The Role of Transposable Elements in TDP-43-Mediated Neurodegeneration.

A Dissertation Presented

by

Lisa Krug

NIH Training Grant Scholar

to

The Watson School of Biological Sciences

at

Cold Spring Harbor Laboratory

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Biological Sciences

Cold Spring Harbor Laboratory

March 2016

TABLE OF CONTENTS:

List of Abbreviations:

A – Adenosine

ADAR – Adenosine deaminase that acts of RNA

Ago2 – Argonaute 2

ALS – Amyotrophic Lateral Sclerosis

ATR – Ataxia telangiectasia and Rad3 related

Aub - Aubergine

ChIP – chromatin immunoprecipiation

Chk-2 – Checkpoint kinase 2

CLIP – crosslinked RIP

CNE – conserved non-coding element

CNCS – conserved non-coding sequence

CNS – central nervous system

CTGS – co-transcriptional gene silencing

DCR-2 – Dicer-2

DCV – *Drosophila* C virus

DD – deaminase domain

dDD – *Drosophila* ADAR deaminase domain

hDD – human ADAR deaminase domain

DDR – DNA damage response

DSB – double strand break, particularly regarding DNA double strand breaks

dsRNA – double stranded RNA

eGFP – enhanced GFP

Env – envelope protein

ERV – endogenous retrovirus

FISH – fluorescence *in situ* hybridization

FTLD – Frontotemporal Lobar Degeneration

FXTAS – Fragile X-Associated Tremor/Ataxia Syndrome

G - Guanosine

GFP – green fluorescent protein

GO – gene ontology

GWAS – genome wide association study

hDD – human ADAR2 deaminase domain

HERV-K – human endogenous retrovirus K family

HIV-1 – human immunodeficiency virus type 1

hnRNP – heterogeneous nuclear ribonucleoprotein

hTDP-43 – human TDP-43

I - Inosine

IR – inverted repeat

L1 – long interspersed nuclear element 1; LINE-1

LexAop – LexA operator

LTR – long terminal repeat

MeCP2 – methyl-CpG-binding protein 2

MITE – miniature transposable element

miRNA – micro RNA

mRNA – messenger RNA

MYA – million years ago

NHEJ – non-homologous end joining

NMD – nonsense-mediated decay

NPG – neural progenitor cells

PCR – polymerase chain reaction

piRNA – Piwi-interacting RNA

PIWI - P-element induced wimpy testis

Pol – Polymerase enzyme

PTGS – post-transcriptional gene silencing

qPCR – quantitative RT-PCR

RAG-1/RAG-2 – Recombination-activating gene 1 / 2

RIP – RNA immunoprecipitation

RISC – RNA-induced silencing complex

RITS – RNA-induced initiation of transcriptional silencing complex

RNA – ribonucleic acid

RNAi – RNA interference

RRM – RNA recognition motif

RSS – recognition signal sequence

RT – reverse transcriptase

RTE – retrotransposable elements

SAHF – senescence-associated heterochromatic foci

SINE – small interspersed non-coding elements

siRNA – small interfering RNA

TAR – trans-activation response element

TD – Targeting domain

TDP-43 – TAR DNA-binding protein 43

TE – transposable elements

TEM – transmission electron microscopy

TF – transcription factor

TIR – terminal inverted repeat

TSS – transcription start site

TUNEL – deoxynucleotidyl transferase dUTP nick end labelling

UAS – upstream activation sequence

UTR – untranslated region

W - Tryptophan

XIC – X-inactivation center

XIST – X-inactive specific transcript

XRCC4 – X-ray repair cross-complementing factor 4

γH2Ax – H2A histone family member X, serine 139 phosphorylated

λN – λ-phage N peptide

List of Tables and Figures:

FORWARD:

In this dissertation I will present data I have collected regarding mechanistic testing of the novel hypothesis that TE activity drives neurodegenerative decline in a subset of neurodegenerative diseases. Specifically, I will demonstrate that TDP-43 pathology, which is highly implicated in the etiology of a spectrum of neurodegenerative diseases that spans ALS and FTLD, induces loss of suppression of TEs. This likely occurs via TDP-43-dependent degradation of the somatic, post-transcriptional TE silencing mechanism. I will show that TDP-43-dependent TE activation causally contributes to physiological decline. Further, I will demonstrate that apoptosis induced by TDP-43 expression is largely mediated by Chk-2 signaling following DNA damage, that DNA damage-mediated apoptosis accounts for much of the toxicity experienced by animals that express hTDP-43, and that TE activation at least partially contributes to this DNA damage-mediated apoptosis. Finally, I will use the data presented herein, recent reports from the literature, and other recent work from the Dubnau lab to build a model by which to interpret the pathological activation of TEs in the context of both normal neurological decline in wild type aging and in the case of TDP-43 protein pathologies. I believe this exercise is likely to prove beneficial to understanding neurodegenerative disease more generally, as TEs have recently been reported to be involved in several other non-TDP-43 related neurodegenerative diseases as well. In order to understand how and why TEs run amok in the aged CNS, however, we must first understand the nature of TEs. This requires more than a simple definition of their genomic features; we must explore the relationship between eukaryotic genomes and their TE residents, the role TEs have played in shaping their host genomes and new findings suggesting that the nervous system specifically has made particular use of TEs. Only then can we place TE activity in the context of normal organismal aging, and begin to address the reasons why they may be particularly susceptible to disregulation in the aged nervous system.

CHAPTER 1 – Introduction:

TEs are ubiquitous, characteristic features of eukaryotic genomes - mobile genetic elements that are capable of replicating themselves via translocation to novel positions in the genome, thus ensuring their continued propagation (Kidwell and Lisch 2001; Brookfield 2005). When allowed to transpose freely, they pose a threat to genome stability by insertional mutagenesis, creating double-stranded DNA breaks, and inducing large-scale genomic duplications, inversions, and deletions (Symer et al. 2002). Eukaryotic organisms have therefore developed interleaved co-transcriptional and post-transcriptional mechanisms by which to ensure TE suppression at inappropriate times (Castel and Martienssen 2013). However, both the raw genomic sequence of TEs and the epigenetic mechanisms that suppress them appear to have been co-opted by their host genomes to serve functions in eukaryotic genome architecture and regulation (Slotkin and Martienssen 2007; Feschotte 2008). This section will begin with a brief overview of the classes of TEs and the mechanisms that eukaryotic systems use to suppress them. I will then elaborate how mobile elements may have facilitated the evolution of the linear chromosome by contributing to centromere and telomere function and the role TEs play in X chromosome-inactivation in placental mammals. I will follow up by describing how the genomic revolution has enabled further elucidation of the myriad ways in which TEs have been exapted for genome regulation in eukaryotes, with a particular emphasis on recent advances in understanding the role TE-derived CNCS plays in regulating the deployment of genetic information, and how TE-driven innovation in these regions may influence morphological and neurological diversity within the vertebrate lineage. Finally, I will discuss our current understanding of the ways in which TEs have been implemented to solve unique biological problems at the organismal level with regard to responding to and interacting with the environment. This includes the well-delineated and distinctive V(D)J recombination mechanism that enables the acquired immune system to adapt to a near-endless and unpredictable influx of foreign antigens, and the more newly

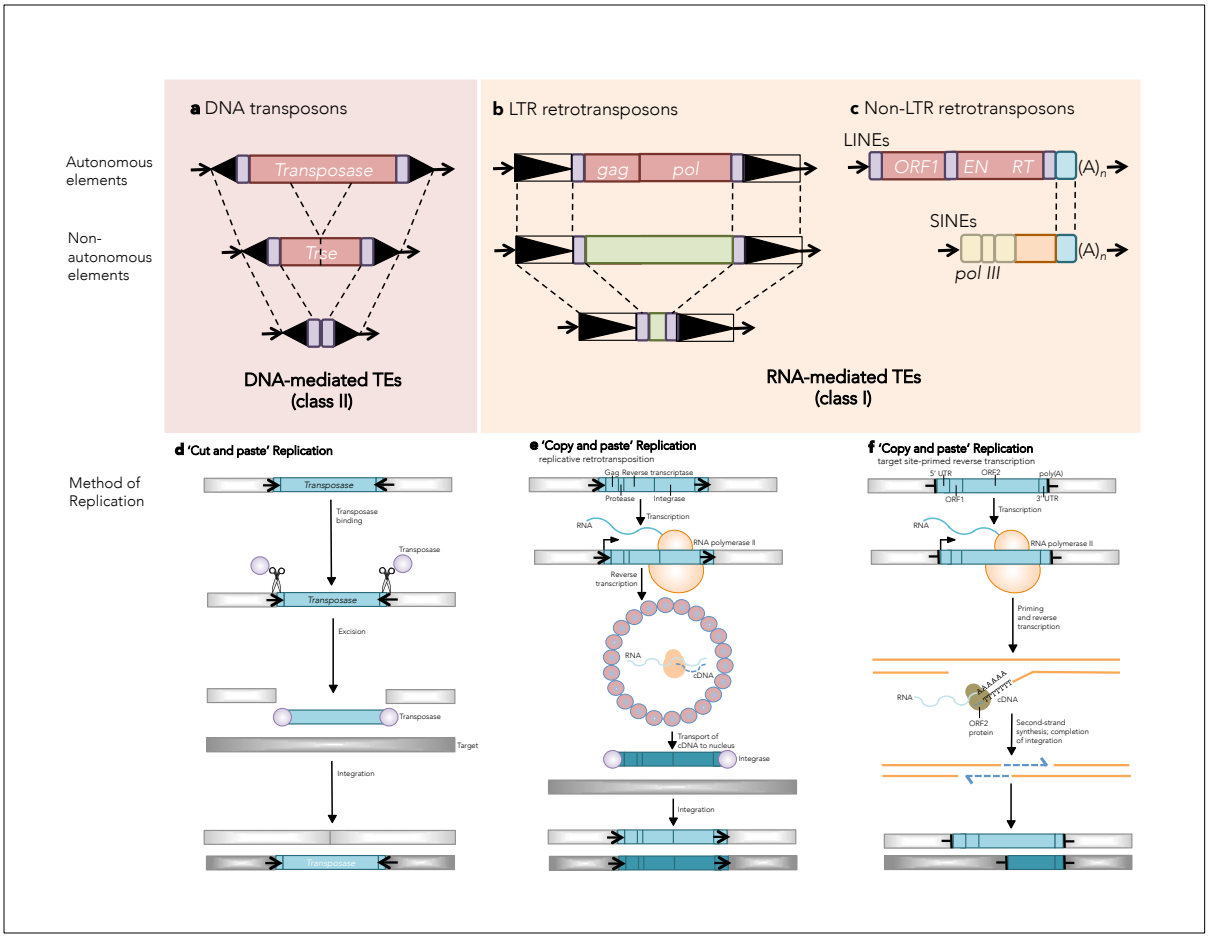


FIGURE 1. Classes of TEs and mechanisms of transposition.
Classes of TEs: (**A**) DNA transposons (Class II), (**B**) LTR and (**C**) Non-LTR RTEs (Class I) [*adapted from*: (Feschotte 2008)].
Mechanisms of transposition: (**D**) 'cut and paste' replication, (**E**) 'copy and paste' replicative transposition, and (**F**) 'copy and paste' target site-primed reverse transcription [*adapted from*: (Levin and Moran 2011)].

described phenomenon of L1 mobilization during neurodevelopment, which results in somatic mosaicism in the adult mammalian brain. As this regulated phenomenon appears to be highly active in brain, with differing levels of somatic transposition in different cell types and *de novo* insertions frequently appearing in close proximity to genes important for nervous system function (Muotri et al. 2005; Coufal et al. 2009; Kuwabara et al. 2009; Muotri et al. 2009; Baillie et al. 2011; Perrat et al. 2013; Upton et al. 2015), it appears that the nervous system has, like so many other facets of eukaryotic biology, made use of the complexity-building functionality of TEs to perform intricate biological tasks.

I. Transposition Mechanisms and Suppression Systems:

TEs can be classified using two sets of criteria. The most common relies on whether the element requires RT activity for mobilization. Elements that require RT activity are called Class I elements, or RTEs (Slotkin and Martienssen 2007) (Figure 1B and 1C). These elements mobilize via a duplicative "copy-and-paste" mechanism in which they pass through an RNA intermediate, generating a new copy of the element with every transposition event (Figure 1E and 1F). Class I elements can be further divided into LTR or non-LTR elements, depending on whether they have LTRs flanking their termini (Figure 1B and 1C). All LTR elements encode Gag and Pol proteins comparable to exogenous retroviruses, but lack the Env protein required for cell exodus. Those RTEs that do encode an Env are denoted ERVs based on their more complete structural homology to exogenous retroviruses (Slotkin and Martienssen 2007). TEs that do not require RT activity for mobilization are termed Class II elements, or DNA TEs. These elements generally encode a transposase enzyme capable of excising the element by recognizing the TIRs at its flanking ends and subsequently inserting the element into a novel position in the genome (Figure 1A). This mechanism is referred to as "cut-and-paste," and frequently leaves behind traces of the element's presence in the form of truncated or duplicated TIR sequences (Slotkin and Martienssen 2007) (Figure 1D). TEs can also be classified based on whether

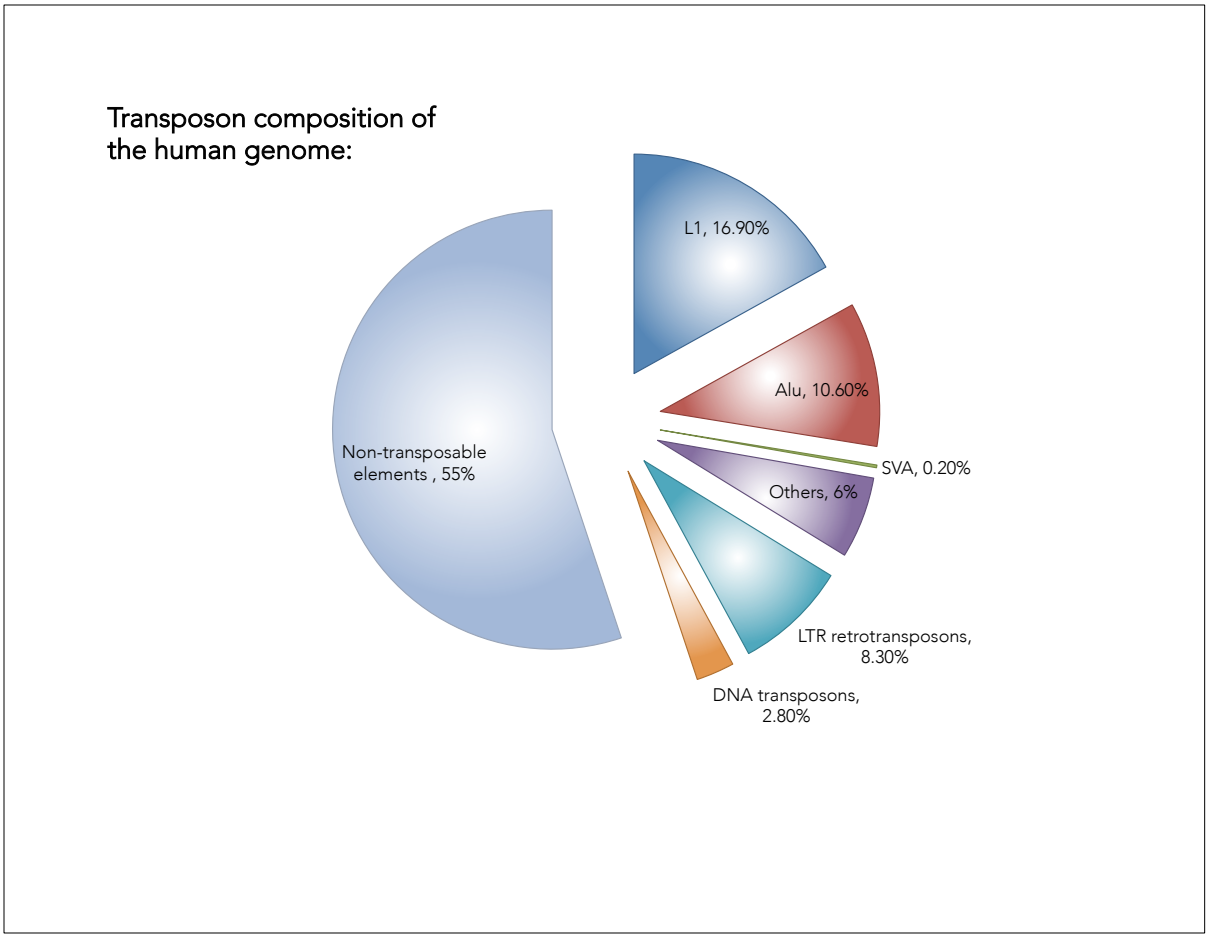


FIGURE 2. TE composition of the human genome.
[*adapted from*: (Cordaux and Batzer 2009)].

they are capable of self-mobilization. The types of elements listed above encode all of the protein machinery required to transpose, and are therefore termed "autonomous elements." Elements containing *cis* regulatory sequences recognized by the machinery of autonomous TEs may also proliferate throughout the genome in *trans* despite their own lack of encoded enzymes (Figure 1A – 1C). Such non-autonomous TEs include SINEs, such as Alu elements in primates (Dewannieux et al. 2003; Garcia-Perez et al. 2007a; Garcia-Perez et al. 2007b) and B2 elements in rodents, and MITEs, which consist of minimal intervening sequence adjoined by two TIRs (Slotkin and Martienssen 2007).

TE mobilization poses an immense threat to the host genome, capable of causing insertional mutations in coding sequences upon arrival, gross chromosomal abnormalities such as large-scale inversions and deletions, and RNA-level toxicity (Kazazian and Goodier 2002; Symer et al. 2002; Kaneko et al. 2011). Given the *trans* activating faculty of transposition machinery and the large amount of real estate they possess in the host genome (Cordaux and Batzer 2009) (Figure 2), it is easy to imagine a scenario in which unchecked run-away transposition completely destroys the host genome. However, such a scenario would be advantageous to neither the host nor to the TEs involved, as their survival relies on being passed from the host to its progeny. As such, eukaryotic organisms have evolved several multilayered mechanisms by which to control and suppress inappropriate transposition. The first consists of epigenetic chromosome-level silencing, including repressive chromatin modifications and even DNA methylation in vertebrates (*see below*). The second line of defense acts post-transcriptionally and usually relies on base complementarity between a small RNA "guide" molecule and the TE transcript to localize RNA degradation machinery to the TE transcript and destroy it before it completes the transposition cycle.

The germ line, which is sequestered early in development (Slotkin and Martienssen 2007), has developed a particular small RNA-based silencing system, collectively called the PIWI-piRNA complex, to shield progeny against specific elements previously experienced within their parental lineage (Figure 3). This system capitalizes on sequence similarity between active TE transcripts and

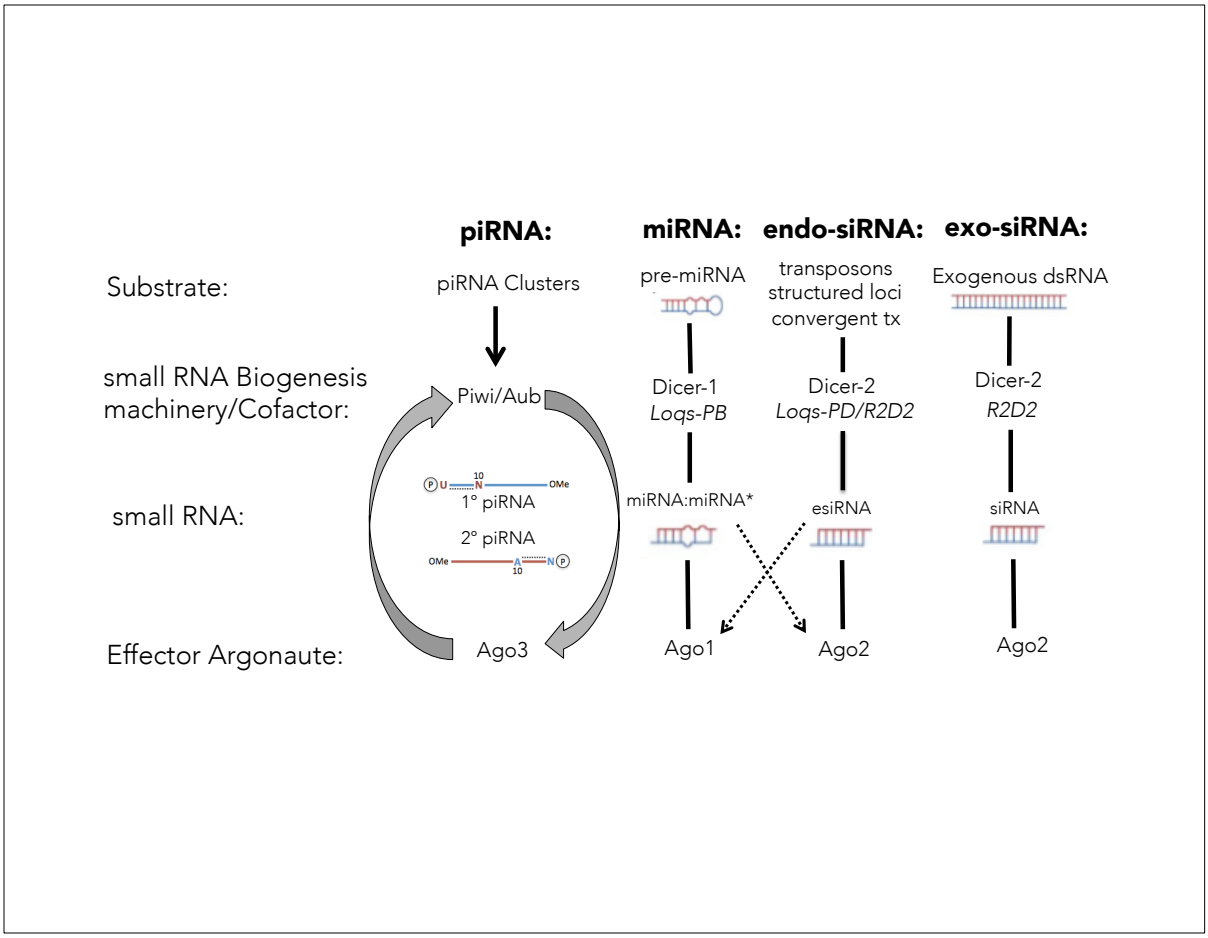


FIGURE 3. small RNA silencing systems in *Drosophila*.
Schematic representations of the components of the piRNA, miRNA, endo- and exo-siRNA pathways in *Drosophila melanogaster* [*adopted from*: (Ghildiyal and Zamore 2009)].

transcripts from piRNA clusters; numerous stretches of genomic DNA of assorted lengths ranging from 1-100 kilobases within which are embedded residual sequence from past transposition events in various orientations (Slotkin and Martienssen 2007). As integration of novel TEs into piRNA clusters results in the creation of new piRNAs (Olovnikov et al. 2013), this system can be thought of as a type of "acquired immunity" for the organism against the TE residents of its genome (Slotkin and Martienssen 2007). In *Drosophila*, antisense transcripts derived from piRNA clusters are loaded onto the PIWI-clade Argonaute protein Aub, which recognizes and cleaves complementary TE transcripts in the cytoplasm, thus recruiting them into the "ping-pong" amplification cycle in conjunction with Argonaute 3 to effectively amplify the antisense piRNA arsenal (Slotkin and Martienssen 2007; Aravin and Hannon 2008; Castel and Martienssen 2013) (Figure 3). This mechanism endows the system with inherent robustness and redundancy that ensures against excessive *de novo* mutagenesis of the next generation. It should be noted that any sequence complementary to the piRNA might operate in this mechanism; therefore a piRNA generated from sequence common to a TE family may silence transcripts from all family members (Kavi et al. 2005). While not identical across species, analogous PIWI-piRNA complex mechanisms have been found across both vertebrates and invertebrates, suggesting that they perform a convergent function (Castel and Martienssen 2013). Indeed, piRNA clusters are known to appear in syntenic locations in the mouse and human genome but are divergent in sequence (Girard et al. 2006), signifying both their conserved function and lineage specificity.

While the PIWI-piRNA complex governs TE protection in the germline, another small RNA system defends somatic tissue against unwanted nucleic acid species. The siRNA or RNAi pathway is triggered by dsRNA, which may be generated by TE transcripts, convergent transcription or transcription of other types of structured loci, or by exogenous virus infection (Figure 3). Perfectly base paired dsRNA is recognized by a Dicer family member, which processively cleaves the dsRNA into 21-30 nucleotide siRNAs (Okamura and Lai 2008). These siRNAs are distinguishable from

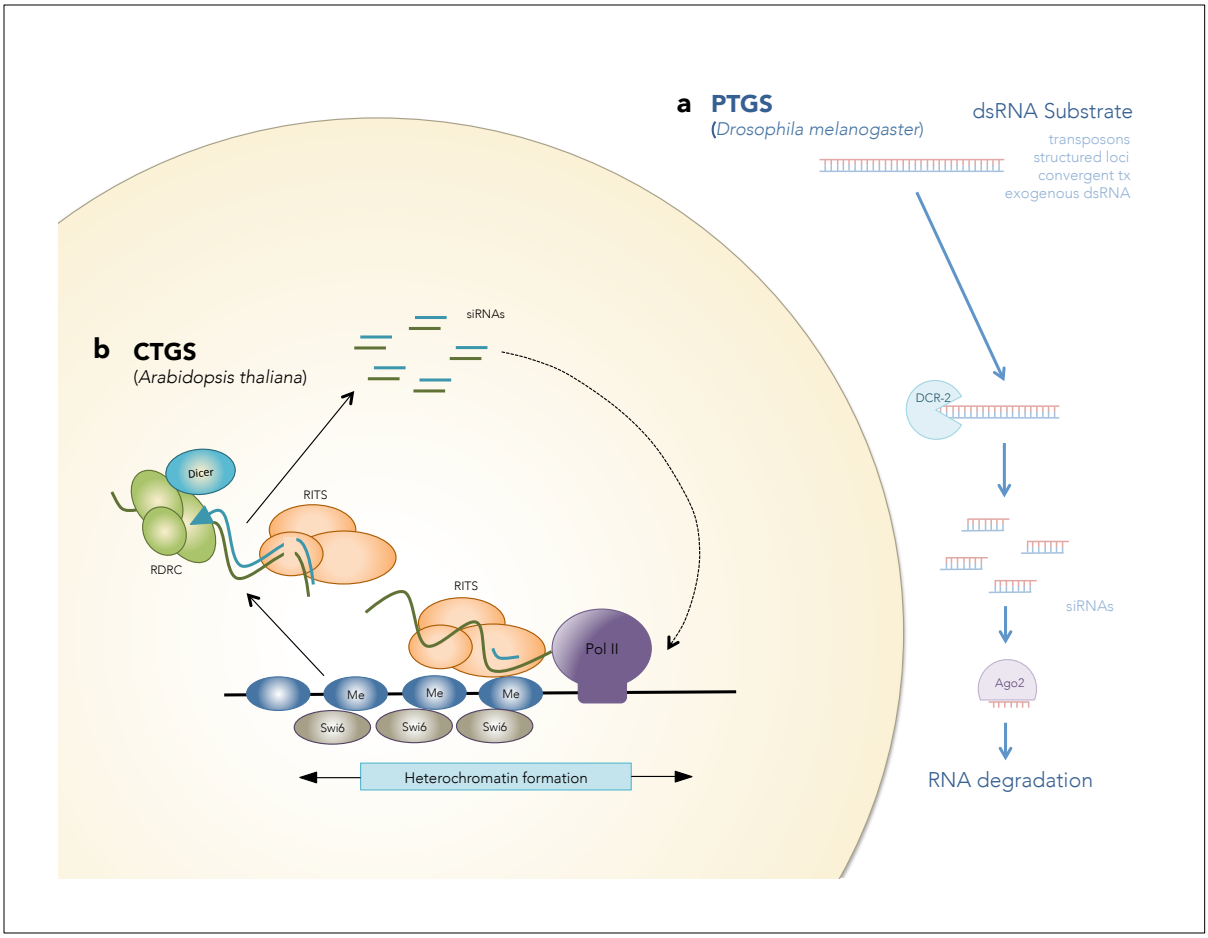


FIGURE 4. **Co-transcriptional and post-transcriptional mechanisms of RNAi.**
(**A**) siRNA-mediated PTGS in *Drosophila melanogaster*; (**B**) siRNA-mediated CTGS as delineated in *Arabidopsis thaliana* [*adopted from*: (Buhler et al. 2006)].

piRNA species based on differences in length and end modifications (Vagin et al. 2006). One strand of this siRNA duplex is then loaded onto an Argonaute family member, forming the core of the RISC, where the siRNA guide molecule recognizes the target RNA species via sequence complementarity and the slicer activity of the Argonaute protein hydrolyzes the target, resulting in RNA degradation (Figure 3). TEs have been observed to produce siRNAs in many species (Vagin et al. 2006; Slotkin and Martienssen 2007; Ghildiyal and Zamore 2009), and TE-siRNA levels have been shown to affect TE activity (Lippman et al. 2003; Sijen and Plasterk 2003; Yang and Kazazian 2006; Slotkin and Martienssen 2007). Disrupting the siRNA pathway results in increased TE transcripts (Svoboda et al. 2004; Czech et al. 2008; Li et al. 2013) as well as increased novel insertions in the genome (Li et al. 2013; Xie et al. 2013). While there have been reports of PIWI family member protein expression in somatic tissue and correlation between expression of these proteins and TE activity (Perrat et al. 2013), the body of evidence indicates that the siRNA pathway is primarily responsible for silencing TEs in somatic tissue. Given that the siRNA-RISC complex acts on cytoplasmic RNA species, this system is generally referred to as PTGS (Castel and Martienssen 2013) (Figure 4A).

Members of the RNAi pathway can also participate in CTGS in which the RITS is recruited to nascent transcripts again via sequence complementarity to a guide siRNA (Castel and Martienssen 2013) (Figure 4B). RITS was first discovered in the fission yeast, *Schizosaccharomyces pombe*, where the RITS complex is composed of an siRNA-loaded Argonaute protein as well as the Argonaute- and chromatin-interacting protein Tas3 and the chromodomain protein Chp2, which assists in localizing RITS to heterochromatin via binding to H3K9me (Verdel et al. 2004; Buhler et al. 2006). RITS in turn recruits factors that generate repressive chromatin structures, including a methyltransferase, a histone deacetylase, and an HP1 ortholog (Buhler et al. 2006). RITS is believed to act via a feed-forward mechanism such that its association with heterochromatin couples with dsRNA synthesis to reinforce or even to spread heterochromatic marks (Buhler et al. 2006; Castel and Martienssen 2013) (Figure 4B). Importantly, both read-through transcription and Argonaute's slicer

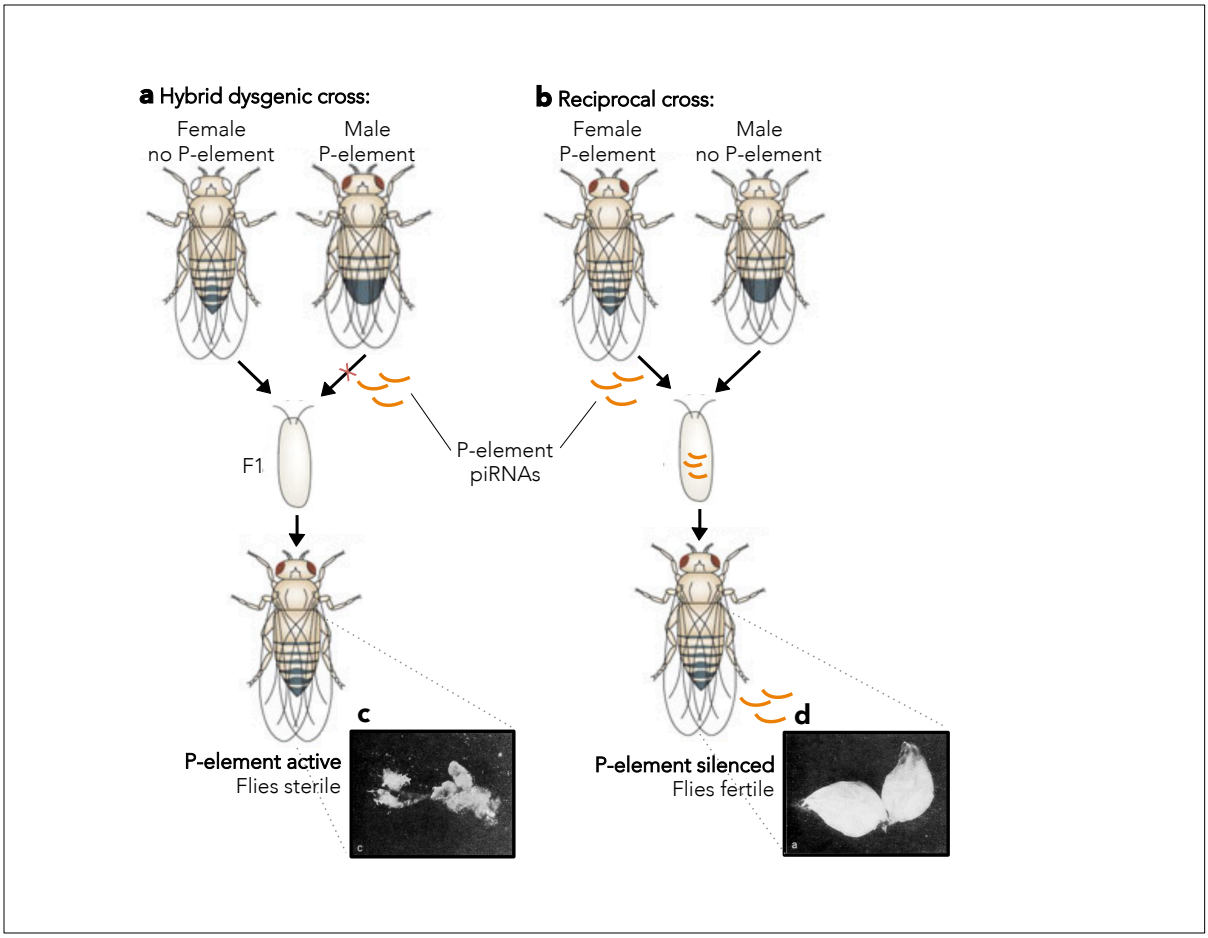


FIGURE 5. The destructive effects of un-checked transposition: P-element hybrid dysgenesis in *Drosophila*.
(**A**) Hybrid dysgenic cross, female F1 progeny are sterile; (**B**) reciprocal cross, female F1 progeny are fully fertile [*adapted from*: (Siomi et al. 2011)]. (**C**) Ovaries of hybrid dysgenic F1 female progeny are completely destroyed, while (**D**) ovaries of the F1 female progeny from the reciprocal cross develop normally [*adapted from*: (Schaefer et al. 1979)].

activity are required for nucleation and spreading of heterochromatin, at least in *S pombe* (Irvine et al. 2006). This mechanism seems to be generally conserved across Eukarya, as components of the piRNA pathway – including Piwi, Aub, and Spindle-E - have all been implicated in heterochromatic gene silencing in *Drosophila* (Pal-Bhadra et al. 1999; Pal-Bhadra et al. 2002; Kogan et al. 2003; Pal-Bhadra et al. 2004; Kanellopoulou et al. 2005), while CTGS of endogenous promoters by synthetic shRNAs has been observed in mammalian cells (Morris et al. 2004; Taira 2006).

A remarkable example of the destructive potential of unchecked TE activity is the phenomenon of hybrid dysgenesis in the fruit fly, *Drosophila melanogaster* (Figure 4). This system relies on the activity of piRNAs in the germline, which are maternally inherited in the fly (Castro and Carareto 2004). As previously mentioned, piRNAs act as a type of acquired immunity for the F1 progeny against those TEs experienced in the mother's genome and her predicessors. Therefore, if a cross is set up in which the mother is carrying a specific TE, in this case the P-element TE, but the male is not, the F1 progeny receive both the P-element and the piRNA immunity against it. Their ovaries develop normally and they are perfectly fertile (Schaefer et al. 1979; Castro and Carareto 2004) (Figure 5B and 5D). However, if the reciprocal hybrid dysgenic cross is set up in which the parental male carries P-element but the parental female does not, the F1 progeny still receive the P-element but do not receive any piRNAs directed against P-element because it was not present in the mother's genome. In this case the female F1 progeny are infertile and, in fact, the tissue of their ovaries is completely destroyed (Schaefer et al. 1979; Castro and Carareto 2004; Brennecke et al. 2008) (Figure 5A and 5C).

Eukaryotes rely on interleaved mechanisms to control TE mobilization, all commonly organized around the basic concept of implementing base pairing interactions of a short RNA species to guide enzyme activity in order to inactivate the element, be it at the level of transcript degradation or repressive chromatin organization. However, mutation of existing elements or acquisition of new elements renders them essentially invisible to the host's surveillance machinery, consequently allowing them to transpose freely. Furthermore, it seems that TEs are not absolutely silenced. The

CTGS process actually requires transcription (Irvine et al. 2006; Castel and Martienssen 2013), while novel germline insertions of TEs are commonly observed in genomic sequences from different individuals (Deininger and Batzer 2002; Ostertag et al. 2003). When this happens, such elements provide the genetic variability substrate to be acted on by natural selection (Zhu et al. 2014).

II. Exaption of Transposable Elements:

Britten and Kohne were the first to demonstrate that eukaryotic genomes are comprised of two distinct subpopulations (Britten and Kohne 1968). This was determined by plotting the kinetics of DNA reassociation on a "COT-curve," where the fraction of reassociated DNA is plotted against the parameter COT, which represents the product of the DNA concentration in the solution times the time of incubation (moles of DNA by seconds per liter). A rapidly reassociating, repetitive fraction and a much more slowly reassociating, non-repetitive fraction were observed. The two authors further explored evolutionary relationships in DNA reassociation rates, finding that reassociated strands of DNA from different species display decreasing thermal stability the longer the two species have been separated in evolutionary time (Britten and Kohne 1968). From this work they determined that mammalian genomes are composed of an estimated 20-35% repetitive DNA – a gross underestimation in comparison to what we know today, but vastly more than was accepted at the time (Marino-Ramirez et al. 2005). Britten and colleagues were some of the first, and loudest, advocates that such repetitive DNA would serve some type of structural or regulatory role, roundly admonishing the conceptualization of repetitive DNA as "junk" (Britten and Kohne 1968; Britten and Davidson 1969; Marino-Ramirez et al. 2005). There was just too much of it, conserved across too many highly divergent species, for it to simply be left behind by neutral evolutionary drift.

The terms "mutable loci," "unstable genes," and "position effects" had all been used in reference to various mysterious genic effects in *Drosophila* before 1950 (McClintock 1984), but the seminal work of Barbara McClintock identified the genetic factors responsible for these phenotypes. She termed them "controlling elements" based on their ability to influence the expression of

neighboring pigment genes in *Zea mays* (McClintock 1950). These elements have been popularized as "selfish" or "parasitic" elements, a reputation that was only imparted to them in the late 1980's as a greater understanding of the dangers these elements can pose to their host began to be further understood (Doolittle and Sapienza 1980; Orgel and Crick 1980; Hickey 1982; Schmid 2003). It is inherent in their nature that their unbounded activity be negatively correlated with host fitness (Slotkin and Martienssen 2007), and that their relative evolutionary success be founded on their replication occurring at a faster rate than the host genome that carries them (Kidwell and Lisch 2001; Brookfield 2005). However, it is also true that they have become dependent on their hosts for their propagation, therefore any system in which their proliferation is allowed to exceed the bounds of host tolerance probably has not withstood the test of evolutionary time.

The dynamics of the host-TE interaction affect our ability to observe and document the process, as current techniques can only capture a static picture of one instant in time. As such, our vision is clouded by recent transposition events, which have yet to be acted on by natural selection and therefore have, as yet, to reveal the ways in which they may prove beneficial to the host; while more ancient events – particularly those derived from now-defunct founder elements – begin to have their identifying features wiped away by neutral mutations until they are all but unrecognizable (Feschotte 2008). However, advances in comparative genomics have gleaned some insight into the multitudinous roles these enigmatic elements play in the regulation of our genome, and their primacy in the evolution of eukaryotic genomes as we observe them today.

II.1: TEs in Genome Evolution:

TEs may provide the substrate for epigenetic regulatory invention in genome evolution as conspicuous similarities can be drawn between epigenetic mechanisms used to repress TEs and those employed by the eukaryotic cell to regulate chromosome function. Epigenetic modifications can occur at one of two levels – the first is at the level of the histone proteins, which package and organize DNA inside the nucleus. Histones have N-terminal tails rich in lysine residues that are

targets of post-translational modifications, including acetylation, methylation, and phosphorylation, which are often collectively referred to as the "histone code" in reference to their combinatorial implementation (Wood et al. 2010; Rothbart and Strahl 2014). Such modifications may then either recruit TFs and polymerases to promote transcription in the case of activating modifications, usually found in areas of "open" euchromatin, or recruit repressive chromatin complexes in the case of repressive modifications in areas of "closed" heterochromatin (Slotkin and Martienssen 2007; Wood et al. 2010; Castel and Martienssen 2013; Rothbart and Strahl 2014). Repressive histone modifications such as H3 lysine 9 methylation (H3K9me) are often found to be associated with TEs, and RNAi components are required for the nucleation and spread of heterochromatin (Slotkin and Martienssen 2007). The second form of epigenetic modification that can occur in many organisms is cytosine methylation of the DNA sequence itself. These types of modification provides a heritable form of "epigenetic memory" due to the ability of symmetrical DNA methylation and histone modifications to be passed on to daughter cells after DNA replication and cell division (Yoder et al. 1997). As TEs are silenced in host cells at both the level of histone modification and DNA methylation, it has been hypothesized that TEs may have acted as the chief driving force in epigenetic regulatory evolution (Henikoff and Matzke 1997; Miller et al. 1999; Wolffe and Matzke 1999; Slotkin and Martienssen 2007), possibly contributing to the two major macroevolutionary transitions in the history of life: chromatin formation at the prokaryotic/eukaryotic transition and DNA methylation at the invertebrate/vertebrate transition (McDonald 1998; Miller et al. 1999).

