Species-specific satellite DNA composition dictates PRC1-mediated pericentric heterochromatin

Piero Lamelza[1], Malena Parrado[1], Michael A. Lampson[1]

[1]Department of Biology, University of Pennsylvania, Philadelphia, PA, 19104

Address correspondence to: plamelza@gmail.com or lampson@sas.upenn.edu

Abstract

Pericentromeres are heterochromatic regions adjacent to centromeres that ensure accurate chromosome segregation. Despite their conserved function, they are composed of rapidly evolving A/T-rich satellite DNA. This paradoxical observation is partially resolved by epigenetic mechanisms that maintain H3K9me3-based heterochromatin, independent of specific DNA sequences. However, these mechanisms can function only after H3K9me3 has already been established, and this mark is absent from paternal chromatin in the mouse zygote. It is unknown how variation in satellite DNA sequence impacts alternative forms of heterochromatin at this earliest stage of life. Here we show functional consequences of satellite DNA variation for pericentric heterochromatin formation, recruitment of the Chromosome Passenger Complex (CPC), and interactions with the mitotic spindle. The AT-hook of Polycomb Repressive Complex 1 (PRC1) directly recognizes A/T-rich satellite DNA and packages it in H2AK119ub1 heterochromatin. By fertilizing *M. musculus* eggs with sperm from other mouse species, we show that divergent satellite sequences differ in their ability to bind PRC1, resulting in differences in H2AK119ub1 heterochromatin formation on mitotic chromosomes. Furthermore, we find that satellites that robustly form H2AK119ub1 inhibit molecular pathways that recruit the CPC to pericentromeres, increasing microtubule forces on kinetochores during mitosis. Our results provide a direct link between satellite DNA composition and pericentromere function and highlight early embryogenesis as a critical point in development that is sensitive to satellite DNA evolution.

Introduction

Satellite DNA is composed of tandemly repeated DNA sequences that make up large non-coding regions of eukaryotic genomes, e.g. 7% and 8% of human and mouse genomes, respectively [1,2]. Satellites are among the most rapidly evolving genomic sequences, with their monomer size, nucleotide sequence, genomic distribution and abundance varying widely between species[3,4]. Despite these rapid changes, they often underlie chromosomal loci with highly conserved functions[5,6]. Recent advances in long-read sequencing technology have made huge progress in precisely characterizing the genetic composition and diversity of repetitive satellite arrays[1,7–9]. However, it is still largely unknown what functional consequences may arise from the natural diversity of these sequences.

Centromeres and pericentromeres, two loci with conserved functions in orchestrating accurate chromosome segregation during cell division, are composed of rapidly evolving satellite arrays in many species. This paradoxical observation is partially resolved by evidence indicating that centromere and pericentromere function, packaging and inheritance rely on epigenetic mechanisms. The histone H3 variant CENP-A (or CenH3) defines centromeres in most eukaryotes[10], while pericentric satellites are enriched with H3K9me3-based constitutive heterochromatin[11]. Both CENP-A and H3K9me3 serve as epigenetic templates for their inheritance through mitotic cell divisions via self-propagating epigenetic loops. CENP-A recruits the machinery for assembling new CENP-A nucleosomes[12], and H3K9me3 recruits histone methyltransferases that catalyze formation of new H3K9me3[13]. Because these epigenetic mechanisms operate independently of DNA sequence, variation in satellite DNA composition was thought to have little or no functional impact[11].

However, a key developmental stage in which pericentric satellites lack H3K9me3 occurs during early embryogenesis in various metazoans[14–20]. In a subset of these species, the lack of H3K9me3 specifically arises from the removal of histones from the paternal genome and their replacement by protamines during spermiogenesis[21–23]. CENP-A is an exception to this rule, as it is retained through spermiogenesis and thus serves as the epigenetic template for centromere inheritance through the male germline[24]. Following fertilization, maternal factors remove protamines and repackage the remainder of the paternal genome with histones initially lacking H3K9me3[25,26], creating a unique developmental window in which paternal pericentric

63 satellites are packaged by alternative forms of chromatin. Indeed, Polycomb Repressive
64 Complexes 1 and 2 (PRC1 and PRC2) form what is typically considered facultative
65 heterochromatin in place of H3K9me3 on paternal satellites in the mouse zygote[14,27,28].
66 Furthermore, PRC1 and PRC2 localize to chromatin by recognition of satellite DNA[27,28],
67 suggesting that natural variation in satellite sequence may alter the binding of these complexes
68 and consequently influence heterochromatin formation and function.

69 In the widely used mammalian model species *Mus musculus* (house mouse), it takes
70 approximately four cell cycles after fertilization for paternal satellites to gradually acquire
71 H3K9me3 heterochromatin. Instead, PRC1 and PRC2 initially package paternal satellites in
72 H2AK119ub1 and H3K27me3, respectively [14,27,28]. Both modifications contribute to the
73 transcriptional silencing of the 240 bp A/T-rich "major satellite" repeats at *M. musculus*
74 pericentromeres[14,29], with PRC1 also crucial for maintaining the integrity of these arrays[30].
75 Moreover, major satellite sequences can direct the recruitment of both PRC1 and PRC2. The
76 Cbx2 subunit of PRC1 contains an AT-hook, a five amino acid motif with nanomolar binding
77 affinity for the relatively narrow minor grooves of double stranded DNA (dsDNA) generated by
78 stretches of tandem A/T nucleotides (A/T runs)[31–35] (Fig 1A). It has therefore been proposed
79 that PRC1 localization to paternal *M. musculus* pericentromeres is mediated by its AT-hook
80 directly binding A/T runs present in major satellite repeats[28]. PRC2 localization to paternal
81 pericentromeres likely requires BEND3, a methylation sensitive DNA-binding protein that can
82 bind unmethylated major satellite DNA *in vitro*[27].

83 Here we ask whether natural variation in satellite sequences among closely related
84 mouse species affects paternal pericentric heterochromatin in the zygote. By fertilizing *Mus*
85 *musculus* eggs with the sperm of mouse species with divergent pericentric satellite DNA
86 repeats, we find that *Mus musculus* pericentromeres form PRC1 heterochromatin whereas
87 *Mus pahari* pericentromeres do not. AlphaFold3 modeling and DAPI binding indicate that the
88 AT-hook of Cbx2 preferentially binds to specific A/T runs (i.e. AAAT) that form some of the
89 narrowest minor grooves[36], which are enriched in satellite arrays found at *M. musculus* but not
90 *M. pahari* pericentromeres. We also find that PRC1 heterochromatin modifies pericentromere
91 function by inhibiting the recruitment of the Chromosome Passenger Complex (CPC) to
92 paternal *M. musculus* pericentromeres. The CPC regulates chromosome attachment to the
93 mitotic spindle by destabilizing kinetochore-microtubule interactions. In zygotes, the distance

between sister kinetochores in metaphase is larger for *M. musculus* vs *M. pahari* paternal chromosomes, indicating increased force pulling sister-kinetochores apart due to reduced CPC recruitment and stabilized kinetochore-microtubule interactions. Our results show that diverse pericentric DNA sequences lead to functionally diverse pericentromeres through differential binding of heterochromatin forming complexes.

Results

Variation in PRC1 localization and heterochromatin formation on divergent satellites

Given that major satellite DNA directs PRC1 and PRC2 heterochromatin formation, we hypothesized that divergent pericentric satellite sequences might alter Cbx2 or BEND3 binding and by extension H2AK119ub1 or H3K27me3 formation, respectively. To compare paternal chromosomes with different pericentric sequences, we developed an experimental system based on fertilizing *M. musculus* eggs with sperm from mouse species harboring genetically divergent pericentric satellites (Fig 1B). We selected three species (*M. musculus, M. caroli* and *M. pahari*) that differ in their satellite monomer sequence and not only in satellite copy number as seen in *M. musculus* subspecies and its sister-species *M. spretus*[2,37]. We injected sperm heads into MII eggs (i.e. intracytoplasmic sperm injection, ICSI[38]) to bypass reproductive barriers between species, arrested the resulting zygotes at the first mitosis by kinesin-5 inhibition, and measured H2AK119ub1 and H3K27me3 by immunofluorescence. Because only maternal chromosomes have H3K9me3 at this stage, we used this mark to distinguish maternal and paternal chromosomes[14]. Major transcriptional activation of the mouse embryo (zygotic genome activation) does not begin until the late two-cell stage[39], so our analyses reflect interactions between satellite sequences and maternally inherited *M. musculus* PRC1/2 complexes rather than a mixture of subunits from both species. Analyzing heterochromatin in the zygote also allows us to test satellite DNA variation from evolutionarily distant species, which would otherwise be inaccessible due to hybrid incompatibilities that typically arise after zygotic genome activation[40–42].

