Genet. Res., Camb. (1987), **49**, *pp.* 121–128 *With 1 text-figure Printed in Great Britain*

Transposition of the I element and *copia* in a natural population of *Drosophila melanogaster*

ANDREW J. LEIGH BROWN* AND JULIE E. MOSS
Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, Scotland

(*Received 19 September 1986 and in revised form 17 November 1986*)

Summary

In order to increase our understanding of the evolutionary dynamics of transposable genetic elements we have studied the chromosal location of copies of 2 element families in 20 X chromosomes extracted from a natural population of *Drosophila melanogaster* from Spain. The I element was localized at a total of 64 chromosomal sites and *copia* at 45 sites in this sample with a mean copy number of 3·2 and 2·3 elements/chromosome respectively. Both elements were highly variable in location, with no site reaching a higher frequency than 4/20 in either case. Comparisons with other data sets suggest that insertion frequencies can be used to detect population structuring.

1. Introduction

The discovery that dispersed, moderately repetitive DNA sequences in *Drosophila* are mobile (Ilyin *et al.* 1978; Potter *et al.* 1979) generated many new evolutionary questions. More recent results indicate that these questions will not be restricted to *Drosophila* but will be of general relevance. There is now strong evidence for an evolutionary relationship between the Ty elements of yeast, *copia*-like elements in *Drosophila* and vertebrate retroviruses. All appear to replicate through an RNA intermediate and make viruses or virus-like particles with which element RNA is associated (Shiba & Saigo, 1983; Boeke *et al.* 1985; Garfinkel, Boeke & Fink, 1985). DNA sequence studies have shown that significant homologies exist at the amino acid level between *copia*, Ty and several vertebrate retroviruses (Mount & Rubin, 1985; Emori *et al.* 1985; Clare & Farabaugh, 1985). All these elements can cause mutations; the majority of the classical spontaneous 'point' mutations in *Drosophila* that have been studied at the molecular level have been shown to be due to transposable element insertions (Rubin, 1983). Ty elements in yeast have generated mutations at several loci (Roeder & Fink, 1983). More recently, endogenous retroviral sequences have been shown to cause mutations at the dilute and agouti loci in the mouse (Jenkins *et al.* 1981; Copeland, Jenkins & Lee, 1983; Copeland, Hutchinson & Jenkins, 1983).

Mobility of such elements is by no means always associated with a detectable mutation, as the first studies in *Drosophila* strains showed. Polymorphism in natural populations due to the insertion of repetitive DNA sequences near cloned genes has been found in Drosophila at the *Adh* and 87A heat shock genes (Langley, Quattlebaum & Montgomery, 1982; Aquadro *et al.* 1986; Leigh Brown, 1983) and recently in the rat (Economou-Pachnis *et al.* 1985).

The distribution of mobile element families among related host species has been studied by more than one group recently (Martin, Wiernasz & Schedl, 1983; Brookfield, Montgomery & Langley, 1984). In general the probability of detecting DNA homology to an element cloned from *D. melanogaster* in genomic DNA of another species increases, the closer the phylogenetic relationship between the two species. Amongst well-characterized elements, only the P element was found exclusively in *D. melanogaster*, although recently it has been described in *D. paulistorum* (Daniels *et al.* 1984); and homology has been detected in other distant species (Lansman *et al.* 1985). Euchromatic copies of the I element are found in *D. simulans* and *D. mauritiana* despite their absence from some strains of *D. melanogaster* (Bucheton *et al.* 1986, J.E.M. and A.J.L.B. unpublished results) and more than one *copia*-like element from *D. melanogaster* has been detected in DNA from *D. pseudoobscura* (Brookfield, Montgomery & Langley, 1984; A.J.L.B., unpublished results) indicating persistence in the genome of both species for as much as 50×10^6 years. We are clearly dealing with an ancient component of the genome. Although the possibility of horizontal transfer of elements between species has been raised

* Corresponding author.

by some studies (Lansman *et al.* 1985), in general the distribution of element families suggests that vertical transmission through the germ line is the rule.

