nature genetics

# Histone H3K9 methylation is dispensable for *Caenorhabditis elegans* development but suppresses RNA:DNA hybrid-associated repeat instability

Peter Zeller[1,2,4], Jan Padeken[1,4], Robin van Schendel[3], Veronique Kalck[1], Marcel Tijsterman[3] & Susan M Gasser[1,2]

**Histone H3 lysine 9 (H3K9) methylation is a conserved modification that generally represses transcription. In *Caenorhabditis elegans* it is enriched on silent tissue-specific genes and repetitive elements. In *met-2 set-25* double mutants, which lack all H3K9 methylation (H3K9me), embryos differentiate normally, although mutant adults are sterile owing to extensive DNA-damage-driven apoptosis in the germ line. Transposons and simple repeats are derepressed in both germline and somatic tissues. This unprogrammed transcription correlates with increased rates of repeat-specific insertions and deletions, copy number variation, R loops and enhanced sensitivity to replication stress. We propose that H3K9me2 or H3K9me3 stabilizes and protects repeat-rich genomes by suppressing transcription-induced replication stress.**



H3K9 is a common target of methylation *in vivo* and can carry one, two or three methyl groups. H3K9me2 or H3K9me3 mark transcriptionally silent heterochromatin in most eukaryotes[1–3]. In mammals, insects and *Schizosaccharomyces pombe,* H3K9 methylation is highly enriched at telomeres, pericentric heterochromatin and interspersed repetitive elements (REs)[4–7].

Ligands that recognize methylated H3K9, such as heterochromatin protein 1 (HP1), mediate transcriptional repression of reporter genes and chromatin compaction near centromeres[2,8]. H3K9me is also implicated in the silencing of genes both during development[9,10] and in pathological states. For instance, tumor-suppressor genes have been found to be transcriptionally silenced by mistargeted H3K9me in cancers[11,12], and H3K9me marks triplet repeat sequences, whose expansion has debilitating consequences in syndromes such as Huntington's or Fragile X[13,14]. Nonetheless, by reducing levels of H3K9me the efficiency of somatic cell reprogramming can be increased[15,16].

It has been difficult to study the function of H3K9me-mediated repression in complex organisms for several reasons. First, there are at least eight documented and partially redundant H3K9 histone methyltransferases (HMTs) in mammals (SUV39h1, SUV39h2, G9a, SETDB1, SETDB2, PRDM2, PRDM3 and PRDM16 in mice). Second, the vast majority of H3K9 methylation is found on extended stretches of REs that cannot be accurately mapped by standard deep sequencing techniques[17]. In some cases the disruption of individual H3K9me HMTs is embryonically lethal, owing in part to compromised mitotic chromosome segregation[18–20]. The loss of SUV39h1, SUv39h2 or their homologs also results in mitotic defects, aneuploidy and chromosomal rearrangements in mice, flies and fission yeast[7,21,22]. This may have masked phenotypes arising from the loss of H3K9me in transcriptional repression during development.

The holocentric nematode *C. elegans* has only two, nonredundant H3K9me-depositing HMTs, MET-2 and SET-25 (refs. 23,24). Here we exploited the finding that mutants lacking both HMTs have no detectable H3K9 methylation[24], and yet produce viable embryos, to study how the loss of this histone modification impacts a multicellular organism.

## RESULTS

### Loss of H3K9me did not impair embryonic differentiation into adult tissues

The HMT MET-2, which catalyzes the mono- and di-methylation of H3K9, is the homolog of mammalian SETDB1, also known as ESET[23]. SET-25, on the other hand, shares considerable SET domain homology with SUV39h1, SUV39h2 and G9a enzymes, and it is the only *C. elegans* enzyme that trimethylates H3K9 (ref. 24). To confirm that *met-2 set-25* double mutant worms lack H3K9 methylation throughout development, we performed immunofluorescence analysis at all stages of worm development (**Fig. 1a**). We found no detectable H3K9me2 or me3 in *met-2 set-25* embryos, second-stage larvae (L2) or gonads of adult worms, confirming our earlier mass spectroscopic analysis of total histones isolated from mutant embryos or larvae[24]. Histone acetylation and other common methylation marks (**Supplementary Fig. 1**) remained intact[24]. Despite this complete absence of H3K9me, the *met-2 set-25* mutant embryos developed into viable adults.

[1]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. [2]Faculty of Natural Sciences, University of Basel, Basel, Switzerland. [3]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands. [4]These authors contributed equally to this work. Correspondence should be addressed to S.M.G. (susan.gasser@fmi.ch).

Received 28 June; accepted 22 August; published online 26 September 2016; doi:10.1038/ng.3672

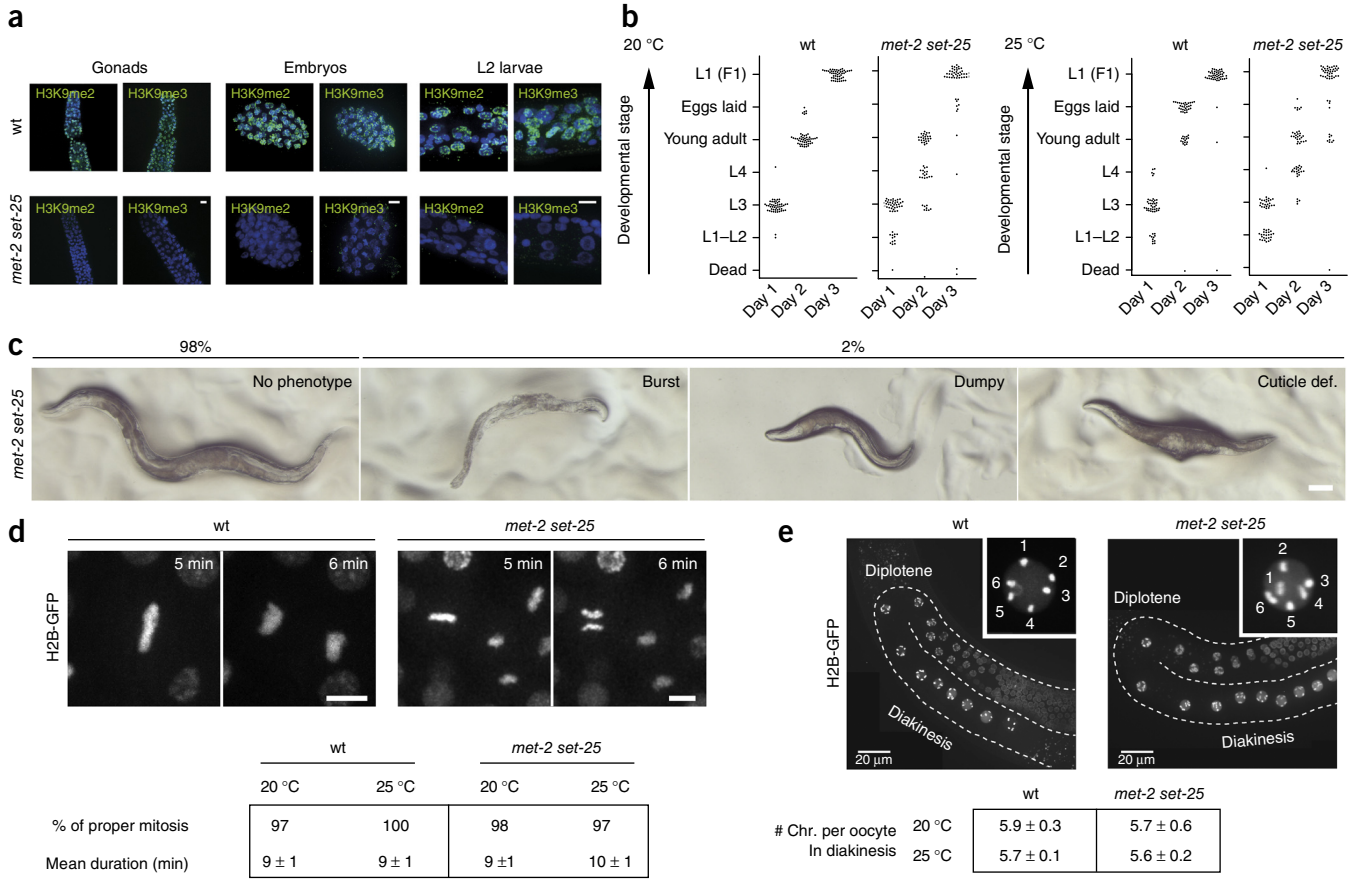


**Figure 1** Worms lacking H3K9me were viable but showed stochastically delayed development. (**a**) Immunofluorescence images using H3K9me2- and H3K9me3-specific antibodies on wild-type (wt) and *met-2 set-25* strains at indicated developmental stages. H3K9me2 and H3K9me3 signals are in green, and DAPI in blue. Scale bars, 5 µm. (**b**) *met-2 set-25* mutation provoked stochastic delays in development from L1 larval stage into fertile adults. Developmental progress of singled mutant and wt L1 larvae monitored every 24 h for 3 d at 20 °C and 25 °C (*N* (number of biological replicates) = 3, *n* (number of animals per replica) = 50). (**c**) Example images of worm morphologies arising in *met-2 set-25* cultures and their frequencies (*N* = 4, *n* = 50). Scale bar, 100 µm. (**d**,**e**) H3K9me2/H3K9me3 was not essential for chromosome segregation in *C. elegans*. Images (**d**) from time-lapse (Δ*t* = 1 min) movies of mitotic cells in embryos expressing H2B-GFP in which mitotic defects were scored (wt, 20 °C *n*= 34; wt, 25 °C *n* = 50; *met-2 set-25*, 20 °C *n* = 45; and *met-2 set-25*, 25 °C *n* = 36). Scale bars, 3 µm. Duration of mitosis reflects minutes from the beginning of chromosome condensation until completion of telophase. The number of bivalent chromosomes (**e**) in wt and *met-2 set-25* worms expressing H2B-GFP counted in oocytes undergoing diakinesis (*N* = 3; wt, 20 °C *n* = 57; *met-2 set-25*, 20 °C *n* = 51; wt, 25 °C *n* = 50; and *met-2 set-25*, 25 °C *n* = 50). Mean and s.d. are shown. Insets, nucleus of an oocyte in diakinesis.

