Genet. Res., Camb. (1998), **72**, *pp.* 1–11. *With* 3 *figures.* Printed in the United Kingdom © 1998 Cambridge University Press

The relationship between the rate of transposition and transposable element copy number for *copia* and *Doc* retrotransposons of *Drosophila melanogaster*

ELENA G. PASYUKOVA[1,2,*], SERGEY V. NUZHDIN[2,3] AND DMITRY A. FILATOV[1]
[1] *Institute of Molecular Genetics of RAS, Moscow 123182, Russia*
[2] *Department of Genetics, North Carolina State University, Raleigh, NC 27695–7614, USA*
[3] *Center of Population Biology, University of California at Davis, Davis, CA 95616, USA*

(*Received 5 January 1998 and in revised form 9 April 1998*)

Summary

We present data on the relationship between the rate of transposition and copy number in the genome for the *copia* and *Doc* retrotransposons of *Drosophila melanogaster*. *copia* and *Doc* transposition rates were directly measured in sublines of the isogenic 2b line using individual males or females, respectively, with a range of *copia* copy numbers from 49 to 103 and *Doc* copy numbers from 112 to 235 per genome. Transposition rates varied from 3×10^{-4} to 2×10^{-2} for *copia* and from 2×10^{-4} to 2×10^{-3} for *Doc*. A positive relationship between transposition rate and copy number was found both for *copia* and for *Doc* when the data were analysed across all the 2b individuals; no significant correlation was found when the data were analysed across the subline means for both retrotransposons tested. Overall, correlation between *copia* and *Doc* transposition rate and their copy number in the genome, if any, was not negative, which would be expected if transposable elements (TEs) self-regulate their copy number. Thus, for *copia* and *Doc* no evidence for self-regulation was provided, and at least for these two TEs this hypothesis is not favoured for explaining the maintenance of the stable copy number that is characteristic for natural populations. The transposition rate of *copia* was measured twice, and a strong positive correlation between copy number and transposition rate both across individuals and subline means was found in 1994, while in 1995 no correlation was found. This fact is in agreement with the hypothesis that a positive correlation between the rate of transposition and TE copy number may be a default starting point for future host–TE coevolution.

1. Introduction

The fact that transposable elements (TEs) are regular genomic components raises the issue of the nature of the forces that determine their persistence in the genome. Experimental and theoretical studies of TEs in *Drosophila* lead to the hypothesis that TEs are maintained as a result of the balance between the process of transposition and natural selection against insertional mutations, deleterious chromosomal rearrangements appearing due to ectopic exchange (Langley *et al.*, 1988; Charlesworth & Langley, 1989; Sniegowski & Charlesworth, 1994), and damage from the transposition process itself (Brookfield, 1991; Nuzhdin *et al.*, 1996). The self-regulation of TEs, for which a lower rate of transposition and a higher rate of excisions occur in individuals with higher TE copy numbers, may also account for the stabilization of TE copy number (Charlesworth & Charlesworth, 1983; Charlesworth & Langley, 1989; Biemont, 1994). It is also evident that the combination of forces operating on TE copy number maintenance depends on the element considered (Biemont *et al.*, 1994), as well as on the region of the genome (Langley *et al.*, 1988; Charlesworth & Langley, 1989).

The systematic study of both the transposition process and selection forces acting on new insertions must be completed for a representative sample of TE families to describe TE population dynamics at a predictive level. Several key parameters should be inferred for a TE family under study (Nuzhdin *et al.*, 1997), one of them being the probability that a given

* Corresponding author. Institute of Molecular Genetics of RAS, Kurchatov sq., Moscow 123182, Russia. Telephone: (7-095)196-8173. Fax: (7-095)196-0221. e-mail: egpas@img.ras.ru.

TE copy transposes to a new site, conditional on the number of copies per genome. The direct measurement of transposition rates in individuals with a range of TE copy numbers is needed to estimate this parameter. To date, only two experiments have been reported in which the relationship between transposition rate and copy number has been measured directly for retrotransposons: for *copia* in *D. melanogaster* laboratory lines (Nuzhdin *et al.*, 1996), and for the retrotransposon *412* in natural populations of *D. simulans* (Vieira & Biemont, 1997). In neither case was a negative association between transposition rate and TE copy number found, as would be expected if TEs self-regulate copy number, and a positive relationship was inferred for the *copia* retrotransposon.

Here, we present data on the relationship between *copia* transposition rate and copy number in another *D. melanogaster* line of independent origin. We also extend the list of TE families for which the relationship between the rate of transposition and TE copy number is known to the *Doc* retrotransposon of *D. melanogaster*. Flies of natural populations may segregate for host genes modifying transposition rate (McDonald *et al.*, 1988; Prud'homme *et al.*, 1995). Using flies with variable TE copy numbers sampled in nature could confound the detection of a consistent relationship between TE copy number and transposition rate (Vieira & Biemont, 1997). To avoid this problem, we used replicate sublines of an isogenic laboratory line. Natural selection is partly relaxed in laboratory lines of flies, and this causes accumulation of TEs (Nuzhdin & Mackay, 1994). Therefore, replicates of a homozygous line initially fixed for a set of TE insertions and host modifiers of transposition rate evolve different numbers of copies. Using replicate sublines of the isogenic 2b line with different numbers of accumulated copies of *copia* and *Doc* (Pasyukova & Nuzhdin, 1993), we demonstrate in direct experiments that: (i) the relationship between *copia* and *Doc* transposition rate and their copy number in the genome, if any, is not negative, and at least for these two retrotransposons the hypothesis of self-regulation is not favoured as explaining the maintenance of the stable TE copy number that is characteristic for natural populations; (ii) the relationship between *copia* transposition rate and copy number is not constant in the same line; and (iii) the strong positive relationship between transposition rate and copy number initially observed for *copia* supports the hypothesis that a positive correlation between the rate of transposition and TE copy number may be a default starting point for future host–TE coevolution. In addition, we demonstrate that 599 *de novo copia* transpositions revealed throughout the experiments are randomly distributed along chromosomes; however, the X chromosome as a whole is a preferential target of transpositions.