At first sight, the persistence of families of mobile elements within the genome for long periods of evolutionary time presents us with a paradox when their characteristic short-term instability in location and numbers is considered (Strobel, Dunsmuir & Rubin, 1979). This question has been addressed in theoretical terms by Brookfield (1982), Langley, Brookfield & Kaplan (1983), Charlesworth & Charlesworth (1983) and Charlesworth (1985). Brookfield (1982) and Charlesworth & Charlesworth (1983) examined the case where increasing element copy number decreases host fitness. They showed that a copy-number distribution that was stable over evolutionary time would not be generated under a multiplicative fitness function, such that each element acts independently on fitness, although Charlesworth & Charlesworth (1983) showed that stability was possible if there were synergistic interactions between elements in their effects on fitness. On the other hand Langley, Brookfield and Kaplan (1983) and Charlesworth & Charlesworth (1983) showed under a neutral model that it was possible to generate a stable copy-number distribution if the elements regulated their own copy number, an aspect which has been explored in greater detail by Charlesworth & Langley (1986). In the model of Charlesworth & Charlesworth (1983) the spectrum of element frequencies at particular chromosomal sites is determined by two parameters: α, the probability of transposition to a new site relative to the total number of sites available (if the total number of occupable sites is assumed to be infinite as in the model of Langley, Brookfield & Kaplan, $\alpha = 0$); and $\beta = 4N_e v$, where v is the probability of excision of an element from a site and N_e is the effective population size. If $\alpha = 0$ then β determines the frequency spectrum. It is equivalent to the single parameter, θ, of Langley, Brookfield & Kaplan (1983).

Until recently only one study had been published which presented the gene frequencies of elements at specific chromosomal sites in natural populations (Montgomery & Langley, 1983). We wished to address two questions raised by that study: first, was the high level of heterozygosity general for all elements, and secondly, was it general for all populations? We present another study on a sample of 20X chromosomes from a Spanish population. Using a biotin-avidin detection system (Langer-Safer, Levine & Ward, 1982), we have determined the chromosomal distribution of sequences homologous to *copia* (Rubin, Finnegan & Hogness, 1976; Finnegan *et al.* 1978) and the I element (Bucheton *et al.* 1986). *Copia* was included in the earlier study made on a North Carolina population (Montgomery & Langley, 1983). The I element is responsible for the I–R system of hybrid dysgenesis (see Bregliano & Kidwell, 1983, for a review of hybrid dysgenesis). Polymorphism in the locations of this element has been studied by Biemont (1986) in a laboratory population established by 50 wild females. We have analysed our data and those of Biemont under a neutral model and we present the predicted frquency spectra for comparison with the data sets.

2. Materials and methods

The population sample was collected by Dr J. S. Jones at Zahara de las Atunas near Cape Trafalgar, south Spain. Single male offspring of wild-inseminated females were mated to virgin females from an attached-X stock: *C(1)DX yf/sn*^{w}*/Y*; Π_2 (Engels & Preston, 1981). This stock was obtained from Dr W. R. Engels and has several P and I elements on its chromosome arms. The induction of transposition by dysgenesis was therefore inhibited. Stocks were maintained at 25 ° on Lewis medium (cornmeal – sucrose – agar).

Male third-instar larvae were dissected, and their salivary glands squashed, using the adaptations developed by Dr J. Lim (given in Shrimpton & Langley, 1986) of the procedure of Pardue & Gall (1975). Squashes were hybridized in 20 μl 2 × SSC containing denaturated herring sperm DNA at 40 μg/ml, 10% w/v sodium dextran sulphate (Pharmacia), 50% formamide (Fluka) and 20 ng probe DNA labelled with bio-11-dUTP by nick translation (see below).

Plasmid DNAs pI407 (Bucheton *et al.* 1984) and pI901, given by Drs D. Finnegan and J. Prosser, were used to detect I-element homology. pC7 (Leigh Brown, unpublished), a copy of a genomic *copia* cloned in pUC8 (Viera & Messing, 1982) was used to detect homology to *copia*. Probe DNA was prepared by the alkaline lysis method as described by Maniatis, Fritsch & Sambrook (1982). DNA was labelled with bio-11-dUTP (BRL) by nick translation. 5 μl of a 0·4 mM solution of biodUTP was used in a reaction volume of 100 μl with up to 100 units of *E. coli* Polymerase I (Anglian Biotechnology) to label 1 μg DNA. dGTP, dATP and dCTP were added to a concentration of 25 μM, and 10 μCi of ^{32}PdCTP were included to follow the incorporation. Reactions which incorporated more than 30% of the radioactivity were used.