We found that all three species' pericentromeres robustly acquire H3K27me3 (Fig S1). In contrast, H2AK119ub1 is enriched on paternal *M. musculus* and *M. caroli* pericentromeres relative to chromosome arms, but not on *M. pahari* pericentromeres (Fig 1C and 1E). To test whether these differences result from differences in PRC1 localization, we measured PRC1's

125 core-catalytic subunit, RNF2, by immunofluorescence. Consistent with the H2AK119ub1
126 enrichment, paternal *M. musculus* pericentromeres are strongly enriched for RNF2 relative to
127 chromosome arms whereas *M. pahari* chromosomes only exhibit a slight enrichment near the
128 very ends of their chromosomes, presumably near centromeres (Fig 1D and 1E). In certain
129 contexts, PRC1 is recruited to chromatin by the chromodomain of its Cbx subunit, which
130 directly binds H3K27me3[43–45], and Cbx2 also contains a chromodomain (Fig 1A). However, the
131 presence of H3K27me3 at *M. pahari* paternal pericentromeres (Fig S1), which lack
132 H2AK119ub1 and RNF2 (Fig 1C-E), indicates that H3K27me3 is not sufficient for PRC1
133 recruitment during the first zygotic mitosis, consistent with previous findings that PRC1 can
134 localize to paternal pericentromeres independent of PRC2 and H3K27me3[14]. Overall, these
135 results demonstrate striking differences in PRC1 recruitment and heterochromatin formation at
136 divergent pericentric satellites, independent of PRC2.
137
138 ***Satellite sequence determinants of PRC1 binding***
139 To explain the differences in pericentric RNF2 and H2AK119ub1, we hypothesized that
140 *M. musculus* and *M. caroli* pericentromeric satellites, but not *M. pahari* satellites, are enriched
141 for narrow DNA minor grooves that bind the AT-hook of Cbx2 to recruit the PRC1 complex.
142 Following a previously described approach[2], we used DAPI staining to compare the amount of
143 narrow minor grooves generated by A/T-runs at each species' pericentromere, as DAPI
144 specifically binds to these regions on the dsDNA molecule[46–48]. Consistent with previous
145 observations[2,9,49], we find increased DAPI staining at *M. musculus* and *M. caroli*
146 pericentromeres relative to chromosome arms, but not at *M. pahari* pericentromeres (Figure
147 2A, 2B and Fig S2A). To exclude the possibility that DAPI enrichment at pericentromeres
148 represents a relatively more compact chromatin state compared to the rest of chromosome, we
149 stained with Sytox Green, an intercalating DNA-dye with no considerable DNA sequence
150 preference[50]. None of the three species show enrichment of Sytox Green at their
151 pericentromeres (Fig 2C, 2D and Fig S2B). Together, these results demonstrate that *M.*
152 *musculus* and *M. caroli* pericentromeres are enriched for narrow A/T-rich minor grooves that
153 can robustly bind the AT-hook of Cbx2, whereas *M. pahari* pericentromeres are not.
154 We considered two possibilities for the underlying differences in pericentric DNA
155 sequence between species, leading to differences in DAPI enrichment and PRC1 localization.

156 We focused on comparisons between *M. musculus* and *M. pahari* because of the long-read
157 sequencing data available for these species. First, *M. musculus* major satellites may have a
158 higher frequency of A/T runs compared to *M. pahari* pericentric satellites (pi satellites)[9].
159 However, major and pi satellite consensus sequences have similar frequencies of A/T runs of
160 lengths ranging from four (4W, the minimal number of nucleotides the AT-hook can span) to
161 ten or more (≥10W) (Fig S2C). To capture nucleotide diversity that may be lost in consensus
162 sequences, we also calculated the frequencies of A/T runs across genomic contigs containing
163 either major or pi satellite arrays. Consistent with consensus sequences, we find that both
164 species' satellite arrays have similar frequencies of A/T runs (~825 vs ~750 per 10 kb window
165 in major satellite or pi satellite, respectively) (Fig 2E). The frequencies are also similar for each
166 length of A/T runs (Fig 2E and Fig S2D), with the exceptions of 6W and ≥10W, which are
167 relatively more abundant in major and pi satellite arrays, respectively. However, it is unclear
168 how differences in these lengths specifically would lead to higher PRC1 and DAPI at *M.*
169 *musculus* major satellite arrays. Therefore, it is unlikely that differences in A/T-run frequency
170 explain the difference in PRC1 heterochromatin formation.
171 The second possibility is that *M. musculus* major satellite arrays might be enriched for
172 specific A/T sequences optimal for Cbx2 AT-hook binding, compared to *M. pahari* satellite
173 arrays. We analyzed the sequence composition of 4W, 5W and 6W A/T runs, which make up
174 ~75% of the A/T stretches in both species' satellite arrays and therefore most potential AT-
175 hook binding sites. The vast majority of 4W sequences in *M. musculus* major satellite arrays
176 are AAAT (114 per 10 kb), AAAA (113 per 10 kb) or TAAA (39 per 10 kb) (Fig 2F and Fig
177 S2E). *M. pahari* pi satellite arrays have a similar frequency of TAAA sites (37 per 10 kb) but
178 much lower frequencies of AAAT and AAAA sites (7.2 and 46 per 10 kb, respectively). Based
179 on these differences, we hypothesized that the Cbx2 AT-hook preferentially binds AAAT and/or
180 AAAA.
181 To test the binding preferences of the Cbx2 AT-hook, we used AlphaFold3[51] to perform
182 *in silico* pairwise competitive binding assays for AAAT and AAAA vs all other possible A/T
183 tetranucleotides (Fig 2G). We first confirmed that AlphaFold3 accurately models AT-hook
184 binding to dsDNA using the well characterized HMGA1 AT-hook bound to a fragment of the
185 IFN-ß promoter[34] (PDB ID 2EZD) as a reference. The predicted structure closely matched the
186 known structure (RMSD=0.48 Å), and in both structures, the AT-hook occupied the minor

187 groove of the A/T run (AAATT) (Fig S2F). Furthermore, the confidence scores of the predicted
188 structure (iPTM=0.48 and PTM=0.47) are similar to the models generated in our *in silico*
189 competitive binding assay (Table S1). The AT-hooks of Cbx2 and HMGA1 share the same
190 core sequence (PRGRP) which binds the minor groove, supporting our use of AlphaFold3 to
191 model the AT-hook binding preference of Cbx2.

192 For each competition assay, we model a single dsDNA molecule encoding two distinct
193 A/T runs flanked by four G:C base pairs (5' GGGGWWWWGGGGWWWWGGGG 3'). These
194 two A/T runs "compete" to bind a single peptide consisting of the core Cbx2 AT-hook along
195 with flanking amino acids (RKRGKR**PRGRP**RKHTVTSS). All AlphaFold models predict that
196 the Cbx2 AT-hook prefers to bind AAAT and AAAA over all other A/T tetranucleotides (Fig 2G).
197 When AAAT and AAAA compete against each other, their relative order on the dsDNA dictates
198 which sequence binds the AT-hook (Fig S2G, see Materials and Methods), indicating that
199 these two sequences have similar affinity for the AT-hook. We obtain similar results when
200 analyzing 5W and 6W A/T runs. AAAAT and AAAAAT are the most abundant 5W and 6W
201 sequences in the *M. musculus* major satellite array (147 per 10 kb and 70 per 10 kb,
202 respectively) (Fig 2F and Fig S2E). These sequences outcompete all other prevalent 5W and
203 6W sequences in both major and pi satellite arrays for Cbx2 AT-hook binding (Fig 2G and Fig
204 S2G). Furthermore, AAAAT and AAAAAT are 8 and 2.5-fold less frequent, respectively, in *M.*
205 *pahari* pi satellite compared to *M. musculus* major satellite arrays (Fig 2F and Fig S2E).

206 These results indicate that the Cbx2 AT-hook prefers to bind stretches of adenine
207 nucleotides that end with a single thymine (A_nT), consistent with previous reports showing that
208 the core motif of the AT-hook (PRGRP) preferentially binds the minor groove of the sequence
209 5' AA(A/T)T 3'. This preference reflects optimal Van der Waals packing when adenine bases
210 are on opposite sides of the A/T run[35]. Our AlphaFold3 models also show that the AT-hook,
211 which can span 4-5 bases, spans AAAT or AAAAT bases within all but one modeled AAAAAT
212 sequence, rather than the overlapping AAAA or AAAAA bases (Fig S2H and Table S1).
213 Additionally, a meta-analysis of PDB structures[36] indicates that AAAT generates one of the
214 narrowest minor grooves compared to other tetranucleotide sequences, which may contribute
215 to preferential binding of both the AT-hook and DAPI. To generalize this analysis to include
216 A_nT sequences (like AAAT) present in longer A/T runs, we quantified the frequency of A_nT
217 sequences of various lengths (i.e. AAAT, AAAAT, AAAAAT, AAAAAAT and AAAAAAAT) in

218 major and pi satellite arrays independent of total A/T run length. We find that major satellite
219 arrays have approximately three times more of these A_nT sequences than *M. pahari* pi satellite
220 arrays (Fig 2H and Fig S2I). Furthermore, pi satellite arrays exhibit A_nT frequencies similar to
221 euchromatin. Overall, these analyses indicate that PRC1 and its H2AK119ub1 modification are
222 enriched at *M. musculus* major satellite arrays because they are enriched for A_nT sequences
223 that preferentially bind the Cbx2 AT-hook. In contrast, *M. pahari* pi satellite arrays exhibit a
224 lower frequency of A_nT sequences like euchromatin and fail to enrich both PRC1 and
225 H2AK119ub1.
226
227 ***Consequences of PRC1 heterochromatin for pericentromere function***
228 Next, we asked if the difference in PRC1-heterochromatin between paternal *M.*
229 *musculus* and *M. pahari* chromosomes has consequences for pericentromere function.
230 Notably, H2AK119ub1 is positioned close to the site of another post-translational modification:
231 phosphorylation of H2AT121 by Bub1 kinase. H2AT121phos recruits the Chromosome
232 Passenger Complex (CPC) to pericentromeres via Sgo1 (Fig 3A) to ensure accurate
233 chromosome segregation by regulating kinetochore-microtubule interactions during mitosis[52–
234 54]. ChIP-seq and *in vitro* assays on reconstituted nucleosomes have shown that H2AK119ub1
235 and H2AT121phos are mutually exclusive[55], suggesting that one modification physically blocks
236 the deposition of the other on the same H2A C-terminal tail. We hypothesized that
237 H2AK119ub1 occludes Bub1 kinase from phosphorylating H2AT121 and thereby reduces CPC
238 binding. This hypothesis predicts low H2AT121phos and low CPC on paternal *M. musculus*
239 pericentromeres (Fig 3A), which have high H2AK119ub1 compared to maternal *M. musculus*
240 and paternal *M. pahari* pericentromeres (Fig 1C and 1E).
241 To test this prediction, we first visualized H2AT121phos, the CPC subunit Survivin, and
242 the PRC1 subunit RNF2. In *M. musculus* ♀ x *M. musculus* ♂ zygotes arrested at metaphase
243 by chemical inhibition of the anaphase promoting complex/cyclosome (APC/C), we found that
244 H2AT121phos is present along the entire chromosome (Fig 3B) as reported in the previous cell
245 cycle in mouse oocytes[56,57]. Moreover, we found an inverse relationship between RNF2 and
246 H2AT121phos, with the presence of RNF2 at paternal pericentromeres coinciding with a
247 reduction in H2AT121phos (Fig 3B). In contrast, maternal pericentromeres with undetectable
248 RNF2 show no reduction in H2AT121phos. Consistently, when we compared regions of

