Transposable Element Activation in Somatic Tissues Impacts *Drosophila* Physiology and Lifespan

---

by

Brian C. Jones

B.S., Fairfield University, 2009

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Division of Biology and Medicine at Brown University

Providence, Rhode Island

May 2017

Signature Page

This dissertation by Brian C. Jones is accepted in its present form by the Department of Molecular Biology, Cellular Biology, and Biochemistry as satisfying the thesis requirements for the degree of Doctor of Philosophy.

Date_____ _____

Stephen L. Helfand, MD, Advisor

Recommended to the Graduate Council

Date_____ _____

Robert A. Reenan, PhD, Committee chair

Date_____ _____

Erica N. Larschan, PhD, Committee member

Date_____ _____

Nicola Neretti, PhD, Committee member

Date_____ _____

Joshua Dubnau, PhD, Outside reader

Date_____ _____

Andrew G. Campbell, PhD, Dean of Graduate School

Curriculum vitae

Brian C. Jones

EDUCATION

Doctoral Training – Brown University **Providence, RI**
Molecular Biology, Cell Biology, and Biochemistry Graduate Program *September 2011 – January 2017*
Anticipated degree: Doctor of Philosophy (PhD)
Expected graduation: May 2017

Undergraduate – Fairfield University **Fairfield, CT**
Degree received: Bachelor of Science in Biology *September 2005 – May 2009*
Minors: Mathematics, Religious Studies
GPA: 3.62 *Cum Laude*
Honors & Awards: Dean's List, Presidential Scholarship, Sigma Xi, Pi Mu Epsilon, Honors Program with High Distinction

RESEARCH EXPERIENCE

Brown University
Department of Molecular Biology, Cell Biology, and Biochemistry **Providence, RI**
Graduate student in the laboratory of *Spring 2012 – January 2017*
Professor Stephen L. Helfand (M.D.)

- Discovered the presence of a somatic piRNA pathway in the *Drosophila* fat body and investigated its role in regulating transposable elements. *Drosophila* with mutations in the piRNA pathway displayed disrupted metabolic homeostasis, compromised fat body function, and shortened lifespan
- Examined the role of RNAi pathways in regulating transposable element expression, preserving somatic tissue function, and maintaining organismal health
- Demonstrated age-associated reactivation of transposable elements in the *Drosophila* fat body. Identified and characterized genetic interventions that suppress this reactivation and extended lifespan

Fairfield University – Department of Biology **Fairfield, CT**
Research assistant in the laboratory of Professor Olivia Harriott (Ph.D.) *Fall 2007 – Spring 2009*

- Designed a protocol to measure biofilm production and characterized capsule and rosette formation under varied culture conditions
- Aided in sequencing of bacterial isolate DNA found in larval marine sponges
- Trained in operation of CEQ 8000 sequencer, BLAST techniques, PCR, and DNA purification

CIEE Tropical Ecology and Conservation Program **Monteverde, Costa Rica**
Independent study *Spring 2008*

- Conducted fieldwork, analyzed, and presented findings on native flora and fauna including species diversity, foraging behavior, and the presence/absence of species symbiosis
- Designed and conducted study of tropical Costa Rican orchid leaf morphology (*Lepanthes helleri*) and the impact of its morphology on microclimate conditions within its curled leaf
- Presented findings of independent study on *L. helleri* at Sigma Xi Poster Session (*Spring 2009*)

New England Biolabs **Beverly, MA**
Intern *January – February 2008*
- Performed column chromatography quality control
- Assisted in the development and testing of prototype chitin beads testing protein binding capacity and bead stability

---

PROFESSIONAL EXPERIENCE

Enzymatics Inc. **Beverly, MA**
Assistant Quality Control Manager *May 2011 – July 2011*
- Supervised and trained QC analysts and technicians
- Founded Enzymatics Journal Club

Quality Control Analyst *March 2011 – May 2011*
- Designed and implemented a revised experimental protocol for a qPCR DNA contamination assay
- Developed new assays for detecting contaminating RNases

Quality Control Technician II *March 2010 – March 2011*
- Worked closely with Research and Development Department and Director in the development of new assays to characterize properties of enzyme products and prototypes
- Designed and optimized protocols and accompanying documentation for new enzyme activity assays
- Researched reaction buffers, temperature conditions, and temperature stability testing for new products
- Aided in defining and quantifying new unit definitions for products

Quality Control Technician I *September 2009 – March 2010*
- Responsible for performance of Quality Control (QC) panel assays to validate and release new product lots assuring purity and adherence to product specifications

Fairfield University – Department of Chemistry **Fairfield, CT**
Chemistry Stockperson *Spring 2007*
- Organized, prepared, and distributed chemicals and related instruments to labs

---

FELLOWSHIPS

Ruth L. Kirschstein National Research Service Award (NRSA) Individual Predoctoral Fellowship (F31)
National Institute on Aging: AG047736-01A1 *2014 – 2016*
- "Dissecting the role of RNAi in the aging fat body of *Drosophila melanogaster*"

NRSA Aging Training Grant (T32) *2012 – 2014*
National Institutes of Health: 1T32AG041688

TEACHING EXPERIENCE

Brown University **Providence, RI**
Undergraduates trained:
Matthew Finn (expected graduation: 2018), *Emily Siegel (2016), Michael Franklin (2015), *Austin Tam (2015), Mark Henriksen (2014), *Gillian Horwitz (2014)
**mentored in honors thesis projects*

Teaching assistant – Microbiology *Spring 2014*

- Reviewed course concepts and materials, graded exams and laboratory reports, led review sections, and provided experimental support during laboratory sections

Sheridan Center Certificate V Program, Research Mentorship Track *Fall 2014 – Spring 2015*

- Participated in undergraduate mentorship training program in conjunction with mentoring and guiding the scientific research projects of three undergraduate students (A. Tam, M. Franklin, and E. Siegel)

PEGASUS Program – La Salle Academy *2012 – Present*

- Instructed middle school students on the basics of fly genetics, molecular biology, and genetic engineering

SPARK – STEM I for Middle School Students *2012 – Present*

- Instructed middle school students on the basics of fly genetics, molecular biology, and genetic engineering

Fairfield University **Fairfield, CT**

Biology & Inorganic Chemistry Tutor *Spring 2007 – Spring 2009*

- Instructed students on course concepts and materials

RELEVANT EXPERTISE and ACTIVITIES

Laboratory/Research skills:

- RNA/smRNA-seq library preparation, SDS-PAGE, RNA gel extraction, qPCR, gel shift assays, bacterial culture and transformation, ^{3}H and ^{32}P based enzyme activity assays, Western blotting, *Drosophila* genetic manipulation, micro-dissection, cryosectioning, immunofluorescence, RNA/DNA isolation, PCR, metabolic profiling, general molecular biology techniques, ImageJ, Photoshop, Illustrator, Microsoft Word/Excel/Powerpoint

ASBMB Hill Day **Washington, D.C.**

Science funding advocate *April 27-28, 2016*

- Served as a representative of Rhode Island and met with Congressional senators and representatives on Capitol Hill to advocate for increased science funding for NIH and NSF

Reviewer for *Aging Cell* *July 2015 – Present*

- Selected as a scientific reviewer for small RNA biology and aging

Brown University **Providence, RI**

MCB Graduate Program Admissions Committee Student Representative *Spring 2013*

- Aided in organizing interview weekend events for prospective graduate students, interviewed candidates, and provided candidate feedback to Admissions Committee

---

AWARDS

Brown University – MCB Graduate Program Retreat **Providence, RI**

Poster winner *September 2, 2015*

- "A somatic piRNA pathway in *Drosophila* fat body regulates metabolic homeostasis"

Poster winner *August 28, 2013*

- "Effects of silencing neuronal transposable elements via RNAi on *Drosophila melanogaster* lifespan"

PUBLICATIONS and PRESENTATIONS

Brian C. Jones, Jason G. Wood, Chengyi D. Chang, Austin Tam, Michael J. Franklin, Emily R. Siegel, and Stephen L. Helfand. (2016). "A somatic piRNA pathway in the *Drosophila* fat body ensures metabolic homeostasis and normal lifespan." *Nature Communications*, 7:13856. DOI: 10.1038/ncomms13856.

Jason G. Wood, **Brian C. Jones**, Nan Jiang, Chengyi Chang, Priyan Wickremesinghe, Meyrolin Garcia, Lucas Burhenn, Nicola Neretti, and Stephen L. Helfand. (2016). "Chromatin modifying genetic interventions suppress age-associated transposable element activation and extend lifespan in *Drosophila*." *PNAS,* 113(40):11277-11282. DOI: 10.1073/pnas.1604621113

Jill A. Kreiling*, **Brian C. Jones***, Jason G. Wood, Marco De Cecco, Steven W. Criscione, Nicola Neretti, Stephen L. Helfand, and John M. Sedivy. (2017). "Contribution of Retrotransposable Elements to Aging", *Human Retrotransposons in Health and Disease.* Cristofari G (Ed.). *Springer International Publishing AG*. (*co-first authors) DOI 10.1007/978-3-319-48344-3_13

Yiannis A. Savva, James E. C. Jepson, Yao-Jen Chang, Rachel Whitaker, **Brian C. Jones**, Georges St. Laurent, Michael R. Tackett, Philipp Kapranov, Nan Jiang, Guyu Du, Stephen L. Helfand, and Robert A. Reenan. (2013). "RNA editing regulates transposon-mediated heterochromatic gene silencing." *Nature Communications*, 4:2745. DOI: 10.1038/ncomms3745

Conference talks and presentations:

Northeastern Glenn Foundation Symposium on the Biology of Aging **New Haven, CT**
Invited speaker *November 4, 2016*
- "A somatic piRNA pathway in the *Drosophila* fat body suppresses transposable elements ensuring metabolic homeostasis"

CSHL Conference on Mechanisms of Aging **Cold Spring Harbor, NY**
Invited speaker *September 26-30, 2016*
- "A somatic piRNA pathway in the *Drosophila* fat body suppresses transposable elements ensuring metabolic homeostasis"

CSHL Conference on Regulatory & Non-Coding RNAs **Cold Spring Harbor, NY**
Invited speaker *August 23-27, 2016*
- "A somatic piRNA pathway in the *Drosophila* fat body suppresses transposable elements ensuring metabolic homeostasis and normal lifespan"

Gordon Research Seminar: Biology of Aging **Newry, ME**
Invited speaker *July 18-19, 2015*
- "Activation of transposable elements across multiple tissues in aging *Drosophila*"

Gordon Research Conference: Biology of Aging **Newry, ME**
Poster presenter *July 19-24, 2015*
- "Activation of transposable elements across multiple tissues in aging *Drosophila*"

Brown University – Providence Area Aging Research Forum **Providence, RI**
Invited speaker *April 14, 2015*
- "A somatic piRNA pathway in *Drosophila* fat body with aging-associated implications"

Invited speaker *March 3, 2014*
- "Characterizing a somatic piRNA pathway in the *Drosophila* fat body"

***Drosophila* Research Conference** **Chicago, IL**
Poster presenter *March 4-8, 2015*

- "Activation of transposable elements across multiple tissues in aging *Drosophila*"

CSHL Conference on Regulatory & Non-Coding RNAs **Cold Spring Harbor, NY**
Invited speaker *August 26-30, 2014*

- "A somatic piRNA pathway in *Drosophila* fat body with aging-associated implications"

Acknowledgements

Thank you to all the members of the Helfand lab, past and present. Thank you to Stephen Helfand for giving me the opportunity to join the lab and the independence to pursue the project ideas that most interested me. Thank you to Jason Wood for your wise words, calming presence, good humor, and technical expertise. I will be forever grateful for your unending patience with all of my questions and for the joy that your friendship has brought me. I will miss our Pokéwalks, long chats, and corny jokes. Thank you to Rachel Whitaker for being one of the best lab mates I could ever ask for. You were the first person I met at Brown and one of the many reasons I chose to pursue my graduate studies here. Thank you for being a constant friend for all of these years. Thank you to Chengyi Chang for being my lab-mother. You have always set me straight when I needed it most and give me the honest guidance I need. You have encouraged and supported me every step of the way and for this I will always be thankful. Finally, thank you to Suzanne Hosier, Nan Jiang, Will Lightfoot, and Jackson Taylor for technical support, good times, and good conversations.

Thank you to my committee members, Robert Reenan, Erica Larschan, and Nicola Neretti for all the advice and guidance you have given me. Thank you also to Joshua Dubnau for being my outside reader. Thank you also to Mark Johnson and especially Judith Bender and Elaine Butler for all of your support and help; it means the world to me.

Thank you to Yiannis Savva for being one of the best mentors I could ask for. You have challenged me to think deeper and to make my science better every step of the way. You taught me what it means to push the envelope and most importantly to always listen to Nature as she speaks, even if I don't always understand what she is saying.

Thank you to all of the undergrads that I have worked with including Emily Siegel, Mike Franklin, Austin Tam, Mark Henriksen, and Gillian Horwitz. Every single one of you made life in lab so much more worthwhile with your presence and your friendship. You have made me laugh, challenged me with your questions, and taught me more so much about myself and about the mentor that I want to become. I will miss you all tremendously.

Thank you to the Rand and Tatar labs for being a second lab family. Thank you to Jim Mossman, Ana Hernandez, Adam Spierer, Eugenia Villa-Cuesta, Aleksandra Norton, and Rochele Yamamoto for years of laughter, good food and drink, expertise, and most of all, your friendship.

Thank you to my professors at Fairfield University for your guidance and mentoring; each of you were instrumental in putting me on this path. Thank you specifically to Tod Osier, my undergraduate advisor, and Angella Biselli for encouraging me to pursue scientific research. Thank you to Olivia Harriott, my undergraduate PI, for giving me my first research experience in her lab. Finally, thank you to Ashley Byun McKay, my

evolutionary biology professor, who first suggested to me (and continually encouraged on numerous occasions) that I should pursue graduate school.

Thank you to Jeff Maycock. In many ways, you were my first scientific mentor and I am so lucky to have been able to work with you, been taught by you, and to share in your friendship. I will always remember fondly the days working beside you at the hot room bench where I earned my “lab hands” to the sounds of Modest Mouse. Thank you for pushing me, supporting me, and for befriending me from the first day we met.

Thank you to my friends, Kristen Girard, Colleen Gibson, Micahel Csorba, Leila Rieder, Bennett Prescott, Selena Gell, Allena Buchholz, and Lior Sapir. You are all more than friends to me; you are my second family. Thank you for all the memories we have made together. Thank you for welcoming me into your lives and for the fits of laughter, deep conversations, camping trips, road trips, board games, late night adventures, dance parties, inside jokes, and Friendsgivings we have shared over the years. I love you all so much.

Thank you to my partner, Marissa Holmbeck. I am so lucky to have met you and am thankful for you every single day. From the first time we met on my interview weekend with you as my guide, I could never have imagined how much you would change my life for the better. I am a better person because of you and I love you more than you know. Thank you for always being there for me and for all of your support.

Thank you to my family: Mom, Nana, Papa, Steve, Mike, and Nancy. You have all been there for me every step of the way and I can't thank you enough. I love you all so much and am so thankful for each of you. You have kept me going with your love and encouragement. You each inspire me to do better, try harder, and love deeper.

Thank you to those who I've lost along the way: Dad, Nana, Father Frank O'Brien, and Max. It is because I lost each one of you that I began this journey studying the biology of aging and I wish you were all here to see this chapter of my life come to a close. I miss you all so much, but I carry each of you with me in my heart and my memory every single day. Thank you for the mark that each of you left on me while you were still here; I will never forget you.

Finally, I give thanks for the two principles that drive every aspect of my life, both personal and scientific: awe and stewardship. Awe for the natural world is what has inspired and thrilled me along every step of this journey. It has also reminded me in more difficult times of the true beauty of scientific work. The wonder and joy that the privilege of scientific discovery brings me is like nothing else in this world, and I am forever thankful for it. I am also thankful for the possibility that my work may one day be of some positive use to humankind. I am proud to call myself a scientist in an endeavor of discovery that gives both extraordinarily beautiful meaning to my own life, but that also manifests the possibility of a better life for others.

Table of Contents

List of Figures

Appendix A:

List of Tables

Chapter 1

General Introduction

Portions of this work were adapted from my own contributions to *Kreiling & ***Jones, et al.** (2017) "Contribution of Retrotransposable Elements to Aging." *Human Retrotransposons in Health and Disease* (see Appendix B).

Transposable elements (TEs) are parasitic agents that rely upon host genomes for their mobility and reproduction. Most active TEs either encode or coopt proteins that allow them to move from one genomic location to another. Due to their high copy number and replicative capacity, TEs contribute significantly to the overall composition of their host genomes[1]. Their ability to mobilize allows TEs to invade and integrate themselves into host genomes, and in doing so directly contributes to genomic damage, gene, cellular, and tissue dysfunction, and even disease states. Thankfully, multiple cellular mechanisms have evolved to effectively silence TEs, including heterochromatization and small RNA pathways. In order to prevent and protect cells from the deleterious effects of TE reactivation, it is essential to understand the mechanisms by which host genomes suppress TEs thereby preserving genomic integrity and cellular function.

TE activity: mobilization and silencing

TEs are classified based upon their method of mobilization. The two main classes of TEs are the DNA transposons and the retrotransposons[2] (Figure 1). Most DNA transposons mobilize via their dependence upon transposase proteins, which recognize target sequences in the TE and excise it from the donor site before again cutting DNA at a new target site and inserting the TE into this new locus (Figure 1). In this sense, DNA transposon mobilization is largely non-replicative, with the exception being target site duplication events caused by a staggered end cleavage from transposase activity. The second class of TEs, retrotransposons, mobilize via an RNA intermediate, which is reverse transcribed into a cDNA copy before being inserted into a

new locus (Figure 1). Retrotransposons are further divided into three main families: the long terminal repeat (LTR) TEs, the long interspersed nuclear elements (LINE), and the short interspersed nuclear elements (SINE). Like DNA transposons, LTR TEs and LINEs encode polycistronic RNAs that code for proteins necessary for their mobilization. Retrotransposons also make up large portions of eukaryotic genomes, including fly and mammalian genomes. These coding TEs are able to mobilize semi-autonomously and depend only upon the host genome's transcriptional and translational machinery. SINEs are non-coding and depend upon proteins encoded by other TEs. Regardless of their mode of replication or mobilization, host genomes are at risk from TE reactivation as they are known to cause insertional mutagenesis, chromosomal rearrangements, double strand DNA breaks (DSBs), RNA toxicity, apoptosis, and tissue dysfunction. Therefore, organisms have evolved two main mechanisms by which cells are able to silence and suppress TE reactivation: heterochromatinization and small RNA pathways.

Heterochromatin silences TE transcription

Chromatin consists of DNA wrapped around a histone protein octamer core that is then packed into higher order structures. The degree of compaction often dictates how accessible a gene is to transcriptional machinery and is often classified as either euchromatic, open and permissible to transcription, or heterochromatic, closed and largely inaccessible to transcription. A region of the genome is able to change from one state to the other and in doing so is a reflection of the impermanence of such epigenetic changes as well as the need for a cell to both maintain and respond to these changing states. Epigenetic changes can be made by the modification of histone tails, methylation

of DNA, or the establishment of chromatin-binding proteins, all of which are able to alter chromatin structure. Establishment of heterochromatin is one way cells are able to silence gene expression, and is essential for prevention of TE reactivation. Cells accomplish this by recruiting protein and DNA modifying enzymes that methylate DNA or establish histone marks such as methylation of histone H3 at lysine 9 (H3K9me) or at lysine 27 (H3K27me). TEs are enriched for these marks and loss of the genes that establish them, such as the histone methyltransferase *Su(var)3-9*, results in elevated levels of TE transcripts[3,4] (Figure 2). In mammals, methylation of cytosine residues (CpG) on the DNA itself can also mark TE sequences for transcriptional silencing. Loss of the DNA methyltransferase, DNMT1, similarly results in loss of silencing and TE reactivation[3,5] as does global DNA methylation[6]. These silencing mechanisms largely alter the physical structure of chromatin, but they can also be reinforced by other non-coding means of transcriptional repression.

RNAi pathways post-transcriptionally suppress TEs

Heterochromatin can also be established by RNA interference (RNAi) pathways. Largely conserved across species from plants to animals, RNAi employs small RNAs (smRNAs) in regulating canonical protein coding genes, viruses, and especially TEs[7]. TEs are repressed by RNAi at two levels: transcriptional regulation via recruitment of repressive heterochromatic marks and at the level of post-transcriptional regulation via cleavage of target mRNA. The two smRNA pathways known to regulate TEs in animals are the short interfering RNA (siRNA) and the Piwi-interacting RNA (piRNA) pathways (Figure 2). Each pathway differs slightly in its effector proteins, smRNA biogenesis, and

modes of silencing. Finally, while both pathways are well known for their ability to silence TEs, the siRNA pathway is active in all tissues and the piRNA pathway is best known for its role in gonadal tissues[3,8].

In flies and mammals, the siRNA pathway employs ~21nt siRNAs derived from cleavage of long double-stranded (dsRNA) substrates by the protein Dicer (Dicer in mammals and Dicer-2 in flies)[9–12] (Figure 2). These siRNAs are then loaded onto an Argonaute (AGO) effector protein, forming an RNA-induced silencing complex (RISC), which uses its siRNA to target and cleave homologous mRNAs in the cytoplasm. While mammals have four Argonaute proteins (AGO1-4) that compete for and bind smRNAs, only AGO2 is catalytically active[13,14]. Mammalian argonautes are also able to bind and use both siRNAs and the similarly-sized micro RNAs (miRNAs) that largely target coding genes for post-transcriptional silencing[15,16]. The siRNA-RISC can promote transcriptional silencing by moving to the nucleus where it recruits chromatin-modifying enzymes promoting the formation of heterochromatin at the site of TE transcription[3,17].

The piRNA pathway, another smRNA pathway known to silence TEs, operates through a similar mechanism of silencing whereby large genomic regions consisting of fragmented and intact TEs, called piRNA clusters, transcribe large single-stranded RNA precursors that are then processed into smaller 23-32nt piRNAs (Figure 2). piRNAs are also loaded into pathway specific Piwi clade argonaute proteins, forming piRNA-RISCs[8,12]. In flies, the three piRNA argonautes, Piwi, Aubergine (Aub), and Argonaute-3 (AGO3), play different roles and operate in two different piRNA pathways: the primary and secondary piRNA pathways. In the primary piRNA pathway, Piwi is loaded with antisense primary piRNAs that largely display a first position nucleotide bias for uracil.

These argonaute-piRNA complexes form piRNA-RISCs, similar to the siRNA pathway, which are then able to target TEs in the nucleus at the site of transcription via recruitment of chromatin-modifying enzymes, such as Su(var)3-9. Primary piRNAs also provide smRNAs for the secondary piRNA pathway, also known as the ping-pong pathway, as they are loaded onto Aubergine in the cytoplasm. Aubergine uses these piRNAs to cleave complementary sense TE transcripts. This cleavage event, due to the 1^{st} position uracil bias of the primary piRNAs, creates a cleaved complementary transcript with a 10^{th} nucleotide bias for adenosine. These cleaved transcripts are further processed into shorter piRNAs called secondary piRNAs, which are then bound by AGO3. AGO3 uses these secondary piRNAs to cleave homologous antisense transcripts, most of which are derived from piRNA clusters. This process forms more antisense piRNAs that are again loaded into Aubergine, forming a heterotypic amplification loop that post-transcriptionally silences TEs. Finally, these pathways are tissue specific in the ovary of flies; the primary piRNA pathway is active in both the somatic follicle cells of the ovary and the germline while the secondary piRNA pathway is active only in the germline. However, recent evidence has begun to suggest that the piRNA pathway may be active in other non-gonadal tissues and have putative roles in these tissues.

A role for the piRNA pathway in somatic tissues

Recent evidence has shown that the piRNA pathway is both present and active in tissues outside of the gonad suggesting new roles for this pathway in regulation of genome integrity and tissue function. Studies have shown examples of piRNA activity in

a diversity of non-gonadal tissues including stem cell populations, cancer types, and normal differentiated tissues[18,19]. For example, multiple animal models with the ability to undergo whole body regeneration, including hydra, planaria, and ascidians, maintain pluripotent and totipotent stem cell populations with active piRNA pathways[19–28]. Interestingly, loss of piRNA argonautes in these models negatively impacts the regenerative capacity of these animals[24,27,28]. In *Hydra*, an immortal model system, loss of the piRNA pathway in the interstitial stem cell population results in loss of epithelial integrity and death[21]. Similarly, loss of piRNA pathway genes in immortal planaria results in a loss of both regenerative capacity and animal viability[29,30]. piRNA pathway genes are also known to increase in expression upon reprogramming of a variety of induced pluripotent stem cell lines[19,31,32]. Mouse mesenchymal stem cells and human stem cell populations such as hematopoietic stem cells[33] also exhibit piRNA gene expression with the later study showing that is lost upon differentiation, suggesting its expression is specific to the undifferentiated state of the cell.

piRNA pathway genes are also known to be expressed in multiple types of human cancers and correlate with poor prognosis[19,34–36]. Expression of piRNA pathway genes has been identified in breast cancer, colon cancer, gastric cancer, renal cancer, seminomas, and endometrial carcinoma[18]. However, it is not yet known what role the piRNA pathway plays in cancer biology. It is possible that the piRNA pathway, arguably the premier genomic defense against TE reactivation in the gonads, is activated as a compensatory response to TE derepression in somatic and cancerous tissues. Interestingly, recent studies in flies have shown that tumorigenesis activates piRNA pathway components[37,38] and loss of *piwi* results in reduced tumor growth suggesting

an active role for the piRNA pathway in cancer cell function.

Finally, signatures of piRNAs and piRNA argonautes have been found in healthy, differentiated, non-gonadal somatic tissues of flies, mice, macaques, and humans[39,40]. However, few studies have included appropriate negative controls to account for possible tissue contamination. Multiple studies have suggested that the piRNA pathway may be active in different tissues of the brain across multiple species[18]. One example in *Aplysia* demonstrated the presence of an active piRNA pathway in neuronal tissues and that Piwi was required for long term memory formation[41]. While Aubergine and AGO3 have also shown to be present in fly brains, there is little evidence for the existence of neuronal piRNAs that are necessary for canonical piRNA pathway gene silencing[42]. Another study also identified an active piRNA pathway in the mouse hippocampus that impacted dendritic spine growth[40]. Much of the early work characterizing the piRNA pathway in differentiated somatic tissues is descriptive at best and experiments determining a mechanistic role for its presence in the soma have yet to be performed.

These smRNA pathways are essential to preventing the genomic damage caused by the reactivation of TEs and may also have additional roles in maintaining normal cell and tissue function. Evidence also suggests that loss of TE suppression, both in somatic and reproductive tissues, may be closely linked to disease and aging phenotypes.

TE reactivation in disease and aging

Recent evidence has begun to suggest that TE reactivation may contribute to a variety of diseases. Of TEs in the human genome, the non-LTR LINE-1 (L1) and the *Alu*

SINE elements are both active and the most abundant making up approximately 76% of all TEs and 31% of the human genome[43]. Specifically, multiple studies have shown reactivation of TEs in cancer[44–47]. In transformed cancer cell lines, RNA Polymerase II was shown to bind more frequently to retrotransposons and this correlated with significant upregulation of L1 RNA expression[48]. Both L1 and *Alu* elements are associated with and are known to contribute to multiple cancer types including breast, ovarian, colon, and prostate cancers[44,49]. Therefore, the mutagenic activity of these elements can directly contribute to tumorigenesis as their mobilization and insertion into oncogenes and proto-oncogenes can result in compromised DNA damage repair or abnormal cell growth.

Over the past few decades, the activity of TEs in neuronal tissues challenged previous dogma regarding neuronal genomes as static[50]. However, additional work has begun to show that TEs directly contribute to diseases associated with neuronal degeneration. For example, the neurodegenerative disorders amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) are both characterized by aggregation of the protein TDP-43, a promiscuous RNA-binding protein whose loss of function contributes to neuron death[50–52]. In one study, tissue from patients with FTLD showed less binding of TDP-43 to transposon sequences relative to controls[53,54] and mice with loss of TDP-43 function showed elevated levels of LINE, SINE, and LTR retrotransposons normally bound by TDP-43[54]. In another example, the neurodegenerative disorder fragile X-associated tremor/ataxia syndrome (FXTAS) is characterized by an expansion of a CGG trinucleotide repeat in the *FMR1* gene and, in the fly model for this disease, TE transcript levels are significantly increased including

the *I-element*, *gypsy*, and *copia* transposons[55]. This study also found that reduced number of CGG repeats correlated with reduced increases of *gyspy* and *copia* transcript levels, suggesting that the observed TE reactivation may be dependent upon number of CGG repeats. Finally, repression of *gypsy* in this model was able to suppress the neurodegenerative eye phenotype observed in expanded CGG repeat flies. TE reactivation may therefore contribute directly to the pathology of multiple neurodegenerative diseases.

TE activity is also known to associate with age-associated diseases and even normal aging. One study examining geographic atrophy, an age-associated form of macular degeneration, showed that *DICER1* was decreased and this correlated with elevated levels of *Alu* expression[56,57]. This phenotype was able to be trangenically phenocopied by *Alu* overexpression or transfection and reversed by *DICER1* overexpression suggesting that *Alu*-induced RNA toxicity mediates this age-dependent disorder[57]. More recently, it has been hypothesized that reactivation of TEs may not be limited to disease states, but also occurs during normal aging and may contribute directly to aging phenotypes[58].

Multiple recent studies suggest that TEs become reactivated in aging somatic tissues[59]. In yeast, chronological aging resulted in increased levels of TE transposition, which correlated with increased levels of chromosomal rearrangements[60]. In *C. elegans*, germ cells exhibit increased TE transcript levels with age[61]. More recently, increases in transcript levels of *gypsy* and *R2* TEs in flies were observed and this correlated with higher rates of transposition of the *gypsy* TE in aged fly neurons[62]. These phenotypes also appear to be conserved in mammals as human replicatively senescent cells and

both aged mouse liver and skeletal muscle show increased transcript levels and transposition of TEs[63,64]. Interestingly, caloric restriction, a robust method known to extend the lifespan of multiple model organisms, delayed the onset of TE reactivation in aged mouse somatic tissues[63] suggesting that the rate of aging may be linked to reactivation of TEs.

Loss of TE suppression pathways correlates with aging

In the past decade, an increasing number of studies have begun to demonstrate that when TE suppression pathways are compromised, this correlates with aging phenotypes[58,59]. Multiple studies have shown genome-wide changes in methylation patterns, with repetitive regions becoming hypomethylated with age[65–67] and a global loss of repressive histone marks[68–71]. Flies also exhibit a similar global, age-dependent loss of silencing marks including H3K9me3 and HP1[72]. In senescent human cell culture, elevated levels of L1 and *Alu* transcripts correlated with heterochromatic regions of the genome becoming more open[64]. Another study also demonstrated that repetitive DNA elements became hypomethylated with age, with some TEs showing loss of the silencing marks H3K9me3 and H4K20me3, suggesting epigenetic derepression of TEs with age[73]. These data suggest a conserved phenotype of aging whereby the genome experiences a loss of repressive marks over TEs with age therefore allowing their activation and a renewed potential for genomic damage.

