Genet. Res., Camb. (2001), **78**, *pp.* 121–136. © 2001 Cambridge University Press
DOI: 10.1017/S0016672301005201 Printed in the United Kingdom

Rates of movement and distribution of transposable elements in *Drosophila melanogaster*: *in situ* hybridization vs Southern blotting data

XULIO MASIDE[1]*, CAROLINA BARTOLOMÉ[1], STAVROULA ASSIMACOPOULOS[2] AND BRIAN CHARLESWORTH[1]
[1] *Institute for Cell, Animal and Population Biology, University of Edinburgh, Ashworth Laboratories, Edinburgh EH9 3JT, UK*
[2] *Department of Pharmacology and Physiology, University of Chicago, Chicago, IL 60637, USA*

(*Received 19 January 2001 and in revised form 5 April 2001*)

Summary

Genomic copy numbers and the rates of movement of nine families of transposable elements (TEs) of *Drosophila melanogaster* were estimated in two sets of mutation accumulation lines: Beltsville and Madrid. Southern blotting was used to screen a large number of samples from both genetic backgrounds for TEs. The Madrid lines were also screened by *in situ* hybridization of TEs to polytene chromosomes, in order to obtain more detailed information about the behaviour of TEs in the euchromatin. Southern blotting data provided evidence of insertions and excision events in both genetic backgrounds, occurring at rates of approximately 10^{-5} and 10^{-6} per element copy per generation, respectively. In contrast, *in situ* data from the Madrid background presented a completely different picture, with no evidence for excisions, and a significantly higher rate of transposition ($1{\cdot}01 \times 10^{-4}$). Direct comparison of the two data sets suggests that the Southern blotting technique had serious deficiencies: (i) it underestimated element abundance; (ii) it revealed less than 30% of the new insertions detected by *in situ* hybridization; and (iii) changes in the size of restriction fragments from any source were spuriously identified as simultaneous insertion–excision events. Our *in situ* data are consistent with previous studies, and suggest that selection is the main force controlling element spread by transposition.

1. Introduction

Transposable elements (TEs) are short fragments of DNA capable of inserting replicas of themselves into novel genomic locations (Berg & Howe, 1989). They are present in most species, and contribute to a significant part of their genome: from 1% to 10% in species of the genus *Drosophila*, up to 40% in humans, or even more than 90% in some plants such as wheat or pine (Flavell, 1986; Berg & Howe, 1989; Capy, 1997; Lander *et al.*, 2001). TEs are a major evolutionary force both at the genetic (Finnegan, 1992*b*) and the cytogenetic level (Steinemann & Steinemann, 1998; Andolfatto *et al.*, 1999; Cáceres *et al.*, 1999), and the study of the forces that determine their abundance is a task of major importance if we want to understand the organization and evolution of the genome (Charlesworth *et al.*, 1994*b*; Capy, 1997).

* Corresponding author. ICAPB, Kings Buildings, University of Edinburgh, Edinburgh, EH9 3JT, UK.
e-mail: xulio.maside@ed.ac.uk

Population genetic models propose that TEs are maintained by an equilibrium between their ability to spread by transposition, and the rate at which they are eliminated, either by self-mediated excision or by selection against the presence of insertions (Charlesworth & Charlesworth, 1983; Langley *et al.*, 1983, 1988; Kaplan *et al.*, 1985; Brookfield, 1996). Data from surveys of TEs on *Drosophila melanogaster* populations suggest that the mean rate of elimination of elements in nature is $\sim 10^{-4}$ per copy per generation, which implies that transposition rates of similar magnitude are expected if elements are at or near equilibrium in these populations (Charlesworth & Langley, 1991).

The models can therefore be tested by obtaining independent estimates of the rate at which elements move in nature. However, very few studies have directly addressed this question (Nuzhdin & Mackay, 1995; Domínguez & Albornoz, 1996; Nuzhdin *et al.*, 1997; Maside *et al.*, 2000), and their results are far from conclusive. These studies consisted of the

quantification of element transpositions and excisions in accumulation experiments by means of *in situ* hybridization of TEs to polytene chromosomes, or of Southern blotting of whole genomic DNA with TE probes. The *in situ* studies produced evidence that transposition rates in the euchromatin are indeed usually very low, of the order of 10^{-4}, and that excision rates are practically negligible (around 2 orders of magnitude lower) (Nuzhdin *et al*., 1997; Maside *et al*., 2000); on some genetic backgrounds, however, certain elements can move at an exceptionally high rate (Biémont *et al*., 1987; Pasyukova & Nuzhdin, 1993; Nuzhdin & Mackay, 1994; Prud'homme *et al*., 1995; Nuzhdin *et al*., 1998).

The results from *in situ* hybridization experiments are broadly compatible with the population genetic models, in particular with the view of TEs as intragenomic parasites (the selfish DNA model) (Doolitle & Sapienza, 1980; Orgel & Crick, 1980), and with selection as the main force controlling their abundance (Charlesworth & Langley, 1991). Interestingly, the only Southern analysis yet available produced a rather different picture (Domínguez & Albornoz, 1996). Estimates of the rates of transposition were significantly lower than the above, of the order of 10^{-6}, and similar to those of excision. The authors suggested that this means that TEs cannot be considered to be mere parasites, and that complex regulatory interactions between the different TE families and the host genome need to be invoked to account for the population data (Domínguez & Albornoz, 1996). This raises the question of the importance of the genetic background for the outcome of accumulation experiments, and the extent to which *in situ* data are representative of the dynamics of TEs in the whole genome.

Here we report the results of a study on the rates of movement of the same nine families of TEs in two different genetic backgrounds, which provide new estimates. We combined Southern blotting and *in situ* hybridization, in order to obtain a meaningful comparison of the results from the two techniques. Our findings are consistent with the selfish DNA hypothesis, and question the utility of the Southern blotting approach for these purposes.

2. Materials and methods

(i) *Genetic stocks*

Element movements were studied in two independent genetic backgrounds. The 'Beltsville' accumulation lines consisted of a set of 60 replicate inbred lines derived from an isogenic stock, *IS-31*, constructed from X, second and third chromosomes independently extracted from the Beltsville natural population (Charlesworth, *et al*., 1994*a*), and kept by single pair matings for 31 generations. The 'Madrid' background consisted of 200 inbred lines derived from a completely isogenic foundation stock (Caballero *et al*., 1991). Inbred lines were kept by double first cousin matings or single brother–sister matings for 47 generations (Santiago *et al*., 1992; Fernández & López-Fanjul, 1996), and by single brother–sister matings for a further 215 generations. In both cases, flies were reared on standard cornmeal–agar–yeast–sucrose medium at 25 ± 1 °C.

(ii) *Probes*

Probes for nine families of retrotransposable elements were used: *297*, *412*, *1731*, *copia*, *mdg-1*, *opus* and *roo* are retrotransposable elements with long terminal repeats (LTRs); *doc* and *jockey* are retrotransposons without LTRs (Finnegan, 1992*a*). DNA fragments internal to each TE family (except *doc*) were isolated from the DNA genomic clones described by Charlesworth *et al*. (1994*a*), by digestion with selected restriction enzymes, followed by purification of the desired bands by agarose gel electrophoresis. Isolated DNA fragments were then subcloned into an appropriate vector (pBluescript 2KSP, Stratagene Cloning Systems), and the subclones checked by direct sequencing from the M13 primers at both ends of the vector. A DNA fragment internal to *doc* was obtained by PCR amplification with specific primers from total genomic DNA from the accumulation lines and subcloned with the Topo-TA cloning kit (Promega). The lengths of the isolated fragments, restriction enzymes and primer sequences used for the construction of the clones are shown in Table 1.

