

Understanding Host Tolerance to Transposable Elements

by

Jyoti Lama

A dissertation submitted to the Department of Biology and Biochemistry,

College of Natural Science and Mathematics

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Biology

Chair of Committee: Erin Kelleher

Committee Member: Brigitte Dauwalder

Committee Member: Tasneem Bawa

Committee Member: William Murphy

University of Houston

December 2021

Abstract

Transposable elements (TE) are mobile genetic parasites, whose unregulated activity in germline causes DNA damage, and results in sterility. Host genomes can avoid these fitness costs of TEs either by regulating TE proliferation, or by altering gametogenesis to tolerate TE-induced DNA damage. Although TE regulation through piRNAs is studied extensively, little is known about mechanisms of gametogenic tolerance to TE activity.

To study tolerance, I took advantage of a unique phenomenon called hybrid dysgenesis in *Drosophila melanogaster* where naïve females devoid of *P*-element DNA transposon typically produce sterile offspring when mated with *P*-element carrying males. However, tolerant individuals are capable of producing viable gametes in spite of transposition. By performing Quantitative Trait Loci (QTL) mapping in a panel of highly recombinant inbred lines, two genomic regions associated with natural tolerance to *P*-element transposition were isolated. Transcriptome analysis of multiple tolerant and sensitive genotypes showed evidence suggesting variation in the Double-strand Break (DSB) repair efficiency. Tolerant genotypes showed increased expression of components of the Tat-interactive Protein 60-kDa (TIP60) complex involved in DSB repair and also exhibited increased chorion gene expression- an indicator of enhanced DSB repair.

By integrating the data from QTL mapping, gene expression, and in-phase SNP analysis, I identified two strong candidate genes that could influence tolerance: *brat* and *Nipped-A*, a member of the TIP60 complex. Loss-of-function mutation in *brat* promoted hybrid dysgenesis by increasing ovarian atrophy in the dysgenic females. Furthermore, tolerant genotypes displayed high resilience to X-ray mediated DNA damage. These results reveal gametogenic regulation and enhanced DSB repair as two potential mechanisms of germline tolerance to TEs.

Acknowledgment

I would like to express my deepest gratitude to my research advisor Dr. Erin Kelleher for her endless support, patience and guidance throughout this journey. She believed in me as a rotation student, let me be a part of her lab, and motivated me in every step of the way. I really appreciate everything she has done for me. I would also like to thank my committee members: Dr. Brigitte Dauwalder, Dr. Tasneem Bawa, and Dr. William Murphy for their helpful suggestions and guidance throughout the course of this project. This dissertation would have been incomplete without the efforts of the undergraduates: Satyam Srivastav, Sadia Tasnim, and Donald Hubbard. I would also like to thank our lab technician, Savanna Hadjipanteli, who helped me with my project. Furthermore, I would like to thank all the lab members: Dr. Shuo Zhang, Dr. Luyang Wang, Lorissa Saiz, and Modupeola Bolaji for their comments and suggestions. I'd also like to thank Dr. Richard Meisel and all the members of Meisel lab: Dr. Jae Hak Son, Dr. Pablo Delclos, Dr. Kiran Adhikari, and Danial Asgari for interesting discussions and constructive comments during our joint lab meetings. Last but not the least, I'd like to thank all my family members, especially my parents who always pushed me to achieve bigger goals in life.

Table of Contents

Chapter 1 Introduction	1
1.1 TE mobilization, diversity, and abundance	2
1.2 Impacts of TE	3
1.3 Host response to TEs	4
1.4 TE regulation through piRNA pathway	5
1.5 Horizontal transfer	7
1.6 P-element and hybrid dysgenesis	8
1.7 De novo piRNA production	10
1.8 Challenges in studying tolerance	11
1.9 Tolerance through promoting stem cell maintenance	11
1.10 Tolerance through efficient DNA-damage repair	13
1.11 Dissertation outline	15
Chapter 2. Investigating host factors conferring tolerance to TEs	16
2.1 Introduction	17
2.2 Result	19
2.2.1 QTL mapping of 2nd chromosome centromere	19
2.2.2 Sensitive and tolerant alleles exhibit differential expression of genes involved in chorion formation and chromatin packaging	21
2.2.3 piRNA clusters in QTL-3d exhibit differential activity that does not translate to TE deregulation	25
2.2.4 Identifying candidate genes influencing tolerance	27
2.3 Discussion	30
2.4 Methods	33
2.4.1 Drosophila strains and husbandry	33
2.4.2 Phenotyping	33
2.4.3 QTL mapping	34
2.4.4 Estimation of founder phenotypes and QTL phasing	34
2.4.5 Identification of in-phase polymorphisms	35
2.4.6 Selection of paired RILs with alternate QTL alleles	35
2.4.7 Small RNA-seq and total RNA-seq	35
2.4.8 Small-RNA analysis	36
2.4.9 Total RNA analysis	36
Chapter 3. Using isogenic lines to investigate the mechanism of tolerance	38
3.1 Introduction	39
3.2 Results	41
3.2.1 Generation of isogenic lines	41

3.2.2 Investigating the role of satellite repeats on tolerance	43
3.2.3 Investigating the role of de novo piRNA production on tolerance	45
3.2.4 Investigating association of irradiation sensitivity with tolerant and sensitive alleles	48
3.2.5 Investigating the role of <i>brat</i> on tolerance	50
3.3 Discussion	51
3.4 Methods	53
3.4.1 <i>Drosophila</i> strains and husbandry	53
3.4.2 Quantification of genomic <i>Responders</i>	54
3.4.3 Egg count assay	54
3.4.5 X-ray irradiation	54
3.4.6 Small-RNA analysis	55
Chapter 4. Conclusion	56
Appendix	62
Bibliography	76

List of Tables

2.1: QTL positions for tolerance in 3 and 21-day old females	20
--------------------------------------------------------------	----

List of Figures

Figure 1.1 piRNA-mediated resistance in <i>Drosophila</i> .	7
Figure 1.2 Variation in germline tolerance to <i>P</i> -elements is revealed by the severity of hybrid dysgenesis among offspring.	10
Figure 1.3 Diagram depicting the two ways the host responds to TE-mediated DNA damage: resistance and tolerance.	14
Figure 2.1 QTL mapping of variation in <i>P</i> -element tolerance.	21
Figure 2.2 Tolerance is associated with upregulated chorion proteins, whereas sensitivity is associated with upregulated replication-dependent histones.	24
Figure 2.3 Tolerance is not determined by differential activity of piRNA clusters or TE deregulation.	27
Figure 2.4 Differential expression and in-phase SNPs identify candidate genes that determine tolerance.	29
Figure 2.5 A model of TE tolerance in Population B RILs.	30
Figure 3.1 Isogenic line generation and phenotyping.	42
Figure 3.2 Responder copy number is associated with but does not determine tolerance.	45
Figure 3.3. Adult induction of <i>P</i> -element activity is not sufficient to trigger hybrid dysgenesis.	48
Figure 3.4. Tolerance is associated with enhanced DNA damage repair.	50
Figure 3.5. Loss-of-function mutation of <i>brat</i> increases severity of hybrid dysgenesis.	51
Figure 4.1. Multiple steps through which TIP60 could confer tolerance	60
Figure 4.2. Two models of TE tolerance to <i>P</i> -elements based on the QTL positions.	61

Chapter 1 Introduction

1.1 TE mobilization, diversity, and abundance

Transposable elements (TEs), also referred to as “jumping genes” are selfish genetic parasites, which are able to propagate themselves by replicating in the host cells. TEs are present in nearly all organisms—both prokaryotes and eukaryotes—where they constitute a major portion of their genome. For example, TEs comprise approximately 45% of the human genome (Consortium, International Human Genome Sequencing 2001), 20% of *D. melanogaster* genome (Mérel et al. 2020; Hill 2019) and >90% of maize genome (Jiao et al. 2017). Following the first discovery of TEs in maize by Barbara McClintok in 1950 (McClintock 1950; McClintock 1953) and the recognition that they are ubiquitous constituents of the eukaryotic genome, enormous research effort has been dedicated to understanding TEs and their manifold consequences to the host.

TEs are classified into two groups according to the mode of transposition: Class I TEs that move through a copy and paste mechanism and Class II TEs that use the cut and paste mechanism of transposition (McClintock 1950; Finnegan 1989). Class I TEs, also known as retrotransposons, transpose via RNA intermediates, which are reverse-transcribed into DNA by TE encoded reverse transcriptase followed by reinsertion into a new genomic location. Class II TEs, on the other hand, are DNA transposons that directly excise themselves and re-integrate elsewhere in the genome. Transposons encode for transposase enzymes that aid in the transposition process. Both Class I and Class II could either be autonomous or non-autonomous. Autonomous elements can mobilize on their own by encoding transposase or reverse transcriptase. Non-autonomous elements do not encode the proteins required for transposition and therefore depend on autonomous elements to provide for the required enzymes *in trans*.

TEs are further grouped into families, which include those elements that share sequence similarity because of their common ancestry. There exist diverse TE families within a genome of

a species, a large fraction of which is unique to that species (McClintock 1956; Kofler et al. 2015; Bourque et al. 2018). Organisms with especially large genomes such as maize, wheat, and barley can carry thousands of different types of TE families (Sanmiguel 1998; Vicient et al. 1999; Morgante 2006; Charles et al. 2008). However, not all TE families within the genome may be transpositionally active or mobilizable. Most older TE families are inactive because all of their genomic copies have lost their ability to mobilize. The proportion of active elements in the genome also varies among species. There are only two active TE families in the human genome (*LINE1* and *Alu*) (Britten and Kohne 1968; Mills et al. 2007), whereas more than 58 TE families are thought to be active in *Drosophila melanogaster* (Kofler et al. 2015).

1.2 Impacts of TE

TEs are a major source of mutations and are known to have diverse consequences for hosts. There are some instances where individual TE insertions are beneficial to the host such as TE-derived *RAG* genes that catalyse V(D)J somatic recombination in the vertebrate immune system (Kapitonov and Koonin 2015; Huang et al. 2016), CENP-B derived from POGO like transposons involved in centromeric chromatin assembly (Casola et al. 2008), and *HeT-A* and *TART* retrotransposons in *Drosophila* involved in telomere maintenance (Pardue and DeBaryshe 2003).

However, the occasional advantages conferred by TEs are unable to mask their parasitic nature, which is largely deleterious to the host. Unrestricted transpositions are highly mutagenic as TEs can either disrupt gene function by inserting within or nearby genomic sequences (McGinnis et al. 1983; Levis et al. 1984; Deininger and Batzer 1999) and/or promote ectopic recombination thereby mediating chromosomal rearrangements such as deletion, duplication, inversion, and translocation (Hedges and Deininger 2007; Han et al. 2008; Ade et al. 2013; Bennetzen and Wang 2014). These properties of TEs make them a powerful transgenic and

mutagenic tool commonly used in functional genetic analysis. Deleterious effects of TEs also arise through their production of Double-strand Breaks (DSBs). In fact, the majority of genomic instability caused by TEs is thought to be the byproduct of rampant DSBs (Vilenchik and Knudson 2003; Friedberg 2005; Hedges and Deininger 2007). DSBs are severe forms of DNA damage that may trigger cell-cycle arrest and even cause cell death (Gasior et al. 2006; Hedges and Deininger 2007).

TEs mobilize in germline as well as in somatic cells, both of which can have detrimental impact on the host. In humans, somatic transposition can lead to diseases such as cancers and psychiatric and neurodegenerative disorders, as well as cause aging and chronic inflammation (De Cecco et al. 2013; Ayarpadikannan and Kim 2014). However, the effects of somatic transposition are limited to the individual and need not have long-term fitness consequences, as these mutations are not transmitted to the offspring. Germline transpositions, which are transmitted to the next generation, impose severe fitness costs. Unrestricted proliferation in germ cells can not only transmit harmful mutations to offspring, it can also cause germ-cell loss and sterility. Germline transposition in humans is known to cause various hereditary diseases and has been associated with infertility (Halling et al. 1999; Hadziselimovic et al. 2015; Hancks and Kazazian 2016; Bourque et al. 2018). In *Drosophila*, DSBs resulting from germline transposition of certain TEs is known to be the cause of a sterility syndrome called hybrid dysgenesis (Kidwell et al. 1977; Bucheton et al. 1984; Brennecke et al. 2008). To avoid these fitness effects, hosts employ mechanisms to regulate TE mobilization or cope up with the harmful consequences of transposition.

1.3 Host response to TEs

Generally, host response to invading parasites, pathogens, and herbivores falls under two distinct categories, resistance and tolerance (Mauricio 2000; Roy and Kirchner 2000;

Råberg 2014). Resistance reduces parasite proliferation, whereas tolerant individuals are able to withstand or repair damage inflicted by the parasite. For instance, in *D. melanogaster*, intestinal infection by bacterial pathogens causes enterocyte cell death (Nehme et al. 2007). The host can either resist by triggering immune response pathways to kill the bacteria, or tolerate by promoting proliferation of intestinal stem cells, thus compensating for the loss of enterocytes (Cronin et al. 2009; Ferrandon 2009).

With respect to TEs, host resistance, whereby TEs are repressed both transcriptionally (histone modification and DNA methylation) and post-transcriptionally (TE transcript modification and degradation), has been well described. In plants, resistance is mediated through small interfering RNAs (siRNAs) (Hamilton et al. 2002; Kasschau et al. 2007). In many metazoans, repression occurs predominantly through the piRNA pathway (Brennecke et al. 2007; Nishida et al. 2007; Ernst et al. 2017). In mammals, TEs are additionally regulated by Kruppel-associated box-containing zinc finger proteins (KRAB-ZFPs) (Wolf et al. 2015; Yang et al. 2017). By contrast, host factors that confer cellular tolerance to TEs have been largely understudied.

1.4 TE regulation through piRNA pathway

In the germline of metazoans, TEs are silenced by the piRNA pathway, a germline-specific small RNA-silencing mechanism (Aravin et al. 2007; Brennecke et al. 2007; Ozata et al. 2018). The targets of piRNA pathway silencing are determined by small RNA molecules called piRNAs (23-29nt) together with Piwi-subfamily Argonaute proteins. The piRNA/protein complexes recognize complementary TE sequences and target them for silencing either transcriptionally by heterochromatin formation or post-transcriptional by endonucleolytic cleavage of TE transcripts (Brennecke et al. 2007; Le Thomas et al. 2013) (**Figure 1.1**). In *Drosophila*, piRNAs are produced in the germline and maternally deposited into the embryo, thereby transmitting silencing of corresponding TEs in the resulting offspring.

piRNAs are transcribed as long precursor transcripts from discrete genomic loci called piRNA clusters (Robert et al. 2001; Ozata et al. 2018) and are subsequently processed into mature piRNAs (23-29). piRNA clusters are enriched with multiple copies of truncated and defective TEs and are usually located in heterochromatic regions, such as the pericentromere and subtelomere (Brennecke et al. 2007; Malone et al. 2009; George et al. 2015). The exact number of piRNA clusters in *D. melanogaster* is unknown but more than 155 piRNA clusters of varying sizes have been reported so far (Brennecke et al. 2007; Malone et al. 2009; George et al. 2015). Among these the largest is known to be the *42AB* piRNA cluster (240 kb), accounting for ~30% of all piRNAs (Brennecke et al. 2007). Dual-stranded piRNA clusters, including *42AB*, are transcribed from both genomic strands and produce sense and antisense piRNAs in the ovarian germ cells. By contrast, uni-strand piRNA clusters produce overwhelmingly antisense TEs from a single genomic strand, and either silence TEs in the soma (e.g., *flamenco*) or in both soma and germline (e.g., *20A*) (Brennecke et al. 2007).

piRNA transcription from dual-stranded clusters is dependent upon the heterochromatic histone modification, histone 3 lysine 9 trimethylation (H3K9me3) (Le Thomas et al. 2014; Mohn et al. 2014). These clusters undergo non-canonical transcription, which do not use promoters but instead rely on H3K9me3 histone modifications (Le Thomas et al. 2014; Mohn et al. 2014). A germline-specific heterochromatin protein 1 paralog, Rhino, initiates transcription at many sites within these clusters by binding to H3K9me3, together with several other protein partners (Klattenhoff et al. 2009; Le Thomas et al. 2014; Mohn et al. 2014; Chang et al. 2019).

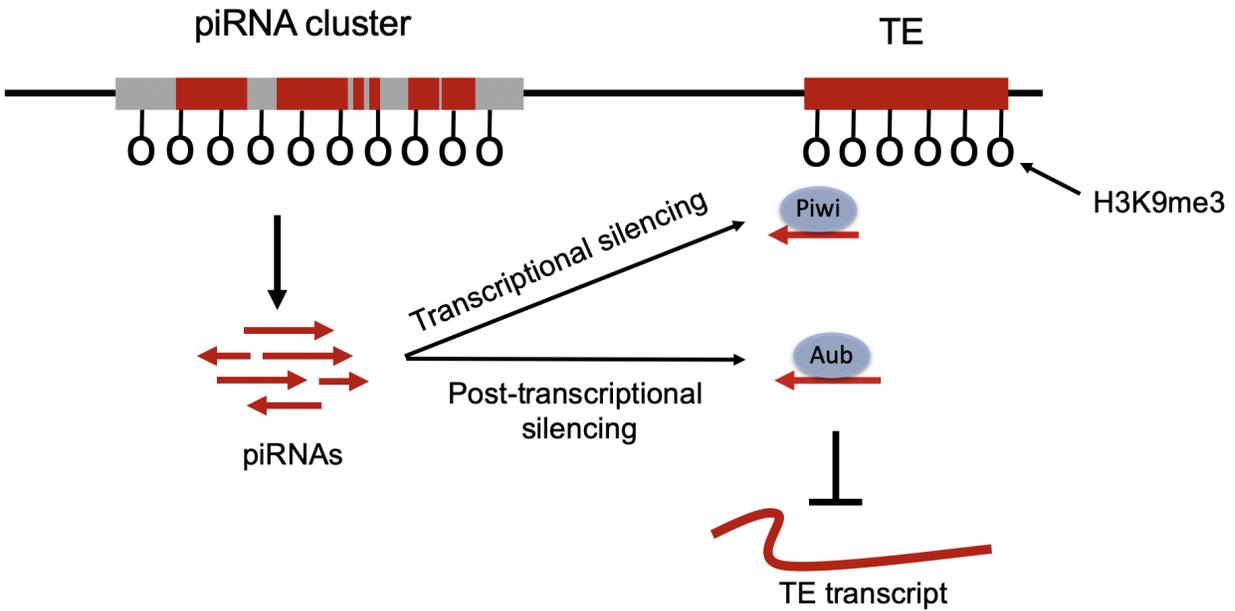


Figure 1.1 piRNA-mediated resistance in *Drosophila*. piRNAs are derived from piRNA clusters that silence complementary TEs either transcriptionally or post-transcriptionally.

1.5 Horizontal transfer

Host regulation of TEs suppresses their proliferation, eventually leading to the degradation and loss of all existing copies of a TE family from a host genome. However, TEs can escape this fate by invading a naive genome that has yet to evolve resistance (Schaack et al. 2010; Panaud 2016). Although TEs, like any genetic material, are inherited from parents to offspring, occasionally TEs transmit horizontally between non-mating individuals with the help of parasites or pathogens as vectors (Schaack et al. 2010; Panaud 2016; Wallau et al. 2018). With the development of new sequencing techniques, TEs, particularly DNA transposons, frequently undergo horizontal transfer (Schaack et al. 2010; Dotto et al. 2015; Panaud 2016; Zhang et al. 2020). For example, comparative genomic analysis of 195 insect species revealed at least 2248 horizontal transfer events during the last 10 million years (Peccoud et al. 2018).

Recurrent TE invasions impose a tremendous mutational burden on the host genome. When a TE invades a naive genome, it takes time for the host to evolve silencing against it

(Meiklejohn and Blumenstiel 2018). Absence of host repression during this time allows for rampant transposition (Anxolabéhère et al. 1988; Kofler et al. 2018), which is associated with severe fitness effects. For instance, unrestricted activity of certain TEs (*pogo*, *hobo*, I element, *P*-element) in *Drosophila* germlines results in a phenomenon called hybrid dysgenesis, characterized by a series of germline abnormalities such as elevated mutation, recombination, and even sterility in severe cases (Bucheton et al. 1976; Kidwell et al. 1977; Schaefer et al. 1979; Bucheton et al. 1984; Yannopoulos et al. 1987; Tudor et al. 1992).

The host's ability to tolerate the new TEs immediately following their invasion, may determine the survival and reproductive fitness of the species. In fact, the repeated successful invasion of TEs despite their detrimental effects, suggest that tolerance may be an important evolutionary strategy for the host. TE invasion throughout evolution may also repeatedly exert strong selective pressure for host tolerance, ultimately driving its adaptive evolution. Therefore, studying tolerance could help us better understand the host-TE interaction and its potential impacts on host gene evolution.

1.6 *P*-element and hybrid dysgenesis

The *P*-element invasion of *D. melanogaster* is one of the best-documented cases of horizontal transfer of a TE. *P*-elements are DNA transposons, which recently invaded *D. melanogaster* genome from *D. willistoni* in the 1950s (Kidwell 1983; Anxolabéhère et al. 1988; Daniels et al. 1990) through an unknown mechanism. Within just a few decades, *P*-elements rapidly spread worldwide, infecting all *D. melanogaster* populations in the wild. At the same time many wild populations also rapidly evolved resistance through the piRNA pathway (Kidwell 1983; Anxolabéhère et al. 1988; Brennecke et al. 2008).

P-element invasion was first detected by crosses between females from laboratory strains and males from natural populations, which produced sterile offspring displaying a suite of

aberrant traits collectively referred to as hybrid dysgenesis (**Figure 1.2**) (Kidwell et al. 1977). However, the reciprocal crosses produce fertile offspring signifying a maternal effect. The lab-maintained strains were collected prior to *P*-element invasion, and therefore lacked *P*-elements (referred to as M or maternal strains). By contrast, flies collected after invasion carried *P*-elements in their genome and were often able to induce dysgenesis (referred to as P or paternal strains) (Kidwell et al. 1977; Rubin et al. 1982). The interaction of the M strain and P strain reveals the detrimental impacts of newly invaded *P*-elements and allows us to further study the variable responses of the host.

The maternal effect of hybrid dysgenesis, in which genetically identical offspring of reciprocal crosses differ in the occurrence of sterility, is explained by the maternal transmission of piRNAs. Because M strain females lack *P*-elements, they are unable to produce *P*-element derived piRNAs that suppress transposition. Lack of maternally deposited *P*-element piRNAs similarly renders the offspring defenseless against the *P*-element activity in their germline. Consequently, the unregulated activity of *P*-elements induces DNA damage, leading to germ-cell degeneration and gonadal atrophy.

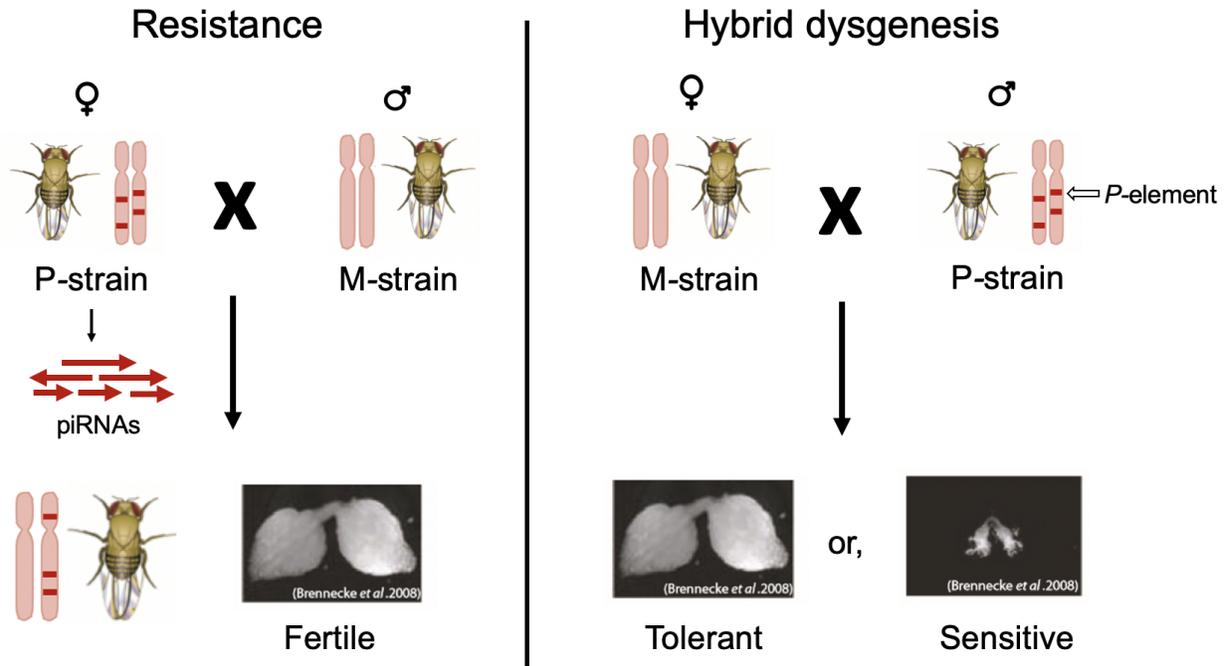


Figure 1.2. Variation in germline tolerance to *P*-elements is revealed by the severity of hybrid dysgenesis among offspring. In the F1 offspring from *P*-elements carrying females and naïve males (left), maternally deposited piRNAs regulate the activity of the *P*-elements in the germline, thus maintaining fertility. The reciprocal cross (right) shows hybrid dysgenesis. The naïve mothers are incapable of producing and transmitting *P*-element piRNAs. Typically, the unregulated transposition of paternally inherited *P*-elements induces germline loss, producing an atrophied ovary showing sensitivity to hybrid dysgenesis. However, dysgenic offspring can also show tolerance, where unregulated transposition of paternally inherited *P*-elements does not induce germline loss and instead produces fertile offspring.