Finally, mutants for genes in smRNA pathways known to suppress TEs, such as the siRNA or piRNA pathways, consistently demonstrate a dramatic upregulation in TE transcript levels[74,75]. This correlates with an increase in transposition and a change in

heterochromatic marks associated with TEs[17,76,77]. In flies, mutants for the siRNA genes *Dcr-2* and *Ago-2* have dramatically shortened lifespans and this correlates with significant reactivation of TEs[10,11,62,78]. Interestingly, while TEs have been shown to reactivate with age across multiple species, the transcript levels of RNAi genes that regulate TEs are not known to decline with age and in fact remain constant[62,79]. Additionally, it remains unknown if the pool of smRNAs that target TEs for silencing changes with age and whether this too might impact suppression of TEs. This suggests that RNAi machinery may prevent further TE reactivation with age and that inhibition or enhancement RNAi effectiveness in silencing TEs may negatively or positively impact lifespan, respectively.

Reactivation of TEs can compromise genomic integrity, disrupt cellular and tissue function, and even contribute to disease. Thankfully, multiple cellular mechanisms are employed by host genomes in order to suppress TEs and prevent such deleterious phenotypes. These mechanisms, including the establishment of heterochromatin and smRNA pathways, are critical to maintaining cellular and organismal health. This thesis explores these themes in two chapters and additional accompanying appendices. Chapter 2 examines the role of a somatic piRNA pathway in suppressing TE reactivation in the fat body of *Drosophila* and characterizes the physiological consequences of its loss. Chapter 3 examines the age-dependent derepression of TEs in somatic *Drosophila* tissues and the different genetic and pharmacological interventions that attenuate TE reactivation. Together, these data provide new insight and define additional novel roles for TE suppression pathways in maintaining somatic tissue function and overall organismal health.

References

1. Cordaux, R. & Batzer, M. a. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **10,** 691–703 (2009).

2. Huang, C. R. L., Burns, K. H. & Boeke, J. D. Active Transposition in Genomes. *Annu. Rev. Genet.* **46,** 651–675 (2012).

3. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8,** 272–85 (2007).

4. Martens, J. H. A. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24,** 800–12 (2005).

5. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosinemethylation. *Nat. Genet.* **20,** 116–117 (1998).

6. Levin, H. L. & Moran, J. V. Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* **12,** 615–27 (2011).

7. Shabalina, S. & Koonin, E. Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* **23,** 578–587 (2008).

8. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10,** 94–108 (2009).

9. Kawamura, Y. *et al.* Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453,** 793–7 (2008).

10. Ghildiyal, M. *et al.* Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320,** 1077–81 (2008).

11. Czech, B. *et al.* An endogenous small interfering RNA pathway in Drosophila. *Nature* **453,** 798–802 (2008).

12. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128,** 1089–103 (2007).

13. Liu, J. *et al.* Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science (80).* **305,** (2004).

14. Meister, G. *et al.* Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Mol. Cell* **15,** 185–197 (2004).

15. Hutvágner, G. & Zamore, P. D. A microRNA in a Multiple-Turnover RNAi Enzyme Complex. *Science (80-.).* **297,** (2002).

16. Martinez, J. *et al.* Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110,** 563–74 (2002).

17. Fagegaltier, D. *et al.* The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 21258–63 (2009).

18. Ross, R. J., Weiner, M. M. & Lin, H. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* **505,** 353–9 (2014).

19. Van Wolfswinkel, J. C. Piwi and Potency: PIWI Proteins in Animal Stem Cells and Regeneration. *Integr. Comp. Biol.* 1–14 (2014). doi:10.1093/icb/icu084

20. Mochizuki, K., Nishimiya-Fujisawa, C. & Fujisawa, T. Universal occurrence of the vasa -related genes among metazoans and their germline expression in Hydra. *Dev. Genes Evol.* **211,** 299–308 (2001).

21. Juliano, C. E. *et al.* PIWI proteins and PIWI-interacting RNAs function in Hydra somatic stem cells. *Proc. Natl. Acad. Sci.* **111,** 337–342 (2014).

22. Lim, R. S. M., Anand, A., Nishimiya-Fujisawa, C., Kobayashi, S. & Kai, T. Analysis of Hydra PIWI proteins and piRNAs uncover early evolutionary origins of the piRNA pathway. *Dev. Biol.* **386,** 237–251 (2014).

23. Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sanchez Alvarado, A. SMEDWI-2 Is a PIWI-Like Protein That Regulates Planarian Stem Cells. *Science (80-.).* **310,** 1327–1330 (2005).

24. Palakodeti, D., Smielewska, M., Lu, Y.-C., Yeo, G. W. & Graveley, B. R. The PIWI proteins SMEDWI-2 and SMEDWI-3 are required for stem cell function and piRNA expression in planarians. *RNA* **14,** 1174–86 (2008).

25. De Mulder, K. *et al.* Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. *Dev. Biol.* **334,** 198–212 (2009).

26. Zhou, X. *et al.* Dual functions of Macpiwi1 in transposon silencing and stem cell maintenance in the flatworm Macrostomum lignano. *RNA* **21,** 1885–97 (2015).

27. Shibata, N. *et al.* Inheritance of a Nuclear PIWI from Pluripotent Stem Cells by Somatic Descendants Ensures Differentiation by Silencing Transposons in Planarian. *Dev. Cell* **37,** 226–237 (2016).

28. Rinkevich, Y. *et al.* Piwi positive cells that line the vasculature epithelium, underlie whole body regeneration in a basal chordate. *Dev. Biol.* **345,** 94–104 (2010).

29. Guo, T., Peters, A. H. F. M. & Newmark, P. A. A bruno-like Gene Is Required for Stem Cell Maintenance in Planarians. *Dev. Cell* **11,** 159–169 (2006).

30. Wagner, D. E., Ho, J. J. & Reddien, P. W. Genetic Regulators of a Pluripotent Adult Stem Cell System in Planarians Identified by RNAi and Clonal Analysis. *Cell Stem Cell* **10,** 299–311 (2012).

31. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454,** 766 (2008).

32. Samavarchi-Tehrani, P. *et al.* Functional Genomics Reveals a BMP-Driven Mesenchymal-to-Epithelial Transition in the Initiation of Somatic Cell Reprogramming. *Cell Stem Cell* **7,** 64–77 (2010).

33. Sharma, A. K. *et al.* Human CD34+ stem cells express the hiwigene, a human homologue of the Drosophila genepiwi. *Blood* **97,** (2001).

34. Sun, G. *et al.* Clinical significance of Hiwi gene expression in gliomas. *Brain Res.* **1373,** 183–188 (2011).

35. Zhao, Y.-M. *et al.* HIWI is associated with prognosis in patients with hepatocellular carcinoma after curative resection. *Cancer* **118,** 2708–2717 (2012).

36. Suzuki, R., Honda, S. & Kirino, Y. PIWI Expression and Function in Cancer. *Front. Genet.* **3,** 204 (2012).

37. Fagegaltier, D. *et al.* Oncogenic transformation of Drosophila somatic cells induces a functional piRNA pathway. *Genes Dev.* **30,** 1623–35 (2016).

38. Sumiyoshi, T. *et al.* Loss of l(3)mbt leads to acquisition of the ping-pong cycle in Drosophila ovarian somatic cells. *Genes Dev.* **30,** 1617–22 (2016).

39. Yan, Z. *et al.* Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Res.* **39,** 6596–607 (2011).

40. Lee, E. J. *et al.* Identification of piRNAs in the central nervous system. *RNA* **17,** 1090–9 (2011).

41. Rajasethupathy, P. *et al.* A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity. *Cell* **149,** 693–707 (2012).

42. Perrat, P. N. *et al.* Transposition-Driven Genomic Heterogeneity in the Drosophila Brain. *Science* **340,** 91–95 (2013).

43. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).

44. Chénais, B. Transposable elements and human cancer: A causal relationship? *Biochim. Biophys. Acta - Rev. Cancer* **1835,** 28–35 (2013).

45. Doucet-O'Hare, T. T. *et al.* LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **112,** E4894-900 (2015).

46. Ewing, A. D. *et al.* Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. *Genome Res.* **25,** 1536–45 (2015).

47. Rodić, N. *et al.* Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. *Nat. Med.* **21,** 1060–1064 (2015).

48. Criscione, S. W., Zhang, Y., Thompson, W., Sedivy, J. M. & Neretti, N. Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics* **15,** 583 (2014).

49. Lee, E. *et al.* Landscape of somatic retrotransposition in human cancers. *Science* **337,** 967–71 (2012).

50. Reilly, M. T., Faulkner, G. J., Dubnau, J., Ponomarev, I. & Gage, F. H. The Role of Transposable Elements in Health and Diseases of the Central Nervous System. *J. Neurosci.* **33,** (2013).

51. Cohen, T. J., Lee, V. M. Y. & Trojanowski, J. Q. TDP-43 functions and pathogenic mechanisms implicated in TDP-43 proteinopathies. *Trends Mol. Med.* **17,** 659–667 (2011).

52. Buratti, E. & Baralle, F. E. The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. *RNA Biol.* **7,** 420–429 (2010).

53. Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14,** 452–458 (2011).

54. Li, W. *et al.* Transposable Elements in TDP-43-Mediated Neurodegenerative Disorders. *PLoS One* **7,** e44099 (2012).

55. Tan, H. *et al.* Retrotransposon activation contributes to fragile X premutation rCGG-mediated neurodegeneration. *Hum. Mol. Genet.* **21,** 57–65 (2012).

56. Kaneko, H. *et al.* DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature* **471,** 325–330 (2011).

57. Tarallo, V. *et al.* DICER1 loss and Alu RNA induce age-related macular degeneration via the NLRP3 inflammasome and MyD88. *Cell* **149,** 847–59 (2012).

58. Gorbunova, V., Boeke, J. D., Helfand, S. L. & Sedivy, J. M. Sleeping dogs of the genome. *Science (80-.).* **346,** (2014).

59. Wood, J. G. & Helfand, S. L. Chromatin structure and transposable elements in organismal aging. *Front. Genet.* **4,** 274 (2013).

60. Maxwell, P. H., Burhans, W. C. & Curcio, M. J. Retrotransposition is associated with genome instability during chronological aging. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 20376–81 (2011).

61. Dennis, S. *et al.* C. elegans Germ Cells Show Temperature and Age-Dependent Expression of Cer1, a Gypsy/Ty3-Related Retrotransposon. *PLOS Pathog.* **8,** e1002591 (1996).

62. Li, W. *et al.* Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat. Neurosci.* **16,** 529–31 (2013).

63. De Cecco, M. *et al.* Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging (Albany. NY).* **5,** 867–83 (2013).

64. De Cecco, M. *et al.* Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell* **12,** 247–256 (2013).

65. Avrahami, D. *et al.* Aging-Dependent Demethylation of Regulatory Elements Correlates with Chromatin State and Improved β Cell Function. *Cell Metab.* **22,** 619–632 (2015).

66. Fernández, A. F. *et al.* H3K4me1 marks DNA regions hypomethylated during aging in human stem and differentiated cells. *Genome Res.* **25,** 27–40 (2015).

67. Sun, D. *et al.* Impacts of Chromatin States and Long-Range Genomic Segments on Aging and DNA Methylation. *PLoS One* **10,** e0128517 (2015).

68. Scaffidi, P. & Misteli, T. Lamin A-Dependent Nuclear Defects in Human Aging. *Science (80-.).* **312,** (2006).

69. Shumaker, D. K. *et al.* Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8703–8 (2006).

70. O'Sullivan, R. J., Kubicek, S., Schreiber, S. L. & Karlseder, J. Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nat. Struct. Mol. Biol.* **17,** 1218–1225 (2010).

71. Zhang, W. *et al.* A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science (80-.).* **348,** (2015).

72. Wood, J. G. *et al.* Chromatin remodeling in the aging genome of Drosophila. *Aging Cell* **9,** 971–8 (2010).

73. Ryu, S. H., Kang, K., Yoo, T., Joe, C. O. & Chung, J. H. Transcriptional repression of repeat-derived transcripts correlates with histone hypoacetylation at repetitive DNA elements in aged mice brain. *Exp. Gerontol.* **46,** 811–818 (2011).

74. Vagin, V. V *et al.* A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313,** 320–4 (2006).

75. Rozhkov, N. V, Hammell, M. & Hannon, G. J. Multiple roles for Piwi in silencing Drosophila transposons. 400–412 (2013). doi:10.1101/gad.209767.112.

76. Jones, B. C. *et al.* A somatic piRNA pathway in the Drosophila fat body ensures metabolic homeostasis and normal lifespan. *Nat. Commun.* **7,** 13856 (2016).

77. Gu, T. & Elgin, S. C. R. Maternal depletion of Piwi, a component of the RNAi system, impacts heterochromatin formation in Drosophila. *PLoS Genet.* **9,** e1003780 (2013).

78. Lim, D.-H. *et al.* The endogenous siRNA pathway in Drosophila impacts stress resistance and lifespan by regulating metabolic homeostasis. *FEBS Lett.* **585,** 3079–85 (2011).

79. Abe, M. *et al.* Impact of age-associated increase in 2'-O-methylation of miRNAs on aging and neurodegeneration in Drosophila. *Genes Dev.* **28,** 44–57 (2014).

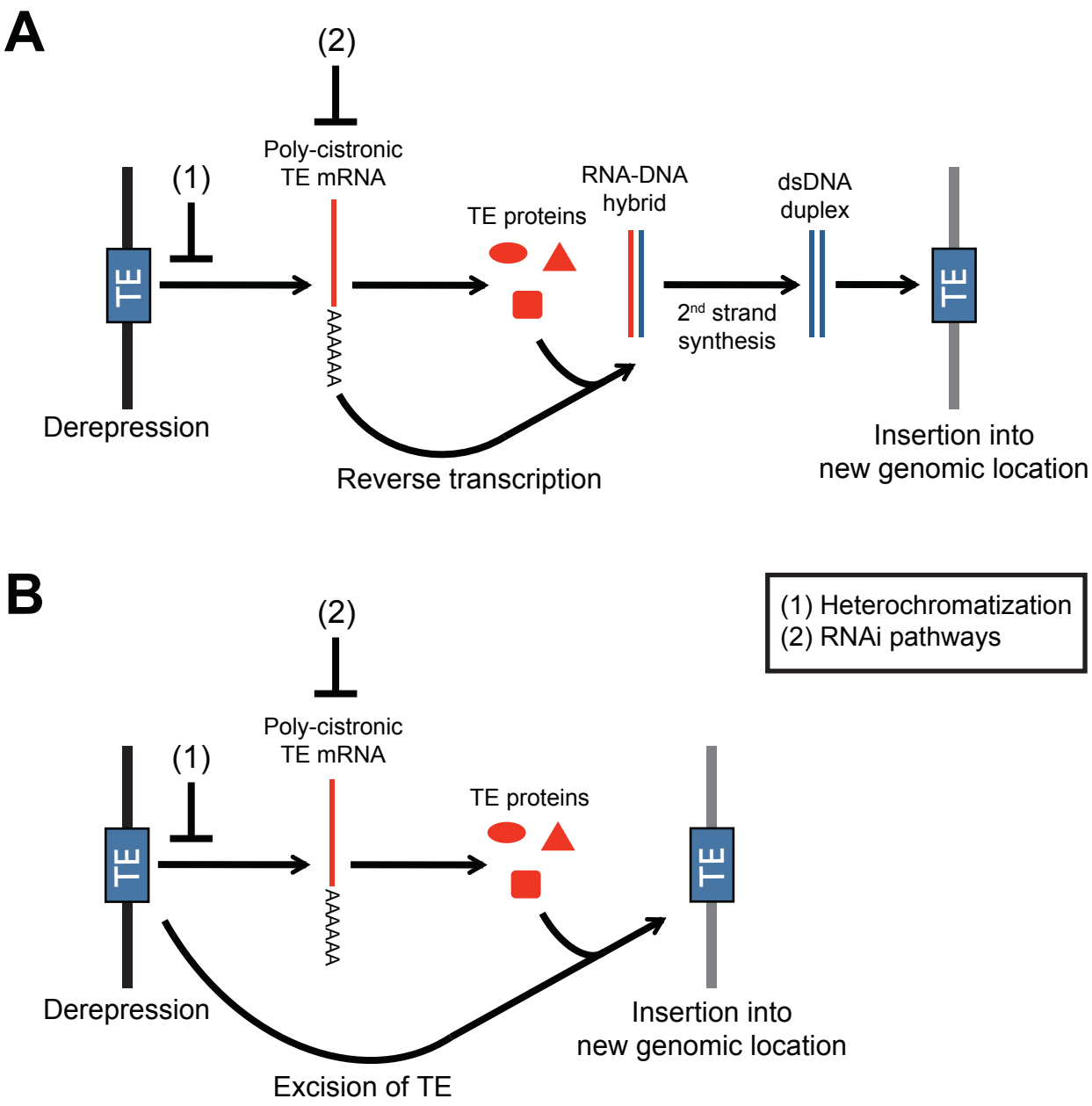


Figure 1. Transposable element mobilization.
(**A**) Class I TEs, retrotransposons, mobilize in two steps. Generally, the TE is first transcribed into a poly-cistronic mRNA copy that codes for proteins necessary for its mobilization. The mRNA is then reverse transcribed into a cDNA copy, which is inserted into a new genomic location.
(**B**) Class II TEs, DNA transposons, are largely non-replicative, but also encode for a polycistronic mRNA. TE proteins recognize TE coding sequence and target it for excision and insertion at a new genomic location.
Both Class I and Class II TEs are targeted by cellular defense pathways for silencing, including heterochromatization (1) and RNAi pathways (2) (see Figure 2).

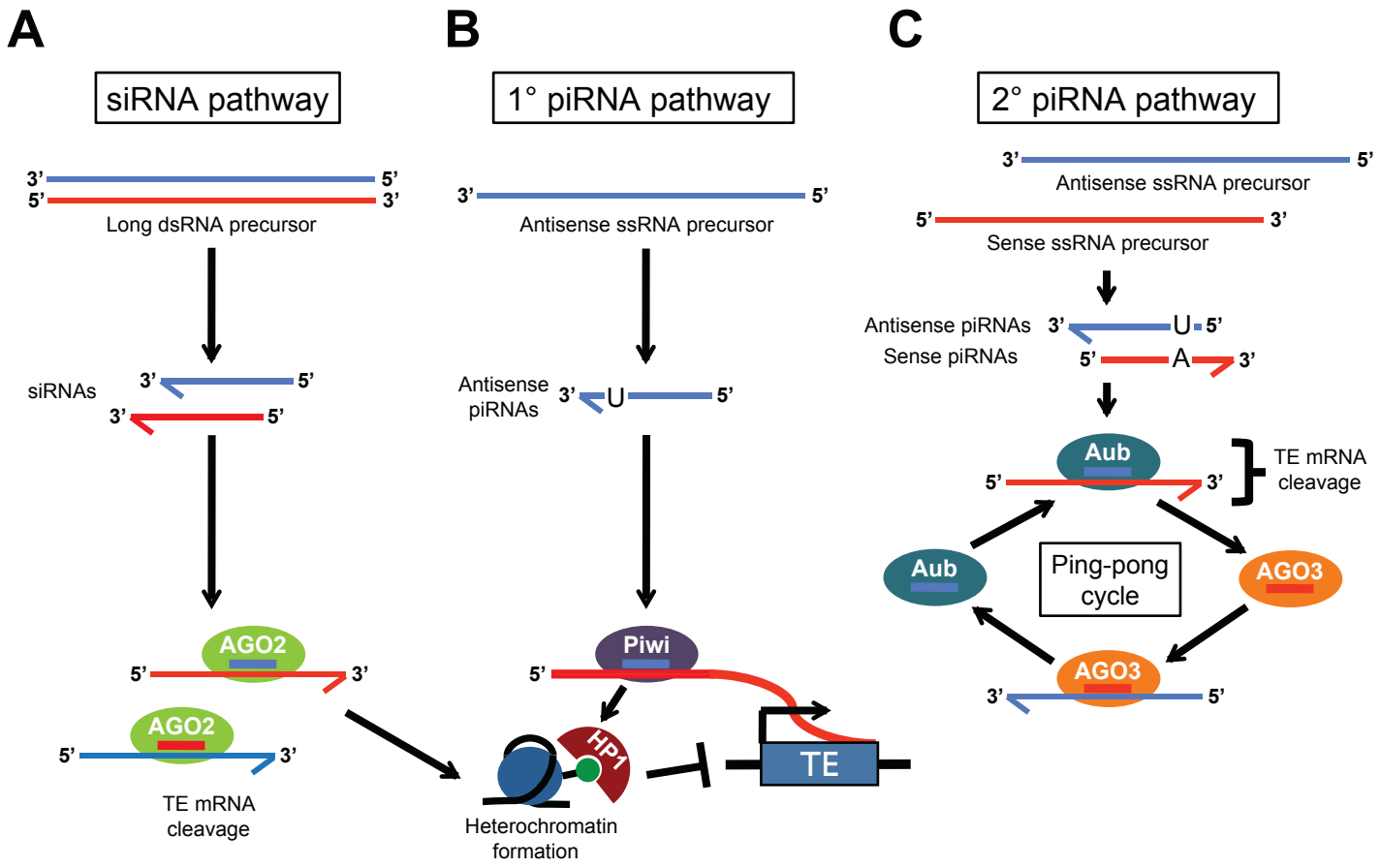


Figure 2. Transposable element suppression pathways in *Drosophila melanogaster*.
(**A**) The siRNA pathway. Long dsRNA substrates are processed into smaller 21nt siRNAs that then partner with AGO2 to post-transcriptionally cleave complementary TE mRNAs in the cytoplasm. AGO2 can also enter the nucleus and direct transcriptional silencing by recruiting Su(var)3-9 which is then able to establish H3K9me3 histone modifications (green dot). HP1 then binds to H3K9me3 marks, thereby fully establishing a heterochromatic state.
(**B**) The primary (1°) piRNA pathway. Long ssRNA precursors are transcribed from genomic regions called piRNA clusters. These precursors are then processed into largely antisense smaller 23-29nt piRNAs and exhibit a first position uracil bias. These piRNAs are then loaded onto the Piwi protein, which uses its piRNAs to direct transcriptional silencing of target TEs, likely via binding to nascent TE transcripts in the nucleus. Piwi then establishes heterochromatin through an unknown mechanism, but one similar to AGO2 in the siRNA pathway.
(**C**) The secondary (2°) piRNA pathway. Similar to the primary piRNA pathway, long dsRNA precursors are transcribed from bidirectional piRNA clusters and processed into smaller piRNAs. Antisense piRNAs, also exhibiting a uracil nucleotide bias for their first position, are loaded into Aubergine (Aub) that then cleaves complementary TE transcripts in the cytoplasm. These cleaved sense transcripts, along with sense precursor-derived piRNAs, form the substrates for piRNAs loaded into AGO3. AGO3-loaded piRNAs also exhibit a tenth position nucleotide bias for adenosine, due to their complementarity to Aub-loaded piRNAs. AGO3 then uses its sense piRNAs to cleave antisense precursor RNAs to create more antisense piRNAs for loading into Aub. This heterotypic amplification loop is commonly referred to as the “ping-pong” cylce.

Chapter 2

A somatic piRNA pathway in the *Drosophila* fat body ensures metabolic homeostasis and normal lifespan

Brian C. Jones, Jason G. Wood, Chengyi Chang, Austin D. Tam, Michael J. Franklin, Emily R. Siegel, Stephen L. Helfand

I conceived and directed the project with contributions from Jason Wood and Stephen Helfand. Chengyi Chang isolated all body segment and pure fat body samples. Jason Wood and Chengyi Chang performed all sequencing library preparation. Jason Wood performed all bioinformatic analysis with my assistance. I performed all immunoblotting. I generated and assayed the *piwi* mutant fat body *gypsy-TRAP* flies. Michael Franklin, Emily Siegel, and I performed all sectioning, immunofluorescence, lipid staining, and microscopy. I quantified the immunofluorescence and lipid droplet staining. I processed all immunoblot and microscopy images for publication. Austin Tam and I performed all metabolic assays. Michael Franklin, Emily Siegel, Austin Tam, and I performed all starvation assays. Emily Siegel and I performed all immune challenge assays. Chengyi Chang, Emily Siegel, and I performed all lifespan assays. I wrote the manuscript with assistance from Jason Wood and Stephen Helfand.

This work has been published in *Nature Communications* (2016). December 21; 7:13856. DOI: 10.1038/ncomms13856.

Abstract

In gonadal tissues, the Piwi-interacting (piRNA) pathway preserves genomic integrity by employing 23-29nt small RNAs complexed with argonaute proteins to suppress parasitic mobile sequences of DNA called transposable elements (TEs). While recent evidence suggests that the piRNA pathway may be present in select somatic cells outside the gonads, the role of a non-gonadal somatic piRNA pathway is not well characterized. Here we report a functional somatic piRNA pathway in the adult *Drosophila* fat body including the presence of the piRNA effector protein Piwi and canonical 23-29nt long TE-mapping piRNAs. *piwi* mutants exhibit depletion of fat body piRNAs, increased TE mobilization, increased levels of DNA damage, and reduced lipid stores. These mutants are starvation-sensitive, immunologically compromised, and short-lived, all phenotypes associated with compromised fat body function. These findings demonstrate the presence of a functional non-gonadal somatic piRNA pathway in the adult fat body that impacts normal metabolism and overall organismal health.

Introduction

TEs parasitize the DNA of their hosts and account for a large portion of eukaryotic genomes[1,2]. To combat the invasion and expansion of TEs, small RNA (smRNA) silencing pathways have evolved to suppress TEs across species from plants to humans[3]. The short interfering RNA (siRNA) pathway suppresses TEs in all tissues of plants and animals, while the activity of the piRNA pathway is thought to be primarily restricted to the gonads of metazoans[4,5]. Loss or decline of these pathways results in

genomic instability and cellular dysfunction caused by TE reactivation and transposition[6–9].

The piRNA pathway is best known for its role in gonadal tissues where it protects against genomic damage caused by TE reactivation[4,5]. The pathway silences TEs by employing complementary small RNAs called piRNAs, generated from large TE-rich genomic regions called piRNA clusters. In flies, these clusters transcribe long single-stranded RNA precursors that are then further processed into smaller 23-29nt piRNAs. These piRNAs partner with argonaute effector proteins (Piwi, Aubergine, or AGO3), which are then able to silence TEs via their homology to TE transcripts[4,10]. This process is accomplished by one of two silencing mechanisms. In the primary piRNA pathway, active in both the germline and ovarian somatic follicle cells, Piwi represses TE transcription by establishing heterochromatin[5,11]. In the secondary piRNA pathway, active only in the germline, Aubergine and AGO3 silence TEs post-transcriptionally in the cytoplasm via mRNA cleavage[4,5,10]. While the role of the piRNA pathway was previously thought to be restricted to the gonads, recent evidence in a diversity of organisms suggests that this pathway may also be present in somatic cells outside of the gonad[12].

Over the past decade, new evidence has begun to reveal non-gonadal examples of the piRNA pathway including a role for the piRNA pathway in stem cell function[12]. In planaria, piRNA pathway proteins are essential in maintaining stem cell pluripotency as well as the regenerative capacity of these animals[13]. piRNA pathway components are active in multiple types of cancer[12], including specific cancers in mammals and flies[12,14–18], and in flies *piwi* has been shown to contribute to malignant tumor growth[19]. Less

information is available for a role of the piRNA pathway in normal differentiated somatic tissues, although evidence for the activity of the secondary piRNA pathway in specific neurons of the adult fly brain has been reported[20]. As more non-gonadal examples of an active RNAi system are discovered, it appears that the piRNA pathway may have other important roles beyond its known functions in gonadal tissue.

Here we show the presence of a functional somatic piRNA pathway in the adult fly fat body. The piRNA pathway in the fat body exhibits all the canonical characteristics of a primary piRNA pathway and actively suppresses TE mobilization in this tissue. We observe that loss of this pathway correlates with compromised fat body function and shortened lifespan. These findings demonstrate a novel role for the piRNA pathway outside of the gonads in a fully differentiated somatic tissue.

Results

The fat body exhibits components of an intact piRNA pathway

Given the recent evidence that the piRNA pathway may also be active in select non-gonadal somatic cells[12], we examined the expression of piRNA pathway genes outside of the gonads. We found that examination of RNA-seq libraries from adult fly eviscerated abdomen, but not heads or thorax, had significant enrichment of piRNA pathway genes relative to other somatic body segments (Fig. 1a, Supplementary Fig. 1, Supplementary Fig. 2, and Supplementary Dataset 1). Immunoblotting of isolated purified fat body, thorax and head demonstrated the presence of Piwi protein in the fat body (Fig. 1b), but not AGO3 or Aubergine. Immunofluorescent microscopy also demonstrated specific localization of the Piwi protein to the nuclei of adult abdominal

and pericerebral fat body cells, but not to nuclei of cells from other tissues outside of the gonads or in fat bodies of homozygous *piwi* null mutants (Fig. 1c; data not shown). The presence of Piwi protein in the nuclei of the fat body suggests the possibility of an intact piRNA pathway in this tissue.

Activated piRNAs in functional piRNA silencing complexes are 2'-*O*-methylated at their 3' termini and can be selectively enriched and detected by using periodate oxidation[4,21]. We found a broad piRNA-like peak of smRNAs ranging from 23 to 29nt in oxidized small RNA-seq (smRNA-seq) libraries from pure adult abdominal fat body (Fig. 1d), suggesting the presence of piRNAs with 2'-*O*-methylation at their 3' termini as is typical of gonadal piRNAs actively loaded into a piRNA-argonaute complex[22]. 49% of oxidized fat body piRNAs (23-29nt smRNAs) mapped to TEs (Fig. 1e) and these reads exhibited a strong antisense bias (Supplementary Fig. 3a) and a canonical first position nucleotide bias for uracil[10] (Fig. 1f) indicating that they likely have the capacity to target TE transcripts for silencing. Oxidized smRNA libraries from purified abdominal fat bodies of two distantly related drosophilids, *D. simulans* and *D. yakuba* (Supplementary Fig. 3b), also showed 23-29nt piRNAs mapping to TEs (Supplementary Fig. 3c-f) with antisense TE-mapping piRNAs exhibiting a strong first position uracil bias (Supplementary Fig. 3g, h) demonstrating that adult fat body piRNAs are conserved across diverse drosophilid species. Together, these data suggest the presence of fat body piRNAs that exhibit canonical piRNA characteristics, are associated with an active piRNA-argonaute silencing complex, and are evolutionarily conserved.

The fat body piRNA pathway suppresses transposable elements

We next examined whether the fat body piRNA pathway exhibits canonical hallmarks of an active piRNA pathway. Loss of Piwi in gonadal tissues results in a dramatic reduction of piRNAs and a de-repression of their corresponding TEs[11,23]. In the fat bodies of *piwi* null mutants, we observed a significant increase in the transcript levels of multiple TEs and a corresponding decrease of their associated piRNAs (Fig. 2a and Supplementary Fig. 4). Total piRNAs also decreased (28.1% decrease) as well as TE-mapping piRNAs (70.5% decrease) (Fig. 2b, side panel). We used a reporter of transposition for the endogenous *gypsy* retrotransposon (*gypsy-TRAP*)[9,24], a known target of Piwi in the ovary[25], to detect transposition in fat body cells and found that *piwi* mutant fat bodies displayed significantly higher levels of *gypsy* transposition relative to controls (Fig. 2c). Together, these data show that a somatic piRNA pathway actively suppresses the expression and mobilization of TEs in adult fat body cells.

piRNAs that target and suppress TEs in ovarian tissues are known to originate from genomic regions called piRNA clusters[10]. Mapping fat body piRNAs to previously annotated fly piRNA clusters showed many map to the *flamenco* cluster, which is known to be specific to somatic follicle cells (Fig. 2d, e)[10,11]. In agreement with studies of ovarian follicle cells[10,11], piRNAs mapping to *flamenco* were depleted in the fat bodies of *piwi* mutants (Fig. 2d, e). These data further support the presence of a functional fat body piRNA pathway where piRNAs produced from somatic piRNA clusters pair with Piwi to suppress TEs as the primary piRNA pathway does in ovarian follicle cells[10,11].

