The First Steps of Transposable Elements Invasion: Parasitic Strategy vs. Genetic Drift

Arnaud Le Rouzic and Pierre Capy¹

Laboratoire Populations, Génétique, Évolution, 91198 Gif-sur-Yvette Cedex, France

Manuscript received May 19, 2004

Accepted for publication September 7, 2004

ABSTRACT

Transposable elements are often considered as selfish DNA sequences able to invade the genome of their host species. Their evolutive dynamics are complex, due to the interaction between their intrinsic amplification capacity, selection at the host level, transposition regulation, and genetic drift. Here, we propose modeling the first steps of TE invasion, *i.e.*, just after a horizontal transfer, when a single copy is present in the genome of one individual. If the element has a constant transposition rate, it will disappear in most cases: the elements with low-transposition rate are frequently lost through genetic drift, while those with high-transposition rate may amplify, leading to the sterility of their host. Elements whose transposition rate is regulated are able to successfully invade the populations, thanks to an initial transposition burst followed by a strong limitation of their activity. Self-regulation or hybrid dysgenesis may thus represent some genome-invasion parasitic strategies.

It is commonly thought that transposable elements (TEs) are universal. Indeed, these autonomous mobile DNA sequences have been found in every organism they were screened for, with very few notable exceptions (GARDNER *et al.* 2002). However, the causes of their universality are still largely misunderstood. Since their discovery, transposable elements have been consecutively thought to be controlling (McCLINTOCK 1984), selfish or parasitic (DAWKINS 1976; DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980), semiparasitic (HICKEY 1982), both parasitic and useful (CAPY *et al.* 2000), or maladaptive (VINOGRADOV 2003). The processes of TE evolution therefore remain questionable, partly because of the complexity of the multilevel ecological niche (genome, individual, population, and species) of these intragenomic inhabitants.

Nevertheless, some evidence of TEs parasitism cannot be questioned. TEs are known to decrease fitness, by inserting themselves into coding DNA sequences and causing chromosomal breaks, deletions, or translocations. They are implied in 50% of deleterious mutations in *Drosophila* (FINNEGAN 1992) and 10% in mice (KAZAZIAN 1998). Several examples of genome invasion by a TE family have been reported; for instance, the whole *Drosophila melanogaster* species has been invaded by *P* elements in a very short period (<40 years) comparative to the evolution timescale (ANXOLABÉHÈRE *et al.* 1988). Moreover, complete genome sequencing has shown that genomes generally contain many inactive TE copies (QUESNEVILLE *et al.* 2003), suggesting that TE families

can colonize genomes and then be partially eliminated in a continuous turnover. Finally, phylogenetic analysis of TE sequences shows inconsistencies between TEs and host phylogenies, suggesting that horizontal transfers (HTs) and invasions of new species appear to be recurrent events in TE evolution (KIDWELL 1992; SILVA *et al.* 2004).

The repetitive successes of TE invasions despite their deleterious effects could be considered as an apparent paradox. The transposition is opposed to several other evolutive constraints: deletion (loss of copies), selection, and regulation (decrease of the transposition rate when the copy number increases). Complex interactions between these forces are not intuitive, and population genetics models have been developed to predict the evolution of these elements (see CHARLESWORTH *et al.* 1994 for review).

Most known models show that the selfish DNA hypothesis is grounded on a solid basis; *i.e.*, a DNA sequence with self-duplication ability is able to invade a genome even if it can be responsible for damage to the host. This colonization is based on transposition and genetic recombination during sexual reproduction, while in asexual organisms the invasion has to be explained in different ways (EDWARDS and BROOKFIELD 2003). Without any evolutive force limiting TE amplification, copy number will increase until the host genome is destroyed. Experimental data generally suggest that selection is the main factor in the limitation of TE copy number increase (CHARLESWORTH *et al.* 1992; HOOGLAND and BIÉMONT 1996), even if the role of self-regulation cannot be ruled out in some situations (COEN *et al.* 1994). However, several key points of TE dynamics remain unclear. For instance, our knowledge about the mecha-

¹Corresponding author: Laboratoire Populations, Génétique, Évolution, Allée de la Terrasse, bât 13, 91198 Gif-sur-Yvette Cedex, France. E-mail: capy@pge.cnrs-gif.fr

nisms by which TEs decrease the fitness of their host is still fragmentary, and TE evolution in situations where several of these selective forces interact is not well understood. Moreover, the long-term evolution of the copy number in a genome after the initial increase is far from being understood either experimentally or theoretically. Indeed, for several species and several TE families, copy number does not seem to vary, and some authors base their expectations on the existence of a state of equilibrium (e.g., CHARLESWORTH and CHARLESWORTH 1983). However, TE dynamics might be slow and unstable in natural populations, and most of the situations observed might not correspond to an equilibrium. In addition, any random demographic, genetic, or ecological event may disturb the state of equilibrium (TSITRONE *et al.* 1999). Generally speaking, the effects of genetic drift on TE dynamics have not yet been fully developed. On one hand, some theoretical results suggest that, in small populations, TE families with several copies per individual can be lost (BROOKFIELD and BADGE 1997), which is contrary to previous analytical expectations. On the other hand, estimations of TE frequency distribution in large natural populations of *Drosophila* tend to confirm that genetic drift is a negligible phenomenon in TE dynamics (BIÉMONT *et al.* 1994).