1.7 *De novo* piRNA production

piRNAs that are produced in the oocyte are maternally deposited and feed forward piRNA defense against complementary TEs in the resulting embryo. Maternally inherited piRNAs are therefore critical for the fertility of the progeny (Le Thomas et al. 2014). M strains lack *P*-element derived piRNAs, thus producing dysgenic/sterile offspring when mated to *P*-element carrying males. However, the offspring of some naïve genotypes can mount piRNA-mediated silencing of *P*-elements even in the absence of maternally deposited piRNAs (Khurana et al. 2011; Casier et al. 2019). These M strains are reported to display

age-dependent recovery of hybrid dysgenesis through *de novo* production of *P*-element piRNAs from paternally inherited piRNA clusters (Khurana et al. 2011; Moon et al. 2018). *De novo* piRNA production is also known to occur under stressful conditions such as when exposed to high temperature (Casier et al. 2019).

1.8 Challenges in studying tolerance

piRNA-mediated silencing presents a challenge to studying tolerance. The repression of transposition masks the action of tolerant variants that work to minimize the impacts of TE transposition on fitness. Therefore, the differential response of germline cells to unbridled transposition is revealed only in dysgenic crosses. Although unregulated mobilization of *P*-elements typically leads to complete loss of germ cells and sterility, the tolerant variants are capable of avoiding this outcome and producing viable gametes (**Figure 1.2**). Multiple studies demonstrate that naive-maternal genotypes show variation in the degree of severity in hybrid dysgenesis, suggesting the existence of tolerant genetic variants (Ignatenko et al. 2015; Funikov et al. 2018; Kelleher et al. 2018; Serrato-Capuchina et al. 2020). However, the mechanisms underlying cellular tolerance to TEs remain unclear. Tolerance could be conferred in two possible ways: either by promoting stem cell maintenance or by a more efficient double-strand break repair system.

1.9 Tolerance through promoting stem cell maintenance

Germline stem cell maintenance may be critical for germline tolerance to *P*-elements. The germ-cell loss arising from *P*-element transposition is due to the loss of adult germline stem cells (GSCs) as well as their larval precursors, the primordial germ cells (PGCs) (Dorogova et al. 2017; Teixeira et al. 2017; Tasnim and Kelleher 2018; Ota and Kobayashi 2020).

Furthermore, TE expression in general is more abundant in GSCs and their immediate progeny

than in late stage germ cells (Story et al. 2019). This may be because TE insertion in GSCs ensures transmission to all the gametes arising from the same progenitor GSC, thus allowing for the ample spread of TEs to the next generation.

The GSCs of *D. melanogaster* are a major model for the investigation of stem cell biology. In *Drosophila* females, GSCs reside in the anterior most portion of germarium, a specialized sub-structure of the ovary. Each germarium contains two to three GSCs attached to terminal filaments and cap cells that together form a niche providing signals for self-renewal to the stem cell. GSCs divide asymmetrically to produce one stem cell and a cystoblast, which later divides and differentiates to form mature oocytes. In this way, GSCs ensure continuous supply of germ cells and ultimately mature gametes. Therefore, if GSCs are lost, the gametes are not formed, which is evident during hybrid dysgenesis.

The larval gonad of *D. melanogaster* harbors over a hundred undifferentiated PGCs (Zhu and Xie 2003), only a subset of which later forms the self-renewing GSCs that will remain in the ovary throughout adulthood (King 1970; Asaoka and Lin 2004). The fate of the PGCs is determined by the niche microenvironment, which provides signals to suppress differentiation. During late larval stages, a somatic niche is formed that limits the contact with PGCs, such that PGCs outside the niche differentiate to cytotoblasts while those in contact give rise to GSCs. The mechanisms underlying the maintenance of PGCs and GSCs are likely similar as they are known to involve the same set of genes (Niki and Mahowald 2003; Gilboa and Lehmann 2004).

GSC/PGCs loss in dysgenic females is most likely driven by the DNA damage resulting from rampant *P*-element transposition (Khurana et al. 2011). Since GSCs are resistant to apoptosis (Xing et al. 2015), the mechanisms underlying their loss are not well understood. However, it is suggested that GSC loss may be due to disruption of signalling pathways that maintain GSCs (Ma et al. 2016; Kelleher et al. 2018). Tolerance could therefore arise by weakening the connection between DNA damage and germline loss, thus making germline robust to DNA damage (**Figure 1.3**). In this scenario, damaged germ cells would persist and

form gametes, but those gametes would carry larger numbers of new mutations.

Recent studies have provided evidence of stem cell maintenance as a source of tolerance. For example, overexpression of the stem cell self-renewal factor *myc* confers tolerance by suppressing loss of damaged PGCs in the dysgenic larval gonads and produces gametes with high mutation loads (Ota and Kobayashi 2020). Similarly, a recent study on natural variation in hybrid dysgenesis by Kelleher et al. (2018) revealed germline differentiation factor, *bruno* as a tolerance factor. Reduced activity of *bruno* is thought to promote retention of GSCs in the niche, thereby preventing loss of germ cells in the face of *P*-element-mediated DNA damage. These observations suggest that germline regulation of stem cell maintenance is a potential mechanism of germline tolerance.

1.10 Tolerance through efficient DNA-damage repair

Alternatively, tolerance could be determined through enhanced repair of TE-induced DNA damage (**Figure 1.3**). DNA damage triggers a cellular response called DNA Damage Response (DDR). The DDR comprises a network of signaling pathways that senses DNA damage and triggers activation of cell-cycle checkpoint and repair machinery. Cell-cycle checkpoints ensure time to repair the DNA damage. However, if the damage is excessive, DDR triggers cell death. But most adult stem cells are resistant to apoptosis (Reya et al. 2001; Liu et al. 2014; Xing et al. 2015) and the cause of their loss remains unclear.

The DNA damage from unrestricted activity of TEs induces cell-cycle arrest and germ-cell loss. Mutations in the DDR components: checkpoint kinase 2 (*chk2*) and *p53* influences the frequency of ovarian atrophy in dysgenic females (Moon et al. 2018; Tasnim and Kelleher 2018). In GSCs, *chk2* triggers cell-cycle arrest in response to DNA damage and *p53* aids in cell-cycle re-entry (Shim et al. 2014; Wylie et al. 2014; Ma et al. 2017). This is distinct from its function in meiotic cells, where *p53* acts downstream of *chk2* to trigger cell-cycle arrest

and apoptosis (Shim et al. 2014). Mutations in *chk2* suppress ovarian atrophy, but the resulting gametes are inviable due to unrepaired DNA damage. On the other hand, mutations in p53 increase ovarian atrophy during hybrid dysgenesis. These observations suggest that DDR represents a potential mechanism of tolerance.

DDR could promote tolerance in three possible ways: **1)** increased activity of repair pathway, **2)** increased duration of arrest to ensure time for repair, and **3)** decreased probability of GSC loss. However, having a more dynamic and effective DNA-damage repair mechanism is likely to be more beneficial to the host as it not only ensures fertility in the face of transposition but also helps preserve genomic integrity for the next generation. Furthermore, stem cells in general are reported to have overexpression of DNA-damage sensing and repair factors (Saretzki et al. 2004; Maynard et al. 2008).

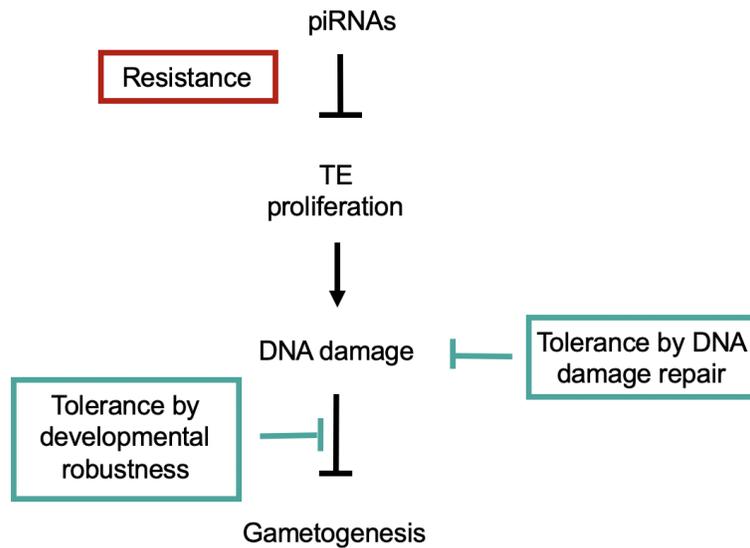


Figure 1.3. Diagram depicting the two ways the host responds to TE-mediated DNA damage: resistance and tolerance. Resistance is conferred by piRNAs that directly suppress TE proliferation. Two possible ways to tolerate consequences of transposition are either through developmental robustness or through DNA-damage repair that could be brought about by more efficient DNA-damage repair.

1.11 Dissertation outline

In this dissertation, I sought **1)** to genetically map host factors conferring germline tolerance to *P*-elements in a panel of recombinant inbred lines derived from *Drosophila* synthetic population resource (DSPR), and **2)** to determine the mechanism of tolerance using the isogenic lines carrying tolerant and sensitive alleles.

To study natural tolerance to TEs, our lab previously performed quantitative trait locus (QTL) analysis in one of the panels of recombinant inbred lines (RILs). It revealed two QTL loci at the pericentromeric region of the 2nd chromosome associated with germline tolerance to *P*-element. In Chapter 2, I used small RNA and total RNA-seq to look at phenotypic differences associated with tolerant and sensitive QTL alleles. I found that tolerance is associated with increased chorion assembly, whereas sensitivity is associated with increased histone and pericentromeric gene expression. To identify candidate genes, I specifically looked for genes that were differentially expressed within the QTL and also harbored in-phase SNPs. Finally, I identified two candidate genes, one in each QTL, that potentially determine the phenotypic differences associated with tolerant and sensitive alleles.

In Chapter 3, I generated isogenic lines carrying either a tolerant or a sensitive QTL allele and tested different hypotheses on the mechanisms of tolerance. I found evidence supporting two of the four hypotheses tested: **1)** enhanced DNA-damage repair promotes tolerance, and **2)** activity of germline differentiation factor *brat* promotes tolerance. These observations point towards two potential mechanisms of tolerance to TEs.

Chapter 2. Investigating host factors conferring tolerance to TEs¹

¹This chapter is a part of a manuscript entitled “Natural tolerance to transposition is associated with Myc-regulation and DNA repair. The preprint is available at <https://doi.org/10.1101/2021.04.30.441852>

2.1 Introduction

Transposable elements (TE) are mobile DNA sequences that spread through host genomes by replicating in germline cells. Although individual TE insertions are sometimes beneficial, genomic TEs are foremost genetic parasites (reviewed in Chuong et al. 2017). Unrestricted transposition not only produces deleterious mutations, but also DSBs that lead to genotoxic stress in developing gametes. Generally, hosts avoid the fitness costs of invading parasites, pathogens, and herbivores by two distinct mechanisms, resistance and tolerance (Mauricio 2000; Roy and Kirchner 2000; Råberg 2014). Resistance reduces parasite proliferation, whereas tolerant individuals experience reduced fitness costs from parasitism. With respect to TEs, host resistance has been the focus of extensive research, and occurs through production of regulatory small RNAs that transcriptionally and post-transcriptionally silence TEs in the germline (Brennecke et al. 2007; Nishida et al. 2007; Malone and Hannon 2009). By contrast, tolerance mechanisms that could ameliorate the fitness costs of transposition during gametogenesis remain largely unstudied.

The absence of research on tolerance in part reflects the ubiquity of resistance. For example, in *Drosophila melanogaster*, where resistance to TEs is extensively studied, all actively transposing TE families are silenced in developing gametes by the Piwi-interacting RNA (piRNA) pathway (Brennecke et al. 2007). In the presence of strong resistance that represses transposition, individual differences in tolerance will not be apparent. Therefore, I made use of the P-M hybrid dysgenesis in *Drosophila melanogaster*, where resistance to *P*-element DNA transposons is short-circuited due to the absence of maternally transmitted piRNAs (reviewed in Kelleher 2016). When males bearing genomic *P*-elements (*P*-strain) are mated to naive females lacking *P*-elements and corresponding piRNAs (*M*-strain), they produce dysgenic offspring that does not regulate *P*-element transposition in germline cells (Brennecke et al. 2008). A range of fertility effects result from *P*-element-induced DNA damage, including the

complete loss of germline cells (Kidwell et al. 1977). The ability of an individual to produce gametes despite *P*-element transposition is therefore a measure of tolerance.

Recent forward genetic studies of dysgenic germline loss have revealed potential mechanisms of *P*-element tolerance. Mutations in checkpoint kinase 2 (*chk2*), a key factor in germline response to DSBs, suppress germline loss in dysgenic females (Moon et al. 2018; Tasnim and Kelleher 2018). While the gametes produced by the dysgenic females are inviable due to unrepaired DNA damage, these observations suggest that enhanced DSB repair in germline cells could provide tolerance. Alternatively, tolerance could arise by weakening the connection between DNA damage and germline loss, allowing dysgenic individuals to maintain gametogenesis but produce gametes with more mutations. For example, overexpression of the stem cell self-renewal factor *myc* is associated with suppressed germline loss in dysgenic males and females, resulting in the production of additional gametes that exhibit more *P-element* transpositions (Ota and Kobayashi 2020).

Natural variation in hybrid dysgenesis provides another opportunity to study tolerance. In particular, the degree of dysgenic sterility differs among M-strains, with germline loss being less prevalent in the offspring of some maternal genotypes (Kidwell et al. 1983; Anxolabéhère et al. 1988; Ignatenko et al. 2015; Kelleher et al. 2018). This suggests the presence of natural tolerance alleles. Using a panel of highly Recombinant Inbred Lines (RILs) from the *Drosophila* Synthetic Population Resource (DSPR, King et al. 2012), our lab recently uncovered a natural tolerance allele through Quantitative Trait Locus (QTL) mapping (Kelleher et al. 2018). Kelleher et al. (2018) further associated tolerance with reduced expression of *bruno*, a female germline differentiation factor whose ectopic expression in stem cells promotes their loss (Parisi et al. 2001; Wang and Lin 2007; Xin et al. 2013). Kelleher et al. (2018) speculated *bruno* tolerance potentially arises by desensitizing gametogenesis to DNA damage in a mechanism analogous to *myc* overexpression.

In this chapter, I report results following the QTL mapping of hybrid dysgenesis in a second, independent panel of DSPR RILs (Population B, King et al. 2012). Our lab (Dr. Kelleher and her group) uncovered two QTL peaks close to the second chromosome centromere that determine tolerance to *P*-element activity in young and old females. Here, I interrogated the tolerance phenotype by contrasting RNA and small RNA expression between nearly isogenic strains (NILs) carrying tolerant and sensitive QTL alleles. Finally, I combined information about expression differences, RIL genotypes, and QTL positions to identify novel candidates for natural variation in tolerance.

2.2 Result

2.2.1 QTL mapping of 2nd chromosome centromere

The DSPR RILs are all *P*-element free M-strains, which were isolated from natural populations before the *P*-element invasion (King et al. 2012). Our lab therefore screened for tolerant alleles among the panel B RIL genomes by crossing RIL females to males from the reference *P*-strain Harwich, and examined the morphology of the F1 ovaries (**Figure 2.1a**). Atrophied ovaries are indicative of germline loss resulting from *P*-element activity, while non-atrophied ovaries are indicative of tolerance (Schaefer et al. 1979; Kelleher et al. 2018). Since some females exhibit age-dependent recovery from *P*-element hybrid dysgenesis (Khurana et al., 2011), our lab phenotyped F1 females at two developmental time points: 3 days and 21 days post-eclosion.

Similar to previous observations with the Population A RILs (Kelleher et al. 2018), our lab found continuous variation in the frequency of ovarian atrophy among dysgenic offspring of different RIL mothers, indicating genetic variation in tolerance. Based on a combined linear model of F1 atrophy among 3 and 21 day-old females, our lab estimated the broad-sense heritability of tolerance in our experiment to be ~42.5%. However, the effect of age on the

proportion of F1 atrophy was significant but minimal ($X^2= 7.03$, $df = 1$, p -value = 0.008) with 3-day-old females showing only 0.7% increase in atrophy as compared to 21-day-old females. Therefore, age-dependent recovery from dysgenic sterility is not common among the genotypes our lab sampled.

To identify the genomic regions associated with genetic variation in germline tolerance, our lab performed QTL analysis using the published RIL genotypes (King et al. 2012). Our lab found a large QTL peak near the 2nd chromosome centromere in both 3 and 21 day-old F1 females (**Figure 2.1b, Table 2.1**). However, the intervals of the major QTL peaks, based on the $\Delta 2$ LOD and Bayes Credible Interval (BCI) methods (Lander and Botstein 1989; Manichaikul et al. 2006), are non-overlapping between the 3 and 21 day-old data sets (**Figure 2.1c, Table 2.1**). The major QTL in 21 day-old females (hereafter, QTL-21d) resides in the euchromatic region and is quite small (990 kb) compared to the major QTL in 3 day-old females (hereafter QTL-3d), which spans the centromere and pericentromeric regions (9.6 Mb). Therefore, there are likely at least two polymorphisms that influence tolerance near the 2nd chromosome centromere, one of which is more important in young 3-day old females, and the other of which is more important in 21 day-old females.

Table 2.1. QTL positions for tolerance in 3 and 21-day old females

Analysis	LOD Score	Peak Position	$\Delta 2$ LOD CI	BCI	% variation
3-day	15.2	2R:6,192,495	2L:20,710,000- 2R:7,272,495	2L:20,820,000- 2R:6,942,495	11.13
21-day	10.13	2L:19,420,000	2L:19,170,000- 20080000	2L:19,010,000- 20000000	9.78

The peak position, $\Delta 2$ LOD drop confidence interval ($\Delta 2$ LOD CI), and the Bayesian Credible Interval (BCI) in dm6 are provided for each analysis.

The presence of two tolerance QTL is further supported by the phenotypic classes our

lab detected among founder alleles (B1-B8) for each of the QTL peaks (**Figure 2.1e**). For QTL-21d, both B2 and B6 founder alleles were sensitive and greatly increased dysgenic ovarian atrophy, while all other founder alleles were tolerant. By contrast for QTL-3d, only the B6 founder allele was associated with increased sensitivity.

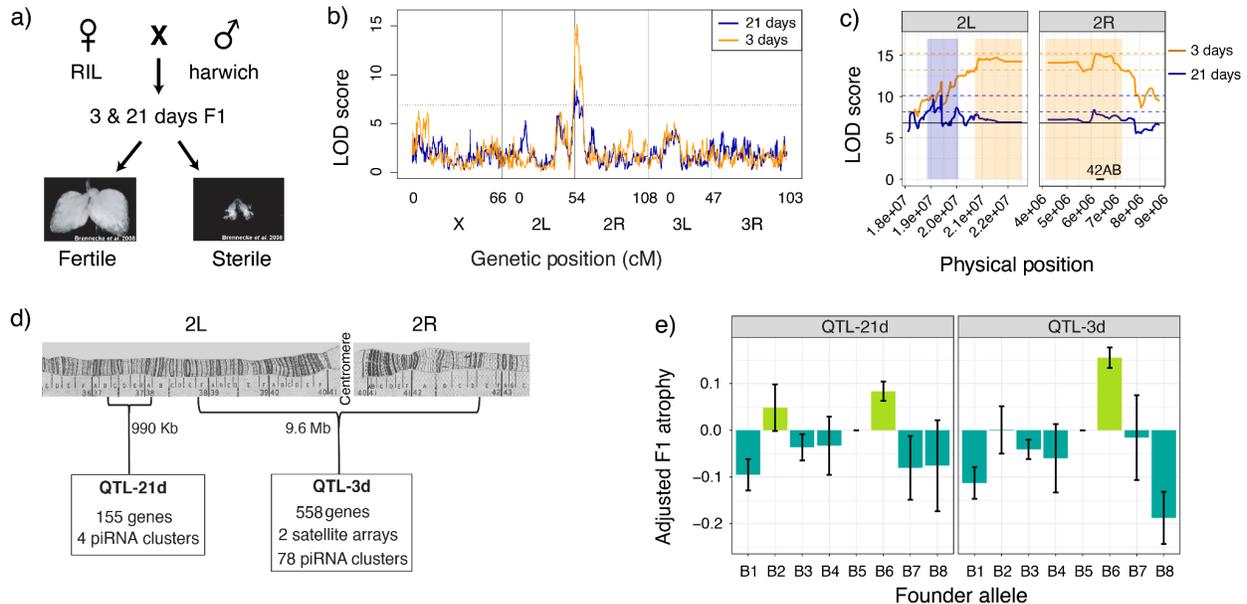


Figure 2.1: QTL mapping of variation in *P*-element tolerance. **a)** Crossing scheme to phenotype the variation in tolerance to *P*-elements among the RILs by screening for ovarian atrophy in 3 and 21 day-old dysgenic F1 females **b)** The log of odds (LOD) plot for QTL mapping of germline tolerance using 3 day-old (orange) and 21 day-old (blue) F1 females. The dotted line is the LOD threshold and x-axis represents the chromosomal positions. **c)** Zoomed-in figure of QTL mapping from 3 days (orange) and 21 days (blue). The colored boxes show the $\Delta 2$ LOD confidence interval of each QTL, and the pairs of dotted lines indicate the LOD peak position and the $\Delta 2$ LOD score that determines the interval. The solid horizontal line is the LOD threshold. **d)** Cytological map depicting the interval of the two QTL peaks (Bridges, 1935; Bridges, 1942). **e)** Graph showing F1 atrophy (y-axis) associated with each of the eight founder alleles (x-axis) at the QTL peaks. All the QTL peaks show two phenotypic classes: a sensitive (light green) and tolerant (dark green) class.

3.2.2 Sensitive and tolerant alleles exhibit differential expression of genes involved in chorion formation and chromatin packaging

Both the QTL regions contain large numbers of protein coding and non-coding RNA

genes, piRNA clusters, and repeats, which could influence tolerance (**Figure 2.1d**). To better understand the tolerance phenotype, I examined differential gene expression between recombinant inbred lines carrying tolerant and the sensitive QTL alleles. I identified three pairs of nearly isogenic lines (NILs), which carried either a sensitive (B6) or tolerant (B4) QTL haplotype across the QTL region (dm6 2L:19,010,000-2R:7,272,495) in an otherwise similar genetic backgrounds. I then performed RNA-seq on ovaries of 3-5 day-old females (three biological replicates). Principal component analysis (PCA) of read counts revealed two independent axes that resolve sensitive and tolerant genotypes, which together accounted for 40% and 16% of variation (**Figure 2.2a**). One biological replicate of RIL 21188 (tolerant) was an outlier, which I excluded from downstream analysis of differentially expressed genes.

I found a total of 530 genes differentially expressed between sensitive and tolerant genotypes (Benjamini-Hochberg adjusted p -value ≤ 0.05 , fold-change > 1.5). The most significantly enriched GO term among genes upregulated in tolerant genotypes was chorion assembly (Bonferroni corrected p -value < 0.01 , **Figure 2.2b**). Indeed, all of the major chorion genes were found to be significantly upregulated in the tolerant genotypes (**Figure 2.2c**, Tootle et al. 2011; Kim et al. 2011). It is unlikely that chorion assembly impacts dysgenic ovarian atrophy, since chorion synthesis occurs in late-stage oocytes (stages 10B-14, Waring 2000), whereas atrophy results from the loss of larval primordial germline cells and subsequent germline stem cells (GSCs) (Dorogova et al. 2017; Teixeira et al. 2017; Tasnim and Kelleher 2018; Ota and Kobayashi 2020). However, chorion genes reside in clusters that undergo amplification (Spradling 1981; Claycomb et al. 2004), a process that relies on the efficient repair of DSBs at the boundaries of an amplified region (Alexander et al. 2015). Therefore, upregulation of chorion genes in tolerant genotypes could indicate more efficient DSB repair, which might off-set the impact of P -element transposition.

