Cross-species incompatibility between a DNA satellite and a chromatin protein poisons germline genome integrity

Cara L. Brand and Mia T. Levine*

Department of Biology and Epigenetics Institute, University of Pennsylvania, Philadelphia, PA

Running title: adaptive evolution to preserve genome integrity

Keywords: DNA satellite, coevolution, chromatin, maternal haploid, 359

*Corresponding Author:
Dr. Mia Levine, University of Pennsylvania, 433 S. University Avenue, Philadelphia, PA 19104.
p: 215-573-9709 e: m.levine@sas.upenn.edu.

ABSTRACT

Satellite DNA spans megabases of eukaryotic genome sequence [1]. These vast stretches of tandem DNA repeats undergo high rates of sequence turnover, resulting in radically different satellite DNA landscapes between closely related species [2-4]. Such extreme evolutionary plasticity suggests that satellite DNA accumulates mutations with no functional consequence. Paradoxically, satellite-rich genomic regions support essential, conserved nuclear processes, including chromosome segregation, dosage compensation, and nuclear structure [5-10]. A leading resolution to this paradox is that deleterious alterations to satellite DNA trigger adaptive evolution of chromatin proteins to preserve these essential functions [11]. Here we experimentally test this model of coevolution between chromatin proteins and DNA satellites by conducting an evolution-guided manipulation of both protein and satellite. We focused on an adaptively evolving, ovary-enriched chromatin protein, called Maternal Haploid (MH) from Drosophila. MH co-localizes with an 11 Mb 359-bp satellite array present in *Drosophila melanogaster* but absent in its sister species, *D. simulans* [12]. Using CRISPR/Cas9-mediated transgenesis, we swapped the *D. simulans* version of MH into *D. melanogaster*. We discovered that *D. melanogaster* females encoding only the *D. simulans mh* ("*mh[sim]*") do not phenocopy the *mh* null mutation. Instead, MH[sim] is toxic to *D. melanogaster* ovaries—we observed elevated ovarian cell death, reduced ovary size, and subfertility in *mh[sim]* females. Using both cell biological and genetic approaches, we demonstrate that MH[sim] poisons oogenesis through a DNA damage pathway. Remarkably, deleting the *D. melanogaster*-specific 359 satellite array from *mh[sim]* females completely restores female germline genome integrity and fertility. This genetic rescue offers experimental evidence that rapid evolution resulted in a cross-species incompatibility between the 359 satellite and MH. These data suggest that coevolution between ostensibly inert repetitive DNA and essential chromatin proteins preserves germline genome integrity.

RESULTS AND DISCUSSION

Satellite-enriched genomic regions evolve rapidly and yet support strictly conserved nuclear functions [2-10]. A classic resolution to this paradox is that DNA satellite-associated proteins evolve adaptively to mitigate deleterious changes to DNA satellite sequence [11]. Repeated bouts of DNA satellite evolution and host chromatin protein adaptation result in exquisitely coevolved satellites and satellite-associated chromatin proteins. This model of coevolution predicts pervasive incompatibilities between satellite DNA and chromatin proteins from closely related species: adaptively evolving chromatin proteins from one species should fail to package or process DNA satellites from another [11, 13, 14].

Evidence for this coevolution model has emerged from engineering “evolutionary mismatches” between the chromatin protein(s) of one species and the DNA satellite landscape of a close relative. Under one approach, a diverged chromatin protein is introduced into a closely related species, generating an evolutionary mismatch between the manipulated protein and one or more DNA satellites [14-17]. Consistent with disrupted DNA satellite:chromatin protein coevolution, the naïve chromatin protein typically perturbs a satellite-mediated function, such as chromosome segregation or nuclear organization [14, 16, 17]. In these cases, however, the incompatible DNA satellites are unknown. A second approach crosses sister species to generate evolutionary mismatches between chromatin proteins and DNA satellites. Consistent with disrupted DNA satellite:chromatin protein coevolution, interspecies hybrid inviability has been linked to DNA satellites [18, 19]. In these cases, however, the incompatible chromatin proteins are unknown. To date, there are no cases of experimental identification of both chromatin protein and satellite putatively engaged in coevolution.

To experimentally probe *both sides* of the coevolution model, we searched for a rapidly evolving DNA satellite that colocalizes with an adaptively evolving chromatin protein. In *Drosophila melanogaster*, the 359-bp satellite occurs in an 11 Mb array at the base of the X chromosome [20, 21]. Close relatives of *D. melanogaster*, including *D. simulans* and *D. erecta*, lack this satellite array on the X chromosome [22]. Instead, these close relatives of *D. melanogaster* have shorter arrays of “359-like” sequence dispersed throughout heterochromatin and euchromatin [3, 23, 24]. Such extreme lineage-restriction to *D. melanogaster* makes this DNA repeat an ideal satellite to test the coevolution model. On the protein side, we identified from the literature Maternal Haploid (MH), a protein that colocalizes with the 359 satellite [12]. MH is ovary-enriched protein and maternally provisioned to the embryo [12, 25]. The name “*maternal haploid*” comes from a striking phenotype in the embryos of *mh* null mothers: paternal chromosomes fail to participate in the first embryonic mitosis. Only 20% of these embryos develop beyond the first division and with only the haploid complement of maternal chromosomes. Moreover, developing “maternal haploid” embryos from *mh* null mothers accumulate fragmented arrays of 359 [12].

If 359 proliferation triggered *mh* to innovate, we should detect evidence of positive selection at *mh* between *D. melanogaster* and *D. simulans*. To determine if *mh* evolves adaptively, we conducted a McDonald-Kreitman test [26] using polymorphism within *D. melanogaster* and

D. simulans and divergence between *D. melanogaster* and *D. simulans* (2.5 million years diverged [27]. This comparison revealed an excess of nonsynonymous fixations, consistent with a history of adaptive evolution (Figure 1A). The dynamic evolution of the 359 satellite and adaptive evolution of a 359-associated protein, MH, raise the possibility that *mh* recurrently evolves to preserve a biological function compromised by 359 satellite proliferation.

To test the possibility of MH-359 coevolution, we first conducted an evolution-guided manipulation of *mh*. We predicted that generating an "evolutionary mismatch" between the *D. simulans mh* and the *D. melanogaster* 359bp X-linked array would compromise an essential nuclear function. To perform this swap, we used CRISPR/Cas9 to integrate into the native *mh* locus either a 3xFLAG-tagged *D. melanogaster mh* coding sequence (our control fly, "*mh[mel]*") or a 3xFLAG-tagged *D. simulans mh* coding sequence (our experimental fly, "*mh[sim]*", Figure 1B). Both the *D. melanogaster* and the *D. simulans* coding sequences were codon-optimized for *D. melanogaster.* We observed equivalent expression of the two transgenes (Figure S1A). Given that MH is ovary-enriched [12, 25], we first quantified female fertility across our two genotypes. *mh[sim]* females produced significantly fewer progeny than the control *mh[mel]* females (Figure 1C). Upon dissecting the ovaries from each genotype, we observed significantly smaller *mh[sim]* ovaries (Figure 1D).

An oogenesis defect was unexpected: previous reports suggest that *mh* null alone yields no ovary phenotype [12, 25]. Consistent with these data, we detected no difference in ovary size of *mh* null mothers compared to heterozygous controls (Figure S1B). To further probe the difference between *mh[sim]* and *mh* null females, we assayed the embryos of *mh[sim]* mothers. If *mh[sim]* behaves as loss of function allele in a *D. melanogaster* genetic background, we expect *mh[sim]* mothers to give rise to elevated "maternal haploid" embryos, similar to *mh* null mothers. To address this prediction, we crossed *mh[mel]* mothers and *mh[sim]* mothers to fathers homozygous for an EGFP-tagged centromere protein, CID. Diploid embryos, but not maternal haploid embryos, inherit the paternally-supplied *EGFP-cid* transgene. We observed no difference in the incidence of EGFP:CID-lacking embryos across *mh[mel]* and *mh[sim]* mothers. (Figure S1C). These data, combined with the unexpected oogenesis phenotype, suggest that *mh[sim]* does not behave as a loss-of-function allele. Instead, MH[sim] might be toxic.