For making the probes used in the experiments described below, element sequences were isolated from the subclones by PCR amplification from the M13 primers derived from the cloning vectors. The resulting products were purified by agarose gel electrophoresis or with the Qiaquick PCR Purification Kit (Qiagen), and labelled with the Renaissance Random Primer Fluorescein Labeling Kit (NEN) for the Southern blotting analyses, or with Biotin High Prime (Boehringer Mannheim) for the *in situ* hybridization analyses.

(iii) *Southern blotting analyses*

Total genomic DNA from 20–25 individuals from each accumulation line was extracted with Puregene Isolation Kit (Gentra). Five micrograms of DNA per line were digested at 37 °C for 5 h (Table 1), sodium acetate–ethanol precipitated, and resuspended in H_2O. Electrophoresis was carried out in a 25×15 cm, 0·6 % agarose gel, at 0·20 V/cm², until the 3 kb marker

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

Table 1. *TE probes and restriction enzymes used for the Southern analysis*

TE family	Source DNA	Restriction enzymes	Fragment size (kb)	Restriction enzyme for the genomic DNA
copia	*cDm5002*	*Hind* III–*Eco* RI	3·1	*Bam* HI
297	*cDm*4006	*Hind* III–*Eco* RV	2·1	*Bam* HI
mdg-1	*cDm2181*	*Bam* HI–*Hind* III	5·2	*Hind* III
jockey	*cDm2161*	*Pst* I–*Xba* I	2·2	*Bam* HI
412	*cDm2042*	*Bam* HI–*Hind* III	5·1	*Bam* HI
opus[a]	*cDm2217*	*Sal* I–*Eco* RI	3·3	*Bam* HI
1731	*cDm2158*	*Pst* I– *Sal* I	2·6	*Hind* III
doc	Genomic DNA	PCR amplification[b]	2·8	*Hind* III
roo	*B104*	*Hind* III–*Sal* I	2·3	*Hind* III

[a] Comparisons of our sequences of the DNA fragment from *opus*, with sequence data from the FlyBase, as well as the corresponding available restriction maps (Finnegan, 1992*a*; Charlesworth *et al*., 1994*a*), revealed that *opus* (Kidd, 1986) and *nomad* (also called *yoyo*) (Whalen, 1998) are the same element.
[b] The primers for the amplification of the *doc* internal sequence were doc-f: 5′ CTT CGT AGG CGT CGT TTA TC 3′ and doc-r: 5′ TTC GGC ATT CCA CAG TCT TC 3′.

reached the bottom of the gel (aprox. 14 h). DNA was then denatured, transferred and UV-crosslinked onto a positively charged nylon membrane (Boehringer Mannheim) using standard methods (Sambrook *et al*., 1989). DNA–DNA hybridization and detection were performed using the Renaissance Random Primer Fluorescein Labeling Kit with Antifluorescein-AP (NEN). In order to obtain a single band per element insertion in the Southern blots, restriction enzymes that do not cut within the portion of the element sequence homologous to the TE probes were selected for digestion of the genomic DNA (Table 1).

Element movements were detected from changes in the pattern of bands in the accumulation lines. In each genetic background, the ancestral pattern was inferred from the bands shared by all the lines from the same background. New bands were considered to represent new insertions and band losses, excision events. With the chosen combination of probes and restriction enzymes, full element insertions were expected to create bands with an average size greater than 7 kb (at least 3 kb of element sequence plus an additional flanking sequence with a mean size 4 kb, if the restriction enzyme cuts within the element). In order to avoid the inclusion of defective element copies in the counts, only bands of 5 kb or larger were considered in the following analyses.

(iv) *In situ hybridization*

Salivary glands from third-instar larvae were dissected and hybridized following previously described protocols (Sniegowski & Charlesworth, 1994), with minor changes: salivary glands were fixed in 1:2:3 solution for 2 min prior to squash; all dehydration times in ethanol 70 % and 100 % were shortened to 1 min and 2 min, respectively; the denaturation step with NaOH 0·07 N was 4 min; the first wash in 2 × SSC after hybridization was performed at 50 °C; and 50 µl of DAB solution per slide was used in the detection step. Slides was examined under a × 40 phase contrast objective, and stained with 5 % pH 7·1 Giemsa solution when chromosome bands were faint or when a large number of signals complicated the readings.

In the salivary glands of *Drosophila* larvae, chromosomes undergo polytenization, a process by which euchromatin and β-heterochromatin (Dimitri, 1997) go through up to nine rounds of replication without separation of the resulting sister chromatids. The remaining chromatin is condensed, non-polytene α-heterochromatin, found as a compact region around the centromere of each chromosome. The β-heterochromatin does not present the banding pattern characteristic of the euchromatin, and can be seen only as an amorphous region at the bases of each chromosome arm. Hybridization signals in the α- or β-heterochromatin should therefore be considered as merely indicative of the presence of TEs, but no further information as to element abundance or location can be extracted. The chromosomal sections included in this analysis were: 1A to 20B (X), 21A to 40F (2L), 41A to 60F (2R), 61A to 80C (3L), 81F to 100F (3R) and 101F to 102F (4), as depicted in the Lefevre photographic maps of the polytene chromosomes (Lefevre, 1976). Element excisions and new insertions were detected from changes in the position of hybridization signals of the different TE probes on the polytene chromosomes. Several nuclei were analysed per slide. Replicate slides were made whenever counts from the different nuclei were inconsistent.

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

3. Results

(i) *Southern blotting analyses*

Element copy numbers in the ancestral state, and numbers of new insertions and excisions detected in both genetic backgrounds for each family, are shown in Table 2. Mean element abundances in the Beltsville and Madrid lines are $22{\cdot}4 \pm 1{\cdot}27$ and $22{\cdot}9 \pm 2{\cdot}57$ copies per family per genome, respectively. Family copy numbers vary between 12 and 39, are very similar between genetic backgrounds, and in general are of the same magnitude as those reported in the literature (Nuzhdin & Mackay, 1995; Domínguez & Albornoz, 1996; Vieira *et al*., 1999; Maside *et al*., 2000). It should be pointed out here that independent Southern estimates of element abundance of *297*, *412*, *copia* and *jockey* in the same Madrid lines, can be found in Domínguez & Albornoz (1996, table 3). There is no systematic difference between the two data sets, and the smaller estimates for *412* and *297* in our study can be explained by: (1) in Domínguez & Albornoz (1996) these two families were expected to generate more than one band per element copy, and (2) we only considered bands larger than 5 kb (see Section 2.iii). *Roo* is the most abundant family, with 27 and 39 copies per genome in the respective backgrounds, although these values are well below the usual abundances reported for this family from *in situ* data on natural or laboratory populations (Nuzhdin & Mackay, 1995; Biémont & Cizeron, 1999; Maside *et al*., 2000).

For each family, the transposition and excision rates per element copy per generation were estimated by dividing the number of events (gains or losses of bands, respectively) by the number of opportunities (calculated as the total number of copies × number of lines × number of generations of accumulation), as in Nuzhdin & Mackay (1994). When no movements were detected for a given family, the 95% upper confidence limits for the transposition and excision rates were calculated, assuming that element movements follow a Poisson distribution (Table 2). The rates of movement of different families and the effect of genetic background were compared by means of a likelihood-ratio test, assuming that twice the difference in log-likelihood of the parameters is χ^2 distributed, with one degree of freedom for pairwise comparisons (Sokal & Rohlf, 1995, chapter 17).

Five movements were detected in the Beltsville background after 31 generations of accumulation: four insertions and one excision (Table 2). The *mdg-1* family, with two transpositions, is the most active. In this background, differences in transposition rates between families are not statistically significant. The pooled transposition rate for all families is $1{\cdot}27 \times 10^{-5}$, 4 times that of excision ($3{\cdot}17 \times 10^{-6}$). As one would expect with such a small total number of events, these rates are non-significantly different from each other.