The initial evolution of TE copy number is not well known and only a few studies have been carried out concerning the increase of the frequency of individuals carrying at least one copy. HICKEY (1982) proposed a simple modeling attempt, based on the existing single-locus segregation distortion model, to describe the evolution of TE frequency in an infinite population. Under these conditions, the element systematically will invade a population provided that its transposition activity (*i.e.*, its overrepresentation in the gametes from heterozygous individuals) counterbalances its loss by natural selection. With a more complex regulation system, like the *P-M* hybrid dysgenesis, the frequency increase might be gradual, and the colonization of the whole population is not systematic if there is no genetic drift (UYENOYAMA 1985). Whatever the population size, genetic drift must influence TE dynamics at least during the first stage of the invasion, when only a few individuals have TE copies. Indeed, after a horizontal transfer, for instance, an element is present with only one copy in one individual. Despite a high transposition ability, the probability of its being lost through genetic drift remains considerable, even in a large population, since the initial frequency of the element $1/2N$ decreases with the population size N . KAPLAN *et al.* (1985) proposed that, in an infinite population, when the elements have no impact on selection, the loss probability p of a single TE copy is the smallest solution of

$$p = e^{-(1-p \cdot e^{-(1-p)u_1})},$$

where u_1 is the transposition rate of an individual carrying one TE copy (Figure 1). However, this result is rather frustrating, since TEs are known to induce deleterious effects, and transposition rates are supposed to vary following the genetic context and the copy number. Therefore, the influence of these two last parameters on TE invasion frequency remains unclear.

The aim of this article is to focus on these first generations after the arrival of a new TE in a diploid outbreeding species. The interaction between a selfish DNA sequence dynamics and random genetic drift indeed raises some important evolutive questions. For instance, we can assume that the higher the transposition rate is, the lower the loss frequency is; but is it always the case? Do TEs invade a population only by chance? If so, most of the transfers should be unsuccessful, and they must be recurrent to explain effective invasions. We propose investigating and discussing the conditions under which TE invasion is more likely, since they may represent an optimized parasitic strategy for frequent horizontally transferred TEs.

MODEL

These questions led us to study a finite size population model. The population size N is considered as constant, with as many males as females, reproducing by panmixia. The genome is made up of several diploid chromosomes that are able to carry a potentially infinite number of transposable elements. The results presented in this article have been obtained with a genome of two chromosomes of 100 cM. The transposition rate is computed for each individual and can vary with TE copy number, parent's genotype, and random factors. Transposition and genetic drift are intrinsically random processes, easily grasped by a stochastic simulation method, implemented in a Monte Carlo simulation program, called DyET and available at <http://www.pge.cnrs-gif.fr/pge/bioinfo/Dyet/>. The random numbers were generated from the Gnu Scientific Library (<http://www.gnu.org/software/gsl/>).

Reproduction: For each of the N offspring, two parents are randomly sampled, one male and one female. The probability for an individual to be selected as a parent is proportional to its fitness. Each parent produces a gamete by recombination. The number of crossings over is chosen following a Poisson process, and their locations are uniformly distributed in the genome. For each chromosome, TEs are copied from a randomly determined parental chromosome to the gamete, and the "matrix" chromosome is swapped when a crossing over occurs. The two haploid genomes of the two gametes are then combined to constitute a new individual.

Fitness computation is based on an additive effect of TE copies. The fitness of an individual i with n elements is $\omega_i = 1 + \sum_{e=1}^n s_e$, where s_e is the selective impact of the element number e . If $s_e < 0$, the element is deleterious for its host. To simplify the interpretation of the results, s_e has been fixed to -10^{-2} for all simulations.

Transposition: The transposable element family is described by several features [selective coefficient (s_e) and

FIGURE 1.—Loss frequency and transposition rate. The solid line represents the analytical result of KAPLAN *et al.* (1985) for an infinite population. Solid circles, simulation results under the same conditions, with a population size $N = 50$; open circles, simulation results when TEs are deleterious ($s = -0.01$). When the transposition rate tends to 0, the expected loss frequencies are respectively $1 - 2u$ for an infinite population, $1 - 1/2N = 0.99$ for the neutral case, and $1 - (1 - \exp(-2s))/(1 - \exp(-4Ns)) = 0.9968$ for the selection case.

basic (u_0) and burst (u_b) transposition rates]. Deletion (loss of copies) is considered as a marginal phenomenon in the first generations and thus has been neglected. Transposition is supposed to occur early in the development (before the germline differentiation), so that all cells are genetically identical. In one generation, the number of transposition events of a precise TE copy is drawn from a Poisson distribution (*i.e.*, a high-transposition rate element can transpose several times). When an individual is determined to be in burst conditions, its TEs transpose with a rate u_b ; otherwise, the transposition rate is u_0 . Five transposition models were followed: no transposition bursts, random transposition bursts, hybrid dysgenesis, threshold self-regulation, and continuous self-regulation. If the bursts occur at random, each individual has a probability of 0.1 to be in burst conditions, regardless of its genome. In the case of hybrid dysgenesis, transposition burst is activated when (i) the female parent has no TE and (ii) the male gamete contains at least one copy. The last two self-regulation models are based on a deterministic setting of the transposition rate following the TE copy number of the individuals (see Figure 2). In the threshold self-regulation model, the burst occurs when the copy number n exceeds a constant k_t , while in the continuous self-regulation model, there is no real burst state, and the transposition rate of each element decreases with the genomic copy number [more precisely, $u = (u_b - u_0) \cdot 2^{-(n-1)/(k-1)} + u_0$, where u_b is the maximum transposition rate, reached in genomes having a single copy; u_0 is the minimum transposition rate, corresponding to an infinite number of copies; and k_c is the number of TE copies for which the transposition rate is $(u_b - u_0)/2$].

Simulations: Simulations were initialized with only one copy randomly disposed in the genome of one adult (reproductive) individual. A repetition consists of letting the system evolve generation after generation until

the end of the simulation, which can be stopped for three reasons: (i) the copy number is nil in all individuals (loss of the element); (ii) the mean fitness is nil (in males, in females, or in both sexes), which leads to population extinction; and (iii) after 1000 generations. For each generation, the mean TE copy number and the mean transposition rate can be computed. As a single repetition represents an example of TE dynamics, several hundred repetitions (1000 in most cases) were run to obtain general trends. More repetitions should have led to more esthetic graphical results, but even if the computation time for one generation is usually < 1 sec (on an Intel Xeon 2400 MHz processor), the numbers of generations, repetitions, and initial conditions tested lead to several days of calculations for each simulation set.