To explore the possibility that MH[sim] is toxic, we first asked if MH[sim] localizes aberrantly in the ovary. We visualized MH[mel] and MH[sim] by staining ovaries with anti-FLAG. We discovered that while MH[mel] localized primarily in the earliest stages of oogenesis (the germarium), MH[sim] localized not only in these early stages of oogenesis but also on the "nurse cell" nuclei of later stage egg chambers and weakly on the somatic follicle cell nuclei (Figure 2A, 2B, S2). The aberrant persistence of MH[sim] during oogenesis, combined with compromised *mh[sim]* ovary development, raised the possibility that MH mislocalization alone might be toxic. To test this hypothesis, we used the UAS-GAL4 system to overexpress MH[mel] in the female germline (driver nos-Gal4-VP16, Figure 2C). In ovaries overexpressing MH[mel], we indeed observed elevated levels and aberrant persistence of the protein in later stage egg chambers. Nevertheless, these females gave rise to abundant progeny (Figure 2C), suggesting that mislocalization alone cannot explain the compromised ovary development of *mh[sim]*

females. In contrast, overexpression of MH[sim] in otherwise wildtype ovaries resulted in female sterility (Figure 2C). These data suggest that MH[sim] – which functions normally in its native *D. simulans* genome – is toxic to oogenesis in *D. melanogaster*. Indeed, in the ovaries expressing the CRISPR-introduced *mh[sim]* transgene under the native promoter (Figure 1B), we discovered an excess of hyper-condensed nuclei, consistent with elevated cell death (Figure 2D [28, 29]).

A classic trigger of cell death is the accumulation of DNA damage [30]. To visualize DNA damage, we stained *mh[mel]* and *mh[sim]* ovaries for the double-stranded break marker, γH2Av. We observed elevated DNA damage signaling in *mh[sim]* ovaries (Figure 3A). These data suggest that MH[sim] triggers DNA damage in the ovary. To address the hypothesis that MH[sim] compromises oogenesis through a DNA repair pathway, we combined *mh[sim]* with a null mutation in a DNA damage checkpoint gene. The gene, *Chk2* (also known as *mnk*), normally blocks egg production in the presence of DNA damage [31, 32]. *Chk2*$^{-/-}$ ovaries bypass this checkpoint, allowing a female to make mature but damaged eggs in the presence of elevated DNA damage. We discovered that *Chk2*$^{-/-}$ restores *mh[sim]* ovaries to *mh[mel]*-like ovary size (Figure 3B). However, the *mh[sim*];*Chk2*$^{-/-}$ females are sterile while *mh[mel]*];*Chk2*$^{-/-}$ females retain fertility (Figure 3C). These data suggest that MH[sim] compromises oogenesis by triggering DNA damage.

Applying these phenotypic data to the coevolution model, we hypothesized that the expanded 11 Mb array of 359 satellite in *D. melanogaster* triggers DNA damage in the presence of MH[sim]. Under this model, MH[sim]-specific residues are incompatible with 359. Removing 359 should restore germline genome integrity and fertility of *mh[sim]* females. To directly test this prediction, we took advantage of a fly strain that lacks the 11 Mb array of X-linked 359 satellite [33]. We recombined this 359-deletion, called *Zygotic hybrid rescue*, or "*Zhr*," onto both the *mh[mel]* and the *mh[sim]* X chromosomes (Figure S3A). If MH[sim]-induced toxicity depends on the presence of the 359 expansion, *mh[sim],Zhr* females should have minimal DNA damage and recover fertility. Remarkably, the 359-deletion completely restores the DNA damage signal γH2Av to wildtype (low) levels (Figure 4A). Consistent with restored germline genome integrity of *mh[sim]* females, we observed no difference in ovary size of *mh[mel]* and *mh[sim]* females that lack 359 (Figure 4B). Finally, the 359-deletion completely restores *mh[sim]* fertility (Figure 4C). These data suggest that 359 mediates MH[sim] toxicity to oogenesis, consistent with a history of coevolution between these two fast-evolving components of the Drosophila genome (Figure 4D, 4E).

The molecular events that trigger 359-dependent DNA damage in *mh[sim]* ovaries remain to be elucidated. We suspect that MH[sim] interferes with a chromatin-mediated pathway required to mitigate 359-induced challenges to the female germline (and possibly other stages of development). The MH human homolog, Spartan, clips DNA-protein crosslinks that accumulate at replication forks using the Spartan-like protease domain [34-36]. Based on this molecular function, we hypothesize that DNA replication challenges at the 11 Mb array of 359 cause the accumulation of chromatin intermediates replete with persistent DNA-protein crosslinks. These crosslinks are normally resolved by other members of the Spartan family, like GCNA [37]. We

speculate that MH[sim] traps these DNA-protein crosslinks, triggering replication fork collapse and DNA damage (Figure S3B). Alternatively, MH[sim] may disrupt the production of 359 transcripts shown previously to recruit essential chromosome segregation factors to the centromere [10].

The deletion of 359 from the *D. melanogaster* X chromosome renders this genomic region "*D. simulans*-like" [22]. The rescue of germline genome integrity upon combining *mh[sim]* and the *D. simulans*-like 359 deletion implicates 359 proliferation as the antagonizing event that triggered MH adaptive evolution (Figure 4E). However, we cannot rule out the possibility that a selection pressure distinct from 359 proliferation triggered MH adaptive evolution. Under this alternative model, the *D. melanogaster*-specific version of MH evolved first, releasing constraint on 359 copy number. Based on the exceptionally rapid rates of DNA satellite evolution, combined with empirical observations that DNA satellites can behave selfishly by gaining a transmission advantage from one generation to the next [38, 39], we favor instead a model under which 359 proliferated first (Figure 4E).

The 359-mediated toxicity to oogenesis highlights the catastrophic functional consequences of DNA satellite evolution. Importantly, 359-mediated toxicity is also apparent in *D. melanogaster*-*D. simulans* hybrid embryos: a distinct, unmapped gene on the *D. simulans* chromosome 2 [40-43] interacts deleteriously with 359 to cause embryonic chromosome mis-segregation, genome instability, and lethality [18, 33, 44]. This interspecies hybrid dysfunction in the embryo, together with the *mh[sim]*-359 toxicity in the ovary reported here, suggests that recurrent bouts of coevolution not only shape essential genome functions within species but also can trigger hybrid incompatibilities between species. Our experimental identification of both satellite and protein engaged in coevolution offers the clearest evidence to date that ostensibly inert DNA repeats trigger adaptive chromatin protein evolution to maintain genome integrity.

MATERIALS AND METHODS

Population genetic and molecular evolution analyses

We conducted population genetic analysis of *mh* using multiple alleles from both *D. melanogaster* and *D. simulans*. We obtained eight *D. melanogaster mh* alleles (coordinates *X*:15472804-15475400, dmel r6.4) from lines collected in Lyon, France [45]. We amplified seven *D. simulans mh* alleles from lines collected in Nairobi, Kenya [46]. We prepared genomic DNA and conducted PCR amplification followed by Sanger sequencing using standard protocols. We aligned the sequences in Geneious using the Geneious Alignment algorithm with default settings (Geneious v11.1.5, Biomatters, Auckland, New Zealand) and confirmed alignment quality by eye. We performed a McDonald-Kreitman test [26] with the *D. melanogaster* and *D. simulans mh* coding sequences.

Fly stock construction and husbandry

Constructing gene swaps

We used CRISPR/Cas9 to generate *D. melanogaster* flies that encode a transgenic *D. melanogaster* allele or a *D. simulans* allele of *mh,* integrated into the native location. We first generated a U6 promoter-driven guide RNA construct by cloning sgRNAs flanking the coding sequence of *mh* (5': GGATTGGCCCAGGATCAACA, 3': CGTGGAGAGCTTCTGCCGCG) into pBFv-U6.2 and pBFv-U6.2B backbones. We shuttled the 3' sgRNA into pBFv-U6.2 to create a dual sgRNA vector (University of Utah Mutagenesis Core). In parallel, we constructed homology directed repair (HDR) plasmids encoding one kilobase (kb) homology arms 5' and 3' of their respective guide RNAs. Between the homology arms we synthesized a codon-optimized (for *D. melanogaster*) *mh* coding sequence from either *D. melanogaster* or *D. simulans* (GenScript, Piscataway, NJ). We N-terminally tagged each sequence with 3xFLAG along with a linker sequence (GGTGGTTCATCA). We injected the dual sgRNA vector and a single HDR plasmid into the Cas9-expressing line, yw; nos-Cas9(II-attP40) (BestGene Inc, Chino Hills, CA).