In the Madrid background, after a much longer period of accumulation (262 generations), 237 elements had moved, with the detection of 207 insertions and 30 excisions (Table 2 and Appendix). *Roo* is the most active family, with an estimated transposition rate of $2{\cdot}08 \times 10^{-4}$, 10 times higher than the next most active family, *doc*, which transposed at a rate of $2{\cdot}10 \times 10^{-5}$ new insertions per copy per generation ($P < 0{\cdot}001$). No movements were detected for *412*. Excision rates were lower than transposition rates for most families, and the pooled rate of excision ($6{\cdot}8 \times 10^{-6}$) was significantly lower than that of transposition ($4{\cdot}69 \times 10^{-5}$; $P < 0{\cdot}001$), as usually observed from *in situ* generated data sets (Nuzhdin *et al*., 1997; Maside *et al*., 2000).

The pooled transposition rate is lower in the Beltsville than in the Madrid background: $1{\cdot}27 \times 10^{-5}$ versus $4{\cdot}69 \times 10^{-5}$, respectively ($P < 0{\cdot}005$). This difference could be interpreted as evidence of host genotype–TE interactions determining the rates of movement in the different genetic backgrounds (Pasyukova & Nuzhdin, 1993; Biémont *et al*., 1997*b*; Nuzhdin *et al*., 1998). The pooled excision rates are not, however, significantly different between backgrounds. If *roo* is excluded from the calculation on the basis that it is an outlier, the difference between the pooled transposition rates from the two genetic backgrounds becomes non-significant, suggesting that host genotype–TE interactions take place in a family-specific manner. This is in agreement with the high heterogeneity observed in the rates of transposition of some other element families in different genetic backgrounds, as have been reported for *doc* and *copia* (Pasyukova & Nuzhdin, 1993). A high rate of transposition of *412* in one out of five sets of independent isogenic lines constructed with chromosomes extracted from the Beltsville population has also been detected by us (data not shown).

(ii) *Element abundances and rates of movement from* in situ *hybridization results*

Here we report data from *in situ* hybridization with probes for nine TE families in 10 inbred lines randomly chosen from the Madrid genetic background previously used for the Southern analysis. The accumulation period was 265 generations. The ancestral state of each TE family was inferred from the sites common to the 10 lines analysed. All elements in the ancestral state as well as the new insertions were mapped onto the polytene chromosome map of Lefevre (1976). Detailed data on element positions will be provided upon request. Two of the chosen lines, M21 and M22, shared two of the six new *roo* insertions found in them. The probability of this being result of chance is

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

Table 2. *Element family abundances and rates of movement in two different genetic backgrounds, estimated from Southern data*

Genetic background	TE family	No. of lines analysed	Element copy no.	No. of new insertions	No. of new excisions	Total no. of movements	Size of the experiment	Insertion rate	Upper 95% CI	Excision rate	Upper 95% CI
Beltsville	*copia*	24	20	1	0	1	14880	$6{\cdot}72\times10^{-5}$			$2{\cdot}01\times10^{-4}$
	297	24	23	0	0	0	17112		$1{\cdot}75\times10^{-4}$		$1{\cdot}75\times10^{-4}$
	mdg1	58	19	2	0	2	34162	$5{\cdot}85\times10^{-5}$			$8{\cdot}77\times10^{-5}$
	jockey	58	23	1	0	1	41354	$2{\cdot}42\times10^{-5}$			$7{\cdot}24\times10^{-5}$
	412	58	27	0	0	0	48546		$6{\cdot}17\times10^{-5}$		$6{\cdot}17\times10^{-5}$
	opus	58	21	0	1	1	37758		$7{\cdot}93\times10^{-5}$	$2{\cdot}65\times10^{-5}$	
	1731	58	16	0	0	0	28768		$1{\cdot}04\times10^{-4}$		$1{\cdot}04\times10^{-4}$
	doc	58	26	0	0	0	46748		$1{\cdot}41\times10^{-5}$		$6{\cdot}41\times10^{-5}$
	roo	58	27	0	0	0	48546		$6{\cdot}17\times10^{-5}$		$6{\cdot}17\times10^{-5}$
	Pooled	50	202	4	1	5	315883	$1{\cdot}27\times10^{-5}$		$3{\cdot}17\times10^{-6}$	
Madrid	*copia*	84	25	5	3	8	550200	$9{\cdot}09\times10^{-6}$		$5{\cdot}45\times10^{-6}$	
	297	84	24	5	4	9	528192	$9{\cdot}47\times10^{-6}$		$7{\cdot}57\times10^{-6}$	
	mdg1	80	24	9	3	12	503040	$1{\cdot}79\times10^{-5}$		$5{\cdot}96\times10^{-6}$	
	jockey	84	23	2	3	5	506184	$3{\cdot}95\times10^{-6}$		$5{\cdot}93\times10^{-6}$	
	412	83	12	0	0	0	260952		$1{\cdot}15\times10^{-5}$		$1{\cdot}15\times10^{-5}$
	opus	83	14	4	2	6	304444	$1{\cdot}31\times10^{-5}$		$6{\cdot}57\times10^{-6}$	
	1731	80	20	5	1	6	419200	$1{\cdot}19\times10^{-5}$		$2{\cdot}39\times10^{-6}$	
	doc	80	25	11	2	13	524000	$2{\cdot}10\times10^{-5}$		$3{\cdot}82\times10^{-6}$	
	roo	78	39	166	12	178	797004	$2{\cdot}08\times10^{-4}$		$1{\cdot}51\times10^{-5}$	
	Pooled	81·8	206	207	30	237	4413710	$4{\cdot}69\times10^{-5}$		$6{\cdot}80\times10^{-6}$	

Table 3. *Rates of transposition in the Madrid genetic background, estimated from* in situ *data*

TE family	No. of lines analysed	Element copy number							No. of new Insertions	Size of the experiment	Insertion rate	Upper 95% CI
		X	2R	2L	3R	3L	4	Total				
copia	10	3	6	4	3	3	0	19	0	50350		$5{\cdot}95\times\ 10^{-5}$
297	10	5	7	10	3	3	1	29	0	76850		$3{\cdot}90\times10^{-5}$
mdg1	10	2	5	4	7	3	0	21	0	55650		$5{\cdot}38\times10^{-5}$
jockey	10	17	7	9	17	9	0	59	0	156350		$1{\cdot}92\times10^{-5}$
412	10	2	3	5	2	4	0	16	0	42400		$7{\cdot}07\times10^{-5}$
opus	10	1	2	5	6	3	0	17	0	45050		$6{\cdot}65\times10^{-5}$
1731	10	4	7	1	6	4	0	22	0	58300		$5{\cdot}14\times10^{-5}$
doc	10	2	12	8	12	9	0	43	0	113950		$2{\cdot}63\times10^{-5}$
roo	10	14	16	9	15	8	1	63	77	166950	$4{\cdot}61\times\ 10^{-4}$	
Pooled	10	50	65	55	71	46	2	289	77	765850	$1{\cdot}01\times10^{-4}$	

The lines included in this study were: M5, M11, M18, M25, M26, M31, M32, M44, M73, M77.

very low, and points towards the possibility of a mixture of the lines during the accumulation process, a hypothesis reinforced by the fact that the two lines have consecutive ordinal numbers. As a precaution, these lines were discarded from the analyses and substituted by two others chosen at random.

