CENTROMERE DRIVE AND SUPPRESSION

BY PARALLEL PATHWAYS FOR RECRUITING MICROTUBULE DESTABILIZERS

Tomohiro Kumon

A DISSERTATION

in

Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation

Michael A. Lampson

Professor of Biology

Graduate Group Chairperson

Brian D. Gregory, Associate Professor of Biology

Dissertation Committee

Mark Goulian, Charles and William L. Day Distinguished Professor in the Natural Sciences

Ben E. Black, Eldridge Reeves Johnson Foundation Professor of Biochemistry and Biophysics

Junhyong Kim, Patricia M. Williams Term Professor and Chair, Biology

Mia Levine, Assistant Professor of Biology

ACKNOWLEDGMENT

I thank Jun Ma and R. Brian Akins for assistance with the cell biological experiments, Derek Stefanik, C. Erik Nordgren, and Junhyong Kim for assistance with the genomic analyses, Mia T. Levine for conceptualization of the evolutionary cell biology project, and Michael A. Lampson for supervision and support in my pursuit of a PhD.

I thank B.E. Black for comments on the manuscript (https://doi.org/10.1101/2020.11.26.400515), Y. Watanabe for the CENP-C and SGO2 antibodies, R. Mallarino (Princeton University) for *Rhabdomys pumilio* animals, J.B. Searle (Cornell University), C. Conroy (Museum of Vertebrate Zoology), and A. Ferguson (Field Museum of Natural History) for collecting Murinae tissue samples, K.G. Boese for assistance with the image analyses, G. Thomas and J. Good (University of Montana) for assistance with the genomic analyses, and the past and present members of the Lampson lab, the Philadelphia Chromosome Club (Black, Good, Grishchuk, and Greenberg labs), my thesis committee, my family and friends for support and advice. The research was supported by the NIH (R35GM122475 to M.A.L.; R35GM124684 to M.T.L.; RM1HG010023 to J.K.) and predoctoral fellowships from the Funai Foundation for Information Technology and University of Pennsylvania (T.K.).

ABSTRACT

CENTROMERE DRIVE AND SUPPRESSION

BY PARALLEL PATHWAYS FOR RECRUITING MICROTUBULE DESTABILIZERS

Tomohiro Kumon

Michael A. Lampson

Selfish centromere DNA sequences bias their transmission to the egg in female meiosis. Evolutionary theory suggests that centromere proteins evolve to suppress costs of this "centromere drive". In hybrid mouse models with genetically different maternal and paternal centromeres, selfish centromere DNA exploits a kinetochore pathway to recruit microtubule-destabilizing proteins that act as drive effectors. We show that such functional differences are suppressed by a parallel pathway for effector recruitment by heterochromatin, which is similar between centromeres in this system. Disrupting the kinetochore pathway with a divergent allele of CENP-C reduces functional differences between centromeres, whereas disrupting heterochromatin by CENP-B deletion amplifies the differences. Molecular evolution analyses using newly sequenced Murinae genomes identify adaptive evolution in proteins in both pathways. We propose that centromere proteins have recurrently evolved to minimize the kinetochore pathway, which is exploited by selfish DNA, relative to the heterochromatin pathway that equalizes centromeres, while maintaining essential functions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Molecular Strategies of Meiotic Cheating and Suppression

1.1. Evolutionary Arms Race at Centromeres

Repetitive DNA comprises the majority of eukaryotic genomes. For example, approximately half of the human genome is composed of repetitive DNA such as transposons and centromeric satellites, which is a sharp contrast to protein coding genes that occupy less than 5% of the genome and gene regulatory sequences that occupy less than 10% of the genome (International Human Genome Sequencing Consortium, 2001; ENCODE Project Consortium, 2020). There is growing evidence that some repetitive DNA is selfish in that it drives, or increases the chance of inheritance at the expense of the host fitness (Burt and Trivers, 2006; Henikoff et al., 2001; Lampson and Black, 2017). While transposons drive by over-replication, centromeric satellites drive by biased segregation in female meiosis (Figure 1.1). The fitness costs imposed by selfish genetic elements are the evolutionary pressure that selects protein variants that suppress costs of drive. Continuous cycles of drive and suppression leads to rapid turnover of repetitive DNA and host suppressor proteins. Here, I focus on evolutionary arms race at centromeres.

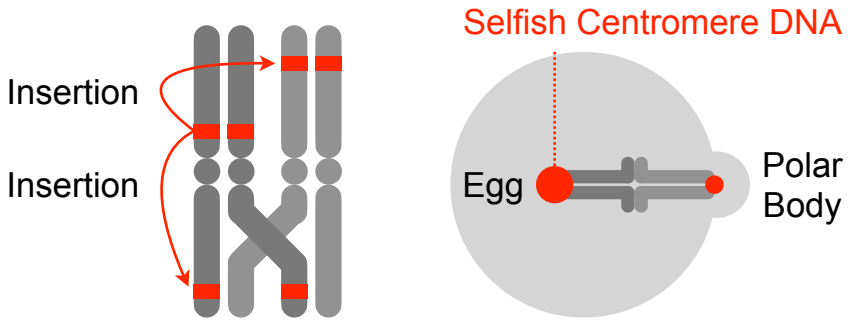


Figure 1.1. Strategies of drive. Transposons drive by over-replication (left). Both intra- and inter-chromosomal insertions can increase the number of gametes that have transposons. Transposon insertion can be deleterious by gene mutation and/or chromosome rearrangement. Centromere DNA drives by biased segregation in female meiosis (right). Selfish centromere DNA increases the chance of segregating into the egg. Selfish centromere DNA can cause chromosome segregation errors.

All eukaryotic cells are the descendants of the preceding cell division and chromosome segregation. Centromeres are the chromosomal regions that assemble kinetochores to attach spindle microtubules for accurate chromosome segregation. Although this centromere function is required for all eukaryotes, forms of centromere architecture, DNA and proteins are diverse. Centromere architectures are distinguished by the position and number of microtubule assembly sites (Figure 1.2). When the spindle microtubules bind kinetochores near telomeres, such chromosomes are called telocentric (also called acrocentric), whereas in other cases chromosomes are called metacentric. When the spindle microtubules assemble at the single CENP-A chromatin locus as in humans and yeasts, such chromosomes are called monocentric. When the spindle microtubules assemble at multiple CENP-A chromatin loci or all over the chromosome, such chromosomes are called polycentric or holocentric, respectively (Cuacos et al., 2015; Mandrioli and Manicardi, 2020).

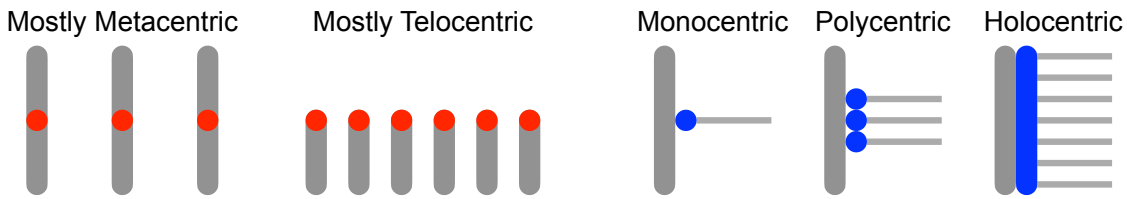


Figure 1.2. Diversity in centromere architectures. Positions and numbers of microtubule assembly sites vary between species. Red circles represent the position of centromere DNA on a chromosome, and blue circles represent centromere-binding proteins that are attached to spindle microtubules. Some species (e.g., humans) have mostly metacentric chromosomes, while others (e.g., mice) have mostly telocentric chromosomes. Depending on the genetic or epigenetic centromere determination, monocentric chromosomes are also called point or regional centromeres, respectively. Polycentric chromosomes have multiple microtubule assembly sites, and holocentric chromosomes attach to microtubules all over the chromosome.

Centromere DNA sequences, which are defined by the presence of functional kinetochores, are often repetitive DNA such as satellite DNA and transposons. Monomer sequences and abundance of centromeric satellites diverge between closely related species, and repeat abundance varies even within species (Arora et al., 2021; Cazaux et al., 2013; Iwata-Otsubo et al., 2017; Langley et

al., 2019; Melters et al., 2013). Transposon enrichment at centromeres varies between species (Chang et al., 2019; Nergadze et al., 2018; Rhind et al., 2011). However, the functional significance of centromere DNA is unclear because of the epigenetic determination of centromere identity (Figure 1.3). Most eukaryotic centromeres are epigenetically defined by CENP-A nucleosomes, as opposed to genetically defined centromeres by cis DNA elements, with notable exception in budding yeast (Malik et al., 2009). Indeed, the position of CENP-A chromatin assembly can change without changing the underlying DNA sequences in a process known as centromere repositioning (Amor et al., 2004; Nergadze et al., 2018).

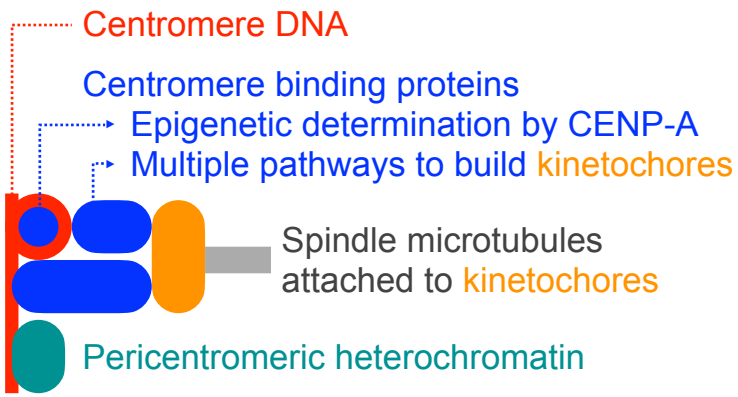


Figure 1.3. Genetic and epigenetic components of centromeres. Centromere DNA (red) is functionally defined by the presence of kinetochores (orange). Centromere binding proteins (constitutive centromere associated network, CCAN proteins; blue) connect centromere DNA and kinetochores. CENP-A nucleosomes epigenetically define the kinetochore assembly position. There are multiple pathways (e.g., CENP-C and CENP-T) to build kinetochores. Pericentromeric heterochromatin (green) flanks CENP-A chromatin.

In addition to centromere DNA, centromere-binding proteins are rapidly evolving. Centromere-binding proteins include constitutive centromere associated network (CCAN) proteins, kinetochore proteins, and inner centromere proteins. Centromeres are more enriched for rapidly evolving proteins than any other subcellular compartments (Kumon et al., 2021), and signatures of adaptive evolution are detected in centromere proteins from multiple eukaryotic lineages (Finseth et al., 2015; Kumon et al., 2021; Malik and Henikoff, 2001; Schueler et al., 2010). One

eukaryotic lineage even uses a distinct set of kinetochore proteins that are not homologous to any other eukaryotic lineages (Akiyoshi et al., 2014). In contrast to prokaryotic chromosome segregation machinery (Badrinarayanan et al., 2015; Barillà, 2016), eukaryotic centromeres are more complex (Figure 1.3). Eukaryotic centromeres have multiple pathways to build a kinetochore, such as the CENP-ACHIKMLN, CENP-TWSX, and CENP-OPQUR pathways (Veld et al., 2016; Pesenti et al., 2018; Weir et al., 2016). Centromeric CENP-A chromatin is flanked with pericentromeric heterochromatin in most eukaryotes (Janssen et al., 2018). The centromere drive hypothesis provides a model to explain paradoxical rapid evolution of complex eukaryotic centromeres.

The centromere drive hypothesis proposes that centromere DNA acts as a selfish genetic element that violates Mendel's law of segregation (Figure 1.1). In female meiosis, homologous chromosomes pair and segregate into the egg or the polar body with 50% probability. The polar body is degraded, so it is an evolutionary dead-end. Thus, any selfish genetic element that increases the chance of segregating into the egg will increase its allele frequency in a population. Centromere DNA is a prime candidate for such a selfish element, as it is the chromosomal region that assembles the segregation machinery. If a centromere DNA sequence evolves to influence interactions with segregation machinery, it can increase its chance of inheritance. This selfish behavior is predicted to have fitness costs, which select centromere-binding protein variants that suppress the costs. This implies that selfish centromere DNA favors genetic centromere determination for its own preferential inheritance, whereas centromere binding proteins favor epigenetic centromere determination for suppressing the fitness costs.

1.2. Mechanisms of Drive

Non-Mendelian segregation of selfish centromeres in female meiosis is studied in mice and monkeyflower (Akera et al., 2017, 2019; Chmátal et al., 2014; Finseth et al., 2021; Fishman et al., 2008; Iwata-Otsubo et al., 2017). Transmission bias of selfish centromere DNA in progeny is reported in monkeyflower (Fishman et al., 2008). In mouse oocytes, selfish centromeres preferentially orient to the egg side of the meiotic spindle before anaphase, implying biased segregation in female meiosis. Due to the available tools for genetic manipulations and cell biology, mechanisms of drive are well characterized in mice (Akera et al., 2017, 2019; Chmátal et al., 2014; Iwata-Otsubo et al., 2017). Conceptually, drive depends on coupling of three asymmetries: fate asymmetry, spindle asymmetry, and centromere asymmetry (Figure 1.4).

In many eukaryotic lineages, meiosis in one sex is asymmetric in that only one cell stores nutrients and produces a functional gamete, whereas the other haploid cells are degraded and therefore evolutionary dead-ends (Gorelick et al., 2016). This meiotic fate asymmetry is maintained in many eukaryotic lineages, likely to increase the genetic diversity in gametes. The number of eggs from a single female is much less than the number of sperm from a single male, and accordingly, female meiosis is less frequent than male meiosis. Because each meiosis creates different meiotic recombination sites, one egg produced per meiosis (rather than four) maximizes genetic diversity. This fate asymmetry creates an opportunity to cheat for any selfish genetic element that increases the chance of segregating into the egg.

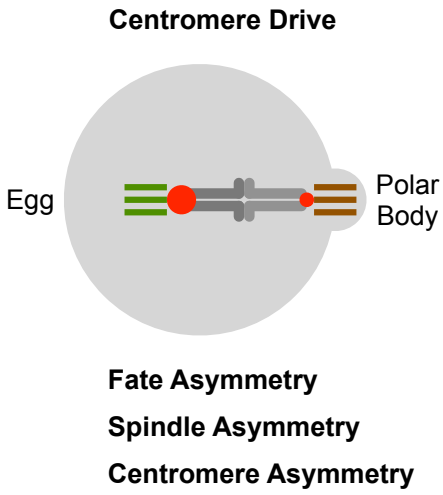


Figure 1.4. Three asymmetries for centromere drive. Selfish centromere DNA creates functional asymmetry in centromeres (as represented by different sizes of red circles). Selfish centromere DNA (larger red circle) preferentially orients on an asymmetric spindle (as represented by different colors in spindle), which orients relative to the cortex where the polar body forms (fate asymmetry).

Spindle asymmetry in female meiosis is present in many species (Akera et al., 2017; Crowder et al., 2015; Hewitt, 1976). In mouse oocytes, this spindle asymmetry is intrinsically coupled with fate asymmetry. As chromosomes migrate to the cortex, due to the RAN^{GTP} activity from the chromosomes, the cortex is polarized and sends $CDC42^{GTP}$ signaling. The $CDC42^{GTP}$ signaling from the cortex creates asymmetry in the post-translational modification on meiotic spindle (Figure 1.5, left) (Akera et al., 2017). The cortical side of the meiotic spindle is enriched for tyrosinated microtubules, whereas the egg side is enriched for detyrosinated microtubules. Asymmetric meiotic spindle morphology is observed in many organisms (Crowder et al., 2015), suggesting that spindle asymmetry is a common feature of meiotic spindles. Functions of spindle asymmetry are unclear, and it is also possible that spindle asymmetry is an unavoidable byproduct of establishing cellular asymmetry (e.g., cortical polarization) necessary for asymmetric cell division. Given that spindle asymmetry is coupled with fate asymmetry, selfish genetic elements can exploit spindle asymmetry to drive. Indeed, asymmetric post-translational

modification of meiotic spindle is required for biased orientation of selfish centromeres to the egg side of the spindle in mice (Akera et al., 2017). Asymmetric morphology of meiotic spindle is observed in grasshopper, and B chromosomes (See Section 3.6) likely exploit this asymmetry for the preferential inheritance (Hewitt, 1976).

Coupling of fate asymmetry, spindle asymmetry, and centromere asymmetry is necessary for drive. How these asymmetries are coupled is best studied in hybrid mouse systems. In these intra-species (cross of different *Mus musculus domesticus* strains) and inter-species (cross of *Mus musculus* and *Mus spretus*) hybrids, homologous centromeres on meiotic bivalents are genetically different, and the centromere that recruits more effector proteins acts selfishly (Akera et al., 2019). Thus, centromeres are asymmetric in that they recruit different amounts of effectors. Effector proteins are microtubule destabilizers that correct erroneous microtubule attachments, but selfish centromere DNA exploits this activity for its preferential orientation to the egg side of the spindle. Molecular details of how selfish centromere DNA preferentially orients to the egg side is still unclear, but several lines of evidence support the idea that selfish centromere DNA recruits more effector proteins that preferentially destabilize interactions with the cortical side of the spindle that would otherwise direct it to the polar body (Figure 1.5, right). First, selfish centromeres that orient to the cortical side of the spindle are likely to flip to the egg side, suggesting microtubule detachment on the cortical side and reattachment on the egg side (Akera et al., 2019). Second, MCAK preferentially destabilizes tyrosinated microtubules (Peris et al., 2009; Sirajuddin et al., 2014). As the cortical side of the spindle microtubules is more tyrosinated, frequent detachment from the cortical side is consistent with the idea that MCAK preferentially destabilizes interactions with the cortical side.

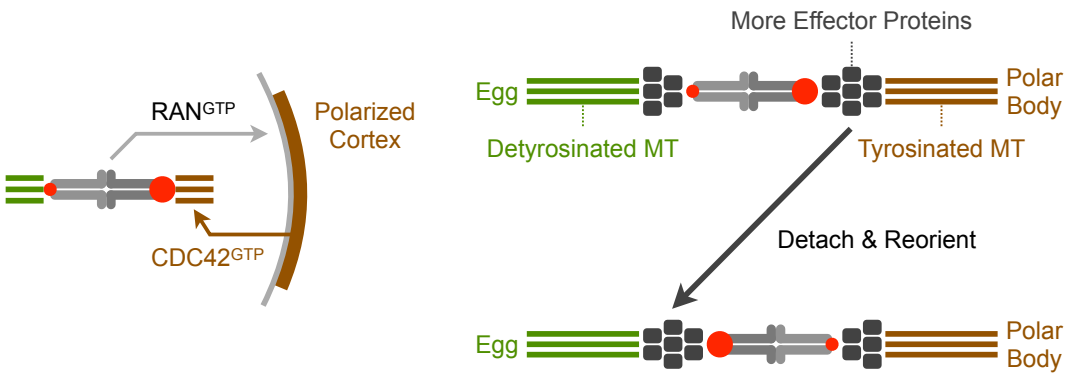


Figure 1.5. Coupling of three asymmetries for centromere drive. As the spindle migrates to the oocyte cortex, The RAN^{GTP} signal from chromosomes polarizes the cortex. The polarized cortex then sends the $CDC42^{GTP}$ signal, which creates the asymmetry post-translational modification in spindle (left). Selfish centromere DNA recruits more effector proteins to drive (represented by the number of black squares). The effector proteins for drive are microtubule destabilizers that are necessary for correcting erroneous kinetochore-microtubule attachments, but selfish centromeres exploit this activity to reorient to the egg side of the spindle (right).

In other eukaryotic lineages with fate and spindle asymmetries, selfish centromeres may use different strategies to interact with spindle microtubules for the preferential segregation to the egg. Selfish centromeres can change the microtubule binding activity as well as microtubule destabilizing activity. Microtubule attachment to kinetochores is necessary for anaphase segregation, and kinetochore-microtubule attachments can be stabilized by, for example, the SKA complex. Microtubule destabilizing activity is also necessary to correct erroneous kinetochore-microtubule attachments, and inner centromere proteins such MCAK and the chromosome passenger complex (CPC) destabilize microtubule attachments. By recruiting more of these microtubule interacting proteins, selfish centromeres can change spindle microtubule attachment and detachment dynamics for preferential inheritance.

Selfish centromeres in mice exploit the microtubule destabilizing activity for their preferential inheritance (Akera et al., 2019). Signals from the kinetochore recruit microtubule destabilizers:

BUB1 at kinetochores phosphorylates pericentromeric histone H2A, phosphorylated H2A recruits Shugoshin-2 (SGO2), and SGO2 recruits microtubule destabilizing proteins (Figure 1.6, kinetochore pathway). There are different strategies to recruit more effector proteins through the kinetochore pathway (Akera et al., 2019). In intra-species *Mus musculus domesticus* hybrids, selfish centromeres with expanded minor satellite DNA repeats assemble more CENP-A chromatin, which forms larger kinetochores with more BUB1 kinase, leading to more effector recruitment (Figure 1.7, intra-species hybrid). In inter-species *Mus musculus* and *Mus spretus* hybrids, selfish *spretus* centromeres with more minor satellites and less major satellites form kinetochores with a similar size to musculus kinetochores. However, *spretus* centromeres recruit more condensin, to form more compact chromatin, leading to more histone substrates for BUB1 and more effector recruitment (Figure 1.7, inter-species hybrid).

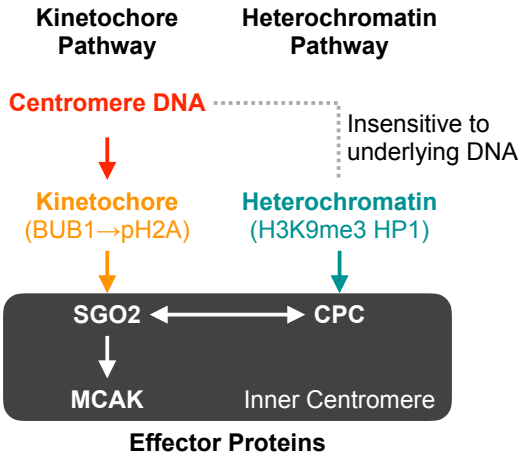


Figure 1.6. Parallel pathways for effector recruitment. Kinetochore-localized BUB1 kinase phosphorylates pericentromeric histone H2A to recruit SGO2. In parallel, pericentromeric heterochromatin also recruits SGO2 via the CPC (chromosome passenger complex) at the inner centromere. In our hybrid mouse model systems, selfish centromere DNA recruits more effector proteins through the kinetochore pathway. In contrast, heterochromatin appears insensitive to the underlying genetic differences.

	Intra-species hybrid (Larger x Smaller)	Inter-species hybrid (Larger x *spretus*)
Genetic Difference	Larger has more minor satellites	*spretus* has more minor satellites *spretus* has less major satellites
Chromatin	Larger has larger CENP-A chromatin Larger forms larger kinetochore More pH2A due to increased BUB1 recruitment	*spretus* recruits more condensin *spretus* chromatin is more compact More pH2A due to compact chromatin

Figure 1.7. Kinetochore pathway asymmetry in hybrid mouse model systems. In the intra-species hybrid where two centromeres have different abundance of minor satellite repeats, larger repeats assemble more CENP-A chromatin, form larger kinetochores, and recruit more effectors. In the inter-species hybrid, two satellite repeats have different abundances. In *musculus* centromeres, centromeric minor satellites are much less abundant than pericentromeric major satellites, whereas *spretus* centromeres have much more abundant minor satellites and much less abundant major satellites compared to *musculus* centromeres. Both centromeres assemble similarly sized kinetochores, but *spretus* centromeres recruit more condensin, leading to more compact chromatin and more histone substrates for BUB1 phosphorylation.

1.3. DNA Evolution to Achieve Preferential Inheritance in Female Meiosis

In the previous examples of centromere drive, genetic differences correlate with functional differences in centromeres, but it remains unclear how selfish centromere DNA creates these functional differences. Due to the epigenetic determination of centromeres, opportunities to genetically recruit centromere binding proteins for preferential inheritance in female meiosis are constrained. Repeat expansion and sequence monomer evolution of centromeric satellite DNA are examples of genetically recruiting more of centromere binding proteins (Figure 1.8 Repeat Expansion and Monomer Sequence Evolution). Throughout this section, recruiting more centromere binding proteins is assumed to benefit selfish centromere DNA for the preferential inheritance. Robertsonian fusion is a unique opportunity to change repeat number and karyotype and will be discussed in Section 3.3. Transposons at centromeres will be discussed in Section 3.5. In addition to the endogenous centromere locus, other genomic loci can evolve to genetically recruit centromere binding proteins for its own inheritance in female meiosis (Figure 1.8 Centromere Repositioning and Alternative MT Interacting Site).

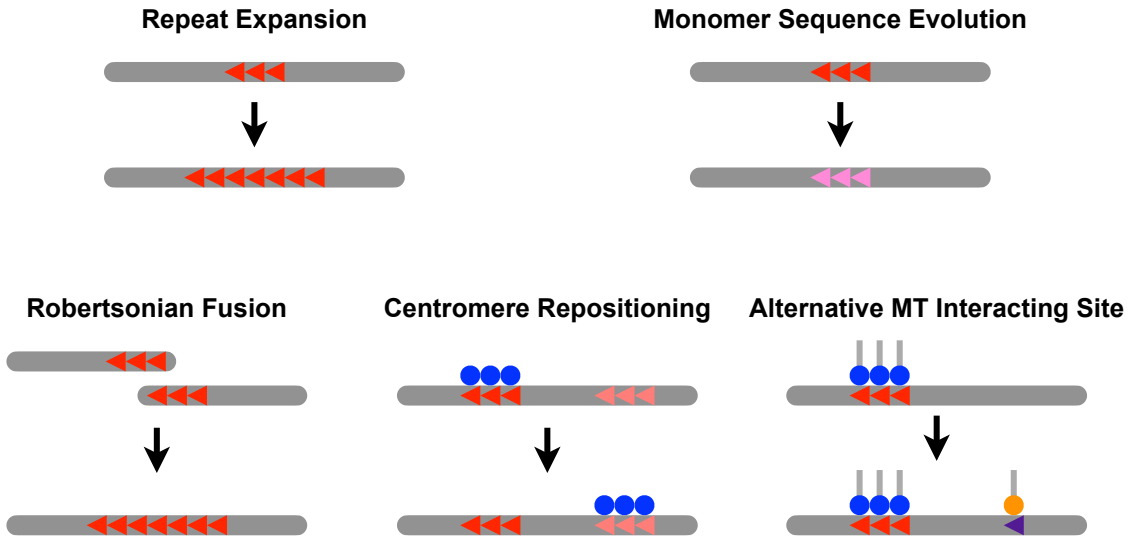


Figure 1.8. DNA evolution to drive in female meiosis. Repeat expansion can provide a space for CENP-A chromatin to drive. Red triangles represent repeat numbers. Monomer sequence evolution can increase affinity with centromere binding proteins. Color change represents

monomer sequence evolution. Robertsonian fusion can lead to increased repeat numbers, but in some cases it leads to reduced repeat numbers. Other genomic loci can acquire centromere function by centromere repositioning. Blue circles represent centromere binding proteins. Other genomic loci can recruit any microtubule binding protein (orange circle).