maternal and paternal chromosomes that both lack RNF2, H2AT121phos intensities were similar, indicating that reduced H2AT121phos depends on RNF2. On average, H2AT121phos is reduced to ~60% at paternal *M. musculus* pericentromeres compared to maternal pericentromeres (Fig 3B). Consistent with paternal *M. musculus* H2AK119ub1 heterochromatin inhibiting H2AT121phos, Survivin intensity on paternal *M. musculus* pericentromeres is reduced to ~65% compared to paternal *M. pahari* and maternal *M. musculus* pericentromeres, both of which form relatively less H2AK119ub1 (Fig 3C). Together, these results show that PRC1 heterochromatin inhibits the acquisition of H2AT121phos and thereby reduces CPC recruitment to pericentromeres of paternal *M. musculus* chromosomes.

The Aurora B kinase subunit of the CPC phosphorylates kinetochore proteins to destabilize kinetochore-microtubule interactions[58]. In tissue culture cells, expression of a non-phosphorylatable mutant of a key Aurora B substrate, Hec1, results in increased stabilization of kinetochore-microtubule interactions and increased forces pulling the sister kinetochores apart at metaphase[59]. Increased inter-kinetochore distance therefore provides a functional readout for reduced Aurora B activity (Fig 3A). To test for functional consequences of CPC differences between pericentromeres in the zygote, we measured the distance between sister kinetochores at metaphase. We found increased inter-kinetochore distances for paternal *M. musculus* chromosomes compared to either maternal *M. musculus* or paternal *M. pahari* chromosomes (Fig 3D). This result is consistent with our hypothesis that PRC1 heterochromatin inhibits H2AT121 phosphorylation, downregulates the CPC, and stabilizes kinetochore-microtubule interactions on paternal *M. musculus* chromosomes. Increased inter-kinetochore distances might also reflect reduced pericentric cohesion because in addition to recruiting the CPC, Sgo1 also recruits phosphatases that counteract cohesin removal that occurs along chromosome arms during mitotic prophase[60,61] (Fig 3A). However, sister chromatids appear in close proximity along most of their lengths during zygotic mitosis[62] (Fig 3D), suggesting that cohesins are not removed during prophase at this stage of development. Therefore, we interpret the increase in inter-kinetochore distance as a consequence of reduced CPC activity.

Discussion

279 Long-read sequencing technology continues to reveal the extensive diversity of satellite
280 DNA sequence composition[8,9,63], but experimental systems that test the functional impact of
281 this diversity remain limited. In the mouse model species *M. musculus,* centromeric and
282 pericentromeric satellite variation is limited to copy-number differences of the same satellite
283 sequences[2,37,56]. Our hybrid embryo system overcomes this limitation by incorporating
284 sequence variation between satellites that is only found in more divergent non-model mouse
285 species. Using this system, we demonstrated that *M. musculus* major satellite sequence is
286 more effective than *M. pahari* pi satellite sequence at recruiting PRC1. AlphaFold3 modeling
287 indicates that this difference results from the AT-hook of Cbx2$^{M.\ musculus}$ preferentially binding
288 the minor groove of A_nT sequences, which are enriched in *M. musculus* major satellite arrays
289 when compared to *M. pahari* pi satellite arrays and the rest of the genome. The sequence
290 preference for the Cbx2 AT-hook is consistent with the known optimal binding sequence of
291 other AT-hook proteins[35]. We propose that the relatively high density of A_nT sequences at *M.*
292 *musculus* major satellite arrays is above a threshold required for efficient PRC1 recruitment
293 and robust H2AK119ub1 formation.
294 Previous findings have shown that the Cbx2 chromodomain, in addition to the AT-hook,
295 contributes to PRC1 localization to chromatin during zygotic interphase, likely by binding
296 H3K27me3 that first forms during S-phase[28]. However, our data demonstrate that during
297 mitosis, paternal *M. pahari* pericentromeres exhibit similar levels of H3K27me3 as paternal *M.*
298 *musculus* pericentromeres (Fig S1A) yet significantly reduced PRC1 and H2AK119ub1 (Figure
299 1C-E). These findings indicate that the Cbx2 chromodomain does not effectively localize PRC1
300 during mitosis. We speculate that Aurora B-mediated phosphorylation of H3S28 during
301 mitosis[64] prevents the Cbx2 chromodomain from binding H3K27me3, similar to how Aurora B-
302 mediated phosphorylation of H3S10 inhibits HP1 chromodomain binding to H3K9me3[65]. As a
303 result, satellite DNA binding via the Cbx2 AT-hook predominately dictates pericentric PRC1
304 localization and H2AK119ub1 heterochromatin during mitosis. Nonetheless, during interphase,
305 the Cbx2 chromodomain likely binds H3K27me3 at *M. pahari* satellites, recruiting PRC1 and
306 establishing H2AK119ub1. Upon mitotic entry, however, PRC1 is displaced from *M. pahari*
307 chromatin and cannot counteract the removal of pericentric H2AK119ub1 by cytosolic
308 deubiquitinating enzymes once the nuclear envelope breaks down[66]. Thus, our analysis of

309 mitotic chromosomes indicates that Cbx2's chromodomain is not required at this phase of the
310 cell cycle, highlighting cell-cycle-specific differences in how PRC1 is recruited to chromatin.
311 Despite the importance of PRC1 pericentric heterochromatin in *M. musculus* zygotes[30],
312 our findings suggest that it is not a conserved feature of mitotic paternal pericentromeres, as it
313 is lacking from *M. pahari* pericentromeres. *M. musculus* major satellites and *M. caroli* satellites
314 may have specific properties that require their packaging in PRC1 heterochromatin, whereas
315 *M. pahari* pi satellites can rely on PRC2-mediated H3K27me3. Alternatively, the *M. pahari* egg
316 cytoplasm may have specific adaptations that allow PRC1 to be recruited to mitotic
317 pericentromeres by a different mechanism. Because the core Cbx2 AT-hook sequence is
318 conserved across Muridae (data not shown), it is unlikely that the AT-hook of Cbx2$^{M.\ pahari}$ has
319 adapted to bind A/T sequences enriched at *M. pahari* pi satellites. Similarly, the Cbx2
320 chromodomain is identical across Muridae (data not shown), indicating that the chromodomain
321 of Cbx2$^{M.\ pahari}$ has not adapted to have a higher binding affinity towards H3K27me3. However,
322 vertebrates encode multiple PRC1 complexes that share the core ubiquitin ligase subunits
323 (RING1A/1B) but differ in associated proteins that diversify PRC1 function, including how the
324 complex is recruited to chromatin[67]. Thus, another chromatin targeting subunit expressed in *M.*
325 *pahari* zygotes might recognize paternal *M. pahari* pericentromeres during mitosis.
326 Although maternal and paternal major satellites are genetically identical, only paternal
327 satellites robustly form PRC1 heterochromatin because H3K9me3-based heterochromatin on
328 maternal satellites inhibits PRC1 recruitment[14,28]. Given this, we propose that in most cell
329 cycles H3K9me3 can serve as a buffer against DNA sequence-based pathways of
330 heterochromatin formation, suppressing functional diversity that might otherwise arise from
331 differences in satellite sequences. In contrast, our findings highlight the first zygotic cell cycle
332 as a unique environment where PRC1 heterochromatin formation and function are sensitive to
333 variation in satellite DNA sequence due to the loss of H3K9me3 from the paternal genome
334 during spermiogenesis. This functional variation occurs at a critical stage of development
335 where chromosome segregation errors would lead to aneuploidy of all subsequent cells.
336 Despite the importance of this first cell division, PRC1 heterochromatin inhibits molecular
337 pathways that ensure accurate chromosome segregation, particularly regulation of
338 kinetochore-microtubule attachments by the CPC. Our findings suggest that expansion of
339 pericentric satellite arrays enriched for A_nT sequences would lead to increased PRC1

