Quantitative Genetic Analysis of *copia* Retrotransposon Activity in Inbred *Drosophila melanogaster* Lines

Sergey V. Nuzhdin,* Elena G. Pasyukova,† Elena A. Morozova† and Andrew J. Flavell‡

**Section of Evolution and Ecology, University of California, Davis, California 95616-5755, †Department of Animal Molecular Genetics, Institute of Molecular Genetics of the Russian Academy of Science, Moscow 123182, Russia and ‡Department of Biochemistry, University of Dundee, Dundee DD1 4HN, Scotland, United Kingdom*

Manuscript received April 9, 1998
Accepted for publication July 6, 1998

ABSTRACT

The rates of transcription and transposition of retrotransposons vary between lines of *Drosophila melanogaster*. We have studied the genetics of differences in *copia* retrotransposon activity by quantitative trait loci (QTL) mapping. Ninety-eight recombinant inbred lines were constructed from two parental lines exhibiting a 10-fold difference in *copia* transcript level and a 100-fold difference in transposition rate. The lines were scored for 126 molecular markers, *copia* transcript level, and rate of *copia* transposition. Transcript level correlated with *copia* copy number, and the difference in *copia* copy number between parental lines accounted for 45.1% of *copia* transcript-level difference. Most of the remaining difference was accounted for by two transcript-level QTL mapping to cytological positions 27B–30D and 50F–57C on the second chromosome, which accounted for 11.5 and 30.4%, respectively. *copia* transposition rate was controlled by interacting QTL mapping to the region 27B–48D on the second and 61A–65A and 97D–100A on the third chromosome. The genes controlling *copia* transcript level are thus not necessarily those involved in controlling *copia* transposition rate. Segregation of modifying genes, rather than mutations, might explain the variability in *copia* retrotransposon activity between lines.

TRANSPOSABLE elements (TEs) are DNA sequences capable of multiplying in their host genome (Berg and Howe 1989). They survive by increasing copy numbers due to transpositions, and natural selection eliminates them because of lower fitness of hosts carrying heavier loads of TEs. The available phylogenetic evidence supports the view that TEs have existed in living organisms for hundreds of millions of years. A fundamental question facing the field is how can equilibrium be attained between transposition and selection that allows these parasitic genetic elements to persist for such a time period (Charlesworth and Langley 1989; Charlesworth *et al.* 1994a)?

Drosophila represents a model species for studying TE population dynamics (Ashburner 1989). Some TE families may be horizontally transferred from distant Drosophila, and other families may lose activity (Bucheton *et al.* 1992; Clark *et al.* 1994; Lohe *et al.* 1995); however, the most common mode of TE transmission is vertical inheritance when transposition-selection are at equilibrium, *i.e.*,

$$-V\frac{\partial \ln \bar{w}}{\partial \bar{n}} = \bar{n}(u - v), \qquad (1)$$

where $\bar{n}$ is the mean population copy number, u is the rate of transposition per element, v is the rate of excision per element, V is the variance in copy number, and $\bar{w}$ is the fitness (Charlesworth and Charlesworth 1983; Kaplan and Brookfield 1983). For an equilibrium to be stable, either v must be an increasing function of n, u must be a decreasing function of n, or fitness must be a sharply decreasing function of n. The first and second cases referred to as self-regulation of TE copy number have been described for TEs, which are causative agents of the phenomenon of hybrid dysgenesis (Bucheton *et al.* 1976; Kidwell *et al.* 1977; Pelisson and Bregliano 1987; Biemont 1994; Lohe and Hartl 1996; Petrov *et al.* 1995). The third case, referred to as copy number control by selection, is the most common type of control in *D. melanogaster* and will be dealt with in detail here.

The description of the balance between transposition and selection requires estimates of both processes. The rate of transposition per element has been inferred as 10^{-5}–10^{-4} from the frequency distribution of TE occupation sites between individuals of natural populations (Montgomery and Langley 1983; Leigh-Brown and Moss 1987; Charlesworth and Lapid 1989; Charlesworth *et al.* 1992a,b; Nuzhdin 1995) and from direct counting of transpositions over time in inbred laboratory lines (Eggleston *et al.* 1988; Harada *et al.* 1990; Nuzhdin and Mackay 1995). Selection washes TEs out of genomes because fitness of hosts with more TEs is decreased. First, TE insertions damage genes either directly if they insert into transcribed regions or indirectly

Corresponding author: Sergey V. Nuzhdin, Section of Evolution and Ecology, University of California, Davis, CA 95616-5755.
E-mail: svnuzhdin@ucdavis.edu

by *cis* effects of TE expression signals (promoters, enhancers, *etc.*) on host gene regulatory elements (Finnegan 1992; McDonald *et al.* 1997). The presence of more TE copies is associated with more deleterious mutations (Mackay *et al.* 1992). Second, TE inserts situated in different positions recombine with each other, inducing deleterious deletion and duplication of sequences situated between copies. Hosts with higher copy number of any TE family may have decreased fitness because the probability of ectopic exchange between homologue copies is increased (Langley *et al.* 1988). Third, the deleterious effect of TE presence in the genome may be mediated through TE expression (Nuzhdin *et al.* 1996), and the harm to the host might increase with TE copy number (Brookfield 1991, 1996). Although the relative contributions of these mechanisms to the TE copy number control have been extensively studied (further discussion may be found in Langley *et al.* 1988; Ajioka and Eanes 1989; Charlesworth and Lapid 1989; Charlesworth 1991; Montgomery *et al.* 1991; Charlesworth *et al.* 1992a,b; Eanes *et al.* 1992; Biemont *et al.* 1994; Charlesworth *et al.* 1994a,b; Sniegowski and Charlesworth 1994; Hoogland and Biemont 1996; Nuzhdin *et al.* 1996), no agreement on which mechanism(s) dominates has been attained (Biemont *et al.* 1997; Charlesworth *et al.* 1997).

The transposition activity of the members of a TE family should be thought of as an evolvable parameter determined by TE sequences (Csink and McDonald 1990, 1995; Kim *et al.* 1994; Pelisson *et al.* 1994) and by alleles of host genes involved in transposition (Kim *et al.* 1990; Pelisson *et al.* 1994). Because the rate of transposition varies, its value may be modified in the process of host-TE coevolution driven by competition between elements and by selection-checking fitness of the host. Different models of TE copy number maintenance predict different ways of coevolution. If TE expression is a main source of lower fitness of a host with higher TE copy number, natural selection will favor less active copies (Brookfield 1991, 1996). If the other selection mechanisms prevail, more active copies will be selected (Charlesworth and Langley 1986). Natural selection always favors alleles of host genes that suppress TE activity. However, the selection is negligible (at least when u is in the order of 10^{-5}–10^{-4}; Charlesworth and Langley 1986) in accordance with the first and second models of TE containment. Selection is strong in accordance with the third mechanism (Brookfield 1991, 1996).

Because the relative contributions of the selection mechanisms are currently unknown, the process of host-TE coevolution remains uncharacterized. Describing host genes involved in the restriction of TE activity, understanding their mechanism of action, and the maintenance of their polymorphism in nature may shed light on the process of TE-host genome coevolution and, correspondingly, on the relative contributions of different selection forces that drive the coevolution. One way of approaching this problem is to use a genetic and molecular analysis of factors that account for differences in TE activity. The rate of *copia* transposition, which we have chosen as a model, is variable between *D. melanogaster* lines. For example, the same set of 28 fixed sites was detected in the Oregon R (Ore) line over 5 yr, giving zero as the estimate for *copia* transposition rate in this line (Nuzhdin *et al.* 1996; Pasyukova *et al.* 1997). However, in one isogenic line, 2b (Pasyukova and Nuzhdin 1992), a high *copia* transposition rate (10^{-3}–10^{-2}) was discovered in 1991 and has continued over a period of 6 yr (Pasyukova *et al.* 1998).

