Title:

ATRX safeguards cellular identity during *C. elegans* development

Authors:

Janie Olver[1, 2], Mariya Shtumpf[1], Karim Hussain[1, 3], Stephen Methot[4], Peter Sarkies[5] and Helder Ferreira[1*]

Affiliations:

[1] School of Biology, University of St Andrews, St Andrews, UK

[2] Currently at Rare and Imported Pathogens Laboratory, UK Health Security Agency, Porton Down, Salisbury, UK

[3] Currently at Ingenza Ltd, Midlothian, UK

[4] Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

[5] Department of Biochemistry, University of Oxford, Oxford, UK

* Corresponding author: hcf2@st-andrews.ac.uk

Acknowledgements

Some strains were provided by the Caenorhabditis Genetics Center, which is funded by the National Institutes of Health Office of Research Infrastructure Programs (P40 OD010440). We thank Prof. Julie Ahringer and Dr Chantal Wicky for helpful discussions. The authors acknowledge Research Computing at the James Hutton Institute for providing computational resources and technical support for the UK's Crop Diversity Bioinformatics HPC (BBSRC grants BB/S019669/1 and BB/X019683/1), use of which has contributed to the results reported within this paper. We thank Peter Thorpe, Marco Ferreira Fernandes and Paul Johnson for help with bioinformatics.

Abstract

ATRX is a member of the SWI/SNF family of ATP-dependent chromatin remodellers. In humans, loss of ATRX function leads to ATRX syndrome, a neurodevelopmental disorder. ATRX mutation in human cell lines is associated with multiple phenotypes including activation of the alternative lengthening of telomere (ALT) pathway, upregulation of retrotransposons and increased sensitivity to replication stress. However, the principal role of ATRX and the reason why its mutation causes such diverse phenotypes is currently unclear. To address this, we studied the role of ATRX in the model organism *Caenorhabditis elegans*. We find that loss of XNP-1, the *C. elegans* homologue of ATRX, recapitulates many human phenotypes. In addition, XNP-1 is required to repress the inappropriate activation of germline genes. Importantly, this germline misexpression correlates with most of the phenotypes observed in *xnp-1* animals. Seemingly distinct *xnp-1* phenotypes such as developmental abnormalities and telomeric defects are both suppressed by mutation of the germline transcription factor *gsox-1*. These findings suggest that the majority of XNP-1-dependent phenotypes stem from its role in maintaining proper cellular identity, offering insights into the functions of ATRX in humans.

Introduction

Chromatin structure impacts all processes that require access to DNA. Thus, proteins that alter chromatin structure, particularly ATP-dependent chromatin remodelling enzymes, are important mediators of DNA transcription, replication and repair. Indeed, many disorders are known to be caused by mutation of human ATP-dependent chromatin remodelling genes (1). One of the best described is the neurodevelopmental disorder ATR-X syndrome, which gave its name to the causal gene, the chromatin remodelling factor, ATRX (2).

Fitting with the ubiquitous roles that chromatin remodelling enzymes play in DNA metabolism, ATRX loss is also strongly associated with cancer, specifically those cancers that use the alternative lengthening of telomere (ALT) pathway (3). ALT uses break-induced replication to maintain telomere length (4,5) and ATRX has been shown to act as a tumour suppressor by repressing the ALT pathway (6). Precisely how ATRX influences both ALT and normal development is unclear. This is compounded by the fact that ATRX has been shown to impact many molecular processes. ATRX interacts with DAXX to deposit the replication-independent histone H3.3 at heterochromatin (7,8) and euchromatin (9). Loss of ATRX leads to increased retrotransposon expression (10-12), increased levels of G-quadruplexes (13), reduced sister chromatid cohesion (14,15) and increased sensitivity to replication stress (16,17). Unpicking the relative importance of these molecular functions within a developmental context has been made harder by the fact that complete loss of ATRX is embryonic lethal in standard model organisms such as fruit flies, mice and zebrafish (18-20). Conditional mouse knockouts have allowed the study of ATRX in specific tissues, particularly neuronal lineages (19,21). However, the role of ATRX in early embryogenesis is particularly poorly understood.

To address this, we used the nematode *C. elegans* as a model system, because the loss of its ATRX homolog, XNP-1, does not result in embryonic lethality. *C. elegans* has a particularly well described and stereotypical order of embryonic development where the fate of every cell is known. These decisions are largely governed by the restricted expression of master transcription factors (22,23). However, chromatin structure also plays a role in controlling tissue-specific expression and therefore cellular identity. This is particularly relevant in the very first cell fate decision in *C. elegans* embryogenesis in which germline versus somatic cell fate is specified (24). The germline is unique as it maintains genetic continuity between generations. Therefore, establishing and maintaining correct germline identity is crucial. Early germline identity in worms is promoted by MES proteins; histone methyltransferases that establish the appropriate genome-wide patterns of H3K27 and H3K36 methylation in germline cells (25). In contrast, in somatic cells the synMuvB proteins, block the activation of germline transcripts (23). SynMuvB proteins include histone deacetylases, histone methyltransferases

(H3K9-specific) and the H3K9me^3 binding factor HPL-2. However, the ATP-dependent chromatin remodelling enzyme LET-418 is also a synMuv B protein. Therefore, it seemed possible that other chromatin remodelling enzymes, such as XNP-1, might also contribute to restricting germline identity.

In this manuscript, we show that *xnp-1* mutants exhibit impaired germline function. We demonstrate that this impairment is due to the aberrant expression of germline genes, which in turn results in increased transposable element expression. Consistent with this hypothesis we identified the germline specific transcription factor *F33H1.4*, which we rename *gsox-1*, as a suppressor of germline defects in *xnp-1*. Together, our work identifies a novel role for ATRX homologs in maintaining cellular identity by regulating germline gene expression.

Materials and Methods

Brood Size

Brood size at 20°C was carried out by singling L4 worms onto OP50 NGM plates and then singling for 3-4 days before counting the total number of progeny. For brood size on plates containing 10 mM hydroxyurea, worms were added to these plates as L1s and then brood size was carried out as above. The worms were maintained on plates containing 10 mM hydroxyurea throughout.

Brood size at 25°C was carried out by shifting L4 worms that were previously grown at 20°C to 25°C. The brood size of their progeny (the F1 or the F2 generation) was recorded in the same way as above.

C-Circle Assay

Animals were synchronised to L1s and approximately 1000 L1s were seeded onto 5-6 OP50 plates (9 cm). Plates were maintained at 20°C or 25°C for 72 hours until animals reached gravid adult stage.

The C-circle assay was performed as described (26) with the difference that the genomic DNA was extracted by bead-beating. In brief, 300 µl of buffer (100 µg/ml RNase A, 10 mM EDTA, 300 mM NaCl) and 400 µl of 0.5 mm glass beads were added to worm pellets and disrupted by mechanical lysis using a cell homogeniser for 3x 20 s at 6 m/s with cooling on ice in between each step. Proteins were denatured by adding 1% SDS and heating the extracts at 65°C for 10 minutes after which the SDS was precipitated with potassium acetate.

The supernatant was collected after spinning and DNA purified using phenol:chloroform:isoamyl alcohol (15:24:1, pH 6.7) extraction and a chloroform back-extraction. DNA was ethanol precipitated and eluted in Tris-EDTA buffer (10 mM Tris-HCl, 100 µM EDTA, pH 7.5). Phi29 polymerase (NEB) was added to 1 µg of genomic DNA in a total volume of 20 µl and incubated at 30°C for 8 hours. This was spotted onto a neutral Hybond-N membrane, 1200J/m2 UV cross-linked, and hybridized with a DIG-labelled $(GCCTAA)_4$ probe at 37°C using DIG Easy Hyb (Roche) according to the manufacturer's instructions.

Telomere length by terminal restriction fragment (TRF) analysis

Genomic DNA was extracted 1x NTE buffer (100 mM NaCl, 50 mM Tris pH 7.4, 20 mM EDTA), 1% SDS and 500 µg/ml Proteinase K overnight at 65°C. DNA was purified by two consecutive phenol-chloroform washes (15:24:1, pH 6.7), followed by a chloroform back-extraction and ethanol precipitation. DNA was dissolved in Tris-EDTA pH 8 and incubated with 50 µg/ml RNase A for 30 minutes at 37°C followed by a final round of phenol-chloroform extraction and ethanol precipitation. 5 µg purified genomic DNA was digested overnight with HinfI and HaeIII (NEB) at 37°C and resolved on 1% agarose gel in 1x TAE. Following a 20-minute depurination in 250 mM HCl, the gel was washed 2x in denaturing buffer (1.5 M NaCl, 0.5 M NaOH) and 2x in neutralising buffer (1.5 M NaCl, 0.5 M Tris-HCl, pH 8) at room temperature. DNA was transferred onto neutral nylon membrane (Hybond-NX, GE Healthcare) by capillary transfer in 10x SSC buffer (1.5M NaCl, 150 mM sodium citrate, pH 7). After briefly rinsing in 2x SSC buffer (300 mM NaCl, 30 mM sodium citrate, pH 7) DNA was UV crosslinked at 1200 J/m^2 and hybridised with a digoxygenin-labelled telomere probe $(GCCTAA)_4$ (1 µg/ml in DIG Easy Hyb™ hybridisation buffer, Roche) for 2 hours at 37°C. Membranes were washed 2x in MS buffer (100 mM maleic acid, 150 mM NCl, pH 7.5), 0.3% Tween-20, blocked in MS buffer, 1% BSA, 1% milk, 0.1% Tween-20, probed with AP-conjugated anti-digoxigenin Fab fragments (Roche) at 1:20 000 dilution in the same buffer (1:20000), followed by 3 washes in MS buffer, 0.3% Tween-20. Membranes were equilibrated in AP buffer (100 mM Tris-HCl, pH 9.5, 100 mM NaCl) followed by chemiluminescent detection with CDP-Star® substrate (Roche).

Quantifying C. elegans growth rates

Post-embryonic growth rates were measured using a luciferase-based assay performed as described previously (27), based on (28). Briefly, wells in luminometer compatible 384-well plates (Berthold Technologies, 32505) were filled with 90 µl S-Basal medium *containing E.*

coli OP50 (OD600 = 1) and 100 µM Firefly D-Luciferin (p.j.k., #102111). Single eggs, from animals expressing luciferase under a constitutive and ubiquitous promoter (xeSi312[eft-3p::luc::gfp::unc-54 3′UTR, unc-119(+)] IV) (29), were then loaded into individual wells and plates were sealed with Breathe Easier sealing membrane (Diversified Biotech, BERM-2000). Development was monitored over 96 hrs by measuring luminescence in every well (every 10 mins for 0.5 secs) using a luminometer (Berthold Technologies, Centro XS3 LB 960) placed in a temperature-controlled incubator set to 18.5°C. Luminescence data were analyzed similarly to before (27), but using the python-based 'PyLuc' tool developed L. Morales Moya (29). The identification of molts (the final phase of each larval stage) was based on a drop in luminescence signal, which results from animals not feeding during this period. For each animal, larval stage durations were then calculated as the time from hatch to the end of the first molt (L1), or from the end of one molt to the end of the subsequent molt (L2 to L4).

RNA Experiments

RNA extraction was performed as described in (30). Approximately 800 ng of RNA was used as input to each reverse transcription reaction (GoScript Reverse Transcriptase, Promega). The qPCR reaction was carried out on a QuantStudioTM Real-Time PCR system machine (Thermo Scientific) with the following reaction steps: 95°C for 60" → 40x(95°C for 15" → 59°C for 30"). The qPCR primers used are given in Supplementary Table 2.

