Experimental evolution reveals hyperparasitic interactions among transposable elements

Émilie Robillard[a], Arnaud Le Rouzic[a], Zheng Zhang[a], Pierre Capy[a], and Aurélie Hua-Van[a,1]

[a]Évolution, Génomes, Comportement, Écologie, CNRS, Institut de Recherche pour le Développement, Université Paris-Sud, Université Paris-Saclay, 91198 Gif-sur-Yvette, France

Transposable elements (TEs) are repeated DNA sequences that can constitute a substantial part of genomes. Studying TEs' activity, interactions, and accumulation dynamics is thus of major interest to understand genome evolution. Here, we describe the transposition dynamics of cut-and-paste *mariner* elements during experimental (short- and longer-term) evolution in *Drosophila melanogaster*. Flies with autonomous and nonautonomous *mariner* copies were introduced in populations containing no active *mariner*, and TE accumulation was tracked by quantitative PCR for up to 100 generations. Our results demonstrate that (*i*) active *mariner* elements are highly invasive and characterized by an elevated transposition rate, confirming their capacity to spread in populations, as predicted by the "selfish-DNA" mechanism; (*ii*) nonautonomous copies act as parasites of autonomous *mariner* elements by hijacking the transposition machinery produced by active *mariner*, which can be considered as a case of hyperparasitism; (*iii*) this behavior resulted in a failure of active copies to amplify which systematically drove the whole family to extinction in less than 100 generations. This study nicely illustrates how the presence of transposition-competitive variants can deeply impair TE dynamics and gives clues to the extraordinary diversity of TE evolutionary histories observed in genomes.

transposable elements | hyperparasitism | *Drosophila* | experimental evolution | invasion dynamics

The evolutionary factors explaining the distribution of transposable elements (TEs) across organisms are still poorly understood (1). TEs are mobile DNA sequences able to invade populations and to duplicate within genomes by various molecular mechanisms (2) and can be found in multiple copies in virtually all living species. However, the nature and abundance of TEs vary substantially throughout the tree of life (3). Although most prokaryotes harbor only a few insertion sequences, large eukaryotic genomes (including plants, amoeba, or animals) may contain up to 80% of TE-derived sequences.

TEs are often considered as selfish-DNA sequences, meaning that they have a greater chance of being transmitted to the progeny than nonselfish sequences (4, 5). In this hypothesis, the ubiquitous presence of TEs can be satisfactorily explained without adaptationist hypotheses (6). The underlying driving mechanism is replicative transposition, which has two combined consequences: (*i*) an inflation of copy number per genome over time; and then, in sexual populations, (*ii*) a tendency of TE copies to be transmitted to the progeny more efficiently than Mendelian factors. Replicative transposition theoretically allows the invasion of populations from a single individual, despite establishment of efficient host regulation; natural selection against deleterious insertions and high TE load; or transposition-related or -unrelated recombination, excision, and deletion (7–9).

Of particular interest is that TE copies from the same family, although derived from a common ancestor, do not necessarily cooperate (10). Whatever the molecular mechanism (e.g., copy-and-paste or cut-and-paste), transposition requires the production of one or several proteins encoded by the TE itself

(11). These proteins may promote the amplification of any similar copies, including those that do not produce any functional transposition machinery. Such nonautonomous copies may thus proliferate, provided that at least one active copy is present in the genome. Nonautonomous copies are often very successful and can even out-compete autonomous copies (12, 13). Because both autonomous and nonautonomous copies compete for the same transposition machinery, it is tempting to speculate that the invasion of autonomous copies may be slowed by the presence of nonautonomous copies. Theoretical models have confirmed that such competition could alter considerably the evolutionary dynamics (14–18), and the presence of nonautonomous competitors may be a major explanatory factor for the fact that a given TE may be extremely successful in some species whereas performing poorly in others.

Interestingly, despite its theoretical relevance to understanding genome evolution, there is very little direct experimental support for such a negative interaction between autonomous and nonautonomous copies. The original cut-and-paste *mariner* transposon, identified first in *Drosophila*, appears as a good model to experimentally test this assumption. Indeed, two distinct *mariner* sequences have been isolated from *Drosophila mauritiana*, a sister species of *Drosophila melanogaster* (19, 20). Both copies are full-length, but one (*peach*) is nonautonomous, unable to promote its own transposition due to nonsynonymous substitutions, whereas the other (*Mos1*) is an autonomous copy able to cross-mobilize *peach* copies. *D. melanogaster* does not naturally carry *Mos1/peach*-related elements, but transgenic lines have been obtained with each of these copies.

Here, we used the *mariner* system in *D. melanogaster* through two series of experiments to study the capacity of *Mos1*-active

Significance

Transposable elements (TEs) are DNA sequences that colonize every genome and have a great impact on the genome evolution and structure. Here, we report experimental evolution results that confirm the intrinsic "selfish" properties of TEs in sexual populations. We also show how different kinds of copies from the same family strongly interfere: cheating nonautonomous copies parasitize autonomous ones, to the extent of endangering the survival of the whole TE family. These results nicely illustrate the "genome-ecology" analogy, according to which genome components are assimilated with interacting species in an ecosystem.

copies to invade the genome of *D. melanogaster* and to decipher the dynamic properties and evolutionary interactions between nonautonomous and autonomous elements. In both experimental setups, we introduced a single "migrant" carrying a few *Mos1* copies among flies deprived of active *mariner* elements but that may contain an inactive *peach* copy. First, we checked the ability of *Mos1* to invade empty populations, free of any kind of *mariner*, and quantified the selfish-DNA properties of this element from several hundred independent experiments, by computing the frequency at which TEs were still present after 10 generations (thereafter, "invasion frequencies" experiments). In the second set of experiments ("experimental evolution"), only a few migration events were reproduced. We either introduced migrant with active *Mos1* only in populations with no *mariner* at all or migrants containing both active and inactive copies in recipient populations containing only one inactive copy. We tracked the number of copies through genomic quantitative (q)PCRs for both *peach* and *Mos1* elements, independently, and when feasible, we followed the dynamics of the invasion process through phenotypic markers, for up to 100 generations.

Results

Invasion Frequencies. We initialized a total of 272 invasion experiments in which a *Mos1*-carrying migrant from two strains (male or nonvirgin female) was introduced into small populations of 9 flies without any *mariner* TEs (*SI Appendix, Supplementary Methods*). Experiments were maintained up to 10 generations. *Mos1* copy number was estimated by qPCR on the migrant, and the presence of *Mos1* was assayed by PCR in the progeny at generation (G)5 and G10 for each experiment separately. For statistical analysis, we only kept experiments in which the migrant copy number was between one and five (i.e., 47% of the initial experiments). There was a significant departure from the expected distribution of copy numbers, both for the mean (around seven copies in average, whereas only five were expected) and for the shape [a significant departure from the theoretical Poisson distribution (21); *SI Appendix, Supplementary Results*]. This result can be explained by replicative transposition during the initial crosses. We also removed experiments with poorly replicable PCRs or inconsistent scenarios [e.g., absence at G5 and presence at G10]. This procedure resulted in 94 experimental lines, representing 34% of the initiated experiments. On average, the *Mos1* element was maintained among 71% of these populations at G10 (Fig. 1; for more details, see *SI Appendix, Supplementary Methods* and *Supplementary Results*).

We used a binomial generalized linear model (GLM) to analyze invasion frequencies at G10 and observed a strong and significant sex effect (invasion frequency higher in females than in males) ($P = 0.002$), as well as an effect of the migrant copy

number ($P = 0.024$). The strain factor was not significant, and the data from the two strains were pooled for further analysis. We compared the invasion frequency of *Mos1* elements with theoretical expectations in absence of transposition, calculated through simulations (see Fig. 1 and *SI Appendix, Supplementary Results*), with two population sizes ($N_e = 10$ and $N_e = 20$). Both a binomial test and an exact distribution test using a Monte Carlo approach revealed that *Mos1* invaded populations more frequently than predicted in absence of transposition, even in the more conservative scenario ($N_e = 20$, two-tailed test, $P = 0.03$ for males and $P = 0.001$ for females). The same analysis at G5 leads to the same trends but lacks statistical power (*SI Appendix, Supplementary Results*) due to the lower frequency of theoretical loss by drift. The data confirm that *Mos1* is well suited to invade *D. melanogaster* naive populations, even when the initial copy number is low. The major factor conditioning the invasion success is the sex of the migrant, because females are very likely to reproduce, whereas the mating success of males is more stochastic. Nevertheless, even a low-copy number male migrant still has a >40% probability to trigger a successful TE invasion in the population (Fig. 1), illustrating the selfish-DNA efficiency of *Mos1* elements.

Experimental Evolution. Transposition is expected to increase both the probability of invasion (as evidenced in the invasion experiments) and the genomic copy number in each individual of the population. We thus ran long-term evolutionary experiments to track the dynamics of the average copy number, initiating populations with two types of copies (autonomous and nonautonomous) in migrants of different sexes (*SI Appendix, Tables S1 and S2*). For discriminating both types of copies (autonomous vs. nonautonomous elements) in the same population, we developed an efficient methodology based on qPCR. In parallel, we also followed the frequency of *Mos1* carriers during the first generations using a phenotypic assay (*SI Appendix, Supplementary Methods*).

Autonomous elements. We first monitored the amplification dynamics of *Mos1* elements alone in *Mos1*-free *D. melanogaster* populations. To increase the chance of successful invasions, we introduced only single wild-type *Mos1*-carrying nonvirgin females (the most favorable scenario for TE invasion) in different *mariner*-free strains [*yellow white* (*yw*) populations (experiment A1) or wild-type populations (experiment A2)], with the hope of detecting the potential influence of the genetic background and/or phenotypic markers. As a control, we ran experiment A3 initialized with 100 initial TE carriers, to study the amplification dynamics in a population already contains TEs. A1 and A2 migrants and A3 flies all resulted from successive backcrosses of the M19 strain with the A2 recipient strain.

