Models of Repression of Transposition in P-M Hybrid Dysgenesis by P Cytotype and by Zygotically Encoded Repressor Proteins

J. F. Y. Brookfield

Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, England

Manuscript received January 10, 1990
Accepted for publication March 4, 1991

ABSTRACT

By analytical theory and computer simulation the expected evolutionary dynamics of *P* transposable element spread in an infinite population are investigated. The analysis is based on the assumption that, unlike transposable elements which move via RNA intermediates, the harmful effects of *P* elements arise primarily in the act of transposition, and that this causes their evolutionary dynamics to be unusual. It is suggested that a situation of transposition-selection balance will be superceded by the buildup of a cytoplasmically inherited repression or by the elimination of active transposase-encoding elements from the chromosomes, a process which may be accompanied by the evolution of elements which encode proteins which repress transposition.

TRANSPOSABLE elements are ubiquitous in the chromosomes of eukaryotic organisms. Their origins and functions, however, remain mysterious. *Drosophila melanogaster*, by virtue of its genetics and the possibility of following changes in transposable element position by *in situ* hybridization, provides a model system for studies of the mechanism and control of transposition. Many families of transposable elements in *D. melanogaster* have been described. Of these, the majority are of the *copia* or retrovirus-like structure (FINNEGAN 1985). These elements move at rates sufficiently high to ensure that there is very great variation in element position between laboratory strains and individuals from wild populations (MONTGOMERY and LANGLEY 1983, MONTGOMERY, CHARLESWORTH and LANGLEY 1987, LEIGH BROWN and MOSS 1987). Furthermore, many spontaneous mutations are known which result from *copia*-like element insertions. However, rates of movement of these sequences are still too low for factors affecting such rates to be profitably studied. A considerable body of theory concerning the evolutionary dynamics of transposable elements has now been generated (CHARLESWORTH and CHARLESWORTH 1983; LANGLEY, BROOKFIELD and KAPLAN 1983, KAPLAN and BROOKFIELD 1983a, b; BROOKFIELD 1982, 1986a, b; CHARLESWORTH 1985, 1988; CHARLESWORTH and LANGLEY 1986; MONTGOMERY, CHARLESWORTH and LANGLEY 1987; LANGLEY *et al.* 1988; OHTA 1984, 1985; SLATKIN 1985). From the most advanced work (CHARLESWORTH and LANGLEY 1986; MONTGOMERY, CHARLESWORTH and LANGLEY 1987; LANGLEY *et al.* 1988), a number of conclusions have emerged. First, a balance between transposition and selection can be achieved but only if the logarithm of the fitness of an individual with i copies of a transposable element declines more than linearly with i. This selection could arise through ectopic recombination with the transposable elements acting as scattered sites of homology in the chromosomes (MONTGOMERY, CHARLESWORTH and LANGLEY 1987). The frequencies of such events will increase with the square of transposable element copy number. There is now indirect experimental evidence for this conclusion in that the numbers of some transposable element families in *D. melanogaster* are elevated in regions of reduced recombination (LANGLEY *et al.* 1988). One interesting aspect is that the selection arises purely because these sequences are interspersed repetitive, and not because of any transposition process which they are undergoing. The second conclusion is that the conditions under which transposable elements can evolve lowered transposition rates are severely limited. A mutation which has a lowered transposition rate will miss out on some of the replicative process by which elements spread but will still experience the process by which elements are eliminated. Thus such a mutation would be expected to be removed from the population of transposable elements (CHARLESWORTH and LANGLEY 1986).

While many transposable elements are known, low rates of transposition make experimental studies of their evolutionary dynamics difficult. Higher rates of movement are seen in elements which undergo hybrid dysgenesis, of which the best studied is the *P* element. In the germline of the F_1 of crosses between males of strains bearing *P* elements and females from strains lacking such sequences (said to be of the "M cytotype") the elements move, probably replicatively (as judged by the net increase in copy number seen in the process). There is evidence that the maize transposable

element, *Activator*, which appears in some ways to be analogous to the *P* element, transposes in a conservative way (BAKER *et al.* 1986; SAEDLER and NEVERS 1985). However, a process which is conservative at the molecular level could nevertheless be replicative in an evolutionary sense, in that transposition could be, for example, from replicated to unreplicated DNA. For P, this transposition persists until a maternally inherited transposition repression system called P cytotype builds up. P cytotype is responsible for the relative lack of movement of *P* elements in crosses within P strains [see ENGELS (1983, 1986, 1989) for reviews].

Autonomous *P* elements ("*P* factors") differ from *copia*-like elements (retrotransposons) in a number of ways. First, they contain introns, which constitute strong evidence that the process of transposition does not go via a mechanism involving RNA. Second, in *D. melanogaster* at least, the population of P-related sequences include a majority of nonautonomous copies formed by internal deletions from the intact sequences. These two properties are not independent. The simplest way to explain the conservation of retrotransposons is to postulate a requirement for intact transposition genes in *cis* for transposition to take place. This requirement probably stems from a process in which a retrotransposon's RNA is packaged into a virus-like particle along with its own translation products. This ensures that, in order to transpose, a retrotransposon has to encode functioning proteins required for reverse transcription and reinsertion. This contrasts with the situation in some retroviruses, which apparently transpose by a very similar mechanism, but which can exist in deleted forms capable of being packaged and replicated by "helper" viruses acting in *trans* (VARMUS 1982). I will return below to reasons why retrotransposons do not seem to have evolved such a strategy. That P transposase$^-$ mutations can be complemented in *trans* may be a reflection of their DNA mode of transposition. A third difference is that *P* factors demonstrate a phenomenon called cytotype. This is the cytoplasmically inherited difference between P activities in strains lacking and possessing intact *P* elements. It is a fine point whether any comparable process affects the retrotransposons. The discovery of cytotype was through hybrid dysgenesis, which only arises when strains of different cytotype are crossed. *In situ* hybridization studies of P cytotype populations and experimental lines show that some P movements occur even in these conditions (AJIOKA and EANES 1989; RONSSERAY and ANXOLABEHERE 1987; EGGLESTON, JOHNSON-SCHLITZ and ENGELS 1988), so one could argue that the only reason why a system analogous to cytotype has not been discovered for retrotransposons such as *copia* is merely that all *D. melanogaster* are the equivalent of P cytotype for *copia*. The phenomenon of hybrid dysgenesis generated by the *P* factor may merely reflect that there are *D. melanogaster* strains which lack this sequence. The P family of transposable elements have apparently invaded *D. melanogaster* wild populations in this century (ANXOLABEHERE, KIDWELL and PERIQUET 1988), and now all wild populations contain many copies of *P* elements. However, despite this, while populations from the Americas have apparently evolved to P cytotype, populations from Europe and Asia are usually M′, possessing *P* elements but showing the M cytotype, or are Q, which fail to produce hybrid dysgenesis when crossed to either P or M strains (ANXOLABEHERE, KIDWELL and PERIQUET 1988; ANXOLABEHERE *et al.* 1985). The evolution of populations toward M′ or Q strains can be duplicated in experimental populations (ENGELS 1989; PERIQUET, RONSSERAY and HAMELIN 1989; PRESTON and ENGELS 1989).

In addition to regulation of P transposition, and its concomitant high temperature sterility, through P cytotype, there is some evidence that there can be zygotic repression, inherited equally through both sexes (KIDWELL 1985). This is probably itself the result of *P* transposable elements of a modified form. The vast majority of strains which show such chromosomally regulated repression have very many copies of a deleted element called the *KP* element (BLACK *et al.* 1987). This element has the capacity to encode a 207-amino acid polypeptide. It has been shown to be capable of causing repression of transposition and of spreading through experimental populations, apparently mimicking its spread through wild populations in Europe and Asia (JACKSON, BLACK and DOVER 1988). It is, however, not clear what is the cause of this rapid spread. Some authors interpret this as merely the consequence of the elements capacity to repress transposition, postulating that the element will spread by natural selection arising from its alleviation of the harmful effects of gonadal dysgenesis (BLACK *et al.* 1987). While it is clear that a restoration of full fertility will give an advantage to an element causing it, such a repressing transposable element will also miss out on the replicative transposition through which normal elements spread. There is a clear need for a quantitative approach to this problem.

A further complexity is introduced by the fact that the harmful effects of transposition are dependent, not only on a supply of transposase, but also on target P sequences which induce chromosome breaks (ENGELS *et al.* 1987). One test of P activity is the reversion of the hypermutable double P insertion in *singed*, sn^w. In this test, a phenomenon called transposase titration is seen, in which the rate of reversion of the mutant in hybrid dysgenesis declines with the increasing number of other *P* elements in the cell (SIMMONS and

BUCHHOLZ 1985). This is not seen with gonadal dysgenesis, which increases with the number of nonautonomous elements in the chromosomes, presumably because the chromosome breaks underlying sterility can be caused by breakage at *P* elements at any location, and which will thus increase uniformly with *P* element number.

I will consider the question of the regulation of *P* transposable element activities, and will suggest ways in which this regulation may differ from situations modeled previously. KAPLAN, DARDEN and LANGLEY (1985) modeled the evolutionary dynamics of a transposable element family, such as *P*, which can exist as intact and deleted forms, in a finite population. They assumed that element regulation was brought about by the balance between deletion and a transposition process which declined in rate with increasing copy number. They concentrated on expected times to stochastic loss of such sequences and drew two main conclusions which are relevant here. First, in a family in which deleted copies can be efficiently transposed through the action of *trans*-acting genes on intact copies, then, if the family is still extant, the majority of sequences are likely to be deleted copies. This contrasts with a sequence which requires all the genes in *cis* for transposition, which will have almost all copies intact. The second conclusion was that a sequence with the majority of copies deleted will have a shorter expected extinction time than one with the majority of copies intact. This model has some similarities to that which I will present below. The main difference is that it is a finite population size model, and a corresponding infinite population size model would allow the stable coexistence of both element types. In the infinite population size models that I outline below, extinction of element types occurs deterministically, driven by the forces of transposition and selection. I will suggest that retrotransposons and *P* elements may both be stabilized by a balance between replicative transposition and selection, but that whereas for retrotransposons the selective consequences result from the effects of the sequences as interspersed DNAs, for the *P* element the harmful consequences of the sequences arise from transposition itself. This dichotomy makes divergent predictions as to the direction of evolution of the sequences. It predicts that, for *P* elements, the balance between transposition and selection will be unstable with respect to likely evolutionary changes in the P sequences themselves. The result of this will be that either the *P* elements will evolve mechanisms of repression such as P cytotype which will be cytoplasmically inherited, or the transposable elements will be replaced by their own nonautonomous products.