Genes upregulated in the sensitive genotypes were enriched for functions in chromatin assembly and transcription, cell division, and translation. A careful inspection of genes

underlying these enriched terms revealed that with the exception of translation, they were majorly driven by the increased expression of replication-dependent (RD) histone gene copies (**Figure 2.2d**). Overexpression of RD histones is associated with increased sensitivity to DNA damage in yeast (Gunjan and Verreault 2003; Liang et al. 2012), mice (Murga et al. 2007) and *Drosophila* (Landais et al. 2014; Ozawa et al. 2016). Therefore, histone upregulation exhibited by sensitive alleles may reduce their tolerance to genotoxic stress resulting from *P*-element activity. Notably, the expression of both histone and chorion genes are increased in late oogenesis (Ambrosio and Schedl 1985; Ruddell and Jacobs-Lorena 1985; Waring 2000; Potter-Birriel et al. 2021), meaning that their inverted differential expression between sensitive and tolerant genotypes cannot be explained by differential abundance of late stage oocytes.

The *D. melanogaster* histone gene cluster is located in the pericentromeric region of QTL-3d and consists of ~100 copies of a 5-kb cluster containing each of the five RD histones (*his1*, *his2A*, *his2B*, *his3*, and *his4*). However, the differential regulation of histones is unlikely to reflect the presence of a *cis*-regulatory variant within the QTL, since the histone gene cluster exhibits coordinated and dosage-compensated regulation in a unique nuclear body called the histone locus body (HLB, McKay et al. 2015). Rather, I postulated that sensitive and tolerant alleles may differ in heterochromatin formation, since many negative regulators of histone gene transcription are also suppressors of position effect variegation (Su(var), Ner et al. 2002; Ozawa et al. 2016). In support of this model, sensitive (B6) genotypes exhibited increased expression of pericentromeric genes, as well as genes on the heterochromatic 4th chromosome (**Figure 2.2e**). I also discovered increased expression of pericentromeric genes associated with the B6 haplotype in a previously published microarray dataset from head tissue (King et al., 2014, **Figure A2.1**), suggesting B6 is unusual among the founder alleles in exhibiting reduced heterochromatin formation.

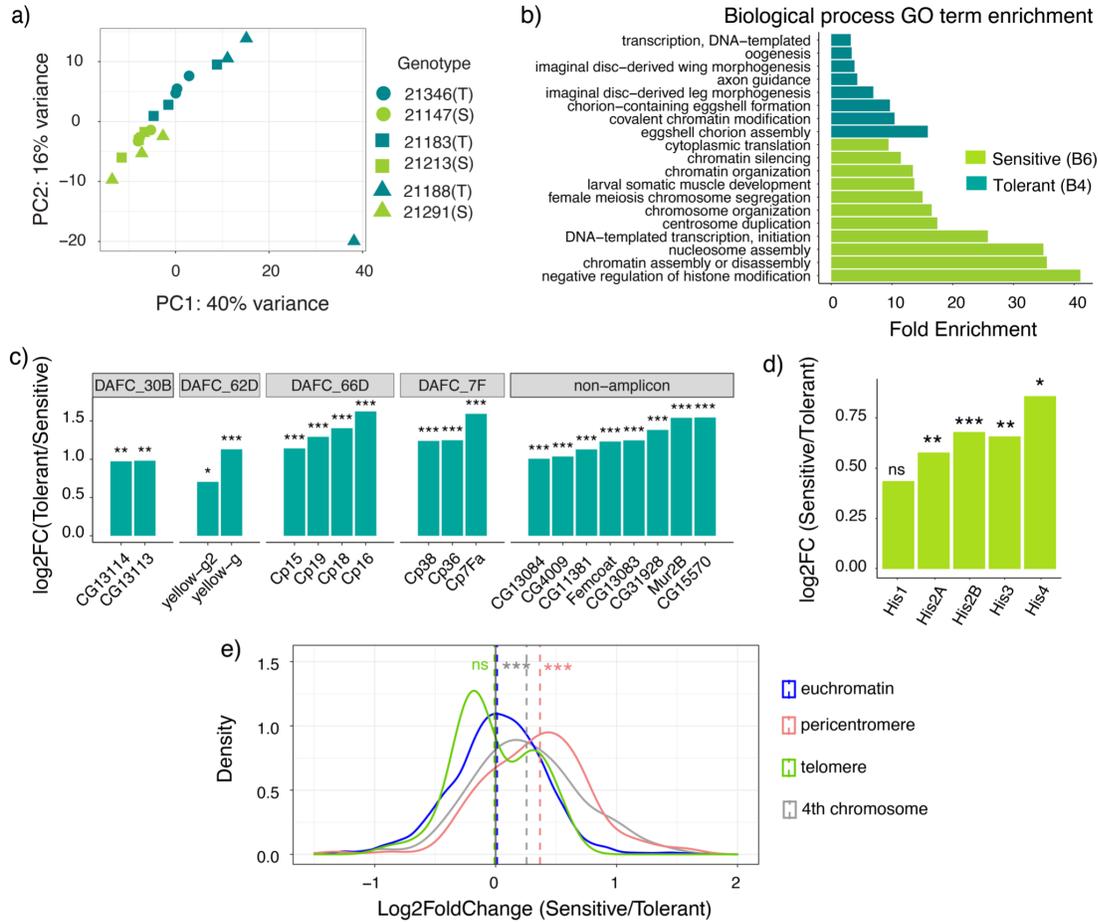


Figure 2.2: Tolerance is associated with upregulated chorion proteins, whereas sensitivity is associated with upregulated replication-dependent histones. **a)** PCA analysis of gene expression data of S/sensitive and T/ tolerant genotypes. Members of the same NIL pair are represented by the same shape. **b)** GO terms enriched among genes upregulated in tolerant and sensitive genotypes. **c)** Log2 fold change increase in expression in tolerant genotypes for chorion genes residing in the four amplicons (*Drosophila* Amplicons in Follicle Cells, DAFCs) as well as outside amplicons (Kim et al. 2011; Tootle et al. 2011). **d)** Log2 fold change increase in RD histone expression in sensitive genotypes. **e)** Probability density plot of log2 fold change values for all euchromatic (blue), pericentromeric (red), telomeric (green) genes, and 4th chromosome (gray) between strains carrying sensitive and tolerant alleles. The mean of each distribution is represented by a dotted line. Sensitive genotypes display significantly higher expression of pericentromeric genes (two-sample t-test, $t_{141} = -9.32$, p -value = $2.335e-16$) and 4th chromosome genes (two-sample t-test, $t_{53} = -4.56$, p -value = $3.014e-05$) when compared to euchromatic genes. For **e)** the x-axis boundaries were confined from (-1.5 to 2) for a better visualization. The pericentromere-euchromatin boundaries were drawn from (Riddle et al. 2011; Hoskins et al. 2015) and subtelomeric-euchromatin boundary coordinates from (Karpen and Spradling 1992; Walter et al. 1995; Yin and Lin 2007).

3.2.3 piRNA clusters in QTL-3d exhibit differential activity that does not translate to TE deregulation

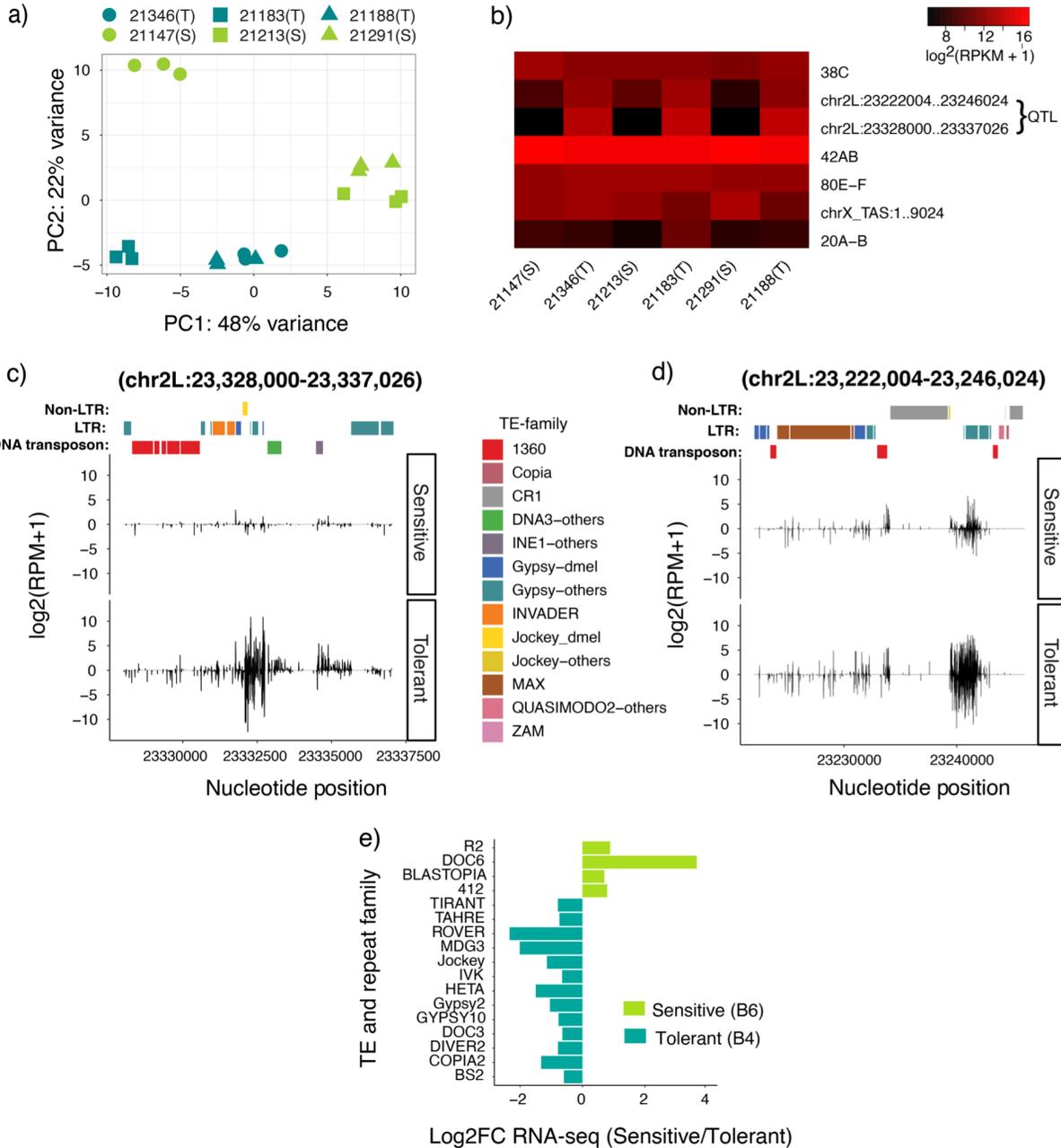
The pericentromeric region is rich in piRNA clusters, and the QTL-3d region itself harbors 78 piRNA clusters. Particularly, the major piRNA cluster, *42AB*, lies very close (~65kb distal) to the QTL-3d peak position. Although the RIL mothers do not produce or transmit *P*-element-derived piRNAs, the *D. melanogaster* genome harbors more than 100 distinct resident TE families (Kaminker et al. 2002; Quesneville et al. 2005) that are also regulated by piRNAs (Brennecke et al. 2007). If sensitive alleles of piRNA clusters within QTL-3d exhibit reduced silencing of resident (non *P*-element) TEs, resulting transposition could enhance genomic instability triggered by *P*-element activity. I therefore evaluated whether tolerant and sensitive alleles differed in the activity of piRNA clusters using small RNA-seq. A PCA of piRNA cluster expression revealed that sensitive and tolerant genotypes are resolved by the second principal component, accounting for 22% variation in expression (**Figure 2.3a**).

I did not find evidence that QTL-3d is explained by the differential activity of piRNA cluster *42AB*, as sensitive and tolerant genotypes exhibited comparable piRNA abundances from this locus. Similarly, piRNA abundance from other major piRNA clusters outside the QTL did not differ between sensitive and tolerant alleles (**Figure 2.3b**). However, I discovered two small pericentromeric piRNA clusters located within QTL-3d that were active in tolerant genotypes but largely quiescent in sensitive genotypes (**Figure 2.3b, c and d; Figure A2.2 and A2.3**). While these piRNA clusters contain no annotated TE insertions in the reference genome (dm6), Repbase Censor (Kohany et al. 2006) revealed that they were largely composed of TE fragments. The majority (~77%) of these TE fragments were relatively divergent from the consensus (0.65-0.95 sequence similarity; **Table A2.4 and A2.5**), and were often most similar to consensus TEs from other (non-*melanogaster*) *Drosophila* species. Given that transpositionally active TE families are often highly similar to the consensus sequence (Bergman and Bensasson

2007), and piRNA silencing is disrupted by mismatches between the piRNA and its target (Post et al. 2014), this suggests that the differential activity of these two piRNA clusters is unlikely to impact the expression of transpositionally active TEs.

To further evaluate if differences in tolerance are related to resident TE regulation, I compared resident TE expression between sensitive and tolerant genotypes in our RNA-seq data. None of the TE families represented in the QTL piRNA clusters were upregulated in sensitive genotypes (**Figure 2.3e**). Furthermore, while some TE families were differentially expressed, there was no systematic increase in TE activity in the sensitive genotypes. Rather, more TE families were upregulated in tolerant genotypes (13 TEs) when compared to sensitive (4 TEs) genotypes. Therefore, despite the conspicuous position of QTL-3d surrounding piRNA-producing regions, as well as evidence for differential heterochromatin formation that could impact piRNA biogenesis (**Figure 2.2b and e**), I found no evidence that tolerance is determined by resident TE silencing.

Figure 2.3: Tolerance is not determined by differential activity of piRNA cluster or TE deregulation. **a)** PCA analysis for piRNA cluster expression data of S/sensitive and T/ tolerant genotypes. The NIL pairs are represented by the same shapes. **b)** Heat map showing the expression of seven piRNA clusters. NIL pairs that are compared are plotted adjacent to each other. **c and d)** represent the piRNA expression between sensitive and tolerant genotypes from one of the NIL pairs along the two QTL piRNA clusters: 2L:23,328,000-23,337,026 and 2L:23,222,004-23,246,024, respectively. Only uniquely mapping piRNAs were considered. The TE families at the top of each figure are represented by different colors. TE-others represent the repeat families coming from sibling species of *D. melanogaster*. Positive value indicates piRNAs mapped to the sense strand of the reference genome and negative value indicates those from the antisense strand. See **Figure A2.2-2.3** for cluster expression in the remaining NIL pair. For **b, c, and d**, piRNA cluster expression levels were estimated by log2 scale transformed of reads per million mapped reads [$\log_2(\text{RPM}+1)$]. **e)** Bar graph depicting differentially expressed TEs (fold change = 1.5, base mean ≥ 100 , adjusted p -value ≤ 0.05) between sensitive and tolerant genotypes.



3.2.4 Identifying candidate genes influencing tolerance

In the absence of an obvious role for piRNA clusters within the QTL in determining tolerance, I sought to identify candidate genes that explain the associated phenotypes. I first identified “in-phase” single nucleotide polymorphisms (SNPs), where the genotypic differences

among the founder alleles are consistent with their tolerance phenotypes (Figure 4b, Long et al. 2014). I identified 64 and 258 genes with in-phase SNPs in QTL-21d and 3d, respectively. These polymorphisms potentially impact either gene expression—by residing within the regulatory/intron region—or affect the activity of the encoded protein through non-synonymous mutations (**Table A2.1, and A2.2**).

To further narrow down the candidates, I similarly identified differentially expressed genes with the QTL. Of 530 genes differentially expressed (**Figure 2.4a**), 43 were within the QTL, representing an approximately five-fold enrichment in the QTL regions compared to the rest of the genome (Pearson's Chi-squared test, $X\text{-squared} = 255.54$, $df = 1$, $p\text{-value} < 2.2e\text{-}16$, **Figure 2.4a**). Ultimately, I identified 5 and 14 differentially expressed genes that also carry in-phase SNPs within the QTL-21d and 3d, respectively (**Figure 2.4c and d; Table A2.1**). These genes, along with those carrying non-synonymous in-phase SNPs, make up the strongest candidate genes influencing tolerance (**Figure 2b; Table A2.2**).

I next scoured our list of candidate genes for those with known functions in heterochromatin formation and DNA-damage response, whose differential function or regulation are plausibly related to gene expression differences associated with sensitive and tolerant alleles. I similarly looked for genes with known functions in germ-cell maintenance or differentiation, which is a critical determinant of the dysgenic phenotype (Ma et al. 2017; Rojas-Ríos et al. 2017; Tasnim and Kelleher 2018). I only found two candidate genes: *brat* within QTL-21d and *Nipped-A* within QTL-3d, that have functions in determining germ-cell fate (Harris et al. 2011; McCarthy et al. 2018). Interestingly, *Nipped-A* is a member of the Tat-interacting Protein 60 kD (TIP60) complex, which has additional roles in DSB repair and heterochromatin formation (Sinclair et al. 1998; Ruhf et al. 2001; Qi et al. 2006; Hanai et al. 2008). Moreover, I found four other members and interactors of TIP60 complexes that were also upregulated in tolerant genotypes (*dRSF-1/CG8677*, *dom*, *E(Pc)* & *DMAP1*) (Kusch et al. 2004; Hanai et al. 2008), and one that was upregulated in sensitive (*yeti*) (Messina et al. 2014). Of

these, *yeti* and *dRSF-1* are located in QTL-3d.

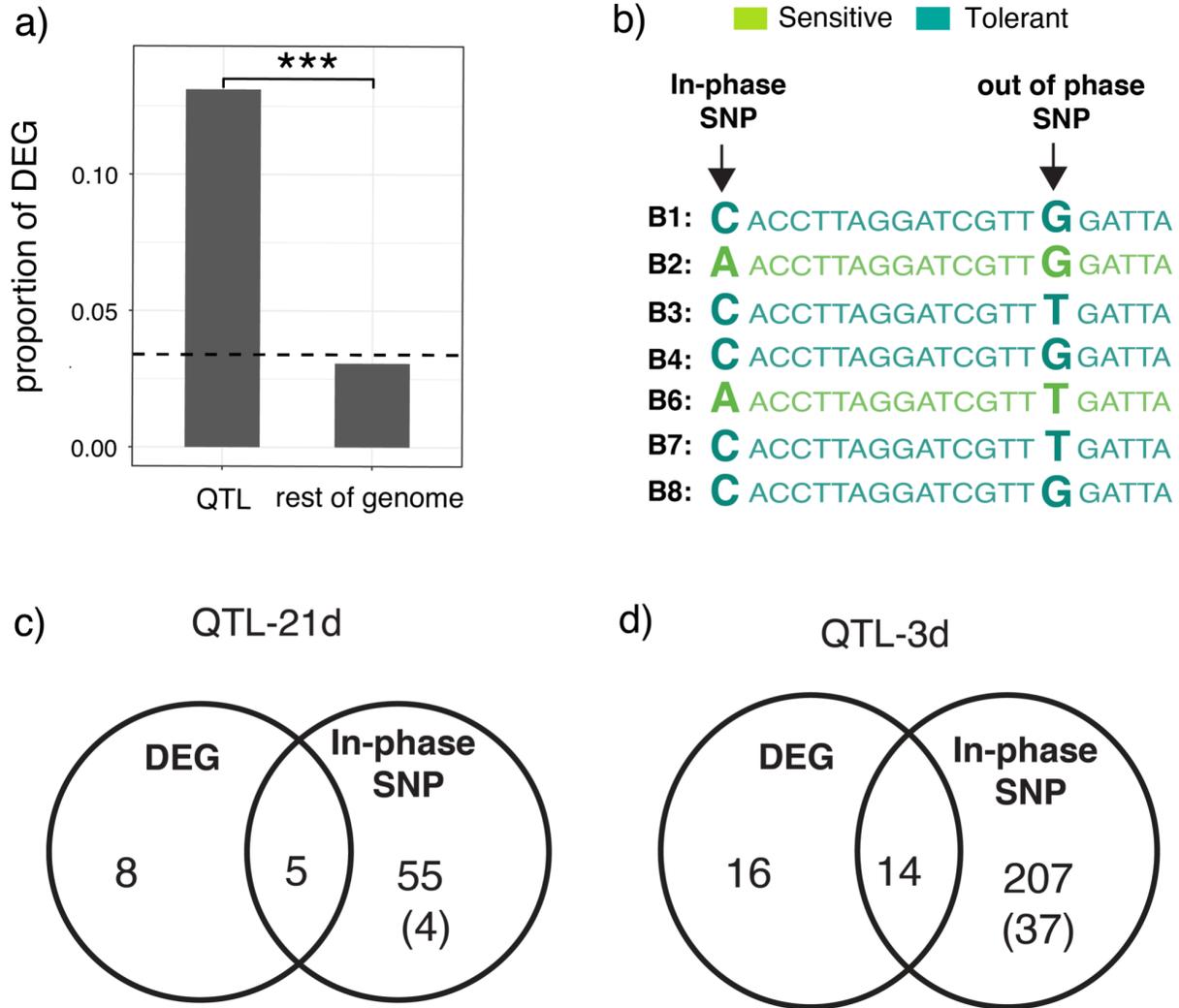


Figure 2.4: Differential expression and in-phase SNPs identify candidate genes that determine tolerance. **a)** Bar graph showing enrichment of differentially expressed genes in QTL. The dotted line is the genome wide average. **b)** Schematics representing the in-phase and out of phase SNPs, where each row represents the genotype of the eight B founder strains and the letters in bold indicate SNP alleles. The founders are colored based on their phenotypic classification, either tolerant or sensitive (**Figure 2.1e**). **c and d)** Venn diagram showing the overlap of differentially expressed genes (DEG) and genes carrying in-phase SNPs for QTL-21d and QTL-3d, respectively. The number within the brackets indicates the genes carrying non-synonymous in-phase SNPs.

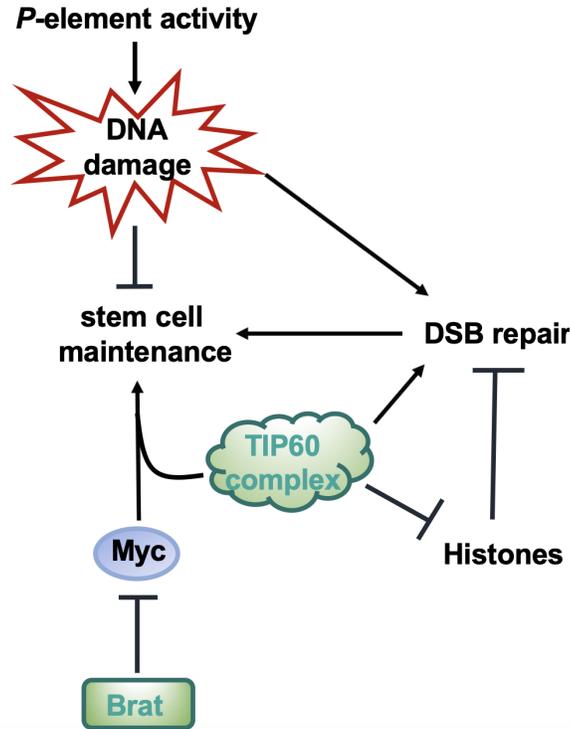


Figure 2.5: A model of TE tolerance in Population B RILs. *brat* and the TIP60 complex (containing Nipped-A) are proposed to determine TE tolerance through the modulation of Myc-dependent stem cell self-renewal or DSB repair (TIP60 only).

2.3 Discussion

Although small-RNA mediated TE regulation is widely studied, little is known about cellular and molecular mechanisms that confer tolerance to transposition. Here, I uncovered natural variation in tolerance to *P*-element DNA transposons, which is associated with two or more loci proximal to the second chromosome centromere in *D. melanogaster*. I further showed that tolerant and sensitive genotypes may differ in their ability to form heterochromatin and enact DNA repair, explaining their differential responses to *P*-element transposition. Finally, I identified candidate genes in each QTL that potentially determine the phenotypic differences between tolerant and sensitive alleles. *Nipped-A*, located in QTL-3d, is a member of the TIP60 complex and has a non-synonymous in-phase SNP that could alter the activity of encoded

protein. By contrast, *brat*, located in QTL-21d, has in-phase SNPs in its intronic and downstream regulatory regions, and is upregulated in the tolerant genotypes.

Nipped-A (QTL-3d) could influence tolerance by promoting the maintenance of larval PGCs or early adult GSCs, which are destabilized by DNA damage (Ma et al. 2016; Ota and Kobayashi 2020). *Nipped-A* is required for female germ-cell maintenance (Yan et al. 2014), as well as maintenance of larval neuroblasts, and adult intestinal and male germline stem cells (Prado et al. 2013; Tauc et al. 2017; Rust et al. 2018). While the functional consequences of the non-synonymous SNP that separates tolerant and sensitive *Nipped-A* alleles is not clear, the upregulation of four other TIP60 members (*dRSF-1*, *dom*, *E(Pc)* & *DMAP1*) suggests increased activity in the tolerant genotypes (**Table A2.3**). Reduced expression of pericentromeric genes in tolerant strains also suggests increased TIP60 activity, since TIP60 is involved in heterochromatin formation (Sinclair et al. 1998; Ruhf et al. 2001; Qi et al. 2006; Hanai et al. 2008).

