Original article

Model of invasion of a population by transposable elements presenting an asymmetric effect in gametes

G Morel [1], R Kalmes [2], G Périquet [2*]

[1] *UFR Sciences et Techniques, Département de Mathématiques, Parc Grandmont, 37200 Tours;*
[2] *UFR Sciences et Techniques, Institut de Biocénotique Expérimentale des Agrosystèmes, URA CNRS 1298, Parc Grandmont, 37200 Tours, France*

Summary – Dynamics of population invasion by transposable elements is analyzed and simulated, using a model with a very large number of transposition sites. The properties of the model are determined in the framework of a conflict between transposition capabilities of the elements and their harmful effects on the host genome. Equations are developed for the mean and the variance of the number of elements at equilibrium. We use simulations to analyze the effects of various parameters on the dynamics of the elements, revealing the importance of the insertion rate and of the self-regulation properties of the elements. Using values obtained from the P-M system of *Drosophila melanogaster*, the simulations show that the invasion of these elements is likely to occur in 100 years, which is an interval compatible with recent ideas on this invasion. Our analysis of chained invasions reveals the possibility of a mean element number gradient occurring, just as has been observed in European wild populations.

transposable elements / hybrid dysgenesis / model of invasion / simulation

Résumé – Modèle d'invasion d'une population par des éléments transposables présentant une action asymétrique selon les gamètes. *L'invasion de populations par des éléments transposables est analysée et simulée en utilisant un modèle à grand nombre de sites de transposition. Les propriétés du modèle sont déterminées dans le cadre du conflit entre les capacités de transposition des éléments et les effets délétères qu'ils induisent sur le génome hôte. Les équations sont développées pour la moyenne et la variance du nombre d'éléments à l'équilibre. Les simulations permettent d'analyser les effets des différents paramètres sur la dynamique des éléments et montrent l'importance du taux d'insertion et des propriétés d'autorégulation des éléments. En utilisant les valeurs obtenues pour le*

---

* Correspondence and reprints

système PM de Drosophila melanogaster, *les simulations montrent que l'invasion de tels éléments est susceptible de se produire en une centaine d'années, intervalle compatible avec les données récentes sur cette invasion. Une analyse d'invasions en chaîne met en évidence la possibilité d'obtenir un gradient de fréquence des éléments dans les populations, similaire à celui actuellement observé dans les populations naturelles européennes.*

éléments transposables / dysgenèse hybride / modèle d'invasion / simulation

INTRODUCTION

About 15% of the eukaryote genome consists of a family of repeated and dispersed DNA sequences. Many of these sequences have been described before, and some of them have been found capable of mobility (review in Berg and Howe, 1989).

Several models have been proposed to characterize the distribution laws of these transposable elements in populations as a function of different variables such as their transposition and excision rate, and the selective values given to carrier individuals (reviewed in Charlesworth, 1985, and Brookfield, 1986 and 1991). Generally speaking, in all sexed organisms, these models have shown that a family of elements could be kept in stable equilibrium by the opposed effects of replicative transposition and selection against the harmful carriers.

However, much experimental research has proved the existence of self-regulation mechanisms by which the probability of transposition of an element decreases as a function of the number of elements of the same family present in the host genome (reviewed in Berg and Howe, 1989). Different models have shown that such self-regulation could also lead to a state of stable equilibrium for the distribution law of a given family of elements (Charlesworth and Charlesworth, 1983; Langley *et al*, 1983, Charlesworth and Langley, 1986; Langley *et al*, 1988; Rio, 1990).

In *Drosophila melanogaster*, the research on hybrid dysgenesis induced by families of *I*, *P* and *hobo* elements (reviewed in Berg and Howe, 1989; and in Berg and Spradling, 1991) has generated a set of data by which more specific basic models can be conceived. Such models have been proposed for describing the evolution of such systems, in consideration of some of their characteristics, but either by dealing only with the case of a single transposition site (Ginzburg *et al*, 1984; Uyenoyama, 1985) or with an infinite number of sites, and analyzing the selective values at the individual level (Brookfield, 1991).

In the present article, we analyze a model for a family of transposable elements whose transposition and excision rates are functions of the copy number, and whose dysgenic effects depend on the type of crossing. The invasion conditions of these elements in a population are determined analytically. When a state of internal equilibrium exists, the mean and the variance of the distributions are found.

Simulations are used to verify the equations, and as a basis for discussing invasion rates of this type of element in populations.

DESCRIPTION OF THE MODEL

The model developed here is based on the opposing actions of a transposase and of a repressor, whose reciprocal concentrations, and thereby the effets, depend on the element copy number. The equilibrium or disequilibrium existing between these 2 components depends on the direction of crossing, and is established in the zygote. Adults have a selective value linked to the possible dysgenetic effects affecting the zygote they come from.

In the model the number of sites (T) is supposed large enough to allow any transposition into an empty site. The gametes are characterized by the number (ranging from 0 to T) of active elements they contain. For the ovum, this number is taken as an index of the concentration of repressor, as we assume the rate of transposition is simply controlled by the repressor present in this gamete. Considering T as very large, the frequency of occupied sites does not appear as a pertinent parameter and we address here only the distribution of copy number per gamete.

The element copy number·distributions in the spermatozoa and in the ova of generation t are considered to be identical. We define them by $(p_t(0), p_t(1), \ldots p_t(T))$, with:

$$\sum_{n=0}^{T} p_t(n) = 1.$$

The zygote obtained by crossing an ovum containing i elements and a spermatozoon containing j, is denoted (i, j).

We must distinguish between 3 types of crossing (table I).

Table I. Selective values of zygotes, obtained as a function of the composition in transposable elements of the gametes concerned.

			Spermatozoa			
	♂			No. of elements		
♀	0	1	...	j	...	T
0				$W(j) = 1 - S(j)$		
Ova 1						
No. of elements i	$w(i,0) = 1$			(i,j)		
				$w(j) = 1 - s(j)$		
T						

First case

The spermatozoa contain no elements, so that $j = 0$. We suppose that the transposable elements have no effect on the $(i, 0)$ zygote because the equilibrium between the repressor concentration and the number of elements in the egg is not disturbed.

The selective value of these zygotes is taken as reference, and is therefore set equal to unity, $w(i, 0) = 1$. Finally, we let $(1 - G(i))$ be the frequency of the gametes without elements in the set of gametes produced by the $(i, 0)$ type zygotes. When the sites of the elements are on a single pair of chromosomes, and when there is no recombination possible, we have $G(i) = 1/2$ for $i > 0$.

Second case

The ovum has no repressor $(i = 0)$ and the spermatozoon has elements $(j > 0)$. In this configuration there is a high level of element activity, and we let $A(j)$ be the mean increase in the number of elements for the type $(0, j)$. $A(j)$ is the mean number of elements created, less the mean number of elements lost.

$W(j) = 1 - S(j)$ represents the selective value of these zygotes. $[1 - B(j)]$ is the frequency of gametes without elements in the set of gametes produced from type $(0, j)$ zygotes.

Third case
$i > 0$ *and* $j > 0$.

We define $a(j)$, $w(j) = 1 - s(j)$ and $b(i, j)$. The values a and w are supposed to depend only on j, which induces the disequilibrium between the number of elements and the repressor concentration.

