Genet. Res., Camb. (1988), **52**, *pp.* 223–235 *Printed in Great Britain*

On the role of unequal exchange in the containment of transposable element copy number

CHARLES H. LANGLEY[1]*, ELIZABETH MONTGOMERY[1], RICHARD HUDSON[2], NORMAN KAPLAN[2], AND BRIAN CHARLESWORTH[3]

[1] *Laboratory of Molecular Genetics and* [2] *DBRA, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709*

[3] *Department of Ecology and Evolution, The University of Chicago, 1103 E 57th Street, Chicago, Illinois 60637*

(*Received 12 January 1988 and in revised form 12 April 1988*)

Summary

A population genetics model of the role of asymmetric pairing and unequal exchange in the stabilization of transposable element copy number in natural populations is proposed and analysed. Monte Carlo simulations indicate that the approximations incorporated into the analysis are robust in the relevant parameter ranges. Given several simple assumptions concerning transposition and excision, equal and unequal exchange, and chromosome structure, predictions of the relative numbers of transposable elements in various regions of the *Drosophila melanogaster* genome are compared to the observed distribution of *roo*/*B104* elements across chromosomal regions with differing rates of exchange, and between *X* chromosomes and autosomes. There is no indication of an accumulation of elements in the distal regions of chromosomes, which is expected if unequal exchange is reduced concomitantly with normal crossing over in the distal regions. There is, however, an indication of an excess of elements relative to physical length in the proximal regions of the chromosomes, which also have restricted crossing over. This observation is qualitatively consistent with the model's predictions. The observed distribution of elements between the mid-sections of the *X* chromosomes and autosomes is consistent with the predictions of one of two models of unequal exchange.

1. Introduction

Many of the dispersed repetitive DNA sequences of *Drosophila* are known to be transposable genetic elements (Rubin, 1983). A large proportion of the spontaneous mutations recovered and routinely utilized in the laboratory are apparently caused by the insertion of these elements into or near transcriptional units (Finnegan, 1985). Comparisons of related species of *Drosophila* indicate that the majority of transposable elements are ancient parasitic occupants of the genome (Brookfield, Montgomery & Langley, 1984). The distribution and inferred dynamics of transposable elements in natural populations have been studied at both cytogenetic (Montgomery & Langley, 1983; Leigh Brown & Moss, 1987; Charlesworth, 1988) and DNA levels (Langley, Montgomery & Quattlebaum, 1982; Leigh Brown, 1983; Golding, Aquadro & Langley, 1986; Aquadro *et al.* 1987) in *Drosophila melanogaster*. The picture that has emerged from these studies is that insertions can be found throughout the genomes of wild-caught individuals. Few are found in the transcriptional units of genes, but many can be found in the flanking regions. Although elements have been found at many different chromosomal sites, their individual frequencies are all quite low. In fact few, if any, have been recovered more than once (at the same nucleotide site) from a natural population. This distribution suggests that there are forces opposing the spread of elements by transpositional increase in copy number, perhaps due to selection against the mildly deleterious phenotypic effects of such insertions on adjacent genes (Langley, Brookfield & Kaplan, 1983; Charlesworth & Charlesworth, 1983; Charlesworth, 1985). This hypothesis was recently tested by comparing the numbers of transposable elements on *X* chromosomes and autosomes isolated from a natural population, under the assumption that selection against hemizygous mutations in males would significantly reduce the equilibrium numbers of transposable elements on the *X* chromosomes relative to those found on autosomes. There was no evidence for any reduction in copy number for the *X* chromosomes for two of the three families of elements studied (Montgomery, Charles-

* Corresponding author.

worth & Langley, 1987). These results, and evidence for an association of transposable elements with crossover suppressors (inversions), led to the suggestion that meiotic recombination between transposable elements at nonhomologous sites (giving rise to aneuploid gametes) is responsible for the containment of transposable element copy number in natural populations (Montgomery, Charlesworth & Langley, 1987).

In order to evaluate this hypothesis, a population genetics model that captures its essential aspects is needed. In Section 2 of this paper, a simple model which predicts the distribution of element numbers across genomic regions with differing rates of unequal exchange is presented. The expected distribution of elements between the *X* chromosome and autosomes is also derived, on the hypothesis that increase in copy number is checked by the deleterious consequences of unequal exchange. Section 3 contains the results of a simulation study which documents the adequacy of various assumptions incorporated into the model of Section 2. In Section 4, available data on the distribution of the transposable element *roo* (Meyerowitz & Hogness, 1982; Scherer *et al.* 1982, who named it *B104*), along chromosomes isolated from a natural population of *Drosophila melanogaster*, are examined for evidence supporting or contradicting the predictions of the model.

2. A population genetics model of unequal exchange

(i) *Construction of the model*

An infinitely large, random mating diploid population will be considered. Let n denote the number of copies in a randomly chosen individual in a given generation. It is assumed that not all copies are equally likely to participate in an unequal exchange. Hence, we suppose that there are J different 'types' of copies and that the number of copies of type j in the individual in question is denoted by n_j. The state of a randomly chosen individual is then characterized by the vector $\mathbf{n} = (n_1, n_2, \ldots, n_J)$. Clearly,

$$n = \sum_{j=1}^{J} n_j. \tag{1}$$

For the sake of generality, the meaning of copy type is kept vague for the moment. At the end of this section, an example is described where it is defined by the physical location of the copies in the genome.

The mean number of copies of the jth type in a given generation is the mean of n_j over all individuals, and will be written as $\bar{n}_j$. In a certain range of the relevant parameters of the model (i.e. rates of recombination, insertion and excision), the element frequencies at individual chromosomal sites will be very low, and linkage disequilibrium between element frequencies at different sites can be neglected (see Montgomery & Langley, 1983). In this case, an extension of the argument of Charlesworth & Charlesworth (1983) shows that the distribution of numbers of copies of the jth type between individuals in the population follows a Poisson distribution, with parameter $\bar{n}_j$. Furthermore, the covariance between the n_j for two different values of j is zero, if there is no linkage disequilibrium. The same argument (Charlesworth & Charlesworth, 1983) allows us to write the effect of selection on the number of copies of the jth type in an infinite population as $\bar{n}_j \partial \ln \bar{w}/\partial \bar{n}_j$ [cf. Charlesworth, 1985, eq. (7)], where $\bar{w}$ is the mean fitness of the population. This assumes that element frequencies are low at individual chromosomal sites, so that elements are mostly present in the heterozygous state. The value of this derivative depends on the way in which copy number is assumed to affect fitness (see below). Data on the distribution of elements in natural populations of *Drosophila* are consistent with this assumption of low frequencies for individual elements, and low levels of linkage disequilibrium (Langley *et al.* 1983; Charlesworth, 1988).