All dynamics displayed a similar pattern, with a continuous increase in copy number for 50 generations (Fig. 2 *A–C*). There were significant differences between experiments [analysis of covariance (ANCOVA): $P = 10^{-4}$] but not between replicates of the same experiment ($P = 0.87$). From G10, we observed less *Mos1* copies in A1 than in A2 (differences in the intercept in a linear model; *t* test: $P = 0.012$) or A3 ($P < 0.001$). At G50, there were about 10 copies per haploid genome in A1 vs. about twice as much (20 per haploid genome) in A2 and A3 (Fig. 2 *A–C*). Furthermore, transposition rates calculated from a linear regression of log(copy number) over generations were lower for experiment A1 [0.013 transpositions per copy per generation; 95% confidence interval (CI): 0.006–0.021] than for experiment A2 (0.023; 95% CI: 0.018–0.028) or A3 (0.022; 95% CI: 0.020–0.029), the difference between A1 and A3 being statistically supported (*t* test: $P = 0.004$). More details are available in *SI Appendix, Supplementary Results*.
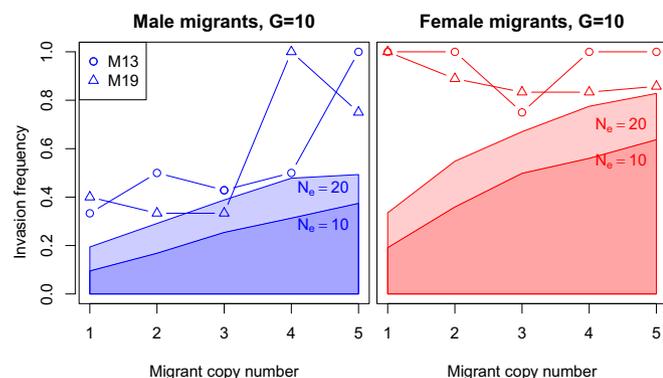
Fig. 1. Theoretical (filled areas) and empirical (symbols) invasion frequencies of *Mos1* brought by migrants carrying one to five copies.
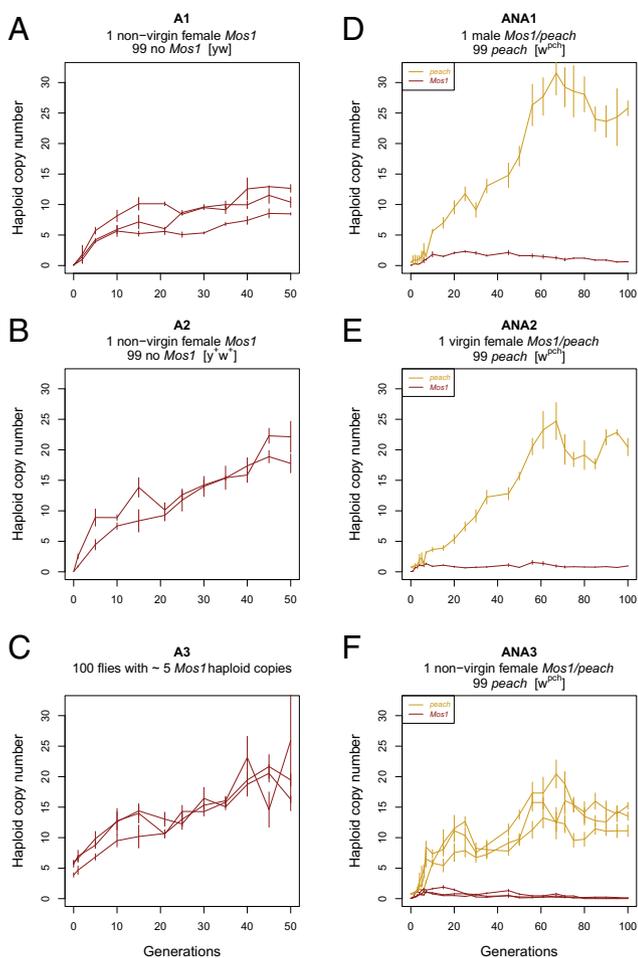
Fig. 2. (*A–C*) Copy number dynamics of the *Mos1* copies when the autonomous element is alone. Bars represent SEs estimated from the qPCR analysis. Graphs differentiate recipient populations with different genetic backgrounds and overlapping curves indicate independent replicates. (*A*) [*yw*] population with no *Mos1* copy. (*B*) [y^+w^+] population with no *Mos1* copy. (*C*) [y^+w^+] population containing *Mos1* copies. (*D–F*) Copy number dynamics when the migrant brings both autonomous *Mos1* (red) copies and nonautonomous *peach* (orange) copies. Recipient populations contained one *peach* copy per genome. Graphs correspond to different migrant categories. (*D*) Population initiated with one male. (*E*) Population initiated with one virgin female. (*F*) Populations initiated with one nonvirgin female.

Nonautonomous along with autonomous elements. We initialized 11 populations, all containing 1 inactive *peach* copy, whereas migrants consisted in 1 male (ANA1), 1 virgin female (ANA2), or 1 nonvirgin female (ANA3), all carrying few copies of both *Mos1* and *peach*. In such conditions, *Mos1* carriers were easily detected by their eye color (*SI Appendix, Supplementary Methods* and Fig. S2), and we could thus readily monitor the invasion success for different kinds of migrants. The invasion of *Mos1* was successful in one replicate of three for experiment ANA1, one of five for experiment ANA2, and three of three for experiment ANA3 (Fig. 2 *D–F*), confirming that starting with one nonvirgin female as a migrant is the most favorable condition. In the same way, as for the autonomous copies alone, there was a significant effect of the experiment for *Mos1* (ANCOVA, $P = 10^{-5}$) but not for *peach* ($P = 0.43$), and the experimental evolution was highly replicable (no effect of the replicate; $P = 0.59$ for *Mos1*; $P = 0.43$ for *peach*; *SI Appendix, Supplementary Results*).

For all experiments, the invasion dynamics of the *Mos1* element was fundamentally altered in presence of nonautonomous *peach* copies. The amplification of *Mos1* stopped rapidly (after less than 10 generations), and the copy number stabilized around 3 copies per haploid genome. Furthermore, the number of *Mos1* elements tended to stabilize or even decrease (Fig. 3*A*), and *Mos1* elements were virtually lost in all time series by G100 (less than one copy per haploid genome in experiments ANA1 and ANA2 and undetectable in all three ANA3 time series). The absence of active *Mos1* was confirmed for all experiments at G120 to G130 by phenotypic test crosses (*SI Appendix, Supplementary Methods*).

Conversely, nonautonomous *peach* copies amplified dramatically, from 1 copy at G0 to 15–30 copies per haploid genome by G60. After G60, the *peach* copy number stabilized, which is likely due to the loss of the source of transposase from *Mos1* (segmented regression: breakpoint at G62.4 ± 5.1; *SI Appendix, Supplementary Results*). Between G10 and G60, transposition rates were 0.040 (95% CI: 0.021–0.059) per *peach* copy per generation in experiment ANA1, 0.039 (95% CI: 0.033–0.045) in experiment ANA2 and 0.014 (95% CI: 0.008–0.021) in experiment ANA3 (Fig. 3*B*). The latter was significantly different from the first ones (*t* test from a linear model: $P < 0.001$). Thus, introducing more initial copies (in ANA3) might have impaired the invasion success of autonomous and nonautonomous copies (both in terms of transposition rate and final genomic TE content).

The major pattern emerging from experimental evolution is that actively transposing elements are *Mos1* copies when alone and *peach* copies when both autonomous and nonautonomous elements are introduced. Indeed, both standalone *Mos1* and nonautonomous *peach* transpose with approximately the same rate (around 0.02 duplication event per copy and per generation) (Fig. 3*B*). A one-way ANOVA considering three groups (ANA/*Mos1*, ANA/*peach*, and A/*Mos1*) highlighted significant differences in transposition rates between *Mos1* and *peach* within ANA (posthoc Tukey test, $P < 0.001$) and between ANA/*Mos1* and A/*Mos1* ($P < 0.001$) but no differences between ANA/*peach* and A/*Mos1* ($P = 0.84$). Interestingly, *Mos1* elements when *peach* are present display a negative transposition rate (i.e., they are more often deleted than amplified) (ANA1: −0.010; ANA2: −0.001; ANA3: −0.034; the negative rate being statistically significant for both ANA1 and ANA3; *SI Appendix, Supplementary Results*).

This result demonstrates the strong interaction between autonomous and nonautonomous copies, because autonomous *Mos1* elements stopped transposing in the presence of nonautonomous *peach* elements.

Early invasion. During the first generations, the estimation of transposition rate is complicated by the fact that some flies do not contain the autonomous element. Indeed, the rise in average copy number of *mariner* elements involves both the increase in copy number within TE-carrier individuals and the increase in frequency of TE carriers. However, we could disentangle both phenomena in ANA experiments, taking advantage of the phenotypic effect of *Mos1*-triggered excision of the *peach* copy, to distinguish between TE-carrier flies and -noncarrier flies (*SI Appendix*, Fig. S2). Indeed, in the presence of *Mos1*, *peach*, originally inserted into the *white* gene, excises, which restores the gene activity. Hence, excision is easily visualized by the eye color, and this system can be used as a phenotypic assay for testing transposition activity (20). Hence, we could estimate both the average copy number in *Mos1* carriers and their frequency in the population and thus estimate the real transposition rate among *Mos1* carriers. Fig. 3*C* shows the copy number of *Mos1* and *peach* copies among *Mos1*-containing flies, as well as theoretical predictions under the hypothesis that there is no transposition, no selection, and assuming random mating (see *SI Appendix, Supplementary Methods* for more details). The discrepancy between
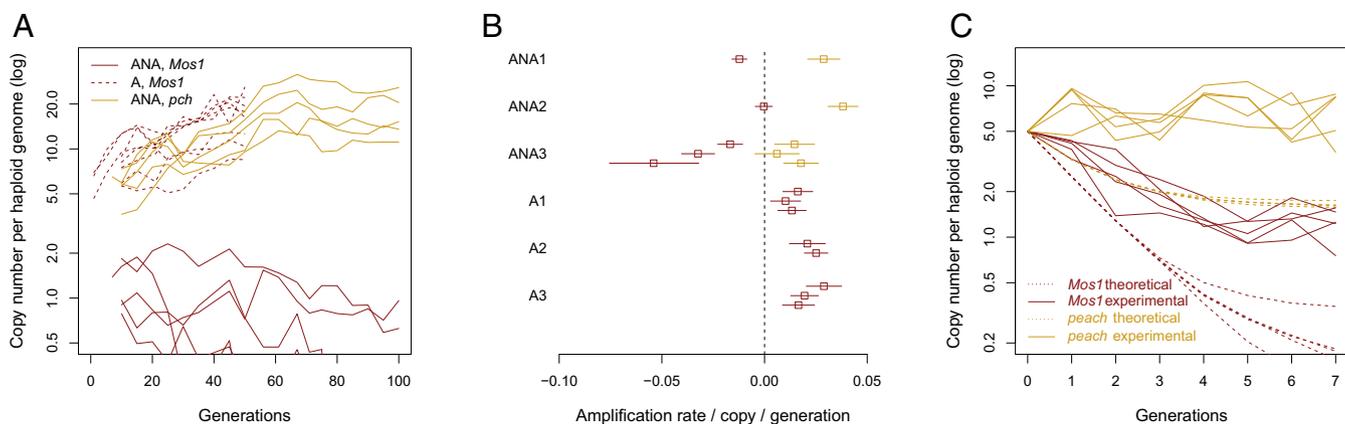
Fig. 3. (A) Dynamics of the number of *Mos1* and *peach* copies from the point where the element has invaded the population (>90% of the population carries at least one copy). The representation on the log scale allows the computation of the transposition rate with a linear regression. (B) Estimate and 95% CI of transposition rates (red: *Mos1* copies; orange: *peach* copies). For *peach* copies, generations posterior to 60 were not considered, because the invasion clearly stops at that point. The very first generations (during which some individuals in the population do not carry the element) were also discarded. (C) Copy number of *Mos1* and *peach* elements per carrying individual during the initial invasion of the element. Dashed lines show the expected number of copies in absence of transposition. The decrease is due to the higher probability of mating with a *Mos1*-empty fly during the first generation when the population is not completely invaded.

observed and theoretical copy numbers suggests a substantial rate of replicative transposition for both *Mos1* and *peach*. Average transposition rates calculated on these first generations were, for *Mos1*, 0.45 per copy and per generation in experiment ANA1 and 0.33 in experiments ANA2 and ANA3. For *peach*, these rates were respectively 0.75, 0.53, and 0.49 for ANA1, ANA2, and ANA3. Therefore, transposition rates in the very first generations of the invasion were at least one order of magnitude larger than their average over long-term experiments. We also checked that natural selection was not responsible for the seemingly elevated transposition rates. Even strong selection (s = 0.5) against *Mos1*-free individuals had a modest impact on transposition rate estimates (25% decrease in the transposition rate of *Mos1* and 7% increase in the transposition rate of *peach*). Hence, the hypothetical effect of natural selection is unlikely to affect our conclusions qualitatively.

Discussion

Our experimental results confirm and expand theoretical expectations on transposable element dynamics. First, we showed that active *mariner* elements behave exactly as expected under the selfish-DNA hypothesis: active transposition promotes both the invasion of the population (the frequency of TE carriers in the population increases deterministically) and the colonization of the genome (the number of copies per individual increases with time). Second, our results demonstrate a strong dynamical interaction between different kinds of copies (autonomous vs. nonautonomous), leading to specific evolutionary patterns depending on the presence of nonautonomous copies, which appear to efficiently act as parasites on autonomous elements.

