Epigenetic, genetic and maternal effects enable stable centromere inheritance

Arunika Das1,2,3,4, Aiko Iwata-Otsubo2,5, Aspasia Destouni1,6, Jennine M. Dawicki-McKenna1, Katelyn G. Boese2, Ben E. Black1,3,4✉ and Michael A. Lampson2,3✉

Centromeres are defined epigenetically by the histone H3 variant CENP-A. The propagation cycle by which pre-existing CENP-A nucleosomes serve as templates for nascent assembly predicts the epigenetic memory of weakened centromeres. Using a mouse model with reduced levels of CENP-A nucleosomes, we find that an embryonic plastic phase precedes epigenetic memory through development. During this phase, nascent CENP-A nucleosome assembly depends on the maternal Cenpa genotype rather than the pre-existing template. Weakened centromeres are thus limited to a single generation, and parental epigenetic differences are eliminated by equal assembly on maternal and paternal centromeres. These differences persist, however, when the underlying DNA of parental centromeres differs in repeat abundance, as assembly during the plastic phase also depends on sufficient repetitive centromere DNA. With contributions of centromere DNA and the Cenpa maternal effect, we propose that centromere inheritance naturally minimizes fitness costs associated with weakened centromeres or epigenetic differences between parents.

Results

Weakened centromeres persist in the male germline and soma.

To test for epigenetic memory of weakened centromeres with reduced CENP-A chromatin in mammals, we generated heterozygous (H) Cenpa+/− mice (Methods). CENP-A chromatin is reduced to 43.0 ± 0.019%, 48.6 ± 0.003% and 53.8 ± 0.004% (mean ± s.e.m.) of control levels in the soma, male gametes and female gametes, respectively, from these animals (Fig. 1a−c and Extended Data Fig. 1, P0 generation). This model system allows us to test two predictions of epigenetic memory between generations. First, weakened centromeres inherited from the gametes should persist in genetically wild-type (WT) animals. In a cross between two Cenpa+/− parents (H×H), Cenpa+/− progeny should maintain reduced CENP-A chromatin (Fig. 1b, F1 generation). Second, memory should be centromere-autonomous, with each centromere remembering its own level, so that inherited differences persist through development. In a cross between a Cenpa+/− mother and a Cenpa+/− father (WT♀×HT♂), Cenpa+/− progeny should maintain a large epigenetic difference between the maternal and paternal centromeres.

For the first prediction, we find reduced CENP-A levels (72.7 ± 0.005%) at centromeres in the male germline of Cenpa+/− progeny of Cenpa+/− parents, relative to controls with WT parents (Fig. 1c,d and Extended Data Fig. 2). Thus, weakened centromeres persist through development of the male germline in the next generation, consistent with epigenetic memory, although the partial recovery suggests that the memory is incomplete. In contrast, the female germline nearly completely recovers centromere chromatin (94.7 ± 0.008% of controls with WT parents), indicating loss of epigenetic memory in one generation (Fig. 1c,d and Extended Data Fig. 3a). This unexpected dichotomy is underscored by analysis of male and female littermates showing differential recovery of centromere chromatin in their germlines (Extended Data Fig. 3b). Our results raise the question of whether weakened centromeres persist in somatic tissues. Using bone marrow as a representative tissue, we find reduced CENP-A levels in both male and female soma, like the male germline (69.7 ± 0.012%, Fig. 2 and Extended Data Fig. 3a). These results are consistent with epigenetic centromere
memory between generations and through mouse development, but the female germline recovers normal CENP-A chromatin levels.

**Zygotic centromere differences are not maintained in adults.** Before testing the second prediction (Fig. 3a), we noted that zygotes from WT × WT crosses exhibit lower CENP-A levels on paternal centromeres identified by the absence of H3K9me3\(^{21–23}\) (paternal/maternal ratio = 0.5; Fig. 3b,c). Mammalian CENP-A nucleosomes are retained robustly in sperm\(^{14–26}\) relative to canonical nucleosomes, which are largely replaced by protamines, and indeed we find no measurable loss of CENP-A nucleosomes during the histone-to-protamine exchange in spermiogenesis (Fig. 3d,e and Extended Data Fig. 4). The difference between maternal and paternal centromeres in the zygote could reflect either some loss of CENP-A nucleosomes during the protamine-to-histone exchange in the zygote and/or excess loading in the oocyte consistent with the recovery observed in the Cenpa\(^{+/−}\) progeny of Cenpa\(^{+/−}\) parents (Fig. 1). As we anticipated, the difference between maternal and paternal centromeres is enhanced (ratio = 0.4) in F1 zygotes from the WT ♀ × H ♂ cross compared to the WT × WT cross (Fig. 3b,c). This result does not depend on the zygotic genotype, because the zygotic genome is not transcribed at this stage. To determine whether this zygotic difference is maintained in adults, as predicted by centromere-autonomous epigenetic memory\(^{12,15}\),
we analysed meiotic bivalents containing one centromere inherited from each parent. In both the female and male germlines, we find that the ratio between the centromeres of paired homologous chromosomes is indistinguishable from controls (Fig. 3f,g), indicating that initial zygotic differences are not maintained. These results from the WT♀×H♂ cross are inconsistent with epigenetic centromere memory, in contrast to the results from the H×H cross.

Centromere strength depends on the maternal Cenpa genotype. The conflicting results from our H×H and WT♀×H♂ crosses suggest that the weakened centromere state in the progeny might reflect a reduced maternal pool of Cenpa gene products, rather than the number of CENP-A nucleosomes inherited in the gametes. To investigate this possibility, we compared reciprocal crosses in which either parent is heterozygous and the other is WT (Fig. 4a) to our original H×H cross. We find that the maternal genotype is key: if the mother is heterozygous, the weakened centromere state persists in the male germline of the F1 progeny, regardless of whether the father has weakened centromeres (72.7±0.005% versus 78.4±0.008%, both relative to control; Fig. 4b–d). Conversely, if the mother is WT, then weakened paternal centromeres completely recover in the male germline of the F1 progeny (104.1±0.024% relative to control; Fig. 4c,d). Consistent with this result, we find that the maternal but not the paternal Cenpa+/− heterozygous genotype has functional consequences for reproductive fitness, with reduced litter size only when the mother is heterozygous (Fig. 4e).

