Title

Centromere drive and suppression by parallel pathways for recruiting microtubule destabilizers

Authors

Tomohiro Kumon[1], Jun Ma[1], Derek Stefanik[1], Erik C. Nordgren[1], R. Brian Akins[1], Junhyong Kim[1], Mia T. Levine[1], Michael A. Lampson[1,2,*]

[1]Department of Biology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Lead Contact
*Correspondence: lampson@sas.upenn.edu

Summary

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 heterochromatin by CENP-B deletion amplifies functional differences between centromeres, whereas disrupting the kinetochore pathway with a divergent allele of CENP-C reduces 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.

Keywords
Meiotic Drive, Selfish Genetic Elements, Evolutionary Arms Race, Centromere, Kinetochore, Heterochromatin

Introduction

Centromere evolution is paradoxical in that both repetitive centromere DNA and centromere-binding proteins evolve rapidly despite the conserved requirement of centromeres for faithful chromosome segregation (Henikoff et al., 2001; Lampson and Black, 2017; Malik and Henikoff, 2001; Melters et al., 2013; Schueler et al., 2010). Centromere DNA repeat monomer sequence and repeat copy number diverge between even closely related species. Repeat copy number also varies within species, for example, between human individuals or between mouse strains (Iwata-Otsubo et al., 2017; Langley et al., 2019). To explain this rapid evolution, the centromere drive hypothesis proposes that selfish centromere DNA sequences (either monomer sequence variants or repeat number expansions) drive in female meiosis by increasing their transmission rate to the egg. Potential deleterious consequences of driving centromeres, such as meiotic segregation errors, would select for centromere-binding protein variants that suppress these fitness costs (Finseth et al., 2020; Fishman and Saunders, 2008; Henikoff et al., 2001). New selfish DNA variants subsequently arise to start another cycle of drive and suppression in a continual evolutionary arms race.

Our previous work leveraged natural variation in mouse centromere DNA to study the molecular mechanisms of centromere drive (Akera et al., 2017, 2019; Chmátal et al., 2014; Iwata-Otsubo et al., 2017). Selfish centromeres in these model systems recruit more effector proteins that destabilize interactions with spindle microtubules, allowing them to detach from microtubules that would otherwise direct them to the polar body. Microtubule detachment and reattachment reorients the selfish centromeres toward the egg side of the meiosis I spindle (Akera et al., 2019) (Figure 1A). This reorientation depends on BUB1 kinase at kinetochores, which phosphorylates pericentromeric histone H2A. Phosphorylated H2A recruits Shugoshin-2 (SGO2), which recruits microtubule destabilizing proteins such as MCAK (Akera et al., 2019) (Figure 1B, kinetochore pathway). In one intra-species *Mus musculus domesticus* hybrid, selfish centromeres with expanded minor satellite DNA repeats assemble more centromere chromatin containing the histone H3 variant CENP-A (Iwata-Otsubo et al., 2017). These expanded centromeres also form larger kinetochores with more BUB1 kinase, leading to more effectors (SGO2 and MCAK) (Akera et al., 2019). In this hybrid, the larger centromeres are from a standard laboratory strain (either CF-1 or C57BL/6J), which is crossed to a wild-derived strain (CHPO) with smaller centromeres. Thus, the centromeres of paired homologous chromosomes within a meiotic bivalent are both genetically and functionally different in the hybrid (Figure 1C). These findings show how selfish centromeres can drive by recruiting more effectors. 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 paired centromeres in meiosis and would therefore be suppressed by reducing these differences.

1 **The parallel pathway model for drive and suppression provides three testable predictions.**

2

3 Based on our finding that selfish centromeres drive by recruiting more effectors, we propose that

4 functional differences between centromeres can be suppressed by equalizing effector recruitment

5 via a second pathway. This equalization would render genetically different centromeres

6 functionally equivalent. This model incorporates previous findings that in addition to the

7 kinetochore pathway, which acts through BUB1 kinase, effectors are also recruited through a

8 heterochromatin pathway. Pericentromeric heterochromatin recruits the chromosome passenger

9 complex (CPC), which recruits SGO2 and MCAK (Figure 1B, heterochromatin pathway) (Abe et

10 al., 2016; Ainsztein et al., 1998; Tsukahara et al., 2010). In our CHPO hybrid model system

11 (Figure 1C), the kinetochore pathway is asymmetric: we observe higher levels of the kinetochore

12 proteins HEC1/NDC80 and CENP-C on larger vs smaller centromeres (Chmátal et al., 2014;

13 Iwata-Otsubo et al., 2017). In contrast, the heterochromatin pathway is symmetric: the

14 heterochromatin mark, H3K9me3, is equal on the two sides of each bivalent (Figure 1D and 1E)

15 (Iwata-Otsubo et al., 2017). These observations suggest that, in this system, selfish centromere

16 DNA exploits the kinetochore pathway to make genetically different centromeres also

17 functionally different, with larger centromeres recruiting more effectors. In contrast, the

18 heterochromatin pathway appears insensitive to selfish DNA, recruiting effectors equally. We

19 propose that the host evolves to suppress functional differences by centromere protein

20 innovations that minimize the contribution of the asymmetric kinetochore pathway to effector

21 recruitment, relative to the symmetric heterochromatin pathway.

22

23 Evolution of the kinetochore pathway is constrained by its indispensable role in mitotic and

24 meiotic chromosome segregation. Nevertheless, proteins may evolve to weaken the pathway by

25 reducing interactions between centromere-binding proteins and DNA or between proteins

26 leading to effector recruitment (Figure 1F). Similarly, evolution of heterochromatin proteins is

27 constrained by numerous vital heterochromatin-dependent cellular functions (Allshire and

28 Madhani, 2017). Inner centromere proteins (such as the CPC) that interact with heterochromatin

29 may evolve, however, to increase effector recruitment. Finally, overall effector levels are also

30 constrained because microtubule destabilizing activity is necessary to correct kinetochore-

31 microtubule attachment errors, but excessive destabilizing activity weakens attachments

32 necessary for anaphase segregation and activates the spindle assembly checkpoint (Godek et al.,

33 2014). According to our parallel pathway model, a new centromere DNA variant can exploit the

34 kinetochore pathway to recruit more effectors by strengthening interactions with any centromere-

35 binding protein that contacts the DNA: CENP-A, the CENP-A chromatin assembly machinery,

36 or other proteins that link centromere chromatin to the kinetochore (e.g., CENP-C or CENP-T).

37 To suppress functional differences between centromeres, the centromere protein network

38 recurrently evolves to minimize the kinetochore pathway relative to the heterochromatin

39 pathway while maintaining essential functions.

40

1   Here we test three predictions from the parallel pathway model. First, when the symmetric
2   heterochromatin pathway is weakened, we predict that the asymmetric kinetochore pathway
3   makes a relatively larger contribution to effector recruitment. Genetically different centromeres
4   in our hybrid model system should therefore become functionally more different. To target
5   pericentromeric heterochromatin, we deleted CENP-B, which is the only centromeric chromatin
6   component that is dispensable for core centromere function. CENP-B is recently acquired in
7   mammals and fission yeast from a pogo-like transposase (Casola et al., 2007; Kipling and
8   Warburton, 1997), and several domesticated transposases regulate heterochromatin (Gao et al.,
9   2020; Jangam et al., 2017; Nozawa et al., 2010; Yang et al., 2017). In mouse and human cultured
10 cells and fission yeast, CENP-B contributes to pericentromeric heterochromatin formation via
11 heterochromatin protein recruitment (Nakagawa et al., 2002; Okada et al., 2007; Otake et al.,
12 2020), so deleting CENP-B should weaken the heterochromatin pathway (Figure
13 1G). Mammalian CENP-B can also contribute to the kinetochore pathway via CENP-C
14 recruitment (Fachinetti et al., 2015), so the functional consequences of CENP-B deletion in our
15 model need to be tested. Second, when the asymmetric kinetochore pathway is weakened, we
16 predict that centromeres become functionally more similar due to the symmetric heterochromatin
17 pathway. We selected CENP-C as a key scaffold protein in the kinetochore pathway that is
18 known to evolve rapidly under positive selection (Klare et al., 2015; Schueler et al., 2010;
19 Talbert et al., 2004). Under the parallel pathway model, CENP-C interfaces have co-evolved
20 with interacting partners to modulate effector recruitment. Thus, introducing a divergent allele of
21 CENP-C in mouse cells (e.g., rat CENP-C, in which 32% of the amino acid sequence is
22 different) is predicted to disrupt such interactions and weaken the kinetochore pathway (Figure
23 1H). Third, if proteins in the kinetochore and heterochromatin pathways have evolved to
24 modulate effector recruitment, we predict signatures of positive selection in multiple protein
25 domains involved in effector recruitment. In contrast, the previous model of an arms race limited
26 to interactions between centromere DNA and DNA-binding proteins only predicts rapid
27 evolution of protein domains involved in DNA binding (Henikoff et al., 2001; Malik and
28 Henikoff, 2001). Our observations are consistent with all three predictions, supporting our
29 parallel pathway model for drive and suppression.
30
31 **Deleting CENP-B weakens the heterochromatin pathway and makes centromeres**
32 **functionally more asymmetric.**
33
34 To determine the contribution of CENP-B to effector recruitment, we created *Cenpb* null mice
35 using CRISPR genome editing (Supplementary Figure 1A-1C). We find that loss of CENP-B
36 weakens both the kinetochore and heterochromatin pathways, as shown by reduced CENP-C and
37 H3K9me3 staining, respectively (Figure 2A). These results are consistent with previous findings
38 that CENP-B contributes to CENP-C recruitment and to formation of pericentromeric
39 heterochromatin (Fachinetti et al., 2015; Okada et al., 2007; Otake et al., 2020). We also find