Repeat expansion can provide a space for CENP-A chromatin expansion (Iwata-Otsubo et al., 2017). However, in intra-species *Mus musculus domesticus* hybrids, despite 6- to 10-fold differences in repeat abundance, CENP-A chromatin on selfish centromere DNA is only 1.5-fold larger than the homologous counterpart. It is possible that CENP-A chromatin expansion is constrained by pericentromeric heterochromatin. Satellite DNA monomer sequence evolves rapidly (Garrido-Ramos, 2017; Melters et al., 2013), due to the lack of constraint (drift by neutral selection) or due to adaptive evolution that increases the chance of inheritance (drive by positive selection). The drive model proposes that new satellite variants that achieve non-Mendelian inheritance will quickly fix in a population. However, the high levels of centromere DNA haplotype diversity in the human populations suggest evolutionary pressures that mitigate rapid fixation of the driving centromere DNA haplotype (Langley et al., 2019). Alternatively, any driving centromere DNA haplotype has not evolved yet in the human populations. Satellite DNA sequences at centromeres are more likely to drive than satellite DNA at other genomic loci where heterochromatin prevents access to the underlying DNA (Janssen et al., 2018). At centromeres, some proteins must interact with DNA, so centromeric satellite DNA can evolve to have different binding affinities or impact the structure of the centromeric nucleosome complex.

Under this model, in sexual species that undergo meiosis, satellite sequences are predicted to be different between populations, but similar within a population. In asexual species, due to the lack of meiosis, satellites sequences are predicted to be as different between populations as they are within a population (Dover, 1986). Satellite DNA sequence diversity observed in the sexual *Bacillus grandii* and parthenogenetic, asexual *Bacillus atticus* is consistent with this prediction

(Luchetti et al., 2003). Furthermore, two homologous centromeres with different abundances of satellite DNA compete in female meiosis, leading to satellite repeat expansion in sexual species. In contrast, due to the lack of meiosis, asexual species are predicted to have less satellite DNA. Indeed, 15-20% of the sexual *Bacillus grandii* genome is composed of satellite DNA, whereas only 2-5% of the asexual *Bacillus atticus* genome is composed of satellite DNA (Mantovani et al., 1997). This observation is consistent with the idea that meiotic conflicts lead to repeat expansion.

Neocentromeres are functional centromeres in ectopic chromosomal regions that are devoid of canonical centromere DNA sequences. Several cases are reported in humans, and some of them are inherited for multiple generations (Hasson et al., 2011). Such neocentromeres will either go extinct or increase their frequency, eventually leading to fixation as evolutionary young centromeres. If neocentromeres are preferentially inherited compared to the homologous counterpart, they will quickly fix (Figure 1.8, Centromere repositioning). Comparative study in mammals revealed that centromere repositioning is frequent (Rocchi et al., 2011), and polymorphism of centromere positions of orangutan chromosomes is reported (Locke et al., 2011). Furthermore, recurrent use of subtelomeric satellite DNA for neocentromere formation is reported. For example, human telocentric chromosomes 14 and 15 are derived from the split of an ancestral metacentric chromosome (Ventura et al., 2003). The ancestral centromere is inactivated, and neocentromeres are formed in subtelomeric regions of both chromosomes. Neocentromeres are often formed in subtelomeres upon endogenous centromere inactivation in fission yeast (Ishii et al., 2008), and it is hypothesized that centromeric repetitive DNA sequences are derived from rapidly evolving subtelomeric repetitive sequences (Villasante et al., 2007). What evolutionary force drives rapid evolution of subtelomeric repetitive DNA is unclear, but recurrent adaptive evolution in telomere and DNA repair proteins is implicated in an evolutionary arms race with

selfish subtelomere DNA (Lee et al., 2017; Saint-Leandre et al., 2020). Inactivated ancestral centromere DNA may have additional functions: the inactivated ancestral centromeric region on human chromosome 15 is associated with a susceptibility factor for anxiety disorders (Gratacòs et al., 2001; Ventura et al., 2003), and recurrent use of ancestral centromeric regions for neocentromere formation is reported (Ventura et al., 2004).

DNA sequences at other genomic loci can recruit microtubule-binding proteins to increase the chance of inheritance in female meiosis (Figure 1.8, Alternative MT interacting site). Non-Mendelian inheritance of abnormal "knob" at the end of maize chromosomes is the earliest discovery of female meiotic drive (Rhoades, 1942). The knob-linked, Kinesin-14A derived Kindr localizes to repetitive DNA on heterochromatic knob region and moves the knob faster than its homologous counterpart to drive (Dawe et al., 2018). Thus, the heterochromatic knob regions acquired kinetochore-like function that do not use canonical segregation machinery. In addition to maize, multiple plant species acquired knobs (Dawe and Hiatt, 2004). In some of these cases, knobs form only in species hybrid. It is hypothesized that knob formation is suppressed in each species, but such suppression mechanisms are compromised in hybrids, leading to knob formation only in hybrids (Dawe and Hiatt, 2004).

1.4. Functionally Different Meiotic and Mitotic Centromeres in Holocentric Species

Localized centromeres are required in meiosis so that homologous chromosomes can recombine and segregate without chromosome breaks, so selfish centromere DNA exploits this constraint to hijack localized centromeres in meiosis. In contrast, mitosis does not have such a constraint, so organisms can have functionally different mitotic centromeres if fitness costs associated with selfish centromere DNA are found in mitosis. Mitotic holocentromeres form kinetochores on the entire chromosome, but meiotic centromeres are constrained by the necessity of meiotic recombination and two-step loss of cohesion, so they form localized centromeres in meiosis (Hughes-Schrader and Schrader, 1961; Monen et al., 2005; Pérez et al., 2000). As mitotic kinetochores are formed all over the chromosomes, the requirement of epigenetic memory for kinetochore assembly may be relaxed. Indeed, many holocentric species have lost CENP-A (Drinnenberg et al., 2014). CENP-A chromatin expansion on selfish centromere DNA is one way to achieve preferential inheritance (Iwata-Otsubo et al., 2017), but such opportunity is lost in holocentric species without CENP-A.

If satellite DNA imposes fitness costs, the host genome can evolve to remove such sequences in mitosis by a process called chromatin diminution. In parasitic nematode *Parascaris univalens*, euchromatic regions are flanked with large blocks of heterochromatic satellite DNA that comprises around 80% of the genome (Goday and Pimpinelli, 1989). Microtubules are attached to the heterochromatic terminal regions during meiosis. In somatic cells, all of the heterochromatic regions are removed by chromatin diminution. During this process, microtubules bind to only euchromatin, resulting in fragmented euchromatic chromosomes (Goday et al., 1992). Another parasitic nematode *Ascaris suum* undergoes chromatin diminution, and genomic regions with reduced CENP-A levels are subsequently removed (Kang et al., 2016). Fragmented chromosomes can be segregated because of mitotic holocentromeres in both species, but fragmented

chromosomes likely impede meiotic recombination, so longer chromosomes are maintained in meiosis.

Holocentromere species still form localized centromeres in meiosis, so selfish DNA can evolve to achieve preferential inheritance in female meiosis. Although a kinetochore-independent mechanism segregates homologous chromosomes during female meiosis in *C. elegans*, kinetochore proteins are required to orient chromosomes properly to the meiotic spindle (Dumont et al., 2010). Thus, there is an opportunity for selfish DNA to bias this orientation process. In another example, *Heteroptera* species form holocentromeres in mitosis, but microtubules attach to one of two ends of a chromosome in meiosis, which is usually randomly determined (Hughes-Schrader and Schrader, 1961; Pérez et al., 2000). However, selfish DNA on either end can evolve to bias this process as well.

1.5. Fitness Costs and Suppressor Evolution

In the original proposal of the centromere drive hypothesis, fitness costs in male meiosis are assumed (Henikoff et al., 2001). However, any fitness costs during the development of female or male gametes, zygotes, or any somatic cells can be a selective pressure for centromere-binding protein variants that suppress the costs. Computational simulation suggests that selfish centromere DNA variants fix much faster than suppressor protein variants (Figure 1.9). This observation implies the existence of fitness costs of selfish centromere DNA in homozygotes to account for the observed adaptive evolution in centromere proteins. In the absence of fitness costs to selfish centromere DNA homozygotes, the time window for suppressor protein variant selection is too limited (Figure 1.10A-B). Consistent with this prediction, male fitness costs as measured by pollen viability and female fitness costs as measured by seed numbers are observed in monkeyflowers that are homozygous for the driving centromere (Fishman et al., 2008; Fishman and Kelly, 2015).

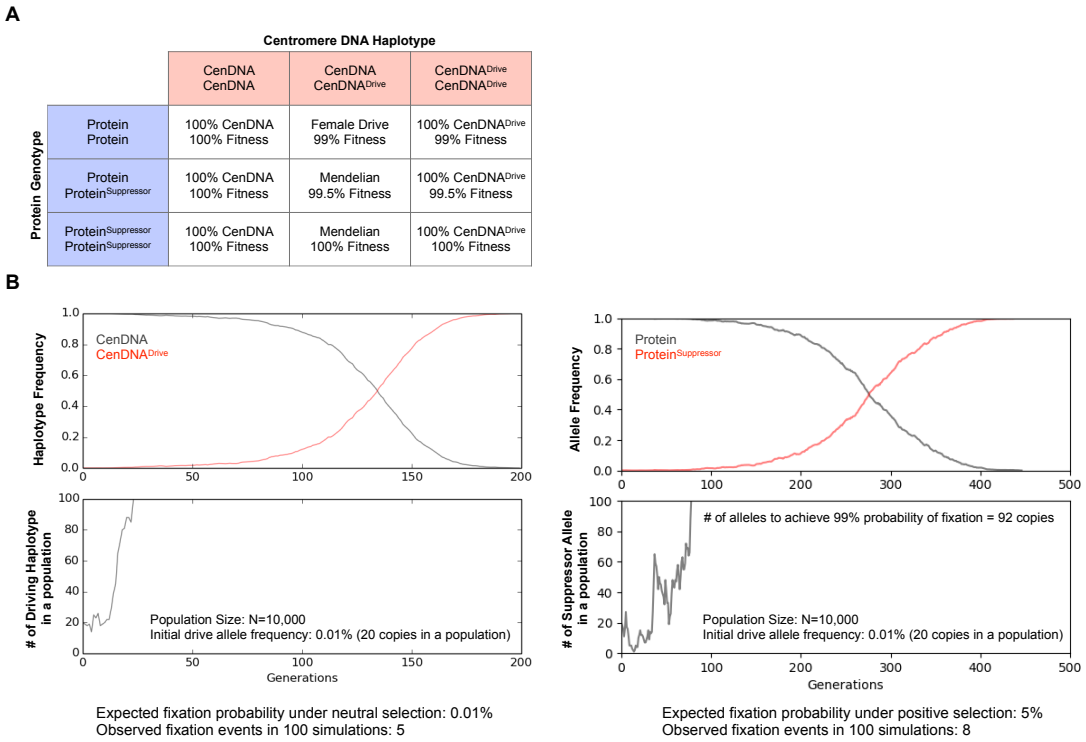


Figure 1.9. (Caption on next page)

Figure 1.9. Selfish centromere DNA fixes faster than suppressor protein variant. (A) Assumptions of simulations. For each combination of centromere DNA haplotype and protein genotype, probability of centromere DNA inheritance and fitness are shown. In the absence of $Protein^{Suppressor}$ allele, $CenDNA^{Drive}$ drives in female meiosis (60% inheritance in female meiosis and 50% inheritance in male meiosis) and imposes fitness costs in both sexes. When an individual is homozygous for $Protein^{Suppressor}$, there is no fitness costs. 60% non-Mendelian inheritance is based on biased orientation of selfish centromeres in hybrid mouse models, and typical population genetics assumptions (s=0.01 and h=0.5) are used for fitness costs. (B) Examples of selfish centromere DNA and suppressor protein variant fixation events. Fixation events of selfish centromere DNA are much more frequent than the expected number of fixation events under neutral selection (fixation by genetic drift). Fixation events of suppressor protein variant is consistent with the expected number of fixation events under positive selection. Selfish centromere DNA fixes rapidly, whereas suppressor protein fixation requires longer time.

Molecular mechanisms of fitness costs are still unclear. When the selfish centromere DNA variant pairs with the preexisting centromere DNA during meiosis, the functional asymmetry between centromeres may cause fitness costs (Figure 1.10.C, left). When the selfish centromere DNA variant spreads to other chromosomes, functional asymmetry between centromeres in meiotic bivalent may cause fitness costs until the selfish centromere DNA variant on these chromosomes fix. Furthermore, having the homozygous selfish centromere DNA variants may cause additional costs (Figure 1.10C, right). For example, selfish centromere DNA variants may sequester centromere binding proteins, and the preexisting centromere DNA may recruit few centromere binding proteins, leading to segregation error. Furthermore, having selfish centromere DNA on its own could be costly, as larger kinetochores are reported to have more merotelic attachments (Drpic et al., 2018).

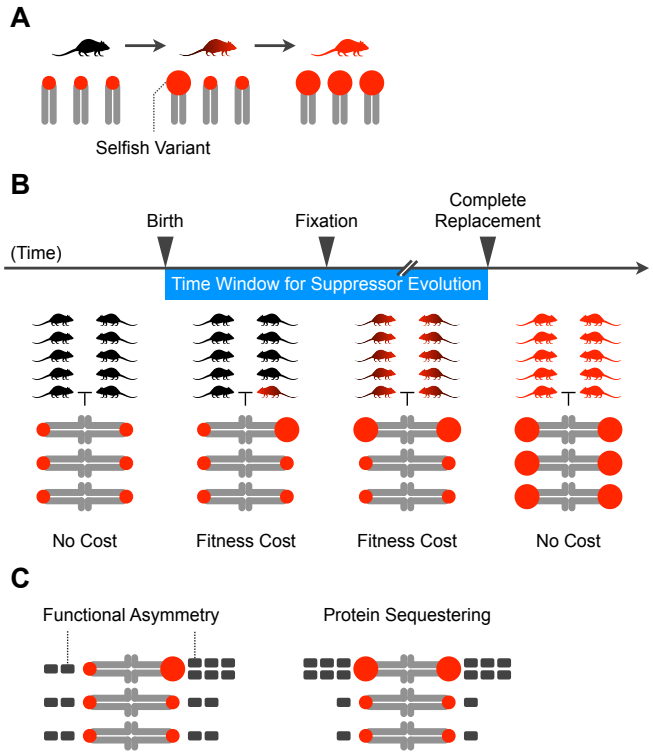


Figure 1.10. Fitness costs of selfish centromere DNA variant. (A) Schematics of chromosomes in gametes. Selfish centromere DNA variant (repeat expansion or sequence evolution that recruits more centromere-binding proteins) first evolves in one of chromosomes. Such selfish DNA variant can spread to other chromosomes. (B) Timescale of selfish centromere DNA variant evolution and suppressor protein evolution. Mice represent individuals in a population, and meiotic bivalents of offspring from two individuals in a population are shown. In the beginning, all individuals have the same centromere DNA, and there is no fitness costs. A selfish centromere DNA variant on one chromosome will drive and fix. A selfish centromere DNA variant can spread to other chromosomes, eventually leading to the complete replacement of centromere DNA. Being heterozygous or homozygous for the selfish centromere DNA variant is predicted to have fitness costs until the selfish variant replaces all centromeres of all chromosomes. Time window for suppressor evolution is between the birth of selfish centromere DNA variant to the complete replacement of all centromeres. (C) Possible mechanisms of fitness costs in heterozygous (left) or homozygous (right) for the selfish centromere DNA variant. Black squares represent centromere binding proteins. When heterozygous for the selfish centromere DNA variant, functional asymmetry between centromeres may cause fitness costs. When homozygous for the selfish centromere DNA variant, selfish centromere DNA variant may sequester centromere binding protein and the preexisting centromere DNA may recruit fewer centromere binding proteins, leading to segregation errors. When the selfish centromere DNA variant replaces all of the centromeres, the available centromere binding proteins are equally recruited to all centromeres so there will be no fitness costs.

How centromere binding proteins can evolve to suppress the costs of drive remains an open question despite being a crucial component of the centromere drive model. Details of the fitness costs are unclear, but they likely depend on functional differences between genetically different centromeres in the shared cytoplasm (Figure 1.10C) and would therefore be suppressed by reducing these differences. As a mouse strain homozygous for selfish centromeres and the preexisting centromeres (as in Figure 1.10 right) is not available, functional differences between paired centromeres (as in Figure 1.10 left) are studied in the following chapter.

Chapter 2. Parallel Pathway Model for Drive and Suppression

2.1. The Parallel Pathway Model Provides Three Testable Predictions

Based on our previous finding that selfish centromeres drive by recruiting more effectors, we propose that functional differences between centromeres can be suppressed by equalizing effector recruitment via a second pathway. This equalization would render genetically different centromeres functionally equivalent. This model incorporates previous findings that in addition to the kinetochore pathway, which acts through BUB1 kinase, effectors are also recruited through a heterochromatin pathway. Pericentromeric heterochromatin recruits the CPC, which recruits SGO2 and MCAK (Figure 1.6, heterochromatin pathway) (Abe et al., 2016; Ainsztein et al., 1998; Higgins and Prendergast, 2016; Kang et al., 2011; Marston 2015). In our intra-species CHPO hybrid model system (Figure 2.1A), the kinetochore pathway is asymmetric: we observe higher levels of the kinetochore proteins HEC1/NDC80 and CENP-C on larger vs smaller centromeres (Chmátal et al., 2014; Iwata-Otsubo et al., 2017). In contrast, the heterochromatin pathway is symmetric: the heterochromatin mark, H3K9me3, is equal on the two sides of each bivalent (Figure 2.1B and C) (Iwata-Otsubo et al., 2017). These observations suggest that, in this system, selfish centromere DNA exploits the kinetochore pathway to make genetically different centromeres also functionally different, with larger centromeres recruiting more effectors. In contrast, the heterochromatin pathway appears insensitive to selfish DNA, recruiting effectors equally. We propose that centromere protein evolution suppresses functional differences by minimizing the contribution of the asymmetric kinetochore pathway to effector recruitment, relative to the symmetric heterochromatin pathway.

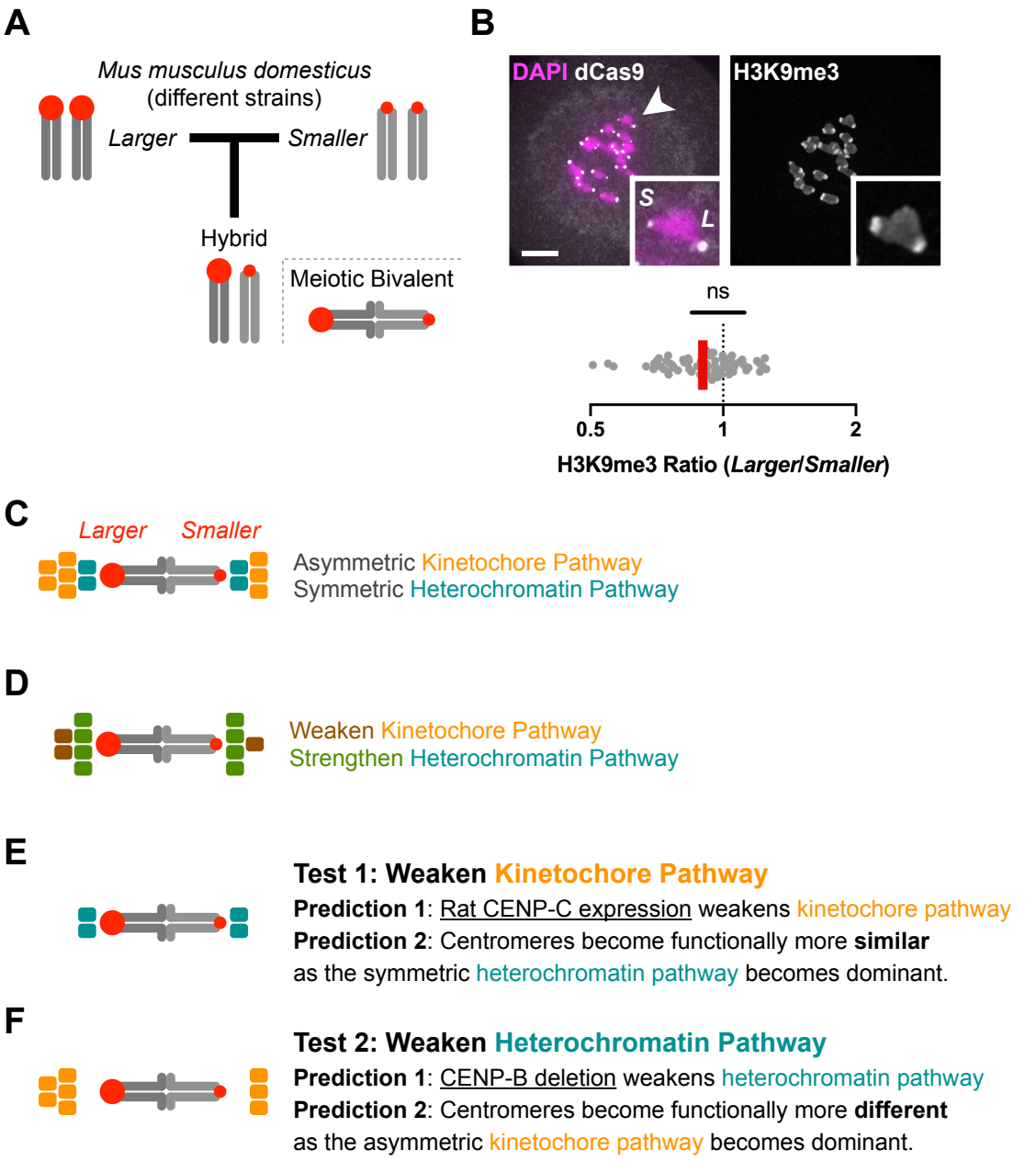


Figure 2.1. Parallel pathway model for drive and suppression.
(A) CHPO hybrid model system. Crossing strains with larger (CF-1) and smaller (CHPO) centromeres generates a hybrid in which genetically different centromeres are paired in meiotic bivalents. Larger red circles indicate more minor satellite centromere DNA repeats.
(B) CHPO hybrid oocytes were microinjected with cRNA for dCas9-EGFP and gRNA targeting minor satellite centromere DNA to distinguish larger (*L*) and smaller (*S*) centromeres, fixed at meiosis I, and stained for H3K9me3; 10µm scale bar, 7.4µm square inset. The H3K9me3 ratio for each pair of larger and smaller centromeres within a bivalent is plotted (n=67 bivalents); red line, geometric mean; ns: no significant deviation from 1.

(C) Asymmetric kinetochore pathway and symmetric heterochromatin pathway in our hybrid model system. Colored boxes represent effector proteins recruited by the kinetochore pathway (orange) or the heterochromatin pathway (green).
(D) Suppression of functional differences between centromeres by recruiting similar amounts of effector proteins on genetically different centromeres. Colored boxes represent changes relative to panel E. Proteins in the kinetochore pathway can adapt by reducing affinity for DNA or for other proteins leading to effector recruitment. Inner centromere proteins can adapt by increasing affinity for heterochromatin or by decreasing their recruitment by the kinetochore pathway.
(E) Introducing a divergent allele of CENP-C (blue boxes) disrupts interactions for effector recruitment and therefore weakens the kinetochore pathway (prediction 1) and makes centromeres functionally more similar (prediction 2).
(F) As CENP-B recruits heterochromatin proteins, deleting CENP-B weakens the heterochromatin pathway (prediction 1), making the asymmetric kinetochore pathway dominant and centromeres functionally more asymmetric (prediction 2).

Evolution of the kinetochore pathway is constrained by its indispensable role in mitotic and meiotic chromosome segregation. Nevertheless, proteins may evolve to weaken the pathway by reducing interactions between centromere-binding proteins and DNA or between proteins leading to effector recruitment (Figure 2.1D). Similarly, evolution of heterochromatin proteins is constrained by numerous vital heterochromatin-dependent cellular functions (Allshire and Madhani, 2017). Inner centromere proteins (such as the CPC) that interact with heterochromatin may evolve, however, to increase effector recruitment. Finally, overall effector levels are also constrained because microtubule destabilizing activity is necessary to correct kinetochore-microtubule attachment errors, but excessive destabilizing activity weakens attachments necessary for anaphase segregation and activates the spindle assembly checkpoint (Godek et al., 2014). According to our parallel pathway model, a new centromere DNA variant can exploit the kinetochore pathway to recruit more effectors by strengthening interactions with any centromere-binding protein that contacts the DNA: CENP-A, the CENP-A chromatin assembly machinery, or other proteins that link centromere chromatin to the kinetochore (e.g., CENP-C or CENP-T). To suppress functional differences between centromeres, the centromere protein network recurrently

evolves to minimize the kinetochore pathway relative to the heterochromatin pathway while maintaining essential functions.

Here we test three predictions from the parallel pathway model. First, when the asymmetric kinetochore pathway is weakened, we predict that centromeres become functionally more similar due to the symmetric heterochromatin pathway. We selected CENP-C as a key scaffold protein in the kinetochore pathway that is known to evolve rapidly under positive selection (Klare et al., 2015; Schueler et al., 2010; Talbert et al., 2004). Under the parallel pathway model, CENP-C interfaces have co-evolved with interacting partners to modulate effector recruitment. Thus, introducing a divergent allele of CENP-C in mouse cells (e.g., rat CENP-C, in which 32% of the amino acid sequence is different) is predicted to disrupt such interactions and weaken the kinetochore pathway (Figure 2.1E). Second, when the symmetric heterochromatin pathway is weakened, we predict that the asymmetric kinetochore pathway makes a relatively larger contribution to effector recruitment. Genetically different centromeres in our hybrid model system should therefore become functionally more different. To target pericentromeric heterochromatin, we deleted CENP-B, which is the only centromeric chromatin component that is dispensable for core centromere function. CENP-B is recently acquired in mammals and fission yeast from a pogo-like transposase (Casola et al., 2007; Kipling and Warburton, 1997), and several domesticated transposases regulate heterochromatin (Gao et al., 2020; Jangam et al., 2017; Nozawa et al., 2010; Yang et al., 2017). In mouse and human cultured cells and fission yeast, CENP-B contributes to pericentromeric heterochromatin formation via heterochromatin protein recruitment (Nakagawa et al., 2002; Okada et al., 2007; Otake et al., 2020), so deleting CENP-B should weaken the heterochromatin pathway (Figure 2.1F). Mammalian CENP-B can also contribute to the kinetochore pathway via CENP-C recruitment (Fachinetti et al., 2015), so the functional consequences of CENP-B deletion in our model need to be tested. Third, if proteins

in the kinetochore and heterochromatin pathways have evolved to modulate effector recruitment, we predict signatures of positive selection in multiple protein domains involved in effector recruitment. In contrast, the previous model of an arms race limited to interactions between centromere DNA and DNA-binding proteins only predicts rapid evolution of protein domains involved in DNA binding (Henikoff et al., 2001; Malik and Henikoff, 2001). Our observations are consistent with all three predictions, supporting our parallel pathway model for drive and suppression.