To monitor the kinetics of somatic development, we compared the timing of wild-type (N2) and *met-2 set-25* organisms as they transitioned from the first larvae stage (L1) to the L1 stage of the next generation. This is a highly synchronous cycle that takes 3 d in wild-type strains grown at 20 °C (**Fig. 1b**). In contrast to wild-type worms, 52% of the *met-2 set-25* mutants showed stochastic delays in stage transitions, even though most mutant embryos reached adulthood (88% became mature adults; **Fig. 1b**). These delays were more pronounced at 25 °C than at 20 °C and were not restricted to one specific stage (**Fig. 1b**). Nonetheless, only 2% of the adult offspring displayed a grossly irregular morphology at 20 °C, i.e., 'dumpy' appearance, partially defective cuticles or bursting as adults (**Fig. 1c**). Such aberrant morphologies were below detection level (<0.1%) in wild-type populations[25].

Chromosome missegregation has been suggested to be a main cause for the phenotypes observed in mutants for H3K9 HMTs in other organisms[7,19,26]. Using histone H2B fusion to GFP (H2B-GFP), we tracked the frequency of mitotic chromosome bridges or lagging chromosomes in wild-type and *met-2 set-25* embryos. The frequency of defective mitoses at either 20 °C or 25 °C was similar in wild-type and mutant embryos (**Fig. 1d**). Moreover, the duration of mitosis was identical, which argues against any mutant-specific spindle checkpoint activation (**Fig. 1d**). To monitor meiotic chromosome missegregation, we followed H2B-GFP-tagged oocytes undergoing meiosis in gonads. Thanks to the chromosome condensation and enlarged nuclei that occur in diakinesis, we could determine bivalent chromosome number per cell. Again there was no detectable difference between *met-2 set-25* and wild-type oocytes at either temperature (**Fig. 1e**). Thus, we excluded aneuploidy and spindle checkpoint activation as triggers for the developmental delay or aberrant morphologies of H3K9me-deficient worms.

### Temperature-dependent sterility of *met-2 set-25* mutant

Brood sizes were notably smaller upon propagation of the double HMT mutant, and worms became completely sterile after two generations at 26 °C (**Supplementary Fig. 2a**). We determined the number of viable progeny of *met-2 set-25* vs. wild-type worms under controlled growth conditions at 15 °C, 20 °C and 25 °C. Although brood size was equal between the *met-2 set-25* and wild-type worms at 15 °C, mutant adults had significantly fewer viable progeny at both 20 °C and

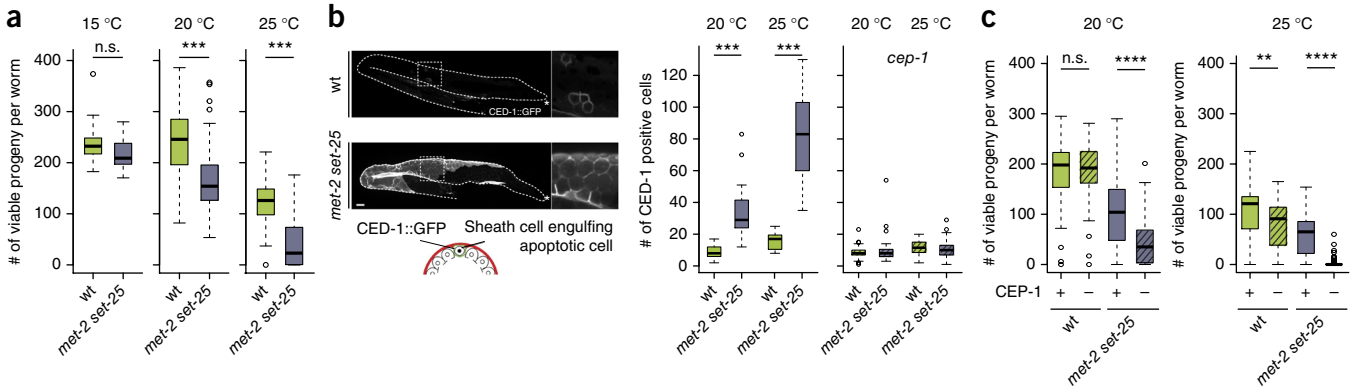


**Figure 2** DNA-damage-checkpoint-dependent increase of apoptotic cells in the germ line of *met-2 set-25* worms. (**a**) Number of viable progeny of wt and *met-2 set-25* mutant per worm at 15 °C, 20 °C and 25 °C ($N = 3$, $n = 75$). (**b**) Example image of a gonad and the quantification of the number of apoptotic cells in worms expressing the apoptosis marker CED-1::GFP in wt and *met-2 set-25* background, with and without CEP-1. Apoptosis rate was determined as the number of cells fully engulfed by CED-1::GFP per gonad arm. CED-1 is a phagocytic receptor, which translocates to the plasma membrane during apoptosis. Asterisks indicate gonad tip and boxes mark enlarged section in the overview image ($N = 3$, $n = 75$). Scale bar, 10 µm. (**c**) Number of viable progeny per worm of wt and *met-2 set-25* with or without CEP-1. At both 20 °C and 25 °C ($N = 3$, $n = 75$) *cep-1* and *met-2 set-25* showed a synthetic loss of viable progeny. Boxplots show median, boxes 50% and whiskers 90% of the group. Two-sided Wilcoxon signed-rank test: n.s. indicates not significant, $^{**}P < 0.005$, $^{***}P < 0.0001$ and $^{****}P < 0.00005$.



25 °C (**Fig. 2a**). A similar temperature-dependent loss of fertility has been observed for mutants of the PIWI pathway[27,28] (**Supplementary Fig. 2b**), a germline-specific small RNA pathway that helps to silence transposable elements[29].

Gonad development *per se* was not impaired in the *met-2 set-25* mutant (**Supplementary Fig. 2c**). However, by scoring the expression of the CED-1::GFP phagocytic receptor, which accumulates on the plasma membrane of apoptotic cells[30], we detected a high level of germline apoptosis (**Fig. 2b**). The level increased when we grew worms at 25 °C. In the double mutant an average of 30 cells per germ line were positive for CED-1 at 20 °C (wild-type: 10 cells), and over 80 cells per germ line at 25 °C (wild-type: 18 cells; **Fig. 2b**). Consistently, RNA sequencing (RNA-seq) of *met-2 set-25* gonads showed an increase in mRNA from various other apoptosis-specific genes[31] (**Supplementary Fig. 2d**).

Although *C. elegans* germline cells are known to be particularly sensitive to DNA damage, germline apoptosis can have multiple causes[32]. To see whether apoptosis in H3K9me-deficient gonads is caused by DNA damage, we deleted the mammalian p53 homolog, CEP-1, and scored CED-1::GFP distribution at 20 °C and 25 °C (ref. 33). In the *met-2 set-25 cep-1* triple mutant and in the strain lacking *cep-1* alone, we detected only background levels of germline apoptosis at both temperatures (**Fig. 2b**). This strongly suggests that the germline apoptosis seen in the absence of H3K9me stemmed from DNA damage. The *met-2 set-25 cep-1* triple mutant was synthetic sterile, as expected (**Fig. 2c**). Of embryos laid at 20 °C, hatching rate dropped from above 95% in the *met-2 set-25* mutant to below 80% when coupled with *cep-1* (**Supplementary Fig. 2e**). This is likely due to an increase in DNA damage in the mutant, because the number of RAD-51 foci per cell, a marker of processed breaks, increased significantly ($P < 0.001$, two-sided Wilcoxon signed rank test), as did the number of cells in the mitotic zone of the germ line with RAD-51 foci (3.4% in the wild type and 14.6% in the double mutant; **Supplementary Fig. 2f**). This suggests that germline cells incur enhanced levels of damage in the absence of H3K9me.

### H3K9me2 marks REs, whereas H3K9me3 marks REs and silent genes

To understand the link between the loss of H3K9 methylation and the observed increase in DNA damage, we first reexamined the sequences reported to be bound by histones bearing H3K9me2 and H3K9me3. We performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) experiments, not unlike those reported by the modENCODE consortium[34,35]. We found a tenfold enrichment of both H3K9me2 and H3K9me3 along the distal arms of the five worm autosomes in early embryos (**Supplementary Fig. 3a**). We did not observe this distribution for other repressive marks, such as H3K27me3, nor for the active mark, H3K4me3. Chromosome arms were similarly enriched for all types of REs (**Supplementary Fig. 3a**[36]). A detailed analysis of the distribution of H3K9me2 versus H3K9me3 in embryos showed that a high proportion of H3K9me2 was on REs (~34% of all H3K9me2), whereas H3K9me3 was present equally on exons and REs (~26% each, **Fig. 3a**).

Distinct classes of repetitive DNA constitute large fractions of the genomes of complex organisms. These include DNA or RNA transposons, which can generate copies of themselves and integrate into the genome, as well as simple repeats, such as tandemly arranged micro- or minisatellites (**Fig. 3b**). Unlike transposons, these latter repeats lack open reading frames (ORFs) and regulatory sequences. Worm genomes contain all classes of REs, although DNA (rather than RNA) transposons are the most abundant transposable elements[37]. Short repetitive sequences are not found as megabase blocks of pericentric satellite sequence in worms, but as short clusters distributed along the chromosome. As a consequence, 87% of the *C. elegans* REs, or roughly ~60,000 discrete elements, can be uniquely mapped to individual sites of the genome by standard next-generation sequencing.

Plotting the enrichment of H3K9me2 and H3K9me3 on all REs in embryos, we found that 24.3% of REs were exclusively enriched for H3K9me2, and 18.1% had either both marks or exclusively H3K9me3 (**Fig. 3c**). This revealed that 42.4% of mappable REs were enriched for H3K9me, with H3K9me2 and H3K9me3 distributed differentially over the three repeat classes. RNA transposons were most strongly correlated with H3K9me3 (58.5%, with 5.7% bearing H3K9me2 only); tandem or simple repeats were more likely to carry H3K9me2 alone (31.6%), and DNA transposons fell into two groups: 25.5% were uniquely dimethylated whereas 17.7% carried H3K9me3 (**Fig. 3c**).