1 reduced effector recruitment, as represented by SGO2 staining (Figure 2A), consistent with the
2 idea that CENP-B recruits effectors through the kinetochore and heterochromatin pathways.
3
4 The known functions of CENP-B suggest two hypotheses for how it might affect centromeres in
5 our CHPO hybrid model system. First, as the only centromere protein known to recognize a
6 specific DNA sequence (the CENP-B box in repetitive centromere DNA) (Masumoto et al.,
7 1989), CENP-B could be exploited by selfish larger centromeres with more CENP-B boxes to
8 increase asymmetry via the kinetochore pathway. Second, CENP-B may suppress functional
9 differences between centromeres by increasing the symmetric heterochromatin pathway. To test
10 these hypotheses, we generated *Cenpb* null mice with paired larger and smaller centromeres
11 through two generations of crosses (Figure 2B and Supplementary Figure 1D), and analyzed
12 kinetochore pathway asymmetry and functional differences between centromeres. To determine
13 the impact of CENP-B on the kinetochore pathway, we analyzed CENP-C in meiotic bivalents
14 with paired larger and smaller centromeres in second-generation hybrid *Cenpb$^{-/-}$* oocytes.
15 CENP-C was reduced to a similar extent on both larger and smaller centromeres (Supplementary
16 Figure 1E) and consistent with this equivalent reduction, the kinetochore asymmetry remained
17 intact (Figure 2C). Therefore, CENP-B does not contribute to asymmetry in the kinetochore
18 pathway, arguing against the first hypothesis that selfish centromere DNA exploits the
19 kinetochore pathway via CENP-B.
20
21 To test the second hypothesis, that CENP-B acts as a suppressor through the symmetric
22 heterochromatin pathway (Figure 1G), we examined functional differences between centromeres
23 in second-generation hybrid oocytes. As a readout, we analyzed chromosome position on the
24 spindle at metaphase I (Figure 2D). Chromosome position is sensitive to differences in
25 interactions with spindle microtubules between centromeres of homologous chromosomes,
26 which are paired in a meiotic bivalent. If the paired centromeres are genetically and functionally
27 similar, then chromosomes align at the spindle equator in a typical metaphase configuration. In
28 our hybrid model systems, paired centromeres are genetically and functionally different, and
29 bivalents are positioned off-center on the spindle, with the larger centromere closer to its
30 attached pole (Akera et al., 2019; Chmátal et al., 2014). Manipulations that make these
31 genetically different centromeres functionally more similar will lead to positioning closer to the
32 spindle equator, as previously shown by manipulating BUB1 kinase to equalize MCAK levels on
33 larger and smaller centromeres (Akera et al., 2019). Conversely, manipulations that make the
34 centromeres functionally more different will position bivalents closer to the poles. We find that
35 asymmetric bivalents with genetically different centromeres are positioned more off-center,
36 closer to the spindle poles, in *Cenpb$^{-/-}$* compared to control *Cenpb$^{+/-}$* oocytes (Figure 2E). This
37 result indicates that paired larger and smaller centromeres are functionally more different in the
38 absence of CENP-B, consistent with the prediction that the symmetric heterochromatin pathway
39 is weakened, making the asymmetric kinetochore pathway relatively more dominant (Figure
40 1G).

5

1

2 **Introducing a divergent CENP-C allele weakens the kinetochore pathway and makes**

3 **centromeres functionally more symmetric.**

4

5 To weaken the kinetochore pathway, we targeted CENP-C because it serves as a hub for

6 recruiting kinetochore proteins. Because it is an essential gene, we introduced a divergent allele

7 rather than deleting it. Our model predicts that CENP-C has co-evolved with interacting partners

8 to modulate effector recruitment, so that an allele from another species will disrupt these

9 interactions and weaken the kinetochore pathway (Figure 1H, Prediction 1). To test this

10 prediction, we selected rat as a model organism close to mouse with divergent centromere DNA

11 and proteins (Gibbs et al., 2004; Takeiri et al., 2013). Because protein interfaces change by

12 genetic drift as well as by selection, an allele from a closely related species minimizes

13 incompatibilities coming from stochastic changes. We find that effector recruitment, as

14 represented by SGO2 staining, is reduced when rat CENP-C is expressed in mouse oocytes,

15 compared to mouse CENP-C (Figure 3A). This result is consistent with our model prediction and

16 could reflect differences between mouse and rat CENP-C in their recruitment to centromeres or

17 in their interactions with other kinetochore proteins. For example, evolution at an interface with

18 CENP-A nucleosomes or with CENP-B may disrupt rat CENP-C recruitment to centromeres.

19 Alternatively, CENP-C evolution might impact the domains that mediate interactions with other

20 kinetochore proteins involved in SGO2 recruitment. We find that mouse and rat CENP-C are

21 equally recruited to mouse centromeres (Figure 3B), indicating functional changes at an interface

22 with other kinetochore proteins.

23

24 Rat CENP-C expression provides an experimental tool to weaken the kinetochore pathway,

25 allowing us to test our prediction that genetically different centromeres become functionally

26 more similar in our hybrid model system (Figure 1H, Prediction 2). Using chromosome position

27 as a functional readout of centromere asymmetry (Figure 2D), we find that expression of rat

28 CENP-C in CHPO hybrid oocytes (Figure 1C) leads to bivalents positioned closer to the spindle

29 equator (Figure 3C). This result indicates that the paired larger and smaller centromeres are

30 functionally more similar, consistent with the prediction that the symmetric heterochromatin

31 pathway becomes relatively more dominant when the asymmetric kinetochore pathway is

32 weakened (Figure 1H, Prediction 2).

33

34 **Proteins in the kinetochore and heterochromatin pathways have signatures of recurrent**

35 **adaptive evolution.**

36

37 The original model of centromere drive and suppression posits an arms race between selfish

38 centromere DNA and DNA-binding proteins such as CENP-A (Henikoff et al., 2001; Malik and

39 Henikoff, 2001). This model predicts adaptive evolution of centromere protein domains that

40 physically interact with DNA, and conservation of domains and other centromere proteins that

1   do not bind DNA. In contrast, our parallel pathway model predicts signatures of recurrent

2   adaptive evolution in protein domains leading to effector recruitment, including those that do not

3   directly contact centromere DNA (Figure 4A). These changes could either weaken the

4   kinetochore pathway or strengthen the heterochromatin pathway to make genetically different

5   centromeres functionally more similar (Figure 1F). Rapid evolution of centromere proteins has

6   been reported in several eukaryotic lineages, but there are no mechanistic studies of drive in

7   these lineages (Finseth et al., 2015; Malik and Henikoff, 2001; Schueler et al., 2010; van der Lee

8   et al., 2017). To analyze centromere protein evolution in a system where we have identified drive

9   effectors, we tested for signatures of positive selection in Murinae. Because the sparseness of the

10   phylogenetic tree of currently available Murinae genomes limits our statistical power to detect

11   positive selection, we sequenced six new genomes (Figure 4B) using linked-read whole genome

12   sequencing (10x Genomics). Each genome was assembled onto the *Mus musculus* reference

13   genome (mm10) with LongRanger and *de novo* assembled with Supernova (see Methods and

14   Supplementary Table 1). Sampling evolutionary time more comprehensively increases our

15   opportunities to observe adaptive changes (and minimize false positives from stochastic changes

16   by genetic drift), especially those adaptive changes that are common to multiple independent

17   lineages. Thus, these genomes provide a valuable resource for molecular evolution approaches in

18   mouse as a mammalian model organism, such as our analyses of centromere proteins discussed

19   below.

20

21   Low rates of nonsynonymous substitutions, which change the encoded amino acid, relative to

22   synonymous substitutions (dN/dS) indicate purifying selection, as deleterious substitutions are

23   selected against. Higher dN/dS indicates either adaptive evolution or loss of constraint,

24   necessitating further analysis to identify signatures of positive selection (Echave et al., 2016;

25   Sironi et al., 2015). We calculated dN/dS for all annotated mouse-rat orthologous genes. We find

26   that multiple genes encoding centromere proteins have high dN/dS relative to the genome overall

27   (Figure 4C), and the average dN/dS for these genes is significantly higher than for any other

28   subcellular compartment (Figure 4D). We selected 46 genes with well-characterized centromere

29   functions to analyze for signatures of positive selection based on phylogenetic analysis, using

30   PAML (Yang, 2007). Consistent with our prediction, we find such signatures at genes in

31   multiple modules of the kinetochore and heterochromatin pathways (Figure 4E).

32

33   Extensive previous studies of centromere organization and function have established functional

34   modules which can recruit drive effectors either directly or indirectly (Figure 4A). To fit our

35   observations into this framework, we assigned genes to these modules. One module is CENP-A

36   chromatin and the constitutive centromere-associated network (CCAN). Selfish centromere DNA

37   can increase effector recruitment by expanding CENP-A chromatin through increased deposition

38   of CENP-A nucleosomes. This process depends on a specialized histone chaperone, HJURP,

39   which is targeted to centromeres by the MIS18 complex though interactions with CENP-C or

40   CENP-I (Dunleavy et al., 2009; Foltz et al., 2009; Fujita et al., 2007; Moree et al., 2011; Shono

1   et al., 2015). We find rapid evolution of HJURP, MIS18BP1, CENP-I, and the domain of CENP-

2   C that interacts with the MIS18 complex (Figure 4E and 5A). In contrast, heterochromatin

3   proteins such as HP1 paralogs and SUV39H1, which are not specific to

4   centromeres/pericentromeres, are highly conserved (Supplementary Table 2), consistent with the

5   idea that heterochromatin broadly suppresses selfish genetic elements regardless of the

6   underlying DNA sequence (Allshire and Madhani, 2017). These findings suggest that selection

7   acts on the CENP-A chromatin assembly pathway to prevent expansion, but selfish centromere

8   DNA does not exploit the heterochromatin pathway, consistent with our observation that

9   genetically different centromeres have symmetric heterochromatin in our intra-species and inter-

10   species hybrids (Figure 1D and our unpublished data).

11

12   Under our model (Figure 1E), selfish centromere DNA can also recruit more effectors through

13   the kinetochore pathway by strengthening direct interactions with CENP-A or with CCAN

14   components, leading to larger kinetochores and more BUB1 kinase. Proteins can subsequently

15   adapt by weakening interactions either with DNA or with other kinetochore proteins (Figure 4A,

16   DNA interface and kinetochore assembly). Within the CCAN, CENP-C and CENP-T sub-

17   pathways connect CENP-A chromatin to kinetochore proteins. The middle part of CENP-C

18   interacts with CENP-A nucleosomes, while the N-terminus interacts with the MIS12 kinetochore

19   complex (Petrovic et al., 2016; Weir et al., 2016). Similarly, the CENP-TWSX nucleosome-like

20   complex contacts centromere DNA, and the other end of CENP-T interact with MIS12 and

21   NDC80 kinetochore complexes (Cortes-Silva et al., 2020; Nishino et al., 2012; Veld et al.,

22   2016). Consistent with our model, we detect signatures of positive selection in the chromatin-

23   interacting domains and the kinetochore-interacting domains of both CENP-C and CENP-T

24   (Figure 5A and 5B). In contrast, the DNA-interacting domain of CENP-B is conserved,

25   consistent with our finding that selfish centromere DNA does not exploit CENP-B. Unlike in

26   other eukaryotic lineages such as monkeyflower, fly, and primates (Finseth et al., 2015; Malik

27   and Henikoff, 2001; Schueler et al., 2010), we do not detect signatures of positive selection in

28   the part of CENP-A that can be aligned in Murinae species, but the N-terminal tail is duplicated

29   in some species and therefore difficult to analyze by standard methods (Supplementary Figure

30   2A). Diversification of the CENP-A N-terminal tail is also observed in plants, where crosses

31   between strains expressing different alleles exhibit zygotic segregation errors and genome

32   elimination (Maheshwari et al., 2015).

