bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The Y chromosome contributes to sex-specific aging in Drosophila

Emily J. Brown & Doris Bachtrog*

Department of Integrative Biology, University of California Berkeley, Berkeley, CA 94720, USA

Heterochromatin suppresses repetitive DNA, and a loss of heterochromatin has been observed in aged cells of several species, including humans and Drosophila [1-3]. Males often contain substantially more heterochromatic DNA than females, due to the presence of a large, repeat-rich Y chromosome, and male flies generally have shorter average life spans than females [4,5]. Here we show that repetitive DNA becomes de-repressed more rapidly in old male flies relative to females, and repeats on the Y chromosome are disproportionally mis-expressed during aging. This is associated with a loss of heterochromatin at repetitive elements during aging in male flies, and a general loss of repressive chromatin in aged males away from pericentromeric regions and the Y. By generating flies with different sex chromosome karyotypes (XXY females; X0 and XYY males), we show that repeat de-repression and average lifespan is directly correlated with the number of Y chromosomes. Thus, sex-specific chromatin differences contribute to sex-specific aging in flies.

The chronic deterioration of chromatin structure has been implicated as one of the molecular signatures of aging [6,7], and an overall loss of heterochromatin and repressive histone marks is observed in many old animals [1-3]. Heterochromatin is enriched at repetitive DNA, and its loss can result in de-repression and mobilization of silenced transposable elements (TEs) [8-13]. The amount of repetitive DNA can differ substantially between sexes, due to the presence of a highly repetitive (and normally poorly assembled) Y or W chromosome in the heterogametic sex. In the fruit fly *Drosophila melanogaster*, males contain a ~40-Mb large completely repetitive Y chromosome, implying that almost half of the male genome is heterochromatic compared to only one-third of the female genome [14]. Males have a shorter average lifespan in many taxa, including humans and most Drosophila species [4]. Indeed, the genetic sex determination system predicts adult sex ratios in tetrapods, with the heterogametic sex being less frequent [15]. Lower survivorship of the sex with the repetitive Y or W chromosome may suggest a link between sex-specific mortality, chromatin and sex chromosomes.

To test for sex-specific heterochromatin loss and a de-repression of repetitive DNA during aging, we assayed chromatin and gene expression profiles in young and aged individuals of a standard lab strain of *D. melanogaster*. Lifespan assays confirm that Canton-S males live significantly shorter than females [16] (**Figure 1A**), consistent with multiple studies on sex-specific lifespan in Drosophila [4,5] (**Figure S1**). We gathered replicate stranded RNA-seq data and ChIP-seq data for a repressive histone modification typical of heterochromatin (H3K9me2) from young 8-day and old 64-68-day *D. melanogaster* males and females, using a 'spike in' normalization method to compare the genomic distribution of chromatin marks across samples [17].

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1B shows the genomic distribution of the repressive histone modification H3K9me2 for young and old male and female flies. As expected, heterochromatin is enriched at repetitive regions, including pericentromeres, the small dot chromosome and the repeat-rich Y. While the genomic distribution of H3K9me2 looks similar between young males and females [17], heterochromatin enrichment changes dramatically in old male but less so in old female flies. In particular, we see a general loss of heterochromatin at repetitive regions in aged males, (**Figure 1B-D; Figure S2**), and males show significantly more regions that lose H3K9me2 signal (1.5-fold or more) during aging compared to females (232 vs. 73, p<2.2e-16, Fisher's exact test; **Figure S3**). Almost all regions that lose heterochromatin are located within the pericentromere or the Y chromosome (**Figure S3,4**). On the other hand, fewer regions in males gain H3K9me2 signal (1.5-fold or more) during aging relative to females (6 vs. 120, p<2.2e-16, Fisher's exact test; **Figure S3**). Genomic regions that gain H3K9me2 signal are enriched on the X of females (p<2.2e-16, Fisher's exact test; **Figure S3,4**), and tend to be located close to the pericentromeric boundary (p=0.04 Fisher's exact test; **Figure 1D**, **Figure S3,4**). This suggests that heterochromatin / euchromatin boundaries are less efficiently maintained in old flies, resulting in spreading of heterochromatin from the repeat-rich pericentromere into neighboring regions.

Sex-specific chromatin changes during aging are associated with sex-specific expression changes. We find that genes located in pericentromeric regions change their expression more during aging compared with genes in chromosomal arms, in both sexes (**Figure S5**). While heterochromatin typically has a repressive effect on gene expression, genes located in normally heterochromatic regions (such as the pericentromere) are known to depend on this repressive chromatin environment for proper transcription [18]. Indeed, the global loss of heterochromatin in pericentromeric regions is associated with reduced expression levels of pericentromeric genes in aged males and females (**Figure S5**). Genes that gain the H3K9me2 mark during aging tend to decrease in expression (**Figure S6,7**).

Overall, we find that gene expression during aging differs between males and females. Of the top 10% of genes that are differentially expressed during aging in males and females, 35.5% show expression changes in both sexes (**Figure S8**)**.** Chromosomal location does not appear to be the main determinant influencing gene expression during aging; of the 464 genes that are most differentially expressed in both males and females, only 17 are located inside or within 1Mb of the pericentromere (we expect 22 by chance). The most differentially expressed genes affect similar functional categories in males and females (including GO categories "antibacterial humoral response", "macromolecule biosynthetic process", "reproduction", and "translation", **Figure S9**), but many GO terms are unique to one sex (13 highly significantly enriched (p<10-9) GO terms shared by both sexes; 13 male-specific and 1 female-specific GO term).

In addition, we find sex-specific differences in repeat reactivation during aging. We mapped our transcriptome data to the consensus repeat library of *D. melanogaster* and detect low levels of expression of repetitive elements in young male and female flies (**Figure 2A**). Aged females maintain efficient repression of TEs, while expression for the major classes of annotated TEs increases during aging for males (**Figure 2A,C**). De-repression of TEs is more pronounced in males both in terms of the number of individual

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

elements that show a significant increase in expression during aging, as well as the fraction of the transcriptome that consists of repetitive transcripts across all repetitive elements. Overall, we find that in females, 6 repetitive elements show a significant increase in expression during aging and 14 a significant decrease (**Figure 2C**), but the total fraction of transcripts derived from repeats increases during aging (the fraction of repetitive reads in all RNA-seq reads is 2.0% at 8 days, vs. 4.6% at 68 days, **Table S1, Figure S10**). The increase in repeat expression is much more pronounced in males, with 32 repetitive elements showing a significant increase in expression during aging and 4 showing a significant decrease (**Figure 2C, Figure S10**), and the total fraction of repetitive reads increases from 1.6% to 5.8% (**Table S1**). The TE showing the highest level of de-repression in both sexes is *copia*, which is expressed 28-fold more in old versus young males, and expressed 15-fold more in old versus young females (**Figure 2C**). H3K9me2 profiles at TE families show that there is a general enrichment of this repressive mark in young male and female flies (**Figure 2B**). Consistent with genome-wide expression profiles showing efficient silencing of repeats in old females, there is no global loss of the repressive chromatin mark at repetitive elements in 68-day old females (in fact, there is a slight increase, **Figure 2B**). However, aged *D. melanogaster* males undergo a general loss of the H3K9me2 histone modification in repetitive elements (**Figure 2B**).