Given that the persistence of weakened centromeres depends on the maternal genotype, we predicted that an epigenetically weakened centromere state in a WT genotype can only last for a single generation. To test this prediction, we crossed Cenpa+/+×Soma to Oocyte.
from spermatocyte centromeres of F1 adult animals (Fig. 6a), which are expected to be unimodal. These results indicate that the first two cell cycles after fertilization represent a phase of plasticity when CENP-A nucleosomes rapidly equalize between parental centromeres to levels determined by the maternal genotype.

A genetic pathway equalizes centromeres in embryos. We next considered how nascent centromere chromatin assembly could be equal on maternal and paternal centromeres, as in our model (Extended Data Fig. 5), despite initial differences in centromere chromatin. Epigenetic memory depends on existing CENP-A nucleosomes directing nascent assembly by binding CENP-C, the centromere component that recruits downstream assembly factors, including the Mis18 complex and the dedicated CENP-A chaperone, HJURP[27,28]. However, we suspected that a genetic contribution might be more important than the epigenetic pathway during the early embryonic cell cycles. The centromeres in all the animals used in our crosses (Figs. 1–5 and 6a) have an identical genetic makeup, with an excess of minor satellite sequences present at each centromere relative to the number of CENP-A nucleosomes. Minor satellite monomer units (120bp) house a preferred assembly site for CENP-A nucleosomes, as well as the binding element (CENP-B box) for the sequence-specific DNA-binding protein, CENP-B[29]. Given that CENP-B contributes to CENP-C recruitment to centromeres[30,31], we predicted that CENP-C might be sensitive to minor satellite DNA rather than to epigenetic differences between paternal and maternal centromeres. Indeed, in zygotes from the WT × WT cross, CENP-C is only slightly different between the paternal and maternal centromeres ( paternal/maternal ratio = 0.80 ± 0.04 for CENP-C versus 0.51 ± 0.04 for CENP-A; Fig. 6b,f). Furthermore, increasing the epigenetic differences in the WT × HT cross has little effect on CENP-C (paternal/maternal ratio = 0.73 ± 0.04 for CENP-C versus 0.38 ± 0.03 for CENP-A; Fig. 6c,f). These findings suggest that the genetic pathway directs CENP-C recruitment and centromere chromatin assembly in the early embryo, leading to epigenetic equalization when centromeres are genetically identical.

We took two approaches to test this hypothesis. First, we took advantage of natural variation between mouse strains to restrict the genetic contribution by reducing the number of minor satellite repeats. The CHPO strain harbours tiny centromere arrays (sixfold to tenfold smaller than C57BL/6) that restrict both CENP-A nucleosome assembly and CENP-B boxes[4]. Owing to these genetic differences, we predicted larger CENP-C differences between paternal and maternal centromeres in zygotes from a WT × CHPO♂ cross compared to our previous WT × WT or WT × HT crosses. Indeed, the CENP-C ratio is significantly reduced (0.58 ± 0.03) in WT × CHPO♂ zygotes relative to the previous crosses, indicating an increase in CENP-C difference between the two parents (Fig. 6d,f). Moreover, this initial difference in CENP-A nucleosomes and CENP-C between maternal and paternal centromeres in

---

**Fig. 3 | Epigenetic differences between parental centromeres are not maintained.** a, Mating scheme to create epigenetic differences between maternal and paternal centromeres in F1. b, Zygotes (one-cell embryos) from WT × WT (control) and WT × HT crosses. Each pair of CENP-A foci represents sister centromeres in mitosis. Insets show 1.5x magnified maternal and paternal centromeres distinguished by H3K9me3. c, Quantifications of maternal and paternal CENP-A intensities are shown in zygotes combined from two independent experiments with ratios designated for each cross; N = 89, 90, 145 and 143 centromeres (top to bottom). The balance symbols indicate the extent of epigenetic differences between parental centromeres. d, Representative images of spermatocyte pachytene (prophase of meiosis I, 4n) and an elongating spermatid (after completing meiosis II and histone-proteamine exchange[30]) in control (Cenpa−/−) animals. e, Quantification of total CENP-A levels per cell; N = 26 spermatocytes or 35 spermatids. The observed reduction to 25% in spermatids (In) compared with prophase I spermatocytes (4n) is expected if there is no loss during the histone-proteamine exchange. f, Diplole spermatoocyte spreads and metaphase I oocytes in F1. During diplole, centromeres of paired homologous chromosomes (marked with SYCP3 in red) can be resolved. Each inset shows a pair of homologous chromosomes (bivalent), and each of the CENP-A foci represents two sister centromeres. g, Quantification of the ratio of CENP-A foci intensities across a meiotic bivalent (brighter/dimmer) in male and female gametes from F1 males and females with weak centromeres, obtained from the H × H cross, to generate F2 animals (Fig. 5a). We find, in line with our expectations, that centromere chromatin almost completely recovers to control levels in the male germline of F2 animals (93.9 ± 0.01%; Fig. 5b,c).

The importance of the maternal genotype suggests that centromere strength is determined during the early embryonic cell cycles, before zygotic genome activation (ZGA), when nascent centromere chromatin assembly would depend on maternally provided protein and/or a pool of messenger RNA (mRNA) rather than the zygotic genotype. Indeed, our experimental results are consistent with simple modelling of this process in the first two embryonic cell cycles, based on three assumptions (Extended Data Fig. 5). First, new assembly is reduced by 50% in early embryos with heterozygous mothers due to the reduced maternal contribution. Second, assembly is equal on maternal and paternal centromeres. Third, weakened centromeres persist by epigenetic memory after the first two cell cycles, even after activation of a WT zygotic genome. This model captures both partial restoration of weakened centromere chromatin (Figs. 1c and 4c) and equalization of the initial differences between the maternal and paternal centromeres (Fig. 3). At the molecular level, mouse oocytes do not harbour a large pool of CENP-A protein[32], although a small pool may exist and suffice. However, we find that the Cenpa 3′ untranslated region (UTR) has hallmark sequences of a dormant maternal mRNA (Extended Data Fig. 6): a class of transcripts stored in a full-grown oocyte and transcribed in the embryo to support cellular functions prior to ZGA[33]. Previous microarray data[34] show an increase in Cenpa transcripts containing long poly(A) tails when oocytes transition to one-cell embryo, consistent with recruitment of dormant maternal mRNAs after fertilization. Thus, we define Cenpa as a maternal effect gene, as the maternal contribution determines centromere strength.