While the specific function of TIP60 in female germ-cell maintenance is not clear, TIP60 is a conserved interactor of Myc: a transcription factor with diverse and well-studied roles in tumorigenesis, cell growth and proliferation, cell competition, and apoptosis (reviewed in Gallant 2013; Grifoni and Bellosa 2015). In larval neuroblasts, TIP60 and Myc coregulate downstream targets that promote stem cell self-renewal (Rust et al. 2018). Similarly, *myc* overexpression confers tolerance in dysgenic larval gonads by suppressing primordial germ cell (PGC) loss (Ota and Kobayashi 2020). Thus, increased TIP60 function in tolerant genotypes may activate *myc*-dependent tolerance in larvae.

Interestingly, *brat* (QTL 21d) is a translational repressor of *myc* that is upregulated in tolerant ovaries (**Figure A2.1**). Conversely, increased expression of translational machinery suggests increased Myc activity in sensitive ovaries as ribosomal proteins are conserved downstream targets of Myc (Figure 2B; Orian et al. 2003). Our data therefore, point to an association between reduced Myc activity and tolerance in adult stages. While puzzling, the

impact of Myc activity on cellular persistence is context and cell type specific. For example, reduced Myc activity confers robustness to X-ray-induced apoptosis in larval eye imaginal discs, while Myc overexpression in the same tissue induces apoptosis (Montero et al. 2008).

Therefore, the modulation of Myc function over the course of development may be a critical determinant of tolerance, with TIP60-dependent regulation of self-renewal factors increasing tolerance in PGCs, while other Myc targets may decrease tolerance in adults (**Figure 2.5**).

Exploring potential interactions between TIP60, Myc and Brat in determining tolerance presents an enticing avenue for future work.

In addition to promoting germ-cell maintenance, *Nipped-A* might also influence tolerance by facilitating repair of DSBs in PGCs or GSCs. The TIP60 complex has a conserved function in DSB repair (Kusch et al. 2004; Sun et al. 2009), and *Nipped-A* in particular promotes the proliferation of intestinal stem cells after DNA damage (Tauc et al. 2017). By contrast, histone upregulation in the sensitive genotypes—which potentially results from reduced TIP60-dependent heterochromatin formation—could inhibit DNA repair. Surplus RD histones are proposed to interfere with DNA-repair machinery, and are considered genotoxic outside of S-phase (Liang et al. 2012; Landais et al. 2014; Kumar et al. 2020). Enhanced repair in tolerant genotypes is further supported by the increased expression of chorion genes, since chorion gene amplification during oogenesis is dependent upon DSB repair (Alexander et al. 2015).

In summary, my work suggests that tolerance to transposition may have a complex architecture, including both the concurrent modulation of Myc-dependent stem cell self-renewal and stem cell loss, and the enhanced repair of DSBs. This complexity contrasts our previous study of natural variation in the population A RILs of the DSPR, which uncovered a major effect of the expression of a single differentiation factor, *bruno*, on tolerance (Kelleher et al. 2018). Furthermore, while DNA-damage signaling is a clear determinant of dysgenic germ-cell loss (Dorogova et al. 2017; Moon et al. 2018; Tasnim and Kelleher 2018), the potential for natural variation DNA repair to offset the mutagenic effects of transposition has never been evaluated.

My observations therefore point to multiple new mechanisms through which cells could withstand the genotoxic effects of unregulated transposition.

2.4 Methods

2.4.1 *Drosophila* strains and husbandry

The recombinant inbred lines (RILs) were generously provided by Dr. Stuart Macdonald. Harwich (#4264) was obtained from the Bloomington *Drosophila* stock center. All flies were maintained in standard cornmeal media.

2.4.2 Phenotyping

Phenotyping was performed as described previously in Kelleher et al. (2018). Briefly, crosses between virgin RIL females and Harwich males were transferred to fresh food every 3-5 days. Since crosses reared at a restrictive temperature (29 °C) result in complete gonadal atrophy in F1 offspring, crosses were reared at a lower permissive temperature (25 °C), which produces an intermediate phenotype that better reveals the variation in severity of dysgenesis (Kidwell et al. 1977; Dorogova et al. 2017; Srivastav and Kelleher 2017; Kelleher et al. 2018). F1 offspring were maintained for 3 days or 21 days, at which point their ovaries were examined using a squash prep (Srivastav and Kelleher 2017). 21 day- old females were transferred onto new food every five days as they aged to avoid bacterial growth. Females who produced 1 or more chorionated egg chambers were scored as having non-atrophied ovaries, and females producing 0 egg chambers were scored as having atrophied ovaries.

Crosses and phenotyping were performed for 673 RILs across 22 experimental blocks for 3 day-old F1 females, and 552 RILs across 18 experimental blocks for 21 day-old F1 females. If fewer than 21 F1 offspring were phenotyped for the same cross, it was discarded

and repeated if possible. In total, our lab phenotyped >20 3-day old and 21 day-old F1 female offspring for 595 RILs and 456 RILs, respectively.

2.4.3 QTL mapping

QTL mapping was performed as described in Kelleher *et al.* (2018). Briefly, for each developmental time point, the arcsine transformed proportion of F1 ovarian atrophy was modeled as a function of two random effects: experimental block and undergraduate experimenter. Regression models were fit using the lmer function from the lme4 package (Bates *et al.* 2015). The residuals were used as a response for QTL mapping using the DSPRqtl package (King *et al.* 2012) in R 3.02 (Team and TRDC 2008). The LOD significance threshold was determined from 1,000 permutations of the observed data, and the confidence interval around each LOD peak was identified by a difference of -2 from the LOD peak position ($\Delta 2$ -LOD), or from the Bayes Confidence Interval (Manichaikul *et al.* 2006). For $\Delta 2$ -LOD intervals, the conservative approach of determining the longest contiguous interval was taken where the LOD score was within 2 of the peak value. Furthermore, the broad sense heritability of ovarian atrophy was calculated as in Kelleher *et al.* (2018).

2.4.4 Estimation of founder phenotypes and QTL phasing

To estimate the phenotypic effect associated with each founder allele at the QTL peak, the distribution of phenotypes from all RILs carrying the founder haplotype at the LOD peak position (genotype probability >0.95%) was considered (King *et al.* 2012; Kelleher *et al.* 2018). QTL were then phased into allelic classes by identifying the minimal number of partitions of founder haplotypes that describe phenotypic variation associated with the QTL peak, as described previously (King *et al.* 2012).

2.4.5 Identification of in-phase polymorphisms

The SNP data of B founders that used to infer in-phase SNPs is based on dm3 (King et al. 2012). To identify in-phase SNPs, I looked for alternate SNP alleles that matched the predicted phenotypic class for each of the QTL peaks. For QTL-21d, I used the criteria: sensitive class (B2, B6) and the tolerant class (B1, B3, B4, B7, B8), whereas for QTL-3d: sensitive class (B6) and the tolerant class (B1, B2, B3, B4, B7, B8).

2.4.6 Selection of paired RILs with alternate QTL alleles

I identified background matched RILs containing either the B6 (“sensitive”) or B4 (“tolerant”) haplotypes from the start position of the QTL-21d confidence interval (2L: 19,010,000) to the end position of QTL-3d confidence interval (2R: 6,942,495) ($P > 0.9$), based on their published HMM genotypes (King et al. 2012). For all possible RIL pairs (B6 and B4), I then calculated the number of 10 Kb genomic windows in which they carried the same RIL haplotype ($P < 0.9$). I selected three pairs of RILs, which carried the same founder genotype for 47% (21213 & 21183), 46% (21147 & 21346) and 44% (21291 & 21188) of genomic windows outside of the QTL.

2.4.7 Small RNA-seq and total RNA-seq

RILs were maintained at 25 °C, and three biological replicates of 20 ovaries were dissected from 3-5 day old females. Ovaries were homogenized in TRIzol and stored at -80 °C until RNA extraction. 50 µg of total RNA from each of 18 biological samples (three biological replicates x three pairs) was size fractionated in a 15% denaturing polyacrylamide gel and the 18-30 nt band was excised. 2S-depleted small RNA libraries for Illumina sequencing were then constructed according to the method of Wickersheim and Blumenstiel (2013). Ovarian small

RNA libraries were published previously (SRP160954, Zhang and Kelleher 2019). Ribodepleted and stranded total RNA libraries were generated from the same ovarian samples using NuGen total RNA kit (TECAN). All 18 small RNA and total RNA libraries were sequenced on an Illumina Nextseq 500 at the University of Houston Seq-N-Edit Core.

2.4.8 Small-RNA analysis

Sequenced small RNAs were separated based on size into miRNAs/siRNAs (18-22 nt) and piRNAs (23-30 nt) (Brennecke et al. 2008). Reads corresponding to contaminating rRNAs, including 2S-rRNA, were removed from each library by aligning to annotated transcripts from flybase (Gramates et al. 2017). To determine the piRNA cluster activity, I first uniquely aligned the piRNAs to reference genome (dm6) using Bowtie1 (-v 1 -m 1) (Langmead et al. 2009). I then used a customized perl script to count reads that mapped to a set of previously annotated piRNA clusters from the same genotypes (497 piRNA clusters, S. Zhang et al. 2020). Read counts normalized to total mapped microRNAs for each library were used to infer differential expression using DESeq2 (Love et al. 2014). Sliding window estimates of piRNA abundance were calculated using bedtools genomecov (Quinlan 2014), normalizing the read counts to total mapped miRNA reads.

2.4.9 Total RNA analysis

Residual ribosomal RNAs (rRNAs) were identified in ribodepleted libraries based on alignment to annotated rRNAs from flybase (Gramates et al. 2017), and excluded from further analysis. Retained reads aligned to the library of consensus satellite and TE sequences from repbase (Bao et al. 2015), plus additional satellite consensus sequences from Larracuente (2014). For TE expression, the total reads mapped to TE sequences were counted using awk commands. Remaining reads that failed to map were aligned to *D. melanogaster* transcriptome

(dm6/BDGP6) using Kallisto with default parameters (Bray et al. 2016). Differentially expressed TEs and genes were identified from a combined analysis in DESeq2 (Love et al. 2014). Genes and TEs with base mean ≥ 100 , Adjusted p -value ≤ 0.05 and whose expression pattern differed (fold change ≥ 1.5) were considered differentially expressed between the B6 and B4 QTL haplotype.

Chapter 3. Using isogenic lines to investigate the mechanism of tolerance

3.1 Introduction

In Chapter 2, I found two natural QTL alleles that differed in tolerance to *P*-element activity using a panel of recombinant inbred lines. However, how these QTL alleles influence tolerance is unclear. Here, I test multiple hypotheses for the mechanistic underpinning of tolerance: **1)** differential silencing of satellite repeats, **2)** differential *de novo* piRNA production, **3)** enhanced DSB repair, and **4)** *brat* function.

DNA damage arising from transposition is proposed to trigger germ-cell loss in the dysgenic germline (Khurana et al. 2011). Therefore, exposure of germ cells to additional agents of DNA damage other than transposition might reduce tolerance. QTL-3d residing in the 2nd chromosome centromere harbors known blocks of satellite repeats- *Responder* (*Rsp*) and *260-bp*. Deregulation of satellite repeats has previously been associated with DNA damage and segregation defects (Ferree and Barbash 2009; Aldrup-MacDonald et al. 2016; Kishikawa et al. 2016; Ichida et al. 2018; Kishikawa et al. 2018). Particularly, *Rsp* mispackaging is proposed to establish a well-known meiotic drive system in the male germline known as Segregation Distorter (SD) (Larracuenta 2012). SD is known to target the *Rsp* block, and spermatids with large *Rsp* blocks become dysfunctional. Additionally, *260-bp* is a member of 1.688 satellite repeats, whose deregulation is proposed to cause hybrid lethality (Ferree and Barbash 2009). Thus, regulation of satellite repeats may be an important determinant of germline tolerance to TEs. I therefore tested whether sensitive genotypes experience deregulation of *260-bp* or *Rsp* repeats.

In Chapter 2, I uncovered that tolerant and sensitive alleles may differ in TIP60 activity, which functions in DSB repair (Sinclair et al. 1998; Ruhf et al. 2001; Qi et al. 2006; Hanai et al. 2008). Since hybrid dysgenesis is characterized by DNA damage (Moon et al. 2018; Tasnim and Kelleher 2018), tolerance could also arise from enhanced repair of DSBs induced by *P*-elements. I therefore tested the hypothesis that tolerance is determined by increased

efficiency of DSB repair.

TIP60 also has a function in heterochromatin formation, and tolerant genotypes exhibit signatures of enhanced heterochromatin formation. Heterochromatin formation is critical for the transcription of piRNA clusters (Le Thomas et al. 2014; Mohn et al. 2014). While I already uncovered that tolerant and sensitive alleles do not differ in regulation of resident TEs in Chapter 2, it is unclear if there is a difference in regulation of *P*-elements through *de novo* piRNA transcription. *De novo* piRNA transcription of paternally inherited piRNA clusters was reported to reduce hybrid dysgenesis with age (Khurana et al. 2011; Moon et al. 2018). Therefore, I examined if tolerance is determined by *de novo* piRNA transcription.

The loss of GSCs and PGCs during hybrid dysgenesis is thought to be caused by disruption of germline stem cell maintenance (Ma et al. 2017). Promoting renewal or retention of GSCs in the dysgenic germline is therefore one potential mechanism of tolerance. For example, overexpression of GSC self-renewal factor, *myc* (Ota and Kobayashi 2020) and reduced expression of germline differentiation factor *bruno* suppresses hybrid dysgenesis (Kelleher et al. 2018). However, contrary to previous studies, I found tolerant alleles are associated with increased expression of a GSC differentiation factor, *brat*, a candidate gene residing within QTL-21d. Therefore, it is puzzling how increased differentiation of GSCs in our genotypes could confer tolerance.

To unravel the mechanisms behind tolerance, I tested these non-mutually exclusive hypotheses using isogenic lines carrying tolerant and sensitive QTL alleles. I found that tolerant genotypes were highly robust to X-ray-induced DNA damage, suggesting that tolerance may be conferred by enhanced DNA-damage repair. Furthermore, loss-of-function mutations in *brat* showed increased ovarian atrophy in dysgenic females, suggesting that tolerance may be conferred by promoting differentiation of GSCs. Hence, our observations point towards DSB repair and *brat* function as two potential mechanisms of tolerance.

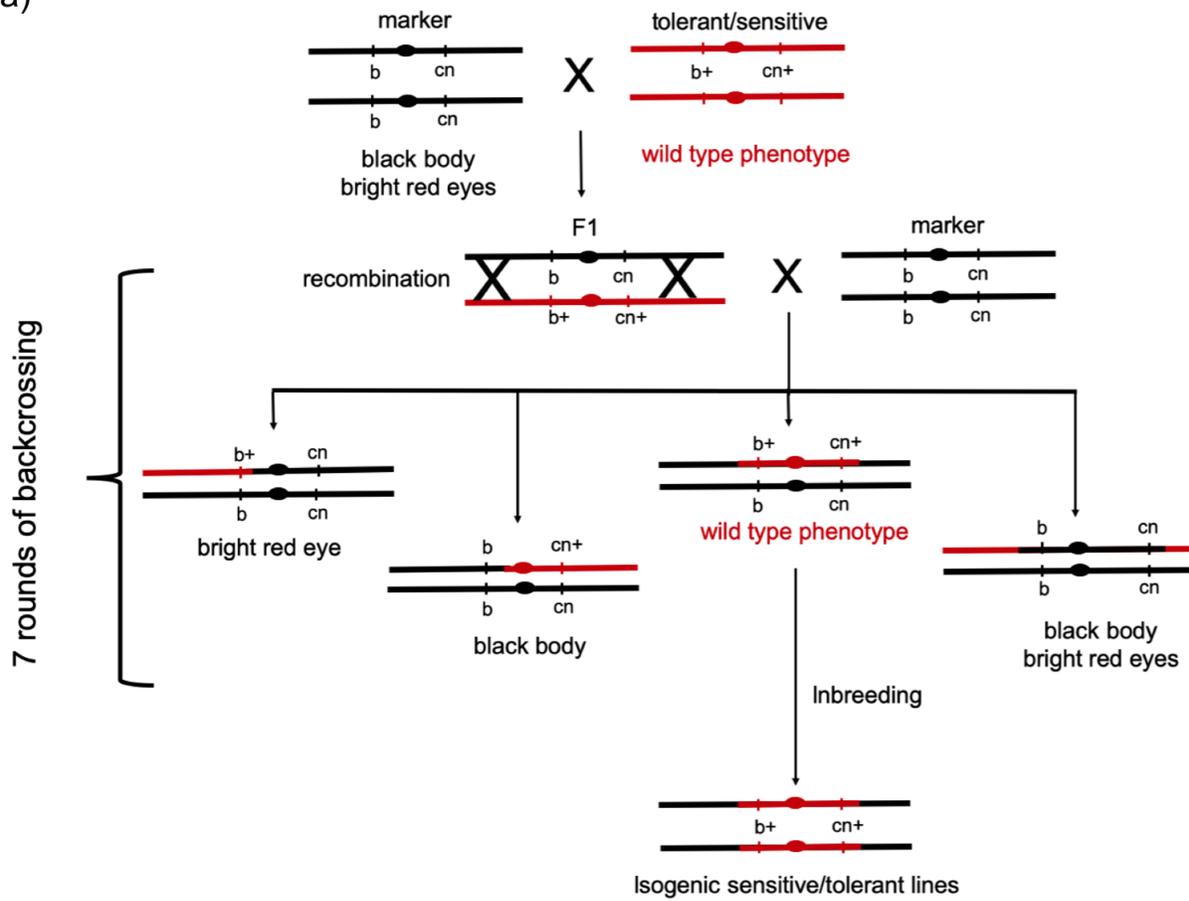
3.2 Results

3.2.1 Generation of isogenic lines

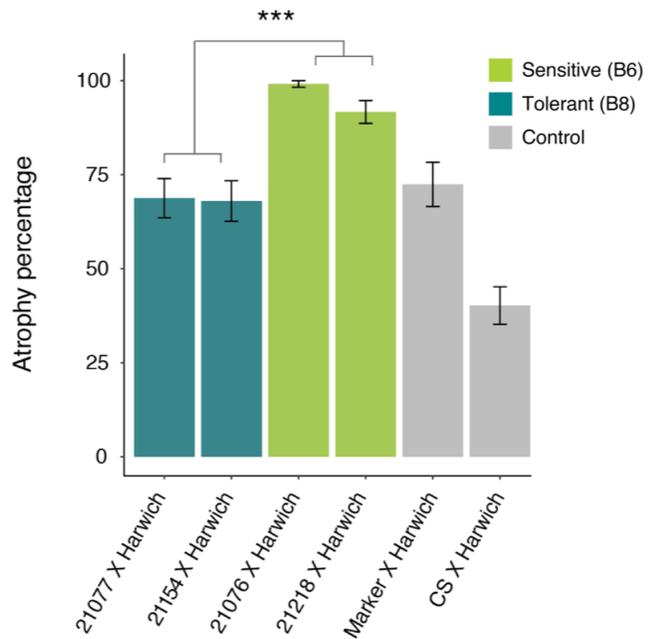
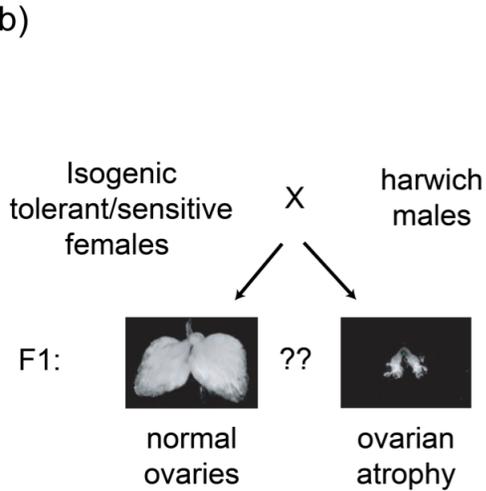
To investigate the mechanisms of tolerance, I generated isogenic lines carrying sensitive or tolerant QTL alleles in an otherwise similar genetic background (**Figure 3.1a**). First, I crossed the tolerant (21077 and 21154) and sensitive (21076 and 21218) strains to a marker strain carrying recessive mutations in black body and cinnabar eyes, which lie at the boundary of the two QTL (**Figure 3.1a**). The F1 was then subjected to seven rounds of backcrossing, each time selecting for wild type phenotype. Finally, I performed inbreeding to generate the homozygous isogenic lines. These lines were verified by PCR and sequencing of the regions within the two QTL. I further confirmed the phenotype of the isogenic lines by mating them to *P*-element carrying males (Harwich) and screening for F1 atrophy in 3 day-old female offspring. Consistent with the previous QTL analysis (**Fig 2.1**), tolerant alleles exhibited 30 percent less F1 atrophy than the sensitive alleles (**Figure 3.1a**).

Figure 3.1. Isogenic line generation and phenotyping. a) Crossing scheme for generating isogenic lines. b) Phenotyping isogenic lines carrying tolerant and sensitive QTL alleles by crossing with Harwich males and screening for ovarian atrophy of F1 females. Bar graph showing the percentage of ovarian atrophy in dysgenic F1 from tolerant (dark green) and sensitive (light green) isogenic lines as well as the control strains (gray).

a)



b)



3.2.2 Investigating the role of satellite repeats on tolerance

Incomplete silencing of satellite repeats could impact tolerance to TEs. Aberrant expression of satellite repeats have been associated with genomic instability, segregation disorders and DNA damage (Ferree and Barbash 2009; Aldrup-MacDonald et al. 2016; Kishikawa et al. 2016; Ichida et al. 2018; Kishikawa et al. 2018). The QTL-3d harbors two unique satellite repeats namely, *Responder* (*Rsp*) and a 260-bp, both of which are predominantly found proximal to the 2nd chromosome centromere. Maternally transmitted piRNAs have been implicated in regulating satellite arrays such as the 359 bp (Usakin et al. 2007; Yuan and O'Farrell 2016) and *Rsp* through heterochromatin formation (Wei et al. 2021). Reduced maternal transmission of satellite-derived piRNAs and siRNAs from sensitive genotypes could result in inefficient packaging of paternally inherited satellite arrays, leading to additional genotoxic stress that enhances ovarian atrophy.

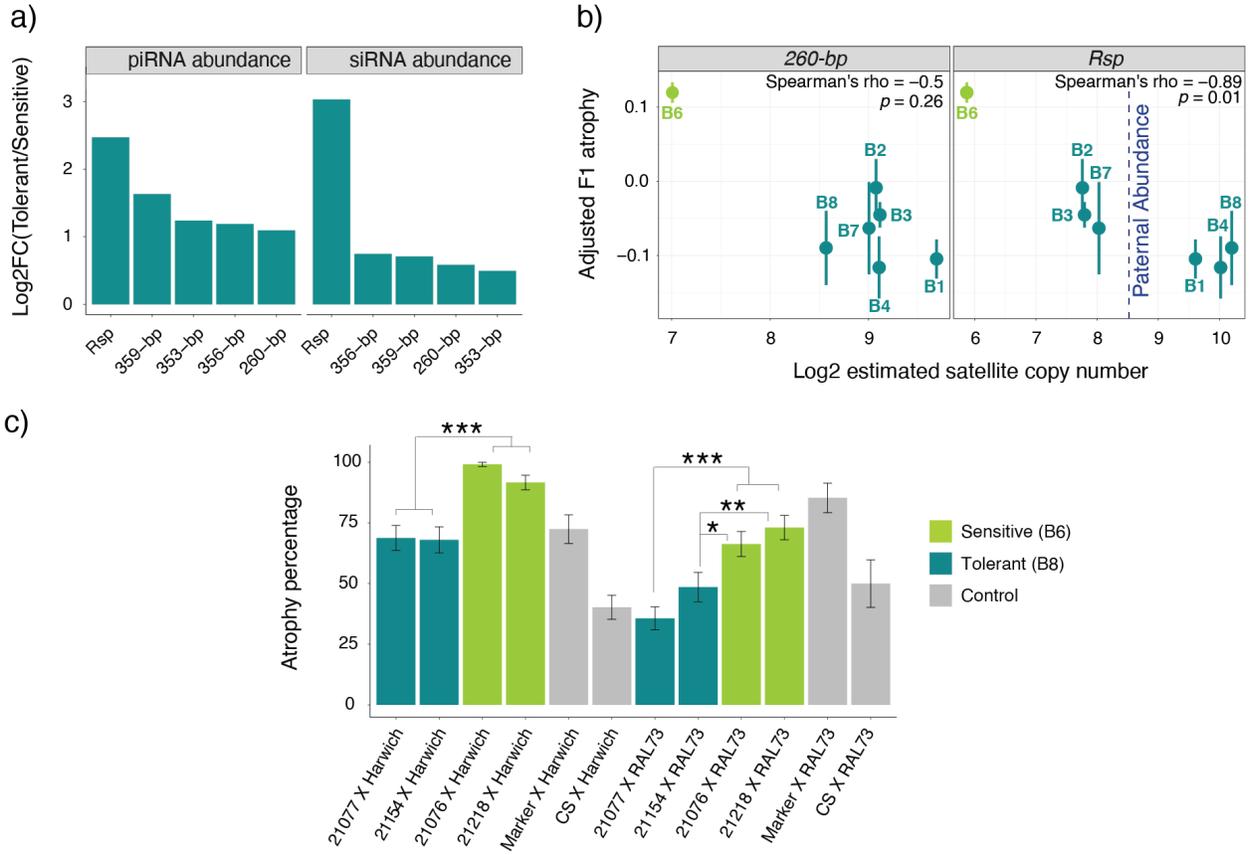
To test this hypothesis, I first asked if small RNAs derived from satellite repeats are less abundant in the sensitive nearly isogenic lines (NILs) that I examined in Chapter 2. *Rsp*-derived small RNAs were 5-fold and 8-fold reduced among sensitive piRNAs and siRNAs, respectively (**Figure 3.2a**). Similarly, 260-bp derived piRNAs and siRNA were 2-fold and 1.5-fold reduced among sensitive alleles. This suggests that sensitive and tolerant NILs may differ in their ability to regulate satellite repeats at 2nd chromosome centromere.