2. Materials and methods

(i) In situ *hybridization*

Insertion sites were determined by *in situ* hybridization of labelled *copia* DNA to polytene salivary gland chromosomes of third instar larvae according to the procedure described in Ashburner (1989). Plasmids containing full-size copies of *copia* and *Doc* retrotransposons (Finnegan, 1992) were labelled with biotin (bio-7-dATP, BRL) by nick-translation. Hybridization was detected using the Elite Vectastain ABC kit (Vector Laboratories, Burlingame, CA) and diaminobenzidine (Sigma). Chromosomal locations of retrotransposons were determined at the level of cytological subsections on the standard Bridges map of *D. melanogaster* (Lefevre, 1976).

(ii) Drosophila *lines*

2b is an isogenic laboratory line of wild phenotype (Pasyukova & Nuzhdin, 1993; Pasyukova *et al.*, 1997). It was synthesized in 1986 and divided into three independently maintained sublines (2b1, 2b2, 2b3) in 1988. 2b1, 2b2 and 2b3 were further divided into a number of sublines (Fig. 1) that were then maintained independently as small mass cultures (about 20 pairs of flies per generation). Two sublines (2b3.iso1 and 2b3.iso2) were isogenic derivatives of 2b3 obtained by standard genetic methods using balancer chromosomes (Pasyukova & Nuzhdin, 1993). Twenty-four euchromatic *copia* sites and 27 euchromatic *Doc* sites were fixed in the original 2b line and were still fixed in all the sublines over 10 years of maintenance. The presence of all the fixed sites was checked in all the individuals analysed by *in situ* hybridization throughout the experiments and was used as a control against contamination. However, in addition to the fixed sites, from five to 30 new polymorphic *copia* sites and from 30 to 85 polymorphic *Doc* sites per haploid genome were found in some sublines. These polymorphic sites presumably appeared due to transposition and were accumulated during the maintenance of sublines. All the sublines of 2b lack P-elements.

Oregon RC-iso is an isogenic derivative of the Oregon RC line described elsewhere (Nuzhdin *et al.*, 1996; Pasyukova *et al.*, 1997). It has 14 and 37 fixed euchromatic insertion sites of *copia* and *Doc*, respectively. The *copia* and *Doc* transposition rates in Oregon RC-iso are less than 10^{-4} per element per generation. It was used as a tester line to measure directly *copia* and *Doc* transposition rates.

(iii) *Direct estimation of germ line transposition rates*

The rate of *copia* transposition was measured in males and the rate of *Doc* transposition was measured in

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

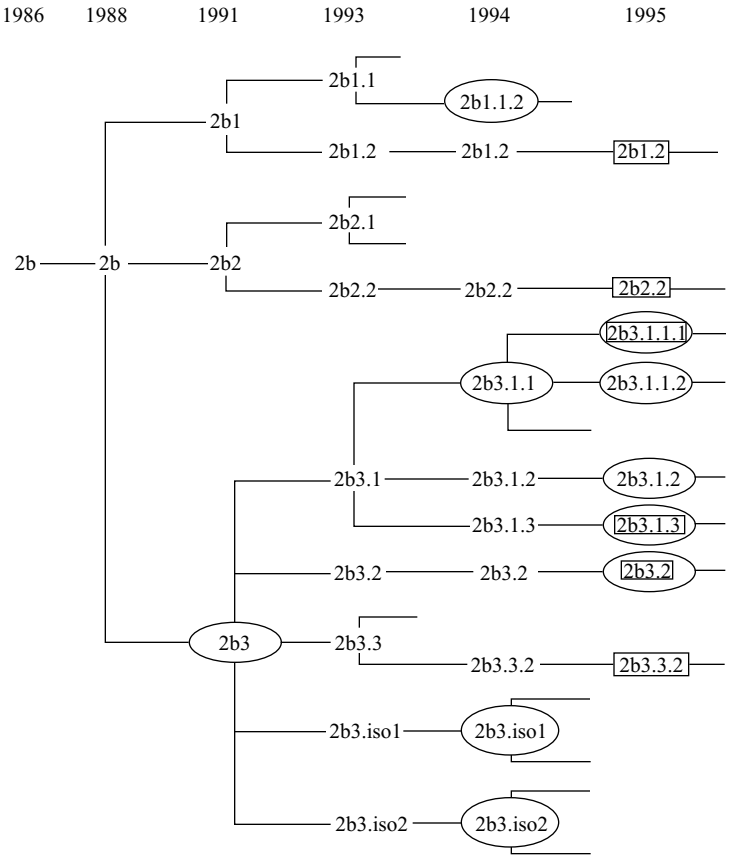

Fig. 1. Sublines of the 2b line. Sublines where *copia* and *Doc* transposition rates were measured are marked by circles and squares, respectively.

females since *copia* transpositions are restricted to males (Pasyukova *et al*., 1997) and *Doc* transpositions are restricted to females (unpublished results).

The experimental design for 2b males was as follows. Individual 2b males were crossed with a tester shock, Oregon RC-iso females, and *copia* insertion sites for between six and 40 progeny larvae of each male were determined by *in situ* hybridization. Each of the F_1 progeny inherited one chromosome set from the maternal Oregon RC-iso line. The Oregon RC-iso insertion pattern was subtracted from the analysis. Additionally, the progeny contained a chromosome set from the 2b male parent. Each 2b parent has its own particular combination of polymorphic sites in addition to fixed original sites shared with its siblings. The paternal *copia* pattern was determined for each parental male by analysis of the segregating *copia* sites in the respective progeny. *copia* transpositions and excisions in the germ lines of tested males were detected by the appearance of non-parental *copia* insertion sites, or the loss of parental insertion sites, respectively, in the progeny larvae. The transposition rate for a given male was calculated as (number of transpositions)/(number of gametes analysed) × (diploid *copia* copy number of a given male) (Pasyukova & Nuzhdin, 1992).