Here, we analyze the genetic factors that modulate the differences in *copia* transposition and transcript levels between 2b and Ore. These are quantitative traits that may be modified by segregating alleles at many loci. Therefore, we have used approaches designed to map quantitative trait loci (QTL). Our results show that the elevated transposition rate in 2b is due to at least three QTL and by the increased *copia* copy number. We also show that at least two of these QTL have little or no effect on *copia* transcript accumulation. The control of *copia* transposition is thus a multifactorial process that may be imposed on different intermediates of the retrotransposition process.

MATERIALS AND METHODS

***in situ* hybridization:** Transposable element insertion sites were determined by *in situ* hybridization of the plasmid cDM5002 containing a full-length *copia* transposable element (Finnegan *et al.* 1978) to polytene salivary gland chromosomes of third-instar larvae (Shrimpton *et al.* 1986). Probes were labeled with biotinylated dATP (bio-7-dATP) (Bio-Rad Labs, Hercules, CA) by nick translation. Hybridization was detected using the Vectastain ABC kit (Vector Labs, Burlingame, CA) and visualized with diaminobenzidine. The element locations were determined at the level of cytological bands on the standard Bridge's map (Lefevre 1976).

Direct measurement of *copia* transposition rate: Transpositions were detected as described by Pasyukova and Nuzhdin (1992) and Nuzhdin *et al.* (1996). Briefly, one male of the tested line was crossed with the Ore females, and *copia* insertion sites in ~40 progeny larvae were scored by *in situ* hybridization. Each of the F_1 progeny inherits one chromosome set from the maternal Ore and the other chromosome set from the tested male parent. The positions of *copia* in the father may thus be reconstructed from the segregation of sites in F_1. *copia* transpositions and excisions in the germ lines of tested males are detected by the appearance of nonparental *copia* insertion sites or the loss of parental insertion sites, respectively, in the progeny larvae. The transposition rate in a given male is calculated as (number of transpositions)/[(number of gametes analyzed) × (*copia* copy number)].

***D. melanogaster* lines:** The isogenic lines 2b (Pasyukova *et al.* 1998) and Ore (Nuzhdin *et al.* 1996) were used as parental stocks. The substitution lines 2b, Ore, Ore; Ore, 2b, Ore; and Ore, Ore, 2b (where the origin of the three major complementation groups is represented by the line names) were made by crosses with the balancer stock *ln(1)sc^{s1L} sc^{8R+l}9n(1)s, sc^{s1} sc^{8}w^{a}B; ln(2LR)SM1, al Cy cn^{2} sp^{2}/ln(2LR)Pm; ln(3LR)Ubx^{130},*

Ubx^{130} *e*^{s}*/Sb; spa*^{pol}. Substitution of the *spa*^{pol} chromosome (Ore line bears *spa*^{pol}) into the 2b line did not influence the *copia* activity (E. Pasyukova and S. Nuzhdin, unpublished results).

copia is fixed in Ore in positions 11C, 21D, 34B, 34F, 42B, 42C, 47A, 52A, 57E, 59D, 69B, 75C, 86E, and 96A, and in 2b in positions 4EF, 18C, 33F, 34EF, 35C, 38A, 40A, 41A, 42A, 42B, 47B, 52D, 53E, 64E, 66C, 71C, 73C, 80A, 82C, 87F, 94E, 96B, 98B, and 99B. In 2b, there are also segregating *copia*s with variable locations due to frequent transpositions. The construction of the recombinant inbred (RI) lines was started from eight individual crosses of 2b males with Ore females. From 9 to 40 F_1 larvae progeny per cross were scored by *in situ* hybridization to reconstruct *copia* positions in the parental males and to infer the rate of *copia* transposition in them (data not shown). Female progeny of the male with homozygous (or hemizygous) *copia* at 5D, 9B, 11A, 16C, 18F, 43A, 50F, 58C, 60D, 94A, 99A, 99E, 100E, and 102B (*copia* positions fixed in 2b are excluded) and heterozygous *copia* at 29D, 30B, 30C, 52D, 65F, 67C, 67D, 67E, 70A, and 85D were backcrossed with another male of the 2b line. Similarly, *copia* positions on the chromosomes of this male were reconstructed from *in situ* analysis of *copia* positions of its progeny at 1F, 3C, 3E, 5D, 6F, 9B, 9F, 11A, 16C, 43A, 50F, 60D, 70A, and 94A (homozygous) and 25A, 29D, 30B, 54D, 55D, 56D, 58A, 99A, and 100E (heterozygous). After four generations of random mating, the full-sib lines were established. Ninety-eight of the full-sib lines survived inbreeding for 25 generations (RI lines) and were maintained as small mass cultures thereafter. The genetic constitution of each RI line was determined by the analysis of 92 *roo* TE polymorphic markers that are fixed within the parental lines but segregate between them (Nuzhdin *et al.* 1997a).

RNA analysis: RNAs were extracted from four to six crawling third-instar larvae by disruption with a mini-pestle in a microfuge tube in 0.5 ml homogenization buffer (100 mm Tris-HCl pH 7.5, 10 mm EDTA, 0.35 m NaCl, 2% SDS, 7 m urea; Savakis *et al.* 1986). Phenol-chloroform (0.5 ml) and TE (0.25 ml) were added, the samples were vortexed and centrifuged, and the supernatant was re-extracted with phenol-chloroform. RNAs were collected by ethanol precipitation and stored as ethanol suspensions until use. Equal volumes of these RNAs (typically 30% of the total preparation) were spun down before use, redissolved in 5 μl H_2O, and subjected to formaldehyde Northern gel analysis (Sambrook *et al.* 1989). Northern gels were blotted to Biodyne A nylon membrane (Pall) using the manufacturer's recommendations. The blots were first probed with a *copia* fragment extending rightward from the *Apa*I site at 283 to the end of the element. The bound *copia* probe was quantitated by phosphorimager (Molecular Dynamics model 400S). The blots were stripped of bound probe, reexposed to ensure signal loss, and then reprobed with a plasmid containing a complete Drosophila protein phosphatase 1a (*PP1a*) cDNA as a loading control (Dombradi *et al.* 1989). The bound *PP1a* probe was quantitated as before. The *copia* transcript values were then normalized to the *PP1a* signal to allow for variations in RNA loading between samples. These values were then expressed as percentages of one of samples extracted from the 2b parental line (this particular RNA was loaded onto all four gels). Most transcription estimates are the averages of 2 measurements, 6 came from 3 measurements and 14 were a single measurement. Transcript levels were highly consistent across measurements (Pearson product-moment correlation, $r = 0.85$, $P < 0.0001$).

Data analysis: Seventy-six cytological markers (17 markers out of 92 were completely linked with neighboring markers and were excluded from the analysis; Nuzhdin *et al.* 1997a) were used for the analysis of trait marker associations by the composite interval mapping technique implemented in the QTL Cartographer program (Version 1.09a, J. C. Basten, B. S. Weir and Z.-B. Zeng, 1996, ftp://esssjp.stat.ncsu.edu/pub/qtlcart/). The parameters were 6 (model), 5 (number of markers included in the multiple regression), and 10 (window size). Recombination distances between markers were calculated by Kosambi transformation of rates calculated as $r = 1/(4/R - 6)$, where R is the proportion of RI lines that are recombinant between neighboring markers (Silver 1985). Because residual heterozygosity of RI lines is 4%, some markers segregated within the lines and were considered missing data.

There were 31 *copia* sites fixed within, but polymorphic between, parental lines 4EF, 11C, 18C, 21D, 33F, 34B, 35C, 38A, 42C, 43A, 47A, 47B, 50F, 53E, 57E, 59D, 60D, 64E, 66C, 69B, 71C, 73C, 75C, 82C, 86E, 87F, 94E, 96A, 96B, 98B, and 99B. They were used as additional molecular markers. *copia* markers were scored in only one individual per RI line, and no information about their homozygosity or heterozygosity within the RI line was available because the site appearance on polytene chromosomes is dominant. These markers were considered fixed within the lines. This incorporated 1.5-fold upbias for the recombination distances, as inferred from *copia* markers compared to *roo* markers. A sample of data on the marker genotypes, *copia* positions, and transcript levels in 10 RI lines are shown in Table 1; the rest of the data will be provided upon request.