RNA sequencing of embryos and L1 was carried out at the Oxford Genomics Centre and for the L4 samples was carried out at Novogene, Cambridge. All genotypes where sequences in biological triplicate. Briefly, reads were trimmed to remove Illumina adaptors using Trim_galore, a wrapper script from Cutadapt (version 0.6.7) (31) using default settings. Trimmed reads were mapped to the reference genome (release WS280, BioProject PRJNA13758) using STAR (version 2.7.10a) (32), without limits on the mapped length and with maximum number of mismatches per pair set to 2. The reads were re-sorted by read name using Samtools (version 1.6) (33). FeatureCounts (Subread, version 2.0.1) (34) was used to generate a count matrix at gene-level. Differential gene expression was conducted in edgeR (version 3.40.0) (35) on a local machine using R (version 4.2.2) applying a minimum gene count of four per replicate and 20 counts across all three replicates per genotype. All scripts can be found in our Github repository ([I will create this one later]). SimpleMine web tool from WormBase was used to convert WormBase gene IDs to gene names. Gene set enrichment analysis was also ran on the WormBase web server and the results were plotted using the bubble plot function provided by the SRplot web server (https://www.bioinformatics.com.cn/srplot) (36). Small non-coding RNA sequencing and

analysis was performed as described (30). Briefly, 1ug total RNA was treated with RppH (NEB) for 1 hour at 37C to remove 5' triphosphates. RNA was then extracted by phenol chloroform and precipitated with 3M sodium acetate and ethanol. The resulting RNA was subjected to small RNA sequencing by Oxford Genomics. Fastq files were processed to remove adaptors using Fastx Toolkit. To obtain genome-wide alignments, all reads were aligned to the reference genome (WS280 as above) using Bowtie with the following parameters: -k 1 -v 0 --best to produce alignment sam files, which were further converted to bam files. Bedtools (37) intersect was used to assign reads to different genomic features. 22G-RNAs aligning to transposable elements were specifically interrogated by first selecting reads that were 22 nucleotides long and with a G as the 5' nucleotide using a custom perl script. These reads were then aligned to the consensus sequences for transposable elements in *C. elegans* downloaded from RepBase (https://www.girinst.org/repbase/) using Bowtie with parameters as above. Antisense reads corresponding to each transposable element were counted using Bedtools genomecov and quantified and visualised using R (version 4.2.2).

Forward genetic screen and DNA sequencing

Mutagenesis was carried out as described (38). Briefly, synchronized L4 *xnp-1(tm678)* worms were exposed to 50mM ethyl methanesulfonate (EMS) for 4 hours at room temperature, before being washed and placed on OP50 at 20°C until the F1 generation was gravid. At this point ~250 worms were singled onto new OP50 plates and kept at 25°C. Three fertile lines were isolated and one of these, *xnp-1(hcf3)* was pursued further. *xnp-1(hcf3)* was backcrossed three times to *xnp-1(tm678)* before genomic DNA was extracted (as described in TRF analysis) and sent for sequencing at Novogene (Cambridge Science Park). Potential mutations were mapped, and homozygous mutations identified using the MiModD (version 0.1.9) package developed in the Baumeister lab (https://mimodd.readthedocs.io/en/latest/index.html). The reference genome used was release WS283 (BioProject PRJNA13758).

Strains

All strains used and their availability are listed in Supplementary Table 1.

Results

XNP-1 loss phenocopies several aspects of ATRX loss in human cells

ATRX is an ATP-dependent chromatin remodelling enzymes that has been linked to multiple cellular functions (39). Loss of ATRX function results in ATRX syndrome in humans. ATRX syndrome mutations cluster in the ATPase domain as well as the ATRX-DNMT3-DNMT3L zinc finger (ADD) domain. The ADD domain binds H3K9me3-modified nucleosomes and is required to recruit human ATRX to heterochromatic regions *in vivo* (40). Curiously, the ADD domain is absent in ATRX homologues from lower eukaryotes such as *C. elegans* (Figure 1A). However, despite the lack of an ADD domain, *C. elegans* XNP-1 still interacts with H3K9me3-rich heterochromatin *in vivo* (Peter Meister, personal communication). Therefore, *C. elegans* XNP-1 possibly requires an accessory protein to recruit it to heterochromatin, as is the case with *D. melanogaster* XNP (41).

Simple genetic model systems have been instrumental in understanding ATP-dependent chromatin remodelling enzyme function (42). These enzymes are defined by the characteristic arrangement of motifs within their helicase-like ATPase domain (43). The importance of this ATPase domain is underscored by domain swap experiments, which show that the ATPase domain alone is sufficient to dictate the activity of a given chromatin remodelling complex (44). *C. elegans* XNP-1 and human ATRX share 70% sequence identity within the catalytic ATPase domain (45,46) suggesting that they likely perform similar functions. Thus, although the recruitment of XNP-1 and ATRX to chromatin may differ, it is likely that they perform similar chromatin remodelling reactions.

To study the effects of XNP-1 loss, we used a null mutant, *xnp-1(tm678)*. This allele contains a 673bp deletion in the *xnp-1* gene and additionally introduces a premature stop codon before the ATPase domain (45,46). As previously described, loss of ATRX in human cells has been linked to several phenotypes such as ALT activation and sensitivity to replication inhibitors (3,12,16). We decided to test whether loss of XNP-1 is sufficient to induce these phenotypes in *C. elegans*.

Human ALT positive tumours are defined clinically by the presence of C-circles: partially single-stranded circles of extrachromosomal telomeric DNA (26). The link between ALT and C-circles extends beyond humans as *C. elegans* ALT-like strains also show increased levels of C-circles (47-49). When synchronised populations of wildtype and *xnp-1* gravid adults were grown at 20°C, we did not detect any significant differences in C-circle levels. However, when grown at 25°C , a known stress condition, we saw significantly elevated C-circle levels in *xnp-1* compared to wildtype strains (Figure 1B). Indeed, the levels of C-circles measured in *xnp-1* were comparable to those seen in a *pot-2(tm1400)* strain. This is noteworthy as *pot-2* mutants are known to form long-term ALT survivors in a telomerase null background (48,49). These findings demonstrate that additional stresses, such as growth at 25°C, is needed to induces

C-circles in *xnp-1* mutant *C. elegans*. This is similar to the situation in humans where additional factors such as telomeric DNA double-strand breaks (50) or protein-DNA adducts (51) are required to induce robust levels of C-circles in ATRX mutant cells.

ATRX mutant cells show signs of replication stress and are sensitive to hydroxyurea (HU) (16). This drug induces replication stress by reducing the levels of deoxyribonucleotides in S phase and thereby uncoupling the movement of DNA polymerases from the replicative CMG helicase. We exposed worms to 10mM hydroxyurea from the L1 stage onwards and quantified its effect by measuring the brood size of the subsequent gravid adults two days later. It is noteworthy that even in the absence of HU, the brood size of *xnp-1* worms is significantly smaller than wildtype (Figure 1C). However, despite the difference in initial brood size, x*np-1* mutant worms displayed increased sensitivity to HU compared to wildtype. Exposure to 10mM HU resulted in an approximately two-fold decrease in wildtype brood size but a more than nine-fold drop in *xnp-1,* to the point where they almost became sterile (Figure 1C).

Besides having a low brood size, *xnp-1* animals are also slow to develop to adulthood. We observed that *xnp-1* mutants took significantly longer than wildtype to develop at every larval stage (Figure 1D). Moreover, the normally tight synchronicity of worm development was weakened in *xnp-1* (Supplementary Figure S1). Overall, our data shows that loss of XNP-1 in worms causes developmental defects as well as phenotypes similar to those seen in ATRX mutant human cells.

XNP-1 represses LTR retrotransposon expression dependent on H3.3

Loss of ATRX leads to increased expression of endogenous retroviruses (ERVs) in mammalian stem cells (12). To test whether increased expression of transposable elements (TE) occurred in *C. elegans* XNP-1 mutants, we used RNA-seq to look at genome-wide changes in TE expression in early embryos and arrested L1s (Figure 2A). A subset of RNA and DNA transposons were upregulated in *xnp-1* mutant embryos and L1 larvae (Figure 2A). The magnitude of transcriptional increase of TEs in *xnp-1* was modest but significant. The PIF-Harbinger family of DNA transposons showed the largest increase in transcription and we also observed significantly increased LTR retrotransposon expression (Figure 2A). Thus, XNP-1 plays a role in repressing TEs in *C. elegans*, similar to the role of ATRX in human cells.

Small interfering RNAs (siRNAs) play a major role in silencing TEs in most species (52). In *C. elegans*, 22G siRNAs play a particularly important role as they are the focal point at which all small RNA pathways converge (53). The same total RNA used for analysing mRNAs was also

subjected to small RNA enrichment and sequencing. We reckoned that one possible reason why the mRNA levels of TEs were higher in *xnp-1* was because the corresponding levels of 22G RNAs targeting these sequences were lower. However, there was no significant difference in 22G-RNA levels targeting TEs between wildtype and *xnp-1* (Supplementary Figure S2). Therefore, the siRNA pathway is unlikely to be involved in the mechanism of XNP-1-dependent TE silencing.

Retrotransposon silencing is especially important in the germline where it largely occurs via a chromatin-based based mechanism (54).We decided to focus on the expression of LTR retrotransposons in worms as they are the TEs most similar to ERVs in humans (55). Of the 17 families of LTR retrotransposons in *C. elegans* (55), we chose Cer10 as there were validated RT-qPCR primers to quantify its expression in worms (56). If the role of XNP-1 in retrotransposon silencing was linked to the germline, we would expect it to be most pronounced in developmental stages that contain a developed germline. Consistent with this, we find that Cer10 levels are not significantly different between *xnp*-1 mutants and wildtype at the embryo or L1 larval stages (Figure 2B). In contrast, by the time the germline is fully formed, at the L4 larval stage, we observed a significantly increased expression of Cer10 in *xnp-1* compared to wildtype, which persists in gravid adults (Figure 2B). This is consistent with previous data showing that the germline is prone to re-expression of LTR retrotransposons (57).

The recruitment of ATRX to retrotransposons in human cells is thought to silence them via two main mechanisms: either working with the H3K9 methyltransferase SETDB1 to promote H3K9Me3 spreading (10) or by incorporating H3.3 in a DAXX-dependent manner (11) (Figure 2C). The SETDB1 homolog, SET-25 is the only enzyme in *C. elegans* to catalyse H3K9 trimethylation (58). In line with previous data (56), we find that *set-25* single mutants show an increase in Cer10 expression (Figure 2C). However, when we looked at a *xnp-1; set-25* double mutant, we found that Cer10 levels were significantly higher than the *set-25* single mutant. This argues against a model where XNP-1 helps the spread of H3K9Me3 across retrotransposons and instead suggests that XNP-1 and SET-25 work in independent pathways to repress retrotransposons.

Worms do not contain a DAXX homolog but have five H3.3 homologs. Of these, HIS-72 shows the highest expression level and is the only homolog to be ubiquitously expressed (59). We did not observe any significant change in Cer10 expression in a *his-72* mutant. However, importantly, Cer10 transcript levels in an *xnp-1; his-72* double mutant were not significantly different to *his-72* (Figure 2C). This indicates that the repression of Cer10 by XNP-1 is

dependent on the replication-independent histone H3.3. Thus, XNP-1 likely silences Cer10 via histone replacement by H3.3 rather than by promoting H3K9Me3 formation.

We wondered whether the reduced brood size in *C. elegans xnp-1* mutants might be due to increased TE expression. If increased TE expression led to defects in brood size we would predict a significant negative correlation between the extent of retrotransposon expression in a particular *xnp-1* strain and its brood size. However, this was not the case (Supplementary Figure S3) indicating that it is unlikely that XNP-1's low brood size is a consequence of increased retrotransposon expression.

XNP-1 represses germline genes during early development in a manner distinct from SynMuvB

To look beyond repetitive elements, we analysed our polyA-enriched RNA sequencing data to compare protein-coding gene expression. Relatively few genes were differentially expressed in *xnp-1* compared to wildtype (Figure 3A and Supplementary Table 3). However, there was a large degree of overlap between differentially expressed genes that were upregulated in both embryos and L1s (Figure 3B). This was not observed in downregulated genes, suggesting that these are more influenced by developmental stage than genotype. Tissue enrichment analysis revealed that upregulated genes were very strongly associated with the germline (Figure 3C). This was surprising as neither embryos nor starved L1 larvae have a germline *per se*. Indeed, the only two germline progenitor cells (Z2 and Z3) in both these developmental stages are maintained in a transcriptionally quiescent state (60). Therefore, the increased level of germline expression we see in *xnp-1* embryos and L1 larvae was most likely to be due to ectopic expression of germline genes from somatic cells.