33

34   In the kinetochore module, proteins can adapt to weaken the kinetochore pathway by reducing

35   either kinetochore assembly or BUB1 binding to the kinetochore (Figure 4A). We find rapid

36   evolution of the kinetochore proteins DSN1, KNL1, and NDC80. DSN1 is a component of the

37   MIS12 complex which assembles onto the CCAN and serves as a platform for binding KNL1

38   and the NDC80 complex (Petrovic et al., 2014). KNL1 contains multiple protein docking motifs,

39   including repeated MELT motifs that recruit BUB1 kinase (Musacchio and Desai, 2017). Thus,

40   changes in DSN1 and KNL1 can regulate kinetochore assembly and BUB1 recruitment.

8

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 5C). 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, components of the heterochromatin pathway, particularly inner centromere proteins, can adapt to increase effector recruitment relative to the kinetochore pathway in our model. In the inner centromere module, which links heterochromatin to effectors (Figure 4A), INCENP is a scaffold component of the CPC that interacts directly with heterochromatin and indirectly with SGO2 (Abe et al., 2016; Ainsztein et al., 1998; Tsukahara et al., 2010). Other CPC components, Borealin and Survivin, regulate SGO2 recruitment and pericentromeric localization (Tsukahara 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 another 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 5D), suggesting that INCENP can adapt to selfish centromere DNA by modulating its localization to pericentromeric heterochromatin and ultimately the recruitment of SGO2 and Aurora B.

Finally, in the effector module we find rapid evolution of SGO2, which is recruited by both the kinetochore and heterochromatin pathways (Figure 1B). This finding suggests that SGO2 can 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 analysis shows 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.

Discussion

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 the host can make centromeres functionally equivalent by minimizing the contribution of the kinetochore pathway relative to the heterochromatin pathway (Figure 1F). This model predicts

9

1 that disruption of either pathway will reduce effector (e.g., SGO2) recruitment, but the functional
2 consequences will depend on which pathway is affected. Centromeres become either functionally
3 more different if the symmetric heterochromatin pathway is weakened, or more similar if the
4 asymmetric kinetochore pathway is weakened. In our experiments, either deletion of CENP-B or
5 introduction of a divergent allele of CENP-C leads to SGO2 reduction to a similar extent (Figure
6 2A and Figure 3A). However, genetically different centromeres in CHPO hybrid oocytes become
7 functionally more different when CENP-B is deleted (Figure 2E), whereas they become
8 functionally more similar when rat CENP-C is expressed (Figure 3C). CENP-B deletion weakens
9 the symmetric heterochromatin pathway, as shown by reduced H3K9me3, making the
10 asymmetric kinetochore pathway more dominant. Loss of CENP-B also reduces CENP-C
11 recruitment but does not affect the asymmetry between larger and smaller centromeres (Figure
12 2C). Complementing these findings, the CENP-C results are consistent with our model
13 prediction that natural selection has acted on CENP-C interfaces involved in effector
14 recruitment, so a divergent rat CENP-C interacts less well with mouse binding partners in the
15 kinetochore pathway. Therefore, expression of rat CENP-C weakens the asymmetric kinetochore
16 pathway, making the symmetric heterochromatin pathway relatively more dominant.
17
18 Our molecular evolution analysis shows adaptive evolution in multiple centromere proteins and
19 in specific domains that interact with CENP-A chromatin or with other proteins leading to
20 effector recruitment (Figure 4 and Figure 5). The previous model of a molecular arms race
21 limited to interactions between centromere DNA and DNA-interacting proteins (such as CENP-
22 A) (Henikoff et al., 2001) does not explain the more widespread recurrent evolution of
23 centromere proteins. An alternative explanation, independent of centromere drive, is that the
24 selective pressure may be related to non-centromere functions. For example, kinetochore
25 proteins are repurposed for neural development in fly and worm (Cheerambathur et al., 2019;
26 Zhao et al., 2019), and KNL1 (also known as CASC5) is implicated in human brain size
27 regulation (Javed et al., 2018; Shi et al., 2016). However, such non-centromere functions have
28 not been identified more broadly in eukaryotes. In contrast, our parallel pathway model predicts
29 recurrent evolution of proteins in both pathways to equalize centromeres by weakening the
30 kinetochore pathway or strengthening the heterochromatin pathway. In our model, selfish
31 centromere DNA evolves to exploit the kinetochore pathway by recruiting more of a protein that
32 ultimately recruits effectors. To suppress functional differences between centromeres, proteins in
33 the kinetochore pathway can adapt to minimize the impact of selfish centromere DNA on
34 kinetochore formation or effector recruitment. Furthermore, proteins in the heterochromatin
35 pathway such as CENP-B and INCENP can adapt to increase effector recruitment equally at all
36 centromeres, or SGO2 can adapt by modulating its recruitment by either pathway (Figure 1F).
37 The acidic domain of CENP-B is implicated in recruiting heterochromatin proteins (Otake et al.,
38 2020), and the number of negatively charged amino acids in this domain is recurrently changed
39 in mammals (Supplementary Figure 2B and 2C). Although these changes are not analyzed in
40 PAML, they suggest that CENP-B may have evolved to regulate pericentromeric

1 heterochromatin. Overall, a protein network for effector recruitment can adapt to minimize
2 asymmetric recruitment by selfish centromere DNA, while maintaining essential functions of the
3 kinetochore and of microtubule destabilizing factors for accurate chromosome segregation.
4
5 Our results suggest an explanation for the conservation of CENP-B in mammals, as well as the
6 presence of its binding sequence, the CENP-B box, at most mammalian centromeres with the
7 notable exception of the Y chromosome. Although CENP-B is the only centromere protein
8 known to bind a specific DNA sequence in mammals, neither the protein nor the binding
9 sequence is essential for centromere function (Amor et al., 2004; Hudson et al., 1998; Kapoor et
10 al., 1998; Logsdon et al., 2019; Perez-Castro et al., 1998). We propose that CENP-B is conserved
11 because it suppresses functional differences between centromeres by strengthening the
12 heterochromatin pathway (Figure 6), consistent with a more general function of heterochromatin
13 in suppressing many selfish genetic elements (Allshire and Madhani, 2017). This CENP-B
14 function is important only when centromeres of homologous chromosomes are different, which
15 would frequently occur in outbred populations. Loss of CENP-B therefore increases functional
16 difference between larger and smaller centromeres in our hybrid model, but does not
17 significantly impair fertility or viability in inbred laboratory strains (Hudson et al., 1998; Kapoor
18 et al., 1998; Perez-Castro et al., 1998). A potential cost of increasing heterochromatin, however,
19 is that its invasion into CENP-A chromatin disrupts centromere function (Ohzeki et al., 2016).
20 We therefore propose that mammalian CENP-B has acquired an additional function to maintain
21 CENP-A chromatin, by recruiting CENP-C and CENP-A chromatin regulators (Fachinetti et al.,
22 2015; Otake et al., 2020) (Figure 6). By regulating both CENP-A chromatin and
23 heterochromatin, alternative functions of CENP-B in different chromatin environments may
24 suppress functional differences between centromeres through heterochromatin while maintaining
25 centromere function. CENP-B can suppress differences between centromeres only if its functions
26 are insensitive to expansion of the number of CENP-B binding sites; otherwise it would
27 contribute to higher levels of effector recruitment by DNA repeat expansions. Indeed, we find
28 that CENP-B does not contribute to asymmetry in CENP-C recruitment between larger and
29 smaller centromeres, despite 6- to 10-fold differences in minor satellite sequences containing
30 CENP-B boxes (Iwata-Otsubo et al., 2017). This result suggests that CENP-B recruits CENP-C
31 only within the CENP-A chromatin domain, so that CENP-B binding outside of this domain does
32 not strengthen the kinetochore pathway (Figure 6). Furthermore, the heterochromatin symmetry
33 between larger and smaller centromeres suggests that although CENP-B contributes to initiating
34 heterochromatin formation, for example by recruiting an H3K9 methyltransferase,
35 heterochromatin spreading does not depend on the number of CENP-B boxes (Figure 6).
36 Initiation of heterochromatin propagation is a common mechanism to regulate heterochromatin
37 formation, as in the example of X inactivation where XIST initiates heterochromatinization of
38 the entire chromosome (Allshire and Madhani, 2017). Thus, CENP-B functions in CENP-A
39 chromatin and heterochromatin are insensitive to repeat expansion. A centromere variant
40 completely lacking CENP-B boxes, however, will lose to an existing centromere in female

1   meiosis because it will recruit less effectors by both the kinetochore and heterochromatin
2   pathways. Therefore, CENP-B boxes are maintained at most centromeres, but this selective
3   pressure does not affect the Y chromosome, which never experiences female meiosis and does
4   not bind CENP-B (Gamba and Fachinetti, 2020).
5
6   Genetic conflict between selfish centromere DNA and centromere-binding proteins potentially
7   explains the complexity of eukaryotic centromeres. Opportunities for selfish genetic elements to
8   exploit the chromosome segregation machinery are not limited to female meiosis, as selfish
9   plasmids (e.g., 2-micron plasmids in budding yeast) benefit by maximizing their transmission to
10 daughter cells in mitosis (Malik and Henikoff, 2009; Rizvi et al., 2017). These opportunities are
11 limited by the strong epigenetic component of most eukaryotic centromeres, which are not
12 defined by specific DNA sequences. Centromeres cannot be completely independent of the
13 underlying DNA sequence, however, because some protein must interact with DNA, so different
14 sequences can have different binding affinities or impact the structure of the centromeric
15 nucleosome complex (Allu et al., 2019). The presence of multiple pathways to form a
16 kinetochore (e.g., via CENP-ACLN and CENP-TWSX connected by CENP-HIKM, or via
17 CENP-OPQUR) (Cortes-Silva et al., 2020; Hamilton et al., 2020; Nishino et al., 2012; Pesenti et
18 al., 2018; Veld et al., 2016; Weir et al., 2016; Yan et al., 2019) allows proteins to adapt by
19 minimizing a pathway that is exploited by a selfish element, while maintaining kinetochore
20 function via other pathways. Consistent with this idea of independent modules for kinetochore
21 formation, CENP-A depletion leads to proportional reduction of centromeric CENP-C, whereas
22 CENP-T and CENP-I persist longer (Fachinetti et al., 2013). In addition, recurrent changes in
23 kinetochore modules are observed throughout eukaryotic evolution, such as changes in the
24 number of MELT motifs in KNL1 and replacement of the SKA complex by the DAM complex
25 (Hooff et al., 2017; Tromer et al., 2015). Regulation of kinetochore-microtubule attachment
26 stability may be another way to suppress selfish genetic elements, as MELT motifs recruit BUB1
27 and SKA and DAM complexes stabilize attachments. Thus, internal conflicts between selfish
28 genetic elements and the chromosome segregation machinery may have shaped complexity in
29 eukaryotic centromeres.
30