Centromeres equalize in early embryogenesis. To test our model prediction that centromeres equalize within the early embryonic cell cycles (Extended Data Fig. 5), we examined four-cell embryos from the WT × WT (control) cross and the WT × HT cross, which maximizes the difference between maternal and paternal centromeres (Fig. 3c). By this stage, paternal chromosomes have gained more CENP-A compared to the WT × WT cross ( Extended Data Fig. 7), and other major chromatin rearrangements have occurred, including broad decoration of chromosomes with nucleosomes harbouring the histone H3.3 variant[35]. In the absence of a cytological marker for paternal versus maternal centromeres at the four-cell stage, we analysed the distributions of CENP-A intensities to determine whether or not two populations of centromeres (low and high CENP-A levels) persist. In one-cell zygotes, we find bimodal distributions of the pooled maternal and paternal centromeres (Fig. 6a and Extended Data Fig. 8), consistent with recruitment of dormant maternal mRNAs after fertilization. Thus, we define Cenpa as a maternal effect gene, as the maternal contribution determines centromere strength.
**WT**♀ × CHPO♂ zygotes is maintained in the adult, leading to asymmetric bivalents that show biased segregation in meiosis. These results indicate that nascent assembly of CENP-A nucleosomes depends on the genetic pathway during the plastic phase, such that centromere chromatin equalizes only when genetically identical.

As a second approach, we eliminated the CENP-B-dependent genetic pathway by crossing *Cenpb*−/− knockout females to WT males to generate zygotes lacking a maternal pool of CENP-B protein. Our equalization model predicts that a potential epigenetic contribution to CENP-C recruitment is masked by the genetic pathway, which is symmetric when maternal and paternal centromeres are genetically identical. In the absence of the genetic pathway, CENP-C asymmetry between maternal and paternal centromeres would increase due to the initial epigenetic asymmetry (Fig. 6c,f).
Fig. 4 | Centromere strength depends on maternally inherited CENP-A. a, Mating scheme to test the Cenpa maternal contribution. b, Prophase I spermatocytes from control and F1 progeny of the H♀ × WT♂ cross. Each of the CENP-A foci represents four centromeres from a pair of homologous chromosomes. Scale bars, 5 μm (main panel) and 1 μm (inset). c, Quantifications of CENP-A foci intensities in F1 spermatocytes from the indicated crosses. Representative images are shown in Fig. 1d († indicates H × H, data replotted for comparison from Fig. 1c), Fig. 3f (WT♀ × H♂) and b (H♀ × WT♂). N = 536, 1,836†, 267 and 604 centromeres (top to bottom). d, Data from c replotted by averaging over all centromeres from spermatocytes in each animal, normalized to controls (dashed line). N = 10, 8 and 6 animals (top to bottom). *P < 0.05; NS, P > 0.05; Wilcoxon signed rank test (two-tailed). e, Litter sizes from the indicated crosses. N = 21, 26, 12 and 12 litters (top to bottom). Mean ± s.e.m. values for each cross are shown next to the graph. **P < 0.0001; NS, P > 0.05; Mann–Whitney U test (two-tailed). Error bars: median ± 95% CI.

Fig. 5 | CENP-A chromatin recovers in adult male F2 progeny from Cenpa<sup>+/–</sup> WT F1 parents. a, Mating scheme to generate the F2 generation from F1 with epigenetically weakened centromeres and WT genotype. b, Prophase I spermatocytes from control and F2 males. Each of the CENP-A foci represents four centromeres from a pair of homologous chromosomes. Scale bars, 5 μm (main panel) and 1 μm (inset). c, Quantification of CENP-A foci intensities. N = 276 (control) and 328 (F2) centromeres. NS, P > 0.05; Mann–Whitney U test (two-tailed). Error bars, median ± 95% CI.
Increased epigenetic difference

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Paternal/maternal ratios:

WT × WT
WT × H
WT × CHPO
Cenpb−/− × WT
Cenpb−/− (KO) × WT

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased epigenetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.

CENP-A
CENP-B
CENP-C

Increased genetic difference

Genetic pathway eliminated, no CENP-B

Minor satellite arrays: note that only the small portion of minor satellite containing CENP-A nucleosomes is drawn, except in CHPO, where CENP-A nucleosomes extend the entire span of the minor satellite arrays.
This epigenetic asymmetry is present in the Cenpb−/−♂×WT♀ cross, although reduced relative to the WT×WT control cross (CENP-A ratio = 0.7 ± 0.02) because CENP-A chromatin is reduced in oocytes from Cenpb−/− females42. Despite this decrease in epigenetic asymmetry, the asymmetry in CENP-C recruitment increased in the Cenpb−/−♂×WT♀ cross relative to the control (CENP-C ratio = 0.66 ± 0.01; Fig. 6e,f and Extended Data Fig. 9). This result demonstrates that equalization depends on CENP-B. Summarizing the results of our two experiments manipulating the genetic pathway, we created a genetic asymmetry in the WT♀×CHPO♂ cross, which increased CENP-C asymmetry relative to the WT×WT control due to the genetic pathway. In contrast, we eliminated the genetic pathway in the Cenpb−/−♂×WT♀ cross, which increased CENP-C asymmetry relative to the control by unmasking the epigenetic pathway.

At the molecular level, a parsimonious explanation for epigenetic and genetic contributions to the results of the four crosses we performed (Fig. 6b–e) would involve two distinct pools of CENP-A nucleosomes: one associated with CENP-B and the other without CENP-B. If CENP-A nucleosomes are present in excess of CENP-C, and CENP-C preferentially binds the subset of CENP-A nucleosomes that are also bound to CENP-B43, then CENP-C recruitment would be dictated by CENP-B (that is, the genetic pathway) (Fig. 6g(i)). Partial reduction of CENP-A nucleosomes on the paternal centromeres, as in our WT♀×Hd♂ cross, would not affect CENP-C recruitment as long as the remaining CENP-A nucleosomes bind CENP-B and are sufficient to bind the available CENP-C (Fig. 6g(ii)). Limiting CENP-B binding to paternal centromeres, as in our WT♀×CHPO♂ cross, increases CENP-C asymmetry because there are fewer paternal CENP-A nucleosomes associated with CENP-B (Fig. 6g(iii)). Finally, in the absence of CENP-B, CENP-C recruitment becomes a simple pairwise interaction with CENP-A, so CENP-C scales relative to the number of CENP-A nucleosomes (Fig. 6g(iv)).

**Discussion**

Together, our findings support a model in which centromere strength is initially determined during a phase of early embryonic plasticity. After the plastic phase, weakened centromeres persist in somatic tissue and the male germline (Fig. 6h), even in genetically WT animals (for example, the Cenpa−/− progeny of Cenpa−/− mothers). We thus provide evidence for epigenetic memory through development as predicted by the established mechanism for centromere propagation in somatic cells. By contrast, our in vivo model uncovers a different paradigm of transmission between generations, with Cenpa acting as a maternal effect gene to determine centromere strength.

