383.

Charlesworth, B., and C. H. Langley, 1989 The population genetics of Drosophila transposable elements. Annu. Rev. Genet. 23: 251–287.

Chen, K., J. W. Wallis, M. D. McLellan, D. E. Larson, J. M. Kalicki *et al.*, 2009 Breakdancer: an algorithm for high-resolution mapping of genomic structural variation. Nature Methods 6: 677–681.

Chen, N., 2004 Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics Chapter 4: Unit 4.10–4.10.14.

Chovnick, A., G. H. Ballantyne, and D. G. Holm, 1971 Studies on gene conversion and its relationship to linked exchange in Drosophila melanogaster. Genetics 69: 179–209.

Chovnick, A., G. H. Ballantyne, D. L. Baillie, and D. G. Holm, 1970 Gene conversion in higher organisms: half-tetrad analysis of recombination within the rosy cistron of Drosophila melanogaster. Genetics 66: 315–329.

Christophorou, N., T. Rubin, and J.-R. Huynh, 2013 Synaptonemal Complex Components Promote Centromere Pairing in Pre-meiotic Germ Cells (R. S. Hawley, Ed.). PLoS Genet 9: e1004012–9.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) 6: 80–92.

Cirulli, E. T., R. M. Kliman, and M. A. F. Noor, 2007 Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*. J Mol Evol 64: 129–135.

Cline, T. W., and B. J. Meyer, 1996 Vive la différence: males vs females in flies vs worms. Annu. Rev. Genet. 30: 637–702.

Cole, F., F. Baudat, C. Grey, S. Keeney, B. de Massy *et al.*, 2014 Mouse tetrad analysis provides insights into recombination mechanisms and hotspot evolutionary dynamics. Nature Publishing Group.

Collins, K. A., J. R. Unruh, B. D. Slaughter, Z. Yu, C. M. Lake *et al.*, 2014 Corolla is a novel protein that contributes to the architecture of the synaptonemal complex of Drosophila. Genetics 198: 219–228.

Comeron, J. M., R. Ratnappan, and S. Bailin, 2012 The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genet 8: e1002905.

Cooper, J. L., E. A. Greene, B. J. Till, C. A. Codomo, B. T. Wakimoto *et al.*, 2008 Retention of induced mutations in a Drosophila reverse-genetic resource. Genetics 180: 661–667.

Craymer, L., 1984 New Mutants Report. DIS 60: 234–236.

Craymer, L., 1981 Techniques for manipulating chromosomal rearrangements and their application to *Drosophila melanogaster*. I. Pericentric inversions. Genetics 99: 75–97.

Curtis, D., and W. Bender, 1991 Gene conversion in Drosophila and the effects of the meiotic mutants mei-9 and mei-218. Genetics 127: 739–746.

Davis, P. S., M. W. Shen, and B. H. Judd, 1987 Asymmetrical pairings of transposons in and proximal to the white locus of Drosophila account for four classes of regularly occurring exchange products. Proc. Natl. Acad. Sci. U.S.A. 84: 174–178.

Do, A. T., J. T. Brooks, M. K. Le Neveu, and J. R. LaRocque, 2013 Double-strand break repair assays determine pathway choice and structure of gene conversion events in *Drosophila melanogaster*. G3 (Bethesda).

Dobzhansky, T., 1930 Translocations involving the third and the fourth chromosomes of *Drosophila melanogaster*. Genetics 15: 347–399.

Duncan, D. M., E. A. Burgess, and I. Duncan, 1998 Control of distal antennal identity and tarsal development in Drosophila by spineless-aristapedia, a homolog of the mammalian dioxin receptor. Genes & Development 12: 1290–1303.

Faust, G. G., and I. M. Hall, 2014 SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics 30: 2503–2505.

Fleming, R. J., T. N. Scottgale, R. J. Diederich, and S. Artavanis-Tsakonas, 1990 The gene Serrate encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in *Drosophila melanogaster*. Genes & Development 4: 2188–2201.

Fraune, J., M. Alsheimer, J.-N. Volff, K. Busch, S. Fraune *et al.*, 2012 Hydra meiosis reveals unexpected conservation of structural synaptonemal complex proteins across metazoans. Proceedings of the National Academy of Sciences 109: 16588–16593.

Gabay, S. J., and J. R. Laughnan, 1973 Recombination at the Bar locus in an inverted attached-*X* system in *Drosophila melanogaster*. Genetics 75: 485–495.

Gilliland, W. D., 2015 A Comment on Fine-Scale Heterogeneity in Crossover Rate in the garnet-scalloped Region of the Drosophila melanogaster X Chromosome. Genetics 201: 1275–1277.

Glaser, R. L., and A. C. Spradling, 1994 Unusual properties of genomic DNA molecules spanning the euchromatic-heterochromatic junction of a Drosophila minichromosome. Nucleic Acids Research 22: 5068–5075.

Goldberg, M. L., J. Y. Sheen, W. J. Gehring, and M. M. Green, 1983 Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the Drosophila melanogaster genome. Proc. Natl. Acad. Sci. U.S.A. 80: 5017–5021.

Goldfarb, T., and M. Lichten, 2010 Frequent and efficient use of the sister chromatid for DNA double-strand break repair during budding yeast meiosis. PLoS Biol 8: e1000520.

Gong, W. J., K. S. McKim, and R. S. Hawley, 2005 All paired up with no place to go: pairing, synapsis, and DSB formation in a balancer heterozygote. PLoS Genet 1: e67.

Gratz, S. J., A. M. Cummings, J. N. Nguyen, D. C. Hamm, L. K. Donohue *et al.*, 2013 Genome Engineering of Drosophila with the CRISPR RNA-Guided Cas9 Nuclease. Genetics 194: 1029–1035.

Green, B. M., K. J. Finn, and J. J. Li, 2010 Loss of DNA replication control is a potent inducer of gene amplification. Science 329: 943–946.

Grell, R. F., and E. B. Lewis, 1956 New mutants report. DIS 30: 71.

Haag-Liautard, C., M. Dorris, X. Maside, S. Macaskill, D. L. Halligan *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila. Nature 445: 82–85.

Hall, J. C., 1972 Chromosome segregation influenced by two alleles of the meiotic mutant c(3)G in Drosophila melanogaster. Genetics 71: 367–400.

Hammonds, A. S., and J. W. Fristrom, 2006 Mutational analysis of Stubble-stubbloid gene structure and function in Drosophila leg and bristle morphogenesis. Genetics 172: 1577–1593.

Hawley, R. S., 1980 Chromosomal sites necessary for normal levels of meiotic recombination in

Drosophila melanogaster. I. Evidence for and mapping of the sites. Genetics 94: 625–646.