Let u be the probability per generation that a given element produces a replicate of itself that is inserted elsewhere in the genome, and let v be the corresponding probability that an element is lost by excision (cf. Charlesworth & Charlesworth, 1983; Langley, Brookfield & Kaplan, 1983). Weak selection due to mutations asociated with insertions of elements can be treated in the same way as excision, to a good approximation (Charlesworth & Charlesworth, 1983; Kaplan & Brookfield, 1983), and the two processes will not be distinguished here. For simplicity, u and v are here assumed to be constants. Let the probability that a newly-transposed element is of type j be p_j. Assuming weak evolutionary forces, so that second-order terms in u, v, etc. can be neglected, the final equation for the change per generation of mean copy number for the jth type of element in an infinite population ($j = 1$ to J) is then

$$\Delta \bar{n}_j = \bar{n}_j \frac{\partial \ln \bar{w}}{\partial \bar{n}_j} + \bar{n} p_j u - \bar{n}_j v, \tag{2}$$

where

$$\bar{n} = \sum_{j=1}^{J} \bar{n}_j.$$

The form of the selection term under the assumption that an individual's fitness is determined by the fraction of euploid gametes that it produces will now be examined. Let κ_{kl} denote the probability that a gamete is aneuploid as a result of an unequal exchange between particular copies of type k and type l. Let us assume that the κ_{kl} are sufficiently small that (to a first approximation) the fitness of an individual with the vector of copy numbers $\mathbf{n} = (n_1, n_2, \ldots, n_J)$ is given by

$$w(n) = 1 - \sum_{1 \leqslant k \leqslant l \leqslant J} \kappa_{kl} n_k n_l. \tag{3}$$

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

It follows that the derivative in equation (2) is given by

$$\frac{\partial \ln \bar{w}}{\partial \bar{n}_j} = -\sum_{k=1}^{J} \kappa_{jk} \bar{n}_k. \tag{4}$$

(ii) *Equilibrium conditions*

At equilibrium, we thus have the relations:

$$\bar{n}_j\left(v + \sum_{k=1}^{J} \kappa_{jk} \bar{n}_k\right) = p_j u \sum_{k=1}^{J} \bar{n}_k. \tag{5}$$

In particular, the equilibrium value of the ratio of mean copy numbers for types i and j is given by

$$\frac{\bar{n}_i}{\bar{n}_j} = \frac{p_i(v+\kappa_j(\bar{n}))}{p_j(v+\kappa_i(\bar{n}))} \quad (i \neq j), \tag{6}$$

where $\kappa_j(\bar{n}) = \Sigma_k \kappa_{jk} \bar{n}_k$. The quantity $\kappa_j(\bar{n})$ equals the probability that a particular element in region j is involved in an unequal exchange.

The model's predictions for the ratios $\bar{n}_i/\bar{n}_j$ depends on the size of v relative to $\kappa_i(\bar{n})$ and $\kappa_j(\bar{n})$. The two extremes are easy to describe. If v dominates, then the ratio $\bar{n}_i/\bar{n}_j$ is approximately equal to p_i/p_j. On the other hand, if v is negligible when compared to $\kappa_i(\bar{n})$ and $\kappa_j(\bar{n})$, then

$$\frac{\bar{n}_i}{\bar{n}_j} = \frac{p_i \kappa_j(\bar{n})}{p_j \kappa_i(\bar{n})}. \tag{7}$$

When v, $\kappa_j(\bar{n})$ and $\kappa_i(\bar{n})$ are all comparable, there is no simple approximation for $\bar{n}_i/\bar{n}_j$, and the equilibria must be found numerically.

Up until now the definition of copy type has been abstract. For the remainder of this section a copy's type is defined by its physical location in the genome. In particular, we assume that the genome is divided into J regions, labelled 1 to J, and that all copies in the same region are of the same type. We also assume that if the jth region is of physical length L_j, then

$$p_j = \frac{L_j}{\sum_{k=1}^{J} L_k} \quad (1 \leqslant j \leqslant J). \tag{8}$$

Hence, the chance that a newly transposed copy is of type j is proportional to the physical length of region j. Furthermore, the sites of newly transposed copies are assumed to be uniformly distributed within each region.

(iii) *Distribution of elements between the X chromosome and autosomes*

Up to now, we have considered only autosomal elements. It was pointed out to us by Dr Michael Turelli that, contrary to the statement on page 38 of Montgomery *et al.* (1987), the model of unequal exchange as a mechanism for checking transpositional increase in copy number may predict a deficiency of elements in the X chromosome relative to random expectation in *Drosophila*. This is because two-thirds of the X chromosomes in a population are present in females, and have the opportunity to participate in meiotic exchange, whereas only one-half of the autosomes are present in females (meiotic exchange is absent in males). The rate of elimination of elements due to the generation of rearrangements by unequal exchange may thus be higher for X-linked elements in *Drosophila*. In order to test fully the hypothesis of a role of unequal exchange in controlling copy number, it is necessary to provide quantitative predictions of the relative abundances of elements on the X and the autosomes from the models described above.

This can be done by combining the theoretical approach of Montgomery *et al.* (1987) with that of the present paper. We adopt the convention that subscript i refers to a region of the X chromosome, j to a region of an autosome, and k to the Y chromosome, treated here as a single entity. As above, we let $\bar{n}_i$, $\bar{n}_j$, and $\bar{n}_k$ denote the mean copy numbers for these genomic regions in diploid individuals. Since only one copy of the X chromosome is present in males, the net mean copy number in males is $\bar{n}_m = \frac{1}{2}\bar{n}_i + \bar{n}_j + \bar{n}_k$, and the mean copy number in females is $\bar{n}_f = \bar{n}_i + \bar{n}_j$. Let p_{im}, p_{jm} and p_{km} be the probabilities that a newly transposed element inserts into the respective regions in a male; the corresponding probabilities for females are p_{if} and p_{jf}. Combining the approach of Montgomery *et al.* [1987, eqns (1) and (2)] with the argument that led to equation (5), we obtain the following equilibrium relations:

$$\bar{n}_i(v+\kappa_i(\bar{n})) = \frac{u}{3}(p_{im}\bar{n}_m + 2p_{if}\bar{n}_f), \tag{9a}$$

$$\bar{n}_j(v+\kappa_j(\bar{n})) = \frac{u}{2}(p_{jm}\bar{n}_m + p_{jf}\bar{n}_f), \tag{9b}$$

where $\kappa_i(\bar{n})$ and $\kappa_j(\bar{n})$ measure the net selective effects of unequal exchange for the X-linked and autosomal regions in question. A similar equation can be written for $\bar{n}_k$, but it turns out that this is not needed in the application of the theory to population data [see Section 4(iii)]. Given functional forms for these quantities, this equation can be used to determine the expected relative copy numbers in these regions. These forms depend on the assumptions made concerning the mechanism of unequal exchange, which will be discussed in the next section.

(iv) *Models of unequal exchange*

Consider first a model in which the probability of unequal exchange between a pair of elements is determined only by the frequencies of recombination events in the regions in which they are located, and is otherwise independent of their genomic locations. One might expect the probability of unequal exchange between two copies to be related to the densities per

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

unit physical length of regular exchange in the regions in which they are located. The simplest model of this kind is to assume a mass-action relationship such that the probability of unequal exchange between a copy in region k and a copy in region l, with densities of regular exchange c_k and c_l, is equal to $\omega c_k c_l$, where ω is a constant of proportionality. Thus

$$\kappa_{kl} = \omega c_k c_l \quad (1 \leqslant k, l \leqslant J). \tag{10}$$

It follows from (10) that

$$\kappa_j(\bar{n}) = \sum_{k=1}^{J} \kappa_{jk} \bar{n}_k = \omega c_j \left[\sum_{k-1}^{J} c_k \bar{n}_k \right]. \tag{11}$$

Substitution into (7) leads to the equilibrium ratio of mean copy numbers for the case of negligible rates of excision

$$\frac{\bar{n}_i}{\bar{n}_j} = \frac{L_i c_j}{L_j c_i}. \tag{12}$$

Hence, the ratio of the mean densities of copies in two regions increases as the ratio of the expected frequencies of exchanges decreases.