We crossed the single males, injected as embryos, to an FM7 (X-chromosome balancer) female. We screened F_1 females to identify positive transformants using forward primer 5'-AAGTGTCGCGCTATTTCACC-3' and reverse primer 5'-TCACCGTCATGGTCTTTGTAGTCCAT-3'. We then backcrossed the positive F_1 females to FM7 males and self-crossed the balanced F_2 progeny to generate lines homozygous for either allele. To confirm that the introduced alleles encoded the expected sequence and in the expected location, we amplified the entire region from homozygous flies using primers that anneal outside of the homology arms (5'-AATGGATTTCGGCAAATGAG-3', 5'-GTCGTTGTAGGAGCCCATGT-3') and then sequenced across the entire region. We also designed primers that amplified the native *mh* locus (5'-GGCCCTGCTCATATCGTATC-3', 5'-AAGAACCTTACTGCGTGCAAC-3') to confirm that our final genotypes were true replacements. Primers can be found in Table S1.

Constructing UAS-mh lines

We used the *ΦC31* integrase-mediated transgenesis system to introduce into the same landing site *mh* from *D. melanogaster* or *D. simulans* downstream of an "upstream activating sequence" or "UAS" [47]. Using the HDR plasmids as a template (see above), we PCR-amplified the 3xFLAG-tagged *mh* coding sequence (either *D. melanogaster* or *D. simulans)* using Phusion High-Fidelity DNA Polymerase (NEB, Ipswich, MA). We cloned the resulting PCR products into *NotI*/*XbaI* sites of the pUASp-attB vector (*Drosophila* Genomics Resource Center, Bloomington, IN). We confirmed the absence of PCR-introduced mutations in the cloned UAS-*mh*[mel] and UAS-*mh*[sim] alleles by direct Sanger sequencing of the constructs (Table S1). We introduced the constructs into *D. melanogaster yw*; PBac[y^+-attP-9A]VK00018 flies, which have an attP transgene landing site at cytological position 75A10 on chromosome 3L, (BestGene Inc, Chino Hills, CA). We next made each transgene homozygous. To overexpress the transgenic alleles, we crossed these stocks to Gal4::VP16-nos (BDSC #64277), which drives germline expression of transgenes downstream of UAS.

Zhr rescue stocks

To generate stocks that encode both the X-linked *mh*-transgene and the X-linked 359 satellite deletion (Zhr^1, BDSC #25140), we first generated trans-heterozygote females. We crossed these trans-heterozygote females to FM7 males and used PCR to assay individual recombinant male progeny for the presence of both the *mh* transgene and *Zhr.* We detected the *mh* transgene with forward primer 5'-AAGTGTCGCGCTATTTCACC-3' and reverse primer 5'-TCACCGTCATGGTCTTTGTAGTCCAT-3'. To detect the *Zhr* mutation (*i.e.*, 359 satellite deletion), we used forward primer 5'-TATTCTTACATCTATGTGACC-3' and reverse primer 5'-GTTTTGAGCAGCTAATTACC-3' [10]. Running this PCR with a 52C annealing temperature for 10 cycles yields a band only in the presence of the 11Mb 359 satellite array. We backcrossed males positive for both the *mh* transgene and *Zhr* mutation to FM7 females to generate a permanent stock.

We used a +/FM7; +/CyO stock to generate flies encoding both the *mh* transgene at the native locus (chromosome *X*) and the $Chk^{-/-}$ (*mnk)* mutation (chromosome 2). The mnk^{l6} stock [32] was a gift from N. Phadnis.

Immunoblotting

To assay protein abundance in the ovary, we dissected 20 ovary pairs in 1X PBS and ground the material in RIPA buffer (Cell Signaling Technology, Danvers, MA), Protease Inhibitor Cocktail (Roche, Basel, Switzerland), and 2X PMSF (Cell Signaling Technology, Danvers, MA). To promote solubility, we incubated the lysate in Benzonase (Sigma Aldrich, St. Louis, MO) for 1 hr at 4C. We used 20 µg of lysate and probed with 1:10,000 anti-FLAG (M2, Sigma Aldrich, St. Louis, MO) or 1:1,000 anti-α-tubulin (Developmental Studies Hybridoma Bank, Iowa City, IA) and 1:1,000 anti-mouse HRP secondaries (Kindle Biosciences, Greenwich, CT). We exposed

blots with Kwikquant Western Blot detection kit and imaged with a Kwikquant imager (Kindle Biosciences, Greenwich, CT).

Fertility assays

Female fertility

To assay female fertility, we first aged virgin females 3-5 days. For each replicate vial, we crossed four virgin females to four w^{1118} males. We conducted all crosses on molasses food at 24°C. We flipped the parents onto new food every three days over the course of nine days and counted all progeny that emerged.

Ovary size

To determine ovary size from the gene swap lines and the mutant line mh^{1} (BDSC #64277), we dissected ovary pairs in 1X PBS and imaged at 8X magnification with a Leica DFC7000 T camera. We quantified the area of each ovary using the polygon tool in FIJI [48] to define the borders of the tissue. For each ovary pair, we used the Measure tool in FIJI to calculate the area (mm^2) within these boundaries for each individual ovary.

Immunofluorescence

We conducted immunofluorescence on ovaries following the protocol described in [49]. We stained ovaries with anti-FLAG (1:3000, M2, Sigma Aldrich, St. Louis, MO) and anti-γH2Av (1:1000, a gift from R. S. Hawley). We mounted ovaries with ProLong Gold Antifade Reagent with DAPI (Thermo Fisher Scientific, Waltham, MA). We imaged slides at 63x magnification on a Leica TCS SP8 Four Channel Spectral Confocal System. For each experiment, we used the same imaging parameters across genotypes.

We conducted immunofluorescence on embryos from *mh[mel]* or *mh[sim]* females crossed to males homozygous for P{gcid.EGFP.cid}III.2 ([50], gift from K. McKim). We followed the protocol described in [51] to fix and stain the embryos with anti-GFP (1:1000, Aves Labs, Tigard, OR). We mounted and imaged the embryos as described above.

Analysis of cytological data

Cell death quantification

To quantify the incidence of cell death, we mounted fixed whole ovaries with ProLong Gold Antifade Reagent with DAPI (Thermo Fisher Scientific, Waltham, MA) and imaged at 63x magnification on a Leica TCS SP8 Four Channel Spectral Confocal System using the tile scanning and merging feature. We identified the number of ovarioles that contained egg chambers with >1 condensed, signal-saturated nurse cell nucleus. We then divided this number

by the total number of ovarioles present in each ovary to determine the fraction of cell death incidence in *mh[mel]* and *mh[sim]* ovaries.

Immunofluorescence quantification

We quantified the average fluorescence of γH2Av in stage four egg chambers in *mh[mel]* and *mh[sim]* ovaries. First, we outlined a representative stage four egg chamber with the Freehand ROI tool in FIJI [48]. We calculated the fluorescent signal intensity following the protocol described by the Keith R. Porter Imaging Facility (https://kpif.umbc.edu/image-processing-resources/imagej-fiji/, UMBC, Baltimore, MD). We normalized the fluorescent signal intensity of *mh[mel]* and *mh[sim]* as well as *mh[mel],Zhr* and *mh[sim],Zhr* to the mean intensity signal of the *mh[mel]* and *mh[mel],Zhr* controls, respectively.

ACKNOWLEDGEMENTS

We thank Isabella Farkas and Courtney Christopher for technical assistance. We also thank the Levine Lab, M. Patel, N. Phadnis, A. Das, and D. Dudka for feedback on the manuscript and the Levine Lab, P. Geyer, H. Malik, and M. Buszczak for discussions about the project. This work was supported by a Life Sciences Research Foundation fellowship to C.L.B. and National Institutes of Health (NIH) NIGMS grant R35GM124684 to M.T.L.