31 **Acknowledgements**
32
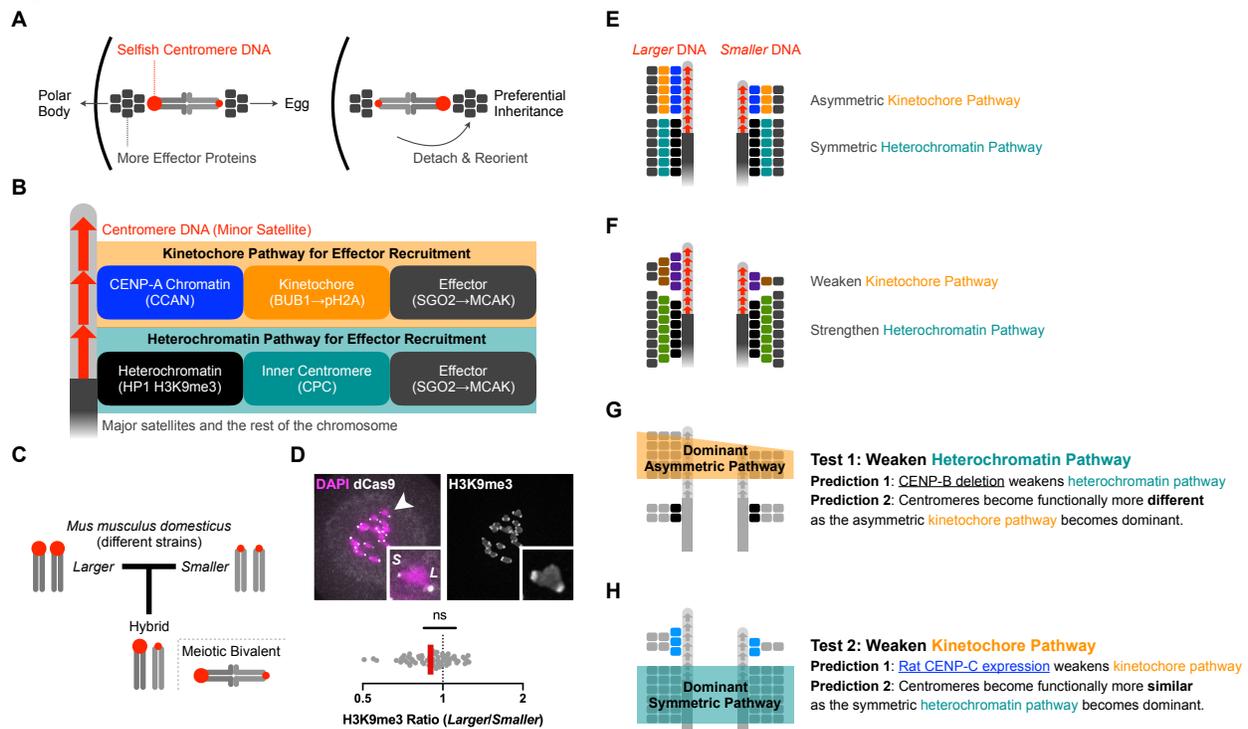
Figure 1: Parallel pathway model for centromere drive and suppression.

(A) Centromere drive by recruiting effector proteins that destabilize interactions with spindle microtubules. Selfish centromeres recruit more effector proteins (SGO2 and MCAK), preferentially detach from microtubules when facing the cortical side of the spindle, and reorient to bias their segregation to the egg. The cortical side of a meitoic bivalent will segregate to the polar body, whereas the other side will segregate to the egg.

(B) Two pathways for effector recruitment. CENP-A and the CCAN (constitutive centromere-associated network) connect centromere DNA to the kinetochore, which assembles during meiosis or mitosis. 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. Although the two pathways are shown as independent here for simplicity, they are not completely separate, as the CPC is also recruited by SGO2 via the kinetochore pathway (Yamagishi et al., 2010). This contribution to effector recruitment depends on the kinetochore, so we consider it part of the kinetochore pathway.

(C) 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.

(D) CHPO hybrid oocytes were microinjected with cRNA for dCas9-GFP and gRNA targeting minor satellite centromere DNA to distinguish larger (*L*) and smaller (*S*) centromeres, fixed at meiosis I, and stained for H3K9me3. 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.

(E) Symmetric heterochromatin pathway and asymmetric kinetochore pathway in our hybrid model system. The amount of recruited proteins is represented by the number of boxes, using the same color code as panel B.

(F) Parallel pathway model for 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 for effectors.

(G) As CENP-B recruits heterochromatin proteins, deleting CENP-B weakens the heterochromatin pathway for effector recruitment (prediction 1), making the asymmetric kinetochore pathway dominant and centromeres functionally more asymmetric (prediction 2).

(H) 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).

14
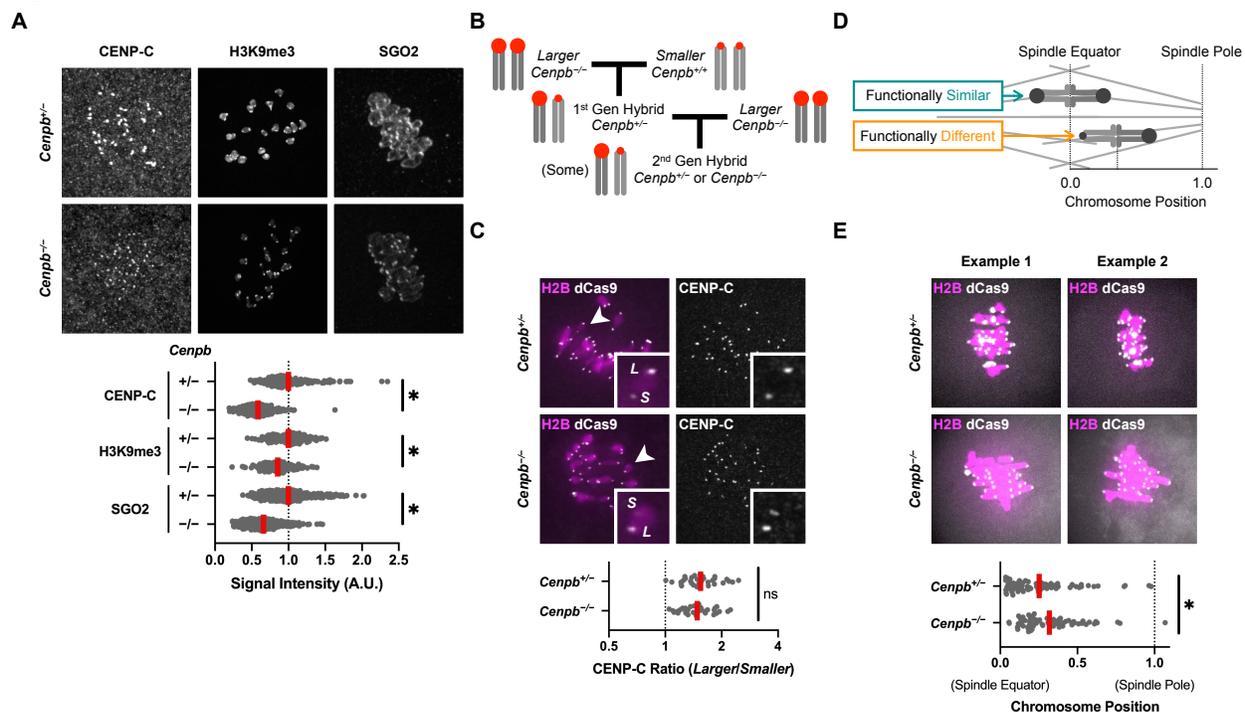
Figure 2: Deleting CENP-B weakens the heterochromatin pathway and makes centromeres functionally more asymmetric.

(A) *Cenpb$^{+/-}$* or *Cenpb$^{-/-}$* oocytes were fixed in meiosis I and stained for CENP-C, H3K9me3, or SGO2. 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 with larger and smaller paired centromeres.

(C) Second generation hybrid oocytes were microinjected with cRNA for dCas9-EGFP and gRNA targeting minor satellite centromere DNA, fixed in meiosis I, and stained for CENP-C. The CENP-C ratio for each pair of larger (*L*) and smaller (*S*) centromeres on a bivalent is plotted (n=34 bivalents for each genotype); red line, geometric mean; ns: not significant.

(D) 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.

(E) Second generation hybrid oocytes were microinjected with cRNAs for dCas9 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); red line, mean.

