Article

Satellite DNA shapes dictate pericentromere packaging in female meiosis

https://doi.org/10.1038/s41586-024-08374-0

Received: 28 February 2024

Accepted: 8 November 2024

Published online: 08 January 2025

Check for updates

Damian Dudka[1], Jennine M. Dawicki-McKenna[2,6], Xueqi Sun[1,6], Keagan Beeravolu[1], Takashi Akera[3], Michael A. Lampson[1,4] ✉ & Ben E. Black[2,4,5] ✉

The abundance and sequence of satellite DNA at and around centromeres is evolving rapidly despite the highly conserved and essential process through which the centromere directs chromosome inheritance[1–3]. The impact of such rapid evolution is unclear. Here we find that sequence-dependent DNA shape dictates packaging of pericentromeric satellites in female meiosis through a conserved DNA-shape-recognizing chromatin architectural protein, high mobility group AT-hook 1 (HMGA1)[4,5]. Pericentromeric heterochromatin in two closely related mouse species, *M. musculus* and *M. spretus*, forms on divergent satellites that differ by both density of narrow DNA minor grooves and HMGA1 recruitment. HMGA1 binds preferentially to *M. musculus* satellites, and depletion in *M. musculus* oocytes causes massive stretching of pericentromeric satellites, disruption of kinetochore organization and delays in bipolar spindle assembly. In *M. musculus* × *spretus* hybrid oocytes, HMGA1 depletion disproportionately impairs *M. musculus* pericentromeres and microtubule attachment to their kinetochores. Thus, DNA shape affects both pericentromere packaging and the segregation machinery. We propose that rapid evolution of centromere and pericentromere DNA does not disrupt these essential processes when the satellites adopt DNA shapes recognized by conserved architectural proteins (such as HMGA1). By packaging these satellites, architectural proteins become part of the centromeric and pericentromeric chromatin, suggesting an evolutionary strategy that lowers the cost of megabase-scale satellite expansion.

Chromatin organization at centromeric regions ensures faithful segregation by spindle microtubules attached to centromere-nucleated kinetochore complexes[6]. Pericentromeric heterochromatin promotes centromere formation[7–9], clusters centromeres in the nucleus[10] and regulates kinetochore–microtubule attachments during cell division[11–13]. Despite these conserved functions, centromeres and pericentromeres typically form on highly repetitive satellite DNA, which evolves rapidly in both sequence and abundance[14,15]. Moreover, either of these chromatin domains can form on different sequences in the same cell[16–18] and functional centromeres can form de novo (neocentromeres) on non-repetitive DNA[19]. These observations suggest that the organization and function of centromeres and pericentromeres is independent of any particular DNA sequence. By contrast, the well-studied example of the house mouse (*Mus musculus*) shows that each domain forms on a dedicated satellite, suggesting that the DNA sequence contributes. Centromeres form on the minor satellite, whereas pericentromeres form on the more abundant major satellite[20,21] (Fig. 1a). The closely related Algerian mouse (*Mus spretus*) has the same two satellites, but the relative abundance is reversed[22–24] (Fig. 1a). Whether pericentromeres form on the major satellite in *M. spretus* as in *M. musculus*, or instead on the more abundant minor satellite, has not been determined. Analyses of *M. musculus* × *spretus* hybrid oocytes reveal differences in chromatin packaging, however, with reduced condensin II localization and greater decondensation at *M. musculus* centromeres and pericentromeres compared with the homologous *M. spretus* chromosomes in the same cell[25,26]. These observations suggest that the two different satellites are packaged by distinct mechanisms, consistent with a dependence on the DNA sequence. Thus, with conflicting lines of evidence, it remains unclear whether and how the DNA sequence contributes to centromere and pericentromere organization and packaging. To address these questions, *M. musculus* and *M. spretus* provide a natural model system to determine the molecular principles that underlie chromatin organization and packaging on different satellites.

Heterochromatin follows repeat abundance

We measured the degree to which the abundance of the minor and major satellite is reversed in the two species using liver samples: the minor satellite accounts for 4.4% of the *M. spretus* but only 0.5% of the *M. musculus* input DNA, whereas the major satellite occupies 7.9% of the *M. musculus* but only 0.2% of the *M. spretus* input DNA (Extended Data Fig. 1a–d and Supplementary Table 1). To test whether satellite identities affect centromere and pericentromere formation, we analysed the underlying DNA sequences using chromatin immunoprecipitation

[1]Department of Biology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [3]Cell and Developmental Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA. [4]Penn Center for Genome Integrity, University of Pennsylvania, Philadelphia, PA, USA. [5]Epigenetics Institute, University of Pennsylvania, Philadelphia, PA, USA. [6]These authors contributed equally: Jennine M. Dawicki-McKenna, Xueqi Sun. ✉e-mail: lampson@sas.upenn.edu; blackbe@pennmedicine.upenn.edu

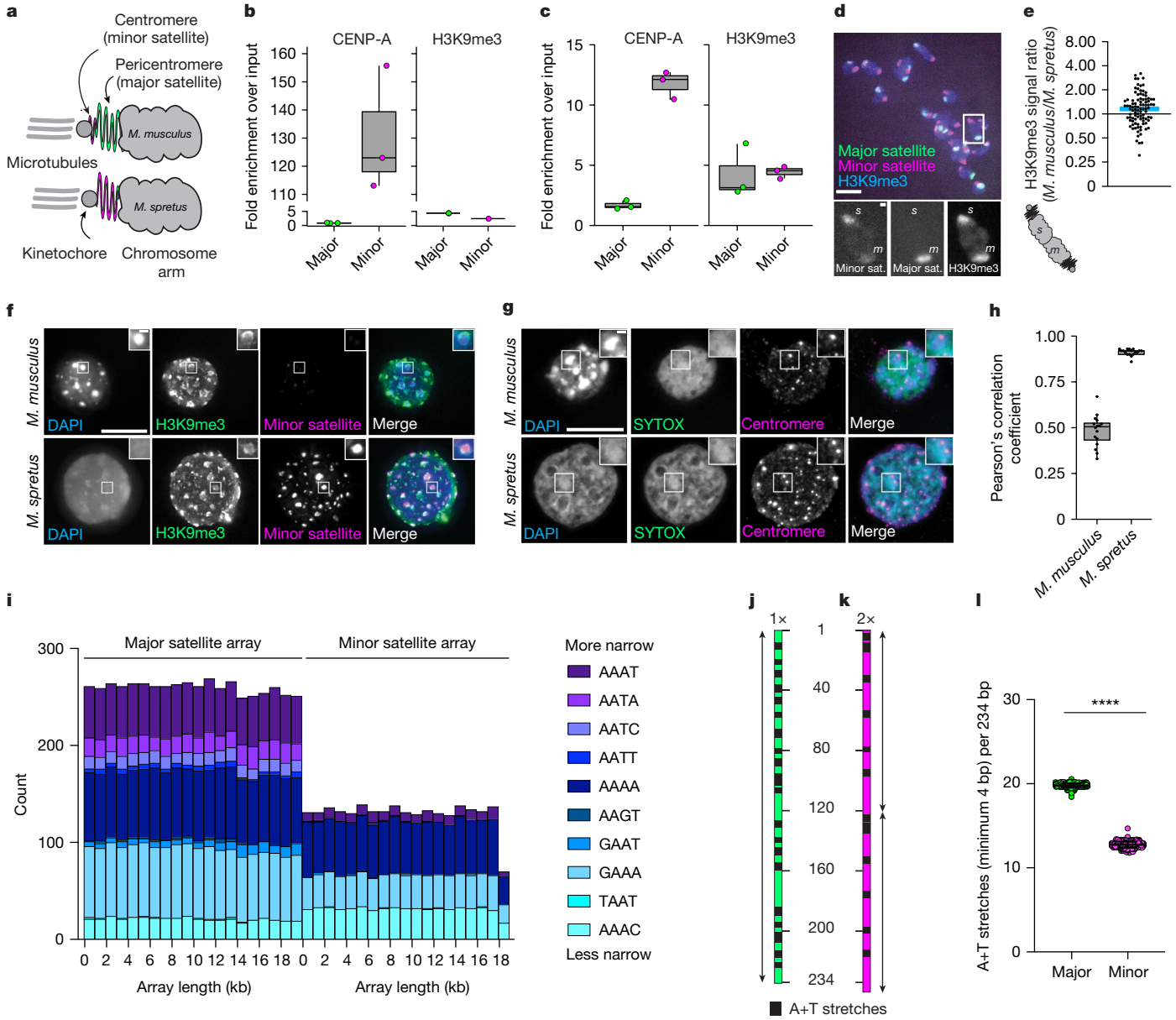


Fig. 1 | Pericentromeric satellites of two closely related mouse species adopt distinct DNA shapes. a, Organization of minor and major satellite DNA in *M. musculus* and *M. spretus*. **b**,**c**, CENP-A or H3K9me3 ChIP results from MNase-digested *M. musculus* (**b**) or *M. spretus* (**c**) chromatin isolated from liver samples. Fold enrichment was calculated as the fraction of reads that align to the major or minor satellite (>80% sequence identity) in the ChIP sample divided by the fraction in the input sample (Tukey box and whiskers plots, $n = 3$ independent experiments except *M. musculus* H3K9me3 ChIP ($n = 1$)). **d**,**e**, Images (**d**) and quantification (**e**) of *M. musculus* (*m*)–*M. spretus* (*s*) hybrid oocytes expressing TALE–mClover (major satellite) and dCas9–mCherry and guide RNA (gRNA) (minor satellite) fixed at metaphase I and stained for H3K9me3. **e**, The H3K9me3 ratio at pericentromeres (*M. musculus*/*M. spretus*) for each bivalent ($n = 95$ bivalents from two independent experiments; blue line, median). **f**, Liver nuclei from *M. spretus* and *M. musculus* stained for H3K9me3. Fluorescence in situ hybridization was performed with probes against CENP-B boxes (minor satellite). A total of 14 cells per condition from two independent experiments yielded similar results. **g**,**h**, Images (**g**) and quantification (**h**) of *M. spretus* and *M. musculus* liver nuclei stained for DNA (SYTOX Green and DAPI) and centromeres (anti-centromere antibody). **h**, The Pearson's correlation coefficient was calculated for nuclear SYTOX Green and DAPI ($n = 16$ and 17 nuclei for *M. musculus* and *M. spretus*, respectively). **i**, A+T-rich tetranucleotides associated with narrow DNA minor groove[35] along representative *M. musculus* major or minor satellite arrays shown side by side. **j**,**k**, Consensus major satellite monomer (green, left) or minor satellite dimers (magenta, right) with stretches of a minimum of four contiguous A or T bases. Despite similar A+T content (63–66%), major and minor satellites differ in how clustered the A and T bases are. **l**, Number of contiguous A or T bases in major or minor satellite arrays per 234 bp (single major satellite repeat length). Data are mean ± s.d. $n = 500$ arrays; two-tailed Mann–Whitney *U*-test. ****$P < 0.0001$. All box plots show the individual values (circles), median, 25th and 75th percentiles (hinges), smallest and largest value at most 1.5 × interquartile range of the hinge (whiskers). Scale bars, 10 µm and 1 µm (insets).

followed by sequencing (ChIP–seq). The two chromatin domains are defined by well-characterized epigenetic marks: the histone H3 variant CENP-A for centromeres and histone H3 trimethylated at K9 (H3K9me3) for pericentromeric heterochromatin. We found that CENP-A chromatin is present mainly on minor satellite arrays, with a similar percentage of CENP-A ChIP–seq reads mapping to minor satellite in both species (Fig. 1b and Supplementary Table 1). In contrast to the allegiance of centromeres to the minor satellite, H3K9me3-defined pericentromeric heterochromatin assembles on whichever satellite is more abundant: the minor satellite in *M. spretus* and the major satellite in *M. musculus* (Fig. 1c and Supplementary Table 1). To determine whether heterochromatin forms to a similar extent on the different satellites, we measured

pericentromeric H3K9me3 using immunofluorescence microscopy on paired homologous chromosomes in *M. musculus* × *spretus* hybrid oocytes. We found similar amounts of H3K9me3 on *M. musculus* and *M. spretus* pericentromeres (Fig. 1d,e). Therefore, although the minor satellite is conducive to centromere specification, possibly due to the presence of CENP-B boxes[27], pericentromeric heterochromatin forms equally well on either satellite. These findings also suggest that *M. spretus* centromeric and pericentromeric chromatin sorts into epigenetically distinct domains at the same type of satellite, perhaps through spatial separation as has recently been observed in human cells using cryo-electron tomography[28].

Divergent satellites differ in DNA shape

Although *M. musculus* and *M. spretus* chromosomes have similar amounts of pericentromeric heterochromatin, only *M. musculus* nuclei form DAPI-rich foci known as chromocentres, which colocalize with H3K9me3-labelled pericentromeric heterochromatin[29] (Fig. 1f). This finding is consistent with DAPI enrichment at *M. musculus* chromosomes compared with *M. spretus*[30] and at the major compared with the minor satellite in *M. musculus*[20]. By contrast, another DNA dye, SYTOX green, shows no distinct foci in either species (Fig. 1g). SYTOX and DAPI staining are well correlated in *M. spretus* nuclei but not in *M. musculus* nuclei (Fig. 1h), suggesting that higher intensities in *M. spretus* simply reflect higher DNA density, but DAPI foci in *M. musculus* must reflect some other property of the DNA. Major and minor satellites are both A+T rich (63–66%), indicating that A+T content cannot explain DAPI enrichment at the major satellite. Because DAPI binds to narrow DNA minor grooves formed by contiguous A or T base pairs[31–33] that deviate from an ideal B-form double helix[34], we hypothesized that the distinct satellites forming pericentromeres in *M. musculus* (major) and *M. spretus* (minor) adopt distinct DNA shapes. To address this, we quantified the abundance of tetranucleotides (including contiguous A or T bases) known to adopt narrow minor groove[35] in major and minor satellites. We found that all nine of the narrowest tetranucleotides are enriched in the major satellite compared with the minor (Fig. 1i, Methods and Extended Data Fig. 2a). Moreover, each major satellite monomer has 20 stretches with a minimum of four contiguous A or T bases (Fig. 1j,l and Extended Data Fig. 2b), whereas the equivalent length of the minor satellite contains only 12 such stretches (Fig. 1k,l and Extended Data Fig. 2c). These results indicate that major satellite repeats contain a higher density of narrow minor DNA grooves than minor satellite repeats, supporting our hypothesis that major and minor satellites differ in DNA shape.

HMGA1 distinguishes divergent satellites

The sequence-dependent shape of DNA creates structural cues for binding by DNA-shape-recognizing proteins[35,36], suggesting that such proteins may bind differentially to pericentromeric satellites in *M. musculus* and *M. spretus*. A primary candidate for differential binding is a chromatin architectural protein, HMGA1, that is enriched at DAPI-rich foci in interphase in primate cells and *M. musculus* pericentromeres[4,5,10]. HMGA1 has three AT-hook motifs that mediate binding to A+T-rich satellite DNA and can bridge major satellite repeats in vitro[5]. Furthermore, HMGA1 competes with DAPI-like small molecules for binding to the DNA minor groove along A+T stretches[37–39] (Extended Data Fig. 3a,b). As AT-hook motifs insert into the minor groove of one face of the DNA double helix, spanning about 4 bp (base pairs)[39], we predicted that HMGA1 would be enriched at the major satellite, which contains more contiguous A+T stretches with a minimum length of 4–5 bp (Fig. 1i and Extended Data Fig. 2d) than the minor satellite, which contains more stretches with a minimum length of 3 bp (Extended Data Fig. 2e). By analysing HMGA1 ChIP–seq data from *M. musculus* somatic cells[40], we find that HMGA1 is enriched twofold at the major satellite compared with the minor satellite (Fig. 2a and Supplementary Table 2). To test whether HMGA1 preferentially binds to *M. musculus* pericentromeres, we analysed its localization in *M. musculus* × *spretus* hybrid oocytes by immunostaining. *Mus musculus* and *M. spretus* HMGA1 proteins have an identical sequence (Extended Data Fig. 4a,b), excluding differences in DNA binding or epitope recognition. Measurements of paired homologous *M. musculus* and *M. spretus* chromosomes show that *M. musculus* pericentromeres recruit more HMGA1 than those of *M. spretus* (Fig. 2b,c). Moreover, HMGA1 levels are similar across hybrid and parental oocytes (Fig. 2b,c). This lack of dependence on genetic background differs from another chromatin packaging protein, condensin II, for which the localization to major satellite is reduced in hybrid *M. musculus* × *spretus* oocytes compared with *M. musculus*, contributing to subfertility in the hybrid[26]. Instead, the threefold binding preference of HMGA1 for *M. musculus* major satellite (Fig. 2c) can be accounted for by the predicted number of minimum 4-bp-long contiguous A+T stretches (2.9-fold; Supplementary Table 1). Moreover, since HMGA1 binds to bulk chromatin in both species (Fig. 2b), its enrichment at *M. musculus* pericentromeres reflects the capacity to bind narrow DNA minor grooves in the major satellite rather than heterochromatin binding. Overall, these findings strongly indicate that HMGA1 recognizes the distinct DNA shape adopted by the major satellite and that DNA shape is a molecular feature that underlies differential packaging of divergent pericentromeres (Fig. 2d).

