ORIGINAL ARTICLE

Genetic variation of *copia* suppression in *Drosophila melanogaster*

W Vu¹ and S Nuzhdin²

¹Department of Molecular and Computational Biology, University of Southern California, Los Angeles, CA, USA and ²Department of Evolution and Ecology University of California at Davis, Davis, CA, USA

Transposable elements (TEs) are genomic parasites that propagate by exploiting its host reproductive machinery. However, some hosts have evolved the ability to silence TE activity, whereas others have not. We are investigating the population dynamics of TE host-silencing pathways, particularly *copia* long terminal repeat retrotransposon in *Drosophila melanogaster*. Here, we identify large effect genes involved in *copia* suppression by using a semi-quantitative analysis to assay levels of *copia* plasmids (believed to be an intermediate of transposition) in 98 recombinant inbred lines constructed from a line exhibiting high *copia* transpositions and a line exhibiting no transpositions. The results revealed that the influence of *copia* copy number and transcription level on *copia* plasmid concentrations are weak and that

genomic factors, presumably encoded by the host, have stronger effects on transposition rates. We mapped a QTL affecting *copia* plasmid concentration within the 33A–43E cytological region of the second chromosome and applied a quantitative deficiency complementation analysis on this chromosomal region. One out of the two large effect deficiencies on *copia* plasmid concentrations corresponded to the *vasa* gene, an important component of the nuage-piwi RNA TE-silencing machinery. We hypothesize that *copia* suppression occurs by the joint action of several post-transcriptional mechanisms with at least one of the blocks taking place in the nuage.

Heredity (2011) 106, 207–217; doi:10.1038/hdy.2010.41; published online 7 July 2010

Keywords: retrotransposon *copia*; QTL mapping; transcription; plasmid DNA; transposition; *Drosophila melanogaster*

Introduction

Transposable elements (TEs) are selfish genetic elements that propagate through the population by exploiting the host sexual reproductive machinery. TEs have existed in the genomes of living organisms for hundreds of millions of years and are found in most plants and animals; with as much as 90% of the genome being TE-derived sequences (Finnegan, 1978). For example, there are ~100 000 copies of LINE 1 elements and about a million copies of the Alu elements in humans (Xing *et al.*, 2007; Lander *et al.*, 2001). As genetic parasites, TEs exploit the host transcriptional system to increase copy number (Nuzhdin *et al.*, 1996). Insertions into new chromosomal locations are generally deleterious because they disrupt gene expression, cause ectopic exchange, and chromosomal breakage (Charlesworth *et al.*, 1994). Although some TE-induced mutations are beneficial, it is clear that their net effect on host fitness is negative, with an estimated fitness cost in excess of 5% (Charlesworth *et al.*, 1994). So how have they continued to parasitize the genome without self-destructing by means of killing the host? One explanation seems to be that TE activity resulting from copy number increase is balanced by selection against hosts with TE-induced deleterious effects.

There are ~50 families of TEs that occupy the genome of *Drosophila melanogaster*, a model species for studying the dynamic relationship between TEs and their host. Horizontal transmission is thought to be the primary mode of TE infection across species and populations. P-elements, for instance, have recently invaded the genome of *D. melanogaster* and have managed to rapidly sweep through the population within the last 50 years (Brookfield *et al.*, 1984; Marzo *et al.*, 2008). With such sweeps, the number of elements within a host might continue to increase and will do so at the expense of host fitness (Charlesworth and Langley, 1989; Mackay *et al.*, 1992; Biémont, 1994; Houle and Nuzhdin, 2004). As TE survival is dependent on the host survival and reproduction, a decrease in host fitness will inevitably interfere with the chances of TEs propagating through the population. This dynamic power play between TEs and their hosts has been shown to ignite co-evolution between these players. Therefore, a self-imposed control mechanism is likely to evolve in TEs to prevent detrimental fitness effects on the host. Indeed, some elements such as Type I transposons can restrict their transposition at high copy number (Biémont, 1994; Lohe and Hartl, 1996). However, self-regulation seems not to be an evolutionarily stable strategy for retrovirus-like elements (Brookfield and Badge, 1997). For instance, *copia* and *Doc* elements do not show reduced transpositions when TE copy number increases (Nuzhdin *et al.*, 1996).

Host fitness loss can be partially rescued if host genes controlling transpositions sweep through the infected

Correspondence: W Vu, Department of Molecular and Computational Biology, University of Southern California, 1050 Childs Way RRI 316, Los Angeles, CA 90089, USA.

E-mail: wvu@usc.edu

Received 15 April 2009; revised 2 March 2010; accepted 8 March 2010; published online 7 July 2010

population. However, these sweeps might not result in fixation of the host transposition suppressors but are likely to remain incomplete due to weak selection (Nuzhdin *et al.*, 1998), which would then result in an overall low transposition rate observed in nature. This hypothesis stems from the fact that *D. melanogaster* genotypes differ in which TE families are active. One genotype, for example, showed transposition of *hobo* and *I* elements but none for *copia* and 412 families (Harada *et al.*, 1990). In another genotype, *copia*, 297, and *FB4* were active but *I*, *mdg1*, *gypsy*, *hobo*, 2161, 2242, and *Doc* elements were not (Eggleston *et al.*, 1988). In yet another genotype, out of the 13 TE families, only *copia*, *I*, *roo*, and *Doc* were found to be actively transposing (Nuzhdin and Mackay, 1994, 1995). These studies collectively suggest that there exists natural variation in TE stability and that the host genetic background is an important factor that influences TE activity and thus copy number maintenance in natural populations.

The *gypsy* element, one of the best-studied retrotransposons, offers a glimpse of genetic variation in TE transmission and host control factors. Two components are required for the activity of *gypsy*: maternally inherited 'rogue' or active elements and the permissive *flamenco* locus (Bucheton, 1995; Prud'homme *et al.*, 1995; Mevel-Ninio *et al.*, 2007). There are two classes of *flamenco* alleles, restrictive and permissive, present in natural populations (Pelisson *et al.*, 1997). The restrictive alleles are dominant and repress the mobility of functional proviruses by suppressing *gypsy* transcripts through the piwi RNA (piRNA)-silencing pathway (Pelisson *et al.*, 2007). The piRNA pathway is a maternally transmitted defense mechanism designed to control the proliferation of a variety of TE families, such as *TAHRE* retroelement (Shpiz *et al.*, 2007) and *P* transposons (Simmons *et al.*, 2007). piRNAs are 29 nucleotide RNA molecules that interact with silencing proteins *Piwi*, *Aubergine*, and *Argonaut*, which are involved in transcriptional and chromatin silencing of retrotransposons in the *D. melanogaster* germ line (Klenov *et al.*, 2007). Worth noting is that piRNA has so far been shown to act through decreasing transcript levels by either direct silencing of transcription or RNA decay (Hutvagner and Zamore, 2002; Grivna *et al.*, 2006).