Two hundred and eighty-nine element copies were detected in the ancestral estate. *Roo*, with 63 copies, is the most abundant family, followed by *jockey* and *doc* with 59 and 43 copies, respectively (Table 3). The mean element abundance per family is $32{\cdot}1 \pm 6{\cdot}1$, in approximate agreement with other *in situ* data reported from studies with different populations of the same species: $22{\cdot}4 \pm 3{\cdot}6$ in a study of 16 families in Nuzhdin & Mackay (1995), $27{\cdot}2 \pm 6{\cdot}9$ in Biémont & Cizeron (1999; averaged from the same subset of families included in the present study) or $23{\cdot}7 \pm 4{\cdot}7$ for a similar set of families in Maside *et al*. (2000). However, it is interesting to note that this estimate is significantly larger than the one obtained in our Southern analysis of the same genetic background ($22{\cdot}9 \pm 2{\cdot}6$; Section 3.i). Although the two approaches yield similar results for most families, particularly for those with low copy number, *in situ* abundance estimates for the most abundant ones tend to be significantly larger than the Southern estimates, being almost twice as large in the case of *doc*, *jockey* and *roo* (Tables 2, 3).

Two main factors associated with the Southern technique may account for this systematic bias. First, based on the expectations of the average size of the restriction fragments corresponding to element insertions, only bands of 5 kb or larger were considered for the Southern analysis (see Section 2.iii). Assuming that the sizes of the genomic sequences flanking the insertions (i.e. the distance between the insertion site and the first site for the restriction enzyme used to digest the genomic DNA in the flanking region) are normally distributed, the probability of finding bands shorter than 5 kb increases with family copy number. Indeed, a few bands between 1·5 and 5 kb were detected in additional Southern blots with probes for *jockey*, *doc* and *roo*, and none for the other families (data not shown). Second, given the level of resolution of agarose gel electrophoresis, it must be borne in mind that Southern blots may become saturated with bands when high copy number families are studied. The *in situ* data showed that *jockey*, *doc* and *roo* have between 40 and 65 copies per genome in the Madrid background, while the highest Southern estimate was 39. This is clear evidence that the Southern blotting was failing to produce enough bands to identify all the euchromatic insertions. The possibility that a significant proportion of the bands detected in the blots correspond to more than one insertion thus cannot be ruled out.

After 265 generations of accumulation, no excisions were detected. With zero events observed and assuming that excisions are Poisson distributed, the upper 95% confidence limit for the excision rate pooled across families is $3{\cdot}91 \times 10^{-6}$. Transpositional activity was detected for only one (*roo*) of the nine families included in the study. The resulting estimates of the rates of transposition are shown in Table 3. *Roo* showed 77 new insertions, transposing at a rate of $4{\cdot}61 \times 10^{-4}$ transpositions per element copy per generation. With the exception of *copia* or *doc* whose transposition rates are significantly lower, the estimates of the rates of movement are in excellent agreement with those reported in the few similar studies available (Biémont *et al*., 1987; Nuzhdin *et al*., 1997; Maside *et al*., 2000).

As with the abundance estimates, these *in situ* estimates of the rates of movement are both quantitatively and qualitatively different from those obtained in our Southern blotting analyses (Section 3.i). While the Southern data produced evidence of both element excision and transposition for eight of the nine families (Table 2), the *in situ* approach revealed transpositional activity only, and from just one of the families (*roo*). Furthermore, the pooled excision rate is significantly higher when estimated from Southern than from *in situ* data ($P < 0{\cdot}005$), while the *in situ* estimate of the pooled insertion rate is higher than the Southern one ($P < 0{\cdot}001$) (Tables 2, 3).

A line-by-line comparison of the results from the 10 lines for which both types of data were available (see note to Table 3) confirmed the profound differences between the results produced by the two techniques. Considering all the families except *roo*, none of the nine changes detected in the Southern data in these lines (five insertions and four excisions: see Appendix) had an equivalent in the *in situ* hybridizations (Table 3). In addition, the number of *roo* transpositions detected in the Southern blots (20) was only one-quarter of that estimated by the *in situ* approach (77), and no excision events were detected by *in situ* hybridization whereas two were inferred from the Southern blots.

There are two possible explanations for these large differences: either the two techniques do not detect the same subsets of elements of each TE family, or they do so but have different sensitivities to the genetic changes occurring during the accumulation period. TE movements in the euchromatin should be detected equally by the two techniques; however, changes in the length of the DNA restriction fragments that include the TEs, caused by insertions, deletions or point mutations altering the pattern of the restrictions sites, will not substantially change the *in situ* hybridization results but are likely to cause shifts of bands on the Southern blots and, consequently, an overestimate of the rates of movement.

The good agreement between abundance estimates

Table 4. *Association between transposition and excision events in the Southern data*

Family	No. of lines with:				*P* (Fisher's exact test)
	no changes	band gain	band loss	gain and loss	
copia	79	2	1	2	$5{\cdot}1 \times 10^{-3}$
297	79	1	1	3	$1{\cdot}7 \times 10^{-4}$
mdg1	70	7	2	1	0·27
jockey	82	0	0	2	$3{\cdot}0 \times 10^{-4}$
412	83	0	0	0	n.a.
opus	79	2	0	2	$1{\cdot}8 \times 10^{-3}$
1731	78	1	0	1	0·03
doc	75	3	0	2	$3{\cdot}2 \times 10^{-3}$
roo	10	57	0	11	0·20
Pooled	635	73	4	24	$6{\cdot}7 \times 10^{-19}$

n.a., not applicable to this set of data.

generated with the two techniques for families with intermediate and low copy number (see Section 3.ii) suggests that they detect roughly the same element insertions. Therefore, the hypothesis of different sensitivities of the techniques was thoroughly checked by: (i) cross-analysing all element movements detected in the Southern analyses of the Madrid background for families other than *roo* (Appendix) by *in situ* hybridization; and (ii) increasing the sample size of the *in situ* study of *roo* with six additional lines randomly selected from the same background (M89, M97, M137, M141, M144 and M197).

From the 59 band changes (41 gains and 18 losses) revealed in the Southern blots for all families other than *roo* (Appendix), only one TE movement (an *mdg-1* transposition in line M97) was detected by *in situ* hybridization. For *roo*, *in situ* hybridization of the six new lines produced evidence of 54 new insertions and no excision events, showing a similar pattern to that found in the former 10 lines (Table 3). Combining these *in situ* data sets, we have a total of 131 *roo* transpositions and no excision events, as opposed to the 35 insertions and 3 excisions inferred from the Southern analyses of the same set of lines (Appendix).

Given that, for most families of retrotransposable elements, there is no evidence for a correlation between insertion and excision events, an alternative way to test further the hypothesis of the different sensitivities of the techniques is to check whether there is a positive association between band gains and losses in the Southern data alone. This would be the expected outcome if the shift of bands observed in the Southern blots was due not to real element movements but to changes in the length of the restriction fragments that include the elements. While there is no evidence for an association within the *in situ* data (no excisions were detected; Table 3), Fisher exact tests of independence show that band gains and losses in the Southern data are indeed strongly associated in most families (Table 4), supporting the hypothesis that the Southern blotting technique detects events other than element movements.

(iii) *Element distributions and rates of recombination within the genome*

To investigate whether there is a correlation between the rate of meiotic recombination and the distribution of elements along the genome (Langley *et al*., 1988), we compared the observed proportion of elements in regions with different rates of recombination with random expectations. To do this, each arm of the major chromosomes was divided into three regions according to the local recombination rates (tip, middle and base, as in Langley *et al*., 1988), and the proportion of TEs in each region was determined. Estimates of the DNA content along the cytogenetic map from Heino *et al*. (1994) were used to calculate the proportion of the arm represented in each region, and the expected element abundances per region were then calculated by multiplying these values by the total element counts on each arm (Table 3).