Hazelrigg, T., and T. C. Kaufman, 1983 Revertants of Dominant Mutations Associated with the Antennapedia Gene Complex of *Drosophila melanogaster*: Cytology and Genetics. Genetics 105: 581–600.

Heil, C. S. S., and M. A. F. Noor, 2012 Zinc Finger Binding Motifs Do Not Explain Recombination Rate Variation within or between Species of Drosophila (N. Singh, Ed.). PLoS ONE 7: e45055.

Hey, J., 2004 What's so hot about recombination hotspots? PLoS Biol 2: e190.

Higashijima, S., T. Kojima, T. Michiue, S. Ishimaru, Y. Emori *et al.*, 1992 Dual Bar homeo box genes of Drosophila required in two photoreceptor cells, R1 and R6, and primary pigment cells for normal eye development. Genes & Development 6: 50–60.

Higashijima, S., T. Michiue, Y. Emori, and K. Saigo, 1992 Subtype determination of Drosophila embryonic external sensory organs by redundant homeo box genes BarH1 and BarH2. Genes & Development 6: 1005–1018.

Hilliker, A. J., and A. Chovnick, 1981 Further observations on intragenic recombination in *Drosophila melanogaster*. Genet. Res. 38: 281–296.

Hilliker, A. J., G. Harauz, A. G. Reaume, M. Gray, S. H. Clark *et al.*, 1994 Meiotic gene conversion tract length distribution within the rosy locus of *Drosophila melanogaster*. Genetics 137: 1019–1026.

Hollingsworth, N. M., and S. J. Brill, 2004 The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. Genes & Development 18: 117–125.

Hoover, M. E., 1938 Cytogenetic analysis of nine inversions in *Drosophila melanogaster*. Z.Ver-erbungslehre 74: 420–434.

Hsieh, T., and D. Brutlag, 1979 Sequence and sequence variation within the 1.688 g/cm3 satellite DNA of *Drosophila melanogaster*. Journal of Molecular Biology 135: 465–481.

Hughes, S. E., W. D. Gilliland, J. L. Cotitta, S. Takeo, K. A. Collins *et al.*, 2009 Heterochromatic threads connect oscillating chromosomes during prometaphase I in Drosophila oocytes. PLoS Genet 5: e1000348.

Hunt, P. A., C. Lawson, M. Gieske, B. Murdoch, H. Smith *et al.*, 2012 Bisphenol A alters early oogenesis and follicle formation in the fetal ovary of the rhesus monkey. Proceedings of the National Academy of Sciences xx–xx.

Itsara, A., G. M. Cooper, C. Baker, S. Girirajan, J. Li *et al.*, 2009 AR TICLE Population Analysis of Large Copy Number Variants and Hotspots of Human Genetic Disease. The American Journal of Human Genetics 84: 148–161.

Jang, J. K., D. E. Sherizen, R. Bhagat, E. A. Manheim, and K. S. McKim, 2003 Relationship of DNA double-strand breaks to synapsis in Drosophila. Journal of Cell Science 116: 3069–3077.

Jin, S., S. Martinek, W. S. Joo, J. R. Wortman, N. Mirkovic *et al.*, 2000 Identification and characterization of a p53 homologue in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 97: 7301–7306.

Johnson-Schlitz, D. M., C. Flores, and W. R. Engels, 2007 Multiple-pathway analysis of double-strand break repair mutations in Drosophila. PLoS Genet 3: e50.

Joyce, E. F., and K. S. McKim, 2009 Drosophila PCH2 is required for a pachytene checkpoint that monitors double-strand-break-independent events leading to meiotic crossover formation. Genetics 181: 39–51.

Joyce, E. F., and K. S. McKim, 2011 Meiotic checkpoints and the interchromosomal effect on crossing over in Drosophila females. Fly (Austin) 5: 134–141.

Joyce, E. F., N. Apostolopoulos, B. J. Beliveau, and C. T. Wu, 2013 Germline Progenitors Escape the Widespread Phenomenon of Homolog Pairing during Drosophila Development (R. S. Hawley, Ed.). PLoS Genet 9: e1004013–11.

Joyce, E. F., B. R. Williams, T. Xie, and C. T. Wu, 2012 Identification of Genes That Promote or Antagonize Somatic Homolog Pairing Using a High-Throughput FISH–Based Screen (R. S. Hawley, Ed.). PLoS Genet 8: e1002667.

Kaminker, J. S., C. M. Bergman, B. Kronmiller, J. Carlson, R. Svirskas *et al.*, 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biology 3: RESEARCH0084.

Keeney, S., C. N. Giroux, and N. Kleckner, 1997 Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88: 375–384.

Kenney, D. E., and G. G. Borisy, 2009 *Thomas Hunt Morgan at the marine biological laboratory: naturalist and experimentalist.* Genetics.

Kojima, T., S. Ishimaru, S. Higashijima, E. Takayama, H. Akimaru *et al.*, 1991 Identification of a different-type homeobox gene, BarH1, possibly causing Bar (B) and Om(1D) mutations in Drosophila. Proc. Natl. Acad. Sci. U.S.A. 88: 4343–4347.

Kulathinal, R. J., S. M. Bennett, C. L. Fitzpatrick, and M. A. F. Noor, 2008 Fine-scale mapping of recombination rate in Drosophila refines its correlation to diversity and divergence. Proc. Natl. Acad. Sci. U.S.A. 105: 10051–10056.

Lachaise, D., J. R. David, F. Lemeunier, L. Tsacas, and M. Ashburner, 1986 The reproductive relationships of *Drosophila sechellia* with *D. mauritiana*, *D. simulans*, and *D. melanogaster* from the Afrotropical region. Evolution 262–271.

Lack, J. B., C. M. Cardeno, M. W. Crepeau, W. Taylor, R. B. Corbett-Detig *et al.*, 2015 The Drosophila Genome Nexus: A Population Genomic Resource of 623 Drosophila melanogaster Genomes, Including 197 from a Single Ancestral Range Population. Genetics 199: 1229–1241.

Lake, C. M., and R. S. Hawley, 2012 The molecular control of meiotic chromosomal behavior: events in early meiotic prophase in Drosophila oocytes. Annu. Rev. Physiol. 74: 425–451.

Lake, C. M., J. K. Holsclaw, S. P. Bellendir, J. Sekelsky, and R. S. Hawley, 2013 The development of a monoclonal antibody recognizing the *Drosophila melanogaster* phosphorylated histone H2A variant (γ-H2AV). G3 (Bethesda) 3: 1539–1543.

Lattao, R., S. Bonaccorsi, X. Guan, S. A. Wasserman, and M. Gatti, 2011 Tubby-tagged balancers for the Drosophila *X* and second chromosomes. Fly (Austin) 5: 369–370.