Such basic roles in chromosome functionality have lead some to suggest that TEs have come to comprise key components of the "regulatory toolkit" of the genome (Slotkin and Martienssen 2007), and perhaps have contributed to evolution of the linear chromosome itself. Indeed, TEs can act as nucleation centers for facultative heterochromatin. This type of heterochromatin can propagate linearly for up to 10 kilobases, potentially interfering with neighboring genes (Sun et al. 2004). Strikingly, the PEV phenomenon (Figure 6A) in which a reporter gene is silenced in some cells and not others due to insertion proximity near heterochromatin, resulting in a variegated phenotype, is

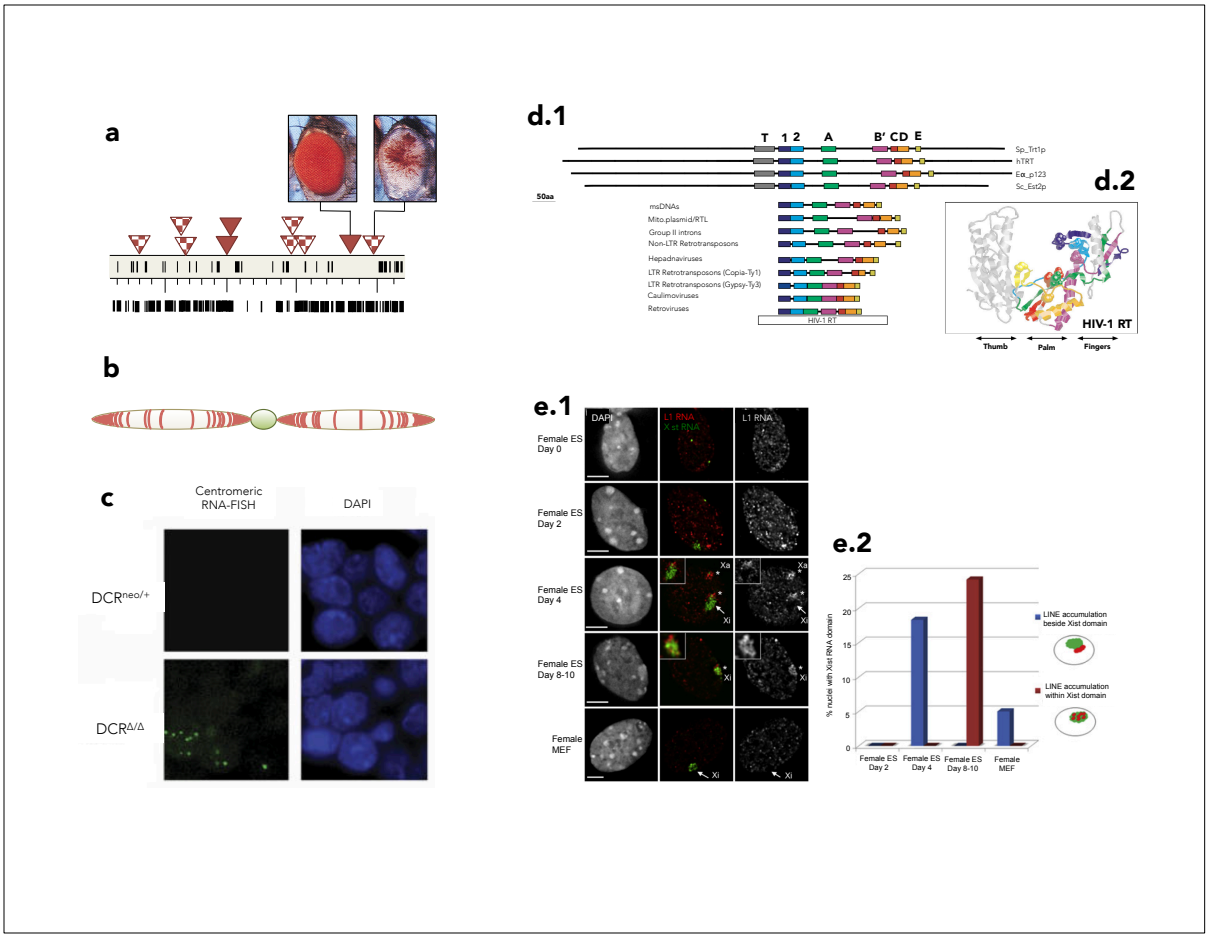


FIGURE 6. The multifaceted roles TEs perform in maintaining global genomic architecture and dosage compensation.
(**A**) Position effect variegation in *Drosophila*; transgene expression is related to insertion site proximity to genomic TEs [*adapted from*: (Slotkin and Martienssen 2007)]. (**B**) Schematic of TE density over the length of the chromosome; high density of TE sequence (red) resides in constitutive pericentromeric and peri-telomeric heterochromatin. (**C**) DICER deficient mouse embryonic stem cells demonstrate decondensation of centromeres; DICER deficient cells display notable binding of RNA FISH probes to centromeric DNA repeat sequences [*adapted from*: (Kanellopoulou et al. 2005)]. (**D**) Deep evolutionary conservation of telomerase in eukaryotes and its relationship to exogenous retroviruses, endogenous RTEs, and group II introns (**D.1**). Yeast *Schizosaccharomyces pombe* Trt1p (Sp_Trt1p); *Homo sapien* TRT (hTRT); ciliate *Euplotes aediculatus* p123 (Eα_p123); and yeast *Schisosaccharomyces cerevisiae* Est2p (Sc_Est2p). (**D.2**) Colored domains from (**D.1**) are mapped onto the structure of HIV-1 RT. [*adapted from*: (Nakamura et al. 1997)]. (**E**) Longitudinal FISH demonstrates close association of L1 transcripts with XIST RNA and subsequent silencing of L1 transcription in female embryonic stem cells (**E.1**). (**E.2**) Quantification of L1 RNA FISH signal near to and within XIST domain over time [*adapted from*: (Chow et al. 2010)].

dependent on both histone methyltransferases and Argonaute proteins, implicating the same underlying RNAi-based silencing mechanisms as those used to silence TEs (Schotta et al. 2003; Haynes et al. 2006; Slotkin and Martienssen 2007). Intriguingly, TEs have become integral components of the constitutive heterochromatic regions flanking both centromeres and telomeres (Figure 6B), whose correct function is absolutely required for the fundamental purposes of trafficking chromosomes during cell division and counteracting chromosome truncation following replication, respectively (Slotkin and Martienssen 2007).

Centromeres generally consist of long tandem arrays of satellite DNA, or simple short sequence repeats, which are bounded by pericentric regions enriched for TEs (Dawe and Henikoff 2006; Slotkin and Martienssen 2007). In humans, the satellite regions are free of TEs while the pericentric areas carry long blocks of LINE and SINE elements (Schueler and Sullivan 2006). The physical DNA composition of pericentric regions differs across species but generally reflects lineage specific TE families; therefore it is the conserved epigenetic context of these regions that imbues them with their functionality (Dawe and Henikoff 2006). Current techniques have been unable to determine whether TEs preferentially target pericentric regions or whether they simply have a higher propensity to be retained there due to a reduced rate of recombination (Dawe and Henikoff 2006; Slotkin and Martienssen 2007). It has been shown in *S pombe* that RNAi mutants accumulate both forward and reverse centromeric transcripts (Volpe et al. 2002) while DICER mutant mammalian cell lines display both condensation and differentiation defects (Kanellopoulou et al. 2005) (Figure 6C), implicating a key role of RNAi in maintaining pericentric heterochromatic structures. Reactivation of silenced centromeric TEs has been shown to espouse both chromosome segregation and meiotic defects in the mouse (Peters et al. 2001; Bourc'his and Bestor 2004; De La Fuente et al. 2006), while an essential role of TEs has been found in maintaining constitutive centromeric heterochromatin in both yeast and plants (Slotkin and Martienssen 2007). It remains possible that TEs could have given rise to the satellite repeats of centromeres themselves. In many species these tandem repeats have

sequence homology to described TEs, although they have lost both their mobility and many structural components likely due to functional constraints of this acquired role (Dawe and Henikoff 2006).

While TEs are also a sizeable component of the constitutive heterochromatin flanking telomeres (Slotkin and Martienssen 2007), the RT activity of the telomerase enzyme itself bears a striking resemblance to RT enzymes from non-LTR RTEs (Nakamura et al. 1997; Slotkin and Martienssen 2007) (Figure 6D.1 and 6D.2). Telomerase displays deep conservation within the eukaryotic lineage, through such evolutionarily distant organisms as protozoans, fungi, and mammals, suggesting that this activity may have been present among even ancestral eukaryotic founders (Nakamura et al. 1997) (Figure 6D.1 and 6D.2). Captivatingly, *Drosophila* carry two domesticated non-LTR RTEs in order to maintain their telomeres, named HeT-A and TART (Pardue et al. 2005). In fact, there are many mechanistic similarities between the RNA-templated activity of the telomerase ribonucleoprotein and non-LTR retrotransposition. Non-LTR elements are typically reverse transcribed directly at the site of integration using nicks in the DNA, usually taking advantage of an exposed 3' hydroxyl to prime reverse transcription directly into the chromosome. Neither telomerase RT nor the HeT-A/TART system require nicked DNA, as they are most likely reverse transcribing directly onto the end of the chromosome (Eickbush 1997). While telomerase templates are generally 5-20 nucleotides in length, *Drosophila* templates consist of the much longer HeT-A and TART sequences, resulting in motley head-to-tail arrays of 5' truncated and full-length elements. The similarities between these two systems, and the deeply conserved homology of telomerase RT, beg the question of whether retrotransposition or telomerase came first in the evolution of the linear chromosome (Eickbush 1997; Miller et al. 1999; Pardue et al. 2005). If, in fact, an ancestral non-LTR RTE gave rise to the cellular machinery we now observe to be prerequisite to the replication of linear chromosomes, this may be the earliest case of molecular domestication shaping eukaryotic genomes. It is curious to consider that Barbara McClintock first described TE activation in response to broken chromosome ends, revealing a specific genetic response to this type of genome stressor (McClintock 1978; Pardue et al. 2005).

TEs have become important players in dosage compensation of X-linked genes as well. In female Eutherian (placental) mammals that carry two X chromosomes, as opposed to an X and a Y chromosome, dosage compensation in achieved via random mosaic inactivation and heterochromatinization of one of the X chromosomes (Ogawa et al. 2008; Minajigi et al. 2015). The development of this system appears to correlate with acquisition of XIST and enrichment for LINE/L1 element density as Metatherian (marsupial) mammals do not carry the XIC responsible for producing XIST and their X chromosomes are actually less TE-dense than their autosomes. Metatherians in turn constitutively inactivate the paternal X chromosomes in females, which is currently believed to be the ancestral therian dosage compensation strategy (Mikkelsen et al. 2007). In the early Eutherian female embryo however, random transcription from the XIC results in the non-coding XIST gene transcript coating one X-chromosome in *cis*, thus initiating silencing of most of the genes on the X chromosome. The high density of LINE elements present on the X chromosome are believed to aid in efficient spreading of heterochromatin away from the XIC (Figure 6E.1 and 6E.2), a theory supported by evidence that if the XIC is translocated to an autosome, the distal spread of heterochromatin is patently less efficient (Lyon 2006). It is possible that low recombination rates between the two sex chromosomes have resulted in the high density of LINE elements observed on the X chromosome, as the Y chromosome displays higher than average LINE density as well (Mikkelsen et al. 2007).

II.2: TEs in Genome Regulation:

Understanding TE-host interactions in the pre-genomic era constituted investigation of individual elements, resulting in hand-curated lists of TEs that had acquired some type of cellular role – a process termed exaptation by Gould and Vrba in 1982 (Gould 1982; Lowe et al. 2007). And the number of different types of roles these elements have been described to fill is truly astounding: from providing distal enhancers (Bejerano et al. 2006; Santangelo et al. 2007) and alternative splice sites (Nekrutenko and Li 2001) to TF binding sites (Thornburg et al. 2006). Like shining a flashlight into a

dark room, this type of investigation began to shed some light on how truly versatile TEs have the capacity to be. The rise of the genomic era has since allowed scientists to observe a whole genome's worth of information from various points along phylogenetic trees, and to begin to determine the extent of TE exaptation by uncovering TE-derived sequences that have become fixed under constrained functional selection over long periods of evolutionary time. Such investigations have proven that TEs can, and frequently are, utilized by the host, and the picture emerging from such investigations paints TEs as a driving force behind evolutionary innovation and speciation (Miller et al. 1999; Marino-Ramirez et al. 2005; Mikkelsen et al. 2007).

On the one hand, TEs can provide pure coding sequence that may come to serve the host as novel genes. A particularly striking phenomenon that illustrates this point is that disparate mammalian species have co-opted Env proteins from different ERVs to serve convergent functions in placental development (Prudhomme et al. 2005; Dunlap et al. 2006). Furthermore, the genomes of many higher organisms contain large quantities of intronless "retrogenes;" pseudogenes which carry the hallmarks of arising via retrotransposition that by some estimates may constitute 25-50% of all active genes. Of particular note, the most abundant intronless genes in higher eukaryotes are potassium channel family and G protein-linked receptor family members, suggesting involvement of retrotransposition in the development of complexity and diversity in the vertebrate nervous system (Betran et al. 2002; Brosius 2003). Additionally, TEs can provide a platform for large-scale genomic duplications, deletions, and inversions during DNA replication due to their repetitive nature and the nature of transposition intermediates (Brosius 2003; Mikkelsen et al. 2007). Indeed, evolutionary patterns of acquired segmental duplications seem to exhibit intervals of relative quiescence punctuated by phases of heightened activity, similar to observed patterns of retrotransposition (Brosius 2003). This pattern has led some to speculate whether periods of active transposition may "regulate the global tempo of phenotypic change" (Rando and Verstrepen 2007). RTEs are also frequently found at the junctions of segmental duplications, as DNA repair processes are capable of capturing and inserting cDNA into DNA damage loci (Moore and Haber 1996; Teng et al. 1996).

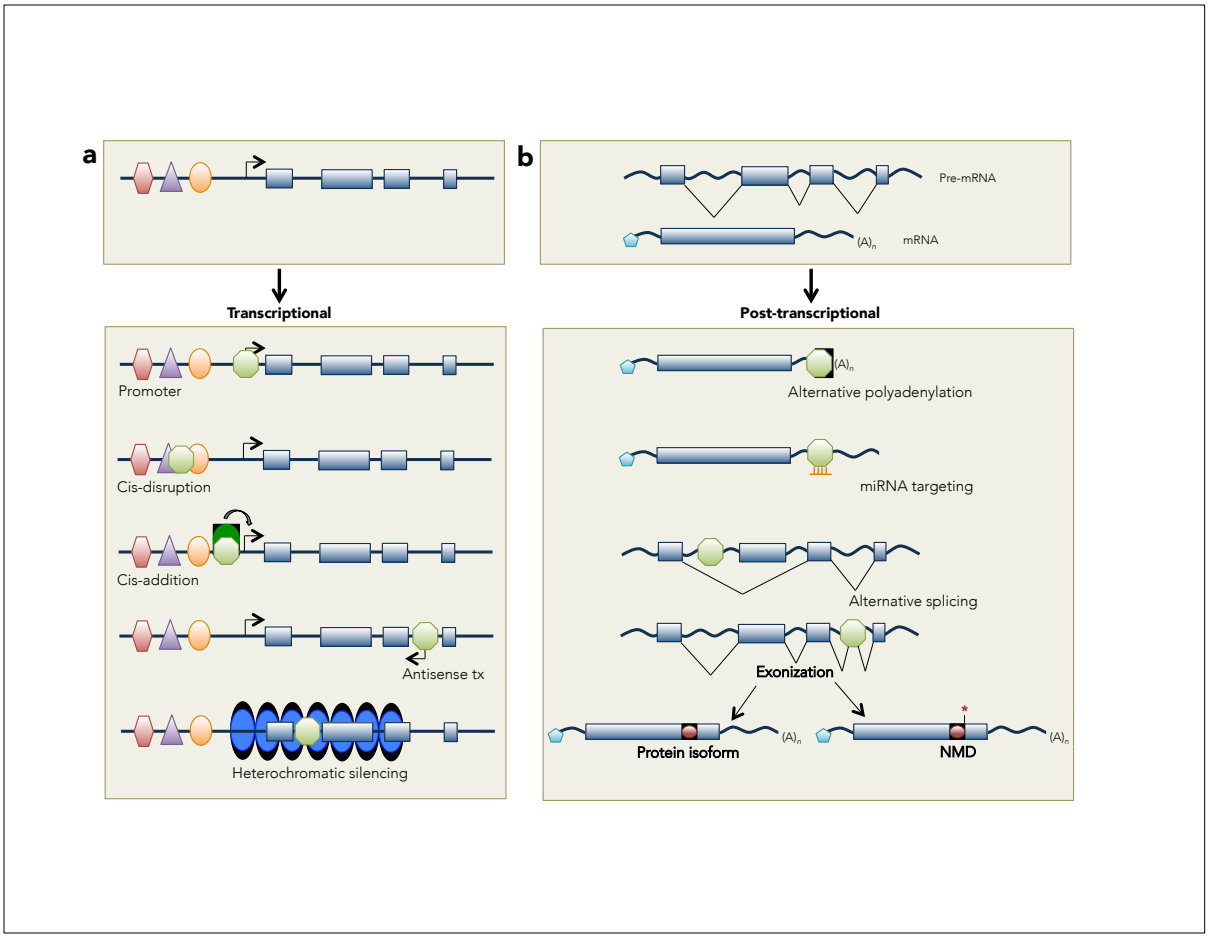


FIGURE 7. Mechanisms by which TEs may exert control over endogenous gene expression. Possible mechanisms by which TEs can influence gene expression at the (A) transcriptional and (B) post-transcriptional level [*adapted from*: (Feschotte 2008)].

Paralogues generated by segmental duplication provide a substrate for genic invention, as one copy is now liberated from functional constraint and free to acquire mutations and explore evolutionary space. Such duplicated genes appear to adhere to the "rapid gene birth and death model" (Consortium 2004; Mikkelsen et al. 2007), either functionally diverging in response to positive selection or promptly disintegrating due to lack of evolutionary benefit. However, this is by far the least numerous class of exapted TEs.

In 1969, Britten and Davidson first put forth their theory of gene network evolution in higher eukaryotes, which prominently featured McClintock's "controlling elements" (McClintock 1950; Britten and Davidson 1969; Feschotte 2008). Specifically, they hypothesized that repetitive elements had the potential to distribute a regulatory cassette throughout the genome, thus causing an assemblage of genes to become coregulated (Britten and Davidson 1969; Britten and Davidson 1971; Lowe et al. 2007; Feschotte 2008). Indeed, evidence for such a relationship between a family of LTR RTEs and p53 master regulator binding sites has been uncovered (Wang et al. 2007). Other work has confirmed that TEs do in fact carry *cis* regulatory elements that are designed to interact with host *trans* factors, such as promoter and enhancer motifs, splicing and polyadenylation signals, and TF binding sites (Feschotte 2008; Batut et al. 2013). Insertion of TEs can influence endogeneous gene expression at the transcriptional level by providing alternative promoters, physically disrupting a coding sequence, introducing a new *cis* regulatory element such as a novel TF binding site, inducing antisense transcription if inserted into an intron, or nucleating heterochromatin formation (Figure 7A). At the post-transcriptional level, TE insertion into coding sequence can introduce novel alternative polyadenylation signals, premature stop codons, miRNA target sites, alter splicing to result in either intron retention or exon skipping, or, in the case of TEs containing cryptic sites, result in exonization leading to potential new protein isoforms (Figure 7B). Documented examples exist for most of these various possibilities (Feschotte 2008). A substantial fraction of 5' and 3' UTRs appear to contain TE-derived sequences (Jordan et al. 2003). However, the contribution of TEs to genome regulation need not necessarily be restricted to insertion into direct coding sequence. The complete genome sequence

of *Caenorhabditis elegans* has demonstrated that the majority of RTEs are located in close proximity to host genes, implying a regulatory role (Ganko et al. 2003), while a survey of RTEs in *S pombe* found a disproportional number of elements associated with pol II promoters (Bowen et al. 2003). In fact, almost 25% of all experimentally characterized human proximal promoter regions ≤500 bp upstream of the TSS contain TE-derived sequences, including empirically validated *cis*-regulatory elements (Jordan et al. 2003).

The genomes of complex metazoans harbor vast expanses of CNCS that, particularly in vertebrates, serve to organize chromatin domains and regulate the expression of neighboring genes. As such, it has been hypothesized that evolution of these regulatory regions underlies the complex morphological diversity observed amongst vertebrates (King and Wilson 1975; Carroll 2005). Examining patterns of CNEs within such sequence that appear to be under purifying selection provides insight into what types of roles repetitive DNA may be co-opted to perform, even once it has devolved past our ability to detect its origin. Comparative genomic analyses between Eutherians have established that CNEs, not protein coding genes, constitute the majority of conserved sequence (Waterston et al. 2002; Gibbs et al. 2004; Consortium 2005; Lindblad-Toh et al. 2005). In 2007, Mikkelson et al. reported the compilation of the first high quality draft of a Metatherian genome, providing a unique, well positioned outgroup to complement ongoing investigations of mammalian genome evolution implementing comparative genomic approaches within the Eutherian lineage. The results of their analyses confirm that the majority of innovation has occurred in CNEs postdating the divergence of Eutheria and Metatheria, with novelties in protein coding genes being rather rare. Recognizable TE sequence was found to account for at least 16% of Eutherian-specific CNEs (Mikkelsen et al. 2007), demonstrating that TEs have had a much greater role in genomic innovation than previously documented. Due to difficulties associated with recognizing TE sequences older than ~100-200 MYA (Waterston et al. 2002; Gentles et al. 2007), this is likely to be a considerable underestimate. Amazingly, 99% of all CNEs were found in both human and opossum, suggesting that they perform such basic functions that they cannot be lost (Mikkelsen et al. 2007).

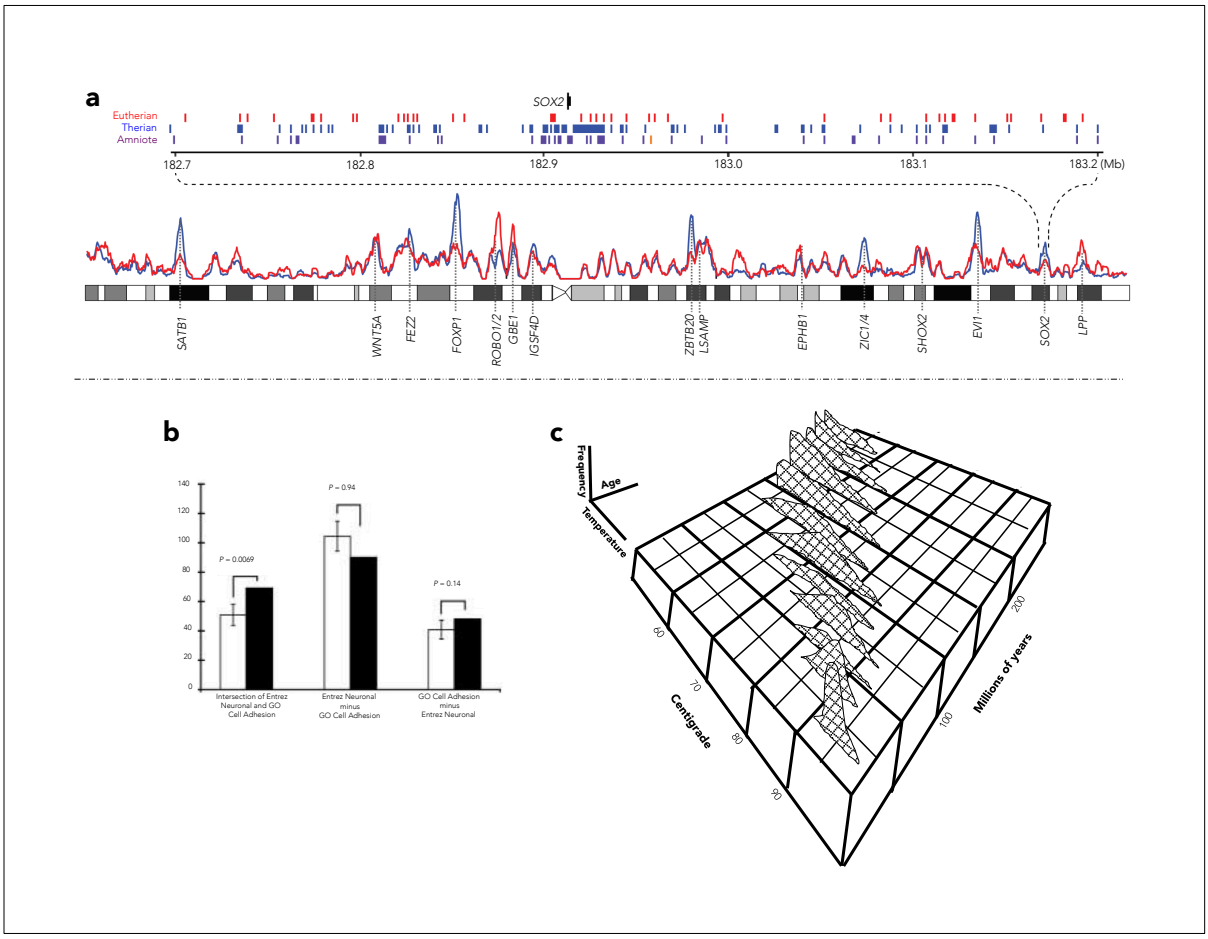


FIGURE 8. Assimilation of TEs into eukaryotic genomes and their lineage-specific residence near developmental genes and genes involved in neuronal cell adhesion.
(**A**) Lineage-specific CNEs near key developmental genes; densities of eutherian CNEs present (blue) or absent (red) in opossum, plotted on a 1-Megabase sliding window across human chromosome 3. Expanded view shows eutherian-specific CNEs (red), eutherian CNEs not overlapping with amniote CNEs (blue) and amniote-specific CNEs (purple) across a 500-kilobase gene desert surrounding the *SOX2* TF gene [*adapted from*: (Mikkelsen et al. 2007)]. (**B**) Evolutionarily-accelerated CNSs disproportionally associated with neuronal cell adhesion genes in humans. CNSs neighboring genes with both Entrez Gene neuronal function and GO cell adhesion function annotations are observed significantly more frequently than expected, while genes only annotated with Entrez Gene neuronal function and no GO cell adhesion function or genes only annotated with GO cell adhesion function and no Entrez Gene neuronal function are not observed more frequently than expected [*adapted from*: (Prabhakar et al. 2006)]. [**C**] The Saltatory Replication model of TE integration into genomes. Each family originates in a sudden event at some time point in the past. Increasing divergence is displayed as the thermal stability of reassociated pairs of strands of DNA formed between members of each family (Centigrade), while time is represented in Millions of years [*adapted from*: (Britten and Kohne 1968)].

While TE sequences are noticeably absent in the regions immediately surrounding the TSS of genes (Polak and Domany 2006), likely due to disruptions in transcription machinery interactions being negatively selected against, a study by Lowe, Bejerano, and Haussler (2007) revealed a noticeable enrichment of TE-derived sequences in gene deserts 0.1-1.0 megabase from the nearest TSS. This study examined CNEs derived from characterized TEs in the human genome that have been under purifying selection since the last boreoeutherian ancestor – at least 100 MYA, or before the human-dog split – representing deep conservation in the Eutherian lineage. A preference for TE-derived sequences residing closest to TF genes or genes involved in development was discovered, particularly those involved in cell adhesion and neural development (Lowe et al. 2007). These results have been echoed in several other studies (van de Lagemaat et al. 2003; Sandelin et al. 2004; Lindblad-Toh et al. 2005; Woolfe et al. 2005; Polak and Domany 2006; Thornburg et al. 2006; Mikkelsen et al. 2007), which additionally emphasize roles for such CNEs in regulating genes for axon guidance receptors and morphogens as well (Figure 8A). Indeed, Prabhakar et al. (2006) report that regions of accelerated evolution in CNEs are similarly enriched in basal lamina GO terms associated with neuronal cell adhesion in both human (Figure 8B) and chimpanzee, but that there is very little overlap between specific elements in the two lineages. The authors remark that this situation would likely result in different consequences for brain development and cognitive function (Prabhakar et al. 2006). As experimental studies of CNEs positioned in such a manner have frequently revealed *cis*-regulatory functionality for neighboring developmental genes (Waterston et al. 2002; Nobrega et al. 2003; de la Calle-Mustienes et al. 2005; Woolfe et al. 2005; Bailey et al. 2006; Pennacchio et al. 2006), it appears that TE-driven evolutionary innovation in these regions would provide the foundation for the observed morphological and neurological diversity within the vertebrate lineage, although it should be noted that enrichment for TEs near neuronal genes may be due to their generally larger gene and intron size. Curiously, correlations have been noted between mouse B1 and human Alu densities in corresponding upstream regions of orthologous genes (Polak and Domany 2006), suggesting conserved functions in the more recently diverged primate and rodent

lineages. Retrotransposition rates are significantly different between rodents and primates, accounting for 10% of the spontaneous mutation rate in mouse and only 0.2% in humans. However, LINE and SINE distributions are much more strongly correlated between orthologous mouse and human loci than their respective base composition (Deininger et al. 2003), suggesting that some type of synonymous pressures may be at play in both lineages (Deininger et al. 2003; Silva et al. 2003; Polak and Domany 2006).

When a novel TE enters a genome, it is free to transpose in an unconstrained manner, as it is effectively unrecognizable to the small RNA-based surveillance machinery of the cell. Newly experienced elements do not display much preference in where they transpose to, as observed with engineered versions of Tc1/Mariner introduced into frogs and fish or human LINE elements introduced into mouse (Ivics et al. 1997; Miskey et al. 2003; An et al. 2006). It appears that these elements actually initiate their own silencing due to the haphazard quality of transposition, since elements both contain their own repetitive sequences and insert in various orientations in the genome without regard to nearby transcription signals, genic context, or antecedent TE presence. Regulation of these elements therefore capitalizes on their tendency to create defunct copies upon arrival (Jensen et al. 1999; Slotkin et al. 2005). Fragmented relics of more ancient TEs accumulate in regions of constitutive heterochromatin due to reduced levels of recombination and selective pressure. Such elements are likely to generate cues to silence their whole family in *trans*. This experimentally documented process is amazingly reminiscent of the "saltatory replication" model (Figure 8C) first hypothesized by Britten and Kohne in 1968 for repetitive DNA (Britten and Kohne 1968). In this light, mobile elements are prime candidates for generating interspecies variation. They are, in fact, largely responsible for differences in the non-coding regions of various inbred populations. Intriguingly, evidence for requisite retrotransposition during cellular proliferation and differentiation (Kuo et al. 1998; Mangiacasale et al. 2003; Lu et al. 2014) has been described, perhaps representing a fascinating case of ontogeny recapitulating phylogeny.

III. Functionalization of Transposable Elements to Interact with the Environment:

The work of the past few decades has established that TEs are functionally important for genomic regulation and evolution at the cellular level. However, work from the past few years has elaborated ways in which TEs may be implemented at the level of the organism. As described above, the structure of TEs confers them with uniquely interesting biology, consequently allowing them to solve uniquely interesting biological problems. Their repetitive and mobile nature, in conjunction with the *cis* regulatory elements and enzymes they encode, allows them build complexity into basic genomic architecture and regulation. This complexity-building capacity may have also become advantageous as the multi-cellular organism began to deploy its genetic information to sense, interpret, and respond to the environment. The two organ systems specifically charged with interacting with the external environment in this manner are the nervous system and the adaptive immune system. And indeed, it is now apparent that both systems make intensive use of TEs. In the case of the immune system, the well-characterized V(D)J recombination mechanism represents domesticated Class II transposition (Slotkin and Martienssen 2007) which serves the purpose of providing a near-endless supply of unique antibodies with which to respond to foreign antigen invasion.

In the mammalian nervous system it is now known that the L1 RTE becomes active and mobilizes during specific phases of neurodevelopment and neuroproliferation (Muotri et al. 2005; Coufal et al. 2009; Muotri et al. 2010), resulting in well-documented somatic mosaicism of both neurons and, to a lesser extent, glia in the adult brain (Coufal et al. 2009; Baillie et al. 2011; Evrony et al. 2012; Perrat et al. 2013; Evrony et al. 2015; Upton et al. 2015). While the precise function of this regulated developmental process for the brain has yet to be revealed, the evolutionary conservation of this process suggests functional benefit.

It is not by chance that the immune and nervous systems display striking similarity, as they are charged with remarkably similar tasks. Both systems are involved in the "*in situ* response to experience" (Edelman 1993; Mattick and Mehler 2008) and as such they both face the cumbersome challenges of data interpretation, storage and retrieval. Both cell types display features of "memory;" in which the system must "recall" a previously experienced event and launch a physical response of the correct speed, magnitude, and duration. Both of these systems have the ability to process a colossal amount of information – the B- and T-cells of the immune system must respond to an unlimited supply of foreign antigens, while a single neuron may make up to thousands of synaptic connections (Habibi et al. 2009; Kioussis and Pachnis 2009). Indeed, the nervous and immune systems display an impressive number of similarities. From a developmental and morphological perspective, both cell types are capable of tracing precise, targeted trajectories in response to chemoattractive and repulsive cues given off either by cells located along their migratory path or at their final destination. In the case of neurons this type of behavior is usually restricted to development, while immune cells retain this behavior throughout the life of the organism. The two systems are both also capable of recognizing environmental cues and transmitting information to and from spatially distant parts of the body, via active and passive immune cell mobility on the one hand and axons and dendrites on the other (Habibi et al. 2009; Kioussis and Pachnis 2009). Additionally, both the immune and nervous systems use concerted physical contact of surface molecules to affect communication of information from one cell to another. While the term "synapse" was first coined to describe these structures between nerve cells (Shaw and Allen 2001; Dustin and Colman 2002), immunobiologists were likely acutely aware of the similarities when they co-opted the term to describe the information sharing platforms of communicating immune cells. The morphological similarities of immune and nervous cells have been observed for quite some time; when Paul Langerhans originally described antigen-presenting dendritic cells in the epidermis, he misconstrued them for cells of the nervous system (Clatworthy et al. 2008). The two systems share a similar supracellular organization as well, with concentrated nuclei responsible for specific processes

intermixed with a more salt-and-peppered interspersion of cell types. This organization has been hypothesized in both cases to give rise to a non-linear integrative functionality essential for the complex data management responsibilities of the two systems (Kioussis and Pachnis 2009). Finally, prominent roles for programmed cell death have been described in the development of both systems (Blaschke et al. 1998; Gao et al. 1998).

While early descriptive studies noted many of the similarities detailed above, more recent work has uncovered a striking overlap in the molecular signatures of the nervous and immune systems as well, possibly reflecting employment of a common chemical language (Hirayama and Yagi 2006; Clatworthy et al. 2008; Habibi et al. 2009; Kerschensteiner et al. 2009; Kioussis and Pachnis 2009; McAfoose and Baune 2009). Cytokines were originally discovered in the immune system and described to effect development and cellular functionality therein, including cell trafficking responsibilities at various stages of the immune response. These molecules were later shown to be important during neurodevelopment, most notably signaling the transition from neurogenesis to gliogenesis (Miller and Gauthier 2007), and are also known to perform prominent roles in such complex cognitive processes as synaptic plasticity and neuromodulation (McAfoose and Baune 2009). It is interesting to note that many of the principal cell-surface receptors of the brain, which are critical for neuronal interactions that dictate migration, survival, axon guidance and synaptic targeting, are members of the immunoglobulin superfamily (Maness and Schachner 2007; Mattick and Mehler 2008; Habibi et al. 2009; Kioussis and Pachnis 2009), leading to hypotheses regarding application of similar somatic diversification strategies in the two systems (Hunkapiller et al. 1989; Chun et al. 1991). In more recent years, immune cells have been found to express neurotransmitters and their associated receptors, as well as neuropeptides, neurotrophins, and glycoproteins classically associated with the growth, guidance and synaptogenesis of neuronal axons (Kikutani and Kumanogoh 2003; Levite 2008; Selmeczy et al. 2008; Yu et al. 2008; Camacho-Arroyo et al. 2009). There appears to be a general concordance between molecules governing process formation, cell motility, and antigen uptake during immunological synapse formation and those

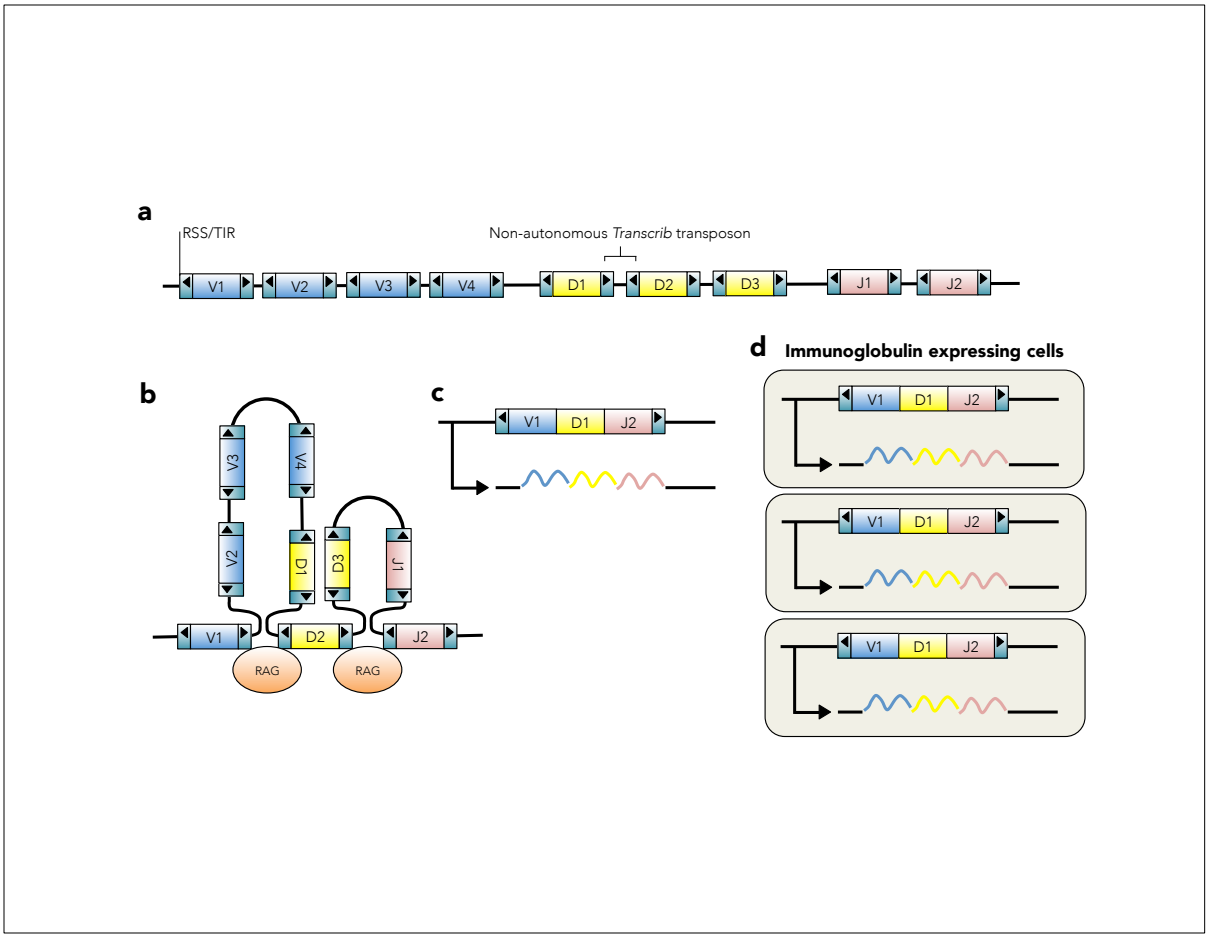


FIGURE 9. The TE origins of V(D)J recombination.
(**A**) Immunoglobulin genes are composed of arrays of variable (V), diversity (D), and joining (J) regions and each segment is flanked with by an RSS that is analogous to the TIR of a DNA transposon. (**B**) Rag1 and Rag2, autonomous immobilized transposases from the Transcrib family of Class II TEs, recognize various combinations of RSS sequences in each B or T cell precursor and (**C**) excise the intervening sequencing. This process can be conceptualized as relating to non-autonomous 'cut-and-paste' transposition. (**D**) As the RSS sequences that interact with RAG proteins are different in each cell individual immunoglobulin cells carry different coding potentials, resulting in production of novel antibody proteins [*adapted from*: (Slotkin and Martienssen 2007)].

regulating neuronal growth and development (Clatworthy et al. 2008). As an example, an interesting study by Yu et al. published in 2008 demonstrated that BASP1 and Plexin B2, among other molecules involved in neurite outgrowth and guidance, can act as molecular markers for B-cells that will terminally differentiate in response to exposure to primed T-cells, as opposed to those B-cells that will abort after brief induction in response to T-independent type 2 antigens (Yu et al. 2008). Overall, it appears that more sophisticated and flexible neural storage and retrieval capacities, immunological responses, and physiological responses to the environment have been fueled by contemporaneous expansions of mechanisms involving genome regulation, somatic plasticity, and cell surface receptor families in the vertebrate lineage (Mattick 2010).

Notably, the most characteristic and well-delineated genetic mechanism responsible for the immune system's ability to respond to environmental stimuli is an exapted transposition mechanism (Slotkin and Martienssen 2007). V(D)J recombination allows B- and T-cells to generate a diverse set of antibodies in order to respond to various antigens. Immunoglobulin genes are comprised of arrays of variable (V), diversity (D), and joining (J) regions, with each segment being bordered by an RSS (Sen and Oltz 2006; Slotkin and Martienssen 2007) (Figure 9A). Rag1 and Rag2, the two enzymes that catalyze V(D)J recombination, are effectively immobilized, autonomous transposases derived from the Transcrib family of DNA elements (Kapitonov and Jurka 2005). As the functional and structural features of RSSs are highly reminiscent of the TIRs of DNA TEs, the process of Rag proteins recognizing RSSs (Figure 9B) and excising the intervening sequence (Figure 9C) is, in essence, non-autonomous Class II transposition (Kapitonov and Jurka 2005; Slotkin and Martienssen 2007). The individual RSSs that interact with the Rag proteins are stochastic and distinct in each cell, generating unique rearrangements of V, D, and J segments that are later transcribed and translated into unique antibodies (Figure 9D). This process is both tissue and stage specific, as these loci are only recombined in specific cell types, in response to certain signals at precise stages of cellular development (Sen and Oltz 2006; Slotkin and Martienssen 2007). Non-domesticated transposons are also known to respond to environmental cues. The yeast Ty5 retrotransposon is known to alter its

target site specificity based on the phosphorylation state of its TD in response to nutrient availability (Dai et al. 2007). When nutrients are readily available the TD is unphosphorylated and can interact with the heterochromatin protein Sir4, effectively targeting transposition to gene-poor heterochromatic regions and reducing its mutagenic potential. In the stressed state, however, the TD becomes phosphorylated, disrupting Sir4 interaction and allowing Ty5 to mutagenize gene-rich, euchromatic regions (Dai et al. 2007).