HMGA1 loss causes satellite stretching

HMGA1 regulates gene expression[41] and bundles multiple pericentromeres into chromocentres in interphase nuclei[10]. However, its enrichment at individual *M. musculus* pericentromeres in female meiosis (Fig. 2b,c) and high expression in proliferating tissues in the embryo[42] suggest additional functions in packaging major satellite in condensed chromosomes. To selectively disrupt HMGA1 function during cell division, we used antibody-mediated protein degradation for acute protein removal during female meiosis (Trim-away[43]; Fig. 3a and Methods). This acute depletion approach minimizes potential compensation by other HMG family members with similar chromatin-packaging roles[44]. The injected antibody does not enter the nucleus[43] (Fig. 3b,c) and the mild decrease in HMGA1 in the intact nucleus is most likely due to protein turnover. Near-complete degradation occurs within 3 h of nuclear envelope breakdown (early meiosis I; Fig. 3d,e). Kinetochores and centromeric chromatin were displaced away from the bulk chromatin in HMGA1-depleted oocytes (Fig. 3d and Extended Data Fig. 5), with massive stretching of DNA corresponding to the major satellite (Fig. 3f,g). Stretching was reduced by treatment with the microtubule-depolymerizing agent nocodazole and restored by nocodazole washout (Extended Data Fig. 6), showing that stretching depends on microtubules. Together, these data demonstrate that HMGA1 packages pericentromeric chromatin to withstand microtubule forces during female meiosis.

To determine whether HMGA1 depletion affects kinetochores, which form on centromeric chromatin adjacent to pericentromeres, we examined highly expressed in cancer 1 (HEC1), a major microtubule-binding protein at kinetochores. We find decreased levels of HEC1 at kinetochores (Fig. 4a,b) and increased kinetochore clustering (Fig. 4c,d), suggesting disrupted kinetochore function. In mouse oocytes, kinetochores facilitate bipolar spindle assembly[45]. Consistent with the disrupted kinetochore function in HMGA1-depleted oocytes, we found delayed spindle bipolarization (Fig. 4e,f). Mouse oocytes typically build a transient radially symmetric apolar spindle[46] (Fig. 3a). Blocking spindle bipolarization using a kinesin-5 inhibitor (STLC) revealed more asymmetric morphology in HMGA1-depleted oocytes (Extended Data Fig. 7a–c), indicating disrupted microtubule organization. HMGA1-depleted oocytes eventually formed bipolar spindles, which coincided with reduced major satellite stretching, but few cells were able to divide (Fig. 4g). Given that spindle forces remain strong during late meiosis I[47], reduced stretching in late meiosis I is probably due to

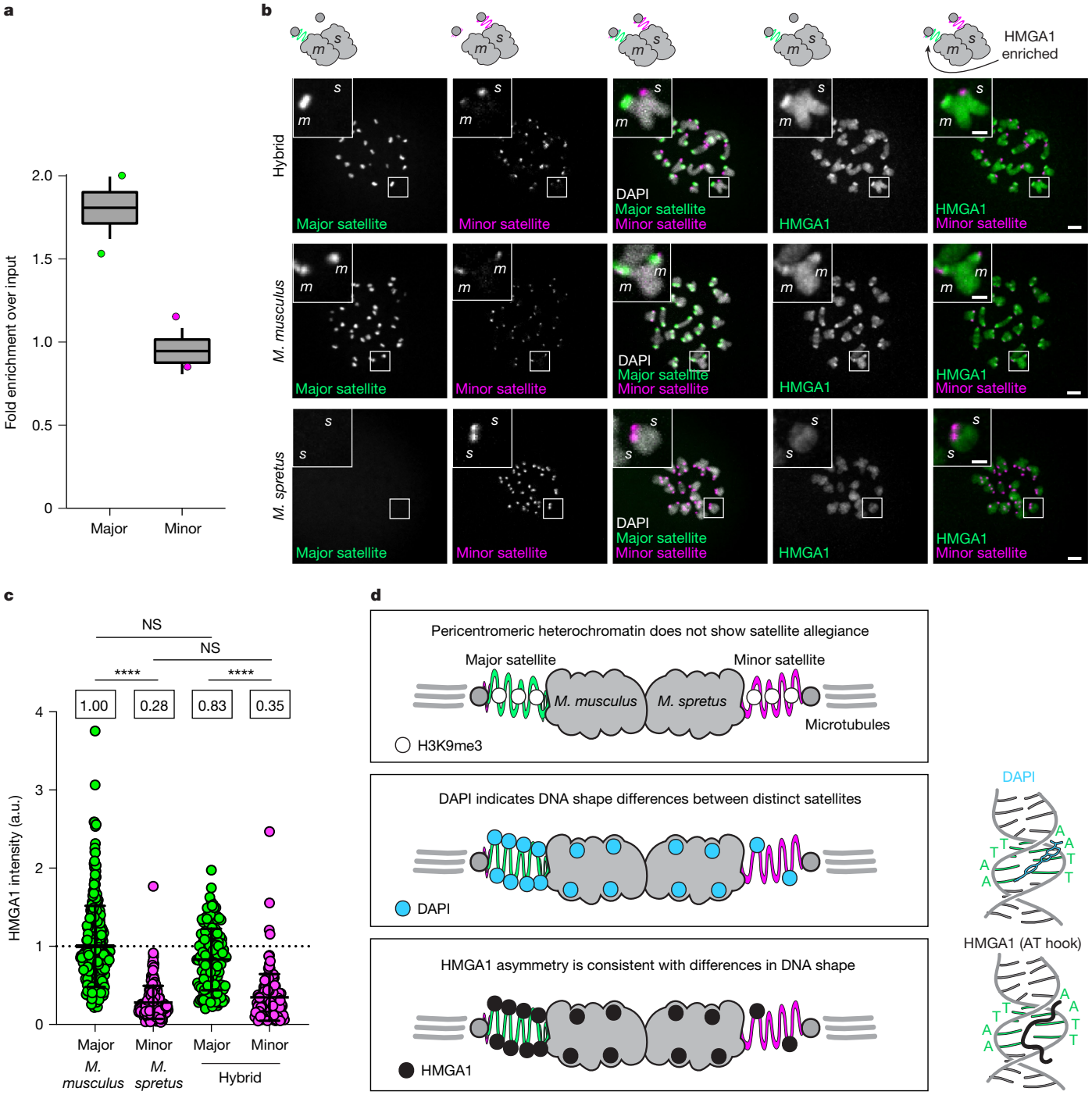


Fig. 2 | HMGA1 is enriched at pericentromeres on the major satellite DNA. **a**, Previously published HMGA1 ChIP–seq data (*M. musculus* embryonic stem cells)[40] were analysed to calculate fold enrichment as the fraction of reads that align to the major or minor satellite (>80% sequence identity) in the ChIP sample divided by the fraction in the input sample (Tukey box and whiskers plots, $n = 2$ independent experiments for each). Box plots indicate individual values (circles), median, 25th and 75th percentiles (hinges), smallest and largest value at most 1.5 × interquartile range of the hinge (whiskers). **b**,**c**, Hybrid *M. musculus* (m) × *spretus* (s) and parental *M. musculus* and *M. spretus* oocytes expressing TALE–mClover targeting the major satellite and dCas9–mCherry in complex with gRNA targeting the minor satellite were fixed and stained for HMGA1. Scale bars, 5 µm and 2 µm (insets). The HMGA1 signal was quantified at major and minor satellites representing *M. musculus* or *M. spretus* pericentromeres. Data are mean ± s.d., with means shown in boxes. $n = 284$ (major satellite in *M. musculus*), 194 (minor satellite in *M. spretus*), 144 (major satellite in hybrid) and 144 (minor satellite in hybrid) pericentromeres from 2–3 independent experiments; Kruskal–Wallis test followed by Dunn's multiple-comparison test. ****$P < 0.0001$; NS, not significant ($P = 0.1059$, major satellite in *M. musculus* versus hybrid; $P = 0.4159$, minor satellite in *M. spretus* versus hybrid). a.u., arbitrary units. **d**, Although both major and minor satellites can form pericentromeric heterochromatin, the major satellite has a higher density of narrow DNA minor grooves as shown by asymmetric DAPI staining. HMGA1 preferentially binds to narrow DNA minor grooves through a mechanism similar to DAPI binding and is therefore enriched on the major satellite. Drawings of the DNA molecule with DAPI and HMGA1 AT hook were based on structures from the Protein Data Bank (PDB): 5T4W and 2EZF, respectively.

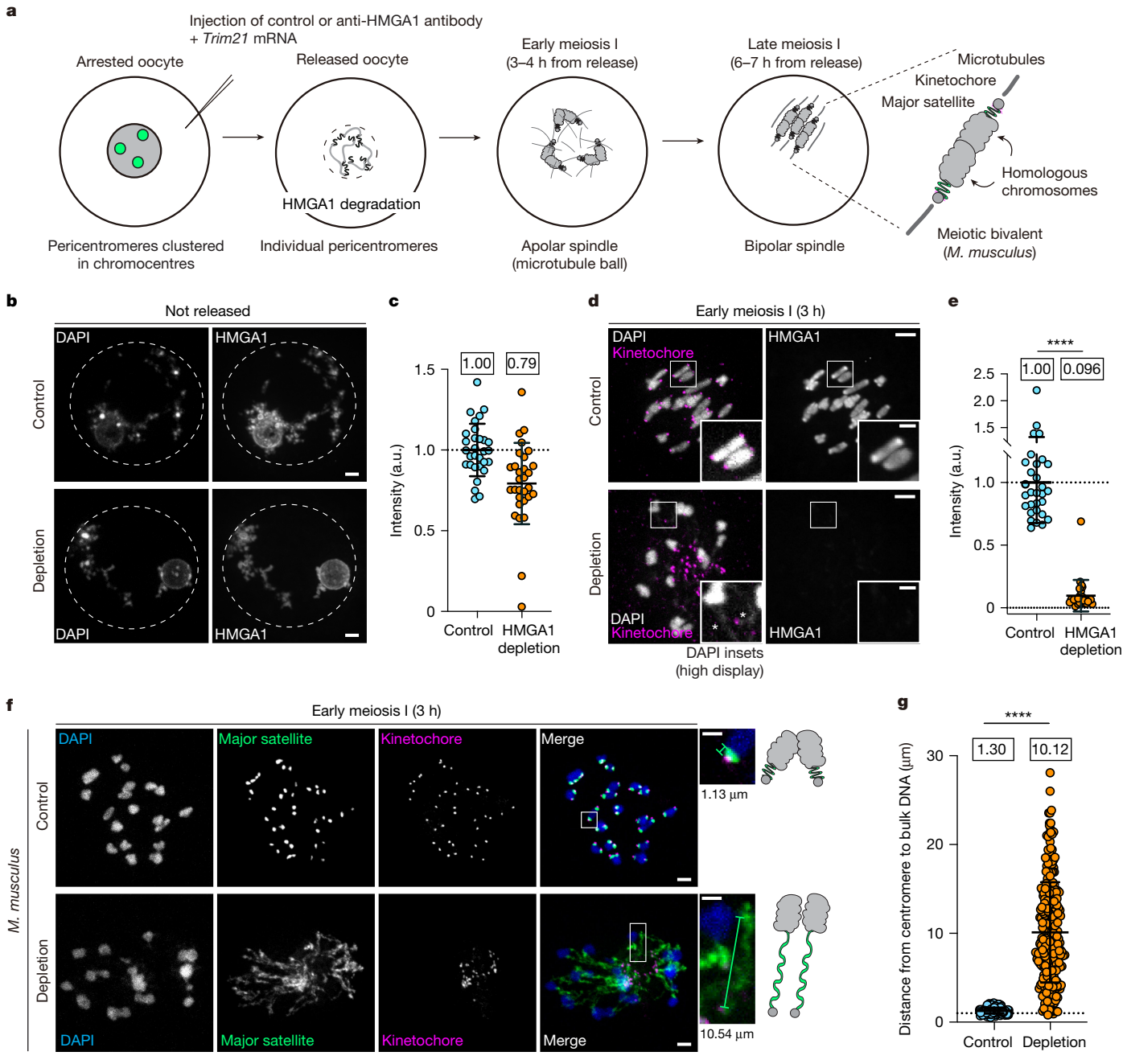


Fig. 3 | HMGA1 packages pericentromeres during female meiosis. **a**, Schematic of meiosis I in *M. musculus* oocytes from milrinone-induced G2/prophase arrest to bipolar spindle assembly. HMGA1 is depleted after release when the nuclear envelope breaks down and an anti-HMGA1 antibody and TRIM21 can access condensing chromosomes (see Methods for details). The major satellite is marked in green. **b**,**c**, Arrested *M. musculus* oocytes were injected with *Trim21* mRNA and anti-HMGA1 (or control) antibodies, incubated for 6 h to enable *Trim21* expression, then fixed and stained for DNA (DAPI) and HMGA1. **b**, Images show nuclei outlined by a dashed line. **c**, HMGA1 intensity was quantified. Data are mean ± s.d., with means shown in boxes. $n = 30$ oocytes per condition from two independent experiments. **d**,**e**, *Mus musculus* oocytes were injected with *Trim21* mRNA and anti-HMGA1 (or control) antibodies, incubated for 3 h, released from arrest, fixed 3 h later (6 h after injection) and stained for DNA (DAPI), HMGA1 and the kinetochore protein HEC1. **d**, Asterisks show DNA threads connecting chromosome arms with kinetochores. DAPI signal is enhanced in the insets to visualize thin threads of DNA. **e**, HMGA1 intensity was quantified. Data are mean ± s.d., with means shown in boxes. $n = 30$ (control) and 28 (depletion) oocytes from two independent experiments; two-tailed Mann–Whitney *U*-test. ****$P < 0.0001$. **f**,**g**, Control or HMGA1-depleted *M. musculus* oocytes expressing TALE–mClover targeting the major satellite were fixed during early meiosis I and stained for DNA (DAPI) and HEC1. Insets show differences in major satellite length (green bars) (**f**). Pericentromere stretching was quantified as the length of major satellite connecting a kinetochore to bulk DNA. **g**, Data are mean ± s.d., with means shown in boxes. $n = 240$ (control) and 270 (depletion) pericentromeres from 48 (control) and 54 (depletion) oocytes from three independent experiments; two-tailed Mann–Whitney *U*-test. ****$P < 0.0001$. Scale bars, 5 µm and 2 µm (insets). Dashed lines in **c**,**e**,**g** indicate average intensity (**c**,**e**) or length (**g**) of the control condition.

compensatory DNA packaging mechanisms (for example, through condensin II[26]). We conclude that HMGA1 ensures pericentromere rigidity both to withstand spindle forces and to promote timely kinetochore-mediated spindle assembly.