Previous studies in flies have shown that primary piRNAs are also derived from the 3'UTRs of coding genes[26,27]. We observed that 17% of fat body piRNAs from the oxidized fat body library mapped to coding genes (Fig. 1e) and that 3'UTR sense-mapping piRNAs were depleted in *piwi* mutants (17.5% decrease; Fig. 2b, side panel). These 3'UTR-derived piRNAs are between 78% (*piwi* $^{-/+}$) and 87% (wt) sense-mapping, as is expected of genic piRNAs (Supplementary Fig. 5a). Fat body piRNAs mapping to the 3'UTR of *traffic jam* (*tj*), a known source of genic piRNAs in ovarian follicle cells[26,27], were depleted in *piwi* mutants (Fig. 2f, g and Supplementary Dataset 2) along with other 3'UTR-derived piRNAs (Supplementary Fig. 5b-e and Supplementary Dataset 2). Together, these data provide evidence of 3'UTR-derived fat body genic piRNAs, another hallmark of the primary piRNA pathway.

In order to rule out the possibility that fat body piRNAs result from contamination by ovarian tissues during dissection and library preparation, we performed smRNA-seq on oxidized fat body libraries isolated from *ovo*^{D1} flies. Due to a dominant female-sterile mutation in the *ovo* gene, these flies exhibit severely degenerated ovaries thus significantly decreasing the likelihood that any piRNAs isolated from the fat body of these animals would be due to contamination from ovarian tissues. We observed in both wild type and *ovo*^{D1} fat bodies 23-29nt smRNAs that mapped uniquely to the *flamenco* locus (Supplementary Fig. 6a, b), suggesting that piRNAs observed in oxidized fat body libraries are in fact originating from this tissue and not a result of ovarian contamination. Additionally, immunoblotting demonstrates the presence of Piwi protein in the eviscerated abdomen and isolated fat body of *ovo*^{D1} flies (Supplementary Fig. 6c). These data, combined with our observation of the Piwi protein in the nuclei of fat body

cells (Fig. 1b and Supplementary Fig. 2), strongly support the presence of a functional canonical piRNA pathway in the fat body.

piRNA mutants display DNA damage and metabolic dysregulation

The replicative mobility of TEs can contribute to mutagenesis via their ability to insert into new genomic loci[6]. TE reactivation and transposition has been shown to correlate with chromosomal rearrangements, double-strand DNA breaks (DSBs), and apoptosis[7,8]. Phosphorylation of the histone variant H2A.v (γ-H2A.v) during DNA repair serves as a reliable marker of DSBs and has been shown to correlate with increased TE activity in the fat body[9,28]. Using immunofluorescent microscopy, we observed an increase in the intensity of γ-H2A.v staining in *piwi* mutants relative to controls (Fig 3a, b). These data suggest that the fat body piRNA pathway normally protects fat body cells from the accumulation of DNA damage that may be caused by TE reactivation.

The fly fat body is a functional analogue of the mammalian liver and adipose tissue, with one of its primary roles being storage of lipids and glycogen[29]. Our observation that *piwi* mutant fat body cells showed increased TE mobilization and elevated levels of DNA damage (Fig 2c and Fig. 3a, b) led us to hypothesize that this could result in disrupted fat body function. Nile Red staining of fat body lipid droplets revealed a reduction of lipid droplet size in *piwi* mutants compared to controls, with larger lipid droplets (>200 μm^2) greatly reduced in their abundance (Fig. 3c-e and Supplementary Dataset 3). These data correlate with a significant reduction of two of the major storage metabolites in fat body, triacylglycerides (TAGs) and glycogen, in both *piwi* and *flamenco* mutants (Fig. 3f-i and Supplementary Fig. 7a, b). We next asked

whether these phenotypes correlated with an altered fat body transcriptome as transcription of TEs alone can cause RNA toxicity and contribute to cellular dysfunction[8]. We generated RNA-seq libraries from *piwi* mutant and control fat bodies and found that many differentially expressed genes in metabolism-associated pathways were significantly changed (Supplementary Fig. 7c). These data support a model in which a loss piRNA pathway function results in a decrease in lipids and stored metabolites, disruption of metabolic homeostasis, and a decline in cellular function, possibly due to reactivation of TEs.

piRNA mutants show altered fat body function and lifespan

The fly fat body plays a major role in resisting stressors such as fasting and, through its role in innate immunity, in resisting pathogenic infections[29]. We observed that *piwi* mutants were highly sensitive to starvation conditions as well as infection by a pathogenic insect bacterium (Fig 4a, b and Supplementary Table 1). Due to the central role the fat body plays in regulating longevity[30,31], we next examined the effect of disrupting the piRNA pathway on lifespan. We found that mutations in either *piwi* or the *flamenco* locus dramatically shorten lifespan (Fig. 4c, d and Supplementary Table 1). Finally, we tested whether the shortened lifespan of piRNA pathway mutants was dependent upon TE activity. Many TEs in Drosophila are retrotransposons, including *gypsy*, and depend upon reverse transcriptase for their replicative mobility[32]. Administration of a known reverse transcriptase inhibitor, 3TC, inhibits the normal age-related increase in *gypsy* mobilization and extends the shortened lifespan of *Dcr-2* mutants, another condition in which de-repression of TEs occurs[9]. We administered

3TC to *flamenco* mutants and observed a significant lifespan extension (Fig. 4e and Supplementary Table 1) suggesting that a shortened lifespan phenotype is at least partially dependent upon TE mobilization. These results suggest that loss of the fat body piRNA pathway and an increase in TE activity and mobilization correlates with compromised fat body function including its ability to otherwise mitigate the detrimental effects of environmental stressors.

Discussion

Here we have shown evidence for a fully functional piRNA pathway in a non-gonadal somatic tissue, the adult fly fat body, which is likely to be necessary for proper tissue function and overall organismal health. These results demonstrate that the adult fat body piRNA pathway exhibits canonical characteristics found in gonadal somatic cells, and its activity likely positively impacts the function of a tissue important to metabolic homeostasis and physiological health. While we are not able to entirely rule out a contribution of the gonadal piRNA pathway to fat body function, many of the phenotypes we observe are opposite to those typically seen in animals with compromised gonadal tissue function and therefore likely represent the effect of a loss of the fat body piRNA pathway. For example, the shortened lifespan and reduced lipid stores in piRNA pathway mutants demonstrates that the piRNA pathway is essential in the health and functioning of non-gonadal somatic tissues, since reduction or ablation of gonadal function in flies often extends lifespan and increases lipid stores rather than decreasing lifespan and fat storage[31]. Recent studies in wild type flies have also demonstrated an important link between TE activity and longevity[9] and our studies

demonstrating partial rescue of the shortened lifespan in *flamenco* mutants upon administration of a reverse transcription inhibitor further support this association.

Interest in a function for the piRNA pathway in the soma has increased recently as new roles for this pathway are being illuminated. The piRNA pathway's association with tissues that maintain a degree of immortalization similar to that exhibited in the germline is of particular interest[12]. For example, the somatic stem cell niches of Hydra maintain an active piRNA pathway that represses TEs, possibly contributing to this organism's remarkably long lifespan[33]. These studies, together with our findings, suggest that the presence of a piRNA pathway in normal somatic tissues may offer an additional cellular defense against TE reactivation and possible somatic genomic damage. Our finding of a role for the piRNA pathway in preserving metabolic homeostasis and the overall health of the fly suggests the potential importance of the piRNA pathway in other somatic tissues. Finally, interventions specifically augmenting the piRNA pathway may provide significant benefits to maintaining genomic integrity, tissue function, and healthy lifespan.

Materials and Methods

Fly Stocks & Husbandry

Drosophila stocks were all maintained on standard media (30.5 g L^{-1} autolysed yeast, 121.8 g L^{-1} sucrose, 52.3 g L^{-1} cornmeal, 8.75 g L^{-1} agar, 2.62 g L^{-1} tegosept, all weight by volume) at 25°C with a 12-hour light/dark cycle at 60% relative humidity. Unless otherwise noted, flies used in all experiments were collected over a 48-hour

period, placed in density-controlled, mixed-sex vials, and aged for 10 days on food containing 150 g L^{-1} autolysed yeast, 150 g L^{-1} sucrose, and 20 g L^{-1} agar, all weight by volume. Unless otherwise noted, all experiments were performed using mated female flies grown under these conditions.

Lab stocks of Canton S or w^{1118} were used for wild type experiments. $flam^{KG}/FM4$ (Bloomington #16453) was obtained from the Bloomington *Drosophila* Stock Center at Indiana University. Stocks of *D. yakuba* and *D. simulans* were provided by the *Drosophila* Stock Center at the University of California, San Diego. The *piwi* mutant fat body *gypsy-TRAP* reporter line was generated from the *gypsy-TRAP* line provided to us by Josh Dubnau[24], *r4-GAL4* (Bloomington #33832), *UAS-GFP* (Bloomington #1522), and the $piwi^2/CyO$ line from Haifan Lin. ovo^{D1} mutants lacking ovaries were generated by crossing males of ovo^{D1} ($ovo^{D1}v^{24}$/C(1)DX, $y^1w^1f^1$) (Bloomington #1309) to virgin Canton S females.

RNA/smRNA-seq Library Preparations

Flies were flash frozen on dry ice and head and thorax body segments were dissected on a -20°C cold block and stored at -80°C. Eviscerated abdomens and pure adult fat bodies were dissected from the abdominal wall in cold PBS and also stored at -80°C.

Total RNA was extracted from relevant fly tissues using the mirVana miRNA Isolation Kit (ThermoFisher Scientific AM1560). For RNA-seq library prep, 100 ng of total RNA was used as input for the Ovation Universal RNA-seq System kit (Nugen #0343), with *Drosophila* rRNA depletion module, according to the manufacturer's

instructions. Three biologically independent libraries were made for each tissue. For smRNA-seq libraries, 2 µg of total RNA was oxidized in 25 mM sodium periodate with 30 mM borax and 30 mM boric acid (pH 8.6) for 30 minutes at room temperature (Phillip Zamore Lab Illumina TruSeq Small RNA Cloning Protocol April, 2014; http://www.umassmed.edu/zamore/resources/protocols/). RNA was then recovered with the RNA Clean & Concentrator-5 kit (Zymo Research R1015) and used as input to the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs E7300), with modifications adapted from[34] (2S rRNA block oligo added before 5' ligation step to decrease rRNA reads in final library).

Bioinformatics

Head, thorax, eviscerated abdomen, and ovary RNA-seq

Reads were mapped to the dm3 genome using Tophat. The Bioconductor R package easyRNASeq was used to count reads (using "geneModels" summarization parameter) and calculate normalized RPKM values. Error bars presented represent standard error of the mean with three independent biological replicates for each condition. FDR values in Supplementary Dataset 1 were calculated using the Bioconductor edgeR package, with a GLM correcting for batch effects as described in the vignette.

smRNA size profiles, pie charts, and nucleotide bias

Oxidized small RNA libraries were first trimmed of adapter sequences using cutadapt, then aligned against *Drosophila* rRNA sequences using bowtie to remove

rRNA aligning reads. For pie charts, reads were first aligned to the dm3 genome, keeping only reads that aligned. Reads were next consecutively mapped to the following genomic compartments using bowtie (-v 1): sno + tRNA, miRNA, transposable elements, exons, and intergenic regions (all sequences downloaded from FlyBase, except transposable elements which came from Repbase). Reads aligning to each compartment were counted to create chart.

For total small RNA size profiles (e.g. Fig. 1d), miRNAs were first removed using bowtie (-v 1), and remaining reads were subsequently fed into a Perl script which counted the number of reads for each different length between 18 and 50 nucleotides. For size profiles of TE-mapping reads (Fig. S2), reads were aligned to the set of Repbase *Drosophila* consensus sequences using bowtie (-v 1 -k 1--best), and sorted into sense/antisense (with respect to TE) using samtools. Alignments were then converted back into FASTQ files using bedtools, and size profiles calculated as described above. For *D. simulans* and *D. yakuba* pie charts and size profiles, analysis was performed as described, using relevant genomic compartments for each species downloaded from Flybase. For transposable elements, the set of all *Drosophila* species transposons from Repbase was used.

For nucleotide bias calculations, 23-29 bp TE-mapping antisense alignments were converted into FASTA format using bedtools and EMBOSS, trimmed to uniform length of 23 bp using FASTX-toolkit, and used as input for the WebLogo 3 program[35].

TE and coding gene analysis

In order to properly account for multi-mapping RNA-seq reads when analyzing TEs, we used the RepEnrich approach[36] to quantify read counts for each TE. This approach combines all instances of each annotated TE in the genome together and counts a read if it aligns to any of them at least once, allowing proper quantitation of reads that would otherwise be discarded due to mapping to multiple locations. Read count tables from RepEnrich were processed with the edgeR package to perform normalization and calculate $\log_2$ fold change. For small RNA-seq libraries, reads were aligned using RepEnrich and counts were normalized using unique alignments to cisNATs and structured loci, as described in [11]. Log_2 fold changes (*piwi* $^{-/-}$ / *piwi* $^{+/-}$) were calculated using normalized read counts. TEs with RNA-seq $\log_2$ fold changes >0.263 (1.2x fold increase) in *piwi* mutants compared to heterozygous controls were plotted on the heat map together with the corresponding piRNA-seq change for each element. Heat maps were generated with the gplots2 package in R.

To create TE alignment profiles (Supplementary Fig. 4), reads were mapped uniquely to the relevant TE Repbase consensus sequence using bowtie (-v 1 -m 1). Read depth across TE consensus sequence was determined using bedtools genomecov command, using total uniquely mapping reads in the library to normalize for library size differences. Plots were created in R.

For normal coding gene analysis, edgeR was used to determine differentially expressed genes whose expression was significantly changed in total RNA-seq libraries of *piwi* mutant fat body relative to heterozygous controls. We then performed a KEGG pathway analysis on these genes using Flymine.org[37].

piRNA cluster analysis

To determine abundance of small RNA reads mapping to annotated piRNA clusters, 23-29 bp size selected reads were aligned uniquely to the genome using bowtie (-v 1 -m 1). piRNA cluster specific reads were extracted using bedtools using the 15 most highly expressing piRNA clusters, as defined in [10,11]. Specific cluster coverage plots were calculated using bedtools genomecov, normalizing to total uniquely aligning reads (excluding snoRNA, tRNA, and miRNA), and plotted with the Sushi R package from Bioconductor[38].

To assay genic piRNA reads, 23-29 bp aligned reads were separated into 5'UTR, CDS, and 3'UTR regions for each gene (Flybase annotations) and counted and sorted sense/antisense using bedtools. For selected genes with high numbers of antisense piRNAs in the 3'UTR (e.g. Fig 2G), coverage was calculated as described above. Coverage and gene models were plotted using Sushi[38].

Immunoblotting

Fly body segments and tissues were dissected as in "RNA/smRNA-seq Library Preparations." Whole cell lysate samples were prepared in RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 1% Triton-X-100, 1% Na-deoxycholate, 0.1% SDS, and 2.5 μg mL^{-1} of Pepstatin A, Leupeptin, Antipain, Aproptinin, and Chymostatin). 15 μg of total protein was loaded and run on a 12% polyacrylamide/SDS gel. Proteins were then transferred to PVDF membrane and the membrane was blocked overnight in 5% TBST-milk at 4°C. Membranes were first cut between a 75 kDa and 50 kDa marker and the lower molecular weight half of the membrane incubated with anti-

Actin (mouse 1:2,000; EMD Millipore MAB1501). The upper molecular weight half of the membrane was incubated with anti-Piwi (mouse 1:200; Santa Cruz sc-390946), anti-Aubergine (rabbit 1:2,000), or anti-AGO3 (rabbit 1:3,000). Before blotting with either anti-Aubergine or anti-AGO3 and between each re-probing, the blot was stripped with stripping buffer (50 mM Tris-HCl pH 6.8, 20 g L^{-1} SDS, and 0.8% β-mercaptoethanol) at 60°C. After treating with each primary, membranes were incubated with either HRP-conjugated anti-mouse (goat 1:5,000; ThermoFisher 31430) or anti-rabbit (goat 1:5,000; ThermoFisher #31460). The anti-AGO3 and anti-Aubergine antibodies were gifts from Phillip Zamore. Uncropped images of the original immunoblots can be found in Supplementary Figure 2.

Immunofluorescence Assays and Lipid Staining

Flies were briefly dipped in EtOH and blotted, fixed in 4% paraformaldehyde in PBS for 20 minutes, washed with PBS on ice, and embedded in TBS Tissue Freezing Medium (Fisher Scientific #15-183-13). Molds were frozen on dry ice and stored at -80°C. Molds were cryosectioned into 10μm sections and gently washed with PBS. For anti-Piwi immunofluorescence, slides were incubated with anti-Piwi antibody (mouse 1:50; Santa Cruz sc-390946) for one hour at room temperature, washed with PBS, incubated with anti-mouse Alexa Fluor-568 conjugated antibody (goat 1:2,000; ThermoFisher A-11031) for one hour at room temperature, and then washed again with PBS. Anti-Piwi primary and secondary antibodies were both diluted in 5% normal goat serum and 0.1% Triton X-100. For lipid staining, sections were incubated with Nile Red (0.5 μg mL^{-1}; ThermoFisher N-1142) for one hour at room temperature and washed with

PBS. For anti-γ-H2A.v immunofluorescence, slides were washed in PBS and blocked for 30 minutes at room temperature in 5% normal goat serum and 0.5% Triton X-100. Slides were then briefly rinsed in PBS, incubated with anti-γ-H2A.v antibody (mouse 1:20; DSHB unc93-5.2.1) for two hours at room temperature, washed with PBS+0.1% Tween-20 (PBST), incubated with anti-mouse Alexa Fluor-568 conjugated antibody (goat 1:1,000; ThermoFisher A-11031) for one hour at room temperature, and washed twice with PBST followed by one final wash with PBS. Anti-γ-H2A.v primary and secondary antibodies were both diluted in 5% normal goat serum and 0.1% Tween-20. All slides were mounted and stained with DAPI (ThermoFisher P36935).

Images of Piwi were acquired with a Zeiss LSM 510 meta-confocal laser-scanning microscope and γ-H2A.v and lipid staining with a Zeiss AxioImager.Z1 ApoTome microscope. Images of representative nuclei, proteins, and lipids were selected and processed using ImageJ and Adobe Photoshop. γ-H2A.v intensity was measured and quantified using ImageJ according to the method published in [28]. Lipid droplet size was measured and quantified using ImageJ from sections of individual flies. For γ-H2A.v fluorescence intensity and lipid droplet size quantification, a Student's two-tailed *t*-test was performed to determine significance.

***gypsy-TRAP* Transposition Reporter**

The *piwi* mutant fat body *gypsy-TRAP* reporter line was generated using the lines mentioned in "Fly Stocks and Husbandry." Briefly, the fat body specific *r4-GAL4* driver line was first recombined with a *UAS-GFP* line generating a new stable line that expresses GFP exclusively in the fat body. This line was then crossed with the *gypsy-*

TRAP reporter line[24], a line containing a *GAL80* driven by a tubulin promoter separated by an *ovo* binding site that attracts the *gypsy* TE, which suppresses GFP expression until TE transposition activates reporter function. Once stable, this new line was then further crossed into a *piwi*^{2}/*CyO* mutant background thereby generating the final *piwi* mutant fat body *gypsy-TRAP* reporter line. Mixed sex heterozygous and homozygous flies from this line were then aged together for 10 days. Flies were then separated by *piwi* genotype and female flies of each *piwi* genotype further separated under a fluorescent dissecting scope into three groups according to the approximate number of GFP cells visibly fluorescing beneath the ventral abdominal wall. Groups were as follows: low (no visible GFP-positive cells), medium (<50% abdominal area showing GFP-positive cells), and high (>50% abdominal area showing GFP-positive cells). A Fisher's exact test was performed to determine significance between the combined medium+high groups and the low group.

As was demonstrated for adult mushroom body neurons[24], we found that the GFP-positive cells in the fat body were dependent upon endogenous gypsy insertion as mutant *ovo* binding sites in the wild type *gypsy-TRAP*/*r4-GAL4::UAS-GFP* reporter line resulted in no GFP-positive cells as expected[9]. All representative images were taken from heterozygous *piwi* mutant fat body *gypsy-TRAP* flies.

Metabolic Assays

Assays were performed as in [39] with the following modifications. Biological replicates for each genotype assayed were generated using 5 whole adult flies. All data were normalized to total protein concentration and calculated as a percent relative to

the control genotype. For each assay, a Student's two-tailed *t*-test was performed to determine significance.

Triacylglycerides (TAG)

Flies were homogenized in 200 µL of cold PBST buffer. The homogenate was heat-treated at 70°C for 5 minutes to inactivate endogenous enzymes. Protein concentration was measured using a Bradford assay (Bio-Rad #5000006). Samples were diluted with PBST and 15 µL of heat-treated homogenate or glycogen standard (Standbio #2103-030) were incubated with 200 µL of Trigylceride Reagent (Thermo Scientific TR22421) for 5 minutes at 37°C. Absorbance was measured at 540nm and TAG content of samples was calculated based on a standard curve of TAG that was run in parallel with experimental samples.

Glycogen

Flies were homogenized in 200 µL of cold PBS. The homogenate was heat-treated at 70°C for 5 minutes to inactivate endogenous enzymes. Samples were centrifuged for 3 minutes at 16,100 xg and the supernatant collected. Protein concentration was measured using a Bradford assay (Bio-Rad #5000006). Samples were diluted with PBS and 90 µL of heat-treated homogenate were incubated with either 20 µL of amyloglucosidase (Sigma-Aldrich A7420) or 20 µL of PBS. To create a glycogen standard curve, 50 µL of glycogen standards (Ambion AM9510) were treated with either 50 µL of amyloglucosidase or 50 µL of PBS. All samples were incubated at 37°C for one hour. 30 µL of each sample was then added to a 96-well plate. 100 µL of Infinity

Glucose Hexokinase reagent (Thermo TR15421) was added to all samples and incubated at room temperature for 15 minutes. The absorbance of samples was then measured at 340 nm and normalized by subtracting the absorbance of the free glucose of untreated samples from the absorbance of the total amount of glucose present in samples treated with amyloglucosidase. Glycogen content was then calculated based on the normalized glycogen standard curve.

Trehalose

Flies were homogenized in 200 µL of cold trehalose buffer (5 mM Tris-HCl pH 6.6, 137 mM NaCl, 2.7 mM KCl). The homogenate was heat-treated at 70°C for 5 minutes to inactivate endogenous enzymes. Samples were centrifuged for 3 minutes at 16,100 xg and the supernatant collected. Protein concentration was measured using a Bradford assay (Bio-Rad #5000006). Samples were diluted with trehalose buffer and 90 µL of heat-treated homogenate were incubated with either 20 µL of trehalase (Sigma-Aldrich T8778) or 20 µL of trehalose buffer. To create a trehalose (Sigma-Aldrich T9531) standard curve, 50 µL of standard were incubated with either 30 µL of trehalase or 30 µL of trehalose buffer. A free glucose (Fisher Scientific 50-99-7) standard curve was also generated. All samples were incubated at 37°C overnight. 30 µL of each sample was then added to a 96-well plate. 100 µL of Infinity Glucose Hexokinase reagent (Thermo TR15421) was added to all samples and incubated at room temperature for 15 minutes. The absorbance of samples was then measured at 340 nm and normalized by subtracting the absorbance of free glucose present in untreated samples from the total amount of glucose present in samples treated with trehalase. Trehalose and free

glucose content were calculated based on standard curves of trehalose and glucose respectively.

Glucose

Flies were homogenized in 200 µL of cold PBS. The homogenate was heat-treated at 70°C for 5 minutes to inactivate endogenous enzymes. Samples were centrifuged for 3 minutes at 16,100 xg and the supernatant collected. Protein concentration was measured using a Bradford assay (Bio-Rad #5000006). Samples were diluted with PBS and 30 µL of supernatant was then added to a 96-well plate. 100 µL of Infinity Glucose Hexokinase reagent (Fisher Scientific TR15421) was added to all samples and incubated at room temperature for 15 minutes. The absorbance of samples was then measured at 340 nm and glucose content was calculated based on a standard curve of glucose (Fisher Scientific 50-99-7).

Starvation Assay

Prior to permanent starvation, flies were synchronized in their feeding by fasting on 2% agar for 4 hours followed by a final period of feeding for 2 hours. Flies were then sorted under CO_2 anesthesia into separate sex vials containing 2% agar at a density of 10 males or 10 females per vial, with a total of 5 vials (n≈50) for each genotype. Dead flies were scored and counted every 6 hours. Starvation analyses were performed and log rank statistics calculated using the online application OASIS[40]. Starvation assays were repeated at least twice.

Immune Challenge Assay

A bacterial culture of *Erwinia carotovora* (gram-negative insect/plant pathogen), a gift from Neal Silverman, was grown overnight in Luria broth, shaking at 220 rpm on a platform shaker at 37°C. The culture was allowed to reach OD_{600nm}≈2.0 before removing 1 mL for centrifugation at 2,000 xg for 2 minutes. The supernatant was removed and the bacterial pellet gently washed with 1 mL of 10 mM $MgSO_4$ to remove traces of culture media. The wash was removed and the pellet resuspended in 10 µL of fresh 10 mM $MgSO_4$. A 0.15 mm needle was then dipped into 80% EtOH and flame sterilized. Flies were sorted under CO_2 anesthesia and inoculated by dipping the needle into either the bacterial suspension, or EtOH for mock infection controls, and inserting the needle midway into one side of the thorax. Flies were then sorted into separate sex vials at a density of 10 males or 10 females per vial, with a total of 5 vials (n≈50) for each condition/genotype, and passed to fresh food at least every other day. Dead flies were scored and counted every 24 hours. Immune challenge analyses were performed and log rank statistics calculated using the online application OASIS[40]. Immune challenge assays were repeated at least twice.

Lifespan Assay

Flies used in lifespan experiments were collected upon eclosion over a 48-hour period and were sorted under CO_2 anesthesia and placed in food vials at a density of 25 males and 25 females per vial, with a total of 10 vials (n≈250) for each genotype. Lifespan food consists of 50 g L^{-1} autolysed yeast, 50 g L^{-1} sucrose, and 20 g L^{-1} agar (all w/v). For the *flamenco* 3TC lifespans, *flamenco* homozygous mutant flies were

collected and aged on either food containing 150 g L^{-1} autolysed yeast, 150 g L^{-1} sucrose, and 20 g L^{-1} agar, (all w/v) or identical food with 10 mM lamivudine (3TC). Flies were passed to fresh food every other day, and dead flies scored and counted. Lifespan analyses were performed and Wilcoxon rank sum test statistics calculated using the online application OASIS[40]. Lifespan assays were repeated twice.

Acknowledgements

We thank Gillian Horwitz, Mark Henriksen, Davis Hartnett, Jackson Taylor and Suzanne Hosier for technical assistance and Will Lightfoot for fly food preparation. We also thank Haifan Lin, Joshua Dubnau, Robert Reenan, Marc Tatar and the Bloomington Drosophila Stock Center for fly stocks; John Sedivy and Neal Silverman for reagents; Chengjian Li, Bo Han and Phillip Zamore for reagents and advice on bioinformatic analysis; Marissa Holmbeck, Leila Rieder, Yiannis Savva, Erica Larschan, Nicola Neretti and Robert Reenan for discussions and comments on the manuscript. This work was supported by a NIA T32 Training Grant AG041688 and a F31 Predoctoral Research Fellowship Award AG047736 to B.C.J. and NIA Grants AG16667 and AG24353, a Glenn/American Federation for Aging Research (AFAR) Breakthroughs in Gerontology award, and Grant P30AI042853 from the Providence/Boston Center for AIDS Research to S.L.H.

References

1. Cordaux, R. & Batzer, M. a. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **10,** 691–703 (2009).

2. Levin, H. L. & Moran, J. V. Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* **12,** 615–27 (2011).

3. Shabalina, S. & Koonin, E. Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* **23,** 578–587 (2008).

4. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10,** 94–108 (2009).

5. Siomi, M. C., Sato, K., Pezic, D. & Aravin, A. a. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **12,** 246–58 (2011).

6. Kaminker, J. S. *et al.* The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biol.* **3,** (2002).

7. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8,** 272–85 (2007).

8. Wallace, N. a, Belancio, V. P. & Deininger, P. L. L1 mobile element expression causes multiple types of toxicity. *Gene* **419,** 75–81 (2008).

9. Wood, J. G. *et al.* Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. *PNAS* **113,** 11277–1282 (2016).

10. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128,** 1089–103 (2007).

11. Malone, C. D. *et al.* Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137,** 522–35 (2009).

12. Ross, R. J., Weiner, M. M. & Lin, H. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* **505,** 353–9 (2014).

13. Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sanchez Alvarado, A. SMEDWI-2 Is a PIWI-Like Protein That Regulates Planarian Stem Cells. *Science (80-.).* **310,** 1327–1330 (2005).

14. Liu, X. *et al.* Expression of hiwi gene in human gastric cancer was associated with proliferation of cancer cells. *Int. J. Cancer* **118,** 1922–1929 (2006).

15. Lee, J. H. *et al.* Stem-cell protein Piwil2 is widely expressed in tumors and inhibits apoptosis through activation of Stat3/Bcl-XL pathway. *Hum. Mol. Genet.* **15,** 201–211 (2006).

16. Grochola, L. F. *et al.* The stem cell-associated Hiwi gene in human adenocarcinoma of the pancreas: expression and risk of tumour-related death. *Br. J. Cancer* **99,** 1083–8 (2008).

17. Sumiyoshi, T. *et al.* Loss of l(3)mbt leads to acquisition of the ping-pong cycle in Drosophila ovarian somatic cells. *Genes Dev.* **30,** 1617–22 (2016).

18. Fagegaltier, D. *et al.* Oncogenic transformation of Drosophila somatic cells induces a functional piRNA pathway. *Genes Dev.* **30,** 1623–35 (2016).

19. Janic, A., Mendizabal, L., Llamazares, S., Rossell, D. & Gonzalez, C. Ectopic expression of germline genes drives malignant brain tumor growth in Drosophila. *Science* **330,** 1824–7 (2010).

20. Perrat, P. N. *et al.* Transposition-Driven Genomic Heterogeneity in the Drosophila Brain. *Science (80-.).* **340,** 91–95 (2013).

21. Saito, K. *et al.* Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev.* **21,** 1603–8 (2007).

22. Horwich, M. D. *et al.* The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* **17,** 1265–72 (2007).

23. Li, C. *et al.* Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* **137,** 509–21 (2009).

24. Li, W. *et al.* Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat. Neurosci.* **16,** 529–31 (2013).

25. Sarot, E., Bucheton, A. & Pe, A. Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. *Genetics* **1321,** 1313–1321 (2004).