Thus, chromatin and gene expression profiles show that TEs lose their epigenetic silencing and become mis-expressed in old male flies. Males have approximately 20% more repetitive sequence than females, due to the repeat-rich Y chromosome. The sex-specific increase in repeat expression may be triggered by the presence of the heterochromatic Y chromosome in males, and the Y indeed shows a dramatic loss of heterochromatin during aging (**Figure 1**). To see if Y-linked repeats are especially prone to mis-regulation during aging, we used *de novo* assembled male-specific and male-biased (putatively Y-linked) repetitive sequences (**Figure S11**) [17]. Indeed, we find that in males, putatively Y-linked repeats are up-regulated more strongly during aging, relative to repeats present in both sexes (p=5.9e-12, Wilcoxon test, **Figure 2D,E**). Overall, we find that 42 Y-linked repeats show a significant increase in expression during aging (and only one a significant decrease; **Figure 2D,E, Figure S12**), and the total fraction of transcripts derived from Y-linked repeats increases more than 9-fold in old males (**Table S2**). Additionally, putatively Y-linked repeats disproportionately lose the repressive histone modification H3K9me2 during aging compared to other repeats (p=3.4e-11, Wilcoxon test, **Figure 2D**). Thus, male-biased and male-specific repeats, i.e. repeats that are located on the Y chromosome, are especially prone to de-repression during aging in males.

To directly test whether the Y chromosome contributes to sex-specific TE de-repression and aging, we generated *D. melanogaster* females containing a Y chromosome (XXY flies), and males with either zero or two Y chromosomes (X0 and XYY flies), by crossing Canton-S flies to different strains with attached-X and attached-X-Y chromosomes (**Figure 3A**, see **Methods** for strain information). We compared sex-specific lifespans of wildtype Canton-S *D. melanogaster* flies, and XXY female and XO and XYY male (**Figure 3B,C**). Note that XO males and XXY females of a given cross are isogenic on the autosomes, but differ in their genomic background from other crosses and from Canton-

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

S, which can contribute to lifespan variation among strains (**Figure 3B,C**). Cumulative survival probabilities show that life span of females that contain a Y chromosome (XXY females) is reduced relative to wildtype females or males that lack a Y chromosome (X0 males) for all crosses assayed (**Figure 3B,C**). Indeed, X0 males show a dramatic increase in life span relative to wildtype males, and even outlive wildtype females (**Figure 3B,C**). X0 males are sterile and have the least amount of repetitive DNA of all karyotypes investigated (~10Mb less than Canton-S females and ~40Mb less than Canton-S males [17]); both of these factors may contribute to increased lifespan. Males with two Y chromosomes (XYY), in contrast, live the shortest (**Figure 3B,C**), and their lifespan is reduced considerably relative to wildtype males, despite both karyotypes being fertile. Thus, survivorship data suggest that the number of Y chromosomes influences organismal survival.

Gene expression changes during aging in the aberrant karyotypes show many of the same patterns as wildtype flies, with similar networks of GO terms enriched in both XO and XYY males, and XXY females, including "reproduction" or "sensory perception of chemical stimulus" (**Figure S9**). Overall, we find that 101 of the top 10% of genes mis-expressed during aging are shared among all 5 karyotypes ($p<<1e-5$, permutation test). These genes do not show any enrichment for a particular GO term, and only 6 of them are located inside or within 1Mb of the pericentromere (expect 4.8 genes).

Genomic location also influences gene expression changes during aging in flies with aberrant karyotypes. As in wildtype males and females, genes located in the pericentromere show a decrease in expression during aging in XXY females and X0 males (**Figure S13;** XYY flies show no significant expression change at pericentromeric genes). Y-linked genes in wild-type males are expressed almost exclusively in male reproductive tissues [19], and we do not detect any expression of Y-linked genes even in very old XY male heads (**Figure S13**). In contrast, we find that Y-linked genes are inefficiently silenced in heads of XXY females and XYY males and are becoming de-repressed even more during aging (6 Y genes are expressed in old XXY females, and 14 in old XYY males, **Figure S13**).

Thus, wildtype males maintain efficient silencing at their Y-linked genes during aging, despite global heterochromatin loss on the Y chromosome and mis-expression of Y-linked repeats. Silencing mechanisms on the Y chromosome of XXY and XYY flies, on the other hand, appear to be generally compromised, even in young individuals. Indeed, we previously showed that the Y chromosome affects global heterochromatin integrity [17]. Young flies with additional Y chromosomes (XXY females or XYY males) show lower levels of H3K9me2 enrichment at their TEs, and a de-repression of Y-linked repeats relative to wildtype flies, while X0 flies showed increased levels of H3K9me2 at repeats [17]. Expression profiles in XXY females and XYY/X0 males demonstrate that the absence or presence of the Y chromosome modulates expression of TEs during aging (**Figure 4**). Expression profiles from aged flies with aberrant sex chromosome karyotypes confirm our expectation that X0 males show less de-repression of TEs during aging relative to wildtype males (**Figure 4A,B**). XXY females, on the other hand, show more mis-expression of repeats during aging compared to wildtype females. In XXY females, 7 elements show a significant increase in expression during aging and 3 elements show a

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

significant decrease in expression (**Figure 4A,B**, compared to 6 /14 elements that increase/decrease expression in wild-type females). The fraction of repetitive transcripts increases more during aging for XXY females (3.1-fold increase in XXY females vs. 2.2-fold in wildtype females, **Table S1**), and a greater proportion of the transcriptome is comprised of repeats in old XXY females relative to wildtype (4.8% in XXY females vs. 4.6% in wild-type females). XYY males show the greatest number of repeats with significantly increased expression during aging (33 elements, **Figure 4A,B**), but not the highest fold change in total fraction of repetitive reads (**Table S1**), partly because young XYY males already show the highest expression of repeats of any of the 5 karyotypes (**Table S1**), and partly because old XYY males are approximately 30 days younger than the other karyotypes (**Figure 3B**).

Mis-expression of repetitive elements in XXY females and XYY males is especially pronounced for repeats found on the Y chromosome. Y-linked repeats show reduced silencing even in young XXY females and XYY males relative to wildtype males (**Figure 4C,D**) [17], and become de-repressed even more during aging in XXY and XYY individuals (**Figure 4C,D**). Overall, wildtype males express 64 putatively Y-linked repeats during their lifespan, while XXY females express 86 and XYY males express 102. In XXY females, 13 repeats significantly increase in expression during aging (and 5 decrease), and 71 repeats significantly increase in expression in XYY males (and 8 decrease; **Figure S12**). Indeed, even at just 37 days old, XYY males already show higher, presumably aberrant, expression of Y-linked repeats compared to 68-day-old wild-type males (**Figure 4C,D**).

