

Transposable elements in *Drosophila*

Tabitha J. McCullers and Mindy Steiniger

Department of Biology, University of Missouri, St. Louis, MO, USA

ABSTRACT

Transposable elements (TEs) are mobile genetic elements that can mobilize within host genomes. As TEs comprise more than 40% of the human genome and are linked to numerous diseases, understanding their mechanisms of mobilization and regulation is important. *Drosophila melanogaster* is an ideal model organism for the study of eukaryotic TEs as its genome contains a diverse array of active TEs. TEs universally impact host genome size via transposition and deletion events, but may also adopt unique functional roles in host organisms. There are 2 main classes of TEs: DNA transposons and retrotransposons. These classes are further divided into subgroups of TEs with unique structural and functional characteristics, demonstrating the significant variability among these elements. Despite this variability, *D. melanogaster* and other eukaryotic organisms utilize conserved mechanisms to regulate TEs. This review focuses on the transposition mechanisms and regulatory pathways of TEs, and their functional roles in *D. melanogaster*.

ARTICLE HISTORY

Received 16 February 2017
Revised 7 April 2017
Accepted 7 April 2017

KEYWORDS

helitrons; LTR retrotransposons; non-LTR retrotransposons; P elements; retrovirus; TEs; TIR transposons; transposons

Introduction

Transposable elements (TEs) exist in the genomes of organisms across all 3 domains of life. Also referred to as “jumping genes,” TEs move, or transpose, to different locations throughout the genomes in which they reside. As mobile genetic elements, TEs are both drivers of evolution and potentially harmful mutagens that may insert within gene-encoding sequences. Interestingly, the C-value paradox, or the lack of correlation between genome size and organism complexity, may be addressed by the presence of TEs, as genome size appears to correlate with TE abundance.^{1,2} One study found significant alteration in genome size among several species of *Drosophila*, which correlated with variation in the amount of repeating sequences, such as TEs.³ Differences in genome size as a result of TE content may have significant functional effects on *Drosophila* and other eukaryotic organisms. Data indicate that genome size correlates with body size, sperm length and duration of development in *Drosophila* species.⁴ This suggests that TEs may indirectly impose highly variable effects on their hosts through genome expansion or contraction, potentially influencing traits of evolutionary significance.

TEs comprise about 45% of the human genome and at least 50% of the maize genome.^{5,6} Transposition of

these elements has been linked to over 75 human diseases including hemophilia A, breast cancer, colorectal cancer, amyotrophic lateral sclerosis, and frontotemporal lobar degeneration.⁷⁻¹¹ Furthermore, TEs also potentially contribute to neurologic development as well as neurologic diseases and disorders.^{12,13} Because of their prevalence and disease-causing potential, it is important to understand how TEs transpose and how their mobilization is regulated in eukaryotic organisms. Most TEs in the human genome, however, are completely inactive, indicating the need for a model organism in which to study these elements. The *Drosophila melanogaster* genome is one of the best studied eukaryotic genomes and while only about 20% of the genome consists of TEs, at least 30% of these elements are full length and believed to be active.^{14,15} As such, *D. melanogaster* is a promising model organism for the study of eukaryotic TEs.

Since the discovery of TEs in maize by Barbara McClintock in the 1940s, it was proposed that these elements be classified into 2 major groups (Fig. 1): DNA transposons (class II elements) and retrotransposons (class I elements).^{16,17} Within these groups are numerous families of TEs, defined primarily by sequence similarity, and still many unclassified TEs. Some TEs are separated into unique subclasses due to

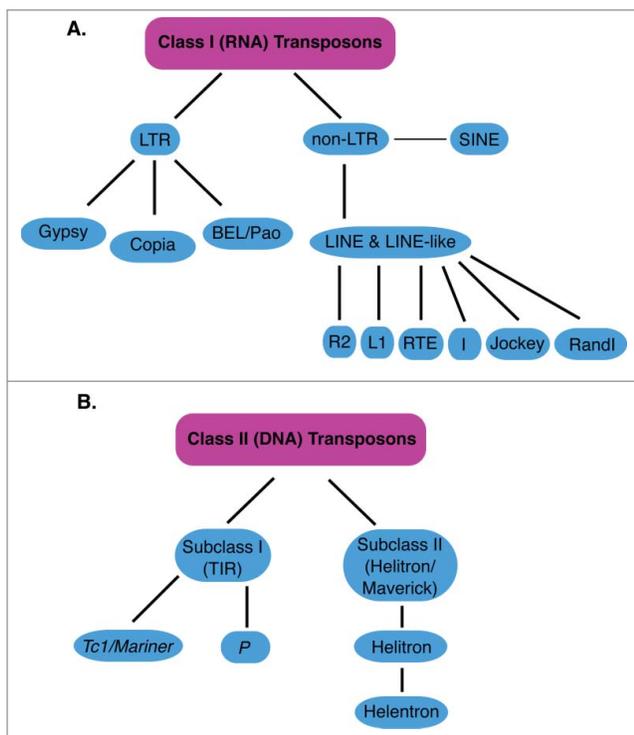


Figure 1. classes of Transposons. Shows classes, subclasses and groups of TEs described in this review. (A) classes, subclasses and groups of class I (RNA) transposons are shown in light blue. LTR is Long Terminal Repeat, SINE is Short Interspersed Nuclear Elements, LINE is Long Interspersed Nuclear Elements (B) classes, subclasses and families of class II (DNA) transposons are shown in light blue. Only the Tc1/*mariner* and P families are shown for simplicity. TIR is Terminal Inverted Repeats.

structural elements or transposition mechanisms that are uncharacteristic of other TEs. TEs may also be classified as autonomous or non-autonomous, depending on whether they transpose independently or require the machinery of autonomous TEs for mobilization. Generally, non-autonomous DNA transposons are regarded as inactive, though this is not always the case, while non-autonomous retrotransposons often utilize the machinery of autonomous retrotransposons for mobilization.

DNA transposons, or terminal inverted repeat (TIR) transposons, consist of a transposase gene flanked by TIRs, and move via a cut-and-paste mechanism. TIRs are repeating sequences found at both ends of these elements, and are inverted with respect to each other. The transposase is responsible for excising the transposon and inserting it into a new location. No active DNA transposons have been identified in humans due to lack of functional transposases, but at least 16% of the DNA transposons in *D. melanogaster* are full length and potentially active, including 1360,

hobo, *Bari1*, *pogo*, and P elements.^{6,14,18-20} Helitrons, a subclass of DNA transposons also present in *D. melanogaster* and other eukaryotic genomes, mobilize by a different mechanism than TIR transposons, using rolling-circle replication with a single stranded DNA intermediate.²¹ The regulation of DNA transposons in somatic cells is poorly understood, though some regulatory mechanisms have been identified for P elements in *D. melanogaster*. Furthermore, regulatory mechanisms have been identified in *Drosophila* germline cells to prevent mobilization of all transposable elements, as harmful transposition events in these cell lines are likely to negatively impact the viability of progeny.

Retrotransposons, or RNA transposons, are classified as either long-terminal repeat (LTR) retrotransposons or non-LTR retrotransposons, depending on the presence or absence of LTRs flanking genes required for element mobilization. The only active TEs identified in humans are non-LTR retrotransposons: Autonomous LINE-1 (long interspersed nuclear element-1) elements, which comprise 17% of the human genome, and non-autonomous *Alu* and SVA elements (Fig. 1).^{22,23} In *Drosophila*, however, at least 21% of non-LTR retrotransposons and 45% of LTR retrotransposons are full-length and potentially active, such as the LINE-like elements *TART*, *jockey* and *Juan* and the LTR retrotransposons *roo*, *copla*, *blood*, *gypsy*, and *mdg1*.^{14,24,25} Retrotransposons use a copy-and-paste mechanism by first generating an RNA intermediate that is then reverse transcribed by an element encoded reverse transcriptase (RT) into a new DNA copy that is inserted elsewhere in the genome. Retrotransposons are regulated in *Drosophila* somatic cells by heterochromatin formation that is mediated by endogenous small interfering RNAs (esiRNAs) generated from retrotransposon-derived double-stranded (ds)RNA precursors by Dicer2 (Dcr2).²⁶ Due to the relative inactivity of TEs in the human genome, this regulatory pathway has not been fully elucidated in human cell lines. This review will explore the mechanisms of transposition and regulation of these TEs in *D. melanogaster*.

Germline regulation of TEs

In *D. melanogaster*, movement of all TEs is tightly regulated in germline cells, where uncontrolled transposition events may impose significant genomic

defects that would be inherited by successive generations. The primary germline regulatory pathway is mediated by PIWI (P-element induced wimpy testis), Aubergine (AUB), Argonaute 3 (AGO3) and small PIWI-interacting RNAs (piRNAs) that regulate TEs via RNAi and epigenetic mechanisms, including heterochromatin formation.^{27,28} The piRNAs derived from TEs, formerly referred to as repeat-associated small interfering RNAs (rasiRNAs), are generated by processing of sense and antisense transcripts from TEs into small RNAs by PIWI and AUB.²⁹ Both the proteins and piRNAs required for this pathway are almost exclusively produced in germ cells.³⁰ The generation of piRNAs is distinct from that of esiRNAs, as this pathway is independent of Dcr2 and relies solely on PIWI proteins not involved in esiRNA biogenesis.²⁹ Most piRNAs are generated from transposon transcripts and form PIWI-piRNA complexes that function to silence transposon transcripts via RNAi.³¹ For a review of piRNA-mediated TE silencing in *Drosophila*, see ref. 31.