Differences in satellite-derived piRNA abundance may arise from differences in satellite copy number. Particularly, the *Responder* (*Rsp*) satellite array is known to differ dramatically in copy number between strains (20-2000 copies) (Lyttle 1991, Larracuente & Presgraves 2012). Differential satellite copy number could also explain the observed differences in small RNA abundance between sensitive and tolerant alleles. Therefore, I next compared satellite copy numbers between the tolerant and sensitive alleles. I discovered that sensitive (B6) alleles carry very few *Rsp* repeats (~60 copies) and 260-bp (~130) when compared to tolerant (B4) alleles

(~1038 *Rsp* copies and ~826 260-bp copies). I further found a correlation between *Rsp* dosage and *P*-tolerance (**Figure 3.2b**), which was lacking in the case of 260-bp. The correlation for the *Rsp* satellite was largely driven by the highly sensitive B6 that has the fewest *Rsp* copies among the founders. These observations suggest that packaging of *Rsp* but not 260-bp may be an important determinant of tolerance.

Figure 3.2: Responder copy number is associated with but does not determine tolerance.

a) Bar graph showing satellite derived piRNA and siRNA abundance between sensitive and tolerant NILs. The x-axis represents the log₂ fold change between tolerant and sensitive alleles and the y-axis represents the different satellite repeats. **b)** Correlation plot between F1 atrophy associated with B founder QTL and log₂ estimated *Responder* and 260-bp copy number at their second chromosome centromere. **c)** Graph showing F1 atrophy percentage in the dysgenic crosses between tolerant, sensitive and control females with Harwich/RAL73 males. Harwich and RAL73 are P-strain carrying a short and long *Rsp* array, respectively. Tolerant crosses show significant reduction in F1 atrophy compared to sensitive crosses with both Harwich (21077 vs. 21076 cross: Pearson's Chi-square test, X-squared = 37.05, df = 1, *p*-value = 1.15e-09; 21077 vs. 21218 cross: X-squared = 13.7, df = 1, *p*-value = 0.0002; 21154 vs. 21076 cross: X-squared = 37.85, df = 1, *p*-value = 7.625e-10; 21154 vs. 21218 cross: X-squared = 14.14, df = 1, *p*-value = 0.0001) and RAL73 (21077 vs. 21076 cross: Pearson's Chi-square test, X-squared = 17.09, df = 1, *p*-value = 3.57e-05; 21077 vs. 21218 cross: Pearson's Chi-square test, X-squared = 24.7, df = 1, *p*-value = 6.68e-07; 21154 vs. 21076 cross: X-squared = 4.83, df = 1, *p*-value = 0.02; 21154 vs. 21218 cross: X-squared = 9.26, df = 1, *p*-value = 0.0023) males. Marker and CS are the controls. Tolerant and sensitive QTL alleles were initially extracted into the marker background.



3.2.3 Investigating the role of *de novo* piRNA production on tolerance

Tolerance may be determined by enhanced production of *P*-element derived piRNAs from paternally derived piRNA clusters in the dysgenic germline. Some dysgenic genotypes recover fertility with age by transcribing piRNAs from paternally inherited piRNA clusters, thereby silencing *P*-elements in the germline (Khurana et al. 2011; Moon et al. 2018). Since, piRNA transcription relies on the heterochromatic mark H3K9me3 (Thomas et al. 2014; Mohn et al. 2014), differences in heterochromatin formation could impact *de novo* production of piRNAs. In Chapter 2, I found tolerant genotypes were associated with increased activity of the TIP60 chromatin remodeling complex, which is involved in heterochromatin formation (Sinclair et al. 1998; Ruhf et al. 2001; Qi et al. 2006; Hanai et al. 2008). Increased heterochromatin could facilitate *de novo* production of *P*-element piRNAs in the dysgenic germline, thus conferring

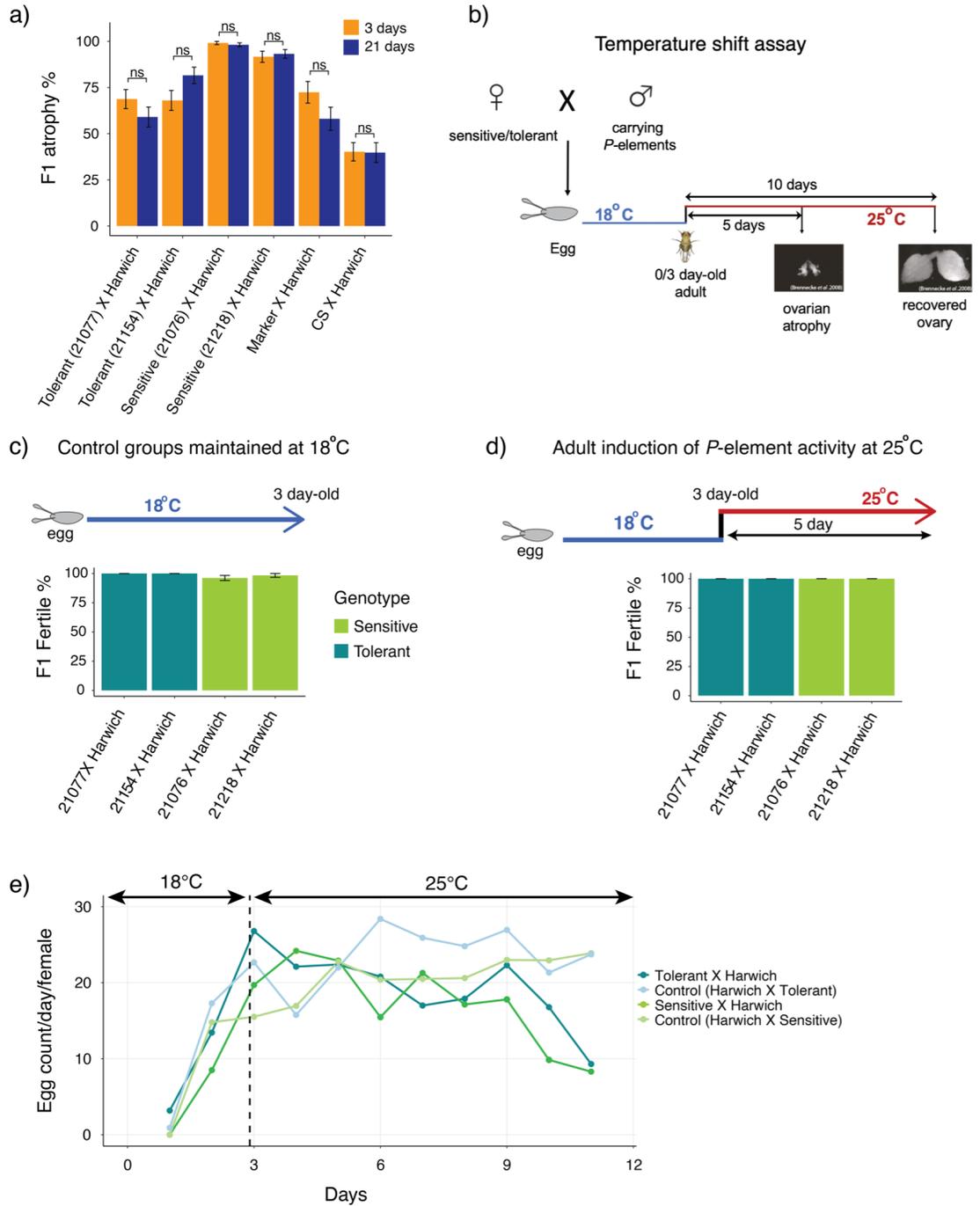
tolerance.

To test this hypothesis, I first used age-dependent recovery of fertility in dysgenic females as a proxy for *de novo* piRNA production. I asked whether the tolerant and sensitive alleles show differences in age-dependent recovery of F1 fertility when mated to Harwich males at 25 °C. If the tolerant genotypes exhibit a greater recovery of fertility from 3 to 21 days relative to sensitive genotypes, it would suggest *de novo* piRNA production as a potential source of tolerance. However, I did not detect fertility recovery in either genotype as there were no significant differences in F1 atrophy between ages (**Figure 3.3a**). This is consistent with our observation that there is a minimal effect of age in Chapter 2. Our observation may either suggest the absence of *de novo* piRNA production in either genotype, or the initiation of *de novo* piRNA production occurs at earlier developmental stages.

Next, I sought to address whether I missed the early acquisition of *de novo* piRNAs in tolerant genotypes. Since examining piRNA production in larval ovaries is technologically challenging, I used an alternative approach that induces dysgenesis in adult stages through temperature shift (Moon et al. 2018). I first maintained the dysgenic offspring (tolerant/sensitive X Harwich) at 18 °C, where *P*-element activity is lower and does not induce ovarian atrophy (Kidwell et al. 1977; Moon et al. 2018). I then transferred newly emerged and 3 day-old adults to 25 °C or 29 °C, where *P*-element activity is higher (Engels and Preston 1979). Using the dysgenic offspring of crosses between *w1* and Harwich, Moon et al. (2018) reported that this treatment results in germ-cell loss and ovarian atrophy at 5 days post-induction and resumption of fertility at 10 days. They further demonstrated that recovery of fertility was associated with *de novo* piRNA production. Therefore, I assayed for ovarian atrophy at 5 days and 10 days post temperature-shift to measure fertility recovery. If *de novo* piRNA production determines tolerance, it is predicted that tolerant genotypes will show higher incidences of fertility recovery post induction.

Unexpectedly, I did not detect any F1 atrophy on the 5th day of *P*-element induction in either genotype (**Figure 3.3 c and d**). Therefore, I was not able to compare piRNA-dependent recovery of fertility between genotypes. Furthermore, I did not detect any differences in the number of the eggs laid by dysgenic daughters of sensitive and tolerant mothers to that of non-dysgenic controls (**Figure 3.3 e**). This suggests that differences between sensitive and tolerant genotypes most likely arise at larval stages. Hence, these results, together with the absence of age-dependent recovery of fertility, make it experimentally intractable to compare *de novo* piRNA production between sensitive and tolerant strains.

Figure 3.3. Adult induction of *P*-element activity is not sufficient to trigger hybrid dysgenesis. **a)** The percentage of ovarian atrophy in 3 day and 21 day-old dysgenic daughters from tolerant and sensitive mothers at 25 °C. The x-axis shows the dysgenic crosses and the y-axis is the F1 ovarian atrophy. **b)** Schematics depicting the steps in the temperature shift assay. The crosses were set at 18 °C till the F1 offspring were 0 or 3 day-old adults. The individuals were then transferred to 25 °C for 5 and 10 days and screened for ovarian atrophy. **c) and d)** Graph showing the percentage fertility **c)** for 3 day-old control daughters raised at 18 °C and **d)** for F1 daughters after 5-day post-induction at 25 °C. The x-axis shows the tolerant and sensitive dysgenic crosses. **e)** Dot plot showing the egg count per day per female of the dysgenic and non-dysgenic F1 daughters at 18 and 25 °C. X-axis shows the no. of days and y-axis shows the total number of eggs laid by a female per day. The colors indicate the tolerant (21077) and sensitive (21076) dysgenic and non-dysgenic (control) crosses.



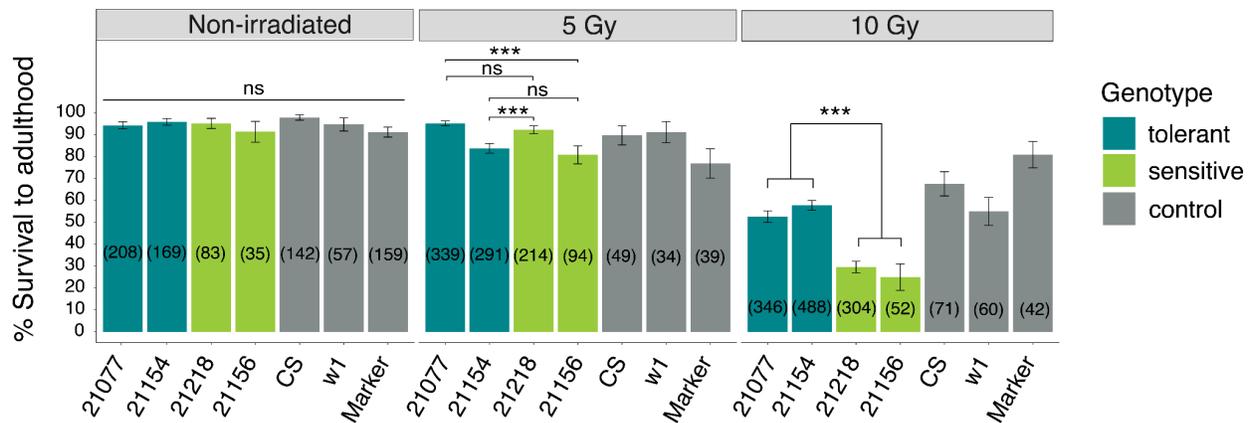
3.2.4 Investigating association of irradiation sensitivity with tolerant and sensitive alleles

My observations from Chapter 2 suggest that sensitive and tolerant alleles may differ in their capacity to repair DSBs. I found that tolerant genotypes display increased expression of

TIP60 components, which is involved in DSB repair (Kusch et al. 2004; Sun et al. 2009) as well as elevated chorion gene expression, a potential indicator of efficient DSB repair. If tolerance is determined by enhanced DSB repair, it is predicted that tolerant genotypes will show higher resilience to radiation. I therefore compared the sensitivity of the two genotypes to X-ray radiation.

I first sought to identify the appropriate radiation dose to reveal the phenotypic difference between the two genotypes. I irradiated 3rd instar larvae at different X-ray doses (5 to 80 Gy) and observed their survival to adulthood for 10 days. I found that doses above 10 Gy showed high lethality, making it difficult to detect differences in radiation sensitivity between the genotypes (**Table A3.1 and A3.2**). Therefore, I compared the response of sensitive and tolerant larvae to radiation doses of 5 Gy and 10 Gy. I observed only modest reduction in larval survival after 5 Gy radiation (76-95% survival to adulthood), with minimal differences between genotypes. By contrast at 10 Gy, considerable rates of lethality were observed for all genotypes (25-80% survival to adulthood). Furthermore, sensitive genotypes had significantly higher lethality than the tolerant genotypes. These results are consistent with differences between sensitive and tolerant alleles in DSB repair.

Figure 3.4. Tolerance is associated with enhanced DNA-damage repair. Bar graph showing the percentage of mock treated and irradiated (5 Gy and 10 Gy) larvae that survived to adulthood for the tolerant, sensitive and the control genotypes. The X-axis represents the different strains with the colors representing the type of genotype. The Y-axis is the percentage of irradiated larvae that survived to adulthood. The numbers in the brackets refer to the sample size. For 5 Gray irradiation, 21077 vs. 21156: Pearson's Chi-squared test, $X\text{-squared} = 15.66$, $df=1$, $p\text{-value} = 0.0008$; 21154 vs. 21218: Pearson's Chi-squared test, $X\text{-squared} = 9.56$, $df=1$, $p\text{-value} = 0.001$. For 10 Gray irradiation, 21077 vs. 21218: Pearson's Chi-squared test, $X\text{-squared} = 34.23$, $df=1$, $p\text{-value} = 0.0001$; 21077 vs. 21156: Pearson's Chi-squared test, $X\text{-squared} = 12.69$, $df=1$, $p\text{-value} = 0.0004$; 21154 vs. 21218: Pearson's Chi-squared test, $X\text{-squared} = 58.6$, $df=1$, $p\text{-value} = 0.0001$; 21154 vs. 21156: Pearson's Chi-squared test, $X\text{-squared} = 19.08$, $df=1$, $p\text{-value} = 0.0001$).



3.2.5 Investigating the role of *brat* on tolerance

To determine the impact of *brat* on tolerance, I looked at the effect of *brat* loss-of-function mutation on hybrid dysgenesis. The causative variant influencing tolerance in the dysgenic hybrid offspring is most likely heterozygous, as it is only present in the maternal genotypes. Therefore, to look at the heterozygous effect of *brat* mutants, I mated the *brat*¹/CyO females to *Harwich* males and compared the incidences of ovarian atrophy among the 3-4 day-old F1 offspring (mutant -/+ vs balancer CyO/+).

In absence of dysgenesis, *brat* loss-of-function alleles impact oogenesis recessively. However, I found that the *brat*¹ heterozygotes showed significantly higher frequency of ovarian atrophy (68.6%) than their balancer control siblings (37.5%) (**Figure 6**). To further prove that this phenotype was not an effect of the 2nd chromosome of the *brat*¹ mutant line, I additionally used three deficiency stocks with deletions overlapping *brat*. I found that two out of three deficiency lines increased ovarian atrophy similar to *brat*¹ mutants (**Figure 6**). The deficiency line (Df(2L)*brat* [ED1231]) with a larger deletion (39 genes) showed modest increase in ovarian atrophy. Whereas another deficiency line (Df(2L)*brat* [ED1200]) also carrying large deletion (31 affected genes) showed no change in the incidences of ovarian atrophy compared to its balancer siblings, which may be due to deletion of genes with opposing function to that of *brat*.

Hence, our results suggest that *brat* activity increases fertility in dysgenic females, which is consistent with our previous observation that tolerant genotypes are associated with increased *brat* expression.

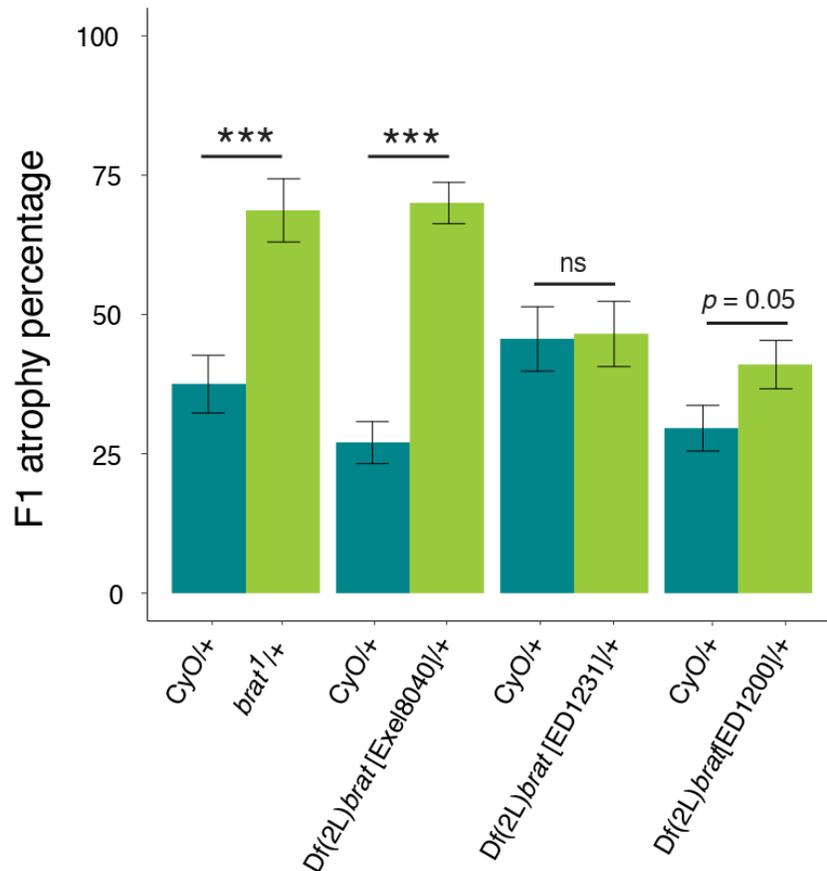


Figure 3.5. Loss-of-function mutation of *brat* increases severity of hybrid dysgenesis.

Graph showing the percentage of F1 ovarian atrophy of control balancer siblings CyO/+, heterozygous *brat*¹ mutants and heterozygous deficiency lines Df(2L)*brat*. *brat*¹ mutant: Pearson's Chi-squared test, $X\text{-squared} = 13.55$, $df=1$, $p\text{-value} = 0.0002$. Df(2L)*brat* [Exel8040]: Pearson's Chi-squared test, $X\text{-squared} = 14.78$, $df=1$, $p\text{-value} = 0.0001$. Df(2L)*brat* [ED1231]: Pearson's Chi-squared test, $X\text{-squared} = 0.06$, $df=1$, $p\text{-value} = 0.8$. Df(2L)*brat* [ED1200]: Pearson's Chi-squared test, $X\text{-squared} = 3.66$, $df=1$, $p\text{-value} = 0.05$.

3.3 Discussion

In Chapter 2, I uncovered two strong candidate genes as well as the phenotypic differences associated with tolerance to *P*-elements. However, the cellular and molecular mechanisms of

tolerance remained unclear. Here, through the use of isogenic lines carrying tolerant and sensitive alleles, I found evidence of two potential mechanisms of tolerance: **1)** through enhanced robustness to DNA damage, and **2)** through increased expression of GSC differentiation factor *brat*.

Since DNA damage underlies hybrid dysgenesis (Ma et al. 2016; Ota and Kobayashi 2020), facilitating DSB repair could minimize the impact of *P*-element activity and confer tolerance. Multiple pieces of evidence from Chapter 2 suggest enhanced DSB repair in tolerant genotypes. Tolerant alleles showed increased expression of TIP60 members, which are involved in DSB repair (Kusch et al. 2004; Sun et al. 2009) as well as show upregulation of chorion gene expression- an indicator of efficient DSB repair (Alexander et al. 2015). Sensitive genotypes also exhibited histone overexpression, which has been associated with DNA damage and therefore could be a possible determinant of the variable DSB repair efficiency. Here, I demonstrated that tolerant larvae are highly resilient to X-ray-induced DNA damage (**Figure 3.4**), which is widely associated with increased activity of DNA repair genes (Staeva-Vieira et al. 2003; Ruike et al. 2006; Uri et al. 2007; Sterpone and Cozzi 2010; Koval et al. 2019). Together these observations suggest that tolerance may be determined by enhanced repair of DSBs induced by *P*-elements.

Recent studies also suggest tolerance could be conferred by promoting GSC maintenance. Ectopic overexpression of GSC self-renewal factor *myc* suppresses hybrid dysgenesis by increased retention of damaged PGCs in larval gonads (Ota and Kobayashi 2020). Similarly, reduced function of the germline differentiation factor *bruno* confers tolerance, potentially through stabilizing damaged GSCs by reducing signals for differentiation (Kelleher et al. 2018). Similar to *bruno*, *brat* is a differentiation factor, which could modulate the response of GSCs and their differentiating daughter to DNA damage (Harris et al. 2011; McCarthy et al. 2018). However, unlike *bruno*, I observed that the *brat* function promotes tolerance. This opposite phenotypic effect might be explained by their activity at distinct developmental stages

and cellular pathways. *bruno* promotes the differentiation of four cell cysts (Parisi et al. 2001; Wang and Lin 2007), whereas *brat* act earlier by regulating single cell cystoblast differentiation through the Bam/Bgcn pathway (Harris et al. 2011; McCarthy et al. 2018). DNA damage in the germline is known to reduce the expression of *bam* and delay cystoblast differentiation (Ma et al. 2016). Therefore, *brat* could confer tolerance by compensating for damage-induced *bam* repression (Harris et al. 2011; McCarthy et al. 2018).

My data highlights the possibility of DSB repair and GSC differentiation as two novel mechanisms of tolerance to *P*-element transposition. While natural variation in irradiation sensitivity has been previously reported (Vaisnav et al. 2014), here, I for the first time demonstrate its association with tolerance to transposition. Furthermore, I show *brat*, a GSC differentiation factor, promotes tolerance, which interestingly is distinct from the Kelleher et al. (2018) report on *bruno*. Together these observations reveal the complex underpinnings of germline tolerance to TEs.

