Review

Transposable elements in natural populations of *Drosophila melanogaster*

Yuh Chwen G. Lee and Charles H. Langley*

Department of Evolution and Ecology, University of California - Davis, Davis, CA 95616, USA

Transposable elements (TEs) are families of small DNA sequences found in the genomes of virtually all organisms. The sequences typically encode essential components for the replicative transposition sequences of that TE family. Thus, TEs are simply genomic parasites that inflict detrimental mutations on the fitness of their hosts. Several models have been proposed for the containment of TE copy number in outbreeding host populations such as *Drosophila*. Surveys of the TEs in genomes from natural populations of *Drosophila* have played a central role in the investigation of TE dynamics. The early surveys indicated that a typical TE insertion is rare in a population, which has been interpreted as evidence that each TE is selected against. The proposed mechanisms of this natural selection are reviewed here. Subsequent and more targeted surveys identify heterogeneity among types of TEs and also highlight the large role of homologous and possibly ectopic crossing over in the dynamics of the *Drosophila* TEs. The recent discovery of germline-specific RNA interference via the piwi-interacting RNA pathway opens yet another interesting mechanism that may be critical in containing the copy number of TEs in natural populations of *Drosophila*. The expected flood of *Drosophila* population genomics is expected to rapidly advance understanding of the dynamics of TEs.

Keywords: ectopic exchange; piRNA; natural selection; RNA interference; epistasis

1. INTRODUCTION

Transposable elements (TEs) are distinct populations of short sequences in the genomic DNA of all cellular organisms (hosts) that survive and reproduce by 'coding' replicative functions that can increase their numbers within their hosts' genomes in the next generation (figure 1). The vast majority of the well-studied 'families' of TEs fall into two distinct classes: those that replicate via an RNA intermediate using reverse transcriptase and those that replicate by excision, insertion elsewhere in the host genome and replicative repair of the excised copy using the sister chromatid as a template. Within each of these two broad classes of TEs, many variations on these two fundamental modes of replicative transposition have evolved.

TEs are widely distributed in genomes of both prokaryotes and eukaryotes. Typically, the genomes of multicellular eukaryotes are populated by many distinct families of TEs. Members of the same TE families usually have similar sequences and transpose via the same mechanism. Sequences derived from TEs constitute a substantial fraction of genomes of most eukaryotes and the presence or absence of various TE families as well as their distributions among individuals of the same species are well documented

in model species such as *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Mus musculus* as well as *Homo sapiens*. Even though researchers have found examples of TE insertions correlated with positive selective sweeps in specific environments (Daborn *et al.* 2002; Aminetzach *et al.* 2005), the great majority of TE insertions have negative impact on host fitness (Mackay 1989; Pasyukova *et al.* 2004) and are widely viewed as 'genomic parasites'. TE insertions in or near genes often change gene expression that result in phenotypic changes of the hosts. Because of this mutagenic property, TEs have been used as efficient mutagens in most model organism systems. For example, in *Zea mays*, large numbers of spontaneous point mutations and chromosomal rearrangements have been associated with 'mobile genetic elements' (McClintock 1950; Feschotte *et al.* 2002). Also, developmentally unstable or variegating mutations have been associated with TE excision.

A typical multicellular eukaryotic host genome contains many copies of apparently transpositionally inactive (defective) copies that became fixed in every genome of the entire host population in the distant past. Some TE families have large populations of mutant nonautonomous copies that depend on the presence of wild-type (autonomous) copies to transpose. Phylogenetic analyses suggest that horizontal transfer (from one host species's genome to that of another) does occur, but is quite rare. The appearance and worldwide spread of the *P* element in *D. melanogaster* is perhaps the best-studied case (Daniels *et al.* 1990). Thus, the inferred most recent common ancestor of

* Author for correspondence (chlangley@ucdavis.edu).

elements with
RNA replicative
intermediates

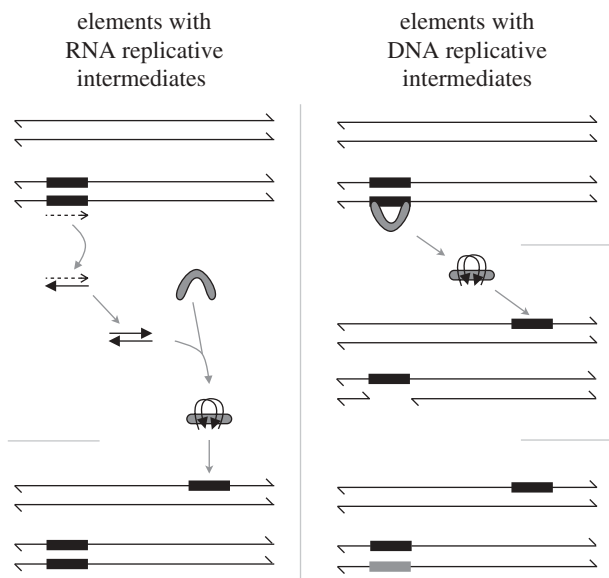elements with
DNA replicative
intermediates

Figure 1. The two modes of replicative transposition of TEs in eukaryotes. The left panel shows the mechanism of retrotransposition in which the TE genome is transcribed into RNA (dashed) and then reverse transcribed into DNA and integrated at a new location by proteins (grey fill), some of which are coded in the TE genome. On the right is a sketch of the mechanism typically used by the second class of TEs in which the double-stranded TE genome is excised from one of the two sister chromatids by the TE-encoded transposase protein (grey fill) that also catalyses the integration elsewhere in the genome of the host's germline. The double-strand break left by the excision is repaired off by the sister chromatid yielding a net increase in copy number. The grey-filled rectangular blocks are the replicated copies of the TEs. The rounded, grey-filled structures depict typical roles of TE proteins in reverse transcription, integration and excision.

all members (defective and active) of a TE family can be quite ancient, while the actively transposing subpopulations of a typical TE family tend to have relatively similar sequences and therefore a much more recent inferred common ancestor (perhaps the best-studied example is the human L1 TE: Sassaman *et al.* 1997; Ostertag & Haig 2001).

Theoretical modelling of TE dynamics in outbreeding host populations has yielded a number of interesting and testable hypotheses. Perhaps the most significant feature of these models is the high fidelity of the vertical transmission in the hosts' genomes. Since each copy of the TE is replicated and transmitted via sexual reproduction to progeny as an integral segment of the host genome, stochastic loss of all transpositionally competent copies of a TE family from all individuals in the host population is extremely unlikely (Kaplan *et al.* 1985). Furthermore, modelling of the evolution of self-regulation indicates that TEs will not evolve mechanisms of a 'prudent predator', but rather it is copy-number-dependent natural selection that must impinge on the TE dynamics if the copy number is to be *contained* (Charlesworth & Langley 1986; see below). Several proposed selective mechanisms for the 'containment' of TE copy number have been analysed and tested experimentally. While a model based on

recombination ('ectopic exchange') leading to strong selection against aneuploid progeny has survived several empirical tests and gained acceptance, there is good evidence that many TE insertions are removed from the population by natural selection on the deleterious phenotypes caused by their insertion into the host's genome. Two recent developments indicate that significant progress on this topic is expected in the near future. First, technological advances in genomic sequencing make it possible to sequence population samples of host genomes, exposing in *complete* detail the numbers, positions and sequences of the TEs in independently sampled hosts. Second, the emerging understanding of RNA interference in the transcriptional expression and replicative transposition of TEs is leading to new functional insights about mechanisms of TE copy-number containment as discussed below.