The rate of *Doc* transposition in 2b females was measured in the same way. Individual females were crossed with a tester stock, Oregon RC-iso males, and from 17 to 23 larvae in the progeny of each female were scored for *Doc* insertion sites. Reconstruction of the *Doc* insertion pattern of a parental female was complicated by recombination. However, transpositions were deduced by the appearance of new *Doc* sites characteristic of only one larvae in the progeny of a tested female (for more details see Pasyukova *et al*., 1997).

(iv) *Estimation of copy numbers*

Characterizing all the sites of a particular retrotransposon in parental males (females) was a necessary step in transposition rate measurement. The diploid copy number of a male was determined as the sum of sites of a particular retrotransposon in its single X chromosome and in two haploid complements of

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

autosomes. The diploid copy number of a female was determined as the sum of sites in its two haploid complements of chromosomes.

(v) *Data analysis*

Correlation and regression analyses were performed using CORR and REG procedures in SAS (SAS Institute, 1988). Kendall and Spearman coefficients of rank correlation were calculated for use in non-parametric tests for association between transposition rate and copy number.

3. Results

(i) *Copy number and transposition rate of* copia

In sublines of the 2b line, the transposition rate of *copia* was measured directly several times: in 1991, 1994 and 1995. Overall, 599 *copia* transpositions were revealed among 1466 larvae in all the experiments with all the sublines of the 2b line. Thus, 0·41 transpositions per gamete per generation were observed on average.

The distribution of 599 *copia* transpositions along 769 X chromosomes, 1466 second, 1466 third and 1466 fourth chromosomes is shown in Fig. 2. Most transpositions were found in only one progeny of a given parental male; however, in 11 cases two transpositions into the same (undistinguishable) cytological site were detected in the progeny of a given parental male (clusters of transpositions; see Fig. 2). Analysis of clusters based on a larger set of data can be found in Filatov *et al.* (1998). The probability of *copia* transposition into a band was low. Overall, the distribution of transpositions on autosomes was not random, the mean of the number of autosomal inserts per cytological subdivision being slightly lower than variance (1·49 and 1·55, respectively; $\chi^2_5 = 113{\cdot}6$, $P < 0{\cdot}001$ for a goodness-of-fit of the autosomal inserts to a Poisson distribution). The non-random distribution of sites of transposition can be partially explained by variation in DNA content. The correlation between the number of *copia* transpositions and estimated DNA content (Heino *et al.*, 1994) per cytological subdivision was highly significant ($r = 0{\cdot}21$, $P < 0{\cdot}001$). The analysis of distribution of *copia* transpositions among the chromosomes demonstrated that the number of transpositions per arm was higher for the X chromosome than for the autosomes ($t = 4{\cdot}4$, $P < 0{\cdot}001$, $t = 3{\cdot}4$, $P < 0{\cdot}001$ for comparison with the second and third chromosome, respectively; Table 1). There is no significant variation in DNA content between the chromosomal arms (Table 1). The number of transpositions per 1 Mb of DNA was significantly higher for the X chromosome than for the autosomes ($t = 3{\cdot}7$, $P < 0{\cdot}001$ for comparison with both second and third chromosomes; Table 1).

In 1991, *copia* transposition rate in the 2b3 subline was equal to 3×10^{-3} (Pasyukova *et al.*, 1997; Table 2). In 1994, *copia* transposition rates were measured in 24 males of four sublines of the 2b line and in 10 F_1 males from a cross between the sublines with the lowest and the highest *copia* copy number (Fig. 1, Table 2). Analysing 643 larvae revealed 349 transpositions. The copy number of *copia* in individual males ranged from 49 to 82 per diploid genome. The positive correlation between *copia* copy number and *copia* transposition rate was highly significant across all the individual males tested in 1994 ($r = 0{\cdot}57$, $P < 0{\cdot}001$ and $r = 0{\cdot}87$, $P < 0{\cdot}001$ for Kendall and Spearman rank-order correlation, respectively). The highest line mean transposition rate (2×10^{-2}) was detected in the 2b3.1.1 subline with the highest *copia*

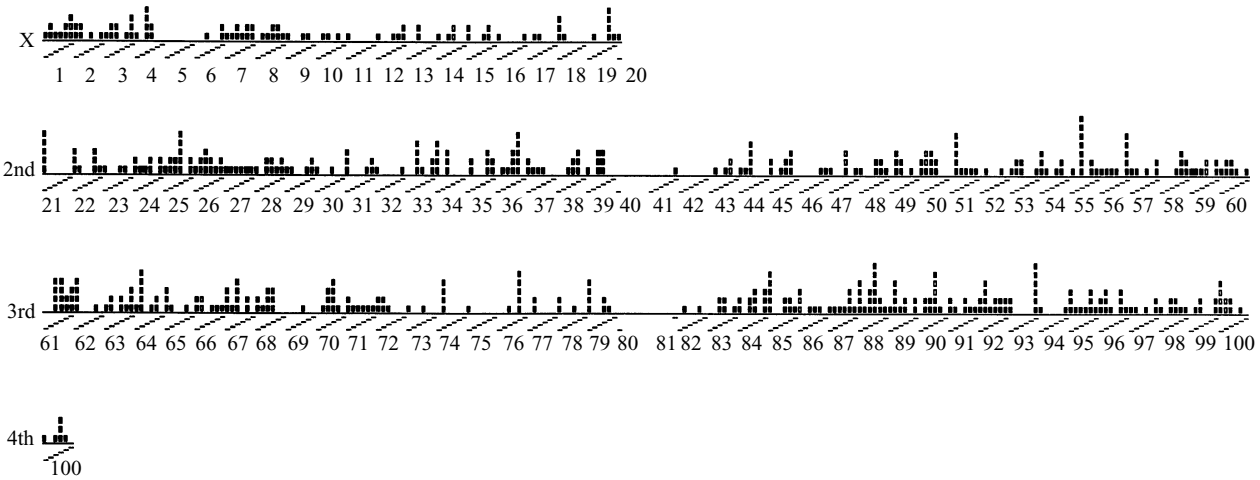

Fig. 2. The chromosomal distribution of *de novo copia* transpositions in sublines of the 2b line. The numbers and ticks show chromosomal sections and subsections, respectively, on the standard cytological map. Only sections and subsections that were scored are included. Transpositions found in only one progeny of a tested male (filled squares) and in several progeny of a tested male (open squares) are shown.