Experimental Design. Testing the selfish-DNA hypothesis consisted of verifying that TEs are able to invade populations better than expected by drift only. We chose to obtain invasion frequencies under conditions (similar genetic backgrounds between migrant and recipient, two different starting strains, two independent replicates), allowing to rule out any confounding drive effect due to alleles that could be present in the migrant strain. Furthermore, indeed we did not detect any effect of the strains or of the replicates. However, this experimental design prevented us to use any neutral markers for estimating the drift force, and we compared the observed frequencies to simulations with arbitrary population sizes ($N_e = 10$ and $N_e = 20$). Although vials

may contain from a dozen to a hundred flies, these figures correspond to conservative assumptions for effective population sizes: simulated values assume a ratio N_e/N of about 0.1–0.2, which remains above empirical estimates in *Drosophila* (22, 23).

The strong effect of the migrant sex on invasion frequency is consistent with the theoretical difference obtained in simulations, due to the combined effect of (*i*) the fact that migrant females are already fertilized by TE-carrying males and (*ii*) the assumption that all females can lay eggs, whereas some males are excluded from the reproduction. Our experimental design makes it impossible to exclude the (likely) hypothesis of a different transposition rate in males vs. females. Furthermore, our results suggest that the number of copies carried by the migrant might be less important for females than males (Fig. 1), but the sex × copy interaction failed to reach statistical significance in the GLM analysis (P = 0.058).

In experimental evolution with competition, TE-invasion tracking was facilitated by phenotypic markers indicating the presence of active TEs in individuals. Although convenient, similar genetic systems have already been suspected to bias the results because TEs might also be driven in populations due to natural selection on marker phenotypes (24, 25). However, we deem it unlikely that the observed patterns could be explained by spurious selection: (*i*) in our system, phenotypic markers are not within the elements, and can then be easily decoupled from TE dynamics within a few generations due to sexual reproduction and recombination; (*ii*) the amplification dynamics and copy number in *yw* populations were never higher than in wild-type populations; and (*iii*) including selection in the formulas used to estimate short-term transposition rates shows that selection has a moderate effect on transposition rate estimates. Consequently, even if we cannot formally exclude a minor quantitative effect of selection, especially in the very first generations, we are confident that the observed dynamics are mainly driven by transposition.

Consistency with Existing Experimental Knowledge and Generalization. TE-invasion experiments in eukaryotes have already been carried out from active copies introduced by transformation, most of the time in *D. melanogaster* or close species (26). To our knowledge, few experiments have been designed to allow TEs to invade freely an empty population (e.g., ref. 27), and no experimental study has focused on the invasion frequency. If TE interactions have already been studied at the functional level (28, 29),

describing the interacting dynamics of several TEs sharing the same transposition machinery at the population level is a unique feature of our experimental design. Overall, the general pattern of active TE invasion is consistent across experiments. *P* elements from recent natural populations introduced into old laboratory populations of *D. melanogaster* tended to multiply up to 50 copies per genome (30), whereas the *hobo* element seemed to stay under 20 copies per genome (31). A decrease in the transposition rate with time is not necessarily observed; for instance, the *roo* retrotransposon has been shown to be able to accumulate more than 80 copies per genome in mutation-accumulation experiments and even more in specific genetic backgrounds (32). Our results suggest that the upper limit for *mariner* is around 30 copies.

Average transposition rates rarely exceed 10^{-3} events per copy and per generation when measured in natural populations (33, 34). Here, we observed replicative transposition rates ranging from 0.3–0.5 per copy and per generation during the very early stages of the invasion and 0.01–0.03 for the 50 subsequent generations. These figures are of the same order of magnitude as for active *P* elements during hybrid-dysgenesis stages recorded in the laboratory (35), although dysgenic symptoms were never observed for *mariner*.

The need to focus on a specific experimental setup, and, in particular, on a specific species–TE pair, necessarily raises issues related to the generality of the results. Here, the choice of *D. melanogaster* as a host species was driven by the facility of transformation and genetic manipulations in this model species. *D. melanogaster* is also known to be susceptible to TE invasion in the wild (36), and three new TEs have very recently (i.e., during historical times) colonized its genome: the *P* element (37), the *hobo* element (38), and the *I* element (39). It has also been shown experimentally that *D. melanogaster*'s *P* element was more efficient than in its sister species *Drosophila simulans* (40). The recent discovery of *P* element in natural populations of *D. simulans* (41) might help to confirm the effect of the host species on TE dynamics in the wild. In addition, *Mos1* is known to be an extremely active copy in *D. melanogaster* (42). In sum, the observed success of the experimental invasions might overestimate the activity compared with an average TE colonization in the wild. However, because observed transposition rates and final genomic copy numbers remain standard for laboratory studies in this species, our results are unlikely to be particularly unrealistic.

Transposition Regulation. In the experimental evolution experiments, we observed a high initial transposition rate during the first generations, followed by a systematic decrease. Rapid changes in transposition rates have also been observed for *P* elements and suggest the involvement of transposition regulation mechanisms. Two types of autoregulation (by copy number or transposase types) have been previously suspected for *Mos1*, based on genetic studies in *Drosophila*. The first is called overproduction inhibition [i.e., formation of inactive aggregates of the transposition machinery when too much transposase is produced (43)]. Although in vitro, cellular, or biochemical studies demonstrated an influence of MOS1 concentration on its cellular localization, and the synaptic complex formation, an effect on transposition rate was never observed (44–46). The second mechanism is dominant-negative complementation (43, 47) between *peach* and *Mos1* that could occur in competition experiments only. Indeed, the *peach* copy (differing from *Mos1* by 11 SNPs) is probably transcribed and translated like *Mos1*, generating inactive transposase monomers. With a large amount of *peach* copies, most active MOS1 monomers could be trapped into inactive dimers, decreasing the transposition efficiency.

Drosophila TEs are also known to be host-regulated by the PIWI pathway, mainly through maternal transmission of cytoplasmic small RNAs [PIWI-interacting (pi)RNAs] able to silence TEs on a sequence-specific basis. As seen for the *P* element-triggered hybrid-dysgenesis syndrome, progeny lacking the silencing maternally transmitted piRNAs displays high transposition rates of the father-transmitted TE and are characterized, for *P* elements at least, by various mutational defects (sterility, lethality, and developmental problems) (48). The silencing piRNAs in nondysgenic progeny emanate from the transcription of maternal genomic pi-clusters containing TE copies (9). pi-clusters containing *Mos1* could be present in our transgenic strain (carrying *Mos1* for about 15 y) and then in the migrants (despite several backcrossed against a *Mos1*-free strain) but not in the recipient population. An elevated transposition rate could then occur in some crosses, before the spread of this hypothetical *Mos1*-containing pi-cluster. Alternatively, a de novo insertion of *Mos1* into a pi-cluster would allow progressive establishment of silencing.

Conclusions

In genomes, some sequences survive by collaborating (such as genes contributing to the survival and reproduction of individuals), whereas others tend to develop conflicts with each other. It has been recently suggested that the relationships between genome components (including genes, transposable elements, or any sequence able to persist over evolutionary time) were similar to the relationships between individuals or species in ecosystems (10, 49, 50), although the possibility to apply ecological formalism to genome evolution remains questionable (51). Here, we brought substantial evidence that the relationships between autonomous and nonautonomous *mariner* TE copies were analogous to parasitism: *Mos1* copies (the "hosts") are able to survive and replicate by themselves, whereas *peach* copies (the "parasites") are unable to transpose without *Mos1* copies. When both copies are present in the same habitat (the genome), parasitic copies amplify, which strongly affects the survival and reproduction activity of the host copies. As active transposable elements themselves are often considered as parasites of the genome (52, 53), this genome-ecology analogy would define nonautonomous copies as hyperparasites.

Using *Drosophila* and *mariner* as an experimental model, we have been able to demonstrate the strong negative interaction between nonautonomous and autonomous copies of the same family. This interaction reveals a potential weakness of the otherwise efficient selfish strategy of TEs such as *mariner*, based on self-amplification and spreading through sexual reproduction. The fact that few mutations in a copy can have such dramatic consequence for the TE family gives some explanatory clues to the huge diversity of TE trajectories observed among species. The rapid loss of transposition activity leaves the genome with inactive copies that may stay for a while, be slowly eliminated by drift or by genome deletion, or occasionally reactivated with the arrival of a new active copy. This view is in accordance with genomic data showing that genomes are often riddled with TE remnants. However, this scenario is also counterintuitive because genomes may also contain numerous active TE lineages. With such a rapid inactivation process/loss of activity, the long-term survival of a TE is not uniquely dependent on its selfishness but also on the opportunity to frequently invade new genomes through horizontal transfers, which have been shown to be especially frequent in *Drosophila* for *mariner*-like elements (54).

Materials and Methods

Invasion Frequencies. Migrant flies containing on average 5 active copies were obtained by 3 successive backcrosses between a *Mos1* strain (about 40 copies) and the empty population. Populations were initiated by introducing 1 single male or female fly (migrant) carrying among 9 flies deprived of *mariner* TEs, keeping an even sex ratio (*SI Appendix, Supplementary Methods* and Table S1). The migrant, marked by cutting a small piece of wing, was

recovered after 3–6 d and analyzed by qPCR to precisely quantify the exact copy number (55). For each generation (every 10–12 d), newly emerged flies (between 30 and a few hundred) were transferred into a new vial with fresh medium, in which they could lay eggs for 2–3 d before being frozen for subsequent molecular analysis. TE persistence after 5 and 10 generations was assessed by PCR.

Experimental Evolution. For the long-term dynamics involving *Mos1* only, migrants contained about five *Mos1* copies per haploid genome. For experimental assays involving w^{pch} populations, migrants carried approximately five *Mos1* and five *peach* copies per haploid genome (*SI Appendix, Supplementary Methods* and Table S1). Invasion dynamics were initialized with 1 migrant individual among 99 recipient flies, in 250-mL bottles raised at 25°C. Flies were allowed to lay eggs for 2 d and then frozen. New emergences were collected 10–12 d later, and 200 progeny flies were used to set up the next generations. Three to five replicates were initialized for all invasion dynamics, although some populations were subsequently lost. qPCR assays were run to quantify the number of *Mos1* and *peach* copies in every generation from G1 to G7 (in *Mos1* carriers and empty flies, separately) and in every five generations afterward.