I will consider initially a simple model in which I consider whether transposable elements with reduced transposition rates would be likely to spread (model A) and then consider more specific models of *P* elements in which known information is incorporated into a theoretical framework (model B). A further set of models (model C) will cast light upon why the P family, but not retrotransposons, will have nonautonomous copies.

MODEL A

This model considers the very simple question of whether a transposable element mutation lowering, in *cis*, the transposition rate, would be expected to spread through a population of transposable elements initially in a transposition-selection balance. It is a familiar result (CHARLESWORTH and CHARLESWORTH 1983) that, if transposition is replicative, a stable equilibrium copy number, resulting from a balance between such transposition and natural selection, can only be achieved if the fitness of an individual declines more than multiplicatively with copy number. Thus, for example, if the fitness of an individual with X copies of a transposable element is given by

$$w_X = w_1^{X^\alpha}$$

where w_1 is the Darwinian fitness of an individual with 1 copy of the transposable element relative to that of an individual with no copies ($w_1 < 1$), and α is some number greater than 1, then a stable copy number will be achieved. The next question concerns whether, given that such a transposition-selection balance has been achieved, a transposable element with a reduced transposition rate will spread. The answer is that it will not (CHARLESWORTH and LANGLEY 1986). However, this conclusion is based upon the assumption that any reduction in transposition has no concomitant effect on selection. In the P case, the evidence is that most of the harm done by the element to the host comes in transposition itself, and a transposable element mutation which lowered transposition rate in *cis* might correspondingly lower the effects of selection. CHARLESWORTH and LANGLEY did consider, with the *P* element in mind, the case in which transposition was accompanied by the formation of dominant lethal insertions and chromsome breaks, and found that under such a model the conditions for the spread of a repressing transposable element were much less stringent. However, they modeled the harmful effects of transposition as increasing linearly with transposition rate, and proposed that it was another form of selection which increased nonlinearly with copy number and thus stabilized the copy number. I will, however, model the situation by proposing that there is only a single selective force operating, and that it is the dominant lethal chromosome breaks induced by transposition which stabilize copy number. Clearly, P sequences are interspersed repeats, and will be subject

to any selection arising from this property, such as the consequences of ectopic recombination. However, I will suppose that the strengths of the selective effects of P sequences arising in transposition will be great enough to swamp their effects as interspersed repeats. Recent work on the fitness consequences of P insertions suggests that the harmful effects of hybrid dysgenesis do not primarily arise from insertions, but rather from chromosomal rearrangements (EANES *et al.* 1988). This makes the analysis different, with interesting and possibly counterintuitive results. We can consider a situation in which initially there is a transposable element present with a duplicative transposition rate of T, and with copy number variance in the population given by a Poisson distribution with a mean given by the mean copy number, and in which the fitness of an individual with X copies of the element is equal to

$$\exp(-s(TX)^2)$$

Thus each generation one can calculate the frequency distribution of numbers with X copies of the transposable element, then multiply this distribution by the corresponding fitness values and renormalize by dividing by mean fitness to give a postselection distribution. Then the mean copy number in the next generation is calculated by multiplying the mean copy number in the postselection distribution by $(1 + T)$. Thus, if the mean element number in generation 0 is $\bar{X}_0$, the initial proportion of individuals with exactly X copies, $P(X)$ is given by

$$P(X) = \frac{e^{-\bar{X}_0}\bar{X}_0^X}{X!}.$$

This distribution is multiplied by the fitness and then normalized by division by the mean fitness, to give a post-selection proportion with X copies, $P'(X)$, of

$$P'(X) = \frac{e^{-s(TX)^2}\,\bar{X}_0^X}{X!\sum_{i=0}^{\infty}\frac{\bar{X}_0^i e^{-s(Ti)^2}}{i!}}.$$

Now the mean X value in the next generation, $\bar{X}_1$, is given by

$$\bar{X}_1 = (1 + T)\sum_{X=0}^{\infty} P'(X).X.$$

By computer simulation it is easy to show that in an infinite population a transposable element initially introduced at low abundance will increase its numbers until a stable equilibrium is reached. We now postulate a mutation introducing a variant form of the transposable element with a new transposition rate, βT. The variant is introduced at low frequency into the population and from then on the number of individuals with X copies of the old element and Y of the new element is given by the product of two independent Poisson terms with means determined by the copy numbers of the two element types, $\bar{X}_0$ and $\bar{Y}_0$. Now the fitness of an individual with X and Y copies, respectively, of the two element types, is given by

$$\exp(-s(XT + Y\beta T)^2).$$

The fitnesses can be used to calculate postselection distributions, and the mean copy numbers of old and new elements in these distributions then multiplied by $(1 + T)$ and $(1 + \beta T)$ respectively. Thus the initial proportion of individuals with X and Y copies of the two element types $P(X,Y)$, is given by

$$P(X,Y) = \frac{e^{-\bar{X}_0-\bar{Y}_0}\bar{X}_0^X\bar{Y}_0^Y}{X!Y!}.$$

After selection the proportion $P'(X,Y)$ is equal to

$$P'(X,Y) = \frac{\bar{X}_0^X\bar{Y}_0^Y e^{-s(XT+Y\beta T)^2}}{X!Y!\sum_{i=0}^{\infty}\frac{\bar{X}_0^i}{i!}\sum_{j=0}^{\infty}\frac{\bar{Y}_0^j e^{-s(X_i+j\beta T)^2}}{j!}}$$

and the values of X and Y in the next generation are given by

$$\bar{X}_1 = (1 + T)\sum_{X=0}^{\infty} X\sum_{Y=0}^{\infty} P'(X,Y)$$

$$\bar{Y}_1 = (1 + \beta T)\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty} P'(X,Y)Y$$

Results: The interesting result common to the extensive computer simulations performed of this process is that the new element always invades and replaces the preexisting one if $\beta < 1$ and is itself eliminated if $\beta > 1$. In other words, the direction of evolution of transposition rate is downward. This is no help to the host, since the other general result of the simulations is that when a transposable element with a lowered transposition rate invades it moves to an equilibrium characterized by a higher copy number. Furthermore, at this new equilibrium host fitness is always less than that of the equilibrium prior to the invasion of the new variant. An illustration of this process is seen in Figure 1, in which the invasion of successive waves of transposable elements is charted.

This result suggests that a transposable element such as *P* which has its harmful effects in transposition may evolve in a fundamentally different way from other elements. It can reach a transposition-selection balance, but the balance is unstable to invasion by modified elements with reduced transposition activity. Such variants would seem, *a priori*, easy to produce by mutation of the element. Thus for elements of this kind, transposition-selection balances would seem likely to lead to cycles of reduced transposition rate, increased transposable element copy number, reduced

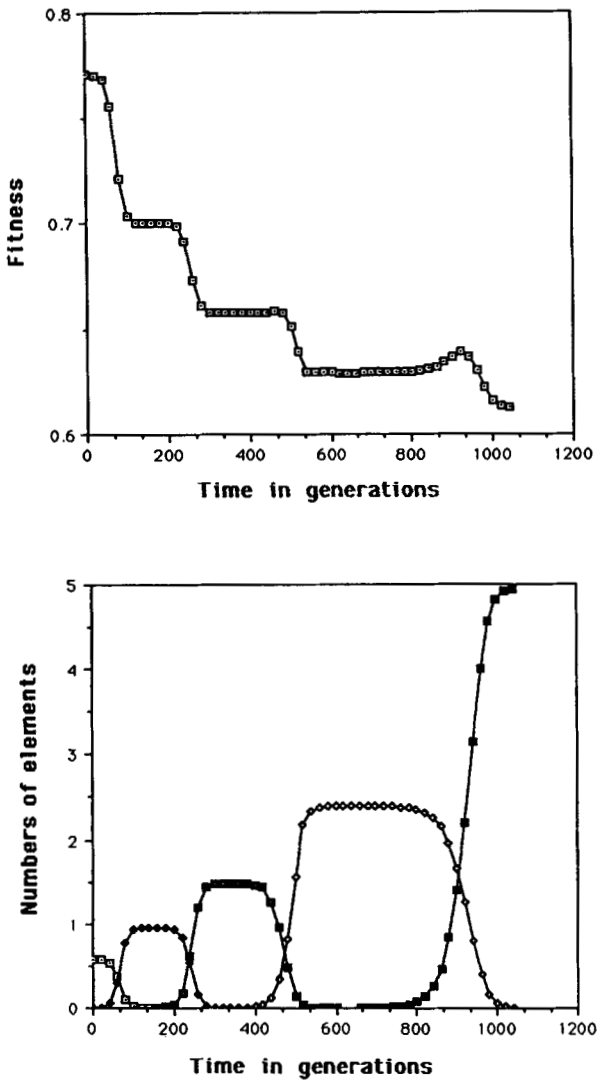

FIGURE 1.—Simulations of model A, in which transposable elements with reduced transposition rates invade a population of transposable elements initially in transposition-selection equilibrium. With $s = 0.25$, the population initially consists of elements with $T = 1$, which reach a transposition selection equilibrium with 0.5773 transposable element per individual. At generation 0 transposable elements with $\beta T = 0.8$ are introduced with initial abundance of 0.001 per individual. There are successive introductions of new less mobile transposable elements, each at 0.001, at generation 150 ($\beta T = 0.6$), generation 350 ($\beta T = 0.4$), generation 650 ($\beta T = 0.2$). The low transposition elements replace the faster transposers even though the mean population fitness is lowered as a result.

host fitness and possible extinction. Thus, the stable end points of transposable element evolution cannot be simply of this form but must show other features. There are two types of effects which have been postulated to occur which could take *P* elements away from their initial transposition-selection balance, but would not be accompanied by reduced host fitness. These are the generation of P cytotype, and the evolution of nonautonomous copies, some of which may have *trans*-acting repressor function. The evolution of *P* elements subject to these processes is investigated in model B below.