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

Table 1. *Distribution of* copia *transpositions on the chromosomal arms*

Chromosome	No. of *copia* transpositions revealed	No. of chromosomes analysed	Approximate DNA content[a] (Mb)	No. of *copia* transpositions per chromosomal arm	No. of *copia* transpositions per 1 Mb of DNA
X	89	768	22·4	$0{\cdot}116 \pm 0{\cdot}02$	5×10^{-3}
Second	235	1466	40·6	$0{\cdot}080 \pm 0{\cdot}01$***	4×10^{-3}***
Third	268	1466	44·3	$0{\cdot}091 \pm 0{\cdot}01$***	4×10^{-3}***

*** $P < 0{\cdot}001$ (X chromosome vs autosomal arm).
[a] DNA content per chromosome was calculated according to Heino *et al.* (1994) as a sum of DNA contents of all the subsections scored (see Fig. 2).

copy number (79), and the lowest rate (3×10^{-4}) in the 2b1.1.2 subline with the lowest *copia* copy number (54). Hybrid F1 males with intermediate copy numbers had intermediate transposition rates. The positive correlation between mean subline *copia* transposition rate and *copia* copy number was significant both when the data on F_1 hybrids were included ($r = 0{\cdot}95$, $P < 0{\cdot}05$ and $r = 0{\cdot}97$, $P < 0{\cdot}01$ for Kendall and Spearman rank-order correlation respectively) or not ($r = 0{\cdot}98$, $P < 0{\cdot}01$ and $r = 0{\cdot}99$, $P < 0{\cdot}001$ for Kendall and Spearman rank order correlation, respectively; Fig. 3). When transposition rate was regressed on copy number and copy number squared simultaneously, both terms were significantly different from zero, whether hybrids were included ($P < 0{\cdot}05$, $P < 0{\cdot}03$, respectively) or not ($P < 0{\cdot}04$, $P < 0{\cdot}03$, respectively).

In 1995, *copia* transposition rates were measured again in 33 males of five sublines, and in eight F_1 males from a cross between the sublines with the lowest and the highest *copia* copy number (Fig. 1, Table 2). Analysing 726 larvae revealed 235 transpositions. Copy number of *copia* in individual males ranged from 56 to 103 per diploid genome. No association between *copia* transposition rate and copy number either across all the tested individual males or across the line means was significant in 1995.

For the pooled data, the correlation between *copia* copy number and transposition rate was significant across all the tested individual males ($r = 0{\cdot}26$, $P < 0{\cdot}01$ and $r = 0{\cdot}39$, $P < 0{\cdot}001$ for Kendall and Spearman rank-order correlation, respectively) and was not significant across the line means. No association between transposition rate and copy number was observed within any subline in either 1994 or 1995.

No excisions were found throughout the experiments.

(ii) *Copy number and transposition rate of* Doc

The rates of *Doc* transposition were directly measured in 16 individual females from six different sublines of the 2b line in 1995 (Fig. 1, Table 3). In total, 63 *Doc* transpositions were detected by analysing 387 larvae. Thus, 0·16 transpositions per gamete per generation were revealed on average. The probability of *Doc* transposition into a band was low. Overall, the distribution of transpositions was not random, the mean of the number of inserts per cytological subdivision being slightly lower than the variance (0·107 and 0·116, respectively; $\chi^2_3 = 29{\cdot}9$, $P < 0{\cdot}01$ for a goodness-of-fit of the inserts to a Poisson distribution). The number of *Doc* transpositions per chromosomal arm did not deviate significantly from random.

Copy numbers of *Doc* in individual females ranged from 112 to 235 per diploid genome. The positive correlation between copy number and the rate of transposition across all the tested females was significant ($r = 0{\cdot}57$, $P < 0{\cdot}01$ and $r = 0{\cdot}69$, $P < 0{\cdot}01$ for Kendall and Spearman rank-order correlation, respectively). The correlation between the subline mean transposition rates and the subline mean copy numbers was not significant.

Excisions of *Doc* could not be detected due to the experimental design.

4. Discussion

(i) *The absence of negative correlation between copy number and transposition rate for the two retrotransposons tested*

Formally, a positive relationship between the transposition rate and copy number was found both for *copia* and for *Doc* retrotransposons in the 2b line, when the data were analysed across all the individuals, of all the sublines. However, taking into account the considerable influence noise can have in individual measurements, we also used subline means to obtain more precise estimates of the correlation. Even when analysed across the subline means, the association between *copia* and *Doc* transposition rates and their copy numbers in the genome, if any, was not significantly negative. The same result was reported for the *412* retrotransposon (Vieira & Biemont, 1997).