To conclude, our data demonstrate that the repeat-rich Y chromosome decreases life span in Drosophila. Loss of heterochromatin in repetitive regions during aging is more pronounced in male flies, and is accompanied by a de-repression of TEs. Y-linked repeats disproportionally lose their repressive marks and become reactivated, and analysis of flies with aberrant sex chromosome configurations demonstrates that the Y has a direct influence on organismal survival. Age-related heterochromatin loss on the repetitive, sex-limited Y or W-chromosome and repeat re-activation can contribute to lower survivorship of the heterogametic sex across taxa [15], including humans. Y chromosomes of Drosophila and humans are known to harbor structural polymorphism in heterochromatic sequences and copy-number variation in repeats [20,21], and polymorphism on the *D. melanogaster* Y has been shown to affect lifespan [22], the formation of heterochromatin [23], and the regulation of TEs and 100s of genes genome-wide [24,25]. More generally, individual humans and flies show extensive variation in their repeat content [26,27], and our results raise the question whether natural variation in repetitive sequences can contribute to genetic variation in longevity among individuals.

Materials & Methods

Drosophila strains. Fly strains were obtained from the Bloomington Stock Center and the Kyoto Stock Center. The following strains were used: Canton-S; Oregon-R; 2549 ($C(1;Y),y^1cv^1v^1B/0$ & $C(1)RM,y^1v^1/0$); 4248 ($C(1)RM, y^1\ pn^1\ v^1/C(1;Y)1, y^1\ B^1/0; sv^{spa-pol}$) from the Bloomington Stock Center, and 100950 ($0/\ C(1)RM, y^1w^{str}/C(t;Y)1, y^1\ y^+\ ac^1\ sc^1\ w^1$) from the Kyoto Stock Center. The crossing scheme used to obtain XO and

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

XYY males and XXY females is depicted in **Fig. 3A**. For chromatin and gene expression analyses, flies were grown in incubators at 25°C, 60% relative humidity, and 12h light for the indicated number of days following eclosion, and were then flash-frozen in liquid nitrogen and stored at -80°C.

Lifespan assays. Lifespan data was collected for all karyotypes in the same rearing conditions as described above. The lifespan assays were conducted as described [28]. Briefly, synchronized embryos were collected on agar plates, mobilized with a cotton swab, washed 3 times with PBS pH 7.4, and 10µl of embryos were pipetted to a fresh vial of standard fly medium. Adult flies were collected over 2 days, and were allowed to mate for 2 more days. Flies were then sexed, and 30 flies were counted into each vial. Vials were then flipped, without using CO_2, every 2-3 days, and fly deaths were recorded. Flies that were observed escaping the vial were censored. To collect samples for the RNA-seq and ChIP-seq experiment, we censored the entire lifespan experiment once it reached 50% survivorship and flash-froze the remaining flies in liquid nitrogen. In total, 8,829 flies in 297 vials were counted for the lifespan assays reported here.

Chromatin Immunoprecipitation and Sequencing. We performed ChIP-seq experiments using a standard protocol adapted from [29]. Briefly, approximately 2ml of adult flash-frozen flies were dissected on dry ice, and heads and thoraces were used to fix and isolate chromatin. Following chromatin isolation, we spiked in 60µl of chromatin prepared from female *Drosophila miranda* larvae (approximately 1µg of chromatin). We then performed immunoprecipitation using 4µl of the H3K9me2 (Abcam ab1220) antibody.

After reversing the cross-links and isolating DNA, we constructed sequencing libraries using the BIOO NextFlex sequencing kit. Sequencing was performed at the Vincent J. Coates Genomic Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. We performed 50bp single-read sequencing for our input libraries, and 100bp paired-end sequencing for H3K9me2 libraries, due to their higher repeat content.

We collected replicate datasets for H3K9me2 enrichment in aged males and females to confirm differences seen between the sexes and between young and old samples (**Fig S2**). Replicate H3K9me2 data for young flies are from [17].

RNA extraction and RNA-seq. We collected replicate RNA samples for aged individuals of all five karyotypes of interest; replicate RNA data for young flies are from [17]. Additionally, we collected 3 replicate samples for aged male Canton-S, aged XXY females, and aged XYY males, and 4 replicate samples for aged female Canton-S and aged XO males. After flash-freezing in liquid nitrogen, we dissected and pooled 5 heads from each sample, extracted RNA, and prepared stranded total RNA-seq libraries using Illumina's TruSeq Stranded Total RNA Library Prep kit with Ribo-Zero ribosomal RNA reduction chemistry, which depletes the highly abundant ribosomal RNA transcripts (Illumina RS-122-2201). We performed single-read sequencing for all total RNA libraries at the Vincent J. Coates Genomic Sequencing Laboratory at UC Berkeley.

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mapping of sequencing reads and data normalization. For all *D. melanogaster* alignments, we used Release 6 of the genome assembly and annotation [30]. For all ChIP-seq datasets, we used Bowtie2 [31] to map reads to the genome, using the parameters "-D 15 –R 2 –N 0 –L 22 –i S,1,0.50 --no-1-mm-upfront", which allowed us to reduce cross-mapping to the *D. miranda* genome to approximately 2.5% of 50bp reads, and 1% of 100bp-paired end reads. We also mapped all ChIP-seq datasets to the *D. miranda* genome assembly [32] to calculate the proportion of each library that originated from the spiked-in *D. miranda* chromatin versus the *D. melanogaster* sample.

To calculate the ChIP signal we first calculated the coverage across 5kb windows for both the ChIP and the input, and then normalized by the total library size, including reads that map to both *D. melanogaster* and the *D. miranda* spike. We then calculated the ratio of ChIP coverage to input coverage for each of the 5kb windows, and normalized by the ratio of *D. melanogaster* reads to *D. miranda* reads in the ChIP library, and then by the ratio of *D. melanogaster* reads to *D. miranda* reads in the input, to account for differences in the ratio of sample to spike present before immunoprecipitation. We describe the validation of this normalization method in [17].

Gene expression analysis. For each replicate of RNA-seq data, we first mapped RNA-seq reads to the ribosomal DNA scaffold in the Release 6 version of the *D. melanogaster* genome, and removed all reads that mapped to this scaffold, as differences in rRNA transcript abundance are likely to be technical artifacts from the total RNA library preparation, which aims to remove the bulk of rRNA transcripts. We then mapped the remaining reads to the Release 6 version of the *D. melanogaster* genome using Tophat2 [33], using default parameters. We then used Cufflinks [34] and Cuffdiff [35] to merge replicates and calculate normalized FPKMs for all samples. GO analysis was performed using GOrilla, using ranked lists of differentially expressed genes [36].

Repeat libraries. We used two approaches to quantify expression of repeats. Our first approach was based on consensus sequences of known repetitive elements that were included in the Release 6 version of the *D. melanogaster* genome and are available on FlyBase. These included consensus sequences for 125 TEs and the 3 largest satellites (359, dodeca, and responder).