DNA transposons

TIR transposons

DNA transposons are often less than 5 kb in length and typically encode a single transposase gene (Fig. 2A).³² DNA transposons are divided into 2 sub-classes based on their transposition mechanisms. Sub-class I elements utilize the canonical cut-and-paste mechanism of TIR transposon transposition, and are divided into several superfamilies: *Tc1/mariner*, *PIF/Harbinger*, *hAT*, *Mutator*, *Merlin*, *Transib*, *P*, *piggyBac*, and *CACTA*.³³ Sub-class II DNA transposons include *Helitron* and *Maverick* elements that utilize unique transposition mechanisms (see *Helitrons* section).³³ DNA transposons are generally regarded as “extinct” in humans and other mammals as most are non-autonomous.^{13,34} *D. melanogaster*, however, has numerous active DNA transposons with full length TIRs and functional transposase genes. For example, the *Bari* elements of the *Tc1/mariner* superfamily of DNA transposons have been identified in *D. melanogaster* and many other *Drosophila* genomes.^{35,36} *D. melanogaster Bari1* elements are autonomous DNA transposons with short TIRs, usually less than 40 nucleotides in length, although non-autonomous *Bari1* elements with long TIRs have been identified in other *Drosophila* species.³⁶ Transposition of *Bari1* elements and other TEs in the

Tc1/mariner superfamily is initiated by interactions between one or more direct repeats in the TIRs and the element encoded transposase (Fig. 2B).²⁰ Dimerization of TIR bound transposases induces cleavage of the element from surrounding sequences (Fig. 2B). Like *Tc1* elements, *Bari1* elements then target TA sites and integration results in the duplication of these nucleotides at both ends (Fig. 2B).³⁵ Target site duplications are characteristic of TIR transposon insertions and may be used to identify transposition events and distinguish between different families of TIR transposons.³⁷ This transposition mechanism is used by most DNA transposons, including 1360, *hobo*, *pogo*, and *P* elements in *D.*

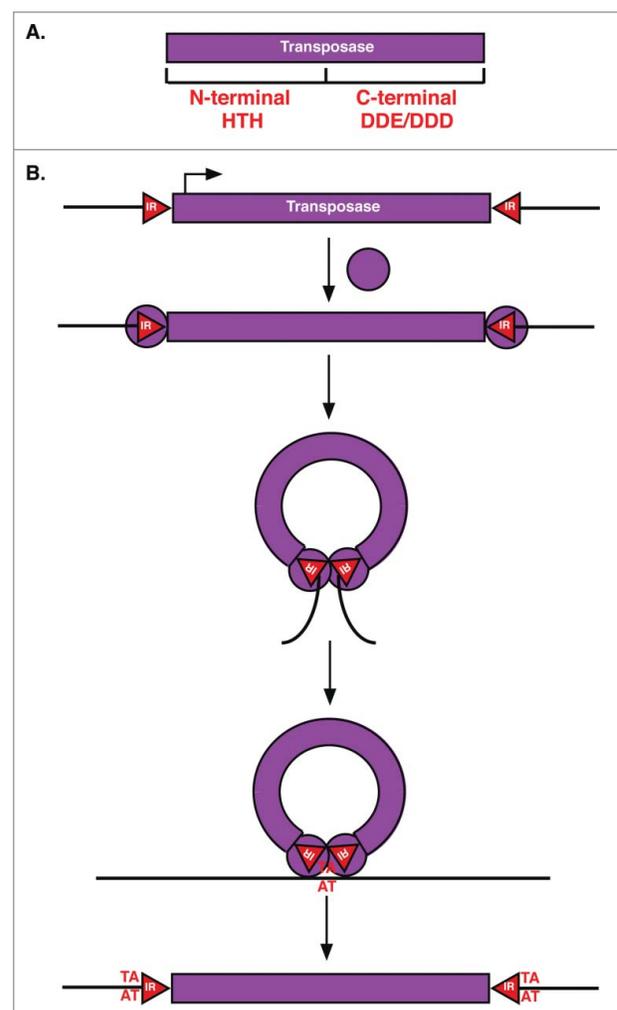


Figure 2. TIR transposase and transposition mechanism. (A) TIR Transposases have an N-terminal DNA binding domain with HTH motifs and a C-terminal DDE or DDD catalytic domain. (B) For transposition, TIR transposases (purple circles) first bind to inverted repeats (red triangles, IR) flanking the element. Bound transposases then dimerize followed by cleavage of the element from surrounding sequences (black lines) and integration into a new target site (AT) resulting in target site duplication.

melanogaster. However, the regulation of these transposition events in somatic cells is still poorly understood.

The *Tc1* and *Bari1* transposase proteins consist of 2 domains: An N-terminal DNA binding domain containing helix-turn-helix motifs and a highly conserved nuclear localization signal, and a C-terminal catalytic domain with a DDE motif (Fig. 2A).^{19,38} The catalytic DDE motif, or DDD motif in some families of transposases, is required for the transposition of DNA transposons in sub-class I.³⁹ These conserved motifs also allow the import of the transposases into the nucleus to bind TIRs, forming a complex that promotes cleavage of the entire double-stranded element.¹⁹

Non-autonomous DNA transposons, such as miniature inverted repeat transposable elements (MITEs), can also be mobilized in eukaryotic genomes. MITEs are short TIR transposons, generally less than 500 bp in length, that do not encode a functional transposase and often lack coding regions entirely.³³ Numerous *Bari*-like MITEs have been detected in 9 *Drosophila* species, and are suspected to have originated from internal deletions in *Bari* elements.³⁶ Many of these *Bari*-like MITEs in *Drosophila* have been amplified in their respective genomes.³⁶ Furthermore, MITEs derived from *mariner* elements have been identified in 20 *Drosophila* species, and make up 23% of all *mariner* elements, supporting the hypothesis that these elements may be mobilized in *trans* by other autonomous TEs in the genome.⁴⁰ Other MITEs may also arise from internal deletions in full-length TIR transposons or shorter non-autonomous TIR transposons and may continue to be amplified in the genome by the machinery of other autonomous TEs. Currently, little is known about the exact mechanisms by which MITEs are mobilized in host genomes, although *trans*-mobilization is the best supported hypothesis.

P elements

P elements are the best-studied DNA transposons in the *D. melanogaster* genome. Full-length autonomous *P* elements are 2.9 kb in length with 31 bp TIRs and 4 exons that encode a transposase when spliced.^{37,41} A similar element in the human genome, THAP9, is a confirmed DNA transposon with the ability to mobilize *P* elements in both *Drosophila* and human cell lines.⁴² Like other TIR transposons, *P* elements utilize a cut-and-paste mechanism of transposition and

create target site duplications upon insertion.⁴¹ *P* elements are unique, however, in their abilities to amplify themselves in *Drosophila* germline cells due to preferential insertion at regions of the genome that bind the origin recognition complex and function as replication origins.⁴³ By transposing during S phase from replicated genomic regions to un-replicated regions, *P* elements are copied, amplifying their presence in the genome with the assistance of the host DNA repair machinery.⁴³

These elements belong to the same class as *pogo* and *hobo* elements, and play a significant role in hybrid dysgenesis syndrome, a phenomenon observed in the progeny of hybrid crosses of certain *Drosophila* strains (Fig. 3A).⁴⁴ *Drosophila* strains are defined as either P type or M type, depending on whether hybrid dysgenesis results from crosses with the paternal or maternal parent.⁴⁵ The phenomena observed in this syndrome include high rates of mutation, recombination, and sterility in the F₁ hybrids of only P type male crosses with M type females (Fig. 3A).⁴⁶⁻⁴⁸ Alternatively, P type female crosses with M type males do not result in hybrid dysgenesis. In P strains, autonomous *P* elements are abundant and tightly regulated in germline cells, a condition referred to as the P cytotype.^{49,50} In M strains, however, there are no autonomous *P* elements nor regulatory pathways in germline cells to prevent transposition of *P* elements (Fig. 3A).^{49,51} Because the cytoplasmic conditions of the P cytotype are exclusively transmitted maternally, crosses between M type females, which lack the P cytotype, and P type males, which cannot pass on their P cytotype, result in unregulated *P* element mobilization in the germlines of the hybrid progeny (Fig. 3A).^{50,51} In addition to *P* elements, *hobo* elements and *I* elements (a family of non-LTR retrotransposons) play a similar role in hybrid dysgenesis when males and females of certain *Drosophila* strains are crossed.⁴⁵

The regulation of *P* elements in *D. melanogaster* is better understood than that of other DNA transposons. *P* element transposition is regulated primarily by alternative splicing of the *P* element transposase mRNA (Fig. 3B).⁵² In germline cells, all 3 transposase introns are spliced out, producing a functional 87 kDa *P* element transposase.^{53,54} Alternatively, in somatic cells, splicing of the third intron is skipped, generating mRNA that encodes a non-functional protein due to an early stop codon in the third intron.^{53,54} The resulting 66 kDa truncated transposase is not only inactive,

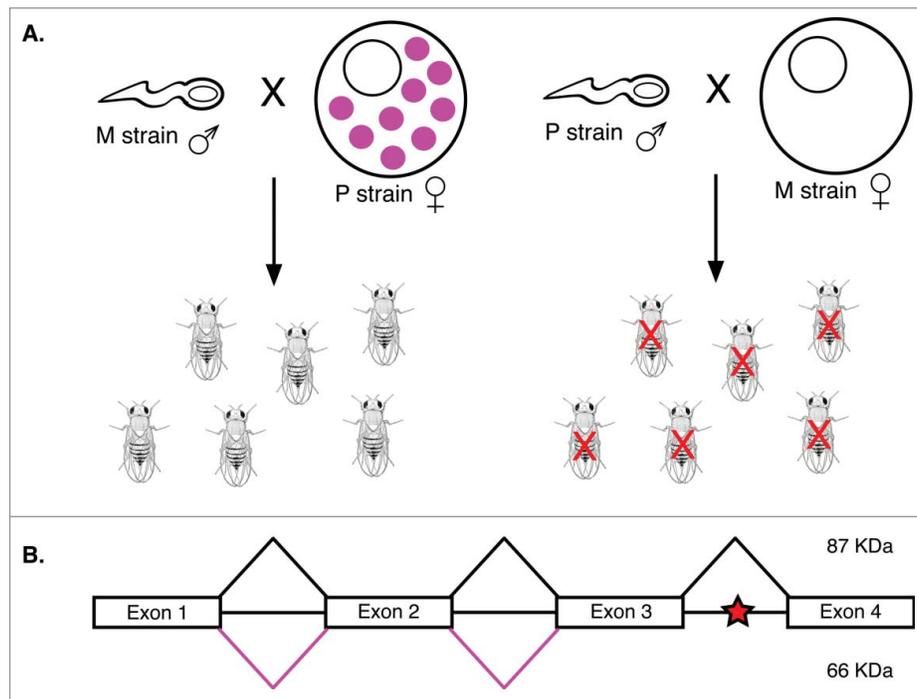


Figure 3. *P* element splicing and hybrid dysgenesis. (A) Hybrid dysgenesis results when M strain females are crossed with P strain males. Because the *P* element repressor (pink circles) is only transmitted by P cytotypic females, progeny of the P strain male-M strain female cross have many mutations caused by germline *P* element transposition. These mutations often result in sterility (red X). (B) Exons 1-4 of *P* element transcripts are spliced to form a functional 87 kDa transposase (black lines). When intron 3 is not properly spliced, a stop codon (red star) generates a 66 kDa truncated repressor of *P* element transposition (pink lines).

but represses the transposition of *P* elements in somatic cells.^{55,56} Some evidence suggests that this transposition repressor may also be generated in the germline during oogenesis, potentially contributing to the exclusively maternal transmission of the P cytotypic in P strains of *D. melanogaster*, as this is one mechanism by which female germline cells may repress *P* element transposition.⁵⁵ Other DNA transposons may also be regulated by a similar splicing mechanism, but this has not been confirmed and is not possible for some elements. The most abundant TIR transposon in *D. melanogaster* euchromatin, 1360 (also called *Hoppel*), is structurally and functionally similar to *P* elements, but lacks introns entirely, preventing regulation by any splicing mechanisms.⁵⁷ However, unlike retrotransposons, DNA transposons likely do not utilize esiRNA mechanisms of repression in somatic cells as very few small RNAs are generated from these elements.⁵⁸ Instead, DNA transposon regulation appears to occur by generation of a non-functional transposase, which may include alternative splicing mechanisms or result from mutations in the nuclear localization signal or the catalytic domain of the protein.^{19,35}