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

Table 2. copia *transposition rate in 2b males*

Subline	Year	Individual	Copy number	Progeny number	No. of transpositions	Transposition rate
2b3[a]	1991	1	52	32	4	0·0024
		2	52	15	4	0·0051
		3	54	9	1	0·0021
		4	54	19	4	0·0039
		5	57	22	2	0·0016
		Total/mean[b]	54	97	15	0·0030 (0·0006)[c]
2b1.1.2	1994	1	51	12	1	0·0016
		2	54	6	0	0
		3	54	6	0	0
		4	54	19	0	0
		5	55	19	0	0
		Total/mean	54	62	1	0·0003 (0·0003)
2b3.1.1	1994	1	74	10	9	0·0122
		2	76	9	25	0·0365
		3	78	10	7	0·0090
		4	79	14	31	0·0280
		5	79	40	50	0·0158
		6	80	12	15	0·0156
		7	82	38	62	0·0199
		8	82	24	31	0·0158
		Total/mean	79	157	230	0·0191 (0·0032)
2b3.iso1	1994	1	54	20	2	0·0018
		2	57	27	7	0·0044
		3	59	19	3	0·0026
		4	60	16	1	0·0010
		5	61	14	0	0
		6	61	23	2	0·0014
		Total/mean	59			0·0019 (0·0006)
2b3.iso2	1994	1	49	14	0	0
		2	53	11	1	0·0017
		3	54	13	0	0
		4	55	21	1	0·0008
		5	57	22	0	0
		Total/mean	54	81	2	0·0005 (0·0003)
2b3.1.1.1[a]	1995	1	81	16	8	0·0062
		2	81	14	7	0·0062
		3	81	9	3	0·0041
		4	82	33	3	0·0011
		5	83	9	6	0·0080
		6	83	15	16	0·0129
		7	83	11	6	0·0066
		8	85	10	17	0·0200
		9	85	9	3	0·0039
		10	86	9	2	0·0026
		11	88	33	17	0·0059
		12	88	54	18	0·0039
		13	88	15	14	0·0106
		14	93	15	7	0·0050
		Total/mean	85	252	127	0·0069 (0·0013)
2b3.1.1.2	1995	1	88	10	2	0·0023
		2	92	12	2	0·0018
		3	95	10	0	0
		4	100	16	1	0·0006
		Total/mean	94	48	5	0·0012 (0·0005)
2b3.1.2	1995	1	86	10	8	0·0093
		2	91	12	7	0·0064
		3	91	12	4	0·0037
		4	92	12	4	0·0036
		5	94	8	1	0·0013
		Total/mean	91	54	24	0·0048 (0·0014)

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

Table 2 (*cont.*)

Subline	Year	Individual	Copy number	Progeny number	No. of transpositions	Transposition rate
2b3.1.3	1995	1	94	24	27	0·0120
		2	96	39	3	0·0008
		3	101	19	10	0·0052
		4	102	42	13	0·0030
		5	103	15	10	0·0065
		Total/mean	99	139	63	0·0055 (0·0019)
2b3.2	1995	1	56	10	0	0
		2	58	8	2	0·0043
		3	58	13	4	0·0053
		4	60	23	2	0·0014
		5	62	14	1	0·0012
		Total/mean	59	68	9	0·0024 (0·0010)
2b3.1.1 × 2b1.1.2	1994	1	65	26	7	0·0040
		2	66	29	11	0·0056
		3	67	13	8	0·0089
		4	68	24	6	0·0036
		5	69	26	4	0·0022
		6	70	18	4	0·0031
		7	72	22	19	0·0117
		8	73	38	21	0·0074
		9	74	10	3	0·0040
		10	77	18	18	0·0127
		Total/mean	70	224	101	0·0063 (0·0018)
2b3.2 × 2b3.1.1.2	1995	1	69	21	0	0
		2	71	14	0	0
		3	72	20	0	0
		4	72	20	2	0·0013
		5	73	30	0	0
		6	75	24	0	0
		7	76	19	5	0·0035
		8	78	17	0	0
		Total/mean	73	165	7	0·0006 (0·0004)
		Total		1466	599	

[a] These data are partially published in Pasyukova *et al.* (1997).
[b] Mean subline copy numbers; total numbers of progeny analysed per subline; total numbers of transpositions found; mean transposition rates are shown.
[c] Standard errors were calculated from the variance among males.

For *copia*, a strong positive relationship between the rate of transposition and copy number was found in the Harwich line (Nuzhdin *et al.*, 1996). Also, we made an attempt to analyse the relationship between transposition rate and copy number for the *roo* retrotransposon, using replicates of the inbred Harwich line (Mackay *et al.*, 1992). We measured directly the rates of *roo* transpositions in eight males of four sublines of the Harwich line. Only six transpositions were detected by scoring 256 larvae. Copy number of *roo* in individual males ranged from 106 to 125 per diploid genome. Unfortunately, no association could be confirmed, due to the limited data on transpositions, and due to the small range of *roo* copy numbers across sublines. However, the distribution of the numbers of accumulated *roo* copies in the independently maintained sublines of the Harwich line reported earlier (Nuzhdin & Mackay, 1994) was not significantly different from random (goodness-of-fit statistic to a Poisson distribution was $\chi^2_8 = 3{\cdot}43$). This is expected under the null hypothesis of no relationship between the number of transpositions per gamete and TE copy number, which suggests that the relationship between *roo* transposition rate and copy number, if any, is also not negative.

Thus, in summary, the data available for four retrotransposons (*copia* in two independent cases, *Doc*, *412* and *roo*) either directly demonstrate or indicate that a negative association between TE transposition rate and its copy number in the genome is not usual. Note that the *copia* and *Doc* copy numbers in the genome of some sublines are very much higher than is typical for laboratory lines and natural populations (Finnegan, 1992; Biemont *et al.*, 1994). A negative association between the TE transposition rate and the copy number would be expected if TEs self-regulate their copy number. Thus, this hypothesis

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

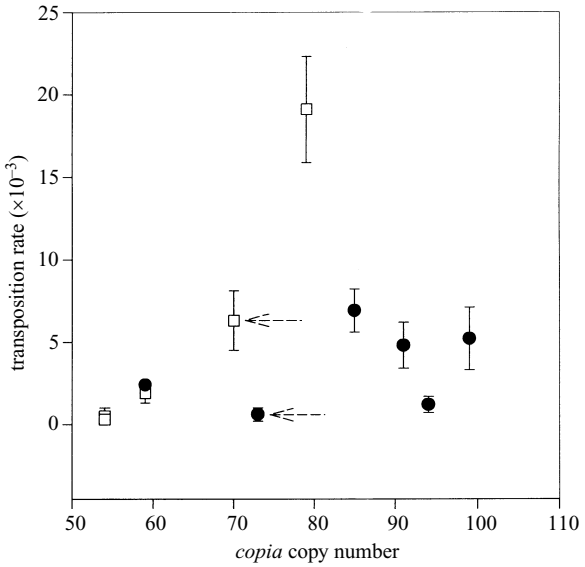

Fig. 3. Rates of *copia* transposition versus diploid *copia* copy number in sublines of the 2b line. Measurement were made in 1994 (open squares) and in 1995 (filled circles). Transposition rates in F_1 males from crosses 2b3.1.1 × 2b1.1.2 and 2b3.1.1.2 × 2b3.2 are included (shown by arrows).

for explaining the maintenance of a stable copy number characteristic for natural populations is not supported for the above-mentioned retrotransposons. The possibility of self-regulation explaining equilibria has already been rejected on the basis of indirect population data (Charlesworth & Langley, 1989), and our direct measurements support this conclusion.