MODEL B

It is possible to devise a simple model incorporating at least some of the known features of P transposition and regulation. one can suppose that there are three types of elements: (i) intact elements encoding transposase—let the number of such elements in an individual be denoted by *X*, (ii) nonautonomous elements which encode a repressor protein, represented in number by *Y*, and (iii) nonautonomous elements which encode no proteins, represented by *Z*.

We can also suppose that in the germline there are transposase proteins with a cellular concentration of x and repressor proteins with a cellular concentration of y. The relationship between *X* and *Y* and x and y will be determined by the cytotype as shown below. I now make the simple assumption that the proteins bind stoichiometrically and reversibly to the DNA at equal rates for both proteins and all DNA elements, but only when the transposase binds will transposition result. Thus both the repressing and nonrepressing deleted elements titrate transposase but repressor proteins also repress by lowering the availability of DNA substrates for transposase. The equivalence of the DNAs and of the proteins allows one to usefully define A ($=X + Y + Z$), and B ($=x + y$) to represent the total concentrations of DNAs and P proteins in the cells as shown below.

The population: The elements are all hypothesized to be at sites at very low population frequencies, and to segregate freely in a randomly mating population. The consequence of this will be the proportion of individuals in the infinite population modeled with *X*, *Y*, and *Z* copies of the three element types will be given by the product of three independent Poisson terms with means given by the instantaneous mean numbers of the three element types in the population, $\bar{X}_0$, $\bar{Y}_0$ and $\bar{Z}_0$. Thus, in the model, all that is carried across from generation to generation are these three mean values (plus some information about the frequency distribution of cytotype, as outlined below). The population is assumed to be infinite in size, since I do not hypothesize that sampling drift plays a significant part in the changes which I am modeling.

Transposition: One can suppose that following binding a protein molecule remains on the DNA for, on average, h time units, irrespective of the nature of the protein and the DNA. I will therefore use simple second order kinetic arguments and suppose that at equilibrium in the cell the rate of binding of proteins to DNA will depend upon the product of the concentrations of unbound protein and unbound DNA. Thus if b represents the proportion of protein molecules

bound to the DNA the rate of binding will be proportional to

$$B(1 - b)(A - Bb)$$

$A - Bb$ represents the amount of DNA without bound protein since Bb (representing the concentration of bound protein molecules) will also represent, since we assume that single proteins bind to individual DNAs, the concentration of bound DNA molecules. If k represents the constant of proportionality then the rate of binding will be $Bk(1 - b)(A - Bb)$.

The rate of unbinding will be Bb/h, and, at equilibrium, will be equal to the rate of binding. Simple algebra gives the solution for b of

$$b = \frac{1 + kh(A + B) - \sqrt{(1 + kh(A + B))^2 - 4k^2h^2AB}}{2khB}. \quad (1)$$

The rate of transposition will be determined by the proportion of times the unbinding of protein is accompanied by transposition. It will thus be bx/h. As repression increases b should go down. This is indeed what is predicted by the above expression, since b will fall, all else being equal, as y rises. The value of b calculated will be a function of X, Y, Z and c, which is the parameter representing the strength of P cytotype in the individual. Thus in the recursion equations below b values will be expressed as $b_{X,Y,Z,c}$ values. Of these transpositions, the DNA sequences replicatively transposed will be proportional to their initial abundances, *i.e.*, to X, Y and Z. However, during transposition mutation occurs from intact to repressor and nonrepressor elements and from repressor to nonrepressor elements at rates of $u1$, $u2$ and $u3$, respectively. Thus if bx/h represents the total transposition then the numbers of copies of new elements of the various types will be

Intacts:

$$\frac{bxX(1 - u1 - u2)}{h(X + Y + Z)} \quad (2)$$

Repressors:

$$\frac{bx(Y(1 - u3) + Xu1)}{h(X + Y + Z)} \quad (3)$$

Nonrepressors:

$$\frac{bx(Z + Xu2 + Yu3)}{h(X + Y + Z)}. \quad (4)$$

Here the denominators are such that the sum of the transpositions remains bx/h.

There is evidence that, contrary to early speculations (O'HARE and RUBIN 1983) the deletions seen in *P* elements are not produced during transposition but rather arise as changes to the elements *in situ* in their chromosomal locations (ENGELS 1989). However, these changes will nevertheless require active transposase and will occur linearly in proportion to the available DNA substrates. Thus modeling mutation in this way will be acceptable. The results of the simulations detailed below do not indicate any processes driven by mutation pressure, and results are qualitatively unaffected by changes in the mutation rates.

The modeling of transposition above is based on a simple reformulation of second order reaction kinetics, and, obviously, the true nature of the process could be much more complicated. However, the qualitative robustness of some of the interesting features of the model, as revealed by the results with models A and C augers well for the possibility that the true details of the transposition mechanism may not much affect the qualitative results of model B.

Selection: The gonadal dysgenesis and embryo lethality seen in hybrid dysgenesis will constitute selection against individuals with high levels of transposition in the germline. One idea for how such selection might work would be to say that cells would be killed through translocations arising through the presence in a cell of two chromosome breaks arising from abortive transpositions. The fitness of an individual might vary with the proportion of its cells escaping such double breaks. ENGELS and PRESTON (1984) have evidence from the structure of inversions generated in hybrid dysgenesis that single transposition events do not simultaneously cause double chromosome breaks. The frequency of double breaks in the model is thus hypothesized to vary with the square of the transposition rate, and the constant of proportionality represented by s, corresponding to the strength of selection. Now the fitness of an individual can be represented by the 0th term of a Poisson distribution with mean

$$s\left(\frac{bx}{h}\right)^2,$$

or

$$\exp\left(-s\left(\frac{bx}{h}\right)^2\right).$$

The use of this formulation in which selection increases more rapidly than transposition with copy number will have the effect of generating a stable transposition-selection equilibrium in the absence of any repression. The consequences of relaxing this assumption for the nature of selection are explored below (models C). There is some evidence that much of the harmful effects of transposition result from single dominant lethal chromosome breaks which could vary linearly with transposition rate. However, a mixture of harmful effects from single and double breaks would be nonlinear overall.

Cytotype: I postulate that while all repressor pro-

teins are equivalent in their effects on transposition, there may exist not only proteins encoded by the repressor elements, but also a repressor protein derived from the incompletely spliced transcripts of the complete elements. I further postulate that these repressor proteins have two further properties. First, they have an effect on the splicing of the transcripts from intact elements, favoring the retention of the intron between open reading frames 2 and 3 and thus the production of more of themselves. Second, they are inherited in the egg nucleus and determine the splicing of zygotic transcripts. In these assumptions I follow SNYDER and DOOLITTLE (1988) and O'HARE (1985). One quantification of this is to suppose that all offspring can be characterized by a concentration, c, of cytotype-determining repressor, and the concentration of transposase molecules in the germline will be equal to

$$Xfe^{-c} \tag{5}$$

which represents the concentration of the translation products of those transcripts of X elements which fail to encounter any of the c repressor molecules and which are thus fully spliced. The term f (which will be slightly less than 1) represents the fidelity of splicing in the absence of cytotype-determining repressor. This has to be less than one in order to initiate the positive feedback process which is the cytotype shift. The prediction of this model is that the amount of cytotype-determining repressor produced in the germline is $X(1 - fe^{-c})$, and a proportion I of this is inherited into the next generation to be used as c in that generation. Thus the distribution of proportions of individuals with c in generation 1, $P(c)$, is calculated from the $P(X)$, $P(Y)$, $P(Z)$ and $P(c)$ values in generation 0 as shown below. The cytotype-determining repressor will also repress at the DNA level, in the sense that it will compete with transposase for DNA, and thereby lower transposition rate. Thus the amount of repressor protein, y, will be given by

$$Y + X(1 - fe^{-c}). \tag{6}$$

Clearly the model is far too complicated for an analytical solution and thus its consequences were explored by computer simulation. Each generation a population was defined by X, Y and Z values, and a frequency distribution of 10 cytotype classes, with c increasing in intervals of 2 from 0 (M cytotype) to 18 (P cytotype). The frequencies of individuals with all possible combinations of X, Y and Z and c values were calculated. These are represented by $P(X,Y,Z,c)$ values and are given by

$$P(X,Y,Z,c) = \frac{e^{-\bar{X}_0-\bar{Y}_0-\bar{Z}_0}\bar{X}_0^X\bar{Y}_0^Y\bar{Z}_0^Z P(c)}{X!Y!Z!}.$$

In such an individual the values for x and y are calculated using (5) and (6). Then the $b_{X,Y,Z,c}$ value is calculated from them using (1). The fitness of such an individual is equal to

$$w_{X,Y,Z,c} = \exp\left(-s\left(\frac{b_{X,Y,Z,c}Xfe^{-c}}{h}\right)^2\right) \tag{7}$$

and the mean population fitness, w, will be

$$w = \sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}\sum_{Z=0}^{\infty}\sum_{c=0}^{18} w_{X,Y,Z,c}P(X,Y,Z,c).$$

Now the new values of X_1, Y_1 and Z_1 are given by the combination of (2), (3), (4), and (7). They are

$$\bar{X}_1 = \frac{1}{w}\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}\sum_{Z=0}^{\infty}\sum_{c=0}^{18} P(X,Y,Z,c)w_{X,Y,Z,c} \cdot \left(X + \frac{b_{X,Y,Z,c}X^2fe^{-c}(1 - u1 - u2)}{h(X + Y + Z)}\right)$$

$$\bar{Y}_1 = \frac{1}{w}\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}\sum_{Z=0}^{\infty}\sum_{c=0}^{18} P(X,Y,Z,c)w_{X,Y,Z,c} \cdot \left(Y + \frac{b_{X,Y,Z,c}Xfe^{-c}(Y(1 - u3) + Xu1)}{h(X + Y + Z)}\right)$$

$$\bar{Z}_1 = \frac{1}{w}\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}\sum_{Z=0}^{\infty}\sum_{c=0}^{18} P(X,Y,Z,c)w_{X,Y,Z,c} \cdot \left(Z + \frac{b_{X,Y,Z,c}Xfe^{-c}(Z + Yu3 + Xu2)}{h(X + Y + Z)}\right).$$

The final part of the model is to define the values of $P(c)$ in the next generation. For each type of individual, as defined by X and c scores (Y and Z being irrelevant) there is a resulting value for the c value used in the next generation of

$$IX(1 - fe^{-c}). \tag{8}$$

The proportion in the next generation with a value of c,$P(c)$, is calculated by summing all the postselection proportions of individuals with the new c value plus or minus one.