Correlations, regressions, and residuals were estimated using SAS procedures CORR, REG, and GLM (SAS Institute, Inc., 1988).

RESULTS

Genetics of the differences in *copia* transposition: First, we assessed the effect of whole chromosomes on transposition rate. If one mutant gene in the 2b line was responsible for switching transpositions on, there would be transpositions in the synthetic line, with just one chromosome carrying the "instability" allele and the rest of the genome from the Ore line. The substitution lines 2b, Ore, Ore; Ore, 2b, Ore; and Ore, Ore, 2b (where the origin of the three major *D. melanogaster* complementation groups is represented by the line names) were constructed. All the lines were maintained as small mass cultures (~20 pairs of flies per generation). After 25 generations *copia* positions were analyzed in 10 larvae per line. *copia* sites in each line did not vary among individuals within lines and represented a combination of all Ore sites on Ore-originated chromosomes and all 2b sites (fixed in 2b plus from 2 to 6 fixed extra sites that segregated in 2b but became homozygous during chromosome substitutions) in 2b-originated chromosomes. In total, there were 38 *copia* copies in the line 2b, Ore, Ore; 44 copies in the line Ore, 2b, Ore; and 46 copies in the line Ore, Ore, 2b. No additional sites that could be interpreted as resulting from transposition were detected. Under the conditions of the above experiment we would have expected to see between 10 and 100 transpositions if the *copia* transposition rate was similar to the one in 2b (10^{-3}–10^{-2}; Pasyukova *et al.* 1998). Because no transpositions were found, we conclude that none of the 2b chromosomes can solely account for frequent *copia* transpositions in the

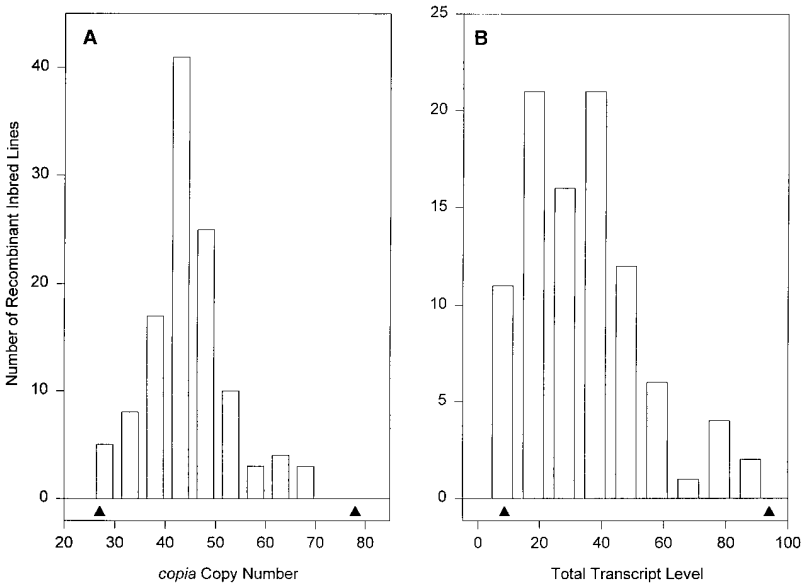

Figure 1.—The distribution of *copia* copy number and transcript level among recombinant inbred lines. Triangles, values characteristic for Ore and 2b lines. *copia* copy number in 2b line is the average between two parental 2b males.

2b line. Factors situated on more than one chromosome must be simultaneously present to allow *copia* transposition.

The power to detect interacting polygenes is maximized with the design of RI lines. We made 98 RI lines, each of which contains a composite genome derived from 2b and Ore. The lines originated from a cross of one 2b male with Ore females and a backcross of F_1 female progeny to another 2b male. These lines were scored for molecular markers to identify the origin of different regions of the genome (Nuzhdin *et al.* 1997a). There was a high rate of *copia* transposition in the germ cells of the first male from which the RI lines originated. The germ-line *copia* transposition rate was not measured for the second male. However, transpositions were found in each of seven other 2b males at the time of these crosses with the average rate 1.9×10^{-2} ($\sigma = 0.8 \times 10^{-2}$). The rate of *copia* transposition in the Ore line is zero (Nuzhdin *et al.* 1996; Pasyukova *et al.* 1997).

To screen the RI lines for *copia* instability, *copia* positions were scored in one individual per RI line and compared with known original positions in the two parental males and in the Ore line (see materials and methods). *copia* copy numbers in RI lines were distributed between the parental values (27 in Ore and 78 in 2b), with more lines having smaller than the midparental number of *copia* copies (Figure 1A). *copia* transpositions (occupation of nonoriginal positions) were found in 7 of the 98 lines. In some of these lines the same new insertions were detected (57C in two lines and 52F in three lines). Because *copia* transpositions demonstrate little specificity (Nuzhdin and Mackay 1994; Pasyukova *et al.* 1998), these instances probably resulted from transpositions in the germ cells of the two parental 2b males. In three lines, new unique positions were found (2B, 12A, 12E, 17A, 44A, 46A, 56F, 59E, 62B, 63C, 63F, 84D, and 85A). In two lines, new *copia* positions were detected in both 2b and Ore-originated genomic regions. The transpositions accumulated in the Ore-originated regions could not happen in parental males, but have happened during or after the RI line construction. Potential instability must be associated with the 2b-derived regions. The lines shared 26F–48D, 61A–65A, 68B–68C, and 97D–100A derived from 2b (the first region segregated in one line that was dropped from further analysis). We took all other RI lines in which the above regions were from 2b (in one of the extra lines, the 68B–68C interval originated from Ore). We hypothesized that *copia* is currently active in all five identified potentially active RI lines (Table 1).

To exclude the chance that transpositions occurred during the establishment of the active lines while they segregated for unknown parts of 2b and Ore genomes, we retested the lines for instability. One male per line was crossed with Ore females, and his 38–42 progeny was scored for *copia* transpositions by *in situ* hybridization (191 in total). Nine *copia* transpositions (12D, 26B, 38F, 57F, 60B, 62B, 88D, 89B, and 95C) were found in three lines (including the line with the interval 68B-68C from Ore). *copia* transpositions were not found in the males of the two other putatively active RI lines. This is not surprising because the rates of transposition in the active RI lines were relatively low, and the occurrence of transpositions in single germ lines is sporadic. Consistent with this, three new *copia* sites (71F, 91B, and 96D) appeared between the time of original screening of *copia* and the measurement of transposition rate in the fourth line. We also screened *copia* transpositions in 34–40 progeny of males of five arbitrarily chosen control lines (181 progeny in total) that did not share the genotype similarity (Table 1). No nonoriginal *copia* positions were found in the latter lines, and no transpositions were detected.