26. Robine, N. *et al.* A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19,** 2066–76 (2009).

27. Saito, K. *et al.* A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature* **461,** 1296–9 (2009).

28. Chen, H., Zheng, X., Xiao, D. & Zheng, Y. Age-associated de-repression of

retrotransposons in the Drosophila fat body, its potential cause and consequence. *Aging Cell* **15,** 542–552 (2016).

29. Arrese, E. L. & Soulages, J. L. Insect fat body: energy, metabolism, and regulation. *Annu. Rev. Entomol.* **55,** 207–225 (2010).

30. Partridge, L., Alic, N., Bjedov, I. & Piper, M. D. W. Ageing in Drosophila : The role of the insulin / Igf and TOR signalling network. *EXG* **46,** 376–381 (2011).

31. Hansen, M., Flatt, T. & Aguilaniu, H. Perspective Reproduction, Fat Metabolism, and Life Span: What Is the Connection? *CMET* **17,** 10–19 (2012).

32. Touret, F., Guiguen, F., Greenland, T. & Terzian, C. In between: Gypsy in Drosophila melanogaster reveals new insights into endogenous retrovirus evolution. *Viruses* **6,** 4914–4925 (2014).

33. Juliano, C. E. *et al.* PIWI proteins and PIWI-interacting RNAs function in Hydra somatic stem cells. *Proc. Natl. Acad. Sci.* **111,** 337–342 (2014).

34. Wickersheim, M. L. & Blumenstiel, J. P. Terminator oligo blocking efficiently eliminates rRNA from Drosophila small RNA sequencing libraries. *Biotechniques* **55,** 269–272 (2013).

35. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: A Sequence Logo Generator. *Genome Res.* **14,** 1188–90 (2004).

36. Criscione, S. W., Zhang, Y., Thompson, W., Sedivy, J. M. & Neretti, N. Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics* **15,** 583 (2014).

37. Lyne, R. *et al.* FlyMine : an integrated database for Drosophila and Anopheles genomics. **8,** (2007).

38. Phanstiel, D. H., Boyle, A. P., Araya, C. L. & Snyder, M. P. Sushi.R: Flexible, quantitative, and integrative genomic visualizations for publication-quality multi-panel figures. *Bioinformatics* **19,** 2808–10 (2014).

39. Tennessen, J. M., Barry, W. E., Cox, J. & Thummel, C. S. Methods for studying metabolism in Drosophila. *Methods* **68,** 105–115 (2014).

40. Yang, J. *et al.* OASIS: Online Application for the Survival Analysis of Lifespan Assays Performed in Aging Research. *PLoS One* **6,** e23525 (2011).

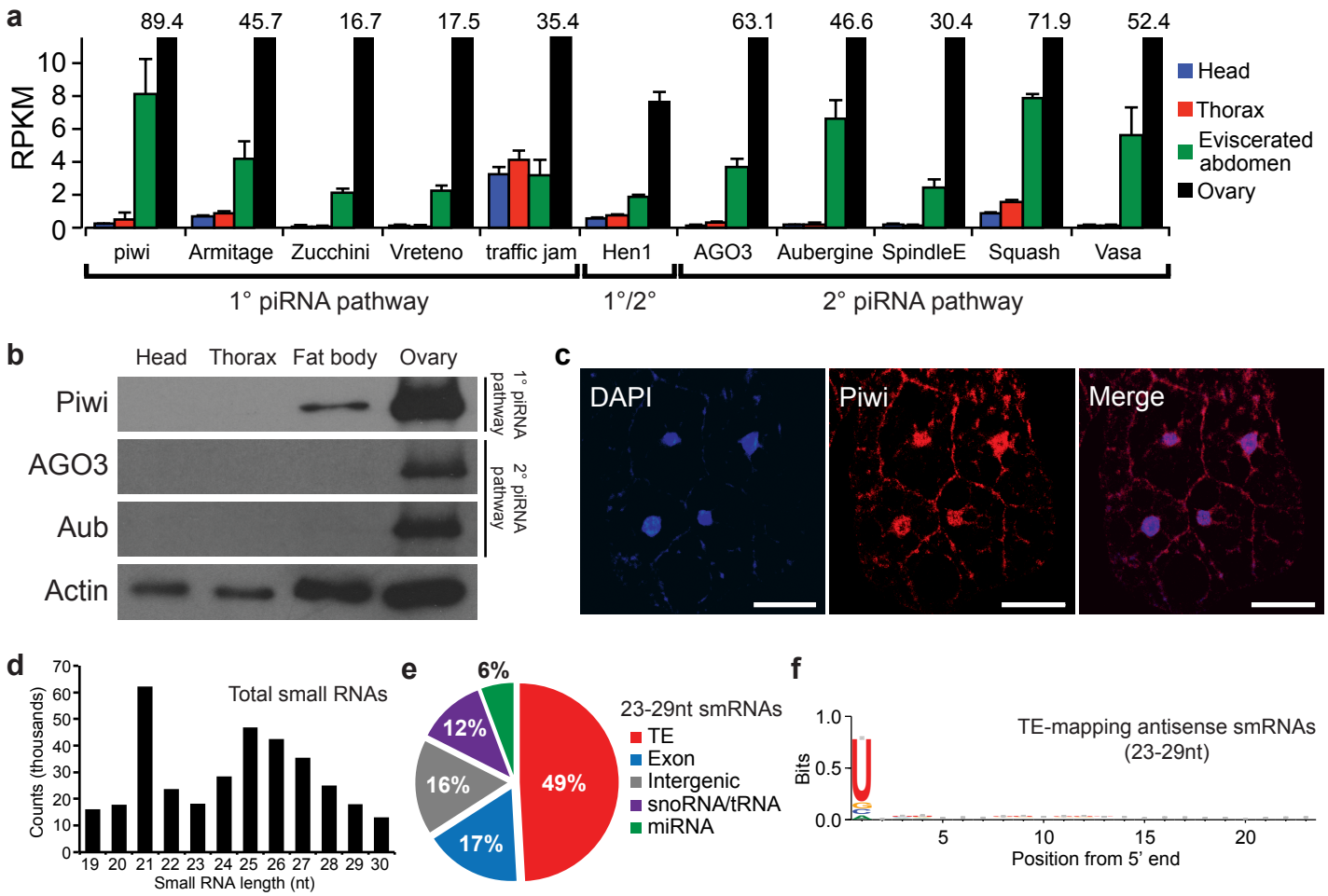


Figure 1: Canonical signatures of a piRNA pathway in the adult fly fat body.
a, Expression of primary and secondary piRNA pathway genes generated from total RNA-seq libraries of head, thorax, eviscerated abdomen, and ovary. piRNA pathway genes are more highly expressed in the eviscerated abdomen than in the head or thorax. Data values for ovary libraries that exceed the range of the plot are shown above each relevant bar. RPKM: Reads Per Kilobase per Million. Error bars represent S.E.M. *n*=3 replicate libraries. In comparing the eviscerated abdomen with head and thorax controls, 10 of 11 genes (excluding *tj*) are statistically significant; $P<0.0001$. See Extended Data Table 1 for statistics.
b, Piwi protein localizes to the nuclei of abdominal fat body cells. DAPI labels fat body nuclei. Staining in the membrane is autofluorescence typical of fat body cells. Scale bars represent 20 μm.
c, Piwi protein is present in the fat body. All piRNA argonautes are present in the ovary samples. Actin serves as a loading control.
d, Fat body smRNA size profile from oxidized smRNA-seq libraries. Oxidation allows for enrichment of 2'-*O*-methylated smRNAs. Peak at 21nt likely represents siRNA population. Broader peak from 23-29nt represents putative fat body piRNAs. Reads aligning to rRNA and miRNA were excluded from analysis.
e, Fat body piRNAs (23-29nt) aligned to the fly genome map primarily to TEs.
f, Sequence composition of TE-mapping fat body piRNAs (23-29nt) displays a first position nucleotide bias for uracil.

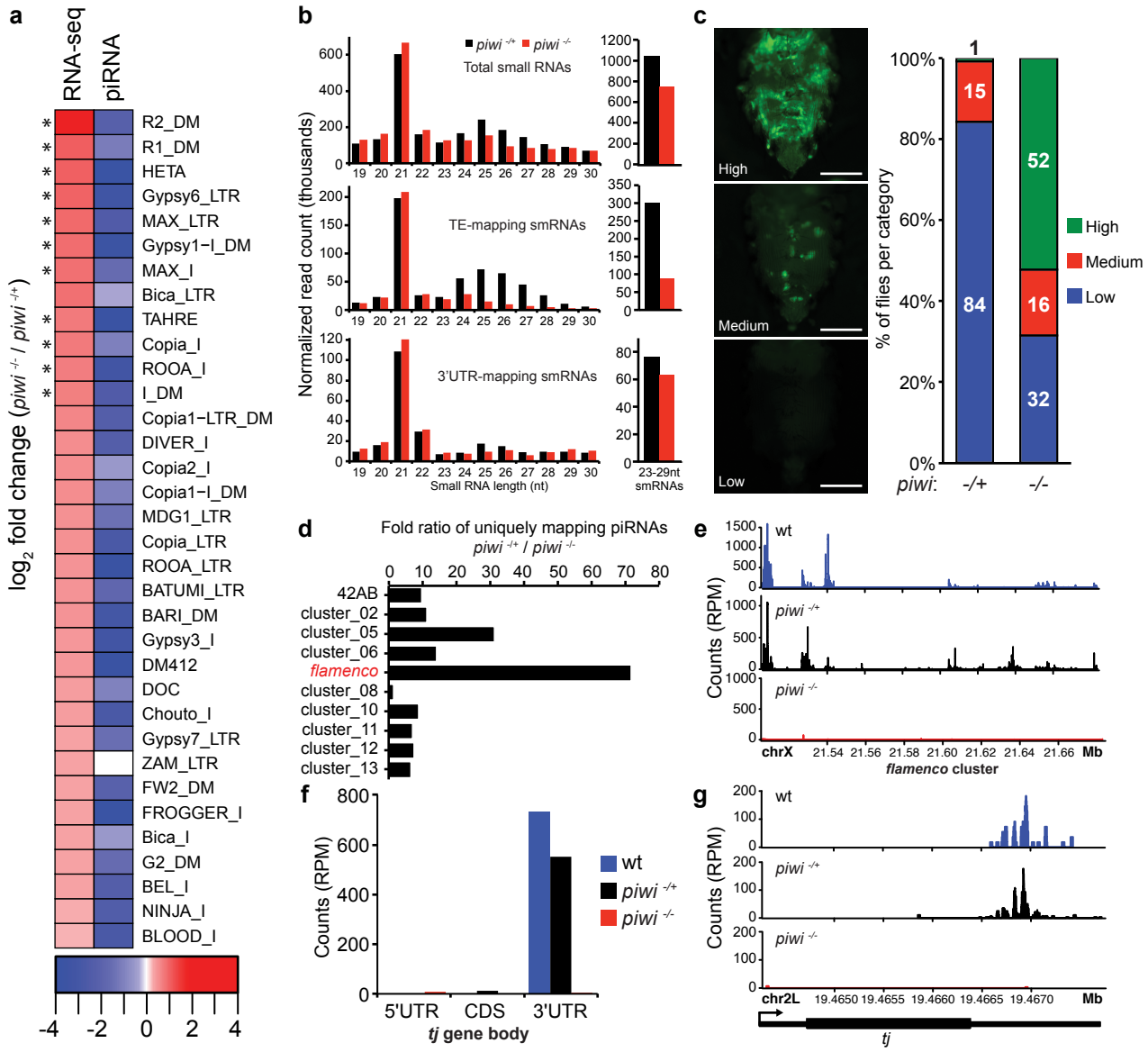


Figure 2: The fat body piRNA pathway is active.

a, Heat map of log2 fold change of TE transcript levels (total RNA-seq) >1.2 fold change and corresponding piRNAs (smRNA-seq) in *piwi* mutant (*piwi* $^{-/-}$) compared to heterozygous control (*piwi* $^{-/+}$) fat bodies. *n*=3 replicate libraires. *FDR<0.05.

b, Fat body smRNA size profile from oxidized smRNA-seq libraries of *piwi* mutant fat bodies (red) and heterozygous controls (black). Shown are total smRNAs (top), TE-mapping smRNAs (middle), and 3'UTR-mapping smRNAs (bottom). Panels at right show levels for 23-29nt smRNAs for each genotype.

c, A fat body-specific transposition reporter line, *gypsy-TRAP/r4-GAL4::UAS-GFP*, in a *piwi* mutant background (see Methods). GFP-positive cells are cells in which a transposition event has activated reporter function. 10 day old *piwi* mutants show elevated levels of GFP-positive cells compared to heterozygous controls. Images are of representative flies exhibiting high, medium, and low levels of GFP-positive cells (left panels). Scale bars represent 150 µm. Panel on right shows the distribution of flies for each group by genotype; numbers within each bar are corresponding percentages for each group. Fisher's exact test compared the combined medium+high groups and the low group of each genotype. *piwi* $^{-/+}$: *n*=140 flies; *piwi* $^{-/-}$: *n*=142 flies; *P*<0.0001.

d, Fold ratio of uniquely mapping cluster piRNAs in *piwi* mutants compared to heterozygous controls. The *flamenco* cluster shows the greatest response to loss of *piwi*.

e, Unique piRNA reads (23-29nt) map to the *flamenco* locus in wt and *piwi* heterozygotes and are lost in *piwi* mutants.

f, **g**, Unique piRNA reads (23-29nt) map to the *tj* gene body in wt and *piwi* heterozygotes and are lost in *piwi* mutants. Thick lines in gene model (**g**) represent coding sequence, thin lines represent 5' and 3'UTRs. See Supplementary Table 1 for raw data.

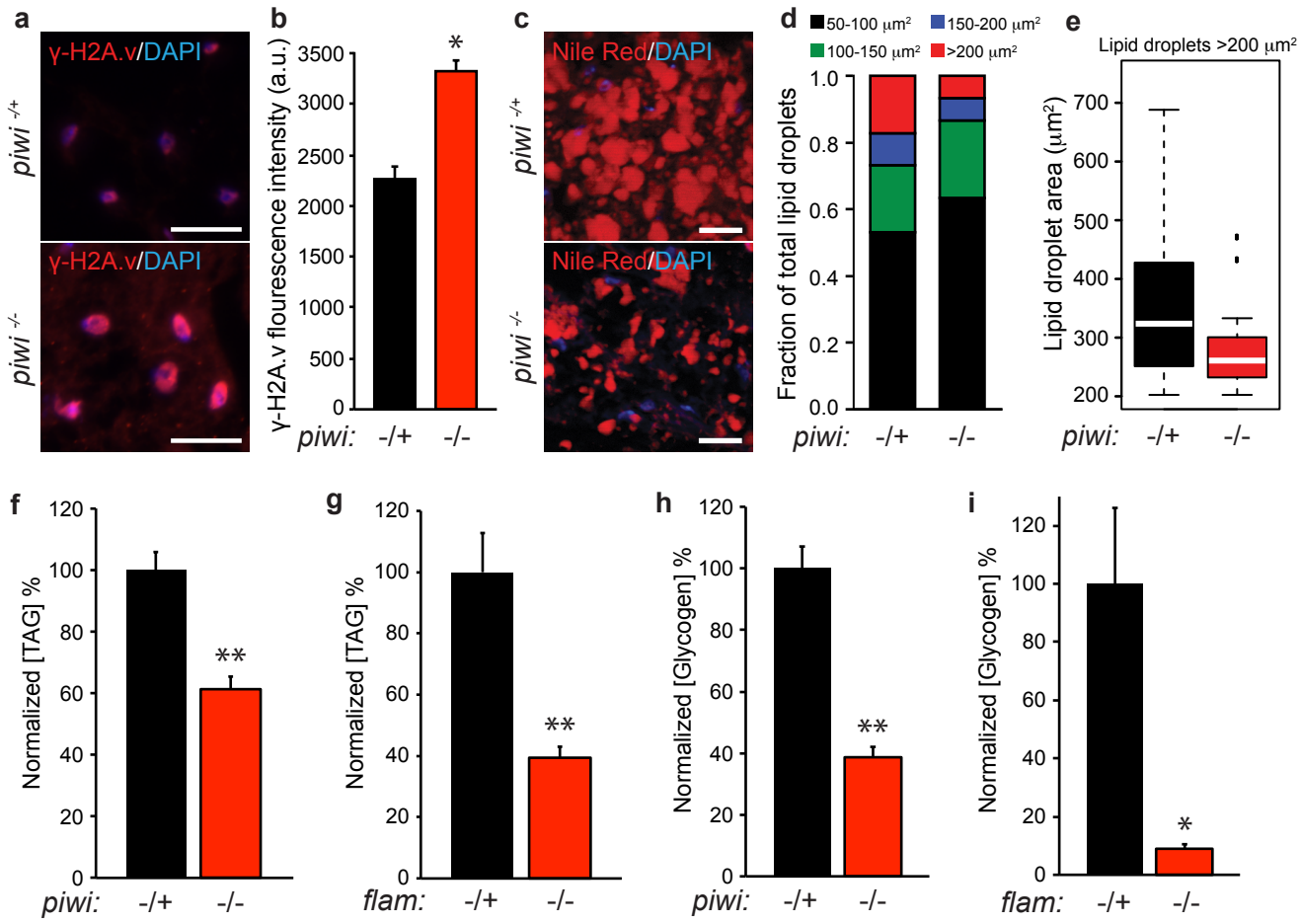


Figure 3: Loss of the piRNA pathway disrupts normal metabolic functions of the fat body.
a, Representative images of γ-H2A.v staining in *piwi* mutants and heterozygous controls. *piwi* mutants exhibit higher levels of γ-H2A.v staining compared to heterozygous controls. Scale bars represent 20 μm.
b, Quantification of γ-H2A.v staining in (**a**). *piwi* $^{-/+}$ and *piwi* $^{-/-}$: Student's two-tailed *t*-test compared to heterozygous control. *n*=104 nuclei per genotype. **P*<0.0001.
c, Representative images of Nile red staining of fat body lipid droplets in *piwi* mutants and heterozygous controls. *piwi* mutants exhibit smaller lipid droplets relative to heterozygous controls. Scale bars represent 20 μm.
d, Quantification of lipid droplet staining in (**c**). *piwi* $^{-/+}$: *n*=6 flies, *piwi* $^{-/-}$: *n*=5 flies. See Supplementary Table 2 for raw data.
e, Box plot showing distribution of lipid droplets >200 μm² from (**d**) comparing *piwi* mutants to heterozygous controls.
f-**i**, Measurements of whole-body adult fly triacylglycerides (TAG) (**f, g**) and glycogen (**h, i**) of *piwi* (**f, h**) or *flamenco* (**g, i**) mutants compared to heterozygous controls. Data were normalized to total protein concentration of each sample and represented as a percent of the heterozygous control. Error bars are S.E.M. For each assay, *n*=5 biological replicates per genotype. Student's two-tailed *t*-test compared to heterozygous control. **P*<0.01; ***P*<0.001.

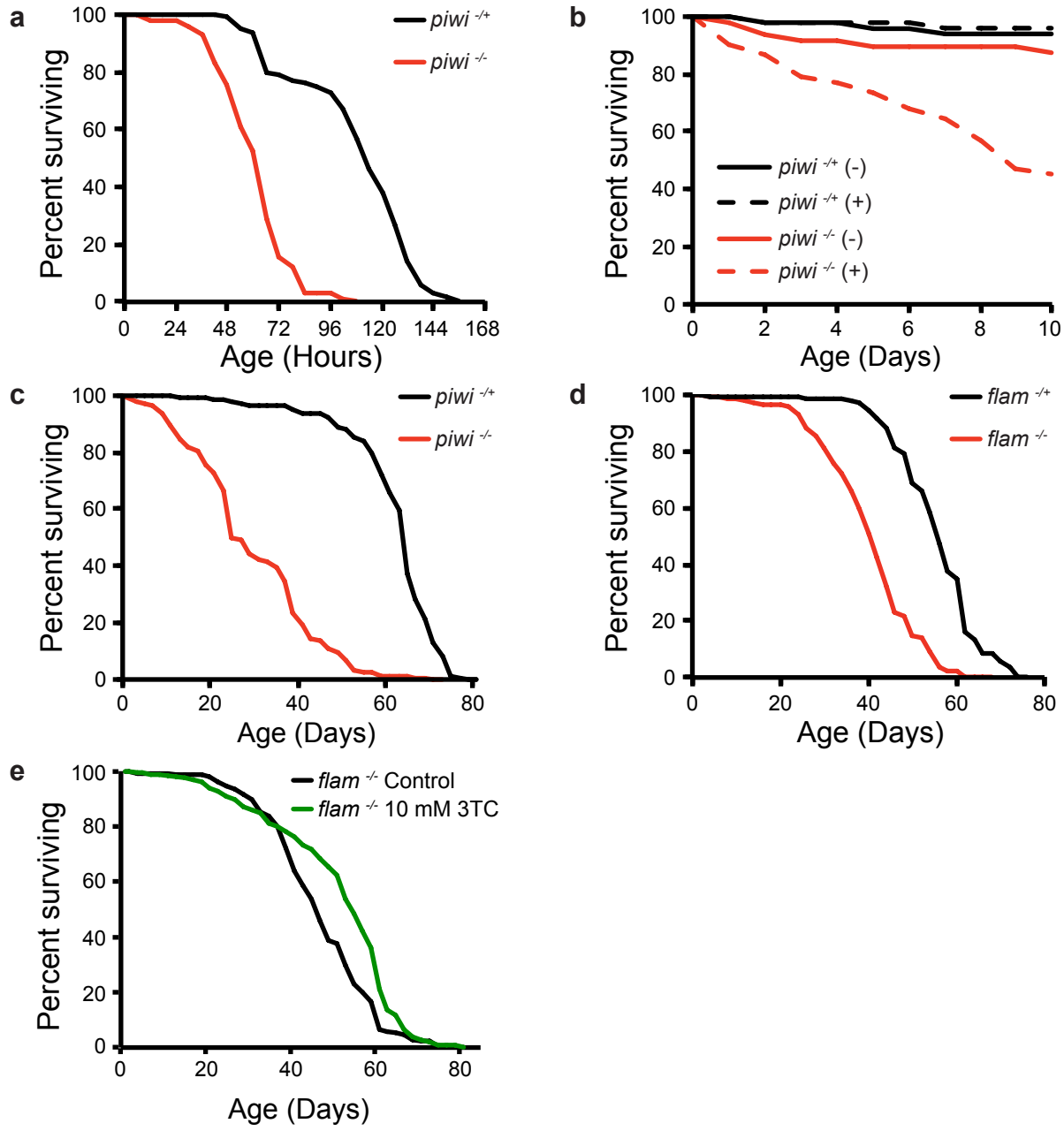


Figure 4: piRNA pathway mutants are stress-sensitive and short-lived.

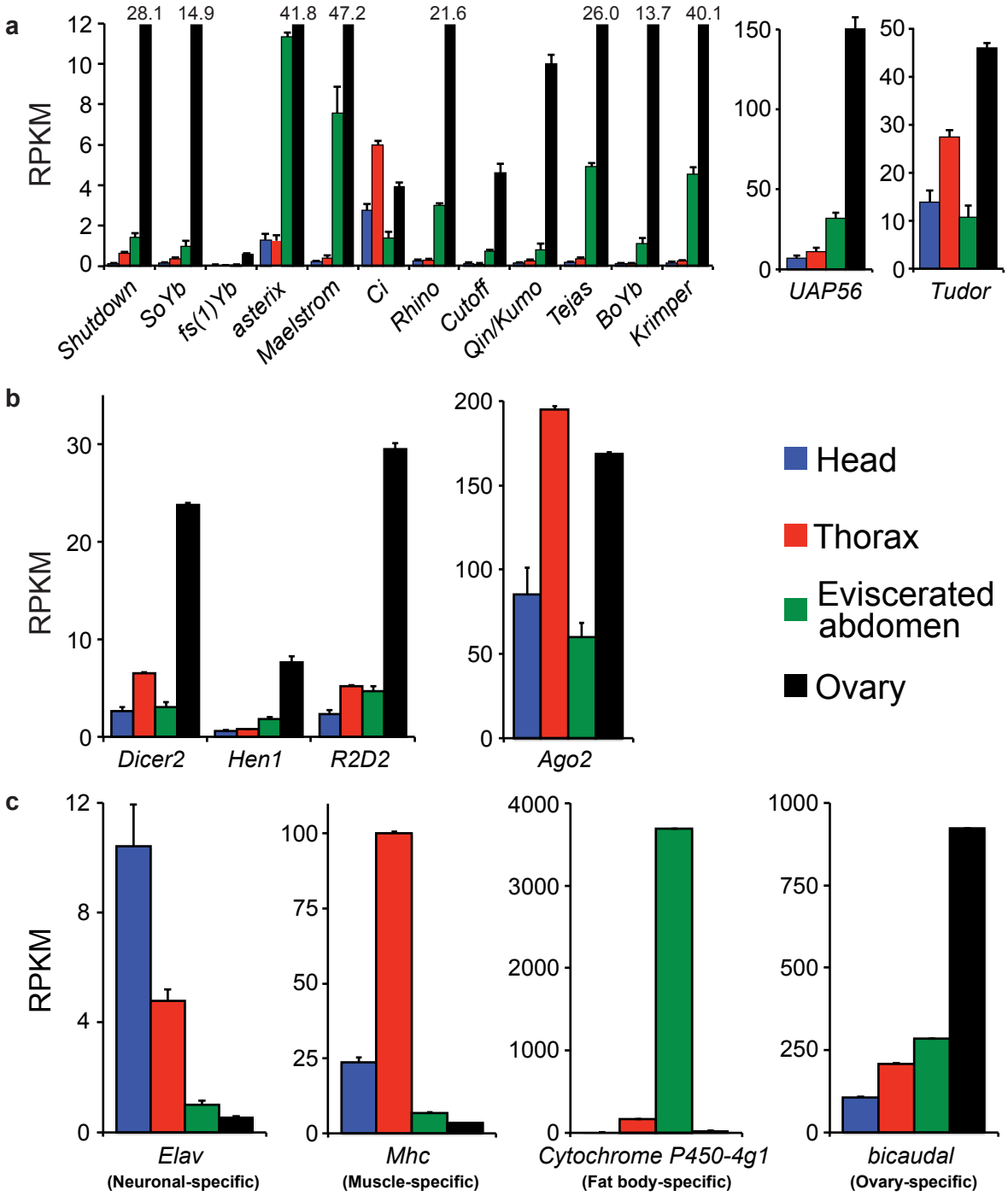
a, Survivorship curves for starvation of *piwi* mutants and heterozygous controls. *piwi* mutants are more sensitive to starvation than heterozygous controls. Log rank test compared to heterozygous controls; n≈50; $P<0.0005$.

b, Survivorship curves for immune challenge of *piwi* mutants and heterozygous controls. *piwi* mutants are more sensitive to infection than heterozygous controls. Flies were either infected with a mock EtOH control (-) or a culture of *E. carotovora* (+). Log rank test compared to heterozygous or mock EtOH control (-); n≈50; $P<0.0005$.

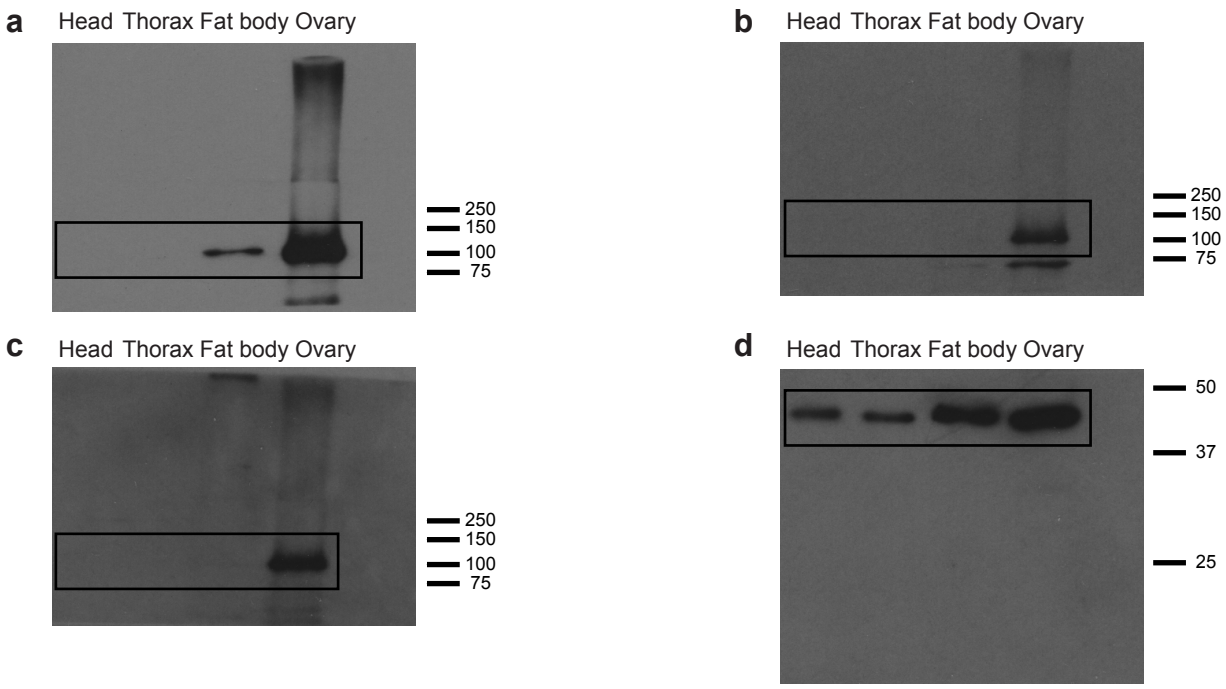
c, Survivorship curves of *piwi* mutants and heterozygous control. *piwi* mutants are shorter lived compared to heterozygous controls. Wilcoxon rank sum test compared to heterozygous control; n≈250; $P<0.0005$.

d, Survivorship curves of *flam* mutants and heterozygous control. *flam* mutants are shorter lived compared to heterozygous controls. Wilcoxon rank sum test compared to heterozygous control; n≈250; $P<0.0005$.