In embryos H3K9me3 was enriched on transcriptionally silent genes (12.0%), where it coated entire ORFs of loci (**Fig. 3d** and

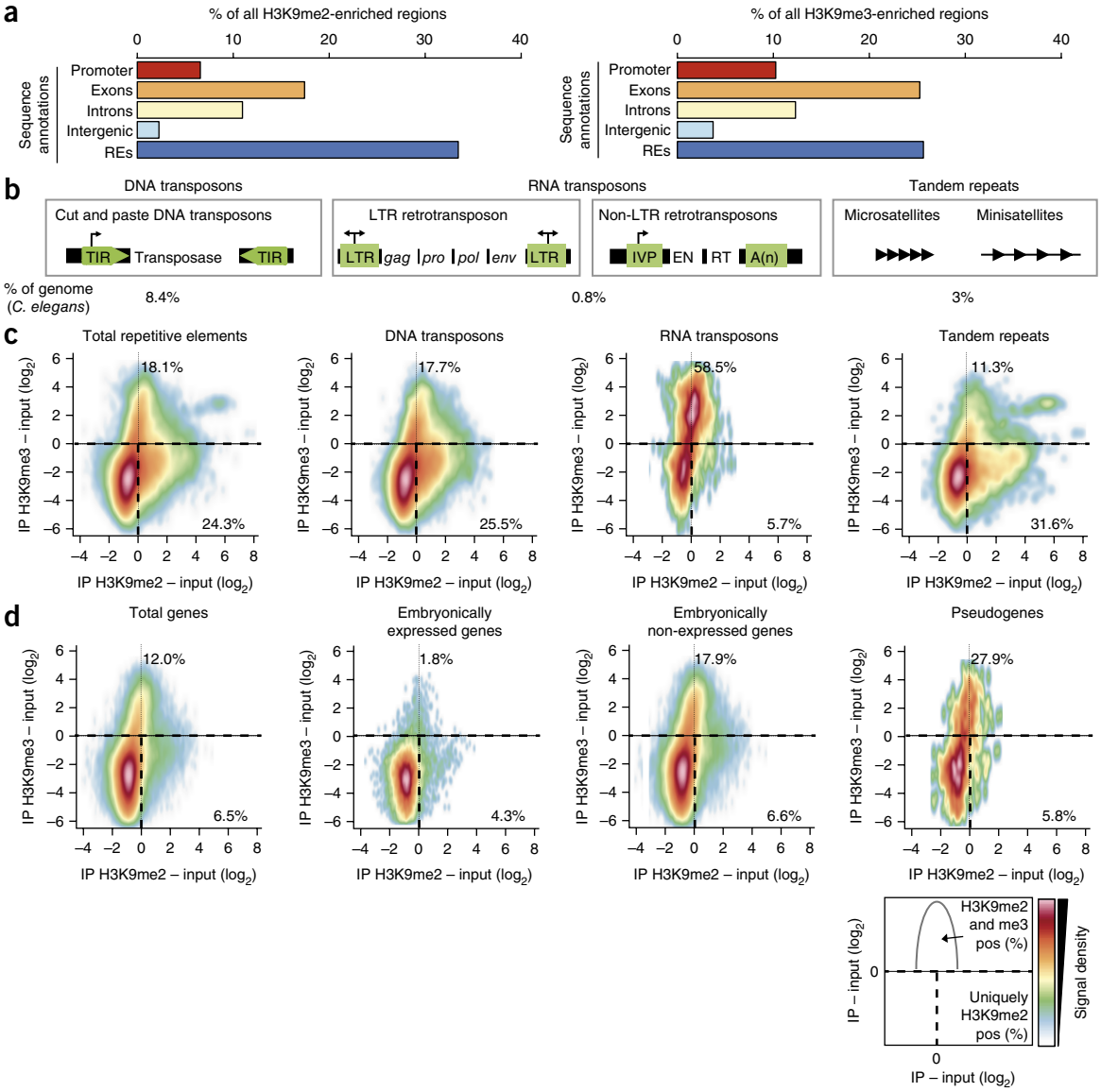


**Figure 3** Differential enrichment of H3K9me2 and H3K9me3 on repeat element classes and gene types. (**a**) Percentage of H3K9me2 and H3K9me3 domains covering promoters, exons, introns, unique intergenic sequences or REs ($N = 2$). H3K9me positive regions were determined from genomic bins of sequences recovered after CHIP-seq using H3K9me2- or H3K9me3-specific antibodies with IP/input > 0. (**b**) Schematic representation of the three major repeat classes. DNA transposons encode a single transposase, which catalyzes all the steps of transposition, flanked by two terminal inverted repeats (TIRs). RNA transposons are either long terminal repeat (LTR) or non-LTR retrotransposon types. As derivatives of ancient retrovirus infections LTR retrotransposons encode *gag* (structural proteins of the virus core), *pol* (reverse transcriptase, integrase), *pro* (protease) and *env* (envelope). Non-LTR transposons encode a reverse transcriptase (RT) and an endonuclease (EN). Retrotransposon flanking regions in both cases supply promoter elements. Tandem repeats are short, noncoding sequence stretches that are repeated in a head-to-tail fashion. (**c**) High-density scatterplots show the enrichment of H3K9me2 and H3K9me3 on REs based on CHIP-seq data. IP, immunoprecipitation. RNA transposons were heavily enriched for H3K9me3 (58.5%), whereas 31.6% of tandem repeats had only H3K9me2. Lines indicate the quadrants of single-positive, double-positive and double-negative elements. (**d**) High-density scatterplots of the H3K9me2 and H3K9me3 enrichment on genes. Nonexpressed genes and pseudogenes were enriched for H3K9me3.

**Supplementary Fig. 3b**), and it was depleted from active genes (1.8%; **Fig. 3d**). Among the H3K9me3-bound genes were many that were expressed only in terminally differentiated tissues and a large fraction of pseudogenes (**Fig. 3d** and **Supplementary Fig. 3c–e**).

### Loss of H3K9me led to the derepression of genes and REs

To determine whether loss of H3K9me affects transcription, we performed RNA-seq on RNA isolated from either gonads or early embryos of wild-type and *met-2 set-25* strains, grown at either 20 °C or 25 °C. In embryos cultured at 20 °C, we observed the reproducible derepression (>2-fold compared to wild-type) of 308 genes. Of these 72.2% (234) were marked by H3K9me in wild-type cells, and are therefore likely to be regulated directly by MET-2 and/or SET-25 (**Fig. 4a** and **Supplementary Fig. 4a**). This set of derepressed genes was only a subset (~9.7%) of all genes bearing H3K9me, arguing that the loss of H3K9me is not always sufficient to activate transcription. Derepression of genes was also temperature-sensitive, with 2.2-fold more genes being upregulated at 25 °C, including 83.8% of those

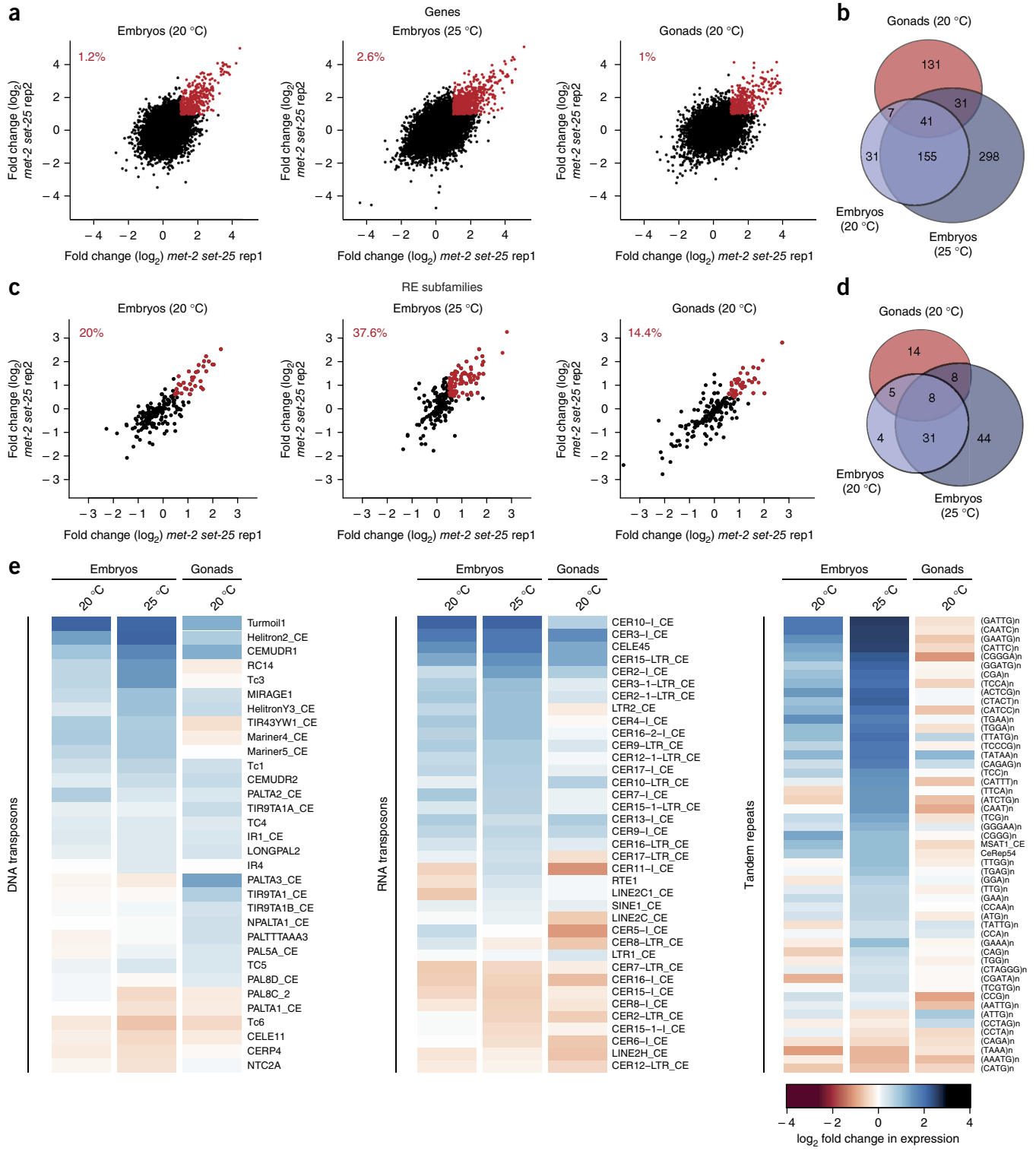


**Figure 4** Temperature-dependent derepression of subsets of genes and repeat families in embryos and gonads in *met-2 set-25* worms. (**a**) Fold change ($\log_2$, *met-2 set-25*/wt) in gene expression of two replicas (rep1 and rep2) of RNA-seq data from embryos (20 °C and 25 °C) and from isolated gonads (20 °C) for each strain. The genes marked in red were consistently >2-fold upregulated ($P < 0.05$, FDR < 0.1), and % of total genes is indicated. (**b**) Venn diagram of the derepressed genes shows that genes affected in gonads were distinct from those upregulated in early embryos, and that derepression was temperature-enhanced in embryos ($N = 3$). (**c**) Scatterplot of the expression changes of H3K9me-enriched RE subfamilies in *met-2 set-25* embryos and gonads compared to wt. The REs marked in red were >1.5-fold derepressed in both replicas ($P < 0.05$, FDR < 0.1). (**d**) Venn diagram of the derepressed REs shows that subfamilies affected in gonads were partially distinct from subfamilies derepressed in embryos, and that RE upregulation was temperature-enhanced in embryos ($N = 3$). (**e**) Heat map of the expression changes in all significantly affected RE subfamilies ($P < 0.05$, FDR < 0.05) in embryos (20 °C and 25 °C) and gonads, sorted by repeat class ($N = 3$).

already derepressed at 20 °C (**Fig. 4a**,**b**). Transcription in gonads was elevated by the loss of H3K9me (210 genes). The affected genes were largely distinct from those derepressed in somatic cells (37.6% overlap; **Fig. 4a**,**b**), arguing that transcription factor availability is critical for transcriptional activity in the absence of repressive chromatin. No essential regulators of meiosis were misregulated.