As in the first case, the repressor concentration of the ovum is assumed to balance its element copy number. In the zygote, the disequilibrium therefore depends only on the number of elements j introduced by the spermatozoon. The values $a(j)$ and $w(j) = 1 - s(j)$ correspond to the mean increase in the number of elements and to the selective value of the zygotes (i, j). In these zygotes, the presence of repressor limits the activity of the j elements introduced, which means $a(j) < A(j)$ and $w(j) > W(j)$. Finally, and as before, we define $(1 - b(i, j))$ to be the frequency of gametes without elements resulting from these zygotes. Table II summarizes the list of parameters used in the model.

Table II. Parameters used in the model.

$p_t(n)$	: Frequency of gametes with n elements at the tth generation
i	: No of elements in the ovum
j	: No of elements in the spermatozoon
(i, j)	: Type of zygote
$G(i)$	: Frequency of gametes harboring elements among the set of gametes produced by the $(i, 0)$ type zygote $(i > 0)$
$B(j)$	: Frequency of gametes harboring elements among the set of gametes produced by the $(0, j)$ type zygote $(j > 0)$
$b(i, j)$	: Frequency of gametes harboring elements among the set of gametes produced by the (i, j) type zygote $(i > 0, j > 0)$
$A(j)$	: Mean increase in the number of elements for the $(0, j)$ type $(j > 0)$
$a(j)$	: Mean increase in the number of elements for the (i, j) type $(i > 0, j > 0)$
$W(j) = 1 - S(j)$	: Selective value of the $(0, j)$ type zygote $(j > 0)$
$w(j) = 1 - s(j)$	: Selective value of the (i, j) type zygote $(i > 0, j > 0)$.

ANALYSIS OF THE MODEL

Analysis of initial element propagation conditions

If we assume panmixia in an infinite population, we get, considering the fertile individuals, $p_{t+1}(0) = D_0/D$, with:

$$D_0 = (p_t(0))^2 + p_t(0) \sum_{i=1}^{T} (1 - G(i)) p_t(i)$$

$$+ p_t(0) \sum_{j=1}^{T} p_t(j) W(j) (1 - B(j))$$

$$+ \sum_{i=1}^{T} \sum_{j=1}^{T} p_t(i) p_t(j) w(j) (1 - b(i, j))$$

and

$$D = (p_t(0))^2 + p_t(0) \sum_{i=1}^{T} p_t(i) + p_t(0) \sum_{j=1}^{T} p_t(j) W(j)$$

$$+ \sum_{i=1}^{T} \sum_{j=1}^{T} p_t(i) p_t(j) w(j)$$

$$= p_t(0) + p_t(0) \sum_{j=1}^{T} p_t(j) W(j) + (1 - p_t(0)) \sum_{j=1}^{T} p_t(j) w(j).$$

D is always positive when at least one of the $w(j)$ is positive.

By replacing $p_t(0)$ with $(1 - \sum_{i=1}^{T} p_t(i))$, $p_{t+1}(0)$ becomes a function of $[p_t(1), \ldots, p_t(T)]$. In order for the element frequency to be able to increase, the function $p_{t+1}(0) - p_t(0) = p_{t+1}(0) - (1 - \sum_{i=1}^{T} p_t(i))$ must be strictly negative in a neighborhood of $(p_t(1) = 0, \ldots, p_t(T) = 0)$.

As $p_{t+1}(0)$ is differentiable in $(0, \ldots, 0)$, we have to compute the partial derivatives in $(0, \ldots, 0)$.

We get for $1 \leqslant k \leqslant T$ *(Appendix 1)*:

$$(\partial p_{t+1}(0)/\partial p_t(k))(0, \ldots, 0) = -G(k) - W(k)B(k).$$

A sufficient condition for an increase of elements starting from a small initial number of gametes possessing elements is therefore $G(k) + W(k)B(k) > 1$ for $1 \leqslant k \leqslant T$. When the gametes introduced have few elements, it is sufficient for the first inequalities alone to be satisfied.

These inequalities are easy to interpret because $G(k)$ [respectively $W(k)B(k)$] is the probability that a $(k, 0)$ type zygote [resp $(0, k)$] that would be viable and nonsterile will produce a gamete containing at least one element.

It will be seen that this model generalizes that of Ginzburg *et al* (1984), which assumes a single insertion site, or that the number of elements has no effect, and on a single pair of chromosomes. In this case, the notation is $G(i) = 1/2$ for $1 \leqslant i \leqslant T$; $S(j) = S$; $s(j) = 0$; $B(j) = \beta + 1/2(1 - \beta) = 1/2 + \beta/2$ for $1 \leqslant j \leqslant T$; $b(i, j) = 1$ for $1 \leqslant i \leqslant T$ and $1 \leqslant j \leqslant T$ (β being the probability that the maternal genome be contaminated by transposable elements in a $(0, j)$ type mating. The T inequalities are identical with $G(k) = 1/2$, $W(k) = 1 - S$ and $B(k) = 1/2 + \beta/2$. So we once again find the necessary and sufficient condition of expansion which they reach in their special case, ie $\beta > S/(1 - S)$.

In the present model, $B(k)$ is not fixed, but rather depends on the transposition process. In the case of only one chromosomal pair and assuming that the k elements of the paternal chromosome are not excised, that the increase in the number of elements is $A(k)$ for any $(0, k)$ zygote, and that any new element is inserted randomly in 1 of the 2 chromosomes, we get $B(k) = 1 - (1/2)^{A(k)+1}$.

The kth inequality is then written

$$1/2 + W(k)[1 - (1/2)^{A(k)+1}] > 1$$
$$\text{or } S(k) < 1 - 1/[2 - (1/2)^{A(k)}].$$

Each inequality yields a relation between $S(k)$ and $A(k)$ for determining the conditions under which the element copy number will increase.

The hatched area of figure 1 corresponds to the values of $S(k)$ and $A(k)$ verifying this inequality. The harmful effect of the transposable elements can increase as the increase in the number of elements created itself becomes greater.
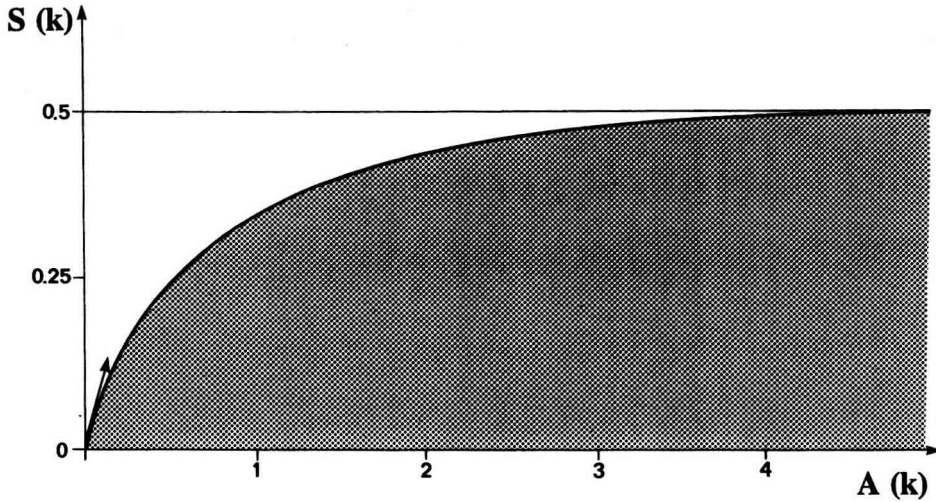
Fig 1. Relation between the selection coefficient of dysgenic zygotes ($S(k)$) and the mean increase in the number of elements in this type of zygote.