Drosophila melanogaster is, of course, a central model organism system for the study of molecular, cellular, developmental and physiological genetics and has long been a system of choice for those investigating fundamental forces shaping variation in outbreeding, natural populations of animals. Because of its rich knowledge base and active research community, *Drosophila* has also played a central role in the study of the dynamics and divergence of TEs in natural populations. Beyond the diverse genetic tools available (markers, balancers, cloned TEs, etc.), *Drosophila* offered the early investigators of the population genetics of TEs two rare and critical advantages. First (for reasons directly related to the topic of this paper), the *Drosophila* genome has not accumulated great numbers of ancient, fixed copies of TEs in the gene-rich, *euchromatic* portion (69%) of its genome. This pattern is dramatically different from the high population frequency of most of the euchromatic TE insertions observed in the human genome. A plausible reason for this discrepancy is that the effective population size of humans is much smaller than that of *D. melanogaster* and the effective population size has an important role in the theory for the population dynamics and distribution of TEs (see below). Second, the technical ability to physically map DNA sequences, such as TEs, on the salivary gland chromosomes of individual *Drosophila* larvae provided an efficient (albeit low-resolution) method to characterize population genetic variation in genomic positions occupied by TEs (figure 2).

Before addressing models of the dynamics of TEs, it is useful to review the basic empirical genomics of *Drosophila* TEs: first in the reference sequence and then in early surveys of population variation. The reference genome of *D. melanogaster* currently provides the most detailed picture of the distribution of TEs in the genome of a single, inbred laboratory strain (Kaminker *et al.* 2002; Bergman *et al.* 2006). There are 121 TE families and 5390 elements annotated in the most recent analysis, together accounting for 5.5 per cent of the genome. Pericentromeric regions of the five large chromosome arms and the small fourth chromosome exhibit a high proportion of TE-derived sequences (Bergman *et al.* 2006). The great majority of sequences derived from TEs in
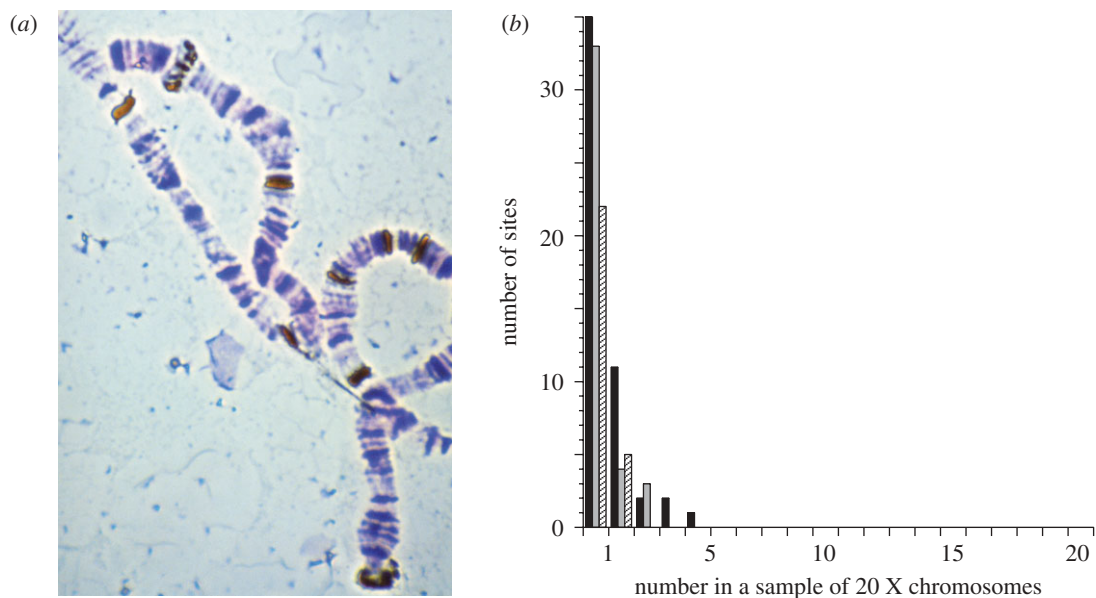
Figure 2. Cytological detection of TEs on *Drosophila* giant salivary gland chromosomes and their frequency spectrum in a sample from a natural population. (*a*) The photomicrograph shows portions of the X and 3R chromosomes in an inbred line from a natural population. The intensely dark bands are sites labelled by histochemically detectable *in situ* hybridization of biotinylated sequences of the *roo* TE family. (*b*) The graph is of the frequency spectra of three TE families among 20 independently sampled X chromosomes from a natural population of *D. melanogaster* (redrawn from Montgomery & Langley 1983). Black bar, 297; grey bar, 412; striped bar, copia.

these pericentromeric regions are found to be incomplete and nested within other TEs. In contrast, more of the complete copies of TEs are found in the gene rich euchromatic regions and are rarely nested (Bergman *et al.* 2006). Only a quarter of TEs are found in regions annotated as transcribed, which accounts for half of the genome, and none appear to be in coding regions of genes (Kaminker *et al.* 2002).

Surveys of population genetic variation in TEs in natural populations of *Drosophila* have taken three forms. The most recent builds on the clear picture from the reference sequence and simply asks which if any of the TEs identified in the sequenced laboratory strain can be found in genomes sampled from natural populations. González *et al.* (2008) used polymerase chain reaction to survey the presence versus absence in six pools of eight to 12 independent strains each. Three hundred and forty-four of the 902 euchromatic TE insertions in the reference genome were not observed in the pools, while 76 per cent of the 902 were observed to be present in low frequency in at least one or more of six surveyed populations. Because of the pooling design, it is difficult to determine if these observations are inconsistent with cytological and southern blot findings (discussed below) that most TE insertions are rare.

While it is possible that forces acting throughout the establishment and maintenance of the sequenced laboratory strain could have yielded the pattern of TEs observed in the reference genome, surveys of TEs in genomes recently sampled from natural populations indicate that the reference strain is not unusual. A number of studies of the distributions of particular TE families were based on cytological localization of *in situ* hybridized labelled TE DNA to giant salivary gland chromosomes (figure 2) of independently

sampled genomes from natural populations of *D. melanogaster* (reviewed by Charlesworth & Langley 1989; Charlesworth & Lapid 1989; Charlesworth *et al.* 1992). Most sites appear to be occupied in only one of the sampled strains (figure 2*b*). While there is certainly a subset of TE insertions that are sampled recurrently by *in situ* surveys, the statistical estimation and inference based on the apparent recurrent occupancy (the frequency spectrum) is confounded and limited by variation in occupancy per base pair below the scale of the cytological resolution (Kaplan & Brookfield 1983).