We show that nascent centromere chromatin assembly in the first embryonic cell cycles depends on maternally provided CENP-A rather than the number of pre-existing CENP-A nucleosomes in the gametes, resetting CENP-A chromatin at the same time that reprogramming occurs for other epigenetic information in the embryo44–46. This maternal effect process suggests a different form of epigenetic memory for transmission of a weakened centromere state to offspring through the female germline. In nature, we envision that, like all genes, Cenpa expression could vary substantially between individuals through epigenetic effects such as differences in promoter methylation. Mothers with attenuated Cenpa expression would therefore transmit weakened centromeres to offspring because of the reduced maternal contribution, even with unattenuated Cenpa expression in the offspring. This maternal effect process limits memory to a single generation, however, and also eliminates epigenetic differences between maternal and paternal centromeres in the embryo. By contrast, epigenetic memory in flies is a paternal effect47, and a genetic contribution to centromere inheritance through sequence-specific DNA-binding proteins is unlikely given that there does not appear to be a counterpart to CENP-B in flies.

We also find that weakened centromeres recover in the female germline, possibly as a mechanism to protect against loss of centromere identity during the prolonged mammalian oocyte prophase arrest, as CENP-A nucleosomes assembled before this arrest last through the reproductive lifespan of the animal48. Recovery may also provide a buffer from potential failure in telomere bouquet protection of centromeres in female meiosis49. By contrast, CENP-A is removed early in female meiosis in holocentric worms50, so de novo assembly is required to re-establish centromere chromatin. Similarly, in worms, Cenpa mutants that disrupt interactions with the assembly machinery are maternal effect lethal as they abort the de novo assembly51.

Epigenetic specification of centromeres may have evolved as a strategy to suppress fitness costs associated with selfish centromere DNA sequences that subvert female meiosis52 (drive) to increase their transmission to the egg. Epigenetic centromeres require a propagation mechanism, which can impose its own costs, however. If pre-existing CENP-A nucleosomes recruit the machinery for nascent assembly, then epigenetic differences between maternal and paternal centromeres in the zygote can lead to differential assembly. Indeed, epigenetic differences in plants cause embryonic aneuploidy due to loss of weaker centromeres or even complete
elimination of one parental genome\textsuperscript{[18,46]}. Our finding of a specialized early embryonic assembly process, directed by centromere DNA sequence rather than pre-existing CENP-A nucleosomes, reveals a mechanism to equalize centromeres to protect against the detrimental consequence of epigenetic asymmetry between the parental genomes. We propose that dual genetic and epigenetic contributions to centromere chromatin assembly represent adaptations to fitness costs arising from either selfish DNA sequences or parental epigenetic asymmetry.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-022-00897-w.

Received: 7 June 2021; Accepted: 16 March 2022; Published online: 9 May 2022

References

1. Xixmoedier, K., Allu, B. K. & Black, B. E. The centromere comes into focus: from CENP-A nucleosomes to kinetochore connections with the spindle. Open Biol. 10, 200051 (2020).

2. Dumont, M. & Fachinetti, D. DNA sequences in centromere formation and function. Prog. Mol. Subcell. Biol. 56, 305–336 (2017).

3. Chimátil, L., Schultz, R. M., Black, B. E. & Lampson, M. A. Cell biology of cheating transmission of centromeres and other selfish elements through asymmetric meiosis. Prog. Mol. Subcell. Biol. 56, 377–396 (2017).

4. Iwata-Otsubo, A. et al. Expanded satellite repeats amplify a discrete CENP-A nucleosome assembly factor sufficient to form a functional de novo kinetochore. J. Cell Biol. 194, 229–243 (2011).

5. Foltz, D. R. et al. HJURP is a CENP-A chromatin assembly factor necessary for centromere maintenance and the fidelity of human centromere function. Dev. Cell 33, 314–327 (2015).

6. Voullaire, L. E., Slater, H. R., Petrovic, V. & Choo, K. H. A. A functional marker sequence rather than pre-existing CENP-A nucleosomes, reveals a mechanism to equalize centromeres to protect against the detrimental consequence of epigenetic asymmetry between the parental genomes. We propose that dual genetic and epigenetic contributions to centromere chromatin assembly represent adaptations to fitness costs arising from either selfish DNA sequences or parental epigenetic asymmetry.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022.
Methods

Animal husbandry and generation of Cenpa−/− heterozygous and Cenpb−/− knockout mice. All animal experiments and protocols were approved by the Institutional Animal Use and Care Committee of the University of Pennsylvania and were consistent with the National Institutes of Health guidelines (protocol #083994). All animals used in this study were within six months of age, and both male and female animals were analysed. Experimental animals were compared to age- and gender-matched controls. Cenpa−/− heterozygous (H) mice were initially generated by mating Cenpa−/−CdflCre/+ conditional Cenpa knockout females with WT males ([C57BL/6]). Jackson Laboratory (006664) and subsequently regrown through either WT × H2, H × H or H2 × WT (crosses), which also generated the experimental Cenpa−/− F1 progeny. F1 control progeny were generated by mating Cenpa−/− females to WT C57BL/6 males. ’CHPO’ males were obtained from Jackson Laboratory (ZALENDE/EiJ, 003392) and then bred in house. For each dataset, at least two to five independent experiments were performed, each having one control and one to two experimental animals that were age- and gender-matched. For embryo collections, five to eight females were mated to five to eight males for each independent experiment. Genotyping for Cenpa was performed using the REDExtract N-AMP kit (Sigma)27 and all animals were sampled twice to confirm their genotype. Cenpa−/− mice were generated through either WT × Cenpa−/− (C57BL/6) females, or WT males with Cenpa−/− females. Each mice was genotyped for age- and gender-matched controls. Genotyping for Cenpa was performed using the REDExtract N-AMP kit (Sigma)27 and all animals were sampled twice to confirm their genotype. Cenpa−/− males were confirmed through either WT × Cenpa−/− (C57BL/6) females, or WT males with Cenpa−/− females. Each mice was genotyped for age- and gender-matched controls.