Given that REs were enriched for H3K9me in wild-type worms, we next examined expression changes for REs, which we analyzed as

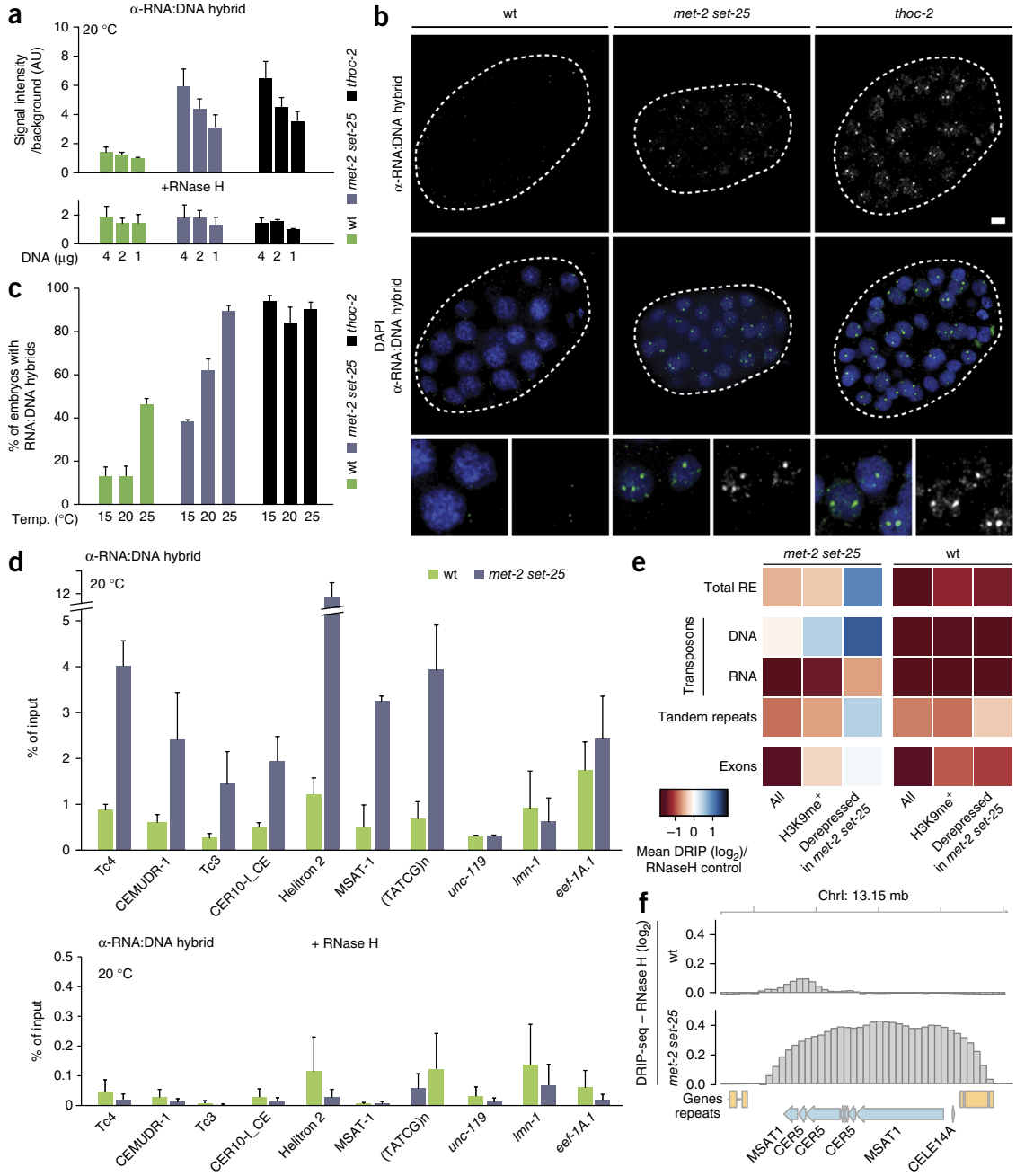


**Figure 5** *met-2 set-25* worms accumulate RNA:DNA hybrids at repeat elements. (**a**) Quantification of multiple dot blots against RNA:DNA hybrids (antibody S9.6, HB-8730, ATCC, $n = 3$) in genomic DNA isolated from gravid adults of wt, *met-2 set-25* and *thoc-2* strains grown at 20 °C. 4 µg, 2 µg and 1 µg of nucleic acids were loaded for each strain. Where indicated, genomic DNA was treated with RNase H before blotting (mean + s.e.m.; $N = 3$). (**b**) Immunofluorescence (IF) images of isogenic wt, *met-2 set-25 and thoc-2* mutant embryos grown at 20 °C, stained with antibody S9.6 to visualize RNA:DNA hybrids (green); DAPI is in blue. Scale bar, 5 µm. (**c**) Quantification of IF signals after S9.6 staining from embryos of indicated strains, grown at 15 °C, 20 °C or 25 °C (mean + s.e.m.; $N = 3$, $n = 15$). (**d**,**e**) Genome-wide distribution of R loops determined by DRIP with antibody S9.6, followed by qPCR or deep sequencing of recovered DNA. DRIP-qPCR (**d**) for seven repeat subfamilies upregulated in *met-2 set-25* worms and for three control loci (*unc-119, lmn-1* and *eef-1A.1*), that were not upregulated (mean and s.e.m.; $N = 2$). Heat map of an S9.6 DRIP-seq experiment (**e**) showing mean $\log_2$ enrichment over the corresponding controls treated with RNase H (samples normalized to total number of reads). Loci were segregated based on indicated sequence criteria, and were further subgrouped based on the presence of H3K9me and response to the *met-2 set-25* mutations ($N = 1$). (**f**) DRIP-seq example showing the R-loop signal over a RE cluster. The IP signal was normalized to the input and the RNase H control values were subtracted.

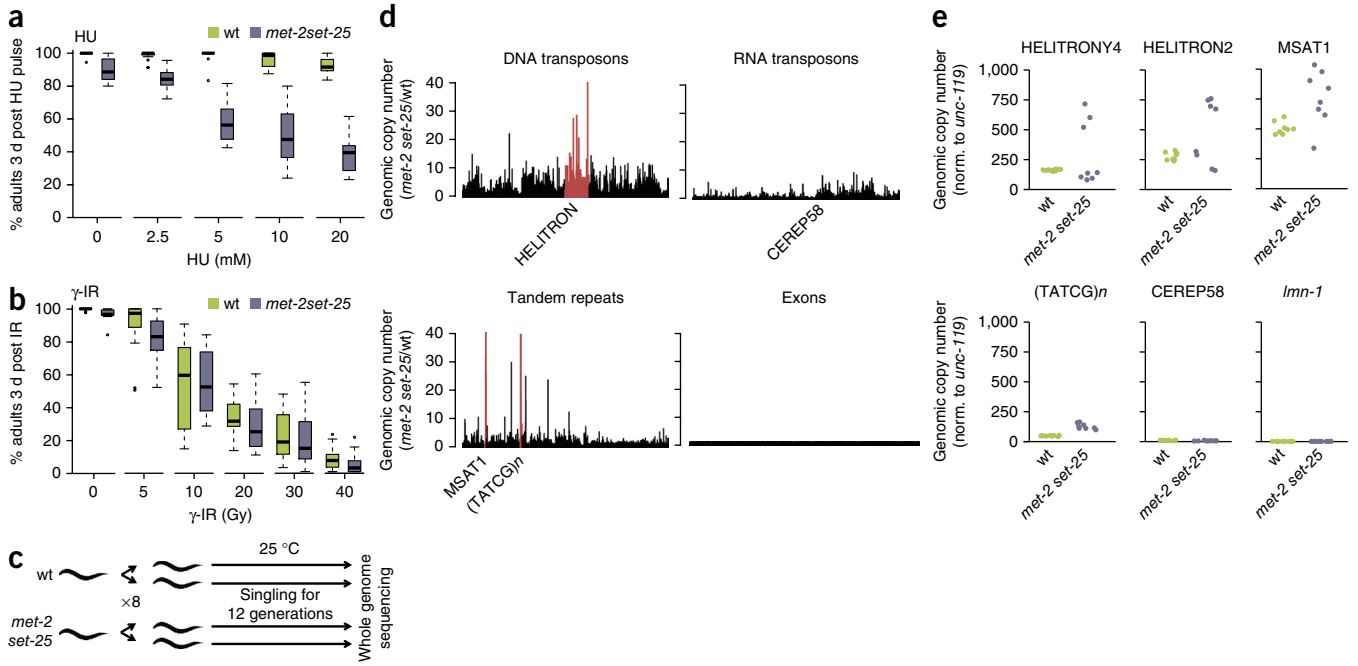


**Figure 6** The *met-2 set-25* strain is hydroxyurea sensitive and accumulates mutations in repeat elements and reactivated transposable elements. (**a**) Worms lacking MET-2 and SET-25 were hypersensitive to hydroxyurea (HU). Synchronized populations of wt and mutant L1 larvae were exposed to 20 mM HU for 16 h, and then the numbers of worms that develop into adults after 3 d were quantified. Statistical analysis shows significant differential loss of viability between *met-2 set-25* and wt worms (Wilcoxon test $P < 0.0001$ at all doses; $N = 6$, $n = 25$). (**b**) Synchronized populations of wt and mutant L1 larvae were exposed to sublethal doses of gamma irradiation (γ-IR), and the numbers of worms that developed into adults after 3 d were quantified. Only at 0 Gy was there enhanced *met-2 set-25* lethality (Wilcoxon test $P < 0.005$; $N = 3$, $n = 15$). Boxplots show median, 50% boxes and 90% whiskers. (**c**) Wt and *met-2 set-25* worms were singled for 12 generations at 25 °C, generating 8 substrains per genotype. Genomic DNA of each substrain was sequenced, and only mutations unique to one of the substrains were considered. (**d**) Copy number of REs and exons in *met-2 set-25* determined by genome sequencing and sorted according to repeat class. REs analyzed by qPCR are indicated in red. (**e**) Analysis of the copy number of REs by qPCR from 8 *met-2 set-25* and 8 wt substrains cultured as described in **c**, shows an increase in CNV of DNA transposons, simple repeats but not RNA transposons or single-copy genes.