1. Ågren JA, Wright SI (2011) Co-evolution between transposable elements and their hosts: A major factor in genome size evolution? *Chromosome Res* 19(6):777–786.
2. Hua-Van A, Le Rouzic A, Maisonhaute C, Capy P (2005) Abundance, distribution and dynamics of retrotransposable elements and transposons: Similarities and differences. *Cytogenet Genome Res* 110(1–4):426–440.
3. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302(5649):1401–1404.
4. Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284(5757):601–603.
5. Orgel LE, Crick FH (1980) Selfish DNA: The ultimate parasite. *Nature* 284(5757):604–607.
6. Hua-Van A, Le Rouzic A, Boutin TS, Filée J, Capy P (2011) The struggle for life of the genome's selfish architects. *Biol Direct* 6:19.
7. Hickey DA (1982) Selfish DNA: A sexually-transmitted nuclear parasite. *Genetics* 101(3–4):519–531.
8. Le Rouzic A, Capy P (2005) The first steps of transposable elements invasion: Parasitic strategy vs. genetic drift. *Genetics* 169(2):1033–1043.
9. Brennecke J, et al. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322(5906):1387–1392.
10. Le Rouzic A, Dupas S, Capy P (2007) Genome ecosystem and transposable elements species. *Gene* 390(1–2):214–220.
11. Wicker T, et al. (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8(12):973–982.
12. Holyoake AJ, Kidwell MG (2003) *Vege* and *Mar*: Two novel *hAT* MITE families from *Drosophila willistoni*. *Mol Biol Evol* 20(2):163–167.
13. Weiner AM (2002) SINEs and LINEs: The art of biting the hand that feeds you. *Curr Opin Cell Biol* 14(3):343–350.
14. Kaplan N, Darden T, Langley CH (1985) Evolution and extinction of transposable elements in Mendelian populations. *Genetics* 109(2):459–480.
15. Brookfield JFY (1996) Models of the spread of non-autonomous selfish transposable elements when transposition and fitness are coupled. *Genet Res* 67(3):199–209.
16. Quesneville H, Anxolabéhère D (2001) Genetic algorithm-based model of evolutionary dynamics of class II transposable elements. *J Theor Biol* 213(1):21–30.
17. Abrusan G, Krambeck HJ (2006) Competition may determine the diversity of transposable elements. *Theor Popul Biol* 70(3):364–375.
18. Le Rouzic A, Capy P (2006) Population genetics models of competition between transposable element subfamilies. *Genetics* 174(2):785–793.
19. Jacobson JW, Hartl DL (1985) Coupled instability of two X-linked genes in *Drosophila mauritiana*: Germinal and somatic mutability. *Genetics* 111(1):57–65.
20. Garza D, Medhora M, Koga A, Hartl DL (1991) Introduction of the transposable element *mariner* into the germline of *Drosophila melanogaster*. *Genetics* 128(2):303–310.
21. Charlesworth B, Charlesworth D (1983) The population dynamics of transposable elements. *Genet Res* 42(1):1–27.
22. Briscoe DA, et al. (1992) Rapid loss of genetic variation in large captive populations of *Drosophila* flies: Implications for the genetic management of captive populations. *Conservat Biol* 6(3):416–425.
23. Frankham R (1995) Effective population-size/adult population-size ratios in wildlife: A review. *Genet Res* 66(2):95–107.
24. Reed SC, Reed EW (1950) Natural selection in laboratory populations of *Drosophila*. II. Competition between a white-eye gene and its wild type allele. *Evolution* 4(1):34–42.
25. Carareto CM, et al. (1997) Testing transposable elements as genetic drive mechanisms using *Drosophila P* elements constructs as a model system. *Genetica* 101(1):13–33.
26. Pyatkov KI, et al. (2002) Penelope retroelements from *Drosophila virilis* are active after transformation of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 99(25):16150–16155.
27. Good AG, Meister GA, Brock HW, Grigliatti TA, Hickey DA (1989) Rapid spread of transposable elements in experimental populations of *Drosophila melanogaster*. *Genetics* 122(2):387–396.
28. Yang G, Nagel DH, Feschotte C, Hancock CN, Wessler SR (2009) Tuned for transposition: Molecular determinants underlying the hyperactivity of a stowaway MITE. *Science* 325(5946):1391–1394.
29. González J, Petrov D (2009) MITEs — The ultimate parasites. *Science* 325(5946):1352–1353.
30. Daniels SB, Clark SH, Kidwell MG, Chovnick A (1987) Genetic transformation of *Drosophila melanogaster* with an autonomous P element: Phenotypic and molecular analysis of long-established transformed lines. *Genetics* 115(4):711–723.
31. Ladeveze V, et al. (2001) Dynamics of the *hobo* transposable element in transgenic lines of *Drosophila melanogaster*. *Genet Res* 77(2):135–142.
32. Diaz-Gonzalez J, Vazquez JF, Albornoz J, Dominguez A (2011) Long-term evolution of the *roo* transposable element copy number in mutation accumulation lines of *Drosophila melanogaster*. *Genet Res (Camb)* 93(3):181–187.
33. Nuzhdin SV, Mackay TFC (1995) The genomic rate of transposable element movement in *Drosophila melanogaster*. *Mol Biol Evol* 12(1):180–181.
34. González J, Petrov D (2012) Evolution of genome content: Population dynamics of transposable elements in flies and humans. *Methods Mol Biol* 855:361–383.
35. 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(5):466–472.
36. Barrón MG, Fiston-Lavier AS, Petrov DA, González J (2014) Population genomics of transposable elements in *Drosophila*. *Annu Rev Genet* 48:561–581.
37. Anxolabéhère D, Kidwell MG, Periquet 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(3):252–269.
38. Bonnivard E, Bazin C, Denis B, Higuet D (2000) A scenario for the *hobo* transposable element invasion, deduced from the structure of natural populations of *Drosophila melanogaster* using tandem TPE repeats. *Genet Res* 75(1):13–23.
39. Kidwell MG (1983) Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 80(6):1655–1659.
40. Montchamp-Moreau C (1990) Dynamics of P-M hybrid dygenesis in P-transformed lines of *Drosophila simulans*. *Evolution* 44(1):194–203.
41. Kofler R, Hill T, Nolte V, Betancourt AJ, Schlötterer C (2015) The recent invasion of natural *Drosophila simulans* populations by the P-element. *Proc Natl Acad Sci USA* 112(21):6659–6663.
42. Medhora MM, MacPeek AH, Hartl DL (1988) Excision of the *Drosophila* transposable element mariner: Identification and characterization of the Mos factor. *EMBO J* 7(7):2185.
43. Lohe AR, Hartl DL (1996) Autoregulation of *mariner* transposase activity by overproduction and dominant negative complementation. *Mol Biol Evol* 13(4):549–555.
44. Tosi LR, Beverley SM (2000) *cis* and *trans* factors affecting *Mos1* mariner evolution and transposition in vitro, and its potential for functional genomics. *Nucleic Acids Res* 28(3):784–790.
45. Jaillet J, Genty M, Cambefort J, Rouault JD, Augé-Gouillou C (2012) Regulation of *mariner* transposition: The peculiar case of *Mos1*. *PLoS One* 7(8):e43365.
46. Bire S, et al. (2013) Transposase concentration controls transposition activity: Myth or reality? *Gene* 530(2):165–171.
47. Lohe AR, Lidholm DA, Hartl DL (1995) Genotypic effects, maternal effects and grand-maternal effects of immobilized derivatives of the transposable element mariner. *Genetics* 140(1):183–192.
48. Bregliano JC, Kidwell MG (1983) Mobile Genetic Elements, ed Shapiro JA (Academic, New York), p 363.
49. Leonardo TE, Nuzhdin SV (2002) Intracellular battlegrounds: Conflict and cooperation between transposable elements. *Genet Res* 80(3):155–161.
50. Venner S, Feschotte C, Biemont C (2009) Dynamics of transposable elements: Towards a community ecology of the genome. *Trends Genet* 25(7):317–323.
51. Linquist S, et al. (2015) Applying ecological models to communities of genetic elements: The case of neutral theory. *Mol Ecol* 24(13):3232–3242.
52. Capy P, Gasperi G, Biémont C, Bazin C (2000) Stress and transposable elements: Co-evolution or useful parasites? *Heredity (Edinb)* 85:101–106.
53. Johnson LJ (2008) Selfish genetic elements favor the evolution of a distinction between soma and germline. *Evolution* 62(8):2122–2124.
54. Wallau GL, Capy P, Loreto E, Le Rouzic A, Hua-Van A (2016) VHICA, a new method to discriminate between vertical and horizontal transposon transfer: Application to the mariner family within *Drosophila*. *Mol Biol Evol* 33(4):1094–1109.
55. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res* 29(9):e45–e45.

Robillard et al.

Supporting Information

from: Experimental evolution reveals hyperparasitic interactions
among transposable elements

by: É. Robillard, A. Le Rouzic, Z. Zhang, P. Capy, and A. Hua-Van

Contents

1 Supplementary Methods

1.1 Drosophila strains

Drosophila melanogaster does not naturally carry *Mos1* or *peach* elements, so all strains are transgenic. The w^{pch} strain is the original strain described in [1] and corresponds to a transformant of a $y\ w^{67c23}$ containing a w^{pch} allele carried by a P transgene on the X chromosome, with *peach* inserted on the 5' untranslated region of *white* gene. This strain has a stable $[w^{pch}]$ phenotype (peach-colored eyes).

The M19 strain is a w^{1118} transformant strain obtained in 1999 with a P-{$Mos1$, miniw$^+$} element inserted on chromosome II. After the transformation, *Mos1* transposed and amplified till about 20 copies per haploid genome. The M13 strain was obtained at the same time as M19, and was subsequently mixed with a $[w^+]$ strain, leading to the loss of the white mutation and of the P transgene. In the M13 strain, *Mos1* haploid copy number is about 19. The M22 strain is another transformant of w^{1118}, obtained at the same time, but that lost the *Mos1* element. All strains used in the experiments are compiled in Tab. S2.

1.2 Genetic crosses for obtaining migrants

In short-term experiments, founder flies containing a reduced copy number of *Mos1* elements were derived from the *Mos1*-carrying M19 or M13 strains (Tab. S2). All were obtained by crossing 5 females with M22 males (carrying no *Mos1*). Female progeny was backcrossed twice more with M22 males to obtain migrants with a theoretical copy number of $20 \times 3/2 = 2.5$ per haploid genome on average.

For the dynamics involving *Mos1* only, the crossing scheme was similar to the test of invasion frequencies except that (i) the starting strain was always M19, and (ii) only two backcross generations were performed (theoretical copy number of 5 per haploid genome), followed by a single generation of random mating for homogenization.

For experimental assays involving w^{pch} populations, "migrants" carried approximately 5 *Mos1* and 5 *peach* elements per haploid genome. They were obtained by crossing M19 females with a balancer strain (M5 ; CyO; Sb/Xasta). F_1 Xasta males were then crossed by w^{pch} females, and F_2 Xasta females crossed with $y\ w^{67c23}$ males. F_3 $[y\ w^{mos}]$ males resulting from X recombination in F_2 females were selected. In absence of transposition during these 3 generations, both *Mos1* and *peach* copies are expected to be located on the X chromosome. These males were mated first with the balancer strain, and then with the F_4 females for isogenization. After a few generations of brother-sister mating, qPCR assays determined that both TEs amplified up to approximately five *Mos1* and five *peach* copies per haploid genome. Flies were then used as "migrants" for introduction in w^{pch} populations .

1.3 Test for *Mos1* presence by PCR

Fly samples were disrupted, and their DNA extracted with Chelex 100 Molecular Biology Grade Resin (Bio-Rad). PCR was performed twice independently for each sample with the Phire Hot Start II DNA Polymerase (Thermo-Scientific), with a 372bp amplicon (mos3in: CCAATTGAGT-GTTTCCAACG, mos5in: AGGAAGTCGTTTTTGCATCG), and *Mos1* presence or absence was checked on agarose gel.

1.4 The discriminant qPCR assay

While the phenotypic *peach* excision assay is convenient to detect the presence of active *mariner* in populations, it presents some dramatic limitations in evolved *D. melanogaster* lab populations because (i) it is not suited for measuring high activity, as in the case when the very active *Mos1* copy is used, since all the progeny constantly exhibits eye mosaicism, (ii) germinal excision gives rise to dominant $[w^+]$ phenotypes that ultimately prevent estimation of the activity, (iii) it detects excision of the white locus, but not reinsertion in the genome, and (iv) it is not quantitative. For monitoring invasion experiments, we thus developed a differential qPCR-based assay that allowed

us to discriminate and quantify the number of each type of copy along generations, independently of the phenotypic assay. In our system, *Mos1* (autonomous) and *peach* (non-autonomous) are very similar in sequences, differing only by 11 positions out of 1286 bp, 4 of them triggering amino-acid changes in the transposase (Fig. S1A). We took advantage of these few differences to develop a discriminant qPCR approach for specific detection of both types of copy. The *peach* and *Mos1* PCRs were made with forward primers which differ by only 1 base at the very 3' end of the primer, corresponding to the SNP at position 132 in *Mos1* sequence (peachSNP116: CACCATAGTTTG-GCGCT, mosSNP116: CACCATAGTTTGGCGCG). The reverse primer was the same for both PCRs (mosDiv5: TTCACAGTTGGTACTTGTTCGC). The amplicons are 184 bp-long and PCR efficiencies are 1.97 and 1.94 for *Mos1* and *peach*, respectively. The normalization was made with the single copy gene RPII140 (CG3180) located on 3R arm using primers RpII140q1F: ATGGTG-GCTTGCGTTTCGGTG and RpII140q1R: ATTGTTGCGCAGATTGGCGATGG with a 157-bp amplicon and a 1.98 PCR efficiency. The cycling was executed on the Bio-Rad CFX96TM Real-Time System with the following program: 3 min 95°C initial denaturation, (30 sec 95°C denaturation, 15 sec 61°C annealing, 30 sec 72°C extension) 40 times, and ending with a dissociation stage 65 to 95°C. The data was collected with the Bio-Rad CFX Manager 2.0 and the copy number was calculated with the $\Delta\Delta Cq$ method. The specificity of each primer pair was verified on strains with known copy number of both types, that were then used for calibration in each experiment (Fig. S1B). Calibration was performed with w^{pch} females for *peach* copy number (1 homozygous *peach* copy on X chromosome) and hsp-mos w^{pch} strain for *Mos1* copy number (1 homozygous immobilized *Mos1* copy on chromosome II). DNA extraction and purification was performed on 40 individual batches at most, with MachereyNagel Nucleospin®Tissue kit, according to the manufacturer's instructions. Samples were first disrupted with 5 mm stainless steel beads in Qiagen tissue Lyzer. The DNA was quantified with the Thermo Scientific NanoDrop 2000c UV-Vis spectrophotometer. The PCRs were performed on 2ng DNA, on a $25\mu L$ reaction volume, with the Bio-Rad iQTM SYBR®Green Supermix. Each sample was tested three to six times, and we reported the mean (and standard errors) of these replicates. qPCRs were run independently on $[w^{mos}]+[w^+]$ vs. $[w^{pch}]$ individuals to get the average copy number within each group.