Le, T., Z. Liang, H. Patel, M. H. Yu, G. Sivasubramaniam *et al.*, 2006 A new family of Drosophila balancer chromosomes with a w- dfd-GMR yellow fluorescent protein marker. Genetics 174: 2255–2257.

Lee, Y. C. G., and C. H. Langley, 2012 Long-term and short-term evolutionary impacts of transposable elements on Drosophila. Genetics 192: 1411–1432.

Lee, Y. C. G., and C. H. Langley, 2010 Transposable elements in natural populations of *Drosophila melanogaster*. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 365: 1219–1228.

Lewis, E. B., 1960 New Mutants Report. DIS 34: 51.

Lewis, E. B., and R. F. Mislove, 1953 New mutants report. DIS 27: 57–58.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Li, R., C. Yu, Y. Li, T.-W. Lam, S.-M. Yiu *et al.*, 2009 SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25: 1966–1967.

Libuda, D. E., S. Uzawa, B. J. Meyer, and A. M. Villeneuve, 2013 Meiotic chromosome structures constrain and respond to designation of crossover sites. Nature 1–14.

Lichten, M., and A. Goldman, 1995 Meiotic recombination hotspots. Annu. Rev. Genet.

Lindsley, D. L., and E. H. Grell, 1967 *Genetic Variations of Drosophila melanogaster*. Carnegie Inst. of Washington.

Lindsley, D. L., and L. Sandler, 1977 The genetic analysis of meiosis in female *Drosophila melanogaster*. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 277: 295–312.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego, CA.

Lohe, A. R., A. J. Hilliker, and P. A. Roberts, 1993 Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. Genetics 134: 1149–1174.

Lucchesi, J. C., and D. T. Suzuki, 1968 The interchromosomal control of recombination. Annu. Rev. Genet.

Lunde, K., 2003 Activation of the knirps locus links patterning to morphogenesis of the second wing vein in Drosophila. Development 130: 235–248.

Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang *et al.*, 2012 SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1: 18.

Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2013 The *Drosophila melanogaster* genetic reference panel. Nature 482: 173–178.

Manzano-Winkler, B., S. E. McGaugh, and M. A. F. Noor, 2013 How hot are Drosophila hotspots? Examining recombination rate variation and associations with nucleotide diversity, divergence, and maternal age in *Drosophila pseudoobscura* (A. Palsson, Ed.). PLoS ONE 8: e71582.

May, H. G., 1917 Selection for higher and lower facet numbers in the bar-eyed race of Drosophila and the appearance of reverse mutations. Biol Bull 33: 361–395.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis *et al.*, 2010 The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20: 1297–1303.

McKim, K. S., and A. Hayashi-Hagihara, 1998 mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes & Development 12: 2932–2942.

McKim, K. S., J. B. Dahmus, and R. S. Hawley, 1996 Cloning of the Drosophila melanogaster meiotic recombination gene mei-218: a genetic and molecular analysis of interval 15E. Genetics 144: 215–228.

McVey, M., D. Radut, and J. J. Sekelsky, 2004 End-joining repair of double-strand breaks in *Drosophila melanogaster* is largely DNA ligase IV independent. Genetics 168: 2067–2076.

Mehrotra, S., and K. S. McKim, 2006 Temporal analysis of meiotic DNA double-strand break formation and repair in Drosophila females. PLoS Genet 2: e200.

Merker, J. D., M. Dominska, and T. D. Petes, 2003 Patterns of heteroduplex formation associated with the initiation of meiotic recombination in the yeast *Saccharomyces cerevisiae*. Genetics 165: 47–63.

Merriam, J. R., 1969 FM7: A “new” first chromosome balancer. DIS 44: 101.

Merriam, J. R., 1968 FM7: first multiple seven. DIS 43: 64.

Merriam, J. R., and C. Duffy, 1972 First Multiple seven: now contains sn[x2] for better balancing. DIS 48: 43–44.

Miller, D. E., K. R. Cook, N. Yeganeh Kazemi, C. B. Smith, A. J. Cockrell *et al.*, 2016 Rare recombination events generate sequence diversity among balancer chromosomes in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences 113 E1352-61.

Miller, D. E., C. B. Smith, N. Yeganeh Kazemi, A. J. Cockrell, A. V. Arvanitakis *et al.*, 2016 Whole-Genome Analysis of Individual Meiotic Events in *Drosophila melanogaster* Reveals that Noncrossover Gene Conversions are Insensitive to Interference and the Centromere Effect. *Genetics* Early online, March 4, 2016; DOI: 10.1534/genetics.115.186486.

Miller, D. E., S. Takeo, K. Nandanan, A. Paulson, M. M. Gogol *et al.*, 2012 A whole-chromosome analysis of meiotic recombination in *Drosophila melanogaster*. G3 (Bethesda) 2: 249–260.

Mohr, S. E., Y. Hu, K. Kim, B. E. Housden, and N. Perrimon, 2014 Resources for Functional Genomics Studies in Drosophila melanogaster. Genetics 197: 1–18.

Montgomery, E. A., S. M. Huang, C. H. Langley, and B. H. Judd, 1991 Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. Genetics 129: 1085–1098.

Muller, H. J., 1936 Bar Duplication. Science 83: 528–530.

Muller, H. J., 1918 Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors. Genetics 3: 422–499.

Muller, H. J., 1926 HJ Muller archive at Indiana University.

Muller, H. J., 1916 The mechanism of crossing-over. The American Naturalist 50: 193–221.

Muller, H. J., and A. A. Prokofyeva, 1934 Continuity and discontinuity of the hereditary material. Dokl. Akad. Nauk SSSR NS 4: 74–83.

Muller, H. J., A. A. Prokofyeva, and K. V. Kossikov, 1936 Unequal Crossing-over in the Bar Mutant as a Result of Duplication of a Minute Chromosome of *Drosophila*.

Muñoz-Fuentes, V., A. Di Rienzo, and C. Vilà, 2011 Prdm9, a major determinant of meiotic

recombination hotspots, is not functional in dogs and their wild relatives, wolves and coyotes. PLoS ONE 6: e25498.

Myers, E. W., G. G. Sutton, A. L. Delcher, I. M. Dew, D. P. Fasulo *et al.*, 2000 A whole-genome assembly of Drosophila. Science 287: 2196–2204.

Norris, E., M. Sanders, V. Crumety, and S. I. Tsubota, 1992 The identification of the Bs breakpoint and of two possible Bar genes. Mol. Gen. Genet. 233: 106–112.

Novak, J. E., P. B. Ross-Macdonald, and G. S. Roeder, 2001 The budding yeast Msh4 protein functions in chromosome synapsis and the regulation of crossover distribution. Genetics 158: 1013–1025.

Novitski, E., and G. Braver, 1954 An analysis of crossing over within a heterozygous inversion in *Drosophila melanogaster*. Genetics 39: 197–209.