2.2. Weakening the Kinetochore Pathway Makes Centromeres Functionally More Symmetric.

To weaken the kinetochore pathway, we targeted CENP-C because it serves as a hub for recruiting kinetochore proteins. Our model predicts that CENP-C has co-evolved with interacting partners to modulate effector recruitment, so that an allele from another species will disrupt these interactions and weaken the kinetochore pathway (Figure 2.1E, Prediction 1). To test this prediction, we selected divergent alleles from rat and from the African striped mouse *Rhabdomys pumilio* as model organisms close to mouse with divergent centromere DNA and proteins (Figure 2.2A) (Cazaux et al., 2013; Gibbs et al., 2004; Mallarino et al., 2018; Takeiri et al., 2013). Because protein interfaces change by genetic drift as well as by selection, alleles from closely related species minimize incompatibilities coming from stochastic changes. We introduced GFP-tagged divergent alleles (or the mouse allele as a control) into mouse oocytes in the presence of endogenous CENP-C (Figure 2.3A-B). We find that rat CENP-C expression reduces effector recruitment, as represented by SGO2 staining, compared to mouse CENP-C expression (Figure 2.2B). This result is consistent with our model prediction and could reflect differences between mouse and rat CENP-C in their recruitment to centromeres or in their interactions with other kinetochore proteins. For example, evolution at an interface with CENP-A nucleosomes or with CENP-B may disrupt rat CENP-C recruitment to centromeres. Alternatively, CENP-C evolution might impact the domains that mediate interactions with other kinetochore proteins involved in SGO2 recruitment. We find that mouse and rat CENP-C are equally recruited and incorporated into mouse centromeres (Figure 2.2C and 2.3C-E), indicating functional changes at an interface with other kinetochore proteins.

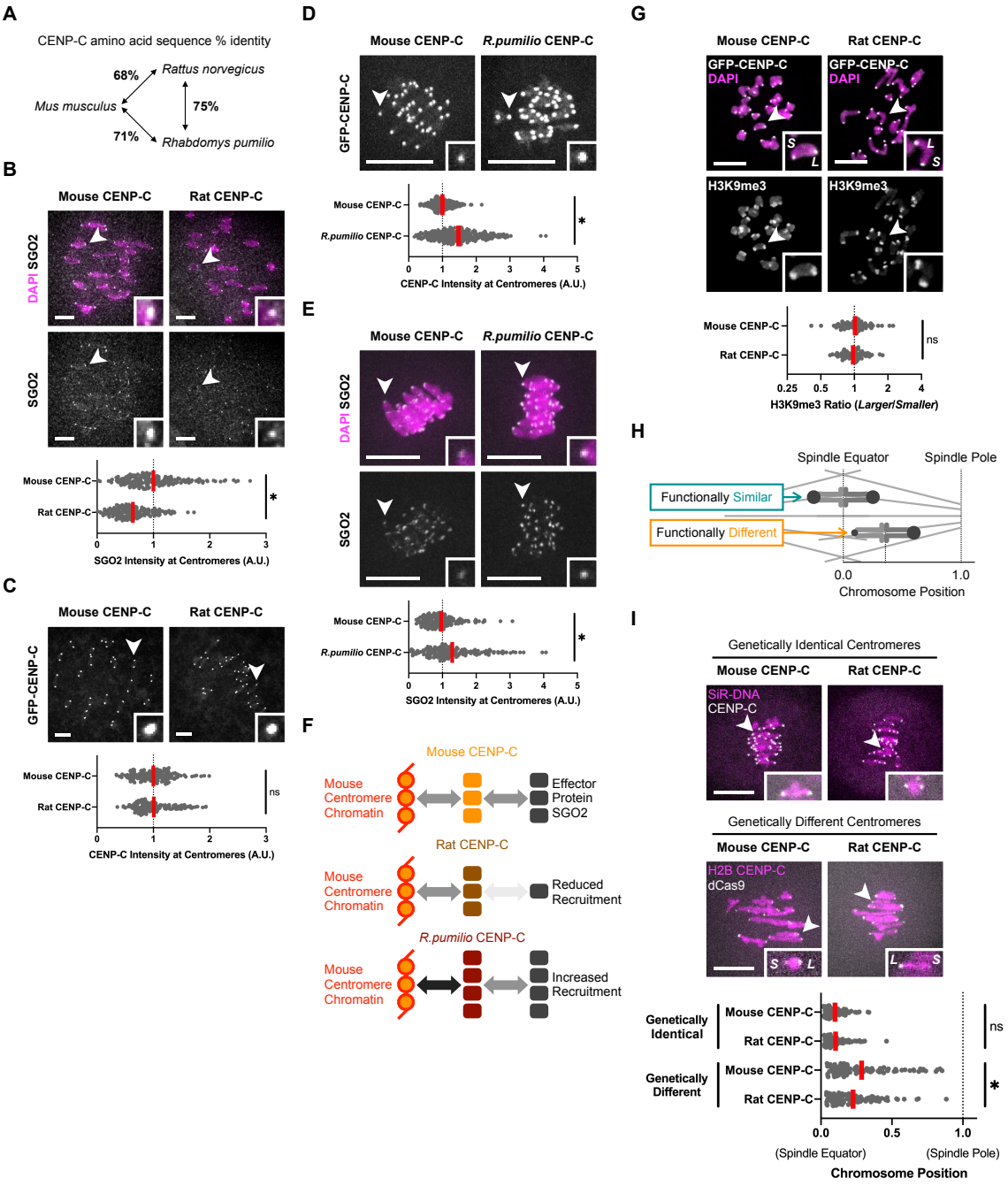


Figure 2.2. (Caption on next page)

Figure 2.2. Introducing rat CENP-C in mouse oocytes weakens the kinetochore pathway and makes centromeres functionally more symmetric.
(A) CENP-C divergence between *Mus musculus* (mouse), *Rattus norvegicus* (rat), and *Rhabdomys pumilio* (a model organism closely related to *Rhabdomys dilectus*, Figure 2.7).
(B and C) CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C and fixed in prometaphase/metaphase I. Cells were stained for SGO2 (A) or analyzed for GFP fluorescence (B). 10μm scale bars, 2.2μm square insets. Plots show centromere signal intensities. Each dot represents a single centromere (n=200 centromeres from 20 oocytes for each construct); red line, mean; *$p<0.05$; ns: not significant.
(D and E) CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or *R. pumilio* CENP-C and fixed in prometaphase/metaphase I. Cells were analyzed for GFP fluorescence (D) or stained for SGO2 (E). 10μm scale bars, 2.2μm square insets. Plots show centromere signal intensities. Each dot represents a single centromere (n≥170 centromeres from ≥22 oocytes for each construct); red line, mean.
(F) Different CENP-C interfaces have changed to modulate effector recruitment. Schematics summarize the results of panels B to E. Compared to mouse CENP-C, rat CENP-C is similarly recruited to mouse centromere chromatin, but downstream effector recruitment is reduced. In contrast, *R. pumilio* CENP-C is recruited at higher levels to mouse centromere chromatin, leading to increased effector recruitment.
(G) CHPO hybrid oocytes (see Figure 2.1A) were microinjected with cRNA for GFP-tagged mouse or rat CENP-C, fixed in prometaphase/metaphase I, and stained for H3K9me3; 10μm scale bar, 5.9μm square inset. The H3K9me3 ratio for each pair of larger (*L*) and smaller (*S*) centromeres on a bivalent is plotted (n≥72 bivalents for each genotype); red line, geometric mean; ns: not significant.
(H) Schematic of chromosome position assay to measure functional differences between paired centromeres. Distance from the spindle pole to the equator is defined as 1 for each cell to normalize for variation in spindle size.
(I) For genetically identical centromeres, CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C, and DNA was visualized with SiR-DNA. For genetically different centromeres, CHPO hybrid oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C, together with cRNAs for GFP-tagged H2B and mCherry-tagged dCas9 and gRNA targeting minor satellite centromere DNA. Cells were imaged live to preserve chromosome positions, measured at late metaphase I. In the plot, each dot represents a single bivalent (n=100 bivalents from 10 CF-1 oocytes and ≥20 CHPO hybrid oocytes for each construct); red line, mean.

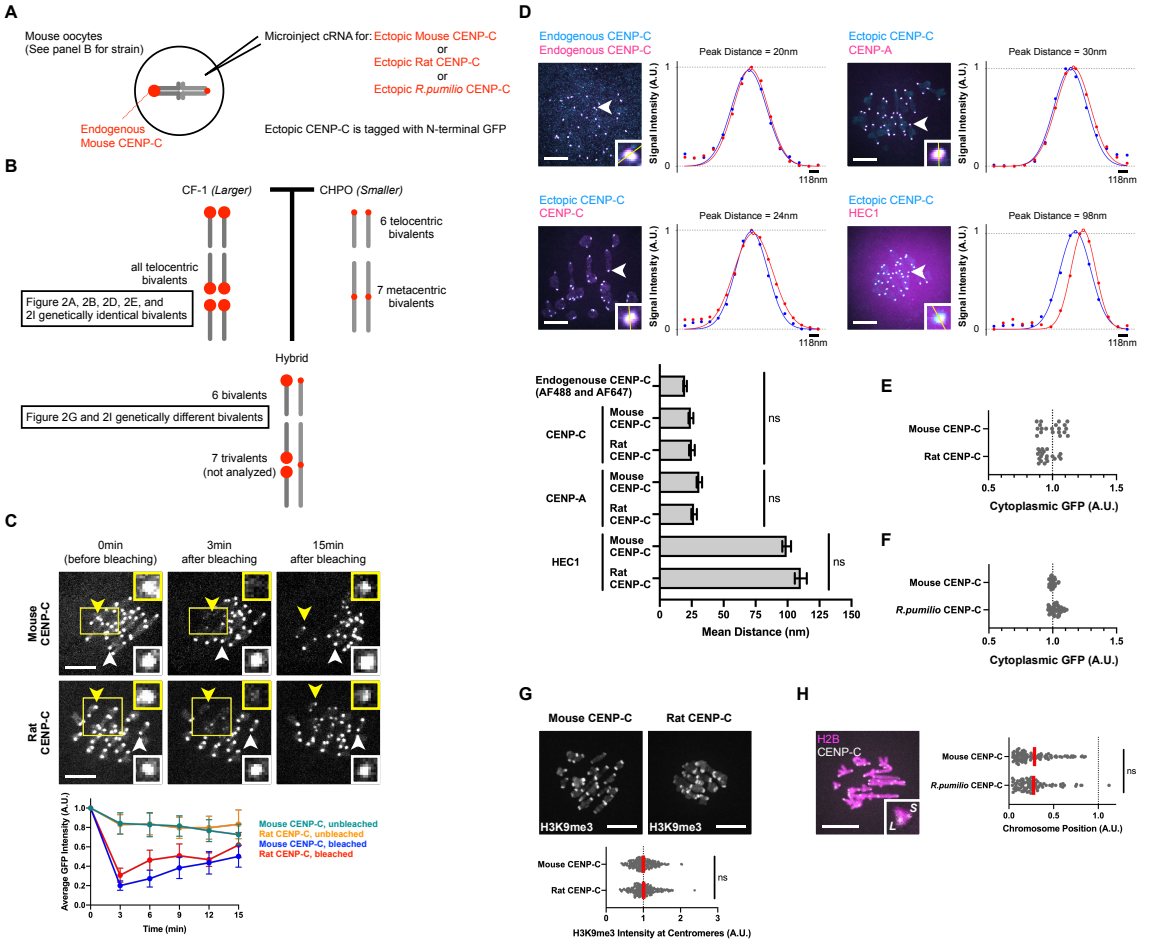


Figure 2.3. Incorporation of ectopically expressed CENP-C.

(A) Schematic of CENP-C expression experiments. GFP-tagged ectopic CENP-C was expressed by cRNA microinjection into mouse oocytes, in the presence of endogenous mouse CENP-C. Different mouse strains were used (see panel B).

(B) Detailed crossing scheme to produce CHPO hybrid mice (related to Figure 2.1A). Crossing *Mus musculus domesticus* strains with larger (CF-1) and smaller (CHPO) centromeres generates a hybrid. The minor satellite monomer is the same in the two strains, but the abundance differs. Due to the metacentric Robertsonian fusion chromosomes in CHPO, hybrid oocytes have 7 trivalents, which are not analyzed in our experiments (see also Figure 2.5D). CHPO hybrid oocytes were used to analyze genetically different bivalents in Figures 2.2F and 2.2G. CF-1 oocytes were used to analyze genetically identical bivalents in Figures 2.2A, B, D, E and G.

(C) CENP-C exchange at centromeres. CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C, and several centromeres (yellow rectangles) were photobleached in meiosis I. Representative centromeres are marked with arrowheads and shown in insets (yellow: bleached, white: unbleached). 10µm scale bar, 4.4µm square insets. Fluorescence recovery after photobleaching (FRAP) was monitored every three minutes. Each data point represents GFP intensity as a fraction of intensity before bleaching, averaged over multiple centromeres (n=10 centromeres from 5 oocytes for each condition). Partially bleached centromeres were not

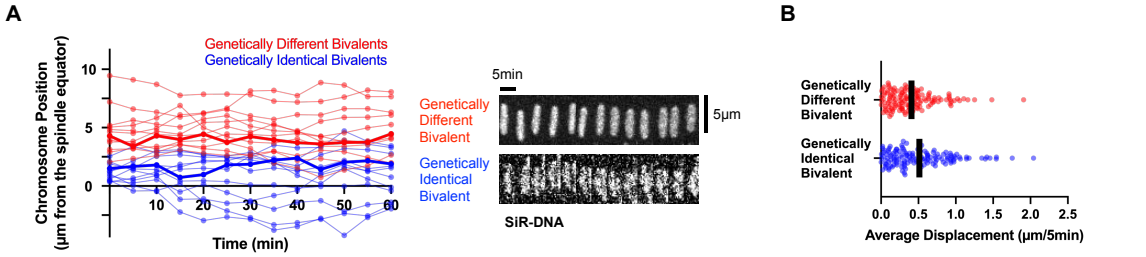
analyzed. Rapid recovery within 15 min is consistent with previous observations for CENP-C in mitosis (Hemmerich et al., 2008). Given that oocytes are cultured overnight after microinjection of CENP-C cRNA, ectopic CENP-C likely replaces most of endogenous CENP-C.
(D) CENP-C position at centromeres relative to CENP-A and HEC1. Protein-protein distances were measured as previously described (Wan et al., 2009). Briefly, line scans of each centromere were taken parallel to the chromosomal axis at a width of ten pixels (yellow lines in insets). Each dot represents signal intensity of one pixel. The data were fit with Gaussian normal distributions to estimate intensity peaks with subpixel resolution, and the distance between the two peaks was calculated. Chromatic aberration was corrected by staining endogenous CENP-C with two secondary antibodies labeled with different colors (AF488 and AF647), and the distance in this condition shows the error range in our analysis. Bar graph shows mean distances and standard error of the mean (n≥90 centromeres for each construct). CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C and fixed in meiosis I. Cells were stained for CENP-C, CENP-A, or HEC1 with AF647-labeled secondary antibody. Images of oocytes expressing mouse CENP-C are shown as examples; 10μm scale bar, 3.5μm square inset. Distances between CENP-C antibody staining and GFP-CENP-C are close to our error range. Distances from CENP-C to CENP-A and CENP-C to HEC1 were consistent with a previous report (11±11nm and 79±10nm, respectively, Suzuki et al., 2014), within our error. We find no significant differences in distance measurements for mouse and rat CENP-C, suggesting that both are incorporated into centromeres similarly.
(E and F) Expression levels are similar between mouse and rat and between mouse and *R.pumilio* CENP-C (related to Figure 2.2). CF-1 oocytes were microinjected with cRNA for the indicated GFP-tagged CENP-C alleles and analyzed for cytoplasmic GFP fluorescence to measure expression levels. Each dot represents one oocyte (n=20 for mouse and rat CENP-C, n≥41 for mouse and *R.pumilio* CENP-C).
(G) H3K9me3 levels are not affected by kinetochore pathway disruption. CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C, fixed in meiosis I, and stained for H3K9me3. Each dot in the plot represents a single centromere (n≥240 centromeres for each construct); red line, mean; ns: not significant. 10μm scale bar.
(H) CHPO hybrid oocytes were microinjected with cRNAs for GFP-tagged *R.pumilio* CENP-C and mCherry-tagged H2B. Cells were imaged live to preserve chromosome positions, measured at late metaphase I. In the plot, each dot represents a single bivalent (n=85 bivalents for *R.pumilio* CENP-C, mouse CENP-C data is from Figure 3G); red line, mean; ns: not significant.

In contrast to our results with rat CENP-C, *R. pumilio* CENP-C is recruited at higher levels to mouse centromeres compared to mouse CENP-C (Figure 2.2D), with similar expression levels as measured by cytoplasmic GFP (Figure 2.3F). Consistent with this result, effector recruitment is also increased in cells expressing *R. pumilio* CENP-C (Figure 2.2E). Together, these findings show that different CENP-C interfaces, with centromere chromatin or with other kinetochore

proteins, have changed through rodent evolution to modulate effector recruitment (Figure 2.2F). Furthermore, differences between *R. pumilio* and mouse CENP-C localization to mouse centromeres suggest that mouse CENP-C has evolved to weaken its interactions with centromere chromatin.

Rat CENP-C expression provides an experimental tool to specifically weaken the kinetochore pathway, without affecting heterochromatin (Figure 2.2G and 2.3G), allowing us to test our prediction that genetically different centromeres become functionally more similar in our hybrid model system (Figure 2.1E, Prediction 2). As a functional readout of centromere asymmetry, we analyzed chromosome position on the spindle at metaphase I (Figure 2.2H and 2.4A-B). Chromosome position is sensitive to differences in interactions with spindle microtubules between centromeres of homologous chromosomes, which are paired in a meiotic bivalent. If the paired centromeres are genetically and functionally similar, then chromosomes align at the spindle equator in a typical metaphase configuration. In our CHPO hybrid model systems, paired centromeres are genetically and functionally different, and bivalents are positioned off-center on the spindle, with the larger centromere closer to its attached pole (Akera et al., 2019; Chmátal et al., 2014). Manipulations that make these genetically different centromeres functionally more similar will lead to positioning closer to the spindle equator, as previously shown by manipulating BUB1 kinase to equalize MCAK levels on larger and smaller centromeres (Akera et al., 2019). Conversely, manipulations that make the centromeres functionally more different will position bivalents closer to the poles. We find that expression of rat CENP-C in CHPO hybrid oocytes (Figure 2.1A) leads to bivalents positioned closer to the spindle equator (Figure 2.2I), without affecting meiotic progression (Figure 2.4C). This result indicates that the paired larger and smaller centromeres are functionally more similar, consistent with the prediction that the symmetric heterochromatin pathway becomes relatively more dominant when the asymmetric

kinetochore pathway is weakened (Figure 2.1E, Prediction 2). We also tested *R. pumilio* CENP-C expression but did not find changes in chromosome position (Figure 2.3H), suggesting that the modest increase in effector recruitment (Figure 2.2E) does not impact microtubule destabilizing activity enough to be detected in our assay.



	5hrs *in vitro* Maturation		7.5hrs *in vitro* Maturation		10hrs *in vitro* Maturation	
	Mouse CENP-C	Rat CENP-C	Mouse CENP-C	Rat CENP-C	Mouse CENP-C	Rat CENP-C
Not Aligned	19 (90%)	18 (95%)	14 (29%)	16 (32%)	4 (10%)	6 (12%)
Aligned	2 (10%)	1 (5%)	29 (59%)	27 (54%)	6 (15%)	11 (22%)
Segregated	0 (0%)	0 (0%)	6 (12%)	7 (14%)	29 (74%)	32 (65%)

Number of Oocytes (%)

	5hrs *in vitro* Maturation		7.5hrs *in vitro* Maturation		10hrs *in vitro* Maturation	
	Cenpb +/–	*Cenpb* –/–	*Cenpb* +/–	*Cenpb* –/–	*Cenpb* +/–	*Cenpb* –/–
Not Aligned	22 (81%)	14 (74%)	15 (35%)	16 (36%)	9 (13%)	6 (12%)
Aligned	5 (19%)	5 (26%)	24 (56%)	27 (61%)	24 (35%)	21 (43%)
Segregated	0 (0%)	0 (0%)	4 (9%)	1 (2%)	35 (51%)	22 (45%)

Number of Oocytes (%)

Figure 2.4. Analyses of chromosome oscillations and meiotic progression.
(A and B) Positions of genetically different bivalents (from CHPO hybrid oocytes) and genetically identical bivalents (from CF-1 oocytes) were measured every 5 min by live imaging of SiR-DNA. Each line in the graph (A) represents movement of a single bivalent, with thicker lines shown as examples in kymographs. In both cases, bivalents stay in a relatively fixed position, as represented by the small average displacement (B, each data point represents displacement of a bivalent between two consecutive time points). Average displacement is similar for genetically identical and genetically different bivalents, suggesting that the differences in chromosome positions are not due to differences in oscillations.
(C and D) Oocytes were collected, matured *in vitro*, and imaged live at different time points to analyze meiotic progression. CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C, and DNA was visualized with SiR-DNA (C). DNA was visualized with SiR-DNA in *Cenpb*$^{+/-}$ or *Cenpb*$^{-/-}$ oocytes (D). Examples of "not aligned", "aligned", and "segregated" are shown (10µm scale bar). The numbers of oocytes at each stage were counted at each time point. We find no obvious differences in meiotic progression between oocytes expressing mouse or rat CENP-C, or between *Cenpb*$^{+/-}$ and *Cenpb*$^{-/-}$ oocytes. Furthermore, genetically identical bivalents are positioned similarly in all cases (Figures 2.2I and 2.6D),

suggesting that differences in chromosome positions for genetically different bivalents are due to functional differences in centromeres rather than differences in meiotic progression.

2.3. Weakening the Heterochromatin Pathway Makes Centromeres Functionally More Asymmetric.

To determine the contribution of CENP-B to effector recruitment, we created *Cenpb* null mice using CRISPR genome editing (Figure 2.5A-C). We find that loss of CENP-B weakens both the kinetochore and heterochromatin pathways, as shown by reduced CENP-C and H3K9me3 staining, respectively (Figure 2.6A). These results are consistent with previous findings that CENP-B contributes to CENP-C recruitment and to formation of pericentromeric heterochromatin (Fachinetti et al., 2015; Okada et al., 2007; Otake et al., 2020). We also find reduced effector recruitment, as represented by SGO2 staining (Figure 2.6A), consistent with the idea that CENP-B recruits effectors through the kinetochore and heterochromatin pathways.

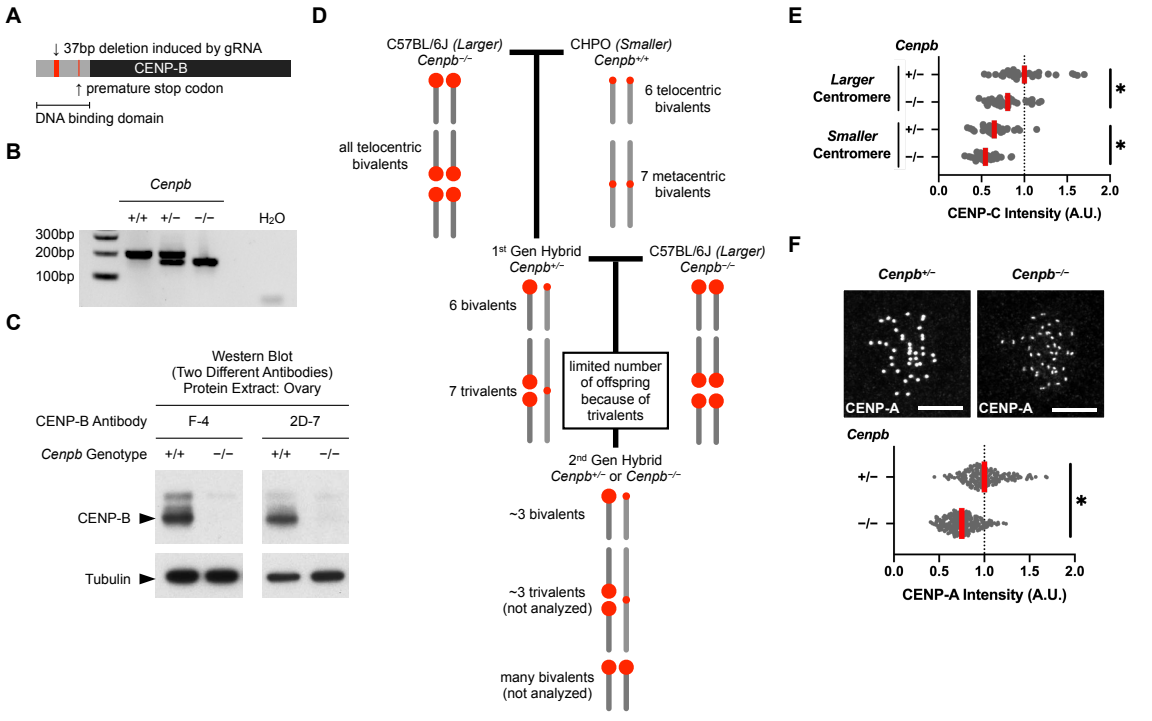


Figure 2.5. (Caption on next page)

Figure 2.5. CRISPR genome editing creates CENP-B null mice.
(A) Summary of CRISPR genome editing, using a gRNA designed to target the DNA binding domain of CENP-B.
(B) *Cenpb* genotyping. As the *Cenpb* mutation is a 37bp deletion, a PCR reaction amplifying the flanking regions can distinguish three *Cenpb* genotypes.
(C) Absence of CENP-B protein in $Cenpb^{-/-}$ mice. Protein extract from ovary was used to detect CENP-B using two different antibodies. Bands that disappear in $Cenpb^{-/-}$ likely correspond to CENP-B with and without post-translational modifications such as SUMOylation (Morozov et al., 2017). Faint bands that also appear in $Cenpb^{-/-}$ are non-specific.
(D) Detailed crossing scheme to produce second-generation hybrid $Cenpb^{-/-}$ mice with larger and smaller paired centromeres (related to Figure 2.6B). The first cross produces first-generation hybrid $Cenpb^{+/-}$ animals with smaller centromeres inherited from CHPO. Because CHPO has six telocentric chromosomes and seven metacentrics formed by Robertsonian chromosome fusions, the first-generation hybrids contain six bivalents in meiosis and seven trivalents, in which a Robertsonian fusion from CHPO pairs with two homologous telocentric chromosomes (Chmátal et al., 2014). Trivalents are associated with meiotic errors (Bint et al., 2011; Daniel, 2002; Pacchierotti et al., 1995), and the first-generation hybrids exhibit low fertility, but some progeny can be obtained in a second cross to $Cenpb^{-/-}$. These second-generation hybrids inherit some smaller centromeres from the first-generation hybrid parent, and 25% are $Cenpb^{-/-}$ females that can be used to collect oocytes for our analyses. Oocytes from the second-generation hybrids do not arrest at metaphase I, likely because they have fewer trivalents that activate the spindle assembly checkpoint (Chmátal et al., 2015). Therefore, we are unable to measure biased orientation of larger centromeres towards the egg side of the spindle, as previously reported in first-generation hybrids (Iwata-Otsubo et al., 2017), because this bias depends on delayed progression through meiosis I (Akera et al., 2019). Due to the limited number of second-generation hybrids and low fertility of these animals, we were also unable to measure transmission bias.
(E) CENP-C reduction in the second-generation hybrid (related to Figure 2.6C). Oocytes from the second-generation hybrid were microinjected with cRNA for GFP-tagged dCas9 and gRNA targeting minor satellite centromere DNA, fixed at metaphase I, and stained for CENP-C. Each dot represents a single centromere (n=34 centromeres for each construct); red line, mean; *$p<0.01$.
(F) $Cenpb^{+/-}$ or $Cenpb^{-/-}$ oocytes were fixed in meiosis I and stained for CENP-A; 10μm scale bar. Plot shows centromere signal intensities, normalized by mean intensity of $Cenpb^{+/-}$ control. Each dot represents a single centromere ($n\geq210$ centromeres); red line, mean; *$p<0.05$. As CENP-B deletion reduces both CENP-A chromatin and heterochromatin, many other proteins are likely affected as well.