A growing body of evidence suggests that the cells of the brain also implement regulated transposition during their development. Chun et al (1991) discovered that Rag1, but not Rag2, is expressed in the embryonic murine CNS in cell body-dense regions (Chun et al. 1991). Two other enzymes, XRCC4 and DNA ligase IV, are involved in DNA repair following Rag1 and Rag2 activity and are expressed in a variety of cell types. It has been reported that XRCC4 and DNA ligase IV deficiency in mice results in late embryonic lethality with massively defective lymphogenesis and neurogenesis. B- and T-cells of these mice arrest development around the progenitor stage where V(D)J recombination begins, while extensive apoptotic cell death occurs in the developing nervous system precisely temporally trailing a wave of neurogenesis (Gao et al. 1998). This remarkable phenotype hints at a susceptibility period at the point at which neuronal precursors transition to post-mitotic neurons (Gao et al. 1998). Such an effect could conceivably be due to an inability to repair DSBs generated by active transposition events. It has also been noted that newly born neurons in rat embryonic cortex display the highest irradiation sensitivity (Gao et al. 1998), again hinting at a pronounced susceptibility to DSBs at this particular developmental stage in neurogenesis.

While a diversity of human tissues express L1 processed transcripts and several different transformed human cell lines support its retrotransposition, a recent body of work has made use of an engineered GFP reporter derived from the element to demonstrate the prevalence of active retrotransposition during neurodevelopment (Muotri et al. 2005; Garcia-Perez et al. 2007b; Coufal et al. 2009; Kuwabara et al. 2009; Muotri et al. 2009; Muotri et al. 2010) (Figure 10A – 10B.2). The engineered L1 mobilizes in rat hippocampal neuronal stem cell-derived NPGs (Muotri et al. 2005)

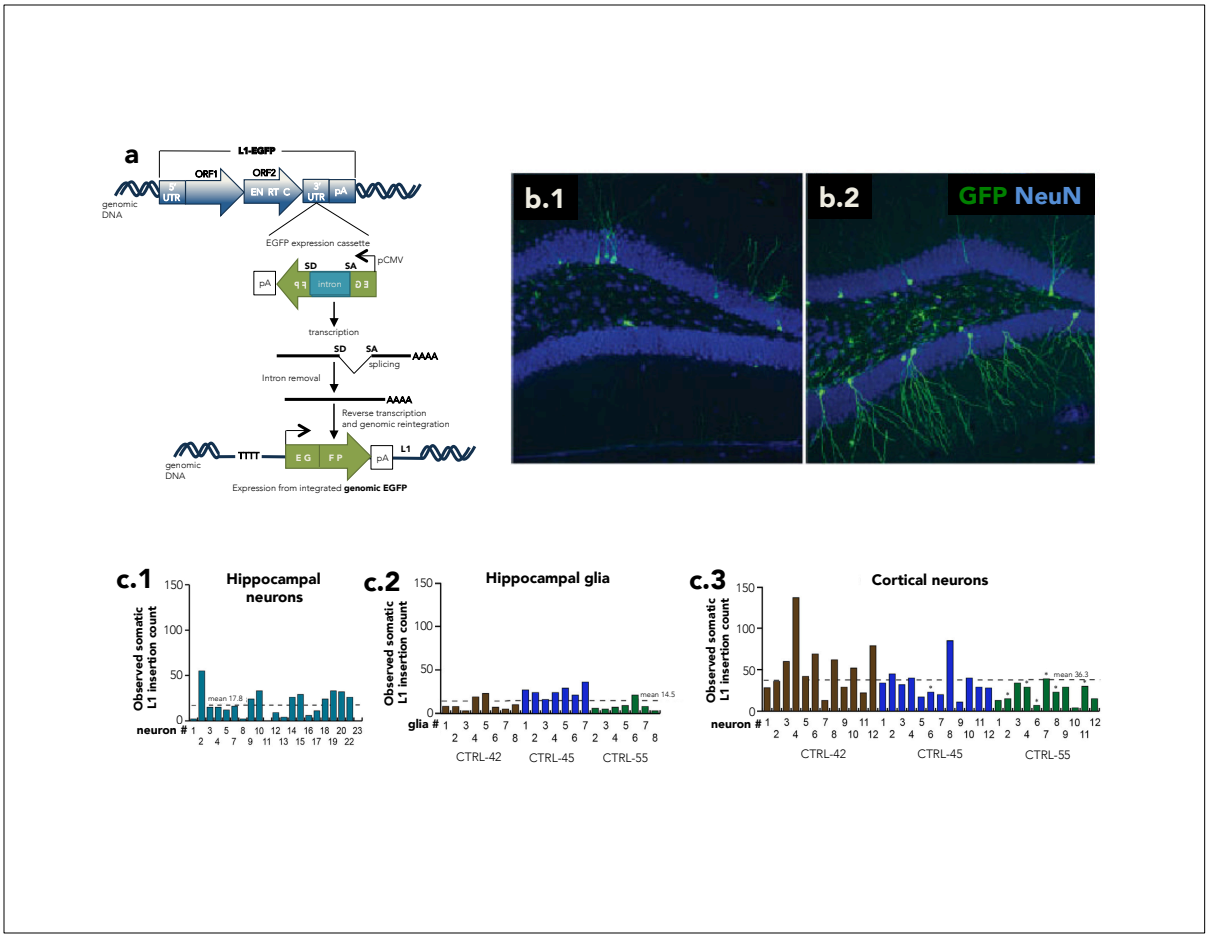


FIGURE 10. Somatic L1 mobilization results in mosaicism in the adult mammalian brain. (**A**) Schematic representation of the L1-EGFP reporter cassette. EGFP has been inserted in the 3' UTR of L1 in a reverse orientation and with an artificial intron embedded in the center of the sequence. Upon transcription the intron is removed, and produces full-length EGFP once the cassette has re-integrated back into the genome [*adapted from*: (Muotri et al. 2009)]. (**B**) Exposure to exercise (wheel running), which has previously been documented to enhance adult neurogenesis in the adult neuroproliferative zone of the dentate gyrus, results in dramatic elevation of L1 mobilization in transgenic mice carrying the L1-EGFP reporter cassette (**B.2**) over sedentary controls (**B.1**) [*adapted from*: (Muotri et al. 2009)]. (**C**) Individual human neurons and glial cells display mosaicism of L1 genomic copy number, with cortical neurons and, to a lesser extent, hippocampal neurons showing an over-all higher genomic copy number than hippocampal glial cells [*adapted from*: (Upton et al. 2015)].

and in transgenic mice, resulting in somatic mosaicism in the adult mouse brain (Kuwabara et al. 2009; Muotri et al. 2009). This process likewise appears to be both tissue and stage specific, as it is mediated by activation of the Wnt signaling pathway, which has various important roles in cell proliferation and migration, cell fate specification, and body axis patterning, and coordinated removal of Sox2, a TF critical for maintaining pleuripotency (Muotri et al. 2005; Kuwabara et al. 2009). L1 activity can induce neural differentiation *in vitro* and has been experimentally validated to affect expression of neuronal genes near *de novo* insertion sites (Muotri et al. 2005). The fact that L1s have been experimentally demonstrated to function as bi-directional promoters suggests that they may indeed act to relay environmental signals to nearby endogenous genic loci (Kuwabara et al. 2009). This strategy is not restricted to embryogenesis, as wheel running in adult transgenic mice, which is known to induce adult neurogenesis, results in increased L1 retrotransposition in the adult neuroproliferative zone of the dentate gyrus (Muotri et al. 2009) (Figure 10B.1 and 10B.2).

NPCs derived from human fetal brain are also capable of supporting retrotransposition and intra-individual brain regions display genomic copy number variations of L1 inserts, with the highest levels observed in the dentate gyrus, frontal cortex, and spinal cord, Importantly, this effect is not observed in other somatic tissues (Coufal et al. 2009). An elegant and rigorous study by Upton et al. (2015) combined single-cell sequencing and modified RC-seq, a methodology which implements sequence capture to enrich for sequences adjoining L1 TE ends, to determine average L1 somatic transposition rates in different cell types of the adult human brain. The authors determined that hippocampal neurons experience 13.7 somatic transposition events on average, while cortical neurons average 16.3 and glia experience 6.5 novel insertions on average per each individual cell (Upton et al. 2015) (Figure 10C.1 – 10C.3). Both Baille et al. (2011) and Upton et al. (2015) have reported that the somatic transposition events from active TE families in humans - L1, Alu, and SVA - have a tendency to insert near protein coding genes that are active and differentially expressed in the brain and which are important for normal brain function (Baillie et al. 2011; Upton et al. 2015). While the open chromatin context of these genes may be responsible for the observed effect, it also enhances the

likelihood for functional relevance of these somatic TE insertion events and may allow mobile elements to preferentially target and modify brain-specific genes. It is possible that somatic TE mobilization during neurodevelopment could result in inter-individual variation, such as is observed in behavioral trait variation in isogenic mouse strains or even between monozygotic human twins (Singer et al. 2010). Such findings demonstrate that it is beyond feasible that the nervous system, just like the immune system, has co-opted TEs to perform critical, complex roles with regard to deploying genetic material to interact with the environment.

IV. Do Transposable Elements Play a Role in Neurodegenerative Disease?

The interplay between TE activity and host response has provided the substrate for massive innovation in the evolution of the linear chromosome, dosage compensation of sex chromosomes, and regulation and employment of genomic information. Domesticated transposition endows the immune system with flexibility in responding to unpredictable invasion of foreign pathogens, and the phenomenon of L1 mobilization and somatic mosaicism in the brain similarly appears to be a cell type-specific and regulated process that in all likelihood endows the brain with some type of positive benefit or function. However, an accumulating body of literature suggests that general chromatin architecture begins to break down with age both *in vitro* and in studies of organismal aging, with a concomitant increase in TE expression (Wood et al. 2010; De Cecco et al. 2013b; Wood and Helfand 2013). Indeed, TE activity has been documented to negatively impact lifespan and age-related senescence phenotypes in a wide-ranging collection of aging models (Driver and McKechnie 1992; St Laurent et al. 2010; Maxwell et al. 2011; Wang et al. 2011; De Cecco et al. 2013b), including recent work from the Dubnau lab that has shown that TEs of disparate families become de-suppressed and actively mobile in head tissue of wild type flies during normal organismal aging (Li et al. 2013). Furthermore work from our lab has shown that loss of function of Ago2, which is highly important for suppression of TEs in somatic tissue such as the brain in *Drosophila*, negatively impacts age-

related brain function as measured by a sensitive learning assay, implying that the consequent TE activity may play a role in age-related decline of neuronal function in normal aging (Li et al. 2013).

It is interesting to note that derepression of TEs has been reported in a sundry assemblage of neurodegenerative diseases as well, including ALS, Rett syndrome, prion disease, macular degeneration, and FXTAS (Lathe and Harris 2009; Muotri et al. 2010; Greenwood et al. 2011; Kaneko et al. 2011; Li et al. 2012; Tan et al. 2012; Li et al. 2015). Activation of TEs has in fact been previously demonstrated to be causal for pathology in macular degeneration (Kaneko et al. 2011). As the type of TE activity described here represents age- or disease-related break down of general TE suppression mechanisms, as opposed to the regulated developmental process addressed above, the broad and general activation of many TE families described in these reports fulfills expectations (Lathe and Harris 2009; Li et al. 2012; Tan et al. 2012; Li et al. 2013). Of particular relevance to the current work, a recent study from the Dubnau and Hammell labs undertook to perform a meta-analysis of publically available data sets from previously published papers regarding RNA sequencing studies of post-mortem tissue from patients with FTLD and cortical tissue from rodent models of TDP-43 pathology (Li et al. 2012). It was found that TDP-43 protein promiscuously binds TE-derived RNA transcripts and that TEs of a wide variety of families become de-suppressed in cortical tissue of rodent models of TDP-43 pathology. Moreover, it was demonstrated that in cortical tissue from human patients with FTLD, a disorder that frequently displays TDP-43 pathology in post-mortem tissue, TDP-43 protein specifically looses its interaction with TE-derived RNA transcripts (Li et al. 2012). Such bioinformatic data, in combination with previous observations that impairing TE suppression mechanisms negatively impacts brain function (Li et al. 2013), suggest that TDP-43 may be having some type of role in suppressing TEs in normal brain tissue and that when this function is disrupted in the disease state the resultant TE activity may actually be having a causal effect in driving neurodegenerative phenotypes. It is the aim of my thesis work to mechanistically test this hypothesis using the genetic tools that are available in *Drosophila*.

CHAPTER 2 – Methods:

Fly Stocks.

All transgenic fly stocks used, with the exception of *w*(IR) and *GMR-Gal4*, were backcrossed into our in-house wild type strain, the Canton-S derivative w^{1118} (*isoCJ1*) (Tully et al. 1994), for at least five generations to homogenize genetic background. The GFP, *OK107*-, *ELAV*-, and *Repo*-Gal4 lines (Qin et al. 2012), as well as the hTDP-43 (Miguel et al. 2011), *gypsy*(IR) (Tan et al. 2012), and *Repo*-LexA::GAD (Lai and Lee 2006) lines are as reported previously. The *GMR-Gal4*, $Gal80^{ts}$, *w*(IR), GFP(IR), and tdTomato lines were acquired from the Bloomington Drosophila Stock Center (stock numbers: 43675, 7019, 25785, 9331, and 32221; respectively), and the *loki*(IR) line was acquired from the Vienna Drosophila Resource Center (stock number: v44980) (Dietzl et al. 2007). Flies were cultured on standard fly food at 22.5 °C.

Bleach Treatment of Embryos.

All fly stocks used for lifespan analysis and longitudinal qPCR experiments were double dechorionated by bleach treatment in order to remove exogenous viral infection (Li et al. 2013). Briefly, 4-hour embryos were collected and treated with 100% bleach for 30 min to remove the chorion. Treated embryos were washed and subsequently transferred to a virus-free room equipped with ultraviolet lights to maintain sterility. This was repeated for at least two successive generations and expanded fly stocks were tested via qPCR of whole flies to ensure DCV levels were below a threshold of 32 cycles.

Lifespan.

Male flies were used for all lifespan assays since the majority of glial-expressing hTDP-43 flies that escape their pupal cases are male. Flies were housed 15 to a vial with a total of 75 flies per genotype and flipped every other day. All vials kept on their side in racks for the duration of the experiment.

Locomotion behavioral assays.

Locomotion behavior was assayed using the classic Benzer counter current apparatus as in Benzer, S, 1967 (Benzer 1967), with the following modifications: freshly eclosed flies were transferred into glass bottles with food and a paper substrate and plugged with foam stoppers. Flies were transferred to fresh bottles every 48 hours until they reached the appropriate age for locomotion assays. The Benzer assay was conducted in a horizontal position to with a fluorescent light source to measure phototaxis.

qPCR and TaqMan Probes.

Tissue preparation, cDNA synthesis and qPCR were performed as previously described (Li et al. 2013) using the Applied Biosystems StepOnePlus Real-Time PCR System. All TaqMan Gene Expression Assays were acquired from Applied Biosystems and used the FAM Reporter and MGB Quencher. The inventoried assays used were: *Act5C* (assay ID Dm02361909_s1), *Dcr-2* (assay ID Dm01821537_g1), *Ago2* (assay ID Dm01805433_g1), *TARDBP* (assay ID Hs00606522_m1), *TBPH* (assay ID Dm01820179_g1), and *loki* (assay ID Dm01811114_g1). All custom TaqMan probes were designed following the vendor's custom assay design service manual and have the following assay IDs and probe sequences: *gypsy ORF2* (assay ID AI5106V; probe: 5'–AAGCATTTGTGTTTGATT TC-3'), *gypsy ORF3* (assay ID AID1UHW; probe: 5'-CTCTAGGATAGGCAATTAA-3'), and *DCV* (assay ID AIPAC3F; probe: 5′-TTGTCGACGCAATTCTT-3′).

Whole mount immunohistochemistry and GFP imaging.

Dissection, fixation, immunolabelling, and confocal imaging acquisition were executed as previously described (Chen et al. 2008). The ENV primary antibody was used as described in Li, *et al*. 2013 (Song et al. 1994; Li et al. 2013). For TUNEL staining, the *In Situ* Cell Death Detection Kit, TMR red (Roche, 12156792910) was used. The same dissection, fixation, and penetration and blocking protocol used for antibody staining was followed (Chen et al. 2008), at which point the brains were transferred to the reaction mix from the kit for 2 hours at 4 °C followed by 1 hour at 37 °C. Brains were then washed, mounted, and imaged as previously described (Chen et al. 2008). For TDP-43 immunohistochemistry, the primary full length hTDP-43 antibody (Protein Tech, 10782-2-AP) was used at a 1:100 dilution, and the primary pSer409 phosphorylated hTDP-43 antibody (Sigma Aldrich, SAB4200223) was used at a 1:500 dilution separately in conjunction with a 1:200 dilution of an Alexa Fluor 488-conjugated secondary antibody (Thermo Fisher Scientific, A-11070). Repo co-labeling was performed using a 1:200 dilution of primary antibody (Developmental Studies Hybridoma Bank, 8D12) and a 1:200 dilution of a Cy3-conjugated secondary antibody (Molecular Probes, A10521). DAPI co-staining was performed after a brief wash in 1x PBS immediately subsequent to secondary antibody staining using DAPI Dilactate (Thermo Fisher Scientific, D3571) as per manufacturer specifications. All brains co-stained with DAPI were imaged on a Zeiss LSM 780 confocal microscope using a UV laser and the Zeiss ZEN microscope software package.

GFP Quantification.

The gain on the confocal microscope was set using the positive control (*Repo* > GFP or *OK107* > GFP) and kept consistent across all subsequent brains imaged. The GFP signal of the median 10 optical sections of the appropriate structures (either the full brain for *Repo* or both lobes of the calyx for *OK107*, respectively) was calculated using ImageJ software, as previously described

(McCloy et al. 2014). These ten values were then averaged, and this number used as a representation for each individual brain. 5-10 brains were analyzed per group.

TUNEL-positive Nuclei Detection and Quantification.

For TUNEL staining, the *In Situ* Cell Death Detection Kit, TMR red (Roche, 12156792910) was used. The same dissection, fixation, and penetration and blocking protocol used for antibody staining was followed (Chen et al. 2008), at which point the brains were transferred to the reaction mix from the kit for 2 hours at 4 °C followed by 1 hour at 37 °C. Brains were then washed, mounted, and imaged as previously described (Chen et al. 2008). For imaging, the gain on the confocal microscope was set using the positive control (*Repo* > hTDP-43) and kept consistent across all subsequent brains imaged. A projection image was generated using the middle 50 optical slices from the z-stack image of the whole brain. This projection image was then thresholded using the maximum entropy technique (See: (Sahoo PK 1988)) via the Fiji plug-in for ImageJ software, and the subsequent binary image was subjected to puncta quantification using ImageJ software. Puncta quantification was performed only for puncta greater than 3 pixels to reduce the likelihood of counting background signal. The total number of puncta counted was then used as a representation for the number of TUNEL-positive nuclei for each brain in subsequent statistical analysis. 7-12 brains were analyzed per group.

Drosophila Eye Imaging.

Flies of the appropriate age and genotype were placed at -70 °C for 25 minutes and then kept on ice until immediately prior to imaging. Imaging was performed using a Nikon SMZ1500 stereoscopic microscope, Nikon DS-Vi1 camera and Nikon Digital Sight camera system, and the Nikon NIS-Elements BR3.2 64-bit imaging software package. The experiment was designed such that each group is balanced for the number of mini-white transgenes and heterozygous for genomic *white*$^{+}$.

TEM.

Drosophila heads were removed, the cuticle removed and the brains fixed overnight in 2% paraformaldehyde and 2% glutaraldehyde in 0.1 mol/L PBS. Samples were rinsed in distilled water and post-fixed for one hour in 1% Osmium tetroxide in 1.5% potassium ferrocyanide in distilled water. Next, the samples were dehydrated in a graded series of ethanol and the final 100% ethanol was replaced with a solution of absolute dry acetone (Electron Microscopy Sciences, Hatfield PA). The samples were then infiltrated with agitation for one hour in an equal mixture of acetone and Epon-Araldite resin, followed by infiltration with agitation overnight in 100% resin. Samples were transferred to embedding capsules with the posterior head facing towards the bottom of the capsule and the resin was polymerized overnight in a vented 60 °C oven. Thin sections were made from the mushroom body region and collected on Butvar coated 2mm x 1mm slot grids (EMS) and the sections were counterstained with lead citrate. Thin sections were imaged with a Hitachi H700 TEM and recorded on Kodak 4480 negatives that were scanned with an Epson V750 Pro Scanner at 2400 DPI.

Cloning.

The three hTDP-43 alleles (hTDP-43 WT, hTDP-43 G294A, and hTDP-43 M337V) were amplified from plasmids generated by the Zador laboratory (idp349, idp350, and idp351, respectively) using the following primers: Forward: 5' – CTCGAGATGTCTGAATATATTCGGGTAAACGAAGATGAG AACGA – 3' and Reverse: 5' – TCTAGACTACATTCCCCAGCCAGAAGACTTAGAATCCATG CTTGAGCC – 3'. These PCR products were double digested with XhoI and XbaI and inserted into pJFRC19. Both the orientation of the insert (5' Sequencing Primer: 5' – AGCAACCAAGTAAATCA ACTGC – 3' and 3' Sequencing Primer: 5' – GAAGGAAAGTCCTTGGGGTC – 3') and the presence or absence of the appropriate point mutation (Internal Sequencing Primer: 5' – GTGGAGA GGACTTGATCATTAAAGG – 3') were confirmed by Sanger sequencing. For the myr-eGFP WT, myr-eGFP W58X, and λN-dDD constructs, inserts were synthesized by IDT and double digested with

XhoI and XbaI. λN-hDD was isolated as an EcoRI fragment from the pPicZa backbone provided by Dr. Rosenthal (Montiel-Gonzalez et al. 2013). All four of these constructs were ligated into pUAST attB. The presence and orientation of the inserts were again confirmed by Sanger sequencing using standard CMV Forward and M13 Reverse primers, as was the presence or absence of the point mutation in the myr-eGFP constructs (Internal Sequencing Primer: 5' – ACGTAAACGGCCACAAG TTCA – 3'). For the pU6-eGFP-X construct, sense (5' – GTCGGGGCCAGGGCACGGGCAGCTTG CGGCCCTGAAAAAGGGCCTGGTGCAGATGAAC – 3') and antisense (5' – CTAGGTTCATCT GCACCAGGCCCTTTTTCAGGGCCGCAAGCTGCCCGTG CCCTGGCCC – 3') oligos were synthesized with appropriate overhangs, phosphorylated and annealed, and then ligated into BbsI/XbaI digested pCFD3. Insertion was verified by Sanger sequencing (5' Sequencing Primer: 5' – CTCAGCCAAGAGGCGAAAAG – 3'). Constructs were injected into recipient embryos and transformant lines were isolated by standard procedures (BestGene).

Statistics.

For qPCR data, the p-values of all data sets with only two groups were calculated using an unpaired t-test. Where an effect of age for more than two time points within one genotype was determined, a one-way ANOVA was performed, and where multiple ages and genotypes are represented a two-way ANOVA was performed; the results are reported in the figure legends. All pairwise comparisons for qPCR reported in the figures were calculated using the Bonferroni method for correction for multiple comparisons. For both the locomotion data and the GFP quantification, p-values were reported using the Sheffé method; ANOVA results are reported in the figure legends. Survival analyses for the lifespan curves were performed using the Kaplan-Meier method, and the Log-rank and Gehan-Breslow-Wilcoxon test were used to compare survival curves. All pairwise comparisons for lifespan curves were corrected using the Bonferroni method.

CHAPTER 3 – Activation of an endogenous retrovirus contributes to neurodegeneration in a *Drosophila* TDP-43 model of Amyotrophic Lateral Sclerosis.

Krug, L., Chatterjee, N., Borges-Monroy, R., Hearn, S., Theodorou, D., and Dubnau, J.

Functional abnormality of TDP-43, an aggregation-prone RNA binding protein, is commonly observed in a spectrum of neurodegenerative diseases that spans motor neuron deterioration and progressive paralysis in ALS to dementia and cognitive decline in FTLD (Ling et al. 2013). We have expressed hTDP-43 in *Drosophila* neurons and glia, a technique that recapitulates important aspects of cellular TDP-43 protein pathology in post-mortem patient tissue when implemented in a variety of model systems and induces progressive locomotor impairment and premature death (Gendron and Petrucelli 2011; Haidet-Phillips et al. 2011; Ling et al. 2013; Meyer et al. 2014; Casci and Pandey 2015; Chen et al. 2015). Here we report that expressing hTDP-43 in both neurons and glia impairs siRNA silencing. The particularly aggressive effects we observe with hTDP-43 expression in glia correlate with early and severe loss of control of a specific RTE, the ERV *gypsy*. We deduce that *gypsy* causes degeneration specifically in these flies because we are able to rescue hTDP-43 toxicity by concomitantly blocking expression of this RTE in glia, but not in neurons. Moreover, we demonstrate that blocking expression of *loki*, the *Drosophila* ortholog of the Chk-2 DDR factor, completely abolishes the rampant apoptosis observed in the CNS of flies that express TDP-43 in glia. This result suggests that the majority of cellular toxicity induced by glial hTDP-43 expression is a result of Chk-2 signaling following DNA damage. RTE activity partially contributes to this DNA damage-induced apoptosis as *gypsy* knockdown partially alleviates the rampant apoptosis observed in the CNS of flies expressing hTDP-43 in glia. Finally, we demonstrate that DNA damage-mediated apoptosis is relevant to physiological decline when hTDP-43 is expressed in both neurons and glia, consistent with broader RTE-mediated effects in both cell types. Our findings build upon recent

reports that the human ERV HERV-K is activated in certain subtypes of ALS, and that over-expression of HERV-K Env results in DNA damage, progressive paralysis, and loss of volume in the motor cortex of transgenic mice (Douville et al. 2011; Li et al. 2015). Furthermore, our findings suggest a novel mechanism in which RTE activity drives neurodegeneration in hTDP-43-mediated disease and potentially implicates RTEs in the etiology of other neurodegenerative disorders as well.

ALS and FTLD are two incurable neurodegenerative disorders that exist on a symptomological spectrum and share both genetic underpinnings and pathophysiological hallmarks (Chen-Plotkin et al. 2010). ALS itself is a muscle wasting disease that presents as progressive paralysis due to death of motor neurons. A notable hallmark of ALS symptomology is its focality and spread; paralysis usually starts in a distal portion of the body such as a hand or a foot and spreads ipsi- and contralaterally such that the usual cause of death is respiratory failure (Ravits and La Spada 2009) (Figure 11C). This characteristic progression suggests a central role for non-cell autonomous effects in perpetuating the disorder. The focality-and-spread mechanism is pronounced in ALS as it affects motor neurons, which are found throughout the body and in the motor cortex. However a similar pattern is observed in various other types of neurodegenerative diseases as well, albeit in more restricted areas of the central brain. While 90% of ALS cases and a large swath of FTLD cases are considered to be sporadic in the sense that they are not precipitated by a known genetic cause, functional abnormality of TDP-43 protein is observed in the vast majority of both familial and sporadic ALS cases (~98% of all ALS cases) and in ~40% of FTLD cases (Ling et al. 2013; Saberi et al. 2015). Both the mechanism that initiates the nucleation of TDP-43 protein pathology in otherwise genetically normal individuals and the mechanism by which cell death occurs is not understood (Chen-Plotkin et al. 2010). Additionally, TDP-43 protein pathology has been documented in the secondary pathology of post-mortem brain tissue from patients diagnosed with a wide array of other

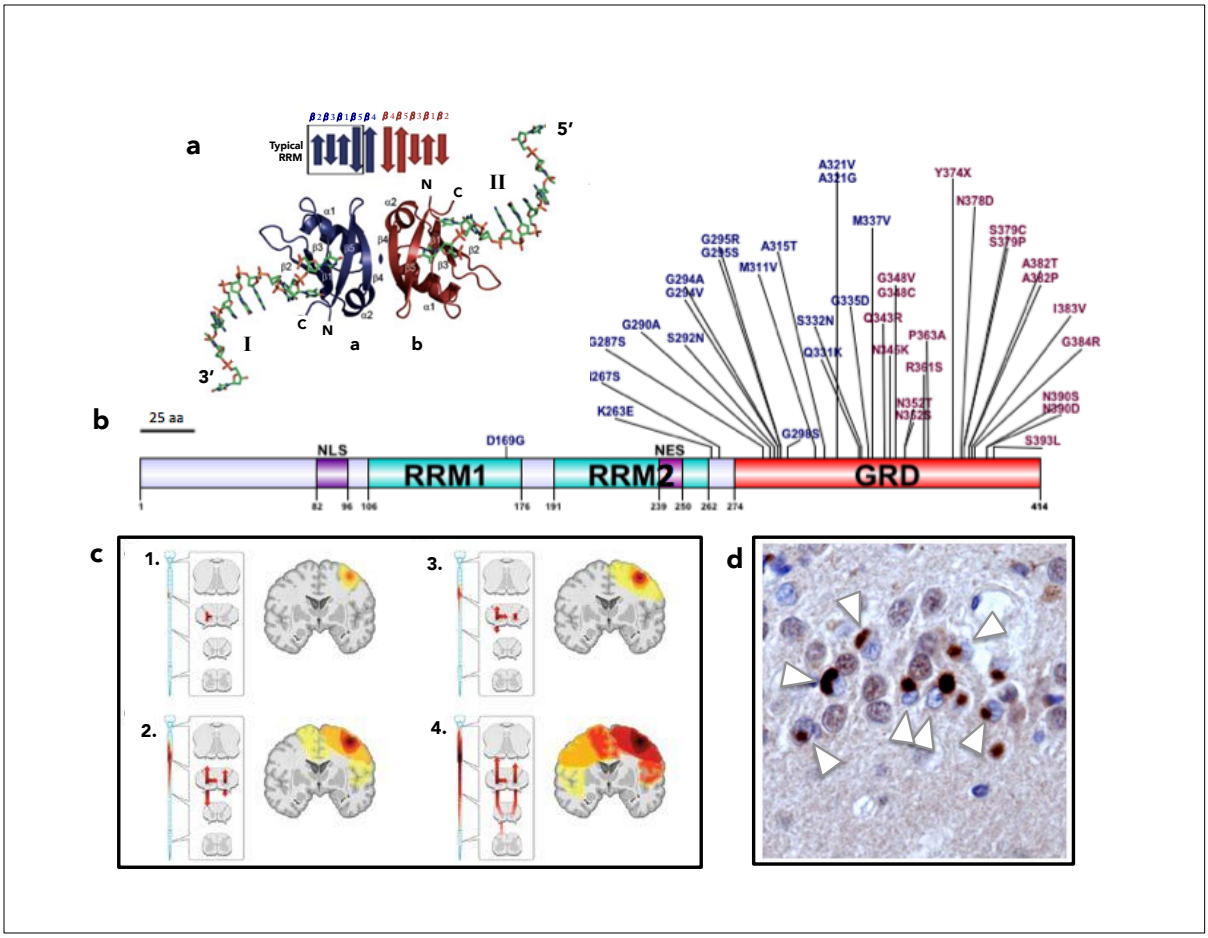


FIGURE 11. The structure of hTDP-43 and its pathology.
(**A**) The crystal structure of hTDP-43 demonstrates that it homodimerizes to bind single stranded RNA and DNA via its N-terminal RNA recognition motifs [*adapted from*: (Kuo et al. 2009)]. (**B**) Disease-causing mutations cluster within the C-terminal glycine rich domain of human TDP-43. (**C**) Focality and spread mechanism of ALS pathological prognosis; warmer colors in the heat map associated with regions of more intense deterioration [*adapted from*: (Ravits and La Spada 2009)]. (**D**) Cellular TDP-43 pathology in post-mortem spinal cord sections from an ALS patient. Arrowheads indicate cells displaying TDP-43 pathology, while other cells present in this section display normal TDP-43 localization. Nuclei are shown in blue; TDP-43 immunoreactivity is shown in brown [*adapted from*: (Neumann 2009)].

neurodegenerative diseases, including hippocampal sclerosis, Alzheimer's disease, corticobasal degeneration, and dementia with Lewy bodies (Amador-Ortiz et al. 2007; Wang et al. 2008).

TDP-43 is a member of the hnRNP family and homodimerizes to bind single stranded RNA and DNA with UG/TG-rich motifs (Kuo et al. 1998) (Figure 11A). This pleiotropic protein was originally identified as a transcriptional repressor that binds to the TAR element of the HIV-1 retrovirus to repress transcription (Ou et al. 1995), and has reported roles in transcriptional regulation, pre-mRNA splicing and mRNA transport, translational regulation, and miRNA biogenesis in the wild type state (Ayala et al. 2005; Kuo et al. 2009; Ling et al. 2013). TDP-43 is capable of shuttling back and forth from the nucleus to the cytoplasm but is predominantly found in the nucleus in healthy cells. In cells that are experiencing TDP-43 protein pathology, the protein accumulates in dense cytoplasmic inclusions that include full-length protein, caspase cleavage products and C-terminal fragments of TDP-43, as well as abnormally phosphorylated and ubiquitinated protein (Arai et al. 2006; Wang et al. 2008; Neumann 2009) (Figure 11D). The vast majority of mutations in the TARDBP gene that encodes TDP-43 and which have been documented to cause human disease cluster within the C-terminal glycine rich domain (Figure 11B). Missense mutations here map to highly conserved residues (Pesiridis et al. 2009; Polymenidou and Cleveland 2011; Ling et al. 2013), suggesting that they are critical for TDP-43 function.

Animal models in which human TDP-43 is transgenically over-expressed reproduce many of the signatures of human disease, including aggregation of TDP-43 protein in cytoplasmic inclusions and downstream neurological effects (Gendron and Petrucelli 2011; Ling et al. 2013; Casci and Pandey 2015). Although such animal models are imperfect representations of what is largely a sporadically occurring disorder, they have enabled the delineation of myriad cellular roles for TDP-43 (Janssens and Van Broeckhoven 2013; Ling et al. 2013). TDP-43 pathology in animal models is now understood to broadly disrupt RNA regulation, resulting in global dysfunction in mRNA stability and splicing, de-repression of cryptic splicing, and biogenesis of some microRNAs

(Chen-Plotkin et al. 2010; Gendron and Petrucelli 2011; Kawahara and Mieda-Sato 2012; Ling et al. 2013; Casci and Pandey 2015; Ling et al. 2015). In principle, any of the cellular impacts of TDP-43 protein pathology could contribute to disease progression either alone or in combination. However, no clear consensus has yet emerged regarding the underlying causes of neurodegeneration.

We advance the novel hypothesis that a morbid loss of control of RTEs contributes to the cumulative degeneration observed with TDP-43 pathology. This RTE hypothesis is founded on a series of observations. First, RTEs are expressed in somatic tissue (Garcia-Perez et al. 2007b; Kazazian 2011) and actively replicate during normal brain development, leading to *de novo* genomic insertions in adult brain tissue (Muotri et al. 2005; Coufal et al. 2009; Muotri et al. 2009; Baillie et al. 2011; Evrony et al. 2012; Perrat et al. 2013; Evrony et al. 2015; Upton et al. 2015). Second, deterioration of RTE suppression – and resultant RTE activity – has been documented with advancing age in a variety of organisms and tissues (Maxwell et al. 2011; De Cecco et al. 2013a; Savva et al. 2013; Zhu et al. 2014; Patterson et al. 2015), including the brain (Li et al. 2013). Moreover, expression of RTEs has been associated with a suite of neurodegenerative diseases (Lathe and Harris 2009; Muotri et al. 2010; Douville et al. 2011; Greenwood et al. 2011; Kaneko et al. 2011; Li et al. 2012; Tan et al. 2012; Li et al. 2015). Indeed, reverse transcriptase biochemical activity has been shown to be present in both serum and cerebrospinal fluid (CSF) of HIV-negative ALS patients (Steele et al. 2005; MacGowan et al. 2007; McCormick et al. 2008; Alfahad and Nath 2013), and a specific RTE, the human ERV HERV-K, is both expressed in post-mortem cortical tissue of ALS patients (Douville et al. 2011; Alfahad and Nath 2013; Li et al. 2015) and can cause motor neuron toxicity when its Envelope (ENV) protein is expressed in transgenic mice (Li et al. 2015). Finally, we have previously demonstrated via meta-analysis of RNA Immunoprecipitation (RIP) and Crosslinked RIP (CLIP) sequencing data that TDP-43 protein binds promiscuously to RTE–derived RNA transcripts in rodent and human brain tissue, and that this binding is selectively lost in cortical tissue of FTLD patients (Li et al. 2012). RTEs inherently act as genome destabilizers; their very nature is to replicate themselves, inducing DNA double strand breaks and inserting themselves into new genomic

locations in the process. Transposon-derived sequence constitutes ~40% of the human genome, a quantity which encompasses a surprisingly large number of functional RTE copies. Even the high levels of RTE RNA transcripts that accumulate when suppression is lost can be potentially cytotoxic, as observed with macular degeneration (Kaneko et al. 2011). The notion that TDP-43:RTE transcript interactions may be lost in the disease state (Li et al. 2012) is therefore particularly attractive when considered in the context of RTEs' notorious destructive capacity, which has been extensively documented in many other biological contexts (O'Donnell and Boeke 2007; Malone and Hannon 2009; Crichton et al. 2014).

RESULTS

gypsy ERV is induced in response to glial expression of hTDP-43

In order to determine whether RTEs mediate TDP-43 pathological toxicity, we implemented an established animal model in which hTDP-43 is transgenically expressed in *Drosophila*. As with other animal models, including mouse, rat, fish, and *C elegans*, such expression reproduces many neuropathological hallmarks of human disease, likely via interference with endogenous protein(s) function (Ash et al. 2010; Kabashi et al. 2010; Gendron and Petrucelli 2011; Vanden Broeck et al. 2014; Casci and Pandey 2015). To test the impact of expressing hTDP-43 on RTE protein or transcript abundance, we hand-selected as a candidate of interest the *Drosophila* RTE, *gypsy*. We chose *gypsy* because we have previously documented this RTE to become aggressively active and mobile in brain tissue of wild type flies with advancing age (Li et al. 2013) and *gypsy* is an ERV with functional similarity to HERV-K, which is expressed in some ALS patients (Douville et al. 2011; Li et al. 2015). In patient tissue, TDP-43 protein pathology is observed in both neurons and glial cells (Chen-Plotkin et al. 2010) and an emerging literature has implicated glial cell toxicity in ALS (Haidet-Phillips et al. 2011; Meyer et al. 2014; Chen et al. 2015). Toxicity of TDP-43 in glia has similarly been documented in animal models, including in *Drosophila*

(Diaper et al. 2013; Estes et al. 2013; Tong et al. 2013; Romano et al. 2015). Disease symptoms nucleate at or after middle age in ALS and FTLD cases, with progressive degenerative effects. We therefore examined the effects of transgenic hTDP-43 expression in the neuronal versus glial compartments of the brain on *gypsy*, as well as the compounding effects of age on pathological prognosis.

We began by performing quantitative RT-PCR (qPCR) for both *ORF2* (*Pol*) and *ORF3* (*ENV*) of *gypsy* on head tissue of flies expressing either pan-neuronal (*ELAV* > hTDP-43) or pan-glial (*Repo* > hTDP-43) hTDP-43 at two relatively young ages (2-4 and 8-10 days post-eclosion). We found an early and dramatic increase in expression of both ORFs (Figures 12A and 12B) specifically in flies expressing hTDP-43 in glia. In contrast, flies expressing neuronal hTDP-43 and genetic controls that do not express hTDP-43 (see also: Li et al., 2013 (Li et al. 2013)) experience a wave of *gypsy* expression at the population level that occurs much later in age (Figure 13A for *ORF3*; similar effects seen for *ORF2,* data not shown). Whole mount immunolabeling of brains using a monoclonal antibody directed against the *gypsy* ENV glycoprotein (Song et al. 1994; Li et al. 2013) likewise shows early (5-8 days post-eclosion) and acute accumulation of strongly immunoreactive puncta particularly in brains of flies expressing glial hTDP-43 (Figure 12C). These intense puncta are observed throughout the superficial regions, which contain the majority of cell somata, as well as in deeper neuropil (Figure 12C and data not shown) and persists into older ages. In contrast, we do not observe neuronal hTDP-43 expression to elevate *gypsy* levels above that seen in wild type flies at any time point with either qPCR or immunolabeling (Figure 12C and 13A). Given that effects of glial hTDP-43 expression on *gypsy* ENV immunoreactivity were so robust in 5-8 day old animals, we examined ENV at earlier time points. We found that in animals expressing hTDP-43 in glia, *gypsy* ENV protein expression appears post-developmentally, with little expression at 0 days (immediately following eclosion) and detectable levels appearing at 3 days post-eclosion. Interestingly, such expression in 3-day old animals appears stochastic in both intensity and spatial location (Figure 13B).