After hybridization biotin-labelled regions were visualized with the avidin–peroxidase ABC system (Vectastain kit pK4000, Vector Labs, Hsu, Raine & Fanger (1981)) in a scaled-down reaction. Then 1 μl each of components A and B were mixed in 125 μl of 50 mM Tris pH 7·6, 4% BSA and left to stand for 10 min; 20 μl of this solution was applied to each slide under a coverslip and incubated at 37 ° for 30 min. After washing three times in Phosphate-buffered saline (PBS) peroxidase activity was stained for with 0·06% v/v solution of hydrogen peroxide and 0·5 mg/ml diaminobenzidine (Sigma) in 50 mM Tris pH 7·6 (20 μl

Table 1. *The distribution of the I element and* copia *in 20 X chromosomes from a Spanish population*

X chromosome division labelled...	1B1	1C	2E	3B	3C8	4A	4B	4C1-4	4D	4F1	5A1-6	5C1	5D1-3	5D4-6	6A	6F1	7A1-5	7B1	7C1	7E1	8A
I element homology detected in	1	—	9	—	6	5	5	10	—	2	11	19	4	11	2	—	5	13	13	1	8
line number	6	—	—	—	—	16	7	—	—	—	—	—	—	—	—	—	—	17	—	11	—
Copia homology detected in	—	19	3	8	—	14	6	—	12	12	—	—	—	—	—	19	—	7	5	—	—
line number	—	—	—	—	—	—	7	—	16	14	—	—	—	—	—	—	—	13	—	—	—
	—	—	—	—	—	—	14	—	—	—	—	—	—	—	—	—	—	17	—	—	—
	—	—	—	—	—	—	16	—	—	—	—	—	—	—	—	—	—	—	—	—	—

X chromosome division labelled...	8B1-4	8C1	8E1-3	8F1	9A	9B1	9D	10A1	10D1	11A1	11A6-9	11C	11D	11E1-4	12A1	12E1	12E8	12F1	13A1	13A5	13B1-4
I element homology detected in	—	14	9	8	—	—	20	13	6	—	6	—	4	11	3	4	9	16	—	2	—
line number	—	—	16	—	—	—	—	—	18	—	—	—	—	—	11	—	20	—	—	3	—
	—	—	—	—	—	—	—	—	—	—	—	—	—	—	16	—	—	—	—	—	—
Copia homology detected in	14	—	10	—	14	11	—	16	—	1	15	17	12	—	11	8	3	19	2	3	2
line number	—	—	—	—	—	—	—	—	—	—	—	—	13	—	20	—	19	—	10	—	15

X chromosome division labelled...	13C5	13F1	14B10	15B1	15D1	15F1	16A	16B1	17A1-2	17A5-6	17B1	17D	18C1-4	19A	19C	19E1	19E8	19F1
I element homology detected in	13	17	9	9	3	7	—	3	—	4	9	12	4	13	12	4	16	—
line number	—	—	—	—	11	—	—	—	—	16	—	—	18	—	—	5	—	—
	—	—	—	—	17	—	—	—	—	—	—	—	—	—	—	11	—	—
	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	14	—	—
Copia homology detected in	—	—	—	—	—	—	11	—	7	—	—	—	—	4	—	16	—	3
line number	—	—	—	—	—	—	—	—	—	—	—	—	—	16	—	—	—	—
	—	—	—	—	—	—	—	—	—	—	—	—	—	18	—	—	—	—

The line number of each of the 20X chromosomes lines scored for each element are given under the chromosome divisions where hybridization was detected with pI 407 (I element) or pC 7 (*copia*). Line number 19 had no *copia* element on its X chromosome.

total volume). The staining reaction was carried out at 37 °C for up to 1½ h.

Slides were washed in PBS and chromosomes stained in 3% Giemsa (Gurr, BDH) before air drying.