Our second approach aimed to specifically assess the repeat content of the Y chromosome. Since the Y chromosome is poorly assembled and repetitive elements on the Y are not annotated, we previously assembled repetitive elements *de novo* from male and female genomic DNA reads using RepARK and identified 101 male-specific repeats comprising 13.7kb of sequence, based on male-specific coverage analysis [17,37].

To assess expression of repetitive elements, we mapped RNA-seq reads to each of the repeat libraries (consensus TEs and putatively Y-linked repetitive elements) using Bowtie2 and the parameters "-D 15 –R 2 –N 0 –L 22 –i S,1,0.50 --no-1-mm-upfront". We then calculated the mean coverage across each repetitive element using Bedtools, and normalized the coverage by the number of uniquely-mapping reads in the sequencing library. We made this calculation independently for each replicate for each time sample

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and karyotype, and then calculated both the average expression value as well as the standard deviation, to assess statistical significance and reproducibility.

To assess H3K9me2 signal in repetitive elements, we took a similar approach as we did for calculating ChIP enrichment profiles across the genome. First, we mapped both ChIP and input sequencing reads to each of the repeat libraries using Bowtie2 and the parameters "-D 15 –R 2 –N 0 –L 22 –i S,1,0.50 --no-1-mm-upfront". We then calculated the mean coverage across each repetitive element using Bedtools, and normalized the coverage by the total library size, including reads that mapped to both the *D. melanogaster* and *D. miranda* genomes. We then calculated the ratio of ChIP coverage to input coverage for each repetitive element, and then normalized by the ratio of *D. melanogaster* reads to *D. miranda* reads in the ChIP library, and then by the ratio of *D. melanogaster* reads to *D. miranda* reads in the input, as described above and in [17]. This method accounts for differences in copy number of the repetitive elements by dividing the ChIP coverage by each repeat's coverage in the input.

References

1. Zhang W, Li J, Suzuki K, Qu J, Wang P, et al. (2015) Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. Science 348: 1160-1163.
2. Larson K, Yan SJ, Tsurumi A, Liu J, Zhou J, et al. (2012) Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis. PLoS Genet 8: e1002473.
3. Haithcock E, Dayani Y Fau - Neufeld E, Neufeld E Fau - Zahand AJ, Zahand Aj Fau - Feinstein N, Feinstein N Fau - Mattout A, et al. (2005) Age-related changes of nuclear architecture in Caenorhabditis elegans. Proc Natl Acad Sci U S A 102: 16690-16695.
4. Yoon JS, Gagen KP, Zhu DL (1990) Longevity of 68 species of Drosophila. . Ohio J Sci 90: 16-32.
5. Tower J, Arbeitman M (2009) The genetics of gender and life span. Journal of Biology 8: 38.
6. O'Sullivan RJ, Karlseder J (2012) The great unravelling: chromatin as a modulator of the aging process. Trends Biochem Sci 37: 466-476.
7. Wood JG, Hillenmeyer S, Lawrence C, Chang C, Hosier S, et al. (2010) Chromatin remodeling in the aging genome of Drosophila. Aging Cell 9: 971-978.
8. Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, et al. (2013) Activation of transposable elements during aging and neuronal decline in Drosophila. Nat Neurosci 16: 529-531.
9. De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, et al. (2013) Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. Aging Cell 12: 247-256.
10. De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM, et al. (2013) Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. Aging (Albany NY) 5: 867-883.
11. Wood JG, Helfand SL (2013) Chromatin structure and transposable elements in

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

organismal aging. Front Genet 4: 274.
12. Wood JG, Jones BC, Jiang N, Chang C, Hosier S, et al. (2016) Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. Proc Natl Acad Sci U S A: in press.
13. Van Meter M, Kashyap M, Rezazadeh S, Geneva AJ, Morello TD, et al. (2014) SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. Nat Commun 5: 5011.
14. Hoskins RA, Smith CD, Carlson JW, Carvalho AB, Halpern A, et al. (2002) Heterochromatic sequences in a Drosophila whole-genome shotgun assembly. Genome Biol 3: RESEARCH0085.
15. Pipoly I, Bokony V, Kirkpatrick M, Donald PF, Szekely T, et al. (2015) The genetic sex-determination system predicts adult sex ratios in tetrapods. Nature 527: 91-94.
16. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Amer Statist Assn 53: 457–481.
17. Brown EJ, Bachtrog D (2016) The Drosophila Y chromosome affects heterochromatin integrity genome-wide submitted.
18. Lu BY, Emtage PC, Duyf BJ, Hilliker AJ, Eissenberg JC (2000) Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. Genetics 155: 699-708.
19. Carvalho AB, Koerich LB, Clark AG (2009) Origin and evolution of Y chromosomes: Drosophila tales. Trends in Genetics 25: 270-277.
20. Lyckegaard EM, Clark AG (1989) Ribosomal DNA and Stellate gene copy number variation on the Y chromosome of Drosophila melanogaster. Proc Natl Acad Sci U S A 86: 1944-1948.
21. Repping S, Skaletsky H, Brown L, van Daalen S, Korver C, et al. (2003) Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. Nat Genet 35: 247 - 251.
22. Griffin RM, Le Gall D, Schielzeth H, Friberg U (2015) Within-population Y-linked genetic variation for lifespan in Drosophila melanogaster. J Evol Biol 28: 1940-1947.
23. Lemos B, Branco AT, Hartl DL (2010) Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. Proceedings of the National Academy of Sciences of the United States of America 107: 15826-15831.
24. Lemos B, Araripe LO, Hartl DL (2008) Polymorphic Y chromosomes harbor cryptic variation with manifold functional consequences. Science 319: 91-93.
25. Sackton TB, Montenegro H, Hartl DL, Lemos B (2011) Interspecific Y chromosome introgressions disrupt testis-specific gene expression and male reproductive phenotypes in Drosophila. Proceedings of the National Academy of Sciences of the United States of America 108: 17046-17051.
26. Ewing AD, Kazazian HH, Jr. (2010) High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. Genome Res 20: 1262-1270.
27. Bosco G, Campbell P, Leiva-Neto JT, Markow TA (2007) Analysis of Drosophila species genome size and satellite DNA content reveals significant differences among strains as well as between species. Genetics 177: 1277-1290.

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

28. Linford NJ, Bilgir C, Ro J, Pletcher SD (2013) Measurement of lifespan in Drosophila melanogaster. J Vis Exp.
29. Alekseyenko A, Larschan E, Lai W, Park P, Kuroda M (2006) High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. Genes Dev 20: 848 - 857.
30. Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, et al. (2015) The Release 6 reference sequence of the Drosophila melanogaster genome. Genome Res 25: 445-458.
31. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357-359.
32. Ellison CE, Bachtrog D (2013) Dosage compensation via transposable element mediated rewiring of a regulatory network. Science 342: 846-850.
33. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, et al. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14: R36.
34. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. 7: 562-578.
35. Trapnell C, Hendrickson D, Sauvageau M, Goff L, Rinn J, et al. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol 31: 46-53.
36. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10: 48.
37. Koch P, Platzer M, Downie BR (2014) RepARK--de novo creation of repeat libraries from whole-genome NGS reads. Nucleic Acids Res 42: e80.