We focus on characterizing a potential host control mechanism for the *copia* element. Similar to the *gypsy* element, *copia* is a long terminal repeat (LTR) retrotransposon. However, unlike the *TAHRE* retroelement and *P* transposons, *copia* activity is suppressed in genotypes exhibiting high *copia* transcript levels (Nuzhdin *et al.*, 1998). Here, we analyze the genetics of this unusual post-transcriptional suppression mechanism. *Copia-Ty1* belongs to the Pseudoviridae retrotransposon family (Peterson-Burch and Voytas, 2002) and is abundant in plants, present in insects (1731 and *copia* families in *Drosophila*), but uncommon in vertebrates. *Copia* transpositions take place primarily in fly spermatocytes, where their transcript level is strongly elevated in the early stages of spermatogenesis (Pasyukova *et al.*, 1997). *Copia* is stable in the majority of genotypes, including Oregon R (Ore), but active in several wild-type strains, such as 2b (Pasyukova and Nuzhdin, 1993). There is a strong positive correlation between copy number and transposition rate, suggesting that *copia* copy number is not self-regulated (Pasyukova *et al.*, 1998). *Copia*, unlike

gypsy, does not require maternal transmission of a 'rogue' copy for transposition to occur (Perdue and Nuzhdin, 2000).

We identified candidate genes and potential control mechanisms that suppress the *copia* element in *D. melanogaster*. According to Boeke and Chapman, 1991 (Figure 1), virus-like particles (VLPs) are assembled, during the late stages of the transposition pathway and serve as sites for RNA reverse transcription into linear and circular DNA transposition intermediates. For this reason, we quantified *copia* extrachromosomal DNA plasmids in 98 recombinant inbred lines (RILs) for QTL analyses. We mapped regions of the genome required for *copia* reverse transcription and conducted a deficiency mapping analysis to localize candidate genes within these QTL regions that interfere with *copia* reverse transcription. The most interesting gene represented in the analysis is the *vasa* gene, which is expressed in a special compartment of the germ cell called the nuage. In the cytoplasm, the nuage interacts with the PIWI protein, which is an important component of the piRNA machinery (Klattenhoff and Theurkauf, 2008). Therefore, we propose an unusual TE suppression mechanism that requires transcript sequestering during critical stages of spermatogenesis.

Materials and methods

Wild-type *D. melanogaster* lines and *copia* transpositions
We used the two isogenic lines 2b (Pasyukova *et al.*, 1998), Ore (Oregon R, Nuzhdin *et al.*, 1996), and a set of RILs derived from them (Nuzhdin *et al.*, 1997). Briefly, in the Ore line, no transpositions have been found over 6 years and *copia* positions are fixed (Nuzhdin *et al.*, 1998). In the 2b line, a high *copia* transposition rate (10^{-3} – 10^{-2}) was observed in 1991 (Pasyukova and Nuzhdin, 1993) and has continued since then (Pasyukova *et al.*, 1998). Although *copia* has fixed in some sites, new transpositions segregate in the 2b line. We generated a set of 98 RILs by crossing one 2b male with an Ore female, then backcrossed F1 females with a 2b male, with repeated full-sib mating of subsequent progeny thereafter (Nuzhdin *et al.*, 1997 for details). We mapped *copia* positions (either fixed or segregating) in two replicate parental 2b males (Table 1). Accordingly, RILs vary not only in genetic composition, but also in the copy number of *copia* elements. Furthermore, several *copia* transpositions were detected during and after the construction of the RILs (see Nuzhdin *et al.*, 1998 for detail). We used transcript level, copy number, and transposition activity generated from Nuzhdin *et al.* (1998) in our analysis of the *copia* element in each of the RILs. Transcript level was evaluated by standard Northern analysis. Cytological positions were obtained by *in situ* hybridization of the plasmid cDM5002 containing a full-length *copia* element (Finnegan, 1978) to polytene salivary gland chromosomes of third instar larvae (Shrimpton *et al.*, 1986). Probes were labeled with biotinylated dATP (bio-7-dATP, BRL) by nick translation. Hybridization was detected using the Vectastain ABC kit (Vector Labs, Burlingame, CA, USA) and visualized with diaminobenzidine. The element locations were determined at the level of cytological bands on the standard Bridge's map (Lefevre, 1976).

Figure 1 *Copia* transposition process. (a) *Copia* transcription. (b) *Copia* mRNA processing and translation of *gag*, *pol*, *int*, and *Rnase* transcripts in the cytoplasm. (c) Translated proteins are delivered into the nucleus and VLPs are assembled from *gag* proteins. (d) Inside the VLPs, reverse transcription of full-length mRNA into DNA. (e) Complex strand exchanges required to fully replicate the two long terminal repeats (LTRs) to produce linear DNA, single and double LTR extrachromosomal plasmids. Finally, integration of a newborn element into the host genome occurs.

Table 1 Effects of deficiencies on *Copia* plasmid concentrations

Deficiency lines	Break points	Parental effect	Effect of deficiency	Parent deficiency interaction
Df(2R)ED1715	43A4–43F1	0.0052	0.0028	0.6317
Df(2R)ED1673	42E4–43D3	0.9185	0.8802	0.5786
Df(2R)ED1618	42C3–43A1	0.0010	0.0011	0.4306
Df(2R)ED1552	42A11–42C7	0.9360	0.1179	0.6425
Df(2R)ED1484	42A2–42A14	0.1357	0.5090	0.6997
Df(2L)ED791	34E1–35B4	0.0871	0.1676	0.6662
Df(2L)ED784	34A4–34B6	<0.0001	0.0599	0.1244
Df(2L)ED780	33E4–34A7	0.7695	0.6655	0.1095
Df(2L)ED761	33A2–33E5	0.0244	0.0216	0.4038
Df(2L)ED1473	39B4–40A5	0.2975	0.0441	0.9268
Df(2L)ED1451	38F5–39E2	0.0013	0.4718	0.4496
Df(2L)ED1315	38B4–38F5	0.0020	0.5130	0.5000
Df(2L)ED1303	37E5–38C6	0.0023	0.0043	0.6030
Df(2L)ED1272	37C5–38A2	0.0337	0.0191	0.4599
Df(2L)ED1203	36F7–37C5	0.0021	0.1246	0.6248
Df(2L)ED1196	36E6–37B1	<0.0001	<0.0001	0.7848
Df(2L)ED1102	35F12–36A10	0.1917	0.7719	0.2454
Df(2L)ED1054	35B10–35D4	<0.0001	<0.0001	0.006

The *P*-values correspond to the effects of parental line (2b versus Ore; chromosome inherited from a deficiency stock—balancer versus deficiency; and their interaction term).