RESULTS

Transposition in a finite population: To determine the loss frequency of a TE copy after an HT, 1000 populations of $N = 50$ individuals were initialized with one element randomly inserted in the genome of one individual. After 1000 generations, the number of repetitions in which the element has been lost is computed. The loss frequency of a single TE copy for a large panel of transposition rates is summarized in Figure 1. When the selective coefficient is 0, the higher the transposition rate is, the lower the loss frequency is. The situation where the transposition rate is very low corresponds to simple genetic drift and the probability of an element being fixed tends to be $\sim 1/2N$. With high transposition rates (above three transpositions per copy per generation), the loss frequency achieves a threshold corresponding mainly to the probability of the first individual carrying the element not transmitting it to its offspring.

FIGURE 2.—Different burst and regulation models used in this work: the transposition rate u_n depends of the copy number n . (a) No transposition burst ($u_n = u_0$); (b) threshold self-regulation ($u_n = u_b$ if $n \leq k_c$; here $k_c = 3$, $u_n = u_0$ otherwise); (c) continuous self-regulation [$u_n = (u_b - u_0) \cdot 2^{-(n-1)/(k_c-1)} + u_0$ with $k_c = 2$].

When the TE copies are deleterious, the loss frequency is higher than that in the nonselection situation: the invasion ability of an element whose insertions are deleterious is less effective whatever its transposition rate. Moreover, the successful invasion frequency becomes very weak for very high transposition rates (more than five transposition events per copy per generation).

Consequently, an optimal transposition rate seems to exist for this element; the transposition rate must be high enough to let a heterozygous individual transmit the element to most of its descendants, but not too high that it would cause sterility in its host.

This striking result can be illustrated by an analysis of a population evolution during the very first generations of the invasion (Figure 3). If the transposition rate is low ($u = 10^{-4}$, Figure 3, a and b), a successful invasion is rare and, when it happens, the TE frequency evolves mainly through genetic drift, while the copy number per individual increases very slowly. When the transposition rate is near the optimum ($u = 1$, Figure 3, c and d), transposition leads to an increase in the copy number per individual and to an increase in TE frequency in the population. When the transposition rate is very high ($u = 10$, Figure 3e), some individuals quickly acquire several dozen elements, but are sterile because of the selection against TEs. They do not reproduce, and the TE disappears.

The “optimal” transposition rate remains very high compared to known transposition rates. Moreover, a TE with such a multiplication capacity will colonize a genome very efficiently, but, after a few generations, the copy number in each individual will be so high that the mean fitness of the population will drop to 0, and no more individuals will reproduce: the population, and then the TE, will be lost. Therefore, two main hypotheses can be proposed:

- The transposition rate of a given family is low and roughly constant, so the colonization of new species is practically fortuitous—horizontal transfers of TEs are a frequent phenomenon and few of them are successful.
- The transposition rate of a TE can vary by several orders of magnitude during the colonization process; it may be high during the first generations following the HT and then decrease until it reaches an equilibrium (or pseudo-equilibrium) value, corresponding to the known measured transposition rates. In this hypothesis, we have to imagine the various processes implied in such “transposition bursts.”

Initial increase of copy number: Different burst models have been tested. For the following results, the basal transposition rate u_0 and the burst transposition rate u_b have been arbitrarily fixed, respectively, to 10^{-4} and five events per copy per generation. Figure 4a presents the evolution of the mean TE copy number in populations where the invasion was successful. If the element is unable to transpose more than its basal transposition rate, the copy number increases very slightly and is lost most of the time because of genetic drift [invasion frequency (IF) is 0.67%]. When the transposition bursts are random and occur in 10% of individuals (IF = 27.1%), the copy number increase is regular and more rapid; but

FIGURE 3.—Examples of copy number dynamics during the first 20 generations with different transposition rates, $u = 10^{-4}$ (a and b), $u = 1$ (c and d), and $u = 10$ (e). Populations were initialized with only one copy. The left (a, c, and e) gives examples of simulations where TEs have been lost, and the right (b and d) gives examples of successful invasions. Over 1000 simulations, no invasion was reported for $u = 10$. For the two other transposition rates, the percentage of TE loss among 1000 repetitions is indicated on the central diagrams. Each small colored rectangle represents a single individual, and the color of the rectangle corresponds to its TE copy number (see the color scale on the bottom). For each of the 20 generations, the 50 individuals of the population are plotted in an arbitrary order. When the transposition rate is low (a and b), the element frequency evolves mainly through genetic drift, and the TE is lost in most cases. With intermediate transposition rates ($u = 1$), the probability of invasion increases; the element can still be lost by genetic drift during the very first generations (c), but the copy number increase due to the high transposition rate leads to a quick diffusion of the TE in all individuals (d). In the last situation ($u = 10$), the TE copy number in the few individuals carrying them is so high (>100) that they are sterile (denoted by *), and the element is not transmitted to the next generation.

FIGURE 4.—(a) Evolution of the mean copy number per individual in successfully invaded populations for the different transposition burst models. Error bars represent standard errors. (b) Evolution of the mean transposition rate in the same set of populations. For both a and b, means and standard errors were computed on 672 successfully invaded populations (over 100,000 repetitions) for the “no-burst” model, 271 populations (1000 repetitions) for the random-burst model, 291 populations (1000 repetitions) for the hybrid dysgenesis model, 464 populations (1000 repetitions) for the threshold self-regulation model ($k_i = 3$), and 518 repetitions (1000 repetitions) for the continuous self-regulation ($k_c = 2$) model. Other simulation parameters are $N = 50$, $u_0 = 10^{-4}$, $u_b = 5$, $v = 0$, and $s = -0.01$.

the stabilization of the copy number is slow and is only due to natural selection against TE accumulation. The last three models (hybrid dysgenesis IF = 29.1%, threshold self-regulation IF = 46.4% and continuous self-regulation IF = 51.8%) give very similar patterns compared to the previous cases, with a sharp initial increase in the mean copy number followed by a stabilization around 10 copies per individual. We can note that the between-repetitions variance is higher in the random-burst model, where, after 100 generations, mean copy number per individual can vary between 12.8 and 47.3. The TE invasion dynamics are more predictable in the other models (Figure 4a).