3.4 Methods

3.4.1 *Drosophila* strains and husbandry

The recombinant inbred lines (RILs) were generously provided by Dr. Stuart MacDonald. Harwich (#4264), RAL73 (#28131), *brat*¹ (#3988), Df(2L)*brat*[Exel8040] (#7847), Df(2L)*brat*[ED1231] (#9174) and Df(2L)*brat*[ED1200] (#9173) were obtained from the Bloomington *Drosophila* stock center. The second chromosome centromere from three recombinant inbred lines carrying B6 QTL allele (#21218, #21076, #21156) and two RILs carrying B8 QTL allele (#21077, #21154) were extracted into a common background by crossing them to multiply marked stocks *b cn* (#44229). After seven rounds of backcrossing followed by inbreeding, the final isogenic lines were generated. RILs were made homozygous for the 2nd chromosome by

inbreeding and selecting for wild type phenotype. The genotype of the isogenic lines were verified through PCR using five different primers within the two QTL. All flies were maintained at room temperature in standard cornmeal media.

3.4.2 Quantification of genomic *Responders*

Paired-end deep sequencing libraries (54 nt) for each founder genome (King et al. 2012) were aligned both to the published *D. melanogaster* Release 6 genome (Hoskins et al. 2015) and our custom repeat library using bowtie 2 (using parameters -sensitive, -a) (Langmead and Salzberg 2012). For the *Responder* and 260-bp array, the estimated copy number was determined based on the number of aligned reads per million mapped, divided by the total number of reads aligning to the founder genome and the length of the satellite repeat, i.e., 120 bp (Wu et al. 1988) and 260 bp, respectively (Abad et al. 2000). The size of satellite arrays in each RIL genome was inferred based on its founder haplotype at the centromeric LOD peak (King et al. 2012).

3.4.3 Egg count assay

To estimate the fertility of the dysgenic female progeny, eggs laid were quantified over a 10-day period. I first mated the females from tolerant and sensitive isogenic lines to Harwich males at 25 °C. Reciprocal crosses were set as controls. The dysgenic and non-dysgenic daughters were then mated to tester males (Canton-S). Each vial contained one male and one female, which were transferred to fresh food after 24 h for 10 days. Eggs laid on the food by each daughter were counted daily.

3.4.5 X-ray irradiation

Third instar larvae were either mock treated or irradiated in a Rad Source RS 1800 X-ray

machine set at 12.5 mA and 160 kV. To obtain 3rd instar larvae, embryos were collected for 24 h and aged for 5 days at 25 °C. The food vials containing larvae were then X-ray irradiated at doses from 5-80 Gray after which an optimal dose that clearly depicts the phenotypic difference was selected. Survival to adulthood was determined by scoring the number of empty and full pupal cases until 10 days.

3.4.6 Small-RNA analysis

Sequenced small RNAs from NILs (described in Section 2.4.7 of Chapter 2) were separated based on size into miRNAs/siRNAs (18-22 nt) and piRNAs (23-30 nt) (Brennecke et al. 2008). Reads corresponding to contaminating rRNAs, including 2S-rRNA, were removed from each library by aligning to annotated transcripts from flybase (Gramates et al. 2017). Small RNAs were then aligned to a custom library of the new consensus sequences for the *Responder* (Larracuente 2014) and 1.688 satellite families (Khost et al. 2017). Reads that could not be uniquely assigned to a single repeat class were discarded.

The raw number of sequenced reads from each repeat class were used to infer differential expression using DESeq2 (Love et al. 2014). To detect repeats whose expression pattern differed depending on the haplotype at the chromosome centromeric QTL region, I employed a linear model that included factors for both the QTL haplotype (B6 or B4) and the matched genotype pair.

Chapter 4. Conclusion

Transposable elements (TEs) threaten the genomic integrity of the host by not only producing deleterious mutations but also by causing DSBs. Accumulation of DSBs in germline cells results in germ-cell loss and sterility. While host regulation of TEs through small RNAs has been extensively studied, tolerance mechanisms that allow germ cells to withstand the disruptive effects of transposition have been largely unstudied. My dissertation investigated natural tolerant factors to *P*-element transposition and explored their mechanisms of action.

I found two linked QTLs (QTL-3d and QTL-21d) associated with natural tolerance to *P*-element transposition and subsequently compared the total RNA and small RNA pools among NILs carrying tolerant and sensitive QTL alleles. Through my observation, I speculated two potential mechanisms of tolerance: **1)** enhanced DSB repair and **2)** increased expression of differentiation factor *brat*. Tolerant genotypes were associated with increased expression of members of TIP60 complex, involved in DSB repair and heterochromatin formation (Sinclair et al. 1998; Ruhf et al. 2001; Kusch et al. 2004; Qi et al. 2006; Hanai et al. 2008). Tolerant genotypes also exhibited increased expression of chorion genes whereas sensitive genotypes exhibited upregulation of histone genes, both of which are reliant on DNA-damage repair. I also found signatures of reduced heterochromatin formation in sensitive genotypes- upregulation of histone as well as pericentromeric gene expression. I further demonstrated that tolerance is associated with reduced sensitivity to X-ray radiation.

Through a combined analysis of QTL mapping, gene expression, and in-phase SNPs, I identified two candidate genes. Within the 3-day old QTL, *Nipped-A*, a member of the TIP60 complex, harbored a non-synonymous in-phase SNP. Whereas within the 21-day old QTL, a GSC differentiation factor, *brat*, carried an in-phase SNP within the intron and regulatory region as well as was upregulated in tolerant genotypes. Finally, I demonstrated that loss-of-function mutations in *brat* reduces tolerance to *P*-elements.

Tolerance could be conferred by stabilizing damaged GSCs and their larval progenitors, PGCs during hybrid dysgenesis. Following transposition, the loss of GSC and PGCs that is

incited by DNA damage is thought to result from disruption of germline stem cell maintenance (Khurana et al. 2011; Ma et al. 2016; Kelleher et al. 2018). Furthermore, overexpression of self-renewal factor *myc* and reduced expression of germline differentiation factor *bruno* allows dysgenic individuals to produce viable gametes, possibly by promoting PGCs and GSCs retention, respectively (Kelleher et al. 2018; Ota and Kobayashi 2020). I found signatures of increased activity of the TIP60 complex, which is involved in male and female GSC maintenance (Prado et al. 2013; Yan et al. 2014). Furthermore, TIP60 components including *Nipped-A*, is a known interactor of *myc* (Rust et al. 2018). Therefore, tolerance could be determined by *myc*-dependent retention of damaged PGCs and GSCs. This is also consistent with my observation that *Nipped-A* is linked to the 3-day old QTL, where tolerance is likely to act in early developmental stages.

However, our data also suggests *brat*, a GSC differentiation factor, could confer tolerance. Tolerant genotypes exhibited increased expression of *brat*. Furthermore, a *brat* loss-of-function mutation exacerbated hybrid dysgenesis (**Figure 3.4**). This provides a puzzling yet intriguing contrast with the existing reports that suggest how sustaining GSCs confers tolerance. DNA damage not only results in the degeneration of GSCs but also delays GSCs/cystoblast differentiation by suppressing germline differentiation factor *bam* (Ma et al. 2016). *brat* could potentially confer tolerance by promoting differentiation of GSCs that undergo arrest following DNA damage. *brat* functions in early stages of GSC differentiation (Harris et al. 2011; McCarthy et al. 2018) and therefore could help the cytoblasts escape arrest, which is triggered by *P*-element-mediated DNA damage. My proposed hypothesis is consistent with the fact that *brat* is linked to a 21-day old QTL, suggesting that it is likely to act during adult developmental stages.

Tolerance could be conferred by efficient repair of DSB generated during rampant transposition. GSC loss in dysgenic germline is proposed to be caused by DSB induced by *P*-elements (Khurana et al. 2011). Furthermore, DNA damage response proteins like *chk2* and

p53 are modulators of hybrid dysgenesis (Moon et al. 2018; Tasnim and Kelleher 2018). Hence, a robust DSB damage repair system is a possible mechanism whereby germline could ameliorate the impacts of transposition. This mechanism would also be highly beneficial to the host as it allows for the production of viable gametes with lower mutation from transposition. Indeed stem cells, in general, are known to adopt superior DNA-damage repair machinery compared to other cell types (Saretzki et al. 2004; Maynard et al. 2008), further bolstering my hypothesis.

TIP60 may also confer tolerance by facilitating DSB repair in dysgenic larvae or young females. The TIP60 complex is involved in multiple steps of the DDR pathway (**Figure 4.1**) starting from the earliest steps through activation of DNA-damage sensor, ataxia telangiectasia-mutated (ATM, Dai and Gu 2010) and later during DSB repair (Sun et al. 2005). Furthermore, the TIP60 complex is known to regulate *p53* activation in mammals (Berns et al. 2004). Tolerant genotypes exhibited signatures of efficient DNA-damage repair activity. Tolerant genotypes not only appear to have increased expression of multiple components of TIP60 complex involved in DSB repair, but also upregulation of chorion genes, whose expression is reliant on DSB repair. Tolerant genotypes also exhibited histone upregulation, which is thought to interfere with DNA-damage repair. Consistent with this model, tolerant genotypes showed higher resilience to irradiation, a phenomenon commonly associated with enhanced DNA repair activity (Staeva-Vieira et al. 2003; Ruike et al. 2006; Uri et al. 2007; Sterpone and Cozzi 2010; Koval et al. 2019). Together, my results suggest enhanced DSB repair as a possible mechanism of tolerance.

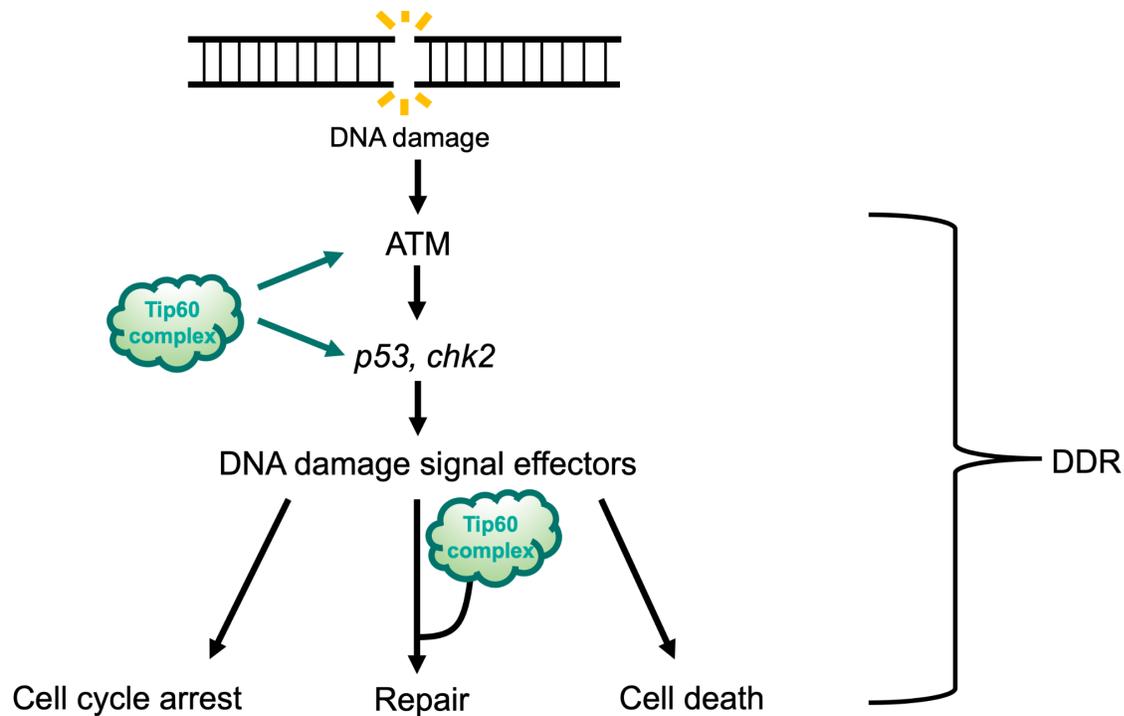


Figure 4.1. Multiple steps through which TIP60 could confer tolerance.

My work points towards two candidate genes with genetic and transcriptional variation associated with tolerance and also reveals the complex underpinnings of tolerance mechanisms. Particularly, my observation that *brat*- a female GSC differentiation factor- promotes tolerance opposes the related Kelleher et al. (2018) study of natural tolerant alleles, where germline differentiation factor *bruno* reduces tolerance. In addition, I for the first time, associate genetic variation in resilience to DNA damage to TE tolerance and the potential for natural variation in DNA-damage repair influencing tolerance. However, further investigation is required to conclusively determine DNA-damage repair and GSC differentiation as mechanisms underlying tolerance. Also, in my study, I measured tolerance based on the individual's capability to make mature eggs. Although there are plenty of studies that report the existence of true tolerance alleles that produce fertile offspring (Kelleher et al. 2018; Ota and Kobayashi 2020), I have yet to study the viability of the gametes produced and the fitness of the resulting progeny.

Tolerance may be an evolutionarily important strategy when new TEs, which the host is unable to silence, invade a genome. Tolerance is also adopted by many tumor types, which experience deregulation of TEs due to loss of silencing (Ardeljan et al. 2020; Grundy et al. 2021). Therefore, my work may be instrumental in understanding host-TE interaction as well as human diseases.

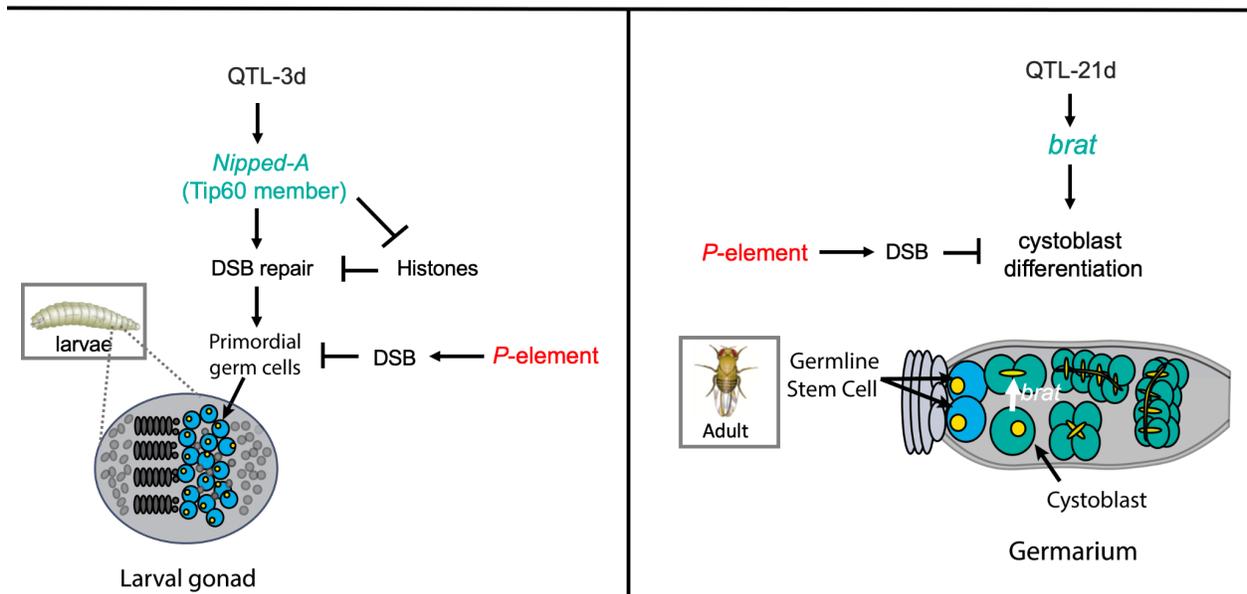


Figure 4.2. Two models of TE tolerance to *P*-elements based on the QTL positions. Since *Nipped-A* resides within the 3-day old QTL, repairing DNA damage may be particularly important in maintaining germline in the dysgenic larvae. On the other hand, since *brat* is linked to 21-day old QTL, maintaining germline as an adult may be dependent upon aiding cystoblasts to escape the cell-cycle arrest imposed by *P*-element-mediated DNA damage. Figure credits: (Kelleher et al. 2020).

Appendix

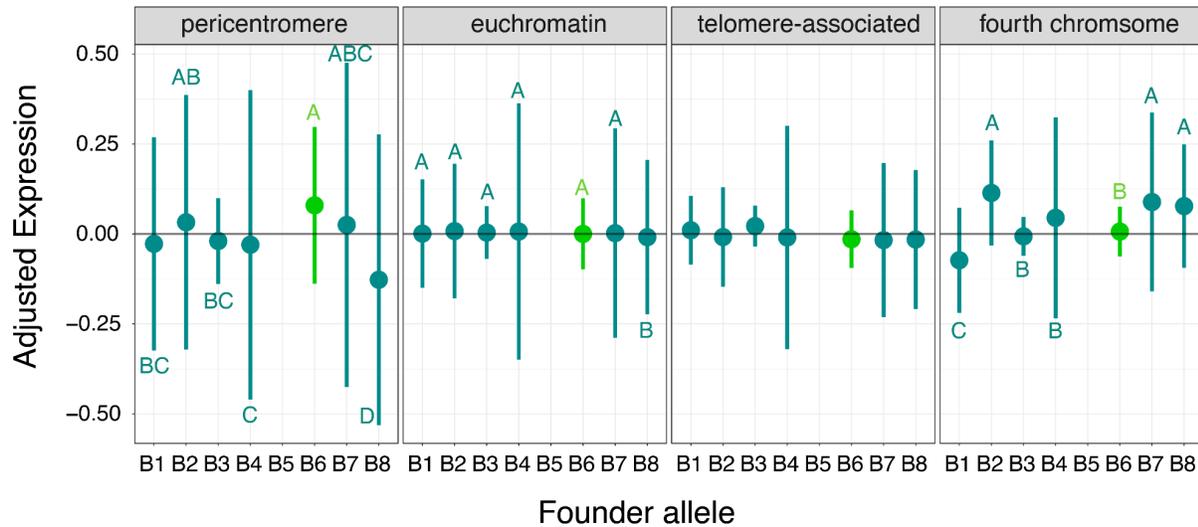


Figure A2.1. Sensitivity is associated with increased expression of pericentromeric genes in the head. a) Mean expression of genes located in the pericentromere, euchromatin, telomere and the fourth chromosome from RILs carrying each of the eight B founder genotypes at the QTL-3d region. Error bars represent the standard deviation among mean expression levels of different genes. The sensitive/B6 (light green) shows high pericentromeric gene expression compared to the tolerant strains (dark green) (Anova; $F_{6,494}=7.775$, $P<5.24e^{-08}$). The letters indicate significantly different expression levels based on Tukey-HSD comparisons between RILs with different founder alleles.

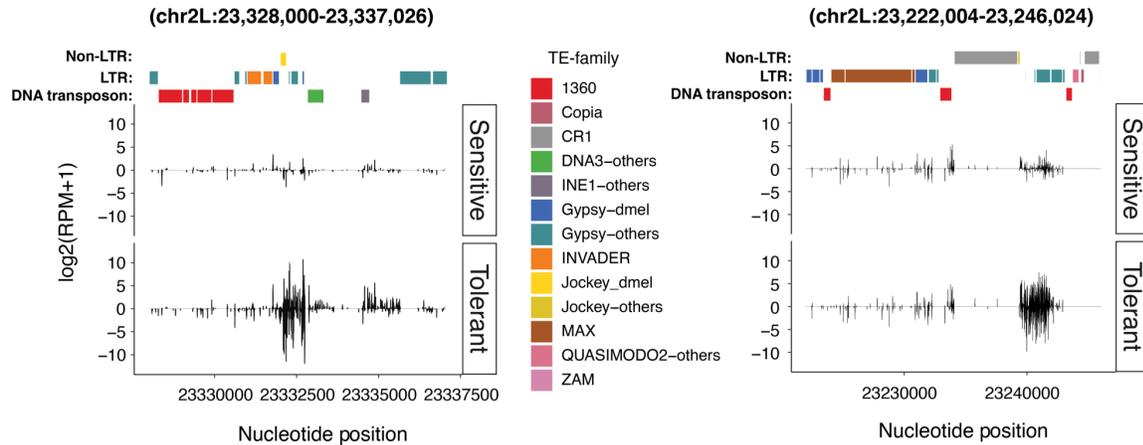


Figure A2.2. Expression profile of QTL piRNA clusters in a sensitive and tolerant NIL pair. The piRNA expression between sensitive and tolerant genotypes from 21188-21291 NIL pairs along the two QTL piRNA clusters: 2L:23,328,000-23,337,026 and 2L:23,222,004-23,246,024, respectively. Only uniquely mapping piRNAs were considered. The TE families at the top of each panel are represented by different colors. TE-others represent the repeat families coming from sibling species of *D. melanogaster*. Positive value indicates piRNAs mapped to the sense strand of the reference genome and negative value indicates those from the antisense strand. The piRNA cluster expression levels were estimated by log₂ scale transformed of reads per million mapped reads [$\log_2(\text{RPM}+1)$].

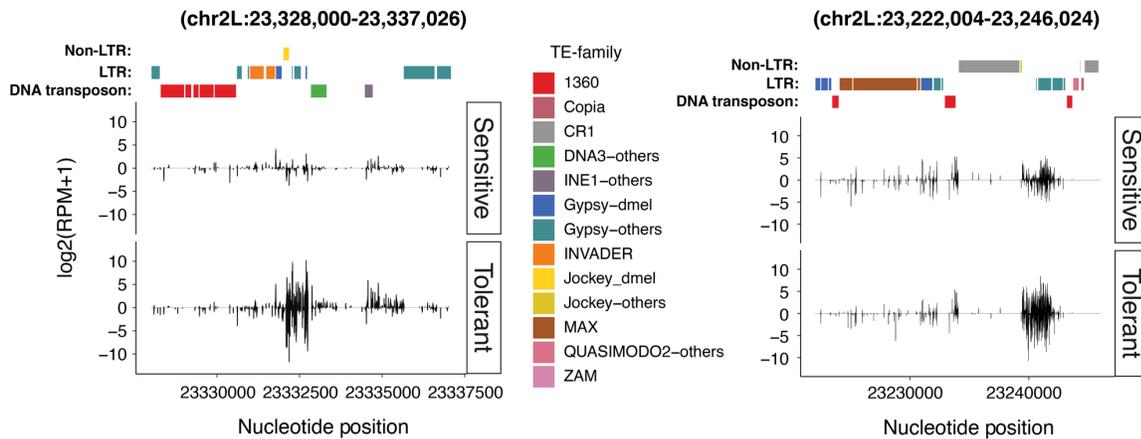


Figure A2.3. Expression profile of QTL piRNA clusters in a sensitive and tolerant NIL pair. The piRNA expression between sensitive and tolerant genotypes from 21346-21147 NIL pairs along the two QTL piRNA clusters: 2L:23,328,000-23,337,026 and 2L:23,222,004-23,246,024, respectively. Only uniquely mapping piRNAs were considered. The TE families at the top of each figure are represented by different colors. TE-others represent the repeat families coming from sibling species of *D. melanogaster*. Positive value indicates piRNAs mapped to the sense strand of the reference genome and negative value indicates those from the antisense strand. The piRNA cluster reads expression levels were estimated by log₂ scale transformed of reads per million mapped reads [$\log_2(\text{RPM}+1)$].

Table A2.1. List of 19 candidate genes that are differential expressed and have in-phase SNPs in their regulatory/intron region

Gene	QTL	Chromosome	No. of in-phase SNPs	Genotype with upregulation	Biological function
brain tumor (brat)	QTL-21d	Chr2L	14	Tolerant	promotes GSC differentiation
CG10492	QTL-21d	Chr2L	3	Tolerant	UNK
gammaTub37C	QTL-21d	Chr2L	8	Sensitive	microtubule nucleation at the centrosomes
lethal(2)37Cg	QTL-21d	Chr2L	5	Sensitive	transcription by RNA poll/III
Leukocyte-antigen-related-like (Lar)	QTL-21d	Chr2L	5	Tolerant	oocyte follicle cell development and patterning
CG31612	QTL-3d	Chr2L	2	Tolerant	replication
CG40006	QTL-3d	Chr2L	2	Sensitive	unknown
CG9336	QTL-3d	Chr2L	18	Sensitive	unknown
His2B:CG33868	QTL-3d	Chr2L	1	Sensitive	Histone 2B-heterochromatin organization
His3:CG33866	QTL-3d	Chr2L	2	Sensitive	Histone 3-heterochromatin organization
Microsomal triacylglycerol transfer protein (Mtp)	QTL-3d	Chr2L	16	Tolerant	phosphatidylcholine transporter involved in lipoprotein metabolism

Table A2.1 continued

CG11163	QTL-3d	Chr2R	15	Sensitive	regulation of sequestering of zinc ion
CG17508	QTL-3d	Chr2R	3	Sensitive	unknown
CG30438	QTL-3d	Chr2R	31	Sensitive	UDP-glycosyltransferase activity
CG34200	QTL-3d	Chr2R	6	Sensitive	unknown
Fission, mitochondrial 1 (Fis1)	QTL-3d	Chr2R	2	Sensitive	mitochondrial organization and fission; autophagy
jing	QTL-3d	Chr2R	79	Tolerant	border follicle cell migration
missing-in-metastasis (mim)	QTL-3d	Chr2R	20	Sensitive	border cell and PGC migration
ubiquitin like (ubl)	QTL-3d	Chr2R	3	Sensitive	Cellular protein modification

The biological function data of genes is derived from Flybase (Thurmond et al., 2018).