Amplification of *Copia* plasmid from parental lines Ore and 2b
During retrotransposition, *Copia* DNA plasmids are produced containing either a single LTR or double LTR (dLTR). Although, single LTR plasmids are thought to be abortive, it is unclear whether dLTR are abortive or direct intermediates of *Copia* transposition (Flavell and Ishorowicz, 1981; Flavell, 1984). According to the *Copia* sequence reported by Csink and McDonald (1995), We designed primers to amplify the junction between the 5' and 3' LTRs of the *Copia* dLTR plasmids (Supplementary

Table 1). Six different primer pairs were designed to represent *Copia* sequence variance (P1/P2, P1/P3, P1/P4, P1/P5, P1/P6, and P1/P7).

DNA samples were prepared from 10 males aged ~1 week using the Puregene Kit (Gentra, Bangkok, Thailand). The DNA concentrations were determined on a Beckman Spectrophotometer Du-65. PCR amplifications were carried out in the 21 μ l reaction containing 200 ng genomic DNA, 20 ng of the primer P1, and one of the primers P2–P7, 1 $\times$ PCR Buffer, 200 μ M dNTP, 3 mM

magnesium chloride, and 2.5 U of *Taq* polymerase (Promega, Madison, WI, USA) 0.15 microCi of γ - ^{33}P (Amersham, Piscataway, NJ, USA) per amplification was incorporated into primer P1 with 0.1 U of T4 Polynucleotide Kinase (New England Biolabs, Ipswich, MA, USA) in a buffer supplied by the manufacturer. Thermal cycle conditions were: 5 min at 94 °C, followed by 35 cycles for 0.5 min at 94 °C and 1 min at 60 °C, followed by a 5 min final extension at 72 °C. The amplification products were visualized after fractionation through 6% acrylamide gel, using Phosphorimaging on the Storm 8600 imager.

Electron microscopy images of VLPs in testes

As VLPs are believed to be primary vehicles for retrotransposition (Boeke and Chapman, 1991), we took electron microscopy images of testes cross-sections to validate this assumption. Testes were dissected in phosphate buffer (pH = 7.4) and fixed in 2% paraformaldehyde and 0.1% glutaraldehyde in phosphate buffer (pH = 7.4) for 12 h at room temperature; then placed in 2% osmium tetroxide and cacodylate buffer for storage at 4 °C. The fixed testes were dehydrated through an ethanol (EtOH) dilution series up to 100% EtOH, infiltrated in a 1:1 EtOH/LR White mixture overnight, embedded in 100% LR White acrylic resin (Ted Pella Inc., Redding CA, USA) in beam capsules, and incubated overnight at 60 °C. The blocks were then ultra thin sectioned (75 nm in thickness) and placed on parlodian-coated nickel grids. Sections on grids were etched with 0.5% sodium metaperiodate 10 min at room temperature to remove excess resin, and then washed five times for 10 min each by drops of 0.025 M Tris buffer (pH = 7.4). Sections were viewed and photographed on a Philips CM10 electron microscope.

Copia dLTR plasmid sequencing

To verify that we are isolating *copia* dLTR plasmids, we sequenced the junction between the LTRs (Figure 1d–e, for dLTR structure). DNA was extracted from individual flies and primers P8 and P9 were designed to amplify the junction between the 5' and 3' LTR of the dLTR plasmid DNA (Supplementary Table 1). PCR reaction mixture consisted of 1 × AmpliTaq Gold PCR Buffer, 2 U of AmpliTaq Gold DNA Polymerase, 100 mM MgCl₂, 40 mM of dNTPs, 1 mM of each primer, and 200 ng of DNA template. Thermo cycling conditions were: 10 min at 94 °C, 15 s at 94 °C, 30 s at 60 °C, and 15 s at 72 °C for 30 cycles. PCR products were visualized on 1.5% acrylamide gel. We cut a band of the expected size (180 bp) from the gel and extracted the product from the gel using the Wizard SV Gel and PCR Clean-Up System (Promega). The extracted product was prepared for sequencing as follows: 6.4 pmol of *copia* P9 primer, 40 ng of template and deionized water to bring total volume to 18 µl. The sample was submitted to USC Norris Core Facility for sequencing.

Semi-quantitative analysis of concentration of *copia* plasmid in each RIL

To get an idea of the relative differences in *copia* dLTR plasmids between the 2b and Ore parental lines, we extracted DNA and semi-quantified relative *copia* plasmid concentrations, as mentioned above. To measure the amount of target dLTR molecules, we compared the

concentration of *copia* plasmids relative to a known quantity of pBB54 plasmid containing two tandemly arranged LTRs separated by 15 bp insertion 5'-AGGT GAAAAGGTTTC-3' (Flavell and Ish-Horowitz, 1983). Both the total DNA and pBB54 plasmid were amplified in the same reaction mixture that included primers P1 and P4 with 0.15 microCi of γ - ^{33}P (Amersham) per amplification. Before amplification, the plasmid was linearized with restriction enzyme Hind III (Promega) at 1 µg per 6000 µl and 1 µg at 60 000 µl. Each PCR amplification contained 200 ng of total DNA, 20 ng of each primer, and 1 µl of the plasmid pBB54. The intensity of the amplified products were recorded after fractionation through 6% acrylamide gel by Phosphorimaging on the Storm 8600 imager. The amount of DNA (^{33}P) in each band was approximated from the signal level in the same area rectangles covering the bands. The relative concentrations of the dLTR circles were estimated as a ratio of the signal intensity of the genomic DNA template band to the plasmid DNA template band generated from the same reaction mixture. Across 98 RILs and 2 parental lines, the average ratio was 0.472 for the plasmid concentration 1/6000 µg/µl, and 4.81 for the plasmid concentration 1/60 000 µg/µl. Correlations between *copia* copy numbers, transcript levels, and plasmid concentrations among RILs were calculated with the CORR procedure on data averaged between observations within each genotype (SAS Institute, 1989).

QTL mapping for *copia* plasmid concentration

Each RIL has been genotyped using 92 *roo* TE with polymorphic markers on the X, second, third, and fourth chromosomes (Nuzhdin *et al.*, 1997) to map factors controlling concentration of *copia* plasmid. We used the marker data available from Nuzhdin *et al.* (1997) to map factors controlling concentration of *copia* plasmid. We replicated statistical techniques and directly linked our new analyses to those described by Nuzhdin *et al.* (1998). Briefly, we used the QTL Cartographer software (Basten *et al.*, 1997) for composite interval mapping (Zeng, 1994) to test whether an interval flanked by two adjacent markers actually contains a QTL affecting the trait of interest. With multiple regression analysis of markers outside of the test interval, we can control for the effects of QTLs that are chromosomally linked. In all, 76 cytological markers (16 markers out of 92 were completely linked with neighboring markers and excluded from the analysis; Nuzhdin *et al.*, 1997) were used for the analysis with parameters 6 (model) and 10 (window size). The conditioning markers were chosen by stepwise forward regression. The likelihood ratio test statistic, LR , is $-2\ln(L_0/L_1)$, where L_0/L_1 is the ratio of the likelihood under the null hypothesis (there is no QTL in the interval) to the alternative hypothesis (there is a QTL in the interval). An empirical distribution of LR test statistics under the null hypothesis of no association between any of the intervals and the trait values were obtained by randomly permuting the trait data 1000 times and calculating the maximum LR statistics across all intervals for each permutation. LR statistics from the original data that were exceeded by the permutation maximum LR statistics <50 times are significant at $P=0.05$. The trait values were the ratios of the band intensities averaged over 2 (3) measurements per line as

described below. The analysis of the log-transformed data yielded the same results (data not shown).