Figure 4b indicates the changes in mean transposition rates observed in populations. With the no-burst model

and the random-burst model, transposition rates remain constant. In the latter case, the transposition rate observed is the mean between a large number of individuals in which transposition is inactive and a few individuals with high-transposition activity. For the other three models, initial transposition rate is high and then decreases. After 20 or 30 generations following the introduction of 1 copy in an empty population, the mean transposition rate falls to usually measured rates ($<10^{-2}$); see Figure 4b. The dynamics of the system in the hybrid dysgenesis model are slower than those in the self-regulation models. This can be because transposition bursts concern all individuals in the self-regulation models, but only half of them in hybrid dysgenesis (since only 50% of the crosses can be dysgenic).

FIGURE 5.—Maintenance frequencies after 1000 generations for the four models. (a) Random burst model; (b) hybrid dysgenesis model; (c) threshold model; (d) continuous regulation model. Two parameters are allowed to vary, the burst transposition rate (u_b) and the basal transposition rate (u_0). Population size N is fixed to 50, and selection coefficient $s = -0.01$.

Loss frequency and invasion capacity: To analyze the interactions between the burst transposition rate and the basal transposition rate, some simulations have been completed for all previously described models. The monitored criterion is the maintenance frequency, which corresponds to the frequency of repetitions where the element is still present after 1000 generations. This excludes (i) the populations in which the TEs have been lost through genetic drift and (ii) the situations where the TE copy number increased until the population became extinct.

When there are no transposition bursts, the maintenance frequency is close to 0 for all transposition rates (results not shown): low-transposition-rate elements are lost through genetic drift, and very active elements can amplify themselves without control and lead to population extinction. The results presented in Figure 5 show very similar profiles for the other models. The conditions required for the TE maintenance are a low basal transposition rate ($u_0 < 10^{-2}$) and an optimal burst transposition rate ($0.5 < u_b < 5$), corresponding roughly to the optimal transposition rate illustrated in Figure 1. Nevertheless, a variation in the invasion efficiency between those models can be detected. With a random-burst model (Figure 5a), the frequency of TE loss by genetic drift is still high

(~80%) even under optimum conditions; on the other hand, the two self-regulation models (Figure 5, c and d) give the higher invasion rate. Hybrid dysgenesis (Figure 5b) shows intermediate results.

DISCUSSION

With this model, we attempted to highlight the early dynamics of a TE that was introduced into a new sexual, diploid species and to provide qualitative information about successful parasitic DNA invasion frequencies. It is unusual to consider that most HTs lead to the loss of the element, since genomes contain only successful “winning” ones. Today, almost all the literature on theoretical TE dynamics and evolution is based on TEs that are already established in the populations. In this work, the first step of the evolution of a newly transferred TE is developed. Our results suggest that this early invasion can represent a challenge for the TE, which has to develop specific abilities, such as a very high initial transposition rate in its new genomic environment. Further investigations in this domain might bring forth important information about the differences among TE family contents in the genomes of living organisms.

Using such a general model to describe a specific

system like TE-host interactions implies some simplifications. The confrontation between the modeling choices and the available data can provide interesting information about the solidity/robustness of the results shown above. However, observations and data mainly concern the genus *Drosophila*, and extrapolation to other groups of organisms can be speculative (KIDWELL and EVGEN'EV 1999; EICKBUSH and FURANO 2002).

Transposable elements: The generic term “transposable elements” usually refers to different self-replicable DNA sequences. These sequences are generally split into two groups: class I elements, which transpose via a RNA intermediate by a “copy and paste” mechanism; and class II elements, whose movements are due to a “cut and paste” process and thus transpose via a DNA molecule (see CRAIG *et al.* 2002 for review). Despite the fundamental structural and mechanistic differences between these two classes, a transposable element can be considered as a genetic entity with a duplication (u) and a deletion (v) ability. This modeling option is generally used in theoretical dynamics studies (CHARLESWORTH and CHARLESWORTH 1983; LANGLEY *et al.* 1983), partly because class I and class II TEs do not seem to show obvious dynamic specificities (HUA-VAN *et al.* 2005). Here, v has been set to 0, because deletion rates are usually negligible compared to duplication rates—at least one order of magnitude below them (EGGLESTON *et al.* 1988; SUH *et al.* 1995).

Moreover, in both classes, several nonautonomous elements exist, as short interspersed elements (WEINER 2002) or miniature inverted-repeat transposable elements (FESCHOTTE *et al.* 2002). In this article, these last elements have not been taken into account, since they are unable to colonize a genome after a horizontal transfer (they are unable to transpose without any corresponding autonomous element in the same genome). Thus, the transposition rate u corresponds here to the probability of an autonomous element duplicating into two autonomous elements. The emergence of such TE “parasites” in a population where autonomous elements are already established is another problem and has been treated in some previous studies (BROOKFIELD 1996; QUESNEVILLE and ANXOLABÉHÈRE 2001).

Demography: To limit the computer simulation time, population sizes have been reduced to 50 individuals. These sizes are clearly underestimated, compared to known population sizes in *Drosophila* species (generally $\sim 10^6$; LI *et al.* 1999; VIEIRA and HOIKKALA 2001) or even in mammals (10^4 in hominoids, for example; YANG 2002). However, when the transposition rate is relatively high (corresponding to burst transposition rates), the influence of population size decreases. Random genetic drift is important in the very first generations, until several individuals carry some copies. According to the burst model, generally <10 individuals with TEs are necessary to decrease the loss probability $<5\%$, whatever the population size. Then, when the elements can no

longer be lost through drift, the dynamics become more predictable (*i.e.*, invasion cannot be stopped).