SynMuvB genes, sometimes referred to as the DREAM complex in other species (61), are the best described inhibitors of ectopic expression of germline genes in worms (62). A hallmark of synMuvB loss of function is increased expression of PGL-1, a P granule component found in all germ cells (62). We monitored *pgl-1* expression in *xnp-1* embryos or L1 larvae at 25°C, a condition in which the increase in PGL-1 in SynMuvB mutants is particularly pronounced (62). However, we found no significant increase in *pgl-1* mRNA levels in *xnp-1* compared to wildtype control (Figure 3D), suggesting that *xnp-1* may not be a SynMuvB gene. Indeed, when we looked beyond *pgl-1* to include all genes that become upregulated in *lin-35,* a key member of the SynMuvB family (62), we saw very little overlap with genes that become upregulated in *xnp-1* (Supplementary Figure S4). Another hallmark of SynMuvB mutants is that they display larval arrest at elevated temperatures, which is suppressed by RNAi of the H3K36

methyltransferase *mes-4* (62). However, loss of XNP-1 did not show a larval arrest phenotype at 25°C (Figure 3E) and *mes-4(RNAi)* did not increase the brood size *xnp-1* at 25°C (Figure 3F and Supplementary Figure S4). Altogether, these data show that XNP-1 and SynMuvB repress germline gene through distinct pathways.

Identification of gsox-1 as a suppressor of xnp-1 sterility at 25°C

C. elegans lacking XNP-1 become completely sterile within two generations when grown at 25°C (Figure 4A). We took advantage of this phenotype to identify the pathway through which *xnp-1* affects germline function by screening for suppressor mutations that restored fertility (Figure 4B). We decided to pursue one suppressor, *xnp-1(hcf3*), which was able to be maintained indefinitely at 25°C and whose brood size was significantly larger than the original *xnp-1(tm678)* strain (Figure 4C). Genome sequencing of this strain identified a locus on chromosome II containing G/C to A/T mutations characteristic of the ethyl methanesulfonate (EMS) mutagenesis used for our screen. Within this locus were four genes (*toe-1*, *mog-5*, *hgap-2* and *F33H1.4*) with mutations within their exons that resulted in protein-coding changes. Using CRISPR we independently introduced these EMS-induced mutations into each of these genes in a wildtype background and then crossed them into *xnp-1(tm678).* Strikingly, the only mutation that was able to suppress the sterility of *xnp-1* at 25°C was in *F33H1.4* (Figure 4D). This gene encodes for a nematode-specific transcription factor that is expressed within the germline and binds to the promoters of germline genes (63,64). Based on this, we re-named *F33H1.4* as *gsox-1* (germline suppressor of *xnp-1*).

Mutation of *gsox-1* suppresses the majority of misregulated germline expression in *xnp-1*

The connection of *gsox-1* to germline gene expression was intriguing as it was reminiscent of the ectopic germline gene expression seen in *xnp-1* mutants at 20°C (Figure 3A-C). We therefore decided to monitor the transcriptional changes in *xnp-1* animals grown at 25°C, a condition that results in rapid sterility. Wildtype and *xnp-1* mutant embryos were hatched at 20°C and then grown from L1 to L4 larvae at 25°C (Figure 5A). Transcriptomic analysis revealed that >6000 genes were significantly altered (FDR <0.05) compared to wildtype (Figure 5A). The number of differentially expressed genes and the magnitude of their fold changes were much greater than those observed in *xnp-1* strains grown at 20°C. (Figure 3A).

More importantly, when we looked at the identity of upregulated genes, we saw that they were again very strongly associated with the germline and the meiotic cell cycle (Figure 5B), similar

to what was seen at 20°C (Figure 3C). This meant that the same types of genes were affected in *xnp-1* worms at both 20°C (relatively normal physiology) and 25°C (profound developmental defects), but that the scale of the transcriptional changes was much larger at 25°C. Altogether, this suggested a link between germline upregulation and sterility in *xnp-1*. Down-regulated genes in *xnp-1* were associated with skin (hypodermis), the cuticle and moulting cycle (Fig 5B). Cuticle formation and the moulting cycle drive the growth of larval stages in *C. elegans*. It is possible that the disruption of these pathways contributes to the slow post-embryonic development of *xnp-1* animals (Figure 1D).

The *gsox-1* allele (R1192Q) we identified is likely to be a hypomorph as genome-wide RNAi experiments show that knockdown of *gsox*-1 is embryonic lethal (65). Indeed, we observed very few significant transcriptional changes in our single *gsox-1* mutant compared to wildtype (Figure 5C), indicating that this allele has little effect in a wildtype background. However, *gsox-1* mutation had a marked effect in an *xnp-1* background. The double *xnp-1; gsox-1* mutant had only 234 differentially expressed genes (DEGs) relative to wildtype (Figure 5D), which is much less than the 6781 DEGs between *xnp-1* and wildtype (Figure 5A). Mutation of *gsox-1* suppresses the majority but not totality of transcriptional changes seen in *xnp-1.* By changing the frame of reference to *xnp-1*, we observed that 63% of the DEGs between *xnp-1; gsox-1* double mutants and *xnp-1* overlap with DEGs between wildtype and *xnp-1* (Figure 5E). However, these data support the correlation between fertility defects in *xnp-1* and germline gene misregulation.

Altered cellular identity likely drives the developmental and telomeric phenotypes of *xnp-1*

We wondered whether *gsox-1* could also suppress other *xnp-1* phenotypes. Indeed, we found that *gsox-1* significantly suppressed the slow growth of *xnp-1* at 20°C. The *xnp-1; gsox-1* double mutant developed significantly faster than *xnp-1* at every larval stage (Figure 6A). This included the first larval stage, L1, which has no significant germline and only two germline progenitor cells per animal. Therefore, *gsox-1* can suppress *xnp-1* developmental phenotypes that are not directly linked to the germline. This view was strengthened when we looked at telomeric phenotypes. We observed that *gsox-1* partially suppressed the long telomere phenotype of *xnp-1* animals (Figure 6B) and almost completely supressed the increased C-circle levels at 25°C (Figure 6C). Similarly, *gsox-1* largely suppressed the upregulation of the Cer10 retrotransposon in *xnp-1* L4 larvae at 25°C (Figure 6D). Our data show that these seemingly unconnected *xnp-1* mutant phenotypes can all be rescued by mutation of a

germline transcription factor. This raises the possibility that these *xnp-1* phenotypes may be driven by altered cellular identity, specifically an increase in a germline-like cell fate.

The one *xnp*-1 phenotype which *gsox-1* could not suppress was sensitivity to HU (Figure 6E). We observed that the brood size of *xnp-1; gsox-1* double mutants on 10mM HU was no higher than *xnp-1*. Instead, *gsox-1* mutants were themselves mildly sensitive to HU. Thus, XNP-1's function in promoting replisome stability (unlike its other roles described here) is independent of GSOX-1 and thus unlikely to be linked to germline identity (Figure 6F).

Discussion

The ATRX gene has long been linked to human disease; it is the causal mutation behind ATRX syndrome (2) and is mutated in 90% of ALT positive cancers (3). In this study, we find that loss of the ATRX homologue, XNP-1 in *C. elegans*, leads to many of the same phenotypes in nematodes as in human cells including: de-repression of the ALT pathway, increased sensitivity to replication stress, de-repressed retrotransposon expression and disrupted development. Given that nematodes and humans diverged over a billion years ago (66), this suggests that ATRX syndrome and upregulation of ALT are caused by disruption of an evolutionarily ancient function of the ATRX gene.

One of the striking aspects of ATRX has been the multiple, seemingly disparate, phenotypes observed when it is mutated (39). It has not been clear how these various ATRX mutant phenotypes are linked, or whether they instead represent separate, independent functions. The fact that we see conservation of so many ATRX phenotypes in *C. elegans*, hints at a common underlying mechanism. This is strengthened by the fact that a single suppressor mutant of *xnp-1* sterility (*gsox-1)* also supressed other *xnp-1* phenotypes that it was not selected for. The simplest explanation for this observation is that these different *xnp-1* phenotypes are all distinct downstream consequences of a common upstream problem.

Given that GSOX-1 is a germline transcription factor (63,64), we propose that the initial trigger in *xnp-1* mutants is misregulated cellular identity, more specifically due to cells taking on a more germline-like fate. We believe that aberrant germline gene expression can explain the majority of phenotypes we see in *xnp-1*. Disruption of gene regulation within the germline is likely to impair its development, which can directly reduce fertility. We also see evidence of increased expression of germline genes in somatic cells. This could also explain the slower development of *xnp-1* animals during developmental stages (L1) where they do not contain a germline *per se* (Figure 1D). Indeed, mutation of *xnp-1* influences gene expression as early as embryogenesis. Retrotransposon expression, normally increases in mammals and *C.*

elegans at the onset of meiosis (67,68). This suggests that some of the retrotransposon misexpression observed in XNP-1/ATRX mutants may be an indirect effect of the increased presence of meiotic factors. The Cer5 and Cer8 retrotransposons are known to be most highly expressed in the germline (57). Our data is consistent with Cer10 expression being increased in *xnp-1* animals because of increased germline identity.

The telomeric phenotypes in *xnp-1* mutants could also be explained by increased germline-like gene expression. Amongst the genes whose expression increases significantly in *xnp-1* L4 larvae at 25°C are the telomerase catalytic subunit (*trt-1*), the shelterin components (*pot-1* and *pot-2)* as well as several DNA repair proteins (Supplementary Table 3). ALT is driven by activation of inappropriate DNA repair pathways at dysfunctional telomeres (5). Thus, aberrant germline gene expression could pre-dispose telomeres to inappropriate lengthening and dysfunction. That the ALT-like phenotype of *xnp-1* animals can be supressed by mutation of a germline transcription factor is intriguing. Most models of ALT invoke aberrant DNA repair at a damaged replication fork (69). However, our data points to a more important role for regulated gene expression than fork stability in *xnp-1* mutants that undergo ALT. Instead, we see that *gsox-1* does not suppress the HU sensitivity of *xnp-1.* This suggests that the fork stability phenotype of *xnp-1* (HU sensitivity) is not linked to other *xnp-1* phenotypes. At first glance this is surprising as expressing meiotic proteins in human cells is sufficient to induce replication stress (70). It is possible that XNP-1 plays a more direct role at the replication fork. Human ATRX associates physically with the MCM complex (71), perhaps pointing to a mechanism by which it travels with the fork to promote replisome stability.

Despite this, we believe that safeguarding correct cellular identity is the key developmental function of XNP-1. Indeed, ATRX is also involved in maintaining cellular identity in mammalian cells. Loss of ATRX leads to altered differentiation of mouse oligodendrocyte and human mesenchymal progenitor cells (72,73). Moreover, a mouse model of ATRX syndrome shows signs of early developmental defects (74). Thus, loss of ATRX leads to changes in stem cell identity and a reduction in their ability to resist differentiation into inappropriate cell fates. This connection might also be relevant to ATRX's role in supressing ALT. A recent study shows that induction of differentiation is a key trigger for ALT in ATRX-deficient stem cells (75).

ALT is more frequent in childhood cancers (76). In these cancers, cells have less time to accumulate additional cancer-driving mutations, suggesting that changes in their initial cellular identity may play a more significant role in their eventual malignant transformation. We suggest that ATRX plays a key role restricting cellular identity, and that loss of ATRX leads to the inappropriate expression of genes linked to other cell fates. In *C. elegans*, this manifests as a large increase in germline gene expression (Figure 3C, 5B). However, this may be

because germline is hypothesized to be the default cell fate in *C. elegans* (77). Thus, XNP-1 may be required to restrict cell fate rather than specifically repress germline genes *per se*. This would be consistent with GSOX-1 not being conserved beyond nematodes whereas XNP-1 is.