Deficiency mapping of *copia* plasmid concentration

We obtained Drosdel deficiency stocks from the Drosophila Genetic Resource Center in Kyoto, Japan. The deficiencies span the 33E–43E cytological region on the second chromosome and maintained with SM6a balancer chromosomes. In choosing a subset of deficiencies, we attempted to minimize their number while maximizing coverage of the QTL controlling the amount of *copia* plasmid as described in the Results section. Virgin females from isogenic 2b and Ore lines were crossed separately with males from each deficiency line. F1 progeny consisted of four classes: Df/2b, SM6a/2b, Df/Ore, and SM6a/Ore (Df corresponds to a deficiency and SM6a corresponds to a balancer chromosome with a dominant curly wing marker). We made two replicates of each cross consisting of five virgin females and three males in 5 ml vials with standard cornmeal-agar-yeast-corn syrup-malt-soy-yeast medium and maintained at 25 °C. F1 males from each genotype were collected and aged for 3–5 days before freezing in liquid nitrogen for DNA extractions.

The relative concentration of *copia* dLTR plasmids was analyzed using DNA samples from F1 males collected from the 2b and Ore deficiency crosses. Replicate DNA samples were extracted from 5 F1 males per genotype. DNA extractions were carried out in a 96-well format using ABI Prism 6100 Nucleic Acid Prepstation (Applied Biosystems, Inc., Foster City, CA, USA) by following the manufacturer's suggested DNA extraction protocol. The concentrations of extracted DNA were spectrophotometrically quantified. The Taqman assay of the real-time PCR detection technique was used to quantify relative concentrations of *copia* plasmids with the comparative method for relative quantification (Applied Biosystems). We chose *mdg3* as the endogenous reference gene, a LTR retro-TE reported to be a stable element in both parental lines (Pasyukova and Nuzhdin, 1993). Amplification efficiencies for both target and reference genes should be between 90 and 105% (Real-Time PCR Applications Guide, Bio-Rad, Hercules, CA, USA). Preliminary experiments consisted of 5 µl of a 10-fold dilution series between 100 and 0.05 ng/µl of DNA extracted from 2b parental lines, targeting both *mdg3* and *copia* plasmids to optimize experimental conditions.

As Taqman-based qPCR assays are very sensitive, the primers used in the semi-quantitative analysis did not produce efficient amplification of *copia* plasmids. We designed a new set of primers maintaining the same conceptual design in this experiment. The primers and probes were designed using Primer3 web-based software (<http://fokker.wi.mit.edu/primer3/input.htm>). *mdg3* primers (P10/P11) and probe (P12) were chosen to flank the region between the 5' and 3' LTR in the genome. *Copia* primers P8 and P9 are described above and the *copia* probe (P13) was designed to hybridize to the complementary strand of the 5' and 3' LTR junction. PCR amplifications were performed in 96-well reaction plates, using separate wells to detect *copia* and *mdg3* sequences. We also included multiple replicates of calibrator samples and a 10-fold dilution series of a sample 2b DNA template for *copia* and *mdg3* sequences to ensure reaction efficiencies are between

90 and 105% for each plate. The reaction mixture consisted of 12.5 µl of Taqman Universal PCR Master Mix (Applied Biosystems), 500 nM of each primer, 250 nM of each Taqman *mdg3* FAM Probe (Applied Biosystems), 30–60 ng of DNA template and deionized water to bring the total reaction volume to 25 µl. The real-time PCR amplification was performed in an MJ Thermocycler (Bio-Rad) with the following conditions: 95 °C, 10 min; 95 °C, 15 s; 60 °C, 1 min; for 50 cycles.

As the genomic DNA samples contain both chromosomal and extrachromosomal DNA, the target *copia* plasmid amplification efficiencies were low. As a result, we verified the efficiency of the experimental conditions using PCR products amplified with P8/P9 primers. 2b, Ore and deficiency parental genotypes were used as calibrators to compare and estimate the threshold cycle (Ct) values (that is the fractional cycle number at which the amount of amplified sequence reaches the threshold, Bio-Rad) for the F1 samples. The amount of target sequence (*copia*) was normalized to the reference sequence (*mdg3*) and compared with the calibrator samples (parental genotypes: 2b/Ore/Def). The calculations to estimate the relative expression ratio of the target and calibrator samples are as follows:

$$\begin{aligned}\Delta Ct &= (\text{target})_{\text{copia}} \text{ Ct} - (\text{ref})_{\text{mdg3}} \text{ Ct} \\ \Delta Ct &= (\text{calibrator})_{\text{copia}} \text{ Ct} - (\text{ref})_{\text{mdg3}} \text{ Ct} \\ \Delta\Delta Ct &= \Delta Ct(\text{target}) - \Delta Ct(\text{calibrator}) \\ \text{Relative expression ratio} &= 2^{-\Delta\Delta Ct}\end{aligned}$$

Results

Copia plasmid in parental stocks

In the process of transposition, *copia* is believed to produce dLTR plasmid products that may be direct intermediates of retrotransposition (Flavell and Ishorowicz, 1981). To test the parental strains 2b and Ore for the presence of dLTR *copia* plasmid, we amplified the junction between the dLTRs. Although *Copia* elements and retroviruses both share similar transposition mechanism along with structural and genetic sequence homology, they differ in the mode of integration and excision. Retroviruses require an additional two base pair overhang for integration, whereas *copia* Ty1 can integrate with blunt ends (Boeke and Chapman, 1991).

Similar to that seen in the *pseudoviridae* retrotransposon, we detected an exact junction without the additional two base pair insertion between the dLTRs for *copia* plasmids (Flavell and Ish-Horowicz, 1983). In the 2b and Ore lines, we used four different combinations of primers and observed the expected size bands of ~260 bp (P1/P4 and P1/P5), 310 bp (P1/P2 and P1/P3), and 330 bp (P1/P6 and P1/P7) in the 2b line in addition to multiple faint bands of different sizes from the amplification products, which may be products of neighboring or scrambled *copia* copies in the genome with LTRs within close proximity (Supplementary Table 2). We also observed multiple faint bands in the Ore samples, some partially overlapping with 2b amplified bands. However, there was no apparent amplification product of the expected size detected in the Ore parental line. These results suggest that dLTR plasmid structures may be intermediates required for *copia* transposition.