In the ancestral state of the Madrid genetic background, elements are not randomly distributed along the chromosomes. Pooled counts in the three chromosomal regions (see above) show that elements tend to accumulate at the base of the arms ($\chi^2 = 10{\cdot}92$, $P < 0{\cdot}01$; Table 5). However, the intensity of this effect does not seem to be the same on all chromosome arms. It is significant on the X and 2R ($P < 0{\cdot}05$), marginal on 3R and 3L (revealed only by the sign of the deviations of the observed with respect to the expected abundances; Table 5) and undetectable on 2L and 4. When element insertions are pooled by family across chromosome arms, only the accumulation of the *copia* elements at the base of the chromosomes reaches statistically significant levels. Yet all TE families display a slightly skewed distribution, with evidence for an excess of elements at

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

Table 5. *Element distribution along the chromosome arms in the Madrid genetic background*

		Ancestral state			
Chromosome arm	Region	Expected	Observed	Deviation	χ^2
X	X-T	5·3	3	−	
	X-M	41·1	35	−	
	X-B	3·6	10	+	12·94**
2L	2L-T	3·4	0	−	
	2L-M	45·4	48	+	
	2L-B	6·2	7	+	3·69
2R	2R-T	5·2	6	+	
	2R-M	51·1	42	−	
	2R-B	8·7	17	+	9·81**
3L	3L-T	2·5	6	+	
	3L-M	38·8	35	−	
	3L-B	4·7	5	+	5·41
3R	3R-T	3·6	6	+	
	3R-M	61·3	56	−	
	3R-B	6·1	9	+	3·41
4	4	2·0	2		
Across chromosome arms					
	Tip	19·6	21	+	
	Middle	236·6	216	−	
	Base[a]	32·8	50	+	10·92**

[a] The fourth chromosome was considered to be a basal region.

the base and a deficiency at the middle regions, and all families except *copia*, *jockey* and *roo* show an excess of elements at the tip of the chromosomes (data not shown). These results are in good agreement with those reported from a natural population survey using a set of families that was similar to the present set (Charlesworth & Lapid, 1989; Charlesworth *et al.*, 1992*b*), where the authors found a tendency for elements to accumulate in the proximal regions of the X and the autosomes.

4. Discussion

(i) *Southern blotting versus* in situ *hybridization data*

Estimates of TE copy numbers and rates of movement generated by Southern blotting and *in situ* hybridization are substantially different, and the discrepancies cannot be accounted for simply by considering the different scopes of the techniques (the former screens the whole genome and the latter is restricted to the euchromatin). Southern estimates of family copy number are smaller than *in situ* ones (Sections 3.i and 3.ii). Given that the Southern method surveys a wider fraction of the genome, the only explanation is that it underestimates TE abundances. The fact that this bias increases with family copy number indicates that the level of resolution of the technique does not satisfy the requirements for this kind of study, and suggests that agarose gels become saturated with bands when highly abundant TE families are involved.

The similarity of the TE abundance estimates produced by the two techniques for the less numerous families suggests that most of the bands in the Southern blots correspond to euchromatic insertions, as proposed by Domínguez & Albornoz (1996). This interpretation is also supported by the evidence that most TE copies in the heterochromatin are very fragmented. Genomic clones from the β-heterochromatin of *D. melanogaster* (Dawid *et al.*, 1981; Vaury *et al.*, 1989; Maside *et al.*, in preparation), and also from some of the centric heterochromatin (Rutsov *et al.*, 1999), show that the heterochromatin generally contains series of short TE fragments, the product of the successive insertion of elements from different families into each other, generating a pattern that resembles that found in the intergenic region flanking the *adh1* gene in maize (SanMiguel *et al.*, 1998). This fragmentation would reduce the detection level of heterochromatic TEs in the Southern blots.

A second outcome from the comparison of both sources of data is the lack of correspondence between Southern and *in situ* estimates of the rates of element movement: (i) Southern estimates of the transposition rate are significantly lower than *in situ* ones, (ii)

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

excision rates are higher when estimated from Southern blots (*in situ* hybridization did not provide evidence of any excisions), and (iii) transposition and excision events are strongly associated when inferred from Southern data only (see Section 3.ii). Point (i) can be explained by the above-mentioned lack of resolution of the Southern blots, as a significant fraction of the new bands may remain undetected, especially when analysing the most abundant families. Points (ii) and (iii) can be explained by the fact that this technique is sensitive to any change in the length of the restriction fragments that harbour the element families under analysis, independently of its cause. Therefore, events such as the insertion of any TE into the relevant restriction fragments would be reflected in the Southern blot as a band shift, and hence spuriously counted as the simultaneous excision and insertion of one element.

The rate at which these apparent double events are expected in a Southern analysis should be greater than or equal to the product of the genomic transposition rate ($\sim 0{\cdot}1$; see Section 4.iii) $\times$ the rate of detection of events (35/131, from the *roo* data in Section 3.ii) $\times$ the probability that a TE inserts into a copy of the element family under study (roughly equivalent to the fraction of the genome made up by the family). The product of the first two factors is 0·027; the third can be estimated from the data as follows. In the Madrid background, the observed rate of double events can be estimated as the product of the pooled Southern transposition rate ($4{\cdot}69 \times 10^{-5}$; Table 2) $\times$ the proportion of transpositions that involve double events (30/207; from Appendix) $\times$ the genomic copy number of the set of families analysed (206; Table 2), which equals $1{\cdot}4 \times 10^{-3}$. Equating this with its expectation, the nine families studied here have to make up to approximately 5% ($1{\cdot}4 \times 10^{-3}/0{\cdot}027$) of the genome in order for TE insertions to account for the double events observed.

Even though we lack a direct estimate of the actual contribution to the genome from this particular set of families, several sources of evidence indicate that this value is close to the above expectation. If the approximately 50 TE families described in *D. melanogaster* (Finnegan, 1992*a*) were equally abundant, the set analysed here would represent around 20% of all TE-derived DNA. However, as some of them are from the most abundant ones (see next section), we expect their contribution to be somewhat higher: between 20% and 30%. Given that C_0t studies suggest that approximately 10% of the genome of *D. melanogaster* is made up of TEs (Manning *et al.*, 1975; Spradling & Rubin, 1981), the families included in this analysis could represent up to 3% of its genome, which implies that TE insertions into the studied families explain the vast majority of observed double events.

Other processes such as DNA loss (Petrov *et al.*, 1996, 2000), short indels and point mutations, may alter the distribution of the restriction sites in the element and/or flanking sequence, and could, thus, contribute to the observed band shifts (Dominguez & Albornoz, 1999). However, given the low rates at which these occur (between 0·25 and 3 times the per nucleotide mutation rate; Petrov *et al.*, 1996; Langley *et al.*, 2000), a simple calculation based on the expected number of restriction sites per band and the estimated per nucleotide mutation rate of $2{\cdot}2 \times 10^{-9}$ per generation (Keightley & Eyre-Walker, 2000) shows that they can cause only approximately 2×10^{-8} band shifts per element per generation, a value which is 2 orders of magnitude lower than that observed in our data ($6{\cdot}8 \times 10^{-6}$).