3. Results

The chromosomal sites at which I element and *copia* DNA were detected are given in Table 1 for all 20 chromosomes for the Zahara population. In most cases the hybridization at each site was localized to a single band on the Bridges map of the polytene chromosomes (Lefevre, 1976). The I element probe used for this study, pI407 (Bucheton *et al.* 1984) included about 800 bp of single copy sequence from the cloning site in the white locus, 3C1-3. The I-element homology is about 5·4 kb in length. Strongly labelled slides accordingly labelled the cloning site as well. This was established by use of the subclone pI 901, from which all genomic sequences have been removed (J. Prosser and D. Finnegan, personal communication).

The total number of sites detected and mean copy number per chromosome for each element are given in Table 2. The distribution of copy number per chromosome did not in either case differ significantly from Poisson expectation. Here we also present the values for θ (see Introduction), estimated from the Zahara data and from the data of Biemont (1986) on the I element, according to the procedure of Kaplan & Brookfield (1983).

The frequency spectrum, or occupancy profile, observed for the Zahara sample of chromosomes is shown in Fig. 1(*a*) for the I element and 1(*b*) for *copia*. In general the frequency at which any site was occupied was very low. The maximum frequency reached for either element was 4/20 (0·2) for sites 19E1 (I element) and 4B (*copia*).

We have obtained predicted frequency spectra for the Zahara data following the procedure of Charlesworth & Charlesworth (1983). These are given in Table 3 along with the observed values. Two cases are presented for each element. The first case ($\alpha = 0$) gives the spectrum under the assumption of an infinite number of occupable sites (Langley, Brookfield & Kaplan, 1983). The second case presents the profiles obtained from the joint estimation of α and β. However, only for the last data set – that obtained by Biemont (1986) on the I element – are there sufficient degrees of freedom to test the fit. For case (*a*) we obtain $\chi^2_5 = 11{\cdot}83$ while for case (*b*) $\chi^2_5 = 2{\cdot}99$. Thus joint estimation of the two parameters seems to give a noticeable improvement to the fit in this case.

Estimation of α and β values jointly allows the total number of occupable sites to be estimated (Charlesworth & Charlesworth, 1983). The estimates for the three data sets are: I element Zahara 159·5, *copia* Zahara 102·9 I element Biemont 260·3.

Table 2. *Number of elements and sites found in 20 X chromosome lines and estimates of* θ

	I element	*copia*
Zahara population		
Total number of elements in 20 chromosomes	64	45
Total number of sites	46	31
Mean number of elements per X chromosome (copy number)	3·2	2·3
θ	21·5	16·9
Data of Biemont (1986)		
Total number of elements in 14 genomes	422	
Mean number of elements per chromosome arm	3·2	
θ, X-linked sites	7·4	
θ, autosomal sites	4·6	
θ pooled	5·1	

θ values were estimated by the procedure of Kaplan & Brookfield (1983) (see text). The data from the study by Biemont (1986) on the I element were taken from his figure 3, ignoring four pooled sites. The 29 X-linked and 122 autosomal sites remaining were analysed separately. It was assumed that equal numbers of male and female larvae were used and the sample sizes were therefore taken as 28 for autosomal loci and 21 for X-linked sites. It was necessary also to make a correction to allow for the possibility of homozygosity at higher frequencies of occupation. Hardy–Weinberg equilibrium was assumed to make this correction. The total number of elements was thus corrected up to 59·5 from 58 for X-linked sites and to 363 from 330 for autosomal sites.

The X chromosome represents a single chromosome arm and thus the mean copy-number values given for the Zahara data and those of Biemont are directly comparable.

https://doi.org/10.1017/S0016672300026914 Published online by Cambridge University Press