Electron microscopy images of VLPs in testes

As VLPs are believed to be the primary vehicle for *copia* plasmid formation and transposition (Boeke and Chapman, 1991), we took electron micrographs of testes cross-sections and found distinct differences between 2b and Ore parental lines (Figure 2). We found VLPs in 2b, but not Ore in the serial cross-sections taken from the distal parts of the testes containing maturing spermatids, specifically after the pre-individualization and before the coiling stages of maturing spermatids. As described in Rachidi *et al.* (2005), we saw spherical structures composed of clusters of well-defined particles ~50 nm in diameter, which appear analogous to A-type particles produced in yeast *Ty1* LTR retrotransposons. This might indicate that VLPs and *copia* plasmids are a necessary part of the transposition pathway.

Copia plasmid in RILs

As plasmids are believed to be *copia* transposition intermediates, we semi-quantified *copia* plasmid concentrations in the 2b and Ore parental lines and found strong differences. We measured approximate concentrations of *copia* plasmids in each of the 98 RILs, to map the regions of the 2b and Ore genomes that underlie these phenotypic differences between the parents. The intensity of the 2b parent was ~45-fold stronger than the known pBB54 plasmid concentration, whereas the Ore parental line and 70% of the RILs displayed amplification intensities similar to the pBB54 concentration (Figure 3). In the remaining 30% of RILs, the amplification intensity of the plasmid product was intermediate between the two parental strains, but closer to the stable Ore parent. Interestingly enough, it was previously observed that there was direct *copia* activity in five of the RILs with intermediate plasmid concentrations (Nuzhdin *et al.*, 1998; Figure 3 shown as black bars), suggesting that these

RILs may contain regions affecting *copia* transposition. Although our assay was only semi-quantitative, the differences were strong enough to allow us to test our hypothesis.

We observed a considerable amount of variation in the RILs, when comparing *copia* plasmid concentrations to copy number and transcript levels. There was a moderate positive Pearson product-moment correlation

Figure 3 Distribution of *copia* plasmid concentrations among parental and recombinant inbred lines. The white and black triangles represent relative *copia* plasmid concentration of the parental Ore and 2b line, respectively. Black bars represent RILs exhibiting *copia* transpositions.

Figure 2 Electron micrographs of testes cross-section parental lines. (a) and (c) Maturing spermatid from the 2b and Ore line, respectively, with its axoneme (Ax), major mitochondrial derivative (M), and minor mitochondrial derivative (m). (b) and (d) A higher magnification of a single spermatid from 2b and Ore, respectively, with virus-like particles indicated by the arrow.

between copy number and plasmid concentration ($r=0.285$, $P=0.004$), shown in Figure 4 (filled circles are RILs with direct transpositions). Interestingly, the *copia* copy number is moderate in some RILs with high concentrations of *copia* plasmids, whereas some lines with high copy number had low plasmid concentrations. Similar to the relationship between *copia* copy number and plasmid concentration, there was a moderate positive correlation between *copia* transcript level and plasmid concentration ($r=0.314$, $P=0.002$). Some RILs exhibiting very high transcript levels yielded little to no detectable plasmids, whereas some RILs exhibiting relatively low transcript levels had high plasmid concentrations. These moderate correlations indicate that genomic factors, presumably encoded by the host, may have additional and possibly stronger effects on *copia* activity than *copia* copies themselves.

QTL mapping of *copia* plasmid concentration

We measured *copia* plasmid concentrations in each of the panel of RILs and mapped QTLs that account for the genetic variation among these RILs. As QTL mapping outcomes might, in principle, be sensitive to the procedures of the analyses, we replicated the exact approach for mapping *copia* transcript levels among the same set of RILs used in Nuzhdin *et al.* (1998). We detected a single QTL in the pericentric region of the second chromosome (33A–43E) with substantial statistical support with an LR of 39, which is much higher than the permutation threshold of 24 (Figure 5). When we compared our current mapping results with those from Nuzhdin *et al.* (1998), the *copia* plasmid concentration QTL was positioned precisely within a QTL required for *copia* transposition. There are at least two (possibly three) regions of the 2b genome required for *copia* transposi-

Figure 4 Correlations of plasmid concentration with copy number and transcript level of *copia* element. Filled circles are RILs with direct transpositions.

Figure 5 Location of QTL for *copia* plasmid concentration with respect to transcript level and transposition rate. Plot from double-log likelihood ratio (LR) of composite interval mapping against recombination distance on the X (a), second (b), and third (c) chromosomes. Horizontal lines, Bonferroni-corrected LR critical value for experiment wise $\alpha=0.05$.

tions (Nuzhdin *et al.*, 1998) but the *copia* plasmid concentration mapped only to one region: the pericentric region between the intervals 33A–43E.

Deficiency mapping of copia plasmid concentration

We used a quantitative deficiency complementation technique to more precisely localize the region directly affecting plasmid concentration within the 33A–43E interval on the second chromosome. We collected 18 deficiency lines covering the candidate pericentric region (33A–43E, Figure 6) to detect genetic interactions with the parental Ore and 2b alleles that influence *copia* plasmid concentrations. As *copia* is only active in males and transposition is not influenced by maternal effects (Nuzhdin *et al.*, 1996), we made separate replicate one-way crosses between males from both parental lines and females from each deficiency line. DNA was extracted from the F1 males of 72 genotypes that fall into four classes: 2b/Def, 2b/SM6, Ore/Def, and Ore/SM6. Two-way analyses of variance were used to estimate the significance of the interaction between the parent of origin and the presence of the deficiency (Table 1). In the

majority of the genotypes, *copia* plasmid concentration was higher in the F1 progeny with the balancer chromosome as opposed to the deficiency chromosome (Supplementary Table 3).

F1 progeny from the crosses of the two deficiency lines, however, showed significant deviations from the patterns described above. The *copia* plasmid concentrations were higher in the 2b/Def and Ore/Def genotype classes compared with the majority of the deficiencies ($P < 0.0001$) in the Df(2L)1196 line, with break points 36E6 and 37B1. Moreover, the deficiency containing F1 genotype classes (2b/Def and Ore/Def) for both parental lines had a much higher plasmid concentration than the progeny containing balancer chromosomes (2b/SM6 and Ore/SM6) of ~4-fold changes on the log-scale, $P < 0.0001$. We conclude that the deficiency had a strong effect on the *copia* reverse transcription process, irrespective of the 2b versus Ore alleles of the parent ($P = 0.78$).

In contrast, Df(2L)1054 with the break points 35B10 and 35D4 exhibited epistatic interactions with the parental alleles. In the F1 progeny from the Ore and Df(2L)1054 cross, the genotypes containing balancers

Figure 6 Black solid lines represent deficiency spanning the pericentric region of the 2R (regions 41–44) and 2L (regions 33–40) chromosome.