FIGURE LEGENDS

Figure 1. MH evolves adaptively to preserve female fertility. (A) Counts of synonymous and nonsynonymous polymorphic and fixed sites within and between *D. melanogaster* and *D. simulans;* χ^2 test, p = 0.04. (B) Swap strategy: the *D. melanogaster* ("*D. mel*", blue) or *D. simulans* ("*D. sim*," yellow) *mh* coding sequence, codon-optimized for *D. melanogaster* and 3xFLAG-tagged, replaced the native *mh* gene on the X chromosome. (C) Total offspring from *mh[mel]* or *mh[sim]* females crossed to wildtype (w^{1118}) males. (D) Ovary size estimates from *mh[mel]* and *mh[sim]* females. (*t*-test: "*" = $p < 0.05$, "***" = $p < 0.001$, scale bar = 100μm)

Figure 2. MH[sim] is toxic to oogenesis. (A) Diagram of a *Drosophila* ovary (above) and a single ovariole (below) with the germline stem cells in the germarium on the left side and the mature eggs on the right side. The dashed box shows the developmental stages represented in images in (B), (C) and (D) (Created using BioRender.com). (B) Anti-FLAG staining to visualize MH[mel] and MH[sim] localization, under the native promoter. Merged images of single nuclei show no MH foci on the DNA (dashed boxes). (C) Anti-FLAG staining to visualize MH[mel] and MH[sim] localization upon overexpression under the UAS-GAL4 system (nanos-Gal4:VP16 driver) (left). Progeny counts of the nanos-GAL4 driven UASp-*mh[mel]* or UASp-*mh[sim]* females (right). Note that wildtype *mh* is present in these genotypes. (D) Incidence of cell death captured by the fraction of ovarioles with condensed nuclei (arrow heads) in *mh[mel]* and *mh[sim]* ovaries. (*t*-test, "***" = $p < 0.001$, scale bar = 25μm)

Figure 3. MH[sim] triggers DNA damage in the ovary. (A) γH2Av signal in *mh[mel]* and *mh[sim]* ovaries (left) and quantification of normalized fluorescent signal intensity (right). Note that the expected γH2Av-positive cells in the germarium in *mh[mel]* are absent under the imaging parameters used but are indeed present, see Figure S2. (B) Ovary size of *mh[mel];Chk2*$^{-/-}$ and *mh[sim];Chk2*$^{-/-}$ females. (C) Progeny counts from *mh[mel];Chk2*$^{-/-}$ and *mh[sim];Chk2*$^{-/-}$ females crossed to wildtype (w^{1118}) males. (*t*-test, "***" = $p < 0.001$, "n.s." $p > 0.05$, scale bar = 25μm)

Figure 4. The 359 satellite deletion rescues *mh[sim]* genome integrity and fertility. (A) γH2Av signal in *mh[mel],Zhr* (*Zhr* is the 359-deletion locus) and *mh[sim],Zhr* ovaries (left) and quantification of normalized fluorescent signal intensity (right) (B) Ovary size of *mh[mel],Zhr* and *mh[sim],Zhr* females. (C) Progeny counts from *mh[mel],Zhr* and *mh[sim],Zhr* females crossed to wildtype (w^{1118}) males. (*t*-test, "n.s." $p > 0.05$, scale bar = 25μm) (D) Cartoon summary showing outcome of adaptive evolution of *mh* along the *D. melanogaster* (blue square) and *D. simulans* (yellow square) lineages and the expansion of 359 along the *D. melanogaster* lineage (blue semicircles). In *D. simulans*, the X pericentromere lacks the 11Mb 359 array and instead harbors other repeats, possibly repeats that MH[sim] specifically packages (yellow triangles). Engineering an evolutionary mismatch between the *D. melanogaster* 359 locus and the *D. simulans* version of MH resulted in compromised genome integrity. (E) Model for MH evolution tracking the 359 satellite evolution.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. MH[sim] is expressed at comparable levels to MH[mel] and does not phenocopy *mh* null. (A) Western Blot of *mh[mel]* and *mh[sim]* ovaries probed with anti-FLAG or anti-tubulin. (B) Ovary size of *mh[1]* homozygote and heterozygote females. (C) Images of embryos from crosses between *EGFP-cid* fathers and *mh[mel]* or *mh[sim]* mothers stained with anti-GFP to identify diploid embryos.

Figure S2. γH2Av staining in *mh[mel]* germarium. Diagram of a *Drosophila* ovariole (above) showing the germarium where meiotic recombination occurs. Germarium of an *mh[mel]* female (below) showing programmed double stranded breaks (arrowheads) occurring in “region 2A” of the germarium detected under higher laser power than used for images displayed in Figure 3A. These double stranded breaks are repaired via meiotic recombination pathways.

Figure S3. Deleting the X-linked 359 array rescues *mh[sim]* genome integrity. (A) A 10-cycle PCR distinguishes between wildtype 359 copy number and the 359-deletion on the X chromosome (*Zhr*). (B) Model of MH[sim] toxicity. In an *mh[mel]* ovary, transient DNA crosslinks accumulate at 359 but are resolved by other factors such as GCNA [37], which shares a Spartan-like domain with MH. In an *mh[sim]* ovary, MH[sim] localizes to 359 and traps crosslinked proteins, compromising germline genome integrity. In a *Zhr* background, MH[sim] has no target to poison because the 11Mb array of 359 is missing. Consequently, genome integrity is restored.

LITERATURE CITED

1. Thakur, J., Packiaraj, J., and Henikoff, S. (2021). Sequence, Chromatin and Evolution of Satellite DNA. Int J Mol Sci *22*.
2. Cechova, M., Harris, R.S., Tomaszkiewicz, M., Arbeithuber, B., Chiaromonte, F., and Makova, K.D. (2019). High satellite repeat turnover in great apes studied with short- and long-read technologies. Mol Biol Evol *36*, 2415-2431.
3. Sproul, J.S., Khost, D.E., Eickbush, D.G., Negm, S., Wei, X., Wong, I., and Larracuente, A.M. (2020). Dynamic Evolution of Euchromatic Satellites on the X Chromosome in Drosophila melanogaster and the simulans Clade. Mol Biol Evol *37*, 2241-2256.
4. Jagannathan, M., Warsinger-Pepe, N., Watase, G.J., and Yamashita, Y.M. (2017). Comparative Analysis of Satellite DNA in the Drosophila melanogaster Species Complex. G3 (Bethesda) *7*, 693-704.
5. Jagannathan, M., Cummings, R., and Yamashita, Y.M. (2018). A conserved function for pericentromeric satellite DNA. Elife *7*.
6. Joshi, S.S., and Meller, V.H. (2017). Satellite Repeats Identify X Chromatin for Dosage Compensation in Drosophila melanogaster Males. Curr Biol *27*, 1393-1402 e1392.
7. Rudd, M.K., and Willard, H.F. (2004). Analysis of the centromeric regions of the human genome assembly. Trends Genet *20*, 529-533.
8. Round, E.K., Flowers, S.K., and Richards, E.J. (1997). Arabidopsis thaliana centromere regions: genetic map positions and repetitive DNA structure. Genome Res *7*, 1045-1053.
9. Kipling, D., Ackford, H.E., Taylor, B.A., and Cooke, H.J. (1991). Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. Genomics *11*, 235-241.
10. Rosic, S., Kohler, F., and Erhardt, S. (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. J Cell Biol *207*, 335-349.
11. Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098-1102.
12. Tang, X., Cao, J., Zhang, L., Huang, Y., Zhang, Q., and Rong, Y.S. (2017). Maternal Haploid, a Metalloprotease Enriched at the Largest Satellite Repeat and Essential for Genome Integrity in Drosophila Embryos. Genetics *206*, 1829-1839.
13. Ferree, P.M., and Prasad, S. (2012). How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. Genet Res Int *2012*, 430136.
14. Jagannathan, M., and Yamashita, Y.M. (2021). Defective satellite DNA clustering into chromocenters underlies hybrid incompatibility in Drosophila. bioRxiv, 2021.2004.2016.440167.
15. Bayes, J.J., and Malik, H.S. (2009). Altered heterochromatin binding by a hybrid sterility protein in Drosophila sibling species. Science *326*, 1538-1541.
16. Kumon, T., Ma, J., Stefanik, D., Nordgren, E.C., Akins, R.B., Kim, J., Levine, M.T., and Lampson, M.A. (2020). Centromere drive and suppression by parallel pathways for recruiting microtubule destabilizers. bioRxiv, 2020.2011.2026.400515.