The precise mechanism by which ATRX safeguards cellular identity remains unclear. As a member of the ATP-dependent family of chromatin remodeling enzymes, ATRX may play a crucial role in maintaining the integrity of the chromatin template. Consistent with this, XNP-1 represses Cer10 in a mechanism dependent on histone H3.3 (Figure 2C). Interestingly, ATRX is also involved in functions beyond histone deposition (78). A recent study found that most genes whose expression increased following ATRX mutation did not show a corresponding increase in H3K4 methylation or chromatin accessibility at their promoters (72). Future research will hopefully reveal the mechanisms through which ATRX regulates gene expression during development.

Figure One – Loss of *C. elegans* XNP-1 mimics several phenotypes of human ATRX mutation

A) XNP-1 is the *C. elegans* homologue of human ATRX. The ADD (ATRX- DNMT3-DNMT3L) domain is not found in ATRX homologues of lower eukaryotes, but the catalytic ATPase domain is highly conserved between humans and worms. (B) *xnp-1* mutant worms display elevated C-circle levels at 25°C. Representative blot from a C-circle assay both at 20°C or 25°C with and without the polymerase Phi29. Quantification of data from five independent C-circle assay experiments using worms synchronised to L1s and then grown for 72 hours at either 20°C or 25°C. Normalised values are presented as percent of *pot-2(tm1400)* signal and the mean across different experiments is shown. Error bars represent the standard deviation from the mean, ns = not significant, * = $p < 5 \times 10^{-2}$ (two-tailed Mann-Whitney U test). (C) *xnp-1* mutant worms have lower brood sizes than wildtype and are sensitive to hydroxyurea (HU). Synchronised L1 wildtype or *xnp-1(tm678)* larvae were placed on either OP50 plates or OP50 plates supplemented with 10 mM HU and maintained at 20°C. The brood sizes of 12 - 20 adults per genotype are displayed as box and whisker plots, ** = $p < 1 \times 10^{-3}$, *** = $p < 1 \times 10^{-5}$ (Welch's test). (D) *xnp-1* mutant worms develop slower than wildtype. Synchronised L1 wildtype or *xnp-1(tm678)* larvae were singled into individual wells and imaged over two days to monitor growth. Box and whisker plots of data from at least 70 worms per genotype indicating the total time to complete each larval stage, *** = $p < 1 \times 10^{-6}$ (two tailed T-test).

Figure Two – XNP-1 represses LTR retrotransposon expression during development in a manner dependent on histone H3.3

(A) Loss of XNP-1 leads to a modest upregulation of repetitive elements in both embryos and L1 larvae. Embryos and L1 animals were grown at 20°C and processed for polyA-RNA sequencing. The expression of the indicated transposons is displayed as a heatmap for each biological replicate. The adjacent graph displays the $\log_2$ fold change (FC) plotted with the data normalised to the *xnp-1(tm678)* strain in a stage-specific manner and to the protein-coding gene reads on the x-axis compared to the $-\log_2$ of the false discovery rate (FDR). Significantly expressed transposon families are marked in purple. (B) The LTR retrotransposon *Cer10* is de-repressed most strongly in larval stages of *xnp-1* animals that contain a germline. RT-qPCR of *Cer10* with *tba-1* used as the reference in wildtype and *xnp-1(tm678)* animals with *Cer10* expression normalised to wildtype. The mean and standard deviation is shown for three biological replicates. ns = not significant, * = $p \leq 5 \times 10^{-2}$, ** = $p \leq 10 \times 10^{-3}$ (two tailed T-test). (C) Model for how ATRX represses endogenous retroviruses in humans and mice. The graph to the right looks at Cer10 expression by RT-qPCR as in panel B except that in all cases, worms were synchronised to the gravid adult stage. The ability of *xnp-1* mutation to increase *Cer10* expression is independent of H3 K9me3 (SET-25) but dependent on H3.3 (HIS-72). ns = not significant , * = $p \leq 5 \times 10^{-2}$ (two tailed T-test).

Figure Three – XNP-1 prevents ectopic germline expression in a different manner to the SynMuvB pathway

(A) Loss of XNP-1 leads to changes to a modest number of differentially expressed genes (DEGs) in both embryos and L1s grown at 20°C. Genes whose expression is significantly altered in *xnp-1* are marked in red within the MA plot, false discovery rate (FDR) < 0.05. (B) Upregulated genes in *xnp-1* are more likely to behave similarly in both embryo and L1 samples compared to downregulated genes. (C) Analysis of genes that are significantly upregulated in *xnp-1* samples (from part A) indicates that they are normally expressed within the germline and typically have functions in DNA repair and meiosis. (D) *pgl-1* is not upregulated in *xnp-1* mutants. RT-qPCR of *pgl-1* normalised to *tba-1* in wildtype and *xnp-1(tm678)* embryos and L1 larvae grown at 25°C. The mean and standard deviation is shown for three biological replicates. ns = $p > 5 \times 10^{-2}$ (two tailed T-test). (E) *xnp-1* animals (in contrast to *lin-35*) do not show L1 larval arrest at 25°C. Synchronised L4s were placed at 25°C and left to produce offspring. The number of offspring scored for an arrest phenotype and the percentage of L1 larval arrest is indicated. (F) The low brood size of *xnp-1* animals at 25°C is not rescued by *mes-4(RNAi).* Synchronised *xnp-1* L1s were placed at 25°C onto control or *mes-4(RNAi)* plates. Once they developed to gravid adults, they were singled and the brood size of those adults (F1) scored. Box and whisker plots from 15 - 20 adults are displayed, ns = not significant.

Figure Four – Identification of *gsox-1* as a suppressor of *xnp-1* sterility at 25°C

(A) Loss of XNP-1 leads to rapid sterility at 25°C. Synchronised *xnp-1* L1s were placed at 25°C onto OP50 plates. Once they developed to gravid adults, they were singled and the brood size of those adults (F1) scored. The subsequent generation at 25°C (F2) were sterile. In contrast, *xnp-1* worms can be maintained indefinitely at 20°C. Box and whisker plots from 15 - 20 adults are displayed, ** = $p \leq 10 \times 10^{-4}$ (two tailed T-test). (B) Schematic of forward genetic screen to isolate *xnp-1* suppressors. (C) Identification of *xnp-1(hcf3)* as a weak suppressor of temperature induced sterility. Box and whisker plot of the number of F2 progeny of wildtype, the original *xnp-1(tm678)* strain and the suppressor line *xnp-1(hcf3)*, * = $p < 5 \times 10^{-2}$ (Welch's test), n > 19 for all genotypes. (D) *gsox-1* is the causal gene behind the fertility of *xnp-1(hcf3)* at 25°C. Box and whisker plot of the number of F2 progeny of wildtype, *xnp-1(tm678)*, *toe-1(syb6057)*, *mog-5(syb5976)*, *hgap-2(syb7188)*, *gsox-1(syb7245)* and double mutants of these with *xnp-1(tm678)*, *** = $p < 1 \times 10^{-4}$ (Welch's test), n > 10 for all genotypes.

Figure Five – Mutation of *gsox-1* suppresses the majority of transcriptional changes in *xnp-1*

(A) Schematic of growth conditions of wildtype and *xnp-1* worms prior to processing for polyA-RNA sequencing. MA plot illustrating that a very large number of transcripts are misrergulated in *xnp-1* relative to wildtype. Differentially expressed genes (DEGs) were defined using a false discovery rate (FDR) of less than 0.05 and a relative fold change (FC) greater than 2. (B) Upregulated and downregulated DEGs are associated with distinct tissues and molecular functions. Upregulated DEGs in *xnp-1* are typically expressed in the germline (similar to what is seen in Figure 3D). In contrast, downregulated DEGs are typically expressed in males and the hypodermis. Enriched gene ontology terms for up- and downregulated DEGs are ordered by enrichment over the background set (all genes detected) and colour coded by their FDR. Upregulated DEGs in *xnp-1* are typically associated with functions in DNA repair and meiotic cell cycle (again similar to what is seen in Figure 3D). In contrast, downregulated DEGs have functions in cuticle formation and moulting. (C) Mutation of *gsox-1* by itself does not cause significant transcriptional disruption. (D) Mutation of *gsox-1* reduces the number of DEGs in an *xnp-1* background. MA plot, as described in part A, showing that the number of DEGs in *xnp-1; gsox-1* is more than 25-fold lower than in *xnp-1*. (E) Loss of *gsox-1* does not suppresses all of the transcriptional changes in *xnp-1*. Venn diagram depicting DEGs in *xnp-1; gsox-1* vs *xnp-1* compared to wildtype vs *xnp-1*. Most but not all DEGs between these sets overlap, indicating that that the majority of transcripts that become misregulated when *xnp-1* is mutated (wildtype -> *xnp-1*) are 'corrected ' when *gsox-1* is subsequently mutated (*xnp-1* -> *xnp-1; gsox-1*).

Figure Six – The developmental and telomeric phenotypes of *xnp-1* are linked and likely due to altered cellular identity

(A) *gsox-1* partially recues the slow development phenotype of *xnp-1* at 20°C. Synchronised L1 larvae were singled into individual wells and imaged over two days to monitor growth. Data for wildtype and

xnp-1 as in Figure 1D. Box and whisker plots of data from at least 70 worms per genotype indicating the total time to complete each larval stage, *** = $p < 1 \times 10^{-6}$ (two tailed T-test). (B) *xnp-1* worms have longer telomeres than wildtype which is partially suppressed by *gsox-1*. Southern blot, probed with a telomere-specific sequence, of an asynchronous population of the indicated strains grown at 20°C. (C) *gsox-1* recues the elevated C-circle levels of *xnp-1* at 25°C. Quantification of data from five independent C-circle assay experiments using worms synchronised to L1s and then grown for 72 hours at either 20°C or 25°C. Data for wildtype and *xnp-1* as in Figure 1B , * = $p < 5 \times 10^{-2}$ (two-tailed Mann-Whitney U test) (D) *gsox-1* partially recues the increased *Cer10* expression observed in *xnp-1.* RT-qPCR of *Cer10* levels normalised to *tba-1* in L4 stage animals of the indicated genotypes grown at 25°C from L1 onwards. The level of wildtype Cer10 expression is set to 1. The graph displays the mean and standard deviation from three biological replicates. * = $p \leq 5 \times 10^{-2}$, ** = $p \leq 5 \times 10^{-3}$ (two tailed T-test). (E) *gsox-1* does not supress the HU sensitivity of *xnp-1.* Synchronised L1 larvae of the indicated genotypes were placed on either OP50 plates or OP50 plates supplemented with 10 mM HU and maintained at 20°C. Data for wildtype and *xnp-1* as in Figure 1C. The brood sizes of 12 - 20 adults per genotype are displayed as box and whisker plots, ns = not significant. (F) Model for how XNP-1 works in *C. elegans*. XNP-1 represses the ectopic expression of germline genes in a manner distinct form the SynMuvB pathway. The inappropriate activation of a germline-like fate is the likely cause of the developmental, telomeric and retrotransposon repression defects seen in in *xnp-1* as all these phenotypes are suppressed by *gsox-1*.