15
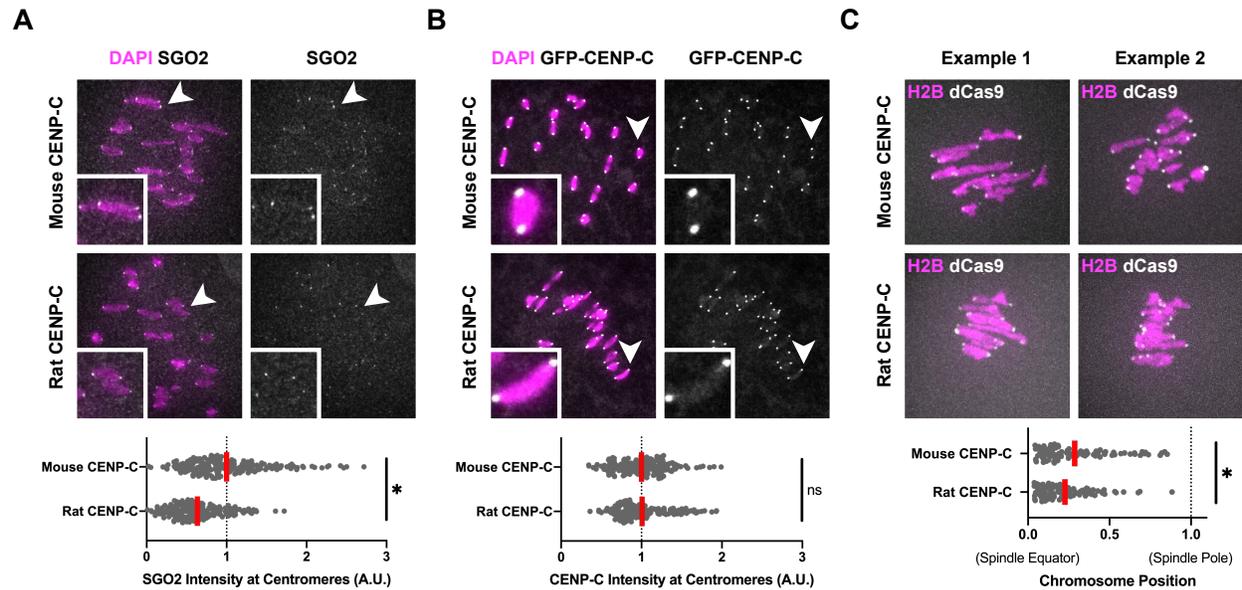
Figure 3

Figure 3: Introducing a divergent CENP-C allele weakens the kinetochore pathway and makes centromeres functionally more symmetric.

(A and B) CF-1 oocytes were microinjected with cRNA for GFP-tagged mouse or rat CENP-C and fixed in meiosis I. Cells were stained for SGO2 (A) or analyzed for GFP fluorescence (B). 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.

(C) CHPO hybrid oocytes (see Figure 1C) 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 ≥20 oocytes for each construct); red line, mean.

16

Figure 4

A

Test 3: Molecular Evolution Analyses

Prediction: Recurrent adaptive evolution in protein domains for effector recruitment

| DNA interface (Figure **5A, B**) | Kinetochore assembly (Figure **5A, B**) | SGO2 recruitment through BUB1 (Figure **5C**) |

CENP-A Chromatin | Kinetochore | Effector

Heterochromatin | Inner Centromere | Effector

| Heterochromatin interface (Figure **5D**) | SGO2 recruitment through CPC (Figure **5D**) |

B Examples of codons in CENP-C

Mus musculus	GGT GGG
Mus spretus	GGT GGG
Mus caroli	GGT GGG
Mus pahari	GGT GGG
Hylomyscus alleni	GGT GGG
Praomys delectorum	GGT GGG
Mastomys natalensis	GCA GGG
Grammomys dolichurus	AAT GGA
Rhabdomys dilectus	AGT GGG
Rhynchomys soricoides	GGT GGG
Rattus norvegicus	AGA GAG

* ns

C

D

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%

17

1   **Figure 4: Proteins in the kinetochore and heterochromatin pathways have signatures of**
2   **recurrent adaptive evolution.**

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

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

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

14   (D) Average dN/dS across subcellular compartments. Red line, mean; *$p<0.05$ for comparison to
15   all other compartments.

16   (E) To test for signatures of positive selection in PAML, the likelihood of models of neutral
17   codon evolution (M1 or M7) are compared to models allowing positive selection (M2 or M8).
18   CENP-A and CENP-B are examples of genes without signatures of positive selection (see
19   Supplementary Table 2 for other genes and Supplementary Figure 2 for further analyses of
20   CENP-A and CENP-B). The number of analyzed codons is less than the total protein length as
21   insertions, deletions, and ambiguous alignment are not analyzed. The number of positive
22   selection sites is the number of codons with $P>0.90$ from Naive Empirical Bayes (NEB) analysis
23   or Bayes Empirical Bayes (BEB) analysis from model 2 or 8.
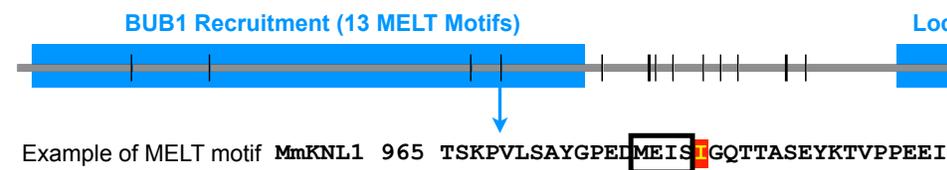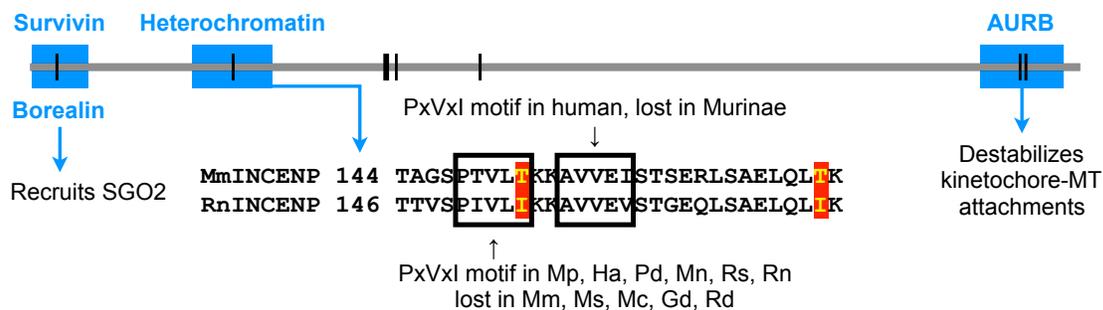
18
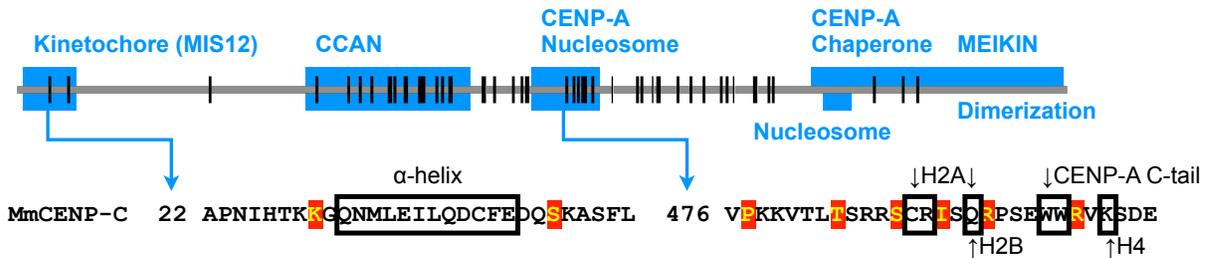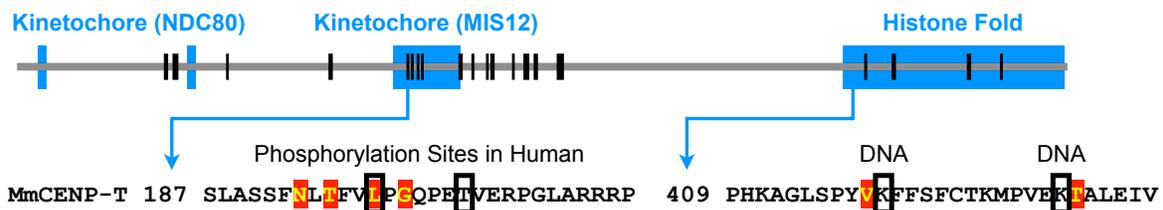
Figure 5

Figure 5: 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 acid residues. 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

19

1   H2A, H2B, H4 and the CENP-A C-terminal tail are indicated. This domain binds CENP-A
2   nucleosomes more specifically than the more C-terminal nucleosome binding domain (also
3   known as CENP-C motif), which also interacts with H3 nucleosomes (Allu et al., 2019; Kato et
4   al., 2013). The CENP-C C-terminus has multiple functions, including M18BP1 recruitment
5   (Dambacher et al., 2012), MEIKIN recruitment (Kim et al., 2015), and dimerization (Sugimoto
6   et al., 1997).

7   (B) Signatures of positive selection are found in the kinetochore interaction domain and histone
8   fold domain of CENP-T. CDK1-dependent phosphorylation at Thr195 and Ser201 in human
9   CENP-T (substituted with Leu and Thr, respectively, in mice) regulates MIS12 recruitment
10   (Rago et al., 2015; Veld et al., 2016). Signatures of positive selection are detected around these
11   regulatory residues for MIS12 recruitment. Some DNA interacting residues within the histone
12   fold domain are shown (Nishino et al., 2012).

13   (C) Signatures of positive selection are found in the domain of KNL1 that recruits BUB1 via
14   repeated MELT motifs (Krenn et al., 2013). One MELT motif is shown as an example.

15   (D) Signatures of positive selection are found in domains of INCENP that interact with
16   Borealin/Survivin, with heterochromatin, and with Aurora B kinase. Heterochromatin recruits
17   INCENP, and Borealin mediates the interaction with SGO2 (Abe et al., 2016; Ainsztein et al.,
18   1998; Tsukahara et al., 2010). Survivin binds cohesin and pH3T3 at pericentromeres (Kelly et
19   al., 2010; Wang et al., 2010; Yamagishi et al., 2010), providing another mechanism to localize
20   the CPC. A PxVxI motif, which interacts with the HP1 chromoshadow domain, is present in
21   some Murinae species and lost in others, shown with *Mus musculus* (Mm) and *Rattus norvegicus*
22   (Rn) as examples. Other species from the phylogenetic tree in Figure 4B: *Mus spretus* (Ms), *Mus*
23   *caroli* (Mc), *Mus pahari* (Mp), *Hylomyscus alleni* (Ha), *Praomys delectorum* (Pd), *Mastomys*
24   *natalensis* (Mn), *Grammomys dolichurus* (Gd), *Rhabdomys dilectus* (Rd), and *Rhynchomys*
25   *soricoides* (Rs).