Restriction-map variation in numerous gene regions of *Drosophila* has been found to be consistent with the results of the *in situ* surveys. In an early survey of restriction-map variation using southern blots in the 16 kbp region surrounding the *Adh* gene, Aquadro *et al.* (1986) observed that all the large insertions (larger than 340 bp) were in fact TE sequences and that they were individually rare if not unique in their sample of 48 alleles. Even the several insertions of the same TE family in a small fragment (approx. 200 bp) proved to be unique (unpublished). Since then many (albeit 'gene-centric') regions of the *D. melanogaster* genome have been surveyed for restriction-map variation. Consistent with the results from the 'reference sequence centric' and the cytological approaches, this method shows that large insertions (likely to be TEs) are typically rare if not unique in small (less than 50) samples (e.g. Miyashita & Langley 1988; Long *et al.* 1998). Most exceptions are observed in pericentromeric regions where crossing over is restricted. The few exceptions in typical euchromatic regions (with normal levels of crossing over per base pair) are, of course, of great inherent interest simply because the recent population genetic dynamics of

these genomic regions is likely to be associated with natural selection acting on variation at or near the TE insertion. This skewed frequency spectrum of euchromatic TE insertions is consistent with the predictions of a number of models in which the selection and/or excision reach equilibrium with replicative transposition such that new insertions are removed from the population well before they can reach intermediate frequencies owing to genetic drift.

2. DYNAMIC MODELS—TRANSPOSITION, EXCISION AND NATURAL SELECTION

In a seminal paper, Charlesworth & Charlesworth (1983) presented a general dynamic model of TE copy-number evolution in terms of transposition, excision and natural selection. The simple and elegant model assumed that, in the germline, each TE replicatively transposes at rate u_n and excises at rate v_n. Selection against diploid hosts is the second force removing TEs from the population. The simplest form of natural selection assumes the fitness of a host, w_n, is a decreasing function of n, the number of TEs of a given family in a diploid host's genome. If mating is random and recombination is sufficiently great, the distribution of TEs will be close to linkage equilibrium. Furthermore, it is reasonable to assume that the number of possible insertion sites in a diploid genome, T, is large compared with the average copy number per genome, $\bar{n}$. Charlesworth & Charlesworth (1983, also see Charlesworth 1985) showed that the approximate change in average copy number per genome per generation under these conditions is given by the following equation:

$$\Delta\bar{n} \approx \bar{n}\left(1 - \frac{\bar{n}}{T}\right)\frac{\partial \ln w_{\bar{n}}}{\partial \bar{n}} + \bar{n}(u_{\bar{n}} - v_{\bar{n}}). \quad (2.1)$$

In the absence of copy-number-dependent regulation of transposition, u is independent of n and is taken as a constant per TE. A similar assumption applies to excision, v. The equilibrium average copy number, $\hat{n}$, is given by setting the above equation to zero. In order for this equilibrium to be stable:

$$-\frac{1}{\hat{n}} < \frac{\partial^2 \ln w_{\hat{n}}}{\partial \hat{n}^2} < 0 \quad (2.2)$$

(Charlesworth & Charlesworth 1983), and therefore the logarithm of fitness declines more rapidly than linearly with average copy number, $\bar{n}$ (e.g. the stippled concave curve of figure 3, also see Brookfield 1982).

The frequency distribution of each TE insertions in a population is obtained by extending the above model to finite-sized populations (Charlesworth & Charlesworth 1983; also see Langley et al. 1983). The steady-state distribution of TE frequency is approximated by a β distribution with the pertinent parameters defined as follows:

$$\phi(x) \approx \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}, \quad (2.3)$$

$$\alpha = \frac{4N_e u\hat{n}}{T - \hat{n}}, \quad (2.4)$$
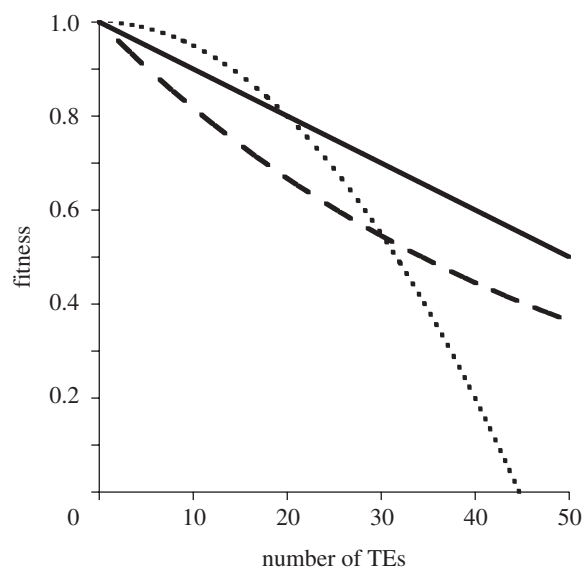
$$\beta = 4N_e(v + s_{\hat{n}}), \quad (2.5)$$

Figure 3. Three forms of interaction among deleterious mutations. In the *additive model* (unbroken line), the fitness decline due to deleterious mutations is the sum of the effects, e.g. $(1 - ns)$. Under the *synergistic epistasis model* (stippled, concave curve), as the number of mutations increases, the deleterious interactions grow stronger, resulting in an ever faster decline in fitness, e.g. $(1 - n^2 s)$. In the *multiplicative model* (dashed convex curve), the deleterious effects of mutations are weaker in individuals with greater numbers of other deleterious mutations, $(1 - s)^n$.

$$s_{\hat{n}} = \frac{\partial \ln w_{\bar{n}}}{\partial \bar{n}} \quad \text{at } \bar{n} = \hat{n}. \quad (2.6)$$

The parameter α describes the effects of genetic drift and transposition while β captures the combined effect of forces removing TEs (excision and selection). As figure 4 shows, the expected frequency of individual TE insertions varies substantially depending on the magnitudes of α and β, and thus u, v, T, $\hat{n}$ and $s_{\hat{n}}$. Most TE insertions in euchromatic regions of the *D. melanogaster* genome have been observed to have low frequencies (see above). This pattern led to estimates of β that are substantially greater than one (Kaplan & Brookfield 1983; Charlesworth & Langley 1989; Charlesworth & Lapid 1989; Charlesworth et al. 1992).