The model ignores the expected weak correlations between the X, Y and Z values and cytotype, but these correlations will probably not be large, and will probably not effect the outcome of the process. The computer simulations of the model introduce some rounding down errors in the values of c, but these will be quantitatively of small effect and will not affect the qualitative outcome of the simulation. The next generation was normalized, as in the above recursion equations, by dividing by the population mean fitness, and the process repeated for a few hundred generations. The populations were initially started from a situation where $\bar{X} = 0.01$, $\bar{Y} = \bar{Z} = 0$.

Negative complementation: One alteration attempted was to hypothesize that the proteins act as

dimers. A repressor element is therefore doubly powerful in that a dimeric protein consisting of a polypeptide encoded by a deleted element in conjunction with one encoded by an intact element will function as a repressor. Thus now, if x and y represent, respectively, the concentrations of transposase and of repressor polypeptides, the concentration of transposase and repressor proteins are given by $x^2/(2(x + y))$ and $y(2x + y)/(2(x + y))$ respectively. It is, however, assumed for simplicity that the process of cytotype-determination is as before, with the splicing of the P transcripts determined by the concentration of the cytotype-determining repressor polypeptide, rather than by the concentration of dimeric proteins including this polypeptide.

This model might also be a useful approximation to a situation in which the repressor protein, while acting as a monomer, has a negative effect on transposase transcription as well as repressing by competing with the transposase for the DNA.

Results: The model makes a series of unjustified assumptions and includes a set of unmeasured and unmeasurable parameters. Its utility will depend upon the extent to which it can make predictions which are qualitatively robust to subtle changes in the underlying parameters and which fit with reality. Figures 2 to 6 give illustrations of the ways in which an experimental population evolves. It can be seen that in the situation in which P cytotype cannot evolve and in which there can be no nonautonomous elements produced, the copy number of the autonomous elements moves to the expected stable equilibrium between transposition and selection (Figure 2). Allowing the buildup of P cytotype quickly changes the situation. The population fitness increases to 1.0 (Figure 3). P cytotype is defined in the figure as the proportion of P transcripts which are incompletely spliced as a result of the interaction with cytotype-determining repressor. With these parameter values there is little effect on the situation if mutation to repressing and nonrepressing nonautonomous transposable elements is allowed. The P cytotype will build up before there is much increase in the various nonautonomous copies (Figure 4). If (by setting f to 1–Figures 5 and 6) P cytotype is prevented from rising, there is now an increase in the numbers of copies of repressing elements (and to a lesser degree of nonrepressors). This increase is always associated with a slow decrease to zero in the number of copies of autonomous elements. Repression apparently does favor the spread of nonautonomous elements in this process. However, as seen on Figure 6, if there are no repressor elements produced, there will still be the elimination of autonomous elements and their replacement with nonautonomous ones. The overproduction of repressors relative to nonrepressors is a result of their advantage

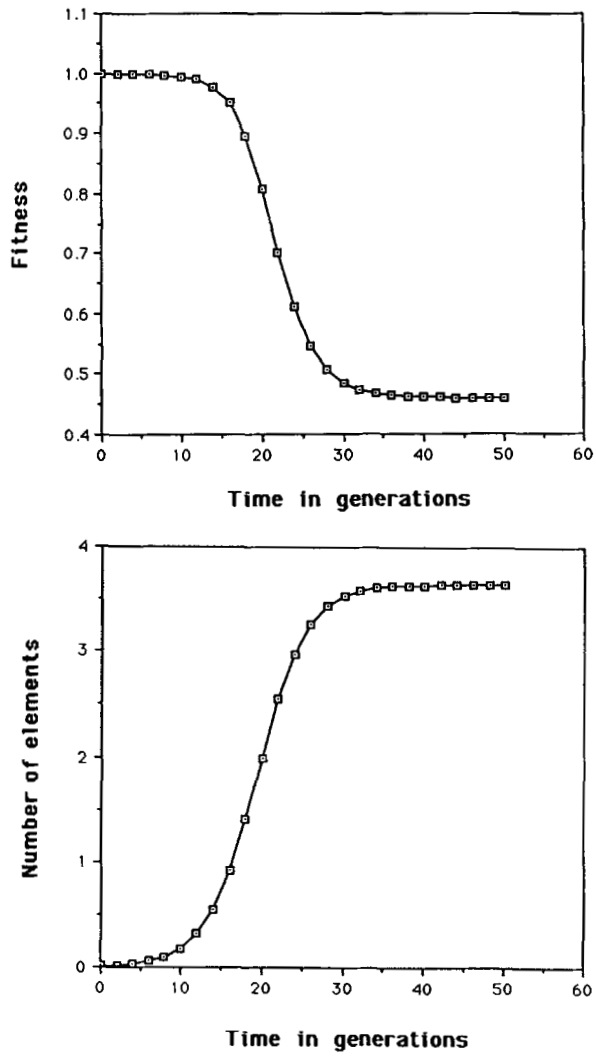

FIGURE 2.—Simulation of model B, with $f = I = k = h = 1$, $u1 = u2 = u3 = 0$, and $s = 0.2$. The population starts with $\bar{X}$ (complete elements per individual) = 0.01, and rises to a transposition-selection equilibrium at $\bar{X} = 3.62$. There is a corresponding diminution of the mean fitness.

when selection is acting. It may be unrealistic in that the process is here simulated with equal mutation rates to repressors and nonrepressors, whereas in reality it seems probable that only a minority of internal deletions will produce repressing elements. Indeed, if the simulation in Figure 4 is repeated, with the mutation rates to nonrepressors kept at 0.0001, but the mutation rate from autonomous copies to repressors lowered to 0.000001, there is replacement of the intact copies by a mixture of other element types dominated (90%) by nonrepressors. These mutation rates are unrealistically low for *P* elements, but I have made them so to show that there is a deterministic spread of nonautonomous elements as a result merely of the interplay of transposition and selection, and to prevent this effect from being conflated in the simulations with a strong mutation pressure.

Thus there seem to be two stable states in this

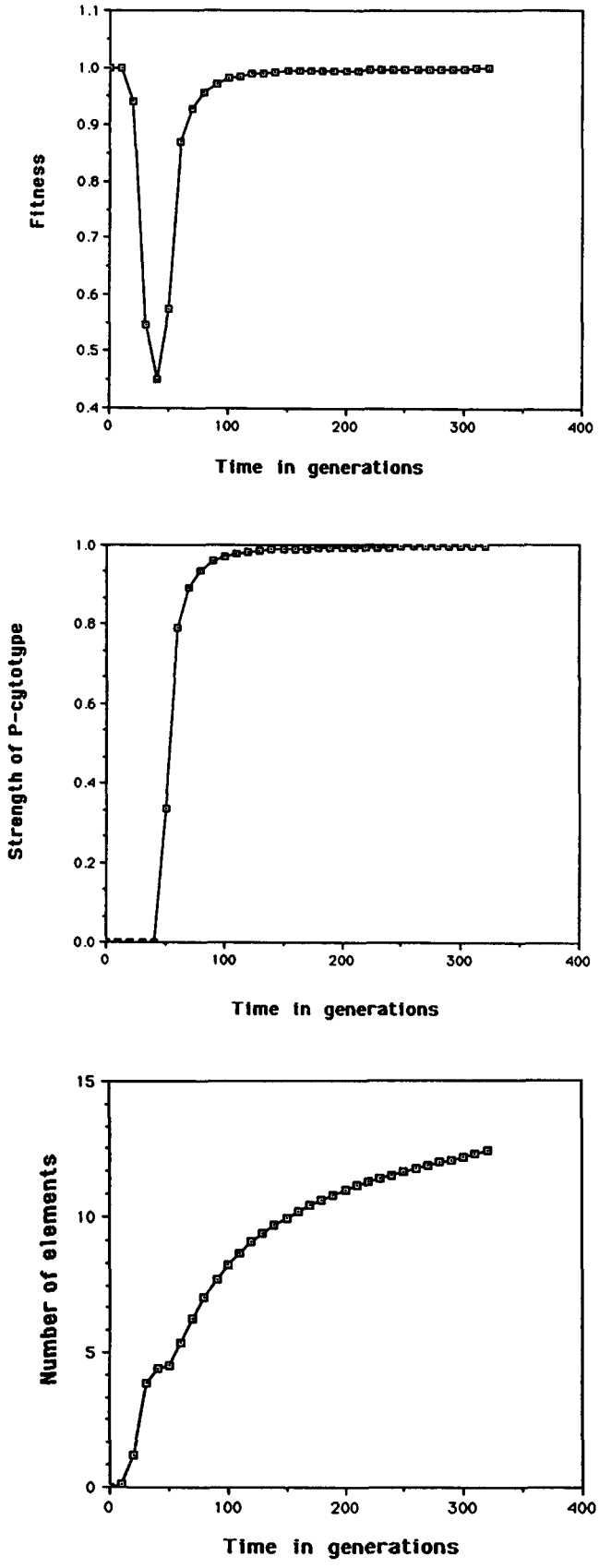

FIGURE 3.—Repetition of Figure 2, but with $f = 0.8$, so that P cytotype can build up. The cytotype shift at around generation 50 causes the restoration of approximately full fitness, and a gradual further increase in $\bar{X}$, which finally tops out at around 12. P cytotype