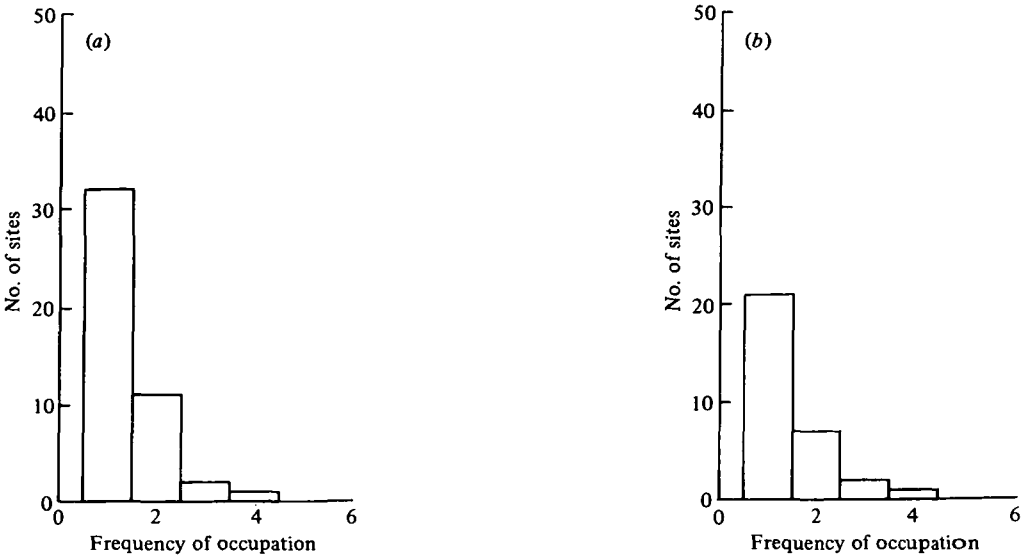

Fig. 1. Frequency spectra for sites occupied by transposable elements in a Spanish population. (*a*) Frequencies of sites occupied by the I element. (*b*) Frequencies of sites occupied by *copia*.

Table 3. *Frequency spectra*

(i) Zahara population, sample size 20 X chromosomes Frequency of occupation	1	2	3	4	5
I element					
Number of sites	32	11	2	1	0
Predicted profiles					
(*a*) $\alpha = 0, \beta = 21{\cdot}5$	33·97	8·17	2·5	0·86	0·3
(*b*) $\alpha = 0{\cdot}87, \beta = 42{\cdot}5$	32·3	9·5	2·7	0·8	0·2
copia					
Number of sites	21	7	2	1	0
Predicted profiles					
(*a*) $\alpha = 0, \beta = 16{\cdot}9$	21·18	5·76	2·04	0·79	0·32
(*b*) $\alpha = 0{\cdot}8, \beta = 35$	20·8	6·7	2·2	0·7	0·2

(ii) I element data of Biemont (1986), sample size 14 individuals Frequency of occupation	1	2	3	4	5	6	7	8	9	10	11	12
Number of sites	60	33	24	10	12	4	3	2	2	0	1	0
Predicted profiles												
(*a*) $\alpha = 0, \beta = 5{\cdot}1$	61·6	26·7	15·4	9·9	6·8	4·8	3·5	2·5	1·9	1·4	1·0	0·7
(*b*) $\alpha = 0{\cdot}72, \beta = 12{\cdot}8$	56·9	34·1	21·3	13·4	8·5	5·4	3·3	2·1	1·3	0·8	0·4	0·04

The procedure of Charlesworth & Charlesworth (1983) was used for fitting the spectra. In each case in (*a*) β was taken as the value of θ given in Table 2. The values for α and β given in (*b*) were the best fit obtained when both were fitted jointly.

The first two refer to a single X chromosome, while the last refers to the whole haploid genome. For comparison the Zahara I-element estimate should be multiplied by a factor of about 5 and would become approximately 800.

4. Discussion

The main motivation for our study of the mobility of transposable elements in natural populations was to assess their mutagenic impact on the host genome. We have studied two elements of very different structure so that our conclusions might be more general. Both showed a high level of variability in their chromosomal location in the Zahara population. For *copia* this is in accord with earlier work (Strobel, Dunsmuir & Rubin, 1979; Montgomery & Langley, 1983). The I element has been studied by Biemont (1986) in a wild-derived laboratory population. Mobilization of the I element is known to be induced by crosses between strains which possess and those which lack euchromatic copies of the full length (5·4 kb) I elements (Bucheton

https://doi.org/10.1017/S0016672300026914 Published online by Cambridge University Press

et al. 1984). Our results and those of Biemont (1986) clearly show that wild-derived I-strains originating from the same population are highly variable in the number and locations of the I element. This suggests that the I element does not require the I–R incompatability in order to be able to transpose. Observations which suggest that the same might apply to the P element have been made by comparisons of lines derived from a single parent stock (Preston & Engels, 1984).