Analysis of the positions of equilibrium

Analysis of the mean

$P_t = (p_t(0), \ldots, p_t(T))$ is an equilibrium point if $p_{t+1}(i) = p_t(i)$ for $0 \leqslant i \leqslant T$.

The modelling described here cannot be used to determine the $p_{t+1}{}^i$ values as a function of P_t for $i > 0$. This is possible only if we know, for each type of zygote, the distribution of the gametes produced as a function of the number of elements they contain. Each of these distributions requires T parameters (the sum of them being less than or equal to unity) in order to be defined. It should be possible to reduce this excess of parameters by adopting assumptions concerning the mode of action of the transposable element. This problem is not addressed in the present paper. Instead, we attempt to obtain the equations for the mean and the variance of the distributions of elements at equilibrium, when it exists. Such equilibria have been found for the corresponding model of Ginzburg *et al* (1984). The mean and variance depend on the parameters previously defined. This way we get $p_{t+1}(0) = D_0/D$ (see *Analysis of initial element propagation conditions*) and the mean $E(X_{t+1}) = E(Y_{t+1})$ for the variables X_{t+1} (resp Y_{t+1}), number of elements in the ova (resp in the spermatozoa), of the $(t+1)$th generation (see *Appendix 2*).

At a point of equilibrium we have:

$$p_{t+1}(0) = p_t(0) \tag{1}$$
$$E(X_{t+1}) = E(X_t). \tag{2}$$

When $p_t(0) \neq 1$ we can use the variable X', which follows the law of X_t conditioned by the gametes containing elements. We get $E(X_t) = (1-p)E(X')$, with $p = p_t(0)$.

For a point of equilibrium $(p, p_1, \ldots, p_T)$ other than $(1, 0, \ldots, 0)$, equation [1] can be written:

$$cp^2 - (c + d + e)p + e = 0$$

with

$$c = E(w(Y')) - E(W(Y'))$$
$$d = E(G(Y')) + E(W(Y')B(Y')) - 1$$
$$e = E(w(Y')(1 - b(X', Y'))).$$

In the case considered before of a species having only one pair of chromosomes, and supposing that the elements are not excised, the (i, j) type yields no gametes without elements; so we have $b(i, j) = 1$ for $i > 0$ and $j > 0$, and therefore $e = 0$, which corresponds to an equilibrium where all of the gametes would possess elements. $p = 0$ is then a solution of the equation [1]. It is the only equilibrium possible if $d > 0$, because $c > 0, (w(j) > W(j))$. This situation occurs in particular when the inequalities related to the element copy number growth conditions are verified.

When there is more than 1 pair of chromosomes, or when the element can be excised, $b(i, j)$ may be other than unity; but it approaches it very quickly as i and j increase. It is therefore not surprising to find populations in equilibrium in which $p(0)$ can be considered zero. If it is, equation [2] is reduced too, and the mean number $E(X')$ of elements per gamete satisfies:

$$E(X')E(w(X')) = E((X' + a(X'))w(X')).$$

If $a(.)$ and $w(.)$ are linear functions $(a(j) = a.j.w(j) = 1 - (s.j))$, this equation is written:

$$E(X')(1 - sE(X')) = (1 + a)(E(X') - s(\text{Var}(X') + (E(X'))^2))$$

in which $\text{Var}(X')$ designates the variance in the number of elements per gamete.

$E(X')$ therefore does not depend only on the mean increase and the selective value, but through the variance of X' it also depends on the dispersion of the insertion-excision process. This variance therefore deserves being analyzed.

Analysis of the variance

Let $\text{Var}(i, j)$ be the variance of the number of elements of gametes produced by type (i, j) zygotes. Even with a deterministic model of the number of transpositions and excisions in these zygotes, $\text{Var}(i, j)$ is not zero.

$\text{Var}(i, j)$ is analyzed in *Appendix 3* for the case of a single pair of chromosomes. These new parameters are introduced in order to calculate the variances $\text{Var}(X_{t+1}) = \text{Var}(Y_{t+1})$ of the number of elements in the gametes of generation $(t + 1)$.

At a point of equilibrium we have $\text{Var}(X_{t+1}) = \text{Var}(X_t)$, which provides a third condition [3] of equilibrium (see *Appendix 4*). This condition depends on the third moment of X'.

Even when $p(0) = 0$ and the functions $a(.)$, $w(.)$ and $\text{Var}(.,.)$ are simple, the simulations (see below) have shown the importance of the third moment, which has in no case been found to be close to zero. Equations [2] and [3] cannot therefore be used to find $E(X')$ and $\text{Var}(X')$.

As the invasion dynamics of the elements are just as interesting as their mean and variance at equilibrium, we chose to simulate the process rather than simplify the equations by approximation. However, the mean increase in the number element per type of zygote is not sufficient and we must take into account the way a (i, j) zygote produces new elements.

SIMULATION AND NUMERICAL ANALYSES

Evolution simulation program

To reduce the simulation program run time and have a first approach to the process, the program computes the case of a single pair of chromosomes without recombination. This has its effect on the numerical results by way of $G(.)$, $B(.)$, $b(.,.)$ and $\text{Var}(.,.)$, but does not change the mean value. Moreover it will allow an introduction to the general features of the phenomena.

The user has to define the element copy number distribution in the gametes of the original generation, as well as the functions $A(.)$, $a(.)$, $W(.)$ and $w(.)$. Table III summarizes the list of parameters used in these simulations.

Table III. Parameters used in simulations for describing the dynamic of transposable elements.

i	: No of elements in the ovum
j	: No of elements in the spermatozoon
(i, j)	: Type of zygote with i and j in $[0, T]$, and $T = 100$
$U.j^D$	: No of elements created per $(0, j)$ type zygote
$V.j^E$	: No of elements lost per $(0, j)$ type zygote
$A(j) = U.j^D - V.j^E$:	Mean increase in the number of elements per $(0, j)$ type zygote
$u.j^d$	: No of elements created per (i, j) type zygote
$v.j^e$	: No of elements lost per (i, j) type zygote
$a(j) = u.j^d - v.j^e$	: Mean increase in the number of elements per (i, j) type zygote
U	: Rate of insertion per element for a $(0, j)$ zygote
V	: Rate of excision per element for a $(0, j)$ sygote
u	: Rate of insertion per element for a (i, j) zygote
v	: Rate of excision per element for a (i, j) zygote
$W(j) = 1 - S.j^F$	: Selective value of the $(0, j)$ zygote
$w(j) = 1 - s.j^f$	: Selective value of the (i, j) zygote
By definition, $w(i, 0) = 1$	

The functions allowed are of the form:

$$A(j) = U.j^D - V.j^E,$$
$$a(j) = u.j^d - v.j^e,$$
$$W(j) = \sup(0, 1 - S.j^F) \text{ and}$$
$$w(j) = \sup(0, 1 - s.j^f).$$

The mean increases are therefore the result of a (U, u) transpose and a (V, v) excision process (Charlesworth and Charlesworth, 1983).

To obtain the $p_t(n)$ frequencies at the tth generation, we have to determine for each (i, j) type zygote the gametes it will produce. However, the knowledge of $A(j)$ and $a(j)$ are not sufficient, and the distribution of the transposed and excised elements around these means is necessary. The program allows the user to choose between a distribution ranging between the two integers to either side of the mean, or a Poisson distribution. From such a distribution the final composition of gametic types is determined, giving to each chromosome produced its number of new elements.