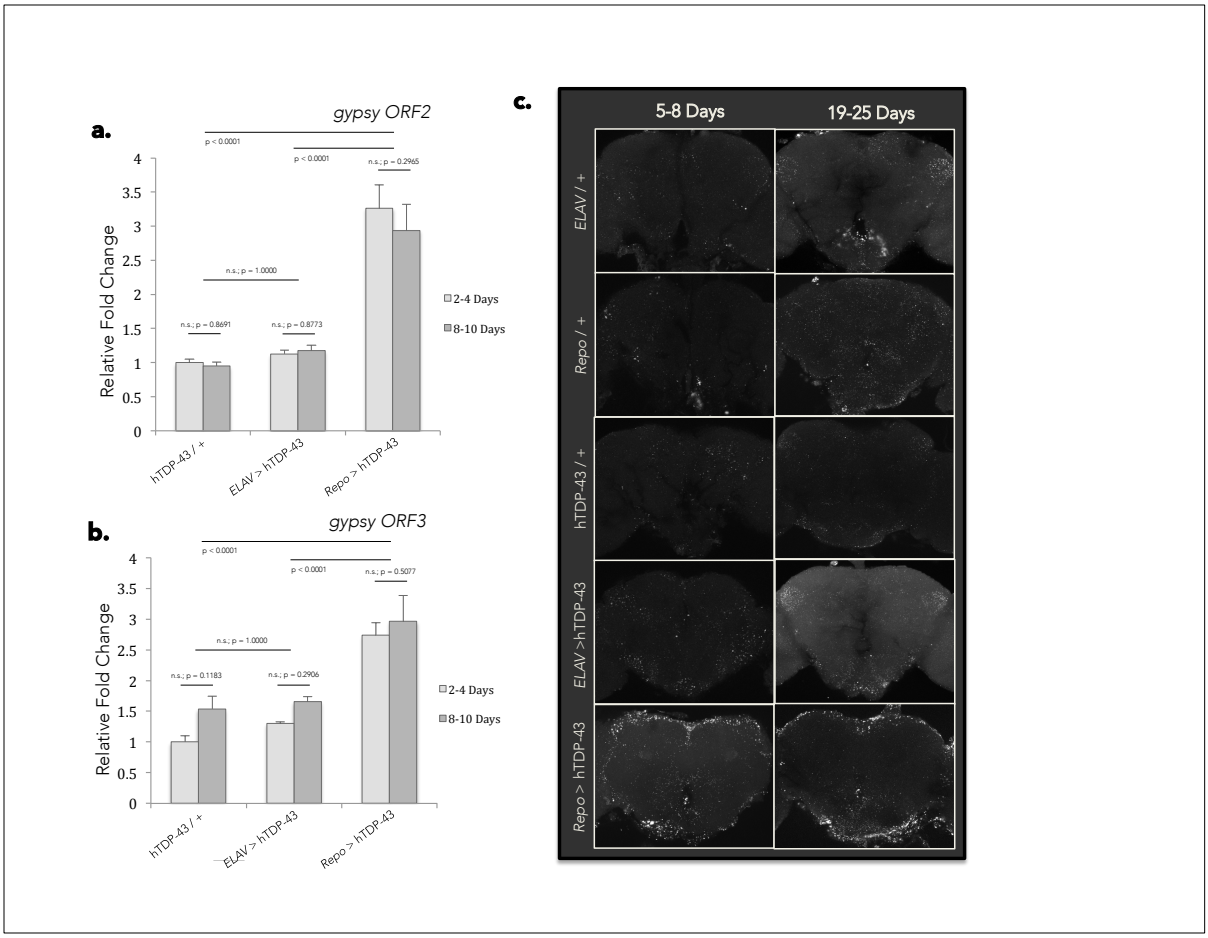
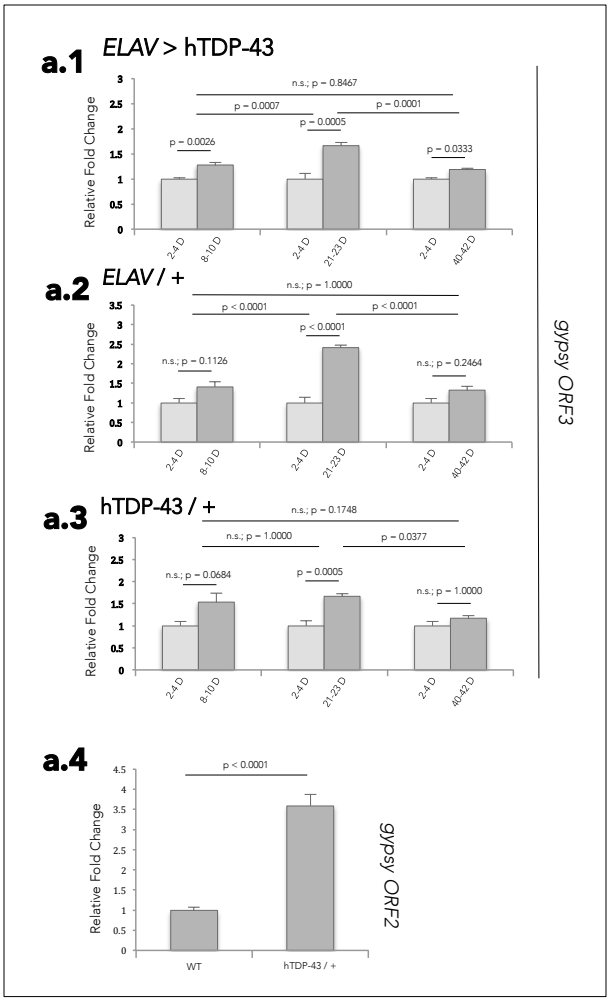
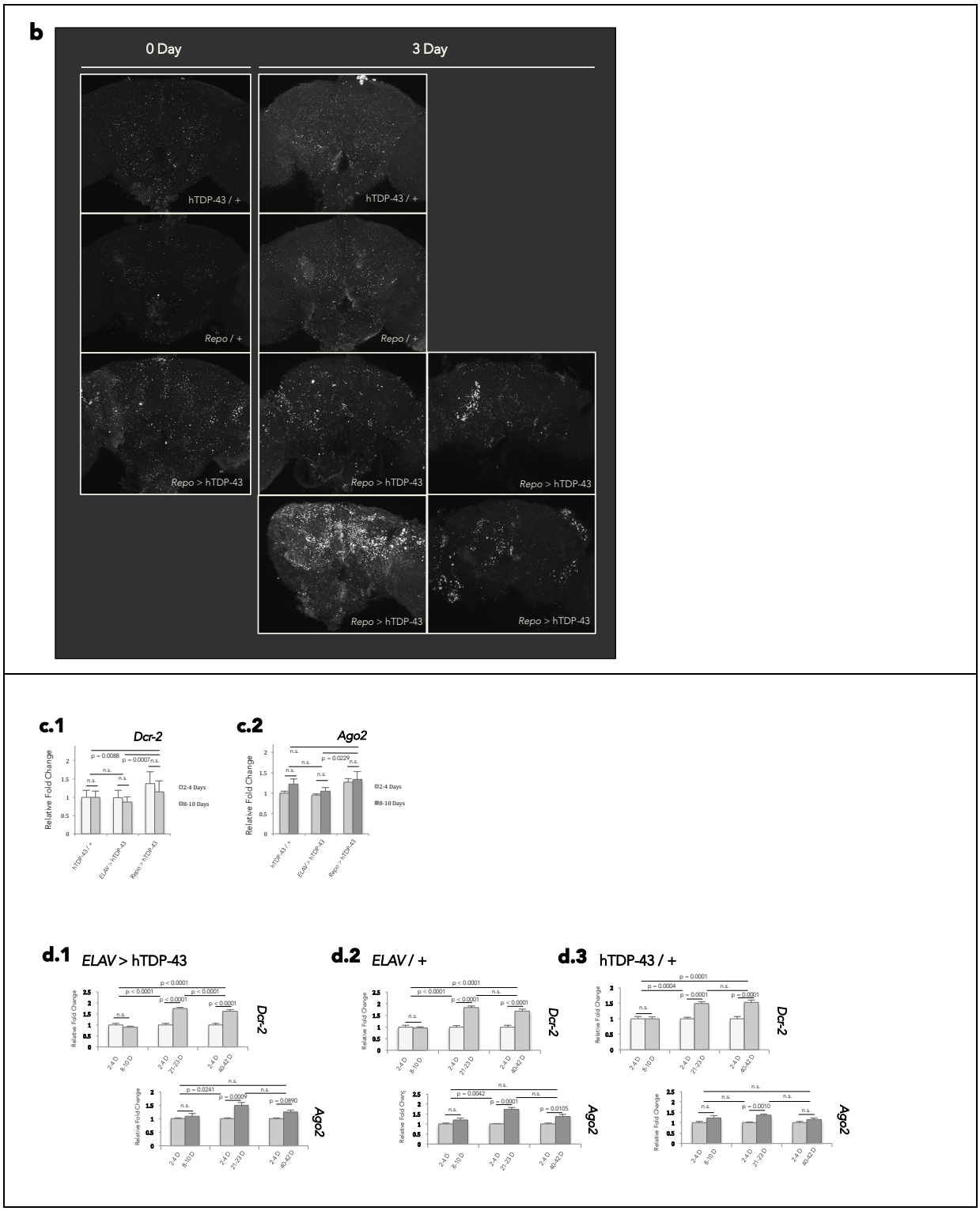


FIGURE 12: Glial hTDP-43 expression results in early and dramatic de-suppression of the *gypsy* ERV.

(**A**) Transcript levels of *gypsy ORF2* (*Pol*) as detected by qPCR in whole head tissue of flies expressing hTDP-43 in neurons (*ELAV* > hTDP-43) versus glia (*Repo* > hTDP-43) at a young (2-4 Day) or aged (8-10 Day) time point. Transcript levels normalized to *Actin* and displayed as fold change relative to flies carrying the hTDP-43 transgene with no Gal4 driver (hTDP-43 / +) at 2-4 Days (means + SEM). A two-way ANOVA reveals a significant effect of genotype ($p < 0.0001$) but no effect of age ($p = 0.5414$). $N = 8$ for all groups. (**B**) An equivalent analysis shows that *gypsy ORF3* (*Env*) likewise displays a significant effect of genotype ($p < 0.0001$) and no effect of age ($p = 0.6530$). $N = 4$ for the 2-4 Day cohort and $N = 5$ for the 8-10 Day cohort. (**C**) Central projections of whole mount brains immunostained with a monoclonal antibody directed against *gypsy* ENV protein reveals dramatic, early accumulation of ENV immunoreactive puncta in brains expressing glial hTDP-43 (5-8 Days) in comparison to both age-matched genetic controls (*ELAV* / + ; *Repo* / + ; hTDP-43 / +) and flies expressing neuronal hTDP-43. This effect persists out to 19-25 Days post-eclosion. *ELAV* / +, 5-8 Day (N= 3), 19-25 Day ($N = 4$); *Repo* / +, 5-8 Day ($N = 3$), 19-25 Day ($N = 3$); hTDP-43 / +, 5-8 Day ($N = 5$), 19-25 Day ($N = 2$); *ELAV* > hTDP-43, 5-8 Day ($N = 2$), 19-25 Day ($N =4$); *Repo* > hTDP-43, 5-8 Day ($N = 7$), 19-25 Day ($N = 8$).


a.1
ELAV > hTDP-43
Relative Fold Change
n.s.; p = 0.8467
p = 0.0007
p = 0.0001
p = 0.0005
p = 0.0026
p = 0.0333
2-4 D
8-10 D
21-23 D
40-42 D
a.2
ELAV / +
n.s.; p = 1.0000
p < 0.0001
p < 0.0001
p < 0.0001
n.s.; p = 0.1126
n.s.; p = 0.2464
a.3
hTDP-43 / +
n.s.; p = 0.1748
n.s.; p = 1.0000
p = 0.0377
n.s.; p = 0.0684
p = 0.0005
n.s.; p = 1.0000
gypsy ORF3
a.4
p < 0.0001
WT
hTDP-43 / +
gypsy ORF2

(*Figure 13*)


b
0 Day
3 Day
hTDP-43 / +
hTDP-43 / +
Repo / +
Repo / +
Repo > hTDP-43
Repo > hTDP-43
Repo > hTDP-43
Repo > hTDP-43
Repo > hTDP-43
c.1
Dcr-2
Relative Fold Change
2-4 Days
8-10 Days
hTDP-43 / +
ELAV > hTDP-43
Repo > hTDP-43
c.2
Ago2
Relative Fold Change
2-4 Days
8-10 Days
hTDP-43 / +
ELAV > hTDP-43
Repo > hTDP-43
d.1 ELAV > hTDP-43
Dcr-2
Ago2
d.2 ELAV / +
Dcr-2
Ago2
d.3 hTDP-43 / +
Dcr-2
Ago2

(*Figure 13*)

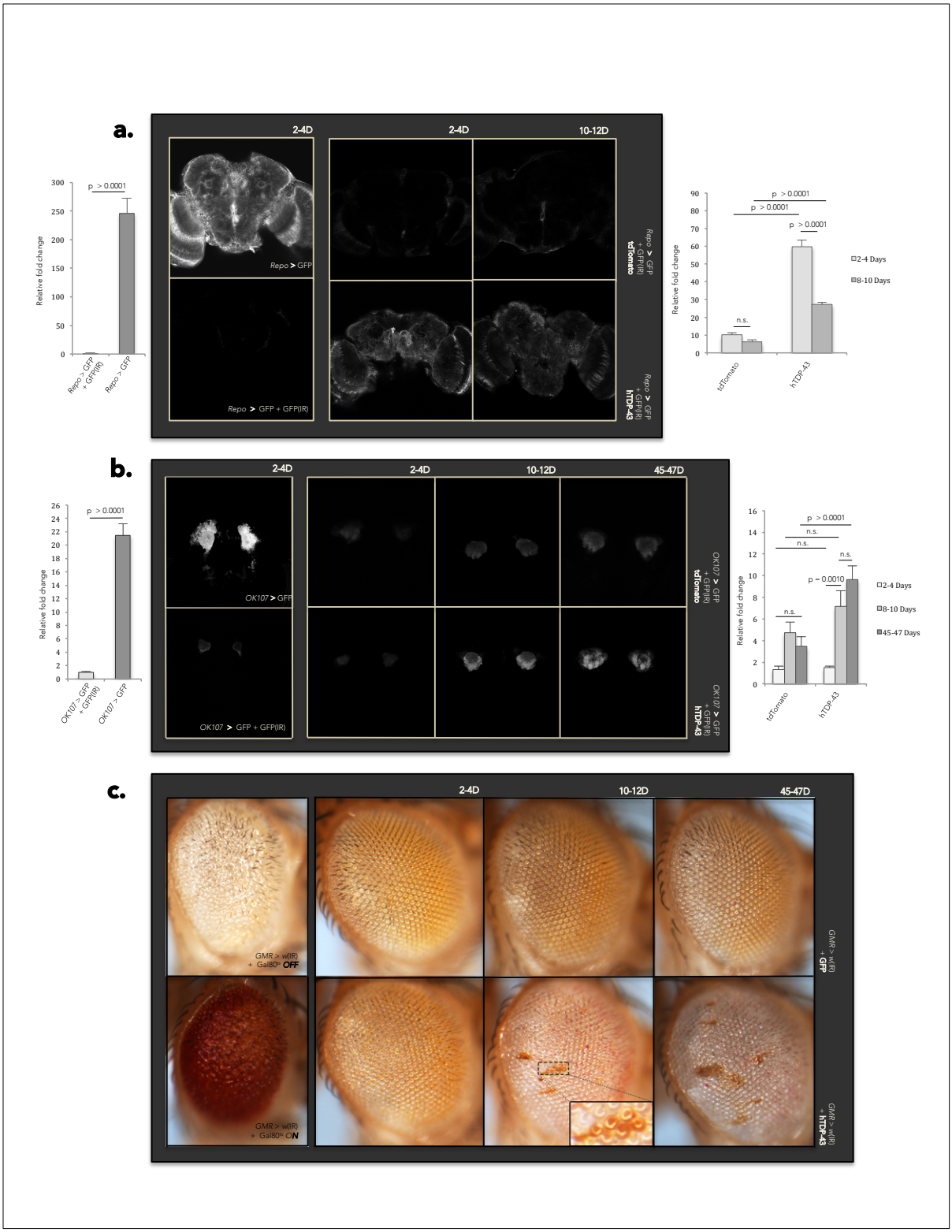
FIGURE 13: *gypsy* expression turns on stochastically in young brains and reaches peak expression in the population at mid-adulthood. Loss of suppression of *gypsy* cannot be explained by hTDP-43- or age-dependent effects on siRNA effector molecules.
(**A**) Transcript levels of *gypsy ORF3* (*Env*) as detected by qPCR on whole head tissue of flies expressing (**13.1**) neuronal hTDP-43 (*ELAV* > hTDP-43), or genetic controls: (**13.2**) *ELAV* / + and (**13.3**) hTDP-43 / +. *gypsy ORF3* transcript levels display an increase by 21-23 days post-eclosion that drops back down by 40-42 days post-eclosion, regardless of genotype. In all cases transcript levels have been normalized to *Actin*, and the aged cohort (8-10 days; 21-23 days; 40-42 days) are represented as a fold change over an appropriate young (2-4 day) cohort that has been processed in parallel. Unpaired t-tests have been used to calculate p-values for each aged cohort with its matched young cohort, while p-values comparing aged cohorts within genotypes have been calculated using the Bonferroni method for multiple comparisons. For all three genotypes a one-way ANOVA shows a significant effect of age on *gypsy ORF3* transcript levels between the aged cohorts (*ELAV* > hTDP-43, $p < 0.0001$; *ELAV* / +, $p < 0.0001$; hTDP-43 / +, $p = 0.0346$). $N = 5$ for all groups. (**13.4**) qPCR of whole head tissue reveals that the presence of the hTDP-43 transgene alone with no Gal4 driver results in elevation of *gypsy ORF2* transcript levels. $N = 6$ for both groups. (**B**) Projections through whole-mount brains immunolabeled with a *gypsy* ENV monoclonal antibody demonstrate that *gypsy* expression turns on post-developmentally, with very little *gypsy* expression immediately following eclosion (0 Days) in genetic controls (hTDP-43 / + and *Repo* / +) and in flies expressing hTDP-43 in glia (*Repo* > hTDP-43). Expression turns on stochastically at 3 days post-eclosion only in the CNS of flies expressing glial hTDP-43. Replicates of 3-day old *Repo* > hTDP-43 brains illustrate the variability of *gypsy* expression at this early time point. $N = 2$ for all 0 Day groups; hTDP-43 / +, 3 Day, $N = 3$; *Repo* / +, 3 Day, $N = 4$; *Repo* > hTDP-43, 3 Day, $N = 7$. (**C**) qPCR of whole head tissue demonstrates that reduced expression of (**13.1**) *Dcr-2* and (**13.2**) *Ago2* cannot account for loss of suppression of *gypsy* in flies expressing glial hTDP-43 (*Repo* > hTDP-43) at either 2-4 or 8-10 days post-eclosion. Transcript levels normalized to *Actin* and displayed as fold change relative to flies carrying the hTDP-43 transgene with no Gal4 driver (hTDP-43 / +) at 2-4 Days (means + SEM). For *Dcr-2*, a two-way ANOVA reveals an effect of age ($p = 0.0006$) but no effect of genotype ($p = 0.1081$); for *Ago2*, a two-way ANOVA also reveals an effect of age ($p = 0.0258$) but no effect of genotype ($p = 0.1591$). $N = 8$ for all groups. (**D**) qPCR of whole head tissue demonstrates that age-dependent changes in expression of *Dcr-2* (**top**) and *Ago2* (**bottom**) cannot account for age-dependent loss of suppression of *gypsy* in flies expressing (**13.1**) neuronal hTDP-43 (*ELAV* > hTDP-43) or genetic controls: (**13.2**) *ELAV* / + and (**13.3**) hTDP-43 / +. All data analyzed as in **13.1-13.3**; one-way ANOVA shows an effect of age across almost all groups (*ELAV* > hTDP-43, *Dcr-2*, $p < 0.0001$, *Ago2*, $p = 0.0269$; *ELAV* / +, *Dcr-2*, $p < 0.0001$, *Ago2*, $p = 0.0051$; hTDP-43 / +, *Dcr-2*, $p < 0.0001$, *Ago2*, $p = 0.3967$). $N = 8$ for all groups.

siRNA-mediated silencing is disrupted by expression of hTDP-43

These findings indicate that the machinery that normally stifles *gypsy* expression in young, healthy brain tissue is eroded by glial hTDP-43 expression. The major post-transcriptional RTE silencing system available in somatic tissue such as the brain is the small interfering RNA (siRNA) pathway (Lee et al. 2004; Aravin and Hannon 2008; Czech et al. 2008; Zhou et al. 2008; Ghildiyal and Zamore 2009; Saito and Siomi 2010). RTEs have been observed to produce siRNAs in many species (Vagin et al. 2006; Slotkin and Martienssen 2007; Ghildiyal and Zamore 2009), and RTE-siRNA levels have been demonstrated to affect RTE activity (Lippman et al. 2003; Sijen and Plasterk 2003; Yang and Kazazian 2006; Slotkin and Martienssen 2007). Moreover, disruptions in the siRNA pathway result in increased TE transcript levels (Svoboda et al. 2004; Czech et al. 2008; Li et al. 2013) as well as novel insertions in the genome (Li et al. 2013; Xie et al. 2013). Indeed, we have previously shown that disruption of the major siRNA pathway effector *Argonaute 2* (*Ago2*) leads to precocious *gypsy* expression in *Drosophila* head tissue and this is accompanied by rapid age-dependent neurophysiological decline (Li et al. 2013). It is noteworthy that TDP-43 has been reported to co-localize with siRNA pathway components in both cell culture and human tissue (Peters and Meister 2007; Robb and Rana 2007; Pare et al. 2009; Freibaum et al. 2010; Liu-Yesucevitz et al. 2010). And while disruption of TDP-43 has been shown to partially impair biogenesis of a subset of miRNAs in human cell culture (Kawahara and Mieda-Sato 2012), the effects of TDP-43 expression on the siRNA pathway are unknown. We therefore designed a genetic reporter system that would inform us as to whether hTDP-43 expression impairs the efficiency of *Dicer-2* (*DCR-2*)/*Ago2*-mediated siRNA silencing in the nervous system *in vivo*.

Our reporter system relied on three components. We co-expressed a Dcr-2 processed inverted repeat (IR) construct directed against GFP (GFP(IR)) with a GFP transgenic reporter. By selecting an effective GFP(IR), we were able to generate substantial silencing of the GFP reporter (Figures 14A and 14B). To test the effects of hTDP-43 on siRNA mediated silencing, we then co-

FIGURE 14: Glial and neuronal hTDP-43 expression erodes siRNA-mediated silencing. (**A**) Representative central projections show that co-expression of the hTDP-43 transgene, but not an unrelated tdTomato control transgene, interferes with the ability of a Dcr-2 processed inverted repeat (GFP(IR)) to silence a GFP transgenic reporter in glial cells using the *Repo-GAL4* driver. Quantification of GFP signal for each group is shown in the appropriate bar graph; values are represented as relative fold change over *Repo* > GFP + GFP(IR) (mean + SEM). A two-way ANOVA reveals significant effects of both genotype ($p < 0.0001$) and age ($p < 0.0001$), and a significant age x genotype interaction ($p < 0.0001$). $N = 5$ for *Repo* > GFP and *Repo* > GFP + GFP(IR); $N = 10$ for all other groups. (**B**) An equivalent analysis demonstrates that hTDP-43 has a similar effect in the neuronal cells of the *Drosophila* mushroom body using the *OK107-Gal4* driver, but with a later age of onset than hTDP-43 expression in glial cells. Quantification of GFP signal for each group is shown in the appropriate bar graph as in **2A**. A two-way ANOVA reveals significant effects of genotype ($p = 0.0054$) and age ($p < 0.0001$), as well as a significant age x genotype interaction ($p = 0.0021$). $N = 5$ for *OK107* > GFP and *OK107* > GFP + GFP(IR); $N = 10$ for all other groups. **C**) Co-expression of hTDP-43, but not GFP, in the photoreceptor neurons of the fly eye under the *GMR-Gal4* driver interrupts the ability of a Dcr-2 processed IR to silence the endogenous *white*$^{+}$ pigment gene with an age of onset similar to that observed with neuronal expression of hTDP-43 in the CNS under *OK107-Gal4*, resulting in characteristic clusters of red-pigmented ommatidia. $N = 5$ for *GMR* > *w*(IR) + Gal80^{ts} *OFF* and *GMR* > *w*(IR) + Gal80^{ts} *ON*; $N = 20$ for all other groups.


a.
2-4D
2-4D
10-12D
Repo > GFP
Repo > GFP + GFP(IR)
Repo > GFP + GFP(IR) tdTomato
Repo > GFP + GFP(IR) hTDP-43
p > 0.0001
Relative fold change
Repo > GFP + GFP(IR)
Repo > GFP
p > 0.0001
p > 0.0001
p > 0.0001
n.s.
2-4 Days
8-10 Days
tdTomato
hTDP-43
b.
2-4D
2-4D
10-12D
45-47D
OK107 > GFP
OK107 > GFP + GFP(IR)
OK107 > GFP + GFP(IR) tdTomato
OK107 > GFP + GFP(IR) hTDP-43
p > 0.0001
OK107 > GFP + GFP(IR)
OK107 > GFP
p > 0.0001
n.s.
n.s.
p = 0.0010
n.s.
n.s.
2-4 Days
8-10 Days
45-47 Days
tdTomato
hTDP-43
c.
2-4D
10-12D
45-47D
GMR > w(IR) + Gal80ts OFF
GMR > w(IR) + Gal80ts ON
GMR > w(IR) + GFP
GMR > w(IR) + hTDP-43

expressed our third component: either hTDP-43 or an unrelated control transgene (tdTomato). This tripartite system was expressed either in all glial cells using the *Repo-Gal4* driver (Figure 14A) or in neurons using the *OK107-Gal4* driver, which provides high levels of expression in the well-defined and easily imaged population of central nervous system (CNS) neurons that constitute the mushroom body (Figure 14B). Brains of young (2-4 day) and middle aged (10-12 days) flies were imaged using confocal microscopy. In the case of neuronal expression we were able to carry the experiment out to old age (45-47 days), but this was not possible with glial expression of hTDP-43 as it results in dramatic reduction in lifespan (see below). What we observed was conspicuously reminiscent of hTDP-43's impact on *gypsy* expression. Glial expression of hTDP-43 causes an early collapse of siRNA silencing, resulting in easily detectable expression of the GFP reporter. Such expression is significant even in 2-4 day old animals (Figure 14A). GFP reporter expression is also seen at the 10-12 day time-point, although these brains are obviously deteriorated (data not shown), which likely explains why GFP levels appear to drop off somewhat. Neuronal expression of hTDP-43 in the mushroom body has a similar but more progressive effect on siRNA-mediated silencing of our GFP reporter, with a later and more gradual onset (Figure 14B). Indeed, when we perform an analogous experiment using an endogenous reporter of siRNA mediated silencing in a separate structure we observe a similar effect. The *GMR-Gal4* driver, which drives high levels of expression in the fly eye, was used to express an IR construct directed against the endogenous *white*$^{+}$ pigment gene in place of GFP as a reporter (Figure 14C and Figure 15). As with mushroom body neurons in the CNS, expression of hTDP-43 in the eye causes a progressive de-repression of the silenced reporter. It is noteworthy that the erosion of siRNA efficacy observed in the eye manifests as clusters of red-pigmented cells, a phenotype which is evocative of the stochastic clusters of ENV immunoreactivity observed early in response to glial hTDP-43 expression (Figure 14C). In contrast, simply turning on expression of *white*$^{+}$ after development results in a uniform darkening of the eye with age (Figure 15B). Taken together, these findings demonstrate that hTDP-43 expression disrupts siRNA-mediated silencing in several tissue types, resulting in robust de-silencing of reporter expression. In neurons

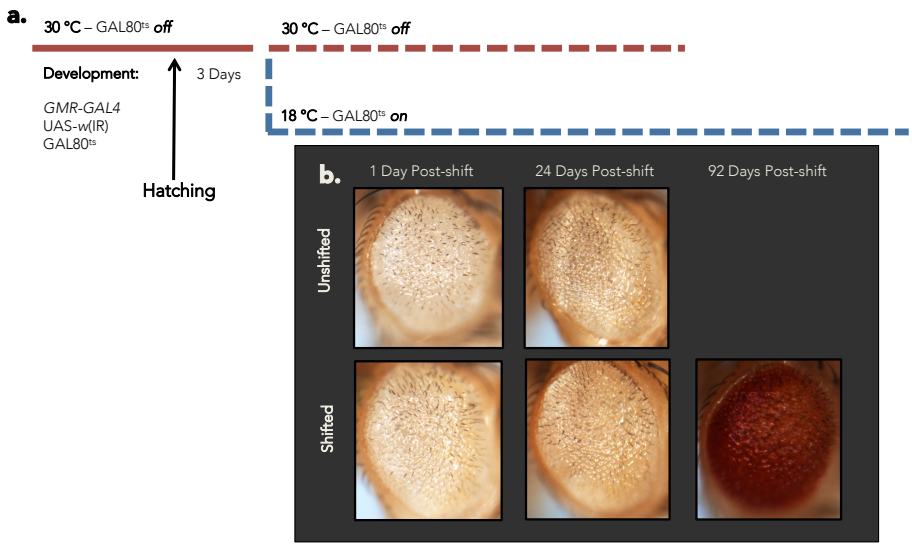


FIGURE 15: Turning on *white*$^{+}$ expression post-developmentally rescues red eye pigmentation in *Drosophila*.
(**A**) Schematic representation of the experimental design. (**B**) Representative images demonstrating that turning off *w*(IR) expression post-developmentally rescues red pigmentation of the Fly eye. $N = 5$ for all groups.

hTDP-43 expression causes age-dependent progressive erosion of siRNA efficacy, while glial expression of hTDP-43 results in more precocious siRNA silencing impairment. In contrast, the *gypsy* ERV is only de-silenced when hTDP-43 is expressed in glia.

Although we have yet to identify which step of the siRNA pathway is disrupted by hTDP-43 expression, it is not simply due to loss of expression of *Dcr-2* or *Ago2*, the two major effectors of siRNA-mediated silencing in *Drosophila* (Lee et al. 2004; Czech et al. 2008; Zhou et al. 2008). qPCR of whole head tissue demonstrated that hTDP-43 expression in both neurons and glial cells does not affect absolute expression levels of *Dcr-2* or *Ago2* at either 2-4 or 8-10 days post-eclosion (Figure 13C). Thus down-regulation of these molecules is not responsible for the de-suppression of *gypsy* we observe with hTDP-43 expression in glia. In fact, in the case of genetic controls and flies expressing hTDP-43 in neurons, *Dcr-2* and *Ago2* levels actually increase with age beginning at 21-23 days post-eclosion and persisting into old age (40-42 days old), suggesting that down-regulation of *Dcr-2* and *Ago2* likewise cannot explain the later elevation of *gypsy* expression observed in these genotypes (Figure 13D).

Age-dependent neurological deterioration with neuronal versus glial hTDP-43 expression

We next examined the relative impact of glial and neuronal hTDP-43 expression on the physiological health of the animal. As previously documented (Diaper et al. 2013; Estes et al. 2013; Romano et al. 2015), we see effects with both neuronal and glial expression. However, we observe differing severity and time courses that mirror the observed effects on siRNA-mediated silencing. Flies expressing hTDP-43 in neurons exhibit significant locomotor impairment at 1-5 days post-eclosion, and flies expressing glial hTDP-43 show more severe locomotor impairment at this same age. This effect is further exacerbated by 5-10 days post-eclosion; at which point the animals expressing hTDP-43 in glia are essentially immobile (Figure 16A). As previously reported

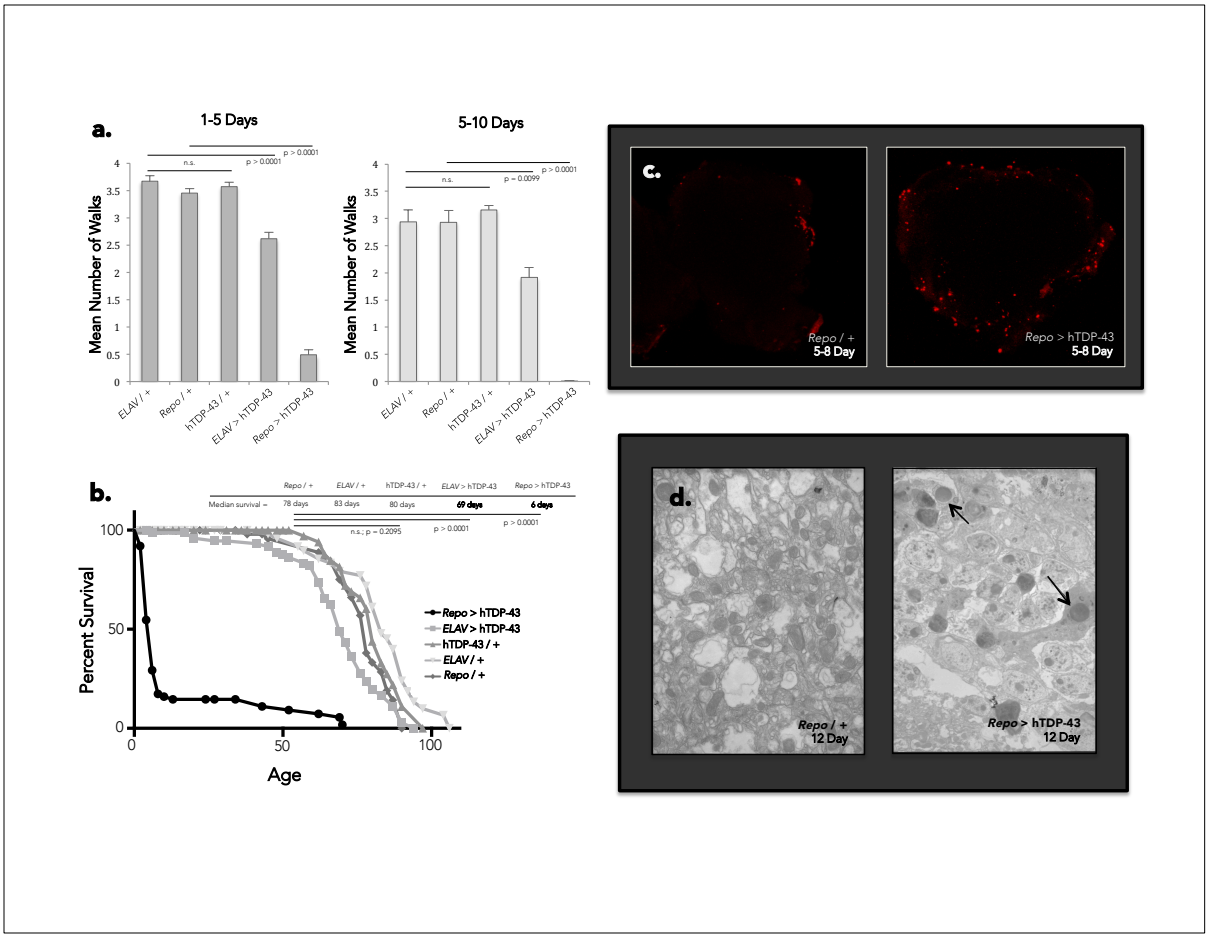
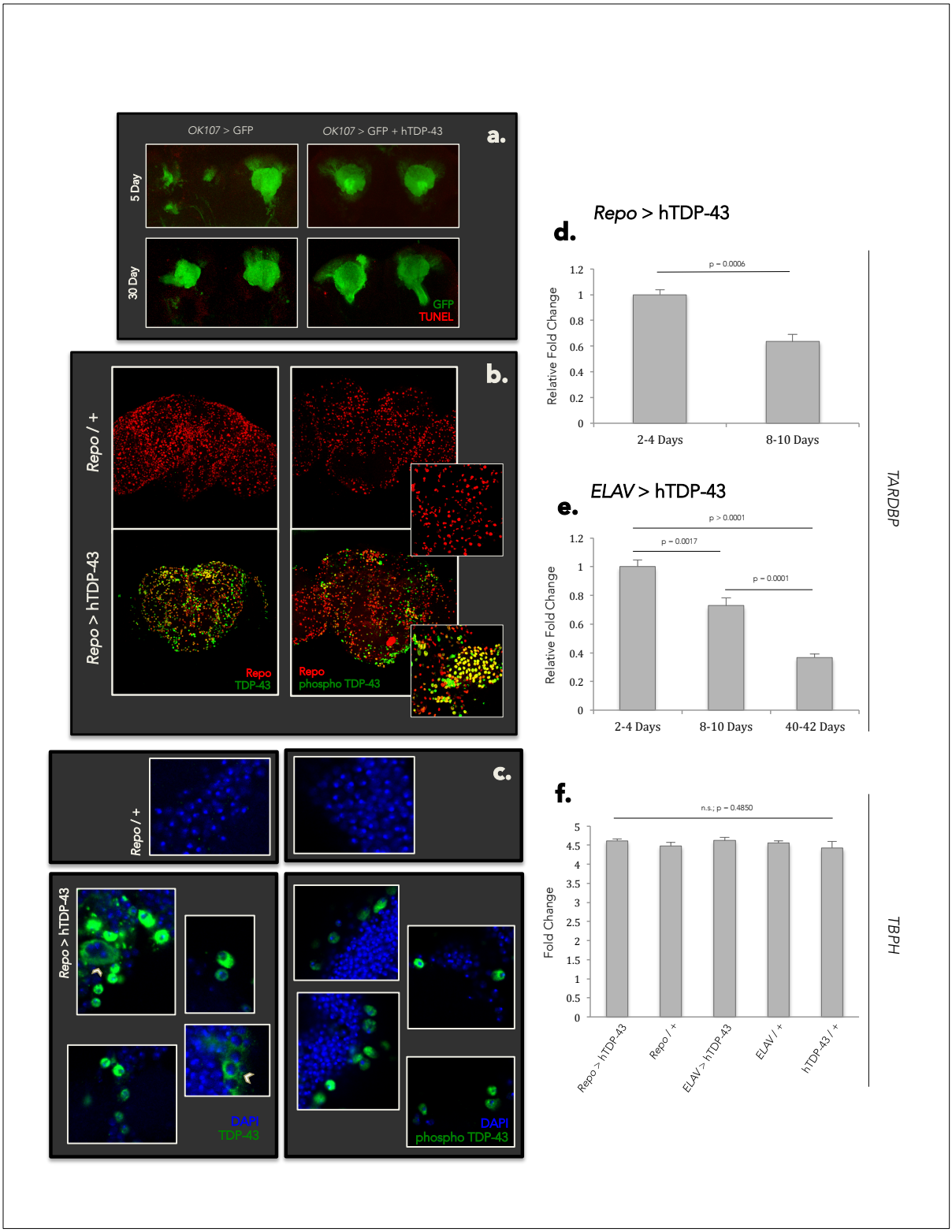


FIGURE 16: Neuronal and glial hTDP-43 expression induces physiological impairment and toxicity with varying severity.
(**A**) Flies expressing glial hTDP-43 display extreme locomotor impairment at 1-5 days post-eclosion in the Benzer fast phototaxis assay, while flies expressing neuronal hTDP-43 demonstrate a slight locomotor deficit in comparison to genetic controls (one-way ANOVA, $p < 0.0001$). This trend continues and is exacerbated by 5-10 days post-eclosion (one-way ANOVA, $p < 0.0001$). Four biological replicates performed for each experiment. (**B**) Lifespan analysis of flies expressing neuronal versus glial hTDP-43 in comparison to genetic controls. (**C**) Central projections of whole-mount brains reveals a stark increase in TUNEL-positive cells in flies expressing glial hTDP-43 in comparison to genetic controls at 5 days post-eclosion. $N = 16$ for *Repo* / + and $N = 18$ for *Repo* > hTDP-43 (**D**) TEM likewise reveals rampant apoptosis in the neuropil of flies expressing glial hTDP-43 at 12 days post-eclosion. Arrowheads indicate pro-apoptotic nuclei, as identified by morphology.

(Hanson et al. 2010; Li et al. 2010; Ritson et al. 2010; Estes et al. 2011; Miguel et al. 2011; Diaper et al. 2013; Estes et al. 2013), flies expressing neuronal hTDP-43 exhibit reduced lifespan in comparison to genetic controls. But flies expressing hTDP-43 in glia display a more severely reduced lifespan with a median survival of only 6 days (Figure 16B). We further observe rampant apoptosis as detected by terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) in the brains of flies expressing hTDP-43 in glia as early as 5 days post-eclosion (Figure 16C). Similarly, we also observe profuse apoptosis in the neuropil of 12 day-old flies expressing glial TDP-43 by transmission electron microscopy (TEM; Figure 16D). In contrast, driving expression of hTDP-43 in mushroom body neurons under *OK107-Gal4* control results in little to no increase in apoptosis (consistent with a previous report (Li et al. 2010)) even when the flies were aged out to 30 days (Figure 17A). The relative expression of hTDP-43 under the two major Gal4 drivers we are using, *Repo-Gal4* and *ELAV-Gal4*, does not differ with age, suggesting that divergent age effects on expression level cannot account for the observed differences in toxicity and impact on physical health (Figure 17D and 17E, respectively). Furthermore, we do not observe any effect of hTDP-43 expression on levels of the endogenous fly ortholog, *TBPH*, regardless of cell type of expression (Figure 17F). Thus, the phenotypes that we observe are not caused by indirect effects on *TBPH* transcript abundance but instead derive from the hTDP-43 transgene itself. As is true in other animal models and in human patients, we cannot readily distinguish whether the effects we observe are due to toxic gain of function, dominant interference with an endogenous protein, or some combination thereof. Importantly, however, we can detect a disease specific phosphorylated isoform of hTDP-43 (Figure 17B) as well as cytoplasmic accumulation and nuclear clearance of the protein (Figure 17C), implying that the human protein is being processed in the CNS of the fly as it is thought to be in the disease state in human tissue.