(ii) *The relationship between TE transposition rate and its copy number in the genome is not constant in the same line*

For the rate of *copia* transposition in 2b we reconstructed the historical pattern, which could be divided into two phases: an earlier period, from 1991 to 1994, when there was a greater than linear relationship between transposition rate and copy number, and a later phase, in 1995, when the transposition rate showed no significant association with copy number. It can be seen from Fig. 3 that, in 1995, the *copia* copy number in some of the 2b sublines is increased compared with the maximal copy number observed in 1994. Despite this increase, the *copia* transposition rate was reduced. Thus, the *copia* transposition rate peaked at a *copia* copy number equal to 80 per diploid genome and then turned downward. These facts can easily be explained if a dominant mutation(s) decreasing *copia* transposition rate occurred in the 2b3.1.1 subline with high *copia* copy number and was(were) inherited by all the progenitors (see Fig. 1). We do not consider self-regulation as a probable

Table 3. Doc *transposition rate in 2b females*

Subline	Individual	Copy number	Progeny number	No. of transpositions	Transposition rate
2b1.2	1	112	25	0	0
	2	116	27	0	0
	3	116	24	0	0
	Total/mean[a]	115	76	0	0
2b2.2	1	152	32	2	0·0004
	2	153	27	0	0
	Total/mean	153	59	2	0·0002 (0·0002)[b]
2b3.2	1	168	28	3	0·0006
	2	173	23	2	0·0005
	Total/mean	171	52	5	0·0006 (0·00005)
2b1.1	1	172	23	4	0·0010
	2	178	29	7	0·0014
	Total/mean	175	52	11	0·0012 (0·0002)
2b3.1.1.1	1	195	17	6	0·0018
	2	199	17	2	0·0006
	3	206	24	6	0·0012
	4	213	24	13	0·0025
	5	220	21	14	0·0030
	Total/mean	207	103	41	0·0018 (0·0004)
2b3.1.3	1	222	21	2	0·0004
	2	235	24	4	0·0007
	Total/mean	229	45	6	0·0006 (0·0002)
	Total		387	65	

[a] Mean subline copy numbers; total numbers of progeny analysed per subline; total number of transpositions found; mean transposition rates are shown.
[b] Standard errors were calculated from the variance among males.

https://doi.org/10.1017/S0016672398003358 Published online by Cambridge University Press

explanation in this case because the copy number at which it could have been switched on is far above the copy numbers typical for natural populations (Biemont *et al*., 1994).

(iii) *A positive greater than linear relationship between* copia *copy number and transposition rate*

For *copia*, a strong positive greater than linear relationship between the rate of transposition and copy number was found in 1994 in two independent lines, Harwich and 2b (Nuzhdin *et al*., 1996, and this study). There are several hypotheses that could explain the observed differences in *copia* transposition rate between the sublines within each line. The first hypothesis is that a mutation affecting *copia* transposition rate occurred in the *copia* elements themselves ('rogue' *copia*s), and these rogue *copia*s are represented by different numbers of copies in the replicates of the line. If two sublines differ in the relative amount of rogue *copia*s, the expected number of transpositions in their hybrids will be strictly additive (for details of the calculation see Nuzhdin *et al*., 1996). The hypothesis of rogue *copia* is not rejected for 2b, as in the 2b3.1.1 × 2b1.1.1 hybrid males *copia* transposition rate is not significantly different from the additive expectation ($t = 0{\cdot}42$, $P > 0{\cdot}05$). The second hypothesis is that host mutations are responsible for the differences in transposition rate between different sublines. For the 2b line, at present we do not have sufficient grounds to either support or reject this. Our preliminary unpublished results show that, if any, host factors controlling *copia* transposition rate in 2b are multiple, interacting and probably localized in more than one complementation group. So far, *gypsy* is the only long terminal repeat retrotransposon of *Drosophila* for which the genetic dissection of transposition rate has been successful. Both the rogue copy and the host mutation hypotheses hold for *gypsy*: two structural variants of *gypsy* differing in transpositional activity were discovered (Lyubomirskaya *et al*., 1990), and the active structural variant does not transpose unless mutation at the X-linked locus, *flamenco* (Prud'homme *et al*., 1995), is present in the line (Kim *et al*., 1994).

However, in the context of this work we would like to stress the existence of the third hypothesis explaining the differences in *copia* transposition rate between sublines of a line. This is that no mutations affecting *copia* transposition rate either appeared or pre-existed in sublines but that *copia* copy number in the genome is itself the factor influencing transposition rate, due to biological properties of the transpositional process. This hypothesis is considered the most probable explanation of the positive greater than linear relationship between *copia* transposition rate and copy number in Harwich sublines (Nuzhdin *et al*., 1996).

There are biological grounds for believing that a positive correlation of the rate of transposition and TE copy number should be a default starting point for future host–TE coevolution. For transposons, for example, the number of transposon transcripts per cell should be proportional to n, where n is the TE copy number (the simplest assumption of no regulation of transposition before host–TE coevolution is adopted here). The concentration of the transposase should therefore also be proportional to n. In this case the probability of the transposase meeting a transposon copy to initiate transposition is proportional to n^2. Thus, the number of transpositions per cell should be proportional to n^2, and the rate of transposition per copy to n (Brookfield, 1991). However, this relationship may be dramatically changed as a result of coevolution (Brookfield, 1991, 1996). One may argue that TEs should have evolved to self-regulate their own transpositions. However, it has been shown theoretically that self-regulation of the rate of transposition is not a favourable evolutionary strategy for TEs. This holds true unless there is a strong deleterious effect associated directly with the TE copy whose activity is responsible for transposition (the mother copy), rather than with its progeny copies (Charlesworth & Langley, 1986). The latter is the case for TEs causing hybrid dysgenesis in *D. melanogaster*, and these elements indeed provide an evidence of self-regulation (Vaury *et al*., 1993; Biemont, 1994).