Disproportional impact of HMGA1 loss

Because HMGA1 is enriched on the major compared with the minor satellite in hybrid *M. musculus* × *spretus* oocytes (Fig. 2b,c) and its

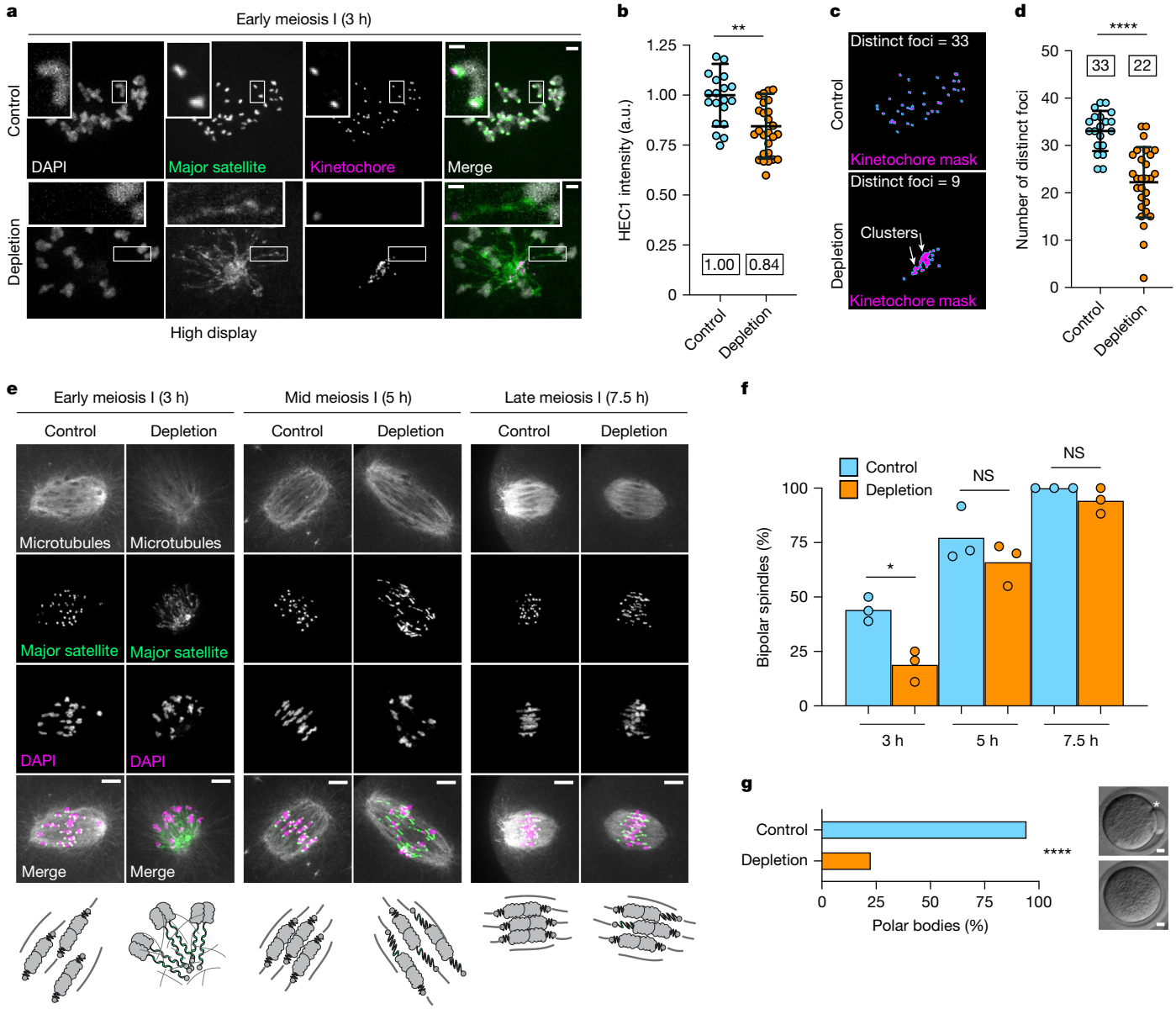


Fig. 4 | HMGA1 depletion disrupts kinetochore organization and delays bipolar spindle assembly. a–**d**, Control or HMGA1-depleted *M. musculus* oocytes expressing TALE–mClover, which targets the major satellite, were fixed during early meiosis I and stained for DNA (DAPI) and HEC1. Scale bars, 5 µm and 2 µm (insets). **a**, Major satellite signal is enhanced in the HMGA1-depleted oocyte to visualize thin threads of DNA. **b**, HEC1 was quantified at individual kinetochores. Data are mean ± s.d., with means shown in boxes. $n = 20$ (control) and 27 (depletion) oocytes from two independent experiments; unpaired two-tailed t-test. $^{**}P = 0.0018$. **c**,**d**, Kinetochore clustering was measured using HEC1 as a mask (**c**) and counting the number of distinct foci per cell (**d**). Data are mean ± s.d., with means shown in boxes. $n = 20$ (control) and 28 (depletion) oocytes from two independent experiments; unpaired two-tailed t-test. $^{****}P < 0.0001$. **e**,**f**, Images (**e**) and quantification (**f**) of control or HMGA1-depleted *M. musculus* oocytes expressing TALE–mClover, which targets the major satellite, were fixed at the indicated timepoints and stained for DNA (DAPI) and microtubules (α-tubulin). Scale bars, 10 µm. **f**, Bipolar spindles were counted to quantify spindle assembly kinetics. Data are means (bars) and individual values (circles). $n = 50$ (3 h, control), 62 (3 h, depletion), 42 (5 h, control), 55 (5 h, depletion), 36 (7.5 h, control) and 57 (7.5 h, depletion) oocytes; ordinary one-way analysis of variance (ANOVA) followed by Tukey's multiple-comparison test. $^{*}P = 0.0187$ (3 h), NS ($P = 0.5306$ (5 h) and $P = 0.9428$ (7.5 h)). **g**, Control or HMGA1-depleted *M. musculus* oocytes were cultured until metaphase of meiosis II. Representative images show a polar body (control; white asterisk) or lack thereof (HMGA1 depleted). Frequency of polar body extrusion was quantified. $n = 60$ oocytes per condition from two independent experiments, Fisher's exact two-sided test. $^{****}P < 0.0001$. Scale bars, 10 µm.

depletion results in massive stretching of the major satellite in *M. musculus* oocytes (Fig. 3f,g), we predict that HMGA1 preferentially packages the major satellite rather than minor satellite. To test this prediction, we depleted HMGA1 in *M. musculus* × *spretus* hybrid oocytes to compare chromosomes with pericentromeres built on different satellites in the same cell. Consistent with our prediction, we found that *M. musculus* pericentromeres were significantly more stretched than *M. spretus* pericentromeres (Fig. 5a,b and Extended Data Fig. 8a–d). To test whether differential disruption of pericentromere packaging also extends to kinetochore function, we examined mitotic arrest deficient 1 (MAD1), a checkpoint protein that localizes to kinetochores that lack stable microtubule attachments during meiosis[48]. We measured MAD1 intensity at individual kinetochores and computed the ratio of *M. musculus* to *M. spretus* kinetochores in each hybrid oocyte (Fig. 5c,d and Extended Data Fig. 9a–c). At an early timepoint (3 h), most kinetochores were MAD1-positive, as expected because kinetochore–microtubule attachments are typically unstable early in metaphase I[49,50], and the *M. musculus*/*M. spretus* ratio was similar in control

and HMGA1-depleted oocytes. At a later timepoint (7 h), MAD1 was twofold higher on *M. musculus* compared with *M. spretus* kinetochores in HMGA1-depleted hybrid oocytes, indicating more unstable attachments at *M. musculus* kinetochores. We could not accurately measure MAD1–GFP levels in control hybrid oocytes at 7 h because most attachments are stable by this time and lack MAD1. Higher MAD1 levels at *M. musculus* kinetochores are not explained by bigger kinetochores, because HEC1 is lower on *M. musculus* compared with *M. spretus* kinetochores in HMGA1-depleted cells (Fig. 5c,e and Extended Data Fig. 9d–f). These results demonstrate a disproportionate role of HMGA1 in the packaging of *M. musculus* pericentromeres, which are built on the major satellite, compared with *M. spretus* pericentromeres built on the minor satellite. Moreover, DNA-shape-mediated chromatin packaging has functional implications for the formation of kinetochore–microtubule attachments during female meiosis.

Discussion

Together, our work provides experimental evidence that the pericentromeric satellite DNA sequence influences its packaging and function during chromosome segregation. We found that the sequence-mediated shape of the DNA molecule is a determinant of pericentromere packaging during female meiosis in mammals. We found that, although both minor and major satellites can form pericentromeric heterochromatin, they differ in their dependence on HMGA1, which recognizes DNA shape (Fig. 2d). The high density of contiguous A+T stretches in major satellite repeats is consistent with an intrinsic capacity to adopt the narrow DNA minor groove recognized by HMGA1. We found that HMGA1 packages individual pericentromeres, probably by bundling DNA via its three AT-hook motifs[5], similar to bundling multiple pericentromeres into chromocentres during interphase[10]. Disrupting pericentromere packaging during meiosis by HMGA1 depletion leads to pericentromere stretching by microtubules, indicating that pericentromeric rigidity is needed to withstand the spindle forces[51,52] that pull and push chromosomes during cell division[53]. These findings also indicate that major satellite packaging in female meiosis requires both condensin II[26] and HMGA1, reflecting the challenges that large satellites pose to the genome. Given that acute HMGA1 depletion results in nearly fivefold longer major satellite stretches compared to the equivalent condensin II depletion[26], HMGA1 might be the more potent packaging mechanism.

Acute HMGA1 depletion most likely disrupts kinetochore organization (Fig. 4a–d) via the massive pericentromere stretching (Fig. 3f,g), consistent with more disrupted *M. musculus* than *M. spretus* kinetochores in hybrid oocytes (Fig. 5c–e). Our results also reveal the importance of pericentromere rigidity for kinetochore interactions with microtubules locally and organization of the spindle more globally (Extended Data Fig. 7d). We propose that the delay in bipolar spindle formation in HMGA1-depleted cells (Fig. 4e,f) is due to kinetochore clustering (Fig. 4c,d), which would change the spatial organization of the key microtubule crosslinker protein regulator of cytokinesis 1 (PRC1)[45]. Disruption of kinetochore–microtubule attachments (Fig. 5c,d) is probably mediated by Aurora A kinase destabilizing attachments near microtubule-organizing centres located in the apolar microtubule ball[48,54] (Extended Data Fig. 7d). Alternatively, stretched pericentromeres may fail to create tension, which typically stabilizes attachments[55]. Altogether, our work using acute depletion in meiosis reveals an important function of HMGA1 distinct between its roles in maintaining chromocentre integrity and in preventing DNA damage and micronuclei formation during interphase[10].

Using the closely related species of *M. musculus* and *M. spretus* as a natural model system with variation in pericentromeric satellites provides insight into the evolution of satellite DNA. Distinct satellites are thought to emerge from a library of common repeats present in closely related species (library hypothesis[56]) but it is unclear how some repeats become more abundant than others. Satellites may have features (such as A+T content or repeat length) that are advantageous, for example, at centromeres or pericentromeres[57–59], and thus favoured by natural selection. However, that scenario does not explain the differential abundance of major and minor satellites in *M. musculus* and *M. spretus*. Although both satellites can form pericentromeric heterochromatin (Fig. 1), major satellite expansion occurred specifically in the *M. musculus* lineage[15,26] (Fig. 5f). Alternatively, satellites may increase their abundance by meiotic drive through biased inheritance during female meiosis[1]. Although expansion of centromeric minor satellite correlates with drive[21,25], there is no evidence that major satellite can drive in meiosis. Our findings suggest a mechanism to facilitate satellite expansion without requiring any particular satellite function (Fig. 5g). We propose that expansion is costly if the satellite is not properly packaged. A satellite that adopts a DNA shape prevalent at other genomic loci (such as enhancers[60]), however, would be efficiently packaged by conserved DNA-shape-recognizing proteins, thus minimizing the cost of expansion. An alternative model is selection for heterochromatin protein variants that bind the expanded satellite to reduce fitness costs[61]. Under our model, adopting a DNA shape recognized by highly conserved HMGA1 facilitated expansion of the major satellite in the *M. musculus* lineage. Consistent with our model, recent work in *M. musculus* × *spretus* hybrid oocytes demonstrated that failure to package the major satellite can lead to chromosome missegregation. That effect is mediated by reduced condensin II at the major satellite in the hybrid background compared to that in *M. musculus* and underlies subfertility in the hybrid[26]. By contrast, we show that HMGA1 recruitment depends on DNA shape and not the genetic background (Fig. 2b,c). This finding implies that HMGA1 is not involved in the subfertility of *M. musculus* × *spretus* hybrid mice and that a readily available HMGA1 pool is present in *M. musculus*, *M. spretus* and hybrid oocytes, probably owing to its highly conserved function in genome organization. Therefore, DNA-shape-mediated satellite packaging by highly conserved and abundant architectural proteins can explain a long-standing paradox: how large satellite arrays rapidly evolve while evading fitness costs associated with their expansion.

It will be important in the future to expand on our model, which is based on the binding patterns of a single DNA-shape-recognizing architectural protein. We propose that this will entail defining the binding patterns of other DNA-shape-recognizing architectural proteins in diverse eukaryotes. We speculate that another protein that is sensitive to DNA sequence and/or shape facilitated minor satellite expansion in *M. spretus*. Candidates include other HMG proteins, histone H1 variants that compete with HMG proteins for DNA binding[62] and the centromeric protein CENP-B[63]. Recent findings in *Drosophila* might offer an additional example as distinct pericentromeric AATAACATAG and GAACAGAACATGTTC satellites are both packaged by highly conserved Prod[64], possibly because each adopted its cognate DNA shape, facilitating expansion in the closely related *Drosophila melanogaster* and *Drosophila simulans* lineages, respectively. Interestingly, most peri/centromeric satellites in eukaryotes are A+T rich[59], suggesting that the general satellite-packaging principles we propose based on our mouse models may be beneficial in diverse species to withstand spindle forces. Furthermore, our model is built on experimentation using acute HMGA1 depletion. Given that *Hmga1* knockout mice are viable, albeit with impaired insulin signalling[65], compensatory mechanisms for packaging the major satellite must arise upon traditional gene disruption. The contribution of these compensatory mechanisms in the female germline could be tested using conditional HMGA1 depletion. Finally, it will be interesting to directly test the impact of DNA shape recognition on major satellite expansion through manipulation of shape. Despite recent advances in the sequencing of repetitive regions[2,3], satellite-manipulation methods remain in their infancy. For now, our model provides a mechanistic explanation for how a specific megabase satellite array could emerge at short evolutionary time scales from a library of repeats present in closely related species[56,59,66].

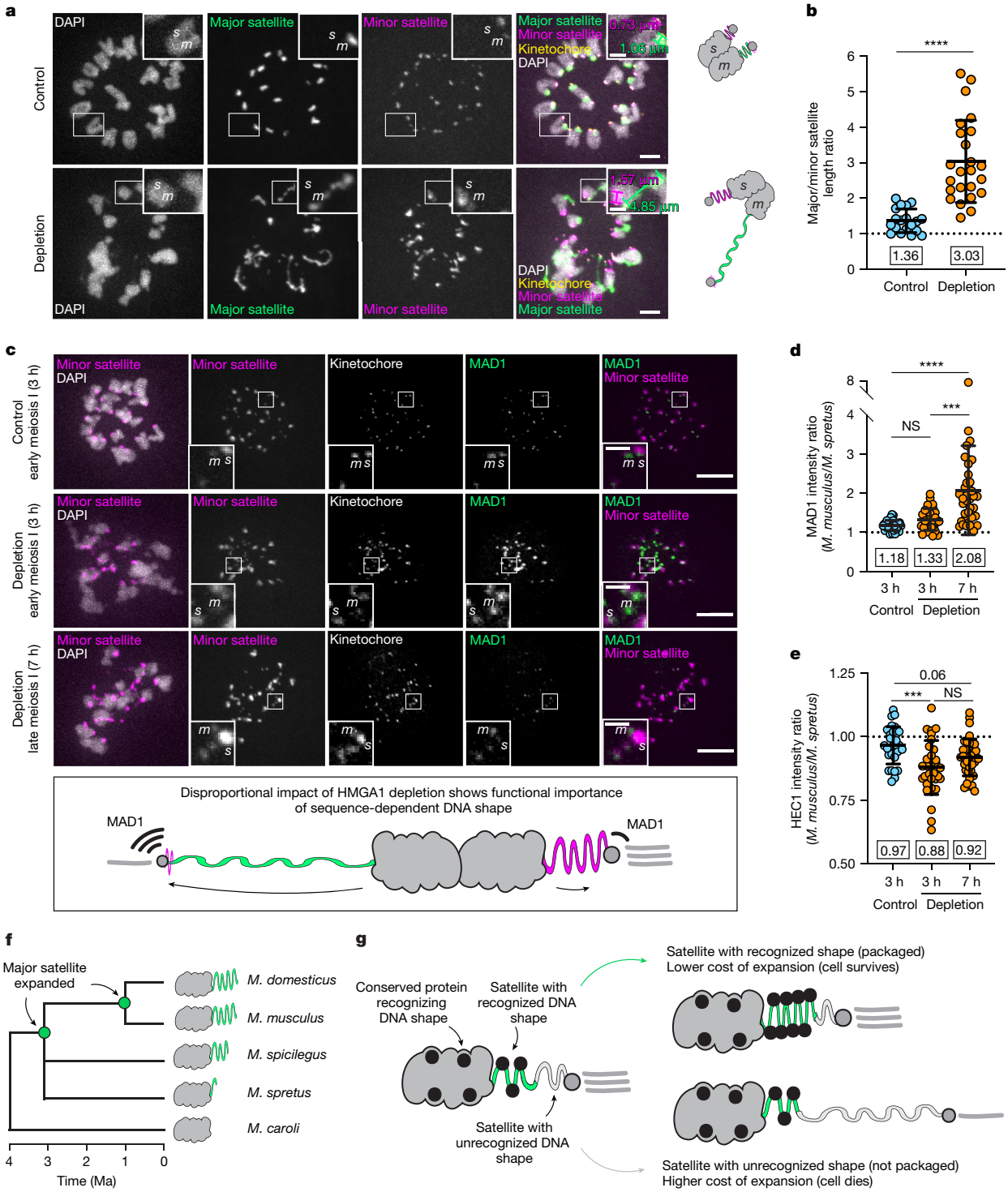


Fig. 5 | HMGA1 depletion disproportionately disrupts *M. musculus* pericentromeres. a,**b**, Control or HMGA1-depleted *M. musculus* (m) × *spretus* (s) hybrid oocytes were fixed during early meiosis I and stained for DNA (DAPI) and kinetochore (HEC1). Cells were expressing TALE–mClover (major satellite) and dCas9–mCherry and gRNA (minor satellite). Scale bars, 5 µm and 2 µm (insets). The ratio of *M. musculus*/*M. spretus* pericentromere stretching was quantified as major satellite length/minor satellite length. **b**, Data are mean ± s.d., with means shown in boxes. $n = 19$ (control) and 24 (depletion) hybrid oocytes from two independent experiments, five pericentromeres per species measured per oocyte, unpaired two-tailed t-test. $^{****}P < 0.0001$. **c**–**e**, Control or HMGA1-depleted *M. musculus* × *spretus* hybrid oocytes were fixed during early or late meiosis I and stained for DNA (DAPI) and kinetochore (HEC1). Cells were expressing MAD1–2×GFP and dCas9–mCherry/gRNA (minor satellite). Scale bars, 10 µm and 2 µm (insets). **d**,**e**, Unattached kinetochores (**d**) and kinetochore size (**e**) were quantified as a ratio between average MAD1 or HEC1 signal at *M. musculus* and *M. spretus* kinetochores. Data are mean ± s.d., with means shown in boxes. $n = 30$ (3 h, control), 31 (3 h, depletion) and 38 (7 h, depletion) hybrid oocytes from two independent experiments, at least 17 kinetochores quantified per oocyte, ordinary one-way ANOVA followed by Tukey's multiple-comparison test. $^{****}P < 0.0001$, $^{***}P = 0.0001$ (MAD1), $^{***}P = 0.0004$ (HEC1), NS ($P = 0.6906$ (MAD1) and $P = 0.1402$ (HEC1 3 h versus 7 h depletion)). **f**, Expansion of major satellite (green) along the *M. musculus* lineage[26]. **g**, Different satellites have distinct DNA shapes. A satellite recognized by existing packaging proteins can be efficiently packaged and expand in the genome without severe fitness costs. Conversely, a satellite that is not recognized is not packaged, leading to a fitness cost of expansion. Note that this model does not predict that DNA-shape-mediated mechanisms evolved to promote satellite expansion but rather that satellites co-opt existing mechanisms that evolved to bind to other parts of the genome.