Helitrons

Helitrons belong to a unique subclass of DNA transposons with a distinct mechanism of transposition. Unlike other DNA transposons, *Helitrons* lack TIRs and encode a DNA helicase and replicator initiator (Rep) protein with nuclease and ligase functions, resembling the machinery of rolling-circle replicons (Fig. 4A).^{59,60} A subclass of *Helitrons*, called *Helentrons*, encode an additional apurinic-aprimidinic endonuclease and may also mobilize non-autonomous *Helentron*-associated interspersed elements (HINEs).^{61,62} *Helitrons* are abundant in plant genomes and have been identified in many other eukaryotic genomes, including *D. melanogaster* in which 1% of the genome consists of non-autonomous *Helitrons*.^{21,59} *Drosophila* interspersed nuclear element-1 (DINE-1), the most abundant TE in the *D. melanogaster* genome, is a non-autonomous *Helentron*, distinct from HINEs due to the presence of unique structural features such as inverted repeats.^{21,63}

Helitrons utilize a rolling-circle replication mechanism of transposition, which has recently been validated by experiments conducted with the *Helraiser*

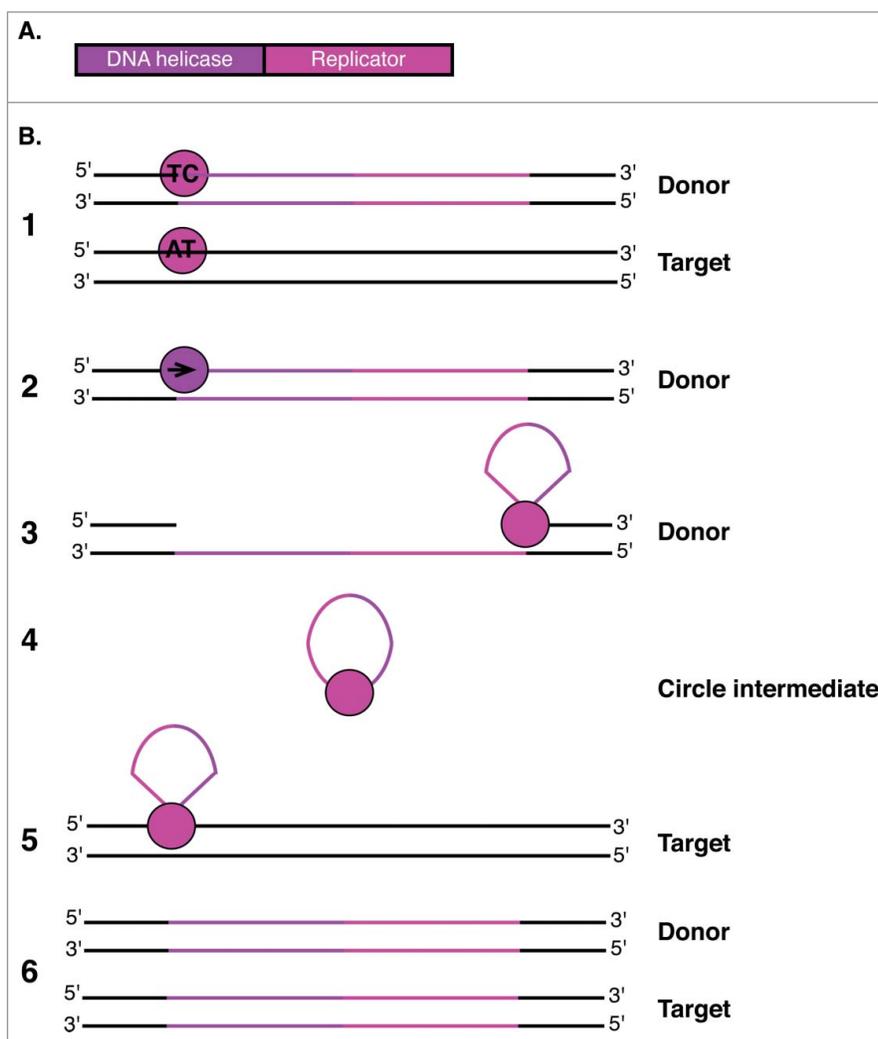


Figure 4. *Helitron* enzymes and transposition mechanism. (A) *Helitron* transposons encode a protein with both DNA helicase and Replicator functions. (B) The *Helitron* is represented with purple and pink lines.¹ The Replicator domain (pink circle) first binds to both donor (TC) and target (AT) creating nicks in both.² The DNA helicase domain (purple circle) then displaces the donor strand.³ The Replicator domain cleaves the 3' end of the element, promoting formation of a circular single-stranded DNA intermediate.⁴ Rep cleaves the circular single-stranded intermediate and promotes covalent bond formation between the 5' and 3' ends of the donor strand and target site.⁵ Host DNA replication generates a second DNA strand at both the donor and target sites.⁶ While the Replicator nicks the other end of the donor and facilitates attachment to the target site. The second strand of the element is generated at both the donor and target sites upon host DNA replication.

Helitron in bats.⁶⁴ This model suggests that tyrosine residues of the Rep protein simultaneously nick the 5' end of one *Helitron* strand at a conserved TC sequence and the AT sequence on the target site.⁶² The *Helitron* donor strand is displaced by the encoded helicase (Fig. 4B).⁶² Rep facilitates cleavage of the donor strand at a conserved hairpin signal in the 3' end, which then attacks the 5' end of the element, generating a circular, single-stranded DNA (ssDNA) intermediate (Fig. 4B).⁶⁴ To complete transposition of the element, Rep cleaves the circular ssDNA intermediate to promote covalent bond formation between the 5' and 3' ends of the *Helitron* donor strand and the nicked

target site.⁶⁴ *Helitrons* preferentially insert at AT target sites, while *Helentrons* preferentially use TT target sites and neither element creates target site duplications upon insertion.⁶² Host DNA replication is responsible for generating the second strand at both the donor and target sites, permitting amplification of these elements.⁶² *Helitrons* and *Helentrons* may also capture host genes during transposition, which complicates their classification due to poor sequence similarity.⁶⁵ This may occur when the 3' end hairpin signal in the *Helitron* is bypassed, and strand displacement continues through nearby gene regions until a new termination signal is reached.⁶⁴ These events are

prominent in maize and have contributed to the evolution of the maize genome, but are not well characterized in *D. melanogaster*.^{60,65}

Retrotransposons

Retrotransposons, or RNA transposons, comprise more than 30% of the human genome and are the most abundant class of TEs in the *D. melanogaster* genome.^{6,14} Retrotransposons include LTR retrotransposons, non-LTR retrotransposons (LINEs and LINE-like elements), short interspersed nuclear elements (SINEs), and other similar TEs.²⁵ Both LTR and non-LTR retrotransposons use similar mechanisms of transposition

and regulation. Retroviruses may also be classified as retrotransposons as they mobilize via similar mechanisms, but are additionally able to infect other cells and organisms by horizontal gene transfer. Retrotransposons are primarily characterized by the presence of *gag* and *pol* genes that may be overlapping and require frameshifting to be translated, but may also be encoded in a single fused ORF (Fig. 5A).^{66,67} Retrotransposon genes resemble those of retroviral genomes in both structure and function, and some retrotransposons contain a third gene encoding the retroviral *env* protein necessary for mobilization of retroelements outside of their host cells (Fig. 5A).⁶⁸ Many of these retrotransposons are classified as endogenous

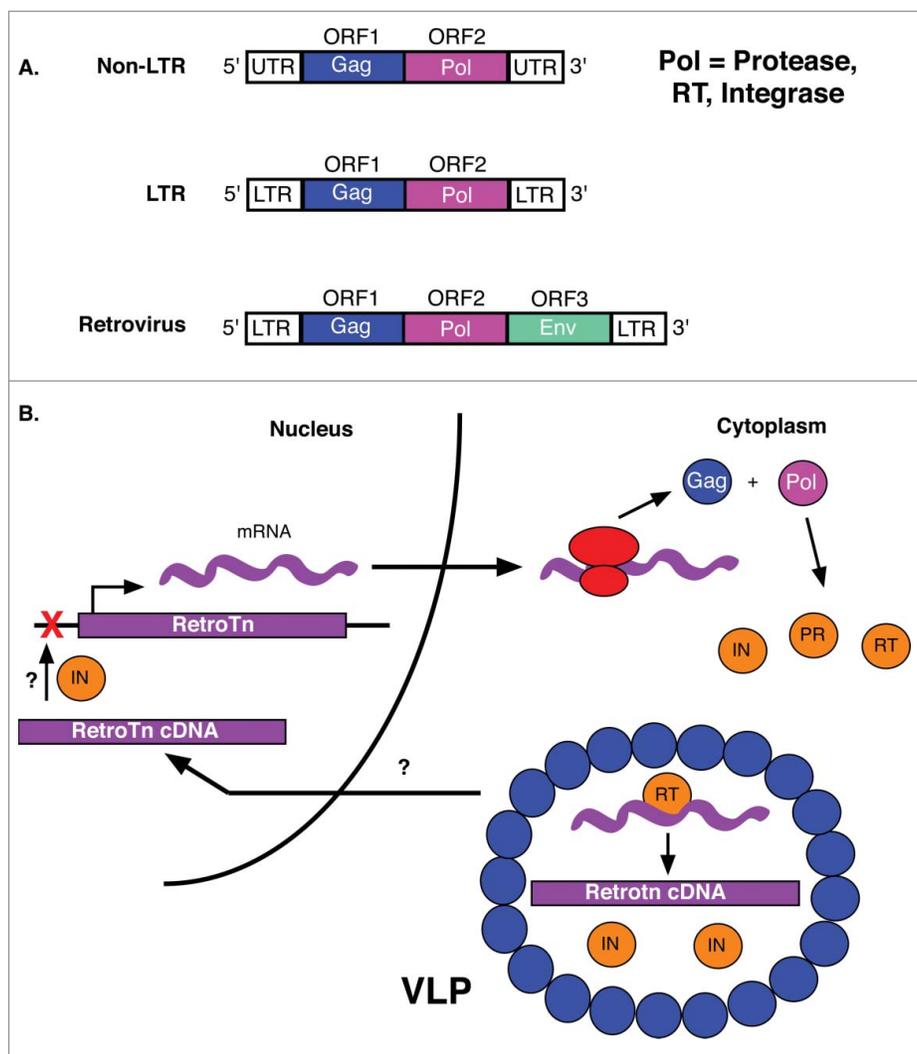


Figure 5. Retrotransposons and the LTR retrotransposition mechanism. (A) Non-LTR transposons encode both *Gag* and *Pol*, but are not flanked by LTRs. LTR retrotransposons contain *gag* and *pol* genes surrounded by LTRs. In addition to *gag* and *pol* genes, retroviruses encode an *env* gene. (B) *Gag* and *pol* of retrotransposon mRNAs are first translated into a polyprotein. The protease (PR) of the *Pol* cleaves the peptide into integrase (IN) and reverse transcriptase (RT) enzymes. The RT, retrotransposon and IN are then packaged into virus-like particles (VLPs) for import into the nucleus where retrotransposon cDNA is integrated into the genome (red X). The mechanisms by which VLP contents are localized to the nucleus and retrotransposon cDNA is integrated into the target site are unknown (?).