subfamilies. We characterized ~84% of all annotated repeats (300 subfamilies), and excluded only very-low-complexity repeat sequences or elements with a single annotated occurrence. In *met-2 set-25* mutant embryos at 20 °C, 20% of the H3K9me-enriched repeat subfamilies were derepressed by at least 1.5-fold, and at 25 °C this value increased to 37.6% (**Fig. 4c** and **Supplementary Fig. 4b**). Gonads isolated from double mutant adults showed an increase of transcription in 14.4% of all H3K9me repeat subfamilies (**Fig. 4c**). This lower number of derepressed REs may reflect germline-specific redundant silencing by the PIWI pathway[38,39]. Indeed, different REs were upregulated in gonads and somatic cells (**Fig. 4d** and **Supplementary Fig. 4c**), with tandem repeats being distinctly underrepresented in the germ line (**Fig. 4e**). We note that each class of repeats includes REs that were not derepressed by loss of H3K9me, which may reflect either the existence of other, H3K9me-independent silencing pathways, or a requirement for transcription factors that are tissue-specific or developmental-stage-specific.

We asked whether the transcriptional landscape of genes surrounding a RE might influence its expression upon loss of H3K9me. This is particularly relevant for simple tandem repeats, which lack recognizable promoter or enhancer sequences[40]. To our surprise, ~50% of the derepressed tandem repeats were not in the proximity of an upregulated gene (data not shown).

### H3K9me-deficient worms accumulated R loops

We next examined the relationship between aberrant RE transcription and the observed DNA damage. Perturbation of the replication fork is a major driver of DNA lesions[41], and a substantial obstacle for its progression is the transcription machinery, in particular when stalled by RNA:DNA hybrids (R loops)[42–45]. In fission yeast, R loops are enriched at repetitive sequences, such as transposons, telomeres or the rDNA[46], and correlated with genetic instability[47,48]. We therefore checked whether the *met-2 set-25* double mutant accumulated R loops, using multiple approaches based on an antibody specific for RNA:DNA hybrids (S9.6, gift of P. Pasero[49]).

We detected an accumulation of R loops in *met-2 set-25* worms that was not detectable in wild-type worm DNA by performing a dot blot analysis of genomic DNA. We also detected significant R-loop occurrence by immunostaining of mutant, but not wild-type, embryos ($P < 0.001$, Student`s t-test; **Fig. 5a–c** and **Supplementary Fig. 5a**). The level was roughly similar to that scored in a mutant strain deficient for the Tho-Trex complex (*thoc-2*), in which RNA:DNA hybrids accumulate owing to impaired RNA processing and export (**Fig. 5a**)[50,51]. To test for antibody specificity, we treated the isolated DNA with RNase H before blotting, to specifically degrade RNA:DNA heteroduplexes. Quantification showed that 60% of the signal (*met-2 set-25*, loading 4 µg; **Fig. 5a**) was lost after treatment with RNase H. Consistent with the elevated level of RE transcription at higher temperatures, the level of R loops increased with temperature, both in the dot blot analysis of adult worm DNA, as well as in the immunostaining of embryos (**Fig. 5c** and **Supplementary Fig. 5a**). The *thoc-2* mutant, on the other hand, reached R-loop saturation even at 15 °C.

To examine formation of R loops in a sequence-dependent manner, we immunoprecipitated RNA:DNA hybrids from wild-type or *met-2*

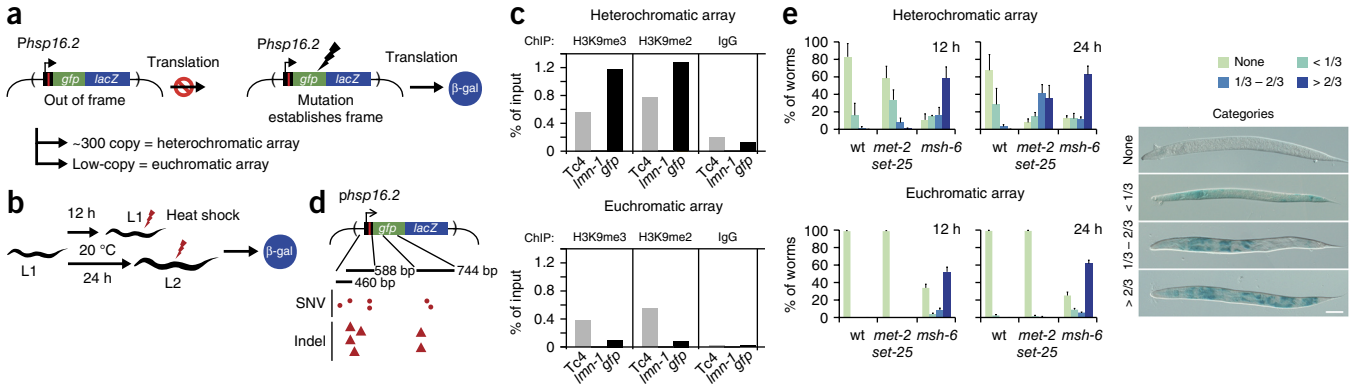


**Figure 7** Somatic accumulation of indels leading to frameshift mutations in *met-2 set-25* mutant larvae. (**a**) Schematic of reporter to monitor mutation frequency in single cells in two different chromatin contexts. A *lacZ* construct containing a frame-shift mutation under the control of a heat shock promoter was integrated either as a high copy (~200–300 copies), or low-copy array (~20 copies). The frameshift prevents *lacZ* translation, which can be reestablished by mutation. (**b**) To quantify the accumulation of mutations, L1 larvae were released into development for 12 h or 24 h at 20 °C before a heat shock and subsequent β-gal staining. (**c**) ChIP-qPCR monitored enrichment of H3K9me2/H3K9me3 on the reporter array by PCR for *gfp*. H3K9me was recovered on the heterochromatic (high-copy) array but not the euchromatic (low-copy) array. The genomic copy of *lmn-1* and *Tc4* served as negative and positive controls. (**d**) Genomic DNA of the heterochromatic array was isolated from either *met-2 set-25*, or wt worms grown for 24 h. Indicated fragments were PCR amplified, subcloned and sequenced by Sanger sequencing. Indels and SNVs that restore the ORF are indicated by triangles and dots, respectively in indicated fragments of the construct that were sequenced ($N = 3$, $n = 50$). (**e**) High frequency of LacZ frameshift mutations was recovered in the heterochromatic reporter in *met-2 set-25* and *msh-6* worms, but not in the euchromatic reporter. Results were categorized according to the proportion of β-gal positive cells per worm (mean and s.e.m.; $N = 3$, $n = 50$). Scale bar, 50 µm.

*set-25* embryos followed by deep sequencing (DRIP-seq) or qPCR (DRIP-qPCR). By qPCR, we found that specific repeat elements that were derepressed in the absence of H3K9me, were enriched for R loops fourfold to ninefold in mutant over wild-type strains. This was not the case for low- or moderate-level transcribed genes (*unc-119* or *lmn-1*), nor was there a *met-2 set-25*-dependent increase in DRIP for a highly transcribed gene (*eef-A.1*), although the levels of R loops did increase at highly transcribed genes in both wild-type and *met-2 set-25* strains (**Fig. 5d**). As proof that the antibody was specific for RNA:DNA hybrids, we note that the DRIP-qPCR signal was highly sensitive to treatment with RNase H (**Fig. 5d**).

On a genome-wide level (DRIP-seq), we detected the most pronounced enrichment of RNA:DNA hybrids in *met-2 set-25* embryos on REs that were derepressed in the double mutant (**Fig. 5e**,**f** and **Supplementary Fig. 5e**). RNA:DNA hybrids were particularly enriched on transcribed DNA transposons and tandem repeats but not on RNA transposons (**Fig. 5e**). Confirming R-loop mapping in other organisms, we observed RNA:DNA hybrids more frequently on highly transcribed genes, telomeres and the rDNA locus, even in wild-type cells (**Supplementary Fig. 5b–d**)[52,53], yet these signals showed no further increase in the *met-2 set-25* mutant.

This high level of RNA:DNA hybrids suggests the presence of replication stress in *met-2 set-25* worms. To monitor their sensitivity to fork stalling, we exposed worms to hydroxyurea, a DNA replication inhibitor that reversibly inhibits ribonucleotide reductase, thereby depleting deoxynucleotide pools and exacerbating replication fork stalling[54]. L1 larvae exposed to 20 mM hydroxyurea for 16 h and allowed to recover for 3 d in absence of the inhibitor, yielded 95 ± 3% (mean ± s.d.) viability (resumption of development), whereas only 43 ± 11% of the *met-2 set-25* larvae survived hydroxyurea exposure (**Fig. 6a**). This hypersensitivity was specific to agents causing replication stress, as treating similarly staged larvae with ionizing radiation did not differentially affect wild-type and *met-2 set-25* strains (**Fig. 6b**). Thus hydroxyurea hypersensitivity correlated with the accumulation of R loops, and suggests that both the developmental delays and sterility detected in H3K9me-deficient worms reflect collisions of replication with unscheduled transcription.

### In the absence of H3K9me, mutations accumulated in REs

Replication stress and formation of R loops have been correlated with both fork instability and double-strand break hotspots in yeast[44,46,55,56]. To determine whether genomes of H3K9me-deficient worms accumulate mutations at elevated rates, we singled 8 wild-type and 8 *met-2 set-25* worms for 12 generations at 25 °C, thereby creating 8 individual substrains per genotype. Sequencing of the genome of each substrain revealed mutations exclusively in one of the 16 genomes (**Fig. 6c**). This allowed us to score the number, nature and location of changes accumulated owing to the *met-2 set-25* mutation.