Acknowledgements. This study was supported by NIH grants to DB (R01GM076007, R01GM093182 and R01GM101255).

Author contributions: DB and EB conceived the study and wrote the paper. EB collected and analyzed the data.

Author information: All RNA-seq and ChIP-seq reads are deposited in GeneBank (Accession number xx). The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to dbachtrog@berkeley.edu.

Figure Legends

Figure 1. Aging and the sex-specific chromatin landscape in Drosophila. **A.** Kaplan-Meier survivorship curves [16] for Canton-S males (blue) and females (red), with the shaded region indicating the upper and lower 95% confidence interval calculated from the Kaplan-Meier curves. Karyotypes of male and female *D. melanogaster* are shown, with heterochromatic regions indicated in blue, and euchromatic regions in gray. **B.** Genome-wide enrichment of H3K9me2 for young (8 days) and old (64 or 68 days) *D. melanogaster*

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

males and females along the different chromosome arms. Enrichment in 5kb windows is shown in red lines (normalized ratio of ChIP to input, see Materials & Methods), and the enrichment in 20kb windows is shown in gray scale according to the scale in the upper left, with the darkest gray corresponding to the highest 5% of values across all windows from all samples, and the lightest gray corresponding to the lowest 10% of values across all windows from all samples. Subtraction plots show the absolute difference in signal of 50kb windows between young and aged flies along the chromosome arms, with each sample further smoothed by subtracting out the median autosomal euchromatin signal, with females in red and males in blue, and the pericentromeric region of each chromosome indicated by the red segment of the line beneath each chromosome. **C.** Box plot showing the smoothed ChIP signal for all 5kb windows in different chromosomal regions (* $p<0.05$, ** $p<1e-6$, *** $p<1e-12$, Wilcoxon test) for males (blue) and females (red), with pericentromere boundaries defined by the Release 6 version of the *D. melanogaster* genome. **D.** Enrichment of H3K9me2 (in 5kb windows) for 1Mb upstream and downstream of the euchromatin/ pericentromere boundary, indicated by the dotted red line, on the 5 major chromosome arms. Subtraction plots show higher H3K9me2 signal in young (blue) or old (red) flies.

Figure 2. Enrichment of H3K9me2 at different repeat families and their expression levels for young and old females and males. **A.** Expression of all repeats from FlyBase consensus library from Release 6 of the *D. melanogaster* genome in young (8day) and old (~68day) male and female Canton-S, averaged across replicates, with significance values calculated using the Wilcoxon test (* $p< 0.05$, ** $p<0.01$, *** $p<1e-5$). Heatmaps are visualized globally according to the scale, with dark red corresponding to the top 5% of all values across all samples and dark blue corresponding to the bottom 5% of all values across all samples. **B.** H3K9me2 enrichment in repeats from FlyBase consensus library in young (8day) and old (64 or 68day) male and female Canton-S, averaged across replicates, with significance values calculated using the Wilcoxon test (* $p<0.05$, ** $p<0.01$, *** $p<1e-5$). The heatmap is scaled in the same manner as in (A.). **C.** Expression for repeat families for old and young males and females, with lines indicating the standard deviation for each estimate of expression across replicates and colors indicating the class of repetitive element. **D.** Expression and H3K9me2 signal in putatively Y-linked repeats in young (8day) and old (64 or 68day) Canton-S males, with significance values calculated using the Wilcoxon test (* $p<0.01$, ** $p<1e-4$, *** $p<1e-10$). Heatmaps are scaled in the same manner as in (A). **E.** Expression of putatively Y-linked repeats for old and young Canton-S males, with lines indicating the standard deviation for each estimate of expression across replicates.

Figure 3. Survivorship of XXY females and X0 and XXY males. **A.** Crossing scheme used to generate flies with aberrant sex chromosomes, with Canton-S used as the wild-type males and females for all crosses, and various lines with C(1)RM and C(1;Y) indicated by the attached X/ XY karyotypes. **B.** Kaplan-Meier survivorship curves for flies with aberrant sex chromosome karyotype, generated with various C(1)RM and C(1;Y) lines as indicated at the top of each survivorship curve. Shaded areas indicate the upper and lower 95% confidence interval calculated from the Kaplan-Meier curves. **C.** Median lifespan for each of the different karyotypes measured, with error bars indicating the upper and lower 95% confidence intervals (estimated by the Kaplan-Meier curves).

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Significance is compared to the wild-type Canton-S of the same sex for each aberrant karyotype, and was calculated using the survdiff package in R (* $p<0.01$, ** $p<1e\text{-}6$, *** $p<1e\text{-}12$).

Figure 4. Expression of repetitive elements in XXY females and X0 and XYY males during aging. **A.** Expression of all repeats from FlyBase consensus library from the Release 6 of the *D. melanogaster* genome, averaged across replicates, with significance values calculated using a Wilcoxon test (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). **B.** Expression of all repeats from the FlyBase consensus library in XO males, XXY females, and XYY males, with lines indicating the standard deviation of each expression value calculated from replicates, and color indicating the class of repetitive element. **C.** Expression for putatively Y-linked (male-specific) repeats in karyotypes with a Y chromosome, averaged across replicates, with significance calculated using the Wilcoxon test (* $p<0.05$, ** $p<0.001$, *** $p<1e\text{-}5$) **D.** Like (B.) for putatively Y-linked repeats.

Figure S1. Kaplan-Meier survivorship curves of line 2549 males and females ((C(1;Y),$y^1cv^1v^1B$/0 & C(1)RM,y^1v^1/0) and Oregon-R wild-type males and females.

Figure S2. Pearson correlation coefficients for replicate H3K9me2 datasets for old males and females, and boxplots of normalized enrichment values for the replicates. Genome-wide plots were generated using replicate data as in Figure 1B. and 1D.

Figure S3. Chromosomal locations of 50kb windows that gain (red) or lose (blue) at least 1.5-fold H3K9me2 signal during aging for males and females. Pericentromeric regions are indicated by the red portion of the line beneath each chromosome.

Figure S4. Chromosomal locations of the top 10% of 50kb windows that gain (red) or lose (blue) H3K9me2 enrichment during aging for males and females. Pericentromeric regions are indicated by the red portion of the line beneath each chromosome.

Figure S5. Expression values of all genes, normalized across replicates, of young and old males and females by chromosome location, as annotated in the Release 6 of the *D. melanogaster* genome. We consider expressed genes as those with FPKM>1, as determined by median intronic FPKM. Significance values are calculated using the Wilcoxon test.

Figure S6. Expression values of genes located in 50kb windows that show either a 1.5-fold loss or 1.5-fold gain of H3K9me2 during aging in males and females. Significance values are calculated using the Wilcoxon test.