Online content

Any methods, additional references, Nature Portfolio 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/s41586-024-08374-0.

1. Henikoff, S., Ahmad, K. & Malik, H. S. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**, 1098–1102 (2001).
2. Altemose, N. et al. Complete genomic and epigenetic maps of human centromeres. *Science* **376**, eabl4178 (2022).
3. Logsdon, G. A. et al. The variation and evolution of complete human centromeres. *Nature* **629**, 136–145 (2024).
4. Strauss, F. & Varshavsky, A. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* **37**, 889–901 (1984).
5. Vogel, B., Loschberger, A., Sauer, M. & Hock, R. Cross-linking of DNA through HMGA1 suggests a DNA scaffold. *Nucleic Acids Res.* **39**, 7124–7133 (2011).
6. Kixmoeller, K., Allu, P. K. & Black, B. E. The centromere comes into focus: from CENP-A nucleosomes to kinetochore connections with the spindle. *Open Biol.* **10**, 200051 (2020).
7. Bernard, P. et al. Requirement of heterochromatin for cohesion at centromeres. *Science* **294**, 2539–2542 (2001).
8. Folco, H. D., Pidoux, A. L., Urano, T. & Allshire, R. C. Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* **319**, 94–97 (2008).
9. Olszak, A. M. et al. Heterochromatin boundaries are hotspots for de novo kinetochore formation. *Nat. Cell Biol.* **13**, 799–808 (2011).
10. Jagannathan, M., Cummings, R. & Yamashita, Y. M. A conserved function for pericentromeric satellite DNA. *eLife* **7**, e34122 (2018).
11. Nonaka, N. et al. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* **4**, 89–93 (2002).
12. Kawashima, S. A. et al. Shugoshin enables tension-generating attachment of kinetochores by loading Aurora to centromeres. *Genes Dev.* **21**, 420–435 (2007).
13. Eckert, C. A., Gravdahl, D. J. & Megee, P. C. The enhancement of pericentromeric cohesin association by conserved kinetochore components promotes high-fidelity chromosome segregation and is sensitive to microtubule-based tension. *Genes Dev.* **21**, 278–291 (2007).
14. Alkan, C. et al. Genome-wide characterization of centromeric satellites from multiple mammalian genomes. *Genome Res.* **21**, 137–145 (2011).
15. Arora, U. P., Charlebois, C., Lawal, R. A. & Dumont, B. L. Population and subspecies diversity at mouse centromere satellites. *BMC Genomics* **22**, 279 (2021).
16. Jagannathan, M., Warsinger-Pepe, N., Watase, G. J. & Yamashita, Y. M. Comparative analysis of satellite DNA in the *Drosophila melanogaster* species complex. *G3* **7**, 693–704 (2017).
17. Chang, C. H. et al. Islands of retroelements are major components of *Drosophila* centromeres. *PLoS Biol.* **17**, e3000241 (2019).
18. Gambogi, C. W. et al. Centromere innovations within a mouse species. *Sci. Adv.* **9**, eadi5764 (2023).
19. DeBose-Scarlett, E. M. & Sullivan, B. A. Genomic and epigenetic foundations of neocentromere formation. *Annu. Rev. Genet.* **55**, 331–348 (2021).
20. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* **166**, 493–505 (2004).
21. Iwata-Otsubo, A. et al. Expanded satellite repeats amplify a discrete CENP-A nucleosome assembly site on chromosomes that drive in female meiosis. *Curr. Biol.* **27**, 2365–2373 (2017).
22. Wong, A. K., Biddle, F. G. & Rattner, J. B. The chromosomal distribution of the major and minor satellite is not conserved in the genus *Mus*. *Chromosoma* **99**, 190–195 (1990).
23. Narayanswami, S. et al. Cytological and molecular characterization of centromeres in *Mus domesticus* and *Mus spretus*. *Mamm. Genome* **2**, 186–194 (1992).
24. Miyanari, Y., Ziegler-Birling, C. & Torres-Padilla, M. E. Live visualization of chromatin dynamics with fluorescent TALEs. *Nat. Struct. Mol. Biol.* **20**, 1321–1324 (2013).
25. Akera, T., Trimm, E. & Lampson, M. A. Molecular strategies of meiotic cheating by selfish centromeres. *Cell* **178**, 1132–1144 (2019).
26. El Yakoubi, W. & Akera, T. Condensin dysfunction is a reproductive isolating barrier in mice. *Nature* **623**, 347–355 (2023).
27. Masumoto, H., Masukata, H., Muro, Y., Nozaki, N. & Okazaki, T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* **109**, 1963–1973 (1989).
28. Kixmoeller, K., Tarasovetc, E. V., Mer, E., Chang, Y. W. & Black, B. E. Centromeric chromatin clearings demarcate the site of kinetochore formation. *Cell* (in the press).
29. Brandle, F., Fruhbauer, B. & Jagannathan, M. Principles and functions of pericentromeric satellite DNA clustering into chromocenters. *Semin. Cell Dev. Biol.* **128**, 26–39 (2022).
30. Matsuda, Y. & Chapman, V. M. In situ analysis of centromeric satellite DNA segregating in *Mus* species crosses. *Mamm. Genome* **1**, 71–77 (1991).
31. Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K. & Dickerson, R. E. The structure of DAPI bound to DNA. *J. Biomol. Struct. Dyn.* **7**, 477–491 (1989).
32. Wilson, W. D. et al. Binding of 4′,6-diamidino-2-phenylindole (DAPI) to GC and mixed sequences in DNA: intercalation of a classical groove-binding molecule. *J. Am. Chem. Soc.* **111**, 5008–5010 (1989).
33. Wilson, W. D. et al. DNA sequence dependent binding modes of 4′,6-diamidino-2-phenylindole (DAPI). *Biochemistry* **29**, 8452–8461 (1990).
34. Hizver, J., Rozenberg, H., Frolow, F., Rabinovich, D. & Shakked, Z. DNA bending by an adenine–thymine tract and its role in gene regulation. *Proc. Natl Acad. Sci. USA* **98**, 8490–8495 (2001).
35. Rohs, R. et al. The role of DNA shape in protein–DNA recognition. *Nature* **461**, 1248–1253 (2009).
36. Joshi, R. et al. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* **131**, 530–543 (2007).
37. Solomon, M. J., Strauss, F. & Varshavsky, A. A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. *Proc. Natl Acad. Sci. USA* **83**, 1276–1280 (1986).
38. Radic, M. Z., Saghbini, M., Elton, T. S., Reeves, R. & Hamkalo, B. A. Hoechst 33258, distamycin A, and high mobility group protein I (HMG-I) compete for binding to mouse satellite DNA. *Chromosoma* **101**, 602–608 (1992).
39. Huth, J. R. et al. The solution structure of an HMG-I(Y)–DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.* **4**, 657–665 (1997).
40. Colombo, D. F., Burger, L., Baubec, T. & Schubeler, D. Binding of high mobility group A proteins to the mammalian genome occurs as a function of AT-content. *PLoS Genet.* **13**, e1007102 (2017).
41. Vignali, R. & Marracci, S. HMGA genes and proteins in development and evolution. *Int. J. Mol. Sci.* **21**, 654 (2020).
42. Chiappetta, G. et al. High level expression of the HMGI (Y) gene during embryonic development. *Oncogene* **13**, 2439–2446 (1996).
43. Clift, D. et al. A method for the acute and rapid degradation of endogenous proteins. *Cell* **171**, 1692–1706 (2017).
44. Federico, A. et al. *Hmga1*/*Hmga2* double knock-out mice display a "superpygmy" phenotype. *Biol. Open* **3**, 372–378 (2014).
45. Yoshida, S. et al. Prc1-rich kinetochores are required for error-free acentrosomal spindle bipolarization during meiosis I in mouse oocytes. *Nat. Commun.* **11**, 2652 (2020).
46. Schuh, M. & Ellenberg, J. Self-organization of MTOCs replaces centrosome function during acentrosomal spindle assembly in live mouse oocytes. *Cell* **130**, 484–498 (2007).
47. Vallot, A. et al. Tension-induced error correction and not kinetochore attachment status activates the SAC in an Aurora-B/C-dependent manner in oocytes. *Curr. Biol.* **28**, 130–139 (2018).
48. Chmatal, L., Yang, K., Schultz, R. M. & Lampson, M. A. Spatial regulation of kinetochore microtubule attachments by destabilization at spindle poles in meiosis I. *Curr. Biol.* **25**, 1835–1841 (2015).
49. Brunet, S. et al. Kinetochore fibers are not involved in the formation of the first meiotic spindle in mouse oocytes, but control the exit from the first meiotic M phase. *J. Cell Biol.* **146**, 1–12 (1999).
50. Kitajima, T. S., Ohsugi, M. & Ellenberg, J. Complete kinetochore tracking reveals error-prone homologous chromosome biorientation in mammalian oocytes. *Cell* **146**, 568–581 (2011).
51. Amor, D. J., Kalitsis, P., Sumer, H. & Choo, K. H. Building the centromere: from foundation proteins to 3D organization. *Trends Cell Biol.* **14**, 359–368 (2004).
52. Vermaak, D. & Malik, H. S. Multiple roles for heterochromatin protein 1 genes in *Drosophila*. *Annu. Rev. Genet.* **43**, 467–492 (2009).
53. Risteski, P., Jagric, M., Pavin, N. & Tolic, I. M. Biomechanics of chromosome alignment at the spindle midplane. *Curr. Biol.* **31**, R574–R585 (2021).
54. Ye, A. A. et al. Aurora A kinase contributes to a pole-based error correction pathway. *Curr. Biol.* **25**, 1842–1851 (2015).
55. Lampson, M. A. & Grishchuk, E. L. Mechanisms to avoid and correct erroneous kinetochore-microtubule attachments. *Biology* **6**, 1 (2017).
56. Fry, K. & Salser, W. Nucleotide sequences of HS-α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* **12**, 1069–1084 (1977).
57. Plohl, M., Mestrovic, N., Bruvo, B. & Ugarkovic, D. Similarity of structural features and evolution of satellite DNAs from *Palorus subdepressus* (Coleoptera) and related species. *J. Mol. Evol.* **46**, 234–239 (1998).
58. Plohl, M., Luchetti, A., Mestrovic, N. & Mantovani, B. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **409**, 72–82 (2008).
59. Melters, D. P. et al. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* **14**, R10 (2013).
60. Parker, S. C., Hansen, L., Abaan, H. O., Tullius, T. D. & Margulies, E. H. Local DNA topography correlates with functional noncoding regions of the human genome. *Science* **324**, 389–392 (2009).
61. Brand, C. L. & Levine, M. T. Functional diversification of chromatin on rapid evolutionary timescales. *Annu. Rev. Genet.* **55**, 401–425 (2021).
62. Postnikov, Y. V. & Bustin, M. Functional interplay between histone H1 and HMG proteins in chromatin. *Biochim. Biophys. Acta* **1859**, 462–467 (2016).
63. Chardon, F. et al. CENP-B-mediated DNA loops regulate activity and stability of human centromeres. *Mol. Cell* **82**, 1751–1767 (2022).
64. Jagannathan, M. & Yamashita, Y. M. Defective satellite DNA clustering into chromocenters underlies hybrid incompatibility in *Drosophila*. *Mol. Biol. Evol.* **38**, 4977–4986 (2021).
65. Foti, D. et al. Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. *Nat. Med.* **11**, 765–773 (2005).
66. Mestrovic, N., Plohl, M., Mravinac, B. & Ugarkovic, D. Evolution of satellite DNAs from the genus *Palorus*–experimental evidence for the "library" hypothesis. *Mol. Biol. Evol.* **15**, 1062–1068 (1998).

Methods

Mice

Mus musculus mouse strains were purchased from The Jackson Laboratory (C57BL/6J, stock no. 000664; used for sequencing experiments) or Envigo/Inotiv (NSA, stock no. 033 corresponding to CF-1; used for oocyte experiments). *Mus spretus* was purchased from Jackson Laboratory (SPRET/EiJ, stock# 001146; used for sequencing experiments) or RIKEN BioResource Research Center (SPR2, RBRC00208; used for oocyte experiments). CF-1 female mice were crossed to SPRET/EiJ male mice to generate hybrids. All mice used in this study were female at 2–4 months old. Mice were housed in controlled room-temperature conditions with minimal disturbances, light–dark cycle of 12 h each and humidity ranging from 30% to 70% depending on the season. All animal experiments were approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania and were consistent with the National Institutes of Health guidelines (protocol no.: 804882).

MNase-digested chromatin and native ChIP

ChIP was performed as previously described[21]. In brief, nuclei were isolated from flash-frozen mouse livers *M. musculus* (C57BL/6J) or *M. spretus*. Livers were homogenized in 4 ml ice-cold buffer I (0.32 M sucrose, 60 mM KCl, 15 mM NaCl, 15 mM Tris-Cl, pH 7.5, 5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM DTT, 0.1 mM PMSF, 1 mM leupeptin–pepstatin, 1 mM aprotinin) per g of tissue by Dounce homogenization. The homogenate was filtered through 100-µm cell strainer (Falcon) and centrifuged at 6,000*g* for 10 min at 4 °C. The pellet was resuspended in the same volume of buffer I. An equivalent volume of ice-cold buffer I supplemented with 0.2% IGEPAL was added and samples were incubated on ice for 10 min. Then 4 ml nuclei were layered on top of 8 ml ice-cold buffer III (1.2 M sucrose, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, 15 mM Tris, pH 7.5, 0.5 mM DTT, 0.1 mM PMSF, 1 mM leupeptin–pepstatin, 1 mM aprotinin) and centrifuged at 10,000*g* for 20 min at 4 °C with no brake. Pelleted nuclei were resuspended in buffer A (0.34 M sucrose, 15 mM HEPES, pH 7.4, 15 mM NaCl, 60 mM KCl, 4 mM $MgCl_2$, 1 mM DTT, 0.1 mM PMSF, 1 mM leupeptin–pepstatin, 1 mM aprotinin), flash-frozen in liquid nitrogen, and stored at −80 °C. Nuclei were digested with MNase (Affymetrix) using 0.05–0.1 U μg^{-1} chromatin in buffer A supplemented with 3 mM $CaCl_2$ for 10 min at 37 °C. The reaction was quenched with 10 mM EGTA on ice for 5 min, and an equal volume of 2× post-MNase buffer (40 mM Tris, pH 8, 220 mM NaCl, 4 mM EDTA, 2% Triton X-100, 0.5 mM DTT, 0.5 mM PMSF, 1 mM leupeptin–pepstatin, 1 mM aprotinin) was added before to centrifugation at 18,800*g* for 15 min at 4 °C. The supernatant containing the MNase-digested chromatin was pre-cleared with 100 µl 50% Protein G Sepharose bead (GE Healthcare) slurry in 1× post-MNase buffer for approximately 2 h at 4 °C with rotation. Beads were blocked in NET buffer (150 mM NaCl, 50 mM Tris, pH 7.5, 1 mM EDTA, 0.1% IGEPAL, 0.25% gelatin and 0.03% NaN_3). The pre-cleared supernatant was divided so that an estimated 250 µg chromatin was used for ChIP (10 µg H3K9me3 antibody, Abcam ab8898; or 10 µg anti-mouse-specific CENP-A antibody[18]) and 12.5 µg was saved as input. ChIP samples were rotated at 4 °C for 2 h. Immunocomplexes were recovered by addition of 100 µl 50% NET-blocked Protein G Sepharose bead slurry followed by overnight rotation at 4 °C. The beads were washed three times with wash buffer 1 (150 mM NaCl, 20 mM Tris-HCl, pH 8, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), once with high salt wash buffer (500 mM NaCl, 20 mM Tris-HCl, pH 8, 2 mM EDTA, 0.1% SDS, 1% Triton X-100) and the chromatin was eluted twice each with 200 µl elution buffer (50 mM $NaHCO_3$, 0.32 mM sucrose, 50 mM Tris, pH 8, 1 mM EDTA, 1% SDS) at 65 °C for 10 min at 1,500 rpm. The input sample was adjusted to a final volume of 400 µl with elution buffer. To each 400 µl input and ChIP sample, 16.8 µl of 5 M NaCl and 1 µl of RNAse A (10 mg ml^{-1}) was added. After 1 h at 37 °C, 4 µl of 0.5 M EDTA and 12 µl proteinase K (2.5 mg ml^{-1}, Roche) were added, and samples were incubated for another 2 h at 42 °C. The resulting proteinase K-treated samples were subjected to a phenol–chloroform extraction followed by purification of DNA with a QIAquick PCR Purification column (Qiagen) in preparation for high-throughput sequencing.