Accumulation of *copia* in new positions and direct

TABLE 1

***Copia* copy number (*n*), number of accumulated copies, *copia* transcript level (*T*) and genotypes of lines**

Line no.	*n*	Accumulated copies[a]	*T* (%)	Genotype[c] 1	11	21	31	41	51	61	71	81	91
36[d]	60	*2B*[b], *12A*[b],*17A*[b], *44A*, 57C, *59E, 62B, 63C, 63F, 85A*[b]	19	oooooooooo	oooooooooo	ooooobbbbb	bbbbbhbbbb	boooooooob	bbboobbooo	oooooooooo	oooooooooo	ooooobbbbb	bbb
56	46		47	ooooobbbbb	bboooooooo	oobbbbbbbb	bbbbbbbbbb	boooobbbbb	bbbbbbbbhh	hhhhhhhhhh	hhhhhhhbbb	bbbbbbbbbb	bbb
57[d]	66		74	ooooobbbbb	bboooooooo	ooooobbbbb	bbbbbbbbbb	boooobbbbb	bbbooooobb	bohooooooo	ooooooobbb	bbbbbbbbbb	bbb
64[d]	49		23	oooooooooo	oooooooooo	ooooobbbbb	bbbbbbbbbb	ooooobbbbb	bbbbbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbb
66[d]	61	*46A*, *84D*	37	oooooooooo	oooooooooo	oooohbbbbb	bbbbbbbbbb	ooooooooob	bbbbbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbb
93	43		12	oooooooooo	oooooooooo	oooobooooo	oooooooooo	oooobooooo	oobbbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbbbbbbbbo	oob
94	55		58	oooooooooo	oooooooooo	oooohbbbhh	hhoohohoho	oooohhhhoo	oobbbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbbbbbbbbo	bbb
95	45		28	hbooohhooo	oooooooooo	oooooooooo	oooooooooo	ooooobbbbb	bbbbbbbbbb	bbbbbbbbbo	bbbbbbbboo	oooobobobo	bbb
97	39		52	oooooooooo	oooooooooo	oooooobbhh	hhoohohoho	oooohbbbbh	hoobooohoh	oohoohohhh	oooohooooo	oobbbbbbbo	oob
98	53		88	oooooooooo	oooooooooo	oobbbbbboo	ohhhhhhhhh	ooooooohoo	ooobbbbbbb	bbbbbbbbbb	bbbbbbbbbb	bbhbbbbboo	oob

[a] Italics, transpositions found in only one RI line.

[b] Transpositions into Ore-R-originated regions.

[c] *X* chromosome markers (1–22): 1B, 3E, 4F, 5D, 6E, 7D, 7E, 9A, 10D, 11C, 11D, 12E, 14C, 15A, 15B, 15F, 16A, 16D, 17C, 18A, 19A, and 19C. Chromosome *2* markers (23–49): 21E, 22F, 27B, 29F, 30A, 30D, 33E, 34EF, 35BC, 38A, 38E, 39A, 43A, 43E, 44C, 44F, 46A, 46C, 48D, 49D, 50B, 50D, 50F, 57C, 57D, 57F, and 60E. Chromosome *3* markers (50–92): 61A, 61D, 63A, 65A, 65D, 67D, 68B, 68C, 69D, 70C, 71E, 72A, 73D, 76A, 76B, 77A, 77E, 78D, 79C, 79E, 82D, 85A, 85F, 87B, 87E, 87F, 88E, 89B, 91A, 91D, 92A, 93A, 93B, 94D, 96A, 96F, 97D, 97E, 98A, 99A, 99B, 99E, and 100A. Chromosome *4* marker (93): *spa*[pol]. o and b, marker alleles from Ore-R and 2b, respectively; h, heterozygous markers. Underline, intervals shared by "active lines".

[d] Numbers of lines in which transpositions were directly observed (see text for more explanations).

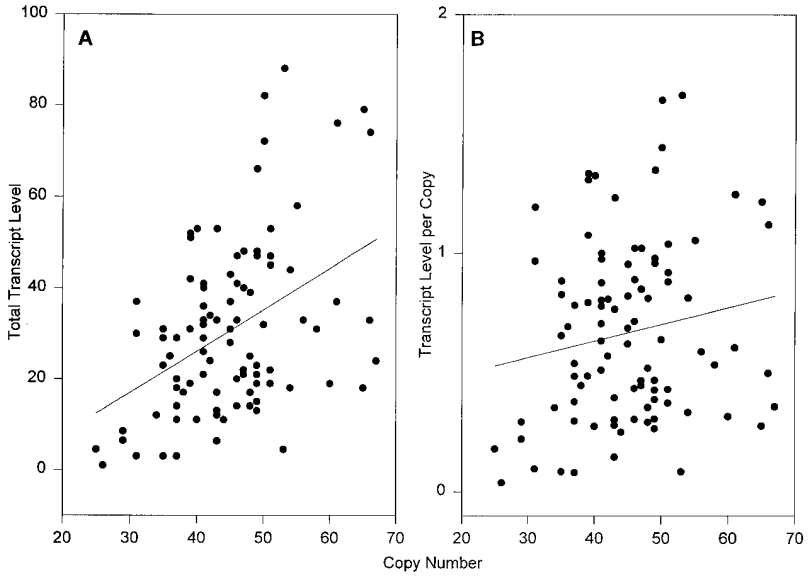

Figure 2.—Correlation between *copia* transcript level and copy number.

evidence of ongoing transpositions were found only in the lines sharing descent of the intervals 27B-48D, 61A-65A, and 97D-100A from the 2b line. Combination of the 2b alleles of QTL located in the second-chromosome interval, and in at least one of the third-chromosome interval, is required for *copia* transposition.

Genetics of the differences in *copia* transcription: Pasyukova *et al.* (1997) reported an approximate 10-fold difference in *copia* transcript level between the parental lines for this study, 2b and Ore. This difference was confirmed when we remeasured *copia* transcript levels. The average transcript level for Ore is 8.6% and for 2b is 94.0% (all RNA values were normalized to one of four duplicate 2b samples, which was given an arbitrary 100% score; see materials and methods).

copia RNA levels were assayed in the RI lines. The transcript level was close to the Ore parent in 11 RI lines and close to the 2b parent in 6 RI lines, with the rest of the lines falling between the two (Figure 1B). There is a positive Pearson product-moment correlation between the *copia* transcript level and copy number across the RI lines ($r = 0.50$, $P < 0.001$, Figure 2A). This is expected because transcription of each *copia* copy inputs into the total transcription. Transcript level per copy (which is a total transcript level divided by the number of *copia* copies in an RI line) positively correlates with the *copia* copy number too, although the significance is marginal ($r = 0.22$, $P = 0.03$, Figure 2B).

We hypothesized that the high degree of scatter around the regression lines in Figure 2 is due to segregation of *copia* transcript level QTL. Because alleles of QTL affecting *copia* transcript level cosegregate with closely linked marker alleles, they may be mapped. However, *copia* transcript level is affected by *copia* copy number, which is strikingly different between the parental lines. Covariance between *copia* copy number and transcript level across the RI lines may mimic segregation of QTL influencing *copia* expression. We therefore analyzed the data in four ways: (i) total *copia* RNA level, (ii) *copia* RNA level per *copia* copy, (iii) the residuals of the regression of the transcript level on *copia* copy number, and (iv) the residuals of the regression of transcript level per copy on copy number. The results were very similar for all four traits (Figure 3). QTL for *copia* RNA level were located in the intervals between markers 27B and 30D and between markers 48D and 57C. An additional *copia* transcription QTL for trait (iii) was found in the interval 85F-87E. Within the intervals, the largest estimates of the QTL effects were obtained for the markers situated at 30D and 49D on the second chromosome and for the marker at 87B on the third chromosome.

The second QTL mapped to the broad gap, which we excluded from the interval mapping because no linkage disequilibrium was detected between the *roo* markers situated at 50F and 57C (see Nuzhdin *et al.* 1997a). To infer a more precise location for this QTL, we used the fixed *copia* occupation positions in the parental lines as an extra set of molecular markers. This *copia*-based mapping identified the same two major QTL. This time we were able to localize the second QTL to marker position 53E (Figure 4). The third QTL was not significant due to the lower power of this analysis, which used a threefold smaller number of markers (data not shown).