Table A2.2. List of candidate genes with non-synonymous in-phase SNPs

Gene	QTL	Chromosome	Biological function
pigeon	QTL-21d	2L	positive regulation of amyloid-beta formation.
CG10492	QTL-21d	2L	unknown
CG17568	QTL-21d	2L	zinc ion binding; nucleic acid binding
CG10700	QTL-21d	2L	apoptosis
ttm3	QTL-3d	2L	mitochondrial membrane organization

Table A2.2 continued

tio	QTL-3d	2L	negative regulation of transcription
tadr	QTL-3d	2L	unknown
Oseg5	QTL-3d	2L	cilium assembly
nolo	QTL-3d	2L	ventral cord development
Mtp	QTL-3d	2L	lipoprotein metabolic process
I(2)05287	QTL-3d	2L	unknown
ltgbn	QTL-3d	2L	apoptotic cell clearance
lr40a	QTL-3d	2L	sensory perception
dtr	QTL-3d	2L	chemical synaptic transmission; cilium assembly
CG9270	QTL-3d	2L	ATPase activity
CG42597	QTL-3d	2L	unknown
CG3651	QTL-3d	2L	unknown
CG31703	QTL-3d	2L	Cdc73/Paf1 complex
CG31702	QTL-3d	2L	Cdc73/Paf1 complex
CG31693	QTL-3d	2L	transmembrane transport
CG31674	QTL-3d	2L	unknown
CG31673	QTL-3d	2L	unknown
CG31601	QTL-3d	2L	unknown

Table A2.2 continued

Tsp42A	QTL-3d	2R	unknown
Src42A	QTL-3d	2R	actin filament bundle assembly; apoptotic cell clearance
sced	QTL-3d	2R	actin filament reorganization involved in cell cycle
Nipped-A	QTL-3d	2R	DNA repair; germline differentiation; heterochromatin organization
mle	QTL-3d	2R	DNA duplex unwinding; dosage compensation; heterochromatin assembly
mim	QTL-3d	2R	border follicle cell migration; germ cell migration
I(2)09851	QTL-3d	2R	unknown
jing	QTL-3d	2R	border follicle cell migration; chromatin organization
Cyp6a2	QTL-3d	2R	response to caffeine; response to DDT
CG7856	QTL-3d	2R	unknown
CG7791	QTL-3d	2R	protein processing involved in protein targeting to mitochondrion
CG43366	QTL-3d	2R	negative regulation of endopeptidase activity
CG3270	QTL-3d	2R	mitochondrial respiratory chain complex I assembly

Table A2.2 continued

CG30431	QTL-3d	2R	regulation of transcription, DNA-templated
CG15237	QTL-3d	2R	regulation of mitotic cell cycle spindle assembly checkpoint
CG15236	QTL-3d	2R	unknown
CG15233	QTL-3d	2R	unknown
CG1344	QTL-3d	2R	phosphorylation
CG30431	QTL-3d	2R	regulation of transcription, DNA-templated

The biological function data of genes is derived from Flybase (Thurmond et al., 2018).

Table A2.3. List of differential expressions of Tip60 members and one of its interactors (RSF)

Gene Symbol	Chromosome	baseMean	log2FoldChange (Sensitive/Tolerant)	padj
Act87E	3R	94.3	-0.05	0.93
E(Pc)	2R	4611	-0.97	6.74E-07
dom	2R	9074.8	-0.66	3.70E-05
Bap55	2R	872.1	0.007	0.97
Tip60	X	711.7	0.36	0.12
CG8677/RSF	2L	2783	-0.83	4.85E-05
MRG15	3R	583.7	-0.11	0.38
Ing3	X	342.8	0.17	0.29
Gas41	2L	270.6	0.47	0.51
YL-1	2L	614.9	0.06	0.72
MrgBP	2R	409.1	0.15	0.36
DMAP1	2R	880.8	-0.22	0.022

Table A2.3 continued

Eaf6	3L	576.7	0.11	0.63
Brd8	3R	488.5	-0.1	0.60
rept	3L	720.3	0.03	0.81
pont	3R	999.7	-0.06	0.63
Nipped-A	2R	3514.8	-0.33	0.16
Yeti	2R	2600.4	0.59	4.16E-07

The positive log₂ fold-change value indicates that the gene is upregulated in the sensitive while negative log₂ fold-change indicates those that are upregulated in the tolerant genotypes. The highlighted ones are those that are significantly differentially expressed. The genomic coordinates for each gene is provided in dm6.

Table A2.4. RepBase Censor table for QTL cluster 1 (chrom2L:23222004..23246024)

Name	From	To	TE	TE_From	TE_To	Class	Dir	Sim
/tmp/censor.2798 0.tmp/data.ori	1	485	Gypsy4	4913	5400	LTR/Gypsy	d	0.9774
/tmp/censor.2798 0.tmp/data.ori	489	1127	Gypsy4	5650	6288	LTR/Gypsy	d	0.989
/tmp/censor.2798 0.tmp/data.ori	1128	1412	Gypsy4	1	287	LTR/Gypsy	d	0.9652
/tmp/censor.2798 0.tmp/data.ori	1413	2030	1360	162	853	DNA/P	c	0.9548
/tmp/censor.2798 0.tmp/data.ori	2031	3176	MAX	976	2022	LTR/BEL	d	0.9224
/tmp/censor.2798 0.tmp/data.ori	3177	8611	MAX	2468	7914	LTR/BEL	d	0.9688
/tmp/censor.2798 0.tmp/data.ori	8615	8890	MAX	1	320	LTR/BEL	d	0.844
/tmp/censor.2798 0.tmp/data.ori	8891	9914	Gypsy12	845	1928	LTR/Gypsy	c	0.7517
/tmp/censor.2798 0.tmp/data.ori	9966	10618	Gypsy-others	387	1110	LTR/Gypsy	d	0.9085

Table A2.4 continued

/tmp/censor.27980.tmp/data.ori	10619	10824	Gypsy-others	4214	4420	LTR/Gypsy	d	0.8889
/tmp/censor.27980.tmp/data.ori	10825	10893	INVADER4	1	70	LTR/Gypsy	d	0.8857
/tmp/censor.27980.tmp/data.ori	10897	11874	1360	30	1153	DNA/P	d	0.9381
/tmp/censor.27980.tmp/data.ori	12068	17244	RT1B	1	5183	NonLTR/R1	c	0.9954
/tmp/censor.27980.tmp/data.ori	17247	17453	Jockey-others	3	209	NonLTR/Jockey	c	0.7212
/tmp/censor.27980.tmp/data.ori	17698	17766	ZAM	1	69	LTR/Gypsy	d	0.8551
/tmp/censor.27980.tmp/data.ori	17767	17817	Gypsy-others	81	131	LTR/Gypsy	d	0.8431
/tmp/censor.27980.tmp/data.ori	17819	17898	Gypsy-others	552	630	LTR/Gypsy	d	0.8125
/tmp/censor.27980.tmp/data.ori	18570	18707	Gypsy-others	2686	2823	LTR/Gypsy	d	0.7194
/tmp/censor.27980.tmp/data.ori	18753	19912	Gypsy-others	1357	2630	LTR/Gypsy	d	0.671
/tmp/censor.27980.tmp/data.ori	19953	20890	Gypsy-others	153	1110	LTR/Gypsy	d	0.9057
/tmp/censor.27980.tmp/data.ori	20891	21096	Gypsy-others	4214	4420	LTR/Gypsy	d	0.8986
/tmp/censor.27980.tmp/data.ori	21097	21165	INVADER4	1	70	LTR/Gypsy	d	0.8857
/tmp/censor.27980.tmp/data.ori	21169	21698	1360	30	625	DNA/P	d	0.9333
/tmp/censor.27980.tmp/data.ori	21699	22237	QUASIMODO2-others	1	497	LTR/Gypsy	c	0.976
/tmp/censor.27980.tmp/data.ori	22270	22377	CR1	4261	4368	NonLTR/CR1	c	0.9259
/tmp/censor.27980.tmp/data.ori	22378	22665	Copia-2	1	288	LTR/Copia	d	0.9688

Table A2.4 continued

/tmp/censor.27980.tmp/data.ori	22666	23904	CR2	3082	4291	NonLTR/CR1	c	0.9469
/tmp/censor.27980.tmp/data.ori	23934	24010	Gypsy-others	6655	6731	LTR/Gypsy	d	0.7273

Column 2 and 3 (From and To) indicate the start and stop position on the piRNA cluster whereas Column 5 and 6 (TE_From to TE_To) indicate the start and stop position on the TE sequence. Dir indicates orientation ('d' for direct, 'c' for complementary) of TE fragment. sim indicates the alignment similarity between the two fragments.

Table A2.5. RepBase Censor table for QTL cluster2 (2L:23328000..23337026)

Name	From	To	TE	TE_From	TE_To	Class	Dir	Sim
/tmp/censor.29235.tmp/data.ori	1	279	Gypsy-others	3304	3582	LTR/Gypsy	d	0.8889
/tmp/censor.29235.tmp/data.ori	280	1016	1360	1	750	DNA/P	d	0.9499
/tmp/censor.29235.tmp/data.ori	1025	1231	1360	1610	1816	DNA/P	d	0.9662
/tmp/censor.29235.tmp/data.ori	1271	1456	1360	2257	2476	DNA/P	d	0.9786
/tmp/censor.29235.tmp/data.ori	1458	1913	1360	2994	3457	DNA/P	d	0.9629
/tmp/censor.29235.tmp/data.ori	1914	2589	1360	3739	4480	DNA/P	d	0.9604
/tmp/censor.29235.tmp/data.ori	2590	2762	Gypsy-others	3576	3769	LTR/Gypsy	d	0.8409

Table A2.5 continued

/tmp/censor.292 35.tmp/data.ori	2916	2987	Gypsy-other s	4366	4448	LTR/Gypsy	d	0.8611
/tmp/censor.292 35.tmp/data.ori	2988	3428	INVADER	2718	3186	LTR/Gypsy	d	0.8787
/tmp/censor.292 35.tmp/data.ori	3475	3770	INVADER	113	414	LTR/Gypsy	d	0.8339
/tmp/censor.292 35.tmp/data.ori	3772	3973	Gypsy-dmel	198	382	LTR/Gypsy	d	0.8053
/tmp/censor.292 35.tmp/data.ori	3998	4193	Jockey-dmel	3794	4000	NonLTR/Jockey	d	0.7929
/tmp/censor.292 35.tmp/data.ori	4253	4305	Gypsy-other s	2	57	LTR/Gypsy	d	0.8364
/tmp/censor.292 35.tmp/data.ori	4327	4553	Gypsy-other s	130	357	LTR/Gypsy	d	0.7588
/tmp/censor.292 35.tmp/data.ori	4666	4739	Gypsy-other s	81	154	LTR/Gypsy	d	0.7838
/tmp/censor.292 35.tmp/data.ori	4831	5332	DNA3-1-oth ers	374	918	DNA	c	0.7202
/tmp/censor.292 35.tmp/data.ori	6463	6730	INE1	1	327	DNA/Helitron	c	0.8982
/tmp/censor.292 35.tmp/data.ori	7644	8603	Gypsy-other s	1660	2709	LTR/Gypsy	d	0.6554
/tmp/censor.292 35.tmp/data.ori	8648	9100	Gypsy-other s	3242	3758	LTR/Gypsy	d	0.7174

Column 2 and 3 (From and To) indicate the start and stop position on the piRNA cluster whereas Column 5 and 6 (TE_From to TE_To) indicate the start and stop position on the TE sequence. Dir indicates orientation ('d' for direct, 'c' for complementary) of TE fragment. sim indicates the alignment similarity between the two fragments.

Table A3.1. Percentage of larvae that survived to adulthood after irradiation at a range of X-ray doses

Strain	Genotype	X-ray Dose	No. of pupae	Empty Pupae	Survival to adulthood%
21218	Sensitive	5 Gray	214	199	92.27
21156	Sensitive	5 Gray	94	76	80.8
21077	Tolerant	5 Gray	339	323	95.2
21154	Tolerant	5 Gray	291	244	85.1
CS	Control	5 Gray	49	44	89.7
w1	Control	5 Gray	34	31	91.97
Marker	Control	5 Gray	39	30	76.9
21218	Sensitive	10 Gray	334	92	27.5
21076	Sensitive	10 Gray	30	5	16.67
21156	Sensitive	10 Gray	52	14	26.9
21077	Tolerant	10 Gray	346	182	52.6
21154	Tolerant	10 Gray	488	282	57.78
CS	Control	10 Gray	71	48	67.6
w1	Control	10 Gray	60	33	55
Marker	Control	10 Gray	42	34	80.9
21218	Sensitive	17 Gray	31	1	3.23
21076	Sensitive	17 Gray	38	3	7.89
21077	Tolerant	17 Gray	28	2	7.14
21154	Tolerant	17 Gray	46	9	19.56
CS	Control	17 Gray	29	6	20.68
w1	Control	17 Gray	50	2	4
21218	Sensitive	25 Gray	11	0	0
21076	Sensitive	25 Gray	29	1	3.4
21077	Tolerant	25 Gray	35	1	2.9
21154	Tolerant	25 Gray	32	0	0
w1	Control	25 Gray	30	0	0

Table A3.1 continued

21218	Sensitive	30 Gray	51	0	0
21076	Sensitive	30 Gray	35	0	0
21077	Tolerant	30 Gray	24	1	4.17
21154	Tolerant	30 Gray	28	0	0
CS	Control	30 Gray	37	0	0
w1	Control	30 Gray	30	0	0
21218	Sensitive	40 Gray	35	0	0
21077	Tolerant	40 Gray	35	1	2.86
21154	Tolerant	40 Gray	35	0	0
CS	Control	40 Gray	35	0	0
w1	Control	40 Gray	35	0	0
21218	Sensitive	60 Gray	35	0	0
21076	Sensitive	60 Gray	35	0	0
21077	Tolerant	60 Gray	70	0	0
21154	Tolerant	60 Gray	35	0	0
CS	Control	60 Gray	35	0	0
w1	Control	60 Gray	35	0	0
21218	Sensitive	80 Gray	35	0	0
21076	Sensitive	80 Gray	35	0	0
21077	Tolerant	80 Gray	35	0	0
21154	Tolerant	80 Gray	35	0	0
CS	Control	80 Gray	35	0	0

Table A3.2. Percentage of mock treated larvae that survived to adulthood

Strain	Genotype	No. of pupae	Empty Pupae	Survival to adulthood%
21218	Sensitive	83	79	95.2
21076	Sensitive	40	38	95
21156	Sensitive	36	32	88.89
21077	Tolerant	208	196	94.23
21154	Tolerant	169	162	95.9
CS	Control	142	139	97.9
w1	Control	57	54	94.7
Marker	Control	159	145	91.2

Bibliography

- Abad JP, Agudo M, Molina I, Losada A, Ripoll P, Villasante A. 2000. Pericentromeric regions containing 1.688 satellite DNA sequences show anti-kinetochore antibody staining in prometaphase chromosomes of *Drosophila melanogaster*. *Mol. Gen. Genet.* 264:371–377.
- Ade C, Roy-Engel AM, Deininger PL. 2013. Alu elements: An intrinsic source of human genome instability. *Curr. Opin. Virol.* 3:639–645.
- Aldrup-MacDonald ME, Kuo ME, Sullivan LL, Chew K, Sullivan BA. 2016. Genomic variation within alpha satellite DNA influences centromere location on human chromosomes with metastable epialleles. *Genome Res.* 26:1301–1311.
- Alexander JL, Barrasa MI, Orr-Weaver TL. 2015. Replication fork progression during re-replication requires the DNA damage checkpoint and double-strand break repair. *Curr. Biol.* 25:1654–1660.
- Amanda M. Larracuent DCP. 2012. The selfish segregation distorter gene complex of *Drosophila melanogaster*. *Genetics* 192:33–53.
- Ambrosio L, Schedl P. 1985. Two discrete modes of histone gene expression during oogenesis in *Drosophila melanogaster*. *Dev. Biol.* 111:220–231.
- Anxolabéhère D, Kidwell MG, Periquet G. 1988. Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile *P* elements. *Mol. Biol. Evol.* 5:252–269.
- Aravin AA, Hannon GJ, Brennecke J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318:371–377.
- Ardeljan D, Steranka JP, Liu C, Li Z, Taylor MS, Payer LM, Gorbounov M, Sarnecki JS, Deshpande V, Hruban RH, et al. 2020. Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. *Nat. Struct. Mol. Biol.* 27:168–178.
- Asaoka M, Lin H. 2004. Germline stem cells in the *Drosophila* ovary descend from pole cells in the anterior region of the embryonic gonad. *Development* 131:5079–5089.
- Ayarpadikannan S, Kim H-S. 2014. The impact of transposable elements in genome evolution and genetic instability and their implications in various diseases. *Genomics Inform.* 12:98–104.
- Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* 6:1–6.
- Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:1–48.
- Bennetzen JL, Wang H. 2014. The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annu. Rev. Plant Biol.* 65:505–530.
- Bergman CM, Bensasson D. 2007. Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.*

S. A. 104:11340–11345.

- Berns K, Hijmans EM, Mullenders J, Brummelkamp TR, Velds A, Heimerikx M, Kerkhoven RM, Madiredjo M, Nijkamp W, Weigelt B, et al. 2004. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428:431–437.
- Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL, Macfarlan TS, et al. 2018. Ten things you should know about transposable elements. *Genome Biol.* 19:1–12.
- Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34:525–527.
- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128:1089–1103.
- Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322:1387–1392.
- Britten RJ, Kohne DE. 1968. Repeated sequences in DNA. *Science* 161:529–540.
- Bucheton A, Lavigne JM, Picard G, L'Heritier P. 1976. Non-mendelian female sterility in *Drosophila melanogaster*: Quantitative variations in the efficiency of inducer and reactive strains. *Heredity* 36:305–314.
- Bucheton A, Paro R, Sang HM, Pelisson A, Finnegan DJ. 1984. The molecular basis of I-R hybrid dysgenesis in *Drosophila melanogaster*: Identification, cloning, and properties of the I factor. *Cell* 38:153–163.
- Casier K, Delmarre V, Gueguen N, Hermant C, Viodé E, Vaury C, Ronsseray S, Brassat E, Teyssat L, Boivin A. 2019. Environmentally-induced epigenetic conversion of a piRNA cluster. *eLife* 8:e39842.
- Casola C, Hucks D, Feschotte C. 2008. Convergent domestication of *pogo*-like transposases into centromere-binding proteins in fission yeast and mammals. *Mol. Biol. Evol.* 25:29–41.
- Chang TH, Mattei E, Gainetdinov L, Colpan C, Weng Z, Zamore PD. 2019. Maelstrom represses canonical polymerase II transcription within bi-directional piRNA Clusters in *Drosophila melanogaster*. *Mol. Cell* 73:291–303.
- Charles M, Belcram H, Just J, Huneau C, Viollet A, Couloux A, Segurens B, Carter M, Huteau V, Coriton O, et al. 2008. Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. *Genetics* 180:1071–1086.
- Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* 18:71–86.
- Claycomb JM, Benasutti M, Bosco G, Fenger DD, Orr-Weaver TL. 2004. Gene amplification as a developmental strategy: Isolation of two developmental amplicons in *Drosophila*. *Dev. Cell* 6:145–155.

- Consortium, International Human Genome Sequencing. 2001. Correction: Initial sequencing and analysis of the human genome. *Nature* 412:565–566.
- Cronin SJ, Nehme NT, Limmer S, Liegeois S, Pospisilik JA, Schramek D, Leibbrandt A, Simoes RM, Gruber S, Puc U, et al. 2009. Genome-wide RNAi screen identifies genes involved in intestinal pathogenic bacterial infection. *Science* 325:340–343.
- Dai C, Gu W. 2010. p53 post-translational modification: deregulated in tumorigenesis. *Trends Mol. Med.* 16:528–536.
- Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A. 1990. Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. *Genetics* 124:339–355.
- De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, Manivannan J, Peterson AL, Kreiling JA, Neretti N, Sedivy JM. 2013. Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell* 12:247–256.
- Deininger PL, Batzer MA. 1999. Alu repeats and human disease. *Mol. Genet. Metab.* 67:183–193.
- Dorogova NV, Bolobolova EU, Zakharenko LP. 2017. Cellular aspects of gonadal atrophy in *Drosophila* P-M hybrid dysgenesis. *Dev. Biol.* 424:105–112.
- Dotto BR, Carvalho EL, Silva AF, Duarte Silva LF, Pinto PM, Ortiz MF, Wallau GL. 2015. HTT-DB: Horizontally transferred transposable elements database. *Bioinformatics* 31:2915–2917.
- Engels WR, Preston CR. 1979. Hybrid dysgenesis in *Drosophila melanogaster*: The biology of female and male sterility. *Genetics* 92:161–174.
- Ernst C, Odom DT, Kutter C. 2017. The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity. *Nat. Commun.* 8:1–10.
- Ferrandon D. 2009. Host tolerance versus resistance and microbial virulence in the host-pathogen equation. *Cell Host Microbe* 6:203–205.
- Ferree PM, Barbash DA. 2009. Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. *PLoS Biol.* 7:e1000234.
- Finnegan DJ. 1989. Eukaryotic transposable elements and genome evolution. *Trends Genet.* 5:103–107.
- Friedberg EC. 2005. Suffering in silence: The tolerance of DNA damage. *Nat. Rev. Mol. Cell Biol.* 6:943–953.
- Funikov SY, Kulikova DA, Krasnov GS, Rezvykh AP, Chuvakova LN, Shostak NG, Zelentsova ES, Blumenstiel JP, Evgen'ev MB. 2018. Spontaneous gain of susceptibility suggests a novel mechanism of resistance to hybrid dysgenesis in *Drosophila virilis*. *PLoS Genet.* 14:e1007400.
- Gallant P. 2013. Myc function in *Drosophila*. *Cold Spring Harb. Perspect. Med.* 3:a014324.

- Gasior SL, Wakeman TP, Xu B, Deininger PL. 2006. The human LINE-1 retrotransposon creates DNA double-strand breaks. *J. Mol. Biol.* 357:1383–1393.
- George P, Jensen S, Pogorelcnik R, Lee J, Xing Y, Brasslet E, Vaury C, Sharakhov IV. 2015. Increased production of piRNAs from euchromatic clusters and genes in *Anopheles gambiae* compared with *Drosophila melanogaster*. *Epigenetics Chromatin* 8:1–21.
- Gilboa L, Lehmann R. 2004. Repression of primordial germ cell differentiation parallels germ line stem cell maintenance. *Curr. Biol.* 14:981–986.
- Gramates LS, Marygold SJ, Santos G dos, Urbano J-M, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, et al. 2017. FlyBase at 25: looking to the future. *Nucleic Acids Res.* 45:D663–D671.
- Grifoni D, Bellosta P. 2015. *Drosophila* Myc: A master regulator of cellular performance. *Biochim. Biophys. Acta* 1849:570–581.
- Grundy EE, Diab N, Chiappinelli KB. 2021. Transposable element regulation and expression in cancer. *FEBS J.* <https://doi.org/10.1111/febs.15722>.
- Gunjan A, Verreault A. 2003. A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in *S. cerevisiae*. *Cell* 115:537–549.
- Hadziselimovic F, Hadziselimovic NO, Demougin P, Krey G, Oakeley E. 2015. Piwi-pathway alteration induces LINE-1 transposon derepression and infertility development in cryptorchidism. *Sex Dev.* 9:98–104.
- Halling KC, Lazzaro CR, Honchel R, Bufill JA, Powell SM, Arndt CAS, Lindor NM. 1999. Hereditary desmoid disease in a family with a germline *Alu I* repeat mutation of the APC gene. *Human Heredity* 49:97–102.
- Hamilton A, Voinnet O, Chappell L, Baulcombe D. 2002. Two classes of short interfering RNA in RNA silencing. *EMBO J.* 21:4671–4679.
- Hanai K, Furuhashi H, Yamamoto T, Akasaka K, Hirose S. 2008. RSF governs silent chromatin formation via histone H2Av replacement. *PLoS Genet.* 4:e1000011.
- Hancks DC, Kazazian HH Jr. 2016. Roles for retrotransposon insertions in human disease. *Mob. DNA* 7:1–28.
- Han K, Lee J, Meyer TJ, Remedios P, Goodwin L, Batzer MA. 2008. L1 recombination-associated deletions generate human genomic variation. *Proc. Natl. Acad. Sci. U. S. A.* 105:19366–19371.
- Harris RE, Pargett M, Sutcliffe C, Umulis D, Ashe HL. 2011. Brat promotes stem cell differentiation via control of a bistable switch that restricts BMP signaling. *Dev. Cell* 20:72–83.
- Hedges DJ, Deininger PL. 2007. Inviting Instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. *Mutat. Res.* 616:46–59.
- Hill T. 2019. Transposable element dynamics are consistent across the *Drosophila* phylogeny, despite drastically differing content. <http://dx.doi.org/10.1101/651059>

- Hoskins RA, Carlson JW, Wan KH, Park S, Mendez I, Galle SE, Booth BW, Pfeiffer BD, George RA, Svirskas R, et al. 2015. The Release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome Res.* 25:445–458.
- Huang S, Tao X, Yuan S, Zhang Y, Li P, Beilinson HA, Zhang Y, Yu W, Pontarotti P, Escriva H, et al. 2016. Discovery of an active *RAG* transposon illuminates the origins of V(D)J recombination. *Cell* 166:102–114.
- Ichida K, Suzuki K, Fukui T, Takayama Y, Kakizawa N, Watanabe F, Ishikawa H, Muto Y, Kato T, Saito M, et al. 2018. Overexpression of satellite alpha transcripts leads to chromosomal instability via segregation errors at specific chromosomes. *Int. J. Oncol.* 52:1685–1693.
- Ignatenko OM, Zakharenko LP, Dorogova NV, Fedorova SA. 2015. *P* elements and the determinants of hybrid dysgenesis have different dynamics of propagation in *Drosophila melanogaster* populations. *Genetica* 143:751–759.
- Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin C-S, et al. 2017. Improved maize reference genome with single-molecule technologies. *Nature* 546:524–527.
- Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, et al. 2002. The transposable elements of the *Drosophila melanogaster* euchromatin: A genomics perspective. *Genome Biol.* 3:1–20.
- Kapitonov VV, Koonin EV. 2015. Evolution of the RAG1-RAG2 locus: Both proteins came from the same transposon. *Biol. Direct* 10:1–8.
- Karpen GH, Spradling AC. 1992. Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome *Dp1187* by single *P* element insertional mutagenesis. *Genetics* 132:737–753.
- Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of *Arabidopsis* siRNAs. *PLoS Biol.* 5:e57.
- Kelleher ES. 2016. Reexamining the *P*-element invasion of *Drosophila melanogaster* through the lens of piRNA silencing. *Genetics* 203:1513–1531.
- Kelleher ES, Jaweria J, Akoma U, Ortega L, Tang W. 2018. QTL mapping of natural variation reveals that the developmental regulator *bruno* reduces tolerance to *P*-element transposition in the *Drosophila* female germline. *PLoS Biol.* 16:e2006040.
- Kelleher ES, Lama J, Wang L. 2020. Uninvited guests: How transposable elements take advantage of *Drosophila* germline stem cells, and how stem cells fight back. *Curr Opin Insect Sci* 37:49–56.
- Khost DE, Eickbush DG, Larracuenta AM. 2017. Single-molecule sequencing resolves the detailed structure of complex satellite DNA loci in *Drosophila melanogaster*. *Genome Res.* 27:709–721.
- Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, Li C, Zamore PD, Weng Z, Theurkauf WE. 2011. Adaptation to *P* element transposon invasion in *Drosophila melanogaster*. *Cell* 147:1551–1563.