(ii) *Element abundances in the genome of* Drosophila melanogaster

Our *in situ* data yielded a pooled estimate of $32{\cdot}1 \pm 6{\cdot}1$ copies per family per haploid euchromatic genome for the present set of families. Considering that there are around 50 TE families in *D. melanogaster* (above), the total element abundance can be calculated as around 1500 TE copies per euchromatic genome. Assuming that the average size of an element is $5{\cdot}4 \pm 0{\cdot}31$ kb (averaged from the sizes of complete sequences of the TE families reported in Flybase), approximately 8 Mb of the 120 Mb that constitute the euchromatic genome of *D. melanogaster* (Adams *et al.*, 2000) correspond to TEs (7·5%). Therefore, if 18 Mb of the *D. melanogaster* genome corresponds to TEs (see above), the 60 Mb of heterochromatin (Adams *et al.*, 2000) must include the remaining TE-derived DNA (~ 10 Mb). This implies that the density of TE-derived DNA in the heterochromatin must be around 17%, almost 3 times higher than in the euchromatin. This estimate is in accord with an observed moderate enrichment of TEs in the non-euchromatic section of the genome, reported from a slot blot study in isogenic lines derived from the Beltsville population (Charlesworth *et al.*, 1994*a*), as well as with other experimental evidence that TEs tend to accumulate in the heterochromatin (Pimpinelli *et al.*, 1995; Junakovic *et al.*, 1998).

Further reinforcing this line of reasoning is the fact that the subset of families chosen for the present study may have led us to underestimate the true element abundance in the heterochromatin. Most of them are among the first-known TE families in *D. melanogaster* (Bingham & Zachar, 1989), probably because they are more active and, hence, more abundant in the euchromatin than other TE families (Nuzhdin *et al.*, 1996; Pasyukova *et al.*, 1998; Maside *et al.*, 2000). In fact, the mean family copy number estimated from data from a broader set of families is $18{\cdot}1 \pm 3{\cdot}2$

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

(Biémont & Cizeron, 1999; Table 1). If this value is used to extrapolate as above, the total euchromatic copy number in this species is approximately 900 elements; following the same reasoning, the proportion of TE-derived DNA in the heterochromatin would then be over 5 times higher than in the euchromatin.

(iii) *Rates of movement of transposable elements*

Our *in situ* results for the Madrid lines yielded a pooled rate of $1{\cdot}01 \times 10^{-4}$ transpositions per element per generation (Table 3), and an excision rate about 2 orders of magnitude lower ($3{\cdot}91 \times 10^{-6}$), in excellent agreement with previous data, particularly with the estimates produced by Maside *et al*. (2000) with a similar set of families in a different genetic background. The experiment of Maside *et al*. (2000) also differed in that only second chromosomes were involved, and they were maintained heterozygous in males, with no opportunity for crossing over. In the present experiments, all the chromosomes were followed, and transpositions could occur in both males and females.

Such good agreement is thus remarkable, especially considering the high levels of between-line variability reported for certain TE families such as *copia* (Nuzhdin *et al*., 1996) and *412* (see Section 3.ii), and suggests that the overall genomic transposition rate tends to have a roughly constant value. Taking into account our estimates of 900 copies per genome (see Section 4.ii), this can be approximated to 0·10 transpositions per generation, in good agreement with Maside *et al*. (2000) but substantially lower than the estimate of 0·20 reported in Nuzhdin & Mackay (1995).

Another important aspect of this new set of data is that, consistent with previous observations (Charlesworth & Langley, 1991; Nuzhdin *et al*., 1997; Maside *et al*., 2000), it suggests that the genomic excision rate is around 2 orders of magnitude lower than that of transposition. This has been interpreted as support for the selfish DNA hypothesis, which proposes that elements are maintained in the populations by an equilibrium between tranpositional spread, and one or more opposing deterministic forces (Charlesworth & Langley, 1991). However, the generality of this conclusion was strongly questioned by the results of Domínguez & Albornoz (1996), who reported unusually high rates of excision (of the same order as the rate of transposition) from a Southern analysis of the same Madrid lines as were used in the present work (Domínguez & Albornoz, 1996, tables 1 and 2). As discussed in Section 4.i, the comparison of Southern and *in situ* data from the same TE families on the same accumulation lines has provided us with enough evidence to question the reliability of Southern analysis as a method for estimating rates of element movement.

This leads us back to the selfish DNA hypothesis. It is generally accepted that selection plays a major role in controlling element spread in natural populations. The nature of the target of selection is, however, a matter of debate (Hoogland & Biémont, 1996; Biémont *et al*., 1997*a*; Charlesworth *et al*., 1997). The insertional mutation model proposes that selection acts against the deleterious effects of element insertions into nearby genes (Charlesworth & Charlesworth, 1983; Langley *et al*., 1983; Charlesworth, 1985). Alternatively, the ectopic exchange model proposes that, even when many insertions may cause small or negligible fitness effects due to insertional mutations, they are prevented from accumulation by the action of selection against the severe deleterious effects of chromosomal rearrangements produced by ectopic recombination between TEs inserted in non-homologous chromosomal locations (Montgomery *et al*., 1987; Langley *et al*., 1988; Charlesworth & Langley, 1991; Charlesworth *et al*., 1994*b*).

Results from theoretical studies suggest that, with TE abundances and rates of movement of the magnitudes reported here, equilibrium would only be possible in a scenario where the selection coefficients of many euchromatic insertions fit within a very narrow range of values, of the order of the transposition rate ($\sim 10^{-4}$), so that not all elements are eliminated from the population by selection. Otherwise, element insertions would be expected to build up rapidly in numbers at neutral or weakly selected sites (Charlesworth, 1991), in contradiction with the very few cases of fixation of an element insertion at a particular site, and the low element frequencies reported almost universally from natural population surveys (Charlesworth *et al*., 1992*a*; Maside *et al*., 2000). Under the insertional mutation model, this can only be explained if there are no neutral or nearly neutral sites in the whole euchromatin. This seems unlikely given the possibility of elements inserting into introns, intergenic regions or duplicated DNA fragments. In contrast, this criticism does not apply to the ectopic exchange model, which postulates that selection coefficients on insertions are a function of two variables: the frequency of the insertion in the population and the local rate of recombination (Langley *et al*., 1988).

(iv) *Element distributions within the genome*

If ectopic exchange is the main force controlling element spread, and the rate of exchange is related to the rate of meiotic recombination, an inverse correlation between local rates of meiotic recombination and element abundance is expected (Langley *et al*.,

https://doi.org/10.1017/S0016672301005201 Published online by Cambridge University Press

1988). Predicted outcomes of this model are (i) a deficit of elements on the X chromosome with respect to random expectations on the basis of the proportion of genome it represents, and (ii) higher element abundances in regions of reduced rate of recombination.

In agreement with these expectations, we have detected a slight deficit of elements on the X chromosome; in the Madrid population it contains 0·17 of the elements from the set of families analysed (data not shown). This value lies precisely between the two alternative versions of the ectopic exchange model: with ectopic exchange between element copies located throughout the genome, or only between elements within the same region (Langley *et al*., 1988). However, there is a high variation in this effect across families, and with the present sample size there is no power to reject the relevant alternative hypotheses of neutrality, or of the mutational effect of the insertions (under which the expected fractions of elements on the X would be 0·20 and 0·13, respectively; Langley *et al*., 1988). A similar situation has been found in other data sets available. Even though the deficit of elements on the X is commonly found for many TE families and genetic backgrounds, it is unclear whether or not it is a consequence of ectopic recombination (Montgomery *et al*., 1987; Langley *et al*., 1988; Charlesworth *et al*., 1992*b*; Biémont *et al*., 1994, 1997*a*; Hoogland & Biémont, 1996; Vieira & Biémont, 1996).

The second prediction of the ectopic exchange model is that elements should be found at higher frequencies/abundances in regions of the genome with a reduced rate of recombination (the tip and the base of the major chromosome arms, and the fourth chromosome). Our *in situ* data revealed a significant excess of elements in the proximal regions of the major arms, but no significant accumulation of elements at the tips of the chromosomes (mainly due to the results on the X and 2L), or on the fourth chromosome, was detected (Table 5). A similar pattern was found in a natural population survey (Charlesworth *et al*., 1992*b*), and the apparent lack of accumulation of elements at the tips of the chromosomes has been interpreted as evidence that ectopic exchange cannot be the only force checking the spread of elements (Charlesworth *et al*., 1992*b*; Hoogland & Biémont, 1996).