References

1. Amberger, J.S., Bocchini, C.A., Schiettecatte, F., Scott, A.F. and Hamosh, A. (2015) OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic Acids Res*, **43**, D789-798.
2. Gibbons, R.J., Picketts, D.J., Villard, L. and Higgs, D.R. (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell*, **80**, 837-845.
3. Lovejoy, C.A., Li, W., Reisenweber, S., Thongthip, S., Bruno, J., de Lange, T., De, S., Petrini, J.H., Sung, P.A., Jasin, M. *et al.* (2012) Loss of ATRX, Genome Instability, and an Altered DNA Damage Response Are Hallmarks of the Alternative Lengthening of Telomeres Pathway. *PLoS genetics*, **8**, e1002772.
4. Zhang, J.M., Yadav, T., Ouyang, J., Lan, L. and Zou, L. (2019) Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. *Cell reports*, **26**, 955-968 e953.
5. Dilley, R.L., Verma, P., Cho, N.W., Winters, H.D., Wondisford, A.R. and Greenberg, R.A. (2016) Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature*, **539**, 54-58.
6. Clynes, D., Jelinska, C., Xella, B., Ayyub, H., Scott, C., Mitson, M., Taylor, S., Higgs, D.R. and Gibbons, R.J. (2015) Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nature communications*, **6**, 7538.
7. Lewis, P.W., Elsaesser, S.J., Noh, K.M., Stadler, S.C. and Allis, C.D. (2010) Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 14075-14080.
8. Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X. *et al.* (2010) Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell*, **140**, 678-691.
9. Truch, J., Downes, D.J., Scott, C., Gur, E.R., Telenius, J.M., Repapi, E., Schwessinger, R., Gosden, M., Brown, J.M., Taylor, S. *et al.* (2022) The chromatin remodeller ATRX facilitates diverse nuclear processes, in a stochastic manner, in both heterochromatin and euchromatin. *Nature communications*, **13**, 3485.
10. Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F.J. and Schotta, G. (2015) Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep*, **16**, 836-850.
11. Elsasser, S.J., Noh, K.M., Diaz, N., Allis, C.D. and Banaszynski, L.A. (2015) Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*, **522**, 240-244.
12. Robbez-Masson, L., Tie, C.H.C., Conde, L., Tunbak, H., Husovsky, C., Tchasovnikarova, I.A., Timms, R.T., Herrero, J., Lehner, P.J. and Rowe, H.M. (2018) The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res*, **28**, 836-845.
13. Law, M.J., Lower, K.M., Voon, H.P., Hughes, J.R., Garrick, D., Viprakasit, V., Mitson, M., De Gobbi, M., Marra, M., Morris, A. *et al.* (2010) ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell*, **143**, 367-378.
14. Ramamoorthy, M. and Smith, S. (2015) Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells. *Cancer Cell*, **28**, 357-369.
15. Ritchie, K., Seah, C., Moulin, J., Isaac, C., Dick, F. and Berube, N.G. (2008) Loss of ATRX leads to chromosome cohesion and congression defects. *J Cell Biol*, **180**, 315-324.

16. Clynes, D., Jelinska, C., Xella, B., Ayyub, H., Taylor, S., Mitson, M., Bachrati, C.Z., Higgs, D.R. and Gibbons, R.J. (2014) ATRX dysfunction induces replication defects in primary mouse cells. *PLoS One*, **9**, e92915.
17. Leung, J.W., Ghosal, G., Wang, W., Shen, X., Wang, J., Li, L. and Chen, J. (2013) Alpha thalassemia/mental retardation syndrome X-linked gene product ATRX is required for proper replication restart and cellular resistance to replication stress. *J Biol Chem*, **288**, 6342-6350.
18. Oppel, F., Tao, T., Shi, H., Ross, K.N., Zimmerman, M.W., He, S., Tong, G., Aster, J.C. and Look, A.T. (2019) Loss of atrx cooperates with p53-deficiency to promote the development of sarcomas and other malignancies. *PLoS Genet*, **15**, e1008039.
19. Berube, N.G., Mangelsdorf, M., Jagla, M., Vanderluit, J., Garrick, D., Gibbons, R.J., Higgs, D.R., Slack, R.S. and Picketts, D.J. (2005) The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. *J Clin Invest*, **115**, 258-267.
20. Bassett, A.R., Cooper, S.E., Ragab, A. and Travers, A.A. (2008) The chromatin remodelling factor dATRX is involved in heterochromatin formation. *PLoS One*, **3**, e2099.
21. Medina, C.F., Mazerolle, C., Wang, Y., Berube, N.G., Coupland, S., Gibbons, R.J., Wallace, V.A. and Picketts, D.J. (2009) Altered visual function and interneuron survival in Atrx knockout mice: inference for the human syndrome. *Hum Mol Genet*, **18**, 966-977.
22. Horner, M.A., Quintin, S., Domeier, M.E., Kimble, J., Labouesse, M. and Mango, S.E. (1998) pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in Caenorhabditis elegans. *Genes Dev*, **12**, 1947-1952.
23. McGhee, J.D., Fukushige, T., Krause, M.W., Minnema, S.E., Goszczynski, B., Gaudet, J., Kohara, Y., Bossinger, O., Zhao, Y., Khattra, J. *et al.* (2009) ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. *Dev Biol*, **327**, 551-565.
24. Wang, J.T. and Seydoux, G. (2013) Germ cell specification. *Adv Exp Med Biol*, **757**, 17-39.
25. Bender, L.B., Cao, R., Zhang, Y. and Strome, S. (2004) The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. *Curr Biol*, **14**, 1639-1643.
26. Henson, J.D., Cao, Y., Huschtscha, L.I., Chang, A.C., Au, A.Y., Pickett, H.A. and Reddel, R.R. (2009) DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol*, **27**, 1181-1185.
27. Meeuse, M.W., Hauser, Y.P., Morales Moya, L.J., Hendriks, G.J., Eglinger, J., Bogaarts, G., Tsiairis, C. and Grosshans, H. (2020) Developmental function and state transitions of a gene expression oscillator in Caenorhabditis elegans. *Mol Syst Biol*, **16**, e9498.
28. Olmedo, M., Geibel, M., Artal-Sanz, M. and Merrow, M. (2015) A High-Throughput Method for the Analysis of Larval Developmental Phenotypes in Caenorhabditis elegans. *Genetics*, **201**, 443-448.
29. Nahar, S., Morales Moya, L.J., Brunner, J., Hendriks, G.J., Towbin, B., Hauser, Y.P., Brancati, G., Gaidatzis, D. and Grosshans, H. (2024) Dynamics of miRNA accumulation during C. elegans larval development. *Nucleic Acids Res*, **52**, 5336-5355.
30. Beltran, T., Pahita, E., Ghosh, S., Lenhard, B. and Sarkies, P. (2021) Integrator is recruited to promoter-proximally paused RNA Pol II to generate Caenorhabditis elegans piRNA precursors. *EMBO J*, **40**, e105564.
31. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10-12.
32. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.
33. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. *et al.* (2021) Twelve years of SAMtools and BCFtools. *Gigascience*, **10**.
34. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923-930.

35. Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139-140.
36. Tang, D., Chen, M., Huang, X., Zhang, G., Zeng, L., Zhang, G., Wu, S. and Wang, Y. (2023) SRplot: A free online platform for data visualization and graphing. *PLoS One*, **18**, e0294236.
37. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841-842.
38. Brenner, S. (1974) The genetics of Caenorhabditis elegans. *Genetics*, **77**, 71-94.
39. Aguilera, P. and Lopez-Contreras, A.J. (2023) ATRX, a guardian of chromatin. *Trends Genet*, **39**, 505-519.
40. Eustermann, S., Yang, J.C., Law, M.J., Amos, R., Chapman, L.M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R.J., Rhodes, D. *et al.* (2011) Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nature structural & molecular biology*, **18**, 777-782.
41. Lopez-Falcon, B., Meyer-Nava, S., Hernandez-Rodriguez, B., Campos, A., Montero, D., Rudino, E., Vazquez, M., Zurita, M. and Valadez-Graham, V. (2014) Characterization of the Drosophila group ortholog to the amino-terminus of the alpha-thalassemia and mental retardation X-Linked (ATRX) vertebrate protein. *PLoS One*, **9**, e113182.
42. Nasmyth, K., Stillman, D. and Kipling, D. (1987) Both positive and negative regulators of HO transcription are required for mother-cell-specific mating-type switching in yeast. *Cell*, **48**, 579-587.
43. Flaus, A., Martin, D.M., Barton, G.J. and Owen-Hughes, T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res*, **34**, 2887-2905.
44. Fan, H.Y., Trotter, K.W., Archer, T.K. and Kingston, R.E. (2005) Swapping function of two chromatin remodeling complexes. *Mol Cell*, **17**, 805-815.
45. Cardoso, C., Couillault, C., Mignon-Ravix, C., Millet, A., Ewbank, J.J., Fontes, M. and Pujol, N. (2005) XNP-1/ATR-X acts with RB, HP1 and the NuRD complex during larval development in C. elegans. *Developmental biology*, **278**, 49-59.
46. Bender, A.M., Wells, O. and Fay, D.S. (2004) lin-35/Rb and xnp-1/ATR-X function redundantly to control somatic gonad development in C. elegans. *Developmental biology*, **273**, 335-349.
47. Lackner, D.H., Raices, M., Maruyama, H., Haggblom, C. and Karlseder, J. (2012) Organismal propagation in the absence of a functional telomerase pathway in Caenorhabditis elegans. *EMBO J.*, **31**, 2024-2033.
48. Cheng, C., Shtessel, L., Brady, M.M. and Ahmed, S. (2012) Caenorhabditis elegans POT-2 telomere protein represses a mode of alternative lengthening of telomeres with normal telomere lengths. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 7805-7810.
49. Shtessel, L., Lowden, M.R., Cheng, C., Simon, M., Wang, K. and Ahmed, S. (2013) Caenorhabditis elegans POT-1 and POT-2 repress telomere maintenance pathways. *G3*, **3**, 305-313.
50. Lovejoy, C.A., Takai, K., Huh, M.S., Picketts, D.J. and de Lange, T. (2020) ATRX affects the repair of telomeric DSBs by promoting cohesion and a DAXX-dependent activity. *PLoS Biol*, **18**, e3000594.
51. Rose, A.M., Goncalves, T., Cunniffe, S., Geiller, H.E.B., Kent, T., Shepherd, S., Ratnaweera, M., O'Sullivan, R.J., Gibbons, R.J. and Clynes, D. (2023) Induction of the alternative lengthening of telomeres pathway by trapping of proteins on DNA. *Nucleic Acids Res*, **51**, 6509-6527.
52. Gebert, D. and Rosenkranz, D. (2015) RNA-based regulation of transposon expression. *Wiley Interdiscip Rev RNA*, **6**, 687-708.
53. Almeida, M.V., Andrade-Navarro, M.A. and Ketting, R.F. (2019) Function and Evolution of Nematode RNAi Pathways. *Noncoding RNA*, **5**.
54. Molaro, A. and Malik, H.S. (2016) Hide and seek: how chromatin-based pathways silence retroelements in the mammalian germline. *Curr Opin Genet Dev*, **37**, 51-58.