FIGURE 17: Characterizing UAS-hTDP-43 expression.
(**A**) TUNEL staining reveals very little apoptotic activity when hTDP-43 is expressed in the mushroom body under *OK107-Gal4*, even when the animals are aged to 30 days post-eclosion. Mushroom bodies marked by co-expression of GFP, shown in green; TUNEL staining shown in red. *OK107* > GFP, 5 Day, *N* = 3; *OK107* > GFP + hTDP-43, 5 Day, *N* = 5; *OK107* > GFP, 30 Day, *N* = 5; *OK107* > GFP + hTDP-43, 30 Day, *N* = 4. (**B**) Full length human TDP-43 (green) can be detected by immunolabelling in the brains of flies expressing glial hTDP-43 under the *Repo-Gal4* driver at 21-23 days post-eclosion, and co-localizes with Repo (red) immunoreactivity (**left**). *Repo* / + (*N* = 4); *Repo* > hTDP-43 (*N* = 4). Immunoreactivity for a disease-specific phosphorylated isoform of the protein (pSer409) can also be readily detected and co-localizes with Repo (**right**). *Repo* / + (*N* = 7); *Repo* > hTDP-43 (*N* = 4). A 63x blow-up is shown in the pop-out. (**C**) Both the full-length (**left**) and disease specific (**right**) isoforms of hTDP-43 (green) are mainly observed in the cytoplasm and vacate the nucleus (visualized by DAPI co-staining, shown in blue). Arrowheads indicate the hTDP-43-filled cytoplasm of a cortical glial cell wrapped around several neuronal nuclei in the neuropil of flies expressing glial hTDP-43. For full-length hTDP-43 antibody, *Repo* / + (*N* = 6), *Repo* > hTDP-43 (*N* = 13); for pSer409 phosphorylated hTDP-43 antibody, *Repo* / + (*N* = 4), *Repo* > hTDP-43 (*N* = 9). (**D**) qPCR of whole head tissue demonstrates that transcript levels of *hTDP-43* diminishes under *Repo-Gal4* from 2-4 days to 8-10 days. Transcript levels normalized to *Actin* and displayed as fold change relative to 2-4 day old flies (means + SEM). *N* = 6 for all groups. (**E**) A similar effect of age on hTDP-43 expression is observed in neurons under *ELAV-Gal4*, and continues to drop off by 40-42 days post-eclosion. A one-way ANOVA shows a significant effect of age ($p < 0.0001$). *N* = 6 for all groups. (**F**) qPCR of whole head tissue demonstrates that expression of hTDP-43 does not effect levels of the Fly ortholog, *TBPH*, regardless of cell type of expression. Transcript levels normalized to *Actin*. *N* = 4 for the hTDP-43 / + group, *N* = 5 for all other groups.


a.
OK107 > GFP
OK107 > GFP + hTDP-43
5 Day
30 Day
GFP
TUNEL
b.
Repo / +
Repo > hTDP-43
Repo
TDP-43
Repo
phospho TDP-43
c.
Repo / +
Repo > hTDP-43
DAPI
TDP-43
DAPI
phospho TDP-43
d.
Repo > hTDP-43
p = 0.0006
Relative Fold Change
2-4 Days
8-10 Days
e.
ELAV > hTDP-43
p > 0.0001
p = 0.0017
p = 0.0001
Relative Fold Change
2-4 Days
8-10 Days
40-42 Days
TARDBP
f.
n.s.; p = 0.4850
Fold Change
Repo > hTDP-43
Repo / +
ELAV > hTDP-43
ELAV / +
hTDP-43 / +
TBPH

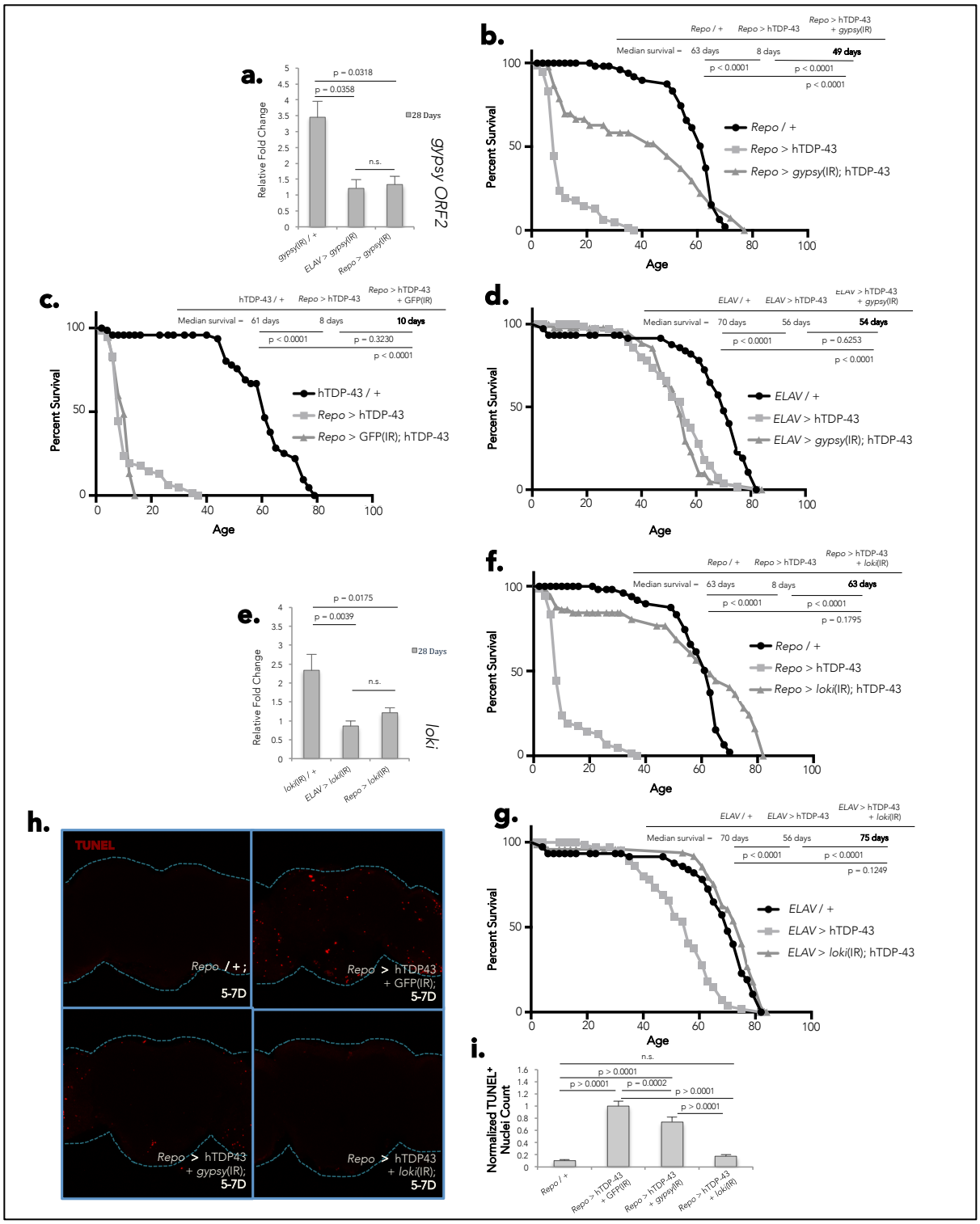
Effects of hTDP-43 on lifespan and apoptosis are mediated by *gypsy* expression and Chk2 activity

We have delineated striking parallels between the age of onset and severity of effects of glial versus neuronal hTDP-43 expression on *gypsy* expression on the one hand and on physiological impairment and apoptosis on the other. These correlative observations, along with the extensively documented toxic effects of loss of control of RTEs (O'Donnell and Boeke 2007; Malone and Hannon 2009; Li et al. 2013; Crichton et al. 2014), suggest that loss of control of *gypsy* might in fact account for the physiological toxicity observed with hTDP-43 expression in glia. To test whether the *Drosophila* ERV *gypsy* causally contributes to the toxic effects of hTDP-43, we used a previously published IR construct (Tan et al. 2012) directed against *gypsy ORF2* (*gypsy*(IR)) that is sufficient to reduce the expression of *gypsy* by approximately 50% in head tissue of 28-day old animals (Figure 18A). We found that co-expression of this *gypsy*(IR) robustly suppresses the extreme lifespan deficit we observe in flies expressing glial hTDP-43 (Figure 18B), an effect which is not observed with a control IR construct (GFP(IR); Figure 18C). Therefore, the expression of *gypsy* caused by hTDP-43 expression in glia is responsible for a significant portion of the toxicity that leads to drastically premature death in these animals. Consistent with our observations that neuronal expression of hTDP-43 does not elevate *gypsy* expression above wild type levels at any given time point over the course of lifespan, co-expression of *gypsy*(IR) likewise does not suppress the lifespan deficit exhibited by animals expressing hTDP-43 in neurons (Figure 18D). As *gypsy*(IR) yields only a partial reduction in *gypsy* expression levels (Figure 18A), we cannot rule out the possibility that *gypsy* also contributes to neuronal toxicity of hTDP-43 below the threshold of the ability of this IR to knock down. However, the glial specificity of *gypsy*(IR) lifespan rescue is consistent with our observation that *gypsy* itself is precociously and aggressively de-silenced strictly in the brains of animals expressing hTDP-43 in glial cells.

FIGURE 18: *gypsy* ERV expression and DNA damage both contribute hTDP-43 mediated toxicity.

(**A**) qPCR on head tissue demonstrates that expressing an IR directed against *gypsy ORF2* (*gypsy*(IR)) in neurons (*ELAV* > *gypsy*(IR)) or glia (*Repo* > *gypsy*(IR)) effectively inhibits age-dependent elevation of *gypsy* transcript levels, and results in an ~2.5-fold reduction at 28 days post-eclosion. *gypsy* transcript levels normalized to *Actin* and displayed as fold change relative to flies carrying the *gypsy*(IR) with no Gal4 driver (*gypsy*(IR) / + ; displayed as means + SEM). A one-way ANOVA shows a significant effect of genotype ($p = 0.0182$). $N = 2$-3 biological replicates generated from heads of 5 mL of flies for each group. (**B**) Lifespan analysis shows that co-expression of *gypsy*(IR) partially rescues the lifespan deficit exhibited by flies expressing glial hTDP-43. (**C**) Co-expression of an unrelated GFP(IR) control does not effect the lifespan of flies expressing glial hTDP-43. (**D**) Co-expression of *gypsy*(IR) has no effect on lifespan in flies expressing neuronal hTDP-43. (**E**) An equivalent analysis as described for **4A** demonstrates that neuronal (*ELAV* > *loki*(IR)) and glial (*Repo* > *loki*(IR)) expression of an IR directed against *loki* (*loki*(IR)) effectively blocks the age-dependent elevation of *loki* transcript levels, resulting in an ~2-fold reduction at 28 days post-eclosion. A one-way ANOVA shows a significant effect of genotype ($p = 0.0039$). $N = 3$-4 biological replicates as in **4A**. (**F**) Lifespan analysis shows that co-expression of *loki*(IR) fully rescues the lifespan deficit exhibited by flies expressing glial hTDP-43. (**G**) Co-expression of *loki*(IR) likewise fully rescues the lifespan deficit exhibited by flies expressing neuronal hTDP-43. (**H**) Central projections of whole-mount TUNEL stained brains reveal a noticeable reduction in the apoptotic activity induced by glial hTDP-43 expression (*Repo* > hTDP-43 + GFP(IR)) when *gypsy* expression is knocked down (*Repo* > hTDP-43 + *gypsy*(IR)), while knocking down *loki* completely alleviates the apoptosis induced by glial hTDP-43 expression (*Repo* > hTDP-43 + *loki*(IR)). (**I**) Quantification of (**H**), normalized to the positive control (*Repo* > hTDP-43 + GFP(IR)). $N = 12$ for *Repo* / +; $N = 9$ for *Repo* > hTDP-43 + GFP(IR); $N = 7$ for *Repo* > hTDP-43 + *gypsy*(IR); and $N = 7$ for *Repo* > hTDP-43 + *loki*(IR).

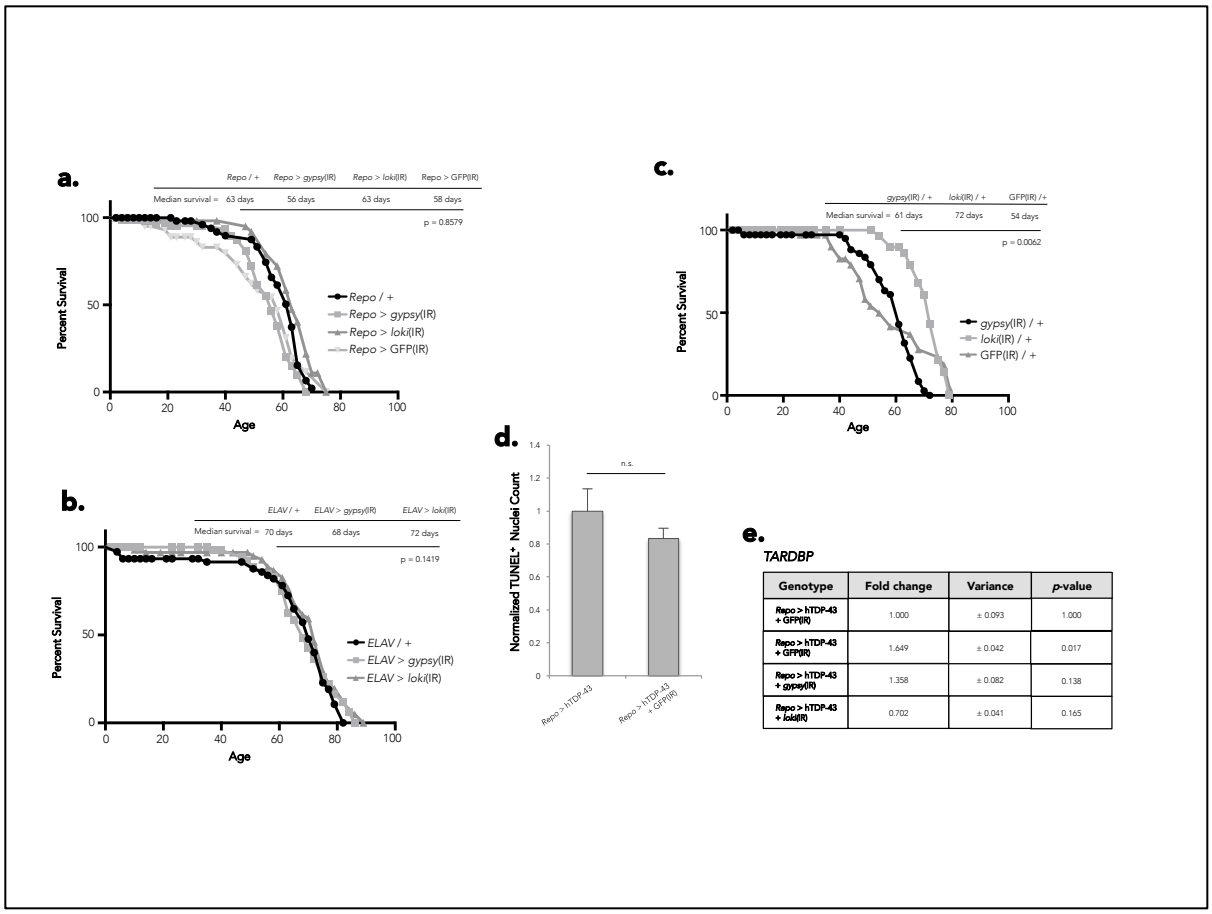
*****All of the lifespans in **Figure 18** and **Figure 19** were performed concurrently in order to ensure comparability across groups. Therefore, appropriate controls are shared across panels.


a.
Relative Fold Change
p = 0.0318
p = 0.0358
n.s.
28 Days
gypsy ORF2
gypsy(IR) / +
ELAV > gypsy(IR)
Repo > gypsy(IR)
b.
Repo / +
Repo > hTDP-43
Repo > hTDP-43 + gypsy(IR)
Median survival = 63 days 8 days 49 days
p < 0.0001
p < 0.0001
p < 0.0001
Percent Survival
Age
Repo / +
Repo > hTDP-43
Repo > gypsy(IR); hTDP-43
c.
hTDP-43 / +
Repo > hTDP-43
Repo > hTDP-43 + GFP(IR)
Median survival = 61 days 8 days 10 days
p < 0.0001
p = 0.3230
p < 0.0001
Percent Survival
Age
hTDP-43 / +
Repo > hTDP-43
Repo > GFP(IR); hTDP-43
d.
ELAV / +
ELAV > hTDP-43
ELAV > hTDP-43 + gypsy(IR)
Median survival = 70 days 56 days 54 days
p < 0.0001
p = 0.6253
p < 0.0001
Percent Survival
Age
ELAV / +
ELAV > hTDP-43
ELAV > gypsy(IR); hTDP-43
e.
Relative Fold Change
p = 0.0175
p = 0.0039
n.s.
28 Days
loki
loki(IR) / +
ELAV > loki(IR)
Repo > loki(IR)
f.
Repo / +
Repo > hTDP-43
Repo > hTDP-43 + loki(IR)
Median survival = 63 days 8 days 63 days
p < 0.0001
p < 0.0001
p = 0.1795
Percent Survival
Age
Repo / +
Repo > hTDP-43
Repo > loki(IR); hTDP-43
g.
ELAV / +
ELAV > hTDP-43
ELAV > hTDP-43 + loki(IR)
Median survival = 70 days 56 days 75 days
p < 0.0001
p < 0.0001
p = 0.1249
Percent Survival
Age
ELAV / +
ELAV > hTDP-43
ELAV > loki(IR); hTDP-43
h.
TUNEL
Repo / + ;
5-7D
Repo > hTDP43 + GFP(IR);
5-7D
Repo > hTDP43 + gypsy(IR);
5-7D
Repo > hTDP43 + loki(IR);
5-7D
i.
Normalized TUNEL+ Nuclei Count
n.s.
p > 0.0001
p > 0.0001
p = 0.0002
p > 0.0001
p > 0.0001
Repo / +
Repo > hTDP-43 + GFP(IR)
Repo > hTDP-43 + gypsy(IR)
Repo > hTDP-43 + loki(IR)

While the effects we observe on *gypsy* expression appear to be specific to hTDP-43 expression in glia, our siRNA reporter assay has revealed that both neuronal and glial expression of hTDP-43 is sufficient to impair Dcr-2/Ago2-mediated silencing. General impairment of siRNA-mediated silencing would be expected to affect RTE expression more broadly, and while the time course of deterioration of siRNA silencing is arguably more gradual in the neuronal subtypes assayed than in glial cells, it is significant and robust (Figures 14B and 14C). This finding is in agreement with our previous observations that TDP-43 protein normally exhibits widespread interactions with RTE transcripts in rodent and human cortical tissue and that these interactions are selectively lost in cortical tissue of FTLD patients (Li et al. 2012), as well as a recent report that knocking out the *C elegans* ortholog of hTDP-43 results in broad accumulation of transposon-derived RNA transcripts (Saldi et al. 2014). It is also compatible with our findings herein that both glial and neuronal hTDP-43 expression in *Drosophila* induces significant negative impacts on physiological health, consistent with previous reports in this model system (Hanson et al. 2010; Li et al. 2010; Ritson et al. 2010; Estes et al. 2011; Miguel et al. 2011; Diaper et al. 2013; Estes et al. 2013). We therefore wondered whether *gypsy*(IR), with its demonstrable glia-specific effects, might neglect to detect the contributions of more widespread RTE activity that is likely to be induced by the general loss of siRNA-mediated silencing exhibited by both neurons and glia expressing hTDP-43. To generate an assay that could inform us of more general RTE activity in response to hTDP-43 expression, we capitalized on the previously documented ability of mutations in the ATR/Chk2 DNA damage response pathway to mask the toxic effects of accumulated RTE-induced DNA damage (Chen et al. 2007; Klattenhoff et al. 2007). When DNA damage accumulates above a threshold at which point repairing the damage is no longer metabolically feasible, cells require ATR/Chk2 signaling to commit to apoptotic cell death. Thus mutations in Chk2 do not prevent accumulation of DNA damage; rather they prevent the signaling required for the cell to recognize that DNA damage has occurred and respond by committing to programmed cell death (Brodsky et al. 2004). We therefore employed an IR construct directed against *loki* (*loki*(IR)), the *Drosophila* ortholog of *chk2*, which is sufficient to

significantly reduce levels of endogenous *loki* mRNA (Figure 18E). Remarkably, co-expression of *loki*(IR) with hTDP-43 is able to fully rescue the lifespan deficit caused by hTDP-43 expression either in glia (Figure 18F) or in neurons (Figure 18G). This effect is not seen with a control IR construct (GFP(IR); Figure 18C). Neither the *gypsy*(IR), GFP(IR), or *loki*(IR) constructs, when expressed individually under *Repo-Gal4* (Figure 19A) or *ELAV-Gal4* (Figure 19B) or present without a Gal4 driver (Figure 19C), has such an effect on lifespan. This finding therefore supports the conclusion that Loki/Chk-2 activity, occurring concomitantly with loss of siRNA silencing, makes a major contribution to the pathological toxicity we observe with both glial and neuronal hTDP-43 expression.

The brains of flies expressing hTDP-43 in glia display rampant apoptosis, seen both with TUNEL staining (Figure 16C) and at the level of TEM (Figure 16D). This is in keeping with the current consensus that cells experiencing TDP-43 pathology in patient tissue die predominantly via apoptosis (Vanden Broeck et al. 2014). It appears that the decision of cells to commit to apoptosis in response to hTDP-43 expression is principally mediated by Loki, as co-expression of the *loki*(IR) which was so effective in suppressing hTDP-43 toxicity in survival analyses also abolishes the dramatic apoptosis observed in the CNS of flies expressing hTDP-43 in glia when these flies are aged to a time point which we have previously documented to display both dramatic *gypsy* expression and apoptosis (5-7 days; Figures 18H and 18I). This effect appears to be specific to *loki*(IR) as co-expression of an unrelated UAS-(IR) construct (*Repo* > hTDP-43 + GFP(IR)) with hTDP-43 in glia does not significantly alter the number of TUNEL positive cells compared to brains of flies expressing hTDP-43 alone under *Repo-Gal4* (Figure 19D). Based on what we know about the cell biological role of Loki/Chk-2, this finding is consistent with the conclusion that the cell death induced by hTDP-43 expression is mediated largely by Loki activity in response to hTDP-43-induced DNA damage. RTE activity does appear to contribute at least in part to the decision of cells to undergo apoptosis in response to hTDP-43 expression in glia, as knocking down *gypsy* partially alleviates the rampant apoptosis observed in the CNS of these animals (Figure 18H and 18I). Importantly, co-



Genotype	Fold change	Variance	*p*-value
Repo > hTDP-43 + GFP(IR)	1.000	± 0.093	1.000
Repo > hTDP-43 + GFP(IR)	1.649	± 0.042	0.017
Repo > hTDP-43 + *gypsy*(IR)	1.358	± 0.082	0.138
Repo > hTDP-43 + *loki*(IR)	0.702	± 0.041	0.165

FIGURE 19: Expression of IR constructs individually does not affect lifespan or hTDP-43 expression.
(**A**) Expression of *gypsy*(IR), *loki*(IR), and GFP(IR) individually in glial cells under the *Repo-Gal4* driver does not significantly alter lifespan. (**B**) Expression of *gypsy*(IR) and *loki*(IR) individually in neurons under the *ELAV-Gal4* driver does not significantly alter lifespan. (**C**) The presence of each of the IR constructs alone without any Gal4 driver only moderately effects lifespan. (**D**) Co-expression of GFP(IR) with hTDP-43 under *Repo*-Gal4 does not significantly alter the number of TUNEL-positive nuclei detected compared to hTDP-43 expression alone under *Repo*-Gal4. $N = 8$ for *Repo* > hTDP-43 and $N = 9$ for *Repo* > hTDP-43 + GFP(IR); data normalized to *Repo* > hTDP-43. (**E**) qPCR for hTDP-43 expression (*TARDBP*) on whole head tissue demonstrates that co-expression of each of the IR constructs with hTDP-43 under *Repo*-Gal4 (*Repo* > hTDP-43 + GFP(IR), *Repo* > hTDP-43 + *gypsy*(IR), and *Repo* > hTDP-43 + *loki*(IR), respectively) does not significantly reduce hTDP-43 expression levels compared to hTDP-43 expression alone under *Repo*-Gal4 (*Repo* > hTDP-43) Fold change is displayed as the mean fold change relative to *Repo* > hTDP-43, while *p*-value represents the *p*-value of a two-tailed Student's t-test in comparison to *Repo* > hTDP-43. $N = 4$ for all groups.

expression of the GFP(IR), *loki*(IR), and *gypsy*(IR) constructs used in both the TUNEL staining and survival analysis with hTDP-43 under *Repo-Gal4* does not significantly reduce the expression of hTDP-43 (*TARDBP*) in whole head tissue, suggesting that reduced expression of hTDP-43 cannot account for the phenotypic rescue induced by *loki*(IR) or *gypsy*(IR) in either of these assays (Figure 19E). This set of observations is in agreement with the well-documented accumulation of DNA double strand breaks induced by unleashing RTEs (Belgnaoui et al. 2006), as well as reports that transgenic expression of the HERV-K ENV protein in mice results in loss of volume in the motor cortex and DNA damage (Li et al. 2015). Our previous findings regarding hTDP-43's broad interactions with RTE-derived RNA transcripts tempts us to postulate that the Chk-2-mediated apoptosis we observe in response to hTDP-43 expression in *Drosophila* glia is due to more general run-away activation of RTEs above and beyond the effects we have documented of hTDP-43 expression on *gypsy*. However, while it is evocative that *gypsy* expression appears to contribute at least in part to the decision of cells to commit apoptosis in response to glial expression of hTDP-43, these results also do not rule out other effects of hTDP-43 expression on DNA damage recognition and repair. These findings lead to a model in which TDP-43 protein pathology in human cells results in a dramatic deterioration of siRNA-mediated silencing accompanied by activation of RTE expression. In this model cells that experience TDP-43 pathology commit to apoptosis via Loki/Chk-2 activity that is the result of accumulation of hTDP-43-induced DNA damage. The DNA damage incurred by hTDP-43 pathology is likely to be caused at least in part by TDP-43's effects on RTE activity.

DISCUSSION

We previously reported bioinformatic evidence for a physical link between TDP-43 protein and RTE RNAs in rodents and in human tissue (Li et al. 2012). Here we provide mechanistic evidence that TDP-43 pathology in flies is associated with a collapse of the siRNA-mediated

silencing system. While we do not know the mechanism by which TDP-43 impacts siRNA silencing, it may involve direct interactions between TDP-43 and the RNAi protein machinery (Kawahara and Mieda-Sato 2012) and our previous findings suggest direct interaction with RTE RNAs (Li et al. 2012). We also demonstrate that the loss of siRNA silencing is accompanied by toxic expression of the *gypsy* ERV. This finding is parsimonious with reports of high levels of reverse transcriptase activity in blood serum and CSF of HIV-negative ALS patients (Steele et al. 2005; MacGowan et al. 2007; McCormick et al. 2008), accumulation of transcripts and protein from HERV-K, a human ERV of the *gypsy* family, in the CNS of ALS patients (Douville et al. ; Li et al.), and severe accumulation of virus-like inclusions detected by electron microscopy in both neurons and glia of the frontal cortex of one ALS patient with extended prolongation of life via artificial lung ventilation (Popova and Sakharova 1982). However, while overexpression of just the ENV protein of HERV-K has been demonstrated to induce progressive motor dysfunction in transgenic mice (Li et al. 2015), this is the first time that the induction of an endogenous RTE has been demonstrated to causally contribute to physiological deterioration in a model of human TDP-43 pathology. Our finding that the siRNA silencing system is compromised by hTDP-43 expression suggests the possibility that other classes of retrotransposons may be similarly activated. In the context of the *Drosophila* model, we provide strong evidence that hTDP-43-induced expression of *gypsy* contributes to DNA damage mediated apoptosis and plays a causal role in the physiological consequences for the animal. Our finding that Chk-2 activity is largely responsible for apoptosis in response to hTDP-43 expression suggests that *gypsy* may be successfully or abortively transposing into genomic DNA, however we are mindful of the fact that increased levels of *gypsy* proteins and RNAs may themselves be cytotoxic, as is observed with *Alu* elements in macular degeneration (Kaneko et al. 2011). These findings lead us to posit the "retrotransposon storm" hypothesis of neurodegeneration. We envision that loss of control of RTE expression and replication leads to a feed-forward mechanism, resulting in massive levels of activity that contribute to toxicity and degeneration in the nervous system. Our findings emphasize the importance of investigating a broader role of RTEs in TDP-43-mediated pathogenesis, and may

indicate a promising common avenue for novel therapeutic targets in both familial and sporadic cases of ALS.

CHAPTER 4 – Tool Development:

Krug, L. and Dubnau, J.

In this section I will briefly describe tools and sample sets I have generated that will enable continued investigation of how TDP-43 pathology affects TEs in the CNS. While performing all the experiments I would like with these tools is beyond the scope of my thesis, they represent a significant portion of my effort during my time in the Dubnau lab and also reflect the ways in which I have been thinking about the phenomenology and the types of follow-up questions I believe will be important to address in the near future. Below I will outline the structure of these tools, the extent to which they have been validated or utilized, and what types of questions they may prove useful in addressing.

I.LexAop-hTDP-43:

Reasoning and Design:

The LexAop system is a secondary yeast operator system that has been modified for use in *Drosophila* (del Valle Rodriguez et al. 2012) (Figure 20A). This system functions very much like the more ubiquitously utilized GAL4-UAS system but will not cross-react with the GAL4-UAS system if employed in the same animal because the two yeast operator proteins are expressed under separate endogenous promoters. Therefore, combined use of the LexAop and GAL4 systems in the same animal allows for cell type-specific expression of different transgenic constructs in two separate cell types (del Valle Rodriguez et al. 2012). Given the well-documented glial toxicity of TDP-43 and other ALS models (Haidet-Phillips et al. 2011; Diaper et al. 2013; Estes et al. 2013; Tong et al. 2013; Meyer et al. 2014; Chen et al. 2015; Romano et al. 2015), as well as the dramatic effects of glial hTDP-43 expression I have presented in Chapter II, the ability to express TDP-43 in glial cells while simultaneously monitoring effects on transgenic reporters in neurons *in vivo* would be highly

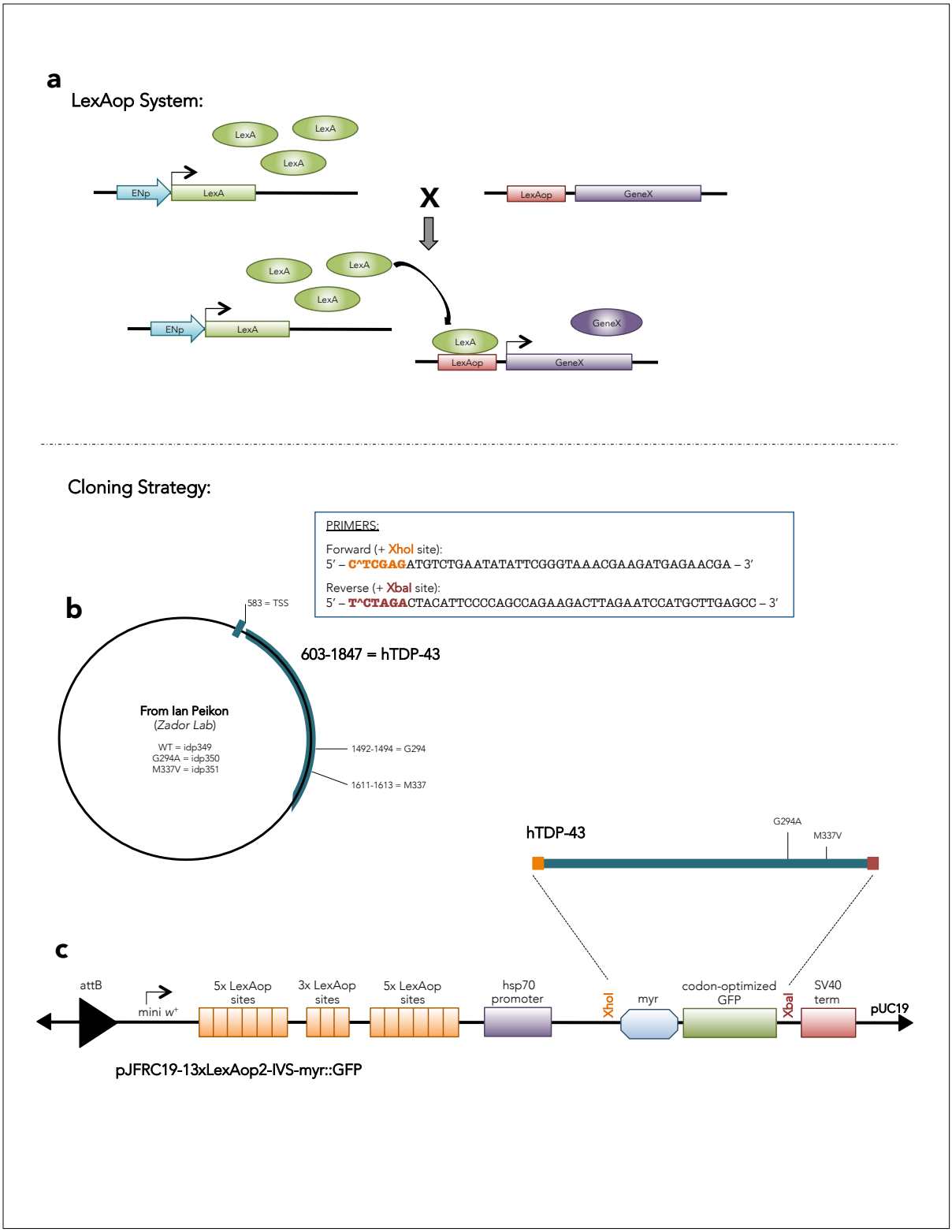


FIGURE 20. Generating LexAop-hTDP-43 transgenic fly lines.
(**A**) Schematic representation of the LexA system in *Drosophila melanogaster*. (**B**) PCR amplification strategy for hTDP-43 WT, hTDP-43 G294A, and hTDP-43 M337V from plasmid constructs. (**C**) Cloning of hTDP-43 inserts into pJFRC19.

advantageous. I have therefore subcloned three alleles of hTDP-43 into a vector that will place them under LexAop control (Figure 20B and 20C). These three alleles include a wild type allele and two C-terminal point mutant alleles that are known to cause human disease and which have been studied in other laboratory contexts (LexAop-hTDP-43 WT, LexAop-hTDP-43 G294A, and LexAop-hTDP-43 M337V). I have generated transgenic fly lines from the LexAop-hTDP-43 WT construct, outcrossed the insertion to our wild type fly line, and validated its expression under *Repo*-LexA by qPCR.

Validation:

Subcloning of all three hTDP-43 alleles was validated by Sanger sequencing, and the LexAop-hTDP-43 WT allele was sent out for injection into *Drosophila* embryos according to standard protocol. Transformants were validated by PCR of genomic DNA. I outcrossed two of these transformant lines to our in-house wild type strain for 5 generations to homogenize genetic background. I then either crossed these two lines to wild type flies or outcrossed flies carrying a *Repo*-LexA construct (Lai and Lee 2006). I performed qPCR on 2-4 day old head tissue of these groups to confirm whether the presence of *Repo*-LexA induced expression of LexAop-hTDP-43 WT in these two lines above background levels. Indeed, expression is elevated ~4-fold over background (Figure 21A and 21B). This is much less than is observed with glial expression of UAS-hTDP-43 under *Repo*-GAL4, which elevates TDP-43 expression ~20-fold over background at the same time point. However, it is on par with the expression levels I observe with UAS-hTDP-43 under *ELAV*-GAL4 in whole head tissue (~3.5-fold elevation over background), and likely reflects higher TDP-43 expression per cell than in the *ELAV*-GAL4 flies as there are orders of magnitude fewer glial cells than neurons in the whole fly brain. This feature may actually prove advantageous for future experiments as the *Repo*-GAL4 phenotype is so severe that it makes collecting tissue extremely cumbersome.

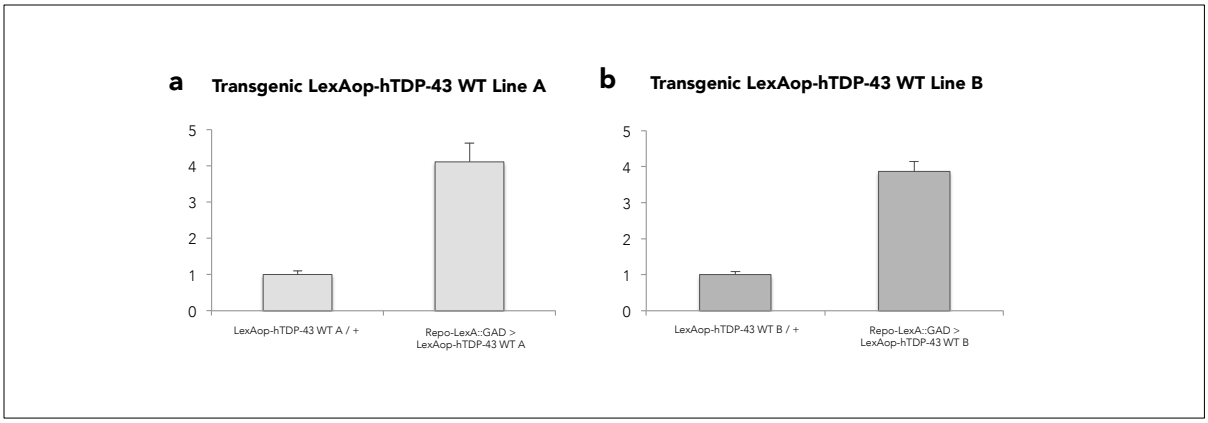


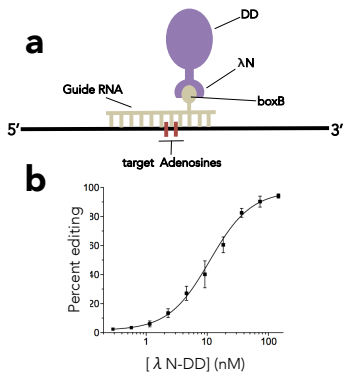
FIGURE 21. Validating LexAop-hTDP-43 expression under *Repo*-LexA::GAD.
(**A**) qPCR validation of expression of the transgenic LexAop-hTDP-43 WT Line A under *Repo*-LexA::GAD in 2-4 day old *Drosophila* head tissue. (**B**) qPCR validation of expression of the transgenic LexAop-hTDP-43 WT Line B under *Repo*-LexA::GAD in 2-4 day old *Drosophila* head tissue. Values for *Repo*-LexA::GAD groups normalized to those for the appropriate control group (i.e.: *Repo*-LexA::GAD > LexAop-hTDP-43 A to LexAop-hTDP-43 WT A / + and *Repo*-LexA::GAD > LexAop-hTDP-43 B to LexAop-hTDP-43 WT B / +, respectively). $N = 4$ for all other groups.

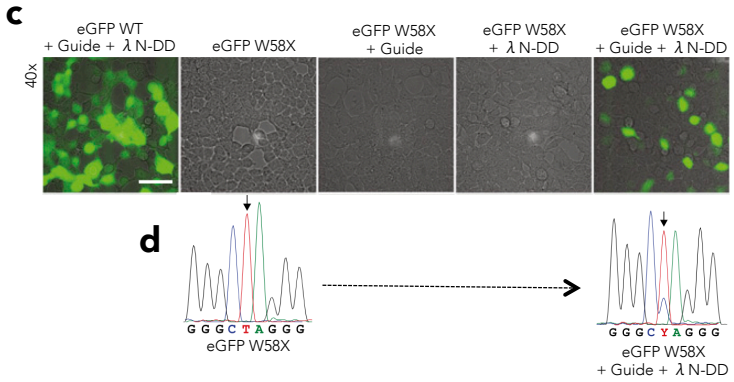
II. λN-Conjugated RNA Editase Constructs:

Reasoning and Design:

The ultimate goal of these tools is to generate a system by which to mark RNA transcripts from different RTE families in a cell type-specific fashion, thereby demarcating the cell type of origin of actively expressed RTEs in preparations from whole *Drosophila* head tissue. The tools described below are intended to harness the activity of natural RNA editing enzymes to alter the sequence of actively transcribed RTE-derived RNA transcripts in such a way that these sequence changes will not alter the natural activity of the element and will be maintained through the transposition cycle all the way through the point where the RTE has re-integrated back into the genome. It is hoped that expressing these RNA editing enzymes with cell type specificity via the GAL-4/UAS system may reveal the cell type of origin of actively transposing RTEs in response to various manipulations by performing high coverage DNA sequencing of preparations of whole head tissue. These tools build on the *in vitro* designs developed by Montiel-Gonzalez, MF, *et al.* in their 2013 *PNAS* paper, in which the authors tethered the catalytic DD of the human ADAR protein to the 22 amino acid λN peptide, which normally mediates binding of the λ-phage N protein to boxB RNA hairpins in order to regulate antitermination during transcription of λ-phage mRNAs. The DD of ADAR normally catalyzes the hydrolytic deamination of A to I, which is read out as G by both translation machinery and sequencing technology. This function makes it an ideal tool by which to mark actively transcribed RNAs of a known sequence. As the λN peptide binds the short (17 nucleotide) boxB hairpin motif with nanomolar efficiency, it makes an excellent tether by which to bring the catalytic activity of the ADAR DD to an RNA substrate. Sequence specificity can be introduced into the system by creating a "guide RNA" that is both complementary to the desired target sequence and contains a boxB hairpin structure (Montiel-Gonzalez et al. 2013). Thus, the guide RNA will localize to actively transcribed RNA species via sequence complementarity, the boxB structure embedded therein will be recognized by the λN peptide in the λN-DD fusion protein, and this will bring the ADAR DD into close

FIGURE 22. The λN-DD system: generating λN-DD constructs for cell type-specific use in *Drosophila* and GFP reporter constructs for functional testing of the λN-DD system *in vivo*. (**A**) Schematic representation of the λN-DD system [*adapted from*: (Montiel-Gonzalez et al. 2013)]. (**B**) Dose-response curve of target editing with regard to concentration of λN-hDD *in vitro* [*adapted from*: (Montiel-Gonzalez et al. 2013)]. (**C**) Fluorescence micrographs demonstrate that λN-DD and guide RNA can restore functional green fluorescence in HEK-293T cells transfected with EGFP W58X [*adapted from*: (Montiel-Gonzalez et al. 2013)]. (**D**) Electropherograms of sequenced RT-PCR products of EGFP W58X cDNA from HEK-293 cells transfected with EGFP W58X alone or in conjugation with λN-DD and guide RNA. Arrows indicated target site [*adapted from*: (Montiel-Gonzalez et al. 2013)]. (**E**) Schematic of λN-dDD and λN-hDD inserts. (**F**) Schematic of myr-eGFP WT and myr-eGFP W58X inserts. (**G**) Insertion into pUAST attB.