The quantitative analysis of our data has shown a good fit to a neutral model of transposition (Table 3). The results for the different elements can be compared in terms of the estimates of θ presented in Table 2. Our estimates of 21·5 for the I element and 16·9 for *copia* are similar in the Zahara sample in contrast to the estimates obtained by Brookfield & Kaplan (1983) on the data of Montgomery & Langley (1983). The three *copia*-like elements they studied yielded θ estimates of 17 (297), 35 (412) and 48 (*copia*) in a sample of 20 X chromosomes from a North Carolina population. The value of θ estimated in this way is strongly dependent on the mean copy number in the sample. The difference between the two estimates for *copia* of 16·9 in Zahara and 48 in North Carolina can be interpreted in this light, as the mean copy number per chromosome were 2·3 and 1·6 respectively.

Caution is required in the interpretation of the gene frequency data. Brookfield & Kaplan (1983) showed that while the North Carolina data appeared to fit their model, it could also be approximated by an alternative hypothesis. Under this no site is in fact occupied more than once, and the apparent multiple occupation arises from the greatly reduced number of sites which can be distinguished by the cytological method adopted. Equally this conclusion could not be ruled out for the Zahara data, so the possibility remains that there is no frequency greater than $\frac{1}{20}$ for either element in the Zahara sample.

Some other results which are relevant to this question have been obtained by mapping insertions of transposable elements in genomic DNA of individuals from natural populations (Leigh Brown, 1983); Aquadro *et al.* 1986; Beech and Leigh Brown, in preparation). In all cases analysed to date, all insertions were found at frequencies of $1/n$ where n is the sample size. In the one instance where this appeared not to hold, different insertions involving the same element occupied very closely adjacent sites and DNA sequencing was needed to demonstrate the differences (Aquadro *et al.* 1986).

The data of Biemont (1986) on the I element provide an interesting exception to the general picture. In his data the frequencies of some occupied sites are high. The overall estimate of θ for his data, 5·1, is much lower than has yet been found for any element in a natural population. In this case the mean number of elements per chromosome arm is the same as in the Zahara data. Restricting the comparison to the X-linked sites in his data by no means removes the anomaly (Table 2). As our Zahara results indicate that this is not a general property of the I element and the data Biemont obtained for mdg- 1 showed similar features, it appears that population history might be responsible. The population that Biemont studied was laboratory culture established from 50 wild females and maintained for 18 months (*ca.* 35 generations) as a mass culture prior to sampling. This restriction of population size has had the effect of increasing the variance in allele frequencies, which is almost 5-fold greater for his X-linked sites than for those in our Zahara population sample. Thus we argue that a significant level of inbreeding, probably arising from the period of laboratory culture, has been revealed by this analysis of the frequency spectrum of chromosomal sites occupied by the I element. It had previously been argued that this might be a generally useful approach to the analysis of population structure (A. Robertson, personal communication) but the data from natural populations of *D. melanogaster* did not bear this out. However, structuring may be virtually absent in these populations, which are characterized by very large effective sizes (Mukai & Yamaguchi, 1974). As this is by no means a feature of all *Drosophila* species (Wright, Dobzhanski & Hovanitz, 1942), it remains possible that this approach will be of use in the analysis of the structure of natural populations for some other species.

We are grateful to Dr D. Finnegan for providing clones containing I element sequences and, together with Dr J. Brookfield, Dr B. Charlesworth and Robin Beech, for discussion and criticism of the manuscript. We also thank Dr J. S. Jones for collecting the wild flies and Dr W. R. Engels for the C(1)DX; Π_2 stock. This project was supported by the Medical Research Council.

References

Aquadro, C. F., Deese, S. F., Bland, M. M., Langley, C. H. & Laurie-Ahlberg, C. C. (1986). Molecular population genetics of the alcohol dehydrogenase gene region of *Drosophila melanogaster*. *Genetics* (In the Press).