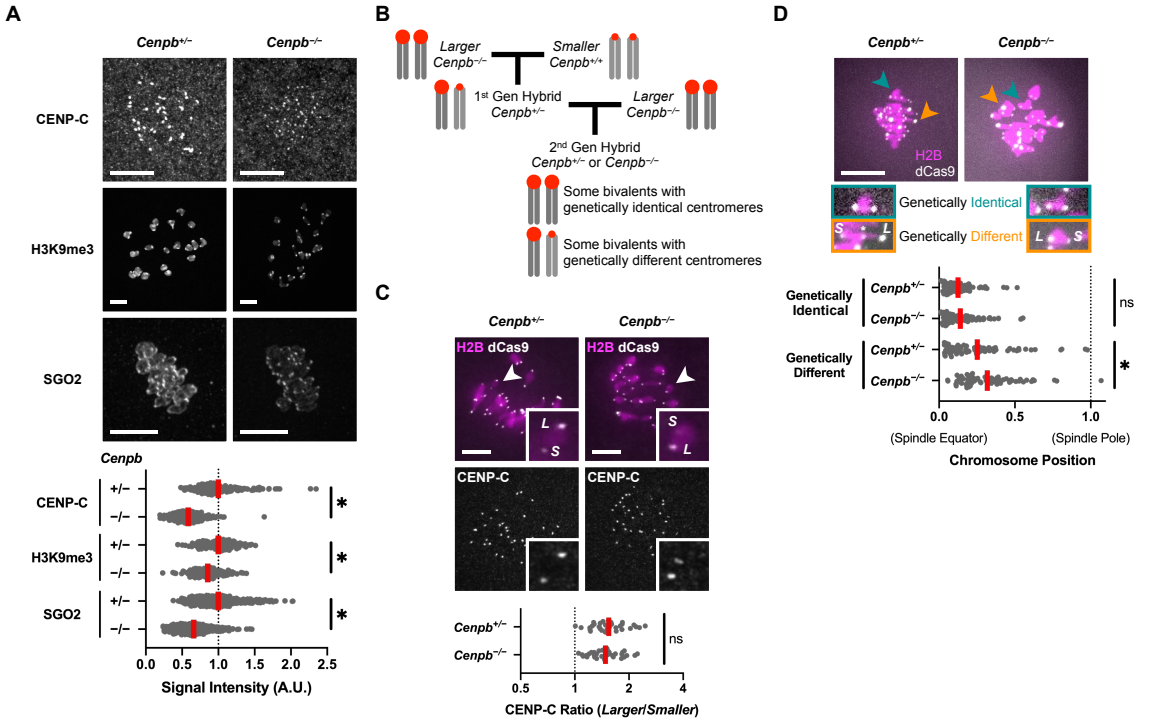


Figure 2.6. Deleting CENP-B weakens the heterochromatin pathway and makes centromeres functionally more asymmetric.

(A) *Cenpb*$^{+/-}$ or *Cenpb*$^{-/-}$ oocytes were fixed in prometaphase/metaphase of meiosis I and stained for CENP-C, H3K9me3, or SGO2; 10μm scale bar. Plot shows centromere signal intensities, normalized by mean intensity of *Cenpb*$^{+/-}$ control for each protein. Each dot represents a single centromere (n≥154 centromeres for each condition); red line, mean; *p<0.05.

(B) Crossing scheme to produce second-generation hybrid *Cenpb*$^{-/-}$ mice. Oocytes from these mice have bivalents with genetically identical centromeres as well as bivalents with genetically different centromeres.

(C) Second-generation hybrid oocytes were microinjected with cRNA for dCas9-EGFP and gRNA targeting minor satellite centromere DNA, fixed in prometaphase/metaphase I, and stained for CENP-C; 10μm scale bar, 5.9μm square inset. The CENP-C ratio is plotted for each pair of larger (*L*) and smaller (*S*) centromeres on bivalents with genetically different centromeres, determined from dCas9-EGFP signals (n=34 bivalents for each genotype); red line, geometric mean; ns: not significant.

(D) Second-generation hybrid oocytes were microinjected with cRNAs for dCas9-EGFP and H2B and gRNA targeting minor satellite centromere DNA. Cells were imaged live to preserve chromosome positions, measured at late metaphase I. In the plot, each dot represents a single bivalent (n≥74 bivalents for each genotype) with either genetically identical or genetically different centromeres, determined from dCas9-EGFP signals; red line, mean.

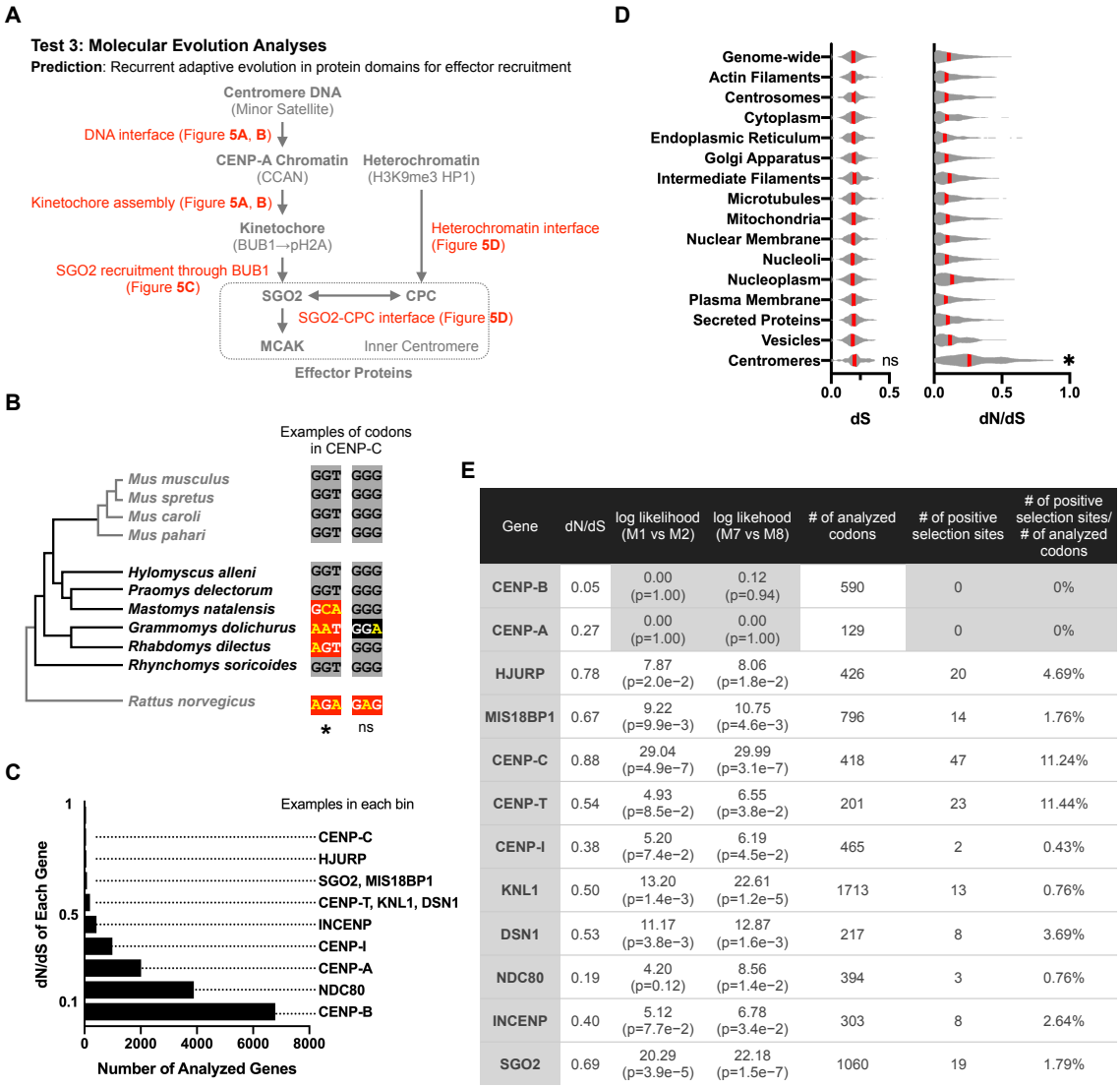
The known functions of CENP-B suggest two hypotheses for how it might affect centromeres in our CHPO hybrid model system. First, as the only centromere protein known to recognize a specific DNA sequence (the CENP-B box in repetitive centromere DNA) (Masumoto et al., 1989), CENP-B could be exploited by selfish larger centromeres with more CENP-B boxes to increase asymmetry via the kinetochore pathway. Second, CENP-B may suppress functional differences between centromeres by increasing the symmetric heterochromatin pathway. To test these hypotheses, we generated *Cenpb* null mice with paired larger and smaller centromeres through two generations of crosses (Figure 2.6B), and analyzed kinetochore pathway asymmetry and functional differences between centromeres. Due to technical limitations (Figure 2.5D), we were unable to measure transmission bias in these animals. To determine the impact of CENP-B on the kinetochore pathway, we analyzed CENP-C in meiotic bivalents with paired larger and smaller centromeres in second-generation hybrid $Cenpb^{-/-}$ oocytes. CENP-C was reduced to a similar extent on both larger and smaller centromeres (Figure 2.5E) and consistent with this equivalent reduction, the kinetochore asymmetry remained intact (Figure 2.6C). Therefore, CENP-B does not contribute to asymmetry in the kinetochore pathway, arguing against the first hypothesis that selfish centromere DNA exploits the kinetochore pathway via CENP-B.

To test the second hypothesis, that CENP-B acts as a suppressor through the symmetric heterochromatin pathway (Figure 2.1F), we examined functional differences between centromeres in second-generation hybrid oocytes, using the chromosome position assay (Figure 2.2H). We find that asymmetric bivalents with genetically different centromeres are positioned more off-center, closer to the spindle poles, in $Cenpb^{-/-}$ compared to control $Cenpb^{+/-}$ oocytes (Figure 2.6D). In contrast, we find no effect on positioning of symmetric bivalents with genetically identical centromeres in the same cells, and meiotic progression is similar in control and $Cenpb^{-/-}$ oocytes (Figure 2.4D). Together these findings indicate that positioning of

asymmetric bivalents closer to spindle poles in *Cenpb*$^{-/-}$ oocytes is due to increased functional differences between paired centromeres. This result is opposite to the result for rat CENP-C expression, which specifically weakens the kinetochore pathway (Figure 2.2B and G) and reduces functional asymmetry (Figure 2.2I). Therefore, although CENP-B deletion also weakens the kinetochore pathway, the dominant effect is to weaken a different pathway that equalizes centromeres, leading to increased functional asymmetry. Several lines of evidence indicate that this equalization pathway acts through heterochromatin: CENP-B is an established regulator of heterochromatin (Nakagawa et al., 2002; Okada et al., 2007; Otake et al., 2020), heterechromatin is similar on larger and smaller centromeres in our hybrid (Figure 2.1B) and reduced in *Cenpb*$^{-/-}$ oocytes (Figure 2.6A), and heterochromatin is an established pathway to recruit effector proteins (Abe et al., 2016; Ainsztein et al., 1998; Higgins and Prendergast, 2016; Kang et al., 2011; Marston 2015). We conclude that CENP-B suppresses functional differences between centromeres through the heterochromatin pathway. It is also possible, however, that CENP-B equalizes centromeres through an uncharacterized pathway independent of kinetochore assembly and heterochromatin.

2.4. Proteins in Both Pathways Have Signatures of Adaptive Evolution.

The original model of centromere drive and suppression posits an arms race between selfish centromere DNA and DNA-binding proteins such as CENP-A (Henikoff et al., 2001; Malik and Henikoff, 2001). This model predicts adaptive evolution of centromere protein domains that physically interact with DNA, and conservation of domains and other centromere proteins that do not bind DNA. In contrast, our parallel pathway model predicts signatures of recurrent adaptive evolution in protein domains leading to effector recruitment, including those that do not directly contact centromere DNA (Figure 2.7A). These changes could either weaken the kinetochore pathway or strengthen the heterochromatin pathway to make genetically different centromeres functionally more similar (Figure 2.1D). Rapid evolution of centromere proteins has been reported in several eukaryotic lineages, but there are no mechanistic studies of drive in these lineages (Finseth et al., 2015; Malik and Henikoff, 2001; Schueler et al., 2010; van der Lee et al., 2017). To analyze centromere protein evolution in a system where we have identified drive effectors, we tested for signatures of positive selection in Murinae. Because the sparseness of the phylogenetic tree of currently available Murinae genomes limits our statistical power to detect positive selection, we sequenced six new genomes (Figure 2.7B) using linked-read whole genome sequencing (10x Genomics). Each genome was assembled onto the *Mus musculus* reference genome (mm10) with LongRanger and *de novo* assembled with Supernova (see Section 2.5 and Table 1). Sampling evolutionary time more comprehensively increases our opportunities to observe adaptive changes (and minimize false positives from stochastic changes by genetic drift), especially those adaptive changes that are common to multiple independent lineages. Thus, these genomes provide a valuable resource for molecular evolution approaches in mouse as a mammalian model organism, such as our analyses of centromere proteins discussed below.



E

Gene	dN/dS	log likelihood (M1 vs M2)	log likehood (M7 vs M8)	# of analyzed codons	# of positive selection sites	# of positive selection sites/ # of analyzed codons
CENP-B	0.05	0.00 (p=1.00)	0.12 (p=0.94)	590	0	0%
CENP-A	0.27	0.00 (p=1.00)	0.00 (p=1.00)	129	0	0%
HJURP	0.78	7.87 (p=2.0e−2)	8.06 (p=1.8e−2)	426	20	4.69%
MIS18BP1	0.67	9.22 (p=9.9e−3)	10.75 (p=4.6e−3)	796	14	1.76%
CENP-C	0.88	29.04 (p=4.9e−7)	29.99 (p=3.1e−7)	418	47	11.24%
CENP-T	0.54	4.93 (p=8.5e−2)	6.55 (p=3.8e−2)	201	23	11.44%
CENP-I	0.38	5.20 (p=7.4e−2)	6.19 (p=4.5e−2)	465	2	0.43%
KNL1	0.50	13.20 (p=1.4e−3)	22.61 (p=1.2e−5)	1713	13	0.76%
DSN1	0.53	11.17 (p=3.8e−3)	12.87 (p=1.6e−3)	217	8	3.69%
NDC80	0.19	4.20 (p=0.12)	8.56 (p=1.4e−2)	394	3	0.76%
INCENP	0.40	5.12 (p=7.7e−2)	6.78 (p=3.4e−2)	303	8	2.64%
SGO2	0.69	20.29 (p=3.9e−5)	22.18 (p=1.5e−7)	1060	19	1.79%

Figure 2.7. Proteins in the kinetochore and heterochromatin pathways have signatures of recurrent adaptive evolution.

(A) Our parallel pathway model predicts that proteins in both pathways will have signatures of recurrent adaptive evolution at interfaces (shown in red) that lead to effector recruitment.

(B) Phylogenetic tree of Murinae species shows previously available genomes in gray and our newly sequenced genomes in black. Example codons show positive selection or neutral changes (mouse CENP-C Gly469 and Gly470). Nucleotide substitutions are shown in yellow, with synonymous and nonsynonymous substitutions highlighted in black or red, respectively. Higher numbers of nonsynonymous substitutions are interpreted as adaptive change under positive selection. PAML analysis: *P>0.99 for positive selection or not significant (ns) indicating a neutral change.

(C) Histogram shows the number of genes in each bin of dN/dS values, with examples of genes in each bin.

(D) Average dS and dN/dS across subcellular compartments. Red line, median; *p<0.05 for comparison to all other compartments.
(E) To test for signatures of positive selection in PAML, the likelihood of models of neutral codon evolution (M1 or M7) are compared to models allowing positive selection (M2 or M8). CENP-A and CENP-B are examples of genes without signatures of positive selection. See Table 2 for other genes, Figure 2.8 for a schematic of centromere proteins grouped by functional modules, and Figure 2.10 for further analyses of CENP-A and CENP-B. The number of analyzed codons is less than the total protein length as insertions, deletions, and ambiguous alignments are not analyzed. The number of positive selection sites is the number of codons with P>0.90 from Naive Empirical Bayes (NEB) analysis or Bayes Empirical Bayes (BEB) analysis from model 2 or 8.

Table 1. Whole Genome Sequencing Statistics

Genome	*Ha*	*Pd*	*Mn*	*Gd*	*Rd*	*Rs*
LongRanger reference-guided assembly						
Mean Molecule Length (kb) [Input DNA Quality]	8.7	11.3	16.7	8.9	13.8	16.8
DNA in Molecules >100kb (%) [Input DNA Quality]	20.9	19.4	24.0	14.7	10.1	12.0
Mean DNA per GEM (kb) [Linked-Reads Quality]	288.1	638.0	593.5	447.5	739.4	654.7
N50 Linked-Reads per Molecule [Linked-Reads Quality]	5.0	4.0	4.0	5.0	6.0	9.0
N50 Phase Block (kb) [Assembly Quality]	10.2	28.2	28.5	42.9	37.6	115.4
Number of Reads (million reads)	1378	1518	1389	1413	1415	1444
Mean Depth	37.7	44.6	38.9	33.9	34.3	39.3
Mapped Reads (%)	69.1	72.3	68.8	63.0	65.3	69.3
Supernova *de novo* assembly						
Number of Reads (million reads)	1378	800	1389	800	1415	1444
Number of Scaffolds ≥10 kb (K)	25.23	66.72	71.61	28.98	8.72	2.54
N50 Contig Size (kb)	13.31	19.47	17.16	12.19	28.42	82.08
N50 Phase Block Size (kb)	4.19	29.01	25.74	7.65	202.77	880.7
N50 Scaffold Size (Mb)	15 kb	31 kb	25 kb	14 kb	2.67	8.73
Assembly Size (Only Scaffolds ≥10 kb) (Gb)	393 Mb	1.72	1.60	410 Mb	2.29	2.26
Effective Coverage	31.2	26.3	33.2	26.3	55.8	56.4
Estimated Genome Size (Gb)	3.6	3.3	4.0	3.2	2.7	2.7

Low rates of nonsynonymous substitutions, which change the encoded amino acid, relative to synonymous substitutions (dN/dS) indicate purifying selection, as deleterious substitutions are selected against. Higher dN/dS indicates either adaptive evolution or loss of constraint, necessitating further analysis to identify signatures of positive selection (Echave et al., 2016; Sironi et al., 2015). We calculated dN/dS for all annotated mouse-rat orthologous genes. We find that multiple genes encoding centromere proteins have high dN/dS relative to the genome overall (Figure 2.7C), and the average dN/dS for these genes is significantly higher than for any other subcellular compartment (Figure 2.7D). We selected 46 genes with well-characterized centromere functions to analyze for signatures of positive selection based on phylogenetic analysis, using PAML (Yang, 2007). Consistent with our prediction, we find such signatures at multiple genes in the kinetochore and heterochromatin pathways (Figure 2.7E).

Extensive previous studies of centromere organization and function have established functional modules which can recruit drive effectors either directly or indirectly (Figure 2.7A). To fit our observations into this framework, we assigned genes to these modules (Figure 2.8). One module is CENP-A chromatin. Selfish centromere DNA can increase effector recruitment by expanding CENP-A chromatin through increased deposition of CENP-A nucleosomes. This process depends on a specialized histone chaperone, HJURP, which is targeted to centromeres by the MIS18 complex though interactions with CENP-C or CENP-I (Dunleavy et al., 2009; Foltz et al., 2009; Fujita et al., 2007; Moree et al., 2011; Shono et al., 2015). We find rapid evolution of HJURP, MIS18BP1, CENP-I, and the domain of CENP-C that interacts with the MIS18 complex (Figure 4E and 5A). In contrast, heterochromatin proteins such as HP1 paralogs and SUV39H1, which are not specific to centromeres/pericentromeres, are highly conserved (Supplementary Table 2), consistent with the idea that heterochromatin broadly suppresses selfish genetic elements regardless of the underlying DNA sequence (Allshire and Madhani, 2017). These findings

suggest that selection acts on the CENP-A chromatin assembly pathway to prevent expansion, but selfish centromere DNA does not exploit the heterochromatin pathway, consistent with our observation that genetically different centromeres have symmetric heterochromatin in our intra-species and inter-species hybrids (Figure 1D and our unpublished data).

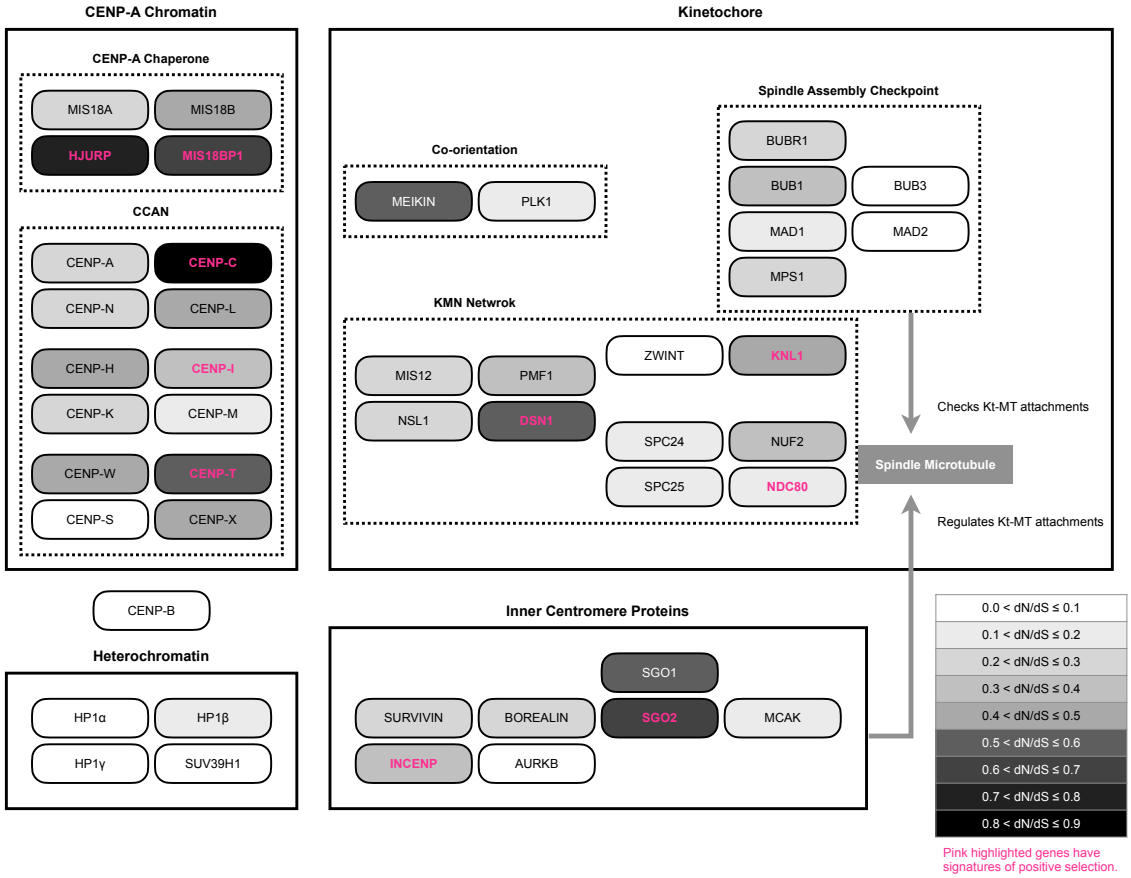


Figure 2.8. Map of rapidly evolving proteins at centromeres.
Proteins are grouped by functional modules corresponding to Figure 2.7A: CENP-A chromatin, heterochromatin, kinetochore, and inner centromere proteins (effector proteins). CENP-B has dual functions in CENP-A chromatin and heterochromatin (Figure 3.2). Within each module, proteins are further grouped by functions: CENP-A chaperone maintains CENP-A chromatin, CCAN connects CENP-A chromatin to the kinetochore, KMN network binds spindle microtubules, spindle assembly checkpoint proteins delays anaphase in response to unattached kinetochores, and meiotic co-orientation creates the meiosis I kinetochore geometry. Proteins are color-coded by dN/dS values (see Table 2). Pink highlighted genes have signatures of positive selection in PAML (see Figure 2.7E and Table 2).

Table 2. PAML Result Summary

Gene	dN/dS	log likelihood (M1 vs M2)	log likelihood (M7 vs M8)	# of analyzed species
CENP-L	0.49	0.00	0.00	7
CENP-N	0.25	0.07	0.71	13*
CENP-W	0.49	0.02	0.24	13*
CENP-S	0.09	0.00	0.17	7
CENP-X	0.41	2.32	2.43	7
CENP-H	0.45	1.89	3.14	13*
CENP-K	0.24	0.00	0.00	8*
CENP-M	0.20	0.10	1.00	7
MIS18A	0.24	0.00	0.00	7
MIS18B	0.49	0.93	1.10	8
MEIKIN	0.55	0.00	0.00	11*
PLK1	0.14	0.84	0.00	10
ZWINT	0.08	2.02	5.60	9
BUB1	0.34	2.04	2.61	13
BUBR1	0.25	0.00	1.70	13*
BUB3	0.00	0.00	0.00	12
MAD 1	0.13	0.00	0.00	12*
MAD 2	0.08	0.00	0.69	9
MPS1	0.30	0.00	0.00	9
MIS12	0.25	2.07	3.33	13
PMF1	0.31	0.00	0.50	13
NSL1	0.27	0.00	0.00	10
NUF2	0.33	0.00	0.00	13*
SPC24	0.20	0.00	0.00	13
SPC25	0.19	0.00	2.34	13
HP1α	0.05	0.00	0.00	11
HP1β	0.18	0.00	0.00	6*
HP1γ	0.01	0.00	0.00	12
SUV39H1	0.10	0.43	0.47	11
BOREALIN	0.22	0.00	0.01	10
SURVIVIN	0.27	0.00	0.86	9
AURKB	0.10	0.07	2.44	11
MCAK	0.17	0.00	3.79	8*
SGO1	0.64	1.02	1.33	10*

Sequences from reference-guided assembly is used for initial screening. Sequences from *de novo* assembly (*marked) is used to confirm the result.