The term “horizontal transfer” is used here as a generic way to refer to the initial introduction of a TE in a species devoid of it. This includes “true” horizontal transfers, mediated by a vector, or the *de novo* appearance of an autonomous intragenomic self-replicator, since it must have the same consequences for the new host species. TEs, like other genes, can also pass from one species to another one by introgression, provided fertile hybrids are present. However, these hybrids often show genomic deregulation symptoms and transposition bursts (LABRADOR *et al.* 1999). Therefore, the invasion of the new species must start under more complex conditions than one copy in one individual. A similar situation is encountered with intraspecific migrations, for which the migrant individual can bear several TEs, which is not the case with horizontal transfer. Moreover, it can bring both TEs and regulator genes, which can disturb the invasion dynamics (QUESNEVILLE and ANXOLABÉHÈRE 1998). In addition, contrary to HTs, migrations are frequent events; even if an element is unable to actively invade a new population, it will certainly colonize it randomly due to recurrent arrivals of individuals with TEs.

Selection: It is widely accepted that TEs have a negative impact on genomes. However, the exact mechanisms implied by this decrease in fitness are not yet perfectly known (NUZHIDIN 1999). Here, we use a simple additive model, which can be considered as an example of the insertion selection model (CHARLESWORTH and CHARLESWORTH 1983). This implies that (i) the impact of the different insertions along the genome is cumulative and linear (no epistasis), and (ii) there is no dominance. This choice appears to be quite arbitrary, since the additivity of selective effects of TEs is far from being demonstrated. However, when the copy number n is low, as is the case after a HT, the additive fitness function ($\omega = 1 + ns$), the “independent” fitness function ($\omega = (1 + s)^n$), or a “log concave” function ($\omega = 1 + sn^c$; *cf.* CHARLESWORTH and CHARLESWORTH, 1983) does not show striking differences, and the initial dynamics of TEs might be very similar in each case. Concerning dominance, it is known that random deleterious mutations are generally recessive, but some evidence tends toward the idea that it may not be the case for transposable elements in *Drosophila* (FRY and NUZHIDIN 2003).

The amount of fitness decrease due to the presence of each copy has been fixed at 1%. This value corresponds to the order of magnitude of mean deleterious effects estimations of spontaneous mutations in *Drosophila* (KEIGHTLEY 1994; CHARLESWORTH *et al.* 2004) and also to the order of magnitude of the insertion effect for P (1%; EANES *et al.* 1988) or *copia* (0.76%; HOULE and NUZHIDIN 2004) elements. However, this value must be considered carefully. On one hand, other studies suggest a much higher value (5.5% in heterozygous and

12.2% in homozygous; MACKAY *et al.* 1992) or much lower mean deleterious effect of TEs ($\sim 2 \times 10^{-4}$ per copy; CHARLESWORTH 1996). On the other hand, fixing a constant deleterious effect for each element is certainly a simplification. Some TE insertions might be lethal, and others almost neutral. In any case, we have to keep in mind that, with an additive deleterious effect of 1% per copy, the maximum number of TEs from all families that are potentially supported by an individual is 100, which is very low compared to known copy numbers (DEININGER and ROY-ENGEL 2002). This suggests that, even if the mean selective effect of a TE is negative, copies persisting in a genome are the least deleterious ones, and some of them can even appear to be beneficial for the host (MASIDE *et al.* 2002; SCHLENKE and BEGUN 2004).

Transposition bursts: Transposition rates in natural and laboratory *Drosophila* populations are relatively low, $\sim 10^{-4}$ transposition events per copy per generation (CHARLESWORTH *et al.* 1992; SUH *et al.* 1995; NUZHIDIN and MACKAY 1995; MASIDE *et al.* 2000), and they do not seem to be very different for class I and class II elements. Deletion rates are often several orders of magnitude below them and are nearly undetectable. However, higher mobilization rates have been obtained under certain conditions: environmental stress, genomic stress, or hybrid dysgenesis, for example. The transposition rates thus measured can reach averages $\sim 10^{-2}$ for *P* elements under dysgenesis conditions (EGGLESTON *et al.* 1988; BIÉMONT 1994), and individual measurements of very active *I* elements can increase $> 10^{-1}$ (SELEME *et al.* 1999; VASILYEVA *et al.* 1999).

With the random-burst model, described, for instance, in NUZHIDIN (1999), TEs never transpose, except in a few genetically disposed individuals, as suggested by observations such as those of GERASIMOVA *et al.* (1984). This represents an attractive model that accounts for the transposition rate variation between laboratory lines (NUZHIDIN 1999) or natural populations (BIÉMONT *et al.* 2003). However, it is not sufficient to explain the invasion after an HT: the proportion of “transpositionally active” individuals in a population in which the mean transposition rate is 10^{-4} (provided $u_0 = 0$ and $u_b = 1$) is 1/10,000. This is too low to account for any effective invasion, since we obtained some fairly weak invasion probabilities ($< 30\%$) with burst frequencies as much as 1/10. Moreover, it supposes a constant mean transposition rate (*cf.* Figure 4b), when transposition rates are supposed to vary with the genomic environment and notably with the TE copy number (NUZHIDIN *et al.* 1996; PASYUKOVA *et al.* 1998).

Hybrid dysgenesis is certainly the most well-documented burst system (BREGLIANO and KIDWELL 1983; BUCHETON 1990). It has been demonstrated only in a few *Drosophila* species with a few TE families (*P*, *hobo*, and *I* elements in *D. melanogaster* and several other elements in *D. virilis*; VIEIRA *et al.* 1998). However, it might

be a more general phenomenon than was previously thought since it can be observed after an artificial transfer of *P* element into *D. simulans* (MONTCHAMP-MOREAU 1990) and might concern more than one TE family in some cases (PETROV *et al.* 1995). Moreover, hybrid dysgenesis can be observed only during the period where some populations have the element when others do not (this period lasted only 40 years for the *P* element in *D. melanogaster*). This period is extremely short and thus reduces the probability of observing this phenomenon in a less-intensively studied species.