For retrotransposons, possible biological mechanisms of transposition rate dependence on copy number could be found. One may be titration of a hypothetical host protein that suppresses *copia* transposition (J. McDonald, personal communication): if the concentration of this host protein is constant, it may effectively suppress transposition at low but not at high *copia* copy numbers. The other follows from the fact that transposition happens via a virus-like particle (VLP) that is formed from full-length *copia* transcript and several *copia*-encoded proteins including integrase–reverse transcriptase. Assuming that the concentration of each VLP component is proportional to *copia* copy number, the probability of VLP formation and correspondingly the transposition rate may depend on *copia* copy number raised to the power that equals the number of VLP components minus one. The greater than linear relationship between *copia* transposition rate and copy number observed in 1994 in both Harwich and 2b sublines is in agreement with this prediction.

The implication of this hypothesis is as follows: if by chance alone *copia* copy number in a line with transposition rate of 10^{-4} exceeds a threshold level, *copia* transposition becomes a self-accelerating process. This threshold level is about 50 euchromatic copies, based on the data on 2b. Note, for example, that *copia* transpositions were not detected in the 2b1

subline in 1991 (Pasyukova & Nuzhdin, 1993) but in 1994 a low *copia* transposition rate was recorded. This could be explained assuming *copia* copy number in 2b1 gradually exceeded the threshold level due to very rare transpositions.

(iv) *Transpositions spectra of retrotransposons*

The majority of *Drosophila* lines with increased transposition rate of TEs were first described as genetically unstable lines, i.e. as lines with a high mutation rate (for example, Mevel-Ninio *et al*., 1989; Engels, 1989; Kim *et al*., 1990; Georgiev *et al*., 1990). Although the mutagenic potential of TEs is not doubted in general, significant increases in the total mutation rate in *Drosophila* lines could be a consequence of the increase of at least 2 orders of magnitude in the transposition rate of a single family of TEs. Indeed, the spontaneous transposition rate for the total population of *D. melanogaster* TEs is estimated as approximately 0·5 transpositions per genome per generation (Nuzhdin & Mackay, 1994). Even in dysgenic crosses, where the mean transposition rate of P-element is about 6·25 transpositions per genome per generation, the transposition rate of the total population of TEs is only 1 order of magnitude higher than the spontaneous level, and frequency of lethals is of the same order of magnitude as the spontaneous rate (for references see Engels, 1989). The frequency of sex-linked lethals is the same in 2b sublines with *copia* transposition rate differing by an order of magnitude (unpublished results).

In this context, the high frequencies of mutations at particular loci observed in unstable lines with high TE transposition rates provokes the thought that some site specificity of TE insertions may exist. Indeed, high variation in the rates at which different loci acquire insertions of P-element (for references see Engels, 1989), *gypsy* (Mevel-Ninio *et al*., 1989) and Stalker (Georgiev *et al*., 1990) has been reported. Hot spots of mdg1 were described at a cytological level (Pasyukova *et al*., 1986). In our experiments we could directly detect all the cytological sites of *copia* transpositions except those that were affected by selection on the level of gametes, embryos and larvae. Though 599 *copia* transpositions and 63 *Doc* transpositions revealed in the 2b line, as well as 492 *copia* transpositions in the inbred Harwich sublines (Nuzhdin *et al*., 1996), were not randomly distributed along cytological subdivisions of polytene chromosomes, the following explanation of this fact seems to be most probable: there is a correlation between the amount of DNA and the number of independent TE insertions per cytological subdivision. The lack of site preference within chromosomes fits the population data based on the analysis of the distributions of TE insertions along chromosomes (for other references see also Biemont *et al*., 1994). However, taking into consideration all the facts, we can not exclude the possibility that the specificity of the transposition spectrum depends on the element considered.

Although *copia* transposition was not site-specific in our experiments, a slight but significant advantage of the whole X chromosome as a target for *copia* transposition was discovered. The cause of this phenomenon is unclear. The transpositions studied took place in males, which have a diploid set of autosomes and a single X chromosome subjected to dosage compensation. We could speculate that this basic mechanism is somehow responsible for a disproportionate ability to acquire *copia* transpositions, as well as the intensification of other processes, such as transposition, in the single X chromosome of males. However, this speculation does not hold, and a mechanism more relevant to particular properties of the 2b line should be proposed, since in the Harwich sublines the distribution of *copia* transpositions on chromosomal arms was random (Nuzhdin *et al*., 1996).

We thank Vladimir Gvozdev and Trudy Mackay for support during the progress of this work, for helpful discussions and comments on the manuscript; Jim Fry for assistance with the statistical analyses; Elena Morozova for technical assistance in experiments; and anonymous reviewers for their helpful comments on the manuscript. This work was supported by a grant from the Ministry of Sciences of Russia and by grants 94-04-11423 and 97-04-48101 from Russian Foundation of Basic Research for EGP; the work of S.V.N. and E.G.P. was also supported by grants NIH GM 45344 and GM 45146 to Trudy Mackay.

References

Ashburner, M. (1989). *Drosophila: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Biemont, C. (1994). Dynamic equilibrium between insertion and excision of P elements in highly inbred lines from an M′ strain of *Drosophila melanogaster*. *Journal of Molecular Evolution* **39**, 466–472.

Biemont, C., Lemeunier, F., Garcia Guerreiro, M. P., Brookfield, J. F., Gautier, C., Aulard, S. & Pasyukova, E. G. (1994). Population dynamics of the *copia, mdg1, mdg3, gypsy* and P transposable elements in a natural population of *Drosophila melanogaster*. *Genetical Research* **63**, 197–212.

Brookfield, J. F. Y. (1991). Models of regulation of transposition in P and 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. *Genetical Research* **67**, 199–211.