An alternative model is to assume that unequal exchanges can only occur between two copies that are sufficiently close to each other in the homologous chromosomes. Hence, κ_{kl} is zero for $k \neq l$. If two copies are present in a region k of length L_k, the probability that they are within a sufficiently small distance of each other to undergo element is proportional to $1/L_k$, assuming a uniform distribution of copies within the region. Given that they are sufficiently close, the probability of an unequal exchange is proportional to c_k. Hence, we can write

$$\kappa_{kk} = \omega \frac{c_k}{L_k}, \tag{13}$$

and

$$\kappa_j(\bar{n}) = \omega \frac{c_j \bar{n}}{L_j}$$

and so (with negligible excision)

$$\frac{\bar{n}_i}{\bar{n}_j} = \frac{L_i}{L_j} \sqrt{\frac{c_j}{c_i}}. \tag{14}$$

Thus, the ratio of the expected densities of copies now decreases as the square root of the ratio of the expected densities of exchanges.

(v) *Dynamics of copy numbers*

The dynamics of change in copy number under these conditions can be investigated as follows. When all the $\bar{n}_j$ are close to zero, it is easy to see from equation (2) that $\Delta\bar{n}_j > 0$ for each j provided that $u > v$, regardless of the values of the κ_{kl}. Furthermore, it follows from equations (2) and (4) that $\Delta\bar{n}_j < 0$ for sufficiently large values of the $\bar{n}_j$ if at least some of the $\kappa_{kl} > 0$ for each k or if $v > 0$. This suggests that copy numbers will normally increase away from zero, and remain bounded, unless v is zero and some rows of the matrix $\{\kappa_{kl}\}$ are zero. A full analytical investigation of the stability properties of the equilibria is difficult, in view of the complexity of the equilibrium equations. However, numerical analysis of equations (2) and (4) confirms this conclusion for the two models of unequal exchange frequencies based on the locations of copies described above. In all cases investigated, mean copy numbers converged to the equilibrium values predicted by the above formulae.

3. Simulation studies

(i) *Simulation model*

Monte Carlo simulations were carried out to examine the validity of the assumptions of the model in a more realistic biological setting, i.e. with a finite population and more detailed genetic mechanisms. In these simulations, each generation is represented by N diploid individuals each consisting of a single pair of chromosomes. The chromosomes are divided into three regions, with physical lengths $L_1 = 0{\cdot}25$, $L_2 = 0{\cdot}5$, and $L_3 = 0{\cdot}25$, respectively. As before, the expected numbers of exchanges (either crossovers or unequal exchanges) per genome per generation per unit physical length in the three regions are denoted c_1, c_2, and c_3. To produce the N individuals of an offspring generation from the parent generation, the following sequence of steps is followed. (1) Two distinct diploid parents are chosen at random with replacement. (2) From each of these two parents, a gamete is produced by recombination, with the probability of an unequal exchange being determined by one of two methods described below. If either of the two gametes is aneuploid as a result of unequal exchange, we go back to step (1). (3) If neither gamete is aneuploid, the two gametes are paired to form a zygote, and each copy is subject to an excision process. That is, each copy, independent of the other copies, is excised from the genome with probability v and is retained with probability $1-v$. (4) Then duplicative transposition occurs, whereby each copy duplicatively transposes to a new site with probability u. The sites of newly transposed copies are assumed to be uniformly distributed on the chromosome, so that the probability that a new site is in region i is L_i. The resulting diploid individual at this point becomes a member of the offspring generation. The sequence of steps (1)–(4) is repeated N times to produce the offspring generation.

Two methods were used to determine the occurrence of unequal exchange in the production of gametes, corresponding to the two models described in Section 2. In method one, it is assumed that the probability that a gamete is aneuploid as a result of unequal exchange is given by equation (1), with $\kappa_{ij} = bc_i c_j$.

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

This form for κ_{ij} is identical to that in equation (10), if b is identified with ω.

In method two, one first determines the locations of the exhanges. The number of exchanges in region i is assumed to be Poisson-distributed with mean $c_i L_i$, and these exchanges are assumed to be uniformly distributed within region i. For each exchange, the maternal and paternal chromosomes are examined for the presence of copies within a distance ω of the exchange. The probability of an unequal exchange occurring is assumed to be kk'/hh', where h and h' are the numbers of copies within distance ω of the exchange on the maternal and paternal chromosomes respectively, and where k and k' are the numbers of heterozygous copies (copies in sites not occupied in the other chromosome) within a distance ω of the element. Thus, an unequal exchange occurs only between heterozygous copies, consistent with the assumptions of the infinite population model of Section 2. If no unequal exchanges occur, then a viable gamete is produced. With this method, the probability of an unequal exchange depends not only on the numbers of copies in each region but also on the spatial distribution of the copies within regions. If the copies are randomly and uniformly distributed within regions, and if ω is sufficiently small, then we expect that equation (7) will apply to the final generation of the simulations.

(ii) *Simulation results*

The simulation results are shown in Tables 1 and 2. All the simulations reported here were run with N equal to 1000. This is a relatively large population for computer simulations, but a large population size is needed if equation (2) is to provide a good approximation. The parameters of the model were chosen so that the number of copies in a random chromosome was similar to that observed (see Table 4). With method one, for all recombination rates considered in the simulations, the mean of the number of copies in a region is approximately equal to the variance of the number of copies in the same region. With method two, even when the expected number of elements is as low as two per chromosome per generation, the variance appears to be only slightly bigger than the mean. Thus the assumption of a Poisson distribution of copy number between individuals for a given class of sites would appear to be valid for the parameter values examined. With larger population sizes, presumably a Poisson distribution would apply even with smaller exchange rates. The means and variances of the total number of copies per chromosome are also nearly equal, as shown in Tables 1 and 2. This is also expected, since the sum of independent Poisson variates is also Poisson. Note that, for the lowest exchange rates in Table 2, the variances of the number of copies in

Table 1. *Simulation results for method one (see text)*

						Theory			Simulation			
b	u	v	c_1	c_2	c_3	n_1	n_2	n_3	n_1	n_2	n_3	$\sum_{j=1}^{3} n_i$
$2{\cdot}5\times10^{-7}$	0·01	0·0001	28·0	96·0	28·0	6·8	4·0	6·8	6·8 (6·7)	4·0 (3·9)	6·8 (6·6)	17·7 (17·4)
$2{\cdot}0\times10^{-6}$	0·01	0·0001	7·0	24·0	7·0	12·2	7·2	12·2	12·0 (11·4)	7·2 (7·1)	12·0 (11·7)	31·3 (30·3)
$2{\cdot}0\times10^{-6}$	0·01	0·001	7·0	24·0	7·0	10·7	7·0	10·7	10·2 (9·9)	7·1 (7·0)	10·6 (10·3)	27·8 (27·2)
$2{\cdot}0\times10^{-7}$	0·02	0·01	28·0	96·0	28·0	5·1	5·6	5·1	4·6 (4·7)	5·3 (5·4)	4·5 (5·4)	14·5 (15·2)

These simulations were carried out with $N = 1000$ and in each case $L_1 = 0{\cdot}25$, $L_2 = 0{\cdot}5$, and $L_3 = 0{\cdot}25$. In the initial generation all chromosomes contain 10 randomly placed copies. After one thousand generations and every five hundred generations after that, the average number of copies in each region and the variance of the number of copies in each region are calculated until the ten thousandth generation. The average of the averages are presented above under the simulation heading. The average of the variances are in parentheses. The predictions of the theory, given above under the heading 'theory', are obtained by solving equation (6) numerically, with $\kappa_{ij} = bc_i c_j$. (Further details are given in the text.)