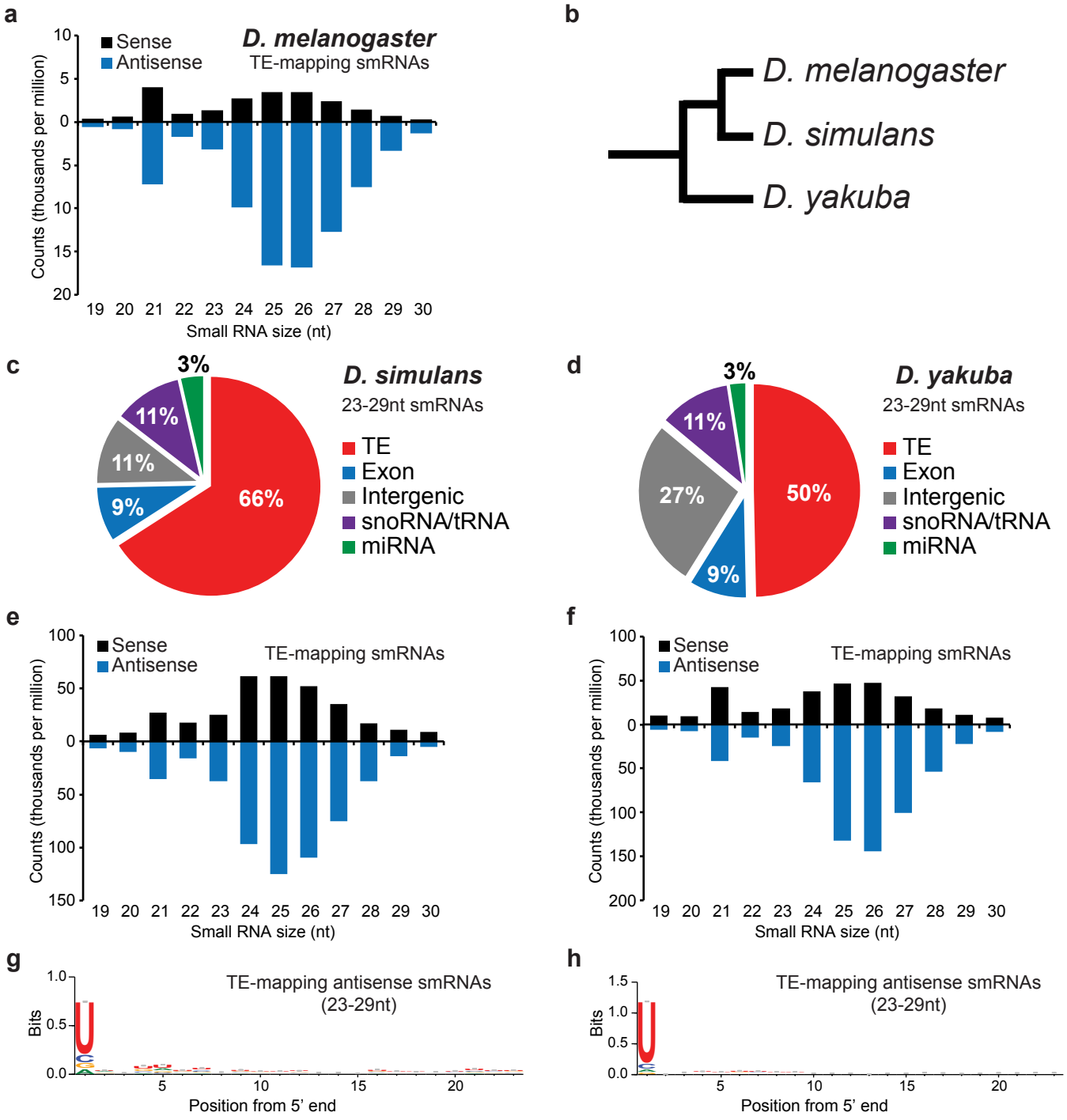
e, Survivorship curves of *flam* mutants treated with 10 mM 3TC. Flies fed 3TC are long-lived compared to untreated controls. Wilcoxon rank sum test compared to control; n≈250; $P<0.0005$.



Supplementary Figure 1: Gene expression profiles of somatic and reproductive tissues.
(**a**) Expression of additional piRNA pathway genes generated from total RNA-seq libraries of head, thorax, eviscerated abdomen, and ovary. Data values for ovary libraries that exceed the range of the plot are shown above each relevant bar.
(**b**) Expression of siRNA pathway genes.
(**c**) Expression of tissue-specific control genes.
RPKM: Reads Per Kilobase per Million. Error bars represent S.E.M. *n*=3 replicate libraries. Of the 14 genes presented, 11 are statistically significantly increased in eviscerated abdomen compared to the head and thorax, one is significantly lower in eviscerated abdomen vs. thorax (*Ci*), and two are not significantly different (*fs(1)Yb* and *Tudor*); *P*<0.0001. See Supplementary Dataset 1 for relevant statistics.



Supplementary Figure 2: Original immunoblots used in Figure 1.
Immunoblots staining for anti-Piwi (**a**), anti-AGO3 (**b**), anti-Aub (**c**), and anti-Actin (**d**). Black blocks indicate the bands selected for the main figure.



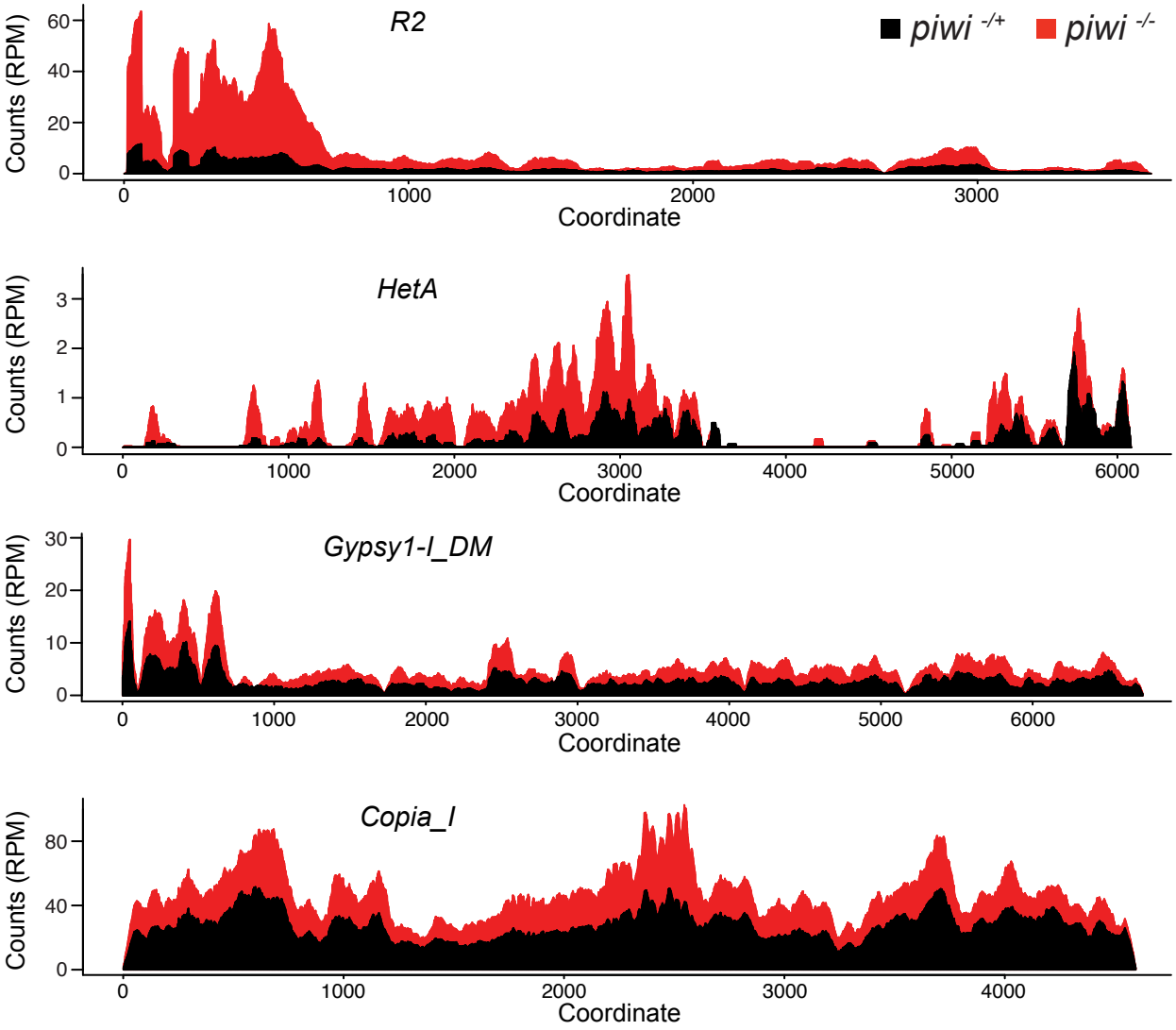
Supplementary Figure 3: Fat body piRNA pathway signatures are conserved across drosophilid species.

(**a**, **e**, **f**,) Fat body smRNA size profile of TE-mapping reads from oxidized smRNA-seq libraries (see **Fig. 1d**) of *D. melanogaster* (**a**), *D. simulans* (**e**), and *D. yakuba* (**f**). Reads are divided by mapping either sense (black) or antisense (blue) to TEs. Peak at 21nt likely represents siRNA population. Broader peak from 23-29nt represents putative fat body piRNAs.

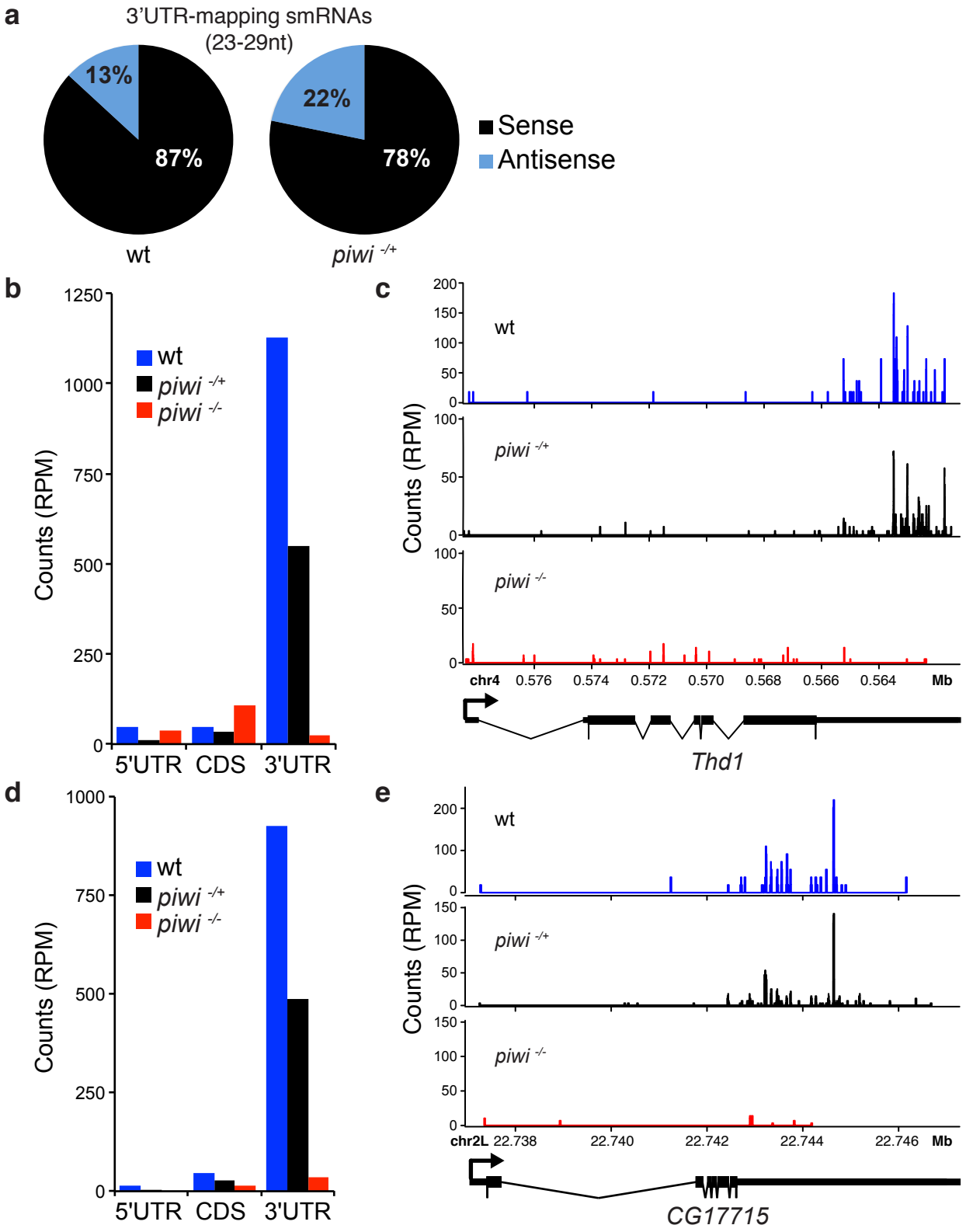
(**b**) Schematic cladogram depicting the evolutionary relationships between three drosophilid species.

(**c**, **d**) Putative fat body piRNAs (23-29nt) aligned to the *D. simulans* (**c**) and *D. yakuba* (**d**) genomes map primarily to TEs.

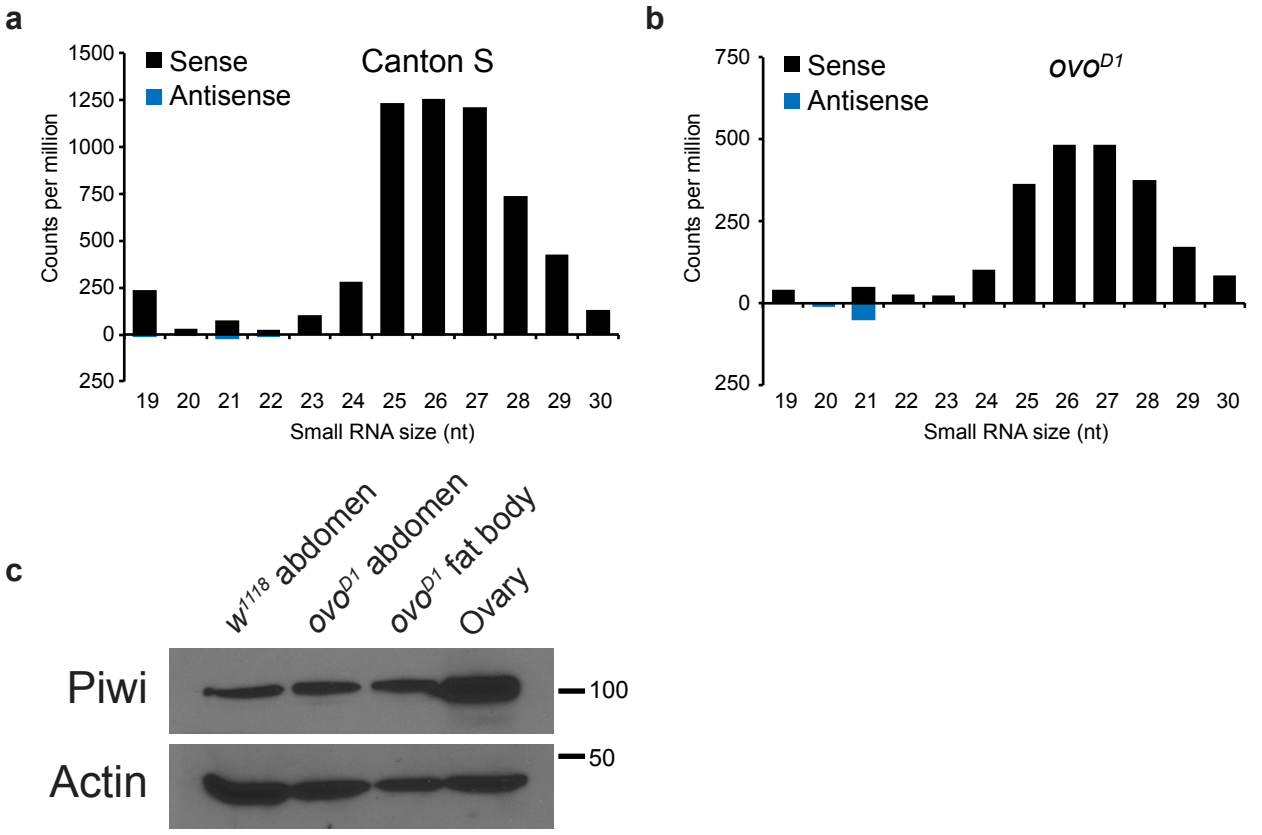
(**g**, **h**) Sequence composition of TE-mapping putative fat body piRNAs (23-29nt) from *D. simulans* (**g**) and *D. yakuba* (**h**) displays a first position nucleotide bias for uracil.



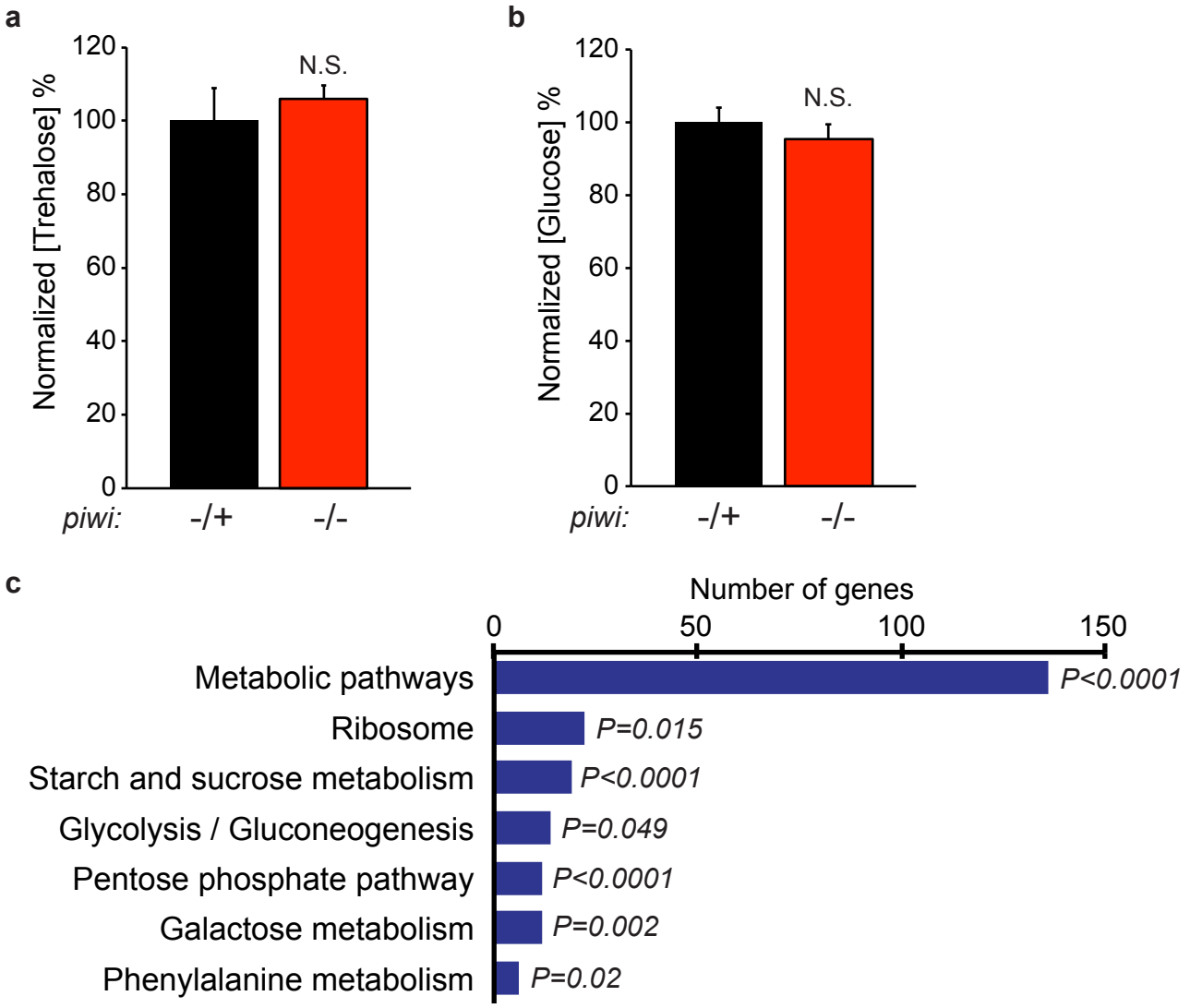
Supplementary Figure 4: Expression profiles of representative upregulated TEs in *piwi* mutant and control fat bodies.
piwi mutants show increased expression of TEs relative to heterozygous controls. Reads from *piwi* mutants (red) and controls (black) are shown spanning the consensus sequences of four representative TEs from **Fig. 2a**. RPM: Reads Per Million. *n*=3 replicate libraries.



Supplementary Figure 5: Characteristics of 3'UTR-mapping fat body piRNAs.
(**a**) 3'UTR-mapping piRNAs (23-29nt) in wt and *piwi* $^{-/+}$ fat bodies are predominantly sense to coding genes.
(**b**-**e**) Representative genes featuring abundant sense 3'UTR-mapping piRNAs. Unique piRNA reads (23-29nt) map to the *Thd1* (**b**, **c**) and *CG17715* (**d**, **e**) gene bodies in wt and *piwi* heterozygotes and are lost in *piwi* mutants. Thick lines in gene model represent coding sequence; thin lines represent 5' and 3'UTRs. RPM: Reads Per Million. See Supplementary Dataset 2 for additional fat body genic piRNAs.



Supplementary Figure 6: Flies lacking ovaries maintain fat body piRNA signatures.
(**a**, **b**) Size profile of small RNA-seq reads from purified fat body mapping uniquely to *flamenco* locus in wild-type Canton S flies (**a**) and *ovo*^{D1} mutant flies lacking ovaries (**b**). Reads were normalized to total uniquely aligning reads in each library.
(**c**) Piwi protein is present in the eviscerated abdomens and pure fat bodies of *ovo*^{D1} mutant flies lacking ovaries. Wild-type *w*^{1118} eviscerated abdomens and ovaries serve as positive controls. Actin serves as a loading control.



Supplementary Figure 7: Metabolic phenotypes of *piwi* mutants.
(**a**, **b**) Measurements of whole-body adult fly trehalose (**a**), and glucose (**b**) of *piwi* mutants compared to heterozygous controls. Data were normalized to total protein concentration of each sample and represented as a percent of the heterozygous control. Error bars are S.E.M. For each assay, *n*=5 biological replicates per genotype.
(**c**) KEGG pathway analysis of differentially expressed coding genes in fat body of *piwi* mutants compared to heterozygous controls. *n*=3 replicate libraries.

piwi Starvation - 092414		Hours at % Mortality		vs. (-/+) Control	
Genotype	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
piwi $^{-/+}$	48	114	132	—	—
piwi $^{-/-}$	48	48	66	77.81	*P*<0.0005

piwi Starvation - 100614		Hours at % Mortality		vs. (-/+) Control	
Genotype	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
piwi $^{-/+}$	100	114	138	—	—
piwi $^{-/-}$	97	66	84	131.39	*P*<0.0005

piwi Immune Challenge - 070215		Days at % Mortality		vs. Respective (-/+) Control		vs. Respective (-) Control	
Genotype/Condition	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**	**X^2**	***P*-value**
piwi $^{-/+}$ (-)	50	N/A	N/A	—	—	—	—
piwi $^{-/-}$ (-)	47	N/A	N/A	1.35	N.S.	—	—
piwi $^{-/+}$ (+)	50	N/A	N/A	—	—	0.21	N.S.
piwi $^{-/-}$ (+)	53	9	N/A	31.08	*P*<0.0005	18.07	*P*<0.0005

piwi Immune Challenge - 073115		Days at % Mortality		vs. Respective (-/+) Control		vs. Respective (-) Control	
Genotype/Condition	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**	**X^2**	***P*-value**
piwi $^{-/+}$ (-)	50	N/A	N/A	—	—	—	—
piwi $^{-/-}$ (-)	50	N/A	N/A	17.43	*P*<0.0005	—	—
piwi $^{-/+}$ (+)	50	N/A	N/A	—	—	6.32	*P*=0.0119
piwi $^{-/-}$ (+)	50	10	N/A	23.56	*P*<0.0005	14.74	*P*<0.0005

piwi Lifespan - 080914		Days at % Mortality		vs. (-/+) Control	
Genotype	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
piwi $^{-/+}$	254	67	75	—	—
piwi $^{-/-}$	245	29	51	389.06	*P*<0.0005

piwi Lifespan - 072414		Days at % Mortality		vs. (-/+) Control	
Genotype	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
piwi $^{-/+}$	238	66	72	—	—
piwi $^{-/-}$	225	30	56	314.39	*P*<0.0005

flam Lifespan - 040815		Days at % Mortality		vs. (-/+) Control	
Genotype	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
flam $^{-/+}$	250	58	68	—	—
flam $^{-/-}$	245	44	56	226.93	*P*<0.0005

flam Lifespan - 051615		Days at % Mortality		vs. (-/+) Control	
Genotype	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
flam $^{-/+}$	263	58	68	—	—
flam $^{-/-}$	258	42	58	221.99	*P*<0.0005

flam $^{-/-}$ 3TC Lifespan - 012816		Days at % Mortality		vs. EtOH Control	
Genotype/Condition	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
flam $^{-/-}$ (Control)	240	47	61	—	—
flam $^{-/-}$ (3TC)	251	55	67	22.48	*P*<0.0005

flam $^{-/-}$ 3TC Lifespan - 051916		Days at % Mortality		vs. EtOH Control	
Genotype/Condition	***n*=X**	**50%**	**90%**	**X^2**	***P*-value**
flam $^{-/-}$ (Control)	246	47	71	—	—
flam $^{-/-}$ (3TC)	254	57	67	16.97	*P*<0.0005

Supplementary Table 1: Parameters and statistical analysis of starvation, immune challenge, and lifespan assays.

Chapter 3

Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend lifespan in *Drosophila*

Jason G. Wood, Brian C. Jones, Nan Jiang, Chengyi Chang, Suzanne Hosier, Priyan Wickremesinghe, Meyrolin Garcia, Davis A. Hartnett, Lucas Burhenn, Nicola Neretti, Stephen L. Helfand

This project was conceived of and designed by Jason Wood, with contributions from Stephen Helfand, Nicola Neretti, and me. Jason Wood and Chengyi Chang performed all sample and sequencing library preparation with assistance from Priyan Wickremesinghe, Meyrolin Garcia, and Lucas Burhenn. Jason Wood performed all bioinformatic analyses with advice from Nicola Neretti and me. I conceived of, designed, and made the fat body *gypsy-TRAP* line and experiment. Stephen Helfand performed the *gypsy-TRAP* experiment. Jason Wood, Nan Jiang, and I designed and prepared all figures. Jason Wood assayed and analyzed the γ-H2A.v experiments with assistant from Davis Hartnett. I performed and analyzed the *Dicer-2* overexpression lifespans. Nan Jiang performed the *Su(var)3-9* overexpression lifespans and qPCR experiment. Suzanne Hosier performed all other lifespans. Jason Wood wrote the manuscript with assistance from Stephen Helfand and myself.

This work has been published in *PNAS* (2016) October 4; 113(40):11277-11282.
DOI: 10.1073/pnas.1604621113.

Abstract

Transposable Elements (TEs) are mobile genetic elements, highly enriched in heterochromatin, that constitute a large percentage of the DNA content of eukaryotic genomes. Aging in *Drosophila melanogaster* is characterized by loss of repressive heterochromatin structure and loss of silencing of reporter genes in constitutive heterochromatin regions. Using next-gen sequencing, we found that transcripts of many genes native to heterochromatic regions as well as TEs increased with age in fly heads and fat bodies. A dietary restriction regimen, known to extend lifespan, repressed the age-related increased expression of genes located in heterochromatin as well as TEs. We also observed a corresponding age-associated increase in TE transposition in fly fat body cells that was delayed by dietary restriction. Furthermore, we found that manipulating genes known to affect heterochromatin structure, including overexpression of *Sir2*, *Su(var)3-9*, and *Dicer-2*, as well as decreased expression of *Adar*, mitigated age-related increases in expression of TEs. Increasing expression of either *Su(var)3-9* or *Dicer-2* also led to an increase in lifespan. Mutation of *Dicer-2* led to an increase in DNA double strand breaks. Treatment with the reverse transcriptase inhibitor 3TC resulted in decreased TE transposition as well as increased lifespan in TE sensitized *Dicer-2* mutants. Together, these data support the retrotransposon theory of aging, which hypothesizes that epigenetically silenced TEs become deleteriously activated as cellular defense and surveillance mechanisms break down with age. Furthermore, interventions that maintain repressive heterochromatin and preserve TE silencing may prove key to preventing damage caused by TE activation and extending healthy lifespan.

Introduction

There is mounting evidence that preservation of the epigenetic structure of chromatin is important to maintain healthy aging, and loss of this chromatin structure may be a hallmark of the aging process (reviewed in (1)). In aging yeast, cells lose characteristic silencing in heterochromatic regions, resulting in deleterious age-related consequences such as sterility, and replicative lifespan is regulated by the histone H4K16 acetylation mark (2, 3). In *C. elegans*, genetic manipulation of chromatin factors affecting histone methylation and acetylation directly affect lifespan (4). In *Drosophila*, the integrity of constitutive repressive heterochromatin declines with age at both the genomic and cellular level (5), with one consequence of these changes being an age-related loss of gene silencing in repressive constitutive heterochromatin regions in multiple tissues of the fly (6). In mice, changes in epigenetic chromatin marks are seen with age, and age-related phenotypes of behavior, memory, and learning are observed upon genetic manipulation of certain histone marks (1). A breakdown in heterochromatin organization has also been observed during replicative senescence of human cells in culture, with expression and copy number of transposable elements (TEs) increasing (7, 8).

TEs are mobile genetic elements that make up a large percentage of eukaryotic genomes, and are particularly enriched in regions of heterochromatin (reviewed in (9)). Because TEs function by either excising or copying themselves from the genome and integrating into new genomic locations, there is a high potential for deleterious effects on cellular function. Overexpressing TE RNAs is sufficient to cause cytotoxicity, and

inhibiting expression of these TEs causes reversal of both senescence and cytotoxic phenotypes (10, 11). TE activation has been observed with age in mice (8, 12), as well as *C. elegans* (13), yeast (14), *Drosophila* (15, 16), and senescent human tissue culture cells (7).

To combat the genomic threat of activation of TEs and loss of heterochromatic gene silencing, somatic cells have evolved mechanisms to suppress TE transcription, primarily through establishment and maintenance of repressive heterochromatin at TE sites (17). In *Drosophila*, the small interfering RNA (siRNA) pathway, through the proteins Dicer-2 and AGO2, recruits the histone H3K9 methyltransferase Su(var)3-9 to catalyze formation of repressive heterochromatin at TE sites (18–20). Mutation of Dicer-2, the double stranded RNase responsible for generating siRNAs, results in TE reactivation and heterochromatin disorganization (19, 20). The RNA-editing enzyme ADAR is thought to modify the effectiveness of siRNA-directed heterochromatinization, and therefore the efficacy of TE silencing, by editing the dsRNA substrates used for siRNA biogenesis, thereby preventing proper processing by Dicer-2 (21). Sir2/Sirt1, the founding member of the sirtuin family of proteins, is a protein deacetylase important in maintaining gene silencing in heterochromatin regions of yeast and flies through deacetylation of specific histone lysines required for establishment and spreading of heterochromatin (3, 6, 22). Increasing Sir2 expression has been shown to repress the normal age-related loss of gene silencing in heterochromatin regions in both yeast and flies (3, 6), as well as the loss of gene silencing seen in mammalian embryonic stem cells undergoing genotoxic stress (23). In addition to their roles in heterochromatin formation and maintenance, these pathways have also been linked to aging.

Decreasing activity of the siRNA pathway through *Ago2* and *Dicer-2* mutants leads to de-repression of TE activity and shortened lifespan (15, 24). Sir2 has a central role in aging with increased gene dosage shown to extend lifespan in yeast (2) and in nine out of ten studies in nematodes and flies (25–28). Hypomorphs of *Adar*, a negative regulator of the siRNA pathway, also exhibit an extension of lifespan (21). These pathways share an ability to impact heterochromatin formation and affect lifespan.