Biemont, C. (1986). Polymorphism of the mdg-1 and I mobile elements in *Drosophila melanogaster*. *Chromosoma* **93**, 393–397.

Boeke, J. D., Garfinkel, D. J., Styles, C. A. & Fink, G. R. (1985). Ty elements transpose through an RNA intermediate. *Cell* **40**, 491–500.

Bregliano, J. C. & Kidwell, M. G. (1983). Hybrid dysgenesis determinants. In *Mobile Genetic Elements* (ed. J. A. Shapiro), pp. 363–410. New York: Academic Press.

Brookfield, J. F. Y. (1982). Interspersed repetitive DNA sequences are unlikely to be parasitic. *Journal of Theoretical Biology* **94**, 281–299.

Brookfield, J. F. Y., Montgomery, E. & Langley, C. H. (1984). Apparent absence of transposable elements related to the P elements of *D. melanogaster* in other species of *Drosophila*. *Nature* **310**, 330–331.

Bucheton, A., Paro, R., Sang, H. M, Pelisson, A & Finnegan, D. J. (1984). The molecular basis of I–R hybrid dysgenesis: identification, cloning and properties of the I factor. *Cell* **38**, 153–163.

Bucheton, A., Simonelig, M., Vaury, C. & Crozatier, M. (1986). Sequences similar to the I transposable element involved in I–R hybrid dysgenesis in *D. melanogaster* occur in other *Drosophila* species. *Nature* **322**, 650–652.

Charlesworth, B. (1985). The population genetics of transposable elements. In *Population Genetics and Molecular Evolution* (ed. T. Ohta and K. Aoki) pp. 213–232. New York: Springer Verlag.

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. *Genetics* **112**, 359–383.

Clare, J. & Farabaugh, P. (1985). Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. *Proceedings of the National Academy of Sciences, USA* **82**, 2829–2833.

Copeland, N. G., Jenkins, N. A. & Lee, B. K. (1983). Association of the lethal yellow (A^y) coat color mutation with an ecotropic murine leukaemia virus genome. *Proceedings of the National Academy of Sciences, USA* **80**, 247–249.

Daniels, S. B., Strausbaugh, L. D., Ehrman, L. & Armstrong, R. (1984). Sequences homologous to P elements occur in *Drosophila paulistorum*. *Proceedings of the National Academy of Sciences, USA* **81**, 6794–6797.

Economou-Pachnis, A., Lohse, M. A. Furano, A. V. & Tsichlis, P. N. (1985). Insertion of long interspersed repeat elements at the *Igh* (immunoglobin heavy chain) and Mlvi-2 (Moloney leukemia virus integration 2) loci of rats. *Proceedings of the National Academy of Sciences, USA* **82**, 2857–2861.

Emori, Y., Shiba, T., Kanaya, S., Inouye, S., Yuki, S. & Saigo, K. (1985). The nucleotide sequences of *copia* and *copia*-related RNA in *Drosophila* virus-like particles. *Nature* **315**, 773–776.

Engels, W. R. & Preston, C. R. (1981). Identifying P factors in Drosophila by means of chromosome breakage hot spots. *Cell* **26**, 421–428.

Finnegan, D. J., Rubin, G. M., Young, M. W & Hogness, D. S. (1978). Repeated gene families in *Drosophila melanogaster*. *Cold Spring Harbor Symposia on Quantitive Biology* **42**, 1053–1063.

Garfinkel, D. J., Boeke J. D. & Fink, G. R. (1985). Ty element transposition: reverse transcriptase and virus-like particles. *Cell* **42**, 507–517.

Hsu, S. M., Raine, L. & Fanger, H. (1981). The use of avidin–biotin–peroxidase complex (ABC) in immunoperoxidase techniques: comparison between ABC and unlabelled antibody (PAP) procedures. *Journal of Histochemistry & Cytochemistry* **29**, 577–581.

Ilyin, Y. V., Tchurikov, N. A., Ananiev, E. V., Ryskov, A. P., Yenikolopov, G. N., Limborska, S. A, Maleeva, N. E., Gvozdev, V. A. & Georgiev, G. P. (1978). Studies on the DNA fragments of mammals and *Drosophila* containing structural genes and adjacent sequences. *Cold Spring Harbor Symposia on Quantitative Biology* **42**, 959–969.