Figure S7. Expression values of genes located in the top 10% of 50kb windows that either gain or lose H3K9me2 during aging in males and females. Significance values are calculated using the Wilcoxon test.

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure S8. Overlap of the top 10% of differentially expressed genes during aging, normalized across replicates, for various combinations of the 5 sex chromosome karyotypes examined.

Figure S9. GO category enrichment of genes differentially expressed during aging in wild-type Canton-S males (**A**.) and Canton-S females (**B**.), and XO males (**C**.), XXY females (**D**.), and XYY males (**E**.). Genes were ranked by their fold change in expression, averaged across replicates, regardless of direction, and submitted to GOrilla [36] for GO category enrichment analysis.

Figure S10. Number of repeats that show a significant increase (red) or decrease (blue) in expression during aging as a fraction of all repeats from the FlyBase consensus repeat library, with significance estimated using standard errors from replicate datasets. Significance is calculated using Fisher's exact test, with red stars indicating significance for repeats that increase in expression, and blue stars indicating significance for repeats that decrease in expression during aging. We also show the estimates of the total fraction of RNA-seq reads that map to the FlyBase consensus repeat library, with error bars calculated from replicate datasets, for young and old samples from each of the 5 karyotypes.

Figure S11. Male vs. female genomic coverage of *de novo* assembled repeats, with putatively Y-linked repeats indicated in blue and purple as those with male-specific or highly male-biased genomic coverage patterns [17].

Figure S12. Number of putative Y-linked repeats that show a significant increase (red) or decrease (blue) in expression during aging as a fraction of all repeats from a male-specific repeat library (see **Figure S11**), with significance estimated using standard errors from replicate datasets. Significance is calculated using Fisher's exact test, with red stars indicating significance for repeats that increase in expression, and blue stars indicating significance for repeats that decrease in expression during aging. We also show the estimates of the total fraction of RNA-seq reads that map to the Y-specific consensus repeat library, with error bars calculated from replicate datasets, for young and old samples from each of the 5 karyotypes.

Figure S13. Expression values of all genes, normalized across replicates, of young and old XO males, XXY females, and XYY males by chromosome location, as annotated in the Release 6 of the *D. melanogaster* genome. Significance values are calculated using the Wilcoxon test.

Figure S14. Estimates of expression values for the FlyBase consensus library for all replicates for all karyotypes.

Figure S15. Estimates of H3K9me2 signal for the FlyBase consensus library for all replicates for all karyotypes.

bioRxiv preprint doi: https://doi.org/10.1101/156042; this version posted June 26, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table S1. Average estimation, across all replicates, of the fraction of all RNA-seq reads that are derived from the FlyBase consensus repeat library, as well as the fold change in repetitive content during aging.

Table S2. Average estimation, across all replicates, of the fraction of all RNA-seq reads that are derived from the putative Y-linked consensus repeat library, as well as the fold change in repetitive content during aging.

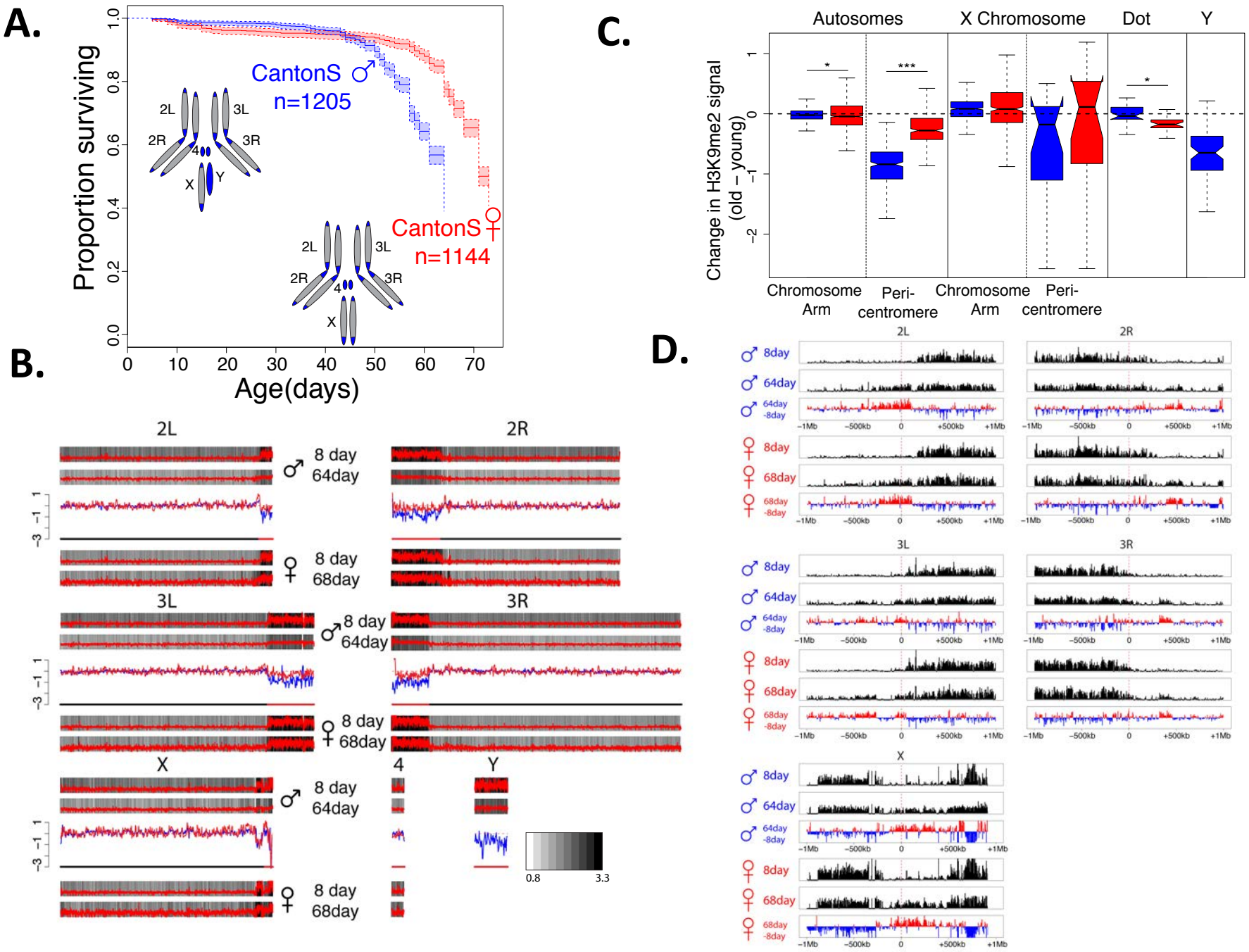
A.
Proportion surviving
Age(days)
CantonS ♂
n=1205
CantonS ♀
n=1144
2L
2R
3L
3R
4
X
Y
B.
2L
2R
3L
3R
X
4
Y
8 day
64day
68day
C.
Autosomes
X Chromosome
Dot
Y
Change in H3K9me2 signal
(old – young)
Chromosome Arm
Peri-centromere
D.
8day
64day
68day
-1Mb
-500kb
+500kb
+1Mb