Both continuous and threshold regulation have been described here as self-regulation models. This means that the transposition rate decrease is considered to be due to an intrinsic feature of the element. However, it is known that the host is also able to repress transposition by various systems, such as gene silencing (JENSEN *et al.* 2002) or RNAi (PLASTERK 2002). Nevertheless, the results presented above are not expected to be affected by the precise regulation machinery, since regulation has been implemented simply as the transposition rate variation following the genomic copy number. Consequently, several regulation systems could fit the regulation functions, and self-regulation refers to a copy-number-dependent regulation rather than to a precise molecular mechanism.

The continuous self-regulation model is the one generally used when self-regulation is modeled (CHARLESWORTH and CHARLESWORTH 1983). Despite some contradictory evidence (NUZHIDIN *et al.* 1996; VIEIRA and BIÉMONT 1997), the control of the transposition activity when the copy number increases appears to be a widespread phenomenon (LABRADOR and CORCES 1997). The putative molecular mechanisms may imply production of repressors, as for the well-documented *P* element (LEMAITRE *et al.* 1993), or overproduction inhibition for the *mariner* element (LOHE and HARTL 1996). Threshold self-regulation is a less widely considered model, even if some results suggest that transposition should be abruptly repressed over a certain threshold by various mechanisms (cosuppression, methylation, repeat-induced point mutations, etc.). However, the limit between continuous and threshold self-regulation remains quite imprecise, experimentally but also theoretically, since these two models give similar theoretical dynamic patterns.

These results lead to the supposition that most TEs (at least those that can “jump” between species) must exhibit burst abilities. The selective pressure on TEs during the first steps of the invasion concerns their burst capacity, but several burst mechanisms should be equally used, since the different burst “strategies” do not show striking efficiency differences, except for the random-burst model, which appears to be less effective. Therefore, each TE family may show various adaptive responses, and hybrid dysgenesis might be only a way to increase the initial transposition rate of an element.

Known TEs, *i.e.*, those found in the genome of present species, are certainly those that were able to implement such an effective invasion strategy, the others being lost through random genetic drift. Nevertheless, after a successful invasion, a TE will evolve without any selective pressure on its burst ability during as many as several thousand generations or even more. In that time, random mutations and intragenomic selection might modify this invasion ability, since no further selection takes place until the next HT. The maintenance of this burst capacity over the evolutive scale thus may appear to be rather paradoxical and remains an open question. Are these transposition bursts consequently an adaptation of TEs to the colonization of new species or populations? It is difficult to answer, because the regulation of the transposition rate implies both the TE and the host. Indeed, both have a "common advantage" to limit the transposition amplification, since TEs that are too aggressive will destroy their host and will disappear with it. On the other hand, species that are too repressive can cause the loss of the mobile elements from their genomes and thus deprive them of an important source of genetic diversity and genomic novelties (KAZAZIAN 2004). Therefore, the transposition patterns actually observed may represent a complex trade-off between host and TE's respective "interests," with an inextricable mix between, on the one hand, immediately advantageous dynamics but without evolutionary future and, on the other hand, long-term stabilizing strategies maintained by clade selection.

We thank Hadi Quesneville for helpful discussions. We also acknowledge two anonymous reviewers who provided useful comments on the earlier version of the manuscript. The English text was reviewed by Malcolm Eden.