Offermann, C. A., and H. J. Muller, 1932 Regional differences in crossing over as a function of the chromosome structure. Proc Sixth Int Congress Genetics 143–145.

Page, S. L., and R. S. Hawley, 2001 c(3)G encodes a Drosophila synaptonemal complex protein. Genes & Development 15: 3130–3143.

Page, S. L., R. S. Khetani, C. M. Lake, R. J. Nielsen, J. K. Jeffress *et al.*, 2008 Corona is required for higher-order assembly of transverse filaments into full-length synaptonemal complex in Drosophila oocytes. PLoS Genet 4: e1000194.

Page, S. L., R. J. Nielsen, K. Teeter, C. M. Lake, S. Ong *et al.*, 2007 A germline clone screen for meiotic mutants in *Drosophila melanogaster*. Fly (Austin) 1: 172–181.

Painter, T. S., 1934 The morphology of the *X* chromosome in salivary glands of *Drosophila melanogaster* and a new type of chromosome map for this element. Genetics 19: 448–469.

Parry, D. M., 1973 A meiotic mutant affecting recombination in female *Drosophila melanogaster*. Genetics 73: 465–486.

Paterson, J., and K. O'Hare, 1991 Structure and transcription of the *singed* locus of *Drosophila melanogaster*. Genetics 129: 1073–1084.

Patterson, J. T., 1933 The mechanism of mosaic formation in Drosophila. Genetics 18: 32–52.

Patterson, J. T., and W. S. Stone, 1935 Some observations on the structure of the scute-8 chromosome of *Drosophila melanogaster*. Genetics 20: 172–178.

Perkins, L. A., L. Holderbaum, R. Tao, Y. Hu, R. Sopko *et al.*, 2015 The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. Genetics 201: 843–852.

Peterson, H. M., and J. R. Laughnan, 1963 Intrachromosomal exchange at the Bar locus in Drosophila. Proc. Natl. Acad. Sci. U.S.A. 50: 126.

Pina, C., and F. Pignoni, 2012 Tubby-RFP balancers for developmental analysis: *FM7c 2xTb-RFP, CyO 2xTb-RFP,* and *TM3 2xTb-RFP*. genesis 50: 119–123.

Radford, S. J., S. McMahan, H. L. Blanton, and J. Sekelsky, 2007 Heteroduplex DNA in meiotic recombination in Drosophila mei-9 mutants. Genetics 176: 63–72.

Ramel, C., 1966 The interchromosomal effect of inversions on crossing-over in relation to non homologous pairing in *Drosophila melanogaster*. Hereditas 54: 293–306.

Remnant, E. J., R. T. Good, J. M. Schmidt, C. Lumb, C. Robin *et al.*, 2013 Gene duplication in the major insecticide target site, *Rdl*, in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences 110: 14705–14710.

Richards, S., Y. Liu, B. R. Bettencourt, P. Hradecky, S. Letovsky *et al.*, 2005 Comparative genome sequencing of Drosophila pseudoobscura: chromosomal, gene, and cis-element evolution. Genome Research 15: 1–18.

Rosenbloom, K. R., J. Armstrong, G. P. Barber, J. Casper, H. Clawson *et al.*, 2015 The UCSC Genome Browser database: 2015 update. Nucleic Acids Research 43: D670–81.

Rosu, S., K. A. Zawadzki, E. L. Stamper, D. E. Libuda, A. L. Reese *et al.*, 2013 The C. elegans DSB-2 Protein Reveals a Regulatory Network that Controls Competence for Meiotic DSB Formation and Promotes Crossover Assurance (J. Sekelsky, Ed.). PLoS Genet 9: e1003674–23.

Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132: 365–386.

Santos, dos, G., A. J. Schroeder, J. L. Goodman, V. B. Strelets, M. A. Crosby *et al.*, 2014 FlyBase: introduction of the Drosophila melanogaster Release 6 reference genome assembly and large-scale migration of genome annotations. Nucleic Acids Research.

Schaeffer, S. W., A. Bhutkar, B. F. McAllister, M. Matsuda, L. M. Matzkin *et al.*, 2008 Polytene Chromosomal Maps of 11 Drosophila Species: The Order of Genomic Scaffolds Inferred From Genetic and Physical Maps. Genetics 179: 1601–1655.

Schultz, J., and H. Redfield, 1951 Interchromosomal effects on crossing over in Drosophila. Cold Spring Harb. Symp. Quant. Biol. 16: 175–197.

Shearn, A., 1980 Reintroduction of y+ onto a TM3 chromosome. DIS 55: 167.

Sidorov, B. N., 1931 A study of step-allelomorphism in *Drosophila melanogaster*. A case of origination of an allelomorph of *scute* producing simultaneously characters of “hairy wing”

(mutation scute-8). Zh. eksp. Biol. Med. 7: 28–40.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus *et al.*, 2011 Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7: 1–6.

Singh, N. D., E. A. Stone, C. F. Aquadro, and A. G. Clark, 2013 Fine-scale heterogeneity in crossover rate in the *garnet-scalloped* region of the *Drosophila melanogaster X* chromosome. Genetics 194: 375–387.

Singhal, S., E. M. Leffler, K. Sannareddy, I. Turner, O. Venn *et al.*, 2015 Stable recombination hotspots in birds. Science 350: 928–932.

Snyder, M. P., D. Kimbrell, M. Hunkapiller, R. Hill, J. Fristrom *et al.*, 1982 A transposable element that splits the promoter region inactivates a Drosophila cuticle protein gene. Proc. Natl. Acad. Sci. U.S.A. 79: 7430–7434.

Stevison, L. S., and M. A. F. Noor, 2010 Genetic and evolutionary correlates of fine-scale recombination rate variation in *Drosophila persimilis*. J Mol Evol 71: 332–345.

Stone, W., and I. Thomas, 1935 Crossover and disjunctional properties of X chromosome inversions in *Drosophila melanogaster*. Genetica 17: 170–184.

Sturtevant, A. H., 1913 A third group of linked genes in *Drosophila ampelophila*. Science 37: 990–992.

Sturtevant, A. H., 2001 *Reminiscences of T. H. Morgan.* Genetics Society of America.

Sturtevant, A. H., 1915 The behavior of the chromosomes as studied through linkage. Z.Ver-erbungslehre 13: 234–287.

Sturtevant, A. H., 1925 The effects of unequal crossing over at the Bar locus in Drosophila. Genetics 10: 117–147.

Sturtevant, A. H., and G. W. Beadle, 1936 The relations of inversions in the *X* chromosome of *Drosophila melanogaster* to crossing over and disjunction. Genetics 21: 554–604.