Jenkins, N. A., Copeland, N. G., Taylor, B. A. & Lee, B. K. (1981). Dilute (*d*) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. *Nature* **293**, 370–374.

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

Langer-Safer, P. R., Levine, M. & Ward, D. C. (1982). Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proceedings of the National Academy of Sciences, USA* **79**, 4381–4385.

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

Langley, C. H., Montgomery, E. & Quattlebaum, W. F. (1982). Restriction map variation in the *Adh* region of *Drosophila*. *Proceedings of the National Academy of Sciences, USA* **79**, 5631–5635.

Lansman, R. A., Stavey, S. M., Grigliatti, T. A. & Brock, H. W. (1985). Sequences homologous to the P mobile element of *Drosophila melanogaster* are widely distributed in the subgenus Sophophora. *Nature* **318**, 561–563.

Lefevre, G. Jr (1976). The polytene chromosomes of *Drosophila melanogaster*. In *The Genetics and Biology of Drosophila*, vol. 1a (ed. M. Ashburner and E. Novitski). New York: Academic Press.

Leigh Brown, A. J. (1983). Variation at the 87A heat shock locus in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* **80**, 5350–5354.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular cloning, a Laboratory Manual*. New York: Cold Spring Harbor.

Martin, G., Wiernasz, D. & Schedl, P. (1983). Evolution of *Drosophila* repetitive-dispersed DNA. *Journal of Molecular Evolution*, **19**, 203–213.

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

Mount, S. M. & Rubin, G. M. (1985). Complete nucleotide sequence of the *Drosophila* transposable element *copia*. Homology between *copia* and retroviral proteins. *Molecular & Cellular Biology* **5**, 1630–1638.

Mukai, T. & Yamaguchi, O. (1974). The genetic structure of natural populations of *Drosophila melanogaster*. XI. Genetic variability in a natural population. *Genetics* **76**, 339–366.

Ohta, T. (1985). A model of duplicative transposition and gene conversion. *Genetics* **110**, 513–524.

Pardue, M. L. & Gall, J. G. (1975). Nucleic acid hybridisation to the DNA of cytological preparations. *Methods in Cell Biology* **10** 1–17.

Potter, S. S., Brorein, W. J. Jr., Dunsmuir, P. & Rubin, G. M. (1979). Transposition of elements of the 412, *copia*, and 297 dispersed repeated gene families in Drosophila. *Cell* **17**, 415–427.

Preston, C. R. & Engels, W. R. (1984). Movement of P elements within a P strain. *Drosophila Information Service* **60**, 169–170.

Roeder, G. S. & Fink, G. R. (1983). Transposable elements in yeast. In *Mobile Genetic Elements* (ed. J. A. Shapiro), pp. 299-328. New York: Academic Press.

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

Rubin, G. M., Finnegan D. J. & Hogness, D. S. (1976). The chromosomal arrangement of coding sequences in a family of repeated genes. In *Progress in Nucleic Acid Research*, vol. 19 (ed W. Cohn and E. Volkin), pp. 221–226. New York: Academic Press.

Shiba, T. & Saigo, K. (1983). Retrovirus-like particles containing RNA homology to the transposable element *copia* in *Drosophila melanogaster*. *Nature* **302**, 119–124.

Shrimpton, A. E. & Langley, C. H. (1986). Association of a transposable DNA sequence with *Om* mutations in *Drosophila ananassae*. *Genetics* **114**, 125–135.

Strobel, E., Dunsmuir, P. & Rubin, G. M. (1979). Polymorphisms in the chromosomal locations of elements of the *412*, *copia* and *297* dispersed repeated gene families in *Drosophila*. *Cell* **17**, 429–439.

Vieira, J. & Messing, J. (1982) The pUC plasmids, M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.

Wright, S., Dobzhansky, Th. & Hovanitz W. (1942). Genetics of natural populations. VII. The allelism of lethals in the third chromosome of *Drosophila pseudo-obscura*. *Genetics* **27**, 363–394.