Under our model (Figure 1.6), selfish centromere DNA can also recruit more effectors through the kinetochore pathway by strengthening direct interactions with CENP-A or with the constitutive centromere-associated network (CCAN), leading to larger kinetochores and more BUB1 kinase. Proteins can subsequently adapt by weakening interactions either with DNA or with other kinetochore proteins (Figure 2.7A, DNA interface and kinetochore assembly). Within the CCAN, CENP-C and CENP-T connect CENP-A chromatin to kinetochore proteins. The middle part of CENP-C interacts with CENP-A nucleosomes, while the N-terminus interacts with the MIS12 kinetochore complex (Petrovic et al., 2016; Weir et al., 2016). Similarly, the CENP-TWSX nucleosome-like complex contacts centromere DNA, and the other end of CENP-T interact with MIS12 and NDC80 kinetochore complexes (Cortes-Silva et al., 2020; Nishino et al., 2012; Veld et al., 2016). Consistent with our model, we detect signatures of positive selection in the chromatin-interacting domains and the kinetochore-interacting domains of both CENP-C and CENP-T (Figure 2.9A-B). In contrast, the DNA-interacting domain of CENP-B is conserved, consistent with our finding that selfish centromere DNA does not exploit CENP-B. Unlike in other eukaryotic lineages such as monkeyflower, fly, and primates (Finseth et al., 2015; Malik and Henikoff, 2001; Schueler et al., 2010), we do not detect signatures of positive selection in the part of CENP-A that can be aligned in Murinae species, but the N-terminal tail is duplicated in some species and therefore difficult to analyze by standard methods (Figure 2.10A). Diversification of the CENP-A N-terminal tail is also observed in plants, where crosses between strains expressing different alleles exhibit zygotic segregation errors and genome elimination (Maheshwari et al., 2015).

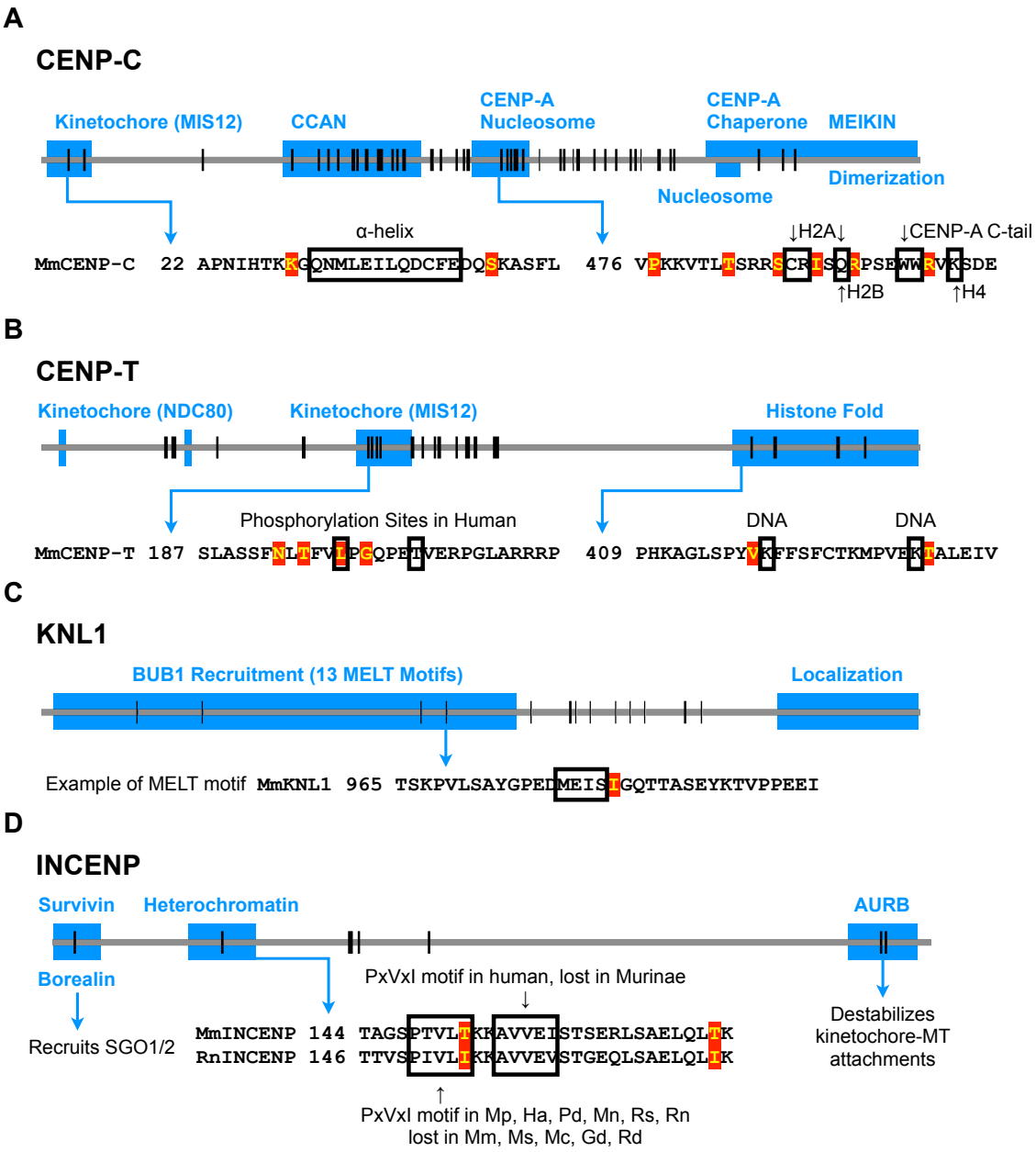


Figure 2.9. Protein domains that lead to microtubule destabilizer recruitment are recurrently evolved.

Each horizontal line represents the entire protein for each gene, and vertical lines represent positions of positively selected amino acids. Blue boxes show known functional domains from previous studies. Amino acid sequences within domains of interest are shown, with positively selected residues highlighted in red and known functional residues outlined in black.

(A) Signatures of positive selection are found throughout CENP-C. In the kinetochore domain, the a-helix interacts with MIS12 (Petrovic et al., 2016). The CCAN domain (also known as PEST domain) interacts with CENP-HIKM (Klare et al., 2015) and CENP-LN (Pentakota et al., 2017), and together forms the CENP-ACHIKMLN complex (Weir et al., 2016). In the domain

interacting with CENP-A nucleosomes (also known as central region), residues interacting with H2A, H2B, H4 and the CENP-A C-terminal tail are indicated. This domain binds CENP-A nucleosomes more specifically than the more C-terminal nucleosome binding domain (also known as CENP-C motif), which also interacts with H3 nucleosomes (Allu et al., 2019; Kato et al., 2013). The CENP-C C-terminus has multiple functions, including M18BP1 recruitment (Dambacher et al., 2012), MEIKIN recruitment (Kim et al., 2015), and dimerization (Sugimoto et al., 1997).
(B) Signatures of positive selection are found in the kinetochore interaction domain and histone fold domain of CENP-T. CDK1-dependent phosphorylation at Thr195 and Ser201 in human CENP-T (substituted with Leu and Thr, respectively, in mice) regulates MIS12 recruitment (Rago et al., 2015; Veld et al., 2016). Signatures of positive selection are detected around these regulatory residues for MIS12 recruitment. Some DNA interacting residues within the histone fold domain are shown (Nishino et al., 2012).
(C) Signatures of positive selection are found in the domain of KNL1 that recruits BUB1 via repeated MELT motifs (Krenn et al., 2013). One MELT motif is shown as an example.
(D) Signatures of positive selection are found in domains of INCENP that interact with Borealin/Survivin, with heterochromatin, and with Aurora B kinase. Heterochromatin recruits INCENP (Abe et al., 2016; Ainsztein et al., 1998; Kang et al., 2011), and Borealin mediates the interaction with SGO1/2 (Tsukahara et al., 2010). Survivin binds cohesin and pH3T3 at pericentromeres (Kelly et al., 2010; Wang et al., 2010; Yamagishi et al., 2010), providing another mechanism to localize the CPC. A PxVxI motif, which interacts with the HP1 chromoshadow domain, is present in some Murinae species and lost in others, shown with *Mus musculus* (Mm) and *Rattus norvegicus* (Rn) as examples. Other species from the phylogenetic tree in Figure 2.7B: *Mus spretus* (Ms), *Mus caroli* (Mc), *Mus pahari* (Mp), *Hylomyscus alleni* (Ha), *Praomys delectorum* (Pd), *Mastomys natalensis* (Mn), *Grammomys dolichurus* (Gd), *Rhabdomys dilectus* (Rd), and *Rhynchomys soricoides* (Rs).

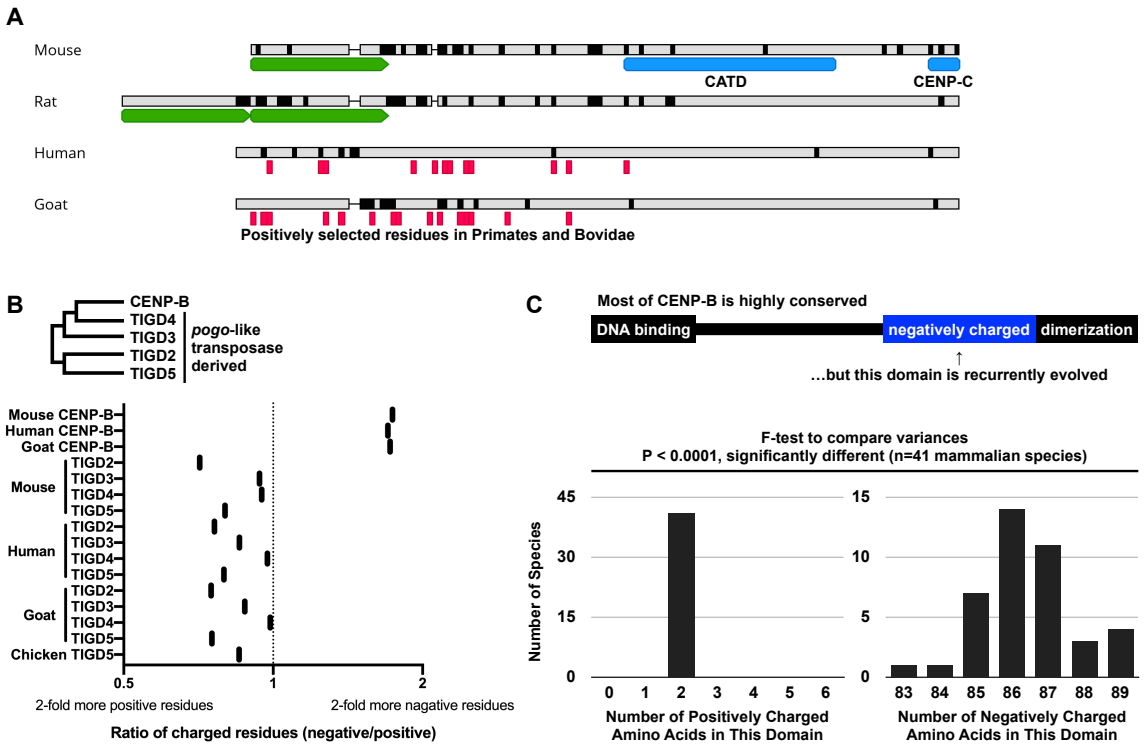


Figure 2.10. Changes in CENP-A and CENP-B are not analyzed by standard methods to detect adaptive evolution.

(A) Changes in CENP-A N-terminal tails. CENP-A amino acid sequences of four mammalian species are aligned. Known domains of CENP-A are shown in blue boxes, deviation from the consensus sequence of all four species is shown in black, and deletions are shown as thin horizontal lines. Signatures of positive selection were previously found in primate CENP-A (Schueler et al., 2010), shown in red boxes in the human sequence. We used bovine genomes (Chen et al., 2019) to detect signatures of positive selection in CENP-A, and the result is shown in the goat sequence. Such signatures are mostly found in the N-terminal tail. The N-terminal tail of Murinae CENP-A is either short (as in mouse) or long with two tandem duplicates (as in rat) (green boxes). Thus, alignment of the Murinae CENP-A N-terminal tail is difficult and removed from our PAML analysis.

(B) CENP-B negatively charged domain. Mouse, human, and goat are shown as examples of genomes with CENP-B and paralogous pogo-like transposases. The ratio of negatively charged to positively charged amino acids is plotted. As pogo-like transposases have fewer negatively charged amino acids than CENP-B, the negatively charged domain is likely unique to CENP-B.

(C) Changes in the CENP-B negatively charged domain. Although most of CENP-B is highly conserved, the number of negatively charged amino acids is variable in mammals. For comparison, the number of positively charged amino acids does not change in this domain. The number of species for each number of positively charged or negatively charged amino acids in this domain is plotted. CENP-B sequences of 41 bovine, primate and rodent species were obtained from the Ensembl genome browser.

In the kinetochore module, proteins can adapt to weaken the kinetochore pathway by reducing either kinetochore assembly or BUB1 binding to the kinetochore (Figure 2.7A). We find rapid evolution of the kinetochore proteins DSN1, KNL1, and NDC80. DSN1 is a component of the MIS12 complex, which assembles onto the CCAN and serves as a platform for binding KNL1 and the NDC80 complex (Petrovic et al., 2014). KNL1 contains multiple protein docking motifs, including repeated MELT motifs that recruit BUB1 kinase (Musacchio and Desai, 2017). Thus, changes in DSN1 and KNL1 can regulate kinetochore assembly and BUB1 recruitment. Consistent with the possibility that these interfaces evolve to modulate effector recruitment, we find signatures of positive selection in the MELT motifs of KNL1 (Figure 2.9C). NDC80 is the major microtubule binding protein in the kinetochore, but we find signatures of positive selection in the coiled-coil domain and not in the microtubule interacting domain. The coiled-coil domain recruits the SKA complex, which stabilizes kinetochore-microtubule attachment (Veld et al., 2019) and could be involved in counteracting destabilizing activities exploited by selfish centromeres.

Although selfish centromere DNA is likely unable to exploit heterochromatin to drive, inner centromere proteins can adapt to increase effector recruitment through the heterochromatin pathway relative to the kinetochore pathway in our model. In the inner centromere module (Figure 2.7A), INCENP is a scaffold component of the CPC that interacts directly with heterochromatin and indirectly with SGO1/2 (Abe et al., 2016; Ainsztein et al., 1998; Kang et al., 2011; Tsukahara et al., 2010). Other CPC components, Borealin and Survivin, regulate SGO1/2 recruitment and pericentromeric localization (Kelly et al., 2010; Tsukahara et al., 2010; Wang et al., 2010; Yamagishi et al., 2010). The catalytic component of the CPC is Aurora B kinase, which phosphorylates kinetochore substrates to destabilize microtubule interactions and is thus a potential drive effector. We find that positive selection shapes the domains of INCENP that

interact with Borealin/Survivin, with HP1, and with Aurora B (Figure 2.9D), suggesting that INCENP can adapt to selfish centromere DNA by modulating its localization to pericentromeric heterochromatin and ultimately the recruitment of SGO1/2 and Aurora B. Furthermore, we find rapid evolution of SGO2, suggesting that it can also tune the relative strength of the two pathways through mutations that modulate its recruitment by either pathway. In comparison, SGO1 is a paralog of SGO2 that does not recruit MCAK (Yao and Dai, 2012) and does not have signatures of positive selection, suggesting that evolutionary pressure to regulate MCAK recruitment shapes SGO2 evolution. Overall, our molecular evolution analyses show signatures of positive selection in both the kinetochore and heterochromatin pathways. We find these changes both in domains that interact directly with DNA and in protein-protein interaction domains leading to recruitment of drive effectors. These results are consistent with our parallel pathway model for drive and suppression, but not with a simpler model of an arms race limited to centromere DNA and DNA binding proteins.

2.5. Materials and Methods.

Mice

Mouse strains were purchased from the Jackson Laboratory (ZALENDE/EiJ, stock #001392 corresponds to CHPO; C57BL/6J, stock# 000664) and from Envigo (NSA, stock# 033 corresponds to CF-1). CHPO males were crossed to CF-1 females to generate hybrids shown in Figure 1c. The CHPO strain contains seven Robertsonian fusions (Rb(1.3), Rb(4.6), Rb(5.15), Rb(11.13), Rb(8.12), Rb(9.14), and Rb(16.17)), each of which pairs with two CF-1 chromosomes in CHPO hybrid meiosis I to form a trivalent (Chmátal et al., 2014). We included only bivalents (chromosome 2, 7, 10, 18, 19, X) in our analyses to avoid complications of trivalents.

In order to generate CENP-B null mice, 1-cell embryos (from female CF-1 and male DBA/2J x C57BL/6J hybrid) were collected and microinjected with Cas9 mRNA (TriLink, CleanCap Cas9 mRNA, L-7606) and gRNA (GAAGAACAAGCGCGCCA) (Thermo Fisher scientific, GeneArt Precision gRNA Synthesis Kit, A29377). Embryos were cultured *in vitro* until blastocyst stage and transferred to pseudopregnant females to produce a founder mouse carrying 37bp deletion (TGAGCACCATCCTGAAGAACAAGCGCGCCATCCTGGC) that produces a premature stop codon at Leu100 in the DNA binding domain. The founder was crossed with C57BL/6J for multiple generations to remove possible off-target mutations. Mice were genotyped by extracting genomic DNA from tail clip (QIAGEN, DNeasy Blood & Tissue Kit, 69504) and amplifying a *Cenpb* fragment (Agilent, Herculase II Fusion DNA Polymerase). To generate *Cenpb* null mice with larger and smaller centromeres, CHPO females were crossed to C57BL/6J *Cenpb* null males to generate first generation hybrid females, which were then crossed to C57BL/6J *Cenpb* null males to generate second-generation hybrid females as shown in Figure 2B and Supplementary Figure 1D. All animal experiments were approved by the Institutional Animal Care and Use Committee and were consistent with the National Institutes of Health guidelines.

Table 3. Reagents

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Alexa Fluor 488 donkey anti-mouse IgG	Molecular Probes	Cat# A-21202; RRID:AB_141607
Alexa Fluor 488 donkey anti-rabbit IgG	Molecular Probes	Cat# A-21206; RRID:AB_141708
Alexa Fluor 594 donkey anti-mouse IgG	Molecular Probes	Cat# A-21203; RRID:AB_141633
Alexa Fluor 594 donkey anti-rabbit IgG	Molecular Probes	Cat# A-21207; RRID:AB_141637
Alexa Fluor 647 donkey anti-mouse IgG	Molecular Probes	Cat# A-31571; RRID:AB_162542
Alexa Fluor 647 donkey anti-rabbit IgG	Molecular Probes	Cat# A-31573; RRID:AB_2536183
Rabbit anti-human H3K9me3	Abcam	Cat# ab8898; RRID:AB_306848
Rabbit anti-mouse CENP-A (C51A7)	Cell Signaling	Cat# 2048; RRID: AB_1147629
Mouse anti-human CENP-B (F-4)	Santa Cruz Biotechnology	Cat# sc-376283; RRID:AB_10988421
Mouse anti-human CENP-B (2D-7)	Santa Cruz Biotechnology	Cat# sc-32285; RRID:AB_627246
Rabbit anti-mouse CENP-C	Yoshinori Watanabe, University of Tokyo; Kim et al., 2015	N/A
Mouse anti-mouse SGO2	Yoshinori Watanabe, University of Tokyo; Kawashima et al., 2010	N/A
Rabbit anti-mouse HEC1	Robert Benezra, Memorial Sloan-Kettering Cancer Center; Diaz-Rodríguez et al., 2008	N/A
Bacterial and Virus Strains		
DH5a subcloning efficiency competent cells	Invitrogen	18265-017
Stellar competent cells	Clontech TAKARA	636763
Chemicals, Peptides, and Recombinant Proteins		
Pregnant Mare Serum Gonadotropin (PMSG)	Calbiochem	367222
CARD HyperOva	Cosmo Bio	KYD-010-EX
Mineral Oil	Sigma Millipore	M5310
Milrinone	Sigma Millipore	M4659
Vectashield with DAPI	Vector laboratories	H-1200
SiR-DNA	Cytoskeleton, Inc.	CY-SC007
λ-phosphatase	New England Biolabs	P0753S
Herculase II Fusion DNA Polymerase	Agilent	600675
In-Fusion HD Cloning Kit	Clontech TAKARA	639648
NucleoSpin Gel and PCR Clean-Up	MACHEREY-NAGEL	740609
NucleoSpin Plasmid	MACHEREY-NAGEL	740588
T7 mScript Standard mRNA Production System	Cell Script	C-MSC100625
GeneArt Precision gRNA Synthesis Kit	Thermo Fisher Scientific	A29377
MEGAclear Transcription Clean-Up Kit	Thermo Fisher Scientific	AM1908
Experimental Models: Organisms/Strains		
Mouse: C57BL/6J	The Jackson Laboratory	664
Mouse: ZALENDE/EiJ (CHPO)	The Jackson Laboratory	1392
Mouse: NSA (CF-1)	Envigo	33
Oligonucleotides		
Primers for *Cenpb* genotyping: FWD: 5'-CAGCTGACGTTCCGGGAGAA-3', REV: 5'-GGGGACAGCTTGTTGGTCTT-3'	This paper	N/A
gRNA target sequence for *Cenpb* null mice: 5'-GAAGAACAAGCGCGCCA-3'	This paper	N/A
gRNA target sequence for minor satellite repeats: 5'-ACACTGAAAAACACATTCGT-3'	This paper	N/A
Recombinant DNA		
H2B-EGFP	Akera et al., 2017	N/A
H2B-mCherry	Akera et al., 2017	N/A
dCas9-EGFP	This paper	N/A
dCas9-mCherry	This paper	N/A
EGFP-MmCENP-C	This paper, cDNA from liver	N/A
EGFP-RnCENP-C	This paper, cDNA from liver	N/A
EGFP-RpCENP-C	This paper, cDNA from liver	N/A
Software and Algorithms		
GraphPad Prism v7	GraphPad	https://www.graphpad.com/
FIJI/ImageJ v2.0.0-rc-61/1.51n	Schindelin et al., 2012; Schneider et al., 2012	https://fiji.sc/
Geneious Prime v2020.1.2	Geneious	https://www.geneious.com/
LongRanger v2.2.2	10x Genomics	https://www.10xgenomics.com/
Supernova v2.1.1	10x Genomics	https://www.10xgenomics.com/
ncbi-blast-2.10.1+	NCBI	https://blast.ncbi.nlm.nih.gov/
anaconda 4.6.14	Anaconda	https://www.anaconda.com/
paml 4.9	Yang, 2007	https://anaconda.org/bioconda/paml
MAFFT 7.407	Katoh and Standley, 2013; Katoh et al., 2002	https://anaconda.org/bioconda/mafft
RAxML 8.2.12	Stamatakis, 2014	https://anaconda.org/bioconda/raxml

Oocyte collection and culture

Female mice (8-14 weeks of age) were hormonally primed with 5U of Pregnant Mare Serum Gonadotropin (PMSG, Calbiochem, cat# 367222) or 0.1mL of CARD HyperOva (Cosmo Bio, KYD-010-EX) 44-48 h prior to oocyte collection. Germinal vesicle (GV)-intact oocytes were collected in M2 medium (Sigma, M7167), denuded from cumulus cells, and cultured in Chatot-Ziomek-Bavister (CZB) medium (Thermo Fisher, MR019D) in a humidified atmosphere of 5% CO_2 in air at 37.8C˚. During collection, meiotic resumption was inhibited by addition of 2.5 mM milrinone. Milrinone was subsequently washed out to allow meiotic resumption. Oocytes were checked for GVBD (germinal vesicle breakdown), and those that did not enter GVBD stage were removed from the culture.

Oocyte microinjection

GV oocytes were microinjected with ~5 pl of cRNAs in M2 medium (with 2.5 mM milrinone and 3mg/mL BSA) at room temperature (RT) with a micromanipulator TransferMan NK 2 (Eppendorf) and picoinjector (Medical Systems Corp.). After the injection, oocytes were kept in milrinone for 16 h to allow protein expression. cRNAs used for microinjections were dCas9-EGFP (dead Cas9 with EGFP at the N terminus) at 1000ng/μL, dCas9-mCherry (dead Cas9 with mCherry at the N terminus) at 1000ng/μL, gRNA that targets minor satellite repeat (ACACTGAAAAACACATTCGT) at 200ng/μL, H2B-EGFP (human histone H2B with EGFP at the C terminus) at 150ng/μL, H2B-mCherry (human histone H2B with mCherry at the C terminus) at 150ng/μL, EGFP-MmCENP-C (mouse CENP-C with EGFP at the N terminus) at 100ng/μL, EGFP-RnCENP-C (rat CENP-C with EGFP at the N terminus) at 100ng/μL, and EGFP-RpCENP-C (*R. pumilio* CENP-C with EGFP at the N terminus) at 100ng/μL. Mouse, rat and *R. pumilio* CENP-C sequences were cloned from cDNA libraries from liver. Mouse and rat CENP-C sequences were verified by mm10 *Mus musculus* and rn6 *Rattus norvegicus* reference

genomes. *R. pumilio* CENP-C sequence was verified by the genome sequence (personal communication with Ricardo Mallarino). cRNAs were synthesized using the T7 mScriptTM Standard mRNA Production System (CELL SCRIPT) or mMESSAGE mMACHINE SP6 Transcription Kit (Thermo Fisher scientific). gRNAs were synthesized using GeneArt Precision gRNA Synthesis Kit (Thermo Fisher scientific A29377).