Table 2. *Simulation results for method two (see text)*

						Theory			Simulation			
ω	u	v	c_1	c_2	c_3	n_1	n_2	n_3	n_1	n_2	n_3	$\sum_{i=1}^{3} n_i$
0·00115	0·01	0·0001	28·0	96·0	28·0	6·1	6·6	6·1	7·2 (6·9)	7·7 (7·8)	6·9 (6·7)	21·9 (22·1)
0·00092	0·03	0·025	6·0	120·0	6·0	8·2	11·8	8·2	7·6 (8·0)	11·9 (12·3)	7·6 (8·1)	27·1 (30·5)
0·00111	0·03	0·025	4·0	60·0	4·0	7·0	10·3	7·0	6·4 (6·9)	10·2 (10·4)	6·4 (6·7)	23·0 (25·5)
0·00533	0·03	0·025	0·25	4·0	0·25	7·0	10·3	7·0	6·8 (8·7)	11·1 (12·0)	6·7 (8·6)	24·6 (32·6)

These simulations were carried out as with those in Table 1, except that the occurrence of unequal exchange is determined according to method two, described in the text. The predictions of the theory, given above under the heading 'theory', are obtained by solving equation (10) numerically with $\kappa_{ij} = 0$ for $i \neq j$, and $\kappa_{ii} = \omega c_i / L_i$. (Further details are given in the text.)

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

each of the regions as well as the total number of copies are somewhat larger than the mean, indicating some correlations between sites. The covariances of the number of copies in different regions were also examined directly for some cases (data not shown) and the results confirm that the Poisson assumption holds approximately for the parameter values used in Tables 1 and 2.

For the first two simulations reported in Table 1, the excision rate is small compared to the effects of unequal exchange, and so we expect that equation (12) should accurately predict the ratios of the mean numbers of copies in the different regions. This is indeed the case. For example, for the first set of parameters, $\bar{n}_1/\bar{n}_2$ is 1·7, and the right-hand side of (11) is (0·25/0·5)(96/28) = 1·71. For the remaining two simulations of Table 1, the excision rates are not small when compared to the $\kappa_j(\bar{n})$, and so the predictions of equation (11) are not as accurate. Of the simulations reported in Table 2, only the first has a relatively low excision rate. For this case, we expect that equation (14) should apply. In fact, $\bar{n}_1/\bar{n}_2$ is 0·94 and the right-hand side of (14) is $(0{\cdot}25/0{\cdot}5)(96/28)^{\frac{1}{2}} = 0{\cdot}93$. For the other cases in Table 2, the rates of unequal exchange and the rates of excision are comparable and equation (13) does not apply, but the predictions of equation (6) are fairly accurate.

4. Data analysis

(i) *General considerations*

The analytical and simulation results in the previous two sections show that the proposed mechanism of unequal exchange between transposable element sequences at nonhomologous sites can stabilize copy number in a random mating population, in the absence of any phenotypic effects of the insertions themselves. If the rate of insertion is uniform and proportional to physical length, then the model predicts that: (1) if excision or selection against insertional mutations is the dominant mechanism for removing copies, then the densities of copies in genomic regions with different rates of exchange should be equal; (2) if unequal exchange is primarily responsible for controlling copy numbers, and is related to exchange as described by the model, then the ratio of densities of copies in different genomic regions is a function of the ratio of the densities of exchange in those regions. The form of the function depends on the particular assumptions about the topologies of pairing.

In order to obtain some empirical information concerning these factors, the distribution of the *copia*-like transposable element, *roo*, in a sample of *X*, *II* and *III* chromosomes from a natural population of *Drosophila melanogaster* was examined using the technique of *in situ* hybridization to salivary gland chromosomes. The methods and source of the chromosomes are described in Montgomery, Charlesworth & Langley (1987). The rationale of the analysis is based on the fact that the frequency of meiotic exchange per nucleotide site is much reduced in the distal and proximal euchromatin of *D. melanogaster* chromosomes (Lindsley & Sandler, 1977); hence, we would expect a disproportionate concentration of elements in these regions if the processes described above were important in controlling element abundances. We chose *roo* as the subject of study, since its mean copy number is much higher than that of the other two elements studied by Montgomery *et al.* (1987), and therefore provides a more sensitive test of non-randomness of the distribution of elements over the salivary chromosome map.

(ii) *Methods for subdividing the salivary chromosome maps*

To examine the relationship between the levels of crossing over and the densities of *roo* copies in the various regions of the chromosomes from a natural population, we have divided each chromosome arm into three arbitrary regions: the base, the tip and the middle. From the above considerations, the model would predict a disproportionate abundance of elements in the tips and bases of the chromosomes, compared with the middle. Table 3 shows the cytological boundaries of these regions for the five chromosome arms. The division of the *X* chromosome follows Rudkin (1965), who measured the DNA content across the *X* and reported values for these specific regions which are also well characterized genetically and cytogenetically. The bases and tips of the autosomal arms were chosen arbitrarily as the proximal three and distal one standard cytological subdivisions. Minor exceptions to this rule were made to incorporate available cytogenetic information in estimating gene densities (see below). Table 4 presents the total numbers of *roo* copies found in the three regions of the sampled sets of chromosomes.

A comparison of the densities of *roo* elements in these different regions with the null hypothesis of a random distribution requires estimates of their physical lengths, which we presume to reflect the chance of insertion. For the *X* chromosome such an estimate exists, as already mentioned, while indirect estimates

Table 3. *Cytogenetic regions used in the analysis. The interval refer to salivary gland chromosome subdivisions (Lefevre, 1976). Each region includes all subdivisions within the designated interval*

	Chromosome arms				
Region	X	2L	2R	3L	3R
Base	18D1-20A4	38-40	41-43E	77-80	81-84
Middle	3A5-18C9	22-37	43F-59D	62B-76	85-99
Tip	1A1-3A4	21	59E-60	61-62A	100

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

Table 4. *Numbers of* roo *transposable elements found in the bases, middles and tips of the* X *chromosomes and autosomal arms from a natural population*

	Chromosome arms				
	X	2L	2R	3L	3R
Sample size	18	13	13	12	11
Base	47	34	38	21	26
Middle	142	132	100	114	122
Tip	17	10	3	12	11

are needed for the autosomes. The density of known mutant loci mapped to the region provides one such estimate. It reflects the mutational target size of the region. The gene densities were estimated from the information on pages 433–470 of Lindsley & Grell (1968), using updated information on cytological locations from Lindsley & Zimm (1986) where necessary. The Lindsley & Grell (1968) mutant list was chosen in preference to the more extensive listing of Lindsley & Zimm (1986), because of the danger of biasing the estimates due to the intensive saturation mapping of specific regions that has been carried out since 1968. Only mutants indicated as extant in 1968 were included in these estimates, in order to avoid dubious assignments of positions. All mutants distal to *l(1)zw1* were included in the tip of the *X*. All mutants proximal to *car* and distal to *bb* were included in the basal region of the *X*. For 2L, all loci distal to *shr* were included in the tip, and all loci proximal to *pr* were included in the base. For 2R, loci distal to *bw* were assigned to the tip, and loci proximal to *ix* were included in the base, ignoring *puf* and *blo* because of uncertainty about their cytological positions. For 3L, loci distal to *eyr* were assigned to the tip, and all loci proximal to *fs(3)G1* were assigned to the base. For 3R, loci proximal to *p* were assigned to the base. Loci distal to *Acph-1* were assigned to the tip of 3R. This probably overestimates the density in this region, since there is uncertainty about the cytological positions of several of these loci. A second indirect method of estimating the relative sizes of the regions of a chromosome arm is to determine the fraction of standard salivary gland chromosome map represented by each region. This was done by measuring the lengths of the regions on the standard photographic maps of the salivary chromosomes (Lefevre, 1976).