17. Maheshwari, S., Tan, E.H., West, A., Franklin, F.C.H., Comai, L., and Chan, S.W.L. (2015). Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. PLoS Genet *11*, e1004970.
18. Ferree, P.M., and Barbash, D.A. (2009). Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLoS Biol *7*, e1000234.
19. Gibeaux, R., Acker, R., Kitaoka, M., Georgiou, G., van Kruijsbergen, I., Ford, B., Marcotte, E.M., Nomura, D.K., Kwon, T., Veenstra, G.J.C., et al. (2018). Paternal chromosome loss and metabolic crisis contribute to hybrid inviability in Xenopus. Nature *553*, 337-341.
20. Lohe, A.R., Hilliker, A.J., and Roberts, P.A. (1993). Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster. Genetics *134*, 1149-1174.
21. Brutlag, D.L. (1980). Molecular arrangement and evolution of heterochromatic DNA. Annu Rev Genet *14*, 121-144.
22. de Lima, L.G., Hanlon, S.L., and Gerton, J.L. (2020). Origins and Evolutionary Patterns of the 1.688 Satellite DNA Family in Drosophila Phylogeny. G3 (Bethesda) *10*, 4129-4146.
23. Abad, J.P., Agudo, M., Molina, I., Losada, A., Ripoll, P., and Villasante, A. (2000). Pericentromeric regions containing 1.688 satellite DNA sequences show anti-kinetochore antibody staining in prometaphase chromosomes of Drosophila melanogaster. Mol Gen Genet *264*, 371-377.
24. Losada, A., and Villasante, A. (1996). Autosomal location of a new subtype of 1.688 satellite DNA of Drosophila melanogaster. Chromosome Res *4*, 372-383.
25. Delabaere, L., Orsi, G.A., Sapey-Triomphe, L., Horard, B., Couble, P., and Loppin, B. (2014). The Spartan ortholog maternal haploid is required for paternal chromosome integrity in the Drosophila zygote. Curr Biol *24*, 2281-2287.
26. McDonald, J.H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in Drosophila. Nature *351*, 652-654.
27. Garrigan, D., Kingan, S.B., Geneva, A.J., Andolfatto, P., Clark, A.G., Thornton, K.R., and Presgraves, D.C. (2012). Genome sequencing reveals complex speciation in the Drosophila simulans clade. Genome Res *22*, 1499-1511.
28. Kerr, J.F., Wyllie, A.H., and Currie, A.R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br J Cancer *26*, 239-257.
29. McCall, K. (2004). Eggs over easy: cell death in the Drosophila ovary. Dev Biol *274*, 3-14.
30. Roos, W.P., Thomas, A.D., and Kaina, B. (2016). DNA damage and the balance between survival and death in cancer biology. Nat Rev Cancer *16*, 20-33.
31. Bakhrat, A., Pritchett, T., Peretz, G., McCall, K., and Abdu, U. (2010). Drosophila Chk2 and p53 proteins induce stage-specific cell death independently during oogenesis. Apoptosis *15*, 1425-1434.
32. Brodsky, M.H., Weinert, B.T., Tsang, G., Rong, Y.S., McGinnis, N.M., Golic, K.G., Rio, D.C., and Rubin, G.M. (2004). Drosophila melanogaster MNK/Chk2 and p53 regulate multiple DNA repair and apoptotic pathways following DNA damage. Mol Cell Biol *24*, 1219-1231.

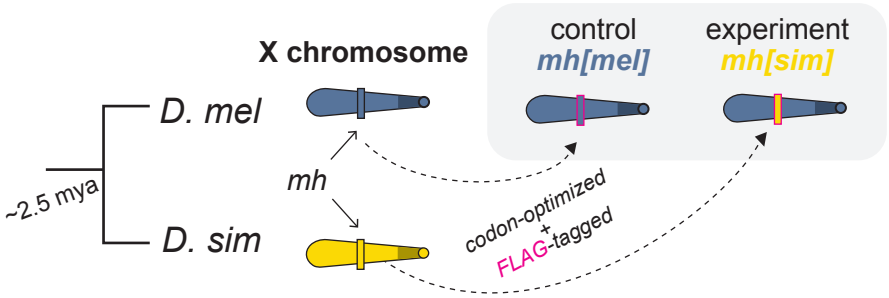
33. Sawamura, K., Yamamoto, M.T., and Watanabe, T.K. (1993). Hybrid lethal systems in the Drosophila melanogaster species complex. II. The Zygotic hybrid rescue (Zhr) gene of D. melanogaster. Genetics *133*, 307-313.
34. Morocz, M., Zsigmond, E., Toth, R., Enyedi, M.Z., Pinter, L., and Haracska, L. (2017). DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA. Nucleic Acids Res *45*, 3172-3188.
35. Stingele, J., Bellelli, R., Alte, F., Hewitt, G., Sarek, G., Maslen, S.L., Tsutakawa, S.E., Borg, A., Kjaer, S., Tainer, J.A., et al. (2016). Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. Mol Cell *64*, 688-703.
36. Vaz, B., Popovic, M., Newman, J.A., Fielden, J., Aitkenhead, H., Halder, S., Singh, A.N., Vendrell, I., Fischer, R., Torrecilla, I., et al. (2016). Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. Mol Cell *64*, 704-719.
37. Bhargava, V., Goldstein, C.D., Russell, L., Xu, L., Ahmed, M., Li, W., Casey, A., Servage, K., Kollipara, R., Picciarelli, Z., et al. (2020). GCNA Preserves Genome Integrity and Fertility Across Species. Dev Cell *52*, 38-52 e10.
38. Chmátal, L., Gabriel, S.I., Mitsainas, G.P., Martínez-Vargas, J., Ventura, J., Searle, J.B., Schultz, R.M., and Lampson, M.A. (2014). Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. Current biology: CB *24*, 2295-2300.
39. Iwata-Otsubo, A., Dawicki-McKenna, J.M., Akera, T., Falk, S.J., Chmátal, L., Yang, K., Sullivan, B.A., Schultz, R.M., Lampson, M.A., and Black, B.E. (2017). Expanded Satellite Repeats Amplify a Discrete CENP-A Nucleosome Assembly Site on Chromosomes that Drive in Female Meiosis. Current biology: CB *27*, 2365-2373.e2368.
40. Carracedo, M.C., Asenjo, A., and Casares, P. (2000). Location of Shfr, a new gene that rescues hybrid female viability in crosses between Drosophila simulans females and D. melanogaster males. Heredity (Edinb) *84 (Pt 6)*, 630-638.
41. Gerard, P.R., and Presgraves, D.C. (2012). Abundant genetic variability in Drosophila simulans for hybrid female lethality in interspecific crosses to Drosophila melanogaster. Genet Res (Camb) *94*, 1-7.
42. Orr, H.A. (1996). The unexpected recovery of hybrids in a Drosophila species cross: a genetic analysis. Genet Res *67*, 11-18.
43. Sawamura, K., Taira, T., and Watanabe, T.K. (1993). Hybrid lethal systems in the Drosophila melanogaster species complex. I. The maternal hybrid rescue (mhr) gene of Drosophila simulans. Genetics *133*, 299-305.
44. Sawamura, K., and Yamamoto, M.T. (1993). Cytogenetical localization of Zygotic hybrid rescue (Zhr), a Drosophila melanogaster gene that rescues interspecific hybrids from embryonic lethality. Mol Gen Genet *239*, 441-449.
45. Pool, J.E., Corbett-Detig, R.B., Sugino, R.P., Stevens, K.A., Cardeno, C.M., Crepeau, M.W., Duchen, P., Emerson, J.J., Saelao, P., Begun, D.J., et al. (2012). Population Genomics of sub-saharan Drosophila melanogaster: African diversity and non-African admixture. PLoS Genet *8*, e1003080.
46. Rogers, R.L., Cridland, J.M., Shao, L., Hu, T.T., Andolfatto, P., and Thornton, K.R. (2014). Landscape of standing variation for tandem duplications in Drosophila yakuba and Drosophila simulans. Mol Biol Evol *31*, 1750-1766.

47. Venken, K.J., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747-1751.
48. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682.
49. McKim, K.S., Joyce, E.F., and Jang, J.K. (2009). Cytological analysis of meiosis in fixed Drosophila ovaries. Methods Mol Biol *558*, 197-216.
50. Schuh, M., Lehner, C.F., and Heidmann, S. (2007). Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. Curr Biol *17*, 237-243.
51. Levine, M.T., Vander Wende, H.M., and Malik, H.S. (2015). Mitotic fidelity requires transgenerational action of a testis-restricted HP1. Elife *4*, e07378.