Live imaging and chromosome position assay

For the chromosome position assay, oocytes were collected and microinjected with the constructs indicated in the figure legends. After inducing meiotic resumption by washing out milrinone, oocytes were placed into 2μL drops of CZB media covered with mineral oil in a glass-bottom tissue culture dish (FluoroDish FD35-100) in a heated environmental chamber with a stage top incubator (Incubator BL and Heating Insert P; PeCon GmBH) to maintain 37C˚. Confocal images were collected with a microscope (DMI4000 B; Leica) equipped with a 63x 1.3 NA glycerol-immersion objective lens, an xy piezo Z stage (Applied Scientific Instrumentation), a spinning disk confocal scanner (Yokogawa Corporation of America), and an electron multiplier charge-coupled device camera (ImageEM C9100-13; Hamamatsu Photonics), controlled by MetaMorph software (Molecular Devices). Excitation was with an LMM5 laser merge module with 488- and 593-nm lasers (Spectral Applied Research) or a Vortran Stradus VersaLase 4 laser module with 405 nm, 488 nm, 561 nm, and 639 nm lasers (Vortran Laser Technology). Confocal images were collected as z stacks at 0.5 μm intervals to visualize the entire meiotic spindle. The position of the spindle near the cortex was confirmed by differential interference contrast images. The spindle equator was determined as a middle of the spindle. The chromosome position of each bivalent was determined as a crossover site and normalized by the distance between spindle equator and spindle poles.

Oocyte immunocytochemistry

After inducing meiotic resumption by washing out milrinone (4.5 hours for prometaphase staining and 7.5 hours for metaphase staining), MI oocytes were fixed in freshly prepared 2% paraformaldehyde in PBS with 0.1% Triton X-100, pH 7.4, for 20 min at RT, permeabilized in PBS with 0.1% Triton X-100 for 15 min at RT, placed in blocking solution (PBS containing 0.3% BSA and 0.01% Tween-20) 15 min RT or overnight at 4C, incubated 1-2 h with primary antibodies in blocking solution, washed 3 times for 15 min each, incubated 1 h with secondary antibodies, washed 3 times for 15 min each, and mounted in Vectashield with DAPI (Vector, H-1200) to visualize chromosomes. Primary antibodies used for this study were rabbit anti-human H3K9me3 (1:500; Abcam, ab8898), mouse anti-mouse SGO2 (1:500, a gift from Yoshinori Watanabe), and rabbit anti-mouse CENP-C (1:2500, a gift from Yoshinori Watanabe). Secondary antibodies were Alexa Fluor 488–conjugated donkey anti-rabbit or donkey anti-mouse, Alexa Fluor 594–conjugated donkey anti-rabbit or donkey anti-mouse, or Alexa Fluor 647–conjugated donkey anti-rabbit or donkey anti-mouse (1:500, Invitrogen). Confocal images were collected as z stacks at 0.5 μm intervals to visualize the entire meiotic spindle, using the spinning disc confocal microscope described above. To quantify centromere signal ratios, optical slices containing centromeres from the same bivalent were added to produce a sum projection using Fiji/ImageJ. Ellipses were drawn around the centromeres, and signal intensity was integrated over each ellipse after subtracting cytoplasmic background. Ratios were obtained for each bivalent by dividing the intensity of the larger centromere by that of the smaller centromere, as determined by dCas9 signal intensity.

Whole Genome Sequencing of Six Murinae Species

Frozen tissue samples from male individuals were obtained from the Museum of Vertebrate Zoology, Berkeley, CA (MZV) and the Field Museum of Natural History, Chicago, IL (FMNH).

Hylomyscus alleni (MVZ Mamm 196246) was captured in Cameroon in 2000, *Praomys delectorum* (MVZ Mamm 221157) was captured in Malawi in 2007, *Mastomys natalensis* (MVZ Mamm 221054) was captured in Malawi in 2007, *Grammomys dolichurus* (MVZ Mamm 221001) was captured in Malawi in 2007, *Rhabdomys dilectus* (FMNH 192475) was captured in Malawi in 2006, and *Rhynchomys soricoides* (FMNH 198792) was captured in The Philippines in 2008. All genomes were sequenced in the Center for Applied Genomics at Children's Hospital of Philadelphia. High molecular weight DNA was extracted following the protocol provided by 10xGenomics (CG000072 Rev B Sample Preparation Demonstrated Protocol, DNA Extraction from Fresh Frozen Tissue). Extracted DNA was quality controlled (CG00019 Rev B Sample Preparation Demonstrated Protocol, High Molecular Weight DNA QC), and all of the samples had a mean length greater than 50kb, and high enough concentration to dilute to 1ng/µL for library preparation. Chromium Genome Reagent Kits v2 from 10xGenomics was used to prepare libraries of 2x150 base reads, with read 1 constituting 10xBarcode (16bp) + nmer (6bp) + genome sequence (128bp) and read 2 constituting genome sequence (150bp). i7 index used 8bp sample index, and i5 index was not used. Sequencing depth was calculated based on putative genome size 3Gb and coverage 56x, following 10xGenomics R&D recommendation, and the libraries were sequenced with Illumina HiSeq. Demultiplexed FASTQ files were analyzed using the LongRanger wgs -basic pipeline. This pipeline gave general QC statistics related to the 10x barcoding and number of read pairs present in the FASTQ files. All sample FASTQs contained more than 688M read pairs and have acceptable barcode diversity/% on whitelist. LongRanger was used to assemble genomes, using the *Mus musculus* (mm10) as reference. In parallel, Supernova was used to assemble *de novo* genomes. See Supplementary Table 1 for assembly statistics. In order to obtain protein coding sequences, mm10 annotation was used to annotate reference-guided assemblies, and translated BLAST (tblastn) was used to pull homologous sequences from de novo assemblies using *Mus musculus* protein sequences as query sequences.

Phylogenetic Tree Construction

The species tree shown in Figure 4B was obtained from maximum likelihood (RAxML) and Bayesian inference (MrBayes). The phylogeny within *Mus* was previously studied (Keane et al., 2011; Thybert et al., 2018). In order to resolve phylogeny in Murinae, the same set of genes that were used to construct a primate phylogenetic tree (Perelman et al., 2011) was aligned by MAFFT (Katoh and Standley, 2013; Katoh et al., 2002). The initial alignment was imported in Geneious Prime, and manually inspected for sequence alignment ambiguity. Ambiguous regions were removed from subsequent analyses. Maximum likelihood tree was constructed with RAxML (Stamatakis, 2014), and Bayesian inference tree was constructed with MrBayes (Huelsenbeck and Ronquist, 2001), with *Peromyscus maniculatus* as outgroup. Both inferences supported the tree topology shown in Figure 4B.

Molecular Evolution Analyses

In order to create a histogram in Figure 4C, alignments of mouse-rat orthologs were filtered for dS below 0.5, as higher dS values indicate misalignment. A list of genes for each subcellular compartment was obtained from Human Protein Atlas. Mouse-human orthologs were used to calculate average dN/dS for each subcellular compartment in Figure 4D. The analysis to identify signatures of positive selection (PAML) is highly sensitive to alignment errors, so automated genome-wide analysis is prone to false positives (van der Lee et al., 2017). To prevent these errors, alignments for selected genes were manually inspected. Coding sequences for each gene were aligned by Geneious Alignment (translation align) implemented in Geneious Prime, and manually inspected for sequence alignment ambiguity. Insertions or deletions as well as their flanking codons were removed from analyses. To test signatures of positive selection, we compared the likelihood of models of neutral codon evolution to models of codon evolution

allowing positive selection, implemented in PAML version 4 (Yang, 2007). The neutral model M1 (fixed dN/dS values between 0 to 1) and M2 (M1 parameters plus dN/dS > 1) were compared in the first test, and the neutral model M7 (dN/dS values fit a beta distribution from 0 to 1) and M8 (M7 parameters plus dN/dS > 1) were compared in the second test, assuming the F3x4 model of codon frequencies. Degree of freedom for each test was 2, and the log likelihood test was significant above 5.99 ($p < 0.05$). We first used the species tree, and signatures of positive selection were confirmed using a gene tree for each gene, created by RAxML.

Quantification and Statistical Analysis

Data points are pooled from at least two independent experiments. The following statistical methods were used: unpaired t test in Figures 2.2B, 2.2C, 2.2D, 2.2E, 2.2G, 2.2I, 2.3D, 2.3G, 2.3H, 2.5E, 2.5F, 2.6A, 2.6C, and 2.6D; Mann-Whitney U test in Figure 2.7D; ordinary one-way ANOVA in Figure 2.3D; chi square test for goodness of fit for deviations from 1 in Figure 2.1B and for statistical models (likelihood-ratio test) in Figure 2.7E and Table 2; Naïve Emprical Bayes (NEB) analysis and Bayes Empirical Bayes (BEB) analysis in Figures 2.7B and 2.7E; F test to compare variance in Figure 2.10C. The exact value of n, what n represents, and definition of center can be found in the figure legends for each experiment. Unpaired t test, Mann-Whitney U test, ordinary one-way ANOVA, and F test were performed using GraphPad Prism; chi square tests were performed using Excel; NEB and BEB analyses were performed using PAML model 2 and 8. *P* value of less than 0.05 was judged as statistically significant.

Chapter 3. Eukaryotic Centromere Evolution

3.1. Summary of Parallel Pathway Model for Drive and Suppression

Here we propose a parallel pathway model for drive and suppression of selfish centromeres: centromere DNA can exploit the kinetochore pathway to increase effector recruitment, and centromere protein evolution can make centromeres functionally equivalent by minimizing the contribution of the kinetochore pathway relative to the heterochromatin pathway (Figure 3.1). This model predicts that disruption of either pathway will reduce effector (e.g., SGO2) recruitment, but the functional consequences will depend on which pathway is affected. Centromeres become either functionally more similar if the asymmetric kinetochore pathway is weakened, or more different if the symmetric heterochromatin pathway is weakened. In our experiments, either introduction of a divergent allele of CENP-C or deletion of CENP-B leads to SGO2 reduction to a similar extent (Figure 2.2B and Figure 2.6A). However, genetically different centromeres in CHPO hybrid oocytes become functionally more similar when rat CENP-C is expressed (Figure 2.2I), whereas they become functionally more different when CENP-B is deleted (Figure 2.6D). The CENP-C results are consistent with our model prediction that natural selection has acted on CENP-C interfaces involved in effector recruitment, so a divergent rat CENP-C interacts less well with mouse binding partners in the kinetochore pathway. Therefore, expression of rat CENP-C weakens the asymmetric kinetochore pathway, making the symmetric heterochromatin pathway relatively more dominant. In contrast, CENP-B deletion weakens the symmetric heterochromatin pathway, as shown by reduced H3K9me3, making the asymmetric kinetochore pathway more dominant. Loss of CENP-B also reduces CENP-C recruitment but does not affect the asymmetry between larger and smaller centromeres (Figure 2.6C).

Parallel Pathway Model

Drive	(1) Selfish centromere DNA evolves to recruit more effectors through the kinetochore pathway (4) Selfish centromere DNA maintains CENP-B boxes because CENP-B recruits effectors
Suppression	(2) CENP-B equalizes centromeres through the heterochromatin pathway (3) Proteins in both pathways evolve to modulate effector recruitment

Figure 3.1. Summary of parallel pathway model.
Selfish centromere DNA recruits more effector proteins through the kinetochore pathway to drive (1). CENP-B equalizes centromeres through the symmetric heterochromatin pathway (2), but the asymmetric kinetochore pathway is dominant in our hybrid model system. Proteins in both pathways evolve to functionally equalize genetically different centromeres by modulating effector recruitment (3). Selfish centromere DNA can evolve again to recruit more effector proteins. However, CENP-B boxes will be maintained because CENP-B recruits effector proteins (4).

Our molecular evolution analyses show adaptive evolution in multiple centromere proteins and in specific domains that interact with CENP-A chromatin or with other proteins leading to effector recruitment (Figure 2.7 and Figure 2.9). The previous model of a molecular arms race limited to interactions between centromere DNA and DNA-interacting proteins (such as CENP-A) (Henikoff et al., 2001) does not explain the more widespread recurrent evolution of centromere proteins. In contrast, our parallel pathway model predicts recurrent evolution of proteins in both pathways to equalize centromeres by weakening the kinetochore pathway or strengthening the heterochromatin pathway. In our model, selfish centromere DNA evolves to exploit the kinetochore pathway by recruiting more of a protein that ultimately recruits effectors. To suppress functional differences between centromeres, proteins in the kinetochore pathway can adapt to minimize the impact of selfish centromere DNA on kinetochore formation or effector recruitment. Indeed, our findings with *R. pumilio* CENP-C indicate that mouse CENP-C is not optimized for maximum binding to mouse centromeres, effectively weakening the kinetochore pathway. Furthermore, proteins in the heterochromatin pathway such as CENP-B can adapt to increase effector recruitment equally at all centromeres, or INCENP and SGO2 can adapt by modulating their recruitment by either pathway (Figure 3.1). The acidic domain of CENP-B is

implicated in recruiting heterochromatin proteins (Otake et al., 2020), and the number of negatively charged amino acids in this domain is recurrently changed in mammals (Figure 2.10B-C). Although these changes are not analyzed in PAML, they suggest that CENP-B may have evolved to regulate pericentromeric heterochromatin. Overall, a protein network for effector recruitment can adapt to minimize asymmetric recruitment by selfish centromere DNA, while maintaining essential functions of the kinetochore and of microtubule destabilizing factors for accurate chromosome segregation.

3.2. CENP-B Evolution

Our results suggest an explanation for the conservation of CENP-B in mammals, as well as the presence of its binding sequence, the CENP-B box, at most mammalian centromeres with the notable exception of the Y chromosome. Although CENP-B is the only centromere protein known to bind a specific DNA sequence in mammals, neither the protein nor the binding sequence is essential for centromere function (Amor et al., 2004; Hudson et al., 1998; Kapoor et al., 1998; Logsdon et al., 2019; Perez-Castro et al., 1998). We propose that CENP-B is conserved because it suppresses functional differences between centromeres by strengthening the heterochromatin pathway (Figure 3.2), consistent with a more general function of heterochromatin in suppressing many selfish genetic elements (Allshire and Madhani, 2017). This CENP-B function is important only when centromeres of homologous chromosomes are different, which would frequently occur in outbred populations. Loss of CENP-B therefore increases functional difference between larger and smaller centromeres in our hybrid model, but does not significantly impair fertility or viability in inbred laboratory strains (Hudson et al., 1998; Kapoor et al., 1998; Perez-Castro et al., 1998). A potential cost of increasing heterochromatin, however, is that its invasion into CENP-A chromatin disrupts centromere function (Ohzeki et al., 2016). We therefore propose that mammalian CENP-B has acquired an additional function to maintain CENP-A chromatin, by recruiting CENP-C and CENP-A chromatin regulators (Fachinetti et al., 2015; Otake et al., 2020) (Figure 3.2). Consistent with this idea, CENP-A chromatin is reduced in *Cenpb* null oocytes (Figure 2.5F). By regulating both CENP-A chromatin and heterochromatin, alternative functions of CENP-B in different chromatin environments may suppress functional differences between centromeres through heterochromatin while maintaining centromere function.

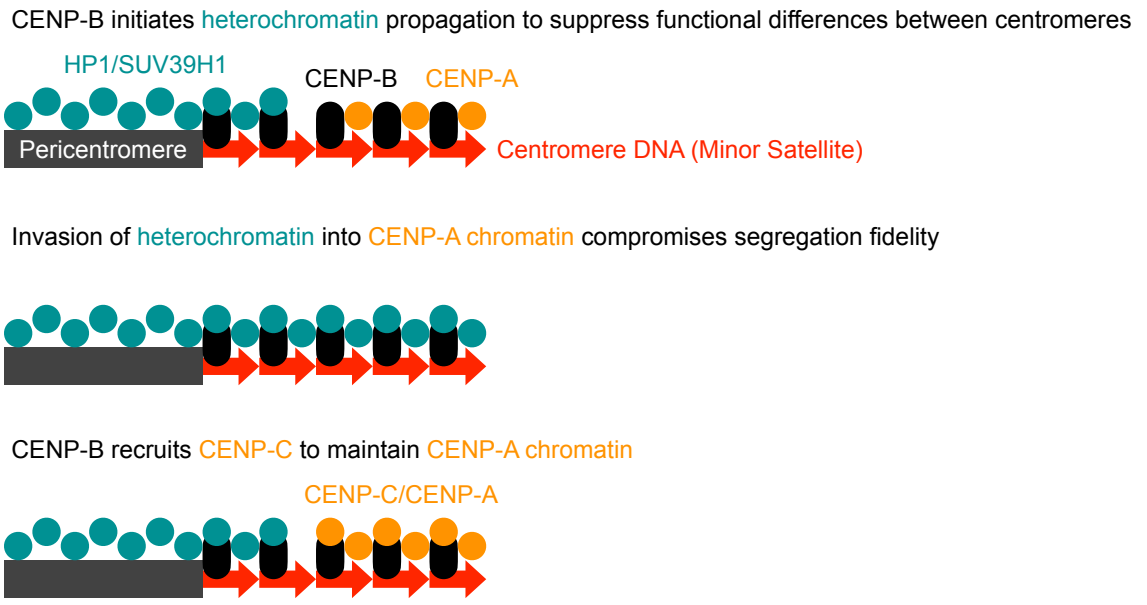


Figure 3.2. Dual functions of CENP-B in heterochromatin and CENP-A chromatin. CENP-B initiates heterochromatin formation to equalize centromeres (top). Despite the difference in CENP-B binding sites, larger and smaller centromeres have similar amounts of H3K9me3 (Figure 1D), indicating that heterochromatin formation is insensitive to CENP-B abundance, likely due to self-propagation of heterochromatin. Invasion of heterochromatin into CENP-A compromises centromere function (middle). To prevent this disruption, we propose that CENP-B has acquired an additional function in CENP-A chromatin (bottom): CENP-B recruits CENP-C but does not contribute to CENP-C asymmetry between larger and smaller centromeres (Figure 2C), suggesting that only CENP-B within CENP-A chromatin recruits CENP-C. Thus, CENP-B functions in heterochromatin and CENP-A chromatin are insensitive to repeat expansion.

CENP-B can suppress differences between centromeres only if its functions are insensitive to expansion of the number of CENP-B binding sites; otherwise it would contribute to higher levels of effector recruitment by DNA repeat expansions. Indeed, we find that CENP-B does not contribute to asymmetry in CENP-C recruitment between larger and smaller centromeres (Figure 2.6C), despite 6- to 10-fold differences in minor satellite sequences containing CENP-B boxes (Iwata-Otsubo et al., 2017). This result suggests that CENP-B recruits CENP-C only within the CENP-A chromatin domain, so that CENP-B binding outside of this domain does not strengthen the kinetochore pathway. Furthermore, the heterochromatin symmetry between larger and smaller centromeres (Figure 2.1B) suggests that although CENP-B contributes to initiating

heterochromatin formation, for example by recruiting an H3K9 methyltransferase, heterochromatin spreading does not depend on the number of CENP-B boxes. Initiation of heterochromatin propagation is a common mechanism to regulate heterochromatin formation, as in the example of X inactivation where XIST initiates heterochromatinization of the entire chromosome (Allshire and Madhani, 2017). Thus, CENP-B functions in CENP-A chromatin and heterochromatin are insensitive to repeat expansion. A centromere variant completely lacking CENP-B boxes, however, will lose to an existing centromere in female meiosis because it will recruit less effectors by both the kinetochore and heterochromatin pathways. Therefore, CENP-B boxes are maintained at most centromeres (Figure 3.1), but this selective pressure does not affect the Y chromosome, which never experiences female meiosis and does not bind CENP-B (Gamba and Fachinetti, 2020).

3.3. Parallel Pathway Model and Karyotype Evolution

The definition of telocentric and metacentric chromosomes is arbitrary, and the position of kinetochore on a chromosome likely has little effect on the processes of chromosome segregation. However, many species have either mostly telocentric chromosomes or mostly metacentric chromosomes, rather than a mixture of telocentric and metacentric chromosomes (Molina et al. 2014; Pardo-Manuel de Villena and Sapienza, 2001). Frequent changes between mostly telocentric karyotypes and mostly metacentric karyotypes are reported (Molina et al. 2014; Pardo-Manuel de Villena and Sapienza, 2001). Our parallel pathway model provides explanations for transition from telocentric to metacentric chromosomes as well as transition from metacentric to telocentric chromosomes.

Centromere drive of Robertsonian (Rb) fusion chromosomes can explain transition of telocentric to metacentric karyotypes. Chromosome fusion of telocentric chromosomes can change centromere satellite repeat number (Figure 1.8, Robertsonian fusion). For example, Rb fusion of telocentric chromosomes occurs in about one in 1000 newborn humans (Hamerton et al., 2008). Although all of the chromosomes are telocentric in standard laboratory mouse strains, some populations of *Mus musculus domesticus* have metacentric chromosomes from Rb fusion events (Britton-Davidian et al., 2000; Garagna et al., 2014; Piálek et al., 2005). When telocentric chromosomes fuse, centromeric satellite DNA repeat number can increase or decrease due to the recombination. Metacentric chromosomes have larger CENP-A chromatin relative to telocentric chromosomes in some strains but have smaller CENP-A chromatin in other strains. Strains with larger CENP-A chromatin on metacentric chromosomes are from populations where Rb fusion events are commonly found (Chmátal et al., 2014). According to our parallel pathway model, having larger CENP-A chromatin means stronger kinetochore pathway for effector recruitment.

When metacentric chromosomes are more likely to be inherited than the homologous telocentric chromosomes, metacentric chromosomes will quickly fix in a population.

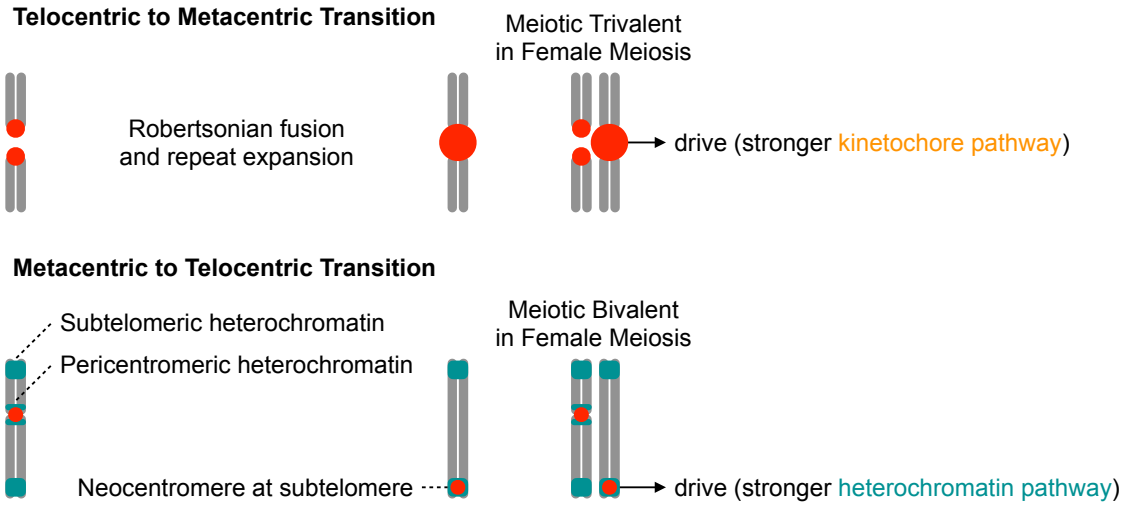


Figure 3.3. Transitions between telocentric and metacentric chromosomes.
Centromeres of Robertsonian fusion chromosomes that increase centromeric satellite repeats will drive in female meiosis (top). Neocentromere formation at subtelomeres will drive if subtelomeric heterochromatin recruit more effector proteins than pericentromeric heterochromatin on the homologous counterpart (bottom).

Although Rb fusion can explain transition from telocentric to metacentric chromosomes, it is not clear how transition in the other direction occurs. Our parallel pathway model proposes that the heterochromatin pathway recruits effector proteins for drive, and this model can explain transition from metacentric to telocentric karyotypes. Cases of neocentromere formation are reported in humans, and comparative study suggest that centromere repositioning is frequent (Rocchi et al., 2011). Neocentromeres at telomeres can drive if they have larger subtelomeric heterochromatin than pericentromeric heterochromatin on the homologous counterpart. Heterochromatin alone is insufficient to recruit effector proteins, but once neocentromeres are formed at telomeres and centromeric histone marks (e.g., phosphorylation at H3T3 and H2AT120) are present, subtelomeric heterochromatin likely contributes to effector recruitment. If the subtelomeric heterochromatin recruits more effectors than the pericentromeric heterochromatin on the

homologous chromosome, neocentromeres at telomeres will drive, leading to metacentric to telocentric transition. In contrast, neocentromeres on chromosome arms tend to have smaller (or very little, if any) heterochromatin. Such neocentromeres will likely lose in female meiosis because of reduced effector recruitment via the heterochromatin pathway. This model also predicts that subtelomeric satellites are also selfish, and that they can evolve to recruit more centromere proteins and/or expand to form a larger subtelomeric heterochromatin region.

3.4. Other Evolutionary Forces that Diversify Centromeres

Centromere drive in female meiosis is not the only evolutionary force that selects centromere protein variants. Indeed, centromeres evolve rapidly in organisms that only undergo symmetric meiosis, where all of the haploid cells from meiosis form functional gametes (Bensasson et al., 2008). This suggests another evolutionary force diversifies centromeres in these organisms.

One possibility is that centromere binding proteins are selected for non-kinetochore functions. In multicellular organisms, stem cell division is a regulatory point for proliferation or differentiation. Chromosome segregation machinery may have additional functions for differentiation and development. In flies and worms, kinetochore proteins are repurposed for neural development (Cheerambathur et al., 2019; Zhao et al., 2019), and in humans, KNL1 regulates brain size (Javed et al., 2018; Shi et al., 2016). Furthermore, anaphase-promoting complex APC/C has additional functions in stem cell differentiation (Oh et al., 2020). As KNL1 indirectly regulates APC/C activity through spindle assembly checkpoint signaling, it is possible that kinetochore proteins regulate cell identity.

Alternatively, selfish genetic elements may hijack centromeres. Transposons and extraneous genetic elements (e.g., plasmids and B chromosomes) are prime candidates for selfish genetic elements that drive rapid centromere evolution. In the following sections, I focus on drive and suppression of transposons and extraneous genetic elements.

3.5. Drive and Suppression of Transposons at Centromeres

Transposons are often inserted in centromeres. In *Drosophila* centromeres, CENP-A chromatin is located on islands of transposons (Chang et al., 2019). Transposons are found in satellite-free, evolutionary young centromeres (Nergadze et al., 2018) and in human neocentromeres (Chueh et al., 2009). Transposons achieve non-Mendelian inheritance by over-replication (Figure 1.1, drive by over-replication), and in principle they can be inserted anywhere on chromosomes. However, transposons benefit themselves in centromeres by two means. First, centromeres have no essential genes but are transcriptionally active. Transposon insertion at transcriptionally active genes is often deleterious, so it is selected against. Transposon insertion at transcriptionally inactive sites prevents transposition to new sites, as transposons need to be transcribed to transpose. Despite the absence of coding genes, centromeric chromatin is transcriptionally active, so transposons at centromeres are transcribed and can be transposed to other genomic loci. Second, if transposons at centromeres recruit centromere proteins, they can achieve preferential inheritance over the homologous counterpart without transposons in female meiosis (Figure 1.1, drive by biased segregation). Thus, centromeres provide unique opportunities for non-Mendelian inheritance of transposons.