340 heterochromatin and a corresponding decline in CPC recruitment, making paternal
341 chromosomes more prone to segregation errors. Although it is difficult to test for potentially
342 subtle differences in error rates between paternal and maternal chromosomes in mouse
343 zygotes, our hypothesis predicts selection against large major satellite arrays in natural
344 populations. Consistent with this idea, wild-caught *M. musculus* have roughly ten times fewer
345 major satellite repeats than inbred laboratory strains, where natural selection is relatively
346 weak[37]. Despite potentially deleterious effects, however, pericentric PRC1 heterochromatin is
347 required for the transcriptional silencing and stability of paternal major satellite arrays[28,30],
348 suggesting an evolutionary trade-off between proper packaging of major satellites and
349 chromosome segregation fidelity.
350
351 **Materials and Methods**
352
353 ***Mouse strains***
354 The mouse strain representing *Mus musculus* in our experiments is FVB/NJ, purchased
355 from Jackson Laboratory (strain#001800). FVB/NJ was chosen because its MII eggs can
356 survive intracytoplasmic sperm injection and because it is an inbred strain. *M. caroli*
357 (CAROLI/EiJ, strain#000926) and *M. pahari* (PAHARI/EiJ, strain #002655) were also
358 purchased from Jackson Laboratory. We maintain our own colony of *M. pahari* because
359 Jackson Laboratory no longer carries this species. All animal experiments and protocols were
360 approved by the Institutional Animal Use and Care Committee of the University of
361 Pennsylvania and are consistent with National Institutes of Health guidelines (protocol
362 #804882).
363
364 ***Intracytoplasmic sperm injection (ICSI)***
365 To prepare sperm for ICSI, the epididymis and vas deferens were removed from males
366 and dissected in PBS to release mature sperm. Sperm were allowed to swim-out for 10 mins
367 on a 37°C slide warmer. After swim out, sperm were pelleted by centrifugation at 700xg at 4°C
368 for 5 min. The PBS supernatant was removed, and sperm were then washed twice with ice-
369 cold Nuclear Isolation Media (NIM, 123 mM KCl, 2.6mM NaCl, 7.8 mM Na_2PO_4, 1.4 mM
370 KH_2PO_4, 3 mM EDTA, pH adjusted to 7.2 using 1M KOH) with 1% poly vinyl alcohol (PVA).

371 After the second wash, sperm were resuspended in 100 μL of NIM 1% PVA and then
372 sonicated using a Branson Sonic Bath Model 1210 for 15-20 second intervals until at most
373 30% of sperm had their heads detached from their tails. Sonicated sperm were washed twice
374 with NIM 1% PVA. After the second wash, the sperm pellet was resuspended in a 1:1 solution
375 of glycerol and NIM 1% PVA and placed at -20°C until the day of injection (the following day or
376 1 week later).
377 Females were super-ovulated by injection with 5 U PMSG (Peptides International)
378 followed by injection with 5 U hCG (Sigma) 48 hours later. MII eggs were collected from
379 females 14-15 hours post-hCG injection and placed in M2 media (Sigma) with hyaluronidase
380 (0.15 mg/mL) to remove cumulus cells. MII eggs were then washed through four drops of M2
381 media supplemented with 4 mg/mL BSA (M2+BSA) and left on a 37°C plate warmer until
382 injection. 10-20 μLs of sonicated sperm was diluted in 200 μL of NIM 1% PVA, gently vortexed,
383 and then washed twice with NIM 1% PVA. After the last wash, the sperm pellet was
384 resuspended in 100 μL of NIM 1% PVA and left on ice until injection.
385 Batches of 10 MII eggs were injected with sperm heads at a time in room temperature
386 drops of M2+BSA. After injection, each batch was moved to a drop of M2+BSA on a 37°C
387 plate warmer and allowed to recover for 1 hour. After recovery, each batch was washed
388 through four drops of pre-equilibrated AKSOM (Millipore Sigma) and placed in at 37°C
389 humidified incubator with 5% CO_2 until injections were complete. Depending on the
390 experiment, at least three hours after injection, eggs were then moved into a drop of AKSOM
391 with either 10 μM STLC (kinesin-5 inhibitor, Sigma) (to arrest cells in mitosis with monopolar
392 spindles) or 5 μM proTAME (APC/C inhibitor, R&D Systems) (to arrest zygotes in metaphase),
393 then allowed to develop overnight in a 37°C incubator with 5% CO_2.
394
395 ***Fixing and staining***
396 The morning after ICSI, arrested zygotes were fixed in 37°C 2% paraformaldehyde in
397 PBS for 20 minutes, washed through three pools of blocking solution (PBS containing 0.5%
398 BSA, 0.01% Tween-20) and then left at 4°C overnight. The next morning cells were
399 permeabilized in PBS containing 0.5% Triton X-100 (Sigma) for 15 minutes at room
400 temperature, quickly washed through two pools of blocking solution and then allowed to block
401 for 20 minutes at room temperature. For H3K9me3 and H3K27me3 staining, the cells were

402 treated with lambda-phosphatase (1600 U, NEB) for 1 hour at 37°C in a humidified chamber.
403 Embryos were then quickly washed through two pools of blocking solution and then incubated
404 with primary antibodies for 1 hour in a dark humidified chamber at room temperature (or
405 overnight at 4°C for H2AT121phos staining). Afterwards, embryos underwent three 15-minute
406 washes in blocking solution and then were incubated with secondary antibodies for 1 hour in a
407 dark humidified chamber at room temperature. Cells then underwent three 15-minute washes
408 in blocking solution and were mounted in Vectashield with DAPI (Vector) to stain
409 chromosomes. For cells that were stained with Sytox Green (Invitrogen), the first wash after
410 incubation with secondary antibodies included 1 uM Sytox Green and after the next two
411 washes were mounted in Vectashield without DAPI. Primary antibodies used are rabbit anti-
412 H2AK119ub1 (1:800, Cell Signaling, D27C4), rabbit anti-H3K27me3 (1:700, Cell Signaling,
413 C36B11), mouse anti-H3K9me3 (1:200, Active Motif, 39285), rabbit anti-H3K9me3 (used when
414 co-staining against RNF2, 1:500, Active Motif, 39162), rabbit anti-Survivin (1:500, Cell
415 Signaling, 71G4B7), rabbit anti-H2AT120phos (1:2500, Active Motif, 39392), mouse anti-
416 RNF2/Ring1B (1:500, Active Motif, 39664) and rabbit anti-CENP-C[68] (1:500). Secondary
417 antibodies used are donkey anti-rabbit Alexa Fluor 488, donkey anti-mouse Alexa Fluor 594,
418 donkey anti-rabbit Alexa Fluor 594, and donkey anti-mouse Alex Fluor 647. All secondary
419 antibodies are purchased from Invitrogen and used at a dilution of 1:500.
420
421 ***Microscopy***
422 Some confocal images were collected as z-stacks with 0.5 μm intervals, using a
423 microscope (DMI4000 B; Leica) equipped with a 63× 1.3 NA glycerol-immersion objective lens,
424 an xy piezo Z stage (Applied Scientific Instrumentation), a spinning disk confocal scanner
425 (Yokogawa Corporation of America), an electron multiplier charge-coupled device camera
426 (ImageEM C9100–13; Hamamatsu Photonics), and either an LMM5 (Spectral Applied
427 Research) or Versalase (Vortran Laser Technology) laser merge module, controlled by
428 MetaMorph software (Molecular Devices, v7.10.3.294). Confocal images were also collected
429 using a Leica TCS SP8 Four Channel Spectral Confocal System with a 63x objective lens. A z-
430 stack of 0.3 μm intervals was used to collect images of embryos used to calculate sister-
431 kinetochore distances. All samples in an experiment were imaged using the same laser
432 settings.

433

434 ***Image quantification and statistical analysis***

435 All image analysis was carried out using ImageJ/Fiji[69]. To measure signal intensities
436 along single chromosomes, we selected chromosomes that were mostly in the XY plane and
437 generated max intensities projections of them. Beginning from the pericentric end, we then
438 used the segmented line tool to manually draw a linear ROI (roughly the width of the
439 chromosome) along the chromosome's length and measured the average signal intensity per
440 unit distance along the ROI. Average background was then subtracted from these values. To
441 more easily compare the distributions of DAPI and Sytox Green intensity, their values were
442 normalized to the average intensity within each cell/embryo using custom R scripts.

443 To measure Survivin intensity at pericentromere ends, we also selected chromosomes
444 that were mostly in the XY plane and generated maximum intensity projections of them. We
445 then use the rectangular selection tool to generate square ROIs that encompassed pericentric
446 ends of chromosomes and measured average intensity within the ROI. Average background
447 was then subtracted from these values. The resulting values were then normalized to maternal
448 chromosomes in each cell/embryo using custom R scripts.

449 To measure the distance between sister-centromeres, we uniquely labelled each
450 centromere and extracted its 3D coordinates using 3D maxima finder and 3D Roi Manager
451 (based on CENP-C signal). We then manually assigned sister-centromeres with the help of
452 BigDataViewer to "rotate" the image stack in cases where the spindle pole axis was not in the
453 XY plane and then used custom R scripts to calculate the distance between them.

454 *P*-values were determined using R. We performed non-parametric Kruskal-Wallis tests
455 followed by Dunn's test and used the Bonferroni method to correct *P*-values for multiple
456 testing. Plots were made in the ggplot2 package in R.