(Ore/SM6) had higher concentrations of *copia* plasmid than the ones containing the deficiency chromosome (Ore/Def). This is much like the typical pattern observed for all the deficiencies described above. However, the F1 genotype with the 2b parental and Df(2L)1054 chromosome (2b/Def) displayed elevated levels of *copia* plasmid concentration, which indicate that this region might be involved in *copia* plasmid regulation. The main effect of the parent ($P < 0.0001$), deficiency versus balancer ($P < 0.0001$), and the interaction term between them ($P = 0.006$) were highly significant. This indicates that there is an interaction between the deficiency and the 2b-originated alleles defining *copia* plasmid concentrations. Out of ~50 genes represented in the deficiency (Supplementary Table 4), the *vasa* gene seemed to be the most interesting, because it has a pivotal function in germ cell development (Styhler *et al.*, 1998).

To test whether *vasa* directly influences plasmid concentrations, we conducted a complementation test using a heterozygote loss-of-function *vasa* mutant maintained over a CyO balancer (<http://flybase.org/reports/FBst0000284.html>). Following the same crossing scheme for the deficiency crosses mentioned in the Materials and methods section, we generated four genotype classes: 2b/*vas*–, 2b/CyO, Ore/*vas*–, and Ore/CyO. As expected, the F1 genotype of the two possible heterozygote combinations of the 2b chromosome (2b/*vas*– and 2b/CyO) displayed high levels of *copia* plasmids (Supplementary Table 3). However, the F1 progeny derived from the Ore parental line (Ore/*vas*–) did not show the same patterns observed in the deficiency crosses. We previously observed lower levels of plasmid concentrations in Ore-derived F1 males containing a deficiency chromosome (Ore/def) than the ones with a balancer chromosome (Ore/SM6), but the complementation results show that all Ore-derived genotypes (Ore/*vas*– and Ore/CyO) displayed approximately the same level of plasmids. Overall, 2b-derived genotypes (2b/*vas*– and 2b/CyO) had higher plasmid concentrations compared with the Ore-derived genotypes (Ore/*vas*– and Ore/CyO), which is a similar pattern observed for the Df(2L)1054 deficiency cross. The failure to completely complement may be due to the fact that Df(2L)1054 and *vasa* mutant lines carry different balancer chromosomes, which makes direct comparisons challenging. Another complementation test using a loss-of-function *vasa* mutant with an SM6 balancer chromosome comparable to the deficiency mutants is required to validate that *vasa* directly influences *copia* activity. Although the deficiencies available to us had some gaps in coverage of the QTL support interval, they were sufficient to identify strong candidate regions that interact with alleles from the unstable 2b and stable Ore parental lines. At present, we can use these regions to infer possible genetic interactions that interfere with the general process of *copia* transposition.

Discussion

In this study, we explore the co-evolution between *copia* retrotransposon and the host genome to understand the potential mechanism that alleviates the detrimental effects caused by retrotranspositions. The piRNA pathway is thought to serve as a pre-adaptive defense mechanism against viruses and TEs (Aravin *et al.*, 2004, 2007; Blumenstiel and Hartl, 2005). Novel TEs spread

unsuppressed until the element inserts itself into either a heterochromatic or euchromatic piRNA generating loci (Brennecke *et al.*, 2007). This results in the production of piRNA, which then elicits a silencing response by the host that is generally mediated through direct transcriptional silencing or degradation of TE transcripts (see recent reviews by Peters and Meister, 2007; Buchon and Vauray, 2006; Seto *et al.*, 2007; O'Donnell and Boeke, 2007; Klattenhoff and Theurkauf, 2008). However, *copia* control seems to deviate from this general mechanism. Nuzhdin *et al.* (1998) showed that *copia* transcript level QTLs did not coincide with the *copia* transposition QTLs, indicating that *copia* transposition is not restricted by transcriptional silencing but instead by a post-transcriptional mechanism. As *copia* transpositions are limited to male spermatocytes, we propose an alternative piRNA suppression mechanism that involves spermatogenesis-specific events.

During spermatocyte development in *D. melanogaster*, transcription ceases during early stages of spermatogenesis, with limited transcription observed after the first meiosis (see Parvinen, 2005 for a helpful review). Accordingly, the transcripts required for completion of spermatogenesis must be produced during the early stages and preserved for use in the later stages. This kind of translational regulation is mediated by a special membrane-free microtubule-based structure called the nuage, also known as the chromatoid body in mammals (Parvinen, 2005). The nuage is comprised of a variety of germ cell related proteins with the most important being the *vasa* protein, a DEAD-box RNA helicase, which acts synergistically with the components from the piRNA system *Piwi*, *AGO3*, and *Argonaut*. The nuage is thought to mediate translational regulation, silencing male-specific transcripts and transcripts originating from TEs through complex interactions among different families of piRNAs (Aravin *et al.*, 2007; Kotaja and Sassone-Corsi, 2007; Lim and Kai, 2007). Therefore, we hypothesize that *copia* transcripts are suppressed in the nuage by a post-transcriptional mechanism during spermatogenesis.

Here, we show that the QTL mapping to the 33A–43E region associated with *copia* transposition (Nuzhdin *et al.*, 1998) contributes to blocking one or more of the post-transcriptional stages of *copia* retrotransposition. The concentration of *copia* plasmids purported to be an intermediate for *copia* transposition (Flavell, 1984) is strongly increased in the RILs with 34A–43E pericentrical region originating from the active 2b parental line. We also establish a strong affect on *copia* plasmid concentration that is partially accounted for by the 35B10–35D4 segment (Supplementary Table 4). Among the 54 genes represented within this region, *vasa* stands to be a promising candidate because this protein is a critical component of the nuage-piRNA processing machinery. *Vasa* is a hallmark for germ cell development and belongs to a class of proteins that act as RNA chaperones (Parvinen, 2005). *Vasa* expression is restricted to testes and ovaries and null mutations of this gene results in complete sterility (Styhler *et al.*, 1998). Furthermore, there were strong fertility problems in the RILs with the pericentrical region inherited from 2b (Nuzhdin *et al.*, 1997). Coincidentally, this commonly observed fertility problem, along with abdomen abnormalities (Nuzhdin and Pasyukova, 1991; Yoshioka *et al.*, 1992), are among

the phenotypes typical for *vasa* mutants (<http://flybase.org/reports/FBgn003970.html>). However, future studies examining the copy number variation, structure, and expression patterns of *vasa* is necessary to validate the direct role of *vasa* in *copia* activity of the 2b line.

Unfortunately, the deficiency lines for the most potent pericentrical regions on the second chromosome were not available in the co-isogenic deficiency collection (<http://www.drosdel.org.uk/coverage.php>). Accordingly, we were unable to test one interesting prediction of this hypothesis: the failure to complement 2b deficiency in *copia* suppression by one of the piRNA generating loci. At present, our results provide a candidate post-transcriptional pathway that involves the nuage-piRNA machinery for *copia* suppression, but how this suppression is mediated remains unclear.