1.5 Phenotypic identification of *Mos1*-carriers

In the w^{pch} strain, *peach* is inserted in the promoter of the *white* gene, resulting in a stable peach-colored eye phenotype when *Mos1* is absent. In presence of *Mos1*, excision of *peach* occurs and restores the *white* gene activity. Two phenotypes can then be observed in the progeny of flies with both types of copies: mosaic-eyed $[w^{mos}]$ flies (red spots on a peach background), which reflects somatic excision of *peach* during the development, and $[w^+]$ revertant flies due to the germinal excision of *peach* [2, 3]. The *peach* copy can then be used as a reporter of *Mos1* activity or presence. In the dynamics described here, we considered $[w^{pch}]$ flies as having never contained *Mos1* copy, and $[w^{mos}]$ and $[w^+]$ as *Mos1*-carriers (Fig. S2). In theory, the $[w^+]$ phenotype could also be observed in absence of *Mos1* elements (MOS1 activity in an ancestor followed by the loss of the element), but we considered that this possibility was unlikely enough to be neglected in the first generations. We also used this test to confirm the presence/absence of *Mos1* in the ultimate generations (males from each experiment were crossed with w^{pch} females, and male progeny scored for mosaicism). The systematic absence of mosaicism indicated the loss of transposition activity in all ANA dynamics.

1.6 Copy number prediction in the absence of transposition, under random mating

In the first generations of ANA experiments, the autonomous element is present only in a subset of flies that are descendants of the initial migrant. Calculating copy number increase only in these flies better reflects the real transposition rate, since non-autonomous copies cannot multiply in absence of autonomous copies, and global copy number can also increase due to the invasion of the population by the descendant of the migrant flies. The two types of flies are easily distinguished by the eye phenotype, since *Mos1*-carriers are $[w^+]$ or $[w^{mos}]$, whereas *Mos1*-free flies are $[w^{pch}]$). The variation of the copy number observed in *Mos1*-carriers between two successive generations

was then compared to theoretical prediction assuming no transposition, no selection, and random mating between *Mos1*-carriers and *Mos1*-free individuals. In the very first generations, only few *Mos1*-carriers exist, so they have a better chance to mate with *Mos1*-free individuals. In absence of transposition, the amount of *Mos1* copies in mos-carriers (n^{mos}) is divided by two compared to the *Mos1*-carrying parent. On the other hand, when almost all the population contains *Mos1*-copies, crosses between two *Mos1*-carriers are more frequent, and in absence of transposition, the copy number in the progeny is assumed to be the same as in parents. Hence at each generation we computed the proportion of different crosses, assuming random mating, and considering the observed frequencies of *Mos1*-free flies $F_t = \mathrm{freq}([w^{\mathrm{pch}}])$. The proportion of $[w^+] \times [w^+]$ crosses (whose offspring contain n^{mos} copies on average) is $(1 - F_t)^2$, and the proportion of $[w^+] \times [w^{\mathrm{pch}}]$ crosses ($n^{\mathrm{mos}}/2$ copies) is $2F_t(1 - F_t)$. The frequency of *Mos1*-carriers at the next generation will be $F_{t+1} = (1 - F_t)^2 + 2F_t(1 - F_t)$. On average, the *Mos1*-carriers at the next generation will thus have $n_{t+1}^{\mathrm{mos}} = n_t^{\mathrm{mos}} \times (1 - F_t)^2/F_{t+1} + (n_t^{\mathrm{mos}}/2) \times 2F_t(1 - F_t)/F_{t+1} = n_t^{\mathrm{mos}}/(1 + F_t)$ copies. The theoretical dynamics of *peach* copies were calculated using the same method, except that a cross with a *Mos1*-free individual also brings 0.75 *peach* copies per gamete on average (1 per X chromosome): $n_{t+1}^{\mathrm{pch}} = (n_t^{\mathrm{pch}} + 3F_t/2)/(1 + F_t)$.

1.7 Transposition rates estimation in early generations of ANA experiments

Replicative transposition rates (u) were estimated as the rate of increase in copy number (n): $n_{t+1} = n_t(1 + u)$. Transposition rates during the first generations were computed considering the copy number estimates among *Mos1* carriers and the frequency of *Mos1* non-carriers in the population:

$$u_t^{\mathrm{mos}} = \frac{1 + F_t}{n_t^{\mathrm{mos}}} n_{t+1}^{\mathrm{mos}} - 1, \tag{1}$$

$$u_t^{\mathrm{pch}} = \frac{1 + F_t}{n_t^{\mathrm{pch}}} \left(n_{t+1}^{\mathrm{pch}} - \frac{3}{2} F_t \right) - 1. \tag{2}$$

In addition, we considered the possibility that the phenotypic marker could be associated with a selective cost s, in such a way that the relative fitness of $[w^{\mathrm{pch}}]$ flies is $1 - s$. The corresponding estimates of transposition rates become:

$$u_t^{\mathrm{mos}} = \frac{1}{n_t^{\mathrm{mos}}} \frac{1 + F_t - 2sF_t}{1 - sF_t} n_{t+1}^{\mathrm{mos}} - 1, \tag{3}$$

$$u_t^{\mathrm{pch}} = \frac{1}{n_t^{\mathrm{pch}}} \left(\frac{1 + F_t - 2sF_t}{1 - sF_t} n_{t+1}^{\mathrm{pch}} - \frac{3F_t(1 - s)}{2(1 - sF_t)} \right) - 1. \tag{4}$$

Note that these formulas coincide with the previous ones when $s = 0$. These transposition rates were computed for the 7 first generations and averaged out.

1.8 Transposition rates estimation in late generations

After 7 to 9 generations, the active *Mos1* element was either present in virtually all individuals, or definitely lost. Assuming that there are no more non-carriers ($F_t = 0$) and that the transposition rate is constant, the number of copies is expected to change between two consecutive generations as $n_{t+1} = n_t(1 + u)$. Negative u (decrease in copy number in the course of time) corresponds to situations in which the deletion rate is larger than the transposition rate. This change in copy number is cumulative, as $n_t = n_0(1 + u)^t$. A log transformation, $\log(n_t) = \log(n_0) + t\log(1+u)$, shows the expected linear relationship between time and the logarithm of copy number. Transposition rates were thus estimated as $u = \exp(b) - 1$, where b is the slope of the regression of $\log(n_t)$ over time. Confidence intervals of u were estimated from the confidence intervals of the slope (`confint.lm` function in `R` [4]), using the same formula.

Table S1: Summary of the invasion experiments.

Experiment	Migrant	Recipient strain
	Invasion frequencies	
M13	M2 (0.5 to 2.5 *Mos1* copies)[a]	9 M22 (no *mariner*)
M19	M2 (1 to 5 *Mos1* copies)	9 M22 (no *mariner*)
	Invasion dynamics	
A1	1 fertilized female M5 ($\sim$5 *Mos1*)	99 y^{w67c23}(no *mariner*)
A2	1 fertilized female M5 ($\sim$5 *Mos1*)	99 M22 (no *mariner*)
A3	100 M5 ($\sim$5 *Mos1* and $\sim$5 *peach*)	-
ANA1	1 male MP5 ($\sim$5 *Mos1* and $\sim$5 *peach*)	99 w^{pch} (1 *peach*)
ANA2	1 virgin female MP5 ($\sim$5 *Mos1* and $\sim$5 *peach*)	99 w^{pch} (1 *peach*)
ANA3	1 fertilized female MP5 ($\sim$5 *Mos1* and $\sim$5 *peach*)	99 w^{pch} (1 *peach*)

[a] All copy numbers are provided by haploid genome, meaning that the total copy number per diploid individual is actually twice more.

Table S2: Summary of the strains used in the experiments.

	Role	$Mos1$[1]	*peach*	Origin	Phenotype
Strain					
M13	*Mos1* donor	$\sim$19	0	P-transformed w^{1118} with a *Mos1* copy	$[y^+ \ w^+]$
M19	*Mos1* donor	$\sim$20	0	P-transformed w^{1118} with a *Mos1* copy	$[y^+ \ w^+]$
M22	Recipient empty strain	0	0	P-transformed w^{1118} with no *Mos1* copy	$[y^+ \ w^+]$
w^{pch}	*peach* donor and recipient strain	0	1	Transgenic	$[y \ w^{pch}]$
$y \ w^{67c23}$	Recipient empty strain	0	0		$[y \ w]$
Migrant					
M2	Migrant in invasion frequency	$\sim$2.5	0	M22 x M19—M13, back-crossed twice by M22	$[y^+ \ w^+]$
M5	Invasion dynamics migrant and control (A)	$\sim$5	0	M22 x M19, backcrossed once by M22	$[y^+ \ w^+]$
MP5	Invasion dynamics migrant in w^{pch} populations (ANA)	$\sim$5	$\sim$5	M19 and w^{pch}	$[y \ w^{mos}]$

1.9   Simulations

Theoretical invasion frequencies from multiple copies in absence of transposition were estimated with individual-based simulations. In simulations, females mate randomly with males, without re-mating. Females contribute equally to the next generation. Each parent gives a random number of TEs to the offspring, drawn in a Poisson distribution whose mean is half the number of parental TEs. Migrant females were considered as non-virgin, and thus mate with a virtual male having as many copies. Simulations were run 10,000 times with 1 to 5 copies in male and female migrants, and we computed the frequency of simulations in which at least one TE copy was present at generations 5 and 10. Two runs of simulations were carried out, with $N_e = 10$ and $N_e = 20$ from generation 2. Simulations scripts were written in R and are provided as Supplementary material.

For statistical testing, due to the small number of replicates for one given copy number, all experiments with 1 to 5 copies were pooled to compute the average invasion frequency (see the Supplementary Results section). This frequency was compared to the theoretical invasion frequency by drift only ($N_e = 20$), computed from the simulations, weighted by the initial copy number frequencies in each experiment (two-tailed binomial test; H_0: no difference between data and theory, H_1: invasion rate different from the simulations).