457

458 ***Satellite sequence analysis***

459 *M. musculus* assemblies containing major satellite arrays were found by BLASTing the
460 major satellite consensus sequence[37] on NCBI BLAST server. These assemblies are from a
461 long-read *M. musculus* genomes generated by the Mouse Genome Project and Darwin Tree of
462 Life Project (https://www.nhm.ac.uk/our-science/research/projects/darwin-tree-of-
463 life.html#:~:text=The%20Darwin%20Tree%20of%20Life%20project%20aims%20to%20genera

464 te%20DNA,fungi%20within%20the%20British%20Isles.). The pi satellite array containing
465 contigs are from a published *M. pahari* long-read assembly[9]. These contigs were chosen
466 because they represent relatively complete assemblies of functional *M. pahari* centromeres
467 and pericentromeres based on CENP-A and H3K9me3 ChIP-seq. We chose to study major
468 and pi satellite arrays because they are the most abundant pericentric satellites and because
469 PRC1 is known to associate with major satellite[14].
470 To quantify the frequency of A/T runs of various lengths along a contig, we used the
471 matchPattern() function in the Biostrings package in R
472 (https://bioconductor.org/packages/Biostrings). Specifically, we searched for A/T runs that
473 were at least four nucleotides long (i.e. matched "WWWW"), with overlapping matches
474 eventually being merged into one longer A/T run by using the reduce() function. To analyze the
475 composition of 4W and 5W A/T runs, we used the matchPattern() function to find all possible
476 4W and 5W sequences and then only kept matches that overlapped with 4W or 5W A/T run
477 lengths, respectively. To analyze the composition of 6W A/T runs, we used the 6W A/T run
478 coordinates to extract sequences within the satellite array using the Genomic Ranges
479 package[70] in R. Sequences that are reverse complements of each other are redundant and
480 therefore were merged into one. The average number of times a particular sequence occurred
481 per 10 kb bin within an array was calculated by dividing the total number of occurrences of a
482 particular sequence within the array by the length of the array and then multiplying by 10,000.
483 We defined the range of the satellite arrays for each contig based on coordinates from satellite
484 consensus sequence BLAST results. Finally A_nT sequences were also found along contigs
485 using the matchPattern() function. To prevent over-counting of smaller AnT sequences (e.g.
486 overlap of AAAT and AAAAT), only the longest of overlapping A_nT sequences were counted.
487

488 ***AlphaFold3 modeling***

489 All sequences were modeled using the AlphaFold3 webserver (https://alphafoldserver.com/).
490 To model a single dsDNA molecule encoding two distinct A/T runs flanked by four G:C base
491 pairs, we entered a single forward and a single reverse strand (e.g. forward:
492 5'GGGGAAAAGGGGAAATGGGG 3', reverse: 5'CCCCATTTCCCCTTTTCCCC 3'). We also
493 modeled a single peptide encoding the AT-hook of Cbx2 with flanking amino acids
494 (RKRGKR**PRGRP**RKHTVTSS). Alphafold3 outputs five model structures (0-4), with 0 scored

495 as the best and the fourth scored as the worst. Each model was visually inspected using
496 Pymol (The Pymol Molecular Graphics System, Version 3.0 Schrodinger, LLC). All five models
497 for each competition assay were consistent with one another (Table S1). To rule out the
498 possibility that the relative position along the dsDNA molecule dictates where the AT-hook
499 binds, we flipped the relative positions of the two A/T sequences along the dsDNA. For
500 example, we modeled both 5'GGGGAAAAGGGGAAATGGGG 3' and
501 5'GGGGAAATGGGGAAAAGGGG3' sequences (Figure S2G).
502

503 **Acknowledgements**

504 We thank Dr. Colin Conine for training in intracytoplasmic sperm injection and Dr. Ben Black
505 and Dr. Damian Dudka for thoughtful conversations on AT-hook proteins. We also thank Dr.
506 Black for initially suggesting that pericentric H2AK119ub1 may be incompatible with
507 H2AT121phos. We thank Dr. Beth L. Dumont for sending us PAHARI/EiJ mice from Jackson
508 Laboratory. We thank Emma Hamlin for help with data analysis. We thank Drs. Mia Levine,
509 Ben Black and Richard Schultz for helpful discussions and comments on the manuscript.
510

511 **Author Contributions**

512 Conceptualization: P. Lamelza and M. A. Lampson, Methodology: P. Lamelza and M. Parrado,
513 Investigation: P. Lamelza, and M. Parrado, Writing: P. Lamelza and M. A. Lampson, Funding
514 Acquisition: P. Lamelza and M.A. Lampson, Resources: M.A. Lampson
515

516 **Conflict of Interest**

517 The authors declare no competing financial interests.
518

519 **Funding**

520 This work was supported by: The Helen Hay Whitney Foundation and National Institutes of
521 Health grant 1K99GM152835 awarded to P. Lamelza and National Institutes of Health grant
522 R35GM122475 to M. A. Lampson.
523

524 **Figure Legends**

525

526 **Figure 1. Natural variation in pericentric H2AK119ub1 heterochromatin and PRC1**
527 **recruitment. A)** Schematic showing PRC1 recruitment to DNA via the AT-hook of its Cbx2
528 subunit, which binds into the minor groove of A/T runs present in satellite DNA. Chromodomain
529 labelled as "CD". **B)** Mouse zygotes are generated by injecting a *M. musculus* egg arrested in
530 meiosis II with the sperm of mouse species with divergent pericentric satellite DNAs. **C-E)**
531 Zygotes generated with sperm from the indicated species were arrested in mitosis with a
532 kinesin-5 inhibitor (STLC, C) or by APC/C inhibition (proTAME, D), then fixed and stained for
533 H2AK119ub1 or RNF2 (green), H3K9me3 (magenta) to mark maternal *M. musculus*
534 chromosomes, and DAPI (gray). Arrowheads (D) highlight the slight enrichment of RNF2 near
535 *M. pahari* centromeres. Graphs (E) show average H2AK119ub1 (n=103-164) and RNF2
536 (n=152-160) intensities along *M. musculus* (red), *M. caroli* (yellow) and *M. pahari* (blue)
537 paternal chromosomes, starting from pericentric ends. S.E.M. is indicated by light band
538 surrounding the mean line. Both *M. musculus* and *M. caroli* chromosomes have significantly
539 higher H2AK119ub1 intensity than *M. pahari* chromosomes within the first two microns (gray
540 box), (*** P <0.001). Statistical significance was calculated by a Kruskal-Wallis test, followed
541 by Dunn's post hoc test with Bonferroni correction. Images are max intensity z-projections,
542 scale bars 10 μm or 1 μm (insets).
543
544 **Figure 2. *M. musculus* major satellite arrays are enriched for A_nT sequences that**
545 **preferentially bind the Cbx2 AT-hook**. **A-D)** Zygotes generated with *M. musculus* or *M.*
546 *pahari* sperm were arrested in mitosis with a kinesin-5 inhibitor (STLC), then fixed and stained
547 as indicated. Survivin (CPC subunit) marks pericentromeres, H3K9me3 marks maternal
548 chromosomes, H3K27me3 mark paternal pericentromeres, and DAPI or Sytox Green label
549 DNA. Arrowheads in insets point to paternal pericentromeres as indicated by presence of
550 Survivin and absence of H3K9me3 (A) or enrichment of H3K27me3 (C). Images are max
551 intensity z-projections, scale bars 10 μm or 1 μm (insets). Graphs show average DAPI (B,
552 n=72-165) or Sytox Green (D, n=44-66) intensity along *M. musculus* (red), *M. caroli* (yellow)
553 and *M. pahari* (blue) paternal chromosomes, starting from pericentric ends. Both *M. musculus*
554 and *M. caroli* chromosomes have significantly higher DAPI intensity than *M. pahari*
555 chromosomes within the first two microns (B, gray box), (*** P <0.001). Statistical significance
556 was calculated by a Kruskal-Wallis test, followed by Dunn's post hoc test with Bonferroni

557 correction. **E)** Histograms show the number of various A/T run lengths per 10 kb bin along
558 portions of two genomic contigs containing arrays of either *M. musculus* major satellite (left) or
559 *M. pahari* pi satellite (right), indicated by tandem arrows. **F)** The average number per 10 kb bin
560 of all possible 4W (left) and 5W (right) sequences within the major and pi satellites arrays
561 shown in E. **G)** Schematic (left) shows *in silico* competition assays between different A/T run
562 sequences for the binding of a single peptide encoding the AT-hook of Cbx2, using
563 AlphaFold3. Results (right) show AT-hook binding preferences. **H)** Histograms show the
564 number of various A_nT sequences per 10 kb bin along the same genomic contigs as in E.
565
566 **Figure 3. PRC1 pericentric heterochromatin inhibits CPC recruitment. A)** Schematic
567 shows how H2AK119ub1 can inhibit H2AT121phos and CPC recruitment, resulting in
568 increased distance between metaphase sister kinetochores at paternal *M. musculus*
569 pericentromeres (right) vs *M. pahari* pericentromeres (left). **B)** Pure *M. musculus* zygotes were
570 arrested at metaphase by APC/C inhibition (proTAME), then fixed and stained for RNF2
571 (PRC1 subunit, magenta) and H2AT121phos (green). Top and bottom insets show paternal
572 (RNF2 positive) and maternal (RNF2 negative) chromosomes, respectively; arrowheads point
573 to pericentromeric ends of chromosomes. Graphs show average H2AT121phos and RNF2
574 intensity along maternal (n=56) and paternal (n=54) chromosomes, starting from pericentric
575 ends. Statistical differences between maternal and paternal chromosomes from 0-1 and 2-3
576 microns along their length (gray boxes) were calculated. **C)** Zygotes generated with *M.*
577 *musculus* or *M. pahari* sperm were arrested in mitosis with a kinesin-5 inhibitor (STLC), then
578 fixed and stained for Survivin (green) and H3K9me3 (magenta). Arrowheads in insets point to
579 pericentric Survivin staining. Graph shows the mean Survivin intensity of maternal and paternal
580 pericentromeres in both types of zygotes. Each point represents a single pericentromere
581 (n=253-374 for each group) and boxes represent interquartile ranges. **D)** Zygotes generated
582 with *M. musculus* or *M. pahari* sperm were arrested at metaphase by APC/C inhibition
583 (proTAME), then fixed and stained for CENP-C (green) and H3K9me3 (magenta). Arrowhead
584 points to a pair of paternal *M. musculus* sister kinetochores. Graph shows distances between
585 sister kinetochores for maternal and paternal chromosomes in both types of zygotes. Each
586 point represents a single pair (n=230-284 for each group). All *P*-values were calculated by a