Selection against TEs can result from the well-established fact that TE insertions in or near genes frequently change gene structure or expression (deleterious insertion model). If this mutational effect of TE insertions is synergistic (figure 3) so that the fitness decline increases faster than linearly with copy number, a stable equilibrium is possible. While synergistic epistasis in the fitness effects of mutations has theoretical appeal and support in the results and interpretation of some mutation accumulation experiments (Mukai 1968), it remains controversial (Fry 2004). If the fitness effects of TE insertions are similar to those of other deleterious mutations, the criterion for stable copy-number equilibrium may well be met. However, the genetic and molecular mechanisms of epistatic interactions between such TE-associated mutations are not well understood. Whether these deleterious fitness effects of TEs are sufficiently
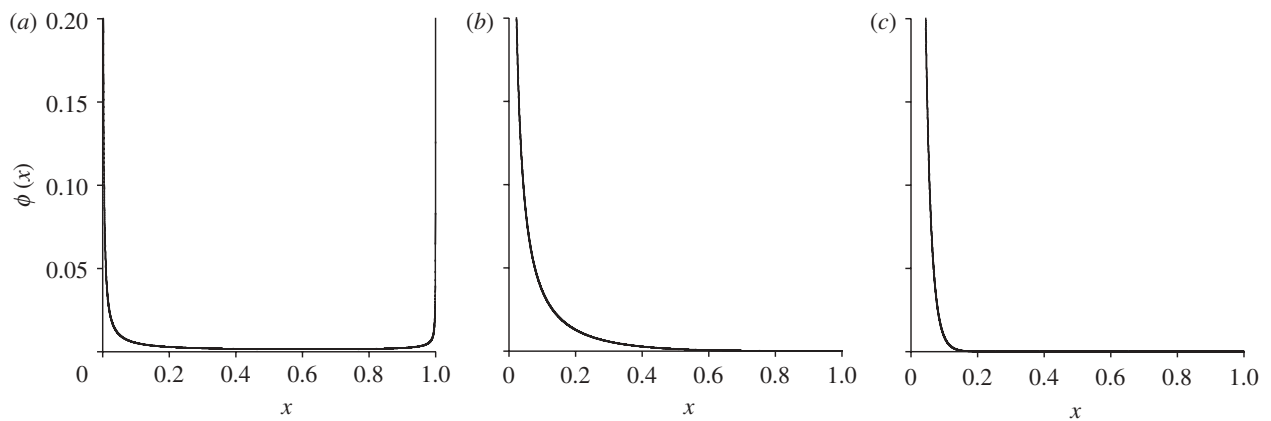
Figure 4. Predicted frequency spectra of TE insertions. If the number of possible insertion sites in the genome, T, is much larger than TE copy number ($n \ll T$) and $\alpha < 1$, the frequency spectrum is dominated by β (see text and equations (2.4)–(2.6) for definitions). Note that at equilibrium $\beta = 4N_e(v + s_n) = 4N_e u$. These three figures depict the impact of β as population size, N_e, goes from 10^4 ($\beta = 0.4$) to 10^5 ($\beta = 4$) and then to 10^6 ($\beta = 40$). The probability a TE insertion will occur at intermediate or high frequencies increases dramatically. Other assumed parameters are $u = 10^{-5}$, $T = 4 \times 10^4$ and $\hat{n} = 50$. (a) $\alpha = 0.0005$, $\beta = 0.4$; (b) $\alpha = 0.005$, $\beta = 4$ and (c) $\alpha = 0.05$, $\beta = 40$.

synergistic to contain the growth of TE copy number is still an open question.

Given the low transposition rates observed for *Drosophila* TEs, approximately 10^{-4} per copy per generation (Nuzhdin & Mackay 1995; Maside *et al.* 2001, 2002), the selection coefficient on each TE insertion is expected to be of a similar order of magnitude at equilibrium. However, the estimated fitness effect of naturally occurring heterozygous deleterious mutations is around 1–2%, which is too strong to maintain TEs in the population (Charlesworth & Langley 1989; Charlesworth 1991). Charlesworth (1991) relaxed the assumption of homogeneous selection coefficients against TE insertions to investigate the possibility of reaching a reasonable equilibrium copy number. When two classes of sites are assumed, selected and neutral sites, the equilibrium is very sensitive to the level of excision. In most of the parameter space, TEs accumulate to unrealistically high frequency in the neutral sites. Accumulation of TEs in the neutral sites is also predicted when the model is extended by allowing the strength of selection to vary at the selected sites, either strongly or weakly selected sites. Biologically reasonable equilibria are reached under the scenarios when copies in neutral sites are transpositionally suppressed (perhaps relevant to TE insertions in heterochromatin) or when there are no neutral sites. If the fitness effects of TE insertions are partially recessive, the equilibrium density on the X chromosome is expected to be reduced relative to that on the autosomes. But this was not well supported by the survey data (Montgomery *et al.* 1987; Langley *et al.* 1988). These theoretical and empirical investigations suggest that selection against the mild zygotic fitness TE insertions may not be the major force involved in TE copy-number containment, thus motivating investigations of alternative models, some with specific predictions for the genomic distributions of TEs (see below).

The basic TE selection model above assumes that the density of TEs per cM is sufficiently low (and

the frequency spectrum sufficiently skewed) that linkage disequilibrium can be ignored. However, it is widely observed that *Drosophila* TEs accumulate at regions with low meiotic crossing over per base pair (Bartolome *et al.* 2002; Kaminker *et al.* 2002; Rizzon *et al.* 2002; Bergman *et al.* 2006). *In situ* hybridization surveys routinely observe TE sequence accumulation in centromere proximal regions with lowered crossing over (Rubin 1983; reviewed by Charlesworth & Langley 1989; Charlesworth & Lapid 1989; Charlesworth *et al.* 1992), on the fourth chromosome (Bartolome & Maside 2004), and near the breakpoints of polymorphic chromosome inversions (Eanes *et al.* 1992; Sniegowski & Charlesworth 1994). An important obvious exception is the low density of TEs near the telomeres where crossing over is also strongly suppressed (see below). Natural selection against deleterious mutations can strongly influence the dynamics of linked variation, a phenomenon known as background selection (Charlesworth *et al.* 1993). In genomic regions where the crossing over per gene is lower, background selection can reduce the equilibrium levels of neutral polymorphism. Charlesworth (1996) incorporated TEs as an additional source of deleterious mutation. Even though the selection coefficient against TE insertions is relatively weak compared with that estimated for heterozygous deleterious mutations, he found the model incorporating both deleterious mutations and TE insertions accounts well for the lower genetic diversity in genomic regions of low crossing over per base pair, such as the distal X chromosome, the pericentromeric regions of major chromosomes (X, 2L, 2R, 3L, 3R) and the fourth chromosome.

Another prediction of selection against deleterious mutations is that the efficacy of selection is lowered at linked sites, the Hill–Robertson effect (Hill & Robertson 1966). In regions with restricted crossing over, interference with the natural selection against TEs owing to their deleterious effects on zygotic fitness should result in relatively higher rates of
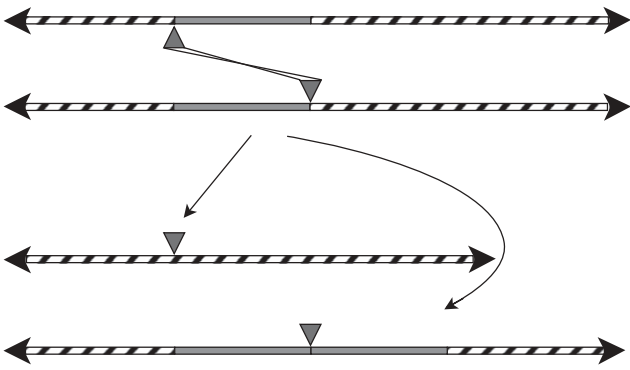
Figure 5. Ectopic exchange leading to a duplicated and deleted chromosome. TEs at different locations along homologous chromosomes can exchange during meiosis to yield daughter chromosomes in the gametes that are complementarily duplicated (grey regions) and deleted. Single 'ectopic' exchanges between nonhomologous chromosomes or in an antiparallel direction produce even greater aneuploidy in the gametes.

accumulation of TEs in low crossing over regions. However, the multilocus dynamics under the Hill–Robertson effect afford little in the way of analytic predictions. Dolgin & Charlesworth (2008) conducted extensive Monte Carlo simulations to investigate the relative level of accumulation of TEs in regions of the genome with different levels of crossing over. The accumulation of TEs in regions with low crossing over was found to be confined to a surprisingly small part of the parameter space: where the synergistic fitness effects between TEs are weak, the host population size is small, the rate of crossing over is extremely low and the excision rate is essentially zero. Natural populations of *Drosophila* are not expected to meet this narrow specification. In contrast, the necessary conditions for the alternative model based on ectopic exchange (see below) are more liberal and may well accommodate the observed patterns.