The simulation stops at the tth generation when the frequencies $p_t(n)$ and $p_{t-1}(n)$ are within 10^{-6} of each other $(0 \leqslant n \leqslant T)$.

The stability of the mean and the variance of the tth generation is verified by computing the mean and the variance of the $(t+1)$th generation from the formulae that led to equations [2] and [3].

Examinations of a few special cases

The examples considered here are based on linear functions $(D = E = d = e = F = f = 1)$, and the increase of elements is distributed over 2 consecutive integers.

Using the notation of the model, we get for the average increases in the number of elements:

$$A(j) = U.j - V.j, \ a(j) = u.j - v.j,$$

and for the selective values of the zygotes:

$$W(j) = 1 - S.j \text{ and } w(j) = 1 - s.j.$$

Using the available experimental data (from Bingham *et al*, 1982; Engels, 1988; Berg and Spradling, 1991) for the P-M system of *Drosophila melanogaster*, orders of magnitude were defined along with rates of insertion, excision and selective values.

A first series of simulations, carried out as a check, shows as expected that, when there is no deleterious effect $(S = s = 0)$ the mean of the number of elements increases indefinitely, at a rate that depends on the insertion and excision rates. In a second series of simulations, the relations between the harmful effects of the elements and their regulation capacities were examined.

Variations with counterselection and self-regulation of elements

When the mobility of the elements causes harmful side effects, the variations depend on the ratios between the various parameters. For a mean increase of the order of

0.25 per element in dysgenic mating ($(0,j)$ zygote), and considering an excision rate 100 times smaller than the insertion rate (Engels, 1988), we get $U = 0.252$ and $V = 0.002$. If the self-regulation phenomena did not exist in the (i,j) zygotes, the parameters u, v and s would be equal to U, V and S, respectively. Under these conditions (fig 2), the simulations show that the transposable elements may invade the population rather quickly (250 generations) when the deleterious effect is not too great ($S = 0.05$), but can only set in once S reaches a threshold value, which is $S = 0.11$ here. For intermediate values of S ($S = 0.08$), the invasion time will be greater (500 generations) and only a part of the gametes would have transposable elements.

Fig 2. Dynamics of transposable elements with time, as a function of their deleterious effects, and without the self-regulating effect of the transposable elements.
Example 1: $U = 0.252$, $D = 1$, $V = 0.002$, $E = 1$, $S = 0.05$, $F = 1$, $u = 0.0504$, $d = 1$, $v = 0.0004$, $e = 1$, $s = 0.024$, $f = 1$.
Example 2: $U = 0.252$, $D = 1$, $V = 0.002$, $E = 1$, $S = 0.05$, $F = 1$, $u = 0.252$, $d = 1$, $v = 0.002$, $e = 1$, $s = 0.05$, $f = 1$.
Example 3: $U = 0.252$, $D = 1$, $V = 0.002$, $E = 1$, $S = 0.08$, $F = 1$, $u = 0.252$, $d = 1$, $v = 0.002$, $e = 1$, $s = 0.08$, $f = 1$.
The initial generations satisfy $p_0(0) = 0.999$, $p_0(1) = 0.001$.

The self-regulation effect in the (i,j) zygotes can then be analyzed by assigning values 10 times smaller to u and v, or $u = U/10$ and $v = V/10$, and choosing a low deleterious effect of $s = 0.012$. The simulation is initialized with 1 gamete among 1 000 carrying a single element, or $p_0(1) = 10^{-3}$. The curves obtained for different values of S (0.05 and 0.08) are given in figure 3. Here, the elements can totally invade the population, while the selection coefficient S against the dysgenic zygotes is < 0.11. The invasion is slower than before and the population reaches a stable equilibrium in 1 500 to 2 000 generations, but with a higher mean number of elements (13.4 on the average, and with a SD of 4.8). On the other hand, the frequency of the gametes without elements rapidly diminishes and becomes practically zero in < 350 generations.
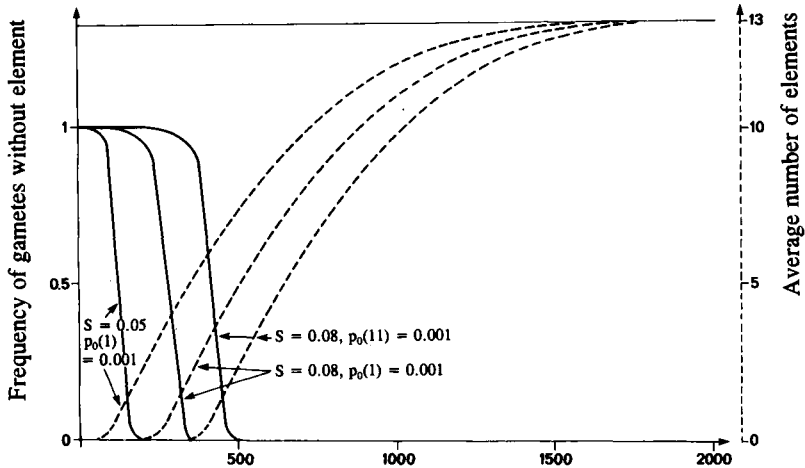
Fig 3. Dynamics of transposable elements with time, as a function of their deleterious effects and of their self-regulating capacity. For these examples, $U = 0.252$, $D = 1$, $V = 0.002$, $E = 1$, $F = 1$, ($S = 0.05$ or $S = 0.08$; in which $S = 0.11$, there is no invasion), $u = 0.0252$, $d = 1$, $v = 0.0002$, $e = 1$, $f = 1$, $s = 0.012$. The founding generation contains 99.9% gametes with no element ($p_0(0) = 0.999$).

It will be noted that the invasion is not necessarily assisted if the rare founder gametes carry a large number of elements. This is due to the fact that the majority of the zygotes in the first generations will be dysgenic and very highly counterselected. When the spermatozoa introduced carry more than $1/S$ elements, the $(0, j)$ zygotes provided will be all dysgenetic ($W(j) = \sup(0, 1 - S.j) = 0$) and there will be no invasion. However, the invasion might occur for lesser values. Such a case is illustrated in figure 3 with $p_0(11) = 10^{-3}$ and $S = 0.08$. We then observe that the evolution proceeds in a manner similar to the previous cases, but a little less rapidly.

Effects of variations in the insertion-excision and self-regulation rates

As the excision rates under dysgenic conditions are known only approximately (to within a factor of 10; see Engels, 1988), a set of simulations were run to get an idea of its role in the case of our model. Using the conditions of the preceding example, of a mean increase of 0.25, we chose to increase the excision rate 10-fold, or:

$$U = 0.27, \quad u = U/10$$
$$V = 0.02, \quad v = V/10$$

and $s = 0.012$.