(ii) *Results: distribution of* roo *within chromosomes*

The relative amounts of DNA (Rudkin, 1965), the estimates of relative gene density, the proportions of the salivary gland chromosome maps, and the frequencies of salivary gland chromosome sites to which *roo* DNA probes hybridized are presented in Table 5 for each of the five chromosome arms. Notice that the relative gene densities in the three regions of the *X* agree well with the relative DNA contents. Statistical tests for the departure of the element abundances were carried out by means of χ^2 comparisons of the observed numbers of elements with the numbers expected if elements were distributed across in proportion to their size, for the size estimates based on DNA content and on salivary map measurements. For the gene density estimates, contingency χ^2 tests

Table 5. *The proportions of DNA, genes, salivary gland chromosome map and* roo *elements in the bases, middles and tips of X, 2L, 2R, 3L and 3R (see text for explanation)*

Region		DNA	Gene density (gene numbers)		Map	Copies
X	Base	0·07****	0·07****	(15)	0·06****	0·23
	Middle	0·81	0·79	(182)	0·85	0·69
	Tip	0·12	0·14	(32)	0·09	0·08
2L	Base	—	0·12	(10)	0·17	0·19
	Middle	—	0·78	(65)	0·78	0·75
	Tip	—	0·10	(8)	0·05	0·06
2R	Base	—	0·17	(17)	0·16****	0·27
	Middle	—	0·65	(63)	0·78	0·71
	Tip	—	0·18****	(18)	0·06	0·02
3L	Base	—	0·18	(10)	0·14	0·14
	Middle	—	0·73	(41)	0·78	0·78
	Tip	—	0·09	(5)	0·08	0·08
3R	Base	—	0·05**	(4)	0·16	0·16
	Middle	—	0·91	(75)	0·80	0·77
	Tip	—	0·04	(3)	0·04	0·07

Significance levels for χ^2 tests (1 D.F.) of the numbers of observed *roo* elements in a given region against the random expectation based on the indicated criterion.
** $P < 0{\cdot}01$; **** $P < 0{\cdot}001$.

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

were performed to compare the distributions of mutant genes and elements across regions.

The density of copies at the base of the *X* is significantly greater than expected from the DNA contents, gene densities or the salivary gland chromosome map. There is a slight but non-significant deficit of copies in the tip compared with the random expectation on any of the three criteria. The distribution of *roo* copies along 2L fits the random expectation based on the gene density or the salivary gland chromosome map. The base of 2R contains significantly more copies than expected from the salivary gland chromosome map, and the tip is deficient in copies on either criterion. 3L shows no significant departures from the expectation based on either method of estimation, while the base of 3R shows a significant excess of copies relative to gene density. However, the numbers of copies at the base of 3R is exactly that predicted from the salivary gland chromosome map.

In summary, there is clear evidence for an excess of copies at the base of the *X* chromosome, of the order of three-fold, but no evidence for an excess near the tip. The data for the other chromosome arms are less clear, since the gene density and salivary gland chromosome map criteria yield somewhat different statistical results. In each case, however, there is either a substantial excess of copies at the base or close agreement between observed frequencies and random expectation. Either criterion yields a highly significant excess at the base for one autosomal arm (2R on salivary gland chromosome map length, and 3R on gene density). This suggests that there is, indeed, a tendency for *roo* copies to accumulate at the base of the euchromatin but the effect is usually not very large for the autosomes. There is no evidence for an accumulation of copies at the tips of the autosomes; rather, there is a slight tendency for copies to be less frequent there than expected (strongly significant in the case of 2R, on the gene density criterion).

Since the data for the *X* chromosome are the strongest for rejecting the hypothesis that excision is the primary force for removing copies, we will focus on these data. In Table 6, the ratios of the densities of copies in the different regions of the *X* chromosome are compared to the ratios and to the square roots of the ratios of the densities of meiotic crossing over as estimated from Lindsley & Grell (1968). It is clear from the table that the data do not support the quantitative predictions of the model given by equations (12) and (14).

Table 6. *The observed and predicted densities of copies on the* X *chromosome*

$(i,j)^a$	$\frac{n_i/L_i}{n_j/L_j}$	$\frac{c_j}{c_i}^b$	$\sqrt{\frac{c_j}{c_i}}^b$
(1, 2)	3·8	1·56	1·25
(1, 3)	4·7	0·17	0·41
(2, 3)	1·23	0·11	0·33

[a] *i* and *j* refer to the regions of the *X* chromosome. Region 1 is the base, region 2 is the middle, and region 3 is the tip of the *X*. $L_1 = 0{\cdot}07$, $L_2 = 0{\cdot}81$, and $L_3 = 0{\cdot}12$, as in Table 3.
[b] These ratios are calculated using available estimates of the densities of equal exchanges in the three regions: $c_1 = 1{\cdot}6 \times 10^{-5}$, $c_2 = 2{\cdot}5 \times 10^{-5}$ and $c_3 = 2{\cdot}7 \times 10^{-6}$. (Crossovers per kilobase.)

(iii) *Results: distribution of* roo *between mid-sections of X chromosomes and autosomes*

We have also compared the relative abundances of *roo* across the *X* chromosome and autosomes. The base and tip regions were eliminated from the study, in order to be able to compare regions with similar amounts of genetic exchange per unit physical length. From the map positions of the cytogenetic delineations of the tip, mid and base portions of the chromosomes in Table 3 (Lindsley & Zimm, 1986), the average map length of the mid-section of the autosomes is 49·9, while the map length of the mid-section of the *X* chromosome is 61·5. These values yield a ratio $c_j/c_i = 0{\cdot}811$, for use in determining the functions $\kappa_i(\bar{n})$ and $\kappa_j(\bar{n})$ in equation (9). With the assumption that the rate of excision v is negligible, it follows from equation (9) that the ratio of the mean numbers of elements in the mid portions of the *X* chromosome and a given autosomal arm ($\bar{n}_i$ and $\bar{n}_j$ respectively) satisfies

$$\frac{\bar{n}_i \kappa_i(\bar{n})}{\bar{n}_j \kappa_j(\bar{n})} = \frac{2(p_{im}\bar{n}_m + 2p_{if}\bar{n}_f)}{3(p_{jm}\bar{n}_m + p_{jf}\bar{n}_f)}. \quad (15)$$

The ratio $\kappa_i(\bar{n})/\kappa_j(\bar{n})$ corresponds to s_X/s_A in equation (2) of Montgomery *et al.* (1987), and the right-hand side of equation (15) is identical to that of their equation (2), with the appropriate change of notation. In determining the value of this ratio, we first note that, on the assumption that unequal exchange is meiotic, no exchange will occur in males. Hence, only the copy numbers for females contribute to $\kappa_i(\bar{n})$ and $\kappa_j(\bar{n})$. In the course of this study, an error in equation (1*a*) of Montgomery *et al.* (1987) was detected. The correct formula is given in the appendix, and was used in our present analysis.