The joint effect of the segregation of alleles of these three QTL and the *copia* copies on the total transcript level was calculated by multiple regression of the transcript level on marker alleles at 30D, 49D, 87B, and *copia* copy number. 2b-originated alleles of the first two QTL increased the transcript level by 11.5% ($P < 0.0029$) and 30.4% ($P < 0.0001$), respectively. The 2b-originated allele of the third QTL decreased the transcript level by 9.0% ($P < 0.017$) relative to the Ore originated allele. Each *copia* copy increased the transcript level by 0.9%

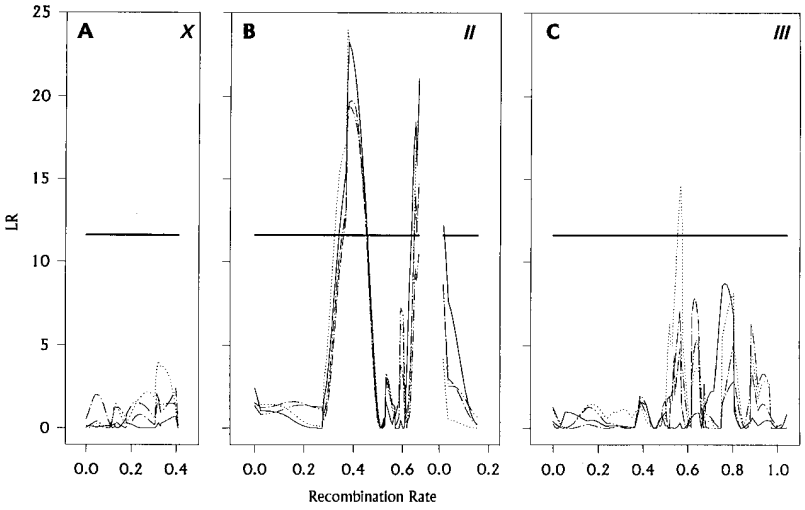

Figure 3.—Location of QTL for *copia* transcription. Plot of double-log likelihood ratio (LR) from composite interval mapping against recombination distance on the *X* (A), second (B), and third (C) chromosomes. Solid line (i), dashed line with one dot (ii), dotted line (iii), and dashed line with two dots (iv), four means of data analysis (see text). Horizontal lines, Bonferroni-corrected LR critical value for experiment-wise $\alpha = 0.05$.

($P < 0.0001$). Taking into account that there are 51 more *copia* copies in 2b compared to Ore, we accounted for 92% of the observed difference in transcript level between parental lines.

DISCUSSION

Understanding the evolution of genome size and the rate of mutations per genome from transpositions requires the description of the genetics of the transposition rate variability. Here, we have tried to understand which genetic factors are responsible for the ability of *copia* to transpose in one Drosophila line (2b) but not in another (Ore).

Genetics of the differences in *copia* transposition: Our first step in dissecting the genetics of the differences in *copia* transposition between parental lines was to assess the effect of whole chromosomes on transposition rate. *copia* was stable in all lines carrying one chromosome from the 2b line, implying that the joint presence of factors situated on at least two chromosomes of the 2b line was necessary to switch transpositions on.

These factors affecting *copia* transposition could be either host genes or *copia* copies themselves. Earlier, we showed that the rate of *copia* transposition strongly correlates with *copia* copy number (Nuzhdin *et al.* 1996; Pasyukova *et al.* 1998), and we hypothesized that the relationship was causal. The low (or zero) rate of transpositions in the Ore line and in the chromosome substitution lines could be simply explained by the small number of *copia* copies in each of them.

To analyze this problem further, we made RI lines from 2b and Ore and scored *copia* transpositions in them. If transpositions primarily accumulated in the RI lines with the highest *copia* copy number irrespective to the line genotype, the above hypothesis would be supported. In contrast, *copia* transpositions were accumulated in only three lines, and two of them had a small number of original *copia* copies (34 and 42, the total number of copies, including accumulated trans-

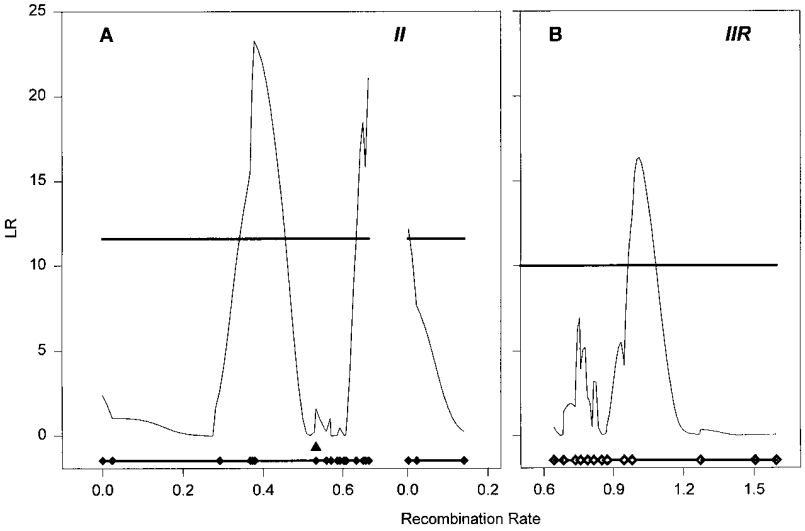

Figure 4.—Location of chromosome *2* QTL for *copia* transcription. Plot of LR statistics along the second chromosome based on *roo* markers depicted as ◆ (A) and along the right arm of this chromosome based on *copia* markers depicted as ◇ (B; see materials and methods for the origin of the differences in recombination distances between panels). Triangle, centromere position.

positions, was 37 and 60, respectively). We therefore adopted an alternative hypothesis, that a combination of 2b alleles in *copia* transposition QTL needs to be present in all RI lines in which *copia* transposes. Accumulation of *copia* transpositions was found only in the lines sharing descent of the intervals 27B-48D, 61A-65A, and 97D-100A from the 2b line. We conclude that *copia* transposition requires a combination of the QTL alleles located in these intervals.

This QTL analysis has been complicated by two factors. First, only a few RI lines were fixed for the 2b-originated alleles of the QTL of both second and third chromosomes. This is due in part to selection against the 2b-originated interval covering the first QTL (Nuzhdin *et al.* 1997). Second, the active RI lines showed low rates of *copia* transposition, which is probably explained by the low-*copia* copy number in them (Nuzhdin *et al.* 1996; Pasyukova *et al.* 1997). At present it is impossible to say whether there are multiple- or single-instability QTL within these chromosome intervals.

Genetics of the differences in *copia* transcription: The genetics of *copia* transcription is important for our study because transcription is the starting point of the transposition process. Retrotransposon RNA is both the message for the transposition machinery and the template from which new DNA copies are synthesized. It therefore seemed likely to us that transcription is a control point at which transposition is restricted.

Mutations in several host genes influence *copia* transcription, including *engrailed*, *even skipped*, *fushi tarazu*, *lola*, and *zerknault* (Cavarec and Heidmann 1993; Cavarec *et al.* 1997); *Doa*, *Inr B*, *Lip*, *Low*, *Msu*, and *Wow* (Birchler and Hiebert 1989; Heibert and Birchler 1992; Csink *et al.* 1994a,b; Bhadra *et al.* 1997); *dunce* (Yun and Davis 1989); and *DmC/EBP* (Wilson *et al.* 1998). Perhaps alleles of these genes segregate in natural populations, causing the major differences in *copia* transcript levels observed between flies (Csink and McDonald 1990).

The results presented here show that approximately one-half of the 10-fold higher *copia* transcript level in the line 2b in comparison with Ore is accounted for by two QTL situated on the second chromosome. We doubt that the third-*copia* transcript level QTL 2b-originated allele, with decreased transcript level, is real. Although this QTL is significant with one model out of four ($P < 0.017$), the significance becomes marginal when testing of multiple (4) hypotheses is taken into account. Could any of the host candidate genes be responsible for identified QTL effects? None of these genes is situated within the support interval for the first QTL but four genes are within or very close to the support interval of the second QTL (Table 2). From those, *even skipped*, *engrailed*, and *lola* are homeotic genes that have been tested for their effects on *copia* transcription, because the *copia* 5′-untranslated region contains a sequence for transcriptional regulation by homeoproteins (Cavarec *et al.* 1994, 1997). Modification of *copia* transcript level by *Inr B* and *Low* was discovered, because these mutations suppress the *white-apricot* phenotype, which is due to a *copia* insertion into the *white* locus (Bhadra *et al.* 1997). These genes may thus account in part for the higher *copia* transcript level in 2b.