Microscopy. Confocal images were collected as z-stacks with 0.5-μm intervals, using a microscope (DM4000 B, Leica) equipped with a x63 1.3-NA glycerol-immersion objective lens, an x-y piezo Z stage (Applied Scientific Instrumentation), a spinning disk confocal scanner (Yokogawa Corporation of America), an electron-multiplying charge-coupled device camera (ImageEM C9100-13; Hamamatsu Photonics) and either LMD5 (Spectral Axiom Research) or Versalase (Vortran Laser Technology) laser module equipped, controlled by MetaMorph software (Molecular Devices, v7.10.3.294). Images were acquired using the same laser settings and all images in a panel were scaled the same. Single channels are shown wherever quantifications were performed.

Oocyte collection and culture. Female mice were hormonally primed with 5 μl of pregnant mare serum gonadotropin (PMSG, Peptides International) 44–48 h before oocyte collection. Germinal vesicle intact oocytes were collected in bicarbonate-free minimal essential medium (MEM, Sigma) supplemented with cumulus cells, and cultured in Chato-Ziomek-Bavister (CZB, Fisher Scientific) medium covered with 5% CO2 in humidified air and 37 °C. 1 h later with primary antibodies in blocking solution, washed three times for 15 min, and mounted with Vectashield with DAPI (Vector) on a 24-well cover glass. From each slide, primary spermatocytes at either pachytene stage (overall CENP-A levels) or diplotene stage (bivalent analysis of meiotic prophase I) were selected based on the distinct SYCP3 staining pattern (paired and threadlike in pachytene and X-shaped in diplotene after synaptonemal complex disassembly) and imaged using the confocal microscope described in the Microscopy section.

Bone marrow collection and immunocytochemistry. Bone marrow was collected from the tibiae (b) and femur(s) by inserting a 26-G syringe needle into the cut end of the marrow cavity. Cells were flushed out into 1 ml of warm ethylenediaminetetraacetic acid (EDTA) buffer (8 g sodium chloride, 0.2 g potassium dihydrogenphosphate, 0.2 g potassium chloride, 1.15 g sodium dihydrogenphosphate, 0.2 g EDTA, dissolved in 11 of deionized water) with 0.025% colchicine (Sigma) and incubated for 3 h at 37 °C. Cells were then diluted 50 times in 0.5% potassium chloride solution on ice for 20 min to swell. Stained were subsequently prepared on Superfrost Plus slides using a double funnel on a Cytopsin 4 cytospinette (ThermoFisher) at 600 r.p.m., high acceleration for 5 min, then rinsed briefly in PBS and fixed in 4% formaldehyde solution for 20 min r.t., permeablized in 0.5% Triton X-100/PBS and blocked for 20 min (0.3% BSA, 0.01% Tween-20). The slides were incubated with anti-CENP-A antibody (1:200) for 1 h r.t., washed three times with PBST (PBS/0.01% Tween-20), incubated with Alexa Fluor 488 (1:500, Invitrogen) for 1 h at r.t., washed three times in PBST for 5 min each and once in distilled water, and then mounted in Vectashield with DAPI to visualize chromosomes.

Collection of embryos and in vitro fertilization. C57BL/6 or Cenpa−/− females were hormonally primed with 5 μl of PMSG (Peptides International) and the oocytes were matured in vivo with 5 μl of human chorionic gonadotropin (bCG, Sigma) before mating with either C57BL/6 WT or Cenpa−/− males. Because the Cenpa−/− mice have low mating efficacy they were fed a special low soy diet (SGl4 irradiated diet, Labdiet) before mating. Embryos were collected 14–16 h post hCG in M2 containing hyaluronidase (0.3 mg ml−1) to remove cumulus cells and subsequently washed in M2 (Sigma) and cultured in EmbryoMax Advanced KSOM (Aksam, Millipore Sigma) with humidified air and 5% CO2. Next, 5 μM proTAME (R&D systems) was added at 4 h or ~32–34 h post collection to arrest embryos at one-cell or four-cell mitosis, respectively. Chromosomes were dispersed by generating a monopolar spindle using 10 μM β-tubulin-l-cysteine (STLC; Sigma) for 3 h. For experiments using males with limited efficiency of overnight mating (CPO or Cenpa−/−), we obtained embryos in vitro fertilization (IVF), modified from a previously established protocol.11 Notably, paternal/maternal CENP-A ratios for controls from IVF and in vivo fertilized embryos were comparable. Briefly, sperm from the cauda epididymis were collected from two- to four-month-old males in 500 μl of EmbryoMax human tubal fluid (HTF, Millipore Sigma) and allowed to swim out for 15 min. Sperm were capacitated for 2 h in 2 ml swim up tubes in HTF, before fertilization. Females were primed with PMSG and hCG as described above. MI1 eggs were collected 14–15 h post hCG into M2 and then transferred into a 50-μl fertilization drop of HTF: sperm were added to a final concentration of 100,000 sperm per drop for IVF. Fertilized eggs were washed three times in M2 containing HTF and cultured overnight at 37 °C in humidified air with 5% CO2. Embryos were arrested in mitosis as described above.

Embryo immunocytochemistry. Embryos were fixed in 2% formaldehyde solution in PBS with 0.1% Triton X-100 for 20 min at r.t., placed in blocking solution (PBS containing 0.3% BSA and 0.01% Tween-20) overnight at 4°C. Treated with 10-μm sections, incubated with primary antibody and secondary antibody for three times for 10 min each, incubated for 1 h with secondary antibody, washed three times for 10 min each, and mounted in Vectashield with 4′,6-diamidino-2-phenylindole (DAPI; Vector) to visualize the chromosomes. The primary antibody was rabbit anti-mouse CENP-A (1:2000, Cell Signaling). CENP-A was visualized with donkey anti-rabbit Alexa Fluor 488 (1:500, Invitrogen). Secondary antibodies used were donkey anti-rabbit Alexa Fluor 488 (1:500, Invitrogen) and donkey anti-mouse Alexa Fluor 594 (1:100, Invitrogen) and donkey anti-rabbit Alexa Fluor 488 (1:2000, Invitrogen) secondary antibodies. The slides were then washed three times for 10 min each, with 0.4% PhotoTox/PBS and once with 0.4% PhotoTox/CMC for 10 min and mounted with Vectashield with DAPI (Vector) to a 24-well cover glass. From each slide, primary spermatocytes at either pachytene stage (overall CENP-A levels) or diplotene stage (bivalent analysis of meiotic prophase I) were selected based on the distinct SYCP3 staining pattern (paired and threadlike in pachytene and X-shaped in diplotene after synaptonemal complex disassembly) and imaged using the confocal microscope described in the Microscopy section.
Statistics and reproducibility. All statistical tests for significance were performed in GraphPad Prism 9 or R. P values were calculated at a significance (α) level of 0.05 (95% confidence interval (CI)) and all tests performed were two-tailed. No statistical method was used to predetermine sample size. Randomization is built into our experiments as each animal was chosen from a different litter and mating pair, no data were excluded and all cells were imaged at random. Samples were designated as control or experiment according to their genotyping data. Animals were genotyped twice. Within a genotype, animals were randomly picked. Because our experiment was with a heterozygous single deletion within the genome, we did not have any covariates to consider. Investigators were not blinded for data collection (imaging) and quantifications (data analysis) as the phenotype automatically reports on the genotype consistently and is very penetrant. Similarly, statistical analysis of bimodality did not require blinding because the rotated kernel density (violin) plots being analysed showed obvious bi/unimodality in each cross. Graphs were created with GraphPad Prism 9 or R.