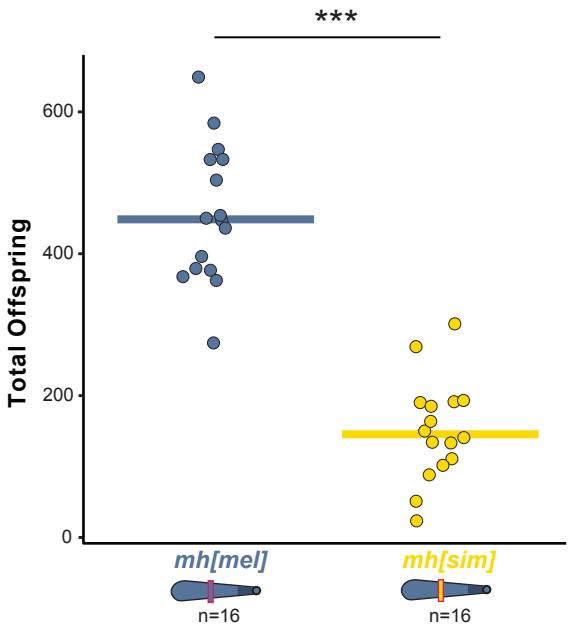
A

	Poly-morphism	Fixation
Syn	**21**	**65**
Non-syn	**12**	**83**

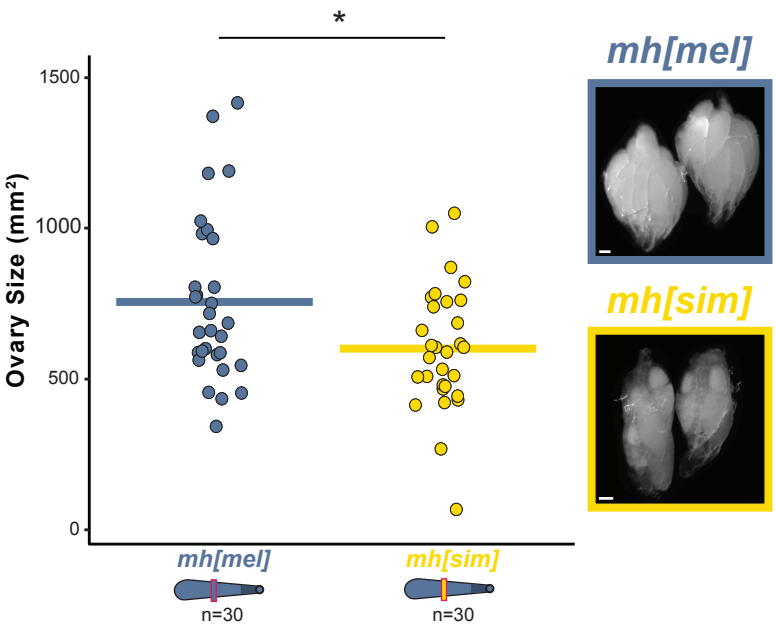
B



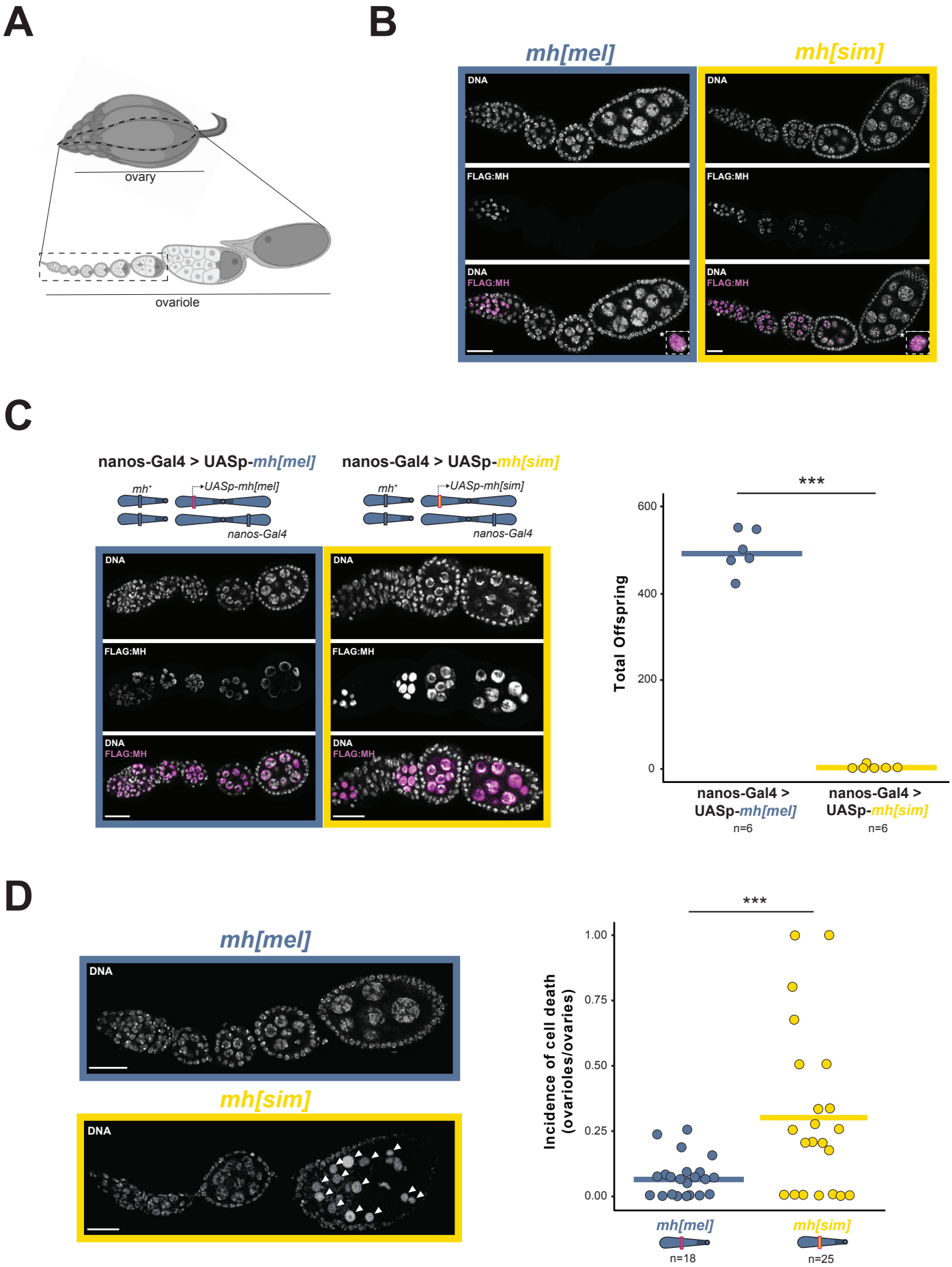
C



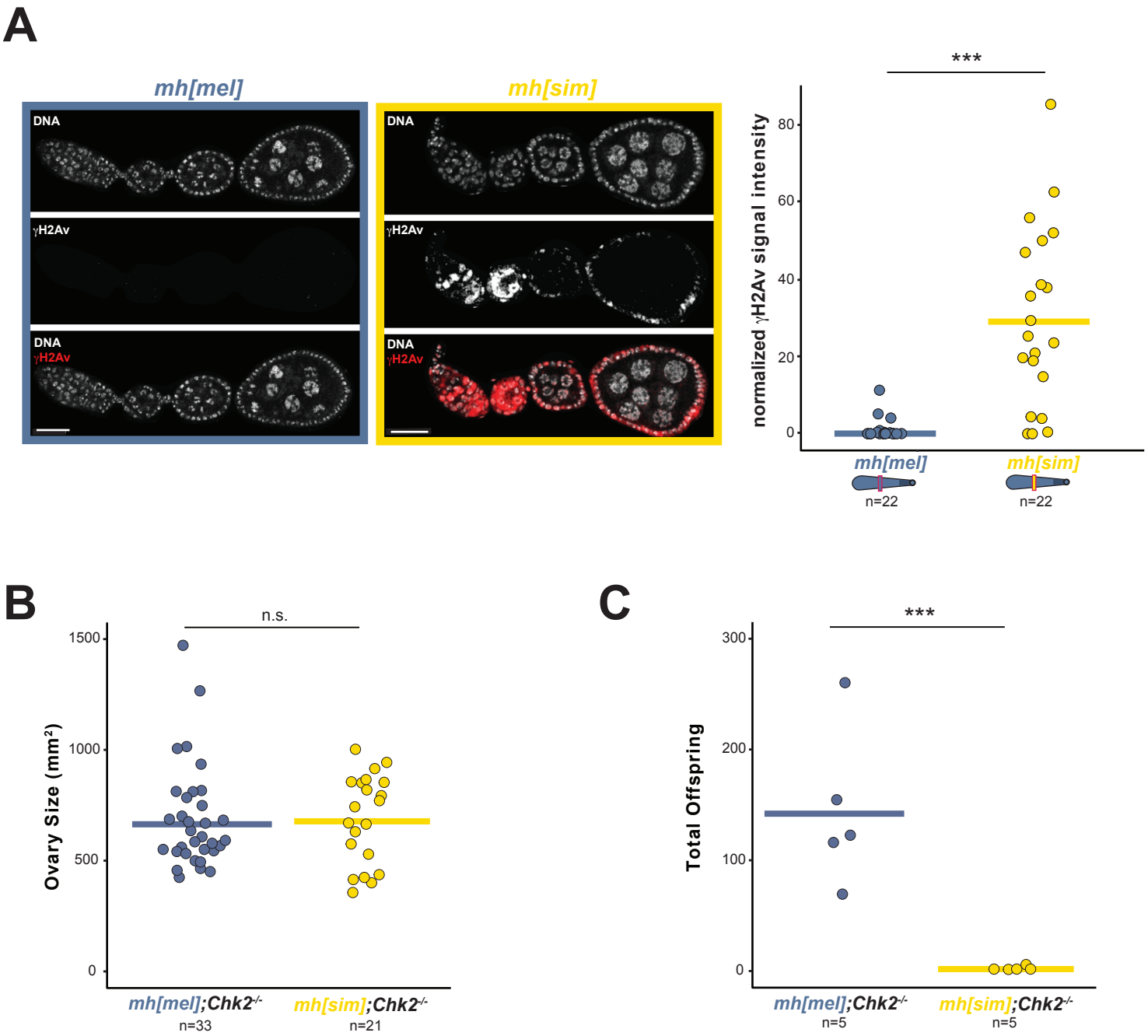
D



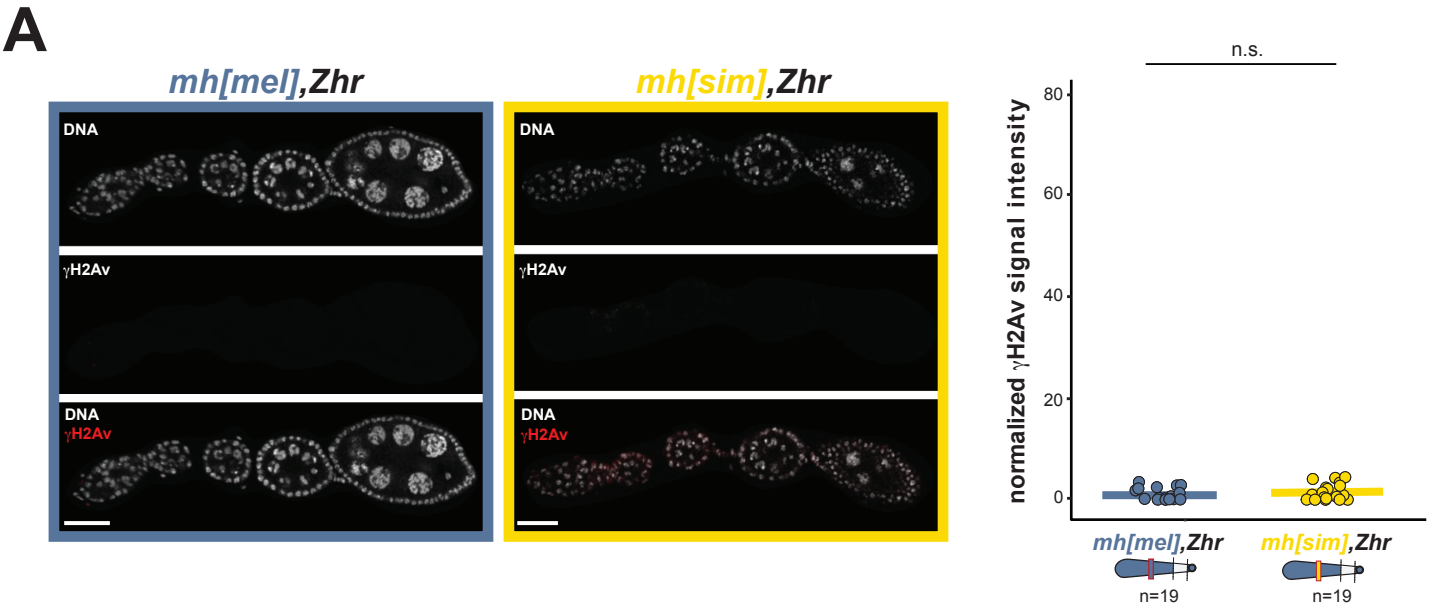