Despite the facts that the reduction in the rate of meiotic recombination generally affects the three most proximal chromosome sections of each autosomal arm, and that the reduction in that rate is of the same magnitude at least along the first two of them (starting from the centromere) (Charlesworth, 1996), the accumulation of elements in these regions is strongly correlated with the distance from the insertion to the recombining portion of the genome (the average densities of elements in the three most proximal divisions of the autosomes from the present *in situ* data are 0·045, 0·014 and 0·005, respectively; data not shown). It is also worth noting that the regions of reduced recombination at the tips of the autosomes are significantly shorter, both physically and in terms of their DNA content, than those at the bases ($1{\cdot}3 \pm 0{\cdot}2$ Mb vs $2{\cdot}3 \pm 0{\cdot}2$ Mb per region per chromosome arm, respectively; from Heino *et al*., 1994). These observations are compatible with a simple scenario in which the reduction in the rate of ectopic exchange is only effective if both elements involved are in a region of reduced recombination. That condition would not easily be met at the telomeres or on the fourth chromosome, which are of the same approximate length, providing an explanation as to why the accumulation of elements is more evident at the bases of the major chromosomes.

These data on the distribution of elements along the genome, along with other evidence for accumulation of elements in regions of low recombination, such as chromosomal inversions found at low frequency in natural populations (Eanes *et al*., 1992; Sniegowski & Charlesworth, 1994), strongly suggest that the distribution of elements is consistent with the ectopic exchange model.

We are indebted to Carlos López Fanjul and Aurora García Dorado for kindly sending us the Madrid mutation accumulation lines. We thank Patricia Pignatelli for her help with fly maintenance and *in situ* hybridization, and Helen Borthwick and Helen Cowan for media preparation and washing up. This work was funded by a BBSRC grant and grant DEB-9527717 from the US National Science Foundation to BC. X. M. was supported by a Marie Curie Fellowship.

Appendix

Profile of movements of elements from nine TE families detected by Southern analysis in the Madrid genetic background

Line	*copia*	*297*	*mdg1*	*jockey*	*412*	*opus*	*1731*	*doc*	*roo*	Pooled	
										Insertions	Excisions
5									+3	3	0
9								+1	+3, −1	4	1
11									0	0	0
13									+3	3	0
18		+2			+1, −1				+1, −1	4	2
21			+1	+2				+2, −1	+1	6	1
22									0	0	0
25				−1					+4	4	1
26									+3	3	0
31									+1	1	0
32									+1	1	0
35									+3	3	0
38									+3	3	0
44						+1, −1			+1	2	1
45					n.a.	n.a.			+2	2	0
52									+3	3	0
54		+1, −1							+5	6	1
59				+1, −1					+2	3	1
60			+1, −1						+1	2	1
62									0	0	0
63								+1	+3	4	0
66									+4	4	0
69			+1						+1, −1	2	1
70									+3, −1	3	1
71									+4	4	0
73				−1					+3	3	1
75									+1, −1	1	1
76									+3	3	0

Appendix (Continued)

Line	*copia*	*297*	*mdg1*	*jockey*	*412*	*opus*	*1731*	*doc*	*roo*	Pooled	
										Insertions	Excisions
77			+1						0	1	0
78									+1	1	0
79									+2	2	0
81									+4	4	0
82									+3	3	0
85									+4	4	0
87									+2	2	0
89	−1								+4	4	1
91									+3, −1	3	1
96									+1	1	0
97			+1			+1, −1			+2	4	1
100	+1								+4	5	0
101						+1			+2, −1	3	1
111									0	0	0
113									+4	4	0
114									+1	1	0
117									+1	1	0
119									+2, −1	2	1
124									+3	3	0
125			+1						+1	2	0
126							+1, −1		+3	4	1
127							+4	+6, 1	+4, −2	14	3
128			n.a.				n.a.	n.a.	n.a.		
129			n.a.				n.a.	n.a.	n.a.		
130			n.a.				n.a.	n.a.	n.a.		

Appendix (Continued)

Line	*copia*	*297*	*mdg1*	*jockey*	*412*	*opus*	*1731*	*doc*	*roo*	Pooled	
										Insertions	Excisions
136			n.a.				n.a.	n.a.	n.a.		
137						+1			+3	4	0
138									0	0	0
139									0	0	0
140									+1	1	0
141	+1, −1								+4	5	1
143									+6	6	0
144			+1						+3	4	0
148									+2, −1	2	1
149									0	0	0
158									+6	6	0
159	+1, 1	−1							n.a.	1	2
160									+1	1	0
161		+2, −1							n.a.	2	1
163									+1	1	0
166									0	0	0
167								+1	+1	2	0
169									0	0	0
171									+3	3	0
172									+2	2	0
175									+2	2	0
177									+2	2	0
181									+1	1	0
183									+3	3	0
187									+4	4	0
188									+1	1	0
190			+1						+2	3	0
191									+1	1	0
193									+1	1	0
197				+1, −2					+2, −1	3	3
199									+1	1	0
Total per family	+5, −3	+5, −4	+9, −3	+2, −3	0, 0	+4, −2	+5, −1	+11, −2	+166, −12	207	30

+ stands for band gain, − for band loss, and n.a. for data not available; blank cells indicate that no changes were detected.

References

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., *et al*. (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195.

Andolfatto, P., Wall, J. D & Kreitman, M. (1999). Unusual haplotype structure at the proximal breakpoint of *In*(2L)*t* in a natural population of *Drosophila melanogaster*. *Genetics* **153**, 1297–1311.

Berg, D. E. & Howe, M. M. (eds.) (1989) *Mobile DNA*. Washington, DC: American Society for Microbiology.

Biémont, C. & Cizeron, G. (1999). Distribution of transposable elements in *Drosophila* species. *Genetica* **105**, 43–62.

Biémont, C., Aouar, A. & Arnault, C. (1987). Genome reshuffling of the *copia* element in an inbred line of *Drosophila melanogaster*. *Nature* **329**, 742–744.

Biémont, 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.

Biémont, C., Tsitrone, A., Vieira, C. & Hoogland, C. (1997*a*). Transposable element distribution in *Drosophila*. *Genetics* **147**, 1997–1999.

Biémont, C., Vieira, C., Hoogland, C., Cizeron, G., Loevenbruck, C., Arnault, C. & Carante, J. P. (1997*b*). Maintenance of transposable element copy number in natural populations of *Drosophila melanogaster* and *D. simulans*. *Genetica* **100**, 161–166.

Bingham, P. M. & Zachar, Z. (1989). Retrotransposons and the FB transposon from *Drosophila melanogaster*. In *Mobile DNA* (ed. D. E. Berg & M. M. Howe), pp. 485–502. Washington, DC: American Society for Microbiology.

Brookfield, J. F. (1996). Models for the spread of non-autonomous selfish transposable elements when transposition and fitness are coupled. *Genetical Research* **67**, 199–210.

Caballero, A., Toro, M. A & López-Fanjul, C. (1991). The response to artificial selection from new mutations in *Drosophila melanogaster*. *Genetics* **128**, 89–102.

Cáceres, M., Ranz, J. M., Barbadilla, A., Long, M. & Ruíz, A. (1999). Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**, 415–418.

Capy, P. (ed.) (1997). *Evolution and Impact of Transposable Elements*. Dordrecht: Kluwer.

Charlesworth, B. (1985). The population genetics of transposable elements. In *Population Genetics and Molecular Evolution* (ed. T. Otha & K. Aoki), pp. 213–232. Berlin: Japan Sci. Soc. Press/Springer.