Sturtevant, A. H., and T. H. Morgan, 1923 Reverse mutation of the Bar gene correlated with crossing over. Science 57: 746–747.

Sutton, E., 1943 A Cytogenetic Study of the Yellow-Scute Region of the X Chromosome in *Drosophila melanogaster*. Genetics 28: 210–217.

Sym, M., and G. S. Roeder, 1994 Crossover interference is abolished in the absence of a synaptonemal complex protein. Cell 79: 283–292.

Syrzycka, M., L. A. McEachern, J. Kinneard, K. Prabhu, K. Fitzpatrick *et al.*, 2007 The pink gene encodes the Drosophila orthologue of the human Hermansky-Pudlak syndrome 5 (HPS5) gene. Genome 50: 548–556.

Tartof, K. D., and I. G. Dawid, 1976 Similarities and differences in the structure of *X* and *Y* chromosome rRNA genes of Drosophila. Nature 263: 27–30.

Thacker, D., N. Mohibullah, X. Zhu, and S. Keeney, 2014 Homologue engagement controls meiotic DNA break number and distribution. Nature 510: 241–246.

Theurkauf, W. E., and R. S. Hawley, 1992 Meiotic spindle assembly in Drosophila females: behavior of nonexchange chromosomes and the effects of mutations in the *nod* kinesin-like protein. The Journal of Cell Biology 116: 1167–1180.

Thorvaldsdottir, H., J. T. Robinson, and J. P. Mesirov, 2013 Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14: 178–192.

Tice, S. C., 1914 A new sex-linked character in Drosophila. Biol Bull.

Tinderholt, V., 1960 New Mutants Report. DIS 34: 53–54.

True, J. R., J. M. Mercer, and C. C. Laurie, 1995 Differences in crossover frequency and distribution among three sibling species of Drosophila. Genetics 142: 507–523.

Tsubota, S. I., D. Rosenberg, H. Szostak, D. Rubin, and P. Schedl, 1989 The cloning of the Bar region and the B breakpoint in *Drosophila melanogaster*: evidence for a transposon-induced rearrangement. Genetics 122: 881–890.

Vincenten, N., L.-M. Kuhl, I. Lam, A. Oke, A. R. Kerr *et al.*, 2015 The kinetochore prevents centromere-proximal crossover recombination during meiosis. eLife 4: e10850.

Watanabe, Y., A. Takahashi, M. Itoh, and T. Takano-Shimizu, 2009 Molecular spectrum of spontaneous de novo mutations in male and female germline cells of *Drosophila melanogaster*. Genetics 181: 1035–1043.

Waterhouse, A. M., J. B. Procter, D. M. A. Martin, M. Clamp, and G. J. Barton, 2009 Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25: 1189–1191.

Weinstein, A., 1918 Coincidence of crossing over in *Drosophila melanogaster* (*Ampelophila*). Genetics 3: 135–172.

Whitby, M. C., 2005 Making crossovers during meiosis. Biochem. Soc. Trans. 33: 1451.

Xiang, Y., D. E. Miller, E. J. Ross, A. Sánchez Alvarado, and R. S. Hawley, 2014 Synaptonemal

complex extension from clustered telomeres mediates full-length chromosome pairing in Schmidtea mediterranea. Proceedings of the National Academy of Sciences 111: E5159-68.

Yamamoto, M., and G. L. Miklos, 1977 Genetic dissection of heterochromatin in Drosophila: the role of basal *X* heterochromatin in meiotic sex chromosome behaviour. Chromosoma 60: 283–296.

Yamamoto, M., and G. L. Miklos, 1978 Genetic studies on heterochromatin in *Drosophila melanogaster* and their implications for the functions of satellite DNA. Chromosoma 66: 71–98.

Yazdani, U., Z. Huang, and J. R. Terman, 2008 The glucose transporter (GLUT4) enhancer factor is required for normal wing positioning in Drosophila. Genetics 178: 919–929.

Zeleny, C., 1921 The direction and frequency of mutation in the bar-eye series of multiple allelomorphs of Drosophila. J. Exp. Zool. 34: 202–233.

Zhang, L., E. Espagne, A. de Muyt, D. Zickler, and N. E. Kleckner, 2014 Interference-mediated synaptonemal complex formation with embedded crossover designation. Proceedings of the National Academy of Sciences 111: E5059–68.

Evolution of genes and genomes on the Drosophila phylogeny., 2007 Evolution of genes and genomes on the Drosophila phylogeny. Nature 450: 203–218.

Appendix A: Publications authored prior to graduate school

SAIDE: A Semi-Automated Interface for Hydrogen/Deuterium Exchange Mass Spectrometry

Villar, M. T., D. E. Miller, A. W. Fenton, and A. Artigues, 2010 SAIDE: A Semi-Automated Interface for Hydrogen/Deuterium Exchange Mass Spectrometry. Proteomica 6: 63–69.

My contribution to this project was to help with the use and troubleshooting of the flow control apparatus described in this paper. I also helped collect some of the MS data reported in the manuscript that was used to demonstrate the accuracy of the cooling system.

Abstract

Deuterium/hydrogen exchange in combination with mass spectrometry (DH MS) is a sensitive technique for detection of changes in protein conformation and dynamics. Since temperature, pH and timing control are the key elements for reliable and efficient measurement of hydrogen/ deuterium content in proteins and peptides, we have developed a small, semiautomatic interface for deuterium exchange that interfaces the HPLC pumps with a mass spectrometer. This interface is relatively inexpensive to build, and provides efficient temperature and timing control in all stages of enzyme digestion, HPLC separation and mass analysis of the resulting peptides. We have tested this system with a series of standard tryptic peptides reconstituted in a solvent containing increasing concentration of deuterium. Our

results demonstrate the use of this interface results in minimal loss of deuterium due to back exchange during HPLC desalting and separation. For peptides reconstituted in a buffer containing 100% deuterium, and assuming that all amide linkages have exchanged hydrogen with deuterium, the maximum loss of deuterium content is only 17% of the label, indicating the loss of only one deuterium molecule per peptide.

HDXFinder: Automated Analysis and Data Reporting of Deuterium/Hydrogen Exchange Mass Spectrometry

Miller, D. E., C. B. Prasannan, M. T. Villar, A. W. Fenton, and A. Artigues, 2011 HDXFinder: Automated Analysis and Data Reporting of Deuterium/Hydrogen Exchange Mass Spectrometry. J. Am. Soc. Mass Spectrom. 23: 425–429.

I worked on this project as an intern in the laboratory of Aron Fenton, working closely with Antonio Artigues at the KU Mass Spectrometry core facility. This paper was published during my first year of medical school. My contribution to this project was to write, from scratch, the software used to identify and characterize isotopic envelope changes in Hydrogen/Deuterium exchange experiments.