The only other well-studied element in the *pseudoviridae* family of retroelements is 1731, which offers insight as to how the nuage-piRNA machinery might silence TE transcripts. Much like *copia*, transcription of 1731 is strongly upregulated in the early stages of spermatogenesis but 1731 VLPs are not assembled immediately after translation (Haoudi *et al.*, 1995). Rather, the VLPs are observed in the nuclei during the late stages of spermatogenesis, in which chromatin condensation would likely make transpositions impossible. Rachidi *et al.* (2005) interpreted these observations by reasoning that, perhaps, 1731 transcripts are silenced or sequestered in the nuage to alleviate detrimental effects of transposition during critical stages of germ cell development. Overall, this picture is consistent with the post-transcriptional blocks of *copia* transpositions that we have analyzed here.

Our experimental results suggest that translational silencing of *copia* elements through the nuage-piRNA machinery seems to be a likely hypothesis for host suppression. In our quantitative study of *copia* plasmid concentration (Nuzhdin *et al.*, 1996), we detected *copia* plasmids in the active 2b but detected little to no *copia* plasmids in the inactive Ore line. Consistent with these results, we saw VLPs present in the 2b line but none were observed in the Ore line in the testes electron microscopy images. Both *copia* and 1731 VLPs are self-assembled in the nucleus through an autocatalytic process (Yoshioka *et al.*, 1990), which serves as a site for mRNA reverse transcription into linear and plasmid DNA. Therefore, *copia* silencing is likely to occur at the post-transcription blocks of spermatogenesis, in which *copia* transcripts may be sequestered in the nuage by complex interactions among piRNAs and its associated proteins.

In the experiments reported here, we have clarified the different stages at which transpositions of *pseudoviridae* TEs may be blocked by providing an analogous pathway to better understand the potential *copia* control mechanism. We mapped regions of the host genome controlling *copia* transpositions, and identified candidate genes and potential pathways that could explain how *copia* transposition is controlled during critical stages of germ cell development. We know that at least some of the mechanisms of TE suppression are post-reverse-transcriptional because *copia* transpositions are blocked at multiple stages with several regions from different chromosomes that interact to suppress or enable *copia* mobility.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Andrew Flavell for the gift of plasmid with 15b insertion between *copia* LTRs, and Kristy Harmon for excellent technical assistance. This work was supported by NSF DEB-0316513 and GM076051.

References

- Aravin AA, Hannon GJ, Brennecke J (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arm race. *Science* **318**: 761–764.
- Aravin AA, Klenov MS, Vagin VV, Bantignes F, Cavalli G, Gvozdev VA (2004). Dissection of a natural RNA silencing process in the *Drosophila melanogaster* germ line. *Mol Cell Biol* **24**: 6742–6750.
- Basten C, Weir BS, Zeng ZB (1997). *QTL Cartographer: A Reference Manual and Tutorial for QTL Mapping*. Department of Statistics, North Carolina State University: Raleigh, NC (<http://statgen.ncsu.edu/qtlcart/>).
- Biémont C (1994). Dynamic equilibrium between insertion and excision of P elements in highly inbred lines from an M' strain of *Drosophila melanogaster*. *J Mol Evol* **39**: 466–472.
- Boeke JD, Chapman KB (1991). Retrotransposition mechanisms. *Curr Opin Cell Biol* **3**: 502–508.
- Brookfield JF, Badge RM (1997). Population genetics models of transposable elements. *Genetica* **100**: 281–294.
- Brookfield JF, Montgomery E, Langley CH (1984). Apparent absence of transposable elements related to the P elements *D. melanogaster* in other species of *Drosophila*. *Nature* **310**: 330–332.
- Blumenstiel JP, Hartl DL (2005). Evidence for maternally transmitted small interfering RNA in the repression of transposition in *Drosophila virilis*. *PNAS* **102**: 15965–15970.
- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R *et al.* (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**: 1089–1103.
- Bucheton A (1995). The relationship between the flamenco gene and gypsy in *Drosophila*: how to tame a retrovirus. *Trends Genet* **11**: 349–353.
- Buchon N, Vauray C (2006). RNAi: a defensive RNA-silencing against viruses and transposable elements. *Heredity* **96**: 195–202.
- Charlesworth B, Langley CH (1989). The population genetics of *Drosophila* transposable elements. *Annu Rev Genet* **23**: 251–287.
- Charlesworth B, Sniegowski P, Stephan W (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**: 215–220.
- Csirik AK, McDonald JF (1995). Analysis of copia sequence variation within and between *Drosophila* species. *Mol Biol Evol* **12**: 83–93.
- Eggleston WB, Johnson-Schlitz DM, Engels WR (1988). P-M hybrid dysgenesis does not mobilize other transposable element families in *Drosophila melanogaster*. *Nature* **331**: 368–370.
- Finnegan DJ (1978). Repeated gene families in *Drosophila melanogaster*. *Cold Spring Harb Symp Quant Biol* **42**: 1053–1063.
- Flavell AJ (1984). Role of reverse transcription in the generation of extrachromosomal *copia* mobile genetic elements. *Nature* **310**: 514–516.
- Flavell AJ, Ishorowicz D (1981). Extrachromosomal circular copies of the eukaryotic transposable element *copia* in cultured *Drosophila* cells. *Nature* **292**: 591–595.
- Flavell AJ, Ish-Horowicz D (1983). The origin of extrachromosomal circular *copia* elements. *Cell* **34**: 415–419.