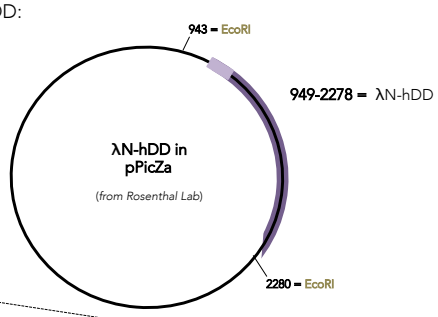
Cloning Strategy:

CMV Forward and M13 Reverse Primer Binding Sites
XhoI / XbaI
5′ FLAG / 6x His
λN Peptide
Drosophila DD homologous to hADAR2 AAs 299-701

λN-dDD:

```
CGCAAATGGGCGGTAGGCGTGC^TCGAGATGGACTACAAGGACGACGATGACAAACTAGT
AAACGCACGAACACGACGACGTGAGCGTCGCGCTGAGAAACAAGCTCAATGGAAAGC
TGCAAACCCAATCGACGACTCGAAGTCGTCATCGATGGAGTTGCCTCAGATACACGCGGAT
ACGATTGGTCGGTTGGTCTTAGAAAAGTTCATGGAAGTAATCAAGGGCCAGGAGGCTTAC
TCGCGTCGAAAGGTATTAGCGGGCATTGTAATGACTGAAAACATGAATTTTTGTGAAGCC
AAAGTTATTTCAGTTTCGACGGGCACCAAGTGTGTCAGCGGTGAGCATATGAGTGTGAACG
GAGCTGTCCTAAATGATTCCCATGCTGAAATAGTCTCCAGGCGTTGTCTTCTCAAATATTT
ATATGCACAGCTGGACCTTCAGTGCAATCAGGCCACAGCATATCAGTCGATTTTCGTGAGG
AATACTGATGGGCAATACCCTTATAAACTAAAATCCGGGGTACATTTCCATTTGTATATAA
ATACAGCACCTTGTGGGGATGCACGGATATTTAGTCCTCACGAAAACGACACTGGTGTTGA
TAAACATCCAAATAGAAAAGCTCGTGGCCAGCTGCGCACCAAATCGAGTCCGGTGAGGG
GACGATTCCAGTGAAAAGCAGTGATGGTATCCAGACGTGGGATGGCGTATTGCAGGGCCA
ACGCCTGCTAACAATGTCGTGCTCGGATAAAATTGCCCGTTGGAACATCGTGGGCATCCAA
GGCTCACTATTGTCTTCCATAATTGAACCGGTGTACCTGCATTCGATTGTGCTGGGAAGCCT
GTTGCATCCGGAGCACATGTACCGCGCAGTTTGCGGCCGAATTGAGAAGTCCATTCAAGGC
CTGCCACCGCCGTACCATCTGAACAAGCCGCGCCTGGCACTAGTCACTTCGGCCGAGCCGCG
GAATCAGGCCAAGGCTCCCAACTTCGGGATCAATTGGACCATTGGGGACACCGAGCTGGA
AGTGGTGAACTCGCTAACCGGTCGAACCATTGGCGGCCAGGTGTCTCGCATCACAAAGCAG
GCGTTTTTCGTTAAGTATGGATTTCTAATGGCAAACTTGCCCGGTATTTTAGTCCGCAAAG
TAACCACTGACTATGGGCAAACCAAAGCTAACGTTAAAGGACTATCAGATCGCAAAGCTA
GAATTGTTCTCTGCATTCAAGCGAGAAGACCTTGGCAGCTGGCTGAAGAAACCCATTGAA
CAAGACGAGTTCGGTCTTGCCGAACCACCATCACCATCACCATTCTAG^ATCCTGTGTGAAA
TTGTTATCCGCT
```

λN-hDD:



e

f

CMV Forward and M13 Reverse Primer Binding Sites
XhoI / XbaI
Myr
Polymorphism

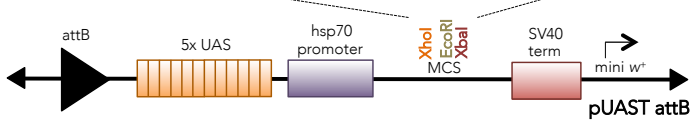
myr-eGFP WT:

```
CGCAAATGGGCGGTAGGCGTGC^TCGAGATGGGCAACAAATGCTGCAGCAAGCG
ACAGGATCAGGAACTGGCACTGGCCTATCCCACTGGGGGCTACAAGAAATCCGA
CTACACCTTTGGCCAGACGCACATCAACAGCAGCGGCGGCGGCAACATGGGCGGC
GTTCTTGGCCAGAAGCATAACAACGGTGGCTCGCTGGACTCGCGCTACACGCCCG
ATCCCAATCATCGGGGTCCGTTGAAAATCGGCGGAAAGGGCGGCGTTGACATCA
TCAGACCACGCGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT
GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCC
GGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCA
CCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGC
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGT
CCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGG
CAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAA
GCTGGAGTACACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGT
GCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG
CTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACG
AGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCT
CGGCATGGACGAGCTGTACAAGTCTAG^ATCCTGTGTGAAATTGTTATCCGCT
```

myr-eGFP W58X:

```
CGCAAATGGGCGGTAGGCGTGC^TCGAGATGGGCAACAAATGCTGCAGCAAGC
GACAGGATCAGGAACTGGCACTGGCCTATCCCACTGGGGGCTACAAGAAATCCG
ACTACACCTTTGGCCAGACGCACATCAACAGCAGCGGCGGCGGCAACATGGGCG
GCGTTCTTGGCCAGAAGCATAACAACGGTGGCTCGCTGGACTCGCGCTACACGC
CCGATCCCAATCATCGGGGTCCGTTGAAAATCGGCGGAAAGGGCGGCGTTGAC
ATCATCAGACCACGCGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGG
GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGC
GTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTC
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGACCCACCCTCGTGACCACCCTGA
CCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT
CTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAG
GACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTG
GTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACACTACAACAGCCACAACGTCTATATCATGGCCGAC
AAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGA
CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGC
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAG
ACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCG
GGATCACTCTCGGCATGGACGAGCTGTACAAGTCTAG^ATCCTGTGTGAAATTG
TTATCCGCT
```

g



proximity with the targeted transcript (Figure 22A). The deaminase activity of DD will then convert the appropriate A to I within the target of interest, which will be read out as a G.

The authors functionally tested their system in HEK-293T cells using a GFP reporter. The W58X point mutation of eGFP encodes a premature stop codon where a W normally would reside, such that no full-length eGFP protein is produced when it is expressed. Importantly, this is due to a base conversion of the third "wobble" position in the codon from UGG to UGA such that editing of just one base in the mRNA transcript by ADAR DD converts the premature stop codon back to the appropriate codon for W, resulting in translation of full-length eGFP protein. After careful testing of the parameters of guide RNA positioning with respect to DD catalytic efficiency, the authors were able to generate robust expression of full-length eGFP protein when eGFP W58X was co-transfected with λN-DD and an eGFP-X guide RNA *in vitro* (Montiel-Gonzalez et al. 2013) (Figure 22C and 22D).

I have undertaken to develop the λN-DD fusion protein from the original study, which employs the DD from human ADAR (λN-hDD), and an analogous fusion protein that employs the DD from *Drosophila* ADAR (λN-dDD), for use in the GAL-4/UAS system such that they may be implemented *in vivo* with cell type specificity. I have subcloned the construct from (Montiel-Gonzalez et al. 2013) into the pUAST vector, which places it under UAS control. I have generated a new construct that replaces the hDD with dDD, and have subcloned this into pUAST as well (Figure 22E and 22G). The idea behind using the two different DDs is simply to couch for the possibility that the dDD may work better in the *Drosophila* cellular milieu. In order to functionally test these two constructs *in vivo*, I have likewise subcloned both eGFP W58X and a full length, wild type eGFP positive control into pUAST (UAS-myr-eGFP W58X and UAS-myr-eGFP WT, respectively) (Figure 22F and 22G). Both of these GFP constructs carry an N-terminal myristoylation sequence that will target the translated protein to the membrane for clean delineation of the morphology of the cell types in which they are expressed by confocal microscopy. Finally, I have made use of a vector designed to express CRISPR guide RNAs under the *Drosophila* U6 promoter for the generation of a pU6-eGFP-X

guide RNA construct using the eGFP-X guide sequence that was so effective *in vitro* in (Montiel-Gonzalez et al. 2013) (Figure 23A). The U6 promoter is used here because it is a strongly expressing RNA Pol III promoter, which facilitates the expression of uncapped and non-polyadenylated RNAs. Once a Gal4 driver, a λN-DD construct, an eGFP construct, and the eGFP-X guide RNA are all present in the same animal, the λN-DD and eGFP constructs will both be expressed in the same cells by the Gal4 driver, while the guide RNA will be ubiquitously expressed under pU6.

Validation:

I have subcloned the two deaminase constructs, λN-hDD and λN-dDD, into pUAST such that they will be under UAS control in the transgenic animal. Subcloning was validated by Sanger sequencing and both constructs were sent out for injection to generate transgenic fly lines according to standard protocol. I further subcloned the myr-eGFP WT and myr-eGFP W58X constructs for functional testing of the two deaminase constructs into pUAST as well. These were also validated by Sanger sequencing and sent out for injection. Both the pair of deaminase constructs and the pair of eGFP constructs were attP/attB integrated to minimize expression level differences within each pair due to variations in the genomic context of the insertion site. Transformants were again validated by PCR of genomic DNA. I have gone on to outcross two transformant lines of each the λN-hDD and λN-dDD constructs to our in-house wild type fly strain for 5 generations, such that if any of these lines proves effective they will be ready for use in sequencing experiments.

I attempted to generate a pU6-eGFP-X guide RNA transformant line with which to test the effectiveness of the λN-hDD and λN-dDD constructs *in vivo* in combination with the UAS-myr-eGFP W58X reporter. Given that my personal communication with Dr. Rosenthal revealed that expression levels of the guide RNA appeared to be rate-limiting in *in vitro* tests of this system, I used the most highly expressing pU6 expression vector I could find. I succeeded in generating a pU6-eGFP-X guide RNA construct with this vector, however transformant lines were near impossible to generate and the only one that the injection facility was ever able to make never was confirmed by PCR validation. I

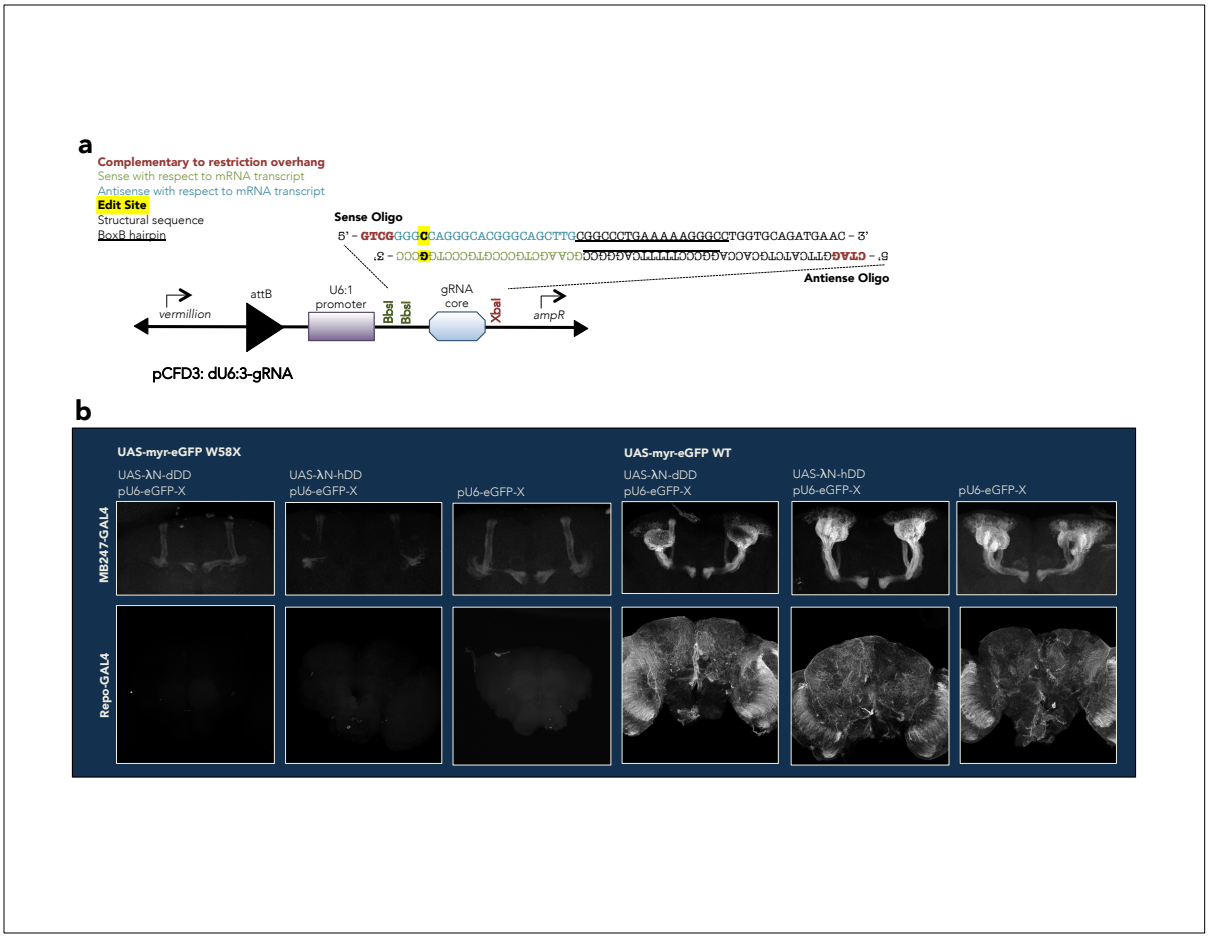


FIGURE 23. **Generating a pU6-eGFP-X guide RNA for use in *Drosophila* and functional testing of the λN-DD system in neurons and glial cells of the fly CNS.**
(**A**) Schematic representation of pU6-eGFP-X guide RNA cloning strategy. (**B**) Functional testing of λN-DD system in neurons (under *MB247*-GAL4) and glia (under *Repo*-Gal4) of the *Drosophila* CNS. The two GFP constructs appear to be working as expected, as there is a high level of GFP signal in the positive control groups (UAS-myr-eGFP WT) and very little GFP signal in the groups carrying GFP with a premature stop codon (UAS-myr-eGFP W58X). However, no editing of the GFP W58X reporter is detectable by either λN-dDD or λN-hDD in either neurons (*MB247*-GAL4 > UAS-λN-dDD + pU6-eGFP-X + UAS-myr-eGFP W58X and *MB247*-GAL4 > UAS-λN-hDD + pU6-eGFP-X + UAS-myr-eGFP W58X) or glia (*Repo*-GAL4 > UAS-λN-dDD + pU6-eGFP-X + UAS-myr-eGFP W58X and *Repo*-GAL4 > UAS-λN-hDD + pU6-eGFP-X + UAS-myr-eGFP W58X). N = 3-4 per group.

tested this single guide RNA transformant line with the eGFP and deaminase constructs under two separate Gal4 drivers, and while is appears that UAS-myr-eGFP WT and UAS-myr-eGFP W58X are working properly, I was not able to detect any editing of UAS-myr-eGFP W58X by either λN-hDD or λN-dDD (Figure 23B). My intuition is that the guide RNA transformant line was a rare escaper that does not actually carry pU6-eGFP-X, and that we are running up against an RNA toxicity issue with the high levels of expression of this short guide RNA from the constitutive U6 promoter. I am in possession of another pU6 vector that does not express at quite such high levels, and it would be a simple matter to place the eGFP-X guide sequence into this vector and send it out for injection. Finally, I have generated 2 guide RNA sequences targeting highly conserved A residues in *gypsy* Gag and 2 in *gypsy* Pol based on the parameters for guide RNA positioning laid out in (Montiel-Gonzalez et al. 2013), and am in possession of the oligonucleotides necessary to clone these sequences into the appropriate pU6 vector for testing. I therefore believe that if any of the λN-hDD or λN-dDD transformant lines prove effective in the eGFP assay, it should be relatively quick and easy to move directly to testing efficacy on marking transcripts derived from the *gypsy* RTE in a cell type-specific fashion via sequencing technology.

III. Tissue Collections for RNA Sequencing:

Reasoning and Design:

Given the glial specificity of hTDP-43's effects on *gypsy* presented in Chapter 3, we are interested in the possibility of cell type-specific activation of different TE classes in different cell types in response to TDP-43 pathology. As *loki*(IR) provides such dramatic rescue of hTDP-43's effects on lifespan in both neurons and glia, while *gypsy* expression only appears to be induced by hTDP-43 expression in glia, this kind of cell type specificity seems to be a strong possibility. Such a phenomenon could reflect the different epigenetic environments or different somatic TE composition of different cell types, or another as-of-yet unimagined explanation. In order to address whether this possibility is in fact a reality, I have collected large quantities of flies expressing hTDP-43 in either neurons or glia along with appropriate control groups at both young and aged time points (Figure 24A). This sample set is intended to generate large-scale RNA preps from whole head tissue with which to perform both RNA sequencing and small RNA sequencing. In addition to addressing the question of whether hTDP-43 expression induces the expression of different TE types in different cell types, small RNA sequencing from these groups would provide insight as to whether hTDP-43 expression interferes with siRNA silencing of TEs by interrupting siRNA biogenesis or somewhere downstream, such as the localization of loaded Ago2/RISC to the target RNA.

Validation:

We have a tried-and-true protocol for isolating large quantities of *Drosophila* heads from whole frozen flies in the Dubnau lab. Lisa Prazak, a post-doc in our lab, is currently working with Nikolay Rozhkov in Molly Hammell's lab to generate the small RNA libraries diagrammed in Figure 24B. Dr. Rozhkov has extensive experience with both small RNA sequencing and RNA sequencing from *Drosophila* tissue from his previous experience in Greg Hannon's lab, and once the libraries have been sequenced they will be analyzed by Molly Hammell's lab as part of our standing collaboration.

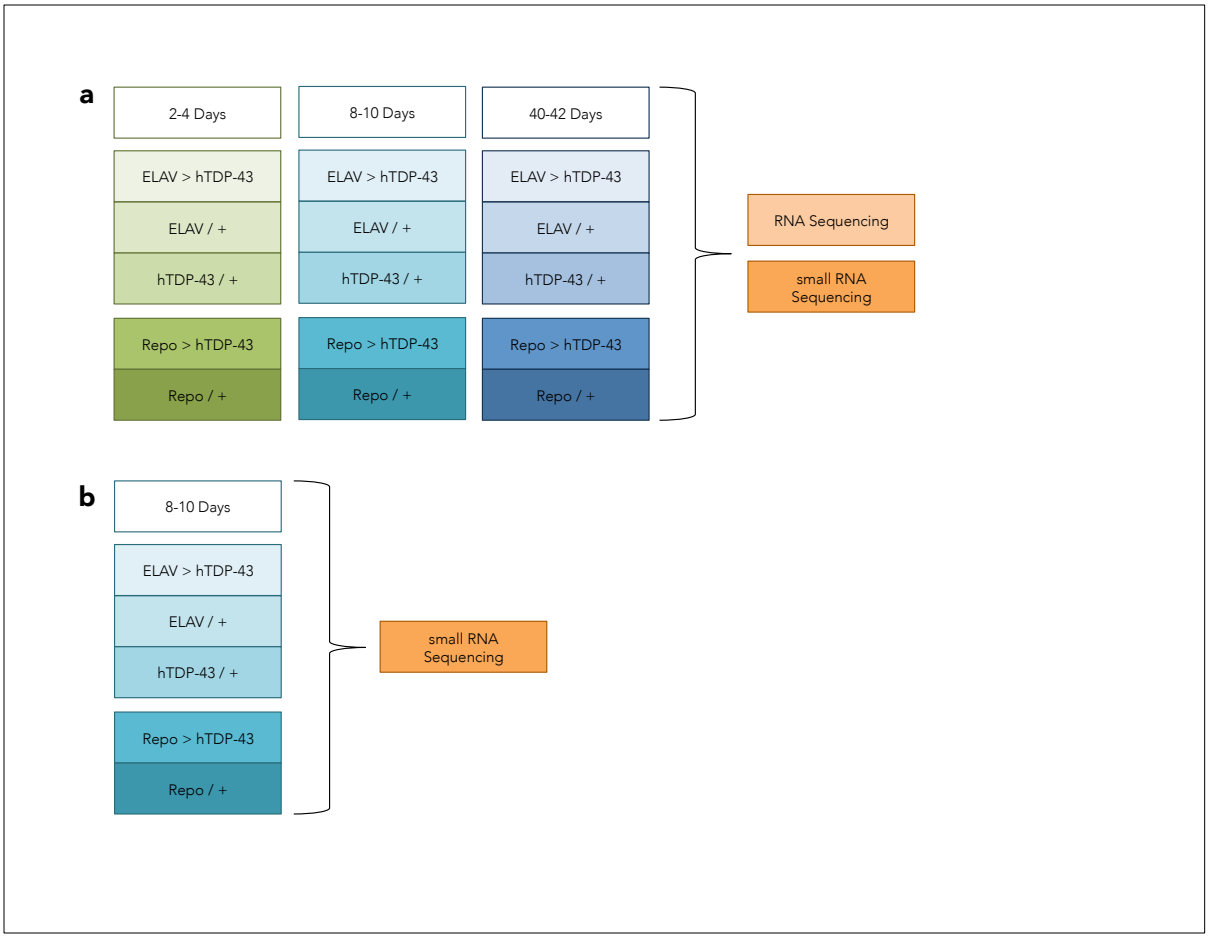


FIGURE 24. Sample sets for RNA sequencing comparing neuronal versus glial expression of hTDP-43 in *Drosophila* head tissue.
(**A**) Schematic representation of the sample set for large-scale RNA preps from head tissue of flies expressing hTDP-43 in neurons versus glia and genetic controls at young and aged time points. (**B**) Small RNA sequencing library preps currently being processed by Dr. Prazak.

CHAPTER 5 – Conclusions and Perspectives:

In the Introduction I have discussed the ways in which TEs have exerted a powerful influence over the evolution of their host genomes. In so doing they have been co-opted to perform critical roles in genome regulation, likely contributing to morphological and neurological diversity within the mammalian lineage (van de Lagemaat et al. 2003; Sandelin et al. 2004; Lindblad-Toh et al. 2005; Woolfe et al. 2005; Polak and Domany 2006; Thornburg et al. 2006; Lowe et al. 2007; Mikkelsen et al. 2007), and endowing the acquired immune system with the flexibility to respond to unpredictable invasion of foreign antigens via the V(D)J recombination mechanism (Kapitonov and Jurka 2005; Sen and Oltz 2006; Slotkin and Martienssen 2007). We now know that they are active during brain development as well (Muotri et al. 2005; Garcia-Perez et al. 2007b; Coufal et al. 2009; Kuwabara et al. 2009; Muotri et al. 2009), and that this activity is likely to be of functional relevance as novel somatic L1 insertions tend to transpose near genes important for nervous system function (Muotri et al. 2005; Baillie et al. 2011; Upton et al. 2015). Indeed, removing the ability to repair DNA damage in mice results in embryonic lethality with massive defects in both lymphogenesis and neurogenesis (Gao et al. 1998), revealing a critical window in the development of these two systems in which the repair of DSBs is of paramount importance. However, we also know that the unbridled activity of even one active TE class is enormously destructive, as in the case of hybrid dysgenesis (Schaefer et al. 1979; Castro and Carareto 2004). Here, introduction of one novel TE into the genome results in sterility and complete destruction of the germline tissue (Schaefer et al. 1979). I have also presented data to the effect that activation of RTEs by TDP-43 protein pathology causally contributes to physiological deterioration and that the cellular toxicity induced by TDP-43 pathology is mediated by Chk2 signaling following DNA damage. We know that TDP-43 protein pathologies arise sometime at or just after middle age, and that this happens, for the most part, in individuals with no known genetic lesion precipitating pathology (Arai et al. 2006; Neumann 2009; Ling et al. 2013). In fact, this is a

common theme for most neurodegenerative diseases. So what is it about the aged somatic environment that makes the nervous system so susceptible to pathology? And how can TEs, whose regulated activity appears to be constructive during neuronal development, become a driving destructive force once someone has reached middle age? Below, I will discuss the evolutionary theory of aging and how TEs and TE control mechanisms fit into this framework. Finally, I will discuss how this relates to what we know about TEs in the nervous system and I will use both my thesis work and other recent work from the Dubnau lab to build a model of how TDP-43 pathology fits into the context of normal neurological decline in wild type aging. I believe this exercise has the potential to provide insight into the pathological prognosis of neurodegenerative disease in general, as TEs have been linked to many different types of neurodegenerative disease in recent years.

I. What is Senescence?

Selective pressure has shaped all of the beneficial functions that TEs have been described to participate in thus far, including the evolution of the linear chromosome, regulation of the deployment of genetic information, and adaptability of the acquired immune system. By its very nature, selective pressure finds fitness maxima – perhaps in individuals with better immune systems or more complex nervous systems that are better equipped to survive in their environments and provide better care for their mates and their offspring. However, it is readily apparent that as an organism ages, it will inevitably become less fit until it succumbs to environmental stressors and passes away. The first theory of aging was published in 1891 by August Weissman, in which he postulated that death of older individuals of the population may have evolved for the benefit of the species such that younger individuals may have more access to resources, thus enabling the evolutionary process (Weismann 1891). This theory implies that individuals undergo a type of "programmed death mechanism" to ensure a limited life span. Contemporary arguments against this theory held two major objections, namely: (1) animals living in the wild rarely live long enough to die of old age. Therefore, the

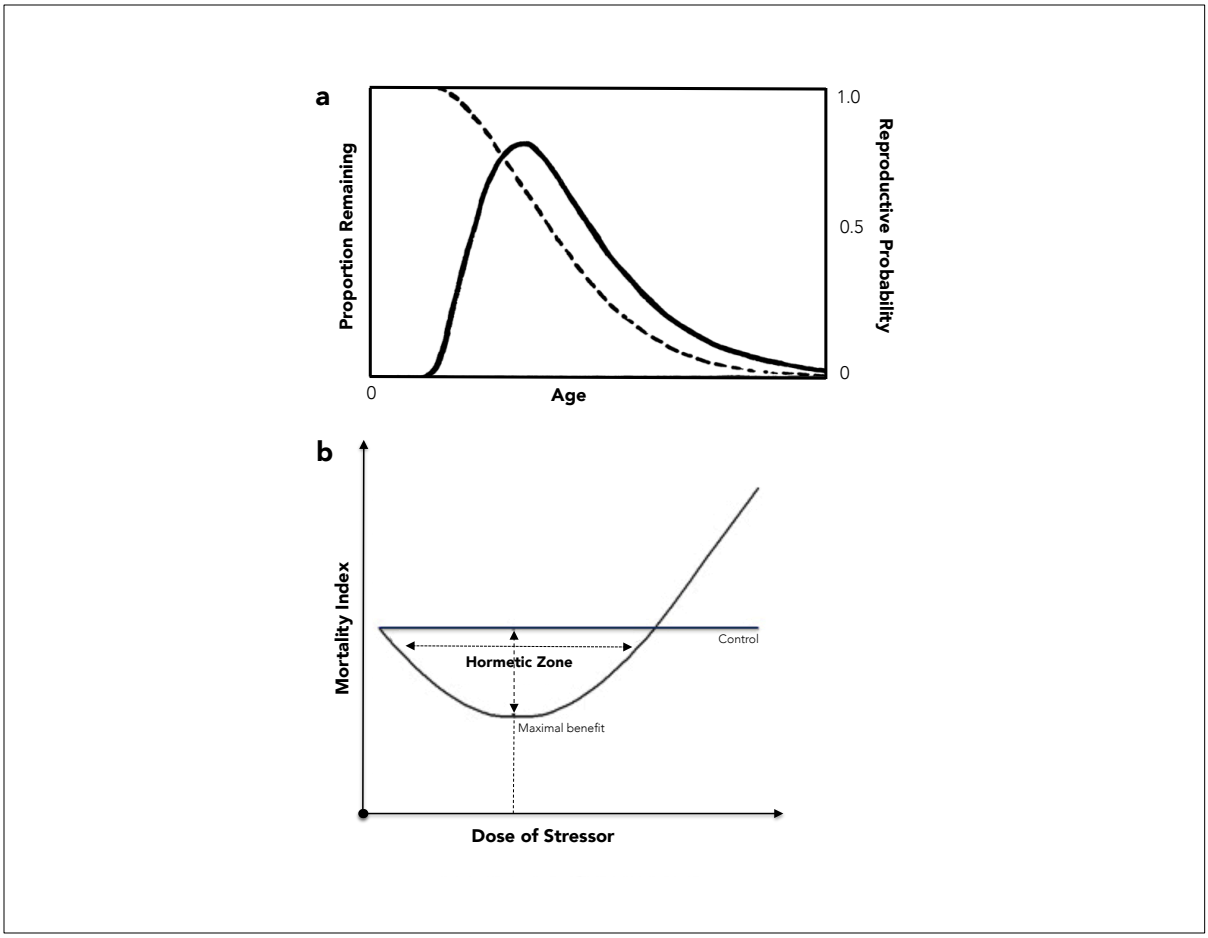


FIGURE 25. The reproductive probability distribution and the response of populations exposed to stressors.
(**A**) The relationship of age to the reproductive probability distribution. The solid line is the reproductive probability distribution; the dashed line indicates the proportion of the total probability that remains after any given age [*adapted from*: (Williams 1957)]. (**B**) Hormesis is experienced in populations with a low dose of a given stressor [*adapted from*: (Parsons 2005)].

mechanism would scarcely, if ever, have an opportunity to operate, raising questions as to how and why it would have a chance to evolve in the first place. (2) Darwinian selection is, by definition, survival of the fittest. Any organism that lives longer and deteriorates less readily over time will leave behind more offspring, thus actively negating the concept of programmed death (Comfort 1956; Williams 1957).

However, we undoubtedly observe a progressive aging process, a phenomenon that has been termed "senescence." As senescence is subject to genetic variability (Gonzales 1923; Kallmann 1948), later theorists began to conceptualize it as a negative trait. However, its pervasiveness dictated that a secondary force must be at play that "favors its [senescence's] development in such a way that the observed variations in senescence reflect variations in the balance between these two forces" (Williams 1957). By the mid-1950's, theorists were beginning to understand the aging process as resulting from depreciation of selective pressure with increasing age (Bidder 1932; Haldane 1941; Medawar 1953; Comfort 1954; Medawar 1955; Comfort 1956; Williams 1957), an effect that makes it appear highly variable and stochastic. Both Peter Medawar and George Williams further elaborated on this idea using the concept of a reproductive probability distribution (Figure 25A), which is based on the equation $W = (1 + m_1p_1) + (1 + m_2p_2) + \ldots (1 + m_np_n)$; where m = the magnitude of the effect and p = the relevant proportion of the reproductive population (Medawar 1953; Medawar 1955; Williams 1957).

The reproductive probability distribution takes into account that there is always a cumulative probability of death even in the absence of senescence, as survival is always more likely to age A than to age $A+1$. Since reproductive probability at a given age is a function of survival to that age, this results in a natural decay in reproductive probability. Therefore, whenever there is a conflict of interest natural selection will always favor youth over old age, as any genic effect that arises in old age will inherently have the smaller p-value. As p begins to drop off at reproductive maturation, theoretically this point should denote the onset of senescence processes (Medawar 1953; Medawar 1955; Williams 1957). As each new genic effect in turn is evaluated based on the current reproductive

probability distribution, a previously established senescence factor actually enables the establishment of additional senescence factors, resulting in a self-perpetuating exacerbation of senescent deterioration (Medawar 1953; Medawar 1955; Williams 1957). In this manner, natural selection prefers maximizing vigor in youth, and this is intrinsically accompanied by a progressive decline in vigor, or senescence, in adult life into old age (Medawar 1953; Medawar 1955; Williams 1957). It is important to note that such a model in fact argues *against* any one archetypal "senescence gene" acting within a population, as its presence would induce positive selection for genes that would mitigate or postpone its adverse effects (Williams 1957).

This model leads to two predictions about the nature of the senescence process. The first, advocated by Peter Medawar, is that senescence reflects either the loss of beneficial genic effects or gain of harmful genic effects (or some combination thereof) in later life due to a relaxation in selective pressure to either maintain or remove them (Medawar 1953; Medawar 1955). The second, put forth by George Williams, states that senescence reflects pleiotropic genic effects at different ages since the selective value of all genes are based on their effects on total reproductive probability (Williams 1957). In this case, both the inflection and the timing of the effect are important. An advantage bestowed before or during the period of maximum reproductive probability would increase the total reproductive probability more than an equivalent disadvantage in later life would reduce it (Williams 1957). The latter of these two predictions is commonly referred to as negative or antagonistic pleiotropy (Williams 1957; Parsons 2007).

Organisms in the wild are faced with energetically or nutritionally inadequate, harsh environments, thus their survival during the aging process is mainly dictated by their ability to efficiently respond to the metabolic and energetic challenges presented to them (Capy et al. 2000; Parsons 2005; Parsons 2007). A fitness-stress continuum is usually employed to express variation in the severity of environmental stressors on natural populations, wherein the reciprocal of the stress intensity relates to the average fitness of an individual within the population (Parsons 2005; Parsons 2007). If stressors are too extreme, the organism cannot survive. Fitness maxima are observed under

conditions of mild stress, with commonly cited examples including caloric restriction, physical activity, temperature, et cetera. Such maxima result in prolonged longevity, an effect that is titled hormesis (Capy et al. 2000; Parsons 2005; Parsons 2007; Zhu et al. 2014) (Figure 25B). Senescence, then, is a product of benign environments. Perhaps the best evidence to test predictions derived from the reproductive probability distribution come from humans living in modern, civilized societies where healthcare and technology have effectively relaxed selection for stress resistance, resulting in increased survival and reproduction of relatively unfit individuals who would have been poorly equipped to survive in hunter-gatherer environments, and increased median age and life expectancy, resulting in massive expansion of our post-reproductive population. Understanding senescence is, therefore, the most pressing medical issue of our age (Tucker et al. 1999; Martin 2007).

II. TEs and Senescence:

There is no need to believe that the theories presented above are restricted solely to the effects of protein coding genes. If the genomic revolution has taught us anything, it is that we should look more comprehensively, to understand the non-coding and epigenetic context of the coding portions of our genome in order to come to a full appreciation of genetic-environmental interactions. Many theories regarding functional decline during aging invoke the progressive inability to maintain cellular structure (Macieira-Coelho 1991; Kennedy et al. 1995; Oberdoerffer and Sinclair 2007; Wilson et al. 2008; Wood and Helfand 2013). Indeed, age effects on chromatin structure are well documented (Wood and Helfand 2013). While the signs of gains or losses of individual activating or repressive histone marks vary in different studies and different systems, recent evidence depicts global degradation of chromatin architecture as a hallmark of aging, with dramatic implications for both gene and TE activity (Wood et al. 2010; De Cecco et al. 2013b; Wood and Helfand 2013).

One of the earliest systems developed for cellular aging studies is yeast replicative senescence. The phenotypic read out for this system is the number of times a mother cell can divide to

produce a daughter cell before entering into a senescent state (Wood and Helfand 2013). Early studies of yeast replicative senescence described heterochromatin loss at ribosomal DNA repeats, the mating type locus, and telomeres in senescent cells (Kim et al. 1996; Smeal et al. 1996; Kennedy et al. 1997). Such observations prompted a preliminary heterochromatin loss model of aging wherein a weakening or breaking down of critical constitutive heterochromatin results in detrimental changes to cellular homeostasis; with increased transcriptional noise obscuring important cellular functions and precious cellular energy being taken up synthesizing unimportant, or even detrimental transcripts or repairing other types of damage (Smeal et al. 1996; Villeponteau 1997; Tsurumi and Li 2012). Indeed, total histone protein levels are diminished in senescent yeast cells, and overexpressing histone proteins can artificially suppress senescence (Feser et al. 2010).

A more recent model of cellular senescence has emerged which exploits the replicative senescence of cells in tissue culture. Cells will only divide a finite number of times under these conditions before entering into a senescent state (Wood and Helfand 2013). Various senescence-dependent histone mark alterations have been described in this system, and reductions in total histone protein levels are also observed (O'Sullivan et al. 2010; O'Sullivan and Karlseder 2012; Ivanov et al. 2013; Shah et al. 2013). These cells also develop characteristic γH2Ax-positive non-pericentromeric SAHF (Narita et al. 2003; Zhang et al. 2005; Kosar et al. 2011; Chandra et al. 2012). A global study by DeCocco et al (2013) implemented the FAIRE technique (Giresi and Lieb 2009) on normal human diploid fibroblasts to map genome-wide conformational change on chromatin state in senescent versus actively growing cells (De Cecco et al. 2013b). It is important to note here that FAIRE can only mark activation not repression, due to the fact that formaldehyde cross-linking is more efficient on nucleosome-bound DNA, resulting in sequencing of non-crosslinked DNA that is predominantly open chromatin. The results were quite remarkable – FAIRE profiles of senescent cells were noticeably smoothened in comparison to growing cells (Figure 26A.3), whose genomes are replete with topography of both closed and open chromatin. This effect is mediated by both FAIRE signal loss in active, gene-rich regions of chromatin (Figure 26A.2) and FAIRE signal gain in

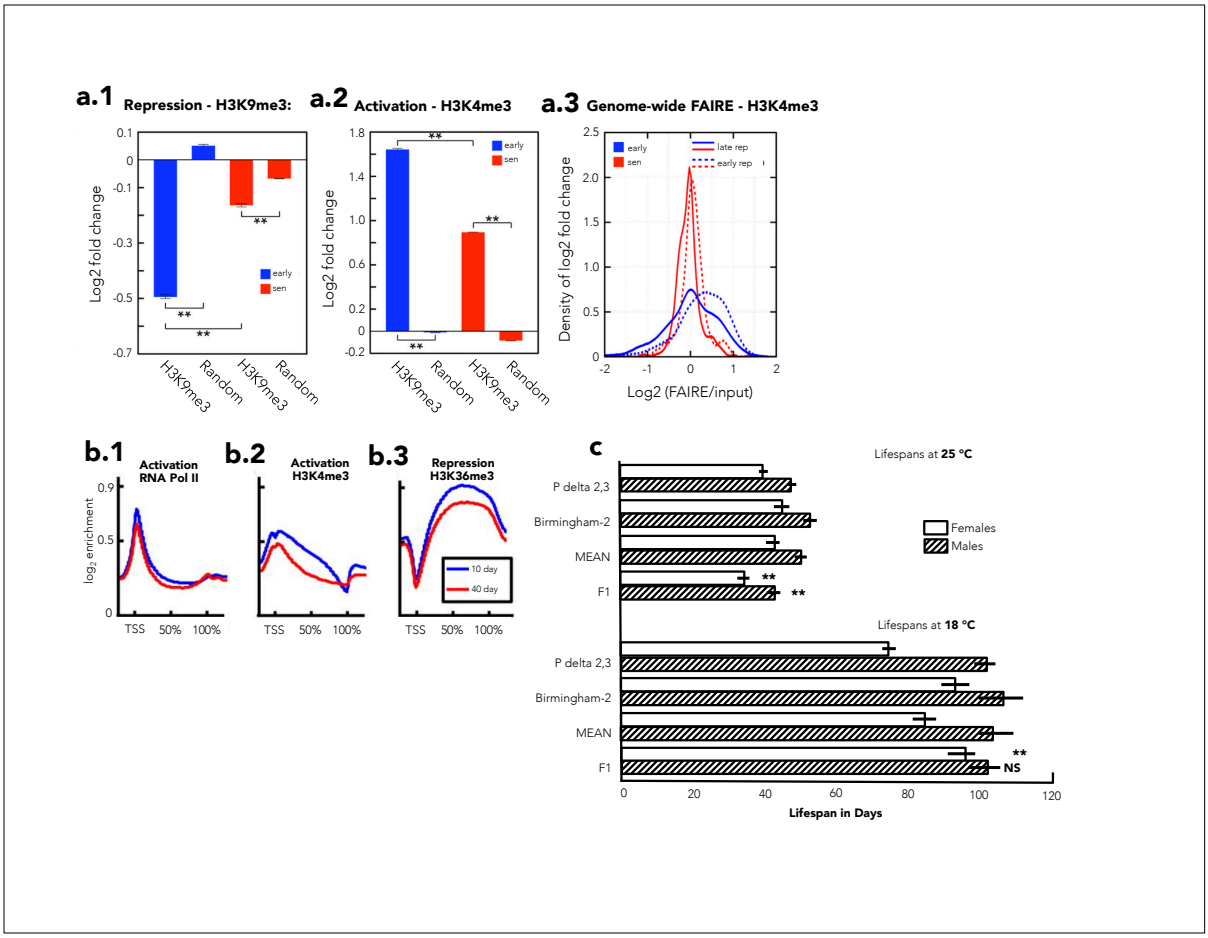


FIGURE 26. Activating and repressive chromatin structure 'smoothens' with age in an *in vitro* replicative senescence model and *Drosophila* organismal aging, while TE activity negatively impacts lifespan in *Drosophila*.

(**A**) FAIRE signal 'smoothing' in senescent normal human diploid fibroblasts. (**A.1**) Gain of FAIRE enrichment in areas marked by repressive H3K9me3 histone modification, with simultaneous (**A.2**) loss of FAIRE enrichment in regions marked with activating H3K4me3 histone modification in senescent (red) versus early passage (blue) cells as compared to randomized controls. (**A.3**) The genome-wide FAIRE enrichment profile with the activating histone mark H3K4me3 demonstrates a dramatic reduction in FAIRE signal enrichment in senescent cells compared to their early passage counterparts [*adapted from*: (De Cecco et al. 2013b)]. (**B**) Chromatin 'smoothing' in aged *Drosophila melanogaster*. The mean $\log_2$ ChIP signal is shown on a composite gene representing the exonic regions of all ~14,000 genes in the *Drosophila melanogaster* genome, scaled from the TSS to the transcriptional stop (100%). Loss of activating H3K4me3 histone modifications (**B.2**) and gain of repressive H3K36me3 histone modifications (**B.3**) are observed in 40 day old flies (red) compared to their 10 day old counterparts (blue) [*adapted from*: (Wood et al. 2010)]. (**C**) Activating the P-element transposon significantly shortens lifespan in *Drosophila melanogaster*. Lifespans are shown at 18 °C (bottom) and 25 °C (top) for females (white bars) and males (hatched bars) of parental strains carrying the second chromosome of the Birmingham strain, with 17 non-autonomous P elements (Birmingham-2) or a P-element with an in-phase deletion that renders its transposase inactive at 18 °C but somatically active and unable to act on its own element at 25 °C (P delta 2,3) in the Canton-S genetic background. The mean of the two parental strains for each sex is also displayed (MEAN), as is that for the F1 progeny (F1) at each temperature [*adapted from*: (Driver and McKechnie 1992)].

heterochromatic, gene-poor regions (De Cecco et al. 2013b) (Figure 26A.1). Centromere structures relax and enlarge in the senescent state, with a concomitant increase in transcription of satellite sequences. Similarly, the heterochromatin surrounding Alu, SVA, and L1 elements becomes relatively more open, with a more pronounced effect on more evolutionarily recent elements. This effect is accompanied by increased transcription of these elements, and even active transposition, as an increase in genomic copy number of specific elements is observed in late-stage senescent cells (De Cecco et al. 2013b). Some of these effects arise early in senescence, while others such as TE mobilization only become prevalent at later stages, perhaps due to the prerequisite of overcoming multiple redundant repressive mechanisms (De Cecco et al. 2013b).