55. Ganko, E.W., Fielman, K.T. and McDonald, J.F. (2001) Evolutionary history of Cer elements and their impact on the C. elegans genome. *Genome Res*, **11**, 2066-2074.
56. Zeller, P., Padeken, J., van Schendel, R., Kalck, V., Tijsterman, M. and Gasser, S.M. (2016) Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. *Nat Genet*, **48**, 1385-1395.
57. Ni, J.Z., Kalinava, N., Mendoza, S.G. and Gu, S.G. (2018) The spatial and temporal dynamics of nuclear RNAi-targeted retrotransposon transcripts in Caenorhabditis elegans. *Development*, **145**.
58. Towbin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P. and Gasser, S.M. (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell*, **150**, 934-947.
59. Delaney, K., Mailler, J., Wenda, J.M., Gabus, C. and Steiner, F.A. (2018) Differential Expression of Histone H3.3 Genes and Their Role in Modulating Temperature Stress Response in Caenorhabditis elegans. *Genetics*, **209**, 551-565.
60. Wong, M.M., Belew, M.D., Kwieraga, A., Nhan, J.D. and Michael, W.M. (2018) Programmed DNA Breaks Activate the Germline Genome in Caenorhabditis elegans. *Dev Cell*, **46**, 302-315 e305.
61. Hoareau, M., Rincheval-Arnold, A., Gaumer, S. and Guenal, I. (2024) DREAM a little dREAM of DRM: Model organisms and conservation of DREAM-like complexes: Model organisms uncover the mechanisms of DREAM-mediated transcription regulation. *Bioessays*, **46**, e2300125.
62. Petrella, L.N., Wang, W., Spike, C.A., Rechtsteiner, A., Reinke, V. and Strome, S. (2011) synMuv B proteins antagonize germline fate in the intestine and ensure C. elegans survival. *Development*, **138**, 1069-1079.
63. Kudron, M., Gevirtzman, L., Victorsen, A., Lear, B.C., Gao, J., Xu, J., Samanta, S., Frink, E., Tran-Pearson, A., Huynh, C. *et al.* (2024) Binding profiles for 954 Drosophila and C. elegans transcription factors reveal tissue specific regulatory relationships. *bioRxiv*.
64. Cao, W., Fan, Q., Amparado, G., Begic, D., Godini, R., Gopal, S. and Pocock, R. (2024) A nucleic acid binding protein map of germline regulation in Caenorhabditis elegans. *Nature communications*, **15**, 6884.
65. Maeda, I., Kohara, Y., Yamamoto, M. and Sugimoto, A. (2001) Large-scale analysis of gene function in Caenorhabditis elegans by high-throughput RNAi. *Curr Biol*, **11**, 171-176.
66. Wang, D.Y., Kumar, S. and Hedges, S.B. (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc Biol Sci*, **266**, 163-171.
67. van der Heijden, G.W. and Bortvin, A. (2009) Transient relaxation of transposon silencing at the onset of mammalian meiosis. *Epigenetics*, **4**, 76-79.
68. Dennis, S., Sheth, U., Feldman, J.L., English, K.A. and Priess, J.R. (2012) C. elegans germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS Pathog*, **8**, e1002591.
69. Lu, R. and Pickett, H.A. (2022) Telomeric replication stress: the beginning and the end for alternative lengthening of telomeres cancers. *Open Biol*, **12**, 220011.
70. Lingg, L., Rottenberg, S. and Francica, P. (2022) Meiotic Genes and DNA Double Strand Break Repair in Cancer. *Front Genet*, **13**, 831620.
71. Teng, Y.C., Sundaresan, A., O'Hara, R., Gant, V.U., Li, M., Martire, S., Warshaw, J.N., Basu, A. and Banaszynski, L.A. (2021) ATRX promotes heterochromatin formation to protect cells from G-quadruplex DNA-mediated stress. *Nature communications*, **12**, 3887.
72. Fang, Y., Barrows, D., Dabas, Y., Carroll, T.S., Singer, S., Tap, W.D. and Nacev, B.A. (2024) ATRX guards against aberrant differentiation in mesenchymal progenitor cells. *Nucleic Acids Res*, **52**, 4950-4968.

73. Rowland, M.E., Jiang, Y., Shafiq, S., Ghahramani, A., Pena-Ortiz, M.A., Dumeaux, V. and Berube, N.G. (2023) Systemic and intrinsic functions of ATRX in glial cell fate and CNS myelination in male mice. *Nature communications*, **14**, 7090.
74. Quesnel, K.M., Martin-Kenny, N. and Berube, N.G. (2023) A mouse model of ATRX deficiency with cognitive deficits and autistic traits. *J Neurodev Disord*, **15**, 39.
75. Turkalo, T.K., Maffia, A., Schabort, J.J., Regalado, S.G., Bhakta, M., Blanchette, M., Spierings, D.C.J., Lansdorp, P.M. and Hockemeyer, D. (2023) A non-genetic switch triggers alternative telomere lengthening and cellular immortalization in ATRX deficient cells. *Nature communications*, **14**, 939.
76. Claude, E. and Decottignies, A. (2020) Telomere maintenance mechanisms in cancer: telomerase, ALT or lack thereof. *Curr Opin Genet Dev*, **60**, 1-8.
77. Cockrum, C.S. and Strome, S. (2022) Maternal H3K36 and H3K27 HMTs protect germline development via regulation of the transcription factor LIN-15B. *Elife*, **11**.
78. Scott, W.A., Dhanji, E.Z., Dyakov, B.J.A., Dreseris, E.S., Asa, J.S., Grange, L.J., Mirceta, M., Pearson, C.E., Stewart, G.S., Gingras, A.C. *et al.* (2021) ATRX proximal protein associations boast roles beyond histone deposition. *PLoS Genet*, **17**, e1009909.

Figure 1

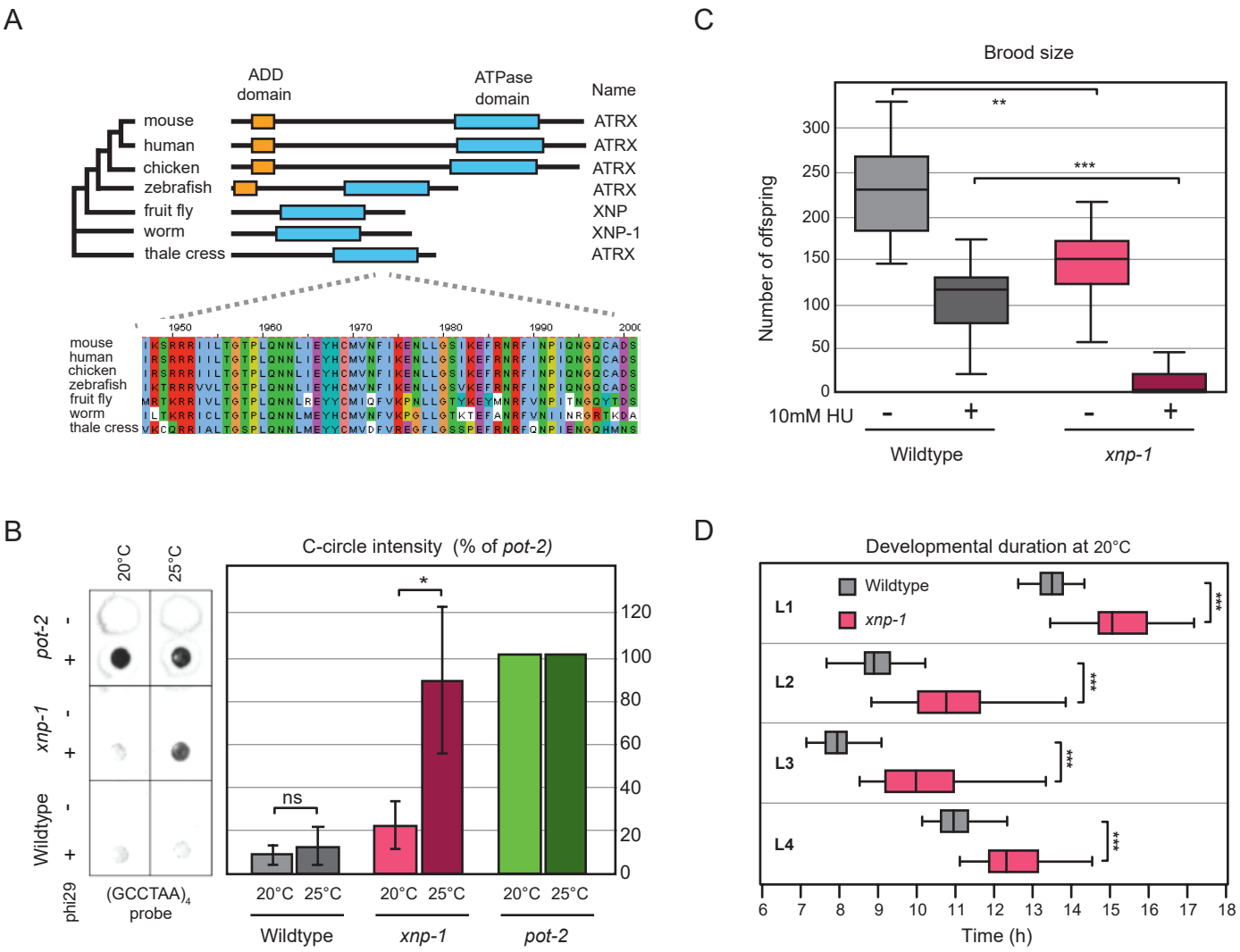
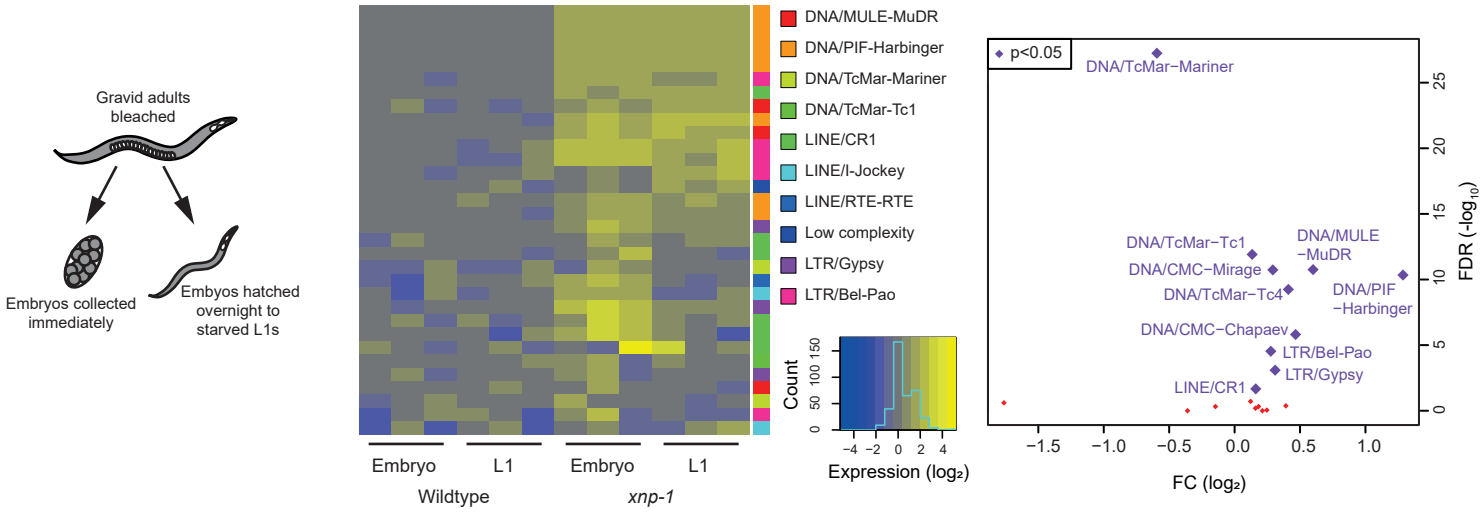
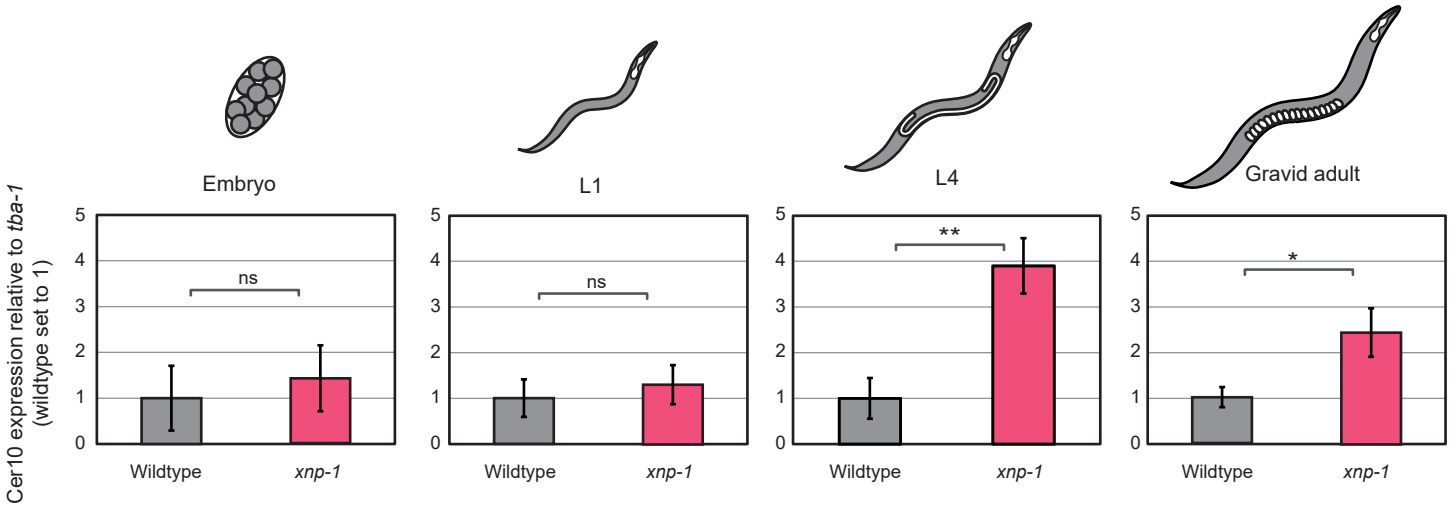