- Grivna ST, Beyret E, Wang Z, Lin H (2006). A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* **20**: 1709–1714.
- Haoudi A, Kim MH, Champion S, Best-Balpomme M, Maisonhaute C (1995). The gag polypeptides of the *Drosophila* 1731 retrotransposon are associated to virus-like particles and to nuclei. *FEBS Lett* **377**: 67–72.
- Harada K, Yukuhiro K, Mukai T (1990). Transposition rates of movable genetic elements in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **87**: 3248–3252.
- Houle D, Nuzhdin SV (2004). Mutation accumulation and the effect of copia insertions in *Drosophila melanogaster*. *Genet Res* **83**: 7–18.
- Hutvagner G, Zamore PD (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**: 2056–2060.
- Klattenhoff C, Theurkauf W (2008). Biogenesis and germline functions of piRNAs. *Development* **135**: 3–9.
- Klenov MS, Lavrov SA, Stolyarenko AD, Ryazansky SS, Aravin AA, Tuschl T *et al.* (2007). Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the *Drosophila melanogaster* germline. *Nucleic Acids Res* **35**: 5430–5438.
- Kotaja N, Sassone-Corsi P (2007). The chromatoid body: a germ-cell-specific RNA-processing center. *Nat Rev* **8**: 85–90.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860–921.
- Lefevre G (1976). A photographic representation and interpretation of the polytene chromosomes of *Drosophila melanogaster* salivary glands. In: *The Genetics and Biology of Drosophila*. Academic Press: London, pp 31–66.
- Lim AK, Kai T (2007). Unique germ-line organelle nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **104**: 6714–6719.
- Lohe AR, Hartl DL (1996). Germline transformation of *Drosophila Virillus* with the transposable element mariner. *Genetics* **143**: 365–374.
- Mackay TFC, Lyman RF, Jackson MS (1992). Effects of P element insertions on quantitative traits in *Drosophila melanogaster*. *Genetics* **130**: 315–332.
- Marzo M, Puig M, Ruiz A (2008). The foldback-like element Galileo belongs to the P superfamily of DNA transposons and is widespread within the *Drosophila* genus. *Proc Natl Acad Sci USA* **105**: 2957–2962.
- Mével-Ninio M, Pelisson A, Kinder J, Campos AR, Bucheton A (2007). The *flamenco* locus controls the gypsy and ZAM retroviruses and is required for *Drosophila* oogenesis. *Genetics* **175**: 1615–1624.
- Nuzhdin SV (1999). Sure facts, speculations, and open questions about the evolution of transposable element copy number. *Genetica* **107**: 129–137.
- Nuzhdin SV, Mackay TFC (1994). Direct determination of retrotransposon transposition rates in *Drosophila melanogaster*. *Genet Res* **63**: 139–144.
- Nuzhdin SV, Mackay TFC (1995). The genomic rate of transposable element movement in *Drosophila melanogaster*. *Mol Biol Evol* **12**: 180–181.
- Nuzhdin SV, Pasyukova EG (1991). New approach to polygene mapping: Location of genes controlling fitness level in related *Drosophila melanogaster* stocks. *Genetika (Russ)* **27**: 849–859.
- Nuzhdin SV, Pasyukova EG, Mackay TFC (1996). Positive association between copia transposition rate and copy number in *Drosophila melanogaster*. *Proc Roy Soc Lond B* **263**: 823–831.
- Nuzhdin SV, Pasyukova EG, Mackay TFC (1997). Accumulation of transposable elements in laboratory lines of *Drosophila melanogaster*. *Genetica* **100**: 167–175.
- Nuzhdin SV, Pasyukova EG, Morozova EA, Flavell AJ (1998). Quantitative genetic analysis of copia retrotransposon activity in inbred *Drosophila melanogaster* lines. *Genetics* **150**: 755–766.
- O'Donnell DA, Boeke JD (2007). Mighty piwis defend the germline against genome intruders. *Cell* **129**: 37–44.
- Pasyukova E, Nuzhdin SV, Li W, Flavell AJ (1997). Germ line transposition of the copia retrotransposon in *Drosophila melanogaster* is restricted to males by tissue-specific control of copia RNA levels. *Mol Gen Genet* **255**: 115–124.
- Pasyukova EG, Nuzhdin S (1993). Doc and copia instability in an isogenic *Drosophila melanogaster* stock. *Mol Gen Genet* **240**: 302–306.
- Pasyukova EG, Nuzhdin SV, Filatov DA (1998). The relationship between the rate of transposition and transposable element copy number for copia and Doc retrotransposons of *Drosophila melanogaster*. *Gen Res* **72**: 1–11.
- Parvinen M (2005). The chromatoid body in spermatogenesis. *Int J Androl* **28**: 189–201.
- Pelisson A, Sarot E, Payen-Groschene G, Bucheton A (2007). A novel repeat-associated small interfering RNA-mediated silencing pathway down regulates complementary-sense gypsy transcripts in somatic cells of the *Drosophila* ovary. *J Virol* **81**: 1951–1960.
- Pelisson A, Teyssset L, Chalvet F, Kim A, Prud'homme N, Terzian C *et al.* (1997). About the origin of retroviruses and the coevolution of the gypsy retrovirus with the *Drosophila flamenco* host gene. *Genetica* **100**: 29–37.
- Perdue S, Nuzhdin SV (2000). Master copy is not responsible for the high rate of copia transposition in *Drosophila*. *Mol Biol Evol* **17**: 984–986.
- Peters L, Meister G (2007). Argonaute proteins: mediators of RNA silencing. *Mol Cell* **26**: 611–623.
- Peterson-Burch BD, Voytas DF (2002). Genes of the pseudoviridae (Ty1/copia) retrotransposons. *Mol Biol Evol* **19**: 1832–1845.
- Prud'homme N, Gans M, Masson M, Terzian C, Butehton A (1995). *Flamenco*, a gene controlling the gypsy retrovirus of *Drosophila melanogaster*. *Genetics* **139**: 697–711.
- Rachidi M, Lopes C, Benichou JC, Heliou R, Maisonhaute C (2005). Virus-like particle formation in *Drosophila melanogaster* germ cells suggests a complex translational regulation of the retrotransposon cycle and new mechanisms inhibiting transposition. *Cyto Gen Res* **111**: 88–95.
- SAS/STAT user's guide, version 6, 4th edn, Vol. 2. Cary, N.C.: SAS Institute, 1989, pp 1027–1069.
- Seto AG, Kingston RE, Lau NC (2007). The coming of age for piwi proteins molecular. *Cell* **26**: 611–623.
- Shrimpton A, Montgomery EA, Langley CH (1986). OM mutations in *Drosophila ananassae* are linked to insertions of transposable elements. *Genetics* **114**: 125–135.
- Simmons MJ, Ryzek DF, Lamour C, Goodman JW, Kummer NE, Merriman PJ (2007). Cytotype regulation by telomeric P elements in *Drosophila melanogaster*: evidence for involvement of an RNA interference gene. *Genetics* **176**: 1945–1955.
- Shpiz S, Kwon D, Uneva A, Kim M, Klenove M, Rozovsky Y *et al.* (2007). Characterization of *Drosophila* telomeric retroelement TAHRE: transcription, transpositions, and RNAi-based regulation of expression. *Mol Biol Evol* **24**: 2535–2545.
- Styhler S, Nakamura A, Swan A, Suter B, Lasko P (1998). *vasa* is required for GURKEN accumulation in the oocyte, and is involved in oocyte differentiation and germline cyst development. *Development* **125**: 1569–1578.
- Xing JC, Witherspoon DJ, Ray DA, Batzer MA, Jorde LB (2007). Mobile DNA elements in primates and human evolution. *Am J Phys Anthro* **134**: 2–19.
- Yoshioka K, Fujita A, Kondo S, Miyake T, Sakaki Y, Shiba T (1992). Production of a unique multi-lamella structure in the nuclei of yeast expressing *Drosophila copia* gag precursor. *FEBS Lett* **302**: 5–7.
- Yoshioka K, Honma H, Zushi M, Kondo S, Togashi S, Miyake T *et al.* (1990). Virus-like particle formation of *Drosophila copia* through autocatalytic processing. *EMBO J* **9**: 535–541.
- Zeng ZB (1994). Precision mapping of quantitative trait loci. *Genetics* **136**: 1457–1468.