For all quantified experiments, the numbers of replicates (animals or independent experiments) are provided in Supplementary Table 1. Bimodality testing was performed using the R package multimode with the function ‘modetest’ (Fig. 6a) using the excess mass statistic with bootstrapping at a significance (α) level of 0.05 (95% CI). A subset from the F1 adult spermatocyte data was used as a representative control unimodal distribution for comparison (Fig. 6a). P < 0.05 from the test indicates that the distributions are significantly non-unimodal. Using results from ‘modetest’, the location of the modes and the density of each mode per distribution were determined and plotted with the ‘locmodes’ function (Extended Data Fig. 8).

Analysis of the 3′ UTR of Cenpa. The consensus sequences for CPEI and CPEII were found by manual evaluation of 3′ UTR sequences of Cenpa from the NCBI database for annotated transcripts. The multiple sequence alignment (MAFFT V7) of the 12 rodent 3′ UTRs were made and annotated with UGene (Unipro V37).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this Article.

Data availability
Previously published microarray data for long poly-(A) tailed Cenpa mRNA in pre-implantation development is available freely on the NCBI Gene Expression Omnibus (GEO) database (accession no. GDSS13 from reference series GSE1749).

The 12 mouse genomes used for Cenpa 3′ UTR analysis are available from the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject) under accession no. PRJNA669840. Data supporting the findings of this study are available from the corresponding author on reasonable request. Source data are provided with this paper.

Code availability
All codes used for statistical and distribution analysis are freely available as part of the R package ‘multimode’, described in ref. 56.

References
53. Stein, P. & Schindler, K. Mouse oocyte microinjection, maturation and ploidy assessment. J. Vis. Exp. 53, 2851 (2011).
54. Chatot, C. L., Ziomek, C. A., Bovister, B. D., Lewis, J. L. & Torres, I. An improved culture medium supports development of random-bred 1-cell mouse embryos in vitro. J. Reprod. Fertil. 86, 679–688 (1989).
55. Dia, F., Strange, T., Liang, J., Hamilton, J. & Berkowitz, K. M. Preparation of meiotic chromosome spreads from mouse spermatocytes. J. Vis. Exp. 129, 55378 (2017).
56. Taft, R. In vitro fertilization in mice. Cold Spring Harb. Protoc. https://doi.org/10.1101/pdb.prot094588 (2017).
57. R: a Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2017).
58. Ameijeiras-Alonso, J., Crujeiras, R. M. & Rodriguez-Casal, A. multimode: An R Package for Mode Assessment. J. Stat. Soft. 97, 1–32 (2021).

Author contributions
A. Das contributed to experiments, quantifications, data analysis and statistical analysis, animal husbandry and genotyping. A.-I.O. carried out experiments and quantification for some of Fig. 3g. J.D.-M. prepared and characterized new reagents and assisted with statistical analysis. A. Destouni performed the initial experimentation in zygotes and early embryos. K.G.B. carried out animal husbandry and genotyping. A. Das, B.E.B. and M.A.L. contributed to experimental design, data interpretation and writing. B.E.B. and M.A.L. provided supervision and sourced funding.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41556-022-00897-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41556-022-00897-w.

Correspondence and requests for materials should be addressed to Ben E. Black or Michael A. Lampson.

Peer review information Nature Cell Biology thanks Hiroshi Kimura and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
Extended Data Fig. 1 | CENP-A chromatin is reduced in the soma of Cenpa<sup>+/−</sup> heterozygous animals in the P<sub>0</sub> generation. **a**, Bone marrow metaphase spreads: each pair of CENP-A foci represents sister centromeres in mitosis. Scale bars: 5 μm (main panel), 1μm (inset). **b**, Quantification of CENP-A foci intensities in control (grey) and P<sub>0</sub> (yellow) generation in soma. N = 166, 170 centromeres (top to bottom). **P < 0.0001, Mann–Whitney U test (two-tailed). Error bars: median ± 95% CI. Source numerical data are available in source data.
Extended Data Fig. 2 | Weakened centromeres in the male germline are independent of meiotic stage. Because oocytes were analysed at metaphase I and spermatocytes at prophase I (Fig. 1), we confirmed that F1 spermatocytes also show weakened centromeres at metaphase I. Images (a) and quantification (b) of F1 spermatocytes show CENP-A reduced to a similar level at metaphase I (70.54 ± 7.1% of control) as prophase I. Each of the CENP-A foci represents four centromeres (a pair of homologous chromosomes, each with two sisters). N = 330 (control), 284 (F1 progeny). Scale bars: 5 μm (main panel), 1 μm (inset). Quantification of SYCP3 foci from the same cells (c) shows no decrease (114.90 ± 5.6% of control). N = 235 (control), 259 (F1 progeny). ** P < 0.001, Mann–Whitney U Test (two-tailed). Error bars: median ± 95% CI. Source numerical data are available in source data.
Extended Data Fig. 3 | Littermate analysis showing that weakened centromeres persist in the male but not female germline. a, Data from Fig. 1c replotted as CENP-A levels per animal, averaged over all centromeres in each animal and normalized to controls (dashed line). N = 10, 10, 10, 9, 7 animals. The F1 male but not the female germline and the male and female soma are significantly lower than the controls **P < 0.001, *P < 0.05 n.s.: P > 0.05, Wilcoxon signed sum rank test (two-tailed). b, CENP-A quantifications in spermatocytes and oocytes from littermates from one set of parents. N = 121, 431, 60, 259, 246, 105 centromeres (top to bottom). Female germline levels are significantly elevated compared to littermate male germline levels. **P < 0.0001, Mann–Whitney U Test (two-tailed). Error bars: median ± 95% CI. Source numerical data are available in source data.
Extended Data Fig. 4 | CENP-A nucleosomes are retained through the replacement of canonical nucleosomes with protamines during spermiogenesis. a, Quantification and b, images showing CENP-A levels are reduced to 42.7 ± 1.5% in spermatids from Cenpa<sup>−/−</sup> males compared to WT males, similar to the reduction measured in prophase spermatocytes (Fig. 1c). N = 20 (control), 32 (Cenpa<sup>−/−</sup>) spermatids. Error bars: median ± 95% CI. Scale bars: 5 μm (main panel), 1 μm (inset). Source numerical data are available in source data.
**Extended Data Fig. 5** | Model to explain equalization of epigenetic differences and subsequent memory. **a,** Assumptions used for the modelling. **1.** New assembly is induced by 50% in zygotes with heterozygous mothers due to reduced maternal contribution in the first cell cycle. **2.** Maternal and paternal centromeres assemble equal amounts. **3.** Weakened centromeres persist by epigenetic memory after the first two cell cycles even after activation of a wild type zygotic genome (ZGA).