The QTL that we have identified here may correspond not to host genes but rather to hyperexpressed *copia* copies. McDonald *et al.* (1997) tested the effect of naturally occurring sequence variations in *copia* long terminal repeats and 5′-untranslated regions on *copia* expression. *copia*s with more copies of enhancer elements had stronger expression. Several *copia*s fixed in the 2b line are situated within or close to the first (33F and 34B) or the second (47A, 47B, and 50F) QTL. One or a few of these copies (as well as copies segregating within the 2b line, see materials and methods for more detail) may hyperexpress. Finer scale QTL mapping, complementation analysis with the *copia* transcription candidate genes, and measuring chromosome site-specific *copia* transcription will help to resolve these hypotheses.

The relationship between factors affecting *copia* transposition and transcription: Earlier, we hypothesized that the difference in *copia* transcript level between the lines 2b and Ore could account for the difference in transposition rate (Pasyukova *et al.* 1997). Mapping QTL for *copia* transcript level and transposition rate was a way of testing this hypothesis and, if correct, narrowing down the support intervals for transposition QTL to the level of "candidate genes" influencing *copia* transcription. In accordance with our expectations, both transcript level QTL map in or close to the second-chromosome transposition factor (Table 2). These may be involved in switching *copia* transposition on and off. The relationship between *copia* transcript and transposition levels is not straightforward, however. *copia* RNA level is not particularly high in the transpositionally active RI lines, compared to the rest of the RI lines in which transpositions were not found (Table 1). Furthermore, the transposition QTL of the third chromosome do not influence *copia* transcript level. These QTL may act at the other stages of the retrotransposition cycle (*i.e.*, post-transcriptional modification, virus-like particle morphogenesis, DNA integration, and DNA repair; Yoshioka *et al.* 1990). Quantitation of the intermediates of *copia* retrotransposition in the RI lines may help to figure out all stages at which *copia* retrotransposition is restricted in the Ore line.

Evolutionary implications: The rates of transposition of TE families vary greatly across lines. For example, *gypsy* and *copia* are stable in the majority of laboratory lines, but *gypsy* is (or was) active in the stocks MS, MG, Uc, tuh^1, and tuh^3 (Pelisson *et al.* 1997), and *copia* is active in many stocks (Biemont *et al.* 1987; Di Franco *et al.* 1992; Pasyukova and Nuzhdin 1992, 1993; Nuzhdin and Mackay 1994; Charlesworth *et al.* 1994b; S.

TABLE 2

Positions of *copia* transcript/transposition level QTLs and candidate genes

Transposition QTL	Transcription QTL	Positions of *copia*[a] within or close to QTL	Transcription candidate gene position	Candidate gene title
			3C	*dunce*[c]
26F–48D	27B–30D	33F, 34EF, 35C, 38A, 40A, 41A, 42A, 42B, 47B		
			39E[b]	*Low*[b]
				even
			46C[b]	*skipped*[b]
			47A[b]	*Iola*[b]
			47A[b]	*Inr-B*[b]
			48A[b]	*engrailed*[b]
	48D–57C	52D, 53E		
			60A	*C/EBP*
61A–65A		64E		
			75C–76F	*Wow*
			83D	*Lip*
			84A	*fushi tarazu*
			88A	*zerknäult*
97D–100A		98B, 99B	98F[b]	*Msu*[b]

[a] Euchromatic sites fixed within parental lines and segregating between them.
[b] Candidate genes within or close to mapped QTL.
[c] See text for references.

Nuzhdin and D. Houle, unpublished results; J. Fry and S. Nuzhdin, unpublished results). It was shown previously for *gypsy* (Kim *et al.* 1990; Pelisson *et al.* 1994, 1997), and we have shown here for *copia*, that stable lines carry restrictive alleles, and unstable lines carry permissive alleles of genes controlling transposition of specific TE families. The large number of stocks with permissive alleles for *gypsy* and *copia* activity makes it improbable that all of them carry new mutations in genes that normally restrict retrotransposition. Indeed, Pelisson *et al.* (1997) showed that permissive alleles of *gypsy* segregate in natural populations.

The rate of transposition in nature, estimated from the distributions of element frequences in natural populations, is similar for different retrotransposon families and lies between 10^{-5} and 10^{-4} (Charlesworth *et al.* 1992a,b). Eggleston *et al.* (1988), Harada *et al.* (1990), Nuzhdin and Mackay (1995), and B. Charlesworth (personal communication) determined the rates of transposition in laboratory lines. These studies showed that the average rate of transposition in these laboratory lines across TE families is similar to the rate seen in nature (the estimates ranged from 1.2×10^{-5} to 5.0×10^{-4}). However, transposition rates vary widely between TE families within lines, for instance from zero to 1.3×10^{-3} averaged over the transposition accumulation replicates of one line (Nuzhdin and Mackay 1994). In addition, different families are active in different lines. Typically, in a given line from one-tenth to one-fifth of TE families transpose with the rate around 10^{-3} and the rest are stable (Nuzhdin *et al.* 1997b). We hypothesize that the segregation of permissive/restrictive alleles in natural populations is responsible for these effects. TE site heterogeneity in nature may be caused by frequent transpositions in the flies with permissive alleles, and no transpositions occur in the flies that carry restrictive alleles. In laboratory lines permissive alleles for some families or restrictive alleles for the other families fix, causing strong heterogeneity of transposition rates between families and lines, but changing the overall rate of transposition averaged across TE families only slightly.

This hypothesis has interesting evolutionary implications. Charlesworth and Langley (1986) showed that the strength of selection for the restrictive allele of a host gene controlling transposition is

$$s \cong -\delta u\left\{\frac{\bar{n}(u - v)}{2\tilde{H}} + \frac{\bar{n}\pi}{2(1 - 2\pi)}\right\}, \qquad (2)$$

where δu is the change in transposition rate due to a restrictive allele (note that δu cannot be larger than u); H is the harmonic mean of the rate of transposition and the rate of recombination between the gene restricting transposition and the insertion site; and π is the fraction of transpositions associated with sterility or lethality of the progeny. Because s is proportional to the squared rate of "unrestricted" transpositions (which were thought to be 10^{-5}–10^{-4}), the selection for the restrictive alleles is negligible unless there is a direct cost of the transposition process (Brookfield 1991, 1996). In accordance with our hypothesis, however, the

rate of unrestricted transpositon is 10^{-3} in permissive backgrounds and zero in restrictive backgrounds. Then s is $\sim 2 \times 10^{-4}$ (it is assumed for simplicity that $v = 0$, $n = 50$, and $\pi = 0$). Thus selection for restrictive alleles is effective.

Brookfield and Badge (1997) concluded that small populations should frequently go extinct, because ineffective selection against TE multiplication leads to TE copy number explosions that drive host fitness to zero. If restrictive / permissive alleles of the host genes controlling transposition do segregate, then the frequency of the restrictive allele should go up when TEs explode, strongly decreasing the rate of transpositions (Pasyukova *et al.* 1998). Accumulation of mutations in stabilized TE copies may contribute to the many cases of transposition-defective TEs in plant (Konieczny *et al.* 1991; Flavell *et al.* 1992), vertebrate (Kazazian *et al.* 1988; Smith 1993), and some invertebrate genomes (Berg and Howe 1989).

If there is selection for restrictive alleles, why do permissive alleles segregate in nature? First, they may be maintained due to the balance between selection for restrictive alleles ($s = 2 \times 10^{-4}$) and mutations to permissive alleles (with the typical per locus mutation rate 10^{-6}; Ashburner 1989). Provided that permissive alleles are recessive (Pelisson *et al.* 1997; E. Pasyukova and S. Nuzhdin, unpublished results), the equilibrium frequency of permissive alleles can be calculated as $\sim$7% (Falconer and Mackay 1996). Second, the restrictive alleles may have pleiotropic, deleterious side effects when homozygous, and natural selection may oppose their fixation.