3. THE ECTOPIC EXCHANGE MODEL

Rearrangements in the genome such as deletions, duplications, inversions and translocations were long suspected to arise through recombination involving nonhomologous but related genomic elements (figure 5). Molecular genetic investigations in a number of systems have born this out and the alternative term 'ectopic' recombination (gene conversion and exchange) was introduced (Lichten *et al.* 1987). It was hypothesized that TE copy number might be contained by selection against the aneuploid progeny produced by ectopic exchange (Langley *et al.* 1988). In *Drosophila*, diverse experiments from Sturtevant's (1925, 1929) classic *Bar* duplication studies to the characterization of putative duplications arising in early intralocus recombination studies at *white* (Green 1959; Judd 1959; Goldberg *et al.* 1983; Davis *et al.* 1987) fostered this hypothesis. Indeed, a systematic study of ectopic exchange indicated that it could be the most common mutation process in female *Drosophila* (Montgomery *et al.* 1991).

As laid out by Langley *et al.* (1988), the impact of the selection against aneuploid progeny is a straightforward extension of the model put forward by Charlesworth & Charlesworth (1983). The fitness reduction of individuals as a function of the TE copy number is:

$$\sum_{l \leq k \leq \mathcal{J}} \kappa_{kl} n_k n_l,$$

where κ_{kl} is the rate of exchange between copies at positions (in the $\mathcal{J}$ regions) k and l, and

$$\frac{\partial \ln w_{\bar{n}}}{\partial \bar{n}} = -\sum_{k=1}^{\mathcal{J}} \kappa_{kl} \bar{n}_k. \tag{3.1}$$

At equilibrium, the ratio of densities of TE copy in two chromosomal regions is a function of the ratio of ectopic exchange rates of two regions. This predicts that TEs tend to accumulate in genomic regions with low ectopic exchange rate. If ectopic exchange rates vary across the genome in parallel with normal crossing over, then the widely observed accumulation of TEs in centromere proximal regions (see above) is consistent with the model. The low density of TEs in the telomeric regions (see above) suggests that the rate of ectopic exchange may be high even though regular homologous exchange is clearly suppressed in this region.

Since the X chromosome experiences proportionally more meiotic recombination, ectopic exchange might be more effective, but the X is also less available as a target for insertion. And the impact of zygotic selection on partially recessive deleterious effects of TE insertions should be more severe on the X than the autosomes, since they would be hemizygous every three generations. Langley *et al.* (1988) fitted a simple model incorporating both zygotic selection and ectopic exchange (on the X and autosomes) and found that the data fitted acceptably well. Two high-copy-number families showed no evidence of zygotic selection against the mutant phenotypes associated with the TE insertions, while the one low-copy-number family did exhibit a significantly lower TE number on the X, consistent with strong hemizygous selection on the X. But this was not a powerful test of the ectopic exchange model. The most elegant test was to compare the density of TEs near the breakpoints of polymorphic inversions to regions far away. Homologous crossing over is known to be strongly suppressed around the breakpoint of rearrangement in heterokaryotypes suggesting that the ectopic exchange might also be inhibited. Sniegowski & Charlesworth (1994) found that indeed the densities of TEs are higher in chromosome regions proximal to the inversion breakpoints. They concluded that the observed patterns of numbers of TEs and their frequency spectra are inconsistent with Muller's ratchet. Sniegowski & Charlesworth (1994) favour the suppression of ectopic exchange between TEs in inversion heterozygotes as the most plausible explanation.

Another important prediction of the ectopic exchange hypothesis for containment of TE copy number is that incomplete copies of the TEs are expected to be less likely to participate in ectopic exchange (Petrov *et al.* 2003) and therefore these smaller members of the TE family will persist longer in the

population. Their results show a significant negative correlation between TE population frequency and their current size. This observation holds for comparisons between TEs from different families as well as TEs from the same family. Petrov *et al.* also observed that TE families with higher genomic copy numbers have frequency spectra with more skew towards rare TE insertions. This is consistent with the expectation that the strength of ectopic exchange at removing TEs grows quadratically with copy number.

4. PIWI-INTERACTING RNA AS A COMPETING MODEL FOR CONTAINMENT

The other possible mechanism for the containment of TE copy number is the regulation of transposition rate, i.e. the transposition rate u_n decreases as the copy number n increases (Charlesworth & Charlesworth 1983; Langley *et al.* 1983). Equilibrium is reached when the transposition rate equals the excision rate. Because most TE insertions appear to be deleterious to the host, it is natural to postulate that TE variants that can repress the transposition rate will be selected because of the relative reduction in the number of deleterious mutations among their hosts' progeny. Charlesworth & Langley (1986) presented a theoretical analysis that showed such self-regulatory TE variants would be unlikely to evolve in an outbreeding and meiotically recombining host, simply because of the weak dynamic coupling (owing to random assortment and recombination) of genetic variation in the self-regulation with the associated mutational effects. Transposition immunity, in which TEs with self-regulatory ability suppress transposition into their genomic neighbourhood, has been observed in prokaryotes but not yet in eukaryotes (Craig 2002). Observations of transposition rate decreasing with increasing copy numbers of *Drosophila* *P* and *I* elements, support the possibility of transposition repression (Bucheton *et al.* 2002; Rio 2002). However, recent discoveries reveal that this regulation of transposition rate may indeed be host mediated.

Several independent investigations indicate that TE transposition rate may be regulated through a newly discovered pathway that includes host-generated small RNA, the so-called piwi-interacting RNA (piRNA), in the *Drosophila* germline (Vagin *et al.* 2006; Brennecke *et al.* 2007; Gunawardane *et al.* 2007; reviewed in Aravin *et al.* 2007). piRNAs are small RNAs (23–29 bp) and found to be enriched with TE sequences, which leads to the hypothesis of their role in the regulation of TEs. piRNAs are associated with the Piwi clade of the Argonaute protein family, including PIWI, AUB (coded by the *Aubergine* locus) and AGO3 (*Argonaute-3*), which are expressed mainly in the *Drosophila* germline. Loss of function mutations in *Piwi* and other genes involved in the piRNA generation pathway disrupt the generation of piRNAs, which leads to increases in expression and/or transposition in several TE families (Sarot *et al.* 2004; Kalmykova *et al.* 2005; Savitsky *et al.* 2006; Chambeyron *et al.* 2008; Li *et al.* 2009; Malone *et al.* 2009). The 'ping-pong' model depicted in figure 6 has been proposed to explain the generation and
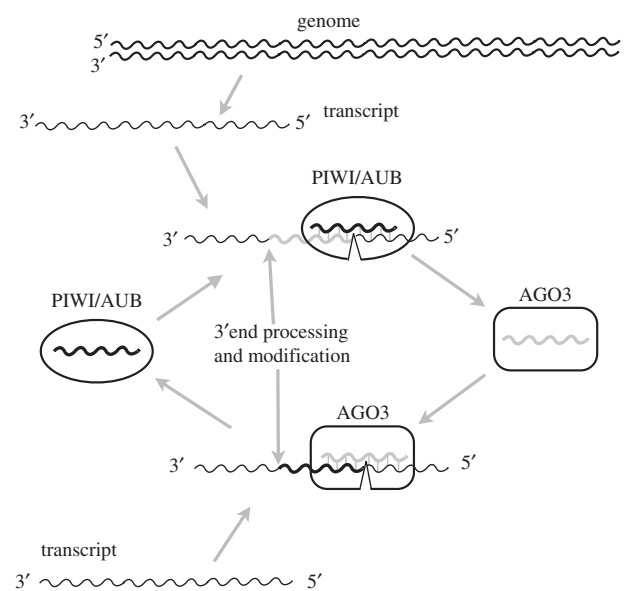