Figure 1

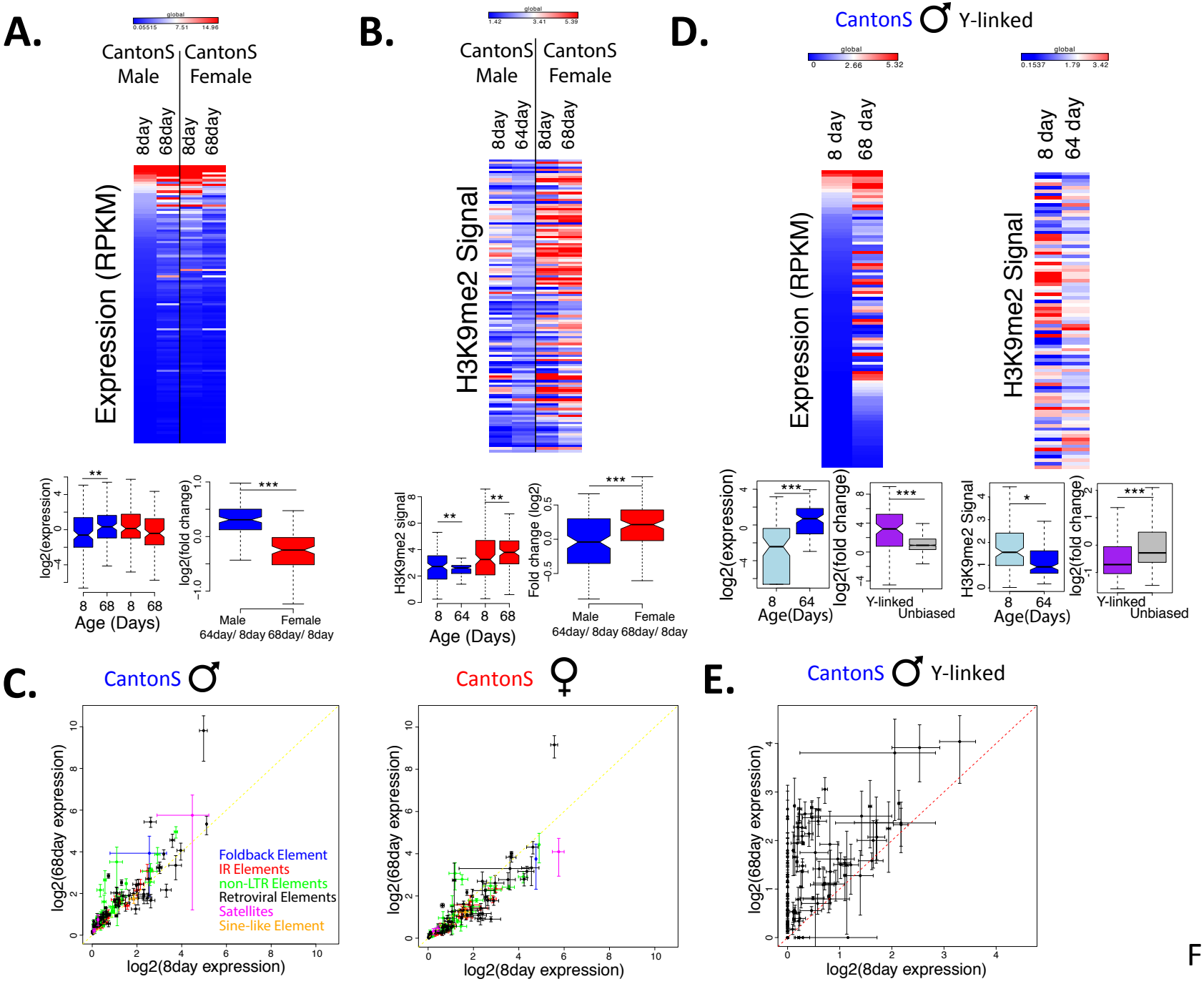
A.
CantonS Male
CantonS Female
8day
68day
8day
68day
Expression (RPKM)
log2(expression)
log2(fold change)
8 68 8 68
Age (Days)
Male
Female
64day/ 8day 68day/ 8day
B.
CantonS Male
CantonS Female
8day
64day
8day
68day
H3K9me2 Signal
H3K9me2 signal
Fold change (log2)
8 64 8 68
Age (Days)
Male
Female
64day/ 8day 68day/ 8day
D.
CantonS ♂ Y-linked
8 day
68 day
8 day
64 day
Expression (RPKM)
H3K9me2 Signal
log2(expression)
log2(fold change)
8 64
Age(Days)
Y-linked
Unbiased
H3K9me2 Signal
log2(fold change)
C.
CantonS ♂
CantonS ♀
log2(68day expression)
log2(8day expression)
Foldback Element
IR Elements
non-LTR Elements
Retroviral Elements
Satellites
Sine-like Element
E.
CantonS ♂ Y-linked
log2(68day expression)
log2(8day expression)