model. One is P cytotype, with large numbers of functional copies and some variable number of nonautonomous ones (with this number depending on the mutation rates and on history), and secondly, populations with no autonomous elements but rather with a mixture of repressing and nonrepressing nonautonomous elements (with this number again depending upon mutation rates and history). These stable states conform rather precisely with the strains collected from America and Europe. The question remains, however, as to why P cytotype does not always build up, as it does here unless prevented from doing so? One interesting and possibly important aspect of the above model for P cytotype is that it involves a positive feedback, in which the repressor protein interacts with the splicing process and favors the production of more of itself. Positive feedback systems have the property that they are structurally unstable—small changes in parameter values can give very divergent outcomes. In the simulation in Figure 4, the population was started with only intact elements, with a mean number of 0.01 per individual. It goes to P cytotype. At generation 100 in this simulation, when P cytotype has built up, there are 7.174 complete elements per individual on average, and also 1.588 repressors and 0.289 nonrepressor. We can now rerun the simulation but initially include repressors and nonrepressors in the same proportions, (*i.e.*, 0.01 complete, 0.00221 repressor, and 0.00040 nonrepressor). This would be a way of simulating the invasion of a large population of M strain flies with a small number of P strain flies, rather than simply invading with only intact *P* elements. As can be seen in Figure 7, the initial dynamics of spread of the *P* elements are similar but the higher number of repressors results in the reduction and loss of complete elements. P cytotype goes up a little, but then vanishes as the complete elements upon which it depends are eliminated. This process thus is consistent with the invasion of Europe by American P strains, their "dilution back" into M cytotype and the evolution toward M′ or Q strains in Europe. This represents, therefore, a situation in which it is subtle changes in historical factors which determine to which of the stable endpoints the population evolves. The structural instability is also consistent with studies of the buildup of P cytotype in artificial populations, in which some populations evolve to P cytotype but other populations, under identical conditions, become M′ (ENGELS 1989; DANIELS *et al.* 1987; ENGELS and PRESTON 1984). In the initial stages of the invasion of a pure M population with a P strain, the assumption of a Poisson variation in copy number will be nowhere near valid, since the elements will be concentrated in

is here defined as the proportion of transposase transcripts which are incompletely spliced as a result of interaction with cytotype-determining repressors.

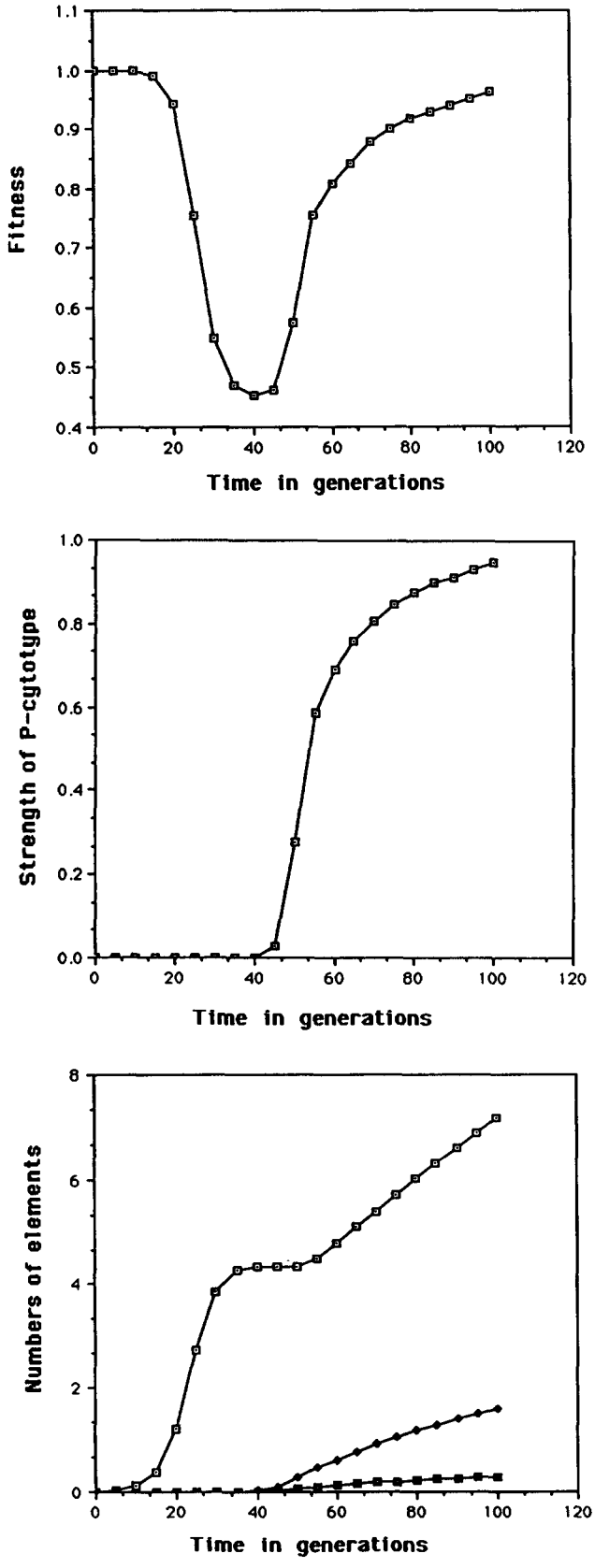

FIGURE 4.—Repetition of Figure 3, but with mutation to repressors (♦) and nonrepressors (■) now allowed ($u1 = u2 = u3 = 0.0001$). At the time of the cytotype shift there has been little buildup of nonautonomous elements, which subsequently lack the selective force required to increase their numbers rapidly.

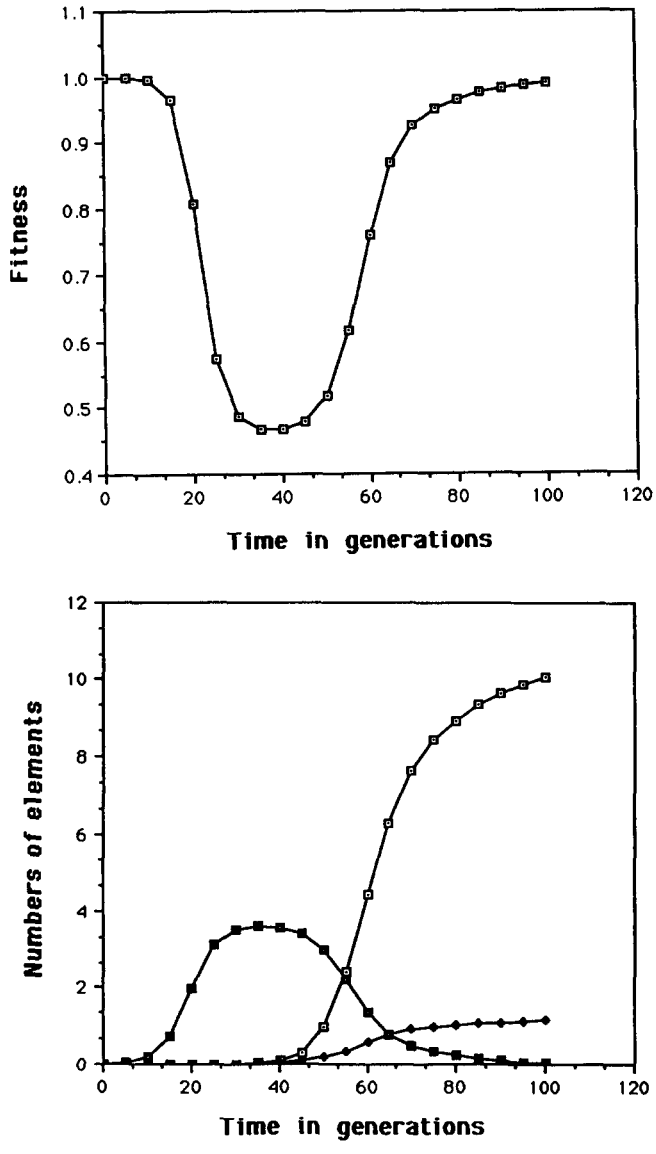

FIGURE 5.—Repetition of Figure 4, but with the cytotype shift prevented ($f = 1$). Now the transposition-selection equilibrium is destroyed by the generation of large numbers of nonautonomous elements (particularly repressors (⊡)), and by the selective elimination of intact copies (□). Fitness is thus restored to 100%.

a few individuals. However, it seems likely that, in such a situation, the strong selection against the offspring suffering hybrid dysgenesis will make the invasion of the M population unlikely. Probably the condition for the successful invasion of the host population with *P* elements will be that the early strong dysgenesis will be succeeded by a phase in which the *P* elements have become scattered at low copy numbers across a number of individuals, and thus the process subsequent to this point can be approximated by the Poisson assumptions.