For the first model of unequal exchange, where the probability of unequal exchange is independent of location, equation (11) implies that $\kappa_i(\bar{n}) = 2\omega c_i \bar{n}_f/3$, since two-thirds of the *X* chromosomes are present in females. Similarly, $\kappa_j(\bar{n}) = \omega c_j \bar{n}_f/2$, since one-half of the autosomes are present in males. Hence, $\kappa_i(\bar{n})/\kappa_j(\bar{n}) = 4c_i/3c_j$. Since $\kappa_i(\bar{n})/\kappa_j(\bar{n})$ is a constant, the ratio $\bar{n}_i/\bar{n}_j$ is the solution of a quadratic equation, and the analysis of Montgomery *et al.* [1987, eqn (2)] can be applied. The only difference is that their estimates for the insertion probabilities in the *X*-chromosomal and autosomal euchromatin (their P_{ij}, P_{im} and P_{jf}, P_{jm}, respectively) must be multiplied by 0·8 in order to

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

obtain the corresponding p values used here, since the mid portions of the X and autosomes constitute approximately 80% of the euchromatin (Table 5).

For the second model of unequal exchange, where unequal exchange occurs only between neighbouring elements, it follows from equations (13) that

$$\frac{\kappa_i(\bar{n})}{\kappa_j(\bar{n})} = \frac{4c_i\bar{n}_i L_j}{3c_j\bar{n}_j L_i}, \tag{16}$$

where L_i and L_j are the total physical lengths of the mid portions of the X and autosomal regions. Assuming that the mid portions of the X chromosome and of each autosomal arm have the same physical length, we have $L_j/L_i = 4$, and so $\kappa_i(\bar{n})/\kappa_j(\bar{n}) = 16c_i\bar{n}_i/3c_j\bar{n}_j$. Substituting this into the equation (15) results in a cubic equation for $\bar{n}_i/\bar{n}_j$ and so the arguments used by Montgomery *et al.* (1987) need to be modified appropriately to obtain the solution.

Table 7 shows the predicted proportions of elements in the mid regions of the X chromosome and the autosomes for four models: (1) the 'null model', in which transposition and elimination of elements take place at equal rates for the X chromosomes and autosomes; (2) the model of natural selection against partially recessive, deleterious effects of insertional mutations (cf. Montgomery *et al.* 1987); (3) the first unequal exchange model in which each element can recombine with other elements located throughout the genome; and (4) the second unequal exchange model in which each element can recombine only with elements located in the same region of the genome.

Since the numbers of autosomes scored varied, and was less than the 18 X chromosomes scored, the expected numbers in Table 7 were adjusted to reflect the fact that the total number of autosomal arms scored was only 49, instead of the 72 expected if the numbers of autosomes had equalled the number of sex chromosomes (i.e. the expected ratio $\bar{n}_i/\bar{n}_j$ is multiplied by 1.469). The results in Table 7 indicate that the null model is marginally inconsistent with the data. The unequal exchange models are consistent with the data. But the insertional mutation model predicts significantly different proportions of elements on the X chromosome from those observed.

5. Discussion

We have described a population genetics model of the evolution of transposable elements in a Mendelian host population that incorporates asymmetric pairing between copies at different genomic sites, and subsequent unequal exchange leading to selectively deleterious aneuploid gametes. The appeal of the mechanism of asymmetric pairing and unequal exchange between transposable elements at non-homologous sites is substantial, both from a theoretical and experimental perspective. First, the probability that a gamete is aneuploid as a result of unequal exchange is a quadratic function of copy number [see equation (3)]. As shown by Charlesworth & Charlesworth (1983), any selective force that is sufficient to stabilize transposable element copy number, in the absence of regulated transposition rates, must be a nonlinear function of copy number. Second, this mechanism does not depend on the details of the nature of the transposable element in question, but rather on sequence similarity between elements belonging to the same family. Thus all repetitive dispersed DNAs, however they spread, will be subject to this process as their copy numbers increase. Third, the observation of Montgomery *et al.* (1987) that the numbers of *roo* and *297* copies on the X chromosomes of flies in natural populations is not reduced in comparison to the numbers on the autosomes, as expected from the effects of natural selection against the mutational effects of insertions, suggests that selection against the mutational effects of element insertions is insufficient. Some other mechanism must be sought. Fourth, the accumulation of copies near the

Table 7. *Comparison of the observed numbers of* roo *elements in the mid-sections of the* X *chromosomes and autosomes with the predictions of the models (see text)*

	Proportions per haploid genome		Numbers		
	X	Autosomes	X	Autosomes	χ^2 D.F. = 1
Observed	0·17	0·84	142	468	—
Models	—	—	(18)[a]		(49)[b]
Null	0·20	0·80	164	446	4·1*
Selective	0·13	0·87	112	498	9·7***
Unequal exchange no. 1	0·16	0·84	131	479	1·1
Unequal exchange no. 2	0·18	0·82	149	461	0.4

[a] Number of X chromosomes scored.
[b] Number of autosomal arms scored.
* $P < 0{\cdot}05$.
*** $P < 0{\cdot}005$.

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press

breakpoints of balanced inversions in old laboratory stocks reported by Montgomery *et al.* (1987) indicates that recombination may be involved in the control of element abundance. The permanent heterozygosity of the inversions in these stocks prevents crossing over, especially near the breakpoints. It may have a similar effect on unequal exchange near these breakpoints, thus allowing copies to accumulate.

In order to test the predictions of the models of unequal exchange, the numbers of *roo* transposable elements found in various regions of a set of chromosomes sampled from a natural population of *Drosophila melanogaster* were determined. The expectation was that elements would be more abundant on average in parts of the genome where crossing over is less frequent, such as the tips and the bases of the euchromatin (Lindsley & Sandler, 1977). The data from chromosome arms 2L and 3L indicate that the densities of copies across different regions are similar, whereas the data from the *X* chromosome and chromosome arms 2R and 3R indicate that elements tend to be more abundant near the bases of chromosomes than expected from a random pattern of insertion. Unpublished data of Charlesworth on ten element families also show a similar excess of the elements *roo*, *297*, *2156*, *2158*, *412* and *2181* at the base of the *X* chromosome, for a sample of 14 chromosomes extracted from a population at Beltsville, Maryland. The elements *roo* and *2181* also showed a smaller, but significant, excess at the tip. The elements *copia*, *2161*, *2210* and *2161* show no significant departure from random expectation. Furthermore, the cloning study of Miklos *et al.* (1984) on the basal euchromatin of the *X* provides evidence for an accumulation of middle repetitive DNA sequences in the same region as our *in situ* studies. On the other hand, the results of Leigh Brown & Moss (1987) on *I* and *copia* in a Spanish population fail to show an excess of elements at the base of the *X*. There is little doubt, that some families of elements tend to accumulate in the basal euchromatin of the *X*, and weaker evidence for an accumulation at the tip.

This pattern of accumulation of elements suggests that unequal exchange may be involved in eliminating copies. The models of unequal exchange predict that the ratio of the densities of copies in two regions of the chromosome is a function of the ratio of densities of exchanges in the two regions, but, as described in Section 3, the data do not fit the models' predictions if it is assumed that the density of unequal exchange events parallels the density of regular exchange. For example, the density of copies in the tip of the *X* chromosome appears to be too small, whereas the density of copies in the base appears to be too large (Table 6). It is interesting to note for the other chromosome arms that, whenever the observed densities do not agree with random expectations, it is because of an excess of copies at the base and not at the tip, except for a deficiency of elements at the tip of 2R on the gene density criterion. The reduction in crossing over at the tips of the autosomes is far less marked for the autosomes than it is for the *X* chromosome (Lindsley & Sandler, 1977), so that the absence of an effect for the autosomal tips is not surprising, but the absence of an effect at the tip of the *X* poses a severe problem for the validity of the unequal exchange hypothesis.