Figure 2

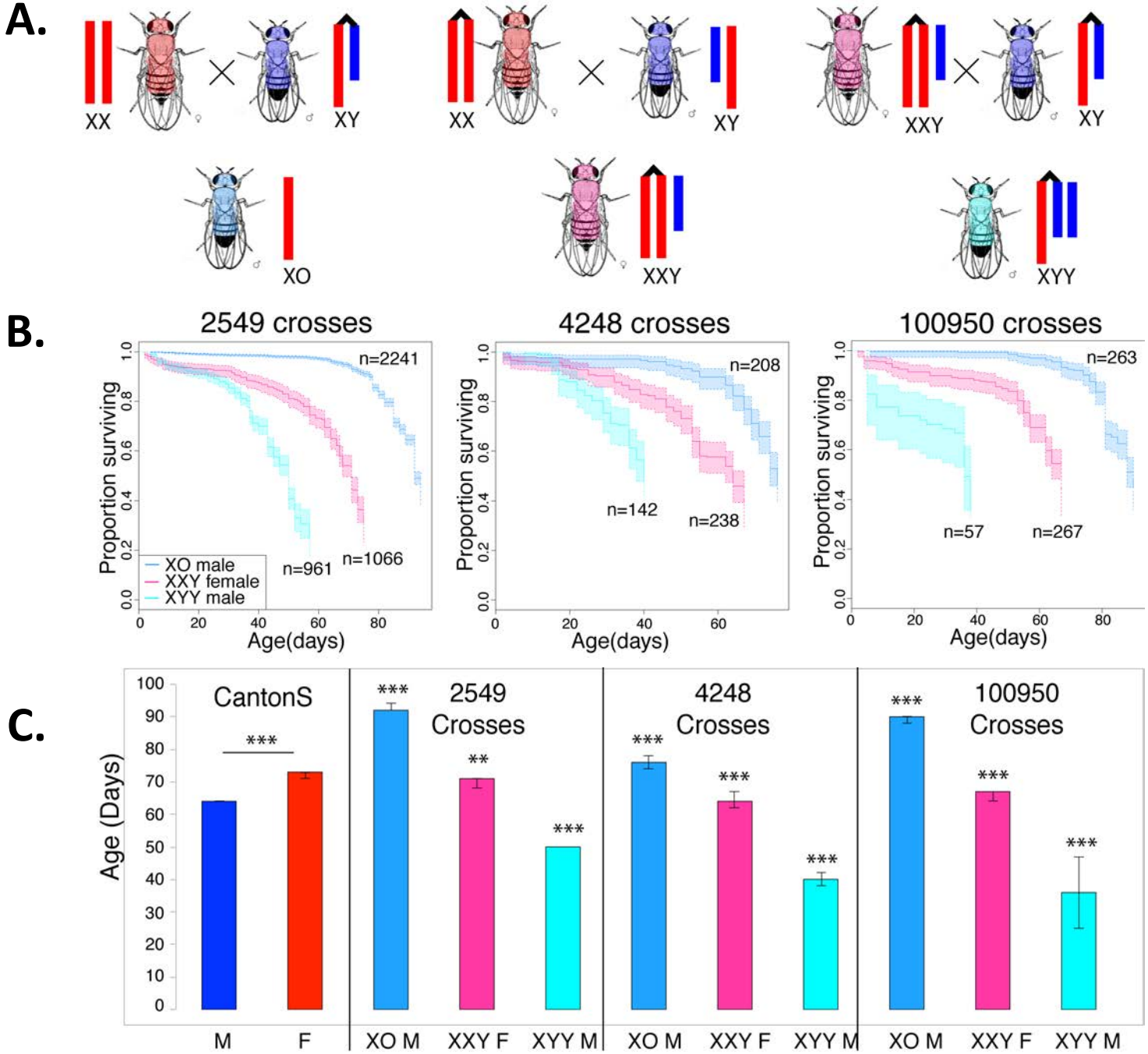
A.
XX
XY
XX
XY
XXY
XY
XO
XXY
XYY
B.
2549 crosses
4248 crosses
100950 crosses
Proportion surviving
Age(days)
n=2241
n=961
n=1066
n=208
n=142
n=238
n=263
n=57
n=267
XO male
XXY female
XYY male
C.
Age (Days)
CantonS
2549 Crosses
4248 Crosses
100950 Crosses
M
F
XO M
XXY F
XYY M

Figure 3

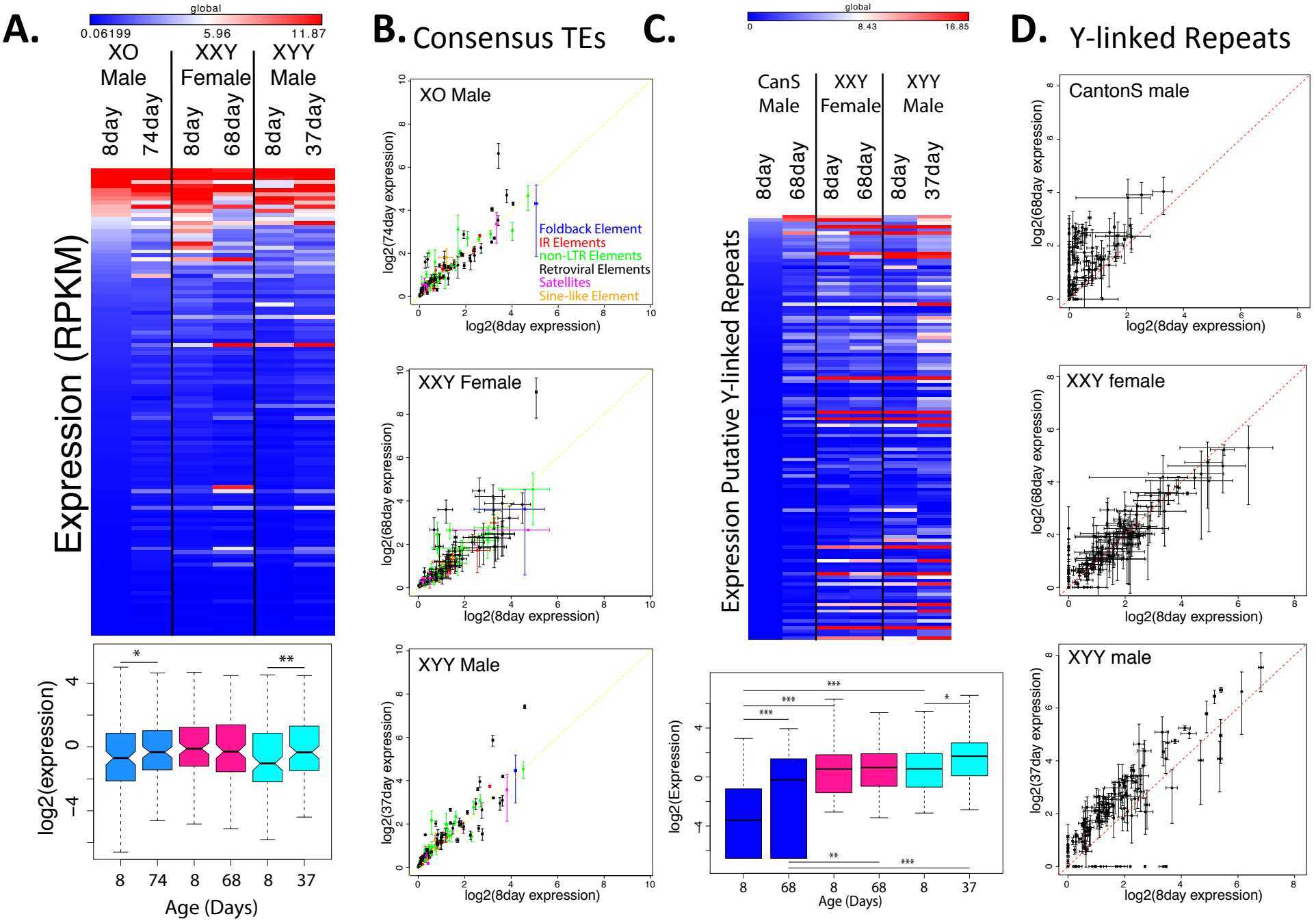
A.
global
0.06199
5.96
11.87
XO Male
XXY Female
XYY Male
8day
74day
68day
37day
Expression (RPKM)
log2(expression)
Age (Days)
B. Consensus TEs
XO Male
XXY Female
XYY Male
log2(74day expression)
log2(68day expression)
log2(37day expression)
log2(8day expression)
Foldback Element
IR Elements
non-LTR Elements
Retroviral Elements
Satellites
Sine-like Element
C.
global
0
8.43
16.85
CanS Male
Expression Putative Y-linked Repeats
log2(Expression)
D. Y-linked Repeats
CantonS male
XXY female
XYY male

Figure 4