There are still three extra considerations. One concerns the *KP* element. The simulations clearly reveal that repressing elements will spread more rapidly than nonrepressing elements (once a transposition-selec-

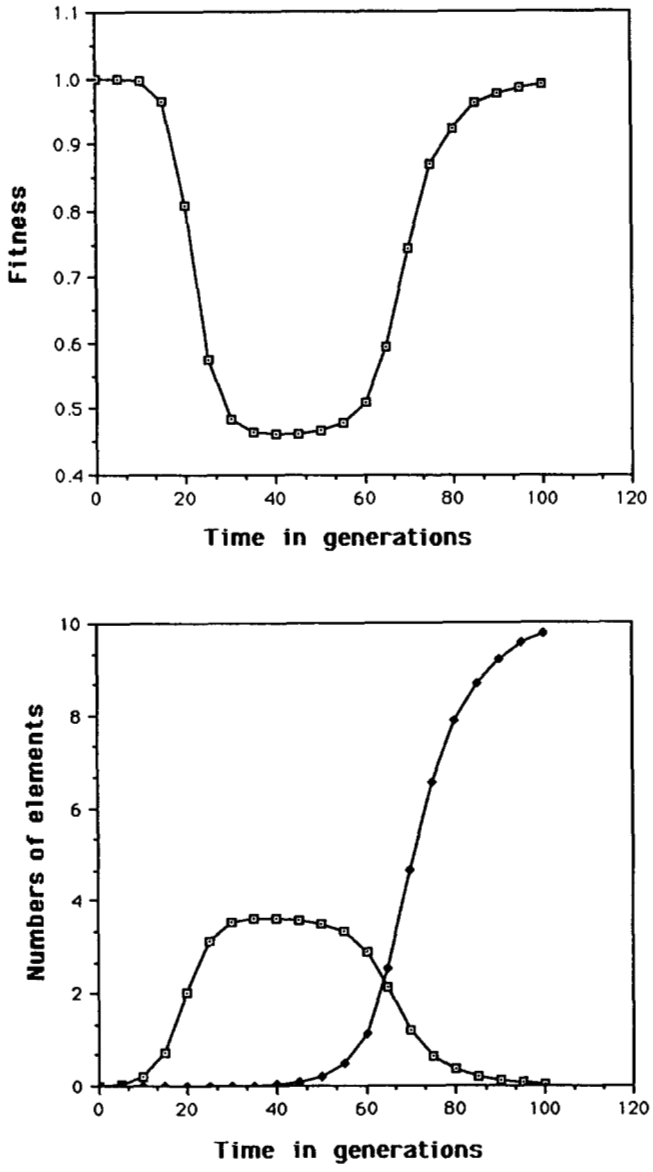

FIGURE 6.—Repetition of Figure 5, but with $u1 = u3 = 0$, so that no repressing elements are created. Now, more slowly, it is the inert elements (♦) which increase greatly in numbers and replace the intact elements.

tion equilibrium has been reached) but this process is never strong enough to be consistent with the rapid spread of *KP* elements seen in some experimental populations (JACKSON, BLACK and DOVER 1988). The model in which the repressors showed negative complementation with the transposase gene was investigated to see if the rate of spread of repressors could now be fitted to the KP situation. The repressors do seem to enjoy an increased advantage in this model. They are capable of spreading, albeit slowly, in conditions of P cytotype and reach a very high number. However, when one remembers that there are mutations to repressors at an equal frequency in these simulations to mutations to inert elements, the spread of the *KP* element still seems too rapid to be accounted for merely by its repressing qualities. It would seem to be most likely that this element is a favored substrate for transposition, and spreads preferentially by this mechanism. If so, its repressing properties, copy for copy, may not be exceptional.

Of course, such an explanation of KP spread is quite general. One could explain the data for any spread of deleted copies by the simpler alternative explanation in which the transposability of a transposable element increases as its size goes down. Evidence for such a phenomenon does exist for *P* elements (SPRADLING 1986). However, such a phenomenon may not be a general result for other types of transposable element. Furthermore, even if there is a transposition advantage for deleted elements, this, in itself, does not predict that they will eliminate intact elements unless selection acts in a manner dependent upon transposition, as shown below in the results for model C.

Second, the buildup of P cytotype is here presented merely as a fact, and it is not investigated whether the mode in which selection acts upon *P* factors will itself promote the evolution of a system of cytotype. It has not been investigated whether a mutation in an element causing it to generate a maternally inherited repression would cause such an element to spread. This is because the assumption that cytotype and the various element copy numbers are uncorrelated will destroy any mechanism whereby cytotype-determining elements might spread, which must require such a correlation. There is a Bayesian argument which says that since (according to model A) *P* elements must evolve some system of repression if their host species are to persist, then were the complete *P* elements to persist in a host (and it must not be forgotten that *D. melanogaster* is not the normal host of the *P* element (STACEY *et al.* 1986)) in which they were subject to an inefficient process of the generation of nonautonomous elements, then some cytotype form of repression must evolve. This argument is rather weak, since it would seem almost inevitable that mutations could occur destroying the *trans*-acting functions of the element. More work clearly needs to be done on this area.

Third, there is the following difficulty. I have argued that, in the absence of cytotype, autonomous *P* elements should be replaced by their nonautonomous products, and I have postulated that this result arises from the fact that the harmful effects of P sequences occur in transposition. This fits with the data on *P* elements, but contrasts with the retrotransposons, which are all autonomous. I explain this through an argument based on the different way in which selection acts on the retrotransposons. However, one could equally argue that the retrotransposons do not get replaced by their deletion products because they simply cannot make deletion products capable of being

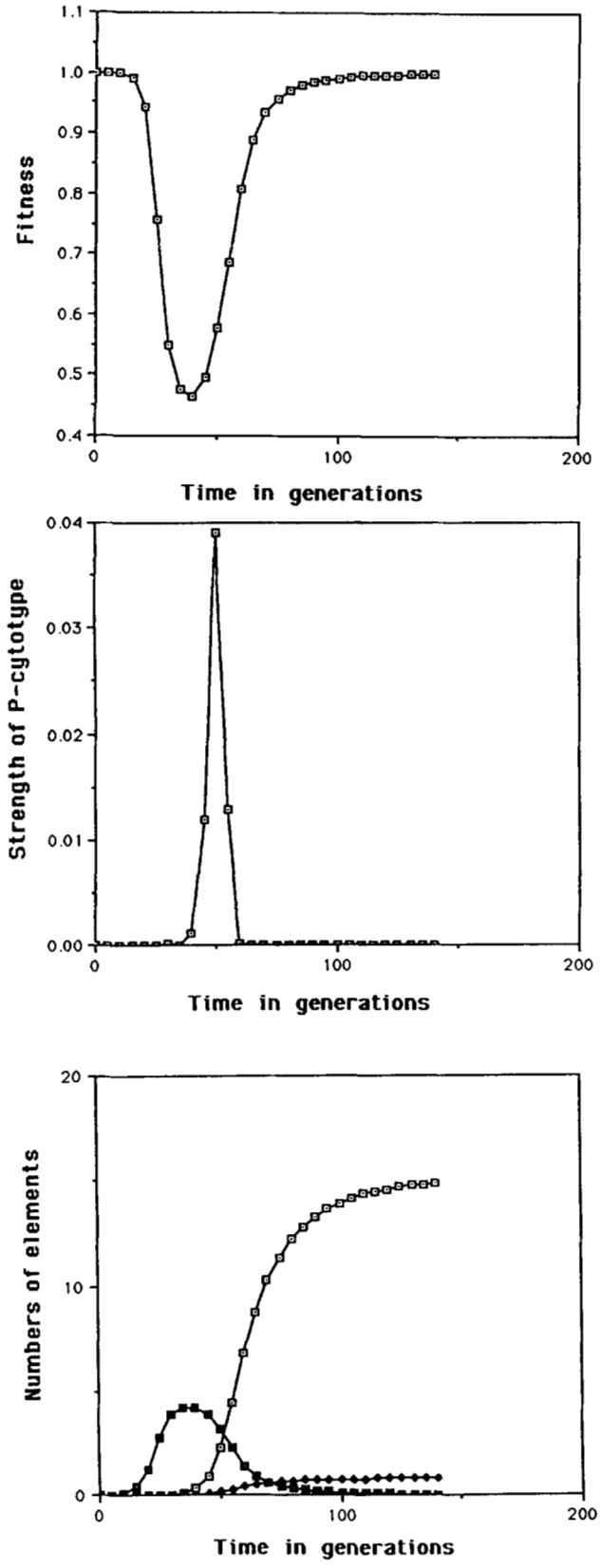

FIGURE 7.—A simulation of model B with the same state parameter values as in Figure 4, but one in which the initial population has $\bar{X} = 0.01$, $\bar{Y} = 0.00221$ and $\bar{Z} = 0.00040$. These small numbers of repressors, in particular, are sufficient to cause the reduction of complete elements such that P cytotype does not build up, but rather the intact copies are eliminated. Comparison of Figures 4 and 7 reveals that the stable outcome reached in this model is sensitive to very subtle changes in initial parameter values.

complemented in *trans* by intact copies. However, there is evidence from helper retroviruses which says that such complementation can occur. Thus in a final set of models (model C) I demonstrate that deleted but still complementable retrotransposons will not be capable of spreading through the population, but complementable deleted P sequences will spread.

MODEL C

Here I use a very simple model system in which I hypothesize a population defined by copy numbers of two different types of elements, intacts and deleteds, with the numbers of copies of each type being defined by independent Poisson distributions. Four models can be defined, differing in the ways in which selection and transposition act. In any individual let X represent the number of intact copies and Y the number of deleted copies. For each model we can define ΔX, the change in X resulting from transposition, and ΔY, which is the change in Y resulting from transposition. We can also define w, which is the fitness of an individual.

These quantities are functions of three constants, T, representing the transposition rate, u representing the mutation rate from intacts to deleteds, and s, representing the strength of selection. Two models will be based on the assumption about P sequences, that changes in transposition have consequential effects on the strength of selection. These models, P1 and P2, differ in the particular form of the relationship between transposition and selection. Two other models, C1 and C2, are inspired by the *copia*-like sequences, or retrotransposons. Thus for P model 1 (P1):

$$\Delta X_{X,Y} = \frac{TX^2(1-u)}{X+Y}$$

$$\Delta Y_{X,Y} = \frac{TX(Y+Xu)}{X+Y}$$

$$w = \exp(-s(TX)^2).$$

For P model 2 (P2):

$$\Delta X_{X,Y} = TX^2(1-u)$$

$$\Delta Y_{X,Y} = TX(Y+Xu)$$

$$w = \exp(-s(TX(X+Y))^2).$$

For *copia* model 1 (C1):

$$\Delta X_{X,Y} = \frac{TX^2(1-u)}{X+Y}$$

$$\Delta Y_{X,Y} = \frac{TX(Y+Xu)}{X+Y}$$

$$w = \exp(-s(X+Y)^2).$$

For *copia* model 2 (C2):

$$\Delta X_{X,Y} = TX^2(1 - u)$$

$$\Delta Y_{X,Y} = TX(Y + Xu)$$

$$w = \exp(-s(X + Y)^2).$$

The $\Delta X_{X,Y}$ and $\Delta Y_{X,Y}$ terms show that these are functions of the X and Y values of the individual. Thus the P and *copia* models differ in the involvement in the P models of the transposition constant, T, in selection. The models also differ in whether the transposition rate depends linearly or quadratically upon copy number (in the model B above, the interaction between P DNA and P protein determining transposition rate has the effect of making transposition increase with the square of copy number). The transposition functions for C1 and P1 are identical, as are those for C2 and P2. In all models, the initial proportion of individuals with exactly X and Y copies of the two types is given by

$$P(X,Y) = \frac{e^{-\bar{X}_0-\bar{Y}_0}\,\bar{X}_0^X\bar{Y}_0^Y}{X!Y!}$$

in which X_0 and Y_0 are the initial mean numbers of the two element types. The postselection distributions are thus,