Figure 6
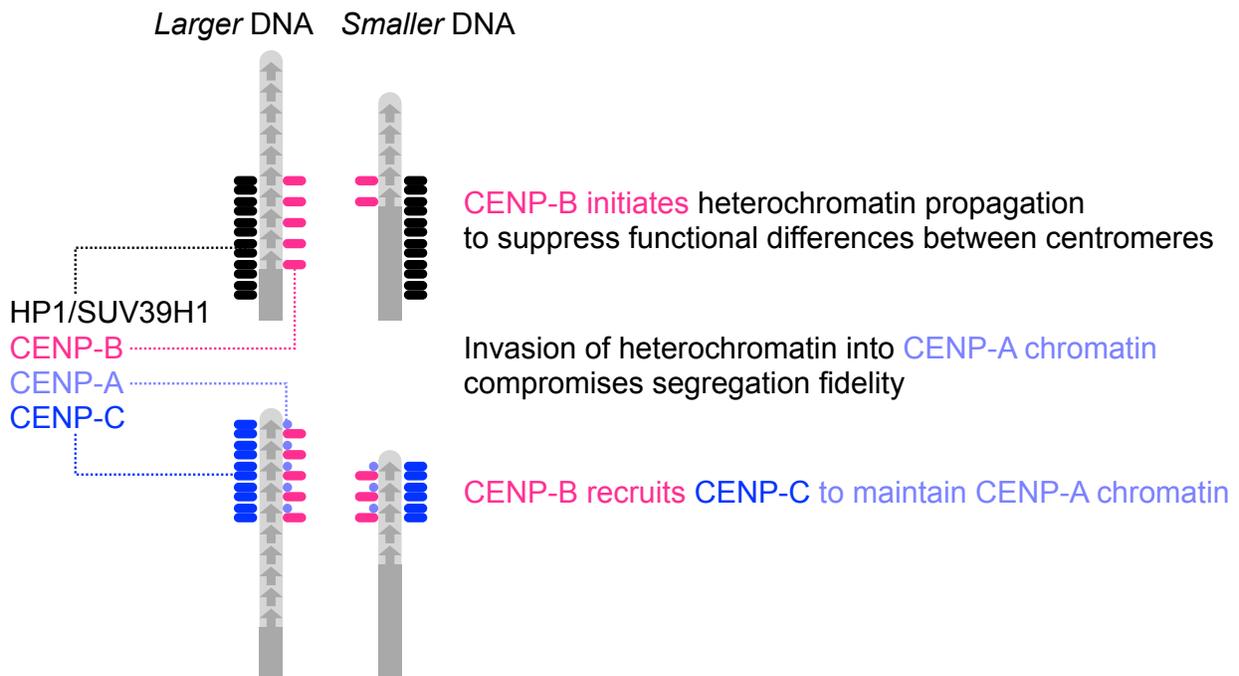
1   **Figure 6: Dual functions of CENP-B in suppressing functional differences in centromeres**
2   **while maintaining CENP-A chromatin.**

3   CENP-B initiates heterochromatin formation to equalize centromeres (top). Despite the

4   difference in CENP-B binding sites, larger and smaller centromeres have similar amounts of

5   H3K9me3 (Figure 1D), indicating that heterochromatin formation is insensitive to CENP-B

6   abundance, likely due to self-propagation of heterochromatin. Invasion of heterochromatin into

7   CENP-A compromises centromere function. To prevent this disruption, we propose that CENP-

8   B has acquired an additional function in CENP-A chromatin (bottom): CENP-B recruits CENP-

9   C but does not contribute to CENP-C asymmetry between larger and smaller centromeres

10   (Figure 2C), suggesting that only CENP-B within CENP-A chromatin recruits CENP-C. Thus,

11   CENP-B functions in heterochromatin and CENP-A chromatin are insensitive to repeat

12   expansion.

21

Supplementary Figure 1
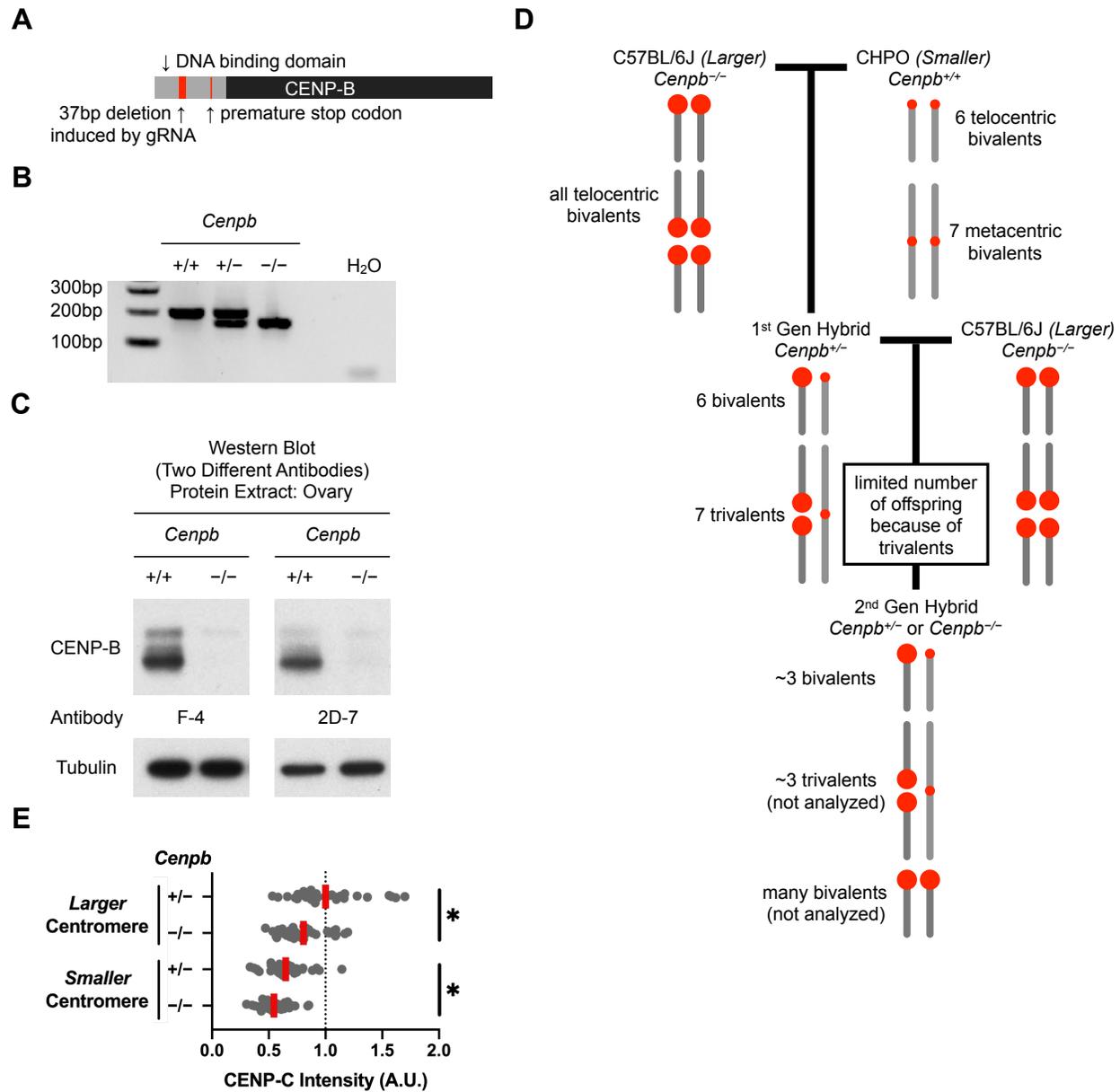
Supplementary Figure 1: CRISPR genome editing creates CENP-B null mice.

(A) Summary of CRISPR genome editing, using a gRNA designed to target the DNA binding domain of CENP-B.

(B) CENP-B genotyping. As the CENP-B mutation is a 37bp deletion, a PCR reaction amplifying the flanking regions can distinguish three genotypes of CENP-B.

(C) Absence of CENP-B protein in *Cenpb*$^{-/-}$ mice. Protein extract from ovary is used to detect CENP-B using two different antibodies.

(D) Detailed crossing scheme to produce second-generation hybrid *Cenpb*$^{-/-}$ mice with larger and smaller paired centromeres (related to Figure 2B). 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).

(E) CENP-C reduction in the second generation hybrid (related to Figure 2C). Oocytes from the second generation hybrid (Figure 2B) 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.
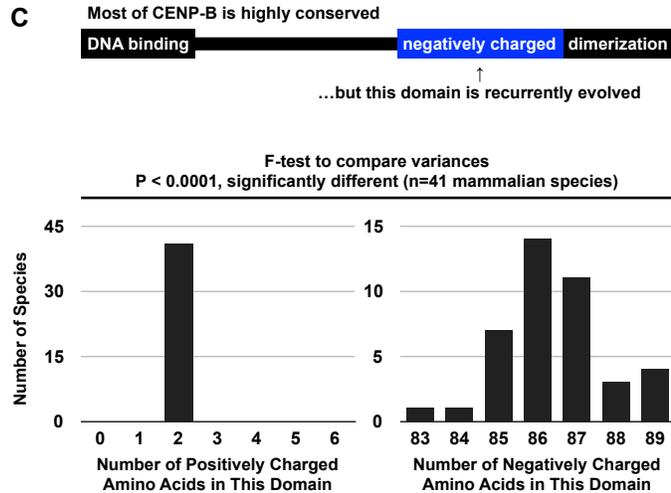
23

Supplementary Figure 2
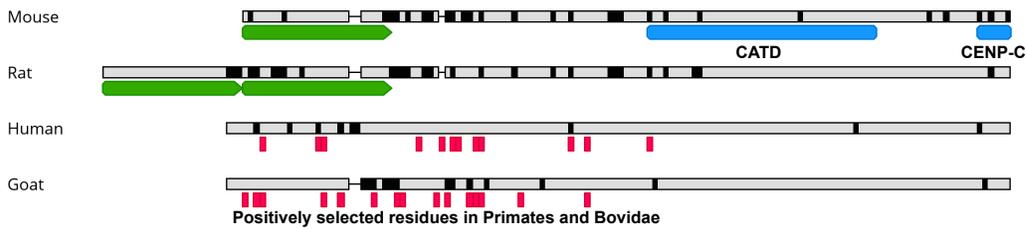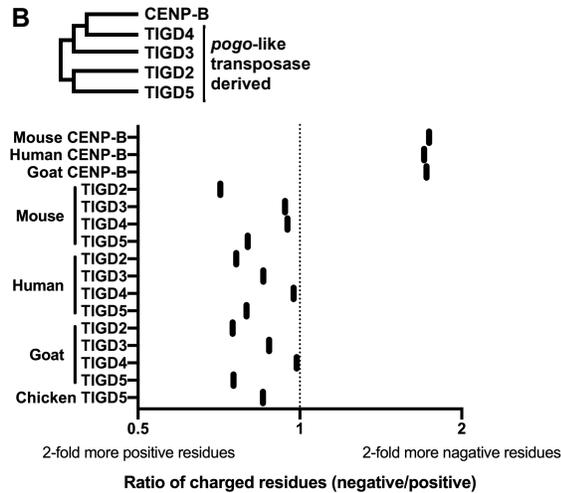
Supplementary Figure 2: Changes in CENP-A and CENP-B cannot be analyzed by standard method to detect adaptive evolution.