Here we report a characterization of the heterochromatic transcriptome in adult *Drosophila* tissues and show activation of TEs and increased transposition with age. We provide evidence that the aging process is characterized by a failure to properly control and repress expression of heterochromatin based genes and TEs in the soma. We show that genetic interventions positively modifying heterochromatin structure or negatively modulating TE activity extend metazoan lifespan. Finally, we show that attenuating TE activity pharmacologically decreases transposition and extends lifespan in *Dicer-2* mutants. This study offers the first comprehensive characterization of the effect of known lifespan extending interventions on heterochromatic regions and explores new interventions for controlling age-associated TE reactivation.

Results

Expression of genes located within heterochromatin increases with age and is attenuated under dietary restriction

Because heterochromatic regions change in structure with age, specifically in female heads and fat bodies (5), we reasoned that genes located within these regions may also show changes in silencing or gene expression with age. We used RNA-seq to

assay the expression of the ~250 genes natively located within the *Drosophila* heterochromatin compartments (29) from young (10 day) and old (40 day) female fly heads and fat bodies. The vast majority of cells in the head are neurons and glial cells. The fly fat body is the functional homolog of mammalian adipose tissue and liver, and a central organ in the fly immune system. We found a number of genes located in heterochromatic regions increased expression with age in both wild-type heads and fat bodies (Fig. 1A-B). Gene ontology analysis reveals no significantly enriched pathways among these genes, suggesting that the age-related increase in expression of these genes is likely to be a consequence of loss of heterochromatic silencing rather than a specific compensatory or correlative response tied to aging.

Dietary restriction (DR) is an intervention known to extend lifespan and delay numerous age-related phenotypes. We previously observed that DR can oppose age-related loss of gene silencing of reporter genes in constitutive heterochromatin (6). Based on this, we examined whether the age-dependent increase in expression of native genes in the constitutive heterochromatin regions were also suppressed under a DR regimen. The vast majority (95% in heads, 89% in fat bodies) of heterochromatin genes that increased expression with age on high calorie food increased less or decreased in expression with age on low calorie food (HC vs. DR) in both heads (Fig. 1A) and fat bodies (Fig. 1B). Although we observed a similar fraction of euchromatin genes that increased expression with age, this pattern of suppression by DR was not maintained in the most highly upregulated euchromatin genes (Fig. S1A-B), suggesting that DR has a selective effect on heterochromatin genes. Similar to previous studies using reporter genes in both yeast (2) and flies (6), these data suggest that gene

expression in heterochromatic regions becomes dysregulated with age, and that a known lifespan extending intervention, DR, can prevent this.

TE expression increases with age and is attenuated under dietary restriction

One of the most notable attributes of heterochromatin is the abundant enrichment of TEs within these regions. The *Drosophila* genome is rich in retroviral-like long terminal repeat (LTR) and LINE elements (30), and activation of both LINEs and LTRs have been observed in the fly nervous system and fat body during aging (15, 16). Since heterochromatin integrity declines with age in multiple model systems (1), we hypothesized that a loss of heterochromatin in regions highly enriched with TEs would show a correlated increase in TE expression with age. To test this hypothesis, we examined the expression profile of TEs with age and diet in female fly heads and fat bodies using whole transcriptome RNA-seq. Similar to our observations for native heterochromatin genes, we also found many TEs that increased their expression with age, in both heads and fat bodies (Fig. 1C-D). We observed a high degree of overlap in the TEs that increased expression most with age between these two tissues, most of which were retrotransposons (Fig. S2A). In both tissue types, as observed with the heterochromatin genes, dietary restriction attenuated the age-related increase in expression observed in high calorie food (Fig. 1C-D, right “DR” columns). TEs that had the highest increase in expression with age on high calorie food tended to increase less or even decrease in expression with age under DR conditions (87% lower in DR relative to HC in both heads and fat bodies; Dataset S1). These data suggest that similarly to

reporter genes in heterochromatin (2, 6) and native heterochromatin genes (Fig. 1A-B), TE silencing is lost with age, and DR opposes this.

TE transposition increases with age and is delayed by dietary restriction

An age-dependent increase in TE expression and transposition in the *Drosophila* brain associated with decline in neural function has been reported (15). Similar increases in TE expression levels and copy numbers have been observed in aging mouse tissue (8). To assay for age-related transposition in our flies, we used the *gypsy*-TRAP reporter system (15). This system is designed such that GFP becomes expressed when an endogenous *gypsy* retrotransposon jumps into an engineered site containing known hotspot sequences from the *ovo* locus (31). We monitored flies as they aged and visually sorted them into either low, medium, or high categories based on the number of GFP-positive cells observed in the fat body (Fig. 2A). Longitudinal fly cohorts were assayed for transposition every seven days on either high calorie or dietary restriction foods (Fig. 2B). We observed a clear age-associated increase in transposition, and DR delayed this increase relative to high calorie controls (Fig. 2B-C). Similar results were seen in two additional genetic backgrounds (Fig. S3A-D). We examined the lifespans of these three different genetic background *gypsy*-TRAP lines and observed that lifespan correlated with the timing of increased transposition. Lines exhibiting earlier transposition showed shorter lifespan and those with later transposition longer lifespan (Fig. S3E). Similarly, the delay in age-related transposition due to DR correlated with the ability of DR to extend lifespan in each line. In order to ensure the increased GFP expression we observed with age was due to transposition rather than

age-related disruption in the GAL80/GAL4/GFP reporter system, we examined a *gypsy*-TRAP line with a mutant *ovo* locus not susceptible to transposition, and observed no significant increase in GFP expression with age (Fig S3F). Rather than a constant linear increase in GFP expression with time, the *gypsy*-TRAP assay showed a clear age-associated increase in TE mobilization, consistent with the TE theory of aging, which posits an age-related loss of normal cellular surveillance mechanisms and consequent increase in deleterious TE activity. The relationship between the timing of increased gypsy transposition and lifespan further supports TE mobilization as an age-associated phenomenon.

Age-related increases in TE expression are attenuated by genetic interventions that affect heterochromatin structure

Our data demonstrate a correlation between changes in heterochromatin structure with age and expression and transposition of TEs during normal aging. This led us to examine whether genetically manipulating heterochromatin structure and maintenance, either directly or indirectly, could also affect the age-related changes in expression of TEs. To test our hypothesis, we genetically increased the activity of two important systems that maintain repressive heterochromatin in somatic cells—the siRNA pathway and Sir2. Using transgenic fly lines, we increased gene dosage of *Sir2*, *Dicer-2*, and *Su(var)3-9*, and also used a *Adar* hypomorph, interventions we expect to stabilize heterochromatin. Each of these genes offers a unique and complementary approach to simultaneously manipulating heterochromatic regions and enhancing TE silencing.

We assayed age-dependent loss of silencing of TEs in each of these various transgenic (*Sir2*, *Su(var)3-9*, and *Dicer-2*) or hypomorphic (*Adar*) fly lines compared with their respective controls. The TEs that showed the greatest increase in expression with age in the control lines were overwhelmingly attenuated or decreased with age by these genetic interventions (Fig. 3A-D; Sir2 72%, Adar 84%, Su(var)3-9 92%, Dicer-2 68%). Interestingly, we again observed significant overlap in the TEs that increased with age among the different genetic backgrounds (Fig. S2B). Together, these data suggest that increased expression of *Sir2*, *Dicer-2*, or *Su(var)3-9*, as well as decreased gene dosage of *Adar*, all prevent the age-related loss of silencing of TEs observed in wild-type and control lines, suggesting a common link between heterochromatin factors and proper maintenance of TE silencing.

TE activation is known to disrupt gene expression and destabilize the genome. Increased DNA damage has been observed in aging fat body nuclei, as assayed by γH2Av immunostaining (16). Our finding of a repression of TE expression with increased *Dicer-2* expression led us to investigate whether genomic integrity was similarly compromised in animals with disrupted RNAi/heterochromatin pathways and increased TE activity. Similar to depletion of AGO2 in fat body cells (16), we observed an increase in double strand breaks in *Dicer-2* mutants as compared to controls in fat body nuclei (Fig. S4A), suggesting deleterious genomic consequences when proper TE silencing is disrupted.

Genetic interventions that affect heterochromatin structure also affect lifespan

In addition to their effects of delaying age-related heterochromatin gene and TE silencing loss, both increased *Sir2* expression and decreased *Adar* expression have been shown to extend lifespan in flies (21, 25), while *Dicer-2* and *Ago2* mutants are short lived (15, 24). Given our findings above, we examined whether increased expression of *Su(var)3-9* or *Dicer-2* also have a positive effect on lifespan. We found that two different *Su(var)3-9* transgenic lines, SV2 and SV3, which constitutively increase *Su(var)3-9* expression to 160-180% of the wild-type level in the absence of a driver (Fig. S4B), extend lifespan significantly when compared to their genetically matched control (Fig. 4A). Similarly, increasing *Dicer-2* expression in adult life in either all tissues (Fig. 4B) or selectively in adult neurons (Fig. 4C) using the conditional inducible GeneSwitch system was able to extend lifespan, with flies fed the gene-activating drug RU486 outliving a genetically identical control cohort fed only the diluent control EtOH. Together with previously published data (21, 25), these observed lifespan extensions suggest a coordinated response integrating maintenance of heterochromatin structure, TE suppression, and longevity.

Pharmacological suppression of transposable element activity delays age-related transposition and extends lifespan

In order to test whether the age-related effects we observed were due to increased TE activity, we sought to directly modulate transposition pharmacologically. We used the reverse transcriptase inhibitor lamivudine (3TC) to directly block TE transposition. Using the *gypsy*-TRAP reporter system in the fat body, we observed that

flies treated with 3TC had a significant delay in age-related transposition compared with controls (Fig. 5A). This effect was observed whether the drug was administered constantly throughout life (Fig. 5A), or started during mid-adulthood at 20 days of age (Fig. S4C).

As our evidence suggests that maintenance of TE silencing is important for ensuring longevity, we tested whether attenuating TE activity directly could extend lifespan. In the *Dicer-2*^{L811fsX} mutant background, which exhibits high levels of TE activity due to defective RNAi silencing (32), we found that administration of 3TC was able to extend lifespan significantly compared to controls (Fig. 5B). Together, these data suggest that heterochromatin-altering pathways that extend lifespan may do so at least in part by suppressing TE activity to ensure healthy longevity, and confirm that suppression of TE activation is likely to be an attractive target mechanism for pharmacological intervention.

Discussion

Aging is characterized by a breakdown in homeostasis throughout multiple cellular and organismal systems. Genomic and epigenomic stability is vital for somatic cells to maintain proper gene expression and silencing. A number of studies have shown changes in chromatin structure during organismal aging in various model systems, especially in heterochromatin (1, 2). Although heterochromatin regions contain some genes, a likely reason for the central importance of maintaining repressive heterochromatin is that heterochromatin is replete with TEs, and unchecked TE activation has particularly damaging ramifications for genomic stability. Indeed, the

expression and mobilization of TEs is disruptive enough that cells have evolved RNAi pathways responsible for silencing and preventing TE activation. These cellular defenses against TE activation include siRNA-directed induction and maintenance of repressive heterochromatin to prevent TE expression, and degradation of target transcripts. siRNA pathway mutants exhibit both an activation of TEs and a shortening of lifespan (15, 24). Here we provide evidence that many heterochromatin genes and TEs do in fact increase expression with age in two different tissues. Increased transposition has been previously reported in the aging fly brain (15). Similarly, we also observe an increase in somatic TE transposition with age in the fat body. DNA double strand breaks increase with age in the fat body (16), and we show this DNA damage also accumulates at a younger age in *Dicer-2* mutant flies with high levels of TEs. The age-related increase in TE transcript levels as well as TE mobilization were suppressed by dietary restriction. Together, these data suggest that an increase in TE expression and transposition is an important hallmark of somatic aging, and preventing this increased TE activity is likely to preserve genomic stability and cellular homeostasis.

We demonstrate in a whole organism model that aging is associated with a loss of silencing of both TEs and genes natively located in heterochromatin regions, with a resultant increase in TE mobilization and an associated increase in DNA damage. It is highly likely that the pathways and processes of heterochromatin maintenance, RNAi, and gene and TE silencing represent potential homeostatic failure points that lead to the phenotypes of aging. Here we have also presented genetic evidence that these pathways are linked to aging, as manipulating the expression of several genes involved in maintaining heterochromatin structure, TE silencing, and RNAi efficacy directly

suppresses age-related TE expression as well as extends lifespan. Furthermore, we show that dietary restriction, an environmental intervention that extends lifespan robustly and delays most age-related phenotypes, also opposes the loss of silencing we observe during normal aging and delays age-dependent transposition, further lending support to the idea that these processes are associated with and may be relevant to aging. We also demonstrate that inhibition of retrotransposon activity is strongly correlated with extension of lifespan. Our data support the retrotransposon theory of aging model, where proper maintenance of silent heterochromatin and prevention of TE activation are key to guaranteeing genomic stability and homeostasis during aging and ensuring healthy longevity. Furthermore, our demonstration that dietary (DR), genetic, and pharmacological interventions that reduce the age-related increases in TE activity can also extend lifespan suggests new and novel pathways for the development of interventions designed to extend healthy lifespan.

Materials and Methods

Fly stocks and culturing

Drosophila stocks were maintained at 25°C with a 12 hour light/dark cycle at 60% relative humidity. Lab stocks of Canton S and w^{1118} were used for wild-type experiments. The *Adar* hypomorph ($dAdar^{hyp}$) as well as its genetically matched control ($dAdar^{loxP}$) were a gift from Robert Reenan and described in (21). The *LacI-Su(var)3-9* lines (SV2 and SV3) were a gift from Kristen Johansen and described in (34). SV2 and SV3 were backcrossed for 20 generations to w^{1118}. Flies expressing *Dicer-2* were generated by crossing Bloomington stock #25750 with tubulin-GeneSwitch or elav-

GeneSwitch, and homozygosing both chromosomes. Experimental flies were generated by crossing these tub-GS/UAS-*Dcr-2* or elav-GS/UAS-*Dcr-2* stocks with w^{1118} such that both the GeneSwitch driver and UAS-*Dcr-2* were heterozygous. *Sir2* expressing EP2300 flies were obtained from Bloomington stock center (#24859). The fat body *gypsy-TRAP* reporter line was generated from the *gypsy-TRAP* line generously provided to us by Josh Dubnau (15), *r4-GAL4* (Bloomington #33832), and *UAS-GFP* (Bloomington #1522). Briefly, the fat body specific *r4-GAL4* driver line was first recombined with a *UAS-GFP* line generating a new stable line that expresses GFP exclusively in the fat body and balanced on the second chromosome with *If*/*CyO*. This line was then crossed with the *gypsy-TRAP* reporter line, which contains a *Tubulin-ovo-GAL80* that suppresses GFP expression until TE transposition disrupts reporter function (15). For the line shown in Fig. 2, *If* F_1 flies from this cross were taken (w^{1118}; *If*/+; r4-GAL4, UAS-GFP/tub-ovo-GAL80). *CyO* F_1 flies were taken to generate *CyO* line #1 in Fig. S2 (w^{1118}; *CyO*/+; r4-GAL4, UAS-GFP/tub-ovo-GAL80). For *CyO* line #2 in Fig. S2, *piwi2*/*CyO* flies were crossed to the *gypsy*-TRAP reporter line, and *CyO* F_1 flies were taken (w^{1118}; *CyO*/+; r4-GAL4, UAS-GFP/tub-ovo-GAL80). The mutant *ovo gypsy*-TRAP line was constructed as described above, except using a *gypsy*-TRAP line with mutated *ovo* sites that were no longer hotspots for *gypsy* transposition.

Gypsy-TRAP assay

Flies from the *gypsy*-TRAP line were aged for 5 days and initially examined for the number of GFP-positive cells, then examined every 7 days until age 49 days using a fluorescent dissecting microscope. Flies were separated into three categories based on

visual inspection according to how many fat body cells were GFP positive: low (<20 cells), medium (20-100), and high (>100). The same cohort of flies was examined longitudinally throughout each of the experiments.

RNA-seq

Total RNA from the indicated genotypes was collected from ~100 fly heads or ~50 dissected fat bodies using the miRvana RNA prep kit (Ambion). 100ng of total RNA was then used as input for stranded RNA-seq library construction using the Ovation Universal RNA-seq kit (Nugen) with Drosophila rRNA depletion module. Libraries were sequenced on an Illumina HiSeq 2500 in 1x50bp mode. All samples were prepared using at least three individually isolated biological replicates for each condition.

Bioinformatics

Standard RNA-seq analysis was performed using Tophat 2.0 for sequence alignment, followed by gene based counting using the easyRNASeq R package from Bioconductor. Finally, TMM normalization and fold change calculation was performed using edgeR, with a GLM correcting for batch effects as described in the edgeR vignette. To account for repetitive sequencing reads and perform TE quantification, RepEnrich (33) was used to generate a count table of reads for each TE and, edgeR was used to normalize, calculate fold changes, and perform differential expression testing. The Repbase annotation was used as input for RepEnrich.

Heatmaps were generated by ranking the TEs or genes in order of $\log_2$ fold change of expression with age (old/young) for the reference (control/high cal) condition,

then plotting log_2 fold change with age for all conditions. In order to show consistent trends in our data and include genes/TEs that behaved consistently but did not meet the threshold of statistical significance, we displayed genes or TEs where log_2 fold change was greater than 0.322 (1.25x fold; TEs) or 0.263 (1.2x fold; heterochromatin genes). Plots were generated using the heatmap.2 function in the gplots2 R package.

Lifespan assays

Food composition for lifespan experiments was as follows: for normal/high calorie food, 15% Dextrose (w/v), 15% autolysed yeast, 2% cornmeal, and 2% agar; for DR/low calorie food, 5% Dextrose, 5% autolysed yeast, 2% cornmeal, and 2% agar. For experiments using GeneSwitch drivers, either RU486 dissolved in EtOH at the indicated concentration or EtOH alone (for controls) was added to the food at 20mL/L. Upon eclosion, flies were sorted under CO2 anesthesia and placed in food vials at a density of 25 males and 25 females per vial, with a total of at least 10 vials (n=250) for each condition. Flies were passed to fresh food every other day, and dead flies scored and counted. Lifespan statistics and Log-Rank P values were determined using the web-based OASIS tool (35).

Double strand break assay

Sectioned female flies were stained with α-γH2Av antibody and imaged with a Zeiss Axiovision epifluorescence scope at a constant exposure time of 500ms (AlexaFluor 568) or 100ms (DAPI). Images from slices containing fat body nuclei were used to quantify fluorescence levels using the technique described in (16) (n=50 nuclei).

Acknowledgements

We thank Will Lightfoot and Gillian Horwitz for technical assistance and Christoph Schorl and the Brown University Genomics Core Facility for sequencing. We also thank Joshua Dubnau, Robert Reenan, Kristen Johansen, and the Bloomington *Drosophila* Stock center for fly stocks and Robert Reenan and John Sedivy for helpful discussions. This work was supported in part by National Institute on Aging (NIA) T32 Training Grant AG41688 and Fellowship Award AG47736 to B.C.J.; NIA Grants AG16667 and AG24353, a Glenn/American Federation for Aging Research (AFAR) Breakthroughs in Gerontology award, NIH Program Project Grant AG51449, and Grant P30AI042853 from the Providence/Boston Center for AIDS Research to S.L.H.; Ellison/AFAR Postdoctoral Fellowship awards to J.G.W. and N.J.; a Nathan Shock Center Pilot Grant to J.G.W.; and NIH Grants AG28753 and AG50582 to N.N.

References

1. Wood JG, Helfand SL (2013) Chromatin structure and transposable elements in organismal aging. *Front Genet* 4(December):1–12.

2. Feser J, Tyler J (2011) Chromatin structure as a mediator of aging. *FEBS Lett* 585(13):2041–8.

3. Dang W, et al. (2009) Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* 459(7248):802–7.

4. Benayoun BA, Pollina EA, Brunet A (2015) Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nat Rev Mol Cell Biol* 16(10):593–610.
5. Wood JG, et al. (2010) Chromatin remodeling in the aging genome of Drosophila. *Aging Cell* 9(6):971–978.

6. Jiang N, et al. (2013) Dietary and genetic effects on age-related loss of gene silencing reveal epigenetic plasticity of chromatin repression during aging. *Aging (Albany NY)* 5(11):813–24.

7. De Cecco M, et al. (2013) Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell* 12(2):247–56.

8. De Cecco M, et al. (2013) Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging (Albany NY)* 5(12):867–883.

9. Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8(4):272–85.

10. Wang J, et al. (2011) Inhibition of activated pericentromeric SINE/Alu repeat transcription in senescent human adult stem cells reinstates self-renewal. *Cell Cycle* 10(17):3016–3030.

11. Ambati J, Fowler BJ (2012) Mechanisms of age-related macular degeneration. *Neuron* 75(1):26–39.

12. Van Meter M, et al. (2014) SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat Commun* 5:5011.

13. Dennis S, Sheth U, Feldman JL, English KA, Priess JR (2012) C. elegans germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS Pathog* 8(3):e1002591.

14. Maxwell PH, Burhans WC, Curcio MJ (2011) Retrotransposition is associated with genome instability during chronological aging. *Proc Natl Acad Sci U S A* 108(51):20376–81.

15. Li W, et al. (2013) Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat Neurosci* 16(1):529–531.

16. Chen H, Zheng X, Xiao D, Zheng Y (2016) Age-associated de-repression of retrotransposons in the *Drosophila* fat body, its potential cause and consequence. *Aging Cell*:1–11.

17. Holoch D, Moazed D (2015) RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* 16(2):71–84.

18. Schotta G, et al. (2002) Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J* 21(5):1121–1131.

19. Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. *Nat Rev Genet* 10(2):94–108.

20. Castel SE, Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* 14(2):100–12.

21. Savva YA, et al. (2013) RNA editing regulates transposon-mediated heterochromatic gene silencing. *Nat Commun* 4:1–11.

22. Rosenberg MI, Parkhurst SM (2002) Drosophila Sir2 is required for heterochromatic silencing and by euchromatic Hairy/E(Spl) bHLH repressors in segmentation and sex determination. *Cell* 109(4):447–58.

23. Oberdoerffer P, et al. (2008) SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. *Cell* 135(5):907–18.

24. Lim DH, et al. (2011) The endogenous siRNA pathway in Drosophila impacts stress resistance and lifespan by regulating metabolic homeostasis. *FEBS Lett* 585(19):3079–3085.

25. Whitaker R, et al. (2013) Increased expression of Drosophila Sir2 extends life span in a dose-dependent manner. *Aging (Albany NY)* 5(9):682–91.

26. Viswanathan M, Guarente L (2011) Regulation of Caenorhabditis elegans lifespan by sir-2.1 transgenes. *Nature* 477(7365):E1–E2.

27. Burnett C, et al. (2011) Absence of effects of Sir2 over-expression on lifespan in C. elegans and Drosophila. *Nature* 477(7365):482–485.

28. Hoffmann J, Romey R, Fink C, Yong L, Roeder T (2013) Overexpression of Sir2 in the adult fat body is sufficient to extend lifespan of male and female Drosophila. *Aging (Albany NY)* 5(4):315–27.

29. Riddle NC, et al. (2011) Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin. *Genome Res* 21(2):147–63.

30. Charlesworth B, Langley CH (1989) The population genetics of Drosophila transposable elements. *Annu Rev Genet* 23:251–287.

31. Dej KJ, Gerasimova T, Corces VG, Boeke JD (1998) A hotspot for the Drosophila gypsy retroelement in the ovo locus. *Nucleic Acids Res* 26(17):4019–4025.

32. Ghildiyal M, et al. (2008) Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* 320(5879):1077–81.

33. Criscione SW, Zhang Y, Thompson W, Sedivy JM, Neretti N (2014) Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics* 15(1):583.

34. Boeke J, et al. (2010) Phosphorylation of SU(VAR)3-9 by the chromosomal kinase JIL-1. *PLoS One* 5(4):e10042.

35. Yang JS, et al. (2011) OASIS: Online application for the survival analysis of lifespan assays performed in aging research. *PLoS One* 6(8). doi:10.1371/journal.pone.0023525.

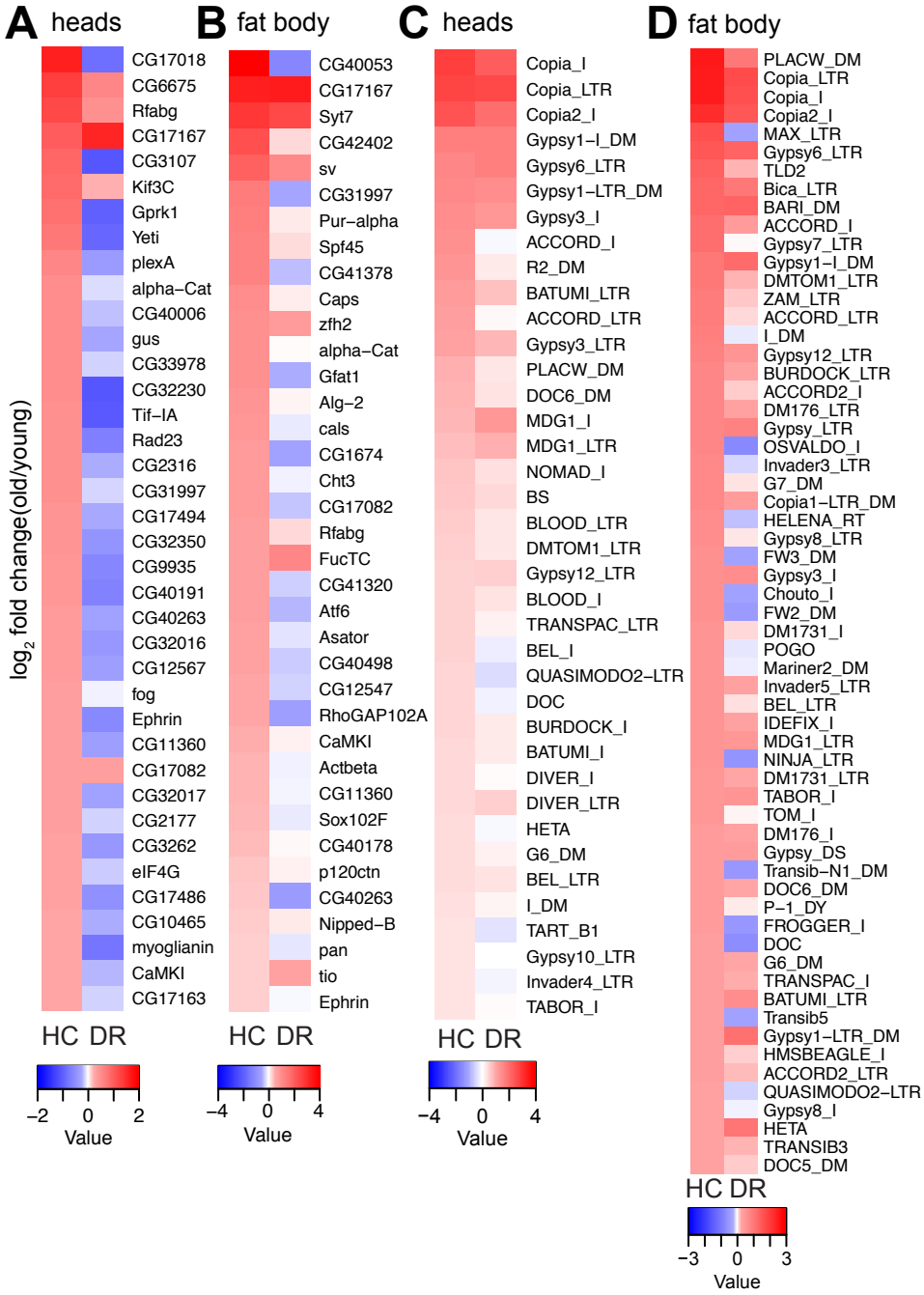


Figure 1. Dietary restriction attenuates age-related loss of silencing in heterochromatin genes and transposable elements (TEs).

(**A**, **B**) Heatmap shows log2 fold change (old/young) values for heterochromatin genes in both (**A**) heads and (**B**) fat bodies. Left column (HC) indicates $\log_2$ fold change under high calorie conditions, right column (DR) under dietary restriction. Heatmap shows genes with $\log_2$ fold changes greater than 0.263 (1.2x fold change) under high calorie conditions that are located within annotated heterochromatin regions (29). 95% of shown genes in heads and 89% in fat bodies are suppressed by DR.

(**C**, **D**) Heatmap shows $\log_2$ fold change (old/young) of TE expression for both high calorie (HC; left columns) and dietary restriction (DR; right columns) diet in both (**C**) heads and (**D**) fat bodies. Heatmap shows TEs with $\log_2$ fold changes greater than 0.322 (1.25x fold change) under high calorie conditions. 87% of shown TEs were suppressed by DR in both heads and fat bodies.