Further studies on the population genetics and molecular genetics of permissive / restrictive alleles, and a theoretical consideration of their segregation in nature are required to test our hypothesis and if it is validated, to describe its implications upon the maintenance of TE copy number.

We thank Trudy F. C. Mackay and Vladimir A. Gvozdev in the labs of which a part of our experiments was done. We thank Brian Charlesworth, Charles H. Langley, and two anonymous reviewers for many ideas and critical comments. The work of S.V.N. in the lab of Trudy F. C. Mackay was supported by National Institutes of Health grants GM45344 and GM45146. The work of E.G.P., A.J.F., and S.V.N. was supported by Russian Fund for Basic Research grant 97-04-48101, Medical Research Council grant PG97-0448101, UC Davis start-up funds, and a New Faculty Research Grant Award, correspondingly.

LITERATURE CITED

Ajioka, J. W., and W. F. Eanes, 1989 The accumulation of *P*-elements on the tip of the *X* chromosome in populations of *Drosophila melanogaster*. Genet. Res. **53:** 1–12.

Ashburner, M., 1989 *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Berg, D. E., and M. M. Howe, 1989 *Mobile DNA*. American Society for Microbiology, Washington, DC.

Bhadra, U., M. Pal-Bhadra and J. A. Birchler, 1997 A sex-influenced modifier in Drosophila that affects a broad spectrum of target loci including the histone repeats. Genetics **146:** 903–917.

Biemont, 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.

Biemont, C., A. Aouar and C. Arnault, 1987 Genome reshuffling of the *copia* element in an inbred line of *Drosophila melanogaster*. Nature **329:** 742–744.

Biemont, C., F. Lemeunier, M. P. Garcia Guerreiro, J. F. Brookfield, S. Gautiers *et al.*, 1994 Population dynamics of the *copia*, *mdg1*, *mdg3*, *gypsy*, and *P* transposable elements in a natural population of *D. melanogaster*. Genet. Res. **63:** 197–212.

Biemont, C., A. Tsitrone, C. Vieira and C. Hoogland, 1997 Transposable element distribution in Drosophila. Genetics **147:** 1997–1999.

Birchler, J. A., and J. C. Hiebert, 1989 Interaction of the *Enhancer of white-apricot* with transposable element alleles at the *white* locus in *Drosophila melanogaster*. Genetics **122:** 129–138.

Brookfield, J. F. Y., 1991 Models of repression of transposition in *P-M* hybrid dysgenesis by *P* cytotype and by zygotically encoded repressor proteins. Genetics **128:** 471–486.

Brookfield, J. F. Y., 1996 Models of the spread of non-autonomous selfish transposable elements when transposition and fitness are coupled. Genet. Res. **67:** 199–210.

Brookfield, J. F., and R. M. Badge, 1997 Population genetics models of transposable elements. Genetica **100:** 281–294.

Bucheton, A., J.-M. Lavige, G. Picard and P. L'Heritier, 1976 Non-Mendelian female sterility in *Drosophila melanogaster* : quantitative variation in the efficiency of inducer and reactive strains. Heredity **36:** 305–314.

Bucheton, A., C. Vaury, M.-C. Chaboissier, P. Abad, A. Pelisson *et al.*, 1992 *I* elements and the Drosophila genome. Genetica **86:** 175–190.

Cavarec, L., and T. Heidmann, 1993 The Drosophila *copia* retrotransposon contains binding sites for transcriptional regulation by homeoproteins. Nucleic Acids Res. **21:** 5041–5049.

Cavarec, L., S. Jensen and T. Heidmann, 1994 Identification of a strong transcriptional activator for the *copia* retrotransposon responsible for its differential expression in *Drosophila hydei* and *melanogaster* cell lines. Biochem. Biophys. Res. Commun. **203:** 392–399.

Cavarec, L., S. Jensen, J.-F. Casella, S. A. Cristescu and T. Heidmann, 1997 Molecular cloning and characterization of a transcription factor for the *copia* retrotransposon with homology to the BTB-containing *lola* neurogenic factor. Mol. Cell. Biol. **17:** 482–494.

Charlesworth, B., 1991 Transposable elements in natural populations with a mixture of selected and neutral insertion sites. Genet. Res. **57:** 127–135.

Charlesworth, B., and D. Charlesworth, 1983 The population dynamics of transposable elements. Genet. Res. **42:** 1–27.

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

Charlesworth, B., and C. H. Langley, 1989 The population genetics of Drosophila transposable elements. Annu. Rev. Genet. **23:** 251–287.

Charlesworth, B., and A. Lapid, 1989 A study of ten transposable elements on *X* chromosomes from a population of *Drosophila melanogaster*. Genet. Res. **54:** 113–125.

Charlesworth, B., A. Lapid and D. Canada, 1992a 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.

Charlesworth, B., A. Lapid and D. Canada, 1992b The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. II. Inferences on the nature of selection against elements. Genet. Res. **60:** 115–130.

Charlesworth, B., P. Sniegowski and W. Stephan, 1994a The evolutionary dynamics of repetitive DNA in eukaryotes. Nature **371:** 215–220.

Charlesworth, B. P. Jarne and S. Assimacopoulos, 1994b The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. III. Element abundances in heterochromatin. Genet. Res. **64:** 183–197.

Charlesworth, B., C. H. Langley and P. D. Sniegowski, 1997 Transposable element distributions in Drosophila. Genetics **147:** 1993-1995.

Clark, J. B., W. P. Maddison and M. G. Kidwell, 1994 Phylogenetic analysis supports horizontal transfer of *P* transposable elements. Mol. Biol. Evol. **11:** 40–49.

Csink, A. K., and J. F. McDonald, 1990 *copia* expression is variable among natural populations. Genetics **126:** 375–382.

Csink, A. K., and J. F. McDonald, 1995 Analysis of *copia* sequence variation within and between Drosophila species. Mol. Biol. Evol. **12:** 83–93.

Csink, A. K., R. Linsk and J. A. Birchler, 1994a The *Lighten up* (*Lip*) gene of *Drosophila melanogaster*, a modifier of retroelement expression, position effect variegetion and *white* locus insertion alleles. Genetics **138:** 153–163.

Csink, A. K., R. Linsk and J. A. Birchler, 1994b *Mosaic suppressor*, a gene in Drosophila that modifies retrotransposon expression and interacts with *zeste*. Genetics **136:** 573–583.

Di Franco, C., D. Galuppi and N. Junakovic, 1992 Genomic distribution of transposable elements among individuals of an inbred Drosophila line. Genetica **86:** 1–12.

Dombradi, V., J. M. Axton, D. M. Glover and T. W. Cohen, 1989 Cloning and chromosomal localization of Drosophila cDNA encoding the catalytic subunit of protein phosphatase 1a. Eur. J. Biochem. **183:** 603–610.

Eanes, W. F., C. Wesley and B. Charlesworth, 1992 Accumulation of *P* elements in minority inversions in natural populations of *Drosophila melanogaster*. Genet. Res. **59:** 1–14.

Eggleston, W. B., D. M. Johnson-Schlitz and W. R. Engels, 1988 *P-M* hybrid dysgenesis does not mobilize other transposable element families in *Drosophila melanogaster*. Nature **331:** 368–370.

Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*. Addison-Wesley, Reading, MA.

Finnegan, D. J., 1992 Transposable elements, pp. 1096–1107 in *The Genome of Drosophila melanogaster*, edited by D. L. Lindsley and G. Zimm. Academic Press, New York.

Finnegan, D. J., G. M. Rubin, M. W. Young and D. S. Hogness, 1978 Repeated gene families in *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. **42:** 1053–1063.

Flavell, A. J., D. B. Smith and A. Kumar, 1992 Extreme heterogeneity of *Ty-copia* group retrotransposons in plants. Mol. Gen. Genet. **231:** 233–242.