**b** CENP-A nucleosome assembly during the cell cycle. **50% CENP-A** Partitioned between sisters by replication. **50% CENP-A** New assembly restores levels to WT.

**C** Example calculation for projected values in graph for WT x WT:

| CENP-A levels | Initial | Diluted by replication | Reload equal amounts |
|---------------|---------|-------------------------|---------------------|
| G1            | 100     | 50                      |                     |
| S             | 50      | 25                      |                     |
| G1            | 100     | 75                      |                     |
| S             | 50      | 37.5                    |                     |
| G1            | 100     | 87.5                    |                     |

**d** Graphs from similar calculations as b, for the designated crosses. Initial CENP-A levels are set to 50 for maternal centromeres from Cenpa\^+/− mothers and 40 for paternal centromeres from Cenpa\^−/− fathers, based on our measurements (Fig. 1c and Fig. 3c). Arrows indicate equal assembly on maternal and paternal centromeres. In the first cell cycle, assembly is from a maternal pool of 100 (black arrows) or 50 (yellow arrows) for WT or Cenpa\^−/− mothers, respectively. In the second cell cycle, assembly is from a zygotic pool of 100 (purple arrows), reflecting a WT zygotic genotype. Calculations show equalization by the four-cell stage in all crosses. Furthermore, crosses with reduced maternal contribution (H♀) equalize to a lower level, which is then remembered through development. Source numerical data are available in source data.
Extended Data Fig. 6 | 3’ UTR of Cenpa message has hallmarks of dormant maternal mRNA. a, Polyadenylation (addition of a poly (A) tail) of mRNA is a mechanism to control gene expression. Nuclear polyadenylation is an essential part of post-transcriptional processing of most mRNAs, dictated by the ubiquitous cis-element 3’ UTR hexameric motif AATAAA (nuclear polyadenylation element, NPE). Dormant maternal mRNAs are deposited in the oocyte with short poly(A) tails and are translationally inactive. After fertilization, these maternal mRNAs undergo translation by elongation of the poly(A) tail, controlled by a cytoplasmic polyadenylation element (CPE) usually present within 100 nt upstream of the NPE. We find conserved CPEs in the mouse, human and frog Cenpa 3’ UTRs (CPE I = TTTTAT or CPE II = TTTTAA) upstream of the NPE as expected for dormant maternal mRNAs. b, Analysis of 12 sequenced rodent species reveals that CPEs (CPE I in bold boxes and CPE II in dashed boxes) are present upstream of the NPE in every species as expected for a maternal effect gene.
Extended Data Fig. 7 | Symmetric distribution of H3K9me3 at the four-cell stage. Representative cell from four-cell embryos for each of the two denoted crosses with H3K9me3 (red), CENP-A (green) and DNA (blue). H3K9me3 is present on both maternal and paternal chromatin at this stage, in contrast to zygotes (Fig. 3b and Fig. 6b–e). Scale bars: 5 μm.
Extended Data Fig. 8 | See next page for caption.
Extended Data Fig. 8 | CENP-A intensity distribution changes from bimodal to unimodal in early embryogenesis. Graphs show locations of the modes in each distribution from Fig. 6a. a, The WT x WT and WT♀ x H♂ zygote distributions contain two modes (dashed lines) on either side of a central antimode (dip, pointed lines) characteristic of bimodal distributions. The separation between the two modes is greater in the WT♀ x H♂ cross as expected. In addition, the ratios of the values of the two modes (x-axis) denoted under each cross agree well with the ratios of paternal to maternal centromere intensities calculated in Figs. 3c and 6f. b,c, Similar plots of four-cell embryos (b) from the same crosses show a single central mode characteristic of a unimodal population, like the F1 adult spermatocytes (c), which represents a uniform centromere population. The ratio of the modes in bimodal or the value of the mode in unimodal distribution is indicated below the graphs. Source numerical data are available in source data.
Extended Data Fig. 9 | Genetic pathway for centromere equalization. a, Quantifications of maternal (pink) and paternal (blue) CENP-A and CENP-C intensities in zygotes from a WT x WT control for the Cenpb<sup>−/−</sup> strain<sup>38</sup>, with average paternal/maternal CENP-A or CENP-C ratios above; N = 46, 42, 237, 231 centromeres (left to right). Error bars: median ± 95% CI. Although these animals are in a CF-1/C57BL/6J/DBA/2J background, CENP-A and CENP-C ratios in WT zygotes using mothers from this background are consistent with those of C57BL/6J alone (Fig. 6b,f). Source numerical data are available in source data.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

☐ For each experimental group/condition, given as a discrete number and unit of measurement
☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐ The statistical test(s) used AND whether they are one- or two-sided
☐ Only common tests should be described solely by name; describe more complex techniques in the Methods section.
☐ A description of all covariates tested
☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient)
☐ AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
☐ Give P values as exact values whenever suitable.
☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

☐ Data collection: Metamorph (Molecular devices, v7.10.3.294) for image acquisition

Data analysis

☐ Statistical analyses were performed with Prism (GraphPad Software Inc., versions 7.0e, 8.4.3 & 9.0.0) and R (open source, https://www.rproject.org); images were analyzed with Fiji/ImageJ (open source at https://imagej.net/, version 2.1.0/1.53c); MAFFT version 7 and Unipro version 37 was used for multiple sequence alignment.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Previously published microarray data for long poly-(A) tailed Cenpa mRNA in pre-implantation development is available freely on NCBI Gene Expression Omnibus (GEO) database (Accession no: GDS813 from reference series: GSE1749). The 12 mouse genomes used for Cenpa 3'UTR analysis are available at NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject) under accession number PRJNA669840. Source data have been provided in Source Data. All other data supporting the findings of this study are available from the corresponding author on reasonable request.
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size
No statistical method was used to determine sample sizes. Sample sizes for all experiments were determined by the current standard used for mouse model systems in epigenetic experiments, based on the minimal amount of mice required to detect significance with an alpha rate set at .05 in a standardly powered experiment. The number of animals and replicates were determined based on the consistency of the phenotype. In most cases, the phenotype was consistent and penetrant as is obvious from the agreement between multiple independent replicates, validating the sample size is sufficient.