A
ADD domain
ATPase domain
Name
mouse
human
chicken
zebrafish
fruit fly
worm
thale cress
ATRX
ATRX
ATRX
ATRX
XNP
XNP-1
ATRX
C
Brood size
Number of offspring
10mM HU
Wildtype
xnp-1
B
C-circle intensity (% of pot-2)
20°C
25°C
pot-2
xnp-1
Wildtype
phi29
(GCCTAA)4 probe
ns
Wildtype
xnp-1
pot-2
D
Developmental duration at 20°C
Wildtype
xnp-1
L1
L2
L3
L4
Time (h)

Figure 2

A


Gravid adults bleached
Embryos collected immediately
Embyos hatched overnight to starved L1s
DNA/MULE-MuDR
DNA/PIF-Harbinger
DNA/TcMar-Mariner
DNA/TcMar-Tc1
LINE/CR1
LINE/I-Jockey
LINE/RTE-RTE
Low complexity
LTR/Gypsy
LTR/Bel-Pao
Count
Expression (log2)
Embryo
L1
Embryo
L1
Wildtype
xnp-1
p<0.05
DNA/TcMar−Mariner
DNA/TcMar−Tc1
DNA/MULE −MuDR
DNA/CMC−Mirage
DNA/TcMar−Tc4
DNA/PIF −Harbinger
DNA/CMC−Chapaev
LTR/Bel-Pao
LTR/Gypsy
LINE/CR1
FDR (-log10)
FC (log2)

B


Embryo
L1
L4
Gravid adult
Cer10 expression relative to tba-1 (wildtype set to 1)
ns
ns
**
*
Wildtype
xnp-1

C

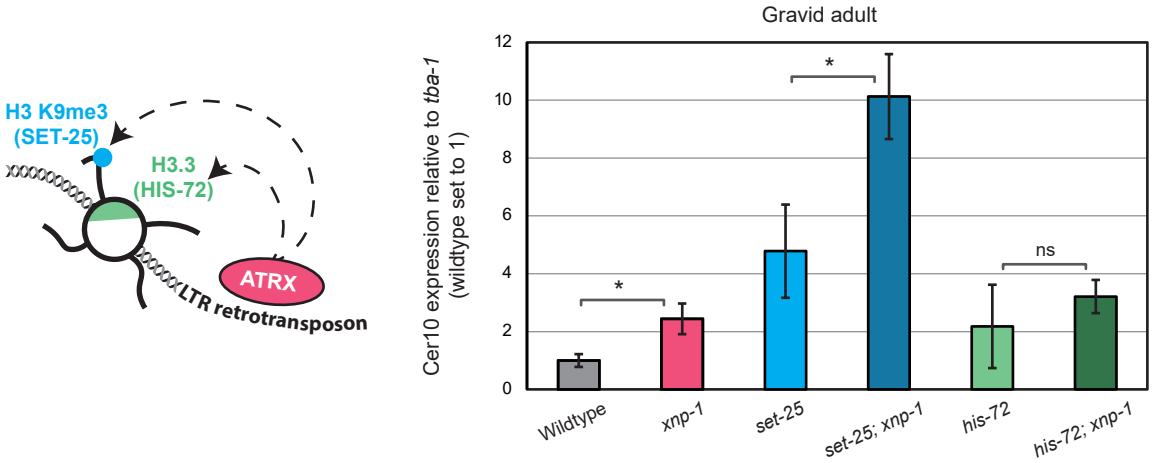

H3 K9me3 (SET-25)
H3.3 (HIS-72)
ATRX
LTR retrotransposon
Gravid adult
Cer10 expression relative to tba-1 (wildtype set to 1)
*
*
ns
Wildtype
xnp-1
set-25
set-25; xnp-1
his-72
his-72; xnp-1

Figure 3

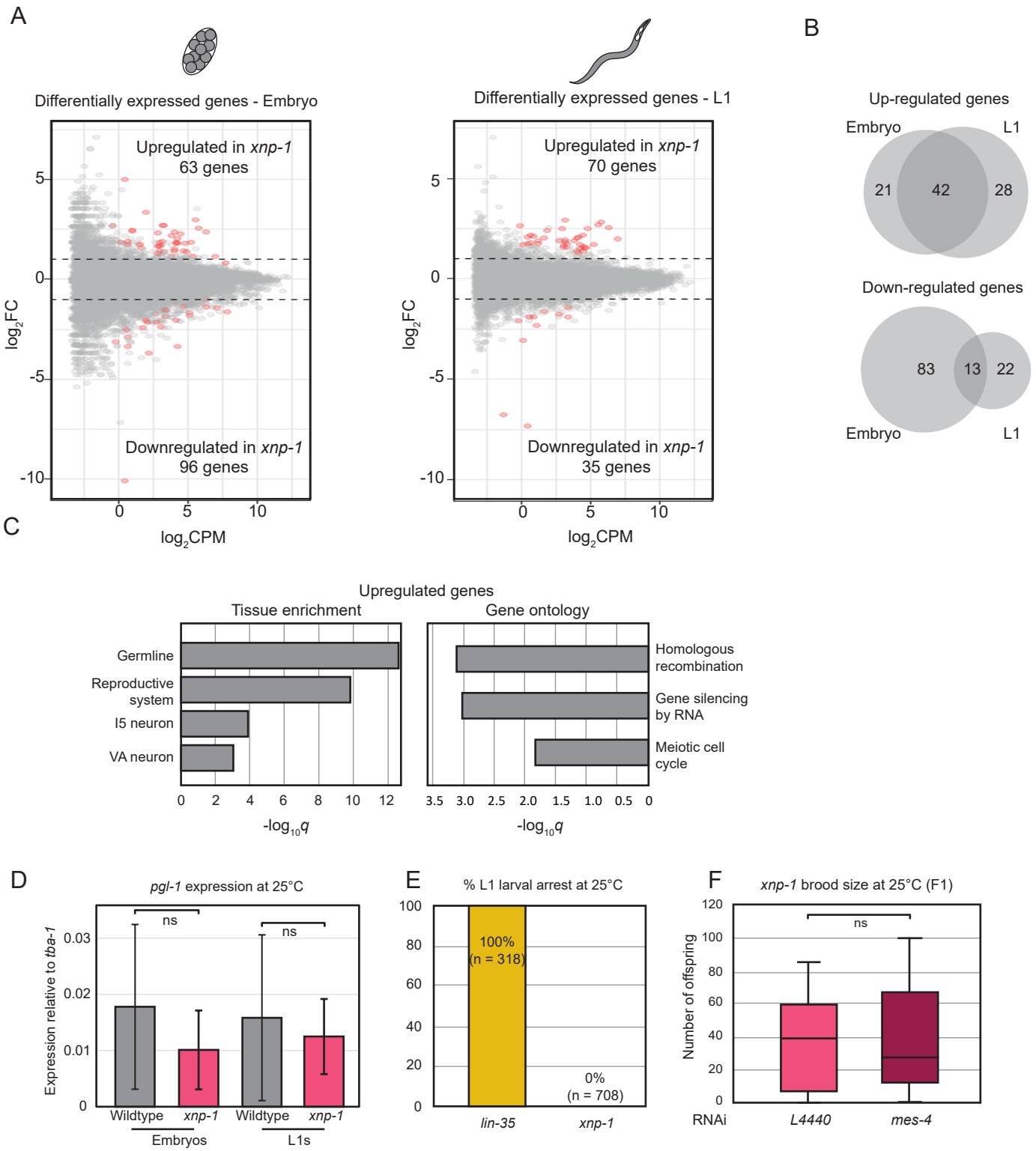

A
Differentially expressed genes - Embryo
Upregulated in xnp-1
63 genes
Downregulated in xnp-1
96 genes
log2FC
log2CPM
Differentially expressed genes - L1
Upregulated in xnp-1
70 genes
Downregulated in xnp-1
35 genes
B
Up-regulated genes
Embryo
L1
21
42
28
Down-regulated genes
83
13
22
Embryo
L1
C
Upregulated genes
Tissue enrichment
Gene ontology
Germline
Reproductive system
I5 neuron
VA neuron
Homologous recombination
Gene silencing by RNA
Meiotic cell cycle
-log10q
D
pgl-1 expression at 25°C
Expression relative to tba-1
ns
ns
Wildtype
xnp-1
Wildtype
xnp-1
Embryos
L1s
E
% L1 larval arrest at 25°C
100%
(n = 318)
0%
(n = 708)
lin-35
xnp-1
F
xnp-1 brood size at 25°C (F1)
Number of offspring
ns
RNAi
L4440
mes-4

Figure 4

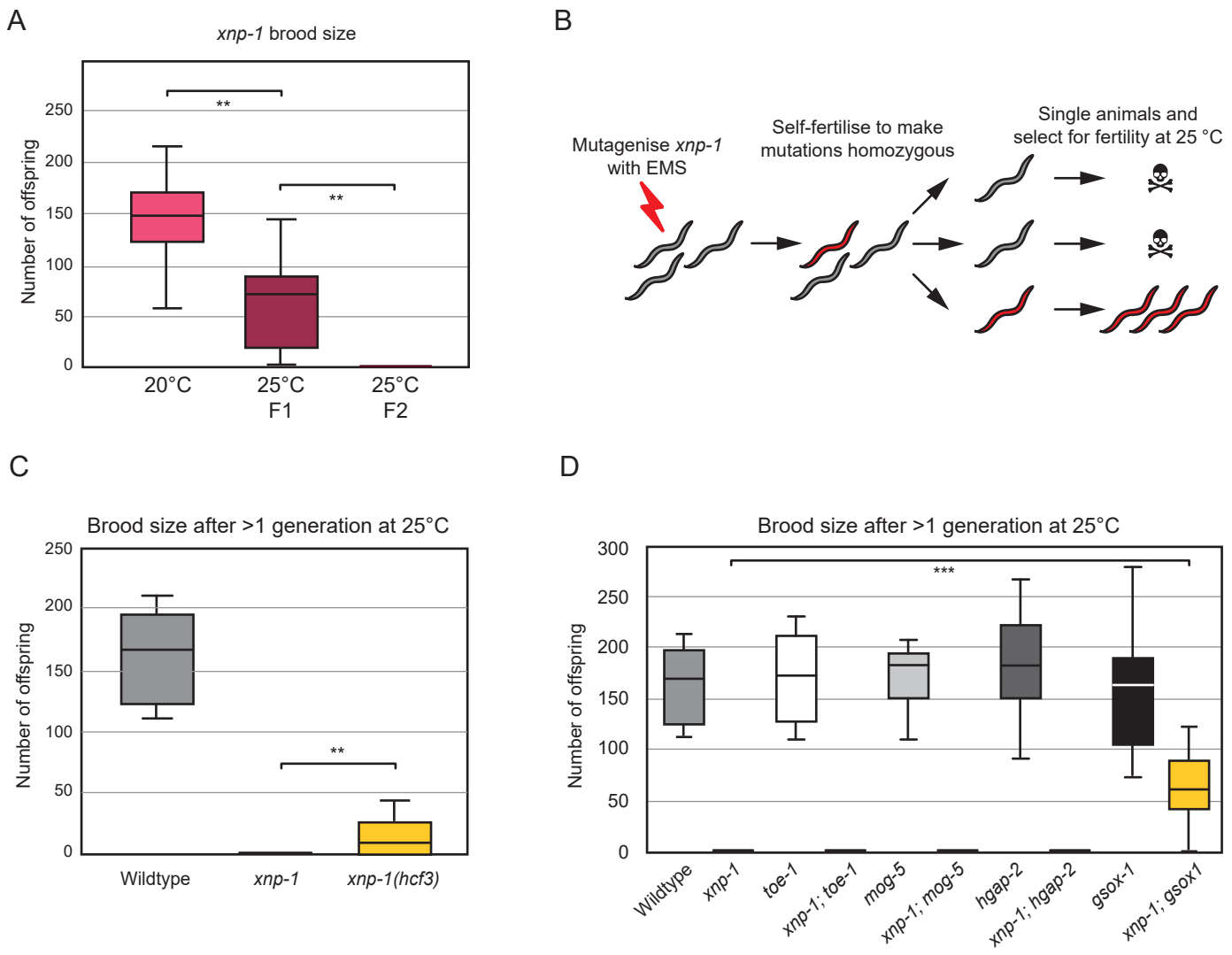

A
xnp-1 brood size
Number of offspring
250
200
150
100
50
0
**
**
20°C
25°C F1
25°C F2
B
Mutagenise xnp-1 with EMS
Self-fertilise to make mutations homozygous
Single animals and select for fertility at 25 °C
C
Brood size after >1 generation at 25°C
Number of offspring
250
200
150
100
50
0
**
Wildtype
xnp-1
xnp-1(hcf3)
D
Brood size after >1 generation at 25°C
Number of offspring
300
250
200
150
100
50
0
***
Wildtype
xnp-1
toe-1
xnp-1; toe-1
mog-5
xnp-1; mog-5
hgap-2
xnp-1; hgap-2
gsox-1
xnp-1; gsox1