For P1

$$P'(X,Y) = \frac{\bar{X}_0^X\bar{Y}_0^Y e^{-s(TX)^2}}{X!Y!\sum_{i=0}^{\infty}\frac{\bar{X}_0^i}{i!}\sum_{j=0}^{\infty}\frac{\bar{Y}_0^j e^{-s(Ti)^2}}{j!}}.$$

For P2

$$P'(X,Y) = \frac{\bar{X}_0^X\bar{Y}_0^Y e^{-s(TX(X+Y))^2}}{X!Y!\sum_{i=0}^{\infty}\frac{\bar{X}_0^i}{i!}\sum_{j=0}^{\infty}\frac{\bar{Y}_0^j e^{-s(Ti(i+j))^2}}{j!}}.$$

For C1

$$P'(X,Y) = \frac{\bar{X}_0^X\bar{Y}_0^Y e^{-s(X+Y)^2}}{X!Y!\sum_{i=0}^{\infty}\frac{\bar{X}_0^i}{i!}\sum_{j=0}^{\infty}\frac{\bar{Y}_0^j e^{-s(i+j)^2}}{j!}}.$$

For C2

$$P'(X,Y) = \frac{\bar{X}_0^X\bar{Y}_0^Y e^{-s(X+Y)^2}}{X!Y!\sum_{i=0}^{\infty}\frac{\bar{X}_0^i}{i!}\sum_{j=0}^{\infty}\frac{\bar{Y}_0^j e^{-s(i+j)^2}}{j!}}.$$

The values for $\bar{X}_1$ and $\bar{Y}_1$ are given, for all models, by

$$\bar{X}_1 = \sum_{X=0}^{\infty}\sum_{Y=0}^{\infty} P'(X,Y)(X + \Delta X_{X,Y})$$

$$\bar{Y}_1 = \sum_{X=0}^{\infty}\sum_{Y=0}^{\infty} P'(X,Y)(Y + \Delta Y_{X,Y}).$$

Results: It is an almost entirely general result that deleted copies invade the population and replace the intact copies in the two P models, but cannot invade in the two *copia* models. An illustration of this is shown in Figures 8 and 9, in which the four models are applied with $T = 0.5$, and the s values adjusted so that with intact copies only the population equilibrium is at 2.475 copies per individual. In both P cases the deleteds invade. They can not in the *copia* models. In the *copia* models intact sequences can invade and replace the deleted sequences. Thus it is the differences in the way selection acts which determine the

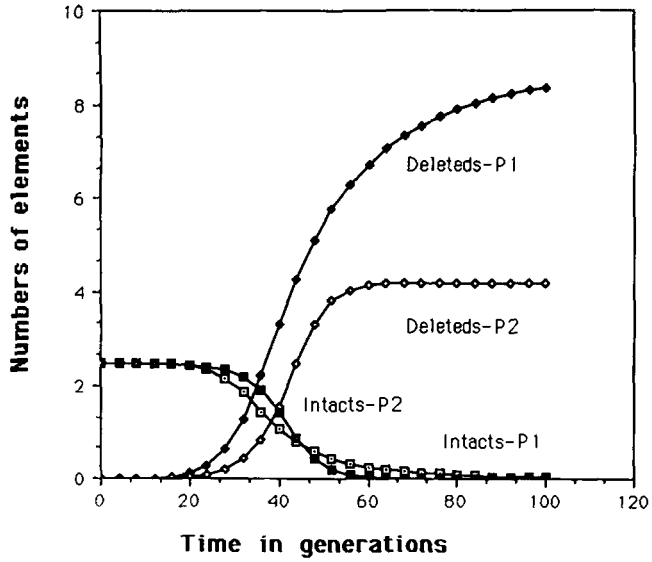

FIGURE 8.—Simulations of models P1 (with $T = 0.5$, $s = 0.4$) and P2 (with $T = 0.5$, $s = 0.1$). Each population was started with the equilibrium number of 2.475 intact copies per individual, to which were added deleted copies at an abundance of 0.001 copy per individual. The deleted copies rapidly replace the intact copies in each model.

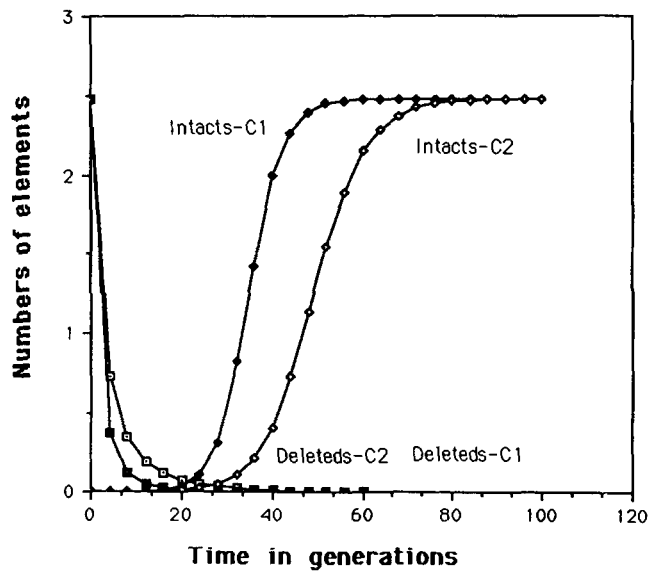

FIGURE 9.—Simulations of models C1 (with $T = 0.5$, $s = 0.1$) and C2 (with $T = 0.5$, $s = 0.213$). Each population was started with 2.475 copies of the deleted element, which were rapidly eliminated by selection. The complete elements, introduced at 0.001 copy per individual, spread to their equilibrium number.

outcomes of the models, not quantitative differences in the transposition process. (For P model 2, there is a small window of parameter space, corresponding to high T and s values, which does not allow invasion of deleted copies. The T and s values are so high that the equilibrium number of complete copies is well below 1. No such window exists for P model 1, in which deleteds will always invade.)

The question as to the effect of an advantage in transposition for deleted copies can be addressed by simulation using C1 and C2. The increases due to transposition in the numbers of copies of deleted elements can be replaced by,

for C1:

$$\Delta Y_{X,Y} = \frac{TX(\gamma Y + Xu)}{X + Y}$$

and for C2:

$$\Delta Y_{X,Y} = TX(\gamma Y + Xu)$$

where γ is the advantage for the deleted copies (thus $\gamma = 1$, is where there is no advantage).

Simulations of these modified models with the same initial conditions as above ($T = 0.5$ and $s = 0.1$ for C1 and $s = 0.213$ for C2) showed that, for model C1, in order for a deleted copy to invade from an initial low frequency γ has to be greater than about 2.1. With higher γ values, the two types are maintained in a stable equilibrium. Even with γ as high as 10, however, at equilibrium 19.6% of the elements are intact. For model C2, again the deleted copies can increase their number when introduced at low frequency, provided that γ is about 2.1 or larger. They again come to a stable equilibrium, and, even when $\gamma = 10$, 24.9% of the elements are intact. Thus, even if deleted transposable elements have a transposition advantage, they will not replace intact copies if selection does not act through transposition.

The general conclusion thus is that while it is an observational fact that retrotransposons exist in autonomous forms, this fact itself can be potentially explained by a consideration of the way in which selection acts on these sequences.

DISCUSSION

The calculations outlined here suggest that it is possible to divide transposable elements into two classes on the basis of the way in which selection acts upon them. Those which have their effects in transposition itself, such as the *P* element in hybrid dysgenesis, will move toward a system of maternally inherited repression or will be replaced by their own deletion products. Other sequences, the main harmful effects of which are as sites of ectopic recombination, will remain intact and stable constituents of the genome. The former class of sequences will move to a situation in which there is little if any transposition going on. They will now be subject only to the weak selective force of their being sites of homology and thus there will be selection upon their copy number. Such selection will tend to reduce their number. However, the process will be slow. As the force of transposition comes out of the system the site occupancies will drift to 0 and 1 and selection will not find much copy number variance on which to act. The only source of such variation will be excisions of copies by recombinations between the short direct target site repeats at their termini. Such recombinations will be rare in heterochromatic regions, such as the chromocenter. It is thus interesting that the *P* elements in *Drosophila nebulosa*, which are nonautonomous as a result of multiple nonsense substitutions and microdeletions, are located in the chromocenter (LANSMAN *et al.* 1987). The prediction is thus that elements of this type will be rapidly lost from the euchromatin, but may leave behind deleted copies in the chromocenter.

The I factor seems to form an intermediate between the classes considered so far (FINNEGAN 1988). This produces the inducer cytotype, but apparently moves via an RNA intermediate. Its conservation is consistent with its RNA mode of transposition. However, in hybrid dysgenesis the harmful effects of the element are correlated with its enhanced transposition rate. One might thus suppose, following the modeling presented here, that in the long run deleted copies might spread. The differences between P and I may represent differences in the rates at which mutations can produce nonautonomous copies which are still capable of being complemented by intact copies. Nonautonomous copies of I are found at the chromocenter in both I and R strains (SIMONELIG *et al.* 1988).