Similar age-dependent chromatin effects have been reported at the organismal level in *Drosophila*. Wood and colleagues (2010) performed ChIP on young and aged flies for the activating marks RNA Pol II and H3K4me3 and the repressive mark HP1, followed by whole genome tiling microarrays. A similar "chromatin smoothing" was observed in aged flies, although due to the way these data were generated and normalized, absolute values could not be measured and therefore it was impossible to conclude whether H3K4me3 or HP1 were gained or lost in heterochromatic regions with age (Wood et al. 2010) (Figure 26B.1 – 26B.3). Since neither of these two studies have determined the cause of this "chromatin smoothing" phenotype, they raise questions regarding whether the phenotype reflects general regulatory decay with age or some type of adaptive or compensatory effect, such as intensified efforts to maintain TE repression "spilling over" into the rest of the genome (Wood et al. 2010). If this phenotype indeed reflects general regulatory decay, it does seem to fit quite well with Peter Medawar's prediction that senescence reflects loss of beneficial genic functions or gain of harmful genic functions due to relaxation of selective pressures with age. If, on the other hand, it reflects intensified efforts to suppress TEs spilling over into the rest of the genome, this effect could be interpreted as mitigating action intended to suppress the archetypal antagonistic pleiotropic activity of TEs in old age.

DNA methylation and DSB repair processes also begin to decay in old age (Li et al. 2008; De Cecco et al. 2013b). Both *in vivo* tissues and cell culture replicative senescence models display reductions in total genomic DNA methylation (Sedivy et al. 2008), likely due to a reduction in levels of the DNA methyltransferase DNMT1 as cells approach senescence (Young et al. 2003). Curiously, these effects are mostly due to reductions in methylation of repetitive DNA (Sedivy et al. 2008). NHEJ becomes more error-prone and less efficient in senescent cells (Seluanov et al. 2004), and defects in NHEJ have been observed in both the brains of aging rats (Ren and Pena de Ortiz 2002; Vyjayanti and Rao 2006) and in Alzheimer's patients (Shackelford 2006). Moreover, reports from the literature suggest that an organism's age as a function of its lifespan appears to predict frequency of chromosomal aberrations more accurately than strict chronological time (Crowley and Curtis 1963; Martin et al. 1985; Ramsey et al. 1995; Tucker et al. 1999; Li et al. 2007; Li et al. 2008). Such observations leave us with a type of "chicken-or-the-egg" problem in determining the causality of TE disregulation in cellular senescence: either a slow down of global cellular regulatory processes ultimately proves toxic to cells during senescence, with TE activation simply being an annoying side effect; or the destructive activation of TEs, as a consequence of this general biological slow down, feeds forward to contribute to cellular decline and eventually overwhelms the system, ultimately causing the cell to succumb.

A series of observations from highly divergent organisms cast an incriminating light on TEs and implicate mobile elements in causally contributing to cellular senescence. Over 20 years ago, Driver and McKechnie (1992) reported that post-developmentally activating the P-element TE significantly reduces lifespan in *Drosophila* (Figure 26C). The authors also noted increased genomic copy number of P-element and Copia TEs in aged wild type flies (Driver and McKechnie 1992). More recently, a study by Maxwell and colleagues (2011) reported that the yeast LTR element Ty1 exhibits elevated transposition in very old populations, with new transposition events associated with chromosomal rearrangements in an *S cerevisiae* chronological aging model. Treatments and mutations that reduce Ty1 retrotransposition were found to attenuate age dependent increases in

chromosome loss and loss of heterozygosity events (Maxwell et al. 2011). Finally, a study by Wang et al. (2011) found that elevated *Alu* transcription induces loss of efficient DNA repair in pericentromeric chromatin as well as formation of persistent SAHF-like γH2Ax-positive DNA damage foci in *ex vivo* aging. Amazingly, stable suppression of *Alu* transcription by shRNA knock-down was shown to cause cells to revert to an "*iPS*-like" phenotype, restore their proliferative faculties, and elevate transcript levels of Nanog and Oct4, two critical pluripotency regulators (Wang et al. 2011).

TE derepression alone can be massively mutagenic and deleterious to the cell (Belgnaoui et al. 2006; St Laurent et al. 2010). This destructive behavior has been experimentally validated in an *in situ* system designed to recapture *de novo* TE insertions of a liberated L1 element in somatic tissue culture. Using this system, Symer et al. (2002) documented "numerous L1 element inversions, exon deletions, extra nucleotide insertions, a chromosomal inversion, and comobilization of flanking sequence" (Symer et al. 2002); recapitulating many types of genomic instability long attributed to TEs. High copy numbers of DNA elements have been shown to induce DSBs in human cell lines (Belgnaoui et al. 2006), while overexpression of L1 from muticopy plasmids not only produces elevated quantities of DSBs, but also results in γH2Ax foci reminiscent of SAHF (Gasior et al. 2006). Furthermore, *in vitro* overexpression of the L1 ORF2 endonuclease results in DNA damage, cellular senescence, and apoptosis in human cells (Belgnaoui et al. 2006; Wallace et al. 2008). As mutations that alter lifespan commonly occur in genotoxic stress-response and DNA repair pathways, the capacity to repair DNA damage is a key player in determining lifespan (Capri et al. 2006; St Laurent et al. 2010). If TEs are even partially responsible for such DNA damage, they may indeed be the spark that catalyzes deterioration, and ultimately demise, in senescence processes (St Laurent et al. 2010). Such effects may be particularly salient to aging nervous systems, which are composed of mostly post-mitotic cells. Taken in this light, TEs, not protein-coding genes, may act as the quintessential antagonistic pleiotropic elements.

III. The Two-Hit Hypothesis of Neurodegeneration:

In his definitive essay on natural selection and aging, George Williams states four preconditions for his theory of antagonistic pleiotropy: "(1) A soma that is essential to reproductive success but no part of which is passed on in either sexual or asexual reproduction. (2) Natural selection of alternative alleles in a population. (3) Pleiotropic genes of a special sort. It is necessary to postulate genes that have opposite effects on fitness at different ages, or, more accurately, in different *somatic environments*. (4) Decreasing probability of reproduction with increasing age" (Williams 1957). All four of these preconditions directly describe TE activity with regard to nervous system function in aging human populations. (1) Properly functioning neural cells and nervous systems collectively are on the whole a pre-requisite for reproductive success in modern human populations, but no part of the nervous system is passed on by either sexual or asexual reproduction. (2) As described in the Introduction, we know that natural selection does in fact act on allelic variations of TEs; indeed this phenomenon is so robust as to have resulted in massive innovations in basic genome structure and function. Given the importance of neuronal function to reproductive success in human populations, the effects of TE mobilization on brain function should be subject to selective pressures as well. (3) TEs' activity with regard to the nervous system fits William's definition of a 'special sort' of pleiotropy amazingly well. As described in the Introduction, cell type-specific, regulated mobilization of specific classes of TEs to genomic regions proximal to nervous system genes is observed in the somatic environment of neurodevelopment and neuroproliferation, leading many to postulate a regulatory role for the resulting somatic mosaicism that is likely to endow the nervous system with some positive benefit or function. However in the somatic environment of the aging brain (see below) more general expression of many types of TEs is observed, likely due to age-related deterioration of TE suppression mechanisms. This type of TE activity potentially contributes to age-related neuronal decline in wild type animals and has been shown to causally contribute to toxicity in

neurodegenerative disease. (4) Human populations are, of course, subject to decreasing probability of reproduction with increasing age.

A recent paper from the Dubnau lab has demonstrated that TEs do behave according to Williams' negative pleiotropy model in the nervous system of normally aging, wild type *Drosophila* (Li et al. 2013). This work showed that transcript levels of TEs of different classes are elevated in head tissue of aged flies, and that in the case of the *gypsy* ERV, both age-dependent accumulation of Env protein and *de novo* genomic re-integration events are observed in the brains of aged animals. This effect appears to be mediated by the siRNA system, as loss of function of Ago2 (Figure 3) accelerates TE expression in the brain and shortens lifespan. These results are consistent with observations from replicative senescence, chronological aging, and *ex vivo* aging studies (Driver and McKechnie 1992; Wood et al. 2010; Maxwell et al. 2011; Wang et al. 2011; De Cecco et al. 2013b). Furthermore, aging effects on TE expression in the *Drosophila* nervous system appears to have functional consequences *in vivo*, as loss of function of Ago2 was demonstrated to exacerbate age-dependent memory impairment as measured by a sensitive learning assay (Li et al. 2013). This result indicates that the loss of suppression of TEs induced by loss of function of Ago2 may contribute to age-dependent neuronal decline.

The idea that unchecked TE activity can negatively impact nervous system function is underscored by the discovery that a reduction in MeCP2, a protein involved in global DNA methylation mutations in which are responsible for Rett's disease, can induce high levels of both L1 element transcription and retrotransposition (Muotri et al. 2010). Rett's syndrome is a post-natal neurological disorder that occurs almost exclusively in females, is characterized by repetitive arm movements, and is frequently found to be co-morbid with epileptic seizures. The disease negatively impacts diverse neurological functions such as the ability to speak, walk, and socially interact (Dolce et al. 2013). The global symptomology, as well as early onset and progressive decline of Rett's disease appear to be consistent with the immediate negative impacts of global decreases in repressive DNA methylation due to abnormal levels of MeCP2 (Lyst and Bird 2015). Accordingly, higher levels

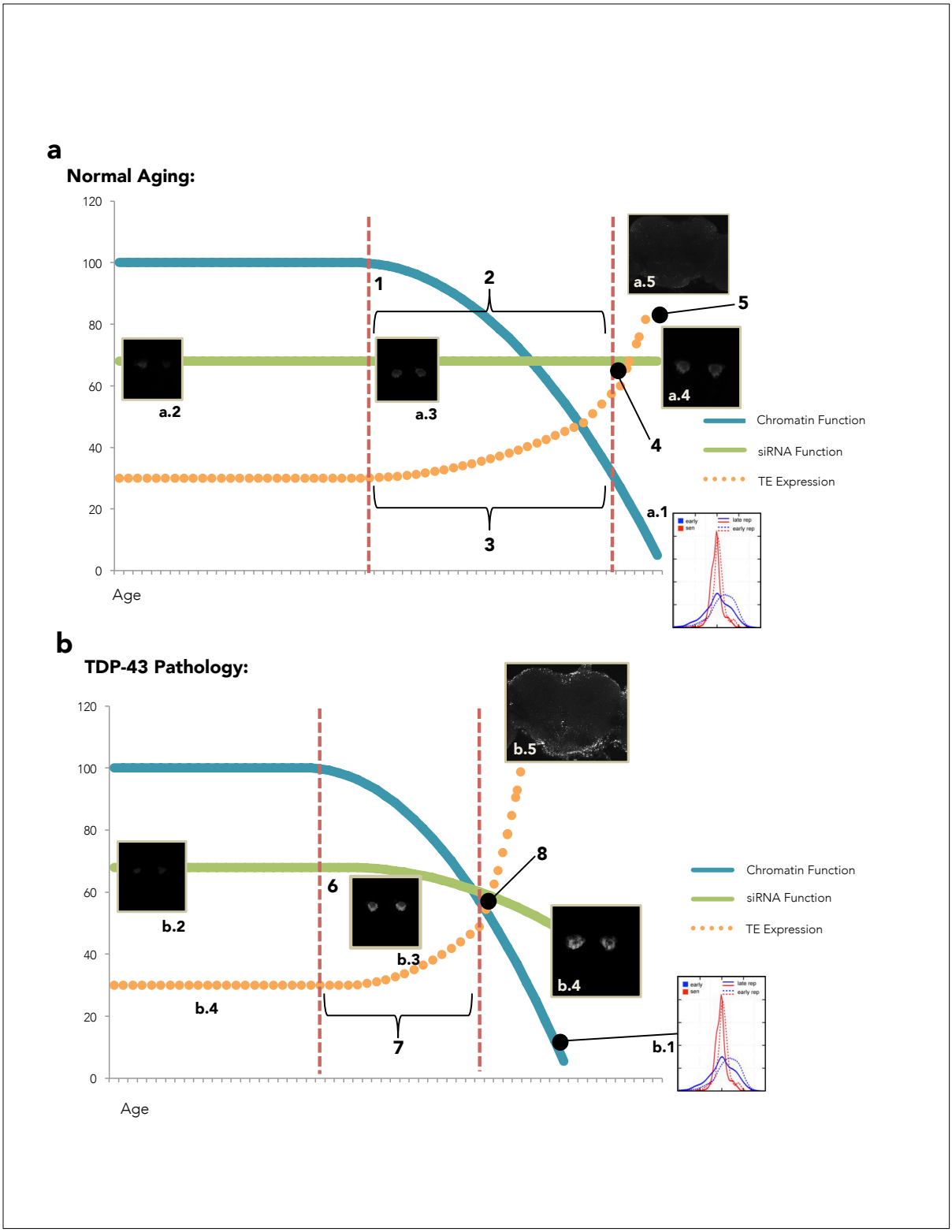
of L1 transcripts and *de novo* transposition events are observed in both brain tissue taken from, and NPCs derived from, patients diagnosed with Rett's disease who carry documented mutations in MeCP2 in comparison to age-matched controls. This effect was not observed in fibroblasts or heart tissue of the same patients, demonstrating that TEs can be affected in a tissue specific manner in response to disease-causing mutations and that the nervous system may be particularly susceptible (Muotri et al. 2010).

The fact that neurodegenerative diseases display an onset of symptoms at or soon after middle age, and therefore after the peak of maximum reproductive potential, suggests that similar to the majority of cancer cases they are on the whole a result of hapless collapse of function during senescent deterioration. The genetic and symptomological diversity of this set of diseases, and the persistent fact that age itself remains the greatest risk factor, strongly indicate that antagonistic pleiotropy at least partially accounts for neurodegenerative etiology. Further emphasizing this point, genetic predisposition accounts for only a minor percentage of most neurodegenerative diseases while the vast majority fall under the 'sporadic' category, with no known disruption in protein-coding genes contributing to pathogenesis. That we observe this suite of diseases separate from neurodevelopmental diseases speaks to a singular susceptibility of the nervous system to the progressively dilapidated cellular milieu and unstable genetic climate intrinsic to advanced age. As noted in the previous section, a body of evidence assembled from studies of normal aging in a variety of systems argues that TEs, in fact, may act as the consummate antagonistic pleiotropic factor. TE involvement has been reported in a diverse set of neurodegenerative diseases (Lathe and Harris 2009; Muotri et al. 2010; Greenwood et al. 2011; Kaneko et al. 2011; Li et al. 2012; Tan et al. 2012; Li et al. 2015) and has beenn shown to causally contribute to toxicity in a subset of neurodegenerative diseases (Kaneko et al. 2011; Tan et al. 2012), suggesting that it may represent a common underlying factor. As the vast majority of studies have either actively or passively ignored TEs by design, it is likely that the existing literature only scratches the surface.

Recent work from the Dubnau lab, and my thesis work presented herein, use complimentary approaches to directly address the role of TEs in TDP-43 protein pathologies. From these two studies we can glean some understanding of how TDP-43 pathology, which precipitates rapid and incurable neurodegenerative decline in mostly otherwise genetically normal individuals around middle age, fits into the context of normal brain aging. Bioinformatically profiling hTDP-43's TE binding partners and determining how they are effected in both rodent genetic models of TDP-43 protein pathology and human patient tissue (Li et al. 2012) allows us to determine how prevalent and robust effects of TDP-43 pathology are with respect to TEs. Mechanistic studies in *Drosophila* provide a deeper understanding of how and why such effects result in rapid deterioration and, ultimately, cell death. Combining these analyses allows us to build a framework with which to interpret both normal neurological decline in wild type aging (Figure 27A) and gain new insights into the pathological prognosis of TDP-43 protein pathologies (Figure 27B), with broader implications for many other neurodegenerative diseases as well.

As the individual ages, general chromatin structure begins to deteriorate, as observed in both replicative senescence and chronological and organismal aging models (Figure 27A: 1). However, there is a secondary PTGS system at play in somatic tissue such as the brain that is capable of suppressing TEs. For a very long time in a healthy individual, the siRNA system (Figure 27A: 2) is able to compensate for loss of heterochromatic silencing with regard to suppressing TEs (Figure 27A: 3). We know this from two pieces of data: first, in CNS neurons of control flies the fidelity of a functional genetic reporter of siRNA-mediated silencing remains robust out to old age *in vivo*, and second, levels of DCR-2 and Ago2 also remain high – if anything they display a slight but statistically significant increase in head tissue of old flies in comparison to their younger counterparts (Figure 13D.1 – 13D.3). Indeed, the upward inflection I observe in DCR-2 and Ago2 levels in the brains of old flies may be indicative of this compensatory mechanism. However, as heterochromatin continues to deteriorate in the healthy individual, persistent, simmering low-level TE activity eventually reaches

FIGURE 27: The two-hit hypothesis of neurodegeneration.
(A) Normal neurological decline in wild type aging. (**A.1**) Chromatin architecture degrades with age [*taken from*: Figure 26A.3]. (**A.2-1.4**) siRNA silencing remains robust out to old age [*taken from*: Figure 14B, *OK107* > GFP + GFP(IR) +tdTomato]. (**A.5**) TEs remain suppressed out to aged time points [*taken from*: Figure 12C, *ELAV* > hTDP-43, 19-25 Days].
(**B**) TE activation in TDP-43 pathology. (**B.1**) Chromatin architecture degrades with age [*taken from*: Figure 26A.3]. (**B.2-B.4**) siRNA silencing is eroded with hTDP-43 expression [*taken from*: Figure 14B, *OK107* > GFP + GFP(IR) + hTDP-43]. (**A.5**) TEs are precociously and aggressively de-repressed in response to simultaneous loss of heterochromatin and siRNA silencing [*taken from*: Figure 12C, *Repo* > hTDP-43, 19-25 Days].
Numbers in (**A**) and (**B**) correlate to references in the text.


a
Normal Aging:
120
100
80
60
40
20
0
Age
1
2
3
4
5
a.1
a.2
a.3
a.4
a.5
Chromatin Function
siRNA Function
TE Expression
b
TDP-43 Pathology:
120
100
80
60
40
20
0
Age
6
7
8
b.1
b.2
b.3
b.4
b.5
Chromatin Function
siRNA Function
TE Expression

a boiling point at which it overwhelms the ability of the siRNA system to keep it suppressed (Figure 27A: 4). This results in increased expression of TEs with age and, eventually, *de novo* transposition events. We observe this effect as detectable new inserts, as described in CNS neurons of aged wild type flies by (Li et al. 2013), or the notable increases in genomic TE copy number and gross chromosomal abnormalities frequently reported with extreme old age in cells or organisms that are otherwise completely genetically normal (Driver and McKechnie 1992; Li et al. 2008; St Laurent et al. 2010; Maxwell et al. 2011; De Cecco et al. 2013b; Wood and Helfand 2013). The normal loss of control of general TE activity once the siRNA system is finally overwhelmed (Figure 27A: 5) could easily contribute to the multifaceted negative heath effects we are accustomed to expect to accompany extreme old age (Tucker et al. 1999; Martin 2007).

In normal brain tissue of both rodents and humans, TDP-43 protein binds promiscuously to RNA transcripts derived from many different classes of TEs (Li et al. 2012). Both overexpressing hTDP-43, which has been documented to replicate cellular protein pathology similar to that observed in human post-mortem patient tissue, and knocking down the mouse ortholog of TDP-43 in rodent brain tissue result in a massive and general increase in transcription of a wide variety of TEs (Li et al. 2012). We are therefore able to deduce that TDP-43 is involved in TE suppression in normal adult mammalian brain function. Moreover, it appears that TDP-43 pathology is likely to have its effects in a manner similar to a toxic dominant negative, with overexpression or nucleation of protein pathology in human patients resulting in the sequestration of functional TDP-43 in cytoplasmic inclusions (Vanden Broeck et al. 2014). In cortical tissue from human FTLD patients, a disease that frequently presents with TDP-43 protein pathology, TDP-43 protein specifically loses its interaction with TE-derived RNA transcripts, while its interaction with the majority of its other RNA binding partners is left largely intact. Intriguingly, there is a general concordance between the TE families that lose their interaction with TDP-43 in cortical tissue of human FTLD patients, and those whose transcription increases in brain tissue of rodent TDP-43 pathology models (Li et al. 2012). While these observations suggest that TDP-43 pathology may involve loss of suppression of TEs, it would take

the genetic tools available in a more tractable model organism to understand why this would be the case and whether this effect has an impact on cellular toxicity and physiological health.

The absolute levels of the two major effector molecules of the siRNA system in *Drosophila*, DCR-2 and Ago2 (Figure 3), do not appear to be affected by hTDP-43 expression in either neurons or glia. However, implementation of a functional genetic reporter system demonstrates that hTDP-43 expression causes erosion of the ability of the siRNA system to silence a transgenic reporter *in vivo* in both CNS neurons and glial cells that express hTDP-43. This suggests that hTDP-43 functionally impairs the siRNA system, and the robust and early effect observed here in glial cells is consistent with the early and aggressive expression of the *gypsy* RTE observed in head tissue of flies expressing hTDP-43 in glia. In neuronal structures, loss of siRNA silencing appears more gradual over age, with a later age of onset. While the precise factor that induces nucleation of TDP-43 pathology in otherwise genetically normal individuals is still not fully understood, it does happen. And when it does, these observations suggest that the effectiveness of siRNA silencing begins to be rapidly degraded in effected cells (Figure 27B: 6). This means that the siRNA system swiftly loses its ability to compensate for on-going loss of heterochromatic silencing of TEs, resulting in a precipitously brief window (Figure 27B: 7) between the onset of protein pathology and the inflection point at which TE activity overwhelms the system (Figure 27B: 8). This feature of the model is in fact reflected by the alarmingly rapid prognosis for an individual diagnosed with ALS, as most patients are expected to succumb to the disease within 3-5 years of the onset of symptoms (Ravits and La Spada 2009). Moreover, there is compelling feedback between the siRNA system and heterochromatin (as discussed in the Introduction) that could conceivably accelerate this process (Castel and Martienssen 2013), and the kinetics of decline may be further exacerbated by the exponential function of activating multi-copy, self-replicating mobile elements. Taken in this light, the onset of symptoms in TDP-43 protein pathologies reflects two “hits” with regard to suppression of TEs: the first being normal age-related decline of heterochromatic suppression and the second being the rapid loss of siRNA silencing induced by the onset of TDP-43 protein pathology. The ensuing swift decline in the

health of an individual affected with TDP-43 protein pathology relates to the concomitant deterioration of both of these mechanisms and the self-multiplying functionality inherent to the resulting activation of TEs.

The direct predictions of this model are that TDP-43 pathology will: a) result in loss of suppression of TEs, and that this loss of control of TEs will b) negatively influence measures of physiological health and c) drive cellular toxicity. As described above, manipulating TDP-43 in rodent cortical tissue results in increased expression of many different TEs (Li et al. 2012). In *Drosophila*, overexpressing hTDP-43 in glia results in dramatic, early induction of the *gypsy* ERV at both the RNA and protein level. This effect turns on post-developmentally and stochastically in the brains of flies expressing hTDP-43 in glia and both lifespan and locomotor behavior are drastically reduced in these flies. Knocking down *gypsy* suppresses this severe lifespan deficit, strongly implicating *gypsy* RTE activity as causally contributing to the premature death induced by glial expression of hTDP-43 in *Drosophila*. Importantly, no rescue of the less severe lifespan deficit induced by neuronal expression of hTDP-43 is achieved by *gypsy* knock-down, which is consistent with the observation that neuronal expression of hTDP-43 does not elevate *gypsy* expression above wild type background levels at any given time point. These effects are in accordance with recent reports that knockout of the *C. elegans* TDP-43 ortholog results in broad accumulation of general TE-derived RNA transcripts (Saldi et al. 2014) and that HERV-K, a human ERV of the *gypsy* family, is activated in a certain subset of patients with ALS (Douville et al. 2011; Li et al. 2015). Indeed, overexpression of just the ENV protein of HERV-K is capable of inducing progressive motor dysfunction in transgenic mice (Li et al. 2015). Therefore the first two direct predictions of the Two-Hit Hypothesis model appear to be fulfilled, as TDP-43 pathology does indeed result in the loss of suppression of TEs in genetic models and subsequent TE activity contributes to deterioration of the physical health of the animal (Figure 28C). The final prediction of the model, regarding whether TDP-43-dependent TE activity contributes to cellular toxicity, is perhaps the most important question

to answer with respect to palliative care and treatment of human patients affected with TDP-43 pathologies.

The brains of flies expressing TDP-43 in glia display rampant apoptosis, both with TUNEL staining and at the level of TEM. This is in keeping with the current consensus that cells experiencing TDP-43 pathology in patient tissue die largely by apoptosis (Vanden Broeck et al. 2014). As described in the previous section, TE activity is well documented to induce prolific DNA damage. Importantly, I have demonstrated that knocking down *loki*, the *Drosophila* ortholog of Chk-2, completely alleviates the rampant apoptosis observed in the CNS of flies expressing TDP-43 in glia. Chk-2 is a DNA-damage response factor that signals cells to undergo apoptosis if DNA damage levels accumulate above a given threshold (Brodsky et al. 2004; Norbury and Zhivotovsky 2004), therefore loss of function of Chk-2 does not stop DNA damage from happening, it simply removes the ability of the cell to detect it and respond by undergoing programmed cell death (Brodsky et al. 2004). Therefore, my results show that the apoptosis induced by glial expression of TDP-43 is in large part induced by Chk-2 signaling following DNA damage (Figure 28A - 28B and 28D – 28F). TE activity does appear to contribute at least in part to the DNA-damage mediated apoptosis observed in response to TDP-43 expression in glia, as knocking down *gypsy* partially alleviates the rampant apoptosis observed in the CNS of these animals. This set of observations is consistent with reports that transgenic expression of the HERV-K Env protein in mice results in loss of volume in the motor cortex and DNA damage (Li et al. 2015) (Figure 28C - 28D). DNA damage-mediated apoptosis appears to be relevant to the general physiological health of the animal as well, as *loki* knockdown completely rescues the lifespan deficits observed with both glial and neuronal hTDP-43 expression. The fact that *loki* knock-down more fully suppresses apoptosis observed in the CNS of flies expressing TDP-43 in glia, as well as more completely suppresses the lifespan deficit induced by both glial and neuronal expression of TDP-43, hints that other TEs beyond *gypsy* may contribute to TDP-43 induced cellular and physiological toxicity via DNA damage-mediated apoptosis. This observation is consistent with previous findings that TDP-43 promiscuously interacts with many types of TEs and

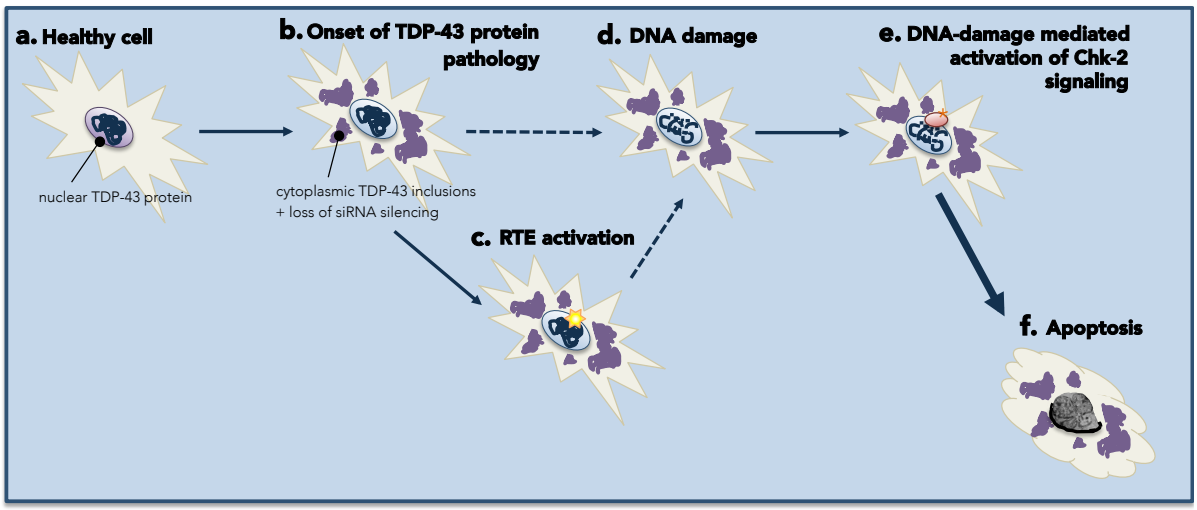


FIGURE 28: A unifying model: schematic illustration of cellular toxicity in TDP-43 protein pathology.
(**A**) In normal, healthy cells, TDP-43 is mainly localized to the nucleus. However, in individuals that experience TDP-43 pathology, around middle age or later cytoplasmic inclusions of TDP-43 nucleate and this is associated with clearance from the nucleus (**B**). This results in loss of siRNA silencing, which precipitates activation of TEs [*see also*: (Li et al. 2012) and (Saldi et al. 2014)] (**C**). TEs at least partially contribute to DNA damage [*see also*: (Li et al. 2015)]. (**D**), which in turn activates Chk-2 signaling (**E**). Chk-2 signaling ultimately results in apoptosis (**F**).

that TDP-43 pathology results in broad de-repression of TEs (Li et al. 2012; Saldi et al. 2014), as well as my findings that expressing TDP-43 in both neurons and glia negatively affects measures of physiological health and general siRNA-mediated silencing. These results indicate that TDP-43-dependent TE activity contributes at least in part to cellular toxicity, once again pointing to TEs a decisive antagonistic pleiotropic entity.

These findings stimulate several new lines of questioning. Is the siRNA system the only TE silencing system negatively affected by TDP-43 protein pathology? Does TDP-43 pathology affect biogenesis of siRNAs or localization of loaded RISC to TE transcripts? What role does TDP-43's binding to TE-derived RNA transcripts play in this regard, and is there negative feedback between pathology-induced malfunction of the siRNA system and heterochromatic stifling of TE expression? Are different TEs indeed activated in response to TDP-43 pathology in different cell types, and is there something particular to glia in this respect? Does activation of RTEs and/or ERVs in particular, with their unique capacity for cellular exodus, contribute to non-cell autonomous spread of pathology in the prognosis of TDP-43 protein pathologies? Indeed, Li et al (2015) note that the localization of HERV-K Env protein to the neuronal cell body in post-mortem tissue from ALS patients would account nicely for the stereotyped anatomical spread of ALS from brain region to adjacent brain region, a pattern that is plainly distinct from that observed with the transsynaptic spread characteristic of rabies virus (Li et al. 2015). However many other transmissible agents remain possible, such as protein or RNA species transferred via exosomes or other types of secretion into extracellular fluid or direct cytoplasmic transfer. Of course, further confirmation of this model by probing for its signatures in post-mortem human patient samples remains to be performed, and is sure to be greatly informative. This subject is incredibly fertile, and has the capacity to shed exciting new information on basic mechanisms of CNS function and its deterioration in both normal aging and neurodegenerative disease.

Direct investigation of the role of TEs in TDP-43 protein pathologies permitted the development of the models described above, which in turn begin to reveal a comprehensive picture of

the role of TEs in neurodegeneration and in the context of normal brain aging. In this picture we see that normal aging erodes the topology of activating and repressive genomic architecture, which in advanced age results in the eventual derepression and mobilization of TEs. As previously elaborated, this represents a classic case of George William's antagonistic pleiotropy by these non-coding genetic elements. While their activity may have been selected for based on certain advantages they confer to the host during neurodevelopment and adult nervous system function, this also exposes them to being particularly sensitive to the relaxation of selective pressure observed in old age, an effect that is amplified in human populations that have been predominantly liberated from natural selection. To further exacerbate the issue the brain is enriched in post-mitotic cells, with neurons being some of the perdurable cells in the body, making it duly susceptible to the accrual of genetic damage from TEs accumulated over the lifetime of the organism. The damaging action of these unregulated TEs has operative consequences for brain function in advanced age, resulting in slow, monotonic decline over the post-reproductive lifespan of the individual. In neurodegeneration collectively, we observe an age-dependent malfunction in at least one of the multi-tiered, interleaved mechanisms the brain employs to maintain its precarious TE equilibrium. The timing, localization, and cell types of the brain that are primarily affected in this diverse suite of diseases are likely dictated by the specificity of the initial malfunctioning component or process. This effectively trips a wire, interrupting the convoluted, finely tuned waltz of genetic information within the neuron and prematurely liberating the reservoir of destructive potential represented by mobile elements. After this initial activating event, the majority of these diseases exhibit astonishingly rapid prognosis. At least in the case of TDP-43 protein pathologies we know that once this balance is disrupted, TE activity ultimately contributes to cellular toxicity and functional decline.

In short, we hypothesize that the brain is singularly tuned to utilize TEs, which also renders it uniquely susceptible to their antagonistic pleiotropic activity in late age. This effect has been exposed in civilized human societies by advances in health care and technology that have artificially elevated median age and life expectancy, and may explain why we observe a suite of neurodegenerative

diseases separate from neurodevelopmental diseases. The development of the model of TDP-43 protein pathology I have presented above was enabled by specific, mechanistic analysis regarding TDP-43's effects on TEs and opens up new avenues for both palliative care and treatment of affected individuals. Given that TEs have now been implicated in such a wide variety of other neurodegenerative diseases, I believe that its underlying principles will likely prove beneficial to the understanding of the pathological prognosis of neurodegenerative disease in general. However, more specific analysis of each of these disorders will be required to determine the individual characteristics that may aid in the development of treatment. I believe that there is now sufficient evidence to consider TEs a fundamental antagonistic pleiotropic force in aging, and it seems foolish to continue to disregard them as we strive for better understanding and treatment for our aging population.

IV. Looking Forward:

Before the cracking of the genetic code there was a noticeable focus among biologists on protein-coding genes. This was an obvious starting place as these units of genetic material have physical products that are tractable to see, purify, and study, and such study has been invaluable in understanding the structure and physiology of cells and organisms as a whole. However, the genomic revolution has provided us with infinitely more information regarding what makes living things the way they are, with protein coding and related untranslated exons comprising at most 2% of the human genome. Indeed, the CNEs we have discussed easily fall under the original definition of a gene as "a unit of heredity that is transferred from parent to offspring and determines some characteristic of the offspring." And yet, most current techniques actively ignore or discard repetitive sequences that would allow us to elucidate further the role of TEs and related CNEs in both normal physiological function and disease mechanisms. The pronounced inter-individual variation in non-coding DNA renders GWAS ineffectual for this type of investigation. It is likely that such approaches could not reach enough statistical power to detect variable pathological polymorphisms in non-coding regions

even if they were embedded in conserved CNEs. Moreover, many GWAS take a candidate gene-biased approach. Microarray techniques, unless otherwise specified, implicitly ignore non-coding sequence. Whole genome sequencing approaches capture all of the information, yet most choose to throw away the fraction represented by repetitive sequence - including that contributed by TEs - by repeat masking. While it was understandable that figuring out how to map reads back to the genome may have excluded some types in order to solve the problem in the beginning of the sequencing era, reliable methods now exist with which to map repetitive reads. The un-mined wealth of discarded information already in existence in published and publically accessible sequencing datasets is astounding, and could benefit us greatly in re-analysis. There is simply no need to continue to cling to our coding sequence bias. The time has come to stop ignoring transposable elements.

Acknowledgements:

We would like to thank Ian Peikon and Anthony Zador for the 349, 350, and 351 plasmid constructs, Joshua Rosenthal for the pPicZa-λN-hDD construct and helpful conversation, Jef Boeke for the 7B3 hybridoma cell line, Carmelita Bautista at the CSHL shared resources for ascites production, and Stephen Hearn, also in CSHL shared resources, for aid in preparing and imaging TEM samples. We thank Magalie Lecourtois for the hTDP-43 fly line, Peng Jin for the *gypsy*(IR) line, and Tzumin Lee for the Repo-LexA::GAD line. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) and the Vienna Drosophila Resource Center (Dietzl et al. 2007) were also used in this study. We would also like to thank Rob Martienssen, Molly Hammell, Steve Shea, Lisa Prazak, Meng-Fu Shih, Yung-Heng Chang, and Nikolay Rozhkov for helpful discussion. This work was supported by NIH R01 grant NS067690 and NIH R01 grant NS091748 awarded to JD, and by an external research award to JD from DART neuroscience LLC. In addition, LK was supported by NIH training grant 5T32GM065094.

REFERENCES:

Alfahad T, Nath A. 2013. Retroviruses and amyotrophic lateral sclerosis. *Antiviral Res* **99**: 180-187.

Amador-Ortiz C, Lin WL, Ahmed Z, Personett D, Davies P, Duara R, Graff-Radford NR, Hutton ML, Dickson DW. 2007. TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. *Annals of neurology* **61**: 435-445.

An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, Ye P, Triplett C, Boeke JD. 2006. Active retrotransposition by a synthetic L1 element in mice. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 18662-18667.

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y et al. 2006. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochemical and biophysical research communications* **351**: 602-611.

Aravin AA, Hannon GJ. 2008. Small RNA silencing pathways in germ and stem cells. *Cold Spring Harbor symposia on quantitative biology* **73**: 283-290.

Ash PE, Zhang YJ, Roberts CM, Saldi T, Hutter H, Buratti E, Petrucelli L, Link CD. 2010. Neurotoxic effects of TDP-43 overexpression in C. elegans. *Human molecular genetics* **19**: 3206-3218.

Ayala YM, Pantano S, D'Ambrogio A, Buratti E, Brindisi A, Marchetti C, Romano M, Baralle FE. 2005. Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *Journal of molecular biology* **348**: 575-588.

Bailey PJ, Klos JM, Andersson E, Karlen M, Kallstrom M, Ponjavic J, Muhr J, Lenhard B, Sandelin A, Ericson J. 2006. A global genomic transcriptional code associated with CNS-expressed genes. *Experimental cell research* **312**: 3108-3119.

Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, Brennan PM, Rizzu P, Smith S, Fell M et al. 2011. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* **479**: 534-537.

Batut P, Dobin A, Plessy C, Carninci P, Gingeras TR. 2013. High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression. *Genome research* **23**: 169-180.

Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR, Rubin EM, Kent WJ, Haussler D. 2006. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* **441**: 87-90.

Belgnaoui SM, Gosden RG, Semmes OJ, Haoudi A. 2006. Human LINE-1 retrotransposon induces DNA damage and apoptosis in cancer cells. *Cancer cell international* **6**: 13.

Benzer S. 1967. BEHAVIORAL MUTANTS OF Drosophila ISOLATED BY COUNTERCURRENT DISTRIBUTION. *Proceedings of the National Academy of Sciences of the United States of America* **58**: 1112-1119.

Betran E, Wang W, Jin L, Long M. 2002. Evolution of the phosphoglycerate mutase processed gene in human and chimpanzee revealing the origin of a new primate gene. *Molecular biology and evolution* **19**: 654-663.