retroviruses, or errantiviruses in *Drosophila* and other insects, as they either arose from retroviruses that lost infectivity or LTR retrotransposons that acquired *env* genes from exogenous sources.⁶⁹

The retrotransposon *pol* gene encodes a polyprotein, typically consisting of a protease, an integrase, and a reverse transcriptase (RT) with an RNase H domain and DNA polymerase activity (Fig. 5A).⁷⁰ The RT is common to all autonomous retrotransposons, as it is required for reverse transcription of the RNA intermediate to generate DNA copies of these TEs. The protease is involved in processing of precursor proteins, such as the Pol polyprotein. The integrase is required for insertion of cDNA into the host genome. Gag is the primary component of virus-like nucleocapsid particles, formed by polymerization of Gag monomers, which provide a structural coat for components involved in the reverse transcription event of retrotransposon mobilization (Fig. 5B).⁷¹

Retrotransposon regulation

In *D. melanogaster*, the regulation of retrotransposons in somatic cells is mediated by esiRNAs, which are generated by Dcr2 cleavage of long dsRNA precursors derived from convergent sense and antisense transcription of retrotransposons in the genome.²⁶ Data support a model in which esiRNAs regulate retrotransposons in the nucleus via heterochromatin formation in *D. melanogaster* and other eukaryotic organisms, such as *S. pombe*.^{72,73} The mechanism by which this occurs has not been fully elucidated, but likely involves recruitment of heterochromatin-inducing factors by esiRNA complexes that may recognize their target RNA during active transcription of the TE. The use of small RNAs to induce heterochromatin formation is a common motif in transposon regulation, as both DNA transposons and retrotransposons are regulated in the *Drosophila* germline by the piRNA pathway. A similar siRNA-mediated pathway has been reported in humans to regulate LINE-1 retrotransposons via RNAi, but the dsRNA precursors are generated by a different mechanism than in *D. melanogaster*.⁷⁴

LTR retrotransposons

LTR retrotransposons are abundant in *Drosophila melanogaster*, as well as in humans. In *D. melanogaster*, there are 3 recognized groups of LTR

retrotransposons (Gypsy, Copia, and BEL/Pao), consisting of 8 clades and at least 35 families.^{75,76} The Gypsy group is the largest, consisting of 27 families, separated into 5 subgroups: *gypsy*, *ZAM*, *Idefix*, *412*, and *blastopia*.^{25,77} The Copia and BEL/Pao groups consist of just 4 and 5 families, respectively.²⁵ The Gypsy and BEL/Pao groups may be distinguished from the Copia group by the arrangement of their *pol* ORFs. The protease and RT are followed by the integrase in Gypsy and BEL/Pao while the protease and integrase are followed by the RT in Copia.²⁵ Mechanisms of transposition may vary slightly between these groups, but all contain LTRs, a feature also common to retroviruses. LTRs play a significant functional role in the mobilization of these elements. For both retrotransposons and retroviruses, LTRs interact directly with specific integrase domains for insertion into target regions of the genome.⁷⁸ Additionally, LTRs are processed by the integrase before insertion. Joining of the LTR ends to the chromosomal DNA generates target site duplications much like those of DNA transposon insertions.⁷⁷ Due to their structure, LTRs also permit recombination events in regions of the genome with high recombination rates and between similar elements in close proximity, often resulting in spontaneous mutations and the remainder of a single LTR (solo-LTR) at the recombination site.^{79,80} As a result, greater numbers of retrotransposon copies are detected in genomic regions with low recombination rates due to selection against these mutagenic events.^{81,82} Notably, several precise excisions of Gypsy elements from the *D. melanogaster* genome have been detected in strains known for mutations and spontaneous reversions, indicating a mechanism of retrotransposon excision other than recombination as no trace of the original element is left behind.^{83,84} One study concluded that these excisions are the result of the element's integrase directly removing the element from the genome and restoring the initial target site, potentially illuminating a new mechanism of retrotransposon mobilization.⁸³

Because a large majority of LTR retrotransposons accumulate within the inaccessible heterochromatin of *Drosophila* chromosomes, their LTR sequences may be analyzed to approximate when these elements were inserted.^{75,85} Analyses of the *D. melanogaster* genome indicate that most heterochromatic copies of LTR retrotransposons integrated after the divergence of this species from *D. simulans*, approximately 5 million

years ago.^{25,86,87} More recent analyses indicate that heterochromatic copies of these elements integrated in the *D. melanogaster* genome within the last 100,000 y while non-LTR retrotransposons are estimated to have integrated much earlier, some even millions of years before the divergence of *D. melanogaster* from *D. simulans*.^{75,86}

LTR retrotransposon insertions are not limited to heterochromatic regions, and may even occur within protein-coding regions of the genome. One study found that one-third of the LTR retrotransposons in *D. melanogaster* euchromatin are integrated either directly in a gene or within 1000 bp of a gene.⁸⁸ Most of these gene-associated insertions have occurred relatively recently, and because they tend to occur in highly conserved genes essential to cell survival, are selected against over time.⁸⁸ Interestingly, more of these LTR retrotransposon insertions are associated with genes involved in signal transduction, morphogenesis, behavior and responses to external stimuli than genes involved in cell differentiation and metabolism.⁸⁸ In addition to the functional effects of gene-associated LTR retrotransposon insertion, gene expression can be affected if these elements are inserted near promoter elements. A recent study found that several solo-LTR elements of the *roo* family are inserted near the transcription start site (TSS) of a candidate cold resistance gene (*CG18446*) in several strains of *D. melanogaster*, contributing cis-regulatory elements to the promoter of this gene and affecting transcription factor binding sites.⁸⁹ One of these inserted solo-LTRs, *FBti0019985*, generated a new TSS for *CG18446*.⁸⁹ Strains carrying this insert demonstrated upregulated expression of the gene in embryos as well as increased viability under both cold-stress and non-stress conditions relative to control strains.⁸⁹ These observations demonstrate the potential of retrotransposon insertions to not only cause detrimental mutations in the genome, but to also contribute adaptive functions to their hosts.

Several LTR retrotransposons contain a third gene downstream of *gag* and *pol*, the *env* gene of retroviral genomes, potentially permitting horizontal transmission to other cells and organisms. The *env* gene is often non-functional in LTR retrotransposons, although this is not always the case. Because these elements strongly resemble proviruses (retroviruses that have integrated in host genomic DNA) they can be difficult to classify. For example, the *gypsy* retrotransposon of *D.*

melanogaster was the first identified endogenous retrovirus in invertebrates, as it can be horizontally transmitted.^{90,91} However, despite characterization of this element as an errantivirus, phylogenetic analyses based on RT sequences still group *gypsy* with other LTR retrotransposons.^{18,92} Several other LTR retrotransposons in *Drosophila* are similarly characterized due to the presence of the *env* gene, including *297*, *17.6*, *tom*, *Idefix*, *ZAM*, and *tirant*.^{93,94} One study concluded that the Gypsy group of errantiviruses in *Drosophila* obtained their *env* genes from insect baculoviruses, DNA viruses that exclusively infect insects and arthropods.⁹⁴ This conclusion is supported by the observation that *TED*, a member of the *gypsy* family of LTR retrotransposons, can integrate into the genome of the baculovirus *Autographa californica*, permitting capture of baculoviral elements.⁹⁵ Events like these may be responsible for the unusually high occurrence of horizontal transfer of TEs in *Drosophila*.

LTR retrotransposon transposition

Retroelements are first transcribed into *gag-pol* fusion transcripts followed by translation into Gag-Pol fusion protein products, sometimes by programmed translational frameshift. Gag-Pol peptides are then rapidly cleaved into individual protein products by the retroelement encoded protease (Fig. 5B).⁷¹ Programmed translational frameshift occurs in many retrotransposon transcripts near the end of the *gag* ORF due to a rare codon awaiting the arrival of its corresponding tRNA.⁶⁶ Resulting translational pausing permits recognition of a more common, but frameshifted, codon, allowing efficient translation of the *pol* ORF in frame without interruption by the *gag* ORF stop codon.⁶⁶ Following processing of the fusion protein, post-translationally modified Gag proteins polymerize to generate virus-like particles (VLPs) in the cytoplasm, capturing the retrotransposon transcript, RT, and integrase (Fig. 5B).⁹⁶⁻⁹⁸ A smaller, likely unmodified Gag of the LTR retrotransposon *1731* in *D. melanogaster* localizes to the nucleus, potentially contributing to the transfer of reverse transcription products from VLPs in the cytoplasm to the nucleus for insertion (Fig. 5B).⁹⁹ However, nuclear localization signals have also been identified in integrases of several eukaryotic retrotransposons, much like the transposases of DNA transposons, and may play a role in the delivery of VLP contents to the nucleus.¹⁰⁰⁻¹⁰²

The same mRNA molecule can be both translated in the cytoplasm and captured in a VLP for use as the RT template. Some retrotransposons in *D. melanogaster* contain extended 5' untranslated regions (UTRs) that regulate the transition of mRNA molecules from translating to packaging states. For example, LTR retrotransposon *Idefix* has a long 5' UTR that generates an internal ribosome entry site in its mRNA to promote translation, but its translation can also be downregulated by Gag binding to the 5' UTR to promote capture within VLPs.¹⁰³ Within VLPs, single-stranded mRNAs are reverse transcribed to generate double-stranded DNA copies (Fig. 5B).¹⁰³ The RT cleaves the RNA in RNA-DNA hybrids and also has DNA-dependent DNA polymerase activity (Fig. 5B).¹⁰³ In yeast, the reverse transcription reaction of the LTR retrotransposon Ty1 is primed by the initiator methionine tRNA.¹⁰⁴ However, other priming mechanisms have been observed, such as the self-priming mechanism of the Tf1 LTR retrotransposon.¹⁰⁵ Much of this process is unclear and probably varies for different types and families of retrotransposons.

Several mechanisms have been proposed for integration of retrotransposon cDNA into the host genome (Fig. 5B). While many retrotransposons demonstrate no specificity for target insertion sites, elements of the *gypsy* family of LTR retrotransposons in *D. melanogaster* show some target site preference.⁷⁷ Integrase sequences are generally highly conserved among copies of the same retrotransposon, despite the atypically high variability of retrotransposon sequences, demonstrating their significant role in mobilization of these elements.⁷⁷ Mechanisms by which integrases interact with genomic DNA for retrotransposon insertion are unclear, although chromatin accessibility and other structural features appear to play a role. A significant amount of research regarding integrase functions has been performed with the retroviral integrase of Human Immunodeficiency Virus type 1 (HIV-1) and the integrases of Ty1 and Ty3 LTR retrotransposons in yeast.

Non-LTR retrotransposons

Non-LTR retrotransposons, or LINE-like elements, have been classified into over 100 families, separated into 28 clades and 6 groups: R2, L1, RTE, I, Jockey and RandI.^{106,107} Non-LTR retrotransposons are structurally similar to LTR retrotransposons, but often lack some of the ORFs and protein domains encoded

by LTR retrotransposons and do not contain LTRs at their 3' and 5' ends (Fig. 6A).¹⁰⁸ The absence of LTRs suggests that these elements interact with their encoded proteins differently than LTR retrotransposons and may utilize different mechanisms of transposition (Fig. 6B). While non-LTR and LTR retrotransposons encode similar proteins and often generate target site duplications, the reverse transcription and integration events of non-LTR retrotransposon mobilization are unique, at least for the R2 group of elements that often lack promoters and only encode a single ORF with RT and endonuclease activities (Fig. 6A).