Brand and Levine Figure 1

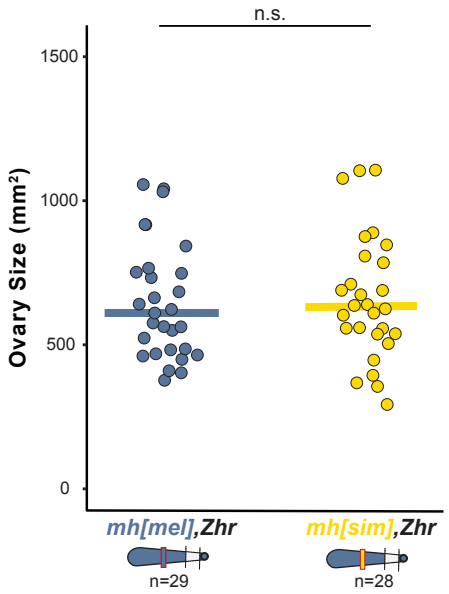


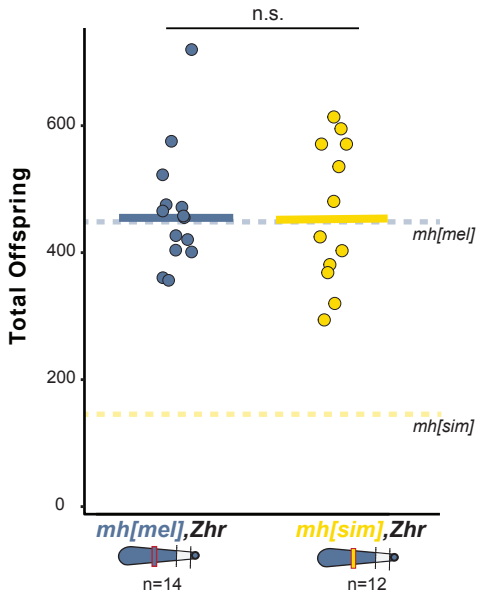
Brand and Levine Figure 2

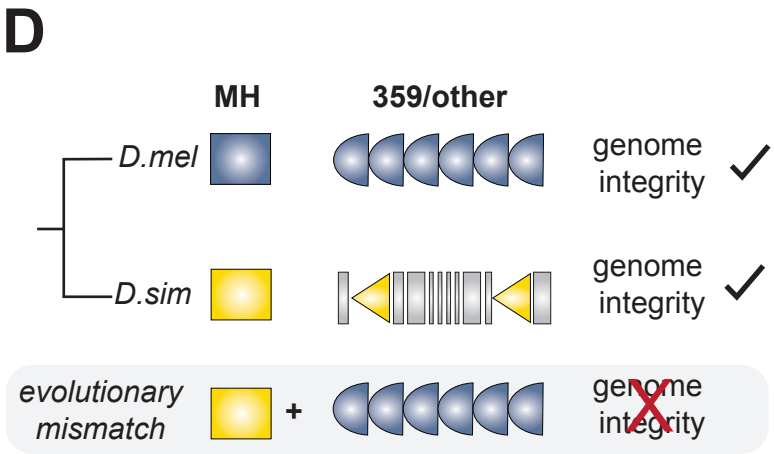


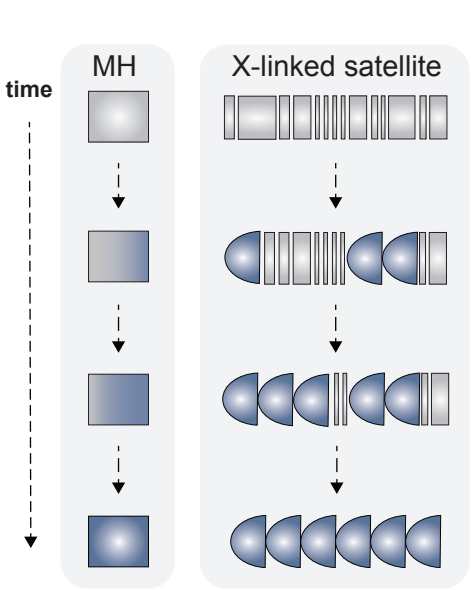
Brand and Levine Figure 3