There are several possible explanations for the inconsistencies between the data on the distribution of the *roo* element in different regions of the *X* chromosome and the models' predictions. The assumption that the probability of insertion in a region is proportional to the physical length of that region may not be valid. For example, a newly transposed copy may be more likely to be inserted in the base of the chromosome than elsewhere, although there is no obvious reason why this should be the case. The assumption of a uniform rate of excision or selection against insertional mutations may also not apply, although again it is not clear why. [An absence of genes that are vulnerable to mutational insertions, yielding a lower chance of an element's experiencing selection, seems unlikely, since the density of mutable loci is not abnormally low in these regions (Table 5)]. Finally, the assumptions about the mechanism of unequal exchange may be invalid, e.g. the parameter ω may be region-specific.

It is also possible that the abundance of elements in regions of restricted crossing over may be affected by factors other than unequal exchange. As pointed out by Charlesworth (1985), if excision rates are sufficiently low in relation to population size, a build-up of elements in genomic regions where meiotic exchange is rare in comparison to the reciprocal of population size is expected as a result of the operation of the stochastic process known as Muller's ratchet. Over a long period of evolution time, this could cause an excess of elements in regions where crossing over is suppressed. However, it seems unlikely that Muller's ratchet could be responsible for the discrepancy just noted, since crossing over on the *X* chromosome is more strongly suppressed near the tip of the *X* than near the base (Lindsley & Grell, 1968; Lindsley & Sandler, 1977). The operation of the ratchet is highly sensitive to the occurrence of even a low level of crossing over (Felsenstein & Yokoyama, 1976), and so it would be expected to act more effectively on the tip than the base of the *X*.

Another source of discrepancy may be that the standard estimates of the densities of equal exchange are not appropriate for natural populations. It is well known that the interchromosomal effect on crossing over caused by rearrangement heterozygosity in other chromosomes is not uniform over different regions of the genome (Sturtevant, 1919; Lucchesi & Suzuki, 1968). The increase in the frequency of crossing over in the *X* caused by inversion heterozygosity in the two autosomes is primarily at the base

and tip of the chromosome. In some cases (e.g. for the effect of the *In(Cy)* inversions on *X*-chromosome crossing over), the largest effect is at the tip. Other inversions may have a larger effect on the base (Schultz & Redfield, 1951). Since many *D. melanogaster* populations have high heterozygosity for autosomal inversions, it is conceivable that the discrepancy in Table 6 may be due to too low an estimate for c_3. Experiments using laboratory stocks suggest that 2nd and 3rd chromosome inversion heterozygosity increases the frequency of unequal exchange in the *white* locus region (Montgomery, E. A., Huang, S.-M., Langley, C. H., Judd, B. H., unpublished results). Data on the effects of second and third chromosome inversions that occur naturally in the population studied here are needed to resolve this question.

One of the motivations for considering unequal element between transposable elements as a mechanism to contain copy number was the observation that two out of the three elements studied previously (Montgomery *et al.* 1987) showed no apparent reduction of copy number on the *X* chromosome relative to the autosomes, as would be expected if natural selection against insertional mutations were acting. Since proportionally more *X* chromosomes than autosomes are in females, and thus undergo crossing over, the effects of unequal exchange are also expected to be more effective in reducing copy number for the *X* chromosomes than for the autosomes. The two models of unequal exchange were both investigated theoretically [Section 2(iii)], and their predictions for the distribution of elements between the *X* chromosomes and the autosomes compared with the data on *roo* (see Table 7). Only the mid 80% of the chromosome arms (excluding the tip and base) were used to compare with the predictions of the various models, in order to avoid pooling regions with very different rates of exchange per unit length.

As described in Section 4(iii), the observed proportion of *roo* elements on the *X* chromosome is inconsistent with the prediction of the model of natural selection against mildly deleterious mutational effects of element insertions [as previously reported by Montgomery *et al.* (1987) for the entire euchromatic portions of these chromosomes]. The data are also marginally inconsistent with the prediction of the first unequal exchange null model. However, the models of unequal exchange make predictions of the proportion of elements carried on the *X* chromosome that are very close to the observed value (see Table 7).

Unequal exchange between repeated elements in yeast and *Drosophila* has been observed to occur mainly at meiosis when regular exchange (crossing over) is also much more frequent. In yeast, repeated sequences may recombine with each other throughout the genome (Mikus & Petes, 1982; Roeder, 1983; Lichten, Borts & Haber, 1987). The evidence from *Drosophila* suggests that local events may be favoured (Green, 1959; Goldberg *et al.* 1983; Judd, 1959; Davis *et al.* 1987). This difference may simply reflect the difference in the sizes of the genomes. In any case it appears that the second unequal exchange model could account for the proportion of *roo* elements on the mid portions of the *X* chromosome relative to the mid portions of the autosomes.

Throughout this analysis, we have assumed that all copies are heterozygous in large populations, as is consistent with the *Drosophila* data (Montgomery & Langley, 1983; Leigh Brown & Moss, 1987; Charlesworth, 1988). In mammals, elements such as *Alu* and *L1* appear to be fixed at most sites where they are present (Brookfield, 1986; Bellis *et al.* 1987). In one instance, these sequences have been implicated in the origins of deleterious rearrangements (Lehrman *et al.* 1987). An important question is whether homozygous copies (i.e. ones that can pair symmetrically) also participate in asymmetric pairing and unequal exchange, or are they somehow inert? If homozygous copies pair symmetrically, then there may be a curious 'under-dominance' through which copies are selected out of the population when they are rare, but are at a relative advantage when they are common. Random genetic drift in small ancestral populations may have allowed copies to rise to high frequency. Selection against heterozygotes via asymmetric pairing and unequal exchange may have pushed such frequent copies to fixation, and maintained a selective pressure against loss of the copies. This may account for the differences in density per kilobase and frequency spectra of transposable elements in populations of mammals and *Drosophila*. Unfortunately, there is no information on the propensity of homozygous copies to undergo asymmetric exchange in *Drosophila*, not to mention in mammals (Maeda & Smithies, 1986). In yeast there appears to be little effect of heterozygosity or any other structural relationships on unequal exchange (Mikus & Petes, 1982; Roeder, 1983; Lichten, Borts & Haber, 1987). An understanding of the effect of homozygosity on unequal exchange will require further research.

This work was supported by the following grants to B.C.: BSR-8-16629 (NSF), GM 36405-01 (PHS), and a grant from the Louis Block fund of The University of Chicago.

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* **114**, 1165–1190.

Bellis, M., Jubier-Maurin, V., Dod, B., Vanlerberghe, F., Laurent, A.-M., Senglat, C., Bonhomme, F. & Roizès, G. (1987). Distribution of two recently inserted long interspersed elements of the L1 repetitive family at the Alb and βg3 loci in wild mice populations. *Molecular Biology and Evolution* **4**, 351–363.