High-throughput sequencing

Purified, unamplified input or ChIP DNA (see 'MNase-digested chromatin and native ChIP' section) was quantified using a 2100 Bioanalyzer High Sensitivity kit (Agilent). DNA libraries were prepared for multiplexed sequencing according to Illumina recommendations as described[67] with minor modifications using NEB enzymes. In brief, 5 ng input or ChIP DNA was end-repaired and A-tailed. TruSeq adaptors (Illumina) were ligated, libraries were size selected to exclude polynucleosomes and adapter-modified DNA fragments were enriched by PCR using KAPA polymerase. Libraries were assessed by Bioanalyzer and the degree of nucleosome digestion for each experiment was assessed to avoid any potentially over digested samples. Libraries were submitted for 150-bp, paired-end sequencing on a NextSeq 500 instrument (Illumina).

Paired-end sequencing analysis

Paired-end sequencing analysis was performed as previously described[21] with some modifications. In brief, SeqPrep (https://github.com/jstjohn/SeqPrep) was used to remove adapters and merge paired-end Illumina reads using a minimum overlap parameter adjusted for read length (-q 25 -L 25 -o 35). FastQC (v.0.11.8) was used before and after SeqPrep to assess read quality. To analyse minor and major satellite DNA, we used a custom tandem repeat analysis as described[67] with the following modifications. Joined reads were aligned to a trimerized mouse minor satellite consensus (GenBank: X14464.1)[68] to a dimerized mouse major satellite consensus (GenBank: V00846.1)[69] or to the reverse complement of those tandem consensus sequences using the MATLAB localign function (MATLAB R2020a, MathWorks). Joined reads that aligned with at least 80% identity were chosen for further analysis. To calculate the percentage of total reads, the number of joined reads aligning to the consensus sequence in either the forward or reverse complement orientation (without double counting any joined read) was divided by the total number of joined reads. ChIP fold enrichment was calculated as the fraction of reads mapping to the minor (or major) satellite from the ChIP divided by the fraction of reads mapping to the minor (or major) satellite in the input.

Analysis of HMGA1 ChIP–seq data

HMGA1 ChIP–seq data (Illumina HiSeq 2500, 50 bp read length, single end) from *M. musculus* embryonic stem cells were downloaded from the Sequence Read Archive (SRA; accession numbers: SRR5749536, SRR5749537, SRR5749546 and SRR5749545)[40]. Single-end sequencing analysis was performed as follows. TrimGalore (https://github.com/FelixKrueger/TrimGalore) was used for adapter and quality trimming (--illumina -e 0.1 --stringency 3 --quality 20). To analyse minor and major satellite DNA, we used the custom tandem repeat analysis described in 'Paired-end sequencing analysis' section. Furthermore, the HMGA1 ChIP–seq data were mapped to high-fidelity long-read sequencing (LRS) reads containing continuous minor and major satellite arrays[70]. LRS data were downloaded from the SRA (accession number SRR11606870; minor satellite SRR11606870.111923, major satellite SRR11606870.2342980). Sequence alignment was performed using Bowtie2 (bowtie2 --end-to-end --very-sensitive --no-mixed --no-discordant -q--phred33)[71].

A+T stretch and tetranucleotide analysis

Publicly available long reads generated using a PacBio Sequel II System with HiFi sequencing from C56BL/6J *M. musculus* strain[72] were analysed using Python 3.11 to count the number of A+T stretches per repeat in major and minor satellite. First, reads containing only major

or minor satellite motifs were isolated (major satellite: 37,744 reads with GAAAACTGAAAA motif; minor satellite: 4,045 reads with CENP-B box ATTCGTTGGAAACGGGA). Second, 500 reads per satellite (15–20 kb long) were isolated via BLAST 2.6.0 package using representative major (SRR11606870.2342980) or minor (SRR11606870.111923) satellite homogenous type I contiguous arrays as queries[70]. Third, an average of minimum 3-, 4- and 5-bp-long A+T stretches (non-overlapping) or an average of tetranucleotides associated with narrow minor groove[35] (overlapping) was computed for each of 500 reads per satellite.

HMGA1 sequence analysis

The *M. musculus* canonical *Hmga1* coding sequence (ENSMUST00000118599.9) was used as a query to identify *M. spretus Hmga1* in the non-annotated genomic assembly (GCA_001624865.1) using BLAST (tblastn) and synteny analysis.

Immunofluorescence microscopy and fluorescence in situ hybridization

Liver nuclei from *M. musculus* or *M. spretus* were cytospun onto Shandon SuperFrost Plus slides (Thermo Fisher Scientific) in a cytofunnel (Thermo Fisher Scientific) at 1,500 rpm for 5 min prior to fixation in 4% formaldehyde in PBS for 10 min, permeabilized with 0.2% Triton X-100 in PBS for 5 min, blocked in immunofluorescence blocking solution (2% FBS, 2% BSA and 0.1% Tween-20 in PBS) for 20 min and then incubated in primary (45 min) and secondary (20 min) antibodies diluted in immunofluorescence blocking solution with washes in 0.1% Tween in PBS after each antibody incubation. Rabbit anti-H3K9me3 (1 μg ml^{-1}; Abcam; ab8898) followed by fluorophore-conjugated secondary antibody Cy5 anti-rabbit (1:200; Jackson ImmunoResearch Laboratories; 711-175-152) were used. After immunofluorescence treatment, cells were post-fixed in 4% formaldehyde in PBS for 10 min and further processed for fluorescence in situ hybridization using CENP-B-Cy3 PNA probes (Panagene) for minor satellite detection following the manufacturer's protocol. In brief, cells were pretreated with 100 μg ml^{-1} RNase A followed by 0.005% pepsin, fixed in 4% formaldehyde, dehydrated, air-dried, hybridized with 200 nM PNA probe, washed and counterstained with DAPI before mounting in Vectashield (Vector Laboratories).

SYTOX staining

Liver nuclei were cytospun onto slides as above, and fixed in 4% formaldehyde in PBS for 10 min. Immunofluorescence staining was performed as above, using anti-centromere antibody (1:500; Antibodies Incorporated 15-235) and Cy5 anti-human antibody (Jackson ImmunoResearch Laboratories; 709-175-149) to detect the centromeres. Cells were again fixed in 4% formaldehyde in PBS for 10 min, treated with 100 μg ml^{-1} RNase A in 2× SSC for 1 h at 37 °C, counterstained with DAPI in 2× SSC for 5 min followed by 167 nM SYTOX Green (Invitrogen) in 2× SSC for 30 min, with rinses in 2× SSC after each stain. The slides were rinsed in distilled water prior to mounting. The Pearson's correlation coefficient between SYTOX intensity and DAPI intensity was quantified from deconvolved images of nuclei (20 *z* slices, elliptical region of interest of constant size across images using the ImageJ[73] plug-in Coloc 2 (point spread function 2 pixels)).

Oocyte collection and culture

Female mice were primed with pregnant mare somatic gonadotropin (*M. musculus*; 5 U per mouse; 367222; Calbiochem) or CARD HyperOva (*M. spretus* and *M. musculus* × *spretus* hybrid; 150 μl per mouse; Cosmo Bio USA; KYD-010-EX-X5) injected into the intraperitoneal cavity 38–48 h prior to oocyte collections to induce superovulation. The ovaries were isolated in M2 medium (M7167; Sigma-Aldrich Milipore) with 2.5 mM phosphodiesterase 3 inhibitor milrinone (M4659; 2.5 mM; Sigma-Aldrich Milipore) to block maturation. Germinal vesicle oocytes were collected, denuded mechanically from cumulus cells using a mouth pipette with plastic 100-μm-diameter stripper tips (CooperSurgical; MXL3-100) and incubated for at least 1 h on a hot plate (38 °C) in mineral oil before microinjection (9305; FUJIFILM Irvine Scientific).

Plasmid and RNA preparation

A pGEMHE-Trim21-OLLAS plasmid used for Trim-away experiments was generated by subcloning *Trim21* from the pGEMHE–mCherry–Trim21 plasmid (kind gift from M. Schuh; Addgene; no. 105522; ref. 43) into a pGEMHE backbone with C-terminal OLLAS tag. pGEMHE–Borealin–OLLAS was generated by cloning *Borealin* (also known as *Cdca8*) from a cDNA library of mouse testis and subcloning it into a pGEMHE backbone with C-terminal OLLAS tag. Inserts and backbones were linearized by PCR reaction using KAPA Polymerase HiFi HotStart ReadyMix (KK2602; Roche) and ligated using In-Fusion Snap Assembly (638948; Takara). Primers were designed using SnapGene 7.1.1 (Dotmatics) software. Other plasmids were generated elsewhere: pIVT–dCas9–mCherry expressing catalytically dead Cas9, which targets the minor satellite through a gRNA that recognizes the 5′-ACACTGAAAAACACATTCGT-3′ sequence[74]; pTALYM3–TALE–mClover (kind gift from M.-E. Torres-Padilla; Addgene; no. 47874) expressing a TALE protein that recognizes the 5′-TGCCATATTCCACGT-3′ sequence of the major satellite repeat[24], pIVT–Mad1–2×GFP[48]. All plasmids were amplified in *Escherichia coli* Stellar Competent Cells (636763; Takara) and purified using a NucleoSpin kit (740588.25; Takara). mRNA was prepared using T7 mScript Standard mRNA Production System (C-MSC100625; CellScript) in vitro transcription kit.

Oocyte microinjection

Oocytes were microinjected with approximately 5 pl of mRNAs in M2 medium with 2.5 mM milrinone and 3 mg ml^{-1} BSA at room temperature with a micromanipulator TransferMan NK 2 or TransferMan 4r (Eppendorf) and picoinjector (Medical Systems Corp). Oocytes were then incubated in 30–50 μl drops of Chatot–Ziomek–Bavister (CZB) medium (MR019D; Thermo Fisher Scientific) in mineral oil (M5310; Sigma-Aldrich Milipore) at 37.8 °C and 5% CO_2 (Airgas) for 16 h to allow protein expression. To deplete HMGA1, the Trim-away approach was used[43]. In brief, milrinone-arrested oocytes were microinjected with approximately 5 pl mix containing *Trim21–OLLAS* mRNA and 0.5 mg ml^{-1} of purified polyclonal normal rabbit IgG (used as control; 12-370; Sigma-Aldrich Milipore) or recombinant rabbit monoclonal anti-HMGA1 antibodies (EPR7839 clone; ab226112; Abcam). The mix was supplemented with the following reagents (see captions of Figs. 1–5): TALE–mClover 200 ng $μl^{-1}$; Mad1–2×GFP 100 ng $μl^{-1}$; Trim21–OLLAS 300–600 ng $μl^{-1}$; dCas9–mCherry 120–160 ng $μl^{-1}$; gRNA 20–30 ng $μl^{-1}$; Borealin–OLLAS 100 ng $μl^{-1}$. Oocytes were then incubated for at least 3 h to enable mRNA expression before the cells were released by washing out milrinone. This was achieved by passing cells through five 100-μl drops of CZB medium. Antibodies were passed through Amicon Ultra-0.5 100-K columns (Sigma-Aldrich Milipore, UFC5100625) to remove preservatives and glycerol following an established protocol[75]. IGEPAL at 0.05% final concentration was added to facilitate injections of viscous mRNA and protein solutions.

Oocyte immunofluorescence

Oocytes were fixed in 2% paraformaldehyde dissolved in PBS (pH 7.4) for 20 min at room temperature with 0.1% Triton X-100 for 15 min at room temperature, placed in blocking solution (PBS with 0.3% BSA and 0.01% Tween-20) overnight at 4 °C, incubated for 1 h with primary antibody in blocking solution, washed three times for 15 min each, incubated for 1 h with secondary antibody, washed three times for 15 min each and mounted in Vectashield with DAPI (H-1200; Vector) to visualize chromosomes. Kinetochores were labelled using mouse anti-HEC1 antibody (C-11; sc-515550; Santa Cruz; 1:100–200). Recombinant rabbit monoclonal anti-HMGA1 antibody (EPR7839 clone; ab226112; Abcam; at 0.1 μg/ml concentration) was used to determine HMGA1 enrichment at pericentromeres and AlexaFluor 488-conjugated anti-HMGA1 antibody

(EPR7839 clone; ab204667; Abcam) was used to assess depletion efficiency in meiosis I (1:500) and to test binding in netropsin-treated germinal vesicle oocytes (1:5,000). Rabbit anti-H3K9me3 (1:500; Abcam; ab8898) antibody was used to measure heterochromatin in hybrid *M. musculus* × *spretus* oocytes. Rat monoclonal anti-OLLAS (1:100; clone L2; NBP1-06713; Novus Biologicals) antibody was used to detect Borealin-OLLAS. To stain microtubules, oocytes were fixed in preheated 2% paraformaldehyde dissolved in PEM buffer (200 mM PIPES; 20 mM EGTA; 2 mM $MgCl_2$; pH 6.9) with 0.5% Triton X-100 for 20 min at 38 °C, washed with 0.5% Triton X-100 in PEM buffer for 15 min at 38 °C, placed in blocking solution (PBS with 0.3% BSA and 0.01% Tween-20) overnight at 4 °C, incubated for 1 h with mouse monoclonal anti-α-tubulin (1:1,000; DM1A clone; T6199; Sigma-Aldrich Milipore) antibody in blocking solution, washed three times for 15 min each, incubated 1 h with secondary antibody and washed three times for 15 min each before mounting. The secondary antibodies used were: donkey anti-rabbit AlexaFluor 488 (1:500, Invitrogen), anti-mouse AlexaFluor 555 (1:500, Invitrogen), anti-mouse AlexaFluor 594 (1:500, Invitrogen), anti-mouse AlexaFluor 647 (1:500, Invitrogen) and anti-rat AlexaFluor 647 (1:500, Invitrogen). To test HMGA1 competition for narrow DNA minor groove, germinal vesicle *M. musculus* oocytes were treated with 200 µM netropsin (Sigma-Aldrich Milipore; N9653-5MG) for 1 h prior to fixation and subsequent staining with AlexaFluor-488-conjugated anti-HMGA1 antibody (EPR7839 clone; ab204667; Abcam) performed as described above.

Image acquisition

Images of liver cells were captured on an inverted fluorescence microscope (DMI6000 B; Leica) equipped with a charge-coupled device camera (ORCA AG; Hamamatsu Photonics) and a 100×, 1.4 numerical aperture (NA) oil-immersion objective. Images were collected as 0.2 µm *z* sections and *z* series were deconvolved using LAS-AF 3.6.0.20104 (Leica). Images of oocytes were collected as *z* stacks (31 or 41 slices) at 0.5-µm intervals to visualize the entire meiotic spindle using a confocal microscope (DMI4000B; Leica) equipped with a 63×, 1.3 NA glycerol-immersion objective lens, an *xy*-piezo *z* stage (Applied Scientific Instrumentation), a spinning disk (Yokogawa Corporation of America) and an electron multiplier charge-coupled device camera (ImageEM C9100-13; Hamamatsu Photonics), controlled by MetaMorph 7.5 software (Molecular Devices). Excitation was carried out with a Vortran Stradus VersaLase 4 laser module with 405-, 488-, 561- and 639-nm lasers (Vortran Laser Technology). Panels of microscopy images were prepared using ImageJ 1.53t software[73]. Oocytes extruding polar bodies were imaged using transmitted light with a Nikon Eclipse TE2000-U inverted light microscope equipped with a 10× objective using 1.5× magnification, RS RoperScientific 7360-0001 camera and MetaMorph 7.5 software (Molecular Devices).