Two forces which have not been dealt with in these models are excision and selection of a nonstabilizing kind arising from harmful DNA insertions. If excisions of transposable elements occur constantly, and independently of the transposition rate, and affect deleted and intact copies equally, then such excisions will, qualitatively, make it harder for deleted copies to spread. There will be a certain level of transposition required just to overcome excision and the repressing elements or nonrepressing nonautonomous elements may not achieve it. However, most of the evidence indicates that the excision of elements requires transposase and is proportional to transposition, and thus can be accurately represented merely by a reduction in the net transposition rate, at least in a model of this kind with infinitesimal site frequencies. Selection of a nonstabilizing kind can also be represented as a reduction in the transposition rate. An insertion which makes a harmful mutation will be rapidly eliminated by natural selection, when compared to transposable

elements in "safe" locations. It can be viewed as an abortive transposition.

A final simplification of the modeling here presented is that it fails to include a fourth type of element which can be envisaged, the cytotype-determining repressor. This would repress and also feedback with the splicing system to cause P cytotype. There is some evidence that there can be such elements (ENGELS 1989; SIMMONS *et al.* 1987). Modeling the dynamics of elements such as these will require more sophisticated treatments, as alluded to above, which allow correlations between generations in the elements' distributions and in cytotype.

I thank BRIAN CHARLESWORTH and KATE EXLEY for numerous helpful comments on an earlier draft of this manuscript.

LITERATURE CITED

AJOIKA, J. W., and W. F. EANES, 1989 The accumulation of P-elements on the tip of the X chromosome in populations of *Drosophila melanogaster*. Genet. Res. **53:** 1–6.

ANXOLABEHERE, D., M. G. KIDWELL and G. PERIQUET, 1988 Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile P elements. Mol. Biol. Evol. **5:** 252–269.

ANXOLABEHERE, D., D. NOUAUD, G. PERIQUET and P. TCHEN, 1985 P element distribution in Eurasian populations of *Drosophila melanogaster*: a genetic and molecular analysis. Proc. Natl. Acad. Sci. USA **82:** 5418–5422.

BAKER, B., J. SCHELL, H. LORZ and N. FEDEROFF, 1986 Transposition of the maize controlling element "Activator" in tobacco. Proc. Natl. Acad. Sci. USA **83:** 4844–4848.

BLACK, D. M., M. S. JACKSON, M. G. KIDWELL and G. A. DOVER, 1987 KP elements repress P induced hybrid dysgenesis in *D. melanogaster*. EMBO J. **6:** 4125–4135.

BROOKFIELD, J. F. Y., 1982 Interspersed repetitive DNA sequences are unlikely to be parasitic. J. Theor. Biol. **94:** 281–299.

BROOKFIELD, J. F. Y., 1986a A model for DNA sequence evolution within transposable element families. Genetics **112:** 393–407.

BROOKFIELD, J. F. Y., 1986b The population biology of transposable elements. Phil. Trans. R. Soc. Lond. B **312:** 217–226.

CHARLESWORTH, B., 1985 The population genetics of transposable elements, pp. 213–232 in *Population Genetics and Molecular Evolution*, edited by T. OHTA and K. AOKI. Springer-Verlag, Berlin.

CHARLESWORTH, B., 1988 The maintenance of transposable elements in natural populations, in *Plant Transposable Elements*, edited by O. E. NELSON. Council for Research Planning, Washington D.C.

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

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

DANIELS, S. B., S. H. CLARK, M. G. KIDWELL and A. CHOVNICK, 1987 Genetic transformation of *Drosophila melanogaster* with an autonomous P element: phenotypic and molecular analysis of long-established transformed lines. Genetics **115:** 711–723.

EANES, W. F., C. WESLEY, J. HEY, D. HOULE and J. W. AJOIKA, 1988 The fitness consequences of P element insertion in *Drosophila melanogaster*. Genetics **52:** 17–26.

EGGLESTON, W. B., D. M. JOHNSON-SCHLITZ and W. R. ENGELS, 1988 P-M hybrid dysgenesis does not mobilize other transposable element families in *Drosophila melanogaster*. Nature **331:** 368–370.

ENGELS, W. R., 1983 The P family of transposable elements in Drosophila. Annu. Rev. Genet. **17:** 315–344.

ENGELS, W. R., 1986 On the evolution and population genetics of hybrid-dysgenesis-causing transposable elements in Drosophila. Phil. Trans. Roy. Soc. Lond. B. **312:** 205–215.

ENGELS, W. R., 1989 P elements in Drosophila, in *Mobile DNA*, edited by D. BERG and M. HOWE. ASM Publications, Washington, D.C.

ENGELS, W. R., and C. R. PRESTON, 1984 Formation of chromosome rearrangements by P factors in Drosophila. Genetics **107:** 657–678.

ENGELS, W. R., W. K. BENZ, C. R. PRESTON, P. L. GRAHAM, R. W. PHILLIS and H. M. ROBERTSON, 1987 Somatic effects of P element activity in *Drosophila melanogaster*: pupal lethality. Genetics **117:** 745–757.

FINNEGAN, D. J., 1985 Transposable elements in eukaryotes. Int. Rev. Cytol. **93:** 281–326.

FINNEGAN, D. J., 1988 I factors in *Drosophila melanogaster* and similar elements in other eukaryotes. Symp. Soc. Gen. Microbiol. **43:** 271–285.

JACKSON, M. S., D. M. BLACK and G. A. DOVER, 1988 Amplification of KP elements associated with the repression of hybrid dysgenesis in *Drosophila melanogaster*. Genetics **120:** 1003–1013.

KAPLAN, N. L., and J. F. Y. BROOKFIELD, 1983a Transposable elements in Mendelian populations. III. Statistical results. Genetics **104:** 485–495.

KAPLAN, N. L., and J. F. Y. BROOKFIELD, 1983b The effect of selective differences between sites of transposable elements on homozygosity. Theor. Popul. Biol. **23:** 273–280.

KAPLAN, N., T. DARDEN and C. H. LANGLEY, 1985 Evolution and extinction of transposable elements in Mendelian populations. Genetics **109:** 459–480.

KIDWELL, M. G., 1985 Hybrid dysgenesis in *Drosophila melanogaster*: nature and inheritance of P element regulation. Genetics **111:** 337–350.

LANGLEY, C. H., J. F. Y. BROOKFIELD and N. L. KAPLAN, 1983 Transposable elements in Mendelian populations. I. A theory. Genetics **104:** 457–472.

LANGLEY, C. H., E. A. MONTGOMERY, R. HUDSON, N. KAPLAN and B. CHARLESWORTH, 1988 On the role of unequal exchange in the containment of transposable element copy number. Genet. Res. **52:** 223–236.

LANSMAN, R. A., R. O. SHADE, T. A. GRIGLIATTI and H. W. BROCK, 1987 Evolution of P transposable elements: sequences of *Drosophila nebulosa* P elements. Proc. Natl. Acad. Sci. USA **84:** 6491–6495.

LEIGH BROWN, A. J., and J. E. MOSS, 1987 Transposition of the I element and *copia* in a natural population of *Drosophila melanogaster*. Genet. Res. **49:** 121–128.

MONTGOMERY, E., B. CHARLESWORTH and C. H. LANGLEY, 1987 A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. Genet. Res. **49:** 31–41.

MONTGOMERY, E. A., and C. H. LANGLEY, 1983 Transposable elements in Mendelian populations: II Distribution of three *copia*-like elements in a natural population of *Drosophila melanogaster*. Genetics **104:** 473–483.

O'HARE, K., 1985 The mechanism and control of P element transposition in *Drosophila melanogaster*. Trends Genet. **1:** 250–254.

O'HARE, K., and G. M. RUBIN, 1983 Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34:** 25–35.

OHTA, T., 1984 Population genetics of transposable elements. IMA J. Math. Appl. Med. Biol. **1:** 17–29.

OHTA, T., 1985 A model of duplicative transposition and gene conversion for repetitive DNA families. Genetics **110:** 513–524.

PERIQUET, G., S. RONSSERAY and M. H. HAMELIN, 1989 Are *Drosophila melanogaster* populations under a stable geographical differentiation due to the presence of P elements? Heredity **63:** 47–58.

PRESTON, C., and W. R. ENGELS, 1989 Spread of P transposable elements in inbred lines of *Drosophila melanogaster*. Prog. Nucleic Acid Res. Mol. Biol. **36:** 71–85.

RONSSERAY, S., and D. ANXOLABEHERE, 1987 Chromosomal distribution of P and I transposable elements in a natural population of *Drosophila melanogaster*. Chromosoma **94:** 433–440.

SAEDLER, H., and NEVERS, 1985 Transposition in plants: a molecular model. EMBO J. **4:** 585–590.

SIMMONS, M. J., and L. M. BUCHHOLZ, 1985 Transposase titration in *Drosophila melanogaster*: a model for cytotype in the P-M system of hybrid dysgenesis. Proc. Natl. Acad. Sci. USA **82:** 8119–8123.

SIMMONS, M. J., J. D. RAYMOND, M. J. BOEDIGHEIMER and J. R. ZUNT, 1987 The influence of nonautonomous P elements on hybrid dysgenesis in *Drosophila melanogaster*. Genetics **117:** 671.

SIMONELIG, M., C. BAZIN, A. PELISSON and A. BUCHETON, 1988 Transposable and nontransposable elements similar to the I factor involved in inducer-reactive (IR) hybrid dysgenesis in *Drosophila melanogaster* coexist in various *Drosophila* species. Proc. Natl. Acad. Sci. USA **85:** 1141–1145.

SLATKIN, M., 1985 Genetic differentiation of transposable elements under mutation and unbiased gene conversion. Genetics **110:** 145–158.

SNYDER, M., and W. F. DOOLITTLE, 1988 P elements in *Drosophila*: selection at many levels. Trends Genet. **4:** 147–149.

SPRADLING, A. C., 1986 P element mediated transformation, pp. 175–198 in *Drosophila, A Practical Approach*, edited by D. B. ROBERTS. IRL Press, Washington, D.C.

STACEY, S. N., R. A. LANSMAN, H. W. BROCK and T. A. GRIGLIATTI, 1986 Distribution and conservation of mobile elements in the genus *Drosophila*. Mol. Biol. Evol. **3:** 522–534.

VARMUS, H., 1982 Form and function of retroviral proviruses. Science **216:** 812–821.

Communicating editor: R. R. HUDSON