Studies in both *D. melanogaster* and *Bombyx mori* (silkworm) have demonstrated that non-LTR retrotransposons utilize target primed reverse transcription (TPRT) for integration of new retrotransposon copies in the genome (Fig. 6B).¹⁰⁷⁻¹¹² Because R2 non-LTR retrotransposons are co-transcribed with their flanking 28S rRNA sequences (Fig. 6A), these elements target 28S rRNA genes for insertion. TPRT is initiated by single-stranded nicking of the target DNA by the element's encoded endonuclease (Fig. 6B).¹⁰⁸ The generated 3' hydroxyl group then primes reverse transcription of the RNA intermediate before cleavage of the second target DNA strand (Fig. 6B).¹⁰⁸ Cleavage of the second strand may occur at the same location as the first strand, or 2 base pairs upstream or downstream of this site.¹⁰⁷ Second strand cleavage location determines whether the target site is unchanged, deleted, or duplicated. The RT/EN encoded by R2 is responsible for both cleavage of the target DNA and reverse transcription of the element.¹⁰⁸ R2 RT/EN also generates the complementary R2 strand at the target site as this enzyme demonstrates DNA-dependent DNA polymerase activity and can displace the RNA template (Fig. 6B).¹⁰⁹ The 3' UTR (Fig. 6A) of these elements is required for TPRT and is inserted during reverse transcription of the element, while integration of the 5' end is highly variable and thought to involve DNA repair and homologous recombination with regions of the 28S rRNA gene.¹¹³ In the absence of upstream homologous 28S rRNA sequences, integration of these elements is either prevented or results in truncations of the 5' UTR.^{112,113} Because the endonuclease interacts with both ends of the integrating RNA, protein dimerization may be required for R2 transposition.¹¹⁴

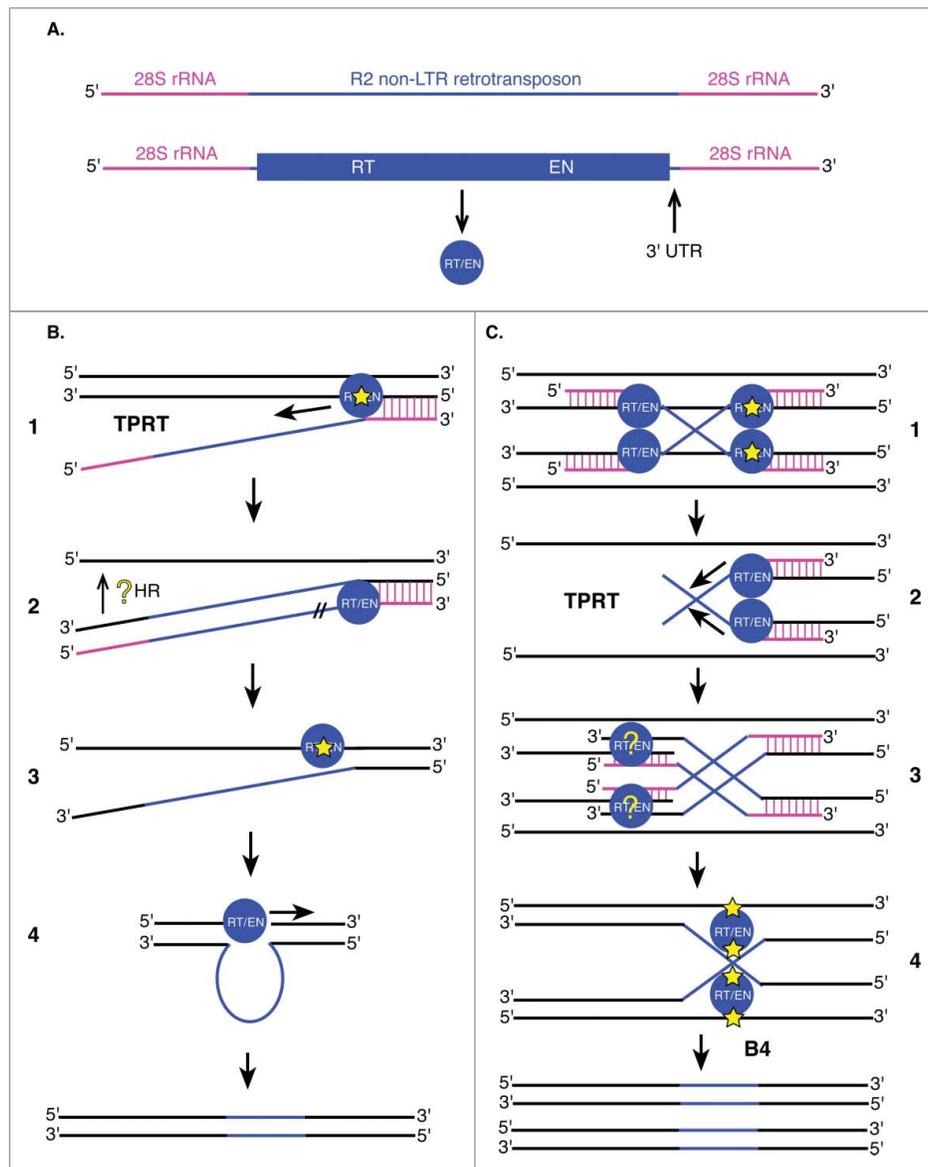


Figure 6. Non-LTR retrotransposons utilize target primed reverse transcription (TPRT) for integration. (A) R2 non-LTR retrotransposons (blue) are flanked by 28S rRNA genomic sequences (pink). The single R2 ORF encodes an enzyme with reverse transcriptase (RT) and endonuclease (EN) activities (blue circle). Other non-LTR retrotransposons may encode these enzymes as 2 separate proteins (RT and integrase with EN activity). The 3' UTR is important for integration of R2 retrotransposons into 28S rRNA genes. (B and C) Proposed models of non-LTR retrotransposon (B) and R2 (C) insertion. DNA is shown in black (including reverse transcribed flanking sequences), 28S rRNA sequences in pink and retrotransposon sequences (mRNA and DNA) in blue, following the color scheme in (A). (B1) Non-LTR retrotransposon transcripts first hybridize to 28S rRNA sequences (vertical pink lines) followed by initiation of TPRT by single-stranded nicking of the target DNA (yellow star) by the element's encoded endonuclease. (B2) Following reverse transcription of the element, element mRNA is degraded by R2 RT/EN (//). Integration of the 5' end of the element is not well understood (yellow ?). (B3) Cleavage of the second strand (yellow star) may occur at the same location as the first strand, or 2 base pairs upstream or downstream of this site. (B4) R2 RT/EN also generates the complementary R2 strand at the target site to fully transpose the element. (C1, C2, C3) The initial steps of this alternative mechanism are identical to those described in B1 and B2 except they take place on 2 homologous targets simultaneously resulting in a Holliday junction intermediate (C4). The Holliday junction intermediate is resolved by R2 RT/EN (C4) followed by second strand synthesis resulting in fully-integrated R2 non-LTR retrotransposons in 2 new locations.

A recent study in *D. melanogaster* found that R2 endonuclease domains are homologous to those of *FokI* restriction enzymes and Holliday-junction resolvases. These associations led the authors to propose a

new model of TPRT transposition for these elements (Fig. 6C).¹⁰⁷ Two R2 elements may utilize their flanking 28S rRNA sequences to bind regions of homologous chromosomes, generating a Holliday junction

structure that is resolved by a dimer of the elements' endonucleases (Fig. 6C).¹⁰⁷ While target site preferences are variable among non-LTR retrotransposons, the TPRT and Holliday junction mechanisms may also be used by other non-LTR retrotransposons.

A few closely related non-LTR retrotransposons present in all *Drosophila* genomes, *HeT-A*, *TART*, and *TAHRE*, play a significant role in telomere maintenance and use a unique mechanism to localize during transposition.¹¹⁵⁻¹¹⁷ These elements are targeted to telomeres by their encoded Gag proteins, permitting generation of telomeric tandem repeats and serving functions similar to that of telomerase.¹¹⁸ Unlike most non-LTR retrotransposons, *HeT-A* elements encode 2 overlapping, frameshifted *gag* ORFs and no *pol* ORF and can therefore only be mobilized *in trans* by the enzymes of other retroelements.¹¹⁷ Alternatively, *TART* and *TAHRE* elements contain 2 non-overlapping ORFs encoding both *gag* and *pol* genes, but are less abundant in *Drosophila* telomeres than *HeT-A*.¹¹⁶ The *gag* ORFs of *HeT-A* elements are generally highly variable in both length and sequence, but contain some conserved motifs present in the other telomeric retrotransposons, such as the zinc knuckle (CCHC) in the *gag* ORF, which may be repeated several times within the element.¹¹⁹ The Gag proteins of *HeT-A* and *TART* elements localize to the nucleus, where *HeT-A* Gags aggregate in telomeric regions during interphase, forming Het dots.¹¹⁸ *TART* Gags, though localized to the nucleus, only associate with telomeres in the presence of *HeT-A* Gags, demonstrating the significance of *HeT-A* Gags in telomere-targeting and potentially signifying *TART* as the RT donor for *HeT-A* transposition.^{118,120} The Gag protein of *TAHRE* shares sequence similarity with the Gag of *HeT-A*, and generally localizes to the nucleus, but only localizes to the Het dots around telomeres in the presence of *HeT-A* Gags.¹¹⁸ Researchers hypothesize that the telomere-targeting specificity of the Gag proteins contributes to the abundance of telomeric *HeT-A*, relative to *TART* and *TAHRE*, despite the absence of *pol* genes in *HeT-A* elements.¹¹⁸ Once targeted to telomeres, telomeric retrotransposons presumably integrate via TPRT, like other non-LTR retrotransposons, and may also be regulated by small RNA pathways to induce heterochromatin formation in these regions.¹²¹⁻¹²³ Observations of telomeric retrotransposon regulation in the germline indicate that regulation of these elements is significant to development as sequences targeted by

piRNAs in the *HeT-A* 3' UTR are highly conserved in *D. melanogaster* and closely related species.¹²²

Discussion

The mechanisms by which TEs mobilize in *Drosophila* and other eukaryotic genomes reveal several common features. For example, the primary integrating enzymes encoded by these elements, transposases and integrases/endonucleases, permit cleavage of target sites in the host genome and promote insertion of TEs into these genomic locations. These enzymes also require interactions with specific structural elements flanking the ORFs of their respective TEs, such as TIRs, LTRs, and 3' or 5' UTRs. Furthermore, many TEs, including DNA transposons, have demonstrated an ability to amplify upon mobilization, either through an RNA intermediate in the case of retrotransposons or through timing transposition events with events of the host cell cycle in the case of some DNA transposons. The similarities between LTR and non-LTR retrotransposon mobilization are also apparent, such as the formation of VLPs in the cytosol via polymerization of encoded Gag proteins, an event resembling a stage of the retroviral life cycle. The connection between LTR retrotransposons and retroviruses is further characterized by the presence of LTRs in all of these elements and the presence of the retroviral *env* gene in many *D. melanogaster* LTR retrotransposons. However, the relationship between retrotransposons and retroviruses remains unclear as studies have demonstrated the propensity of retrotransposons to acquire *env* genes and function as retroviruses, yet the presence of these elements in eukaryotic genomes in the first place may be the result of horizontal transfer from ancient viruses or retroviruses.¹²⁴