Image analysis

To quantify relative HMGA1 intensity at pericentromeres in *M. musculus*, *M. spretus* and hybrid bivalents, sum intensity projections were used. Circular six-pixel-diameter regions were drawn to measure intensities at major (*M. musculus*) or minor (*M. spretus*) satellites and the nearby cytoplasm. The mean HMGA1 background intensity (cytoplasm) was subtracted from the mean HMGA1 intensity at pericentromeres. Values were obtained for ten randomly selected, non-overlapping hybrid bivalents for each cell. HMGA1 depletion efficiency was calculated on the basis of maximum projection images. A DAPI-based mask was created for each image using identical intensity thresholds for control and HMGA1-depleted cells and the ImageJ 1.53t software plug-in Analyze Particles (size 3-inifinity; circularity 0.00–3.00). The mean HMGA1 intensity in the mask was measured, and a 50 × 50-pixel box drawn at the cell periphery was used to subtract the cytoplasmic background. HMGA1 intensity at the chromocentres in control and netropsin-treated *M. musculus* oocytes was calculated using custom-size circular regions of interest on the basis of the mean intensity projection of each chromocentre. First, the mean cytoplasmic HMGA1 signal (background) was subtracted from the mean chromocentre HMGA1 signal and mean non-chromocentre chromatin HMGA1 signal. Second, the intensity ratio of the chromocentre HMGA1 signal over the non-chromocentre chromatin HMGA1 signal was computed. Pericentromeric length in *M. musculus* and hybrid oocytes was measured by connecting chromosome arms and the nearest kinetochore signal linked by a thread of a major (*M. musculus*) or minor (*M. spretus*) satellite on the basis of maximum projection images. Owing to poor kinetochore staining, the pericentromeric length in *M. spretus* oocytes was measured as the longest axis of the minor satellite signal on the basis of maximum projection images. The ratio of major to minor satellite length was reported to estimate the relative stretching of *M. musculus* and *M. spretus* pericentromeres. To measure the kinetochore size in oocytes, circular five-pixel-diameter regions were drawn, and mean HEC1 intensities were calculated and corrected for local cytoplasmic background on the basis of sum intensity projection images. Intensities for eight randomly selected distinct kinetochores were obtained for each cell and the mean intensity was calculated for each cell. Kinetochore clustering in oocytes was calculated using the Analyze Particles ImageJ plug-in (size 3-inifinity; circularity 0.00–1.00) on the basis of maximum intensity projection images. Aster eccentricity in oocytes arrested with 10 µM S-trityl-L-cysteine; 164739 (STLC; Sigma-Aldrich Milipore) was measured using a modified MTAster pipeline (CellProfiler v.2.1.1)[76]. In brief, maximum intensity projection images were used to identify microtubule asters as objects of 80–500 pixels in diameter. Objects outside the diameter or adjacent to the border of the image were discarded. The threshold strategy and method were 'adaptive' and 'otsu', respectively, with two-class thresholding, weighted variance, smoothing method selected automatically and correction factor 4.0. The lower and upper bounds on the threshold were 0.0 and 1.0. The image size was used to calculate the adaptive window size. The remaining parameters were left unchanged. Kinetochore–microtubule attachment stability in hybrid oocytes was measured by quantifying the MAD1 intensity at the kinetochores using the automated Python-built Centrocalc tool[77]. In brief, kinetochores were identified automatically as spots with a radius of 150 nm and local maxima were found using the HEC1 signal. Up to 38 spots were selected, separated by a minimum of 2 pixels and 3D ellipsoid regions of 4 × 4 × 3 pixels were drawn. HEC1 and MAD1 intensities were calculated as mean greyscale pixel values corrected for local cytoplasmic background. Each kinetochore was manually assigned as *M. musculus* or *M. spretus* on the basis of the size of the minor satellite signal. The mean intensity of each cell was calculated for each species. Only spots associated with a minor satellite signal were analysed.

Statistics

All experiments were carried out in two to three independent biological replicates; exact sample sizes are listed in the captions of Figs. 1–5. The analyses were performed using Microsoft Excel and GraphPad Prism 10. All graphs were generated in GraphPad Prism 10. Two-tailed *t*-tests and one-way ANOVAs were used for analysing data that followed a normal distribution. Two-tailed Mann–Whitney *U*-tests and Kruskal–Wallis tests were used to analyse data that did not follow a normal distribution. Corrections for comparing multiple variables and exact *P* values are listed in the captions of Figs. 1–5. Sample size choice was dictated by the ability to derive meaningful statistics, typically by selecting at least ten oocytes or liver cells for each condition per experiment. The investigators were not blinded during data collection and quantification because the phenotypic changes were visibly distinguishable for analysed conditions. Cells analysed were selected at random.

Reporting summary

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

Data availability

The ChIP–seq data generated in this study have been deposited in the SRA under BioProject accession code PRJNA1074844. The analysed datasets are publicly available under SRA Experiment accession codes: *M. musculus* input, SRX2939766, SRX2939768, SRX2939769; *M. musculus* CENP-A ChIP–seq, SRX2939767, SRX2939772, SRX2939773; *M. musculus* H3K9me3 ChIP–seq, SRR27928404; *M. spretus* input, SRR27922556, SRR27922557, SRR27922558; *M. spretus* CENP-A ChIP–seq, SRR27922553, SRR27922554, SRR27922555; *M. spretus* H3K9me3 ChIP–seq, SRR27922552; input and HMGA1 ChIP–seq, SRR5749536, SRR5749537, SRR5749546 and SRR5749545; long-read *M. musculus* genome assembly, SRR11606870. All other data needed to evaluate our conclusions are included in the paper and its Supplementary Information. Source data are provided with this paper.

Code availability

Python code used to analyse A+T stretches and tetranucleotides is available at Github (https://github.com/DDudka9/DNA-shape.git).

67. Hasson, D. et al. The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* **20**, 687–695 (2013).
68. Wong, A. K. & Rattner, J. B. Sequence organization and cytological localization of the minor satellite of mouse. *Nucleic Acids Res.* **16**, 11645–11661 (1988).
69. Horz, W. & Altenburger, W. Nucleotide sequence of mouse satellite DNA. *Nucleic Acids Res.* **9**, 683–696 (1981).
70. Packiaraj, J. & Thakur, J. DNA satellite and chromatin organization at mouse centromeres and pericentromeres. *Genome Biol.* **25**, 52 (2024).
71. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie2. *Nat. Methods* **9**, 357–359 (2012).
72. Hon, T. et al. Highly accurate long-read HiFi sequencing data for five complex genomes. *Sci. Data* **7**, 399 (2020).
73. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
74. Kumon, T. et al. Parallel pathways for recruiting effector proteins determine centromere drive and suppression. *Cell* **184**, 4904–4918 (2021).
75. Clift, D., So, C., McEwan, W. A., James, L. C. & Schuh, M. Acute and rapid degradation of endogenous proteins by Trim-Away. *Nat. Protoc.* **13**, 2149–2175 (2018).
76. Grenfell, A. W. et al. A versatile multivariate image analysis pipeline reveals features of *Xenopus* extract spindles. *J. Cell Biol.* **213**, 127–136 (2016).
77. Dudka, D., Akins, R. B. & Lampson, M. A. FREEDA: an automated computational pipeline guides experimental testing of protein innovation. *J. Cell Biol.* **222**, e202212084 (2023).

Acknowledgements We thank J. Ma for assistance, C. Brand for comments on the manuscript, members of the Philadelphia Chromosome Club for discussions on our findings, A. Das for experimental advice and Y. Yang for preliminary HMGA1 Trim-away experiments. Funding was from NIH grants GM130302 (B.E.B.), GM122475 (M.A.L.) and HD058730 (M.A.L. and B.E.B.).

Author contributions J.M.D.-M., M.A.L. and B.E.B. conceived the project. D.D., X.S., K.B. and T.A. performed oocyte experiments. J.M.D.-M. performed somatic cell and sequencing experiments. All authors analysed data. D.D. and X.S. generated reagents. D.D., M.A.L. and B.E.B. wrote the paper. D.D., J.M.D.-M., X.S. and T.A. edited the paper. M.A.L. and B.E.B. supervised the project.

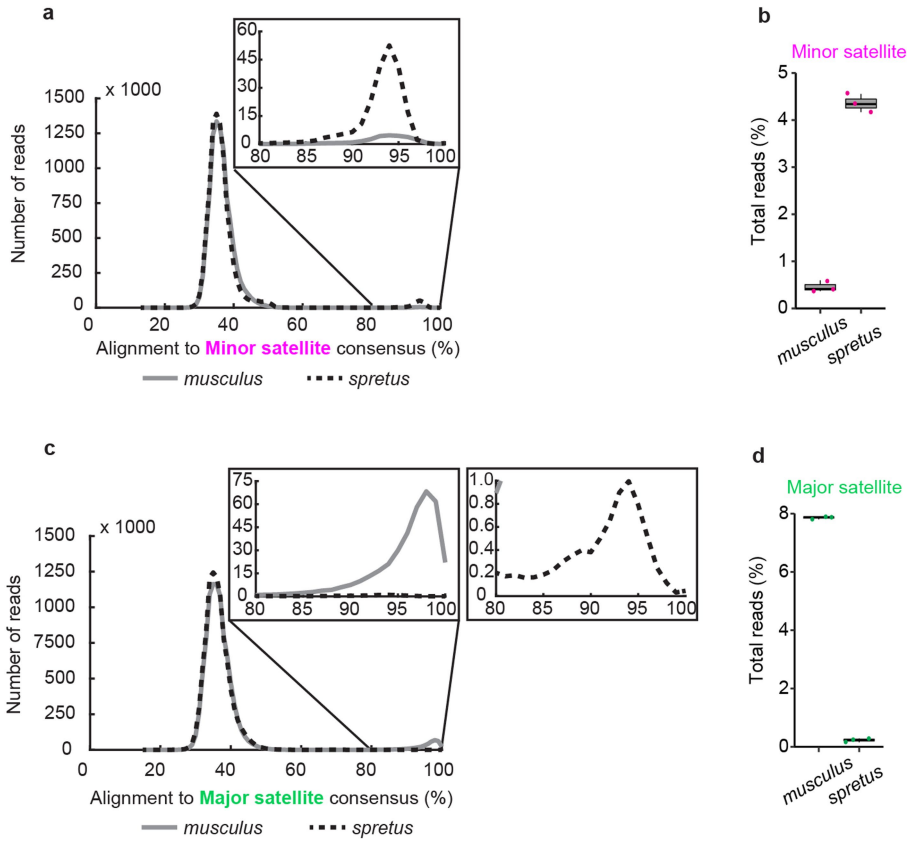
Competing interests The authors declare no competing interests.

Additional information

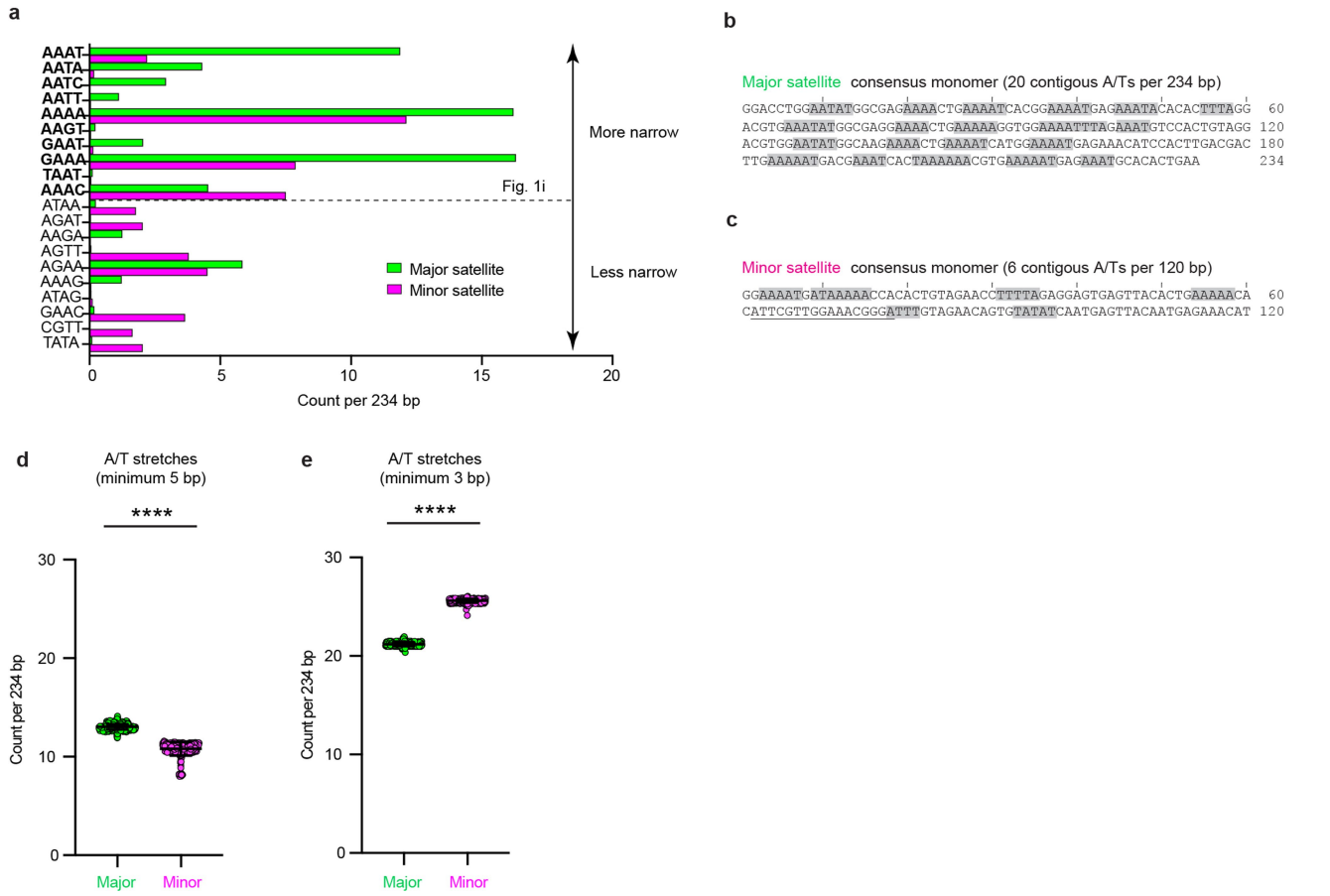
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08374-0.

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

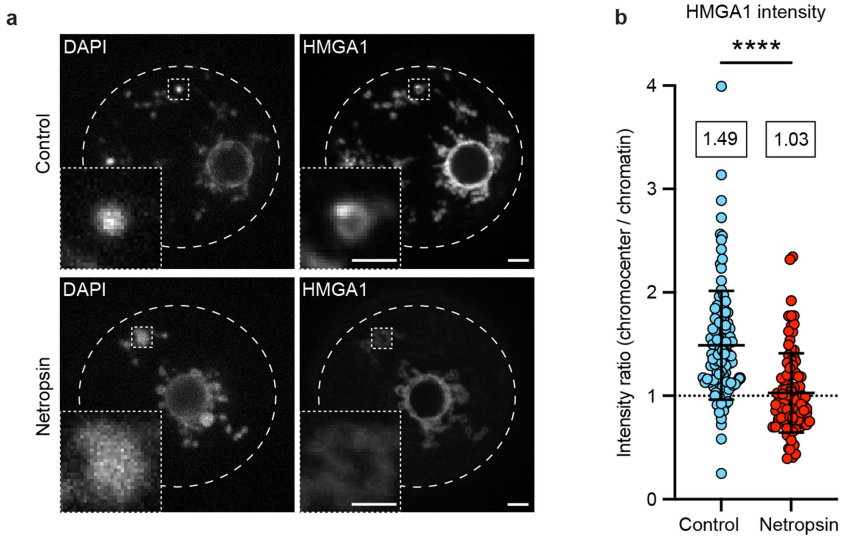
Peer review information *Nature* thanks Nicolas Altemose, Pragya Sidhwani and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