Charlesworth, B. & Charlesworth, D. (1983). The population dynamics of transposable elements. *Genetical Research* **42**, 1–27.

Charlesworth, B. & Langley, C. H. (1986). The evolution of self-regulated transposition of transposable elements. *Genetical Research* **42**, 1–27.

Charlesworth, B. & Langley, C. H. (1989). The population genetics of *Drosophila* transposable elements. *Annual Review of Genetics* **23**, 251–287.

Engels, W. R. (1989). P elements in *Drosophila melanogaster*. In *Mobile DNA*, ed. D. E. Berg & M. M. Howe, pp. 437–484. Washington, DC: American Society for Microbiology.

Filatov, D. A., Morozova, T. V. & Pasyukova, E. G. (1998). Age-dependent *copia* transposition rate is positively associated with *copia* transcript abundance in a *Drosophila melanogaster* inbred line. *Molecular and General Genetics*, in press.

Finnegan, D. J. (1992). Transposable elements. In *The Genome of* Drosophila melanogaster, ed. D. L. Lindsley & G. G. Zimm, pp. 1096–1107. San Diego: Academic Press.

Georgiev, P. G., Kiselev, S. L., Simonova, O. B. & Gerasimova, T. I. (1990). A novel transposition system in *Drosophila melanogaster* depending on the *Stalker* mobile genetic element. *Journal of the European Molecular Biology Organization* **9**, 2037–2044.

Heino, T. I., Saura, A. O. & Sorsa, V. (1994). Maps of the salivary gland chromosomes of *Drosophila melanogaster*. *Drosophila Information Service* **73**, 621–738.

Kim, A., Belyaeva, E. S. & Aslanian, M. M. (1990). Autonomous transposition of *gypsy* mobile elements and genetic instability in *Drosophila melanogaster*. *Molecular and General Genetics* **224**, 303–308.

Kim, A., Lubomirskaya, N. V., Belyaeva, E. S., Shostack, N. G. & Ilyin, Y. V. (1994). The introduction of a transcriptionally active copy of retrotransposon *gypsy* into the Stable Strain of *Drosophila melanogaster* causes genetic instability. *Molecular and General Genetics* **242**, 472–477.

Langley, C. H., Montgomery, E. H., Hudson, R., Kaplan, N. & Charlesworth, B. (1988). On the role of unequal exchange in the containment of transposable element copy number. *Genetical Research* **52**, 223–236.

Lefevre, G. (1976). A photographic representation of the polytene chromosomes of *Drosophila melanogaster* salivary glands. In *The Genetics and Biology of* Drosophila, vol 1a, ed. M. Ashburner & E. Novitski, pp. 31–36. London: Academic Press.

Lubomirskaya, N. V., Arkhipova, I. R., Ilyin, Y. V. & Kim, A. I. (1990). Molecular analysis of the *gypsy* (*mdg4*) retrotransposon in two *Drosophila melanogaster* strains differing by genetic instability. *Molecular and General Genetics* **223**, 305–309.

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

McDonald, J. F., Strand, D. J., Brown, M. R., Paskewitz, S. M., Csink, A. K. & S. H. Voss, A. K. (1988). Evidence of host-mediated regulation of retroviral element expression at the posttranscriptional level. In *Eukaryotic Transposable Elements as Mutagenic Agents*, ed. E. M. Lambert, J. F. McDonald & I. B. Weinstein, pp. 219–234. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mevel-Ninio, M., Mariol, M. C. & Gans, M. (1989). Mobilization of the *gypsy* and *copia* retrotransposons in *Drosophila melanogaster* induces reversion of the ovo^D dominant female-sterile mutations: molecular analysis of revertant alleles. *Journal of the European Molecular Biology Organization* **8**, 1549–1558.

Nuzhdin, S. V. & Mackay, T. F. C. (1994). Direct determination of retrotransposon transposition rates in *Drosophila melanogaster*. *Genetical Research* **63**, 139–144.

Nuzhdin, S. V., Pasyukova, E. G. & Mackay, T. F. C. (1996). Positive association between *copia* transposition rate and copy number in *Drosophila melanogaster*. *Proceedings of the Royal Society of London, Series B* **263**, 823–831.

Nuzhdin, S. V., Pasyukova, E. G. & Mackay, T. F. C. (1997). Accumulation of transposable elements in laboratory lines of *Drosophila melanogaster*. *Genetica*, in press.

Pasyukova, E. G. & Nuzhdin, S. V. (1992). Mobilization of retrotransposon *copia* in the genome of *Drosophila melanogaster*. *Genetika* **28**, 5–18.

Pasyukova, E. G. & Nuzhdin, S. V. (1993). *Doc* and *copia* instability in an isogenic *Drosophila melanogaster* stock. *Molecular and General Genetics* **240**, 302–306.

Pasyukova, E. G., Belyaeva, E. S., Kogan, G. L., Kaidanov, L. Z. & Gvozdev, V. A. (1986). Concerted transpositions of mobile genetic elements coupled with fitness changes in *Drosophila melanogaster*. *Molecular Biology and Evolution* **3**, 299–312.

Pasyukova, E. G., Nuzhdin, S. V., Li, W. & Flavell, A. J. (1997). Germ line transposition of the *copia* retrotransposon in *Drosophila melanogaster* is restricted to males by tissue-specific control of *copia* RNA level. *Moecular and General Genetics* **255**, 115–124.

Prud'homme, N., Gans, M., Masson, M., Terzian, C. & Bucheton, A. (1995). *Flamenco*, a gene controlling the *gypsy* retrovirus of *Drosophila melanogaster*. *Genetics* **139**, 697–711.

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

Vaury, C., Pelisson, A., Abad, P. & Bucheton, A. (1993). Properties of transgenic strains of *Drosophila melanogaster* containing *I* transposable elements from *Drosophila teissieri*. *Genetical Research* **61**, 81–90.

Vieira, C. & Biemont, C. (1997). Transposition rate of the *412* retrotransposable element is independent of copy number in natural populations of *Drosophila simulans*. *Molecular Biology and Evolution* **14**, 185–188.