In addition to similar mechanisms of transposition, TEs are regulated by a common mechanism in *D. melanogaster*: small RNA biogenesis. While DNA transposons do not appear to be regulated by the esiRNA pathway which regulates retrotransposons in somatic cells, the piRNA regulatory pathway of germline cells regulates all TEs in *D. melanogaster*.²⁶ Both regulatory pathways rely on convergent sense and antisense transcription of TEs to generate the precursors to the siRNAs of these pathways. While these pathways utilize distinct proteins for the processing of siRNA precursors, the generated siRNAs regulate TEs via similar

mechanisms, such as heterochromatin formation.¹²⁵ The factors that regulate mobilization of TEs in host genomes may significantly influence genome size evolution, as TE abundance correlates with genome size and is directly dependent upon the efficiency of transposition regulation and selection against deleterious transposition events over evolutionary time.¹²⁶ Furthermore, the interplay between host silencing of TEs and transposition demonstrates a host-parasite co-evolution in which familiar TEs are better regulated by their hosts than newly introduced TEs. This relationship is exemplified by hybrid dysgenesis in *Drosophila* and dysregulated transposition of TEs resulting from other hybrid crosses in eukaryotes, as hybrid hosts are maladapted to the newly introduced TEs in their genomes.¹²⁶

Functionally, TEs may have a broad range of impacts on their hosts. Most deleterious integrations of TEs into host genomes are negatively selected against over time, while some TE insertions may provide adaptive functions to their hosts, such as the insertion of the solo-LTR *FBti0019985* in a candidate cold stress response gene of *D. melanogaster*.⁸⁹ Furthermore, the role of TEs in *Drosophila* telomere maintenance demonstrates the ability of these elements to develop significant functional roles in their hosts to positively influence genome stability. This has strong implications for the role of TEs in the evolutionary development of host genomes, as selective forces act on these transpositional events, influencing the coevolution of the genome and its TEs. As more is learned about the origin of TEs and their regulation by host genomes, the evolution and roles of TEs in eukaryotic genomes will become better defined.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

References

- [1] Pagel M, Johnstone RA. Variation across species in the size of the nuclear genome supports the junk-DNA explanation for the C-value paradox. *Proc Biol Sci* 1992; 249:119-24; PMID:1360673; <https://doi.org/10.1098/rspb.1992.0093>
- [2] Vieira C, Nardon C, Arpin C, Lepetit D, Biemont C. Evolution of genome size in *Drosophila*. Is the invader's genome being invaded by transposable elements? *Mol Biol Evol* 2002; 19:1154-161; PMID:12082134
- [3] Bosco G, Campbell P, Leiva-Neto JT, Markow TA. Analysis of *Drosophila* species genome size and satellite DNA content reveals significant differences among strains as well as between species. *Genetics* 2007; 177:1277-90; PMID:18039867; <https://doi.org/10.1534/genetics.107.075069>
- [4] Gregory TR, Johnston JS. Genome size diversity in the family Drosophilidae. *Heredity (Edinb)* 2008; 101:228-38; PMID:18523443; <https://doi.org/10.1038/hdy.2008.49>
- [5] SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, et al. Nested retrotransposons in the intergenic regions of the maize genome. *Science* 1996; 274:765-8; PMID:8864112
- [6] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409:860-921; PMID:11237011; <https://doi.org/10.1038/35057062>
- [7] Lannoy N, Hermans C. Principles of genetic variations and molecular diseases: Applications in hemophilia A. *Crit Rev Oncol Hematol* 2016; 104:1-8; PMID:27296059; <https://doi.org/10.1016/j.critrevonc.2016.04.005>
- [8] Solyom S, Ewing AD, Rahrman EP, Doucet T, Nelson HH, Burns MB, Harris RS, Sigmon DF, Casella A, Erlanger B, et al. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res* 2012; 22:2328-38; PMID:22968929; <https://doi.org/10.1101/gr.145235.112>
- [9] Chen L, Dahlstrom JE, Chandra A, Board P, Rangasamy D. Prognostic value of LINE-1 retrotransposon expression and its subcellular localization in breast cancer. *Breast Cancer Res Treat* 2012; 136:129-42; PMID:23053642; <https://doi.org/10.1007/s10549-012-2246-7>
- [10] Batzer MA, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet* 2002; 3:370-9; PMID:11988762; <https://doi.org/10.1038/nrg798>
- [11] Li W, Jin Y, Prazak L, Hammell M, Dubnau J. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PLoS One* 2012; 7:e44099; PMID:22957047; <https://doi.org/10.1371/journal.pone.0044099>
- [12] Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, Brennan PM, Rizzu P, Smith S, Fell M, et al. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 2011; 479:534-7; PMID:22037309; <https://doi.org/10.1038/nature10531>
- [13] Reilly MT, Faulkner GJ, Dubnau J, Ponomarev I, Gage FH. The role of transposable elements in health and diseases of the central nervous system. *J Neurosci* 2013; 33:17577-86; PMID:24198348; <https://doi.org/10.1523/JNEUROSCI.3369-13.2013>
- [14] Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, et al. The transposable elements of the *Drosophila melanogaster* euchromatin: A genomics

- perspective. *Genome Biol* 2002; 3:1-20; http://www.fruitfly.org:9005/about/pubs/genome_biology/gb-2002-3-12-research0084.pdf
- [15] Barrón MG, Fiston-Lavier A-S, Petrov DA, González J. Population genomics of transposable elements in *Drosophila*. *Annu Rev Genet* 2014; 48:561-81; PMID:25292358; <https://doi.org/10.1146/annurev-genet-120213-092359>
- [16] McClintock B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A* 1950; 36:344-55; PMID:15430309; <https://doi.org/10.1073/pnas.36.6.344>
- [17] Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends Genet* 1989; 5:103-7; PMID:2543105; [https://doi.org/10.1016/0168-9525\(89\)90039-5](https://doi.org/10.1016/0168-9525(89)90039-5)
- [18] Lerat E, Capy P. Retrotransposons and retroviruses: Analysis of the envelope gene. *Mol Biol Evol* 1999; 16:1198-207; PMID:10486975
- [19] Muñoz-López M, García-Pérez JL. DNA transposons: Nature and applications in genomics. *Curr Genomics* 2010; 11:115-28; PMID:20885819; <https://doi.org/10.2174/138920210790886871>
- [20] Palazzo A, Marconi S, Specchia V, Bozzetti MP, Ivics Z, Caizzi R, Marsano RM. Functional characterization of the Bari transposition system. *PLoS One* 2013; 8:e79385; PMID:24244492; <https://doi.org/10.1371/journal.pone.0079385>
- [21] Kapitonov VV, Jurka J. Helitrons on a roll: Eukaryotic rolling-circle transposons. *Trends Genet* 2007; 23:521-9; PMID:17850916; <https://doi.org/10.1016/j.tig.2007.08.004>
- [22] Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* 2003; 100:5280-5; PMID:12682288; <https://doi.org/10.1073/pnas.0831042100>
- [23] Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 2009; 10:691-703; PMID:19763152; <https://doi.org/10.1038/nrg2640>
- [24] Sheen FM, Levis RW. Transposition of the LINE-like retrotransposon TART to *Drosophila* chromosome termini. *Proc Natl Acad Sci U S A* 1994; 91:12510-4; PMID:7809068
- [25] Bowen NJ, McDonald JF. *Drosophila* euchromatic LTR retrotransposons are much younger than the host species in which they reside. *Genome Res* 2001; 11:1527-40; PMID:11544196; <https://doi.org/10.1101/gr.164201>
- [26] Russo J, Harrington AW, Steiniger M. Antisense transcription of retrotransposons in *Drosophila*: An origin of endogenous small interfering RNA precursors. *Genetics* 2016; 202:107-21; PMID:26534950; <https://doi.org/10.1534/genetics.115.177196>
- [27] Yin H, Lin H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in *Drosophila melanogaster*. *Nature* 2007; 450:304-8; PMID:17952056; <https://doi.org/10.1038/nature06263>
- [28] Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. *Mol Cell* 2013; 50:736-48; PMID:23665228; <https://doi.org/10.1016/j.molcel.2013.04.006>
- [29] Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 2006; 313:320-4; PMID:16809489; <https://doi.org/10.1126/science.1129333>
- [30] Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 2007; 128:1089-103; PMID:17346786; <https://doi.org/10.1016/j.cell.2007.01.043>
- [31] Siomi MC, Miyoshi T, Siomi H. piRNA-mediated silencing in *Drosophila* germlines. *Semin Cell Dev Biol* 2010; 21:754-9; PMID:20080197; <https://doi.org/10.1016/j.semcdb.2010.01.011>
- [32] Pritham EJ. Transposable elements and factors influencing their success in eukaryotes. *J Hered* 2009; 100:648-55; PMID:19666747; <https://doi.org/10.1093/jhered/esp065>
- [33] Muñoz-López M, García-Pérez JL. DNA transposons: Nature and applications in genomics. *Curr Genomics* 2010; 11:115-28; PMID:20885819; <https://doi.org/10.2174/138920210790886871>
- [34] Chalopin D, Naville M, Plard F, Galiana D, Volff J-N. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. *Genome Biol Evol* 2015; 7:567-80; PMID:25577199; <https://doi.org/10.1093/gbe/evv005>
- [35] Caizzi R, Caggese C, Pimpinelli S. Bari-1, a new transposon-like family in *Drosophila melanogaster* with a unique heterochromatic organization. *Genetics* 1993; 133:335-45; PMID:8382176
- [36] Palazzo A, Lovero D, D'Addabbo P, Caizzi R, Marsano RM. Identification of Bari transposons in 23 sequenced *Drosophila* genomes reveals novel structural variants, MITEs and horizontal transfer. *PLoS One* 2016; 11:e0156014; PMID:27213270; <https://doi.org/10.1371/journal.pone.0156014>
- [37] Karess RE, Rubin GM. Analysis of P transposable element functions in *Drosophila*. *Cell* 1984; 38:135-46; PMID:6088058
- [38] Brillet B, Benjamin B, Bigot Y, Yves B, Augé-Gouillou C, Corinne A-G. Assembly of the Tc1 and mariner transposition initiation complexes depends on the origins of their transposase DNA binding domains. *Genetica* 2007; 130:105-20; PMID:16912840; <https://doi.org/10.1007/s10709-006-0025-2>
- [39] Yuan Y-W, Wessler SR. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. *Proc Natl Acad Sci U S A* 2011; 108:7884-9; PMID:21518873; <https://doi.org/10.1073/pnas.1104208108>
- [40] Wallau GL, Capy P, Loreto E, Hua-Van A. Genomic landscape and evolutionary dynamics of mariner