The results obtained (fig 4) show that the evolutions are quite similar to those obtained previously, with equilibrium achieved in 1 500 to 2 000 generations with a mean number of elements of 12.6 and a SD of 4.7. In fact, when the insertion rate is much greater (10-100-fold) than the excision rate, variations in the latter have little influence on the evolution dynamics.
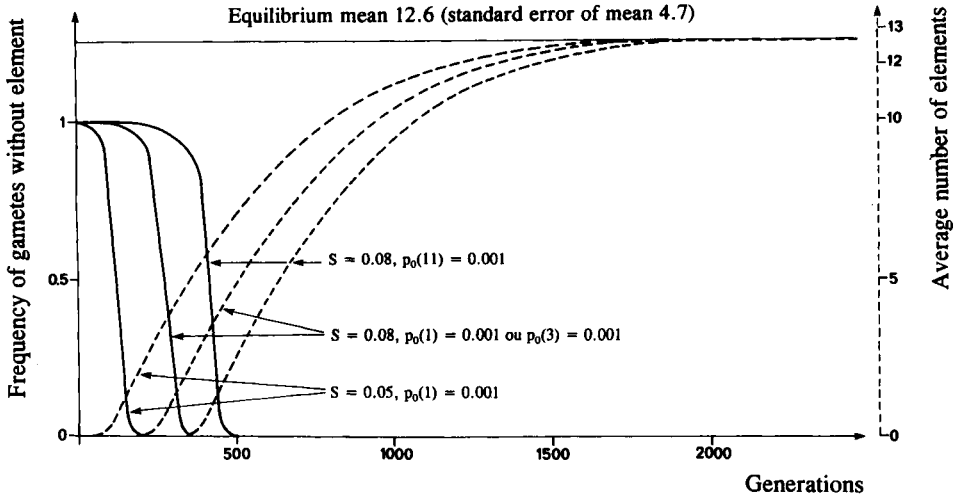
Fig 4. Dynamics of transposable elements with time as a function of the variation of the excision rate. In the examples chosen, the excision rate is 10-fold greater than in the simulations presented in figure 2. For these examples, $U = 0.27$, $D = 1$, $V = 0.02$, $E = 1$, $F = 1$, ($S = 0.05$ or $S = 0.08$; in which $S = 0.11$, with no invasion), $u = 0.027$, $d = 1$, $v = 0.002$, $e = 1$, $f = 1$, $s = 0.012$. The founding generation contains 99.9% gametes with no element ($p_0(0) = 0.999$).

In the same way, the role of element mobility regulation and the dysgenic effects observed in (i, j) type zygotes can be evaluated by modifying the values of the corresponding parameters. We examined the case of relatively ineffective regulation in this way, by increasing the values of u and v, and generally the simulations show that the invasion time is shorter.

In the examples given in figure 5 (with $U = 0.252$, $V = 0.002$, $u = U/2$, $v = V/2$ and $s = S/2$), the invasion process is possible only for $S < 0.11$, and the equilibria are achieved at between 250 and 500 generations. In the limiting case, this situation tends toward the one presented before (fig 2), in which self-regulation did not exist in the (i, j) zygotes.

Finally, a last set of simulations was undertaken, to estimate the effect of very high transposition rates that could lead to a mean increase A of one element. The values chosen (fig 6) of $U = 1.5$, $V = 0.5$, $u = U/10$, $v = V/10$ and $s = 0.03$ when the invasion takes place ($S \in [0, 0.24]$) showed that the invasion is faster (from 200 to 800 generations for the 3 examples considered) but leads to equilibrium values very similar to the previous ones (average of 13.7 and a SD of 4.9).

Examination of an invasion in a sequence of stages

The model proposed can also be used to study an invasion occurring in successive waves. An original population A is invaded by transposable elements and then, after a certain number of generations, a part of its individuals emigrate into a fresh population B having none of these elements. Under these conditions, the distri-
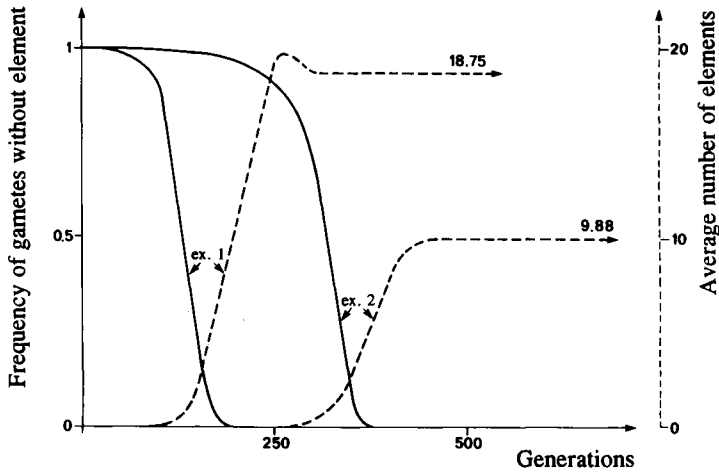
G Morel *et al*

Fig 5. Dynamics of transposable elements. Effect of low self-regulation rates.
Example 1: $U = 0.252$, $D = 1$, $V = 0.002$, $E = 1$, $S = 0.05$, $F = 1$, $u = 0.126$, $d = 1$, $v = 0.001$, $e = 1$, $s = 0.025$, $f = 1$.
Example 2: $U = 0.252$, $D = 1$, $V = 0.002$, $E = 1$, $S = 0.08$, $F = 1$, $u = 0.126$, $d = 1$, $v = 0.001$, $e = 1$, $s = 0.04$, $f = 1$. The founding generations satisfy $p_0(0) = 0.999, p_0(1) = 0.001$.
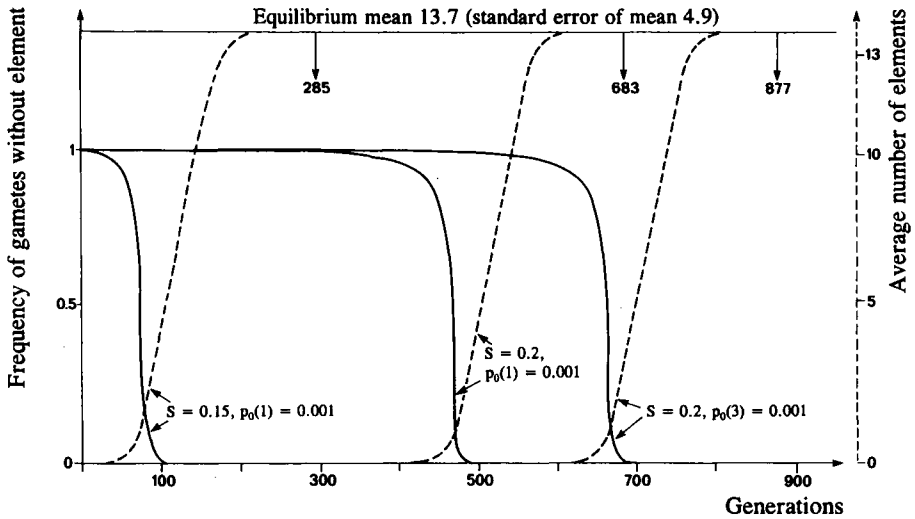

Fig 6. Dynamics of transposable elements. Effect of high transposition rates. For these examples, $U = 1.5$, $D = 1$, $V = 0.5$, $E = 1$, $F = 1$, $u = 0.15$, $d = 1$, $v = 0.05$, $e = 1$, $f = 1$, $s = 0.03$. The founding generations contains 99.9% gametes with no element $p_0(0) = 0.999$.

bution of elements in the gametes introduced into population B will be more or less close to the equilibrium distribution of population A, but usually very different from the initial distribution introduced into this same population A.

The simulations were carried out to create mixed populations from 90% gametes without elements and 10% gametes originating from a parent population in equilibrium. In all cases, the mixed populations evolved toward the equilibrium state of the parent population while the parameters remained unchanged.