- Kidwell MG. 1983. Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 80:1655–1659.
- Kidwell, M.G., Frydryck, T., and Novy, J.B. 1983. The hybrid dysgenesis potential of *Drosophila melanogaster* strains of diverse temporal and geographical natural origins. *Drosophila Information Service* 59:59–63.
- Kidwell MG, Kidwell JF, Sved JA. 1977. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* 86:813–833.
- Kim JC, Nordman J, Xie F, Kashevsky H, Eng T, Li S, MacAlpine DM, Orr-Weaver TL. 2011. Integrative analysis of gene amplification in *Drosophila* follicle cells: Parameters of origin activation and repression. *Genes Dev.* 25:1384–1398.
- King EG, Merkes CM, McNeil CL, Hooper SR, Sen S, Broman KW, Long AD, Macdonald SJ. 2012. Genetic dissection of a model complex trait using the *Drosophila* synthetic population resource. *Genome Res.* 22:1558–1566.
- King RC. 1970. Ovarian development in *Drosophila melanogaster*. New York: Academic Press
- Kishikawa T, Otsuka M, Suzuki T, Seimiya T, Sekiba K, Ishibashi R, Tanaka E, Ohno M, Yamagami M, Koike K. 2018. Satellite RNA increases DNA damage and accelerates tumor formation in mouse models of pancreatic cancer. *Mol. Cancer Res.* 16:1255–1262.
- Kishikawa T, Otsuka M, Yoshikawa T, Ohno M, Ijichi H, Koike K. 2016. Satellite RNAs promote pancreatic oncogenic processes via the dysfunction of YBX1. *Nat. Commun.* 7:13006.
- Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, et al. 2009. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* 138:1137–1149.
- Kofler R, Nolte V, Schlötterer C. 2015. Tempo and mode of transposable element activity in *Drosophila*. *PLoS Genet.* 11:e1005406.
- Kofler R, Senti K-A, Nolte V, Tobler R, Schlötterer C. 2018. Molecular dissection of a natural transposable element invasion. *Genome Res.* 28:824–835.
- Kohany O, Gentles AJ, Hankus L, Jurka J. 2006. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* 7:1–7.
- Koval L, Proshkina E, Shaposhnikov M, Moskalev A. 2019. The role of DNA repair genes in radiation-induced adaptive response in *Drosophila melanogaster* is differential and conditional. *Biogerontology* 21:45–56.
- Kumar K, Moirangthem R, Kaur R. 2020. Histone H4 dosage modulates DNA damage response in the pathogenic yeast *Candida glabrata* via homologous recombination pathway. *PLoS Genet.* 16:e1008620.
- Kusch T, Florens L, Macdonald WH, Swanson SK, Glaser RL, Yates JR, Abmayr SM, Washburn MP, Workman JL. 2004. Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* 306:2084–2087.

- Landais S, D'Alterio C, Jones DL. 2014. Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in *Drosophila melanogaster*. *Cell Rep.* 7:859–870.
- Lander ES, Botstein D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357–359.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:1–10.
- Larracuent AM. 2014. The organization and evolution of the Responder satellite in species of the *Drosophila melanogaster* group: Dynamic evolution of a target of meiotic drive. *BMC Evol. Biol.* 14:1–12.
- Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* 27:390–399.
- Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen Y-CA, Luo Y, Sachidanandam R, Toth KF, et al. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev.* 28:1667–1680.
- Levis R, O'Hare K, Rubin GM. 1984. Effects of transposable element insertions on RNA encoded by the white gene of *Drosophila*. *Cell* 38:471–481.
- Liang D, Burkhart SL, Singh RK, Kabbaj M-HM, Gunjan A. 2012. Histone dosage regulates DNA damage sensitivity in a checkpoint-independent manner by the homologous recombination pathway. *Nucleic Acids Res.* 40:9604–9620.
- Liu JC, Lerou PH, Lahav G. 2014. Stem cells: Balancing resistance and sensitivity to DNA damage. *Trends Cell Biol.* 24:268–274.
- Long AD, Macdonald SJ, King EG. 2014. Dissecting complex traits using the *Drosophila* synthetic population resource. *Trends Genet.* 30:488–495.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:1–21.
- Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137:522–535.
- Malone CD, Hannon GJ. 2009. Small RNAs as guardians of the genome. *Cell* 136:656–668.
- Manichaikul A, Dupuis J, Sen S, Broman KW. 2006. Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 174:481–489.
- Mauricio R. 2000. Natural selection and the joint evolution of tolerance and resistance as plant defenses. *Evol. Ecol.* 14:491–507.

- Ma X, Han Y, Song X, Do T, Yang Z, Ni J, Xie T. 2016. DNA damage-induced Lok/CHK2 activation compromises germline stem cell self-renewal and lineage differentiation. *Development* 143:4312–4323.
- Ma X, Zhu X, Han Y, Story B, Do T, Song X, Wang S, Zhang Y, Blanchette M, Gogol M, et al. 2017. Aubergine controls germline stem cell self-renewal and progeny differentiation via distinct mechanisms. *Dev. Cell* 41:157–169.
- Maynard S, Swistowska AM, Lee JW, Liu Y, Liu S-T, Da Cruz AB, Rao M, de Souza-Pinto NC, Zeng X, Bohr VA. 2008. Human embryonic stem cells have enhanced repair of multiple forms of DNA damage. *Stem Cells* 26:2266–2274.
- McCarthy A, Deilulo A, Martin ET, Upadhyay M, Rangan P. 2018. Tip60 complex promotes expression of a differentiation factor to regulate germline differentiation in female *Drosophila*. *Mol. Biol. Cell* 29:2933–2945.
- McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U. S. A.* 36:344–355.
- McClintock B. 1953. Induction of instability at selected loci in maize. *Genetics* 38:579–599.
- McClintock B. 1956. Controlling elements and the gene. *Cold Spring Harbor Symposia on Quant Biol* 21:197–216.
- McGinnis W, Shermoen AW, Beckendorf SK. 1983. A transposable element inserted just 5' to a *Drosophila* glue protein gene alters gene expression and chromatin structure. *Cell* 34:75–84.
- McKay DJ, Klusza S, Penke TJR, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW, Tatomer DC, Lieb JD, et al. 2015. Interrogating the function of metazoan histones using engineered gene clusters. *Dev. Cell* 32:373–386.
- Meiklejohn CD, Blumenstiel JP. 2018. Invasion of the *P* elements: Tolerance is not futile. *PLoS Biol.* 16:e3000036.
- Mérel V, Boulesteix M, Fablet M, Vieira C. 2020. Transposable elements in *Drosophila*. *Mob. DNA* 11:1–20.
- Messina G, Damia E, Fanti L, Atterato MT, Celauro E, Mariotti FR, Accardo MC, Walther M, Verni F, Picchioni D, et al. 2014. *Yeti*, an essential *Drosophila melanogaster* gene, encodes a protein required for chromatin organization. *J. Cell Sci.* 127:2577–2588.
- Mills RE, Andrew Bennett E, Iskow RC, Devine SE. 2007. Which transposable elements are active in the human genome? *Trends Genet.* 23:183–191.
- Mohn F, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* 157:1364–1379.
- Montero L, Müller N, Gallant P. 2008. Induction of apoptosis by *Drosophila* Myc. *Genesis* 46:104–111.
- Moon S, Cassani M, Lin YA, Wang L, Dou K, Zhang ZZ. 2018. A robust

- transposon-endogenizing response from germline stem cells. *Dev. Cell* 47:660–671.e3.
- Morgante M. 2006. Plant genome organisation and diversity: The year of the junk! *Curr. Opin. Biotechnol.* 17:168–173.
- Murga M, Jaco I, Fan Y, Soria R, Martinez-Pastor B, Cuadrado M, Yang S-M, Blasco MA, Skoultchi AI, Fernandez-Capetillo O. 2007. Global chromatin compaction limits the strength of the DNA damage response. *J. Cell Biol.* 178:1101–1108.
- Nehme NT, Liégeois S, Kele B, Giammarinaro P, Pradel E, Hoffmann JA, Ewbank JJ, Ferrandon D. 2007. A model of bacterial intestinal infections in *Drosophila melanogaster*. *PLoS Pathogens* 3:e173.
- Ner SS, Harrington MJ, Grigliatti TA. 2002. A role for the *Drosophila* SU(VAR)3-9 protein in chromatin organization at the histone gene cluster and in suppression of position-effect variegation. *Genetics* 162:1763–1774.
- Niki Y, Mahowald AP. 2003. Ovarian cystocytes can repopulate the embryonic germ line and produce functional gametes. *Proc. Natl. Acad. Sci. U. S. A.* 100:14042–14045.
- Nishida KM, Saito K, Mori T, Kawamura Y, Nagami-Okada T, Inagaki S, Siomi H, Siomi MC. 2007. Gene silencing mechanisms mediated by Aubergine piRNA complexes in *Drosophila* male gonad. *RNA* 13:1911–1922.
- Orian A, van Steensel B, Delrow J, Bussemaker HJ, Li L, Sawado T, Williams E, Loo LWM, Cowley SM, Yost C, et al. 2003. Genomic binding by the *Drosophila* Myc, Max, Mad/Mnt transcription factor network. *Genes Dev.* 17:1101–1114.
- Ota R, Kobayashi S. 2020. Myc plays an important role in *Drosophila* P-M hybrid dysgenesis to eliminate germline cells with genetic damage. *Commun Biol* 3:1–8.
- Ozata DM, Gainetdinov I, Zoch A, O’Carroll D, Zamore PD. 2018. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* 20:89–108.
- Ozawa N, Furuhashi H, Masuko K, Numao E, Makino T, Yano T, Kurata S. 2016. Organ identity specification factor WGE localizes to the histone locus body and regulates histone expression to ensure genomic stability in *Drosophila*. *Genes Cells* 21:442–456.
- Panaud O. 2016. Horizontal transfers of transposable elements in eukaryotes: The flying genes. *C. R. Biol.* 339:296–299.
- Pardue M-L, DeBaryshe PG. 2003. Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu. Rev. Genet.* 37:485–511.
- Parisi MJ, Deng W, Wang Z, Lin H. 2001. The arrest gene is required for germline cyst formation during *Drosophila* oogenesis. *Genesis* 29:196–209.
- Peccoud J, Cordaux R, Gilbert C. 2018. Analyzing horizontal transfer of transposable elements on a large scale: Challenges and prospects. *Bioessays* 40:1700177.
- Post C, Clark JP, Sytnikova YA, Chirn G-W, Lau NC. 2014. The capacity of target silencing by *Drosophila* PIWI and piRNAs. *RNA* 20:1977–1986.

- Potter-Birriel J, Gonsalvez GB, Marzluff WF. 2021. A region of *Drosophila* SLBP distinct from the histone pre-mRNA binding and processing domains is essential for deposition of histone mRNA in the oocyte. *J Cell Sci.* 134:jcs251728
- Prado JRM, Srinivasan S, Fuller MT. 2013. The histone variant His2Av is required for adult stem cell maintenance in the *Drosophila* Testis. *PLoS Genet.* 9:e1003903.
- Qi D, Jin H, Lilja T, Mannervik M. 2006. *Drosophila* reptin and other TIP60 complex components promote generation of silent chromatin. *Genetics* 174:241–251.
- Quesneville H, Bergman CM, Andrieu O, Autard D, Nouaud D, Ashburner M, Anxolabehere D. 2005. Combined evidence annotation of transposable elements in genome sequences. *PLoS Comput. Biol.* 1:166–175.
- Quinlan AR. 2014. BEDTools: The Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinformatics* 47:11.12.1–34.
- Råberg L. 2014. How to live with the enemy: Understanding tolerance to parasites. *PLoS Biol.* 12:e1001989.
- Reya T, Morrison SJ, Clarke MF, Weissman IL. 2001. Stem cells, cancer, and cancer stem cells. *Nature* 414:105–111.
- Riddle NC, Minoda A, Kharchenko PV, Alekseyenko AA, Schwartz YB, Tolstorukov MY, Gorchakov AA, Jaffe JD, Kennedy C, Linder-Basso D, et al. 2011. Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Res.* 21:147–163.
- Robert V, Prud'homme N, Kim A, Bucheton A, Pélisson A. 2001. Characterization of the flamenco region of the *Drosophila melanogaster* genome. *Genetics* 158:701–713.
- Rojas-Ríos P, Chartier A, Pierson S, Simonelig M. 2017. Aubergine and piRNAs promote germline stem cell self-renewal by repressing the proto-oncogene. *EMBO J.* 36:3194–3211.
- Roy BA, Kirchner JW. 2000. Evolutionary dynamics of pathogen resistance and tolerance. *Evolution* 54:51–63.
- Rubin GM, Kidwell MG, Bingham PM. 1982. The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. *Cell* 29:987–994.
- Ruddell A, Jacobs-Lorena M. 1985. Biphasic pattern of histone gene expression during *Drosophila* oogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 82:3316–3319.
- Ruhf ML, Braun A, Papoulas O, Tamkun JW, Randsholt N, Meister M. 2001. The domino gene of *Drosophila* encodes novel members of the SWI2/SNF2 family of DNA-dependent ATPases, which contribute to the silencing of homeotic genes. *Development* 128:1429–1441.
- Ruike T, Takeuchi R, Takata K-I, Oshige M, Kasai N, Shimanouchi K, Kanai Y, Nakamura R, Sugawara F, Sakaguchi K. 2006. Characterization of a second proliferating cell nuclear antigen (PCNA2) from *Drosophila melanogaster*. *FEBS J.* 273:5062–5073.
- Rust K, Tiwari MD, Mishra VK, Grawe F, Wodarz A. 2018. Myc and the Tip60 chromatin

- remodeling complex control neuroblast maintenance and polarity in *Drosophila*. *EMBO J.* 37:e98659.
- Sanmiguel P. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Ann. Bot.* 82:37–44.
- Saretzki G, Armstrong L, Leake A, Lako M, von Zglinicki T. 2004. Stress defense in murine embryonic stem cells is superior to that of various differentiated murine cells. *Stem Cells* 22:962–971.
- Schaack S, Gilbert C, Feschotte C. 2010. Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* 25:537–546.
- Schaefer RE, Kidwell MG, Fausto-Sterling A. 1979. Hybrid dysgenesis in *Drosophila melanogaster*: Morphological and cytological studies of ovarian dysgenesis. *Genetics* 92:1141–1152.
- Serrato-Capuchina A, Wang J, Earley E, Peede D, Isbell K, Matute DR. 2020. Paternally inherited *P*-element copy number affects the magnitude of hybrid dysgenesis in *Drosophila simulans* and *D. melanogaster*. *Genome Biol. Evol.* 12:808–826.
- Shim HJ, Lee E-M, Nguyen LD, Shim J, Song Y-H. 2014. High-dose irradiation induces cell cycle arrest, apoptosis, and developmental defects during *Drosophila* oogenesis. *PLoS One* 9:e89009.
- Sinclair DA, Clegg NJ, Antonchuk J, Milne TA, Stankunas K, Ruse C, Grigliatti TA, Kassis JA, Brock HW. 1998. Enhancer of polycomb is a suppressor of position-effect variegation in *Drosophila melanogaster*. *Genetics* 148:211–220.
- Spradling AC. 1981. The organization and amplification of two chromosomal domains containing *Drosophila* chorion genes. *Cell* 27:193–201.
- Srivastav SP, Kelleher ES. 2017. Paternal induction of hybrid dysgenesis in *Drosophila melanogaster* is weakly correlated with both *P*-element and *hobo* element dosage. *G3* 7:1487–1497.
- Staeva-Vieira E, Yoo S, Lehmann R. 2003. An essential role of DmRad51/SpnA in DNA repair and meiotic checkpoint control. *EMBO J.* 22:5863–5874.
- Sterpone S, Cozzi R. 2010. Influence of *XRCC1* genetic polymorphisms on ionizing radiation-induced DNA damage and repair. *J. Nucleic Acids* 2010:780369.
- Story B, Ma X, Ishihara K, Li H, Hall K, Peak A, Anoja P, Park J, Haug J, Blanchette M, et al. 2019. Defining the expression of piRNA and transposable elements in *Drosophila* ovarian germline stem cells and somatic support cells. *Life Sci. Alliance* 2:e201800211.
- Sun Y, Jiang X, Chen S, Fernandes N, Price BD. 2005. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proc. Natl. Acad. Sci. U. S. A.* 102:13182–13187.
- Sun Y, Jiang X, Xu Y, Ayrappetov MK, Moreau LA, Whetstine JR, Price BD. 2009. Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60. *Nat. Cell Biol.* 11:1376–1382.

- Tasnim S, Kelleher ES. 2018. p53 is required for female germline stem cell maintenance in *P*-element hybrid dysgenesis. *Dev. Biol.* 434:215–220.
- Tauc HM, Tasdogan A, Meyer P, Pandur P. 2017. Nipped-A regulates intestinal stem cell proliferation in *Drosophila*. *Development* 144:612–623.
- Team, TRDC. 2008. The R Project for Statistical Computing.
- Teixeira FK, Okuniewska M, Malone CD, Coux R-X, Rio DC, Lehmann R. 2017. piRNA-mediated regulation of transposon alternative splicing in the soma and germ line. *Nature* 552:268–272.
- Tootle TL, Williams D, Hubb A, Frederick R, Spradling A. 2011. *Drosophila* eggshell production: identification of new genes and coordination by Pxt. *PLoS One* 6:e19943.
- Tudor M, Lobočka M, Goodell M, Pettitt J, O'Hare K. 1992. The *pogo* transposable element family of *Drosophila melanogaster*. *Mol. Gen. Genet.* 232:126–134.
- Uri A, Martha K, Veronika B-I, Anna B, Trudi S. 2007. An essential role for *Drosophila hus1* in somatic and meiotic DNA damage responses. *J. Cell Sci.* 120:1042–1049.
- Usakin L, Abad J, Vagin VV, de Pablos B, Villasante A, Gvozdev VA. 2007. Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in *Drosophila melanogaster* ovaries. *Genetics* 176:1343–1349.
- Vaisnav M, Xing C, Ku H-C, Hwang D, Stojadinovic S, Pertsemliadis A, Abrams JM. 2014. Genome-wide association analysis of radiation resistance in *Drosophila melanogaster*. *PLoS One* 9:e104858.
- Vicient CM, Kalendar R, Anamthawat-Jónsson K, Schulman AH. 1999. Structure, functionality, and evolution of the *BARE-1* retrotransposon of barley. *Genetica* 107:53–63.
- Vilenchik MM, Knudson AG. 2003. Endogenous DNA double-strand breaks: Production, fidelity of repair, and induction of cancer. *Proc. Natl. Acad. Sci. U. S. A.* 100:12871–12876.
- Wallau GL, Vieira C, Loreto ÉLS. 2018. Genetic exchange in eukaryotes through horizontal transfer: connected by the mobilome. *Mob. DNA* 9:1–16.
- Walter MF, Jang C, Kasravi B, Donath J, Mechler BM, Mason JM, Biessmann H. 1995. DNA organization and polymorphism of a wild-type *Drosophila* telomere region. *Chromosoma* 104:229–241.
- Wang Z, Lin H. 2007. Sex-lethal is a target of Bruno-mediated translational repression in promoting the differentiation of stem cell progeny during *Drosophila* oogenesis. *Dev. Biol.* 302:160–168.
- Waring GL. 2000. Morphogenesis of the eggshell in *Drosophila*. *Int. Rev. Cytol.* 198:67–108.
- Wei X, Eickbush DG, Speece I, Larracuente AM. 2021. Heterochromatin-dependent transcription of satellite DNAs in the *Drosophila melanogaster* female germline. *eLife* 10:e62375
- Wickersheim ML, Blumenstiel JP. 2013. Terminator oligo blocking efficiently eliminates rRNA

- from *Drosophila* small RNA sequencing libraries. *Biotechniques* 55:269–272.
- Wolf G, Greenberg D, Macfarlan TS. 2015. Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Krüppel-associated box zinc finger protein family. *Mob. DNA* 6:1–20.
- Wu CI, Lyttle TW, Wu ML, Lin GF. 1988. Association between a satellite DNA sequence and the *Responder of Segregation Distorter* in *D. melanogaster*. *Cell* 54:179–189.
- Wylie A, Lu W-J, D’Brot A, Buszczak M, Abrams JM. 2014. p53 activity is selectively licensed in the *Drosophila* stem cell compartment. *eLife* 3:e01530.
- Xing Y, Su TT, Ruohola-Baker H. 2015. Tie-mediated signal from apoptotic cells protects stem cells in *Drosophila melanogaster*. *Nat. Commun.* 6:7058.
- Xin T, Xuan T, Tan J, Li M, Zhao G, Li M. 2013. The *Drosophila* putative histone acetyltransferase Enok maintains female germline stem cells through regulating Bruno and the niche. *Dev. Biol.* 384:1–12.
- Yan D, Neumüller RA, Buckner M, Ayers K, Li H, Hu Y, Yang-Zhou D, Pan L, Wang X, Kelley C, et al. 2014. A regulatory network of *Drosophila* germline stem cell self-renewal. *Dev. Cell* 28:459–473.
- Yang P, Wang Y, Macfarlan TS. 2017. The role of KRAB-ZFPs in transposable element repression and mammalian evolution. *Trends Genet.* 33:871–881.
- Yannopoulos G, Stamatis N, Monastiriotti M, Hatzopoulos P, Louis C. 1987. *hobo* is responsible for the induction of hybrid dysgenesis by strains of *Drosophila melanogaster* bearing the male recombination factor 23.5MRF. *Cell* 49:487–495.
- Yin H, Lin H. 2007. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. *Nature* 450:304–308.
- Yuan K, O’Farrell PH. 2016. TALE-light imaging reveals maternally guided, H3K9me2/3-independent emergence of functional heterochromatin in *Drosophila* embryos. *Genes Dev.* 30:579–593.
- Zhang H-H, Peccoud J, Xu M-R-X, Zhang X-G, Gilbert C. 2020. Horizontal transfer and evolution of transposable elements in vertebrates. *Nat. Commun.* 11:1–10.
- Zhang S, Pointer B, Kelleher ES. 2020. Rapid evolution of piRNA-mediated silencing of an invading transposable element was driven by abundant de novo mutations. *Genome Res.* 30:566–575.
- Zhu C-H, Xie T. 2003. Clonal expansion of ovarian germline stem cells during niche formation in *Drosophila*. *Development* 130:2579–2588.