We note that the rate and nature of single nucleotide variants (SNVs) did not differ between wild-type and *met-2 set-25* worms, which allowed us to exclude generation time as a confounding factor in the analysis (**Supplementary Fig. 6a**). However, 6 of the 8 *met-2 set-25* sub-strains acquired at least one insertion or deletion (indel) (with a total of 9 different observed indels; **Supplementary Table 1**). In contrast, only one wild-type strain incurred small deletions (3-base pair (bp) and 5-bp). The average indel in the *met-2 set-25* substrains covered 5.3 kilobases (kb) (the largest being 33.5 kb), all *met-2 set-25* indels occurred at sites enriched for H3K9me3, and 8 of the 9 *met-2 set-25* indels occurred in REs whose majority showed enhanced transcription upon loss of H3K9me.

Confirming the existence of large and stable germline changes, we detected a 10-kb inversion flanked on one side by a 1-kb deletion by whole genome sequencing and PCR of *met-2 set-25* worms that had been cultivated for several months. The inversion was immediately adjacent to an excised *Tc3* transposon, and opposite the inversion was a *de novo Tc3* transposon insertion unique to the cultivated H3K9me-deficient strain (**Supplementary Fig. 6b–d**). The excised *Tc3* element carried H3K9me3 in the wild-type strain, and was transcriptionally activated in *met-2 set-25*.

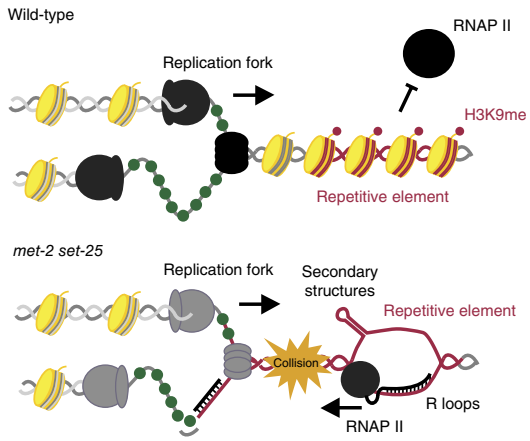


**Figure 8** Transcribed REs in H3K9me-deficient strains can exacerbate replication stress provoking genomic instability. A model illustrates how the loss of H3K9me could lead to the formation of secondary DNA structures that engender replication stress specifically at heterochromatic repeats, to perturb genome integrity.

We next checked wild-type and *met-2 set-25* genomes for copy number variations (CNVs) in repeat families with multiple members. Two DNA transposons (HELITRON2 and HELITRONY4) and two tandem repeats (MSAT1 and $(TATCG)_n$) showed high CNV uniquely in the *met-2 set-25* substrains. In contrast, the RNA transposon CEREP58 and the single-copy gene *lmn-1*, like telomeric repeats and the rDNA, remained stable (**Fig. 6d**,**e** and **Supplementary Fig. 6e**,**f**)[57]. RNA transposons, which also failed to accumulate R loops, did not show CNV. We conclude that *met-2 set-25* germ lines accumulated indels at sites bearing H3K9me in wild-type strains as well as changes in DNA-transposon and tandem-repeat copy number.

### A reporter incurred frequent indels in H3K9me-deficient somatic cells

This sequence analysis monitored stable germline changes in the worm population, and selected against any mutation that would perturb meiotic genome transmission. To visualize the mutation rate in somatic cells, we used a heterochromatic reporter with a *lacZ* gene placed out of frame to the ATG start codon, generating multiple premature stop codons in the first 100 bp of the transcript. Insertions or deletions between the ATG and the ORF are necessary to enable the translation of the *lacZ* mRNA into a functional β-galactosidase enzyme (**Fig. 7a**)[58]. This allowed us to compare mutation rates of wild-type and *met-2 set-25* worms by microscopy, following a colorimetric stain for heat-shock-induced β-galactosidase expression. By comparing two time points during somatic development (12 h and 24 h after L1) we could differentiate mutations that might have been present in the fertilized egg from mutations incurred during somatic development (**Fig. 7b**). To compare the mutation rate of repetitive heterochromatic and unique euchromatic sequences, we made use of the observation that transgenes integrated as high–copy number arrays induce the formation of H3K9me-containing heterochromatin[59] (i.e., enriched for H3K9me2 and H3K9me3; **Fig. 7c**). We compared this reporter with the same reporter construct integrated as a low–copy number array, which remains unmethylated and euchromatic. We classified phenotypes by the extent of β-galactosidase expression on a worm-by-worm basis, and sequenced the constructs amplified from worms at the 24-h time point (**Fig. 7d**,**e**).

In the wild-type background after 24 h of cultivation, the heterochromatic reporter produced functional β-galactosidase in only around 3 ± 2% (s.e.m.) of the worms (**Fig. 7e**; >1/3 expressing in-frame *lacZ*). In the *met-2 set-25* mutant, the fraction of worms expressing in-frame *lacZ* increased to 78 ± 8% (>1/3 staining blue, $P = 0.01$). In contrast, the euchromatic reporter did not express in-frame *lacZ* in *met-2 set-25* worms (1 ± 1%, any level of in-frame *lacZ*). We used a mutant of the mismatch repair machinery *msh-6* as a positive control[60]. The *met-2 set-25* mutant primarily showed an increase in the β-galactosidase-positive phenotype by 24 h, and not by 12 h, unlike the *msh-6* mutant (**Fig. 7e**), suggesting that the *met-2 set-25*-induced mutations occurred during differentiation. The types of mutations monitored were confirmed by batch-wise cloning of single reporter units and Sanger sequencing. We indeed detected small insertions and deletions in the *met-2 set-25* worms, enabling in-frame translation of β-galactosidase (**Fig. 7d**). Thus, like the germline changes scored by genome sequencing, sequences with H3K9me in wild-type backgrounds accumulated indels at high rates during somatic cell division in H3K9me-deficient worms.

## DISCUSSION

H3K9 methylation is the defining histone modification for heritably silent chromatin and is conserved as such from fission yeast to humans. *C. elegans* mutants lacking H3K9me are viable, despite the enrichment of H3K9me2/H3K9me3 on silent tissue-specific genes, on pseudogenes and on RE. In contrast to the case in other species[18,19,61], we found no defects in chromosome segregation upon loss of H3K9me. However, we observed a temperature-dependent sterility, which coincided with an increase in DNA-damage-induced apoptosis and stochastic delays in development. Correlating with these phenotypes, we detected derepression of ~20% RE, from all repeat classes, a value that increased at elevated temperatures. Expression of these RE was accompanied by the accumulation of RNA:DNA hybrids, CNV and a hypersensitivity to replication stress. This correlation suggests that it is either the transcription of the repetitive sequence alone, or transcription coupled with the inherent pairing nature of repeats, that generates insertions and deletions within REs in the absence of H3K9me. Of note, DNA transposons and tandem repeats showed higher levels of R loops and CNV than RNA transposons, although all classes were derepressed upon loss of H3K9me.

The damage incurred in the germ line leads to extensive apoptosis and Rad51 focus accumulation, suggesting that these cells accumulated double-strand breaks as well as indels. There may be additional sources of damage in the germ line, other than those that correlate with replication-fork-associated damage, R loops and indels scored by genome sequencing. We note that RNA polymerase-DNA polymerase collision has been reported to generate fragile sites of breakage[44,62], which in worm germline cells would provoke an apoptotic response[32,63].

It is likely that the genomic mutations we detected in the *met-2 set-25* strain arise from replication fork perturbation. This can be triggered by enhanced stalling of the replication fork generated by R-loop formation, which in turn allows hairpin or fold-back structures to form in repeats as they are being replicated. Hairpin or fold-back structures can also arise from breaks in the single-stranded DNA that accumulate either at R loops or behind the fork, owing to perturbed coordination between leading- and lagging-strand polymerases (**Fig. 8**). The passage of the replication fork through REs itself can lead to hairpin structures[41]. However, we propose that in *met-2 set-25* cells, unprogrammed transcription of REs enhances R loops, which may in turn enhance aberrant structures to such a degree that the cellular machineries that normally relieve such stress, can no longer cope with their abundance.

Unscheduled collisions of the replication and transcription machineries appear to generate breaks as well as other forms of genome instability[42–46,62]. Damage is often attributed to the presence of RNA:DNA hybrids[64–67], yet torsional stress, which can arise from high levels of bidirectional transcription[68,69], may also contribute to genomic instability. We consider it notable not only that the derepression of RE generated genomic mutations and R loops, but that both these events mapped to REs that are normally marked by H3K9me in wild-type cells, and which became derepressed in a temperature-enhanced manner in the *met-2 set-25* mutant. We propose that the crucial role of H3K9 methylation in suppressing transcription on a genome-wide level is not to program cell differentiation, but to stabilize repetitive sequences that accumulate in higher eukaryotic genomes.

Several studies have suggested the use of inhibitors for H3K9me HMTs in the treatment of cancer (for example, lung, prostate, hepatocellular and pancreatic cancer)[19,61], and preclinical studies have been considered promising so far[70]. These same inhibitors have been used to show that hypomethylation of H3K9 increases the rate of induced pluripotent stem cell generation[15,16]. We argue that there are clear drawbacks to such therapies, given the genomic instability provoked by loss of H3K9me shown here. Whereas mammals additionally silence through meCpG, it has been documented that DNA methylation can be targeted by H3K9me or its HMTs[71]. Thus, the findings presented in this study are likely to have implications for protocols that attempt to manipulate the mammalian epigenome.

**URLs.** http://www.bioconductor.org/packages/3.1/bioc/html/QuasR.html.

## METHODS

Methods and any associated references are available in the online version of the paper.

**Accession codes.** All data from this study have been deposited in the Sequence Read Archive (SRA) under accession SRP080806.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

### ACKNOWLEDGMENTS

A number of strains were provided by the *Caenorhabditis* Genetics Center (CGC), which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We thank R. Ciosk and P. Pasero for reagents and materials, I. Katiç and members of the Friedrich Miescher Institute Genomics and Microscopy facilities for advice and discussion, and P. Ginno and L. Constantino for advice on R-loop detection. J.P. is supported by a long-term EMBO fellowship. S.M.G. thanks the Swiss National Science Foundation as well as the Novartis Research Foundation for support.

### AUTHOR CONTRIBUTIONS

P.Z. and J.P. planned and executed most experiments, evaluated results and wrote the paper; S.M.G. planned experiments, evaluated results and wrote the paper; R.v.S. and M.T. helped with evaluation of genome mutations and provided the *LacZ* mutagenesis assay; and V.K. provided invaluable technical help.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.