Abstract

Hydrogen/deuterium exchange in combination with mass spectrometry (H/D MS) is a sensitive technique for detection of changes in protein conformation and dynamics. However, wide application of H/D MS has been hindered, in part, by the lack of computational tools necessary for efficient analysis of the large data sets associated with this technique. We report a novel web-based application for automatic analysis of H/D MS experimental data. This application relies on the high resolution of mass spectrometers to extract all isotopic envelopes before correlating these envelopes with individual peptides. Although a fully automatic analysis is

possible, a variety of graphical tools are included to aid in the verification of correlations and rankings of the isotopic peptide envelopes. As a demonstration, the rate constants for H/D exchange of peptides from rabbit muscle pyruvate kinase are mapped onto the structure of this protein.

A whole-chromosome analysis of meiotic recombination in *Drosophila melanogaster*.

Miller, D. E., S. Takeo, K. Nandanan, A. Paulson, M. M. Gogol *et al.*, 2012 A whole-chromosome analysis of meiotic recombination in *Drosophila melanogaster*. G3 (Bethesda) 2: 249–260.

I worked on this project as an intern and technician in the lab of R. Scott Hawley. The paper was published during my second year of medical school.

Abstract

Although traditional genetic assays have characterized the pattern of crossing over across the genome in Drosophila melanogaster, these assays could not precisely define the location of crossovers. Even less is known about the frequency and distribution of noncrossover gene conversion events. To assess the specific number and positions of both meiotic gene conversion and crossover events, we sequenced the genomes of male progeny from females heterozygous for 93,538 X chromosomal single-nucleotide and InDel polymorphisms. From the analysis of the 30 F1 hemizygous X chromosomes, we detected 15 crossover and 5 noncrossover gene conversion events. Taking into account the nonuniform distribution of polymorphism along the chromosome arm, we estimate that most oocytes experience 1 crossover event and 1.6 gene conversion events per X chromosome pair per meiosis. An extrapolation to the entire genome would predict approximately 5 crossover events and 8.6 conversion events per meiosis. Mean gene conversion tract lengths were estimated to be 476 base pairs, yielding a per nucleotide

conversion rate of 0.86 x 10^5 per meiosis. Both of these values are consistent with estimates of conversion frequency and tract length obtained from studies of rosy, the only gene for which gene conversion has been studied extensively in Drosophila. Motif-enrichment analysis revealed a GTGGAAA motif that was enriched near crossovers but not near gene conversions. The low-complexity and frequent occurrence of this motif may in part explain why, in contrast to mammalian systems, no meiotic crossover hotspots have been found in Drosophila.

Appendix B: Publications authored during graduate school

Bisphenol A and the primate ovary

Miller, D. E., and R. S. Hawley, 2012 Bisphenol A and the primate ovary. Proceedings of the National Academy of Sciences 109: 17315–17316.

This was an invited commentary on the manuscript “Bisphenol A alters early oogenesis and follicle formation in the fetal ovary of the rhesus monkey” (Hunt *et al.* 2012). I wrote the first draft of the commentary and worked through revisions with Scott.

Binding of Drosophila Polo kinase to its regulator Matrimony is noncanonical and involves two separate functional domains

Bonner, A. M., S. E. Hughes, J. A. Chisholm, S. K. Smith, B. D. Slaughter, J. R. Unruh, K. A. Collins, J. M. Friederichs, L. Florens, S. K. Swanson, M. C. Pelot, D. E. Miller, M. P. Washburn, S. L. Jaspersen, and R. S. Hawley, 2013 Binding of Drosophila Polo kinase to its regulator Matrimony is noncanonical and involves two separate functional domains. Proceedings of the National Academy of Sciences 110: E1222–31.

For this manuscript I interpreted the mass spectrometry data shown in table S2 that demonstrated that amino acid S52 from matrimony expressed in *S. cerevisiae* was phosphorylated.

Abstract

Drosophila melanogaster Polo kinase physically interacts with, and is repressed by, the Matrimony (Mtrm) protein during oogenesis. Females heterozygous for a deletion of the mtrm gene display defects in chromosome segregation at meiosis I. However, a complete absence of Mtrm results in both meiotic catastrophe and female sterility. We show that three phosphorylated residues in an N-terminal region in Mtrm are required for Mtrm::Polo binding. However, this binding is noncanonical; it does not require either a complete S-pS/pT-P motif in Mtrm or key residues in the Polo- box domain of Polo that allow Polo to bind phosphorylated

substrates. By using fluorescence cross-correlation spectroscopy to characterize the Mtrm::Polo interaction in vivo, we show that a sterile α-motif (SAM) domain located at the C terminus of Mtrm increases the stability of Mtrm::Polo binding. Although Mtrm's C-terminal SAM domain is not required to rescue the chromosome segregation defects observed in mtrm/+ females, it is essential to prevent both meiotic catastrophe and the female sterility observed in mtrm/mtrm females. We propose that Polo's interaction with the cluster of phosphorylated residues alone is sufficient to rescue the meiosis I defect. However, the strengthening of Mtrm:: Polo binding mediated by the SAM domain is necessary to prevent meiotic catastrophe and ensure female fertility. Characterization of the Mtrm::Polo interaction, as well as that of other Polo regulators, may assist in the design of a new class of Polo inhibitors to be used as targeted anticancer therapeutic agents.

Discovery of supernumerary B chromosomes in *Drosophila melanogaster*

Bauerly, E., S. E. Hughes, D. R. Vietti, D. E. Miller, W. McDowell, and R. S. Hawley, 2014 Discovery of supernumerary B chromosomes in *Drosophila melanogaster*. Genetics 196: 1007–1016.

This manuscript reported the discovery and characterization of supernumerary B chromosomes in *Drosophila melanogaster*. My role in this project was to design and help perform the qPCR experiments that validated that B chromosomes did not contain 4^{th} chromosome euchromatin, as well as to perform the modeling which suggested scenarios in which B chromosomes may be lost or maintained in stocks.

Abstract

B chromosomes are small, heterochromatic chromosomes that are transmitted in a non-Mendelian manner. We have identified a stock of Drosophila melanogaster that recently (within the last decade) acquired an average of 10 B chromosomes per fly. These B chromosomes are transmitted by both males and females and can be maintained for multiple generations in a wild-type genetic background despite the fact that they cause high levels of 4th chromosome meiotic nondisjunction in females. Most curiously, these B chromosomes are mitotically unstable, suggesting either the absence of critical chromosomal sites or the inability of the meiotic or mitotic systems to cope with many additional chromosomes. These B

chromosomes also contain centromeres and are primarily composed of the heterochromatic AATAT satellite sequence. Although the AATAT sequence comprises the majority of the 4th chromosome heterochromatin, the B chromosomes lack most, if not all, 4th chromosome euchromatin. Presumably as a consequence of their heterochromatic content, these B chromosomes significantly modify position-effect variegation in two separate reporter systems, acting as enhancers of variegation in one case and suppressors in the other. The identification of B chromosomes in a genetically tractable organism like D. melanogaster will facilitate studies of chromosome evolution and the analysis of the mech- anisms by which meiotic and mitotic processes cope with additional chromosomes.