(A) primers location

(B) qPCR
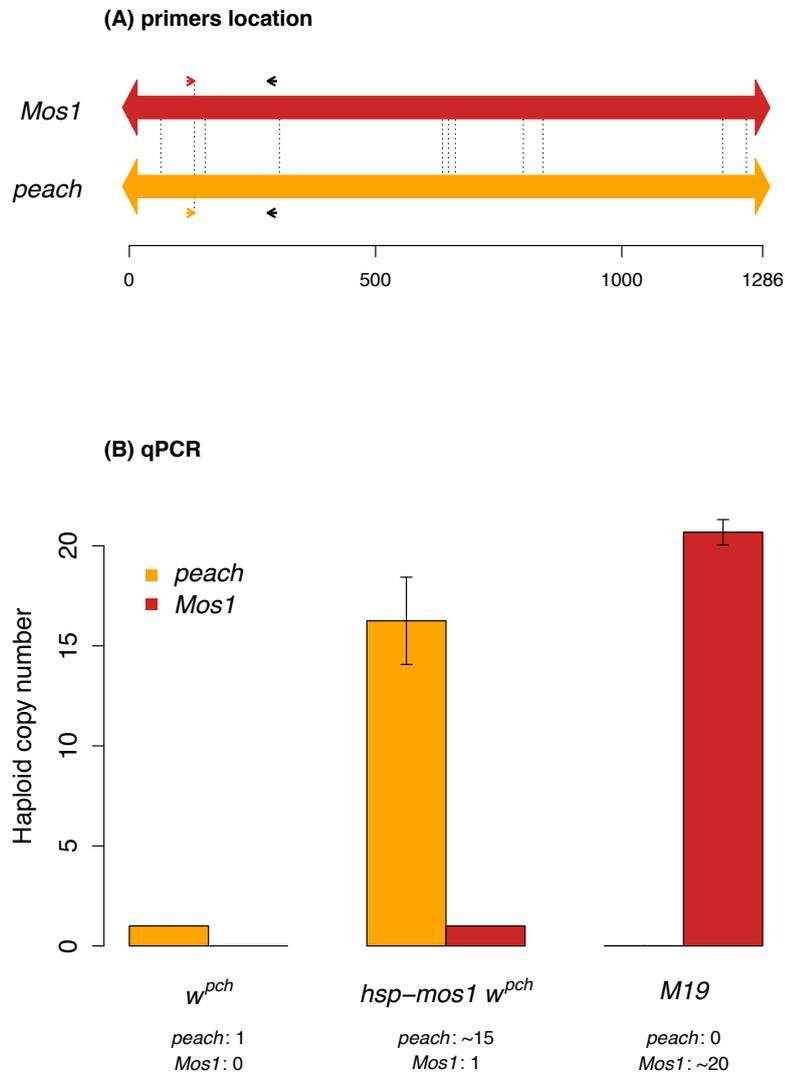
Figure S1: A. *Mos1* and *peach* copies differ by ten substitutions and one deletion (vertical dotted lines). One of this SNP (position 132) was used to anchor a specific primer (red and orange small arrows), whereas the reverse primer is common to both PCR. (B) qPCR results on strains with known copy number. The expected copy numbers (as determined by Southern Blot) are indicated below each strain. High specificity of the two primer pairs are visible from the absence of amplification with *Mos1* primers in w^{pch} strain, and from the absence of amplification with *peach* primers in the M19 strain. w^{pch} and hsp-mos w^{pch} contain 1 copy of *peach* and 1 copy of *Mos1*, respectively, and were constantly used for PCR calibration. Data shown here results from 26 independent experiments.
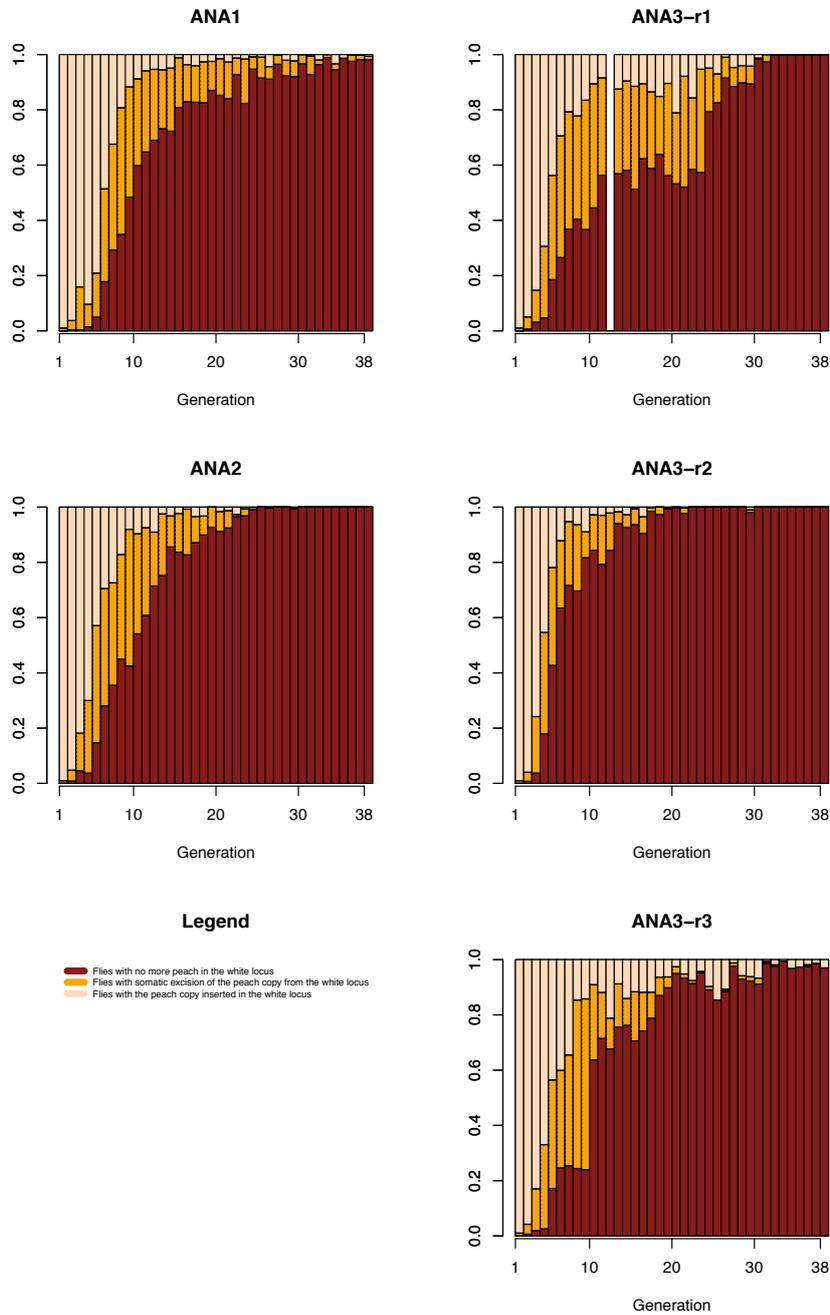
Figure S2: Invasion of the *Mos1* copy in populations as revealed by the phenotypically-detectable excision of *peach* from the *white* locus. Somatic and germinal excisions of the *peach* copy inserted in the *white* locus occur as soon as an active *Mos1* copy is present, and result in mosaic-eyed or red-eyed flies. On the other hand, peach-eyed flies do not contain any active *Mos1* copy. In all dynamics, the frequency of these naive flies rapidly decreases, and at generation 10, most flies contain *Mos1* and/or are descendant of a *Mos1*-containing ancestor.

7

2 Supplementary Results

2.1 Invasion frequencies

Raw data

Table S3

rep	sex	strain	expm[a]	withcopy[b]	$\leq 5^c$	N(G5)[d]	G5[e]	N(G10)	G10[e]
1	female	M13	34	26	7	7	7	7	6
2	female	M13	34	29	6	6	6	5	5
1	male	M13	32	27	17	14	9	14	8
2	male	M13	32	27	6	6	4	6	3
1	female	M19	29	25	14	13	13	14	12
2	female	M19	32	30	19	19	19	17	15
1	male	M19	26	21	16	16	11	16	11
2	male	M19	34	15	9	9	1	5	0

[a]: Two replicates (2×34) were carried on for each condition, but some strains have been lost or PCR/qPCR assays failed or lead to inconsistent results.
[b]: Experiments for which copies were detected by qPCR in the migrant.
[c]: Experiments in which the migrant carried between 1 and 5 copies.
[d]: Sample size with conclusive PCR results.
[e]: Presence of at least one copy at generation 5 and 10, respectively.

Initial copy number

Table S4

rep	sex	strain	mean[a]	var[b]	mean w/o 0[c]	var w/o 0
1	female	M13	5.59	19.70	7.31	12.94
2	female	M13	9.06	47.39	10.62	38.67
1	male	M13	4.50	14.19	5.33	12.31
2	male	M13	8.59	41.02	10.19	32.08
1	female	M19	5.00	15.50	5.80	13.25
2	female	M19	6.12	38.95	6.53	38.88
1	male	M19	3.58	8.09	4.43	6.16
2	male	M19	2.12	13.26	4.80	17.46

[a]: mean copy number in migrants determined by qPCR.
[b]: variance of copy number in migrants. The expected Poisson distribution is featured by mean = variance. [c]: mean (and variance) of copy numbers excluding 0s.

The distribution appears to be zero-inflated and overdispersed for both sexes in both populations. The red curve in the figure corresponds to the theoretical Poisson distribution.

Figure S3
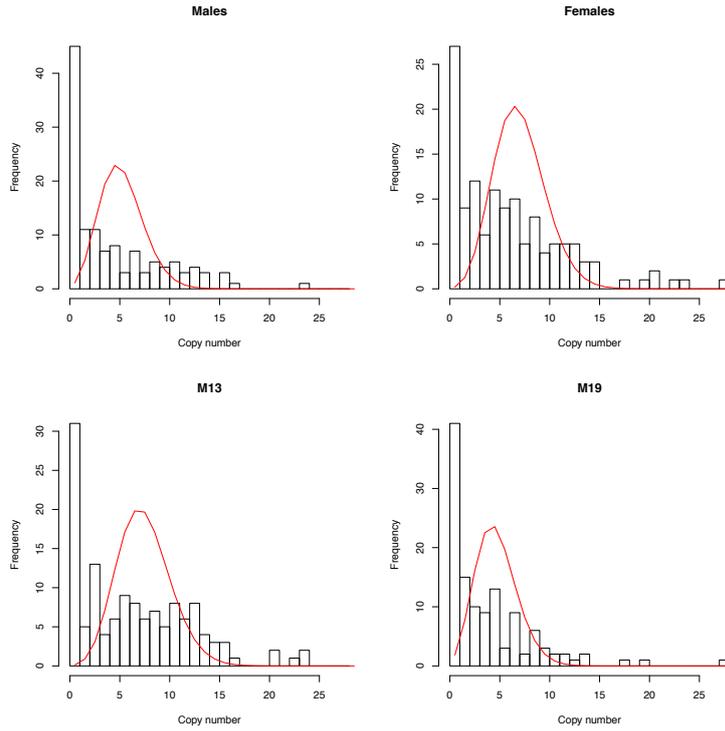
Goodness-of-fit tests clearly exclude Poisson distribution (function `goodfit` in package `vcl`, Maximum-likelihood ratio test) in all categories, both when including or excluding the zero counts.

Table S5

sex	strain	P(Pois)	P(Pois) w/o 0
female	M13	$1.32 \ 10^{-38}$	$9.09 \ 10^{-11}$
male	M13	$7.33 \ 10^{-31}$	$2.64 \ 10^{-14}$
female	M19	$1.30 \ 10^{-20}$	$5.04 \ 10^{-15}$
male	M19	$7.59 \ 10^{-20}$	$1.20 \ 10^{-03}$

A generalized linear model, featured by a quasi-Poisson family (log link function) in order to account for overdispersion, highlights a significant sex effect (around 30% more copies in females migrants than in males) and a significant strain effect (around 36% more copies in the M13 population than in M19). The overdispersion parameter (variance-to-mean ratio) was 4.81.