1. Bannister, A.J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
2. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
3. Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. & Grewal, S.I. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**, 110–113 (2001).
4. Allshire, R.C., Javerzat, J.P., Redhead, N.J. & Cranston, G. Position effect variegation at fission yeast centromeres. *Cell* **76**, 157–169 (1994).
5. Cam, H.P. *et al.* Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* **37**, 809–819 (2005).
6. Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192–1200 (2003).
7. Peters, A.H. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
8. Eissenberg, J.C. *et al.* Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position–effect variegation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **87**, 9923–9927 (1990).
9. Epsztejn-Litman, S. *et al. De novo* DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15**, 1176–1183 (2008).
10. Wissmann, M. *et al.* Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor–dependent gene expression. *Nat. Cell Biol.* **9**, 347–353 (2007).
11. Chen, M.W. *et al.* H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. *Cancer Res.* **70**, 7830–7840 (2010).
12. Hua, K.T. *et al.* The H3K9 methyltransferase G9a is a marker of aggressive ovarian cancer that promotes peritoneal metastasis. *Mol. Cancer* **13**, 189 (2014).
13. Kumari, D. & Usdin, K. The distribution of repressive histone modifications on silenced *FMR1* alleles provides clues to the mechanism of gene silencing in fragile X syndrome. *Hum. Mol. Genet.* **19**, 4634–4642 (2010).
14. Lee, J. *et al.* Epigenetic regulation of cholinergic receptor M1 (CHRM1) by histone H3K9me3 impairs $Ca^{2+}$ signaling in Huntington's disease. *Acta Neuropathol.* **125**, 727–739 (2013).
15. Loh, Y.H., Zhang, W., Chen, X., George, J. & Ng, H.H. Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. *Genes Dev.* **21**, 2545–2557 (2007).
16. Shi, Y. *et al.* A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* **2**, 525–528 (2008).
17. Treangen, T.J. & Salzberg, S.L. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.* **13**, 36–46 (2011).
18. Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H. & Li, E. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.* **24**, 2478–2486 (2004).
19. Kondo, Y. *et al.* Downregulation of histone H3 lysine 9 methyltransferase G9a induces centrosome disruption and chromosome instability in cancer cells. *PLoS One* **3**, e2037 (2008).
20. Tachibana, M. *et al.* G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16**, 1779–1791 (2002).
21. Mellone, B.G. *et al.* Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. *Curr. Biol.* **13**, 1748–1757 (2003).
22. Peng, J.C. & Karpen, G.H. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat. Cell Biol.* **9**, 25–35 (2007).
23. Andersen, E.C. & Horvitz, H.R. Two *C. elegans* histone methyltransferases repress *lin-3* EGF transcription to inhibit vulval development. *Development* **134**, 2991–2999 (2007).
24. Towbin, B.D. *et al.* Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**, 934–947 (2012).
25. Harris, J. *et al.* Mutator phenotype of *Caenorhabditis elegans* DNA damage checkpoint mutants. *Genetics* **174**, 601–616 (2006).
26. Bernard, P. *et al.* Requirement of heterochromatin for cohesion at centromeres. *Science* **294**, 2539–2542 (2001).
27. Batista, P.J. *et al.* PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell* **31**, 67–78 (2008).
28. Ketting, R.F., Haverkamp, T.H., van Luenen, H.G. & Plasterk, R.H. Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99**, 133–141 (1999).
29. Sijen, T. & Plasterk, R.H. Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* **426**, 310–314 (2003).
30. Zhou, Z., Hartwieg, E. & Horvitz, H.R. CED-1 is a transmembrane receptor that mediates cell corpse engulfment in *C. elegans*. *Cell* **104**, 43–56 (2001).
31. Greiss, S., Schumacher, B., Grandien, K., Rothblatt, J. & Gartner, A. Transcriptional profiling in *C. elegans* suggests DNA damage dependent apoptosis as an ancient function of the p53 family. *BMC Genomics* **9**, 334 (2008).
32. Gartner, A., Milstein, S., Ahmed, S., Hodgkin, J. & Hengartner, M.O. A conserved checkpoint pathway mediates DNA damage–induced apoptosis and cell cycle arrest in *C. elegans*. *Mol. Cell* **5**, 435–443 (2000).
33. Schumacher, B., Hofmann, K., Boulton, S. & Gartner, A. The *C. elegans* homolog of the p53 tumor suppressor is required for DNA damage–induced apoptosis. *Curr. Biol.* **11**, 1722–1727 (2001).
34. Gerstein, M.B. *et al.* Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* **330**, 1775–1787 (2010).
35. Liu, T. *et al.* Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Res.* **21**, 227–236 (2011).

36. Cangiano, G. & La Volpe, A. Repetitive DNA sequences located in the terminal portion of the *Caenorhabditis elegans* chromosomes. *Nucleic Acids Res.* **21**, 1133–1139 (1993).
37. Padeken, J., Zeller, P. & Gasser, S.M. Repeat DNA in genome organization and stability. *Curr. Opin. Genet. Dev.* **31**, 12–19 (2015).
38. Das, P.P. *et al.* Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* **31**, 79–90 (2008).
39. Tabara, H. *et al.* The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99**, 123–132 (1999).
40. Wicker, T. *et al.* A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973–982 (2007).
41. Aguilera, A. & García-Muse, T. Causes of genome instability. *Annu. Rev. Genet.* **47**, 1–32 (2013).
42. Azvolinsky, A., Giresi, P.G., Lieb, J.D. & Zakian, V.A. Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. *Mol. Cell* **34**, 722–734 (2009).
43. Deshpande, A.M. & Newlon, C.S. DNA replication fork pause sites dependent on transcription. *Science* **272**, 1030–1033 (1996).
44. Hoffman, E.A., McCulley, A., Haarer, B., Arnak, R. & Feng, W. Break-seq reveals hydroxyurea-induced chromosome fragility as a result of unscheduled conflict between DNA replication and transcription. *Genome Res.* **25**, 402–412 (2015).
45. Merrikh, H., Machón, C., Grainger, W.H., Grossman, A.D. & Soultanas, P. Co-directional replication–transcription conflicts lead to replication restart. *Nature* **470**, 554–557 (2011).
46. Wahba, L., Amon, J.D., Koshland, D. & Vuica-Ross, M. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* **44**, 978–988 (2011).
47. Nakamori, M., Pearson, C.E. & Thornton, C.A. Bidirectional transcription stimulates expansion and contraction of expanded (CTG)*(CAG) repeats. *Hum. Mol. Genet.* **20**, 580–588 (2011).
48. Prak, E.T. & Kazazian, H.H. Jr. Mobile elements and the human genome. *Nat. Rev. Genet.* **1**, 134–144 (2000).
49. Boguslawski, S.J. *et al.* Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J. Immunol. Methods* **89**, 123–130 (1986).
50. Castellano-Pozo, M., García-Muse, T. & Aguilera, A. The *Caenorhabditis elegans* THO complex is required for the mitotic cell cycle and development. *PLoS One* **7**, e52447 (2012).
51. Castellano-Pozo, M. *et al.* R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* **52**, 583–590 (2013).
52. Balk, B. *et al.* Telomeric RNA–DNA hybrids affect telomere-length dynamics and senescence. *Nat. Struct. Mol. Biol.* **20**, 1199–1205 (2013).
53. Wahba, L., Costantino, L., Tan, F.J., Zimmer, A. & Koshland, D. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* **30**, 1327–1338 (2016).
54. Rosenkranz, H.S. & Levy, J.A. Hydroxyurea: a specific inhibitor of deoxyribonucleic acid synthesis. *Biochim. Biophys. Acta* **95**, 181–183 (1965).
55. Gan, W. *et al.* R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* **25**, 2041–2056 (2011).
56. Zhang, H. & Freudenreich, C.H. An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in *S. cerevisiae*. *Mol. Cell* **27**, 367–379 (2007).
57. Collins, J., Saari, B. & Anderson, P. Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. *Nature* **328**, 726–728 (1987).
58. Pothof, J. *et al.* Identification of genes that protect the *C. elegans* genome against mutations by genome-wide RNAi. *Genes Dev.* **17**, 443–448 (2003).
59. Towbin, B.D., Meister, P., Pike, B.L. & Gasser, S.M. Repetitive transgenes in *C. elegans* accumulate heterochromatic marks and are sequestered at the nuclear envelope in a copy-number- and lamin-dependent manner. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 555–565 (2010).
60. Tijsterman, M., Pothof, J. & Plasterk, R.H.A. Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient *Caenorhabditis elegans*. *Genetics* **161**, 651–660 (2002).
61. Wagner, T. & Jung, M. New lysine methyltransferase drug targets in cancer. *Nat. Biotechnol.* **30**, 622–623 (2012).
62. Helmrich, A., Ballarino, M. & Tora, L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol. Cell* **44**, 966–977 (2011).
63. Vermezovic, J., Stergiou, L., Hengartner, M.O. & d'Adda di Fagagna, F. Differential regulation of DNA damage response activation between somatic and germline cells in *Caenorhabditis elegans*. *Cell Death Differ.* **19**, 1847–1855 (2012).
64. Chan, Y.A., Hieter, P. & Stirling, P.C. Mechanisms of genome instability induced by RNA-processing defects. *Trends Genet.* **30**, 245–253 (2014).
65. El Hage, A., Webb, S., Kerr, A. & Tollervey, D. Genome-wide distribution of RNA–DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.* **10**, e1004716 (2014).
66. Kilchert, C., Wittmann, S. & Vasiljeva, L. The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* **17**, 227–239 (2016).
67. Skourti-Stathaki, K., Proudfoot, N.J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
68. Kim, N. & Jinks-Robertson, S. Transcription as a source of genome instability. *Nat. Rev. Genet.* **13**, 204–214 (2012).
69. Kouzine, F., Levens, D. & Baranello, L. DNA topology and transcription. *Nucleus* **5**, 195–202 (2014).
70. Yuan, Y. *et al.* Gossypol and an HMT G9a inhibitor act in synergy to induce cell death in pancreatic cancer cells. *Cell Death Dis.* **4**, e690 (2013).
71. Liu, S. *et al.* Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev.* **28**, 2041–2055 (2014).

## ONLINE METHODS

***C. elegans* cultures and strains. Supplementary Table 2** lists the strains used in this study. Strains were made by backcrossing deletion alleles and reporter strains obtained from the *C. elegans* knockout consortium to the GW638 strain (met-2(n4256) III; set-25(n5021) III) at least five times. Worms were grown at 20 °C, except where specifically indicated.

**Immunofluorescence analysis, antibodies and live microscopy, including apoptosis assay.** IF analysis was carried out as previously described[24] by freeze-cracking and fixation in 1% paraformaldehyde followed by short postfixation in methanol (for embryos and gonads[72]) or methanol followed by acetone (for larval stages). Staining was performed in PBS with 0.1% TritonX-100 and 2% milk powder. For live-cell imaging, larvae were mounted on slides coated with 2% agarose. Microscopy was carried out on a spinning disc confocal microscope (SD1, W1, Visitron, Puchheim). Stacks of images were analyzed using ImageJ.