Host genomes have evolved mechanisms to suppress transposon activity at centromeres. Transcriptional silencing by heterochromatin (characterized by H3K9me3 histone marks, HP1-mediated chromatin compaction, and DNA methylation) is the predominant strategy to suppress transposons (Janssen et al., 2018). Briefly, RNA-based silencing and protein-based silencing can initiate heterochromatin formation. RNA-based silencing is universal among eukaryotes (Gutbrod and Martienssen, 2020). The RNAi pathways in the last eukaryotic common ancestor involve Ago-like and Dicer-like proteins for siRNA silencing as well as Piwi-like proteins for piRNA silencing. The ancestral function of such RNAi pathways is likely to repress transposons and

viruses, and some eukaryotic lineages acquired an additional gene regulatory function by miRNA silencing. The role of RNAi in chromosome segregation is also conserved in multiple eukaryotic lineages, suggesting that it is also an ancestral function. Deletion of RNAi machinery compromises pericentromeric heterochromatin formation, leading to chromosome segregation errors.

However, RNAi alone seems insufficient to completely purge transposons from centromeres, as transcripts are required to initiate silencing. Indeed, fission yeast *S. japonicus* centromeres are mostly transposons despite the presence of RNAi machinery that targets transposons (Rhind et al., 2011). Protein-based silencing provides an additional layer of transposon silencing. After the divergence from *S. japonicus*, fission yeast species (e.g., *S. pombe*, *S. octosporus* and *S. cryophilus*) acquired CENP-B homologs, and their centromeres are largely transposon-free, despite the loss of RNAi machinery in *S. octosporus* and *S. cryophilus* (Upadhyay et al., 2017). Similarly, *S. pombe* RNAi machinery targets repetitive DNA, instead of transposons, suggesting that the RNAi has been repurposed from its ancestral function of transposon silencing to pericentromeric heterochromatin formation (Rhind et al., 2011). Absence of functional transposons at fungal centromeres correlates with the absence of RNAi (Yadav et al., 2018), suggesting that, once active transposons are lost, the RNAi machinery becomes dispensable.

The *pogo-like* transposase is one of the most widespread DNA transposons found in animals, plants, fungi and protozoans (Plasterk et al., 1999), and several eukaryotic lineages have domesticated *pogo-like* transposases that have lost transposition activity but been repurposed for other cellular processes (Gao et al., 2020; Mateo et al., 2014). CENP-B in yeasts and mammals is one such example (Casola et al., 2007; Kipling, 1997), and CENP-B in both lineages regulate heterochromatin formation (Kumon et al., 2021; Nakagawa et al., 2002; Okada et al., 2007; Otake

et al., 2020), suggesting that the heterochromatin formation is the ancestral function of CENP-B. Heterochromatin formation by yeast CENP-B prevents retrotransposon insertion (another class of transposons that use RNA as an intermediate for insertion) (Cam et al., 2007). It is not known whether mammalian CENP-B also prevents transposon insertion.

In summary, centromeres provide unique opportunities for transposon drive. Eukaryotic genomes have conserved RNA-based transposon silencing machinery, but transposons often escape from the RNAi silencing. Some eukaryotic lineages have evolved a protein-based transposon silencing which has successfully purged transposons from centromeres in some species. Transposons can also be domesticated to silence other types of transposons, and yeast CENP-B is one such example.

3.6. Plasmids and B Chromosomes

Plasmids and B chromosomes are extraneous genetic information that is usually dispensable for the host, but they exploit the host replication and segregation machinery for their inheritance. The 2µm plasmid is an example of a selfish plasmid found in budding yeasts (Rizvi et al., 2017). This plasmid does not encode proteins beneficial to the host but has *STB* centromere-like DNA and Rep1/2 proteins that bind STB as well as proteins for over-replication. Centromere-like *STB* recruits endogenous CENP-A in a Rep1/2-dependent manner and binds microtubules for segregation. In the absence of *STB*, 2µm plasmids are quickly lost due to the extreme mother bias of plasmid inheritance during mitosis. Analogous to the centromere drive hypothesis, an evolutionary arms race between selfish plasmids and centromere proteins can lead to rapid evolution of both. Under this model, centromere-like DNA on selfish plasmids evolve to hijack endogenous centromere proteins. As selfish plasmids impose significant load to the host, centromere protein variants that are not recruited by selfish plasmid DNA are selected.

It is hypothesized that genetically defined centromeres on budding yeast chromosomes are originally from selfish plasmids (Malik and Henikoff, 2009). Similar to the 2µm plasmids, *S. cerevisiae CDEIII* centromere DNA recruits CENP-A in a CBF3 protein complex-dependent manner. Another budding yeast *N. castellii* has a distinct point centromere sequence that is different from the consensus *CDE* sequences (Kobayashi et al., 2015). CBF3 still binds N. castellii centromere DNA, but the DNA-binding domain of a CBF3 protein complex is rapidly evolving (Kobayashi et al., 2015). No homologous proteins of CBF3 proteins are found outside budding yeasts, consistent with the idea that the *CDE* sequence and CBF3 proteins are brought by rapidly evolving selfish plasmids. Origin of centromere-specific DNA motifs can be from selfish plasmids or transposons: the fungus species *Mucor circinelloides* has a centromere-specific DNA motif, which is flanked by lineage-specific transposons (Navarro-Mendoza et al., 2019). The

Mucor circinelloides genome does not have CENP-A or CENP-C genes, but CENP-T and kinetochore proteins are localized at centromeres. Presence of these lineage-specific transposons correlates with the absence of CENP-A/C in Mucoromycotina species, implying that transposons which recruit CENP-T take over the canonical centromeres.

Diversity of chromosome segregation machinery is not limited to centromere/kinetochore. The nuclear envelope is one of the hallmarks of prokaryote-eukaryote difference. The nuclear envelope remains intact in the closed mitosis of budding yeasts, but it breaks down in the open mitosis of mammals. It is speculated that transitions of open and closed mitosis are driven by selfish genetic elements (Sazer et al., 2014). Plasmids and viruses first enter the cytoplasm, whereas transposons are transcribed in the nucleus. If plasmids and viruses in the cytoplasm are the immediate threat to the host genome, closed mitosis may prevent these elements from entering the nucleus, whereas if many transposons are transcribed, releasing the transcripts to the cytoplasm by open mitosis may reduce transposon insertion.

B chromosomes are extraneous, dispensable chromosomes that are not homologous to any of the canonical sets of "A" chromosomes. B chromosomes drive by biased segregation toward the germline stem cell (Figure 3.3), or biased segregation toward the germ cells (Figure 1.1, drive by biased segregation) (Jones, 1991). In both animals (e.g., grasshopper *Calliptamus palaestinensis*) and plants (e.g., sunflower *Crepis capillaris*), most somatic cells have a single copy of B chromosome, but the germline stem cells predominantly have two copies of B chromosomes, due to directional nondisjunction to the stem cells (Figure 3.3). Biased segregation of B chromosomes to the egg in female meiosis is observed in animals and plants (Hewitt, 1976; Jones, 1991). Although little is known about the mechanisms of B chromosome drive, repetitive DNA on B chromosomes likely biases the segregation. B chromosomes are devoid of coding genes and

mostly composed of tandem repeats such as satellite DNA and ribosomal DNA (Camacho et al., 2000). Biased segregation of rDNA in *Drosophila* male germline stem cells is reported (Watase et al., 2018), raising the possibility that ribosomal DNA is also selfish (Burt and Trivers, 2006).

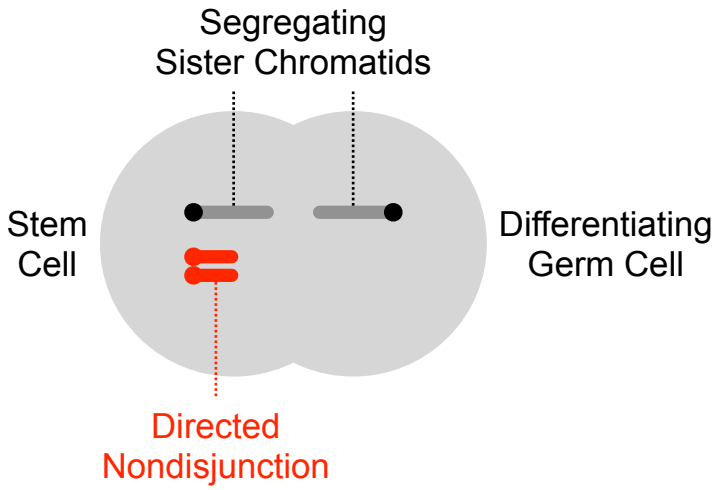


Figure 3.4. Biased segregation toward the germline stem cells.
A single copy of B chromosomes usually results in a 50% chance of inheritance. Directional nondisjunction toward germline stem cells leads to accumulation of germline stem cells with two copies of B chromosomes. Consequently, more germ cells with B chromosomes are produced.

3.7. Selfish Genetic Elements and Speciation

Genetic conflict between selfish genetic elements and centromere-binding proteins potentially explains the complexity of eukaryotic centromeres. Opportunities for selfish genetic elements to exploit the chromosome segregation machinery are not limited to female meiosis, as selfish transposons, plasmids and B chromosomes benefit by maximizing their inheritance (See Section 3.5 and 3.6). These opportunities are limited by the strong epigenetic component of most eukaryotic centromeres, which are not defined by specific DNA sequences. Centromeres cannot be completely independent of the underlying DNA sequence, however, because some proteins must interact with DNA, so different sequences can have different binding affinities or impact the structure of the centromeric nucleosome complex (Allu et al., 2019). The presence of multiple pathways to form a kinetochore (e.g., via CENP-ACLN and CENP-TWSX connected by CENP-HIKM, or via CENP-OPQUR) (Cortes-Silva et al., 2020; Hamilton et al., 2020; Nishino et al., 2012; Pesenti et al., 2018; Veld et al., 2016; Weir et al., 2016; Yan et al., 2019) allows proteins to adapt by minimizing a pathway that is exploited by a selfish element, while maintaining kinetochore function via other pathways. Consistent with this idea of independent modules for kinetochore formation, CENP-A depletion leads to a proportional reduction of centromeric CENP-C, whereas CENP-T and CENP-I persist longer (Fachinetti et al., 2013). In addition, recurrent changes in kinetochore modules are observed throughout eukaryotic evolution, such as changes in the number of MELT motifs in KNL1 and replacement of the SKA complex by the DAM complex (Hooff et al., 2017; Tromer et al., 2015). Regulation of kinetochore-microtubule attachment stability may be another way to suppress selfish genetic elements, as MELT motifs recruit BUB1 and SKA and DAM complexes stabilize attachments. Thus, internal conflicts between selfish genetic elements and chromosome segregation machinery may have shaped complexity in eukaryotic centromeres.

Selfish genetic elements that propagate at the expense of organismal fitness are universally found in all lineages of life. Why have only eukaryotes developed complex centromeres, even though transposons and plasmids are also present in bacteria and archaea? Although the inheritance of genetic information is an essential process for all life, it is proposed that prokaryotic chromosomes can spontaneously segregate by physical forces without sophisticated segregation machineries (Jun and Mulder, 2006; Jun and Wright, 2010). In *E. coli*, the cis-DNA element *migS* helps promote the bipolar segregation of origins, but this sequence is not essential for chromosome segregation (Fekete and Chattoraj, 2005; Wang and Sherratt, 2010; Yamaichi and Niki, 2004). Furthermore, although bacterial *parABS* segregation machinery helps stabilize inheritance of low-copy plasmids (Austin and Abeles, 1983; Ogura and Hiraga, 1983), *E. coli* does not encode the *parABS* system (Livny et al., 2007). In contrast, eukaryotes require chromosome segregation machinery for meiosis. Meiotic recombination is the predominant way to exchange genetic information in eukaryotes, and homologous chromosomes must pair and segregate each generation (Lenormand et al., 2016). This requirement makes chromosome segregation machinery indispensable for all eukaryotes. In contrast, bacteria and archaea can abandon the chromosome segregation machinery if it is exploited by selfish genetic elements that are harmful to the host.

Selfish genetic elements are constantly evolving, and if a new selfish genetic element achieves non-Mendelian inheritance, it will quickly spread in a population. Thus, just a brief period of population isolation might be sufficient to generate a population-specific selfish genetic element. Population-specific suppressor protein variants will then be selected. This rapid diversification of drive suppression mechanisms may cause hybrid incompatibility when the diverged populations hybridize, leading to reproductive isolation and speciation (Presgraves, 2010; Werren, 2011).

3.8. Future Directions

Effector recruitment was reduced when rat CENP-C was expressed in mouse oocytes, whereas *pumilio* CENP-C expression leads to increased effector recruitment (Figure 2.2). Although rat CENP-C and *pumilio* CENP-C are similarly divergent from mouse CENP-C, the effector recruitment phenotype was the opposite. As CENP-C sequence changes are found all over the protein (Figure 2.9), it is unclear which amino acid changes are responsible for effector recruitment change. Further investigation, such as domain swap experiments, will elucidate which functional domains are responsible for changes in effector recruitment.

As CENP-B deletion disrupts both the kinetochore and heterochromatin pathways for effector recruitment (Figure 2.6), the molecular detail of the CENP-B dependent centromere equalization pathway is still unclear. The negatively charged domain of CENP-B recruits many CENP-B interacting proteins including SUV39H1 and CENP-C (Otake et al., 2020), so deletion of this domain will also disrupt both pathways. Expression of chimeric proteins that have CENP-B DNA binding domain and a component of each pathway (e.g., SUV39H1 or CENP-C) is one possible way to separate two pathways. Yeast CENP-B homologs silence transposons (Cam et al., 2007; Upadhyay et al., 2017), but it is unclear whether mammalian CENP-B also silences transposons. Investigation of transposon insertion will provide insights into yet another function of mammalian CENP-B.

Although repeat expansion correlates with CENP-A chromatin expansion (Iwata-Otsubo et al., 2017), it is unclear which proteins are genetically recruited by selfish centromere DNA. When a divergent allele of a “hijacked” protein is expressed, such protein is predicted to be recruited less to mouse centromeres. Positively selected genes identified from the PAML analyses are candidate proteins that may be exploited by selfish centromere DNA. Expression of rat or *pumilio* proteins

in mouse cells will identify proteins that are recruited less to mouse centromeres and therefore likely exploited by selfish centromere DNA. In this dissertation, selfish centromere DNA is assumed to directly recruit centromere binding proteins, but it is also possible that selfish centromere DNA indirectly recruits centromere binding proteins through RNA transcripts. Centromeres are transcriptionally active, and functions of centromere transcripts are still unclear. Transcription inhibitor can be added to see if protein recruitment is RNA transcription dependent.

Having genetically different centromeres in the same cytoplasm is predicted to have fitness costs (Figure 1.10), and it is critical to test this prediction. Trivalents in intra-species hybrids and hybrid incompatibility genes in inter-species hybrids made it difficult to study fitness costs of having genetically different centromeres. After introgression of smaller centromeres or *spretus* centromeres in a larger *musculus* strain, chromosome segregation errors can be studied. If fitness costs are found, expression of a divergent allele of candidate suppressor proteins is another direction.

BIBLIOGRAPHY

Abe, Y., Sako, K., Takagaki, K., Hirayama, Y., Uchida, K.S.K., Herman, J.A., DeLuca, J.G., and Hirota, T. (2016). HP1-assisted Aurora B kinase activity prevents chromosome segregation errors. Developmental Cell *36*, 487–497.

Ainsztein, A.M., Kandels-Lewis, S.E., Mackay, A.M., and Earnshaw, W.C. (1998). INCENP centromere and spindle targeting: identification of essential conserved motifs and involvement of heterochromatin protein HP1. Journal of Cell Biology *143*, 1763–1774.

Akera, T., Chmátal, L., Trimm, E., Yang, K., Aonbangkhen, C., Chenoweth, D.M., Janke, C., Schultz, R.M., and Lampson, M.A. (2017). Spindle asymmetry drives non-Mendelian chromosome segregation. Science *358*, 668–672.

Akera, T., Trimm, E., and Lampson, M.A. (2019). Molecular strategies of meiotic cheating by selfish centromeres. Cell *178*, 1132–1144.

Akiyoshi, B., and Gull, K. (2014). Discovery of unconventional kinetochores in kinetoplastids. Cell *156*, 1247–1258.

Allshire, R.C., and Madhani, H.D. (2017). Ten principles of heterochromatin formation and function. Nature Reviews Molecular Cell Biology *19*, 229–244.

Allu, P.K., Dawicki-McKenna, J.M., Eeuwen, T.V., Slavin, M., Braitbard, M., Xu, C., Kalisman, N., Murakami, K., and Black, B.E. (2019). Structure of the human core centromeric nucleosome complex. Current Biology *29*, 2625–2639.

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., and Choo, K.H.A. (2004). Human centromere repositioning “in progress.” Proceedings of the National Academy of Sciences *101*, 6542–6547.

Arora, U.P., Charlebois, C., Lawal, R.A., and Dumont, B.L. (2021). Population and subspecies spanersity at mouse centromere satellites. BMC Genomics *22*, 279.

Austin, S., and Abeles, A. (1983). Partition of unit-copy miniplasmids to daughter cells. Journal of Molecular Biology *169*, 353–372.

Badrinarayanan, A., Le, T.B.K., and Laub, M.T. (2015). Bacterial chromosome organization and segregation. Annual Review of Cell and Developmental Biology *31*, 171–199.

Barillà, D. (2016). Driving apart and segregating genomes in Archaea. Trends in Microbiology *24*, 957–967.

Barnhart-Dailey, M.C., Trivedi, P., Stukenberg, P.T., and Foltz, D.R. (2017). HJURP interaction with the condensin II complex during G1 promotes CENP-A deposition. Molecular Biology of the Cell *28*, 54–64.

Bensasson, D., Zarowiecki, M., Burt, A., and Koufopanou, V. (2008). Rapid evolution of yeast centromeres in the absence of drive. Genetics *178*, 2161–2167.

Bint, S.M., Ogilvie, C.M., Flinter, F.A., Khalaf, Y., and Scriven, P.N. (2011). Meiotic segregation of Robertsonian translocations ascertained in cleavage-stage embryos — implications for preimplantation genetic diagnosis. Human Reproduction *26*, 1575–1584.

Britton-Davidian, J., Catalan, J., da Graça Ramalhinho, M., Ganem, G., Auffray, J.-C., Capela, R., Biscoito, M., Searle, J.B., and da Luz Mathias, M. (2000). Rapid chromosomal evolution in island mice. Nature *403*, 158–158.

Bureš, P., and Zedek, F. (2014). Holokinetic drive: centromere drive in chromosomes without centromeres. Evolution *68*, 2412–2420.

Burt, A., and Trivers, R. (2006). Genes in conflict: the biology of selfish genetic elements (Harvard University Press).

Cam, H.P., Noma, K., Ebina, H., Levin, H.L., and Grewal, S.I.S. (2007). Host genome surveillance for retrotransposons by transposon-derived proteins. Nature *451*, 431–436.

Camacho, J.P.M., Sharbel, T.F., and Beukeboom, L.W. (2000). B-chromosome evolution. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences *355*, 163–178.

Casola, C., Hucks, D., and Feschotte, C. (2007). Convergent domestication of pogo-like transposases into centromere-binding proteins in fission yeast and mammals. Molecular Biology and Evolution *25*, 29–41.

Cazaux, B., Catalan, J., Justy, F., Escudé, C., Desmarais, E., and Britton-Davidian, J. (2013). Evolution of the structure and composition of house mouse satellite DNA sequences in the subgenus *Mus* (Rodentia: Muridea): a cytogenomic approach. Chromosoma *122*, 209–220.

Chang, C.-H., Chavan, A., Palladino, J., Wei, X., Martins, N.M.C., Santinello, B., Chen, C.-C., Erceg, J., Beliveau, B.J., Wu, C.-T., et al. (2019). Islands of retroelements are major components of *Drosophila* centromeres. PLOS Biology *17*, e3000241.

Cheerambathur, D.K., Prevo, B., Chow, T.-L., Hattersley, N., Wang, S., Zhao, Z., Kim, T., Gerson-Gurwitz, A., Oegema, K., Green, R., et al. (2019). The kinetochore-microtubule coupling machinery is repurposed in sensory nervous system morphogenesis. Developmental Cell *48*, 864–872.

Chen, L., Qiu, Q., Jiang, Y., Wang, K., Lin, Z., Li, Z., Bibi, F., Yang, Y., Wang, J., Nie, W., et al. (2019). Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science *364*, eaav6202.

Chmátal, L., Gabriel, S.I., Mitsainas, G.P., Martínez-Vargas, J., Ventura, J., Searle, J.B., Schultz, R.M., and Lampson, M.A. (2014). Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. Current Biology *24*, 2295–2300.

Chmátal, L., Yang, K., Schultz, R.M., and Lampson, M.A. (2015). Spatial regulation of kinetochore microtubule attachments by destabilization at spindle poles in meiosis I. Current Biology *25*, 1835–1841.

Chueh, A.C., Northrop, E.L., Brettingham-Moore, K.H., Choo, K.H.A., and Wong, L.H. (2009). LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. PLOS Genetics *5*, e1000354.

Cortes-Silva, N., Ulmer, J., Kiuchi, T., Hsieh, E., Cornilleau, G., Ladid, I., Dingli, F., Loew, D., Katsuma, S., and Drinnenberg, I.A. (2020). CenH3-independent kinetochore assembly in Lepidoptera requires CCAN, including CENP-T. Current Biology *30*, 561–572.

Crowder, M.E., Strzelecka, M., Wilbur, J.D., Good, M.C., von Dassow, G., and Heald, R. (2015). A comparative analysis of spindle morphometrics across Metazoans. Current Biology *25*, 1542–1550.

Cuacos, M., H. Franklin, F.C., and Heckmann, S. (2015). Atypical centromeres in plants—what they can tell us. Frontiers in Plant Science *6*, 913.

Dambacher, S., Deng, W., Hahn, M., Sadic, D., Fröhlich, J., Nuber, A., Hoischen, C., Diekmann, S., Leonhardt, H., and Schotta, G. (2012). CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. Nucleus *3*, 101–110.

Daniel, A. (2002). Distortion of female meiotic segregation and reduced male fertility in human Robertsonian translocations: Consistent with the centromere model of co-evolving centromere DNA/centromeric histone (CENP-A). American Journal of Medical Genetics *111*, 450–452.

Dawe, R.K., and Hiatt, E.N. (2004). Plant neocentromeres: fast, focused, and driven. Chromosome Research *12*, 655–669.

Dawe, R.K., Lowry, E.G., Gent, J.I., Stitzer, M.C., Swentowsky, K.W., Higgins, D.M., Ross-Ibarra, J., Wallace, J.G., Kanizay, L.B., Alabady, M., et al. (2018). A kinesin-14 motor activates neocentromeres to promote meiotic drive in maize. Cell *173*, 839-850.

de Villena, F.P.-M., and Sapienza, C. (2001). Female meiosis drives karyotypic evolution in mammals. Genetics *159*, 1179–1189.

Dover, G.A. (1986). Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends in Genetics *2*, 159–165.

Drinnenberg, I.A., deYoung, D., Henikoff, S., and Malik, H.S. (2014). Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. eLife *3*, e03676.

Drpic, D., Almeida, A.C., Aguiar, P., Renda, F., Damas, J., Lewin, H.A., Larkin, D.M., Khodjakov, A., and Maiato, H. (2018). Chromosome segregation is biased by kinetochore size. Current Biology *28*, 1344-1356.

Dumont, J., Oegema, K., and Desai, A. (2010). A kinetochore-independent mechanism drives anaphase chromosome separation during acentrosomal meiosis. Nature Cell Biology *12*, 894–901.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485–497.

Echave, J., Spielman, S.J., and Wilke, C.O. (2016). Causes of evolutionary rate variation among protein sites. Nature Reviews Genetics *17*, 109–121.

ENCODE Project Consortium, Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A., Kaul, R., et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature *583*, 699–710.

Fachinetti, D., Folco, H.D., Nechemia-Arbely, Y., Valente, L.P., Nguyen, K., Wong, A.J., Zhu, Q., Holland, A.J., Desai, A., Jansen, L.E.T., et al. (2013). A two-step mechanism for epigenetic specification of centromere identity and function. Nature Cell Biology *15*, 1056–1066.

Fachinetti, D., Han, J.S., McMahon, M.A., Ly, P., Abdullah, A., Wong, A.J., and Cleveland, D.W. (2015). DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. Developmental Cell *33*, 314–327.

Fekete, R.A., and Chattoraj, D.K. (2004). A cis-acting sequence involved in chromosome segregation in Escherichia coli. Molecular Microbiology *55*, 175–183.

Finseth, F.R., Dong, Y., Saunders, A., and Fishman, L. (2015). Duplication and adaptive evolution of a key centromeric protein in Mimulus, a genus with female meiotic drive. Molecular Biology and Evolution *32*, 2694–2706.

Finseth, F.R., Nelson, T.C., and Fishman, L. (2021). Selfish chromosomal drive shapes recent centromeric histone evolution in monkeyflowers. PLOS Genetics *17*, e1009418.

Fishman, L., and Kelly, J.K. (2015). Centromere-associated meiotic drive and female fitness variation in Mimulus. Evolution *69*, 1208–1218.

Fishman, L., and Saunders, A. (2008). Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. Science *322*, 1559–1562.

Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates, J.R., Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. Cell *137*, 472–484.

Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., and Yanagida, M. (2007). Priming of centromere for CENP-A recruitment by human hMis18α, hMis18β, and M18BP1. Developmemtal Cell *12*, 17–30.

Gamba, R., and Fachinetti, D. (2020). From evolution to function: Two sides of the same CENP-B coin? Experimental Cell Research *390*, 111959.

Gao, B., Wang, Y., Diaby, M., Zong, W., Shen, D., Wang, S., Chen, C., Wang, X., and Song, C. (2020). Evolution of pogo, a separate superfamily of IS630-Tc1-mariner transposons, revealing recurrent domestication events in vertebrates. Mobile DNA *11*, 25.

Garagna, S., Page, J., Fernandez-Donoso, R., Zuccotti, M., and Searle, J.B. (2014). The Robertsonian phenomenon in the house mouse: mutation, meiosis and speciation. Chromosoma *123*, 529–544.

Garrido-Ramos, M. (2017). Satellite DNA: an evolving topic. Genes *8*, 230.

Gibbs, R.A., Weinstock, G.M., Metzker, M.L., Muzny, D.M., Sodergren, E.J., Scherer, S., Scott, G., Steffen, D., Worley, K.C., Burch, P.E., et al. (2004). Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature *428*, 493–521.

Gillespie, J.H. (2004). Population Genetics: A Concise Guide (Johns Hopkins University Press).

Goday, C., and Pimpinelli, S. (1989). Centromere organization in meiotic chromosomes of *Parascaris univalens*. Chromosoma *98*, 160–166.