(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 lines. Signatures of positive selection were previously found in primate CENP-A (Schueler et al., 2010), shown in red boxes in the human sequence. Bovine genomes (Chen et al., 2019) are used 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

24

1   number of species for each number of positively charged or negatively charged amino acids in

2   this domain is plotted.

Supplementary Table 1

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

Supplementary Table 2

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
MAD1	0.13	0.00	0.00	12*
MAD2	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.

1   KEY RESOURCES TABLE

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 Flour 594 donkey anti-mouse IgG	Molecular Probes	Cat# A-21203; RRID:AB_141633
Alexa Flour 594 donkey anti-rabbit IgG	Molecular Probes	Cat# A-21207; RRID:AB_141637
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
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
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	N/A
EGFP-RnCENP-C	This paper	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

1   LEAD CONTACT AND MATERIALS AVAILABILITY

2

3   Further information and requests for resources and reagents should be directed to and will be

4   fulfilled by the Lead Contact, Michael A. Lampson (lampson@sas.upenn.edu).

5

6   EXPERIMENTAL MODEL AND SUBJECT DETAILS

7

8   **Mice**

9

10   Mouse strains were purchased from the Jackson Laboratory (ZALENDE/EiJ, stock #001392

11   corresponds to CHPO; C57BL/6J, stock# 000664) and from Envigo (NSA, stock# 033

12   corresponds to CF-1). CHPO males were crossed to CF-1 females to generate hybrids shown in

13   Figure 1c. The CHPO strain contains seven Robertsonian fusions (Rb(1.3), Rb(4.6), Rb(5.15),

14   Rb(11.13), Rb(8.12), Rb(9.14), and Rb(16.17)), each of which pairs with two CF-1

15   chromosomes in CHPO hybrid meiosis I to form a trivalent (Chmátal et al., 2014). We included

16   only bivalents (chromosome 2, 7, 10, 18, 19, X) in our analyses to avoid complications of

17   trivalents.

18

19   In order to generate CENP-B null mice, 1-cell embryos (from female CF-1 and male DBA/2J x

20   C57BL/6J hybrid) were collected and microinjected with Cas9 mRNA (TriLink, CleanCap Cas9

21   mRNA, L-7606) and gRNA (GAAGAACAAGCGCGCCA) (Thermo Fisher scientific, GeneArt

22   Precision gRNA Synthesis Kit, A29377). Embryos were cultured *in vitro* until blastocyst stage

23   and transferred to pseudopregnant females to produce a founder mouse carrying 37bp deletion

24   (TGAGCACCATCCTGAAGAACAAGCGCGCCATCCTGGC) that produces a premature stop

25   codon at Leu100 in the DNA binding domain. The founder was crossed with C57BL/6J for

26   multiple generations to remove possible off-target mutations. Mice were genotyped by extracting

27   genomic DNA from tail clip (QIAGEN, DNeasy Blood & Tissue Kit, 69504) and amplifying a

28   *Cenpb* fragment (Agilent, Herculase II Fusion DNA Polymerase). To generate *Cenpb* null mice

29   with larger and smaller centromeres, CHPO females were crossed to C57BL/6J *Cenpb* null males

30   to generate first generation hybrid females, which were then crossed to C57BL/6J *Cenpb* null

31   males to generate second generation hybrid females as shown in Figure 2B and Supplementary

32   Figure 1D. All animal experiments were approved by the Institutional Animal Care and Use

33   Committee and were consistent with the National Institutes of Health guidelines.

34

35   METHOD DETAILS

36

37   **Oocyte collection and culture**

38

39   Female mice (8-14 weeks of age) were hormonally primed with 5U of Pregnant Mare Serum

40   Gonadotropin (PMSG, Calbiochem, cat# 367222) or 0.1mL of CARD HyperOva (Cosmo Bio,

1   KYD-010-EX) 44-48 h prior to oocyte collection. Germinal vesicle (GV)-intact oocytes were

2   collected in M2 medium (Sigma, M7167), denuded from cumulus cells, and cultured in Chatot-

3   Ziomek-Bavister (CZB) medium (Thermo Fisher, MR019D) in a humidified atmosphere of 5%

4   CO_2 in air at 37.8C˚. During collection, meiotic resumption was inhibited by addition of 2.5 mM

5   milrinone. Milrinone was subsequently washed out to allow meiotic resumption. Oocytes were

6   checked for GVBD (germinal vesicle breakdown), and those that did not enter GVBD stage were

7   removed from the culture.

8

9   **Oocyte microinjection**

10

11   GV oocytes were microinjected with ~5 pl of cRNAs in M2 medium (with 2.5 mM milrinone

12   and 3mg/mL BSA) at room temperature (RT) with a micromanipulator TransferMan NK 2

13   (Eppendorf) and picoinjector (Medical Systems Corp.). After the injection, oocytes were kept in

14   milrinone for 16 h to allow protein expression. cRNAs used for microinjections were dCas9-

15   EGFP (dead Cas9 with EGFP at the N terminus) at 1000ng/µL, dCas9-mCherry (dead Cas9 with

16   mCherry at the N terminus) at 1000ng/µL, gRNA that targets minor satellite repeat

17   (ACACTGAAAAACACATTCGT) at 200ng/µL, H2B-EGFP (human histone H2B with EGFP

18   at the C terminus) at 150ng/µL, H2B-mCherry (human histone H2B with mCherry at the C

19   terminus) at 150ng/µL, EGFP-MmCENP-C (mouse CENP-C with EGFP at the N terminus) at

20   100ng/µL, and EGFP-RnCENP-C (rat CENP-C with EGFP at the N terminus) at 100ng/µL.

21   cRNAs were synthesized using the T7 mScriptTM Standard mRNA Production System (CELL

22   SCRIPT) or mMESSAGE mMACHINE SP6 Transcription Kit (Thermo Fisher scientific).

23   gRNAs were synthesized using GeneArt Precision gRNA Synthesis Kit (Thermo Fisher

24   scientific A29377).

25

26   **Live imaging and chromosome position assay**

27

28   For the chromosome position assay, oocytes were collected and microinjected with the

29   constructs indicated in the figure legends. After inducing meiotic resumption by washing out

30   milrinone, oocytes were placed into 2µL drops of CZB media covered with mineral oil in a

31   glass-bottom tissue culture dish (FluoroDish FD35-100) in a heated environmental chamber with

32   a stage top incubator (Incubator BL and Heating Insert P; PeCon GmBH) to maintain 37C˚.

33   Confocal images were collected with a microscope (DMI4000 B; Leica) equipped with a 63x 1.3

34   NA glycerol-immersion objective lens, an xy piezo Z stage (Applied Scientific Instrumentation),

35   a spinning disk confocal scanner (Yokogawa Corporation of America), an electron multiplier

36   charge-coupled device camera (ImageEM C9100-13; Hamamatsu Photonics), and an LMM5

37   laser merge module with 488- and 593-nm diode lasers (Spectral Applied Research) controlled

38   by MetaMorph software (Molecular Devices). Confocal images were collected as z stacks at 0.5

39   µm intervals to visualize the entire meiotic spindle. The position of the spindle near the cortex

40   was confirmed by differential interference contrast images. The spindle equator was determined

1   as a middle of the spindle. The chromosome position of each bivalent was determined as a

2   crossover site and normalized by the distance between spindle equator and spindle poles.

3

4   **Oocyte immunocytochemistry**

5

6   After inducing meiotic resumption by washing out milrinone (4.5 hours for prometaphase

7   staining and 7.5 hours for metaphase staining), MI oocytes were fixed in freshly prepared 2%

8   paraformaldehyde in PBS with 0.1% Triton X-100, pH 7.4, for 20 min at RT, permeabilized in

9   PBS with 0.1% Triton X-100 for 15 min at RT, placed in blocking solution (PBS containing

10   0.3% BSA and 0.01% Tween-20) 15 min RT or overnight at 4C, incubated 1-2 h with primary

11   antibodies in blocking solution, washed 3 times for 15 min each, incubated 1 h with secondary

12   antibodies, washed 3 times for 15 min each, and mounted in Vectashield with DAPI (Vector, H-

13   1200) to visualize chromosomes. Primary antibodies used for this study were rabbit anti-human

14   H3K9me3 (1:500; Abcam, ab8898), mouse anti-mouse SGO2 (1:500, a gift from Yoshinori

15   Watanabe), and rabbit anti-mouse CENP-C (1:2500, a gift from Yoshinori Watanabe).

16   Secondary antibodies were Alexa Fluor 488–conjugated donkey anti-rabbit or donkey anti-

17   mouse or Alexa Fluor 594–conjugated donkey anti-rabbit, or donkey anti-mouse (1:500,

18   Invitrogen). Confocal images were collected as z stacks at 0.5 µm intervals to visualize the entire

19   meiotic spindle, using the spinning disc confocal microscope described above. To quantify

20   centromere signal ratios, optical slices containing centromeres from the same bivalent were

21   added to produce a sum projection using Fiji/ImageJ. Ellipses were drawn around the

22   centromeres, and signal intensity was integrated over each ellipse after subtracting cytoplasmic

23   background. Ratios were obtained for each bivalent by dividing the intensity of the larger

24   centromere by that of the smaller centromere, as determined by dCas9 signal intensity.