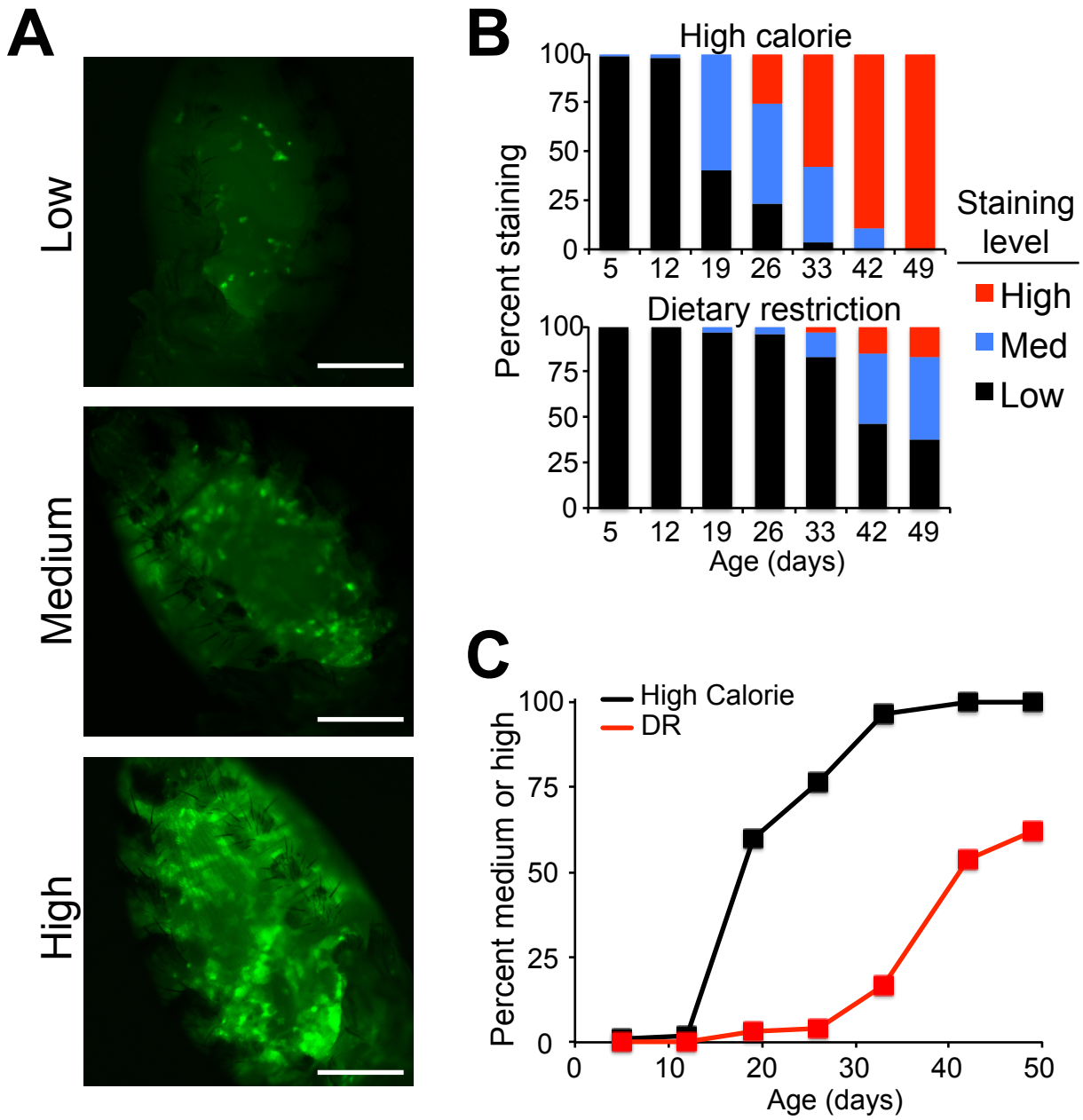


Figure 2. Dietary restriction delays age-related increase in TE transposition.
(**A**) *Gypsy-TRAP* flies (*w*^{1118}*; lf/+; r4-GAL4, UAS-GFP/tub-ovo-GAL80*) were categorized as low (<20 GFP-positive cells), medium (20-100), or high (>100) based on number of GFP-positive cells observed in fat body (representative images shown). The demarcation between low and medium categories is clear; most flies have either fewer than 10 cells staining, or more than 50. Scale bar represents 200μm.
(**B**) Percentage of observed flies falling into each category at indicated timepoint for high calorie (top) and dietary restriction (bottom) diets (*n*=~120 per cohort).
(**C**) Transposition (defined here as medium or high category GFP staining) increases with age and is delayed by calorie restriction. Percentage of flies falling into either medium or high categories is displayed with time for both high calorie and dietary restriction diets. Log-Rank $P<10^{-10}$.

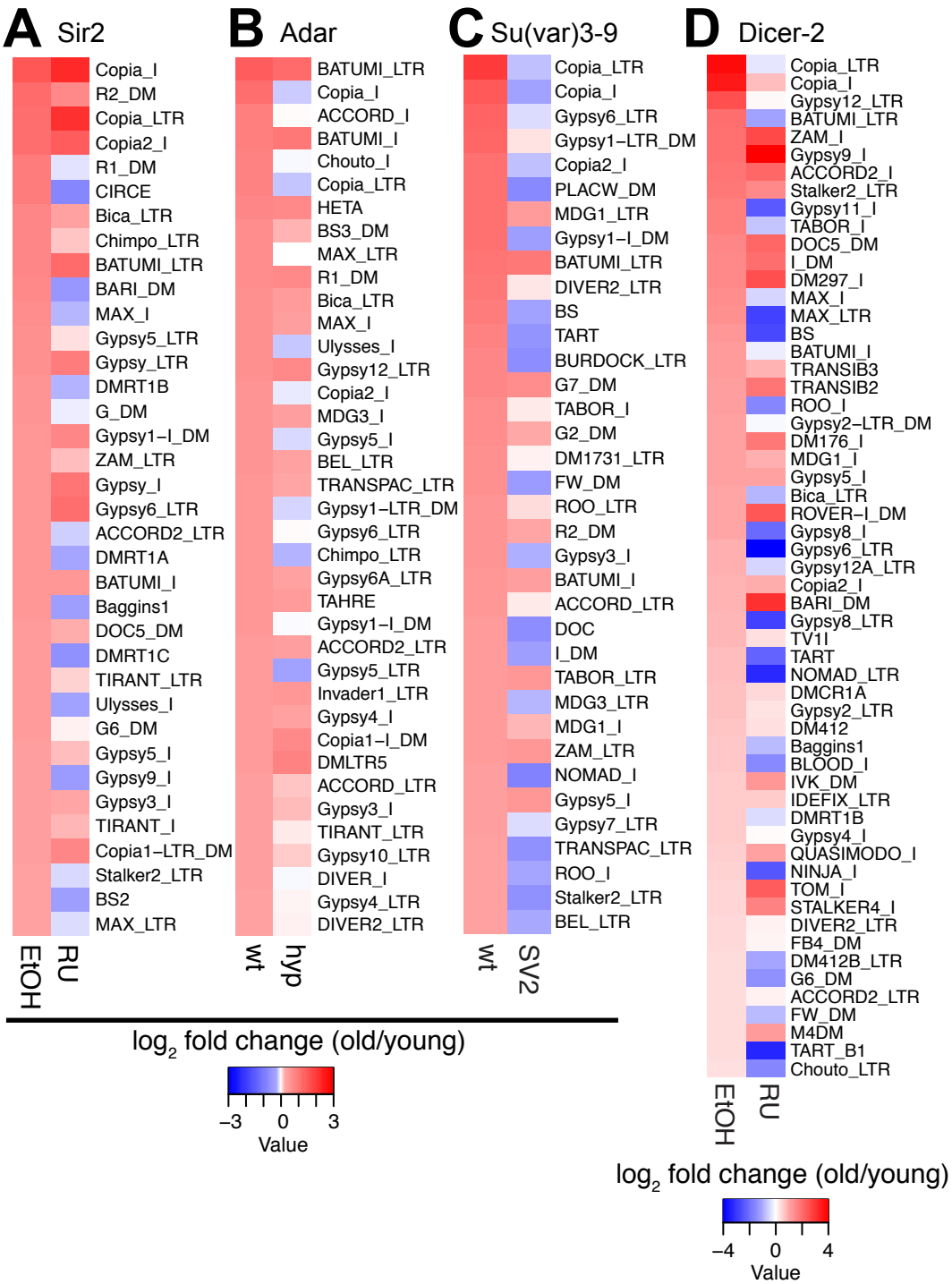


Figure 3. Manipulation of genes in heterochromatin pathways affects TE expression.
Heatmaps of log2 fold change (old/young) of TEs for the following genotypes and their respective controls: (**A**) *elav-GeneSwitch* > *UAS-EP2300* (Sir2 overexpression); (**B**) Adar hypomorph vs. wild-type; (**C**) *Su(var)3-9* overexpression (SV2 and SV3 lines overexpress *Su(var)3-9* from a basally expressed UAS promoter 160-180% above control; Fig. S4B); and (**D**) *tubulin-GeneSwitch* > *UAS-Dicer-2* (*Dicer-2* overexpression). In each case the control condition is shown in the left column, with the transgenic, hypomorph, or RU486-induced condition in the right column(s). The percentage of shown TEs that are suppressed by the indicated genotype are: (**A**) 72%, (**B**) 84%, (**C**) 92%, and (**D**) 68%. All TEs with $\log_2$ fold change (old/young) greater than 0.322 (1.25x fold change) in the control condition are shown.

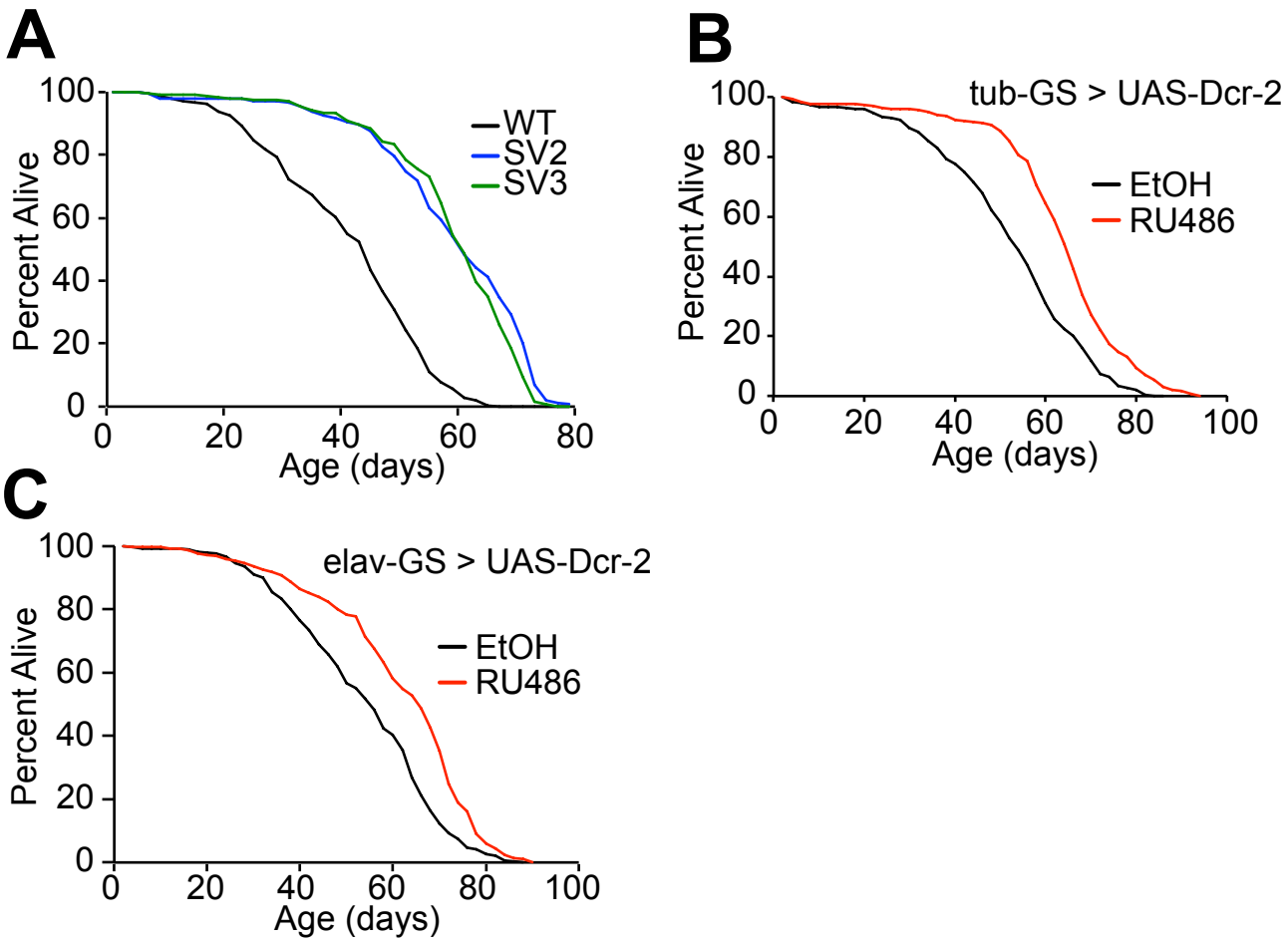


Figure 4. Increasing expression of Su(var)3-9 or Dicer-2 increases lifespan.
(**A**) SV2 and SV3 lines (overexpressing Su(var)3-9) are long lived compared to controls. Survivorship of adult female flies is shown (*n*=250). WT mean lifespan=40.6, median=44. SV2 mean=58.5, median=60, 36% median extension, Log-Rank $P<10^{-10}$. SV3 mean=58.3, median=60, 36% median extension, $P<10^{-10}$.
(**B**) Whole adult animal *Dicer-2* overexpressing flies are long lived compared to controls. Survivorship of *tubulin-GeneSwitch* > *UAS-Dicer-2* adult females (*n*=250) is shown, either with EtOH (uninduced) or 200μM RU486 (induced expression). EtOH mean=52.3, median=54. RU486 mean=63.5, median=66, 22% median extension, $P<10^{-10}$.
(**C**) Flies expressing *Dicer-2* in adult neurons are long lived compared to controls. Survivorship of *elav-GeneSwitch* > *UAS-Dicer-2* adult females is shown, for EtOH (uninduced) or 500μM RU486 (induced) (n=250). EtOH mean=53.8, median=56. RU486 mean=61.7, median=66, 18% median extension, $P<10^{-10}$.
All lifespan experiments were repeated with similar results; representative experiments shown.

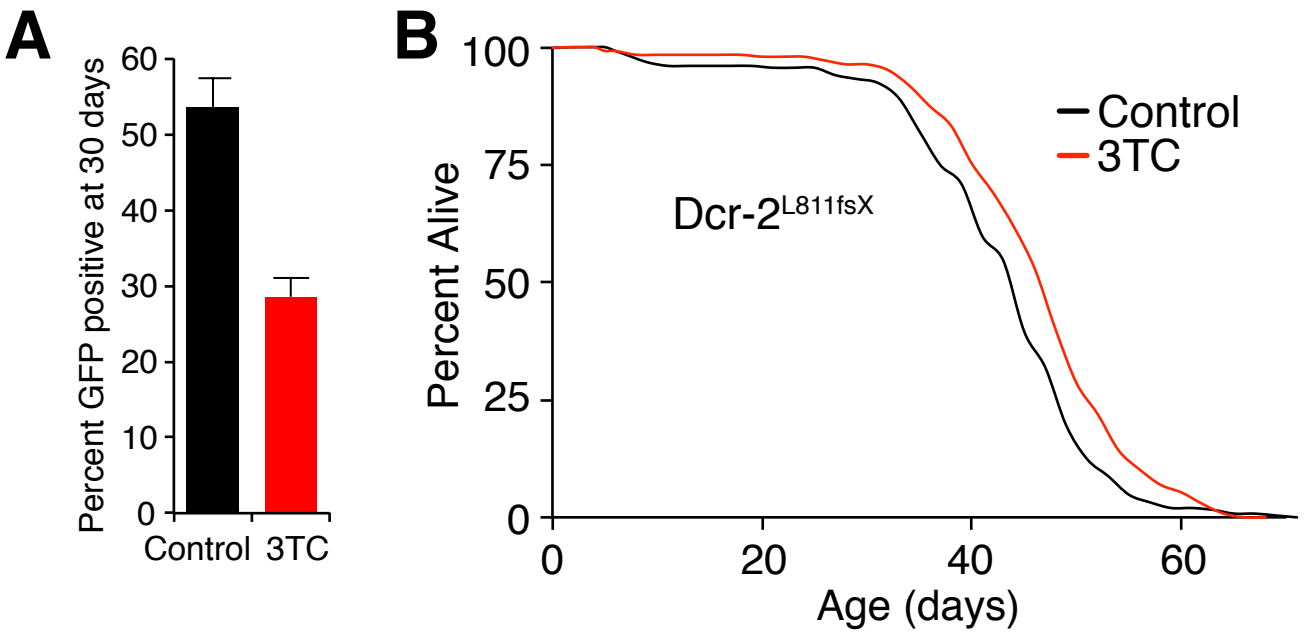


Figure 5. Administration of lamivudine (3TC) suppresses retrotranposition and extends *Dicer-2* mutant lifespan.

(**A**) Percentage of *gypsy-TRAP* flies with more than 30 fat body cells positive for GFP expression at 30 days of age for both untreated (control) and 10μM 3TC treated fly cohorts. Fisher's exact test $P<0.0001$.

(**B**) 10μM 3TC treated flies are long lived compared to untreated controls in the *Dicer-2*^{L811fsX} genetic background. Survivorship of adult female flies is shown ($n=250$). Control mean=43.0, median=45. 3TC mean=46.4, median=48, 6.6% median extension, Log-Rank $P<0.0001$. Lifespan was repeated with similar results.

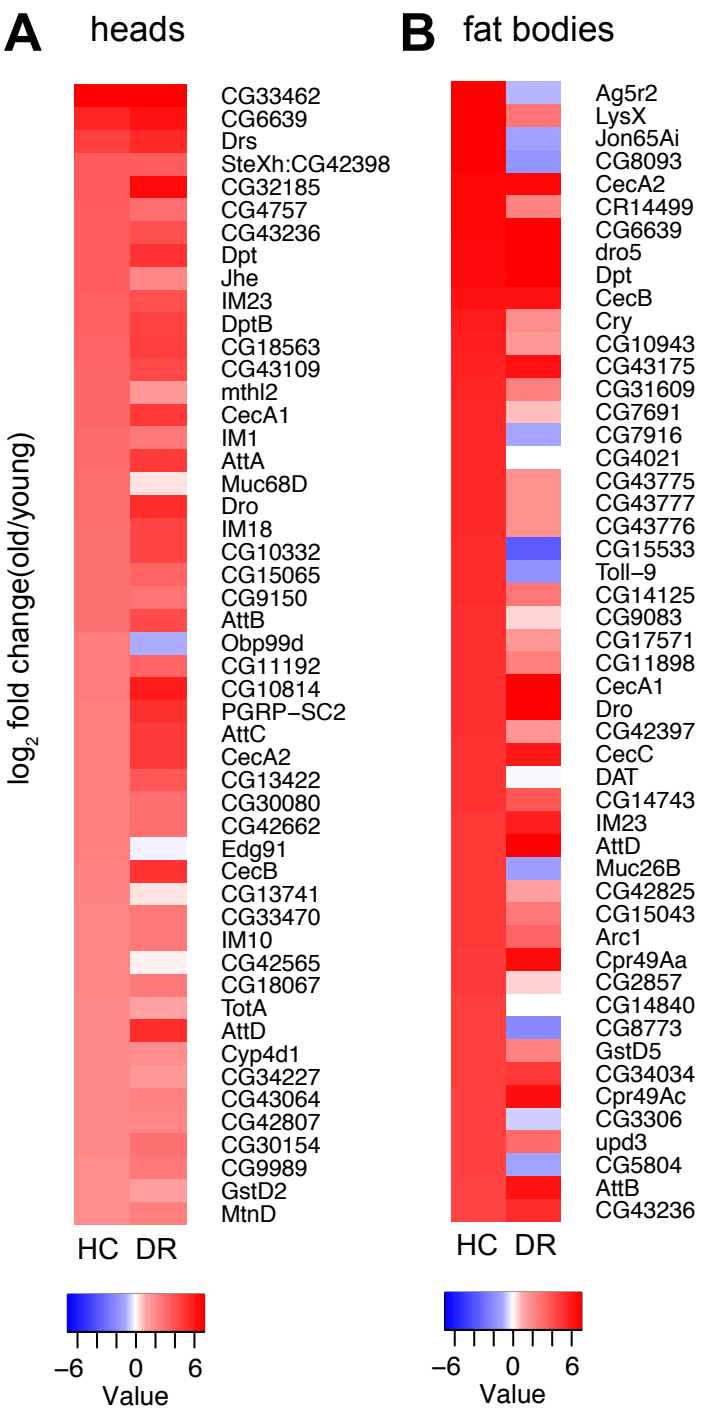


Figure S1. Effect of dietary restriction on age-related loss of silencing in euchromatin genes.

(**A**, **B**) Heatmap shows $\log_2$ fold change (old/young) of top 50 euchromatin genes upregulated with age in heads (**A**) and fat bodies (**B**). Left column (HC) is high calorie diet, right column (DR) is dietary restriction. Genes located in euchromatin regions do not show the same consistent DR suppression of age-related increase in expression observed in heterochromatin genes (**Fig. 1A**, **B**).

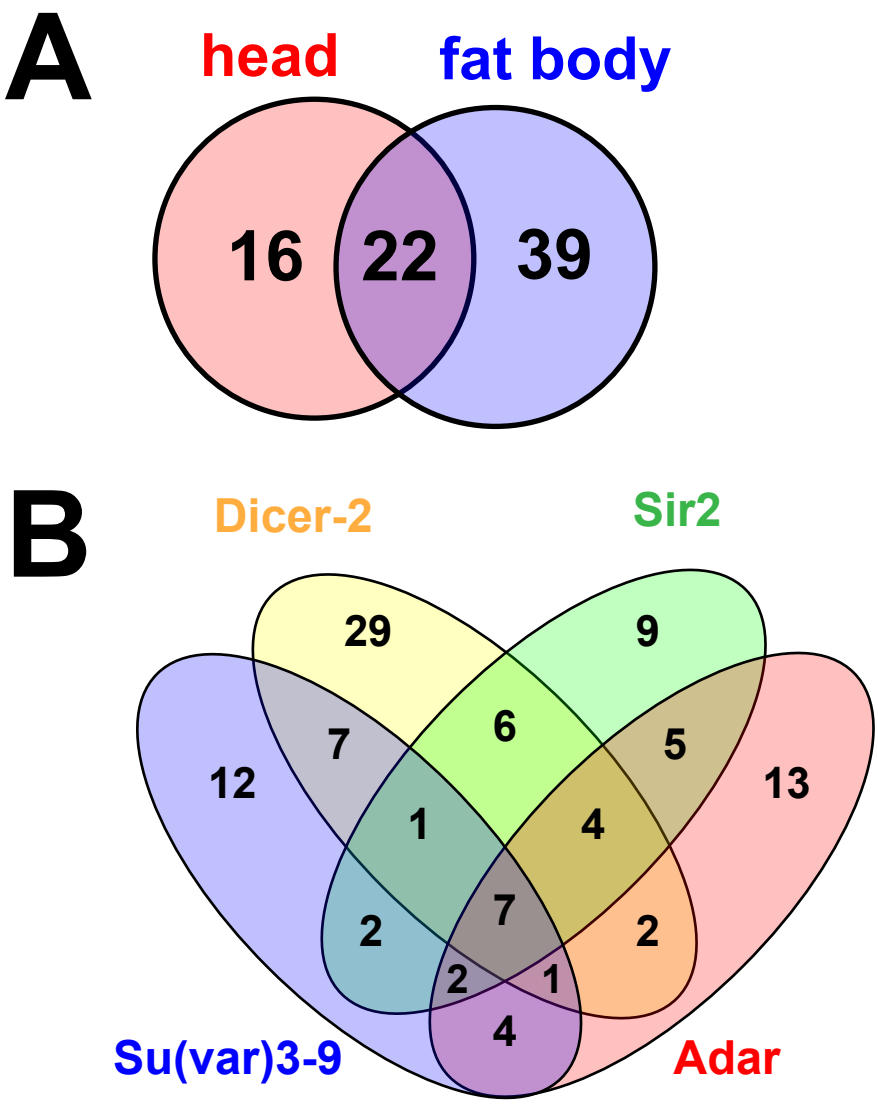


Figure S2. Spectrum of upregulated TEs shows significant overlap among different tissues and genotypes.

(**A**) Venn diagram indicating overlap of TEs that increase > 0.322 $\log_2$ fold change in head and fat body libraries (*P*=0.0395, Fisher's exact test). The LTR retrotransposons *copia*, *gypsy6*, and *accord* were among the most highly upregulated TEs with age in both backgrounds. *Copia* is the most highly expressed TE in both our head and fat body datasets, and additionally copia is known to be active in *D. melanogaster*.

(**B**) Four way Venn diagram showing overlap of TEs displayed in four genotypes of **Fig. 3A-D**. TEs shown are those that increase with age in the control or uninduced condition. Five TEs (seven sequences including *Copia_LTR, Copia_I, Gypsy6_LTR, Copia2_I, BATUMI_LTR, BATUMI_I, Gypsy5_I*; internal and LTR portions of sequences are annotated separately) appear in all four genetic backgrounds. The fact that there are also differences between the genotypes indicates that effects on TE expression are likely dependent on genetic background and strain specific genomic location of TEs.

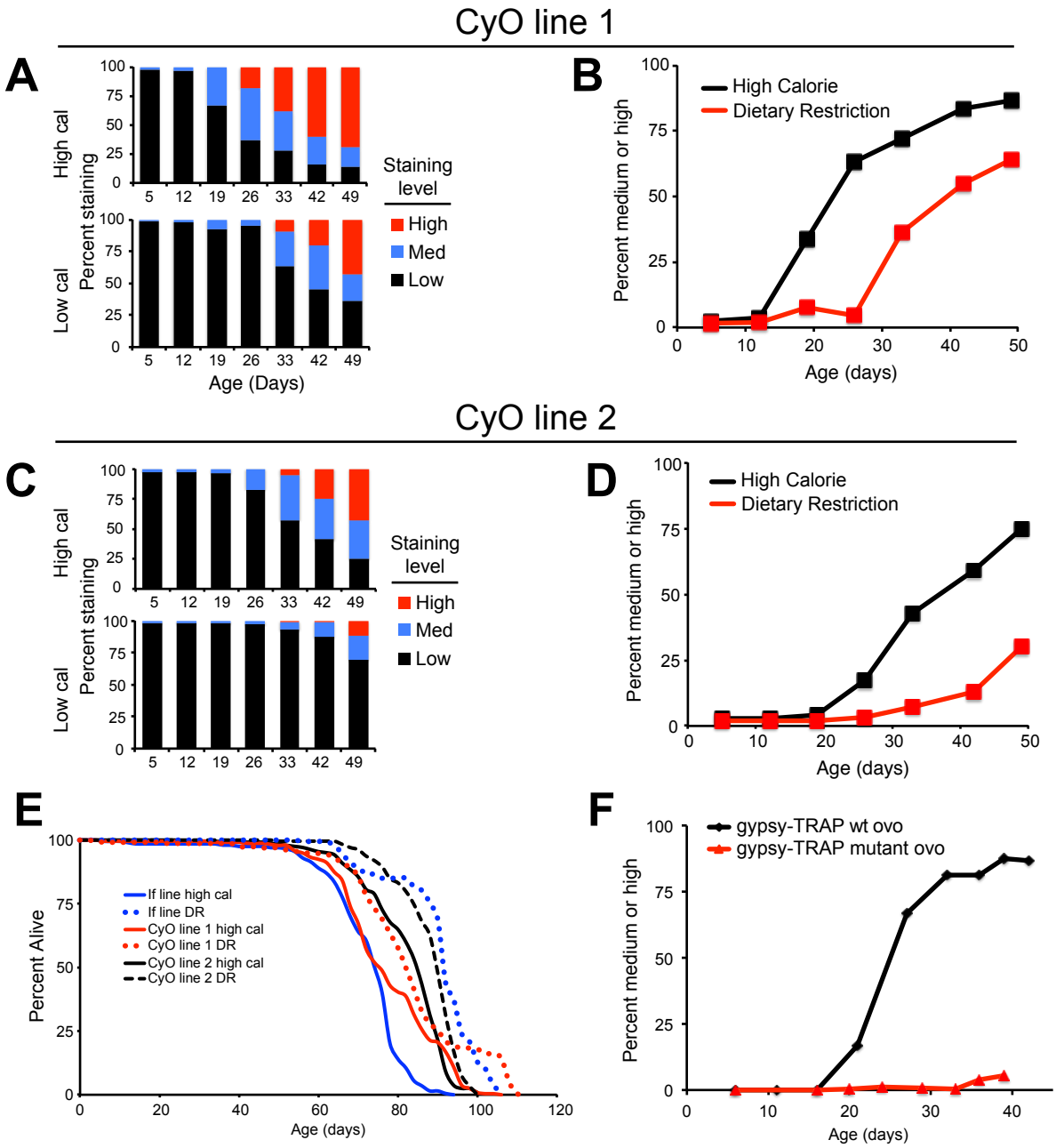


Figure S3. Dietary restriction delays age-related increase in TE transposition in alternate genetic backgrounds.
(**A**, **C**) Percentage of *gypsy-TRAP* flies in two independently derived *CyO* lines (w^{1118}*; CyO/+; r4-GAL4, UAS-GFP/tub-ovo-GAL80*; see supplemental methods for further details) falling into each GFP staining category at indicated time point, for both high calorie (top) and dietary restriction (bottom) diets. Data is presented analogously to **Fig. 3B** (*n*=120 HC, 80 DR).
(**B**, **D**) Percentage of flies falling into medium or high staining categories in these two gypsy-TRAP lines with time, for both high calorie and dietary restriction diets, analogous to **Fig. 3C** (*n*=123 HC, 129 DR). Log-Rank $P<10^{-10}$ for both lines.
(**E**) Lifespan of individual *gypsy-TRAP* lines correlates with onset of increased transposition. Survivorship curves for the three independent *gypsy-TRAP* genetic backgrounds shown in (**A-D**) as well as **Fig 2B-C**. Adult females (*n*=250) are shown for each line on high calorie (solid lines) and dietary restriction (dotted) food types. *If* line: high cal mean=72.3, median=76; DR mean=90.9, median=92. *CyO* line 1: high cal mean=77.5, median=76; DR mean=83.1, median=82. *CyO* line 2: high cal mean=82.4, median=86; DR mean=88.6, median=90. All DR lifespans are significantly different compared to their respective high cal cohort at $P<10^{-6}$ (Log-Rank). All high cal lifespans are also significantly different from each other ($P<10^{-6}$ for *If/CyO1* and *If/CyO2*; $P=0.03$ for *CyO1/CyO2*).
(**F**) Observed increases in GFP expression are not due to non-specific age-related changes in *GAL80, GAL4*, or *GFP* expression. Percentage of fat body cells staining positive for GFP with age in the normal *gypsy-TRAP* line (black) as well as a *gypsy-TRAP* line containing a mutated *ovo* locus that is not sensitive to *gypsy* transposon integration (red). Note the mutant *gypsy-TRAP* line shows no increase in GFP expression with age, indicating GFP expression is reporting increased transposition rather than general age-related changes.

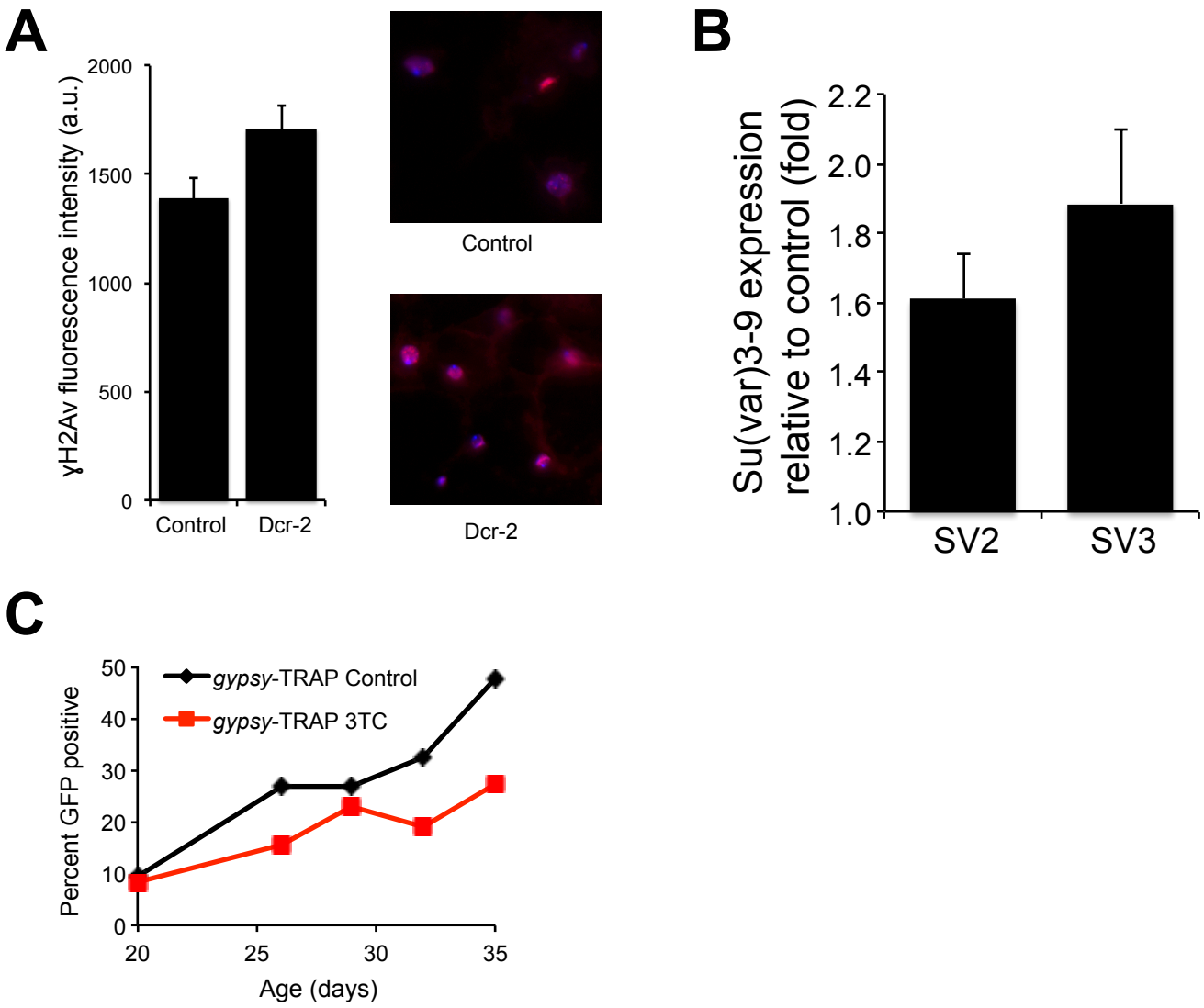


Figure S4. Additional data figures.