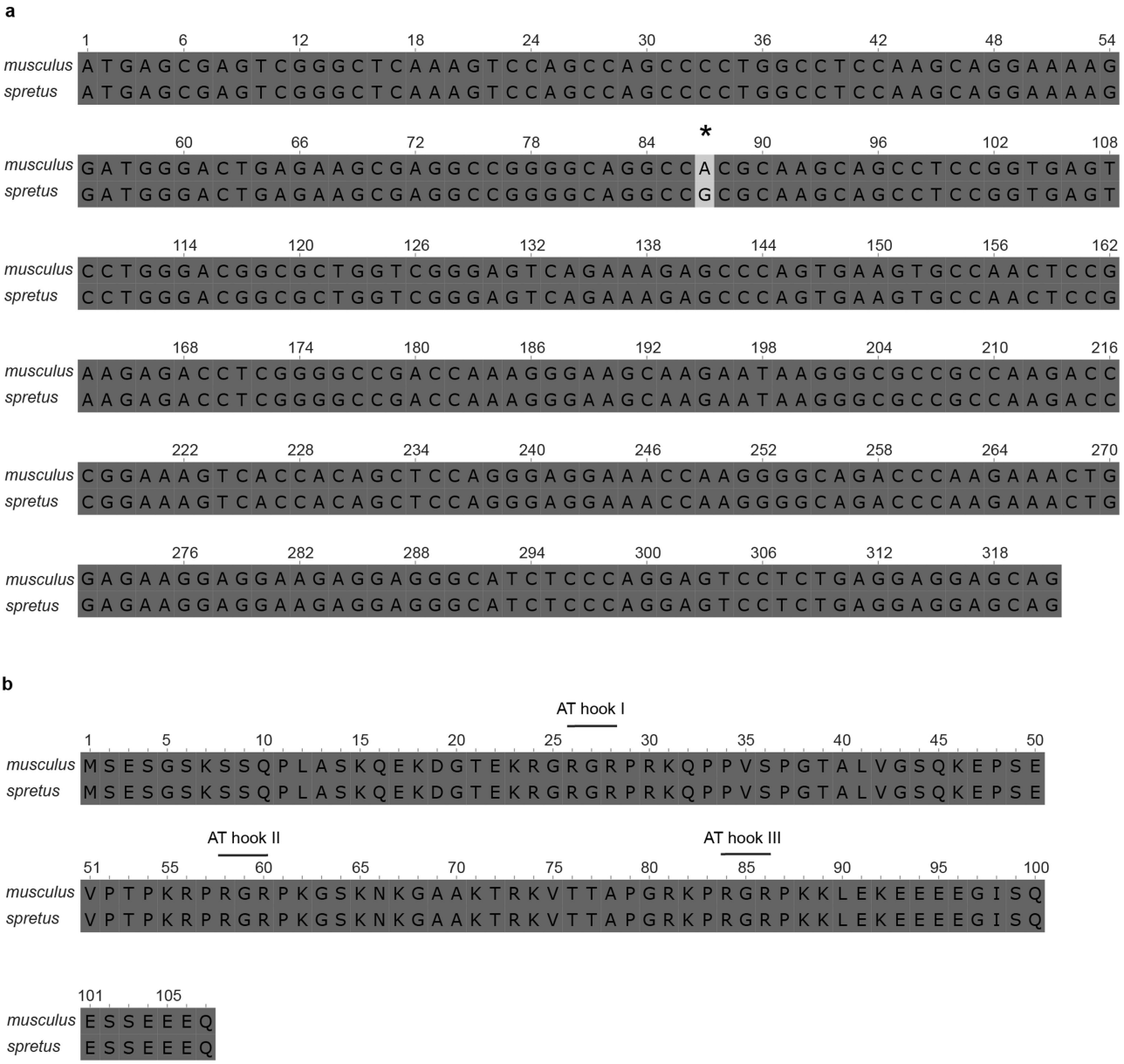
Extended Data Fig. 1 | *M. musculus* and *M. spretus* harbor different amounts of major and minor satellites. a-d, High-throughput sequencing of MNase-digested *M. spretus* or *M. musculus* chromatin. Reads were aligned to a trimer of minor satellite (**a**,**b**) or a dimer of major satellite (**c**,**d**) consensus sequences (see Methods). Histograms show distribution of reads aligning to minor (**a**) or major (**c**) satellite, with 80–100% range expanded in insets. The percent of reads that aligned with ≥ 80% identity to minor (**b**) or major (**d**) satellite is plotted as a Tukey box and whiskers plot (N = 3 independent experiments). Boxplots indicate median, 25th and 75th percentiles (hinges), smallest and largest value at most 1.5 * interquartile range of the hinge (whiskers).



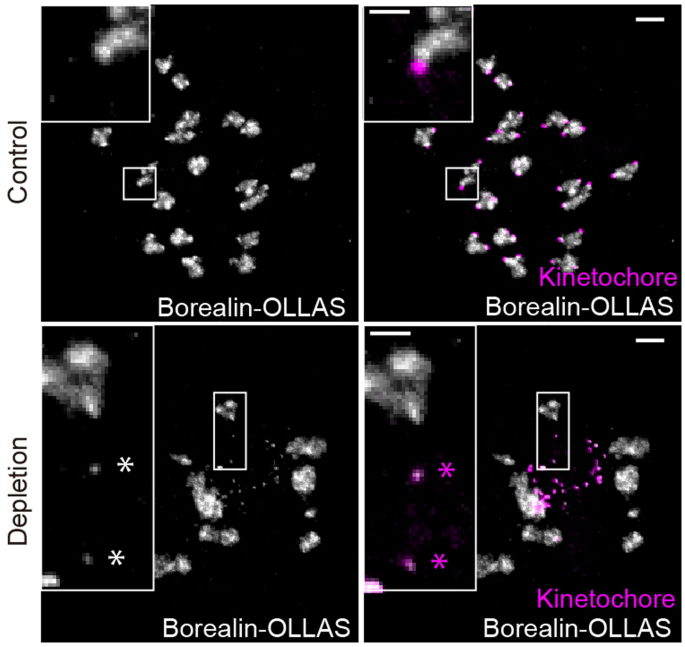
Extended Data Fig. 2 | *M. musculus* and *M. spretus* pericentromeric satellites differ in AT-rich tetranucleotides and frequencies of contiguous A/Ts. a, Number of AT-rich tetranucleotides associated with narrow DNA minor groove[35] per length of a single major satellite repeat (234 bp) averaged from 500 *M. musculus* major or minor satellite arrays. Note that the ten tetranucleotides with narrowest minor grooves (top, bold) are counted along representative arrays in Fig. 1i. **b**,**c**, Major (**b**) or minor (**c**) satellite consensus monomer sequences with stretches of 4 or more consecutive A/Ts shaded in grey; the 17-bp CENP-B box is underlined. **d**,**e**, Number of minimum 5 bp- (**d**) or 3 bp-long (**e**) contiguous A/T stretches in major or minor satellite arrays per 234 bp length of a single major satellite repeat). Mean and standard deviation indicated, n = 500 arrays; two-tailed Mann–Whitney test, ****$P < 0.0001$.



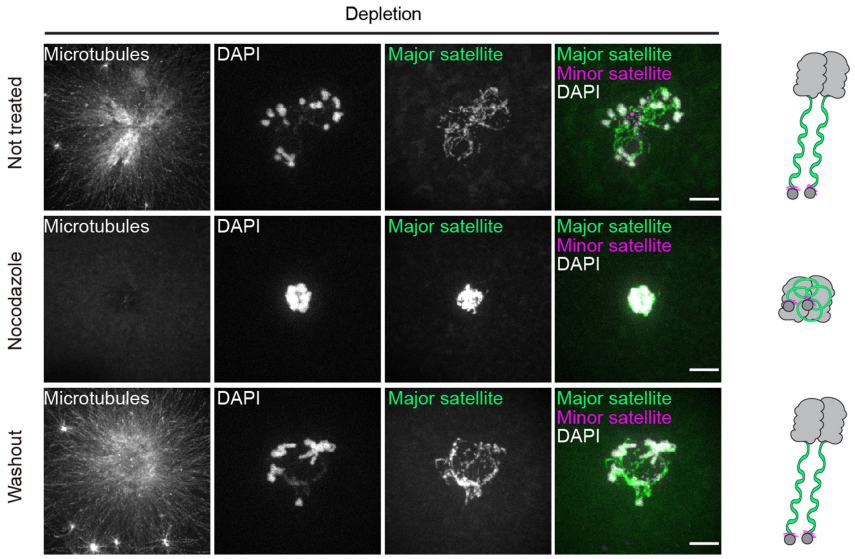
Extended Data Fig. 3 | HMGA1 competes with netropsin for binding to major satellite. a,**b**, Control or netropsin-treated *M. musculus* oocytes were fixed in germinal vesicle stage and stained for DNA (DAPI) and HMGA1 (**a**). Scale bars, 5 µm and 2 µm (insets). The ratio of HMGA1 signal at chromocentres to non-chromocentre chromatin was quantified (**b**). Mean and standard deviation indicated, means shown in boxes, n = 40 (control) and 30 (netropsin) oocytes from 3 independent experiments, with an average of 3 chromocenters measured in each oocyte; two-tailed Mann–Whitney test, ****$P < 0.0001$.



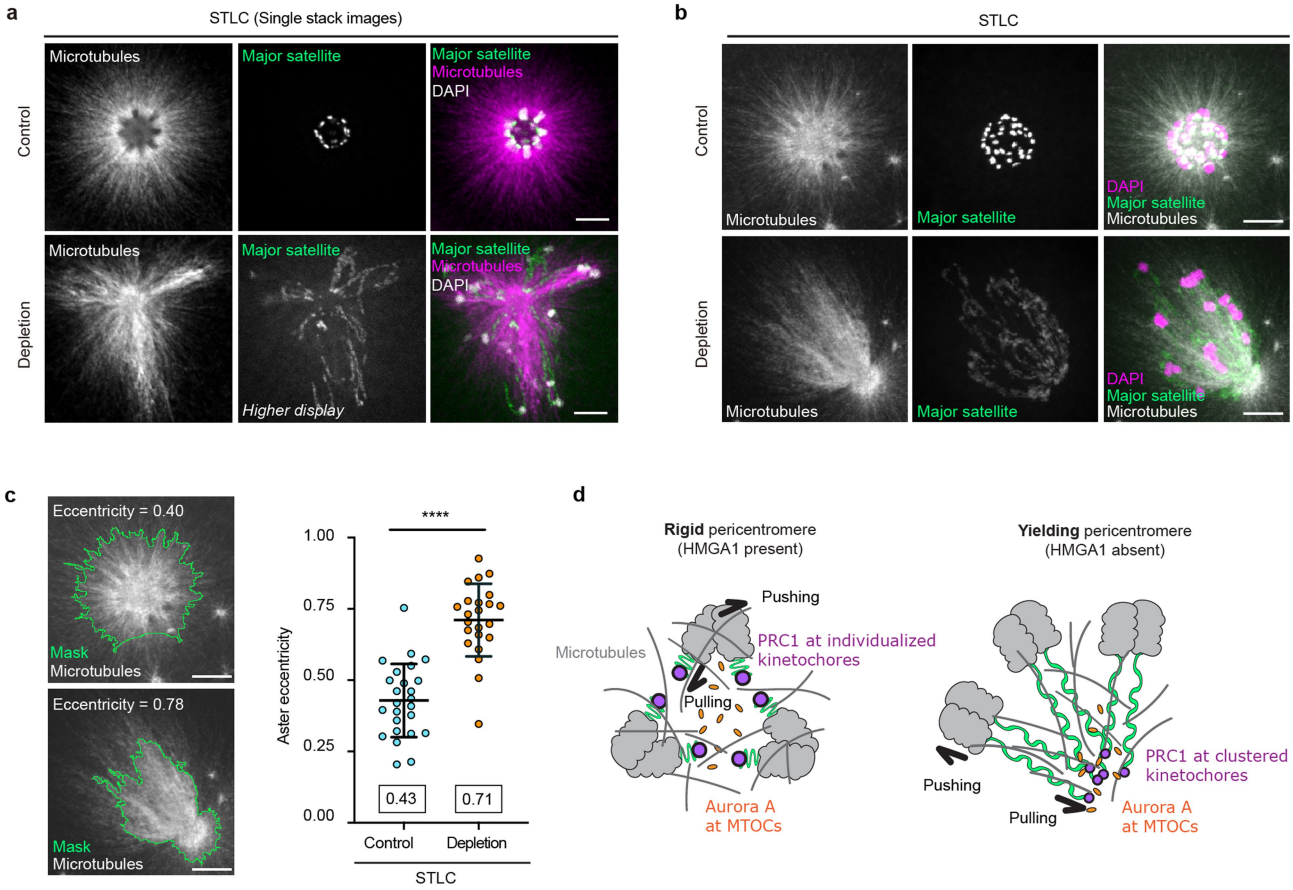
Extended Data Fig. 4 | HMGA1 protein sequence is identical between *M. musculus* and *M. spretus*. **a**,**b**, Alignment of *M. musculus* and *M. spretus* HMGA1 coding (**a**) and amino acid (**b**) sequences. Asterisk (**a**) marks a single synonymous substitution. AT hooks (RGR motifs) are shown (**b**) based on[5].



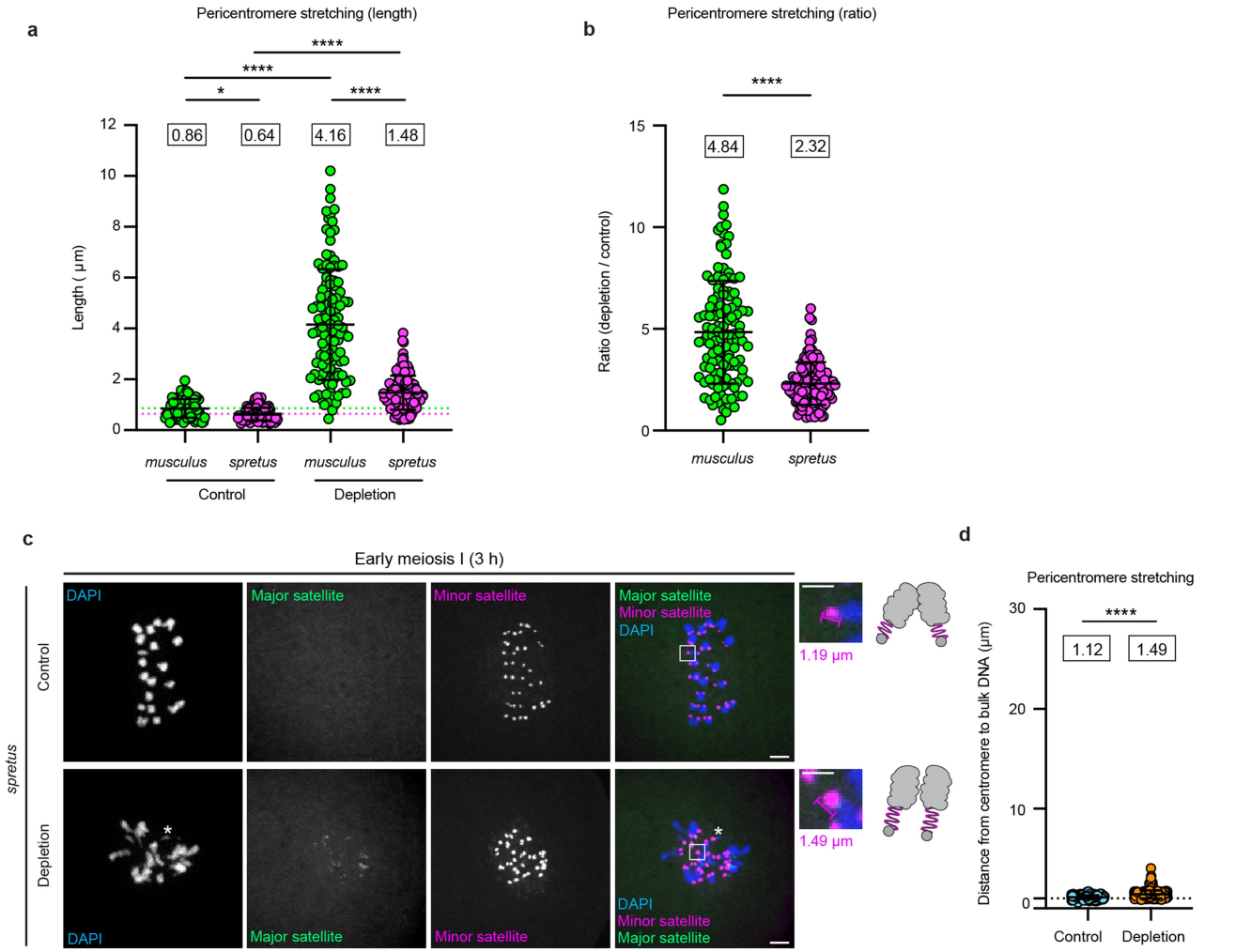
Extended Data Fig. 5 | Kinetochores displaced from chromosome arms remain associated with centromeric chromatin. Immunofluorescence images of control or HMGA1-depleted *M. musculus* oocytes represented in Fig. 3f. Cells were expressing Borealin tagged with the OLLAS epitope to visualize centromeric chromatin. Colocalization of Hec1 (kinetochore; magenta asterisks) and centromeric Borealin-OLLAS is shown (white asterisks). Note that Borealin also localizes to bulk chromatin. Scale bars 5 µm and 2 µm (insets). 3 independent experiments showed the same result.