Table S6

	Estimate	Std. Error	t value	Pr($>$ |t|)	
Intercept: female M13	2.07971	0.08695			
male	-0.31740	0.11824	-2.684	0.00775	**
M19	-0.50410	0.12151	-4.149	$4.58 \ 10^{-5}$	***

Observed invasion frequencies   Distribution of the number of experiments (N), TE presence at G5 and G10 in both strains and both sexes (M and F), decomposed for each number of initial copies in the migrant (qPCR estimate rounded to the closest integer).

Table S7

	females								males							
	M13				M19				M13				M19			
Copy nb	N_5	G_5	N_{10}	G_{10}	N_5	G_5	N_{10}	G_{10}	N_5	G_5	N_{10}	G_{10}	N_5	G_5	N_{10}	G_{10}
0	13	0	13	0	6	0	6	0	10	0	10	0	24	0	24	0
1	3	3	3	3	5	5	5	5	5	4	4	2	6	1	4	1
2	1	1	1	1	7	7	8	7	4	3	4	2	7	3	6	2
3	5	5	4	3	7	7	6	5	5	2	7	3	3	1	3	1
4	1	1	1	1	5	5	5	4	3	2	2	1	4	4	4	4
5	3	3	3	3	8	8	7	6	3	2	3	3	5	3	4	3
>5	42	42	39	38	21	20	22	19	30	18	28	13	11	5	10	4

The relative influence of each factor was tested with a binomial GLM (logit link function) with three fixed factors: initial copy number (treated as a continuous factor, and shifted such as the intercept of the model is at 1 copy, since no invasion is expected when starting from 0), sex, and strain. Interestingly, the strain factor is never significant, which makes it reasonable to pool both strains to gain statistical power. There was no significant interaction between sex and the number of copies.

At generation 5, 100% of the female-migrant experiments still carried the TE, which lead to convergence issues with the GLM (the sex effect could not be estimated on the logit scale). This issue was addressed by turning artificially a missing data point into a failure event, which may slightly underestimate the difference between sexes.

The corresponding analyses of deviance confirm the absence of replicate effects. Note that the replicate effect is nested within strains (replicates 1 and 2 are independent between strains M13 and M19), but not within sexes (as males and females come from the same backcross).

GLM, Generation 5

	Estimate	Std. Error	z value	Pr(> |t|)	
Intercept: 1 copy, male, M13	0.2146	0.5881			
copy	0.2638	0.2207	1.196	0.2319	
female	3.2289	1.4737	2.191	0.0285	*
M19	-0.7984	0.6145	-1.299	0.1939	
copy × sex	0.3926	0.9351	0.420	0.6746	

Table S8

Analysis of deviance, Generation 5

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
sex	1	4.06	4.06	30.71	0.0000
strain	1	0.23	0.23	1.74	0.1902
replicate in strain	2	0.48	0.24	1.81	0.1705
Residuals	86	11.38	0.13		

Table S9

GLM, Generation 10

	Estimate	Std. Error	z value	Pr(> |t|)	
Intercept: 1 copy, male, M13	-0.8220	0.6375			
copy	0.6111	0.2699	2.264	0.02358	*
female	3.5092	1.1134	3.152	0.00162	**
M19	-0.2600	0.5794	-0.449	0.65362	
copy × sex	-0.8242	0.4354	-1.893	0.05834	

Table S10

Analysis of deviance, Generation 10

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
sex	1	2.53	2.53	14.19	0.0003
strain	1	0.02	0.02	0.13	0.7186
replicate in strain	2	0.51	0.25	1.42	0.2483
Residuals	79	14.08	0.18		

Table S11

Comparison with the simulations

Expected frequencies of TE persistence over 5 and 10 generations in absence of transposition, obtained by simulations with $Ne = 20$ (see the main text for details)

Copies in migrant	male-G5	male-G10	female-G5	female-G10
1	0.31	0.18	0.56	0.36
2	0.47	0.32	0.79	0.57
3	0.55	0.41	0.88	0.70
4	0.60	0.47	0.93	0.79
5	0.62	0.51	0.95	0.84

Table S12

Several strategies were used to test whether observed invasion frequencies deviate significantly from the expected persistence rates under genetic drift. In all cases, the null hypothesis is H_0: the TE persists as frequently as expected by drift, and H_1: the TE persists at a different rate than expected by drift. The fact that the real alternative hypothesis is that the TE should persist more frequently (direct test of the selfish DNA hypothesis) increases the robustness of the two-tailed test.

Goodness-of-fit A goodness-of-fit test based on the chi-square distance between theoretical simulations and observations was run in order to detect a potential discrepancy. In practice, each data point (specific sex with a specific migrant copy number) is featured by a success rate (presence), a failure rate (absence), to be contrasted with a theoretical invasion probability under the null hypothesis. The distance between theory and observation can be computed by a traditional chi-square, and chi-square measurements were summed over all conditions, leading to a global distance score. The empirical distribution of distance scores under H_0 was determined by 10,000 Monte-Carlo simulations, and the associated p-values were calculated as the frequency of simulated scores which distance was larger than the observations.

Table S13

p-values from Monte-Carlo resampling	G5	G10
female	0.06	0.00
male	0.51	0.11
both	0.12	0.00

General binomial test As an alternative, we computed the average theoretical invasion frequency (weighted by the occurrence of each migrant copy number) and ran a two-tailed and a one-tailed exact binomial test (function `binom.test` in R).

Table S14

|	success	tot	theor	1-tailed P	2-tailed P
male-G5	25	45	0.49	0.238	0.457
female-G5	45	45	0.83	0.000	0.000
male-G10	22	41	0.37	0.022	0.034
female-G10	38	43	0.65	0.001	0.001

Exact distribution The previous binomial test lacks power because the real distribution is a mixture distribution involving several binomials of different probabilities. Although there exist complex approximations of distributions resulting from summing binomial variables, it appeared simpler to reconstitute the resulting distribution by a Monte Carlo approach.

Exact distributions of the number of expected invasions under the null hypothesis computed by Monte-Carlo sampling (100,000 replicates) in mixtures of five binomial distributions. The vertical red lines indicate the observed invasion counts.

Figure S4
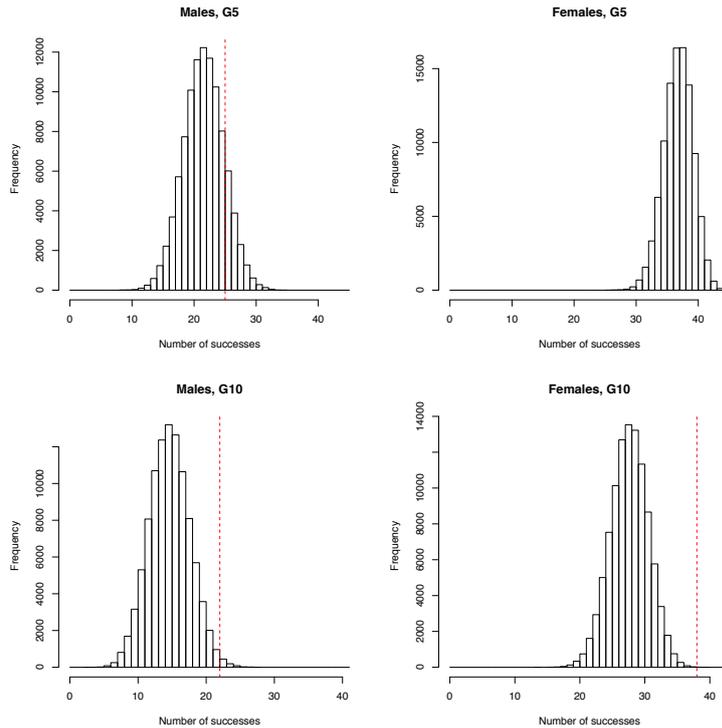
One- and two-tailed p-values estimated from the mixture distributions. The one-tailed p-values were obtained by summing up the probabilities of counts larger or equal than the observations. The two-tailed p-values were obtained by summing up the probabilities of all counts which are as likely or less likely than the observation.

Table S15

	1-tailed P	2-tailed P
male-G5	0.231	0.440
female-G5	0.000	0.000
male-G10	0.019	0.030
female-G10	0.000	0.000

2.2   Experimental evolution

2.2.1   Autonomous elements

In the first experimental evolution setting, the number of autonomous copies can be influenced by three factors: (i) the generation (the number of copies is expected to change with time), the experimental condition (A1, A2, A3), and the replicate nested into each experimental condition. An analysis of variance (more exactly, an analysis of covariance, as the generation is treated as a numerical covariable) leads to:

Table S16

log(*Mos1* copy number)	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
generation	1	57.21	57.21	61.92	$9.12\ 10^{-12}$	***
experiment	2	17.72	8.86	9.59	$1.73\ 10^{-4}$	**
replicate:experiment	5	1.72	0.34	0.37	0.87	
residuals	87	80.38	0.92			

The linear model accounting for generation and experiment effect (as well as their interactions) leads to:

Table S17

log(*Mos1* copy number)	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept: G_{10} A1	1.843	0.075			
generation A1	0.013	0.003	4.20	0.000	***
A2	0.305	0.118	2.58	0.012	**
A3	0.335	0.087	3.86	0.000	***
G:A2	0.010	0.005	1.93	0.057	
G:A3	0.011	0.004	2.95	0.004	**

The corresponding linear regression can be represented as follows, thick black lines standing for the linear model pedictions:

Figure S5
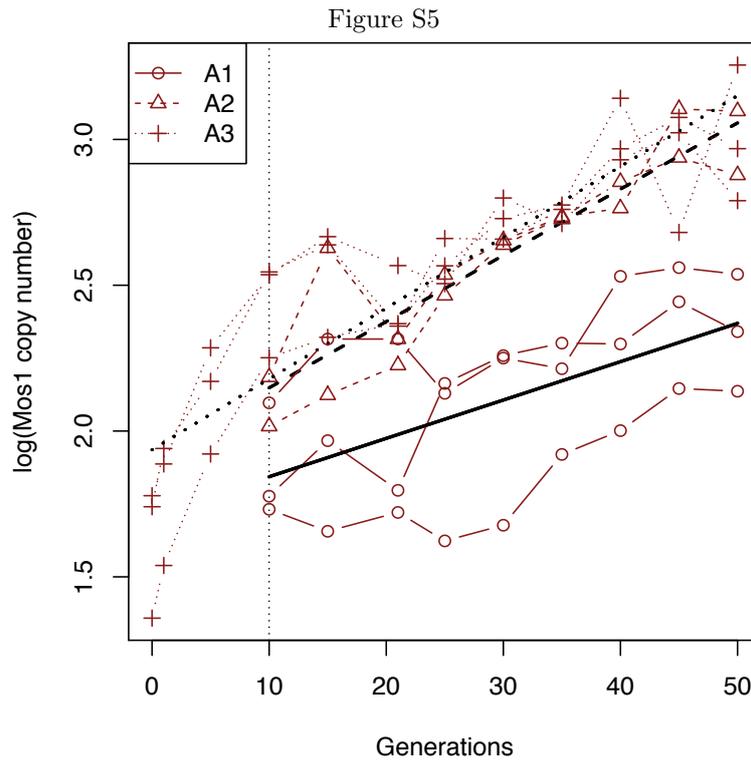
This analysis confirms:

- A significant (and positive) generation effect (the copy number increases along with time);

- A significantly lower rate of increase (transposition) in experiment A1 than in A3, A2 being intermediate.

- At generation 10, A1 has significantly less copies than A2 and A3.