587 Kruskal-Wallis test followed by Dunn's test using Bonferroni correction for multiple testing (***
588 $P < 0.001$). Images are max intensity z-projections; scale bars 10 μm or 1 μm (insets).
589
590
591 **References**

592 1. Altemose, N. *et al.* Complete genomic and epigenetic maps of human centromeres. *Science*
593 **376**, eabl4178 (2022).
594 2. Dudka, D. *et al.* Satellite DNA shapes dictate pericentromere packaging in female meiosis.
595 *Nature* **638**, 814–822 (2025).
596 3. Thakur, J., Packiaraj, J. & Henikoff, S. Sequence, Chromatin and Evolution of Satellite DNA.
597 *IJMS* **22**, 4309 (2021).
598 4. Melters, D. P. *et al.* Comparative analysis of tandem repeats from hundreds of species
599 reveals unique insights into centromere evolution. *Genome Biol* **14**, R10 (2013).
600 5. Henikoff, S. The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA.
601 *Science* **293**, 1098–1102 (2001).
602 6. Saint-Leandre, B. & Levine, M. T. The Telomere Paradox: Stable Genome Preservation with
603 Rapidly Evolving Proteins. *Trends in Genetics* **36**, 232–242 (2020).
604 7. Nurk, S. *et al.* The complete sequence of a human genome. *Science* **376**, 44–53 (2022).
605 8. Logsdon, G. A. & Eichler, E. E. The Dynamic Structure and Rapid Evolution of Human
606 Centromeric Satellite DNA. *Genes* **14**, 92 (2022).
607 9. Gambogi, C. W. *et al.* Centromere innovations within a mouse species. *Science Advances*
608 **9**, eadi5764 (2023).
609 10. Black, B. E. & Cleveland, D. W. Epigenetic Centromere Propagation and the Nature of
610 CENP-A Nucleosomes. *Cell* **144**, 471–479 (2011).

611 11. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: Guardian of the
612 Genome. *Annu. Rev. Cell Dev. Biol.* **34**, 265–288 (2018).
613 12. French, B. T., Westhorpe, F. G., Limouse, C. & Straight, A. F. Xenopus laevis M18BP1
614 Directly Binds Existing CENP-A Nucleosomes to Promote Centromeric Chromatin
615 Assembly. *Developmental Cell* **42**, 190-199.e10 (2017).
616 13. Fioriniello, S., Marano, D., Fiorillo, F., D'Esposito, M. & Della Ragione, F. Epigenetic
617 Factors that Control Pericentric Heterochromatin Organization in Mammals. *Genes* **11**, 595
618 (2020).
619 14. Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive
620 heterochromatin in early mouse embryos. *Nat Genet* **40**, 411–420 (2008).
621 15. Heras, S., Smits, K., Leemans, B. & Van Soom, A. Asymmetric histone 3 methylation
622 pattern between paternal and maternal pronuclei in equine zygotes. *Analytical Biochemistry*
623 **471**, 67–69 (2015).
624 16. Hou, J. *et al.* Epigenetic modification of histone 3 at lysine 9 in sheep zygotes and its
625 relationship with DNA methylation. *BMC Dev Biol* **8**, 60 (2008).
626 17. Pichugin, A. *et al.* Dynamics of constitutive heterochromatin: two contrasted kinetics of
627 genome restructuring in early cloned bovine embryos. *REPRODUCTION* **139**, 129–137
628 (2010).
629 18. Atinbayeva, N. *et al.* Inheritance of H3K9 methylation regulates genome architecture in
630 Drosophila early embryos. *EMBO J* **43**, 2685–2714 (2024).
631 19. Tang, R. *et al.* H3K14ac facilitates the reinstallation of constitutive heterochromatin in
632 *Drosophila* early embryos by engaging Eggless/SetDB1. *Proc. Natl. Acad. Sci. U.S.A.* **121**,
633 e2321859121 (2024).

634 20. Fukushima, H. S., Ikeda, T., Ikeda, S. & Takeda, H. Cell cycle length governs
635 heterochromatin reprogramming during early development in non-mammalian vertebrates.
636 *EMBO Rep* **25**, 3300–3323 (2024).
637 21. Van Der Heijden, G. W. *et al.* Transmission of modified nucleosomes from the mouse
638 male germline to the zygote and subsequent remodeling of paternal chromatin.
639 *Developmental Biology* **298**, 458–469 (2006).
640 22. Bao, J. & Bedford, M. T. Epigenetic regulation of the histone-to-protamine transition
641 during spermiogenesis. *Reproduction* **151**, R55–R70 (2016).
642 23. Dubruille, R. *et al.* Histone removal in sperm protects paternal chromosomes from
643 premature division at fertilization. *Science* **382**, 725–731 (2023).
644 24. Das, A., Smoak, E. M., Linares-Saldana, R., Lampson, M. A. & Black, B. E. Centromere
645 inheritance through the germline. *Chromosoma* **126**, 595–604 (2017).
646 25. Lin, C.-J., Koh, F. M., Wong, P., Conti, M. & Ramalho-Santos, M. Hira-Mediated H3.3
647 Incorporation Is Required for DNA Replication and Ribosomal RNA Transcription in the
648 Mouse Zygote. *Developmental Cell* **30**, 268–279 (2014).
649 26. Loppin, B. *et al.* The histone H3.3 chaperone HIRA is essential for chromatin assembly
650 in the male pronucleus. *Nature* **437**, 1386–1390 (2005).
651 27. Saksouk, N. *et al.* Redundant Mechanisms to Form Silent Chromatin at Pericentromeric
652 Regions Rely on BEND3 and DNA Methylation. *Molecular Cell* **56**, 580–594 (2014).
653 28. Tardat, M. *et al.* Cbx2 Targets PRC1 to Constitutive Heterochromatin in Mouse Zygotes
654 in a Parent-of-Origin-Dependent Manner. *Molecular Cell* **58**, 157–171 (2015).
655 29. Santenard, A. *et al.* Heterochromatin formation in the mouse embryo requires critical
656 residues of the histone variant H3.3. *Nat Cell Biol* **12**, 853–862 (2010).

657 30. Liu, Z. *et al.* SUMO ylated PRC 1 controls histone H3.3 deposition and genome integrity
658 of embryonic heterochromatin. *EMBO J* **39**, (2020).
659 31. Kawaguchi, T., Machida, S., Kurumizaka, H., Tagami, H. & Nakayama, J.
660 Phosphorylation of CBX2 controls its nucleosome-binding specificity. *The Journal of*
661 *Biochemistry* **162**, 343–355 (2017).
662 32. Battista, S. *et al.* Binding to the Other Side: The AT-Hook DNA-Binding Domain Allows
663 Nuclear Factors to Exploit the DNA Minor Groove. *IJMS* **25**, 8863 (2024).
664 33. Reeves, R. & Nissen, M. S. The A.T-DNA-binding domain of mammalian high mobility
665 group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *Journal*
666 *of Biological Chemistry* **265**, 8573–8582 (1990).
667 34. Huth, J. R. *et al.* The solution structure of an HMG-I(Y)–DNA complex defines a new
668 architectural minor groove binding motif. *Nat Struct Mol Biol* **4**, 657–665 (1997).
669 35. Reeves, R. Structure and Function of the HMGI(Y) Family of Architectural Transcription
670 Factors. *Environmental Health Perspectives* **108**, 803–809 (2000).
671 36. Rohs, R. *et al.* The role of DNA shape in protein–DNA recognition. *Nature* **461**, 1248–
672 1253 (2009).
673 37. Arora, U. P., Charlebois, C., Lawal, R. A. & Dumont, B. L. Population and subspecies
674 diversity at mouse centromere satellites. *BMC Genomics* **22**, 279 (2021).
675 38. Stein, P. & Schultz, R. M. Chapter 14 - ICSI in the Mouse. in *Methods in Enzymology*
676 (eds. Wassarman, P. M. & Soriano, P. M.) vol. 476 251–262 (Academic Press, 2010).
677 39. Jukam, D., Shariati, S. A. M. & Skotheim, J. M. Zygotic Genome Activation in
678 Vertebrates. *Developmental Cell* **42**, 316–332 (2017).

679 40. Ferree, P. M. & Barbash, D. A. Species-Specific Heterochromatin Prevents Mitotic
680 Chromosome Segregation to Cause Hybrid Lethality in Drosophila. *PLOS Biology* **7**,
681 e1000234 (2009).
682 41. Josefsson, C., Dilkes, B. & Comai, L. Parent-Dependent Loss of Gene Silencing during
683 Interspecies Hybridization. *Current Biology* **16**, 1322–1328 (2006).
684 42. Gibeaux, R. *et al.* Paternal chromosome loss and metabolic crisis contribute to hybrid
685 inviability in Xenopus. *Nature* **553**, 337–341 (2018).
686 43. Wang, L. *et al.* Hierarchical Recruitment of Polycomb Group Silencing Complexes.
687 *Molecular Cell* **14**, 637–646 (2004).
688 44. Cao, R. *et al.* Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing.
689 *Science* **298**, 1039–1043 (2002).
690 45. Blackledge, N. P., Rose, N. R. & Klose, R. J. Targeting Polycomb systems to regulate
691 gene expression: modifications to a complex story. *Nat Rev Mol Cell Biol* **16**, 643–649
692 (2015).
693 46. Breusegem, S. Y., Clegg, R. M. & Loontiens, F. G. Base-sequence specificity of
694 Hoechst 33258 and DAPI binding to five (A/T)4 DNA sites with kinetic evidence for more
695 than one high-affinity Hoechst 33258-AATT complex. *Journal of Molecular Biology* **315**,
696 1049–1061 (2002).
697 47. Ren, J. & Chaires, J. B. Sequence and Structural Selectivity of Nucleic Acid Binding
698 Ligands. *Biochemistry* **38**, 16067–16075 (1999).
699 48. Tanious, F. A., Veal, J. M., Buczak, H., Ratmeyer, L. S. & Wilson, W. D. DAPI (4',6-
700 diamidino-2-phenylindole) binds differently to DNA and RNA: minor-groove binding at AT
701 sites and intercalation at AU sites. *Biochemistry* **31**, 3103–3112 (1992).