Charlesworth, B. (1991). Transposable elements in natural populations with a mixture of selected and neutral insertion sites. *Genetical Research* **57**, 127–134.

Charlesworth, B. (1996). Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genetical Research* **68**, 131–149.

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

Charlesworth, B. & Lapid, A. (1989). A study of ten families of transposable elements on X chromosomes from a population of *Drosophila melanogaster*. *Genetical Research* **54**, 113–125.

Charlesworth, B. & Langley, C. H. (1991). Population genetics of transposable elements in *Drosophila*. In *Evolution at the molecular level* (ed. R. K. Selander, A. G. Clark & T. S. Whittam), pp. 150–176. Sunderland, MA: Sinauer.

Charlesworth, B., Lapid, A. & Canada, D. (1992*a*). The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. I. Element frequencies and distribution. *Genetical Research* **60**, 103–114.

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

Charlesworth, B., Jarne, P. & Assimacopoulos, S. (1994*a*). The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. III. Element abundances in heterochromatin. *Genetical Research* **64**, 183–197.

Charlesworth, B., Sniegowski, P. D. & Stephan, W. (1994*b*). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**, 215–220.

Charlesworth, B., Langley, C. H. & Sniegowski, P. D. (1997). Transposable element distributions in *Drosophila*. *Genetics* **147**, 1993–1995.

Dawid, I. B., Long, E. O., DiNocera, P. P. & Pardue, M. L. (1981). Ribosomal insertion-like elements in *Drosophila melanogaster* are interspersed with mobile sequences. *Cell* **25**, 399–408.

Dimitri, P. (1997). Constitutive heterochromatin and transposable elements in *Drosophila melanogaster*. *Genetica* **100**, 85–93.

Domínguez, A. & Albornoz, J. (1996). Rates of movement of transposable elements in *Drosophila melanogaster*. *Molecular and General Genetics* **251**, 130–138.

Domínguez, A. & Albornoz, J. (1999). Structural instability of *297* element in *Drosophila melanogaster*. *Genetica* **105**, 239–248.

Doolitle, W. F. & Sapienza, D. (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**, 601–603.

Eanes, W. F., Wesley, C. & Charlesworth, B. (1992). Accumulation of P elements in minority inversions in natural populations of *Drosophila melanogaster*. *Genetical Research* **59**, 1–9.

Fernández, J. & López-Fanjul, C. (1996). Spontaneous mutational variances and covariances for fitness-related traits in *Drosophila melanogaster*. *Genetics* **143**, 829–837.

Finnegan, D. (1992*a*). Transposable elements. In *The Genome of* Drosophila melanogaster (ed. D. L. Lindsley & G. G. Zimm), pp. 1097–1107. San Diego: Academic Press.

Finnegan, D. J. (1992*b*). Transposable elements. *Current Opinion in Genetics and Development* **2**, 861–867.

Flavell, R. B. (1986). Repetitive DNA and chromosome evolution in plants. *Philosophical Transactions of the Royal Society of London, Series B* **312**, 227–242.

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

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

Junakovic, N., Terrinoni, A., Di Franco, C., Vieira, C. & Loevenbruck, C. (1998). Accumulation of transposable elements in the heterochromatin and on the Y chromosome of *Drosophila simulans* and *Drosophila melanogaster*. *Journal of Molecular Evolution* **46**, 661–668.

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

Keightley, P. D. & Eyre-Walker, A. (2000). Deleterious mutations and the evolution of sex. *Science* **290**, 331–333.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al*. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

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

Langley, C. H., Montgomery, E., Hudson, R., Kaplan, N. & Charlesworth, B. (1988). On the role of unequal exchange in the containment of transposable element copy number. *Genetical Research* **52**, 223–235.

Langley, C. H., Lazzaro, B. P., Phillips, W., Heikkinen, E. & Braverman, J. M. (2000). Linkage disequilibria and the site frequency spectra in the *su*(*s*) and *su*(*w*(*a*)) regions of the *Drosophila melanogaster* X chromosome. *Genetics* **156**, 1837–1852.

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.

Manning, J. E., Schmid, C. W. & Davidson, N. (1975). Interspersion of repetitive and nonrepetitive DNA sequences in the *Drosophila melanogaster* genome. *Cell* **4**, 141–155.

Maside, X., Assimacopoulos, S. & Charlesworth, B. (2000). Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster*. *Genetical Research* **75**, 275–284.

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

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

Nuzhdin, S. V. & Mackay, T. F. (1995). The genomic rate of transposable element movement in *Drosophila melanogaster*. *Molecular Biology and Evolution* **12**, 180–181.

Nuzhdin, S. V., Pasyukova, E. G. & Mackay, T. F. (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. (1997). Accumulation of transposable elements in laboratory lines of *Drosophila melanogaster*. *Genetica* **100**, 167–175.

Nuzhdin, S. V., Pasyukova, E. G., Morozova, E. A. & Flavell, A. J. (1998). Quantitative genetic analysis of *copia* retrotransposon activity in inbred *Drosophila melanogaster* lines. *Genetics* **150**, 755–766.

Orgel, L. E. & Crick, F. H. C. (1980). Selfish DNA: the ultimate parasite. *Nature* **284**, 604–607.

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., Nuzhdin, S. V. & Filatov, D. A. (1998). The relationship between the rate of transposition and transposable element copy number for *copia* and *Doc* retrotransposons of *Drosophila melanogaster*. *Genetical Research* **72**, 1–11.

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

Petrov, D. A., Sangster, T. A., Johnston, J. S., Hartl, D. L. & Shaw, K. L. (2000). Evidence for DNA loss as a determinant of genome size. *Science* **287**, 1060–1062.

Pimpinelli, S., Berloco, M., Fanti, L., Dimitri, P., Bonaccorsi, S., Marchetti, E., Caizzi, R., Caggese, C. & Gatti, M. (1995). Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. *Proceedings of the National Academy of Sciences of the USA* **92**, 3804–3808.

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.

Rutsov, N. B., Alekseenko, A. A., Beliaeva, E. S., Volkova, E. I., Kokoza, E. B., Makunin, I. V., Moshkin Iu., M., Shestopal, S. A. & Zhimulev, I. F. (1999). Microcloning and characteristics of DNA from regions of the centromeric heterochromatin of *Drosophila melanogaster* polytene chromosomes. *Genetika* **35**, 55–61.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (eds.) (1989) *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

SanMiguel, P., Gaut, B. S., Tikhonov, A., Nakajima, Y. & Bennetzen, J. L. (1998). The paleontology of intergene retrotransposons of maize. *Nature Genetics* **20**, 43–45.

Santiago, E., Albornoz, J., Domínguez, A., Toro, M. A. & López-Fanjul, C. (1992). The distribution of spontaneous mutations on quantitative traits and fitness in *Drosophila melanogaster*. *Genetics* **132**, 771–781.

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

Sokal, R. R. & Rohlf, F. J. (eds.) (1995). *Biometry*. New York: W. H. Freeman.

Spradling, A. C. & Rubin, G. M. (1981). *Drosophila* genome organization: conserved and dynamic aspects. *Annual Review of Genetics* **15**, 219–264.

Steinemann, M. & Steinemann, S. (1998). Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution. *Genetica* **102–103**, 409–420.

Vaury, C., Bucheton, A. & Pelisson, A. (1989). The beta heterochromatic sequences flanking the *I* elements are themselves defective transposable elements. *Chromosoma* **98**, 215–224.

Vieira, C. & Biémont, C. (1996). Selection against transposable elements in *D. simulans* and *D. melanogaster*. *Genetical Research* **68**, 9–15.

Vieira, C., Lepetit, D., Dumont, S. & Biémont, C. (1999). Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Molecular Biology and Evolution* **16**, 1251–1255.