Figure 6. The ping-pong model: piRNA generation process is coupled with RNA interference mediated by *Piwi* proteins (PIWI, AUB and AGO3) (redrawn from fig. 2 of Aravin *et al.* 2007). PIWI and AUB preferentially bind to antisense piRNAs. Guided by piRNAs, PIWI and AUB cleave TE transcripts having complementary sequences to the piRNA sequences, which lead to the inactivation of TE transcript as well as the generation of new sense piRNAs. AGO3 mainly binds sense piRNAs to target antisense TE transcripts, generate new antisense piRNA and complete the other half of the amplification cycle. The source of the primary piRNAs that initiate the amplification cycle may be maternally inherited (Blumenstiel & Hartl 2005; Brennecke *et al.* 2008; Malone *et al.* 2009). Recent studies suggest a modified picture, namely that PIWI is responsible for piRNA generation through another pathway in the somatic tissue of *Drosophila* ovaries, while AUB and AGO3 are involved in the above ping-pong model in germline tissues (Li *et al.* 2009; Malone *et al.* 2009).

amplification of piRNAs in the *Drosophila* (Brennecke *et al.* 2007; Gunawardane *et al.* 2007).

Piwi proteins are guided by piRNAs to RNA or DNA targets having complementary sequence to piRNAs. Targeted TE RNAs are cleaved by Piwi-family proteins (Saito *et al.* 2006; Gunawardane *et al.* 2007). Another proposed mechanism is the formation of heterochromatin guided by the small RNAs to the targeted genomic regions containing TE-derived sequences (Grewal & Elgin 2007; Klenov *et al.* 2007; Slotkin & Martienssen 2007). In this compact chromatin environment, targeted sequences are expected to be less accessible to the transcription machinery. Through either mechanism, the expression potential of the TEs is suppressed. We propose that the piRNA pathway not only suppresses the expression of TE RNAs in the germlines of individuals, but also may well lead to containment of TE copy number in the natural outbreeding host populations, such as *Drosophila*. This model based on piRNA interference shares two critical elements with the ectopic exchange model that make TE copy-number containment through piRNA such an attractive and robust model for TE copy-number containment. First, piRNA depends directly on sequence homology (not

phenotypes expressed via translation). The mechanism is thus general to all families and independent of any specific TE phenotype (other than transcription). Secondly, the ping-pong amplification of piRNAs should grow quadratically or even exponentially with TE copy number (figure 6) and thus potentially provide the built-in synergism necessary to strongly stabilize the distribution of TE copies in host populations. Definitive studies directly linking piRNA to copy-number-dependent transposition rates are still required. A study by Jensen *et al.* (1999) can be taken as indirect evidence that transposition rate decreases as TE copy number increases. They introduced *I* elements into *D. melanogaster* genomes devoid of such elements and showed that the transposition activity decreases as the number of introduced elements to the genome increases. Their results demonstrated that the repression of *I* element transposition depends on the transcription of, and the sequence homology between, elements, but not on translation. This is consistent with properties of transposition rate regulation through piRNAs. Even though some studies suggest the possibility that some piRNAs are mainly generated from a subset of TE copies (Brennecke *et al.* 2007), there are not enough data to be conclusive about whether every TE copy contributes differently to the piRNA pool.

We note that the TEs that are diverged from the dominant sequences in the piRNA pool are expected to enjoy less suppression and thus more replicative transposition. Also, several studies of the molecular evolution of genes in the piRNA pathway have reported evidence for adaptive evolution (Heger & Ponting 2007; Obbard *et al.* 2009). The divergence-based escape to these two sequence-similarity-based mechanisms (piRNA pathway and ectopic exchange) has the potential to drive an additional TE–host coevolutionary arms race directly at the TE sequences.

The suppression of TE expression through piRNA can only explain how the transposition rate is regulated. However, this does not account for the possible evolutionary forces removing TEs and the observations that most of the TE insertions are rare in a population. As mentioned above, piRNA has been proposed to suppress expression for TEs through heterochromatin formation of the genomic sequences containing TEs. When a euchromatic gene is translocated near heterochromatin, the propagation of heterochromatin marks from the nearby heterochromatin frequently results in the unstable expression of the gene, a phenomenon known as position effect variegation (PEV; see Talbert & Henikoff 2006). It was shown that an artificially introduced tandem array of TEs can initiate heterochromatin formation within the euchromatic regions, which results in PEV of the nearby reporter gene in *Drosophila* (Dorer & Henikoff 1994) and naturally occurring PEV caused by TE insertions near genes have been documented (reviewed in Slotkin & Martienssen 2007). We thus propose that PEV caused by TEs on nearby genes may provide another plausible form of zygotic selection against TEs. The amplification of piRNAs once again provides the necessary synergism of stable equilibrium in the dynamics of TE containment.

5. CONCLUSIONS

The dynamics of TEs in natural populations of *Drosophila* can be studied effectively via comparisons of predictions of models to population samples. Active TEs and their recent descendents can readily be detected and analysed. The mechanism(s) that may be critical to the 'containment' of TE copy numbers in outbreeding populations of hosts such as *Drosophila* remains an open question. Simple zygotic selection against the deleterious phenotype of the mutations associated with TE insertions clearly influences the observed patterns, e.g. TEs are seldom found within transcription units. But a number of model-based tests have suggested that selection against the zygotic fitness reductions associated with TE insertion is unlikely to be the critical mechanism in the containment of TE copy number. Ectopic exchange between nonhomologous copies of a TE family is a well-documented meiotic phenomenon that has the potential to provide a robust explanation for the observed distributions of TEs in samples from natural populations. The recently discovered germline-specific RNA interference shares attractive features with the ectopic exchange model, i.e. being dependent only on genomic sequence similarity (not on encoded protein structure and function) and on having a potentially greater than linear response with increasing copy number. The piRNA pathway may prove to be a critical element of TE copy-number containment. Virtually, all the questions concerning the population dynamics of *Drosophila* TEs raised in the literature and discussed here will be open to detailed re-examination and powerful testing based on the anticipated samples of fully sequenced *Drosophila* genomes.

REFERENCES

Aminetzach, Y. T., Macpherson, J. M. & Petrov, D. A. 2005 Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. *Science* **309**, 764–767. (doi:10.1126/science.1112699)

Aquadro, C. F., Deese, S. F., Bland, M. M., Langley, C. H. & Laurie-Ahlberg, C. C. 1986 Molecular population genetics of the *alcohol dehydrogenase* gene region of *Drosophila melanogaster*. *Genetics* **114**, 1165–1190.