However, the analysis of several families of transposable elements has revealed the formation of deleted elements in the course of the generations, which might play a role in the dynamic regulation of the invasion (Black *et al*, 1987; Jackson *et al*, 1988; Périquet *et al*, 1990; Raymond *et al*, 1991).

The parameters then have to be modified, as the regulatory process decreases the mobility of the elements and thereby their dysgenic effects. Considering the importance of the insertion rate, a series of simulations was made with only this parameter modified, in order to represent the invasion of a series of 3 populations. The parent population is defined as $U = 0.27$, $V = 0.02$, $S = 0.05$, $u = U/10$, $v = V/10$ and $s = 0.012$.

At equilibrium, 10% of the gametes are introduced into a population with no elements, which then evolves with $U = 0.22$. The same process will be repeated in a third population for which $U = 0.15$.

The results given in figure 7 show that each population then reaches a specific state of equilibrium, with lower and lower average numbers of elements: 12.6, 6.5 and then 3.0. It is also seen that these states of equilibrium are achieved more and more quickly. Such a process of successive colonizations can therefore lead to the establishment of a geographical differentiation in the distribution of the number of elements. It should be pointed out, though, that considering the formation of deleted elements, each population will contain a mixture of different elements. The values obtained by our simulations concern only complete and active elements.

DISCUSSION

The model and dynamic simulations of transposable elements presented here are based on a genetic approach to the phenomena of hybrid dysgenesis, described mainly for *D melanogaster*. The model leads to a generalization of the model of Ginzburg *et al* (1984) for a large number of transposition sites.

The equations for the mean and the variance of the number of elements at equilibrium depend on the third moment, however, which cannot be neglected. The validity of the model has been confirmed by simulations, mainly by examining the population invasion dynamics. The main results of these simulations can now be discussed and compared with knowledge acquired on P and *hobo* elements in *D melanogaster*, although the model is an oversimplified version for the complex mechanisms of regulation known for these elements.

The parameter values used in our simulations come from the insertion and excision rates observed in dysgenic matings of the P-M system. Although these values are only orders of magnitude, they do show the impact of the various model parameters. The element invasion conditions is given by a system of inequalities stating the relation between the selective coefficient of the dysgenic zygotes and
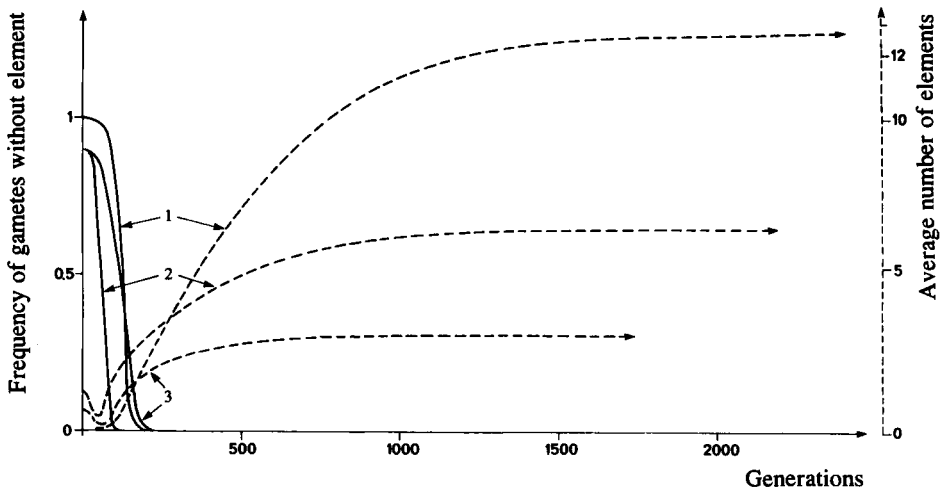
Fig 7. Dynamics of transposable elements in the case of an invasion in waves by 3 populations. Population $A = 1$, population $B = 2$, population $C = 3$.

1. $U = 0.27$, $V = 0.02$, $S = 0.05$, $u = 0.027$, $v = 0.002$, $s = 0.012$. Founding generation consisting of 99.9% gametes with no element, 0.001 with one element.

2. $U = 0.22$, $V = 0.02$, $S = 0.05$, $u = 0.022$, $v = 0.002$, $s = 0.012$. Founding generation consisting of 90% gametes of a population with no element, and 10% gametes from population 1, at equilibrium.

3. $U = 0.15$, $V = 0.02$, $S = 0.05$, $u = 0.015$, $v = 0.002$, $s = 0.012$. Founding generation consisting of 90% gametes of a population without elements, and 10% gametes from population 2, at equilibrium.

the net increase in the element copy number A. When it occurs, the invasion is slower with increasing S. With the values used, the critical threshold of S is of the order of 0.10. The conflict between the deleterious effect of the elements and their invasion can be decided either by increasing the transposition rate or by decreasing their deleterious effect, or by acquiring a self-regulation process.

In the case of P and *hobo* dysgenic systems, the harmful effect due to sterility is smaller at the temperatures usually encountered by the individuals, and a complex self-regulation system was developed. The present model does not account for cytoplasmic type repression effects, but clearly shows the role of chromosomal regulation which, while diminishing the harmful effects when the egg has elements ($s < S$), favors the survival of the carriers and allows the effective invasion of the population by a large number of elements.

It is particularly interesting to note that, for a species like *D melanogaster*, the values used lead to a state of equilibrium in 1 500 generations, for a population that had one gamete to begin with out of 1 000 carrying a single element. Considering the species, such an invasion process would be of the order of some 100 years, which is compatible with the hypothesis of recent invasion by P and *hobo* elements (Kidwell, 1983; Anxolabéhère *et al*, 1988; Pascual and Périquet, 1991).

So the importance of setting up self-regulation mechanims seems to be primordial in the evolution of these systems. We can then understand that mutant elements capable of participating in the global regulation process can be retained by selection.

This would be the case of the *KP* elements and other deleterious elements in the P-M system, as well as the *Th* element in the *hobo* system. A model has been developed recently that includes the effect of this type of element (Brookfield, 1991). The author shows that deleterious elements can indeed be favored by selection if they are a favored substrate for transposition, and that the populations then consist of a combination of complete and deleted elements. In wild P strains, the total number of element is $\approx$ 50. For those analyzed, some one-third are complete, or $\approx$ 15 of these elements, which corresponds to the values obtained in our simulations.

Finally, the impact of the deleted elements on the dynamics of the whole also appears to be important in the case of invasions in series. In our simulations, we saw that invasion was not facilitated simply when the gametes of the original population contained more than one element. This suggests that the diffusion of elements starting from individuals that are part of a population that has already been invaded would be based rather on those individuals containing the fewest complete elements.

Moreover, the presence of incomplete elements, limiting the activity of active ones, leads the receiving population to a state of equilibrium at a lower number of complete elements. When the process is repeated from population to population with a low elements transfer flow, it can lead to the onset of a decreasing gradient of the element copy number in all of the populations. This process is thus consistent with the invasion of European strains by American P ones and their dilution into M cytotype, leading to the gradual variation currently observed. Here again, the values obtained in our simulations, from 12.6 to 3.0 complete elements, are compatible with data observed in wild populations: from 35 elements in all for the French populations, 7 or 8 for those of central Asia (Périquet *et al*, 1989). It will be noted that, if the element transfer flow continued in the course of generations, all of the populations would then tend to homogenize, which is the case of the *I* and *Hobo* elements, for which there is no evidence of geographical differentiation (reviewed in Bregliano and Kidwell, 1983; Pascual and Périquet, 1991). For the P-M system, for which the invasion in *D melanogaster* appears to be the most recent, the homogenization process would be under way and recent observations on the French population do in fact show a trend in this direction over the last decade (Fleuriet *et al*, 1992).