LITERATURE CITED

- ANXOLABÉHÈRE, 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.
- BIÉMONT, C., 1994 Dynamic equilibrium between insertion and excision of *P* elements in highly inbred lines from an *M'* strain of *Drosophila melanogaster*. *J. Mol. Evol.* **39**: 466–472.
- BIÉMONT, C., F. LEMEUNIER, M. P. GARCIA GUERREIRO, J. F. BROOKFIELD, C. GAUTIER *et al.*, 1994 Population dynamics of the *copia*, *mdg1*, *mdg3*, *gypsy*, and *P* transposable elements in a natural population of *Drosophila melanogaster*. *Genet. Res.* **63**: 197–212.
- BIÉMONT, C., C. NARDON, G. DECELIÈRE, D. LEPETIT, C. LOEVENBRUCK *et al.*, 2003 Worldwide distribution of transposable element copy number in natural populations of *Drosophila simulans*. *Evolution* **57**: 159–167.
- BREGLIANO, J.-C., and M. G. KIDWELL, 1983 Hybrid dysgenesis determinants, p. 363 in *Mobile Genetic Elements*, edited by J. A. SHAPIRO. Academic Press, San Diego.
- BROOKFIELD, J. F. Y., 1996 Models and spread of non-autonomous selfish transposable elements when transposition and fitness are coupled. *Genet. Res.* **67**: 199–209.
- BROOKFIELD, J. F. Y., and R. M. BADGE, 1997 Population genetics models of transposable elements. *Genetica* **52**: 281–294.
- BUCHETON, A., 1990 *I* transposable elements and *I-R* hybrid dysgenesis in *Drosophila*. *Trends Genet.* **6**: 16–21.
- CAPY, P., G. GASPERI, C. BIÉMONT and C. BAZIN, 2000 Stress and transposable elements: Co-evolution or useful parasites? *Heredity* **85**: 101–106.
- CHARLESWORTH, B., 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genet. Res.* **68**: 131–149.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1983 The population dynamics of transposable elements. *Genet. Res.* **42**: 1–27.
- CHARLESWORTH, B., A. LAPID and D. CANADA, 1992 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. ii: Inferences on the nature of selection against elements. *Genet. Res.* **60**: 115–130.
- CHARLESWORTH, B., P. SNIEGOWSKI and W. STEPHAN, 1994 The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**: 215–220.
- CHARLESWORTH, B., H. BORTHWICK, C. BARTOLOMÉ and P. PIGNATELLI, 2004 Estimates of genomic mutation rate for detrimental alleles in *Drosophila melanogaster*. *Genetics* **167**: 815–826.
- COEN, D., B. LEMAITRE, M. DELATTRE, H. QUESNEVILLE, S. RONSSERAY *et al.*, 1994 *Drosophila P* element: transposition, regulation and evolution. *Genetica* **93**: 61–78.
- CRAIG, N. L., R. CRAIGIE, M. GELLERT and A. M. LAMBOWITZ, 2002 *Mobile DNA II*. American Society for Microbiology Press, Washington, DC.
- DAWKINS, R., 1976 *The Selfish Gene*. Oxford University Press, London/New York/Oxford.
- DEININGER, P. L., and A. M. ROY-ENGEL, 2002 Mobile elements in animals and plants, pp. 1074–1002 in *Mobile DNA II*, edited by N. CRAIG, R. CRAIGIE, M. GELLERT and A. LAMBOWITZ. American Society for Microbiology Press, Washington, DC.
- DOOLITTLE, W., and C. SAPIENZA, 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603.
- EANES, W. F., C. WESLEY, J. HEY and D. HOULE, 1988 The fitness consequences of *P* element insertion in *Drosophila melanogaster*. *Genet. Res.* **52**: 17–26.
- EDWARDS, R., and J. BROOKFIELD, 2003 Transiently beneficial insertions could maintain mobile DNA sequences in variable environments. *Mol. Biol. Evol.* **20**: 30–37.
- EGGLESTON, W., D. JOHNSON-SCHLITZ and W. ENGELS, 1988 *P-M* hybrid dysgenesis does not mobilize other transposable element families in *Drosophila melanogaster*. *Nature* **331**: 368–370.
- EICKBUSH, T., and A. FURANO, 2002 Fruit flies and humans respond differently to retrotransposons. *Curr. Opin. Genet. Dev.* **12**: 669–674.
- FESCHOTTE, C., X. ZHANG and S. R. WESSLER, 2002 Miniature inverted-repeat transposable elements and their relationships to established DNA transposons, pp. 1147–1158 in *Mobile DNA II*, edited by N. CRAIG, R. CRAIGIE, M. GELLERT and A. LAMBOWITZ. American Society for Microbiology Press, Washington, DC.
- FINNEGAN, D. J., 1992 *The Genome of Drosophila melanogaster*, pp. 1096–1107. Academic Press, San Diego.
- FRY, J., and S. NUZHIDIN, 2003 Dominance of mutations affecting viability in *Drosophila melanogaster*. *Genetics* **163**: 1357–1364.
- GARDNER, M., N. HALL, E. FUNG, O. WHITE, M. BERRIMAN *et al.*, 2002 Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**: 498–511.
- GERASIMOVA, T. I., L. J. MIZROKHI and G. P. GEORGIEV, 1984 Transposition bursts in genetically unstable *Drosophila melanogaster*. *Nature* **309**: 714–716.
- HICKEY, D. A., 1982 Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**: 519–531.
- HOOGLAND, C., and C. BIÉMONT, 1996 Chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for maintenance of insertion site number. *Genetics* **144**: 197–204.
- HOULE, D., and S. V. NUZHIDIN, 2004 Mutation accumulation and the effect of *copia* insertions in *Drosophila melanogaster*. *Genet. Res.* **83**: 7–18.
- HUA-VAN, A., A. LE ROUZIC, C. MAISONHAUTE and P. CAPY, 2005 Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenet. Genome Res.* (in press).
- JENSEN, S., M. GASSAMA, X. DRAMARD and T. HEIDMANN, 2002 Regulation of *I*-transposon activity in *Drosophila*: evidence for cosuppression of nonhomologous transgenes and possible role of ancestral *I*-related pericentromeric elements. *Genetics* **162**: 1197–1209.
- KAPLAN, N., T. DARDEN and C. H. LANGLEY, 1985 Evolution and