Antibodies used in this study were mouse anti-H3K9me2, MABI0317 (MBL[73]), mouse anti-H3K9me3, MABI0318 (MBL[73]), mouse anti-RNA:DNA hybrid S9.6, hybridoma HB-8730 (ATCC)[49], rabbit anti-pan-acetyl H4, 06-866 (Merck Millipore) and rabbit anti-RAD-51, 29480002 (Novu Biologics).

**Developmental timing, progeny size and hatching rate.** Worms of indicated genotype were synchronized through bleaching and were then singled onto plates containing OP50 bacteria. For the developmental timing their stage was determined every 24 h. In order to determine the progeny size, adults were transferred to fresh plates once a day for three days to keep generations separate and their complete progeny size was determined after their hatching at the indicated temperature. To determine the hatching rate singled worms were transferred every 8 h to freshly seeded plates. The number of laid embryos was determined directly after transfer, the number of hatched animals was determined on day 3. If not otherwise indicated, worms were grown at the experimental temperature (transferred from 20 °C) for at least two generations before the experiments.

**Chromatin immunoprecipitation experiments.** Early embryonic progeny was harvested after synchronization (60–65 h depending on each strain) for wt and *met-2 set-25* mutant strains in two independent biological replicates. H3K9me2 and H3K9me3 ChIP was performed as previously described[74] using the antibodies mentioned above. In brief, 40 µg of chromatin was incubated overnight with 3–6 µg of antibody coupled to Dynabeads Sheep Anti-Mouse IgG (Invitrogen) or Dynabeads Sheep Anti-Rabbit IgG (Invitrogen), in FA buffer (50 mM HEPES/KOH pH 7.5, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 150 mM NaCl)) containing 1% SDS. Chromatin-antibody complexes were washed with the following buffers: 3 × 5 min FA buffer; 5 min FA buffer with 1 M NaCl; 10 min FA buffer with 500 mM NaCl; 5min with TEL buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0) and twice for 5 min with TE. Complexes were eluted at 65 °C in 100 µl of elution buffer (1% SDS in TE with 250 mM NaCl) for 15 min. Both input and IP samples were incubated with 20 µg of RNAse A for 30 min at 37 °C and 20 µg of proteinase K for 1 h at 55 °C. Crosslinks were reversed overnight at 65 °C. DNA was purified using a Zymo DNA purification column (Zymo Research).

**Library preparation and analysis.** Libraries were prepared from chromatin IP and genomic DNA samples using the NEBNext ultra DNA library prep kit for Illumina (NEB # 7370) and the NEBNext Multiplex Oligos for Illumina (NEB # E7335), according to the manufacturer's recommendations. No size selection was performed during sample preparation and the libraries were indexed and amplified using 12 PCR cycles, using the recommended conditions. After further purification with Agencourt AmPure XP beads (Beckman # A63881), the library size distribution and concentrations were determined using a BioAnalyzer 2100 (Agilent technologies) and Qubit (Invitrogen) instrument, respectively. The final pools were prepared by mixing equimolar amounts of all individually indexed libraries and then sequenced on a HiSeq 2500 (Illumina) in rapid mode (Paired-End 50). Processing of the LEM-2 ChIP-seq data, all paired-end ChIP-seq data (2 × 50 bp) were mapped to the *C. elegans* genome (ce6) with the R package QuasR[75] using the included aligner bowtie[76]. Definitions of REs were taken from Repbase[77]. Repeat subfamilies were built to allow assignment of multimapping reads to all REs and collapsing single elements according to their Repbase ID into families.

Read density along the genome was calculated by tiling the genome into 200-bp windows (non-overlapping) and counting the number of sequence fragments within each window, using the qCount function of the QuasR package (see URLs). To compensate for differences in the read depths of the various libraries, we divided each sample by the total number of mapped reads and multiplied by the average library size. $\text{Log}_2$ expression levels were calculated after adding a pseudocount of 1 ($y = \log_2(x + 1)$). ChIP-seq signals are displayed as average enrichment of IP − input ($\log_2$).

**RNA expression experiments (RNA-seq and qPCR).** For embryos and larvae, RNA was isolated by freeze cracking (four times) followed by phenol-chloroform extraction and isopropanol precipitation. Total RNA was depleted for rRNA using Ribo-Zero Gold kit from Epicentre before library production using Total RNA Sequencing ScriptSeq kit. Gonad RNA was extracted from 50 prepared gonads per replica using the Arcturus pico pure RNA isolation kit followed by library production using the Total RNA-seq NuGen Ovation kit. 50-bp single-end sequencing was done on an Illumina HiSeq 2500. Processing of the RNA-seq data, gene and repeat expression levels from RNA-seq data were quantified as described previously[24] using WormBase (WS190) annotation for coding transcripts and Repbase annotations for REs. Primers used for qPCR experiments are listed in **Supplementary Table 3**.

**Mutation sequencing experiments.** Worms were grown at 25 °C for 1 month, and singled every second generation. Afterward worms were expanded on peptone-rich plates (20 cm) per replica and mixed staged worms were harvested for genomic DNA isolation. DNA was extracted by a standard protocol digesting worms with proteinase K, followed by phenol/chloroform extraction and RNAse treatment. 50-bp paired-end sequencing was done on an Illumina HiSeq 2500. Reads were mapped to the WS190 genome using BWA[78] and converted into BAM files using samtools[79]. Breakpoints were identified with Pindel[80] and SNVs with samtools.

**Southern blot.** Southern blot was performed following a standard protocol using a digoxigenin-labeled probe produced by PCR with primers listed in **Supplementary Table 3**.

**LacZ mutator assay and cloning for somatic mutations.** LacZ mutator assay was adapted from ref. 58. Worms were synchronized and grown for indicated durations on Dh5α containing plates. After a heat shock (heterochromatic array: 5 h (2 h at 33 °C, 1 h at 20 °C and 2 h at 33 °C), euchromatic array: 1 h 20 min (20 min at 33 °C, 10 min at 20 °C, 20 min at 33 °C, 10 min at 20 °C and 20 min at 33 °C) and 2 h recovery at 20 °C, worms were stained for β-galactosidase expression. To identify somatic mutations the indicated regions of the reporter were amplified using a Q5 proofreading polymerase (NEB) and primers listed in **Supplementary Table 3**. PCR products were batch clones into pCR2.1-TOPO sequencing vector (Invitrogen) and Sanger sequencing was performed on 20 clones per replica and region.

**DNA damage sensitivity assays.** Assays were previously described[81]. Recovery from an hydroxyurea (HU) pulse was monitored by soaking L1 larvae in M9 buffer containing indicated concentrations of HU and OP50 bacteria for 16 h before washing and plating on fresh OP50 plates. At day 3, the percentage of viable adults was quantified. To quantify IR sensitivity worms were irradiated (CellRad, Faxitron) at the L1 stage. At day 3, the percentage of viable adults was quantified.

**R-loop detection.** For dot plots, genomic DNA was isolated using phenol-chloroform extraction followed by ethanol precipitation. DNA concentrations were determined using Nanodrop and the indicated amount of DNA was resuspended to a final volume of 50 µl in nuclease-free water after either a 1 h incubation with 5 µl of RNase H (NEB; +RNAse H), or a 1 h mock incubation at 37 °C, and spotted directly onto a nylon GeneScreen Plus membrane (NEF988; PerkinElmer) using a Bio-Dot Microfiltration Apparatus (Bio-Rad). The membrane was UV-crosslinked and blocked with 5% milk in




doi:10.1038/ng.3672

1 × PBS/0.1% Tween-20 before incubation with primary and secondary antibodies. The mouse S9.6 antibody (HB-8730, ATCC, gift of P. Pasero, Montpellier) was used at a 1:500 dilution, and a 10,000× dilution of goat anti-mouse HRP (Bio-Rad) was used as the secondary. The HRP signal was developed with Clarity Western ECL Substrate (Bio-Rad). Imaging was performed using ImageQuant LAS4000 mini und analyzed using ImageJ. Immunofluorescence staining was performed according to the steps described for the *C. elegans* larval stages using 4XSSC-T (0.1% Tween-20) instead of PBS-T. R loops were stained with the S9.6 antibody, diluted 1:100 in SSC-T and 3% BSA overnight at 4 °C.

**DNA:RNA hybridization.** Embryos were lysed by bead beating (MP BIOMEDICALS FastPrep-24 5G Instrument) in G2 buffer (80 mM guanidine HCl, 30 mM Tris pH 8.0, 30 mM 5% Tween, 0.5% TritonX). Genomic DNA was isolated by proteinase K digestion, followed by purification using genomic tips (500/G, QIAGEN). DNA was digested over night with AseI and BstUI at 37 °C (ref. 82). For RNase H control samples, RNase H was added in parallel to digestion. 5 µg of digested DNA per IP was incubated with 10 µl of S9.6 antibody overnight in binding buffer (10 mM $NaPO_4$, 140 mM NaCl, 0.5% Triton). Bound DNA fragments were recovered with 50 µl of Protein-A Dynabeads (Invitrogen), followed by four washes with binding buffer and proteinase K treatment. Samples were purified using DNA Clean & Concentrator-5 (Zymo Research) columns. Samples were sonicated to ~400-bp fragments before library preparation, as described above.


72. Meister, P., Towbin, B.D., Pike, B.L., Ponti, A. & Gasser, S.M. The spatial dynamics of tissue-specific promoters during *C. elegans* development. *Genes Dev.* **24**, 766–782 (2010).
73. Kimura, H., Hayashi-Takanaka, Y., Goto, Y., Takizawa, N. & Nozaki, N. The organization of histone H3 modifications as revealed by a panel of specific monoclonal antibodies. *Cell Struct. Funct.* **33**, 61–73 (2008).
74. Ikegami, K., Egelhofer, T.A., Strome, S. & Lieb, J.D. *Caenorhabditis elegans* chromosome arms are anchored to the nuclear membrane via discontinuous association with LEM-2. *Genome Biol.* **11**, R120 (2010).
75. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M.B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31**, 1130–1132 (2015).
76. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
77. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462–467 (2005).
78. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
79. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
80. Ye, K., Schulz, M.H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).
81. Craig, A.L., Moser, S.C., Bailly, A.P. & Gartner, A. Methods for studying the DNA damage response in the *Caenorhabdatis elegans* germ line. *Methods Cell Biol.* **107**, 321–352 (2012).
82. Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I. & Chédin, F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* **45**, 814–825 (2012).