702 49. Traut, W., Weichenhan, D., Eickhoff, U. & Winking, H. Origin of the Chromosome 1
703 HSR of the House Mouse Detected by CGH. *Chromosome Res* **7**, 649–653 (1999).
704 50. Thakur, S., Cattoni, D. I. & Nöllmann, M. The fluorescence properties and binding
705 mechanism of SYTOX green, a bright, low photo-damage DNA intercalating agent. *Eur*
706 *Biophys J* **44**, 337–348 (2015).
707 51. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with
708 AlphaFold 3. *Nature* **630**, 493–500 (2024).
709 52. Carmena, M., Wheelock, M., Funabiki, H. & Earnshaw, W. C. The chromosomal
710 passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol*
711 **13**, 789–803 (2012).
712 53. Yamagishi, Y., Honda, T., Tanno, Y. & Watanabe, Y. Two Histone Marks Establish the
713 Inner Centromere and Chromosome Bi-Orientation. *Science* **330**, 239–243 (2010).
714 54. Kawashima, S. A. *et al.* Shugoshin enables tension-generating attachment of
715 kinetochores by loading Aurora to centromeres. *Genes Dev.* **21**, 420–435 (2007).
716 55. Aihara, H. *et al.* Histone H2A T120 Phosphorylation Promotes Oncogenic
717 Transformation via Upregulation of Cyclin D1. *Molecular Cell* **64**, 176–188 (2016).
718 56. Akera, T., Trimm, E. & Lampson, M. A. Molecular Strategies of Meiotic Cheating by
719 Selfish Centromeres. *Cell* **178**, 1132-1144.e10 (2019).
720 57. Schober, C. S., Aydiner, F., Booth, C. J., Seli, E. & Reinke, V. The kinase VRK1 is
721 required for normal meiotic progression in mammalian oogenesis. *Mechanisms of*
722 *Development* **128**, 178–190 (2011).
723 58. McVey, S. L., Cosby, J. K. & Nannas, N. J. Aurora B Tension Sensing Mechanisms in
724 the Kinetochore Ensure Accurate Chromosome Segregation. *IJMS* **22**, 8818 (2021).

725 59. DeLuca, J. G. *et al.* Kinetochore Microtubule Dynamics and Attachment Stability Are
726 Regulated by Hec1. *Cell* **127**, 969–982 (2006).
727 60. Kitajima, T. S. *et al.* Shugoshin collaborates with protein phosphatase 2A to protect
728 cohesin. *Nature* **441**, 46–52 (2006).
729 61. Riedel, C. G. *et al.* Protein phosphatase 2A protects centromeric sister chromatid
730 cohesion during meiosis I. *Nature* **441**, 53–61 (2006).
731 62. Yueh, W.-T., Singh, V. P. & Gerton, J. L. Maternal *Smc3* protects the integrity of the
732 zygotic genome through DNA replication and mitosis. *Development* **148**, dev199800 (2021).
733 63. Logsdon, G. A. *et al.* The variation and evolution of complete human centromeres.
734 *Nature* **629**, 136–145 (2024).
735 64. Goto, H. *et al.* Identification of a Novel Phosphorylation Site on Histone H3 Coupled with
736 Mitotic Chromosome Condensation. *Journal of Biological Chemistry* **274**, 25543–25549
737 (1999).
738 65. Fischle, W. *et al.* Regulation of HP1–chromatin binding by histone H3 methylation and
739 phosphorylation. *Nature* **438**, 1116–1122 (2005).
740 66. Rong, Y. *et al.* USP16-mediated histone H2A lysine-119 deubiquitination during oocyte
741 maturation is a prerequisite for zygotic genome activation. *Nucleic Acids Research* **50**,
742 5599–5616 (2022).
743 67. Geng, Z. & Gao, Z. Mammalian PRC1 Complexes: Compositional Complexity and
744 Diverse Molecular Mechanisms. *IJMS* **21**, 8594 (2020).
745 68. Das, A. *et al.* Epigenetic, genetic and maternal effects enable stable centromere
746 inheritance. *Nat Cell Biol* **24**, 748–756 (2022).

747 69. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat*
748 *Methods* **9**, 676–682 (2012).
749 70. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLoS*
750 *Comput Biol* **9**, e1003118 (2013).
751

Figure 1

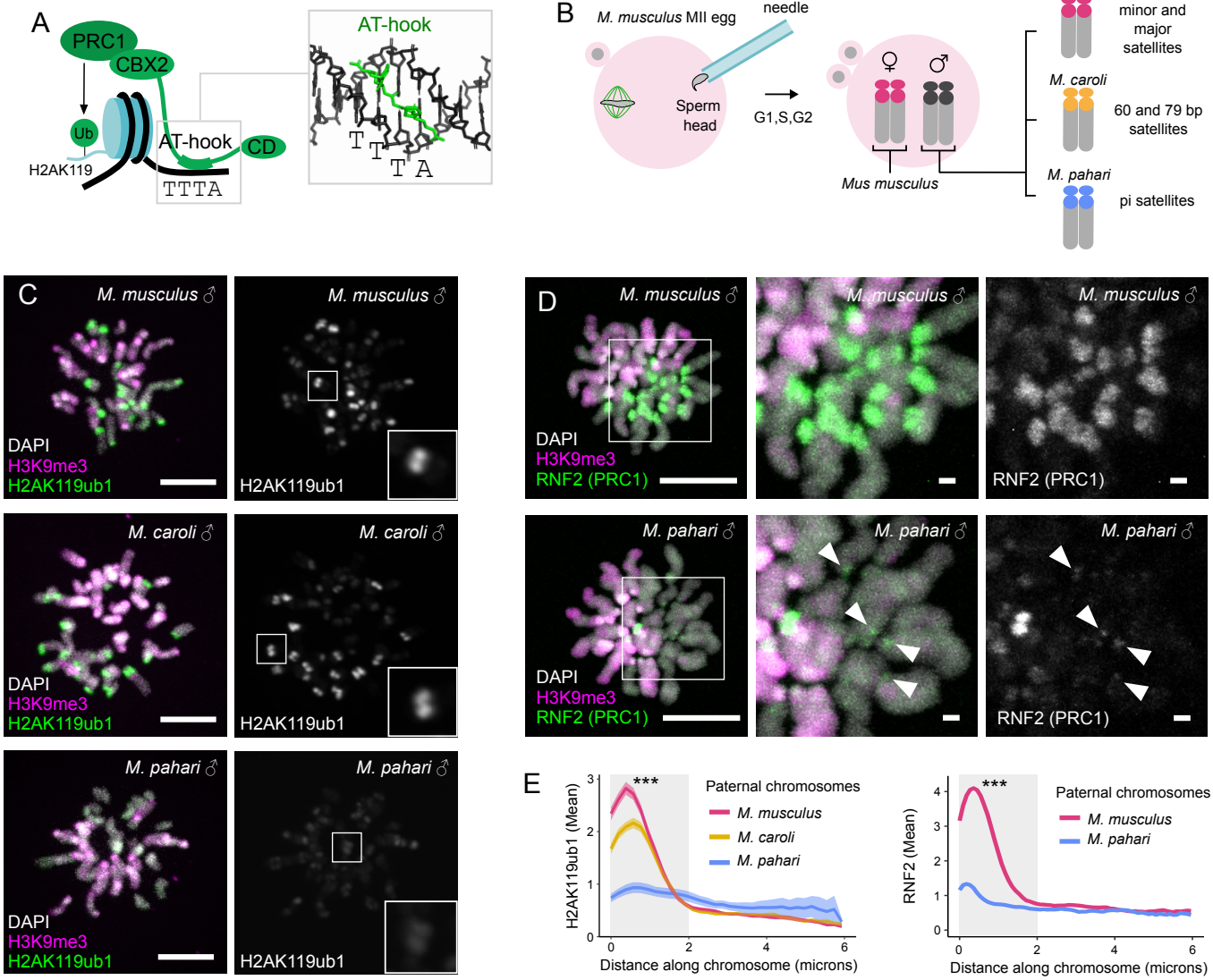

A
PRC1
CBX2
Ub
H2AK119
AT-hook
CD
TTTA
AT-hook
B
M. musculus MII egg
ICSI needle
Sperm head
G1,S,G2
Mus musculus
M. musculus
minor and major satellites
M. caroli
60 and 79 bp satellites
M. pahari
pi satellites
C
M. musculus ♂
DAPI
H3K9me3
H2AK119ub1
M. caroli ♂
M. pahari ♂
D
RNF2 (PRC1)
E
H2AK119ub1 (Mean)
Paternal chromosomes
M. musculus
M. caroli
M. pahari
Distance along chromosome (microns)
RNF2 (Mean)