As the model presented here suggests, this process could extend to all European populations, but this would require a relatively long time. We will be able to answer these questions by analyzing wild populations and comparing the results with an extended version of the present model, introducing recombination and segregation between more than one pair of chromosomes and taking into account some aspects of the mechanisms of regulation of these elements. Such a program is under investigation.

ACKNOWLEDGMENTS

APPENDIX 1

Initial element propagation conditions

According to the panmixia assumption, we have to compute:

$$(\partial p_{t+1}(0)/\partial p_t(k))(0, \ldots, 0) \text{ for } 1 \leqslant k \leqslant T$$

Since $D_0(0, \ldots, 0) = D(0, \ldots, 0) = 1$, we have:

$$(\partial p_{t+1}(0)/\partial p_t(k))(0, \ldots, 0) =$$
$$(\partial D_0/\partial p_t(k))(0, \ldots, 0) - (\partial D/\partial p_t(k)(0, \ldots, 0)$$

with

$$(\partial D_0/\partial p_t(k)) = -2\left(1 - \sum_{i=1}^{T} p_t(i)\right) + \left(1 - \sum_{i=1}^{T} p_t(i)\right)(1 - G(k))$$
$$- \sum_{i=1}^{T} p_t(i)(1 - G(i))$$
$$+ \left(1 - \sum_{j=1}^{T} p_t(j)W(k)(1 - B(k))\right) - \sum_{j=1}^{T} p_t(j)W(j)(1 - B(j))$$
$$+ \sum_{i=1}^{T} p_t(i)w(k)(1 - b(i, k)) + \sum_{j=1}^{T} p_t(j)w(j)(1 - b(k, j))$$

and

$$(\partial D/\partial p_t(k)) = -1 + \left(1 - \sum_{j=1}^{T} p_t(j)\right)W(k) - \sum_{j=1}^{T} p_t(j)W(j)$$
$$+ \left(\sum_{j=1}^{T} p_t(j)\right)w(k) + \sum_{j=1}^{T} p_t(j)w(j)$$

whence:

$$(\partial p_{t+1}(0)/\partial p_t(k))(0, \ldots, 0) = (-2 + (1 - G(k)) + W(k)(1 - B(k)))$$
$$- (-1 + W(k))$$
$$= -G(k) - W(k)B(k).$$

APPENDIX 2

Analysis of the means $E(X_{t+1})$ and $E(Y_{t+1})$

$$E(X_{t+1}) = E(Y_{t+1})$$

$$= 1/D(p_t(0) \sum_{i=0}^{T} (i/2)p_t(i) + p_t(0) \sum_{j=1}^{T} ((j + A(j))/2)p_t(j)W(j)$$

$$+ \sum_{i=1}^{T} \sum_{j=1}^{T} ((i + j + a(j))/2)p_t(i)p_t(j)w(j)).$$

We will now rewrite these equations using the variables X'_t and Y'_t, restrictions of the variables X_t and Y_t on the gametes containing elements. This is obviously possible when $P_t \neq (1,0,\ldots,0)$, but this is a trivial point of equilibrium for which we studied the instability conditions in *Analysis of initial element propagation conditions*. To find the nontrivial points of equilibrium, we notice that, when they exist, X'_t and Y'_t have the same distribution, which is defined by:

$$(p_t(1)/c,\ldots,p_t(T)/c)$$

with $c = 1 - p_t(0) = \sum_{i=1}^{T} p_t(i)$.

Giving this distribution and $p_t(0)$ is equivalent to giving P_t. If we simplify the notation by letting $p_t(0) = p$, $X'_t = X'$ and $Y'_t = Y'$, $p_{t+1}(0)$ and $E(X_{t+1})$ are written:

*
$$p_{t+1}(0) = D_0/D$$

in which $D_0 = p - p(1-p)E(G(X')) + p(1-p)E(W(X')(1 - B(X')))$
$$+ (1-p)^2 E(w(X')(1 - b(X',Y')))$$

and $D = p + p(1-p)E(W(X')) + (1-p)^2 E(w(X'))$.

** $E(X_{t+1}) = (1/2D)(p(1-p)E(X')$
$$+ p(1-p)E((X' + A(X'))W(X'))$$
$$+ (1-p)^2 E(X')E(w(X'))$$
$$+ (1-p)^2 E((X' + a(X'))w(X'))).$$

At a point of equilibrium, we have
(1) $p_{t+1}(0) = p$
(2) $E(X_{t+1}) = (1-p)E(X')$.

A point of equilibrium $(p, p_1, \ldots, p_T)$ other than $(1, 0, \ldots, 0)$ must therefore verify the 2 following conditions:

$$
\begin{aligned}
& p^2(E(w(Y')) - E(W(Y'))) \\
& + p(1 - E(G(Y')) - E(w(Y')) + E(W(Y')(1 - B(Y'))) \\
& \qquad - E(w(Y')(1 - b(X', Y')))) \\
& + E(w(Y')(1 - b(X', Y'))) = 0
\end{aligned}
\tag{1}
$$

and

$$
\begin{aligned}
& p^2 2E(Y')(E(w(Y')) - E(W(Y'))) \\
& + p(E(Y') + 2E(Y')E(W(Y')) - 3E(Y')E(w(Y')) \\
& \qquad - E((Y' + A(Y'))W(Y')) + E((Y' + a(Y'))w(Y'))) \\
& + E(Y')E(w(Y')) - E((Y' + a(Y'))w(Y')) = 0.
\end{aligned}
\tag{2}
$$

APPENDIX 3

Analysis of Var(i,j) in the case of a single pair of chromosomes

When $i = j = 0$, we have Var(0,0)=0.

Let us begin by finding $\mathrm{Var}(i, j)$ in the case where the increase in the number of elements is the result of the loss of nd elements $(nd < i+j)$ and of the transposition of na elements (the increase is thus assumed to be constant for the (i, j) strain). We moreover assume that each element has the same probability of disappearing and that any new element has 1 chance out of 2 of meeting the chromosome from the egg. The distribution of the number of elements lost by the chromosome from the egg then follows a hypergeometric distribution with parameters $N = i + j$, $n = nd$ and $N_I = i$. The distribution of the number of new elements is binomial with parameters $n = na$ and $p = 1/2$.

In the (i, j) strains, the distribution of the number of elements of the gametes obtained from the maternal chromosome has the mean $(i - nd.(i/i+j) + na/2)$ and the variance $((nd.i.j.(i+j - nd)/(i+j)^2.(i+j - 1)) + na/4)$. We conclude that the distribution of the number of elements of gametes obtained from the (i, j) strains has mean $(1/2).(i + j - nd + na)$, and variance:

$$
\begin{aligned}
\mathrm{Var}(i, j) = {} & (nd.i.j.(i+j - nd)/(i+j)^2.(i+j - 1)) + na/4 + (i - j)^2/4 \\
& - nd.(i - j)^2/2.(i+j) + nd^2.(i - j)^2/4.(i+j)^2
\end{aligned}
$$

(If $i + j = 1$, then $i = 0$ or $j = 0$ and the first term must be considered zero).