Harada, K., K. Yukuhiro and T. Mukai, 1990 Transposition rates of movable genetic elements in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **87:** 3248–3252.

Hiebert, J. C., and J. A. Birchler, 1992 Dosage compensation of the *copia* retrotransposon in *Drosophila melanogaster*. Genetics **130:** 539–545.

Hoogland, C., and C. Biemont, 1996 Chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for the maintenance of insertion site number. Genetics **144:** 197–204.

Kaplan, N. L., and J. F. Y. Brookfield, 1983 Transposable element in mendelian populations. III. Statistical results. Genetics **104:** 485–495.

Kazazian, H. H.., C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips *et al.*, 1988 Haemophilia A resulting from *de novo* insertion of *L1* sequences represents a novel mechanism for mutation in man. Nature **332:** 164–166.

Kidwell, M. G., J. F. Kidwell and J. A. Sved, 1977 Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics **86:** 813–833.

Kim, A., E. S. Belyaeva and M. M. Aslanian, 1990 Autonomous transposition of *gypsy* mobile elements and genetic instability in *Drosophila melanogaster*. Mol. Gen. Genet. **224:** 303–308.

Kim, A. I., N. V. Lyubomirskaya, E. S. Belyaeva, N. G. Shostack and Yu. V. Ilyuin, 1994 The introduction of a transpositionally active copy of retrotransposon *gypsy* into a stable strain of *Drosophila melanogaster* causes genetic instability. Mol. Gen. Genet. **242:** 472–477.

Konieczny, A., D. F. Voytas, M. P. Cummings and F. M. Ausubel, 1991 A superfamily of *Arabidopsis thaliana* retrotransposons. Genetics **127:** 801–809.

Langley, C. H., E. A. Montgomery, R. Hudson, N. Kaplan and B. Charlesworth, 1988 On the role of unequal exchange in the containment of transposable element copy number Genet. Res. **52:** 223–236.

Lefevre, G., 1976 A photographic representation of the polygene chromosomes of *Drosophila melanogaster* salivary glands, pp. 31–66 in *The Genetics and Biology of Drosophila*, Vol. 1A, edited by M. Ashburner and E. Novitski. Academic Press, London.

Leigh-Brown, A. J., and J. E. Moss, 1987 Transposition of the *I* element and *copia* in a natural population of *Drosophila melanogaster*. Genet. Res. **49:** 121–128.

Lohe, A. R., and D. L. Hartl, 1996 Reduced germline mobility of a *mariner* vector containing exogenous DNA: effect of size or site? Genetics **143:** 1299–1306.

Lohe, A. R., E. N. Moriyama and D. L. Hartl, 1995 Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. Mol. Biol. Evol. **12:** 62–72.

Mackay, T. F. C., R. F. Lyman and M. S. Jackson, 1992 Effects of *P* element insertions on quantitative traits in *Drosophila melanogaster*. Genetics **130:** 315–332.

McDonald, J. F., L. V. Matyunina, S. Wilson, I. K. Jordan, N. J. Bowen *et al.*, 1997 LTR retrotransposons and the evolution of eukariotic enhancers. Genetica **100:** 111–115.

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

Nuzhdin, S. V., 1995 The distribution of transposable elements on *X* chromosomes from a natural population of *Drosophila simulans*. Genet. Res. **66:** 159–166.

Nuzhdin, S. V., and T. F. C. Mackay, 1994 Direct determination of retrotransposon transposition rates in *Drosophila melanogaster*. Genet. Res. **63:** 139–144.

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

Nuzhdin, S. V., E. G. Pasyukova and T. F. C. Mackay, 1996 Positive association between *copia* transposition rate and copy number in *Drosophila melanogaster* Proc. R. Soc. Lond. Ser B **263:** 823–831.

Nuzhdin, S. V., E. G. Pasyukova, C. A. Dilda, Z.-B. Zeng and T. F. C. Mackay, 1997a Sex-specific quantitative trait loci affecting longevity in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **94:** 9734–9739.

Nuzhdin, S. V., E. G. Pasyukova and T. F. C. Mackay, 1997b Accumulation of transposable elements in laboratory lines of *Drosophila melanogaster*. Genetica **100:** 167–175.

Pasyukova, E. G., and S. V. Nuzhdin, 1992 Mobilization of retrotransposon *copia* in *Drosophila melanogaster* genome. Genetica (Russ.) **28:** 5–18.

Pasyukova, E. G., and S. V. Nuzhdin, 1993 *Doc* and *copia* instability in an isogenic *Drosophila melanogaster* stock. Mol. Gen. Genet. **240:** 302–306.

Pasyukova, E. G., S. V. Nuzhdin, W. Li and A. J. Flavell, 1997 Germ line transposition of the *copia* mobile element in *Drosophila melanogaster* is restricted to males by sex-specific control of *copia* RNA level. Mol. Gen. Genet. **255:** 115–124.

Pasyukova, E. G., S. V. Nuzhdin and D. A. Filatov, 1998 The relationship between the rate of transposition and transposable element copy number for *copia*, *Doc* and *roo* retrotransposons. Genet. Res. (in press).

Pelisson, A., and J. C. Bregliano, 1987 Evidence for rapid limitation of the *I* element copy number in a genome submitted to several generations of *I-R* hybrid dysgenesis in *Drosophila melanogaster*. Mol. Gen. Genet. **207:** 306–313.

Pelisson, A., S. U. Song, N. Prud'homme, P. Smith, A. Bucheton *et al.*, 1994 *Gypsy* transposition correlates with the production of a retroviralenvelope-like protein under the tissue-specific control of Drosophila *flamenco* gene. EMBO J. **13:** 4401–4411.

Petrov, D. A., J. L. Schutzman, D. L. Hartl and E. R. Lozovskaya, 1995 Diverse transposable elements are mobilized in hybrid dysgenesis in *Drosophila virilis*. Proc. Natl. Acad. Sci. USA **92:** 8050–8054.

Pelisson, A., L. Teysset, F. Chalvet, A. Kim, N. Prud'homme *et al.*, 1997 About the origin of retroviruses and the co-evolution of the *gypsy* retrovirus with the Drosophila *flamenco* host gene. Genetica **100:** 29–37.

Petrov, D. A., E. R. Lozovskaya and D. L. Hartl, 1996 High intrinsic rate of DNA loss in Drosophila. Nature **384:** 346–349.

SAS Institute, Inc., 1988 SAS/STAT User's Guide. Release 6.03 Edition (SAS Institute, Inc., Cary, NC).

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Savakis, C., M. Ashburner and J. H. Willis, 1986 The expression of the gene coding for alcohol dehydrogenase during the development of *Drosophila melanogaster.* Dev. Biol. **114:** 194–207.

Shrimpton, A. E., E. A. Montgomery and C. H. Langley, 1986 *Om* mutations in *Drosophila ananassae* are linked to insertions of a transposable element. Genetics **114:** 125–135.

Silver, J., 1985 Confidence limits for estimates of gene linkage based on analysis of recombinant inbred strains. Heredity **76:** 436–440.

Smith, A. F. A., 1993 Identification of a new, abundant superfamily of mammalian LTR-transposons. Nucleic Acids Res. **8:** 1863–1872.

Sniegowski, P. D., and B. Charlesworth, 1994 Transposable element in numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster.* Genetics **137:** 815–826.

Wilson, S., L. V. Matyunina and J. F. McDonald, 1998 An enhancer region within the *copia* untranslated leader contains binding sites for Drosophila regulatory proteins. Gene **209:** 239–246.

Yoshioka, K., H. Honma, M. Zushi, S. Kondo, S. Togashi *et al.*, 1990 Virus-like particle formation of Drosophila *copia* through autocatalytic processing. EMBO J. **9:** 535–541.

Yun, Y., and R. L. Davis, 1989 *copia* RNA levels are elevated in *dunce* mutants and modulated by cAMP. Nucleic Acids Res. **17:** 8313–8326.

Communicating editor: W. Stephan