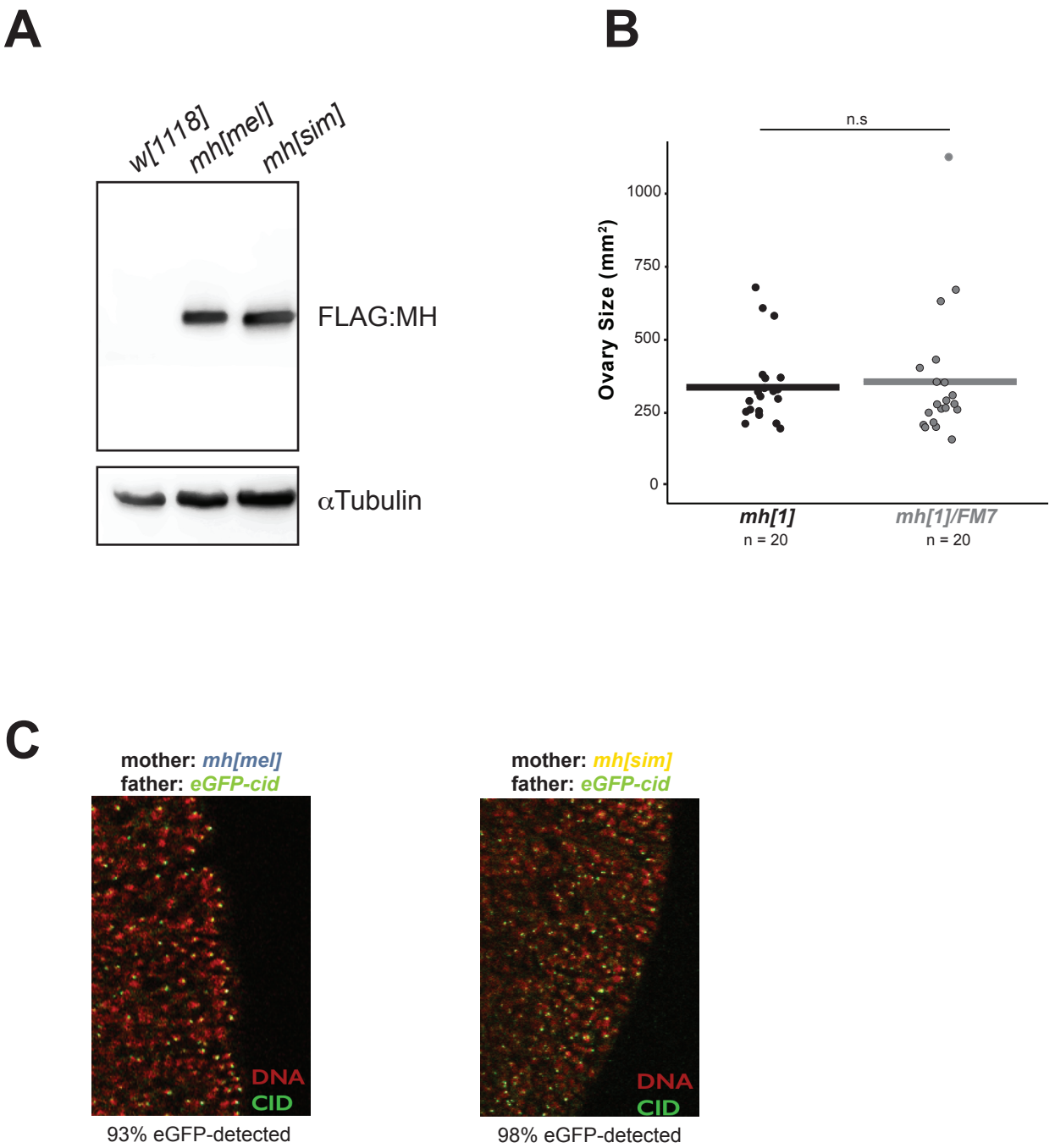




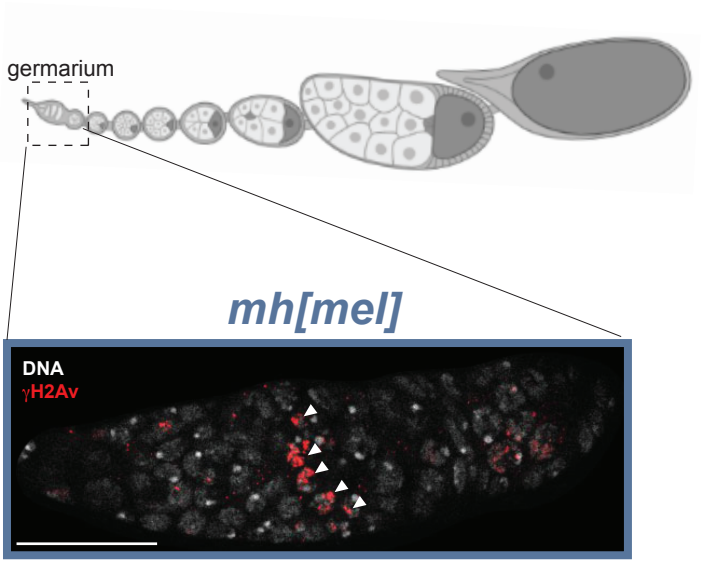




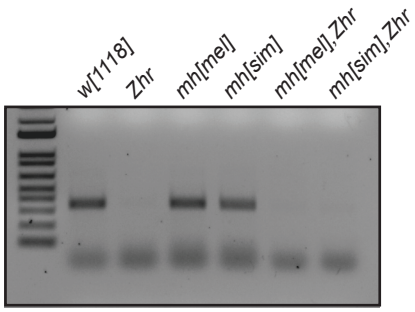
Brand and Levine Figure 4

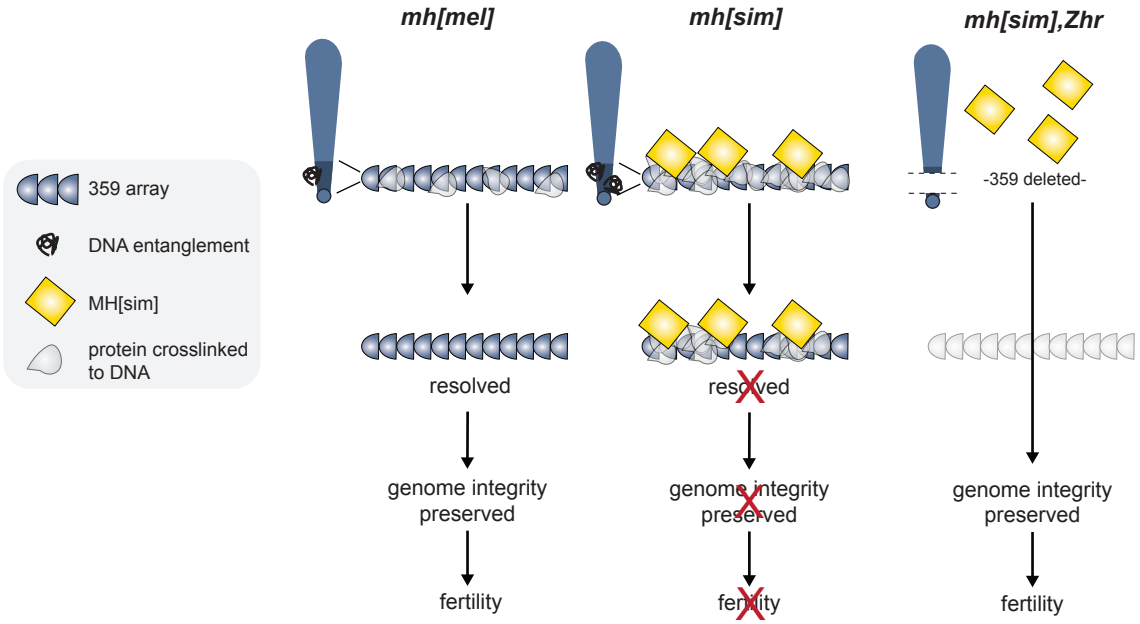


Brand and Levine Figure S1



Brand and Levine Figure S2





Brand and Levine Figure S3