(**A**) *Dicer-2* mutants have increased double strand breaks in nuclei compared with controls. Relative nuclear γ-H2A.v fluorescence intensity in fat body cells from w^{1118} controls or $Dicer\text{-}2^{L811fsX}$ mutants. *Dicer-2* mean fluorescence is 23% higher than control (n=50, t-test P=0.027). Representative images from each genotype are shown.

(**B**) Expression level of *Su(var)3-9* relative to control for SV2 and SV3 lines, assayed by qPCR using *GAPDH1* as a reference. Error bars represent standard deviation.

(**C**) 3TC treatment delays age-related increase in transposition. Time course of GFP expression in *gypsy-TRAP* flies either untreated (control) or treated with 10μM 3TC, starting at day 20 of adulthood. Percentage of GFP positive (>100 cells) fat bodies for each condition is shown.

Chapter 4

General Discussion

Portions of this work were adapted from my own contributions to *Kreiling & ***Jones, et al.*** (2017) “Contribution of Retrotransposable Elements to Aging.” *Human Retrotransposons in Health and Disease* (see Appendix B).

The discovery of TEs as mobile genetic elements by Barbara McClintock in the 1940s upended conventional thought of the genome as static. Even after their discovery, it was long assumed that TEs were ancient relics, or “junk DNA,” with little capacity to influence host genomes or cellular biology. However, with the advent of whole genome sequencing technologies, TEs were later realized to contribute significant portions of the genome and function as significant sources of genomic dynamicity contributing to genomic structure, evolution, and damage. Later, with the discovery of RNA interference in the 1990s by Andrew Fire and Craig Mello, these two transformative fields soon became intertwined when others realized RNAi pathways and heterochromatin actively suppressed TEs. As the role and impact of TEs in cellular processes and disease becomes clearer, a better understanding of TE suppression pathways becomes increasingly important.

Beyond mutagenesis: TEs impact coding gene function via non-canonical means

While much of the early research into the consequences of TE reactivation focused on their ability to mobilize and insert into normal coding genes thereby disrupting gene function (insertional mutagenesis), more studies have begun to show that TEs are novel regulators of gene function. Chapters 2 and 3 of this thesis focus on TE reactivation in the context of the piRNA pathway and heterochromatin loss during aging, respectively. However, while both of these chapters demonstrate TE reactivation and correlated physiological consequences, it remains unknown what direct role TEs have in the context of aging or metabolism. Finally, it is unclear if the altered

transcriptomes observed in somatic tissues with TE reactivation are due to canonical TE jumping or alternative mechanisms.

One alternative means by which TEs can influence coding gene function beyond insertional mutagenesis is via altered gene transcription due to a nearby TE insertion. Multiple papers have shown that in fly ovarian cell culture, novel insertion of TEs can impact nearby coding gene function in a piRNA mutant background[1,2]. In one case, an insertion of the *mdg1* TE near the coding gene *CG2017* was shown to promote heterochromatin spreading, reduced Poll II occupancy, and gene silencing[1]. When the piRNA pathway is lost in these cells, bleeding of downstream transcription occurs up to 15kb at multiple *mdg1* insertion sites, suggesting that TE insertions near normal genes are able to dynamically regulate neighboring transcription[1]. This was also true for other TEs near coding genes and especially in euchromatic insertions[1,2]. This novel mechanism of gene regulation has been broadly implicated in a variety of species resulting in both beneficial and deleterious phenotypes[3].

Although TEs are well known to disrupt gene function by disrupting exon integrity, they can also impact coding gene function via intron insertion. *Alu* elements inserted in introns have been shown to affect splicing of downstream exons[4]. Multiple studies have also shown that TE insertion into introns can result in disease due to gene disruption or alternative splicing events[5]. L1 promoters are also known to produce chimeric transcripts by driving transcription into nearby genes[6,7]. Finally, TE intronic insertion also has the potential to benefit a host. For example, genomic sequencing of the pro-longevity gene *Indy* across different fly strains revealed an intronic insertion of

the *Hoppel* TE that, when heterozygous, correlated with altered *Indy* expression and extended lifespan[8].

While reactivation TEs is well characterized as a source of traditional genomic damage such as insertional mutagenesis and double strand breaks, these examples demonstrate that TE mobility can also indirectly impact coding gene function. In both Chapters 2 and 3, it is possible that the transcriptomic changes observed in piRNA mutants and aged flies are not only due to a general stress response of TE reactivation from heterochromatic regions. It is also possible that systemic dysregulation of euchromatic insertion sites contributes to altered gene expression and disruption of cellular function.

Transposon suppression pathways with alternative functions

RISC complex misloading

Argonautes are the main effector proteins that facilitate RNAi silencing and therefore are critical nexus points of smRNA pathways. The siRNA, piRNA, and miRNA pathways all employ these proteins to target complementary RNAs for silencing[9]. These argonaute proteins also all depend upon size-specific smRNAs for effective silencing. The miRNA pathway utilizes miRNAs (21-22nt long) that often imperfectly base pair with their targets upon association with an argonaute protein. This miRNA-RISC then prevents translation of the target mRNA by one of two methods: prolonged base pairing stalls or even completely blocks ribosome access or the miRNA-RISC cleaves the target mRNA[9]. While mammals have four argonautes (AGO1-4), only one, AGO2, is able to enact RNAi silencing via smRNA association[10,11]. Interestingly, human AGO2 is

permissive in accepting both miRNAs and siRNAs both of which range from 21-22nt[12,13]. This is in contrast with species like *Drosophila* where one argonaute, Ago1, is almost exclusively loaded with miRNAs and another, Ago2, is loaded mostly with siRNAs[14]. However, recent research in flies has shown that miRNAs and siRNAs compete for loading into Ago2, the argonaute responsible for TE silencing in flies[15]. In both flies and mammals, siRNAs are specifically 2'-*O*-methylated at their 3' termini[9]. Research in flies has recently shown that miRNAs are increasingly inappropriately methylated with age allowing them to be loaded into Ago2, instead of Ago1, thereby reducing siRNA occupancy[15]. While this study did not examine the effect of this RISC misloading on the ability of siRNAs to target and silence TEs, it remains a possibility, especially in humans where siRNAs and miRNAs share AGO2 for silencing, that this competition could indirectly impact TE silencing. Chapter 3 of this thesis observed the age-dependent reactivation of TEs with age in somatic tissues, but it is still unknown to what degree RNAi silencing changes with age if at all. Other studies have shown that expression of RNAi genes remain constant with age[15,16]. However, whether smRNAs that may be regulating TEs change at all remains largely unexplored. Therefore, it will be especially interesting for future studies to determine whether smRNA metabolism and the dynamics of the pool of smRNAs able to be loaded into argonaute proteins impacts TE silencing with age.

piRNA Targeting of non-TE sequences

Since piRNAs were first identified and mapped to the genome, it was known that they did not map exclusively to TE sequences[17]. This suggested that piRNAs were

derived from non-TE sources and that they may have the capacity to silence non-TE targets. During the last decade, new evidence has shown that the piRNA pathway has a non-TE role in silencing. During spermatogenesis in mice, two waves of piRNAs are expressed. The first wave, called pre-pachytene piRNAs are derived mostly from TE sequences while the second, called pachytene piRNAs, come primary from non-TE intergenic regions[18]. A possible role for pachytene piRNAs was identified when it was discovered that piRNA-RISCs contained both piRNAs and non-TE coding gene mRNAs[19]. These non-TE mRNAs are targeted by piRNA-RISCs complexed with CAF1 deadenylase that then promote removal of the target transcript's poly-A tail resulting in target mRNA destabilization and degradation[19]. Interestingly, piRNAs derived from TEs normally have near-perfect complementarity to their targets, while pachytene piRNAs that elicited targeting of mRNAs were able to do so via imperfect base-pairing similar to miRNA targeting[20]. Additionally, in flies, there have been individual reported cases where piRNAs are able to directly target coding genes for silencing. This was hypothesized when it was first observed that coding genes, especially 3'-UTRs, were significant sources of piRNAs[21,22]. In flies, piRNAs are derived from the 3'-UTR of the *traffic jam* (*tj*) that are then able to target the transcript of the *FasIII* gene for silencing[22]. In another case, the transcript of the *c-Fos* gene is destabilized by Piwi targeting its 3'-UTR for piRNA biogenesis[23]. In Chapter 2, we observed significant dysregulation of a number of coding genes in *piwi* mutants, many of which were associated with metabolic processes. While this chapter focused primarily on the role of a piRNA pathway in the fat body in suppressing TEs, we cannot rule out that there may be additional genomic targets of the fat body piRNA pathway. This is supported by our observation of multiple

genes generating genic piRNAs. As the field begins to better understand the non-canonical roles of the piRNA pathway beyond suppressing TEs, we can begin to probe how piRNAs may regulate vast gene regulatory networks.

Putative roles for the fat body piRNA pathway

The piRNA pathway is known to have a diversity of roles in both TE suppression as well as maintenance of stemness[24,25]. In Chapter 2, we present evidence for a piRNA pathway in a somatic tissue with the canonical function of TE suppression. However, it is unclear what specific role the piRNA pathway serves by being active in the fat body. For example, what drives the piRNA pathway to be active in the germline and the fat body, but not other somatic tissues? Is there a specific advantage to having a piRNA pathway active specifically in the fat body? What additional functions does the fat body piRNA pathway perform that is not covered by an active siRNA pathway (see Appendix A, Figure 2)? While we do not yet know the answers to these questions, I propose some intriguing hypotheses for possible future investigation.

Interestingly, many TEs in flies are thought to be retroviral-like[26–28] and therefore capable of transmission between cells, tissues, and even individuals. It is possible that TEs are secreted from the fat body into circulating hemolymph. This hemolymph serves as a nutritive source for a variety of cell types and tissues and is taken up by the ovary in large amounts during germ cell and early organismal development. In this manner, the fat body represents a potential reservoir of TEs that could potentially infect ovarian tissues via vertical transmission. In this sense, the fat body piRNA pathway could be similar to the piRNA pathway in the somatic ovarian follicle cells, which are directly

adjacent to ovarian tissues, in that both may serve as upstream defenses against vertical transmission of TEs.

Finally, It is also interesting to note that the fat body and somatic follicle cells of the ovary share a common developmental lineage[29,30], which may explain the origin of a piRNA pathway in both of these tissues. In this way, the piRNA pathway may have been selected to be active in the somatic follicle cells and the fat body as secondary and tertiary defenses, respectively, against TEs that may infect and otherwise threaten the genomic integrity of the germline and future generations.

Enhancing TE suppression pathways to promote longevity

Fundamental work over the last two decades has helped define the physiological relevance of TEs in disease and the aging process. As researchers continue to resolve the molecular pathways involved in TE regulation, we may soon be able to augment these pathways in order to enhance genomic integrity and improve cellular function. Multiple studies have shown age-associated changes in chromatin states as well as the implementation of small molecules to alter the efficacy of chromatin-modifying enzymes resulting in extension of lifespan[31,32]. However, these strategies broadly target chromatin, and do not address the more mechanistic question of whether TEs contribute directly to lifespan or aging.

A growing number of studies have shown intact piRNA pathways in a diversity of stem cell populations and cancer cell types[25,33]. Perhaps not coincidentally, these cases of somatic piRNA pathways in non-gonadal somatic tissues also happen to be in cell types that are highly proliferative and even immortalized. Additionally, animals with

highly proliferative and abundant stem cell populations in which the piRNA pathway is active, such as *Hydra* and planaria, are effectively immortal[34–36]. It is likely that, in addition to the maintenance of stemness, the piRNA pathway in these cell types preserves genome integrity via TE suppression thus maintaining stem cell function. Progress has been made in testing reverse transcriptase inhibitors (RTIs), originally used to suppress HIV, for suppressing the reverse transcriptase encoded for by L1s[37,38]. However, a broader approach that could target multiple TEs for silencing, such as activating the piRNA pathway in somatic tissues, may also be beneficial. One study in mice showed that multiple piRNA pathway genes and piRNA precursors were under transcriptional control of the A-MYB transcription factor[39], suggesting a potential nexus point for pathway attenuation. Interestingly, another study was able to design and employ a small molecule consisting of a histone deacetylase inhibitor coupled with a DNA recognition domain to specifically target piRNA pathway genes for upregulation[40]. However, the small molecule's effect on TE activation was not tested. In Chapters 2 and 3, we examined the effect of feeding flies the HIV RTI lamivudine (3TC) in an effort to pharmacoligically inhibit TEs and, as a result, their deleterious effects. It will be of critical importance to expand on our findings and the work of other groups to see if it is possible to further enhance TE silencing, perhaps by augmenting or even reactivating endogenous TE suppresssion pathways.

Summary

The work described here investigates the role of TE suppression pathways in maintaining tissue function and organismal health. The data highlight the importance of

these pathways in preventing TE reactivation and genomic damage, thereby preserving cellular and tissue homeostasis. Additional studies will further illuminate how TEs contribute to disease states and normal aging and how TE suppression pathways may ameliorate such conditions.

References

1. Sienski, G., Dönertas, D. & Brennecke, J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151,** 964–80 (2012).

2. Post, C., Clark, J. P., Sytnikova, Y. a, Chirn, G. & Lau, N. C. The capacity of target silencing by Drosophila PIWI and piRNAs. 1977–1986 (2014). doi:10.1261/rna.046300.114.2

3. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8,** 272–85 (2007).

4. Lev-Maor, G. *et al.* Intronic Alus Influence Alternative Splicing. *PLoS Genet.* **4,** e1000204 (2008).

5. Hancks, D. C. & Kazazian, H. H. Roles for retrotransposon insertions in human disease. doi:10.1186/s13100-016-0065-9

6. Speek, M. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol. Cell. Biol.* **21,** 1973–85 (2001).

7. Cruickshanks, H. A. & Tufarelli, C. Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter. *Genomics* **94,** 397–406 (2009).

8. Zhu, C.-T., Chang, C., Reenan, R. A. & Helfand, S. L. Indy gene variation in natural populations confers fitness advantage and life span extension through transposon insertion. *Aging (Albany. NY).* **6,** 58–69 (2014).

9. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10,** 94–108 (2009).

10. Liu, J. *et al.* Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science (80-.).* **305,** (2004).

11. Meister, G. *et al.* Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Mol. Cell* **15,** 185–197 (2004).

12. Martinez, J. *et al.* Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110,** 563–74 (2002).

13. Hutvágner, G. & Zamore, P. D. A microRNA in a Multiple-Turnover RNAi Enzyme Complex. *Science (80-.).* **297,** (2002).

14. Förstemann, K., Horwich, M. D., Wee, L., Tomari, Y. & Zamore, P. D. Drosophila microRNAs Are Sorted into Functionally Distinct Argonaute Complexes after Production by Dicer-1. *Cell* **130,** 287–297 (2007).

15. Abe, M. *et al.* Impact of age-associated increase in 2'-O-methylation of miRNAs on aging and neurodegeneration in Drosophila. *Genes Dev.* **28,** 44–57 (2014).

16. Li, W. *et al.* Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat. Neurosci.* **16,** 529–31 (2013).

17. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128,** 1089–103 (2007).

18. Meikar, O. *et al.* An atlas of chromatoid body components. *RNA* **20,** 483–95 (2014).

19. Gou, L.-T. *et al.* Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.* **24,** 680–700 (2014).

20. Zhang, P. *et al.* MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. *Cell Res.* **25,** 193–207 (2015).

21. Robine, N. *et al.* A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19,** 2066–76 (2009).

22. Saito, K. *et al.* A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature* **461,** 1296–9 (2009).

23. Klein, J. D. *et al.* c-Fos Repression by Piwi Regulates Drosophila Ovarian Germline Formation and Tissue Morphogenesis. *PLOS Genet.* **12,** e1006281 (2016).

24. Siomi, M. C., Sato, K., Pezic, D. & Aravin, A. a. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **12,** 246–58 (2011).

25. Ross, R. J., Weiner, M. M. & Lin, H. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* **505,** 353–9 (2014).

26. Touret, F., Guiguen, F., Greenland, T. & Terzian, C. In between: Gypsy in Drosophila melanogaster reveals new insights into endogenous retrovirus evolution. *Viruses* **6,** 4914–4925 (2014).

27. Lecher, P. *et al.* Expression of the Drosophila retrovirus gypsy as ultrastructurally detectable particles in the ovaries of flies carrying a permissive flamenco allele. *J Gen Virol* **78 (Pt 9),** 2379–2388 (1997).

28. Kim, A., Terzian, C., Santamaria, P., Ptlisson, A. & Homme, N. P. Retroviruses in invertebrates: The gypsy retrotransposon is apparently an infectious retrovirus of Drosophila. *PNAS* **91,** 1285–1289 (1994).

29. Riechmann, V., Rehorn, K. P., Reuter, R. & Leptin, M. The genetic control of the distinction between fat body and gonadal mesoderm in Drosophila. *Development* **125,** 713–23 (1998).

30. Moore, L. a, Broihier, H. T., Van Doren, M. & Lehmann, R. Gonadal mesoderm and fat body initially follow a common developmental path in Drosophila. *Development* **125,** 837–44 (1998).

31. Wood, J. G. & Helfand, S. L. Chromatin structure and transposable elements in organismal aging. *Front. Genet.* **4,** 274 (2013).

32. Han, S. & Brunet, A. Histone methylation makes its mark on longevity. *Trends Cell Biol.* **22,** 42–9 (2012).

33. Van Wolfswinkel, J. C. Piwi and Potency: PIWI Proteins in Animal Stem Cells and Regeneration. *Integr. Comp. Biol.* 1–14 (2014). doi:10.1093/icb/icu084

34. Juliano, C. E. *et al.* PIWI proteins and PIWI-interacting RNAs function in Hydra somatic stem cells. *Proc. Natl. Acad. Sci.* **111,** 337–342 (2014).

35. Guo, T., Peters, A. H. F. M. & Newmark, P. A. A bruno-like Gene Is Required for Stem Cell Maintenance in Planarians. *Dev. Cell* **11,** 159–169 (2006).

36. Wagner, D. E., Ho, J. J. & Reddien, P. W. Genetic Regulators of a Pluripotent Adult Stem Cell System in Planarians Identified by RNAi and Clonal Analysis. *Cell Stem Cell* **10,** 299–311 (2012).

37. Jones, R. B. *et al.* Nucleoside analogue reverse transcriptase inhibitors differentially inhibit human LINE-1 retrotransposition. *PLoS One* **3,** e1547 (2008).

38. Dai, L., Huang, Q. & Boeke, J. D. Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. *BMC Biochem.* **12,** 18 (2011).

39. Li, X. Z. *et al.* An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol. Cell* **50,** 67–81 (2013).

40. Han, L. *et al.* A Synthetic Small Molecule for Targeted Transcriptional Activation of Germ Cell Genes in a Human Somatic Cell. *Angew. Chemie* **125,** 13652–13655 (2013).

Appendix A

Additional supplementary data

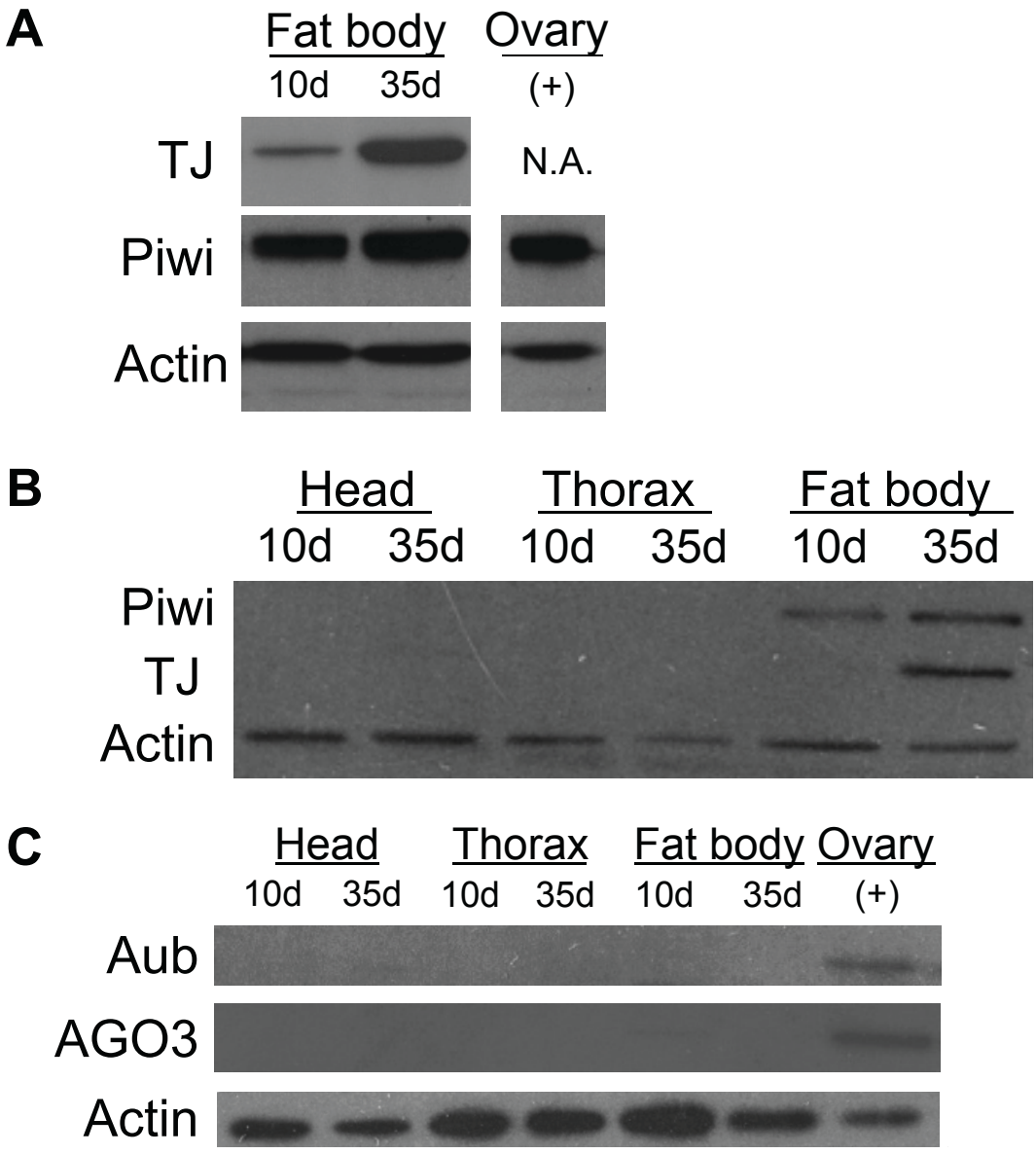


Figure 1. piRNA pathway proteins in aging somatic tissues.
(**A**) Protein of Piwi's transcription factor, Traffic jam (TJ) protein is present in the fat body along with Piwi. TJ and Piwi both increase with age in the fat body. Piwi is present in ovarian samples. TJ is not present in whole ovary samples as TJ is only found in the follicle cells of the ovary and whole ovaries do not give adequate protein signal.
(**B**) TJ and Piwi are not present in other somatic body segments and they do not appear with age.
(**C**) Other piRNA argonautes, Aubergine (Aub) and Argonaute-3 (AGO3), are absent from somatic body segements and tissues and do not appear with age. All piRNA argonautes are present in the ovary samples.
Actin serves as a loading control in all blots.

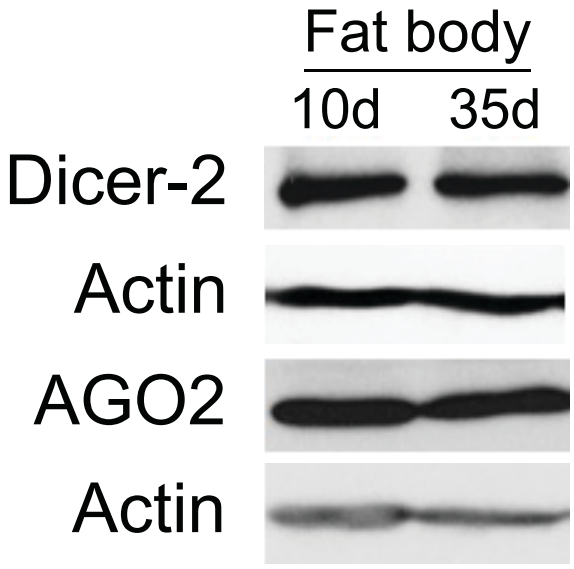


Figure 2. siRNA pathway proteins in aging somatic tissues. Proteins of the siRNA pathway remain consant with age. AGO2= Argonaute-2. Actin serves as a loading control.

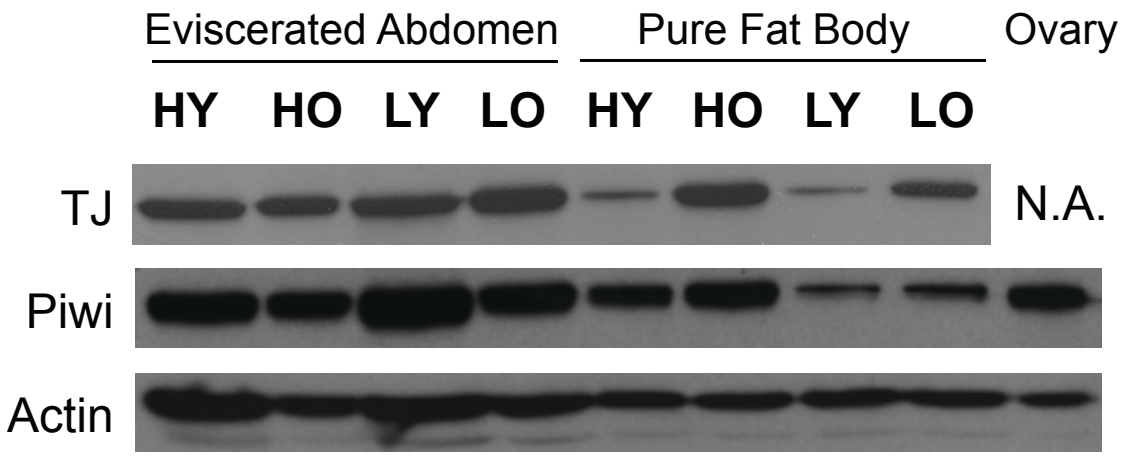


Figure 3. Effects of diet on piRNA pathway protein levels.
piRNA pathway proteins increase with age in pure fat body under both high calorie and low calorie diets. TJ and Piwi protein levels appear to decrease under low calorie diet relative to a high calorie diet. H=high calorie; L=low calorie; Y=young, 10 day old; O=old, 35 day old. Actin serves as a loading control.

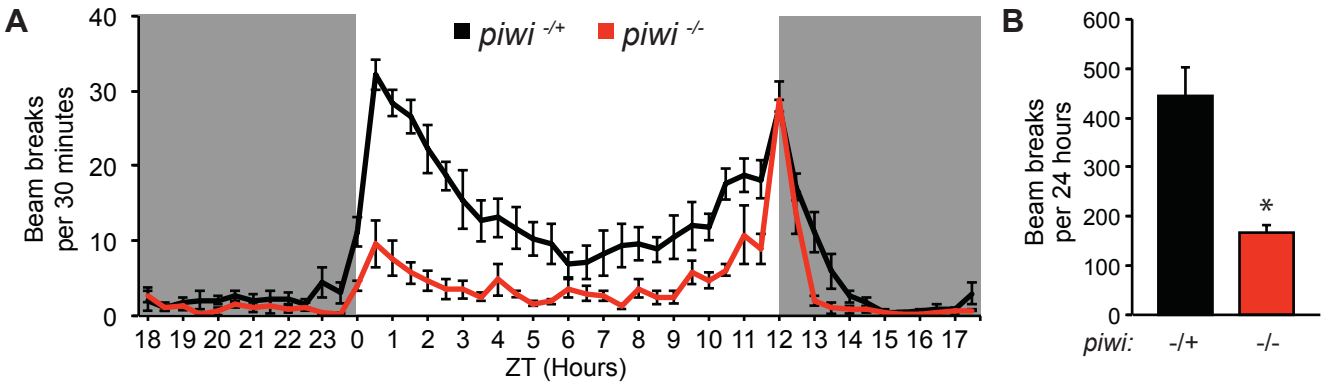


Figure 4. Activity monitor data of *piwi* mutants.
(**A**) Activity profiles of flies over a 24-hour period consisting of a 12-hour light-dark (LD) cycle. Bars represent average activity over 4 days. White background represents light period; grey background represents dark period. ZT: Zeitgeber time. Error bars are S.E.M. *piwi* $^{-/+}$: *n*=16 flies, *piwi* $^{-/-}$: *n*=14 flies.
(**B**) Total activity levels over a 24-hour period calculated from (I). *piwi* mutants are less active than their heterozygous controls. Bars represent average total activity over 4 days. Error bars are S.E.M. **P*<0.005

Appendix B

Contributions to additional published works

*Kreiling JA, ***Jones BC**, Wood JG, De Cecco M, Criscione SW, Neretti N, Helfand SL, and Sedivy JM. (2017) “Contribution of Retrotransposable Elements to Aging.” *Human Retrotransposons in Health and Disease*. Cristofari G (Ed.). *Springer International Publishing AG*. DOI 10.1007/978-3-319-48344-3_13 (*co-first authors)

Savva YA, Jepson JEC, Chang Y, Whitaker R, **Jones BC**, St Laurent G, Tackett MR, Kapranov P, Jiang N, Du G, Helfand SL, Reenan RA. (2013) RNA editing regulates transposon-mediated heterochromatic gene silencing. *Nature Communications* 4:1–11. DOI: 10.1038/ncomms3745