Bidder GP. 1932. Senescence. *Brit Med J* **2**: 583-585.
Blaschke AJ, Weiner JA, Chun J. 1998. Programmed cell death is a universal feature of embryonic and postnatal neuroproliferative regions throughout the central nervous system. *The Journal of comparative neurology* **396**: 39-50.
Bourc'his D, Bestor TH. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**: 96-99.
Bowen NJ, Jordan IK, Epstein JA, Wood V, Levin HL. 2003. Retrotransposons and their recognition of pol II promoters: a comprehensive survey of the transposable elements from the complete genome sequence of Schizosaccharomyces pombe. *Genome research* **13**: 1984-1997.
Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science (New York, NY)* **322**: 1387-1392.
Britten RJ, Davidson EH. 1969. Gene regulation for higher cells: a theory. *Science (New York, NY)* **165**: 349-357.
-. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *The Quarterly review of biology* **46**: 111-138.
Britten RJ, Kohne DE. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science (New York, NY)* **161**: 529-540.
Brodsky MH, Weinert BT, Tsang G, Rong YS, McGinnis NM, Golic KG, Rio DC, Rubin GM. 2004. Drosophila melanogaster MNK/Chk2 and p53 regulate multiple DNA repair and apoptotic pathways following DNA damage. *Molecular and cellular biology* **24**: 1219-1231.
Brookfield JF. 2005. The ecology of the genome - mobile DNA elements and their hosts. *Nature reviews Genetics* **6**: 128-136.
Brosius J. 2003. The contribution of RNAs and retroposition to evolutionary novelties. *Genetica* **118**: 99-116.
Buhler M, Verdel A, Moazed D. 2006. Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125**: 873-886.
Camacho-Arroyo I, Lopez-Griego L, Morales-Montor J. 2009. The role of cytokines in the regulation of neurotransmission. *Neuroimmunomodulation* **16**: 1-12.
Capri M, Salvioli S, Sevini F, Valensin S, Celani L, Monti D, Pawelec G, De Benedictis G, Gonos ES, Franceschi C. 2006. The genetics of human longevity. *Annals of the New York Academy of Sciences* **1067**: 252-263.
Capy P, Gasperi G, Biemont C, Bazin C. 2000. Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85 (Pt 2)**: 101-106.
Carroll SB. 2005. Evolution at two levels: on genes and form. *PLoS biology* **3**: e245.
Casci I, Pandey UB. 2015. A fruitful endeavor: modeling ALS in the fruit fly. *Brain research* **1607**: 47-74.
Castel SE, Martienssen RA. 2013. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature reviews Genetics* **14**: 100-112.
Castro JP, Carareto CM. 2004. Drosophila melanogaster P transposable elements: mechanisms of transposition and regulation. *Genetica* **121**: 107-118.
Chandra T, Kirschner K, Thuret JY, Pope BD, Ryba T, Newman S, Ahmed K, Samarajiwa SA, Salama R, Carroll T et al. 2012. Independence of repressive histone marks

and chromatin compaction during senescent heterochromatic layer formation. *Molecular cell* **47**: 203-214.
Chen G, Li W, Zhang QS, Regulski M, Sinha N, Barditch J, Tully T, Krainer AR, Zhang MQ, Dubnau J. 2008. Identification of synaptic targets of Drosophila pumilio. *PLoS computational biology* **4**: e1000026.
Chen H, Qian K, Chen W, Hu B, Blackbourn LWt, Du Z, Ma L, Liu H, Knobel KM, Ayala M et al. 2015. Human-derived neural progenitors functionally replace astrocytes in adult mice. *The Journal of clinical investigation* **125**: 1033-1042.
Chen Y, Pane A, Schupbach T. 2007. Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in Drosophila. *Current biology : CB* **17**: 637-642.
Chen-Plotkin AS, Lee VM, Trojanowski JQ. 2010. TAR DNA-binding protein 43 in neurodegenerative disease. *Nature reviews Neurology* **6**: 211-220.
Chow JC, Ciaudo C, Fazzari MJ, Mise N, Servant N, Glass JL, Attreed M, Avner P, Wutz A, Barillot E et al. 2010. LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. *Cell* **141**: 956-969.
Chun JJ, Schatz DG, Oettinger MA, Jaenisch R, Baltimore D. 1991. The recombination activating gene-1 (RAG-1) transcript is present in the murine central nervous system. *Cell* **64**: 189-200.
Clatworthy MR, Lyons PA, Smith KG. 2008. The nervous spleen: The nervous spleen. *Immunology and cell biology* **86**: 1-2.
Comfort A. 1954. Biological aspects of senescence. *Biol Rev* **29**: 284-329.
-. 1956. The Biology of Senescence. *Rinehart and Co, Inc*: xiii + 257 pp.
Consortium CSaA. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**: 69-87.
Consortium IHGS. 2004. Finishing the euchromatic sequence of the human genome. *Nature* **431**: 931-945.
Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. *Nature reviews Genetics* **10**: 691-703.
Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. 2009. L1 retrotransposition in human neural progenitor cells. *Nature* **460**: 1127-1131.
Crichton JH, Dunican DS, Maclennan M, Meehan RR, Adams IR. 2014. Defending the genome from the enemy within: mechanisms of retrotransposon suppression in the mouse germline. *Cellular and molecular life sciences : CMLS* **71**: 1581-1605.
Crowley C, Curtis HJ. 1963. THE DEVELOPMENT OF SOMATIC MUTATIONS IN MICE WITH AGE. *Proceedings of the National Academy of Sciences of the United States of America* **49**: 626-628.
Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R et al. 2008. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**: 798-802.
Dai J, Xie W, Brady TL, Gao J, Voytas DF. 2007. Phosphorylation regulates integration of the yeast Ty5 retrotransposon into heterochromatin. *Molecular cell* **27**: 289-299.
Dawe RK, Henikoff S. 2006. Centromeres put epigenetics in the driver's seat. *Trends in biochemical sciences* **31**: 662-669.

De Cecco M, Criscione SW, Peckham EJ, Hillenmeyer S, Hamm EA, Manivannan J, Peterson AL, Kreiling JA, Neretti N, Sedivy JM. 2013a. Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging cell* **12**: 247-256.

De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM, Kreiling JA. 2013b. Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging* **5**: 867-883.

de la Calle-Mustienes E, Feijoo CG, Manzanares M, Tena JJ, Rodriguez-Seguel E, Letizia A, Allende ML, Gomez-Skarmeta JL. 2005. A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. *Genome research* **15**: 1061-1072.

De La Fuente R, Baumann C, Fan T, Schmidtmann A, Dobrinski I, Muegge K. 2006. Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells. *Nature cell biology* **8**: 1448-1454.

Deininger PL, Batzer MA. 2002. Mammalian retroelements. *Genome research* **12**: 1455-1465.

Deininger PL, Moran JV, Batzer MA, Kazazian HH, Jr. 2003. Mobile elements and mammalian genome evolution. *Current opinion in genetics & development* **13**: 651-658.

del Valle Rodriguez A, Didiano D, Desplan C. 2012. Power tools for gene expression and clonal analysis in Drosophila. *Nature methods* **9**: 47-55.

Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. *Nature genetics* **35**: 41-48.

Diaper DC, Adachi Y, Lazarou L, Greenstein M, Simoes FA, Di Domenico A, Solomon DA, Lowe S, Alsubaie R, Cheng D et al. 2013. Drosophila TDP-43 dysfunction in glia and muscle cells cause cytological and behavioural phenotypes that characterize ALS and FTLD. *Human molecular genetics* **22**: 3883-3893.

Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S et al. 2007. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**: 151-156.

Dolce A, Ben-Zeev B, Naidu S, Kossoff EH. 2013. Rett syndrome and epilepsy: an update for child neurologists. *Pediatric neurology* **48**: 337-345.

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601-603.

Douville R, Liu J, Rothstein J, Nath A. 2011. Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. *Annals of neurology* **69**: 141-151.

Driver CJ, McKechnie SW. 1992. Transposable elements as a factor in the aging of Drosophila melanogaster. *Annals of the New York Academy of Sciences* **673**: 83-91.

Dunlap KA, Palmarini M, Varela M, Burghardt RC, Hayashi K, Farmer JL, Spencer TE. 2006. Endogenous retroviruses regulate periimplantation placental growth and differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 14390-14395.

Dustin ML, Colman DR. 2002. Neural and immunological synaptic relations. *Science (New York, NY)* **298**: 785-789.

Edelman GM. 1993. Neural Darwinism: selection and reentrant signaling in higher brain function. *Neuron* **10**: 115-125.
Eickbush TH. 1997. Telomerase and retrotransposons: which came first? *Science (New York, NY)* **277**: 911-912.
Estes PS, Boehringer A, Zwick R, Tang JE, Grigsby B, Zarnescu DC. 2011. Wild-type and A315T mutant TDP-43 exert differential neurotoxicity in a Drosophila model of ALS. *Human molecular genetics* **20**: 2308-2321.
Estes PS, Daniel SG, McCallum AP, Boehringer AV, Sukhina AS, Zwick RA, Zarnescu DC. 2013. Motor neurons and glia exhibit specific individualized responses to TDP-43 expression in a Drosophila model of amyotrophic lateral sclerosis. *Disease models & mechanisms* **6**: 721-733.
Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, Parker JJ, Atabay KD, Gilmore EC, Poduri A et al. 2012. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**: 483-496.
Evrony GD, Lee E, Mehta BK, Benjamini Y, Johnson RM, Cai X, Yang L, Haseley P, Lehmann HS, Park PJ et al. 2015. Cell lineage analysis in human brain using endogenous retroelements. *Neuron* **85**: 49-59.
Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. *Nature reviews Genetics* **9**: 397-405.
Feser J, Truong D, Das C, Carson JJ, Kieft J, Harkness T, Tyler JK. 2010. Elevated histone expression promotes life span extension. *Molecular cell* **39**: 724-735.
Freibaum BD, Chitta RK, High AA, Taylor JP. 2010. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *Journal of proteome research* **9**: 1104-1120.
Ganko EW, Bhattacharjee V, Schliekelman P, McDonald JF. 2003. Evidence for the contribution of LTR retrotransposons to C. elegans gene evolution. *Molecular biology and evolution* **20**: 1925-1931.
Gao Y, Sun Y, Frank KM, Dikkes P, Fujiwara Y, Seidl KJ, Sekiguchi JM, Rathbun GA, Swat W, Wang J et al. 1998. A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* **95**: 891-902.
Garcia-Perez JL, Doucet AJ, Bucheton A, Moran JV, Gilbert N. 2007a. Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. *Genome research* **17**: 602-611.
Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, O'Shea KS, Moran JV. 2007b. LINE-1 retrotransposition in human embryonic stem cells. *Human molecular genetics* **16**: 1569-1577.
Gasior SL, Wakeman TP, Xu B, Deininger PL. 2006. The human LINE-1 retrotransposon creates DNA double-strand breaks. *Journal of molecular biology* **357**: 1383-1393.
Gendron TF, Petrucelli L. 2011. Rodent models of TDP-43 proteinopathy: investigating the mechanisms of TDP-43-mediated neurodegeneration. *Journal of molecular neuroscience : MN* **45**: 486-499.
Gentles AJ, Wakefield MJ, Kohany O, Gu W, Batzer MA, Pollock DD, Jurka J. 2007. Evolutionary dynamics of transposable elements in the short-tailed opossum Monodelphis domestica. *Genome research* **17**: 992-1004.
Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: an expanding universe. *Nature reviews Genetics* **10**: 94-108.

Gibbs RA Weinstock GM Metzker ML Muzny DM Sodergren EJ Scherer S Scott G Steffen D Worley KC Burch PE et al. 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**: 493-521.
Girard A, Sachidanandam R, Hannon GJ, Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**: 199-202.
Giresi PG, Lieb JD. 2009. Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). *Methods (San Diego, Calif)* **48**: 233-239.
Gonzales BM. 1923. Experimental studies on the duration of life. *Amer Nat* **57**: 289-325.
Gould SJaV, E. S. 1982. Exaptation - a missing term in the science of form. *Paleobiology* **8**: 4-15.
Greenwood AD, Vincendeau M, Schmadicke AC, Montag J, Seifarth W, Motzkus D. 2011. Bovine spongiform encephalopathy infection alters endogenous retrovirus expression in distinct brain regions of cynomolgus macaques (Macaca fascicularis). *Molecular neurodegeneration* **6**: 44.
Habibi L, Ebtekar M, Jameie SB. 2009. Immune and nervous systems share molecular and functional similarities: memory storage mechanism. *Scandinavian journal of immunology* **69**: 291-301.
Haidet-Phillips AM, Hester ME, Miranda CJ, Meyer K, Braun L, Frakes A, Song S, Likhite S, Murtha MJ, Foust KD et al. 2011. Astrocytes from familial and sporadic ALS patients are toxic to motor neurons. *Nature biotechnology* **29**: 824-828.
Haldane JBS. 1941. New Paths in Genetics. *Harper* **New York & London**: 206 pp.
Hanson KA, Kim SH, Wassarman DA, Tibbetts RS. 2010. Ubiquilin modifies TDP-43 toxicity in a Drosophila model of amyotrophic lateral sclerosis (ALS). *J Biol Chem* **285**: 11068-11072.
Haynes KA, Caudy AA, Collins L, Elgin SC. 2006. Element 1360 and RNAi components contribute to HP1-dependent silencing of a pericentric reporter. *Current biology : CB* **16**: 2222-2227.
Henikoff S, Matzke MA. 1997. Exploring and explaining epigenetic effects. *Trends in genetics : TIG* **13**: 293-295.
Hickey DA. 1982. Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**: 519-531.
Hirayama T, Yagi T. 2006. The role and expression of the protocadherin-alpha clusters in the CNS. *Current opinion in neurobiology* **16**: 336-342.
Hunkapiller T, Goverman J, Koop BF, Hood L. 1989. Implications of the diversity of the immunoglobulin gene superfamily. *Cold Spring Harbor symposia on quantitative biology* **54 Pt 1**: 15-29.
Irvine DV, Zaratiegui M, Tolia NH, Goto DB, Chitwood DH, Vaughn MW, Joshua-Tor L, Martienssen RA. 2006. Argonaute slicing is required for heterochromatic silencing and spreading. *Science (New York, NY)* **313**: 1134-1137.
Ivanov A, Pawlikowski J, Manoharan I, van Tuyn J, Nelson DM, Rai TS, Shah PP, Hewitt G, Korolchuk VI, Passos JF et al. 2013. Lysosome-mediated processing of chromatin in senescence. *The Journal of cell biology* **202**: 129-143.
Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**: 501-510.

Janssens J, Van Broeckhoven C. 2013. Pathological mechanisms underlying TDP-43 driven neurodegeneration in FTLD-ALS spectrum disorders. *Human molecular genetics* **22**: R77-87.
Jensen S, Gassama MP, Heidmann T. 1999. Taming of transposable elements by homology-dependent gene silencing. *Nature genetics* **21**: 209-212.
Jordan IK, Rogozin IB, Glazko GV, Koonin EV. 2003. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends in genetics : TIG* **19**: 68-72.
Kabashi E, Lin L, Tradewell ML, Dion PA, Bercier V, Bourgouin P, Rochefort D, Bel Hadj S, Durham HD, Vande Velde C et al. 2010. Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Human molecular genetics* **19**: 671-683.
Kallmann FJaG, S. 1948. Twin studies on aging and longevity. *J Hered* **39**: 349-357.
Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, Kleinman ME, Ponicsan SL, Hauswirth WW, Chiodo VA et al. 2011. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature* **471**: 325-330.
Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & development* **19**: 489-501.
Kapitonov VV, Jurka J. 2005. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS biology* **3**: e181.
Kavi HH, Fernandez HR, Xie W, Birchler JA. 2005. RNA silencing in Drosophila. *FEBS letters* **579**: 5940-5949.
Kawahara Y, Mieda-Sato A. 2012. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 3347-3352.
Kazazian HH, Jr. 2011. Mobile DNA transposition in somatic cells. *BMC biology* **9**: 62.
Kazazian HH, Jr., Goodier JL. 2002. LINE drive. retrotransposition and genome instability. *Cell* **110**: 277-280.
Kennedy BK, Austriaco NR, Jr., Zhang J, Guarente L. 1995. Mutation in the silencing gene SIR4 can delay aging in S. cerevisiae. *Cell* **80**: 485-496.
Kennedy BK, Gotta M, Sinclair DA, Mills K, McNabb DS, Murthy M, Pak SM, Laroche T, Gasser SM, Guarente L. 1997. Redistribution of silencing proteins from telomeres to the nucleolus is associated with extension of life span in S. cerevisiae. *Cell* **89**: 381-391.
Kerschensteiner M, Meinl E, Hohlfeld R. 2009. Neuro-immune crosstalk in CNS diseases. *Neuroscience* **158**: 1122-1132.
Kidwell MG, Lisch DR. 2001. Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution; international journal of organic evolution* **55**: 1-24.
Kikutani H, Kumanogoh A. 2003. Semaphorins in interactions between T cells and antigen-presenting cells. *Nature reviews Immunology* **3**: 159-167.
Kim S, Villeponteau B, Jazwinski SM. 1996. Effect of replicative age on transcriptional silencing near telomeres in Saccharomyces cerevisiae. *Biochemical and biophysical research communications* **219**: 370-376.
King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science (New York, NY)* **188**: 107-116.

Kioussis D, Pachnis V. 2009. Immune and nervous systems: more than just a superficial similarity? *Immunity* **31**: 705-710.
Klattenhoff C, Bratu DP, McGinnis-Schultz N, Koppetsch BS, Cook HA, Theurkauf WE. 2007. Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. *Developmental cell* **12**: 45-55.
Kogan GL, Tulin AV, Aravin AA, Abramov YA, Kalmykova AI, Maisonhaute C, Gvozdev VA. 2003. The GATE retrotransposon in Drosophila melanogaster: mobility in heterochromatin and aspects of its expression in germline tissues. *Molecular genetics and genomics : MGG* **269**: 234-242.
Kosar M, Bartkova J, Hubackova S, Hodny Z, Lukas J, Bartek J. 2011. Senescence-associated heterochromatin foci are dispensable for cellular senescence, occur in a cell type- and insult-dependent manner and follow expression of p16(ink4a). *Cell cycle (Georgetown, Tex)* **10**: 457-468.
Kuo KW, Sheu HM, Huang YS, Leung WC. 1998. Expression of transposon LINE-1 is relatively human-specific and function of the transcripts may be proliferation-essential. *Biochemical and biophysical research communications* **253**: 566-570.
Kuo PH, Doudeva LG, Wang YT, Shen CK, Yuan HS. 2009. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic acids research* **37**: 1799-1808.
Kuwabara T, Hsieh J, Muotri A, Yeo G, Warashina M, Lie DC, Moore L, Nakashima K, Asashima M, Gage FH. 2009. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. *Nature neuroscience* **12**: 1097-1105.
Lai SL, Lee T. 2006. Genetic mosaic with dual binary transcriptional systems in Drosophila. *Nature neuroscience* **9**: 703-709.
Lathe R, Harris A. 2009. Differential display detects host nucleic acid motifs altered in scrapie-infected brain. *Journal of molecular biology* **392**: 813-822.
Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW. 2004. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**: 69-81.
Levin HL, Moran JV. 2011. Dynamic interactions between transposable elements and their hosts. *Nature reviews Genetics* **12**: 615-627.
Levite M. 2008. Neurotransmitters activate T-cells and elicit crucial functions via neurotransmitter receptors. *Current opinion in pharmacology* **8**: 460-471.
Li H, Mitchell JR, Hasty P. 2008. DNA double-strand breaks: a potential causative factor for mammalian aging? *Mechanisms of ageing and development* **129**: 416-424.
Li H, Vogel H, Holcomb VB, Gu Y, Hasty P. 2007. Deletion of Ku70, Ku80, or both causes early aging without substantially increased cancer. *Molecular and cellular biology* **27**: 8205-8214.
Li W, Jin Y, Prazak L, Hammell M, Dubnau J. 2012. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PloS one* **7**: e44099.
Li W, Lee MH, Henderson L, Tyagi R, Bachani M, Steiner J, Campanac E, Hoffman DA, von Geldern G, Johnson K et al. 2015. Human endogenous retrovirus-K contributes to motor neuron disease. *Sci Transl Med* **7**: 307ra153.

Li W, Prazak L, Chatterjee N, Gruninger S, Krug L, Theodorou D, Dubnau J. 2013. Activation of transposable elements during aging and neuronal decline in Drosophila. *Nature neuroscience* **16**: 529-531.
Li Y, Ray P, Rao EJ, Shi C, Guo W, Chen X, Woodruff EA, 3rd, Fushimi K, Wu JY. 2010. A Drosophila model for TDP-43 proteinopathy. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 3169-3174.
Lindblad-Toh K Wade CM Mikkelsen TS Karlsson EK Jaffe DB Kamal M Clamp M Chang JL Kulbokas EJ, 3rd Zody MC et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**: 803-819.
Ling JP, Pletnikova O, Troncoso JC, Wong PC. 2015. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science (New York, NY)* **349**: 650-655.
Ling SC, Polymenidou M, Cleveland DW. 2013. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**: 416-438.
Lippman Z, May B, Yordan C, Singer T, Martienssen R. 2003. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS biology* **1**: E67.
Liu-Yesucevitz L, Bilgutay A, Zhang YJ, Vanderweyde T, Citro A, Mehta T, Zaarur N, McKee A, Bowser R, Sherman M et al. 2010. Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PloS one* **5**: e13250.
Lowe CB, Bejerano G, Haussler D. 2007. Thousands of human mobile element fragments undergo strong purifying selection near developmental genes. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 8005-8010.
Lu X, Sachs F, Ramsay L, Jacques PE, Goke J, Bourque G, Ng HH. 2014. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature structural & molecular biology* **21**: 423-425.
Lyon MF. 2006. Do LINEs have a role in X-chromosome inactivation? *Journal of biomedicine & biotechnology* **2006**: 59746.
Lyst MJ, Bird A. 2015. Rett syndrome: a complex disorder with simple roots. *Nature reviews Genetics* **16**: 261-275.
MacGowan DJ, Scelsa SN, Imperato TE, Liu KN, Baron P, Polsky B. 2007. A controlled study of reverse transcriptase in serum and CSF of HIV-negative patients with ALS. *Neurology* **68**: 1944-1946.
Macieira-Coelho A. 1991. Chromatin reorganization during senescence of proliferating cells. *Mutation research* **256**: 81-104.
Malone CD, Hannon GJ. 2009. Small RNAs as guardians of the genome. *Cell* **136**: 656-668.
Maness PF, Schachner M. 2007. Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration. *Nature neuroscience* **10**: 19-26.
Mangiacasale R, Pittoggi C, Sciamanna I, Careddu A, Mattei E, Lorenzini R, Travaglini L, Landriscina M, Barone C, Nervi C et al. 2003. Exposure of normal and transformed cells to nevirapine, a reverse transcriptase inhibitor, reduces cell growth and promotes differentiation. *Oncogene* **22**: 2750-2761.

Marino-Ramirez L, Lewis KC, Landsman D, Jordan IK. 2005. Transposable elements donate lineage-specific regulatory sequences to host genomes. *Cytogenetic and genome research* **110**: 333-341.
Martin GM. 2007. The genetics and epigenetics of altered proliferative homeostasis in ageing and cancer. *Mechanisms of ageing and development* **128**: 9-12.
Martin GM, Smith AC, Ketterer DJ, Ogburn CE, Disteche CM. 1985. Increased chromosomal aberrations in first metaphases of cells isolated from the kidneys of aged mice. *Israel journal of medical sciences* **21**: 296-301.
Mattick JS. 2010. RNA as the substrate for epigenome-environment interactions: RNA guidance of epigenetic processes and the expansion of RNA editing in animals underpins development, phenotypic plasticity, learning, and cognition. *BioEssays : news and reviews in molecular, cellular and developmental biology* **32**: 548-552.
Mattick JS, Mehler MF. 2008. RNA editing, DNA recoding and the evolution of human cognition. *Trends in neurosciences* **31**: 227-233.
Maxwell PH, Burhans WC, Curcio MJ. 2011. Retrotransposition is associated with genome instability during chronological aging. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 20376-20381.
McAfoose J, Baune BT. 2009. Evidence for a cytokine model of cognitive function. *Neuroscience and biobehavioral reviews* **33**: 355-366.
McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United States of America* **36**: 344-355.
-. 1978. Mechanisms that rapidly reorganize the genome. *Stadler Genet Symp* **10**: 25-47.
-. 1984. The significance of responses of the genome to challenge. *Science (New York, NY)* **226**: 792-801.
McCloy RA, Rogers S, Caldon CE, Lorca T, Castro A, Burgess A. 2014. Partial inhibition of Cdk1 in G 2 phase overrides the SAC and decouples mitotic events. *Cell cycle (Georgetown, Tex)* **13**: 1400-1412.
McCormick AL, Brown RH, Jr., Cudkowicz ME, Al-Chalabi A, Garson JA. 2008. Quantification of reverse transcriptase in ALS and elimination of a novel retroviral candidate. *Neurology* **70**: 278-283.
McDonald JF. 1998. Transposable elements, gene silencing and macroevolution. *Trends in ecology & evolution* **13**: 94-95.
Medawar PB. 1953. An unsolved problem in biology. *HK Lewis* **London**: 24 pp.
-. 1955. The definition and measurement of senescence. *Ciba Found Colloq on Agin* **1**: 4-15.
Meyer K, Ferraiuolo L, Miranda CJ, Likhite S, McElroy S, Renusch S, Ditsworth D, Lagier-Tourenne C, Smith RA, Ravits J et al. 2014. Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 829-832.
Miguel L, Frebourg T, Campion D, Lecourtois M. 2011. Both cytoplasmic and nuclear accumulations of the protein are neurotoxic in Drosophila models of TDP-43 proteinopathies. *Neurobiology of disease* **41**: 398-406.
Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, Duke S, Garber M, Gentles AJ, Goodstadt L, Heger A et al. 2007. Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. *Nature* **447**: 167-177.

Miller FD, Gauthier AS. 2007. Timing is everything: making neurons versus glia in the developing cortex. *Neuron* **54**: 357-369.
Miller WJ, McDonald JF, Nouaud D, Anxolabehere D. 1999. Molecular domestication--more than a sporadic episode in evolution. *Genetica* **107**: 197-207.
Minajigi A, Froberg JE, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Payer B, Boukhali M, Haas W et al. 2015. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science (New York, NY)* **349**.
Miskey C, Izsvak Z, Plasterk RH, Ivics Z. 2003. The Frog Prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells. *Nucleic acids research* **31**: 6873-6881.
Montiel-Gonzalez MF, Vallecillo-Viejo I, Yudowski GA, Rosenthal JJ. 2013. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 18285-18290.
Moore JK, Haber JE. 1996. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* **383**: 644-646.
Morris KV, Chan SW, Jacobsen SE, Looney DJ. 2004. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science (New York, NY)* **305**: 1289-1292.
Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **435**: 903-910.
Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH. 2010. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* **468**: 443-446.
Muotri AR, Zhao C, Marchetto MC, Gage FH. 2009. Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus* **19**: 1002-1007.
Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, Lingner J, Harley CB, Cech TR. 1997. Telomerase catalytic subunit homologs from fission yeast and human. *Science (New York, NY)* **277**: 955-959.
Narita M, Nunez S, Heard E, Narita M, Lin AW, Hearn SA, Spector DL, Hannon GJ, Lowe SW. 2003. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113**: 703-716.
Nekrutenko A, Li WH. 2001. Transposable elements are found in a large number of human protein-coding genes. *Trends in genetics : TIG* **17**: 619-621.
Neumann M. 2009. Molecular neuropathology of TDP-43 proteinopathies. *International journal of molecular sciences* **10**: 232-246.
Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. 2003. Scanning human gene deserts for long-range enhancers. *Science (New York, NY)* **302**: 413.
Norbury CJ, Zhivotovsky B. 2004. DNA damage-induced apoptosis. *Oncogene* **23**: 2797-2808.
O'Donnell KA, Boeke JD. 2007. Mighty Piwis defend the germline against genome intruders. *Cell* **129**: 37-44.
O'Sullivan RJ, Karlseder J. 2012. The great unravelling: chromatin as a modulator of the aging process. *Trends in biochemical sciences* **37**: 466-476.

O'Sullivan RJ, Kubicek S, Schreiber SL, Karlseder J. 2010. Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nature structural & molecular biology* **17**: 1218-1225.
Oberdoerffer P, Sinclair DA. 2007. The role of nuclear architecture in genomic instability and ageing. *Nature reviews Molecular cell biology* **8**: 692-702.
Ogawa Y, Sun BK, Lee JT. 2008. Intersection of the RNA interference and X-inactivation pathways. *Science (New York, NY)* **320**: 1336-1341.
Okamura K, Lai EC. 2008. Endogenous small interfering RNAs in animals. *Nature reviews Molecular cell biology* **9**: 673-678.
Olovnikov I, Ryazansky S, Shpiz S, Lavrov S, Abramov Y, Vaury C, Jensen S, Kalmykova A. 2013. De novo piRNA cluster formation in the Drosophila germ line triggered by transgenes containing a transcribed transposon fragment. *Nucleic acids research* **41**: 5757-5768.
Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. *Nature* **284**: 604-607.
Ostertag EM, Goodier JL, Zhang Y, Kazazian HH, Jr. 2003. SVA elements are nonautonomous retrotransposons that cause disease in humans. *American journal of human genetics* **73**: 1444-1451.
Ou SH, Wu F, Harrich D, Garcia-Martinez LF, Gaynor RB. 1995. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *Journal of virology* **69**: 3584-3596.
Pal-Bhadra M, Bhadra U, Birchler JA. 1999. Cosuppression of nonhomologous transgenes in Drosophila involves mutually related endogenous sequences. *Cell* **99**: 35-46.
-. 2002. RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila. *Molecular cell* **9**: 315-327.
Pal-Bhadra M, Leibovitch BA, Gandhi SG, Chikka MR, Bhadra U, Birchler JA, Elgin SC. 2004. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science (New York, NY)* **303**: 669-672.
Pardue ML, Rashkova S, Casacuberta E, DeBaryshe PG, George JA, Traverse KL. 2005. Two retrotransposons maintain telomeres in Drosophila. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **13**: 443-453.
Pare JM, Tahbaz N, Lopez-Orozco J, LaPointe P, Lasko P, Hobman TC. 2009. Hsp90 regulates the function of argonaute 2 and its recruitment to stress granules and P-bodies. *Molecular biology of the cell* **20**: 3273-3284.
Parsons PA. 2005. Environments and evolution: interactions between stress, resource inadequacy and energetic efficiency. *Biological reviews of the Cambridge Philosophical Society* **80**: 589-610.
-. 2007. Antagonistic pleiotropy and the stress theory of aging. *Biogerontology* **8**: 613-617.
Patterson MN, Scannapieco AE, Au PH, Dorsey S, Royer CA, Maxwell PH. 2015. Preferential retrotransposition in aging yeast mother cells is correlated with increased genome instability. *DNA repair* **34**: 18-27.

Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD et al. 2006. In vivo enhancer analysis of human conserved non-coding sequences. *Nature* **444**: 499-502.

Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S. 2013. Transposition-driven genomic heterogeneity in the Drosophila brain. *Science (New York, NY)* **340**: 91-95.

Pesiridis GS, Lee VM, Trojanowski JQ. 2009. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Human molecular genetics* **18**: R156-162.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323-337.

Peters L, Meister G. 2007. Argonaute proteins: mediators of RNA silencing. *Molecular cell* **26**: 611-623.

Polak P, Domany E. 2006. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. *BMC genomics* **7**: 133.

Polymenidou M, Cleveland DW. 2011. The seeds of neurodegeneration: prion-like spreading in ALS. *Cell* **147**: 498-508.

Popova LM, Sakharova AV. 1982. [Virus-like inclusions in the cells of the central nervous system in amyotrophic lateral sclerosis]. *Arkhiv patologii* **44**: 28-33.

Prabhakar S, Noonan JP, Paabo S, Rubin EM. 2006. Accelerated evolution of conserved noncoding sequences in humans. *Science (New York, NY)* **314**: 786.

Prudhomme S, Bonnaud B, Mallet F. 2005. Endogenous retroviruses and animal reproduction. *Cytogenetic and genome research* **110**: 353-364.

Qin H, Cressy M, Li W, Coravos JS, Izzi SA, Dubnau J. 2012. Gamma neurons mediate dopaminergic input during aversive olfactory memory formation in Drosophila. *Current biology : CB* **22**: 608-614.

Ramsey MJ, Moore DH, 2nd, Briner JF, Lee DA, Olsen L, Senft JR, Tucker JD. 1995. The effects of age and lifestyle factors on the accumulation of cytogenetic damage as measured by chromosome painting. *Mutation research* **338**: 95-106.

Rando OJ, Verstrepen KJ. 2007. Timescales of genetic and epigenetic inheritance. *Cell* **128**: 655-668.

Ravits JM, La Spada AR. 2009. ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration. *Neurology* **73**: 805-811.

Ren K, Pena de Ortiz S. 2002. Non-homologous DNA end joining in the mature rat brain. *Journal of neurochemistry* **80**: 949-959.

Ritson GP, Custer SK, Freibaum BD, Guinto JB, Geffel D, Moore J, Tang W, Winton MJ, Neumann M, Trojanowski JQ et al. 2010. TDP-43 mediates degeneration in a novel Drosophila model of disease caused by mutations in VCP/p97. *J Neurosci* **30**: 7729-7739.

Robb GB, Rana TM. 2007. RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Molecular cell* **26**: 523-537.

Romano G, Appocher C, Scorzeto M, Klima R, Baralle FE, Megighian A, Feiguin F. 2015. Glial TDP-43 regulates axon wrapping, GluRIIA clustering and fly motility by autonomous and non-autonomous mechanisms. *Human molecular genetics*.
Rothbart SB, Strahl BD. 2014. Interpreting the language of histone and DNA modifications. *Biochimica et biophysica acta* **1839**: 627-643.
Saberi S, Stauffer JE, Schulte DJ, Ravits J. 2015. Neuropathology of Amyotrophic Lateral Sclerosis and Its Variants. *Neurol Clin* **33**: 855-876.
Sahoo PK SS, and Wong AKC. 1988. A Survey of Thresholding Techniques. *Computer Vision, Graphics, and Image Processing* **41**: 233-260.
Saito K, Siomi MC. 2010. Small RNA-mediated quiescence of transposable elements in animals. *Developmental cell* **19**: 687-697.
Saldi TK, Ash PE, Wilson G, Gonzales P, Garrido-Lecca A, Roberts CM, Dostal V, Gendron TF, Stein LD, Blumenthal T et al. 2014. TDP-1, the Caenorhabditis elegans ortholog of TDP-43, limits the accumulation of double-stranded RNA. *The EMBO journal* **33**: 2947-2966.
Sandelin A, Bailey P, Bruce S, Engstrom PG, Klos JM, Wasserman WW, Ericson J, Lenhard B. 2004. Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. *BMC genomics* **5**: 99.
Santangelo AM, de Souza FS, Franchini LF, Bumaschny VF, Low MJ, Rubinstein M. 2007. Ancient exaptation of a CORE-SINE retroposon into a highly conserved mammalian neuronal enhancer of the proopiomelanocortin gene. *PLoS genetics* **3**: 1813-1826.
Savva YA, Jepson JE, Chang YJ, Whitaker R, Jones BC, St Laurent G, Tackett MR, Kapranov P, Jiang N, Du G et al. 2013. RNA editing regulates transposon-mediated heterochromatic gene silencing. *Nature communications* **4**: 2745.
Schaefer RE, Kidwell MG, Fausto-Sterling A. 1979. Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: Morphological and Cytological Studies of Ovarian Dysgenesis. *Genetics* **92**: 1141-1152.
Schmid CW. 2003. Alu: a parasite's parasite? *Nature genetics* **35**: 15-16.
Schotta G, Ebert A, Reuter G. 2003. SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. *Genetica* **117**: 149-158.
Schueler MG, Sullivan BA. 2006. Structural and functional dynamics of human centromeric chromatin. *Annual review of genomics and human genetics* **7**: 301-313.
Sedivy JM, Banumathy G, Adams PD. 2008. Aging by epigenetics--a consequence of chromatin damage? *Experimental cell research* **314**: 1909-1917.
Selmeczy Z, Vizi ES, Csoka B, Pacher P, Hasko G. 2008. Role of nonsynaptic communication in regulating the immune response. *Neurochemistry international* **52**: 52-59.
Seluanov A, Mittelman D, Pereira-Smith OM, Wilson JH, Gorbunova V. 2004. DNA end joining becomes less efficient and more error-prone during cellular senescence. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 7624-7629.
Sen R, Oltz E. 2006. Genetic and epigenetic regulation of IgH gene assembly. *Current opinion in immunology* **18**: 237-242.

Shackelford DA. 2006. DNA end joining activity is reduced in Alzheimer's disease. *Neurobiology of aging* **27**: 596-605.
Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, Aggarwala V, Cruickshanks HA, Rai TS, McBryan T et al. 2013. Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes & development* **27**: 1787-1799.
Shaw AS, Allen PM. 2001. Kissing cousins: immunological and neurological synapses. *Nature immunology* **2**: 575-576.
Sijen T, Plasterk RH. 2003. Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature* **426**: 310-314.
Silva JC, Shabalina SA, Harris DG, Spouge JL, Kondrashovi AS. 2003. Conserved fragments of transposable elements in intergenic regions: evidence for widespread recruitment of MIR- and L2-derived sequences within the mouse and human genomes. *Genetical research* **82**: 1-18.
Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. 2010. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends in neurosciences* **33**: 345-354.
Siomi MC, Sato K, Pezic D, Aravin AA. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. *Nature reviews Molecular cell biology* **12**: 246-258.
Slotkin RK, Freeling M, Lisch D. 2005. Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nature genetics* **37**: 641-644.
Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nature reviews Genetics* **8**: 272-285.
Smeal T, Claus J, Kennedy B, Cole F, Guarente L. 1996. Loss of transcriptional silencing causes sterility in old mother cells of S. cerevisiae. *Cell* **84**: 633-642.
Song SU, Gerasimova T, Kurkulos M, Boeke JD, Corces VG. 1994. An env-like protein encoded by a Drosophila retroelement: evidence that gypsy is an infectious retrovirus. *Genes & development* **8**: 2046-2057.
St Laurent G, 3rd, Hammell N, McCaffrey TA. 2010. A LINE-1 component to human aging: do LINE elements exact a longevity cost for evolutionary advantage? *Mechanisms of ageing and development* **131**: 299-305.
Steele AJ, Al-Chalabi A, Ferrante K, Cudkowicz ME, Brown RH, Jr., Garson JA. 2005. Detection of serum reverse transcriptase activity in patients with ALS and unaffected blood relatives. *Neurology* **64**: 454-458.
Sun FL, Haynes K, Simpson CL, Lee SD, Collins L, Wuller J, Eissenberg JC, Elgin SC. 2004. cis-Acting determinants of heterochromatin formation on Drosophila melanogaster chromosome four. *Molecular and cellular biology* **24**: 8210-8220.
Svoboda P, Stein P, Anger M, Bernstein E, Hannon GJ, Schultz RM. 2004. RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. *Developmental biology* **269**: 276-285.
Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD. 2002. Human l1 retrotransposition is associated with genetic instability in vivo. *Cell* **110**: 327-338.
Taira K. 2006. Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. *Nature* **441**: 1176.

Tan H, Qurashi A, Poidevin M, Nelson DL, Li H, Jin P. 2012. Retrotransposon activation contributes to fragile X premutation rCGG-mediated neurodegeneration. *Human molecular genetics* **21**: 57-65.
Teng SC, Kim B, Gabriel A. 1996. Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. *Nature* **383**: 641-644.
Thornburg BG, Gotea V, Makalowski W. 2006. Transposable elements as a significant source of transcription regulating signals. *Gene* **365**: 104-110.
Tong J, Huang C, Bi F, Wu Q, Huang B, Liu X, Li F, Zhou H, Xia XG. 2013. Expression of ALS-linked TDP-43 mutant in astrocytes causes non-cell-autonomous motor neuron death in rats. *The EMBO journal* **32**: 1917-1926.
Tsurumi A, Li WX. 2012. Global heterochromatin loss: a unifying theory of aging? *Epigenetics* **7**: 680-688.
Tucker JD, Spruill MD, Ramsey MJ, Director AD, Nath J. 1999. Frequency of spontaneous chromosome aberrations in mice: effects of age. *Mutation research* **425**: 135-141.
Tully T, Preat T, Boynton SC, Del Vecchio M. 1994. Genetic dissection of consolidated memory in Drosophila. *Cell* **79**: 35-47.
Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sanchez-Luque FJ, Bodea GO, Ewing AD, Salvador-Palomeque C, van der Knaap MS, Brennan PM et al. 2015. Ubiquitous L1 mosaicism in hippocampal neurons. *Cell* **161**: 228-239.
Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science (New York, NY)* **313**: 320-324.
van de Lagemaat LN, Landry JR, Mager DL, Medstrand P. 2003. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends in genetics : TIG* **19**: 530-536.
Vanden Broeck L, Callaerts P, Dermaut B. 2014. TDP-43-mediated neurodegeneration: towards a loss-of-function hypothesis? *Trends Mol Med* **20**: 66-71.
Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science (New York, NY)* **303**: 672-676.
Villeponteau B. 1997. The heterochromatin loss model of aging. *Experimental gerontology* **32**: 383-394.
Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science (New York, NY)* **297**: 1833-1837.
Vyjayanti VN, Rao KS. 2006. DNA double strand break repair in brain: reduced NHEJ activity in aging rat neurons. *Neuroscience letters* **393**: 18-22.
Wallace NA, Belancio VP, Deininger PL. 2008. L1 mobile element expression causes multiple types of toxicity. *Gene* **419**: 75-81.
Wang IF, Wu LS, Shen CK. 2008. TDP-43: an emerging new player in neurodegenerative diseases. *Trends Mol Med* **14**: 479-485.
Wang J, Geesman GJ, Hostikka SL, Atallah M, Blackwell B, Lee E, Cook PJ, Pasaniuc B, Shariat G, Halperin E et al. 2011. Inhibition of activated pericentromeric SINE/Alu repeat transcription in senescent human adult stem cells reinstates self-renewal. *Cell cycle (Georgetown, Tex)* **10**: 3016-3030.

Wang T, Zeng J, Lowe CB, Sellers RG, Salama SR, Yang M, Burgess SM, Brachmann RK, Haussler D. 2007. Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 18613-18618.

Waterston RH Lindblad-Toh K Birney E Rogers J Abril JF Agarwal P Agarwala R Ainscough R Alexandersson M An P et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-562.

Weismann A. 1891. On Heredity. *Claredon Press* **Oxford**.

Williams GC. 1957. Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution; international journal of organic evolution* **11**: 398-411.

Wilson DM, 3rd, Bohr VA, McKinnon PJ. 2008. DNA damage, DNA repair, ageing and age-related disease. *Mechanisms of ageing and development* **129**: 349-352.

Wolffe AP, Matzke MA. 1999. Epigenetics: regulation through repression. *Science (New York, NY)* **286**: 481-486.

Wood JG, Helfand SL. 2013. Chromatin structure and transposable elements in organismal aging. *Frontiers in genetics* **4**: 274.

Wood JG, Hillenmeyer S, Lawrence C, Chang C, Hosier S, Lightfoot W, Mukherjee E, Jiang N, Schorl C, Brodsky AS et al. 2010. Chromatin remodeling in the aging genome of Drosophila. *Aging cell* **9**: 971-978.

Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SF, North P, Callaway H, Kelly K et al. 2005. Highly conserved non-coding sequences are associated with vertebrate development. *PLoS biology* **3**: e7.

Xie W, Donohue RC, Birchler JA. 2013. Quantitatively increased somatic transposition of transposable elements in Drosophila strains compromised for RNAi. *PloS one* **8**: e72163.

Yang N, Kazazian HH, Jr. 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nature structural & molecular biology* **13**: 763-771.

Yoder JA, Walsh CP, Bestor TH. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends in genetics : TIG* **13**: 335-340.

Young JI, Sedivy JM, Smith JR. 2003. Telomerase expression in normal human fibroblasts stabilizes DNA 5-methylcytosine transferase I. *J Biol Chem* **278**: 19904-19908.

Yu D, Cook MC, Shin DM, Silva DG, Marshall J, Toellner KM, Havran WL, Caroni P, Cooke MP, Morse HC et al. 2008. Axon growth and guidance genes identify T-dependent germinal centre B cells. *Immunology and cell biology* **86**: 3-14.

Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, Daganzo SM, Erzberger JP, Serebriiskii IG, Canutescu AA, Dunbrack RL et al. 2005. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Developmental cell* **8**: 19-30.

Zhou R, Hotta I, Denli AM, Hong P, Perrimon N, Hannon GJ. 2008. Comparative analysis of argonaute-dependent small RNA pathways in Drosophila. *Molecular cell* **32**: 592-599.

Zhu CT, Chang C, Reenan RA, Helfand SL. 2014. Indy gene variation in natural populations confers fitness advantage and life span extension through transposon insertion. *Aging* **6**: 58-69.