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**, 330–332.
Brookfield, J. F. Y. (1986). Population biology of transposable elements. *Philosophical Transactions Royal Society London* B **312**, 217–226.
Charlesworth, B. (1985). The population genetics of transposable elements. In *Population Genetics and Molecular Evolution* (ed. T. Ohta and K. Aoki), pp. 213–232. Springer Verlag, Berlin.
Charlesworth, B. (1988). The maintenance of transposable elements in natural populations. In *Plant Transposable Elements* (ed. O. E. Nelson). Council for Research Planning, Washington D.C. (in press).
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.
Davis, P. S., Shen, M. W. & Judd, B. H. (1987). Asymmetrical pairing of transposons in and proximal to the white locus of *Drosophila* account for four classes of regularly occurring exchange products. *Proceedings of the National Academy of Science USA* **84**, 174–178.
Felsenstein, J. & Yokoyama, S. (1976). The evolutionary advantage of recombination. II. Individual selection for recombination. *Genetics* **83**, 845–859.
Finnegan, D. J. (1985). Transposable elements in eukaryotes. *International Review of Cytology* **93**, 281–326.
Goldberg, M. L., Sheen, J.-Y., Gehring, W. J. & Green, M. M. (1983). Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the *Drosophila melanogaster* genome. *Proceedings of the National Academy of Science USA* **80**, 5017–5021.
Golding, G. B., Aquadro, C. F. & Langley, C. H. (1986). Sequence evolution within populations under multiple types of mutation. *Proceedings of the National Academy of Science USA* **83**, 427–431.
Green, M. M. (1959). Spatial and functional properties of pseudo-alleles at the white locus in *Drosophila melanogaster*. *Heredity* **13**, 302–315.
Judd, B. H. (1959). Studies on some position pseudoalleles at the white locus region in *Drosophila melanogaster*. *Genetics* **44**, 34–42.
Kaplan, N. & Brookfield, J. F. Y. (1983). Transposable elements in Mendelian populations. III. Statistical results. *Genetics* **104**, 485–495.
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. A. & Quattlebaum, W. F. (1982). Restriction map variation in the Adh region of *Drosophila*. *Proceedings of the National Academy of Science USA* **79**, 5631–5635.
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 and E. Novitski), pp. 31–36. Orlando: Academic Press.
Lehrman, M. A., Goldstein, J. L., Russell, D. W. & Brown, M. S. (1987). Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell* **48**, 827–835.
Leigh Brown, A. J. (1983). Variation at the 87A heat-shock locus in *Drosophila melanogaster*. *Proceedings of the National Academy of Science USA* **80**, 5350–5354.
Leigh Brown, A. J. & Moss, J. E. (1987). Transposition of the I element and *copia* in a natural population of *Drosophila melanogaster*. *Genetical Research* **49**, 121–128.
Lichten, M., Borts, R. H. & Haber, J. E. (1987). Meiotic gene conversion and crossing over between dispersed homologous sequences occurs frequently in *Saccharomyces cerevisiae*. *Genetics* **115**, 233–246.
Lindsley, D. L. & Grell, E. H. (1968). *Genetic Variations of Drosophila melanogaster*. Washington: Carnegie Institute.
Lindsley, D. L. & Sandler, L. (1977). The genetic analysis of meiosis in female *Drosophila*. *Philosophical Transactions Royal Society London* B **277**, 295–312.
Lindsley, D. L. & Zimm, G. (1986). The genome of *Drosophila melanogaster*. Part 2. lethals; maps. *Drosophila Information Service* **64**, 1–158.
Lucchesi, J. C. & Suzuki, D. T. (1968). The interchromosomal control of recombination. *Annual Review of Genetics* **2**, 53–86.
Maeda, N. & Smithies, O. (1986). The evolution of multigene families: human haptoglobin genes. *Annual Review of Genetics* **20**, 81–108.
Meyerowitz, E. M. & Hogness, D. S. (1982). Molecular organization of a *Drosophila* puff site that responds to ecdysone. *Cell* **28**, 165–176.
Miklos, G. L. G., Healy, M. J., Pain, P., Howells, A. J. & Russell, R. J. (1984). Molecular and genetic studies on the euchromatin-heterochromatin transition region of the *X* chromosome of *Drosophila melanogaster*. *Chromosoma* **89**, 218–227.
Mikus, M. D. & Petes, T. D. (1982). Recombination between genes located on nonhomologous chromosomes in *Saccharomyces cerevisiae*. *Genetics* **101**, 473–483.
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.
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.
Roeder, G. S. (1983). Unequal crossing over between yeast transposable elements. *Molecular and General Genetics* **190**, 117–121.
Rubin, G. M. (1983). Dispersed repetitive DNAs in *Drosophila*. In *Mobile Genetic Elements* (ed. J. A. Shapiro), pp. 329–361. Academic Press, New York.
Rudkin, G. T. (1965). The relative mutability of DNA in regions of the *X* chromosome of *Drosophila melanogaster*. *Genetics* **52**, 665–681.
Scherer, G., Tschudi, C., Perera, J., Delias, H. & Pirrotta, V. (1982). B104, a new dispersed repeated gene family in *Drosophila melanogaster* and its analogies with retroviruses. *Journal of Molecular Biology* **157**, 435–451.
Schultz, J. & Redfield, H. (1951). Interchromosomal effects on crossing over in *Drosophila*. *Cold Spring Harbor Symposia on Quantitative Biology* **16**, 175–197.
Sturtevant, A. H. (1919). Contribution to the genetics of *Drosophila melanogaster*. III. Inherited linkage variations in the second chromosome. *Carnegie Institute of Washington Publications* **421**, 305–341.

Appendix

Correction to Montgomery et al. (1987)

Equation (1 *a*) of Montgomery *et al.* (1987) should be replaced with

$$\Delta H_X = -H_X(s_X + v) + \tfrac{2}{3}uP_{Xf}(H_X + H_A + \tfrac{1}{2}D_f) + \tfrac{1}{3}uP_{Xm}(H_X + 2H_A + D_m), \quad \text{(A 1)}$$

where H_X and H_A are the mean copy numbers of the element per haploid euchromatic genome for the X chromosome and autosomes respectively; D_f and D_m are the mean copy numbers of the element per diploid heterochromatic region for females and males respectively; P_{Xf} and P_{Xm} are the probabilities of insertion of a newly transposed element X chromosome in females and males respectively. This equation correctly takes into account the fact that there are two X chromosomes in a female to each one in a male, and that the total mean number of transposition events in a female is $u(2H_X + 2H_A + D_f)$ compared with $u(H_X + 2H_A + D_m)$ in a male.

Using this relation, the coefficients a_i of Montgomery *et al.* (1987) become:

$$a_1 = (P_{Xf} + \tfrac{1}{2}P_{Xm})/(P_{Xf} + P_{Xm}),$$

$$a_2 = P_{Xf}/2(P_{Xf} + P_{Xm}),$$

$$a_3 = P_{Xm}/2(P_{Xf} + P_{Xm}).$$

All other terms remain unchanged. Using these corrected coefficients, the analysis of Montgomery *et al.* (1987) yields the expected proportions of elements carried on the X chromosome as 0·199 for the null hypothesis of no selection against insertional mutations, and 0·133 on the hypothesis of selection against insertional mutations with a dominance coefficient of 0·35. Application of these expected frequencies to the data on the copy numbers of the elements *297*, *412* and *roo* reported by Montgomery *et al.* (1987) yields, as before, good fits of the null hypothesis for *297*, and *roo* (where the frequencies of the elements on the X are 0·20 and 0·19 respectively), but a significant disagreement ($\chi_1^2 = 12{\cdot}65$, $P < 0.001$) for *412* (where the frequency of elements on the X is 0·13). Conversely, the hypothesis of selection against insertional mutations yields significant disagreements for *297* and *roo* ($\chi_1^2 = 15{\cdot}65$ and 30·91 respectively), but gives a good fit for *412* ($\chi_1^2 = 0{\cdot}06$). The evidence for selection against insertional mutations for *412* is thus much stronger than was concluded by Montgomery *et al.* (1987), where the P value for the null hypothesis was only 0·05.

https://doi.org/10.1017/S0016672300027695 Published online by Cambridge University Press