- transposable elements within the *Drosophila* genus. *BMC Genomics* 2014; 15:727; PMID:25163909; <https://doi.org/10.1186/1471-2164-15-727>
- [41] O'Hare K, Rubin GM. Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* 1983; 34:25-35; PMID:6309410
- [42] Majumdar S, Singh A, Rio DC. The human THAP9 gene encodes an active P-element DNA transposase. *Science* 2013; 339:446-8; PMID:23349291; <https://doi.org/10.1126/science.1231789>
- [43] Spradling AC, Bellen HJ, Hoskins RA. *Drosophila* P elements preferentially transpose to replication origins. *Proc Natl Acad Sci* 2011; 108:15948-53; PMID:21896744; <https://doi.org/10.1073/pnas.1112960108>
- [44] Tudor M, Lobočka M, Goodell M, Pettitt J, O'Hare K. The pogo transposable element family of *Drosophila melanogaster*. *Mol Gen Genet* 1992; 232:126-34; PMID:1313144; <https://doi.org/10.1007/BF00299145>
- [45] Kidwell MG, Kidwell JF, Sved JA. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* 1977; 86:813-33; PMID:17248751
- [46] Kidwell MG, Kidwell JF, Nei M. A case of high rate of spontaneous mutation affecting viability in *Drosophila melanogaster*. *Genetics* 1973; 75:133-53; PMID:4202770
- [47] Sved JA. Hybrid dysgenesis in *Drosophila melanogaster*: A possible explanation in terms of spatial organization of chromosomes. *Aust J Biol Sci* 1976; 29:375-88; PMID:828493
- [48] Hiraizumi Y. Spontaneous recombination in *Drosophila melanogaster* males. *Proc Natl Acad Sci U S A* 1971; 68:268-70; PMID:5277066
- [49] Coen D, Lemaitre B, Delattre M, Quesneville H, Ronsseray S, Simonelig M, Higuët D, Lehmann M, Montchamp C, Nouaud D. *Drosophila* P element: Transposition, regulation and evolution. *Genetica* 1994; 93:61-78; PMID:7813918
- [50] Simmons MJ, Meeks MW, Jessen E, Becker JR, Buschette JT, Thorp MW. Genetic interactions between P elements involved in piRNA-mediated repression of hybrid dysgenesis in *Drosophila melanogaster*. *G3 (Bethesda)* 2014; 4:1417-27; PMID:24902606; <https://doi.org/10.1534/g3.114.011221>
- [51] Lemaitre B, Ronsseray S, Coen D. Maternal repression of the P element promoter in the germline of *Drosophila melanogaster*: A model for the P cytotyping. *Genetics* 1993; 135:149-60; PMID:8224815
- [52] Engels WR. The P family of transposable elements in *Drosophila*. *Annu Rev Genet* 1983; 17:315-44; PMID:6320712; <https://doi.org/10.1146/annurev.ge.17.120183.001531>
- [53] Siebel CW, Rio DC. Regulated splicing of the *Drosophila* P transposable element third intron in vitro: Somatic repression. *Science* 1990; 248:1200-8; PMID:2161558
- [54] Laski FA, Rio DC, Rubin GM. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. *Cell* 1986; 44:7-19; PMID:3000622; [https://doi.org/10.1016/0092-8674\(86\)90480-0](https://doi.org/10.1016/0092-8674(86)90480-0)
- [55] Misra S, Rio DC. Cytotype control of *Drosophila* P element transposition: The 66 kd protein is a repressor of transposase activity. *Cell* 1990; 62:269-84; PMID:2164887
- [56] Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK, Robertson HM, Engels WR. Type I repressors of P element mobility. *Genetics* 1993; 135:81-95; PMID:8224830
- [57] Reiss D, Quesneville H, Nouaud D, Andrieu O, Anxolabehere D. Hoppel, a P-like element without introns: A P-element ancestral structure or a retrotranscription derivative? *Mol Biol Evol* 2003; 20:869-79; PMID:12717001; <https://doi.org/10.1093/molbev/msg090>
- [58] Harrington AW, Steiniger M. Bioinformatic analyses of sense and antisense expression from terminal inverted repeat transposons in *Drosophila* somatic cells. *Fly (Austin)* 2016; 10:1-10; PMID:26986720; <https://doi.org/10.1080/19336934.2016.1165372>
- [59] Kapitonov VV, Jurka J. Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* 2001; 98:8714-9; PMID:11447285; <https://doi.org/10.1073/pnas.151269298>
- [60] Lal S, Oetjens M, Hannah LC. Helitrons: Enigmatic abductors and mobilizers of host genome sequences. *Plant Science* 2009; 176:181-6.
- [61] Poulter RTM, Goodwin TJD, Butler MI. Vertebrate helitrons and other novel Helitrons. *Gene* 2003; 313:201-12; PMID:12957391
- [62] Thomas J, Pritham EJ. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol Spectr* 2015; 3; PMID:26350323; <https://doi.org/10.1128/microbiolspec.MDNA3-0049-2014>
- [63] Thomas J, Vadnagara K, Pritham EJ. DINE-1, the highest copy number repeats in *Drosophila melanogaster* are non-autonomous endonuclease-encoding rolling-circle transposable elements (Helitrons). *Mob DNA* 2014; 5:18; PMID:24959209; <https://doi.org/10.1186/1759-8753-5-18>
- [64] Grabundzija I, Messing SA, Thomas J, Cosby RL, Bilic I, Miskey C, Gogol-Döring A, Kapitonov V, Diem T, Dalda A, et al. A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat Commun* 2016; 7:10716; PMID:26931494; <https://doi.org/10.1038/ncomms10716>
- [65] Barbaglia AM, Klusman KM, Higgins J, Shaw JR, Hannah LC, Lal SK. Gene capture by Helitron transposons reshuffles the transcriptome of maize. *Genetics* 2012; 190:965-75; PMID:22174072; <https://doi.org/10.1534/genetics.111.136176>
- [66] Kalmykova A, Maisonhaute C, Gvozdev V. Retrotransposon 1731 in *Drosophila melanogaster* changes retrovirus-like expression strategy in host genome. *Genetica* 1999; 107:73-77; PMID:10952199
- [67] Gao X, Havecker ER, Baranov PV, Atkins JF, Voytas DF. Translational recoding signals between gag and pol

- in diverse LTR retrotransposons. *RNA* 2003; 9:1422-30; PMID:14623998
- [68] Kim FJ, Battini J-L, Manel N, Sitbon M. Emergence of vertebrate retroviruses and envelope capture. *Virology* 2004; 318:183-91; PMID:14972546; <https://doi.org/10.1016/j.virol.2003.09.026>
- [69] Stefanov Y, Salenko V, Glukhov I. *Drosophila* errantiviruses. *Mob Genet Elements* 2012; 2:36-45; PMID:22754751; <https://doi.org/10.4161/mge.19234>
- [70] Engelman A. Retroviral integrase structure and DNA recombination mechanism. *Microbiol Spectrum* 2(6): MDNA3-0024-2014; <https://doi.org/10.1128/microbiolspec.MDNA3-0024-2014>.
- [71] Farabaugh PJ. Programmed translational frameshifting. *Annual Review of Genetics* 1996; 30:507-28.
- [72] Fagegaltier D, Bougé A-L, Berry B, Poisot E, Sismeiro O, Coppée J-Y, Théodore L, Vionnet O, Antoniewski C. The endogenous siRNA pathway is involved in heterochromatin formation in *Drosophila*. *Proceedings of the National Academy of Sciences*; 106(50):21258-21263; <https://doi.org/10.1073/pnas.0809208105>
- [73] Kloc A, Martienssen R. RNAi, heterochromatin and the cell cycle. *Trends Genet* 2008; 24:511-7; PMID:18778867; <https://doi.org/10.1016/j.tig.2008.08.002>
- [74] Yang N, Kazazian HH. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* 2006; 13:763-71; PMID:16936727; <https://doi.org/10.1038/nsmb1141>
- [75] Mugnier N, Gueguen L, Vieira C, Biémont C. The heterochromatic copies of the LTR retrotransposons as a record of the genomic events that have shaped the *Drosophila melanogaster* genome. *Gene* 2008; 411:87-93; PMID:18281162; <https://doi.org/10.1016/j.gene.2008.01.010>
- [76] la Chaux de N, Wagner A. BEL/Pao retrotransposons in metazoan genomes. *BMC Evol Biol* 2011; 11:154; PMID:21639932; <https://doi.org/10.1186/1471-2148-11-154>
- [77] Nefedova LN, Mannanova MM, Kim AI. Integration specificity of LTR-retrotransposons and retroviruses in the *Drosophila melanogaster* genome. *Virus Genes* 2011; 42:297-306; PMID:21369828; <https://doi.org/10.1007/s11262-010-0566-4>
- [78] Khan E, Mack JP, Katz RA, Kulkosky J, Skalka AM. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res* 1991; 19:851-60; PMID:1850126
- [79] Mizrokhi LJ, Georgieva SG, Ilyin YV. Jockey, a mobile *Drosophila* element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. *Cell* 1988; 54:685-91; PMID:2842063; [https://doi.org/10.1016/S0092-8674\(88\)80013-8](https://doi.org/10.1016/S0092-8674(88)80013-8)
- [80] Carbonare BD, Gehring WJ. Excision of copia element in a revertant of the white-apricot mutation of *Drosophila melanogaster* leaves behind one long-terminal repeat. *Mol Gen Genet* 1985; 199:1-6; PMID:2987642
- [81] Kofler R, Betancourt AJ, Schlötterer C. Sequencing of pooled DNA samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in *Drosophila melanogaster*. *PLoS Genet* 2012; 8:e1002487; PMID:22291611; <https://doi.org/10.1371/journal.pgen.1002487>
- [82] Rizzon C, Marais G, Gouy M, Biémont C. Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. *Genome Res* 2002; 12:400-7; PMID:11875027; <https://doi.org/10.1101/gr.210802>
- [83] Nefedova LN, Liubomirskaya NV, Il'in IV, Kim AI. [Precise excision of long terminal repeats of the gypsy (mdg4) retrotransposon of *Drosophila melanogaster* detected in *Escherichia coli* cells is explained by its integrase function]. *Genetika* 2006; 42:1656-63; PMID:17326385
- [84] Kuzin AB, Lyubomirskaya NV, Khudaibergenova BM, Ilyin YV, Kim AI. Precise excision of the retrotransposon gypsy from the forked and cut loci in a genetically unstable *D. melanogaster* strain. *Nucleic Acids Res* 1994; 22:4641-5; PMID:7984412
- [85] Hoskins RA, Smith CD, Carlson JW, Carvalho AB, Halpern A, Kaminker JS, Kennedy C, Mungall CJ, Sullivan BA, Sutton GG, et al. Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome Biol* 2002; 3:RESEARCH0085; PMID:12537574; <https://doi.org/10.1186/gb-2002-3-12-research0085>
- [86] Bergman CM, Bensasson D. Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 2007; 104:11340-5; PMID:17592135; <https://doi.org/10.1073/pnas.0702552104>
- [87] Tamura K, Subramanian S, Kumar S. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol Biol Evol* 2004; 21:36-44; PMID:12949132; <https://doi.org/10.1093/molbev/msg236>
- [88] Ganko EW, Greene CS, Lewis JA, Bhattacharjee V, McDonald JF. LTR retrotransposon-gene associations in *Drosophila melanogaster*. *J Mol Evol* 2006; 62:111-20; PMID:16408244; <https://doi.org/10.1007/s00239-004-0312-4>
- [89] Merenciano M, Ullastres A, de Cara MAR, Barrón MG, González J. Multiple independent retroelement insertions in the promoter of a stress response gene have variable molecular and functional effects in *Drosophila*. *PLoS Genet* 2016; 12:e1006249; PMID:27517860; <https://doi.org/10.1371/journal.pgen.1006249>
- [90] Song SU, Gerasimova T, Kurkulos M, Boeke JD, Corces VG. An env-like protein encoded by a *Drosophila* retroelement: Evidence that gypsy is an infectious retrovirus. *Genes Dev* 1994; 8:2046-57; PMID:7958877; <https://doi.org/10.1101/gad.8.17.2046>
- [91] Kim A, Terzian C, Santamaria P, Pélisson A, Purd'homme N, Bucheton A. Retroviruses in invertebrates: The gypsy retrotransposon is apparently an infectious