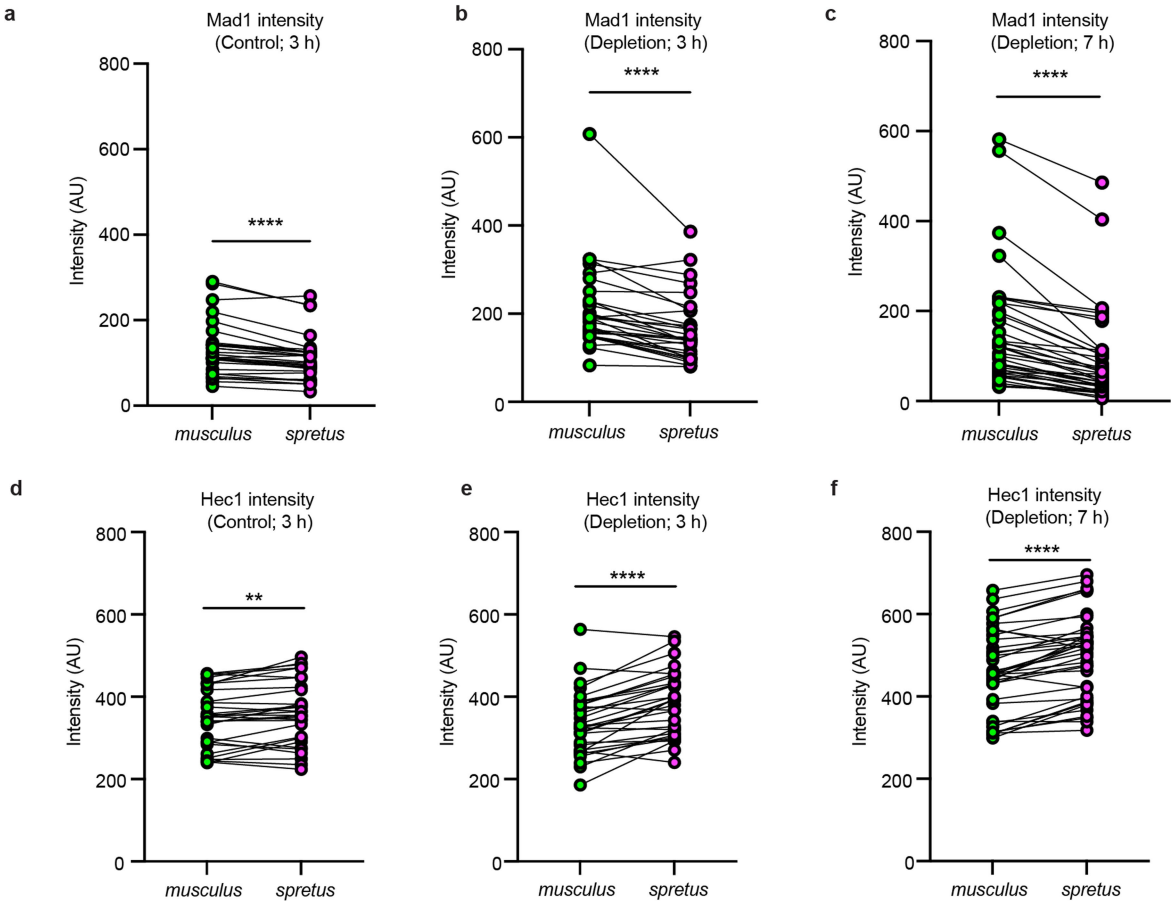
Extended Data Fig. 6 | Microtubules mediate major satellite stretching in HMGA1-depleted oocytes. Immunofluorescence images of HMGA1-depleted *M. musculus* oocytes expressing major-satellite-targeting TALE-mClover and treated with nocodazole to depolymerize microtubules. Oocytes were fixed in early meiosis I before removing microtubules ("not treated"), during treatment ("nocodazole"), or after microtubules were allowed to regrow ("washout"). DAPI stained DNA and anti-α-tubulin antibody-stained microtubules. Scale bar 10 µm. 2 independent experiments showed the same result (17 control- and 15 nocodazole-treated oocytes). Nocodazole washout was done once in 10 oocytes, all showing the same result.



Extended Data Fig. 7 | HMGA1 depletion disrupts microtubule organization. **a-c**, Control and HMGA1-depleted *M. musculus* oocytes expressing major satellite-targeting TALE-mClover were treated with a kinesin-5 inhibitor (STLC) in early meiosis I to arrest cells before spindle bipolarization, fixed, and stained for DNA (DAPI) and microtubules (anti-α-tubulin). 2 independent experiments showed the same result. Single z-slices (**a**) from confocal stacks show the concentric ring formed by chromosomes and radial symmetry of the microtubule aster in control oocytes, but lack of symmetry in HMGA1-depleted oocytes. Note that major satellite signal is enhanced in HMGA1-depleted oocyte to visualize thin threads of DNA. Maximal intensity projections (**b**) show the entire microtubule aster and were used to automatically generate a mask around the aster ((**c**), green outline) using CellProfiler (see Methods). Scale bars 10 µm. Note that A and B show different cells. For control and HMGA1-depleted *M. musculus* oocytes treated with STLC, aster eccentricity was quantified on a scale of 0 (perfect circle) to 1 (straight line) ((**c**), mean and standard deviation indicated, means shown in boxes, n = 25 (control) and 23 (depletion) oocytes from 2 independent experiments; unpaired two-tailed *t*-test, $^{****}P < 0.0001$). **d**, HMGA1 packages major satellite in *M. musculus* oocytes to withstand forces exerted by microtubules and microtubule motors that pull kinetochores toward spindle poles while pushing chromosomes away from the poles. In the absence of HMGA1, major satellite-built pericentromeres lose rigidity and yield under spindle forces, resulting in stretching, kinetochore clustering, and reduction in kinetochore size. Kinetochore clustering may destabilize microtubule attachments because of proximity to Aurora A kinase at microtubule organizing centers (orange, MTOCs). Disrupted kinetochore organization may also interfere with the oocyte spindle assembly pathway dependent on PRC1 at kinetochores (purple).



Extended Data Fig. 8 | Disproportionate impact of HMGA1 depletion on packaging of *M. musculus* and *M. spretus* pericentromeres. a, Quantification of *M. musculus* and *M. spretus* pericentromere length in control or HMGA1-depleted hybrid *M. musculus* × *spretus* oocytes (means and standard deviations are indicated, means shown in boxes, n = 95 (control) and 120 (depletion) pericentromeres from 2 independent experiments; Kruskal–Wallis test followed by Dunn's multiple-comparison test, ****$P < 0.0001$, *$P = 0.0466$). Dotted lines indicate mean *M. musculus* (green) or *M. spretus* (magenta) pericentromere length in control hybrid oocytes. These data were used to calculate major over minor satellite length ratios shown in Fig. 5b. **b**, Data in panel A were used to calculate the depleted over control satellite length ratios for *M. musculus* (major) or *M. spretus* (minor) pericentromeres in hybrid *M. musculus* × *spretus* oocytes (means and standard deviations are indicated, means shown in boxes; two-tailed Mann–Whitney test; ****$P < 0.0001$). **c**,**d**, Control or HMGA1-depleted *M. spretus* oocytes were fixed in early meiosis I and stained for DNA (DAPI). Cells were expressing TALE-mClover (major satellite) and dCas9-mCherry/gRNA (minor satellite). Scale bars 5 µm or 2 µm (insets). Pericentromere stretching was quantified as the longest axis of minor satellite signal. Mean and standard deviation indicated, means shown in boxes, n = 17 (control) and 16 (depletion) *M. spretus* oocytes from 2 independent experiments, up to 20 pericentromeres per satellite measured in each oocyte; two-tailed Mann–Whitney test, ****$P < 0.0001$. Note that some pericentromeres were displaced from bulk chromatin without detectable minor satellite stretching (asterisk).



Extended Data Fig. 9 | Disproportionate impact of HMGA1 depletion on *M. musculus* and *M. spretus* kinetochores. a-**f**, Mad1 (**a**-**c**) or Hec1 (**d**-**f**) signal intensities from Fig. 5d,e plotted as paired measurements in hybrid *M. musculus* × *spretus* oocytes in early meiosis I ((**a**,**d**) – control; (**b**,**e**) – HMGA1 depleted) or in late meiosis I ((**c**,**f**) – HMGA1 depleted) (n = 30 (3 h, control), 31 (3 h, depletion) and 38 (7 h, depletion) hybrid oocytes from 2 independent experiments; paired two-tailed t-test, $^{****}P < 0.0001$; $^{**}P = 0.0039$). Each pair of measurements (connected by a line) represents a single cell. Control oocytes in late meiosis I were not quantified due to the lack of Mad1 signal. Note that *M. musculus* kinetochores harbor marginally more Mad1 than *M. spretus* in control early meiosis I, but *M. musculus* kinetochores are marginally smaller than *M. spretus*.

nature portfolio

Corresponding author(s): Michael A. Lampson and Ben E. Black

Last updated by author(s): Oct 16, 2024

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.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) 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	Metamorph (Molecular Devices) - version 7.5
Data analysis	Python - version 3.11 SnapGene - version 7.1.1. Image J - version 1.53t CellProfiler - version 2.1.1. LAS-AF - version 3.6.0.20104 Centrocalc - release version available here: https://github.com/DDudka9/Centrocalc Code to analyze number of AT stretches - release version available here: https://github.com/DDudka9/DNA-shape.git SeqPrep (https://github.com/jstjohn/SeqPrep) FastQC - version 0.11.8 MATLAB localalign function - version R2020a

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

The ChIP-seq data generated in this study have been deposited in the Sequence Read Archive (SRA) under BioProject accession code PRJNA1074844. The analyzed datasets are publicly available under SRA Experiment accession codes:
musculus input: SRX2939766, SRX2939768, SRX2939769
musculus CENP-A ChIP-seq: SRX2939767, SRX2939772, SRX2939773
musculus H3K9me3 ChIP-seq: SRR27928404
spretus input: SRR27922556, SRR27922557, SRR27922558
spretus CENP-A ChIP seq: SRR27922553, SRR27922554, SRR27922555
spretus H3K9me3 ChIP seq: SRR27922552
input and HMGA1 ChIP-seq: SRR5749536, SRR5749537, SRR5749546, and SRR5749545
long-read musculus genome assembly: SRR11606870.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

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 size for each experiment was determined following best practices in the field, ensuring that a minimal number of mice/cells/pericentromeres/kinetochores were analyzed per each biological replicate required to detect significance using the appropriate statistical analysis. The availability of animals was also considered.
Data exclusions	All the exclusion criteria were pre-determined and based on the abnormal behavior of the control group (e.g., cell death, lack of a clear fluorescent signal).
Replication	Experiments were done in 2-3 biological replicates. All experiments yielded similar results. DAPI/SYTOX green correlation and nocodazole washout experiments were done once with at least 10 oocytes/liver cells analyzed independently.
Randomization	Sample randomization was ensured by the fact that each experimental group (oocytes) originated from the same animals.
Blinding	Blinding during data collection or analysis was not applicable due to the visually distinct traits of the studied phenotype (DNA stretching).

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
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Primary:
1. human Anticentromere Antibody (ACA; 15-235; Antibodies Incorporated)
2. rabbit anti-mouse CENP-A antibody (custom made; Covance)
3. rabbit anti-H3K9me3 (ab8898; abcam)
4. rabbit anti-HMGA1 (clone EPR7839; ab226112 abcam)
5. rabbit anti-HMGA1 conjugated with AlexaFluor488 (clone EPR7839; ab204667; abcam)
6. mouse anti-alpha-tubulin (clone DM1A; T6199; Sigma)
7. mouse anti-Hec1 antibody (C-11; sc-515550; Santa Cruz)
8. rat anti-OLLAS antibody (clone L2; NBP1-06713; Novus Biologicals)

Secondary:
donkey anti-human IgG (H+L) conjugated with Cy5 (709-175-149; Jackson ImmunoResearch Laboratories)
donkey anti-rabbit IgG (H+L) conjugated with Cy5 (711–175-152; Jackson ImmunoResearch Laboratories)
donkey anti-rabbit IgG (H+L) conjugated with AlexaFluor488 (A21206; Invitrogen)
goat anti-mouse IgG (H+L) conjugated with AlexaFluor555 (A21422; Invitrogen)
goat anti-mouse IgG (H+L) conjugated with AlexaFluor594 (A11032; Invitrogen)
donkey anti-mouse IgG (H+L) conjugated with AlexaFluor647 (A31571; Invitrogen)
goat anti-rat IgG (H+L) conjugated with AlexaFluor647 (A21247; Invitrogen)

Validation

1. ACA antibody (Antibodies Incorporated 15-235) for immunocytochemistry in mouse oocytes was validated in: PMID: 34314679 (See Figure 6D; centromere-like dot localization pattern, each dot associated with one homologous chromosome).
2. anti-mouse CENP-A antibody was validated for ChIP in mouse cells in: PMID: 37967185
3. anti H3K9me3 antibody ab8898 for immunocytochemistry in mouse was validated in: PMID: 34433012 (See in Figure 3A the reduction of H3K9me3 signal in mouse knockout of the known heterochromatin regulator CENP-B). This antibody has been used in the field since 2003 resulting in over 1600 citations according to manufacturer. The antibody has been validated for ChIP applications according to the manufacturer.
4. anti-HMGA1 antibody for Trim-Away in mouse oocytes was validated here. Namely, Trim-Away-mediated HMGA1 depletion was confirmed using anti-HMGA1-488 antibody produced from the same clone (See Fig. 3d,e).
5. anti-HMGA1-488 antibody for immunofluorescence in mouse oocytes was validated here by depleting HMGA1 in mouse oocytes and observing no signal upon staining with anti-HMGA1-488 antibody (produced from the same clone; See Fig. 3d,e).
6. anti-alpha tubulin DM1A antibody for immunofluorescence in mouse cells was validated over the nearly 40 years it has been used in the field. This antibody was used in over 2000 publications according to manufacturer. Here, we see a clear loss of microtubule signal in mouse oocytes upon treatment with a known microtubule-depolymerizing agent nocodazole and signal recovery after nocodazole washout.
7. anti-HEC1 antibody C-11 for immnofluorescence in mouse oocytes was validated in: PMID: 31402175 (See Figure 4B; colocalization with a known centromere marker CENP-C).
8. anti-OLLAS antibody for immunofluorescence in mouse oocytes was validated here by microinjecting mRNA encoding OLLAS-tagged Borealin, which resulted in a Borealin-like centromere enrichment and some chromosome arms staining consistent with literature.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

1. Mus musculus (C57BL/6J, stock# 000664; Jackson Laboratory) - sequencing experiments
2. Mus musculus (NSA, stock# 033 corresponding to CF-1; Envigo/Inotiv) - oocyte experiments
3. Mus spretus (SPRET/EiJ, stock# 001146; Jackson Laboratory) - sequencing experiments
4. Mus spretus (SPR2, RBRC00208; RIKEN BioResource Research Center) - oocyte experiments
Mus musculus/Mus spretus hybrid animals (cross of strains 2 and 4)
All animals were 2-4 months old.
Mice were housed in controlled room temperature conditions with minimal disturbances, light/dark cycle of 12 h each, and humidity ranging between 30-70% depending on the season.

Wild animals	Study did not involve wild animals.
Reporting on sex	Only female meiosis was studied, therefore only female animals were used to collect data. Sex was assigned based on morphological reproductive traits.
Field-collected samples	No field samples were collected.
Ethics oversight	All animal experiments were approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania and were consistent with the National Institutes of Health guidelines (protocol: #804882).

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

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1074844 For paired-end sequencing analysis performed for this study, sequences were aligned to a satellite consensus sequence and not to the genome. Therefore, bed files using genomic coordinates are not available.
Files in database submission	BL6-1-H3K9me3_R1.fastq.gz BL6-1-H3K9me3_R2.fastq.gz Spretus-3-INP_R1.fastq.gz Spretus-3-INP_R2.fastq.gz Spretus-3-CENPA_R1.fastq.gz Spretus-3-CENPA_R2.fastq.gz Spretus-2-INP_R1.fastq.gz Spretus-2-INP_R2.fastq.gz Spretus-2-CENPA_R1.fastq.gz Spretus-2-CENPA_R2.fastq.gz Spretus-2-H3K9me3_R1.fastq.gz Spretus-2-H3K9me3_R2.fastq.gz Spretus-1-INP_R1.fastq.gz Spretus-1-INP_R2.fastq.gz Spretus-1-CENPA_R1.fastq.gz Spretus-1-CENPA_R2.fastq.gz Spretus-1-H3K9me3_R1.fastq.gz Spretus-1-H3K9me3_R2.fastq.gz Spretus-3-H3K9me3_R1.fastq.gz Spretus-3-H3K9me3_R2.fastq.gz
Genome browser session (e.g. UCSC)	No genome browser session generated.

Methodology

Replicates	Replicates: 1-3 biological. Replicates were reproduced.
Sequencing depth	All libraries were sequenced as 150-bp, paired-end reads. Below, we list each sample and total # of reads. Reads were not mapped to the genome. Spretus-1-INP, 11098759 Spretus-1-CENPA, 9754224 Spretus-1-H3K9me3, 12487334

	Spretus-2-INP, 9702899 Spretus-2-CENPA, 8626299 Spretus-2-H3K9me3, 8826454 Spretus-3-INP, 11174163 Spretus-3-CENPA, 12686080 Spretus-3-H3K9me3, 9747402 BL6-1-H3K9me3, 7955897
Antibodies	anti-mouse specific CENP-A antibody published in: PMID: 37967185 rabbit anti-H3K9me3 antibody (ab8898; abcam)
Peak calling parameters	No peak calling performed.
Data quality	FastQC was used to assess read quality.
Software	SeqPrep (https://github.com/jstjohn/SeqPrep) FastQC - version 0.11.8 MATLAB localalign function - version R2020a