Aravin, A. A., Hannon, G. J. & Brennecke, J. 2007 The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318**, 761–764. (doi:10.1126/science.1146484)

Bartolome, C. & Maside, X. 2004 The lack of recombination drives the fixation of transposable elements on the fourth chromosome of *Drosophila melanogaster*. *Genet. Res.* **83**, 91–100. (doi:10.1017/S0016672304006755)

Bartolome, C., Maside, X. & Charlesworth, B. 2002 On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol. Biol. Evol.* **19**, 926–937.

Bergman, C. M., Quesneville, H., Anxolabéhère, D. & Ashburner, M. 2006 Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. *Genome Biol.* **7**, R112. (doi:10.1186/gb-2006-7-11-r112)

Blumenstiel, J. P. & Hartl, D. L. 2005 Evidence for maternally transmitted small interfering RNA in the repression of transposition in *Drosophila virilis*. *Proc. Natl Acad. Sci. USA* **102**, 15 965–15 970. (doi:10.1073/pnas.0508192102)

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R. & Hannon, G. J. 2007 Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103. (doi:10.1016/j.cell.2007.01.043)

Brennecke, J., Malone, C. D., Aravin, A. A., Sachidanandam, R., Stark, A. & Hannon, G. J. 2008 An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **322**, 1387–1392. (doi:10.1126/science.1165171)

Brookfield, J. F. Y. 1982 Interspersed repetitive DNA sequences are unlikely to be parasitic. *J. Theor. Biol.* **94**, 281–299. (doi:10.1016/0022-5193(82)90313-7)

Bucheton, A., Busseau, I. & Teninges, D. 2002 I elements in *Drosophila melanogaster*. In *Mobile DNA II* (eds N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz), pp. p796–p812. Herndon, VA: American Society for Microbiology Press.

Chambeyron, S., Popkova, A., Payen-Groschêne, G., Brun, C., Laouini, D., Pelisson, A. & Bucheton, A. 2008 piRNA-mediated nuclear accumulation of retrotransposon transcripts in the *Drosophila* female germline. *Proc. Natl Acad. Sci. USA* **105**, 14 964–14 969. (doi:10.1073/pnas.0805943105)

Charlesworth, B. 1985 The population genetics of transposable elements. In *Population genetics and molecular evolution* (eds T. Ohta & K. Aoki), pp. 213–232. Berlin, Germany: Springer-Verlag.

Charlesworth, B. 1991 Transposable elements in natural populations with a mixture of selected and neutral insertion sites. *Genet. Res.* **57**, 127–134. (doi:10.1017/S0016672300029190)

Charlesworth, B. 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genet. Res.* **68**, 131–149. (doi:10.1017/S0016672300034029)

Charlesworth, B. & Charlesworth, D. 1983 The population dynamics of transposable elements. *Genet. Res.* **42**, 1–27. (doi:10.1017/S0016672300021455)

Charlesworth, B. & Langley, C. H. 1986 The evolution of self-regulated transposition of transposable elements. *Genetics* **112**, 359–383.

Charlesworth, B. & Langley, C. H. 1989 The population genetics of *Drosophila* transposable elements. *Annu. Rev. Genet.* **23**, 251–287. (doi:10.1146/annurev.ge.23.120189.001343)

Charlesworth, B. & Lapid, A. 1989 A study of ten families of transposable elements on X chromosomes from a population of *Drosophila melanogaster*. *Genet. Res.* **54**, 113–125. (doi:10.1017/S0016672300028482)

Charlesworth, B., Lapid, A. & Canada, D. 1992 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. I. Element frequencies and distribution. *Genet. Res.* **60**, 103–114. (doi:10.1017/S0016672300030792)

Charlesworth, B., Morgan, M. T. & Charlesworth, D. 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303.

Craig, N. L. 2002 Tn7. In *Mobile DNA II* (eds N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz), pp. p423–p456. Herndon, VA: American Society for Microbiology Press.

Daborn, P. J. *et al.* 2002 A single p450 allele associated with insecticide resistance in *Drosophila*. *Science* **297**, 2253–2256. (doi:10.1126/science.1074170)

Daniels, S. B., Peterson, K. R., Strausbaugh, L. D., Kidwell, M. G. & Chovnick, A. 1990 Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. *Genetics* **124**, 339–355.

Davis, P. S., Shen, M. W. & Judd, B. H. 1987 Asymmetrical pairings of transposons in and proximal to the *white* locus of *Drosophila* account for four classes of regularly occurring exchange products. *Proc. Natl Acad. Sci. USA* **84**, 174–178. (doi:10.1073/pnas.84.1.174)

Dolgin, E. S. & Charlesworth, B. 2008 The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* **178**, 2169–2177. (doi:10.1534/genetics.106.060434)

Dorer, D. R. & Henikoff, S. 1994 Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77**, 993–1002. (doi:10.1016/0092-8674(94)90439-1)

Eanes, W. F., Wesley, C. & Charlesworth, B. 1992 Accumulation of P elements in minority inversions in natural populations of *Drosophila melanogaster*. *Genet. Res.* **59**, 1–9. (doi:10.1017/S0016672300030111)

Feschotte, C., Jiang, N. & Wessler, S. R. 2002 Plant transposable elements: where genetics meets genomics. *Nat. Rev. Genet.* **3**, 329–341. (doi:10.1038/nrg793)

Fry, J. D. 2004 On the rate and linearity of viability declines in *Drosophila* mutation-accumulation experiments: genomic mutation rates and synergistic epistasis revisited. *Genetics* **166**, 797–806. (doi:10.1534/genetics.166.2.797)

Goldberg, M. L., Sheen, J.-Y., Gehring, W. J. & Green, M. M. 1983 Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the *Drosophila melanogaster* genome. *Proc. Natl Acad. Sci. USA* **80**, 5017–5021. (doi:10.1073/pnas.80.16.5017)

González, J., Lenkov, K., Lipatov, M., Macpherson, J. M. & Petrov, D. A. 2008 High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biol.* **6**, e251. (doi:10.1371/journal.pbio.0060251)

Green, M. M. 1959 Non-homologous pairing and crossing over in *Drosophila melanogaster*. *Genetics* **44**, 1243–1256.

Grewal, S. I. & Elgin, S. C. 2007 Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399–406. (doi:10.1038/nature05914)

Gunawardane, L. S., Nishida, K. M. & Siomi, M. C. 2007 A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315**, 1587–1590. (doi:10.1126/science.1140494)

Heger, A. & Ponting, C. 2007 Evolutionary rate analyses of orthologs and paralogs from 12 *Drosophila* genomes. *Genome Res.* **17**, 1837–1849. (doi:10.1101/gr.6249707)

Hill, W. G. & Robertson, A. 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294. (doi:10.1017/S0016672300010156)

Jensen, S., Gassama, M. P. & Heidmann, T. 1999 Taming of transposable elements by homology-dependent gene silencing. *Nat. Genet.* **21**, 209–212. (doi:10.1038/5997)

Judd, B. H. 1959 Studies on some position pseudoalleles at the *white* locus region in *Drosophila melanogaster*. *Genetics* **44**, 34–42.