- extinction of transposable elements in Mendelian populations. *Genetics* **109**: 459–480.
- KAZAZIAN, JR., H. H., 1998 Mobile elements and disease. *Curr. Opin. Genet. Dev.* **8**: 343–350.
- KAZAZIAN, JR., H. H., 2004 Mobile elements: drivers of genome evolution. *Science* **303**: 1626–1632.
- KEIGHTLEY, P. D., 1994 The distribution of mutation effects on viability of *Drosophila melanogaster*. *Genetics* **138**: 1315–1322.
- KIDWELL, M. G., 1992 Horizontal transfer. *Curr. Opin. Genet. Dev.* **2**: 868–873.
- KIDWELL, M. G., and M. B. EVGEN'EV, 1999 How valuable are model organisms for transposable element studies? *Genetica* **107**: 103–111.
- LABRADOR, M., and V. G. CORCES, 1997 Transposable element-host interactions: regulation of insertion and excision. *Annu. Rev. Genet.* **31**: 381–404.
- LABRADOR, M., M. FARRE, F. UTZET and A. FONTDEVILA, 1999 Interspecific hybridization increases transposition rates of *Osvado*. *Mol. Biol. Evol.* **16**: 931–937.
- LANGLEY, C. H., J. F. Y. BROOKFIELD and N. KAPLAN, 1983 Transposable elements in Mendelian populations. I. A theory. *Genetics* **104**: 457–471.
- LEMAITRE, B., S. RONSSERAY and D. COEN, 1993 Maternal repression of the P element promoter in the germline of *Drosophila melanogaster*: a model for the P cytotyping. *Genetics* **135**: 149–160.
- LI, Y.-J., Y. SATTI and N. TAKAHATA, 1999 Paleo-demography of the *Drosophila melanogaster* subgroup: application of the maximum likelihood method. *Genes Genet. Syst.* **74**: 117–127.
- LOHE, A. R., and D. L. HARTL, 1996 Autoregulation of *mariner* transposase activity by overproduction and dominant negative complementation. *Mol. Biol. Evol.* **13**: 549–555.
- MACKAY, T., R. LYMAN and M. JACKSON, 1992 Effects of P element insertions on quantitative traits in *Drosophila melanogaster*. *Genetics* **130**: 315–332.
- MASIDE, X., S. ASSIMACOPOULOS and B. CHARLESWORTH, 2000 Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster*. *Genet. Res.* **75**: 275–284.
- MASIDE, X., C. BARTOLOME and B. CHARLESWORTH, 2002 Selement insertions are associated with the evolution of the *Hsp70* genes in *Drosophila melanogaster*. *Curr. Biol.* **12**: 1686–1691.
- MCCLEINTOCK, B., 1984 The significance of responses of the genome to challenge. *Science* **226**: 792–801.
- MONTCHAMP-MOREAU, C., 1990 Dynamics of P-M hybrid dysgenesis in P-transformed lines of *Drosophila simulans*. *Evolution* **44**: 194–203.
- NUZHIDIN, S. V., 1999 Sure facts, speculations, and open questions about evolution of transposable elements. *Genetica* **107**: 129–137.
- NUZHIDIN, S. V., and T. F. C. MACKAY, 1995 The genomic rate of transposable element movement in *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**: 180–181.
- NUZHIDIN, S. V., E. PASYUKOVA and T. F. C. MACKAY, 1996 Positive association between *copia* transposition rate and copy number in *Drosophila melanogaster*. *Proc. R. Soc. Lond. Ser. B* **263**: 823–831.
- ORGEL, L. E., and F. H. C. CRICK, 1980 Selfish DNA: the ultimate parasite. *Nature* **284**: 604–607.
- PASYUKOVA, E., S. NUZHIDIN and D. FILATOV, 1998 The relationship between the rate of transposition and transposable element copy number for *copia* and *Doc* retrotransposons of *Drosophila melanogaster*. *Genet. Res.* **72**: 1–11.
- PETROV, D., J. SCHUTZMAN, D. HARTL and E. LOZOVSKAYA, 1995 Diverse transposable elements are mobilized in hybrid dysgenesis in *Drosophila virilis*. *Proc. Natl. Acad. Sci. USA* **92**: 8050–8054.
- PLASTER, R., 2002 RNA silencing: the genome's immune system. *Science* **296**: 1263–1265.
- QUESNEVILLE, H., and D. ANXOLABÉHÈRE, 1998 Dynamics of transposable elements in metapopulations: a model of P elements invasion in *Drosophila*. *Theor. Popul. Biol.* **54**: 175–193.
- QUESNEVILLE, H., and D. ANXOLABÉHÈRE, 2001 Genetic algorithm-based model of evolutionary dynamics of class II transposable elements. *J. Theor. Biol.* **213**: 21–30.
- QUESNEVILLE, H., D. NOUAUD and D. ANXOLABÉHÈRE, 2003 Detection of new transposable element families in *Drosophila melanogaster* and *Anopheles gambiae* genomes. *J. Mol. Evol.* **57** (Suppl 1): S50–S59.
- SCHLENKE, T., and D. BEGUN, 2004 Strong selective sweep associated with a transposon insertion in *Drosophila simulans*. *Proc. Natl. Acad. Sci. USA* **101**: 1626–1631.
- SELEME, M., I. BUSSEAU, S. MALINSKY, A. BUCHETON and D. TENINGES, 1999 High-frequency retrotransposition of a marked I factor in *Drosophila melanogaster* correlates with a dynamic expression pattern of the ORF1 protein in the cytoplasm of oocytes. *Genetics* **151**: 761–771.
- SILVA, J., E. LORETO and J. CLARK, 2004 Factors that affect the horizontal transfer of transposable elements. *Curr. Issues Mol. Biol.* **6**: 57–71.
- SUH, D., E. CHOI, T. YAMAZAKI and K. HARADA, 1995 Studies on the transposition rates of mobile genetic elements in a natural population of *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**: 748–758.
- TSITRONE, A., S. CHARLES and C. BIÉMONT, 1999 Dynamics of transposable elements under the selection model. *Genet. Res.* **74**: 159–164.
- UYENOYAMA, M. K., 1985 Quantitative models of hybrid dysgenesis: rapid evolution under transposition, extrachromosomal inheritance, and fertility selection. *Theor. Popul. Biol.* **27**: 176–201.
- VASILYEVA, L. B., E. V. BUBENSHCHIKOVA and V. A. RATNER, 1999 Heavy heat shock induced retrotransposon transposition in *Drosophila*. *Genet. Res.* **74**: 111–119.
- VIEIRA, C., and C. BIÉMONT, 1997 Transposition rate of the 412 retrotransposable element is independent of copy number in natural populations of *Drosophila simulans*. *Mol. Biol. Evol.* **14**: 185–188.
- VIEIRA, J., and A. HOIKKALA, 2001 Variability levels, population size and structure of American and European *Drosophila montana* populations. *Heredity* **86**: 506–511.
- VIEIRA, J., C. VIEIRA, D. HARTL and E. LOZOVSKAYA, 1998 Factors contributing to the hybrid dysgenesis syndrome in *Drosophila virilis*. *Genet. Res.* **71**: 109–117.
- VINOGRADOV, A., 2003 Selfish DNA is maladaptive: evidence from the plant red list. *Trends Genet.* **19**: 609–614.
- WEINER, A., 2002 SINEs and LINEs: the art of biting the hand that feeds you. *Curr. Opin. Cell Biol.* **14**: 343–350.
- YANG, Z., 2002 Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics* **162**: 1811–1823.

Communicating editor: T. H. EICKBUSH