25

26   **Whole Genome Sequencing of Six Murinae Species**

27

28   Frozen tissue samples from male individuals were obtained from the Museum of Vertebrate

29   Zoology, Berkeley, CA (MZV) and the Field Museum of Natural History, Chicago, IL (FMNH).

30   *Hylomyscus alleni* (MVZ Mamm 196246) was captured in Cameroon in 2000, *Praomys*

31   *delectorum* (MVZ Mamm 221157) was captured in Malawi in 2007, *Mastomys natalensis* (MVZ

32   Mamm 221054) was captured in Malawi in 2007, *Grammomys dolichurus* (MVZ Mamm

33   221001) was captured in Malawi in 2007, *Rhabdomys dilectus* (FMNH 192475) was captured in

34   Malawi in 2006, and *Rhynchomys soricoides* (FMNH 198792) was captured in The Philippines

35   in 2008. All genomes were sequenced in the Center for Applied Genomics at Children's Hospital

36   of Philadelphia. High molecular weight DNA was extracted following the protocol provided by

37   10xGenomics (CG000072 Rev B Sample Preparation Demonstrated Protocol, DNA Extraction

38   from Fresh Frozen Tissue). Extracted DNA was quality controlled (CG00019 Rev B Sample

39   Preparation Demonstrated Protocol, High Molecular Weight DNA QC), and all of the samples

40   had a mean length greater than 50kb, and high enough concentration to dilute to 1ng/µL for

1   library preparation. Chromium Genome Reagent Kits v2 from 10xGenomics was used to prepare

2   libraries of 2x150 base reads, with read 1 constituting 10xBarcode (16bp) + nmer (6bp) +

3   genome sequence (128bp) and read 2 constituting genome sequence (150bp). i7 index used 8bp

4   sample index, and i5 index was not used. Sequencing depth was calculated based on putative

5   genome size 3Gb and coverage 56x, following 10xGenomics R&D recommendation, and the

6   libraries were sequenced with Illumina HiSeq. Demultiplexed FASTQ files were analyzed using

7   the LongRanger wgs -basic pipeline. This pipeline gave general QC statistics related to the 10x

8   barcoding and number of read pairs present in the FASTQ files. All sample FASTQs contained

9   more than 688M read pairs and have acceptable barcode diversity/% on whitelist. LongRanger

10   was used to assemble genomes, using the *Mus musculus* (mm10) as reference. In parallel,

11   Supernova was used to assemble *de novo* genomes. See Supplementary Table 1 for assembly

12   statistics. In order to obtain protein coding sequences, mm10 annotation was used to annotate

13   reference-guided assemblies, and translated BLAST (tblastn) was used to pull homologous

14   sequences from de novo assemblies using *Mus musculus* protein sequences as query sequences.

15

16   **Phylogenetic Tree Construction**

17

18   The species tree shown in Figure 4B was obtained from maximum likelihood (RAxML) and

19   Bayesian inference (MrBayes). The phylogeny within *Mus* was previously studied (Keane et al.,

20   2011; Thybert et al., 2018). In order to resolve phylogeny in Murinae, the same set of genes that

21   were used to construct a primate phylogenetic tree (Perelman et al., 2011) was aligned by

22   MAFFT (Katoh and Standley, 2013; Katoh et al., 2002). The initial alignment was imported in

23   Geneious Prime, and manually inspected for sequence alignment ambiguity. Ambiguous regions

24   were removed from subsequent analyses. Maximum likelihood tree was constructed with

25   RAxML (Stamatakis, 2014), and Bayesian inference tree was constructed with MrBayes

26   (Huelsenbeck and Ronquist, 2001), with *Peromyscus maniculatus* as outgroup. Both inferences

27   supported the tree topology shown in Figure 4B.

28

29   **Molecular Evolution Analyses**

30

31   In order to create a histogram in Figure 4C, alignments of mouse-rat orthologs were filtered for

32   dS below 0.5, as higher dS values indicate misalignment. A list of genes for each subcellular

33   compartment was obtained from Human Protein Atlas. Mouse-human orthologs were used to

34   calculate average dN/dS for each subcellular compartment in Figure 4D. The analysis to identify

35   signatures of positive selection (PAML) is highly sensitive to alignment errors, so automated

36   genome-wide analysis is prone to false positives (van der Lee et al., 2017). To prevent these

37   errors, alignments for selected genes were manually inspected. Coding sequences for each gene

38   were aligned by Geneious Alignment (translation align) implemented in Geneious Prime, and

39   manually inspected for sequence alignment ambiguity. Insertions or deletions as well as their

40   flanking codons were removed from analyses. To test signatures of positive selection, we

1 compared the likelihood of models of neutral codon evolution to models of codon evolution
2 allowing positive selection, implemented in PAML version 4 (Yang, 2007). The neutral model
3 M1 (fixed dN/dS values between 0 to 1) and M2 (M1 parameters plus dN/dS > 1) were
4 compared in the first test, and the neutral model M7 (dN/dS values fit a beta distribution from 0
5 to 1) and M8 (M7 parameters plus dN/dS > 1) were compared in the second test, assuming the
6 F3x4 model of codon frequencies. Degree of freedom for each test was 2, and the log likelihood
7 test was significant above 5.99 (p < 0.05). We first used the species tree, and signatures of
8 positive selection were confirmed using a gene tree for each gene, created by RAxML.
9
10 QUANTIFICATION AND STATISTICAL ANALYSIS
11
12 Data points are pooled from at least two independent experiments. The following statistical
13 methods were used: unpaired t test in Figures 2A, 2C, 2E, 3A, 3B, 3C, and S1E; Mann-Whitney
14 U test in Figure 4D; chi square test for goodness of fit for deviations from 1 in Figure 1D and for
15 statistical models (likelihood-ratio test) in Figure 4E and Supplementary Table 2; Naïve
16 Emprical Bayes (NEB) analysis and Bayes Empirical Bayes (BEB) analysis in Figures 4B and
17 4E; F test to compare variance in Figure S2C. The exact value of n, what n represents, and
18 definition of center can be found in the figure legends for each experiment. Unpaired t test,
19 Mann-Whitney U test, and F test were performed using GraphPad Prism; chi square tests were
20 performed using Excel; NEB and BEB analyses were performed using PAML model 2 and 8. *P*
21 value of less than 0.05 was judged as statistically significant.
22
23 DATA AND CODE AVAILABILITY
24
25 The draft genomes and raw sequencing reads have been submitted to the NCBI BioProject
26 database (https://www.ncbi.nlm.nih.gov/bioproject) under accession number PRJNA669840. In-
27 house scripts and pipelines are available from the authors upon request. Raw imaging data is
28 available from the authors upon request.
29
30 AUTHOR CONTRIBUTIONS
31
32 Conceptualization, T.K., M.T.L. and M.A.L.; Methodology, T.K.; Software, T.K., D.S. and
33 E.C.N.; Investigation, T.K., J.M. and R.B.A.; Writing – Original Draft, T.K.; Writing – Review
34 & Editing, T.K., M.T.L. and M.A.L.; Funding Acquisition, T.K., J.K., M.L.T. and M.A.L.;
35 Resources, J.K. and M.A.L; Supervision, M.A.L.
36
37 DECLARATION OF INTERESTS
38
39 The authors declare no competing interests.

References

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. Dev 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. J Cell Biol 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.e10.

Allshire, R.C., and Madhani, H.D. (2017). Ten principles of heterochromatin formation and function. Nat Rev Mol 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. Curr Biol 29, 2625-2639.e5.

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., and Choo, K.H.A. (2004). Human centromere repositioning "in progress." Proc National Acad Sci 101, 6542–6547.

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. Hum Reprod 26, 1575–1584.

Casola, C., Hucks, D., and Feschotte, C. (2007). Convergent Domestication of pogo-like Transposases into Centromere-Binding Proteins in Fission Yeast and Mammals. Mol Biol Evol 25, 29–41.

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. Dev Cell 48, 864-872.e7.

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. Curr Biology Cb 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. Curr Biology Cb 25, 1835–1841.

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. Curr Biology Cb 30, 561-572.e10.

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). Am J Med Genet 111, 450–452.

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. Nat Rev Genetics 17, 109–121.

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. Nat Cell Biol 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. Dev Cell 33, 314–327.

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. Mol Biol Evol 32, 2694–2706.

Finseth, F.R., Nelson, T.C., and Fishman, L. (2020). Selfish chromosomal drive shapes recent centromeric histone evolution in monkeyflowers.

Fishman, L., and Saunders, A. (2008). Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. Sci New York N Y 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. Dev Cell 12, 17–30.

Gamba, R., and Fachinetti, D. (2020). From evolution to function: Two sides of the same CENP-B coin? Exp Cell Res 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-Uk 11, 25.

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.

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

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.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA. Science 293, 1098–1102.

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 Biol Evol 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. J Cell Biol 141, 309–319.

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

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. Curr Biology Cb 27, 2365-2373.e8.

Jangam, D., Feschotte, C., and Betrán, E. (2017). Transposable Element Domestication As an Adaptation to Evolutionary Conflicts. Trends Genet 33, 817–831.

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.e5.

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.

35

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. Sci New York N Y 340, 1110–1113.

Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 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 Res 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 Genet 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. J Cell Biology 210, 11–22.

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. Curr Biology Cb 24, 29–39.

Lampson, M.A., and Black, B.E. (2017). Cellular and Molecular Mechanisms of Centromere Drive. Cold Spring Harb Sym 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.

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.e19.

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 Genet 11, e1004970.

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

36

Malik, H.S., and Henikoff, S. (2009). Major Evolutionary Transitions in Centromere Complexity. Cell 138, 1067–1082.

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. J Cell Biology 109, 1963–1973.

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 Biol 14, R10.

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. J Cell Biology 194, 855–871.

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. Gene Dev 16, 1766–1778.

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. Nat Cell Biol 12, 719–727.

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. Dev 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. J Cell Sci jcs.243303.

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 Carcinog Mutagen 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 Genet 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. Dev Biol 201, 135–143.

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. Mol Cell 71, 923-939.e10.

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. Mol Cell 53, 591–605.

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.e15.

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. Curr Biol 25, 671–677.

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. Curr Genet 64, 25–42.

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. Mol Biol Evol 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. Hum Genet 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. J Cell Sci 128, 4572–4587.

Sironi, M., Cagliani, R., Forni, D., and Clerici, M. (2015). Evolutionary insights into host-pathogen interactions from mammalian sequence data. Nat Rev 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 Res 5, 132–141.

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. Mutat Res 755, 73–80.

Talbert, P.B., Bryson, T.D., and Henikoff, S. (2004). Adaptive evolution of centromere proteins in plants and animals. J 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 theMus caroliandMus paharigenomes. Genome Res 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 Biol Evol 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.

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 Res 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 kinetochoreassembly 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.

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.

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.

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

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, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol 24, 1586–1591.

Yang, P., Wang, Y., and Macfarlan, T.S. (2017). The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. Trends Genetics Tig 33, 871–881.

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

Zhao, G., Oztan, A., Ye, Y., and Schwarz, T.L. (2019). Kinetochore Proteins Have a Post-Mitotic Function in Neurodevelopment. Dev Cell 48, 873-882.e4.