Corolla Is a Novel Protein that Contributes to the Architecture of the Synaptonemal Complex of Drosophila

Collins, K. A., J. R. Unruh, B. D. Slaughter, Z. Yu, C. M. Lake, R. J. Nielsen, K. S. Box, D. E. Miller, J. P. Blumenstiel, A. G. Perera, K. E. Malanowski, and R. S. Hawley, 2014 Corolla is a novel protein that contributes to the architecture of the synaptonemal complex of Drosophila. Genetics 198: 219–228.

This paper reported the characterization of a novel SC component, Corolla, in Drosophila melanogaster. My specific role was to identify the mutant using whole-genome sequencing data and perform the conservation analysis between Corolla and *C. elegans* SYP-4 seen in Figure S4.A.

Abstract

In most organisms the synaptonemal complex (SC) connects paired homologs along their entire length during much of meiotic prophase. To better understand the structure of the SC, we aim to identify its components and to determine how each of these components contributes to SC function. Here, we report the identification of a novel SC component in Drosophila melanogaster female oocytes, which we have named Corolla. Using structured illumination microscopy, we demonstrate that Corolla is a component of the central region of the SC. Consistent with its localization, we show by yeast two-hybrid analysis that Corolla strongly

interacts with Cona, a central element protein, demonstrating the first direct interaction between two inner-synaptonemal complex proteins in Drosophila. These observations help provide a more complete model of SC structure and function in Drosophila females.

Tetrad analysis in the mouse

Miller, D. E., and R. S. Hawley, 2014 news and views. Nature Genetics 46: 1045–1046.

This was an invited commentary for the paper “Mouse tetrad analysis provides insights into recombination mechanisms and hotspot evolutionary dynamics“(Cole *et al.* 2014). I wrote the first draft of the commentary and worked through revisions with Scott.

Synaptonemal complex extension from clustered telomeres mediates full-length chromosome pairing in *Schmidtea mediterranea*

Xiang, Y., D. E. Miller, E. J. Ross, A. Sánchez Alvarado, and R. S. Hawley, 2014 Synaptonemal complex extension from clustered telomeres mediates full-length chromosome pairing in *Schmidtea mediterranea*. Proceedings of the National Academy of Sciences 111: E5159-68.

This manuscript reported how SC forms in the flatworm *Schmideta mediterranea*. My contribution to this project was to identify the meiosis-specific proteins in *S. mediterranea* characterized in this report and to perform the protein conservation analysis discussed throughout the manuscript and shown in supplemental data.

Abstract

In the 1920s, József Gelei proposed that chromosome pairing in flatworms resulted from the formation of a telomere bouquet followed by the extension of synapsis from telomeres at the base of the bouquet, thus facilitating homolog pairing in a processive manner. A modern interpretation of Gelei's model postulates that the synaptonemal complex (SC) is nucleated close to the telomeres and then extends progressively along the full length of chromo- some arms. We used the easily visible meiotic chromosomes, a well-characterized genome, and RNAi in the sexual biotype of the planarian *Schmidtea mediterranea* to test that hypothesis. By identifying and characterizing *S. mediterranea* homologs of genes encoding synaptonemal

complex protein 1 (SYCP1), the topoisomerase-like protein SPO11, and RAD51, a key player in homologous recombination, we confirmed that SC formation begins near the telomeres and progresses along chromosome arms during zygotene. Although distal regions pair at the time of bouquet formation, pairing of a unique interstitial locus is not observed until the for- mation of full-length SC at pachytene. Moreover, neither full extension of the SC nor homologous pairing is dependent on the formation of double-strand breaks. These findings validate Gelei's speculation that full-length pairing of homologous chromosomes is mediated by the extension of the SC formed near the telomeres. *S. mediterranea* thus becomes the first organism described (to our knowledge) that forms a canonical telomere bouquet but does not require double-strand breaks for synapsis between homologous chromosomes. However, the initiation of SC formation at the base of the telomere bouquet, which then is followed by full-length homologous pairing in planarian spermatocytes, is not observed in other species and may not be conserved.

Dynamics of *Wolbachia pipientis* Gene Expression Across the *Drosophila melanogaster* Life Cycle

Gutzwiller, F., C. R. Carmo, D. E. Miller, D. W. Rice, I. L. G. Newton, R. S. Hawley, L. Teixeira, C. M. Bergman, 2015 Dynamics of *Wolbachia pipientis* Gene Expression Across the *Drosophila melanogaster* Life Cycle. G3 (Bethesda) 5: 2843–2856.

This manuscript reported the presence of the endosymbotic *Wolbachia pipientis* transcriptome in the *Drosophila melanogaster* modEncode data. This finding means that the complete life-cycle of this endosymboint is present in this data—an unprescendented observation. My contribution to this project was the initial observation that Wolbachia was present in the stock used in the modEncode project as well as in helping deterime exactly how the stock used in the modEncode project differed from the stock used for the Drosophila reference genome.

Abstract

Symbiotic interactions between microbes and their multicellular hosts have manifold biological consequences. To better understand how bacteria maintain symbiotic associations with animal hosts, we analyzed genome-wide gene expression for the endosymbiotic a-proteobacteria Wolbachia pipientis across the entire life cycle of Drosophila melanogaster. We found that the majority of Wolbachia genes are expressed stably across the D. melanogaster life cycle, but that 7.8% of Wolbachia genes exhibit robust stage- or sex- specific expression differences when

studied in the whole-organism context. Differentially-expressed Wolbachia genes are typically up-regulated after Drosophila embryogenesis and include many bacterial membrane, secretion system, and ankyrin repeat-containing proteins. Sex-biased genes are often organized as small operons of uncharacterized genes and are mainly up-regulated in adult Drosophila males in an age-dependent manner. We also systematically investigated expression levels of previously-reported candidate genes thought to be involved in host-microbe interaction, including those in the WO-A and WO-B prophages and in the Octomom region, which has been implicated in regulating bacterial titer and pathogenicity. Our work provides comprehensive insight into the developmental dynamics of gene expression for a widespread endosymbiont in its natural host context, and shows that public gene expression data harbor rich resources to probe the functional basis of the Wolbachia-Drosophila symbiosis and annotate the transcriptional outputs of the Wolbachia genome.

Turner syndrome as a model for understanding sex biases in disease

Miller DE, Page DC. Turner syndrome as a model for understanding sex biases in disease. 2016. *Turner Syndrome Society*.