In the (i, j) strains, we no longer assume that the number of elements lost and the number of new elements are constant; they have distributions with means md and ma, respectively, variances Var d and Var a, in which the difference $(ma - md)$ is equal to $A(j)$ if $i = 0$ and to $a(j)$ if $i \neq 0$).

Using the previous case, we deduce that the distribution of the number of elements of gametes obtained from the (i, j) strains has mean $(1/2).(i+j - md + ma)$

and variance:

$$\text{Var}(i,j) = (i.j.((i+j).md - \text{Var } d - md^2)/(i+j)^2.(i+j-1))$$
$$+ ma/4 + (i-j)^2/4$$
$$- md.(i-j)^2/2.(i+j) + (\text{Var } d + md^2).(i-j)^2/4.(i+j)^2$$
$$+ \text{Var } d/4 + \text{Var } a/4.$$

Note : When the distribution of the number of elements lost (resp created) is concentrated on the 2 integers ndi and $ndi+1$ (resp nai and $nai+1$) to either side of md (resp ma), we have:

$$\text{Var } d = (md - ndi).(ndi + 1 - md)$$
$$\text{Var } a = (ma - nai).(nai + 1 - ma).$$

APPENDIX 4

Analysis of the variances $\text{Var}(X_{t+1})$ and $\text{Var}(Y_{t+1})$

$$\text{Var}(X_{t+1}) = \text{Var}(Y_{t+1})$$

$$= (1/D) \left[p_t(0) \sum_{i=0}^{T} p_t(i) \text{ Var}(i,0) + p_t(0) \sum_{j=1}^{T} p_t(j) W(j) \text{ Var}(0,j) \right.$$

$$+ \sum_{i=1}^{T} \sum_{j=1}^{T} p_t(i) p_t(j) w(j) \text{ Var}(i,j)$$

$$+ p_t(0) \sum_{i=0}^{T} p_t(i)(i/2)^2 + p_t(0) \sum_{j=1}^{T} p_t(j) W(j)((j + A(j))/2)^2$$

$$+ \sum_{i=1}^{T} \sum_{j=1}^{T} p_t(i) p_t(j) w(j)((i + j + a(j))/2)^2 \right] - (E(X_{t+1}))^2$$

At equilibrium, we should have $\text{Var}(X_{t+1}) = \text{Var}(X_t)$, which provides a third condition of equilibrium. When $P_t \neq (1,0,\ldots,0)$, the notation of *Appendix 2* leads to the following equation:

$$p^2 E(X')^2 (E(w(Y')) - E(W(Y'))) \qquad\qquad [3]$$
$$+ p((3/4)E(X')^2 + E(X')^2 E(W(Y')) - 2E(X')^2 E(w(Y'))$$
$$- E(\text{Var}(X',0)) - E(W(Y') \text{ Var}(0,Y')) + E(w(Y') \text{ Var}(X',Y'))$$
$$- (1/4)E(W(Y')(Y' + A(Y'))^2)$$
$$+ (1/4)E(w(Y')(X' + Y' + a(Y'))^2))$$
$$+ E(X')^2 E(w(Y')) - E(w(Y') \text{ Var}(X',Y'))$$
$$- (1/4)E(w(Y')(X' + Y' + a(Y'))^2) = 0.$$

REFERENCES

Anxolabéhère D, Kidwell MG, Périquet G (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

Berg DE, Howe MM (1989) *Mobile DNA*, Am Soc Microbiol, Washington, DC

Berg CA, Spradling AC (1991) Studies on the rate and site-specificity of P element transposition. *Genetics* 127, 515-524

Bingham PM, Kidwell MG, Rubin GM (1982) The molecular basis of the P-M hybrid dysgenesis: the role of the P element, a P strain specific transposon family. *Cell* 29, 995-1004

Black DM, Jackson MS, Kidwell MG, Dover GA (1987) KP elements repress P-induced hybrid dysgenesis in *D melanogaster. EMBO J* 6, 4125-4135

Bregliano JC, Kidwell MG (1983) Hybrid dysgenesis determinants. *In: Mobile Elements* (Shapiro J, ed), Academic Press, New York, 363-410

Brookfield JFY (1986) A model for DNA sequence evolution within transposable element families. *Genetics* 112, 393-407

Brookfield JFY (1991) Models of repression of transposition in P-M hybrid dysgenesis by P cytotype and by zygotically encoded repressor proteins. *Genetics* 128, 471-486

Charlesworth B, Charlesworth D (1983) The population dynamics of transposable elements. *Genet Res (Camb)* 42, 1-27

Charlesworth B (1985) *The Population Genetics of Transposable Elements. Population Genetics and Molecular Evolution* (Ohta T, Aoki K eds) Jpn Sci Soc Press, Tokyo; Springer-Verlag, Berlin, 213-232

Charlesworth B, Langley CH (1986) The evolution of self-regulated transposition of transposable elements. *Genetics* 112, 359-383

Charlesworth B, Langley CH (1989) The population genetics of *Drosophila* transposable elements. *Annu Rev Genet* 23, 251-287

Engels WR (1988) P elements in *Drosophila. In: Mobile DNA* (Berg DE, Howe MM eds) ASM Publ, Washington, DC, 437-484

Fleuriet A, Kalmes R, Pascual L, Périquet G (1992) Comparative evolution of P-M system and infection by the sigma virus in French and Spanish populations of *Drosophila melanogaster. Genet Res (Camb)* 60, 77-86

Ginzburg LR, Bingham PM, Yoo S (1984) On the theory of speciation induced by transposable elements. *Genetics* 107, 331-341

Jackson MS, Black DM, Dover GA (1988) Amplification of KP elements associated with the repression of hybrid dysgenesis in *Drosophila melanogaster. Genetics* 120, 1003-1013

Kidwell MG (1983) Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster. Proc Natl Acad Sci USA* 80, 1655-1659

Langley CH, Brookfield JFY, Kaplan N (1983) Transposable elements in Mendelian populations. I. A theory. *Genetics* 104, 457-471

Langley CH, Montgomery EA, Hudson RH, Kaplan NL, Charlesworth B (1988) On the role of unequal exchange in the containment of transposable element copy number. *Genet Res* 52, 223-235

Pasqual L, Périquet G (1991) Distribution of Hobo transposable elements in natural populations of *Drosophila melanogaster. Mol Biol Evol* 8, 282-296

Périquet G, Ronsseray S, Hamelin MH (1989) Are *Drosophila melanogaster* populations under a stable geographical differentiation due to the presence of P elements. *Heredity* 63, 47-58

Périquet G, Hamelin MH, Kalmes R, Eeken J (1990) Hobo elements and their deletion-derivative sequences in *Drosophila melanogaster* and its sibling species *D simulans, D mauritiana* and *D sechellia. Genet Sel Evol* 22, 393-402

Raymond JD, Ojala TA, White J, Simmons MJ (1991) Inheritance of P-element regulation in *Drosophila melanogaster. Genet Res (Camb)* 57, 227-234

Rio CD (1990) Molecular mechanisms regulating *Drosophila* P element transposition. *Annu Rev Genet* 24, 543-578

Uyenoyama MK (1985) Quantitative models of hybrid dysgenesis: Rapid evolution under transposition, extrachromosomal inheritance, and fertility selection. *Theor Popul Biol* 27, 176-201