Figure 5

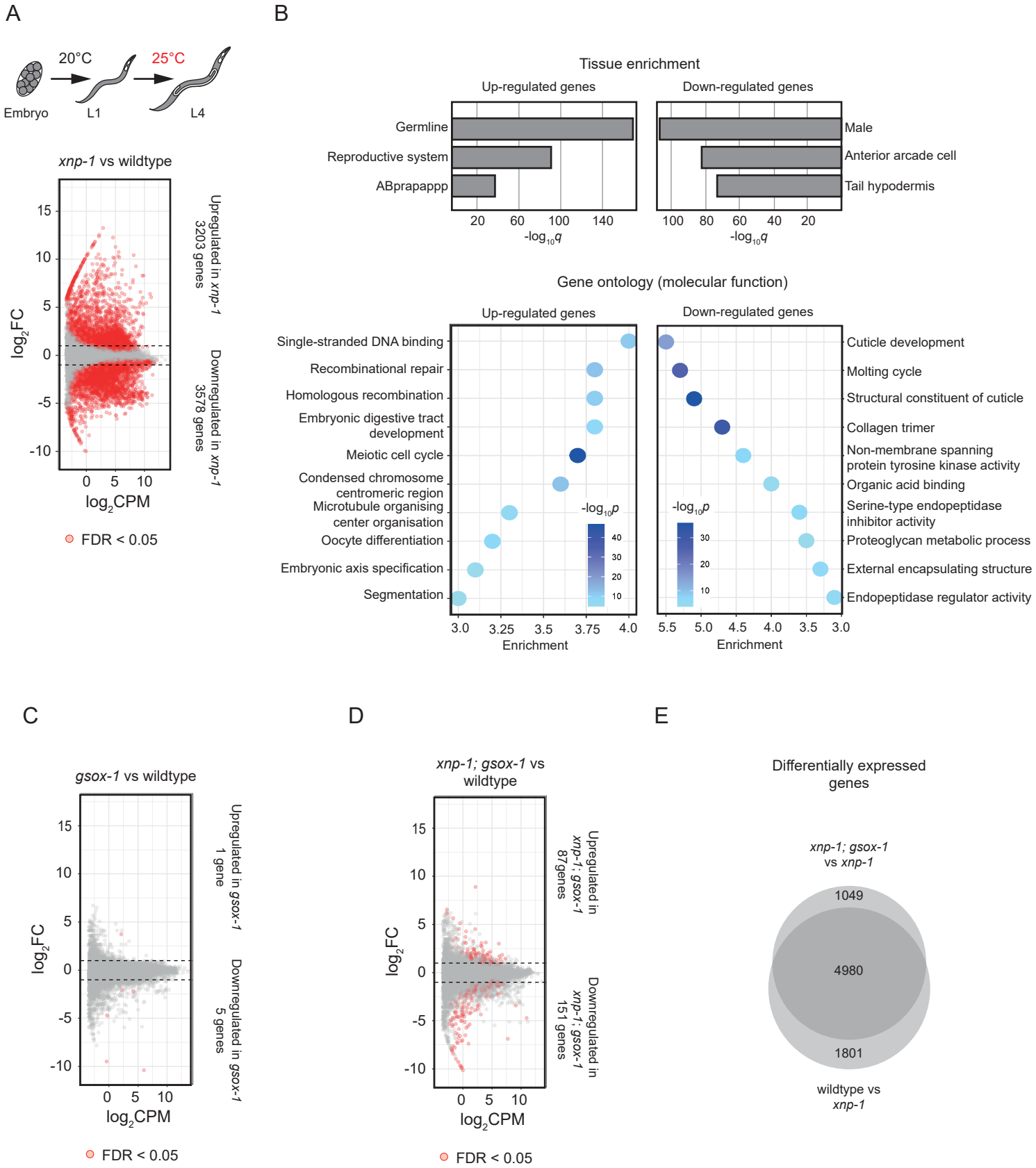
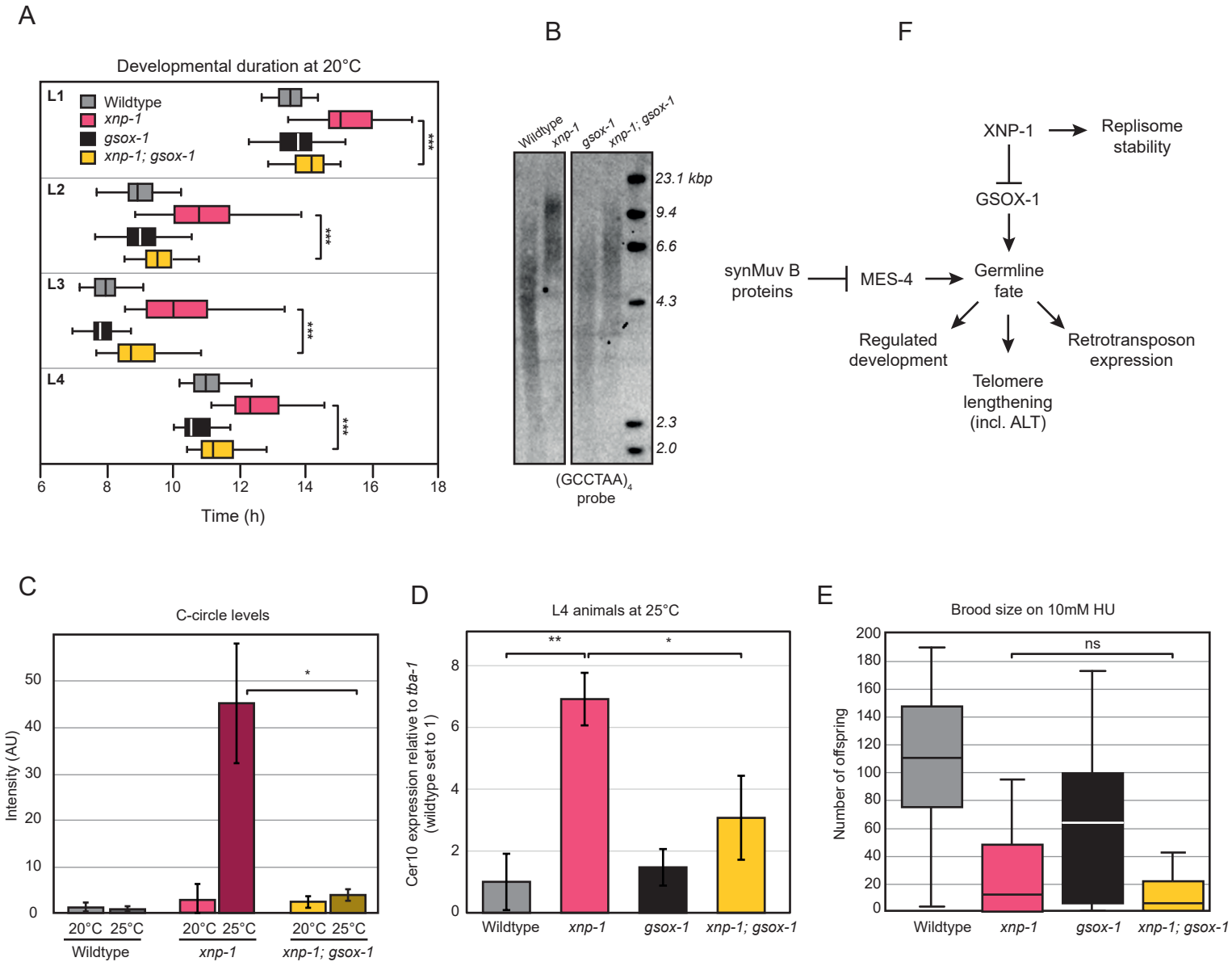

A
20°C
25°C
Embryo
L1
L4
xnp-1 vs wildtype
Upregulated in xnp-1
3203 genes
Downregulated in xnp-1
3578 genes
log2FC
log2CPM
FDR < 0.05
B
Tissue enrichment
Up-regulated genes
Down-regulated genes
Germline
Reproductive system
ABprapappp
Male
Anterior arcade cell
Tail hypodermis
-log10q
Gene ontology (molecular function)
Single-stranded DNA binding
Recombinational repair
Homologous recombination
Embryonic digestive tract development
Meiotic cell cycle
Condensed chromosome centromeric region
Microtubule organising center organisation
Oocyte differentiation
Embryonic axis specification
Segmentation
Cuticle development
Molting cycle
Structural constituent of cuticle
Collagen trimer
Non-membrane spanning protein tyrosine kinase activity
Organic acid binding
Serine-type endopeptidase inhibitor activity
Proteoglycan metabolic process
External encapsulating structure
Endopeptidase regulator activity
-log10p
Enrichment
C
gsox-1 vs wildtype
Upregulated in gsox-1
1 gene
Downregulated in gsox-1
5 genes
log2FC
log2CPM
FDR < 0.05
D
xnp-1; gsox-1 vs wildtype
Upregulated in xnp-1; gsox-1
87genes
Downregulated in xnp-1; gsox-1
151 genes
log2FC
log2CPM
FDR < 0.05
E
Differentially expressed genes
xnp-1; gsox-1 vs xnp-1
1049
4980
1801
wildtype vs xnp-1

Figure 6


A
Developmental duration at 20°C
L1
L2
L3
L4
Wildtype
xnp-1
gsox-1
xnp-1; gsox-1
***
6
8
10
12
14
16
18
Time (h)
B
Wildtype
xnp-1
gsox-1
xnp-1; gsox-1
23.1 kbp
9.4
6.6
4.3
2.3
2.0
(GCCTAA)4
probe
F
XNP-1
Replisome stability
GSOX-1
synMuv B proteins
MES-4
Germline fate
Regulated development
Telomere lengthening (incl. ALT)
Retrotransposon expression
C
C-circle levels
Intensity (AU)
0
10
20
30
40
50
*
20°C 25°C
Wildtype
20°C 25°C
xnp-1
20°C 25°C
xnp-1; gsox-1
D
L4 animals at 25°C
Cer10 expression relative to tba-1 (wildtype set to 1)
0
2
4
6
8
**
*
Wildtype
xnp-1
gsox-1
xnp-1; gsox-1
E
Brood size on 10mM HU
Number of offspring
0
20
40
60
80
100
120
140
160
180
200
ns
Wildtype
xnp-1
gsox-1
xnp-1; gsox-1