This is a book chapter that was an adapation of a talk given by David Page at the Turner Syndrome Society of the United States yearly meeting. My role was to adapt the talk into text, include references, and figures where appropriate.

Phosphorylation of the Synaptonemal Complex Protein Zip1 Regulates the Crossover/Noncrossover Decision during Yeast Meiosis

Chen, X., R. T. Suhandynata, R. Sandhu, B. Rockmill, N. Mohibullah, H. Niu, J. Liang, H. Lo, D. E. Miller, H. Zhou, G. V. Börner, and N. M. Hollingsworth, 2015 Phosphorylation of the Synaptonemal Complex Protein Zip1 Regulates the Crossover/Noncrossover Decision during Yeast Meiosis (D. Durocher, Ed.). PLoS Biol 13: e1002329–35.

This manuscript examined how the decision to repair a DSB as a CO or NCO is determined in *Saccharomyces Cerevisiae*. Specifically, I identified a conserved phosphoratable site containing three conserved residues in the transvserse filament protein from several diverse species, which is shown in Figure 1. Identification of sites conserved in the transverse filament protein is difficult, as there is little conservation among species of any SC component (see Chapter 1 for an example of this difficulty).

Abstract

Interhomolog crossovers promote proper chromosome segregation during meiosis and are formed by the regulated repair of programmed double-strand breaks. This regulation requires components of the synaptonemal complex (SC), a proteinaceous structure formed between homologous chromosomes. In yeast, SC formation requires the “ZMM” genes, which encode a functionally diverse set of proteins, including the transverse filament pro- tein, Zip1. In wild-

type meiosis, Zmm proteins promote the biased resolution of recombina- tion intermediates into crossovers that are distributed throughout the genome by interference. In contrast, noncrossovers are formed primarily through synthesis-dependent strand annealing mediated by the Sgs1 helicase. This work identifies a conserved region on the C terminus of Zip1 (called Zip1 4S), whose phosphorylation is required for the ZMM pathway of crossover formation. Zip1 4S phosphorylation is promoted both by double- strand breaks (DSBs) and the meiosis-specific kinase, MEK1/MRE4, demonstrating a role for MEK1 in the regulation of interhomolog crossover formation, as well as interhomolog bias. Failure to phosphorylate Zip1 4S results in meiotic prophase arrest, specifically in the absence of SGS1. This gain of function meiotic arrest phenotype is suppressed by spo11Δ, suggesting that it is due to unrepaired breaks triggering the meiotic recombination check- point. Epistasis experiments combining deletions of individual ZMM genes with sgs1-md zip1-4A indicate that Zip1 4S phosphorylation functions prior to the other ZMMs. These results suggest that phosphorylation of Zip1 at DSBs commits those breaks to repair via the ZMM pathway and provides a mechanism by which the crossover/noncrossover decision can be dynamically regulated during yeast meiosis.

Rare recombination events generate sequence diversity among balancer chromosomes in *Drosophila melanogaster*

Miller, D. E., K. R. Cook, N. Yeganeh Kazemi, C. B. Smith, A. J. Cockrell, R. S. Hawley, and C. M. Bergman, 2016 Rare recombination events generate sequence diversity among balancer chromosomes in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences 113: E1352-61.

This work is presented in chapter 4 of this thesis.

Abstract

Multiply inverted balancer chromosomes that suppress exchange with their homologs are an essential part of the genetic toolkit in *Drosophila melanogaster*. Despite their widespread use, the organization of balancer chromosomes has not been characterized at the molecular level, and the degree of sequence variation among copies of any given balancer chromosome is unknown. To map inversion breakpoints and study potential diversity in the descendants of a structurally identical balancer chromosome, we sequenced a panel of laboratory stocks containing the most widely used *X*-chromosome balancer, *First Multiple 7* (*FM7)*. We mapped the locations of *FM7* breakpoints to precise euchromatic coordinates and identified the flanking sequence of breakpoints in heterochromatic regions. Analysis of SNP variation revealed megabase-scale blocks of sequence divergence among currently used *FM7* stocks. We present evidence that this divergence arose by rare double crossover events that replaced a female-

sterile allele of the *singed* gene (sn^{X2}) on *FM7c* with wild type sequence from balanced chromosomes. We propose that, although double crossover events are rare in individual crosses, many *FM7c* chromosomes in the Bloomington *Drosophila* Stock Center have lost sn^{X2} by this mechanism on a historical timescale. Finally, we characterize the original allele of the *Bar* gene (B^1) that is carried on *FM7* and validate the hypothesis that the origin and subsequent reversion of the B^1 duplication is mediated by unequal exchange. Our results reject a simple non-recombining, clonal mode for the laboratory evolution of balancer chromosomes and have implications for how balancer chromosomes should be used in the design and interpretation of genetic experiments in *Drosophila*.

Whole-Genome Analysis of Individual Meiotic Events in Drosophila melanogaster Reveals that Noncrossover Gene Conversions are Insensitive to Interference and the Centromere Effect

Miller, D. E., C. B. Smith, N. Yeganeh Kazemi, A. J. Cockrell, A. V. Arvanitakis, J. P. Blumenstiel, S. L. Jaspersen, and R. S. Hawley, 2016 Whole-Genome Analysis of Individual Meiotic Events in Drosophila melanogaster Reveals that Noncrossover Gene Conversions are Insensitive to Interference and the Centromere Effect. *Genetics* Early online, March 4, 2016; DOI: 10.1534/genetics.115.186486.

This work is presented in chapter 2 of this thesis.

Abstract

A century of genetic analysis has revealed that multiple mechanisms control the distribution of meiotic crossover events. In *Drosophila melanogaster*, two significant positional controls are interference and the strongly polar centromere effect. Here, we assess the factors controlling the distribution of crossovers (COs) and noncrossover gene conversions (NCOs) along all five major chromosome arms in 196 single meiotic divisions in order to generate a more detailed understanding of these controls on a genome-wide scale. Analyzing the outcomes of single meiotic events allows us to distinguish among different classes of meiotic recombination. In so doing, we identified 291 NCOs spread uniformly among the five major chromosome arms and 541 COs (including 52 double crossovers and one triple crossover). We find that unlike COs,

NCOs are insensitive to the centromere effect and do not demonstrate interference. Although the positions of COs appear to be determined predominately by the long-range influences of interference and the centromere effect, each chromosome may display a different pattern of sensitivity to interference, suggesting that interference may not be a uniform global property. In addition, unbiased sequencing of a large number of individuals allows us to describe the formation of *de novo* copy number variants, the majority of which appear to be mediated by unequal crossing over between transposable elements. This work has multiple implications for our understanding of how meiotic recombination is regulated to ensure proper chromosome segregation and maintain genome stability.