Kalmykova, A. I., Klenov, M. S. & Gvozdev, V. A. 2005 Argonaute protein PIWI controls mobilization of retrotransposons in the *Drosophila* male germline. *Nucleic Acids Res.* **33**, 2052–2059. (doi:10.1093/nar/gki323)

Kaminker, J. S. *et al.* 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* **3**, research0084.1–0084.20. (doi:10.1186/gb-2002-3-12-research0084)

Kaplan, N. L. & Brookfield, J. F. 1983 Transposable elements in Mendelian populations. III. Statistical results. *Genetics* **104**, 485–495.

Kaplan, N., Darden, T. & Langley, C. H. 1985 Evolution of transposable elements in Mendelian populations. IV. Mutant elements and extinction. *Genetics* **109**, 459–480.

Klenov, M. S., Lavrov, S. A., Stolyarenko, A. D., Ryazansky, S. S., Aravin, A. A., Tuschl, T. & Gvozdev, V. A. 2007 Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the *Drosophila melanogaster* germline. *Nucleic Acids Res.* **35**, 5430–5438. (doi:10.1093/nar/gkm576)

Langley, C. H., Brookfield, J. F. & Kaplan, N. 1983 Transposable elements in Mendelian populations. I. A theory. *Genetics* **104**, 457–471.

Langley, C. H., Montgomery, E., Hudson, R., Kaplan, N. & Charlesworth, B. 1988 On the role of unequal exchange in the containment of transposable element copy number. *Genet. Res.* **52**, 223–235. (doi:10.1017/S0016672300027695)

Li, C. *et al.* 2009 Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* **137**, 509–521. (doi:10.1016/j.cell.2009.04.027)

Lichten, M., Borts, R. H. & Haber, J. E. 1987 Meiotic gene conversion and crossing over between dispersed homologous sequences occurs frequently in *Saccharomyces cerevisiae*. *Genetics* **115**, 233–246.

Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. 1998 Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017.

Mackay, T. F. 1989 Transposable elements and fitness in *Drosophila melanogaster*. *Genome* **31**, 284–295.

Malone, C. D., Brennecke, J., Dus, M., Stark, A., McCombie, W. R., Sachidanandam, R. & Hannon, G. J. 2009 Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* **137**, 522–535. (doi:10.1016/j.cell.2009.03.040)

Maside, X., Bartolome, C., Assimacopoulos, S. & Charlesworth, B. 2001 Rates of movement and distribution of transposable elements in *Drosophila melanogaster*: in situ hybridization vs Southern blotting data. *Genet. Res.* **78**, 121–136. (doi:10.1017/S0016672301005201)

Maside, X., Assimacopoulos, S. & Charlesworth, B. 2002 Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster*. *Genet. Res.* **75**, 275–284. (doi:10.1017/S0016672399004474)

McClintock, B. 1950 The origin and behavior of mutable loci in maize. *Proc. Natl Acad. Sci. USA* **36**, 344–355. (doi:10.1073/pnas.36.6.344)

Miyashita, N. & Langley, C. H. 1988 Molecular and phenotypic variation of the white locus region in *Drosophila melanogaster*. *Genetics* **120**, 199–212.

Montgomery, E. A. & Langley, C. H. 1983 Transposable elements in Mendelian populations II. Distribution of three copia-like elements in a natural population of *Drosophila melanogaster*. *Genetics* **104**, 473–483.

Montgomery, E. A., Charlesworth, B. & Langley, C. H. 1987 A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genet. Res.* **49**, 31–41. (doi:10.1017/S0016672300026707)

Montgomery, E. A., Huang, S. M., Langley, C. H. & Judd, B. H. 1991 Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* **129**, 1085–1098.

Mukai, T. 1968 The genetic structure of natural populations of *Drosophila melanogaster*. VII. Synergistic interaction of mutant polygenes controlling viability. *Genetics* **61**, 749–761.

Nuzhdin, S. V. & Mackay, T. F. 1995 The genomic rate of transposable element movement in *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**, 180–181.

Obbard, D. J., Gordon, K. H. J., Buck, A. H. & Jiggins, F. M. 2009 The evolution of RNAi as a defence against viruses and transposable elements. *Phil. Trans. R. Soc. B* **364**, 99–115. (doi:10.1098/rstb.2008.0168)

Ostertag, E. & Haig, H. K. 2001 Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* **35**, 501–538. (doi:10.1146/annurev.genet.35.102401.091032)

Pasyukova, E. G., Nuzhdin, S. V., Morozova, T. V. & Mackay, T. F. C. 2004 Accumulation of transposable elements in the genome of *Drosophila melanogaster* is associated with a decrease in fitness. *J. Hered.* **95**, 284–290. (doi:10.1093/jhered/esh050)

Petrov, D. A., Aminetzach, A. T., Davis, J. C., Bensasson, D. & Hirsh, A. E. 2003 Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol. Biol. Evol.* **20**, 880–892. (doi:10.1093/molbev/msg102)

Rio, D. C. 2002 P transposable elements in *Drosophila melanogaster*. In *Mobile DNA II* (eds N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz), pp. 484–518. Herndon, VA: American Society for Microbiology Press.

Rizzon, C., Marais, G., Gouy, M. & Biémont, C. 2002 Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. *Genome Res.* **12**, 400–407. (doi:10.1101/gr.210802)

Rubin, G. M. 1983 Dispersed repetitive DNA in *Drosophila*. In *Mobile genetic elements* (ed. J. A. Shapiro), pp. 329–361. New York, NY: Academic Press.

Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H. & Siomi, M. C. 2006 Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes & Dev.* **20**, 2214–2222. (doi:10.1101/gad.1454806)

Sarot, E., Payen-Groschêne, G., Bucheton, A. & Pélisson, A. 2004 Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the *Drosophila melanogaster* flamenco gene. *Genetics* **166**, 1313–1321. (doi:10.1534/genetics.166.3.1313)

Sassaman, D. M., Dombroski, B. A., Moran, J. V., Kimberland, M. L., Naas, T. P., DeBerardinis, R. J., Gabriel, A., Swergold, G. D. & Kazazian, H. H. 1997 Many human L1 elements are capable of retrotransposition. *Nat. Genet.* **16**, 37–43. (doi:10.1038/ng0597-37)

Savitsky, M., Kwon, D., Georgiev, P., Kalmykova, A. & Gvozdev, V. 2006 Telomere elongation is under the control of the RNAi-based mechanism in the *Drosophila* germline. *Gene. Dev.* **20**, 345–354. (doi:10.1101/gad.370206)

Slotkin, R. K. & Martienssen, R. 2007 Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8**, 272–285. (doi:10.1038/nrg2072)

Sniegowski, P. D. & Charlesworth, B. 1994 Transposable element numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster*. *Genetics* **137**, 815–827.

Sturtevant, A. H. 1925 The effects of unequal crossing over at the *bar* locus in *Drosophila*. *Genetics* **10**, 117–147.

Sturtevant, A. H. 1929 A further study of the so-called mutation at the *bar* locus of *Drosophila*. *Genetics* **13**, 401–409.

Talbert, P. B. & Henikoff, S. 2006 Spreading of silent chromatin: inaction at a distance. *Nat. Rev. Genet.* **7**, 793–803. (doi:10.1038/nrg1920)

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V. & Zamore, P. D. 2006 A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**, 320–324. (doi:10.1126/science.1129333)