Figure 2

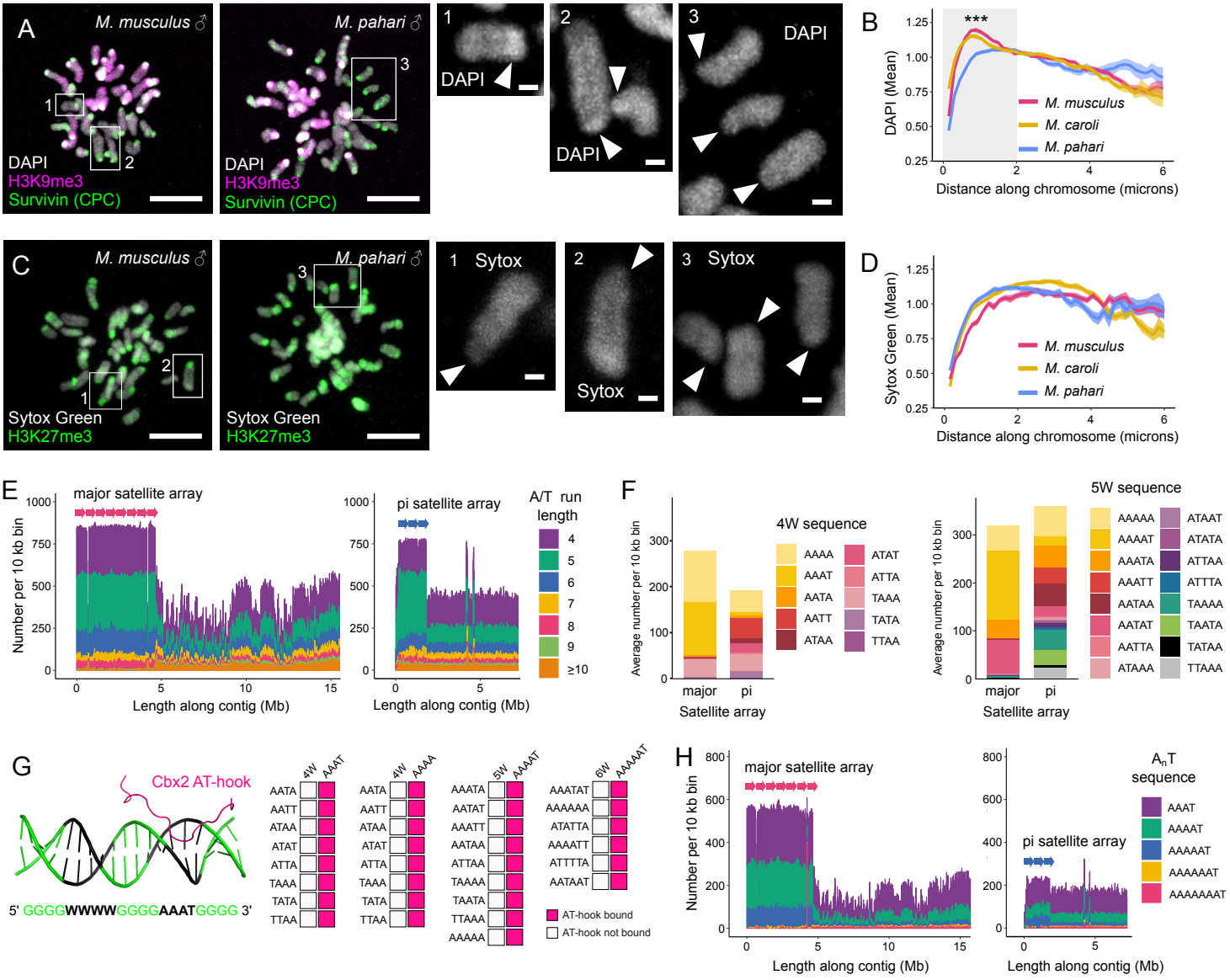
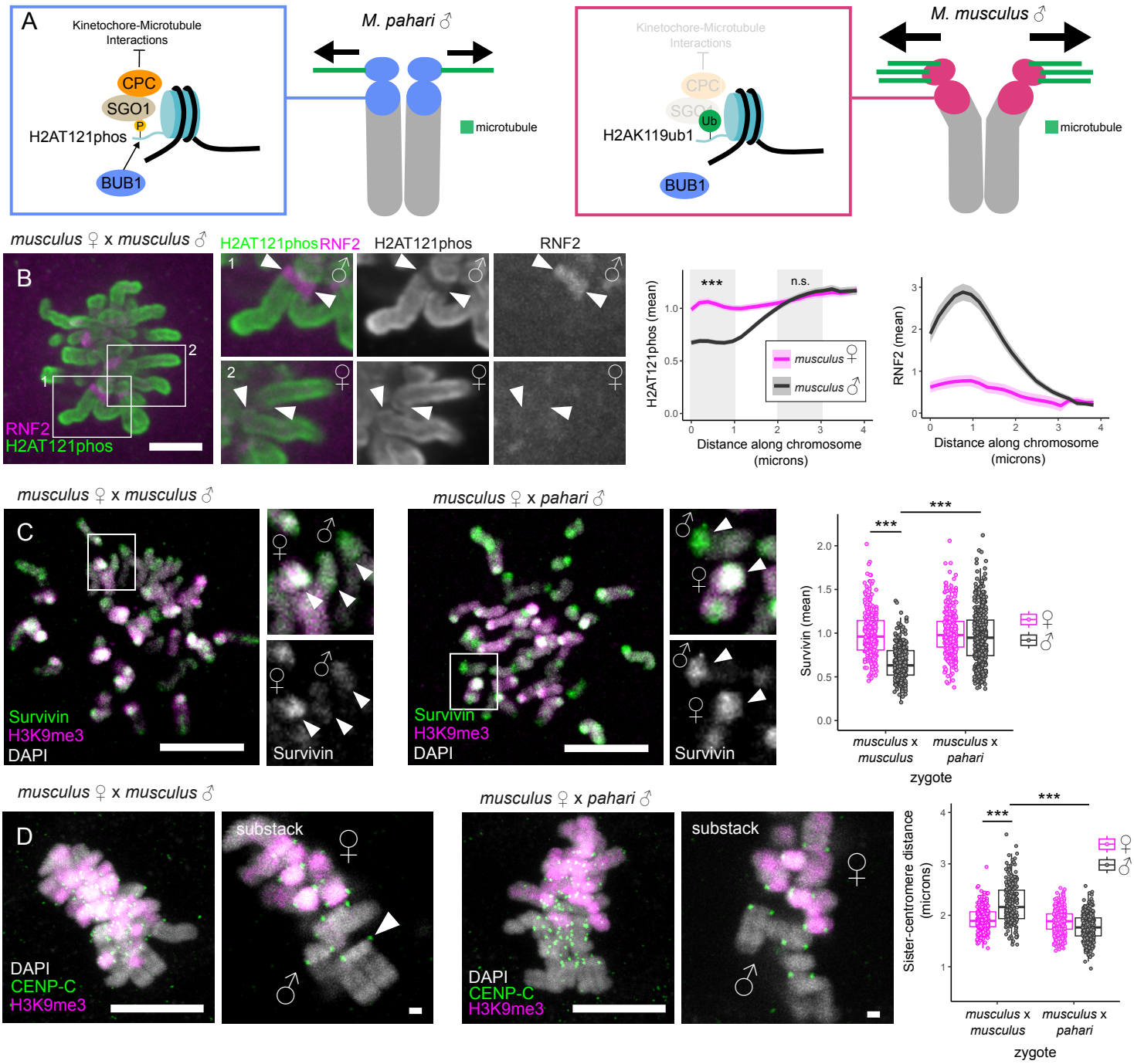

A
M. musculus ♂
M. pahari ♂
DAPI
H3K9me3
Survivin (CPC)
B
***
DAPI (Mean)
M. musculus
M. caroli
M. pahari
Distance along chromosome (microns)
C
Sytox Green
H3K27me3
Sytox
D
Sytox Green (Mean)
E
major satellite array
pi satellite array
Number per 10 kb bin
Length along contig (Mb)
A/T run length
F
4W sequence
5W sequence
Average number per 10 kb bin
major
pi
Satellite array
G
Cbx2 AT-hook
5' GGGGWWWWGGGGAAATGGGG 3'
AT-hook bound
AT-hook not bound
H
A_nT sequence

Figure 3


A
Kinetochore-Microtubule Interactions
CPC
SGO1
P
H2AT121phos
BUB1
M. pahari ♂
microtubule
Kinetochore-Microtubule Interactions
CPC
SGO1
Ub
H2AK119ub1
BUB1
M. musculus ♂
microtubule
musculus ♀ x musculus ♂
B
RNF2
H2AT121phos
H2AT121phos RNF2
H2AT121phos
RNF2
***
n.s.
H2AT121phos (mean)
musculus ♀
musculus ♂
Distance along chromosome (microns)
RNF2 (mean)
Distance along chromosome (microns)
musculus ♀ x musculus ♂
C
Survivin
H3K9me3
DAPI
Survivin
musculus ♀ x pahari ♂
Survivin
H3K9me3
DAPI
Survivin
Survivin (mean)
musculus x musculus
musculus x pahari
zygote
musculus ♀ x musculus ♂
D
DAPI
CENP-C
H3K9me3
substack
musculus ♀ x pahari ♂
DAPI
CENP-C
H3K9me3
substack
Sister-centromere distance (microns)
musculus x musculus
musculus x pahari
zygote