- retrovirus of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 1994; 91:1285-9; PMID:8108403; <https://doi.org/10.1073/pnas.91.4.1285>
- [92] Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* 1990; 9:3353-62; PMID:1698615
- [93] Tanda S, Mullor JL, Corces VG. The *Drosophila* tom retrotransposon encodes an envelope protein. *Mol Cell Biol* 1994; 14:5392-401; PMID:8035817; <https://doi.org/10.1128/MCB.14.8.5392>
- [94] Malik HS, Henikoff S, Eickbush TH. Poised for contagion: Evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* 2000; 10:1307-18; PMID:10984449; <https://doi.org/10.1101/gr.145000>
- [95] Friesen PD, Nissen MS. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol Cell Biol* 1990; 10:3067-77; PMID:1692964; <https://doi.org/10.1128/MCB.10.6.3067>
- [96] Hansen LJ, Sandmeyer SB. Characterization of a positionally active Ty3 element and identification of the Ty3 integrase protein. *J Virol* 1990; 64:2599-607; PMID:2159534
- [97] Champion S, Maisonhaute C, Kim MH, Best-Belpomme M. Characterization of the reverse transcriptase of 1731, a *Drosophila melanogaster* retrotransposon. *Eur J Biochem* 1992; 209:523-31; PMID:1385119
- [98] Haoudi A, Rachidi M, Kim MH, Champion S, Best-Belpomme M, Maisonhaute C. Developmental expression analysis of the 1731 retrotransposon reveals an enhancement of Gag-Pol frameshifting in males of *Drosophila melanogaster*. *Gene* 1997; 196:83-93; PMID:9322744
- [99] Haoudi A, Kim MH, Champion S, Best-Belpomme M, Maisonhaute C. The Gag polypeptides of the *Drosophila* 1731 retrotransposon are associated to virus-like particles and to nuclei. *FEBS Lett* 1995; 377:67-72; PMID:8543022
- [100] Lin SS, Nymark-McMahon MH, Yieh L, Sandmeyer SB. Integrase mediates nuclear localization of Ty3. *Mol Cell Biol* 2001; 21:7826-38; PMID:11604517; <https://doi.org/10.1128/MCB.21.22.7826-7838.2001>
- [101] Moore SP, Rinckel LA, Garfinkel DJ. A Ty1 integrase nuclear localization signal required for retrotransposition. *Mol Cell Biol* 1998; 18:1105-14; PMID:9448008; <https://doi.org/10.1128/MCB.18.2.1105>
- [102] Maertens G, Cherepanov P, Debyser Z, Engelborghs Y, Engelman A. Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75. *J Biol Chem* 2004; 279:33421-9; PMID:15163664; <https://doi.org/10.1074/jbc.M404700200>
- [103] Meignin C, Bailly J-L, Arnaud F, Dastugue B, Vaury C. The 5' untranslated region and Gag product of Idefix, a long terminal repeat-retrotransposon from *Drosophila melanogaster*, act together to initiate a switch between translated and untranslated states of the genomic mRNA. *Mol Cell Biol* 2003; 23:8246-54; PMID:14585982; <https://doi.org/10.1128/MCB.23.22.8246-8254.2003>
- [104] Chapman KB, Byström AS, Boeke JD. Initiator methionine tRNA is essential for Ty1 transposition in yeast. *Proc Natl Acad Sci U S A* 1992; 89:3236-40; PMID:1314382; <https://doi.org/10.1073/pnas.89.8.3236>
- [105] Lin JH, Levin HL. Reverse transcription of a self-primed retrotransposon requires an RNA structure similar to the U5-IR stem-loop of retroviruses. *Mol Cell Biol* 1998; 18:6859-69; PMID:9774699; <https://doi.org/10.1128/MCB.18.11.6859>
- [106] Kapitonov VV, Tempel S, Jurka J. Simple and fast classification of non-LTR retrotransposons based on phylogeny of their RT domain protein sequences. *Gene* 2009; 448:207-13; PMID:19651192; <https://doi.org/10.1016/j.gene.2009.07.019>
- [107] Mukha DV, Pasyukova EG, Kapelinskaya TV, Kagramanova AS. Endonuclease domain of the *Drosophila melanogaster* R2 non-LTR retrotransposon and related retroelements: A new model for transposition. *Front Genet* 2013; 4:63; PMID:23637706; <https://doi.org/10.3389/fgene.2013.00063>
- [108] Luan DD, Korman MH, Jakubczak JL, Eickbush TH. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. *Cell* 1993; 72:595-605; PMID:7679954; [https://doi.org/10.1016/0092-8674\(93\)90078-5](https://doi.org/10.1016/0092-8674(93)90078-5)
- [109] Kurzynska-Kokorniak A, Jamburuthugoda VK, Bibillo A, Eickbush TH. DNA-directed DNA polymerase and strand displacement activity of the reverse transcriptase encoded by the R2 retrotransposon. *J Mol Biol* 2007; 374:322-33; PMID:17936300; <https://doi.org/10.1016/j.jmb.2007.09.047>
- [110] Stage DE, Eickbush TH. Origin of nascent lineages and the mechanisms used to prime second-strand DNA synthesis in the R1 and R2 retrotransposons of *Drosophila*. *Genome Biol* 2009; 10:R49; PMID:19416522; <https://doi.org/10.1186/gb-2009-10-5-r49>
- [111] Luan DD, Eickbush TH. RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element. *Mol Cell Biol* 1995; 15:3882-38; PMID:7540721; <https://doi.org/10.1128/MCB.15.7.3882>
- [112] Fujimoto H, Hirukawa Y, Tani H, Matsuura Y, Hashido K, Tsuchida K, Takada N, Kobayashi M, Maekawa H. Integration of the 5' end of the retrotransposon, R2Bm, can be complemented by homologous recombination. *Nucleic Acids Res* 2004; 32:1555-65; PMID:14999096; <https://doi.org/10.1093/nar/gkh304>
- [113] George JA, DeBaryshe PG, Traverse KL, Celniker SE, Pardue M-L. Genomic organization of the *Drosophila* telomere retrotransposable elements. *Genome Res* 2006; 16:1231-40; PMID:16963706; <https://doi.org/10.1101/gr.5348806>
- [114] Christensen SM, Ye J, Eickbush TH. RNA from the 5' end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site. *Proc Natl*

- Acad Sci U S A 2006; 103:17602-7; PMID:17105809; <https://doi.org/10.1073/pnas.0605476103>
- [115] Mason JM, Biessmann H. The unusual telomeres of *Drosophila*. *Trends Genet* 1995; 11:58-62; PMID:7716808; [https://doi.org/10.1016/S0168-9525\(00\)88998-2](https://doi.org/10.1016/S0168-9525(00)88998-2)
- [116] Abad JP, De Pablos B, Osoegawa K, De Jong PJ, Martín-Gallardo A, Villasante A. TAHRE, a novel telomeric retrotransposon from *Drosophila melanogaster*, reveals the origin of *Drosophila* telomeres. *Mol Biol Evol* 2004; 21:1620-4; PMID:15175413; <https://doi.org/10.1093/molbev/msh180>
- [117] Pardue ML, Danilevskaya ON, Lowenhaupt K, Slot F, Traverse KL. *Drosophila* telomeres: New views on chromosome evolution. *Trends Genet* 1996; 12:48-52; PMID:8851970; [https://doi.org/10.1016/0168-9525\(96\)81399-0](https://doi.org/10.1016/0168-9525(96)81399-0)
- [118] Fuller AM, Cook EG, Kelley KJ, Pardue M-L. Gag proteins of *Drosophila* telomeric retrotransposons: Collaborative targeting to chromosome ends. *Genetics* 2010; 184:629-36; PMID:20026680; <https://doi.org/10.1534/genetics.109.109744>
- [119] Pardue ML, Danilevskaya ON, Lowenhaupt K, Wong J, Erby K. The gag coding region of the *Drosophila* telomeric retrotransposon, HeT-A, has an internal frame shift and a length polymorphic region. *J Mol Evol* 1996; 43:572-83; PMID:8995054; <https://doi.org/10.1007/BF02202105>
- [120] Rashkova S, Karam SE, Pardue ML. Element-specific localization of *Drosophila* retrotransposon Gag proteins occurs in both nucleus and cytoplasm. *Proc Natl Acad Sci U S A* 2002; 99:3621-6; PMID:11891280; <https://doi.org/10.1073/pnas.032071999>
- [121] Shpiz S, Kwon D, Rozovsky Y, Kalmykova A. rasiRNA pathway controls antisense expression of *Drosophila* telomeric retrotransposons in the nucleus. *Nucleic Acids Res* 2009; 37:268-78; PMID:19036789; <https://doi.org/10.1093/nar/gkn960>
- [122] Petit N, Piñeyro D, López-Panadès E, Casacuberta E, Navarro A. HeT-A_pi1, a piRNA target sequence in the *Drosophila* telomeric retrotransposon HeT-A, is extremely conserved across copies and species. Lalueza-Fox C, editor. *PLoS One* 2012; 7:e37405; PMID:22629389; <https://doi.org/10.1371/journal.pone.0037405>
- [123] Shpiz S, Ryazansky S, Olovnikov I, Abramov Y, Kalmykova A. Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the *drosophila* germline. *PLoS Genet* 2014; 10:e1004138; PMID:24516406; <https://doi.org/10.1371/journal.pgen.1004138>
- [124] Malik HS, Eickbush TH. Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. *J Virol* 1999; 73:5186-90; PMID:10233986
- [125] Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev* 2006; 20:2214-22; PMID:16882972; <https://doi.org/10.1101/gad.1454806>
- [126] Ågren JA, Wright SI. Co-evolution between transposable elements and their hosts: A major factor in genome size evolution? *Chromosome Res* 2011; 19:777-86; PMID:21850458; <https://doi.org/10.1007/s10577-011-9229-0>