Data exclusions
No data were excluded from analysis.

Replication
All experiments have 2-5 replicates with multiple animals in each totaling at least 4-8 per experiment per genotype. The animals in each generation were produced from multiple mating pairs. For IVF and embryo collections, at least 3 replicates were performed each having at least 5 females mated to male(s) (5-8 for in vivo fertilization and 1 for in vitro fertilization). Importantly, most replicates showed similar trends and all the data is presented in the paper and in source data files. Also, we confirmed that in vivo and in vitro fertilized embryos show similar ratios (each method repeated 2-3 times) validating our results further. Lastly, all the conclusions in this study are drawn from multiple animals sampled from many litters from different mating pairs for each generation. Detailed replicate numbers for all figures are provided in the Supplementary Table 1 and source data files.

Randomization
Randomization is built into our experiments as each animal was chosen from a different litter and mating pair, no data was excluded and all cells were imaged at random. Samples are designated as control or experiment according to their genotyping data. Animals were genotyped twice. Within a genotype animals are randomly picked. Since our experiment is with a heterozygous single deletion within the genome, we do not have any covariates to consider.

Blinding
Investigators were not blinded for data collection (imaging) and quantifications (data analysis) as the phenotype automatically reports on the genotype consistently and is very penetrant. Likewise, statistical analysis of bimodality did not require blinding since the rotated kernel density (violin) plots being analyzed show obvious bi/uni-modality in each cross. Blinding in cases of an obvious phenotype is unnecessary.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| -   | Antibodies            |
| x   | Eukaryotic cell lines |
| x   | Palaeontology and archaeology |
| -   | Animals and other organisms |
| -   | Human research participants |
| -   | Clinical data         |
| x   | Dual use research of concern |

### Methods

| n/a | Involved in the study |
|-----|-----------------------|
| x   | ChIP-seq              |
| x   | Flow cytometry        |
| x   | MRI-based neuroimaging |

#### Antibodies

Primary antibodies used are as follows:
- Rabbit anti-mouse CENP-A: Cell Signaling, 20485, CS1A7, 4, 1:200 (oocytes, embryos, bone marrow), 1:400 (spermatocytes)
- Mouse anti-mouse H3K9me3: Active motif, 39285, 2AG-6F12-H4, 03019002, 1:500 (embryos)
- Rabbit anti-mouse CENP-C (this study, details are in Methods), 1:1000 in embryos.

Secondary antibodies used are:
- Donkey anti-rabbit Alexa Fluor 488: Invitrogen, A21206, 1113537 1:500 (oocytes, embryos, bone marrow), 1:2000 (spermatocyte spreads)
- Donkey anti-mouse Alexa Fluor 594: Invitrogen, A11032, 1: 500
- Mouse anti-mouse SYCP3: Abcam, ab97672, 10G11/7, 1:200
Validation

All the antibodies used are commercially available except for anti-mouse CENP-C and are validated for use in immunocytochemistry.

Validation information is provided in the links below:
- Rabbit anti-mouse CENP-A: Cell Signaling, 2048S, C51A7, see manufacturer’s website for references and validation information (https://www.cellsignal.com/products/primary-antibodies/cenp-ac51a7-rabbit-mab-mouse-specific/2048)
- Mouse anti-mouse H3K9me3: Active motif, 39285, 2AG-6F12-H4, 03019002, see manufacturer’s website for references and validation (https://www.activemotif.com/catalog/details/39285/histone-h3-trimethyl-lys9-antibody-mabclone-2ag-6f12-h4)
- Rabbit anti-mouse CENP-C (this study, details are in Methods). The rabbit anti-mouse CENP-C was validated for this study through both Western blot and immunocytochemistry on NIH3T3 cells (reactivity) and human cell lysates (no reactivity). The antibody recognizes a band at the expected molecular weight of ~130kDa by immunoblot. It recognizes punctate centromeric foci in both NIH3T3 cells and mouse embryos (1:1000 dilution).

Secondary antibodies used are:
- Donkey anti-rabbit Alexa Fluor 488: Invitrogen, A21206, 1113537, see manufacturer’s website for references and validation (https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21206)
- Donkey anti-mouse Alexa Fluor 594: Invitrogen, A11032, see manufacturer’s website for references and validation (https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11032)
- Mouse anti-mouse SYCP3: Abcam, ab97672, 10G11/7, see manufacturer’s website for references and validation (https://www.abcam.com/scp3-antibody-cor-10g117-ab97672.html)

Animals and other organisms

Policy information about studies involving animals: ARRIVE guidelines recommended for reporting animal research.

| Laboratory animals | 
|---------------------|
| Mouse strains used in this study are as follows: |
| - CenpaFl/Fl;Gdf9Cre/+ females (described in reference 28) |
| - Cenpa+/- heterozygous males and females (this study, see Methods) |
| - CS7B16/I males and females (Jackson Laboratory, 000664) |
| - ZALENDE/EiJ (CHPO) males (Jackson laboratory, 001392) |
| - Cenpb-/- knockout females (described and validated in reference 38) |
| All data was collected from animals that were reproductively mature (at least 6-8 weeks of age) and less than 6-month-old. Both male and female animals were used wherever possible. |
| Mouse colony breeding conditions: |
| Light/dark cycle: 12 hours each |
| Bred at room temperature with minimal disturbance with a range of 30-70% humidity depending on the season. |
| All animal experiments and protocols were approved by the Institutional Animal Use and Care Committee of the University of Pennsylvania and were consistent with National Institutes of Health guidelines (protocol #803994) |

| Wild animals | 
|---------------|
| No wild animals were used in this study. |

| Field-collected samples | 
|-------------------------|
| There were no field samples in this study. |

| Ethics oversight | 
|------------------|
| All protocols were overseen by Institutional Animal Use and Care Committee of the University of Pennsylvania. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.