Goday, C., González-García, J.M., Esteban, M.R., Giovinazzo, G., and Pimpinelli, S. (1992). Kinetochores and chromatin diminution in early embryos of *Parascaris univalens*. Journal of Cell Biology *118*, 23–32.

Godek, K.M., Kabeche, L., and Compton, D.A. (2014). Regulation of kinetochore-microtubule attachments through homeostatic control during mitosis. Nature Reviews Molecular Cell Biology *16*, 57–64.

Gorelick, R., Carpinone, J., and Derraugh, L.J. (2016). No universal differences between female and male eukaryotes: anisogamy and asymmetrical female meiosis. Biological Journal of the Linnean Society *120*, 1–21.

Gratacòs, M., Nadal, M., Martín-Santos, R., Pujana, M.A., Gago, J., Peral, B., Armengol, L., Ponsa, I., Miró, R., Bulbena, A., et al. (2001). A polymorphic genomic duplication on human chromosome 15 is a susceptibility factor for panic and phobic disorders. Cell *106*, 367–379.

Gutbrod, M.J., and Martienssen, R.A. (2020). Conserved chromosomal functions of RNA interference. Nature Reviews Genetics *21*, 311–331.

Hamerton, J.L., Canning, N., Ray, M., and Smith, S. (2008). A cytogenetic survey of 14,069 newborn infants. Clinical Genetics *8*, 223–243.

Hamilton, G.E., Helgeson, L.A., Noland, C.L., Asbury, C.L., Dimitrova, Y.N., and Davis, T.N. (2020). Reconstitution reveals two paths of force transmission through the kinetochore. eLife *9*, e56582.

Hasson, D., Alonso, A., Cheung, F., Tepperberg, J.H., Papenhausen, P.R., Engelen, J.J.M., and Warburton, P.E. (2011). Formation of novel CENP-A domains on tandem repetitive DNA and across chromosome breakpoints on human chromosome 8q21 neocentromeres. Chromosoma *120*, 621–632.

Hemmerich, P., Weidtkamp-Peters, S., Hoischen, C., Schmiedeberg, L., Erliandri, I., and Diekmann, S. (2008). Dynamics of inner kinetochore assembly and maintenance in living cells. Journal of Cell Biology *180*, 1101–1114.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098–1102.

Hewitt, G.M. (1976). Meiotic drive for B-chromosomes in the primary oocytes of Myrmekotettix maculatus (Orthoptera: Acrididae). Chromosoma *56*, 381–391.

Higgins, J.M.G., and Prendergast, L. (2016). Mitotic mysteries: the case of HP1. Developmental Cell *36*, 477–478.

Hooff, J.J. van, Snel, B., and Kops, G.J. (2017). Unique phylogenetic distributions of the Ska and Dam1 complexes support functional analogy and suggest multiple parallel displacements of Ska by Dam1. Genome Biology and Evolution *9*, 1295–1303.

Hudson, D.F., Fowler, K.J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N.G., Kretser, D.M. de, Cancilla, M.R., et al. (1998). Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. Journal of Cell Biology *141*, 309–319.

Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755.

Hughes-Schrader, S., and Schrader, F. (1961). The kinetochore of the hemiptera. Chromosoma *12*, 327–350.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Ishii, K., Ogiyama, Y., Chikashige, Y., Soejima, S., Masuda, F., Kakuma, T., Hiraoka, Y., and Takahashi, K. (2008). Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science *321*, 1088–1091.

Iwata-Otsubo, A., Dawicki-McKenna, J.M., Akera, T., Falk, S.J., Chmátal, L., Yang, K., Sullivan, B.A., Schultz, R.M., Lampson, M.A., and Black, B.E. (2017). Expanded satellite repeats amplify a discrete CENP-A nucleosome assembly site on chromosomes that drive in female meiosis. Current Biology *27*, 2365–2373.

Jangam, D., Feschotte, C., and Betrán, E. (2017). Transposable element domestication as an adaptation to evolutionary conflicts. Trends in Genetics *33*, 817–831.

Janssen, A., Colmenares, S.U., and Karpen, G.H. (2018). Heterochromatin: guardian of the genome. Annual Review of Cell and Developmental Biology *34*, 265–288.

Javed, A.O., Li, Y., Muffat, J., Su, K.-C., Cohen, M.A., Lungjangwa, T., Aubourg, P., Cheeseman, I.M., and Jaenisch, R. (2018). Microcephaly modeling of kinetochore mutation reveals a brain-specific phenotype. Cell Reports *25*, 368–382.

Jones, R.N. (1991). B-chromosome drive. The American Naturalist *137*, 430–442.

Jun, S., and Mulder, B. (2006). Entropy-driven spatial organization of highly confined polymers: Lessons for the bacterial chromosome. Proceedings of the National Academy of Sciences *103*, 12388–12393.

Jun, S., and Wright, A. (2010). Entropy as the driver of chromosome segregation. Nature Reviews Microbiology *8*, 600–607.

Kang, J., Chaudhary, J., Dong, H., Kim, S., Brautigam, C.A., and Yu, H. (2011). Mitotic centromeric targeting of HP1 and its binding to Sgo1 are dispensable for sister-chromatid cohesion in human cells. Molecular Biology of the Cell *22*, 1181–1190.

Kang, Y., Wang, J., Neff, A., Kratzer, S., Kimura, H., and Davis, R.E. (2016). Differential chromosomal localization of centromeric histone CENP-A contributes to nematode programmed DNA elimination. Cell Reports *16*, 2308–2316.

Kapoor, M., Luna, R.M. de O., Liu, G., Lozano, G., Cummings, C., Mancini, M., Ouspenski, I., Brinkley, B.R., and May, G.S. (1998). The *cenpB* gene is not essential in mice. Chromosoma *107*, 570–576.

Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science *340*, 1110–1113.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution *30*, 772–780.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research *30*, 3059–3066.

Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. Nature *477*, 289–294.

Kelly, A.E., Ghenoiu, C., Xue, J.Z., Zierhut, C., Kimura, H., and Funabiki, H. (2010). Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. Science *330*, 235–239.

Kim, J., Ishiguro, K., Nambu, A., Akiyoshi, B., Yokobayashi, S., Kagami, A., Ishiguro, T., Pendas, A.M., Takeda, N., Sakakibara, Y., et al. (2015). Meikin is a conserved regulator of meiosis-I-specific kinetochore function. Nature *517*, 466–471.

Kipling, D., and Warburton, P.E. (1997). Centromeres, CENP-B and Tigger too. Trends in Genetics *13*, 141–145.

Klare, K., Weir, J.R., Basilico, F., Zimniak, T., Massimiliano, L., Ludwigs, N., Herzog, F., and Musacchio, A. (2015). CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. Journal of Cell Biology *210*, 11–22.

Kobayashi, N., Suzuki, Y., Schoenfeld, L.W., Müller, C.A., Nieduszynski, C., Wolfe, K.H., and Tanaka, T.U. (2015). Discovery of an unconventional centromere in budding yeast redefines evolution of point centromeres. Current Biology *25*, 2026–2033.

Krenn, V., Overlack, K., Primorac, I., Gerwen, S. van, and Musacchio, A. (2013). KI motifs of human Knl1 enhance assembly of comprehensive spindle checkpoint complexes around MELT repeats. Current Biology *24*, 29–39.

Kumon, T., Ma, J., Stefanik, D., Nordgren, E.C., Akins, R.B., Kim, J., Levine, M.T., and Lampson, M.A. (2020). Centromere drive and suppression by parallel pathways for recruiting microtubule destabilizers (Cold Spring Harbor Laboratory).

Lampson, M.A., and Black, B.E. (2017). Cellular and molecular mechanisms of centromere drive. Cold Spring Harbor Symposia on Quantitative Biology *82*, 249–257.

Langley, S.A., Miga, K.H., Karpen, G.H., and Langley, C.H. (2019). Haplotypes spanning centromeric regions reveal persistence of large blocks of archaic DNA. eLife *8*, e42989.

Lee, Y.C.G., Leek, C., and Levine, M.T. (2016). Recurrent innovation at genes required for telomere integrity in *Drosophila*. Molecular Biology and Evolution *34*, 467–482.

LeMaire-Adkins, R., and Hunt, P.A. (2000). Nonrandom segregation of the mouse univalent X chromosome: evidence of spindle-mediated meiotic drive. Genetics *156*, 775–783.

Lenormand, T., Engelstädter, J., Johnston, S.E., Wijnker, E., and Haag, C.R. (2016). Evolutionary mysteries in meiosis. Philosophical Transactions of the Royal Society B: Biological Sciences *371*, 20160001.

Livny, J., Yamaichi, Y., and Waldor, M.K. (2007). Distribution of centromere-like *parS* sites in bacteria: insights from comparative genomics. Journal of Bacteriology *189*, 8693–8703.

Locke, D.P., Hillier, L.W., Warren, W.C., Worley, K.C., Nazareth, L.V., Muzny, D.M., Yang, S.-P., Wang, Z., Chinwalla, A.T., Minx, P., et al. (2011). Comparative and demographic analysis of orang-utan genomes. Nature *469*, 529–533.

Logsdon, G.A., Gambogi, C.W., Liskovykh, M.A., Barrey, E.J., Larionov, V., Miga, K.H., Heun, P., and Black, B.E. (2019). Human artificial chromosomes that bypass centromeric DNA. Cell *178*, 624–639.

Luchetti, A., Cesari, M., Carrara, G., Cavicchi, S., Passamonti, M., Scali, V., and Mantovani, B. (2003). Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus* taxa (Insecta Phasmatodea). Journal of Molecular Evolution *56*, 587–596.

Maheshwari, S., Tan, E.H., West, A., Franklin, F.C.H., Comai, L., and Chan, S.W.L. (2015). Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. PLOS Genetics *11*, e1004970.

Malik, H.S., and Henikoff, S. (2001). Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. Genetics *157*, 1293–1298.

Malik, H.S., and Henikoff, S. (2009). Major evolutionary transitions in centromere complexity. Cell *138*, 1067–1082.

Mallarino, R., Pillay, N., Hoekstra, H.E., and Schradin, C. (2018). African striped mice. Current Biology *28*, R299–R301.

Mandrioli, M., and Manicardi, G.C. (2020). Holocentric chromosomes. PLOS Genetics *16*, e1008918.

Mantovani, B., Tinti, F., Bachmann, L., and Scali, V. (1997). The Bag320 satellite DNA family in Bacillus stick insects (Phasmatodea): different rates of molecular evolution of highly repetitive DNA in bisexual and parthenogenic taxa. Molecular Biology and Evolution *14*, 1197–1205.

Marston, A.L. (2015). Shugoshins: tension-sensitive pericentromeric adaptors safeguarding chromosome segregation. Molecular and Cellular Biology *35*, 634–648.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., and Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. Journal of Cell Biology *109*, 1963–1973.

Mateo, L., and González, J. (2014). Pogo-like transposases have been repeatedly domesticated into CENP-B-related proteins. Genome Biology and Evolution *6*, 2008–2016.

Melters, D.P., Bradnam, K.R., Young, H.A., Telis, N., May, M.R., Ruby, J.G., Sebra, R., Peluso, P., Eid, J., Rank, D., et al. (2013). Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. Genome Biology *14*, R10.

Molina, W.F., Martinez, P.A., Bertollo, L.A.C., and Bidau, C.J. (2014). Evidence for meiotic drive as an explanation for karyotype changes in fishes. Marine Genomics *15*, 29–34.

Monen, J., Maddox, P.S., Hyndman, F., Oegema, K., and Desai, A. (2005). Differential role of CENP-A in the segregation of holocentric C. elegans chromosomes during meiosis and mitosis. Nature Cell Biology *7*, 1248–1255.

Moree, B., Meyer, C.B., Fuller, C.J., and Straight, A.F. (2011). CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. Journal of Cell Biology *194*, 855–871.

Morozov, V.M., Giovinazzi, S., and Ishov, A.M. (2017). CENP-B protects centromere chromatin integrity by facilitating histone deposition via the H3.3-specific chaperone Daxx. Epigenetics Chromatin *10*, 63.

Musacchio, A., and Desai, A. (2017). A molecular view of kinetochore assembly and function. Biology *6*, 5.

Nakagawa, H., Lee, J.-K., Hurwitz, J., Allshire, R.C., Nakayama, J., Grewal, S.I.S., Tanaka, K., and Murakami, Y. (2002). Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. Genes Development 16, 1766–1778.

Navarro-Mendoza, M.I., Pérez-Arques, C., Panchal, S., Nicolás, F.E., Mondo, S.J., Ganguly, P., Pangilinan, J., Grigoriev, I.V., Heitman, J., Sanyal, K., et al. (2019). Early diverging fungus *Mucor circinelloides* lacks centromeric histone CENP-A and displays a mosaic of point and regional centromeres. Current Biology *29*, 3791-3802.

Nergadze, S.G., Piras, F.M., Gamba, R., Corbo, M., Cerutti, F., McCarter, J.G.W., Cappelletti, E., Gozzo, F., Harman, R.M., Antczak, D.F., et al. (2018). Birth, evolution, and transmission of satellite-free mammalian centromeric domains. Genome Research *28*, 789–799.

Nishino, T., Takeuchi, K., Gascoigne, K.E., Suzuki, A., Hori, T., Oyama, T., Morikawa, K., Cheeseman, I.M., and Fukagawa, T. (2012). CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. Cell *148*, 487–501.

Nozawa, R.-S., Nagao, K., Masuda, H.-T., Iwasaki, O., Hirota, T., Nozaki, N., Kimura, H., and Obuse, C. (2010). Human POGZ modulates dissociation of HP1alpha from mitotic chromosome arms through Aurora B activation. Nature Cell Biology *12*, 719–727.

Ogura, T., and Hiraga, S. (1983). Partition mechanism of F plasmid: Two plasmid gene-encoded products and a cis-acting region are involved in partition. Cell *32*, 351–360.

Oh, E., Mark, K.G., Mocciaro, A., Watson, E.R., Prabu, J.R., Cha, D.D., Kampmann, M., Gamarra, N., Zhou, C.Y., and Rape, M. (2020). Gene expression and cell identity controlled by anaphase-promoting complex. Nature *579*, 136–140.

Ohzeki, J.-I., Shono, N., Otake, K., Martins, N.M.C., Kugou, K., Kimura, H., Nagase, T., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2016). KAT7/HBO1/MYST2 regulates CENP-A chromatin assembly by antagonizing Suv39h1-mediated centromere inactivation. Developmental Cell *37*, 413–427.

Okada, T., Ohzeki, J., Nakano, M., Yoda, K., Brinkley, W.R., Larionov, V., and Masumoto, H. (2007). CENP-B controls centromere formation depending on the chromatin context. Cell *131*, 1287–1300.

Otake, K., Ohzeki, J.-I., Shono, N., Kugou, K., Okazaki, K., Nagase, T., Yamakawa, H., Kouprina, N., Larionov, V., Kimura, H., et al. (2020). CENP-B creates alternative epigenetic chromatin states permissive for CENP-A or heterochromatin assembly. Journal of Cell Science *133*, jcs243303.

Pacchierotti, F., Tiveron, C., Mailhes, J.B., and Davisson, M.T. (1995). Susceptibility to vinblastine-induced aneuploidy and preferential chromosome segregation during meiosis I in Robertsonian heterozygous mice. Teratogenesis Carcinogenesis Mutagenesis *15*, 217–230.

Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., and Luger, K. (2017). Decoding the centromeric nucleosome through CENP-N. eLife *6*, e33442.

Perelman, P., Johnson, W.E., Roos, C., Seuánez, H.N., Horvath, J.E., Moreira, M.A.M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., et al. (2011). A molecular phylogeny of living primates. PLOS Genetics *7*, e1001342.

Perez-Castro, A.V., Shamanski, F.L., Meneses, J.J., Lovato, T.L., Vogel, K.G., Moyzis, R.K., and Pedersen, R. (1998). Centromeric protein B null mice are viable with no apparent abnormalities. Developmental Biology *201*, 135–143.

Pérez, R., Rufas, J.S., Suja, J.A., Page, J., and Panzera, F. (2000). Meiosis in holocentric chromosomes: orientation and segregation of an autosome and sex chromosomes in *Triatoma infestans* (Heteroptera). Chromosome Research *8*, 17–25.

Peris, L., Wagenbach, M., Lafanechère, L., Brocard, J., Moore, A.T., Kozielski, F., Job, D., Wordeman, L., and Andrieux, A. (2009). Motor-dependent microtubule disassembly driven by tubulin tyrosination. Journal of Cell Biology *185*, 1159–1166.

Pesenti, M.E., Prumbaum, D., Auckland, P., Smith, C.M., Faesen, A.C., Petrovic, A., Erent, M., Maffini, S., Pentakota, S., Weir, J.R., et al. (2018). Reconstitution of a 26-subunit human kinetochore reveals cooperative microtubule binding by CENP-OPQUR and NDC80. Molecular Cell *71*, 923–939.

Petrovic, A., Keller, J., Liu, Y., Overlack, K., John, J., Dimitrova, Y.N., Jenni, S., Gerwen, S. van, Stege, P., Wohlgemuth, S., et al. (2016). Structure of the MIS12 complex and molecular basis of its interaction with CENP-C at human kinetochores. Cell *167*, 1028–1040.

Petrovic, A., Mosalaganti, S., Keller, J., Mattiuzzo, M., Overlack, K., Krenn, V., Antoni, A.D., Wohlgemuth, S., Cecatiello, V., Pasqualato, S., et al. (2014). Modular assembly of RWD domains on the Mis12 complex underlies outer kinetochore organization. Molecular Cell *53*, 591–605.

PIálek, J., Hauffe, H.C., and Searle, J.B. (2005). Chromosomal variation in the house mouse. Biological Journal of the Linnean Society *84*, 535–563.

Plasterk, R.H.A., Izsvák, Z., and Ivics, Z. (1999). Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends in Genetics *15*, 326–332.

Presgraves, D.C. (2010). The molecular evolutionary basis of species formation. Nature Reviews Genetics *11*, 175–180.

Rago, F., Gascoigne, K.E., and Cheeseman, I.M. (2015). Distinct organization and regulation of the outer kinetochore KMN network downstream of CENP-C and CENP-T. Current Biology *25*, 671–677.

Rattani, A., Wolna, M., Ploquin, M., Helmhart, W., Morrone, S., Mayer, B., Godwin, J., Xu, W., Stemmann, O., Pendas, A., et al. (2013). Sgol2 provides a regulatory platform that coordinates essential cell cycle processes during meiosis I in oocytes. eLife *2*, e01133.

Rhind, N., Chen, Z., Yassour, M., Thompson, D.A., Haas, B.J., Habib, N., Wapinski, I., Roy, S., Lin, M.F., Heiman, D.I., et al. (2011). Comparative functional genomics of the fission yeasts. Science *332*, 930–936.

Rhoades, M.M. (1942). Preferential segregation in maize. Genetics *27*, 395–407.

Rizvi, S.M.A., Prajapati, H.K., and Ghosh, S.K. (2017). The 2 micron plasmid: a selfish genetic element with an optimized survival strategy within Saccharomyces cerevisiae. Current Genetics *64*, 25–42.

Rocchi, M., Archidiacono, N., Schempp, W., Capozzi, O., and Stanyon, R. (2011). Centromere repositioning in mammals. Heredity *108*, 59–67.

Saint-Leandre, B., and Levine, M.T. (2020). The telomere paradox: stable genome preservation with rapidly evolving proteins. Trends in Genetics *36*, 232–242.

Sazer, S., Lynch, M., and Needleman, D. (2014). Deciphering the evolutionary history of open and closed mitosis. Current Biology *24*, R1099–R1103.

Schueler, M.G., Swanson, W., Thomas, P.J., Program, N.C.S., and Green, E.D. (2010). Adaptive evolution of foundation kinetochore proteins in primates. Molecular Biology and Evolution *27*, 1585–1597.

Shi, L., Hu, E., Wang, Z., Liu, J., Li, J., Li, M., Chen, H., Yu, C., Jiang, T., and Su, B. (2016). Regional selection of the brain size regulating gene CASC5 provides new insight into human brain evolution. Human Genetics *136*, 193–204.

Shono, N., Ohzeki, J., Otake, K., Martins, N.M.C., Nagase, T., Kimura, H., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2015). CENP-C and CENP-I are key connecting factors for kinetochore and CENP-A assembly. Journal of Cell Science *128*, 4572–4587.

Sirajuddin, M., Rice, L.M., and Vale, R.D. (2014). Regulation of microtubule motors by tubulin isotypes and post-translational modifications. Nature Cell Biology *16*, 335–344.

Sironi, M., Cagliani, R., Forni, D., and Clerici, M. (2015). Evolutionary insights into host-pathogen interactions from mammalian sequence data. Nature Reviews Genetics *16*, 224–236.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Sugimoto, K., Kuriyama, K., Shibata, A., and Himeno, M. (1997). Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C in vitro: role of DNA-binding and self-associating activities in kinetochore organization. Chromosome Research *5*, 132–141.

Suzuki, A., Badger, B.L., Wan, X., DeLuca, J.G., and Salmon, E.D. (2014). The architecture of CCAN proteins creates a structural integrity to resist spindle forces and achieve proper intrakinetochore stretch. Developmental Cell *30*, 717–730.

Takeiri, A., Motoyama, S., Matsuzaki, K., Harada, A., Taketo, J., Katoh, C., Tanaka, K., and Mishima, M. (2013). New DNA probes to detect aneugenicity in rat bone marrow micronucleated cells by a pan-centromeric FISH analysis. Mutation Research *755*, 73–80.

Talbert, P.B., Bryson, T.D., and Henikoff, S. (2004). Adaptive evolution of centromere proteins in plants and animals. Journal of Biology *3*, 18.

Thybert, D., Roller, M., Navarro, F.C.P., Fiddes, I., Streeter, I., Feig, C., Martin-Galvez, D., Kolmogorov, M., Janoušek, V., Akanni, W., et al. (2018). Repeat associated mechanisms of genome evolution and function revealed by the *Mus caroli* and *Mus pahari* genomes. Genome Research *28*, 448–459.

Tromer, E., Snel, B., and Kops, G.J.P.L. (2015). Widespread recurrent patterns of rapid repeat evolution in the kinetochore scaffold KNL1. Genome Biology and Evolution *7*, 2383–2393.

Tsukahara, T., Tanno, Y., and Watanabe, Y. (2010). Phosphorylation of the CPC by Cdk1 promotes chromosome bi-orientation. Nature *467*, 719–723.

Upadhyay, U., Srivastava, S., Khatri, I., Nanda, J.S., Subramanian, S., Arora, A., and Singh, J. (2017). Ablation of RNA interference and retrotransposons accompany acquisition and evolution of transposases to heterochromatin protein CENPB. Molecular Biology of the Cell *28*, 1132–1146.

van der Lee, R., Wiel, L., van Dam, T.J.P., and Huynen, M.A. (2017). Genome-scale detection of positive selection in nine primates predicts human-virus evolutionary conflicts. Nucleic Acids Research *45*, 10634–10648.

Veld, P.J.H.I. 't, Jeganathan, S., Petrovic, A., Singh, P., John, J., Krenn, V., Weissmann, F., Bange, T., and Musacchio, A. (2016). Molecular basis of outer kinetochore assembly on CENP-T. eLife *5*, e21007.

Veld, P.J.H.I. 't, Volkov, V.A., Stender, I.D., Musacchio, A., and Dogterom, M. (2019). Molecular determinants of the Ska-Ndc80 interaction and their influence on microtubule tracking and force-coupling. eLife *8*, e49539.

Ventura, M. (2003). Neocentromeres in 15q24-26 map to duplicons which flanked an ancestral centromere in 15q25. Genome Research *13*, 2059–2068.

Ventura, M. (2004). Recurrent sites for new centromere seeding. Genome Research *14*, 1696–1703.

Villasante, A., Abad, J.P., and Mendez-Lago, M. (2007). Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome. Proceedings of the National Academy of Sciences *104*, 10542–10547.

Wan, X., O'Quinn, R.P., Pierce, H.L., Joglekar, A.P., Gall, W.E., DeLuca, J.G., Carroll, C.W., Liu, S.-T., Yen, T.J., McEwen, B.F., et al. (2009). Protein architecture of the human kinetochore microtubule attachment site. Cell *137*, 672–684.

Wang, F., Dai, J., Daum, J.R., Niedzialkowska, E., Banerjee, B., Stukenberg, P.T., Gorbsky, G.J., and Higgins, J.M.G. (2010). Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. Science 330, 231–235.

Wang, X., and Sherratt, D.J. (2010). Independent segregation of the two arms of the *Escherichia coli* ori region requires neither RNA synthesis nor MreB dynamics. Journal of Bacteriology *192*, 6143–6153.

Watase, G., and Yamashita, Y. (2018). Ribosomal DNA and the rDNA-binding protein Indra mediate non-random sister chromatid segregation in *Drosophila* male germline stem cells (Cold Spring Harbor Laboratory).

Weir, J.R., Faesen, A.C., Klare, K., Petrovic, A., Basilico, F., Fischböck, J., Pentakota, S., Keller, J., Pesenti, M.E., Pan, D., et al. (2016). Insights from biochemical reconstitution into the architecture of human kinetochores. Nature *537*, 249–253.

Werren, J.H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. Proceedings of the National Academy of Sciences *108*, 10863–10870.

Yadav, V., Sun, S., Billmyre, R.B., Thimmappa, B.C., Shea, T., Lintner, R., Bakkeren, G., Cuomo, C.A., Heitman, J., and Sanyal, K. (2018). RNAi is a critical determinant of centromere evolution in closely related fungi. Proceedings of the National Academy of Sciences *115*, 3108–3113.

Yamagishi, Y., Honda, T., Tanno, Y., and Watanabe, Y. (2010). Two histone marks establish the inner centromere and chromosome bi-orientation. Science *330*, 239–243.

Yamagishi, Y., Sakuno, T., Shimura, M., and Watanabe, Y. (2008). Heterochromatin links to centromeric protection by recruiting shugoshin. Nature *455*, 251–255.

Yamaichi, Y., and Niki, H. (2003). migS, a cis-acting site that affects bipolar positioning of oriC on the *Escherichia coli* chromosome. The EMBO Journal *23*, 221–233.

Yan, K., Yang, J., Zhang, Z., McLaughlin, S.H., Chang, L., Fasci, D., Ehrenhofer-Murray, A.E., Heck, A.J.R., and Barford, D. (2019). Structure of the inner kinetochore CCAN complex assembled onto a centromeric nucleosome. Nature *574*, 278–282.

Yang, P., Wang, Y., and Macfarlan, T.S. (2017). The role of KRAB-ZFPs in transposable element repression and mammalian evolution. Trends in Genetics *33*, 871–881.

Yang, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. Molecular Biology and Evolution *24*, 1586–1591.

Yao, Y., and Dai, W. (2012). Shugoshins function as a guardian for chromosomal stability in nuclear division. Cell Cycle *11*, 2631–2642.

Zhao, G., Oztan, A., Ye, Y., and Schwarz, T.L. (2019). Kinetochore proteins have a post-mitotic function in neurodevelopment. Developmental Cell *48*, 873–882.