13

Transposition rates were estimated as described in the supplementary material ($u = e^b - 1$, where b is the slope of the linear regression $\log(n) = bt + \varepsilon$, where t stands for the generation. As detailed in the manuscript, generations in which TE copies were present in less than about 80% of the individuals were discarded. 95% confidence intervals were determined with the `confint.lm` procedure in `R`. Estimated transposition rates (from generation 10 to 50 for experiments A1 and A2, and from generations 1 to 50 for experiment A3) are:

Table S18

Mos1 transposition rates	2.5%	Estim.	97.5%
A1.1	0.006	0.013	0.020
A1.2	0.003	0.010	0.018
A1.3	0.009	0.016	0.023
A1 pooled	0.006	0.013	0.021
A2.1	0.019	0.025	0.031
A2.2	0.012	0.021	0.030
A2 pooled	0.018	0.023	0.028
A3.1	0.011	0.020	0.028
A3.2	0.015	0.022	0.029
A3.3	0.023	0.032	0.041
A3 pooled	0.020	0.025	0.029

2.2.2 Non-autonomous along with autonomous elements

A similar analysis on the non-autonomous set of experiments (experiments ANA1, ANA2, and ANA3), forcing the model intercept at generation 7 (which corresponds approximately to the point where the *Mos1* copies have invaded all populations) lead to the following results:

- The replicate effect within experiments is non-significant for both *Mos1* and *peach* copy numbers, and explains only a tiny part of the total variance.

- There is a modest generation effect for *Mos1* (slighly negative and not significant in the reference ANA1 experiment, and significant only in its interaction with the ANA3 experiment (in which copies are lost at a higher rate).

- This generation effect is larger, positive, and more significant for the *peach* copies. The negative interaction is also negative for the ANA3 experiment (less *peach* copies in ANA3).

Table S19

log(*Mos1* copy number)	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
generation	1	15.03	15.033	12.504	0.000572	***
experiment	2	28.77	14.386	11.966	$1.77 \ 10^{-5}$	***
replicate:experiment	2	1.26	0.631	0.525	0.592784	
residuals	124	149.07	1.202			

Table S20

log(*Mos1* copy number)	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept: ANA1, G_7	0.728	0.314			
generation (ANA1)	-0.010	0.006	-1.70	0.092	
ANA2	-0.756	0.443	-1.70	0.092	
ANA3	-0.410	0.362	-1.13	0.260	
G:ANA2	0.008	0.008	1.02	0.311	
G:ANA3	-0.025	0.007	-3.77	0.000	***

Figure S6
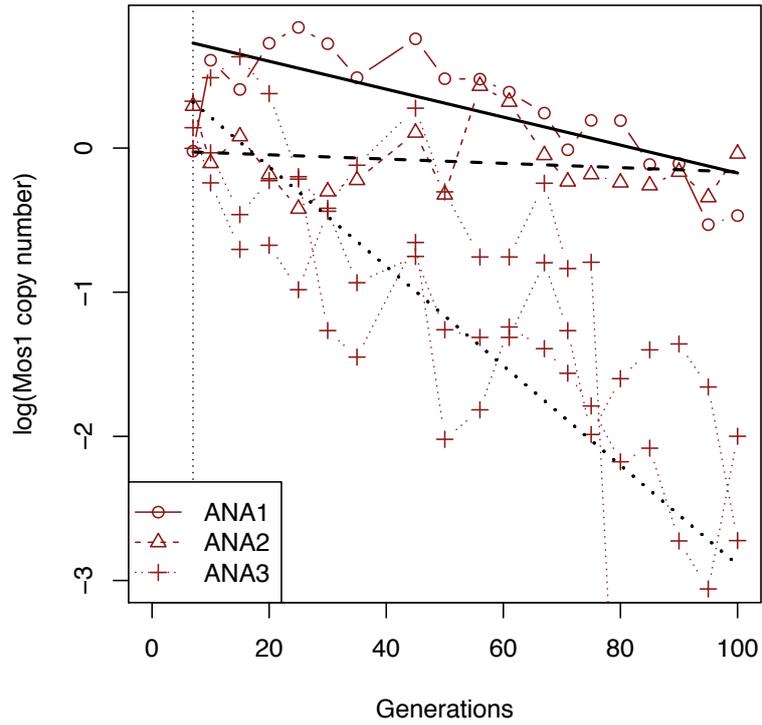
Table S21

log(*peach* copy number)	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
generation	1	107.36	107.36	252.175	$< 2 \; 10^{-16}$	***
experiment	2	0.71	0.36	0.836	0.436	
replicate:experiment	2	0.50	0.25	0.583	0.560	
residuals	124	52.79	0.43			

Table S22

log(*peach* copy number)	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept: ANA1, G_7	1.371	0.154			
generation (ANA1)	0.039	0.006	7.05	0.000	***
ANA2	-0.164	0.217	-0.76	0.454	
ANA3	0.468	0.178	2.64	0.012	
G:ANA2	-0.001	0.008	-0.17	0.867	
G:ANA3	-0.025	0.006	-3.94	0.000	***

Figure S7

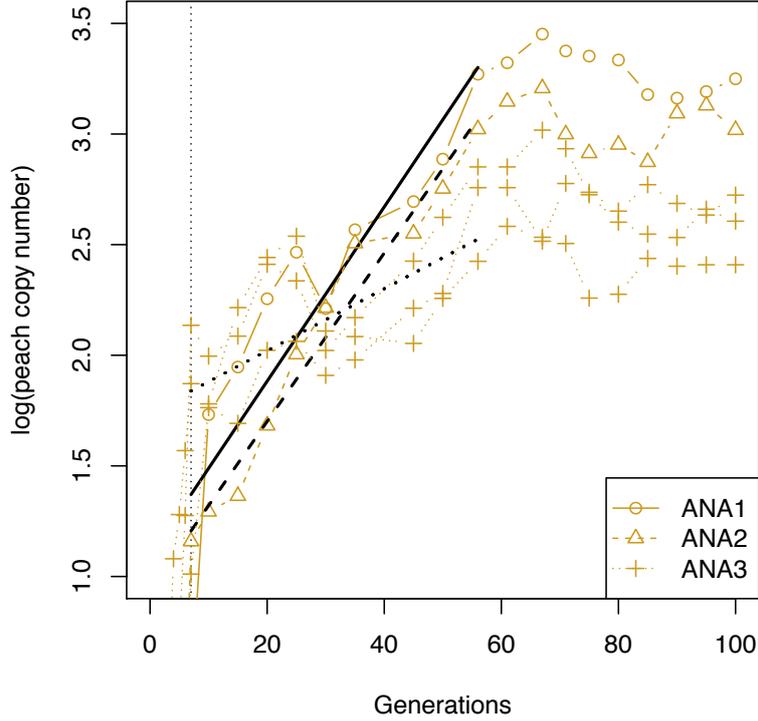The fact that the *peach* copy numbers are rather homogeneous throughout experiments (no experiment effect in the ANCOVA) makes it possible to detect non-linearity with more power than in individual experiments by pooling the whole dataset. We seeked breakpoints in the dynamics of *peach* copy number through time using the `segmented` library in R (model: `log(peach)` $\sim$ `G`). Estimated breakpoints were at $G_1 = 9.5\pm$ s.e. 0.7 and $G_2 = 62.4 \pm 5.1$ generations. This is consistent with our strategy to limit further analysis from generation 7 (end of the *Mos1* invasion) to generation 60 (loss of the *Mos1* copy).

Table S23

Mos1 transposition rates	2.5%	Estim.	97.5%
ANA1	-0.014	-0.010	-0.005
ANA2	-0.005	-0.001	0.003
ANA3.1	-0.071	-0.051	-0.031
ANA3.2	-0.040	-0.032	-0.024
ANA3.3	-0.024	-0.018	-0.012
ANA3 pooled	-0.042	-0.034	-0.026

Table S24

peach transposition rates	2.5%	Estim.	97.5%
ANA1	0.021	0.040	0.059
ANA2	0.033	0.039	0.045
ANA3.1	0.012	0.023	0.034
ANA3.2	-0.005	0.006	0.017
ANA3.3	0.005	0.013	0.021
ANA3 pooled	0.008	0.014	0.021

The full dataset for transposition rates is:

16

Table S25

copy	compet	exp	rate
mos	A	A1	0.013
mos	A	A1	0.010
mos	A	A1	0.016
mos	A	A2	0.025
mos	A	A2	0.021
mos	A	A3	0.020
mos	A	A3	0.022
mos	A	A3	0.032
mos	ANA	ANA1	-0.010
mos	ANA	ANA2	-0.001
mos	ANA	ANA3	-0.051
mos	ANA	ANA3	-0.032
mos	ANA	ANA3	-0.018
pch	ANA	ANA1	0.040
pch	ANA	ANA2	0.039
pch	ANA	ANA3	0.023
pch	ANA	ANA3	0.006
pch	ANA	ANA3	0.013

The three-way ANOVA (`rate ~ copy + compet + exp`) shows that *Mos1*t (72%) of the variance in the transposition rate can be attributed to the "competition" factor:

Table S26

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
copy	1	0.00154	0.00154	20.38	0.0009	***
compet	1	0.00555	0.00555	73.30	0.0000	***
exp	4	0.00195	0.00049	6.46	0.0063	**
Residuals	11	0.00083	0.00008			

In order to distinguish between the effect of the nature of the copy (*Mos1* or *peach*) and the competition effect, the data set was split into three categories (mos:A, mos:ANA, and pch:ANA), and a one-way analysis of variance was run on this factor:

Table S27

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
category	2	0.0071	0.0035	19.0745	0.0001	***
Residuals	15	0.0028	0.0002			

A Tukey post-hoc analysis (thus including corrections for multiple testing) leads to:

Table S28

	diff	lwr	upr	p adj
mos:ANA-mos:A	-0.042	-0.063	-0.022	0.000
pch:ANA-mos:A	0.004	-0.016	0.025	0.844
pch:ANA-mos:ANA	0.047	0.024	0.069	0.000

This confirms that there is no significant difference between *peach* transposition rate in the ANA experiments and *Mos1* transposition rate in the A experiment. In contrast, *Mos1* transposition rate in the ANA experiment is significantly lower than the two others.

3 Simulation script

```
##############################################################################
# Simulations of invasion frequencies without transposition
#
# author: Arnaud Le Rouzic, 2015 <lerouzic@egce.cnrs-gif.fr>
#
# This work is free. It comes without any warranty, to
# the extent permitted by applicable law.
# You can redistribute it and/or modify it under the
# terms of the WTFPL, Version 2.
# See http://www.wtfpl.net/ for more details.
#
##############################################################################

# The main function is simul.full. It can be used in the following way:
#
# simul.full(cpy=2, male=TRUE, N=10, rep=100)
#
# The main parameters are:
# * cpy: the number of copies in the migrant
# * male: if FALSE, the migrant is a female
# * N: effective population size from generation 2
# * rep: number of replicates
# In addition, it is possible to set N0 (the population size at
# generation 1, including the migrant) and Gmax, the max number of
# generations

simul.full <- function(cpy=1, male=TRUE, N0=10, N=N0, Gmax=10, rep=1000) {
  # Full simulation with rep replicates.
  fdt <- replicate(rep, simul.mig(cpy=cpy, male=male, N0=N0, N=N, Gmax=Gmax))
  apply(fdt>0, 1, mean)
}

reproduct.ind <- function(TEmale, TEfemale) {
  # gives the offspring from a male carrying TEmale copies
  # and a female carrying TEfemale copies.
  sum(rpois(2, lambda=c(TEmale/2, TEfemale/2)))
}

reproduct.pop <- function(listpop, N=length(males)+length(females), fmig=FALSE) {
  # Reproduces the population
  males <- listpop$males
  females <- listpop$females
  reprod.males <- sample(males, length(females), replace=TRUE)
  ans <- sapply(rep(seq_along(females), each=N/length(females)),
    function(ffi) {
      if (fmig && ffi==1) reproduct.ind(females[1], females[1])
      else reproduct.ind(reprod.males[ffi], females[ffi]) })
  list(males=ans[1:(length(ans)/2)], females=ans[(length(ans)/2+1):length(ans)])
}

simul.mig <- function(cpy=1, male=TRUE, N0=10, N=N0, Gmax=10) {
  # Runs a single simulation
  popin <- if(male) list(males=c(cpy, rep(0, N0/2-1)), females=rep(0,N0-N0/2))
      else list(males=rep(0, N0/2), females=c(cpy, rep(0, N0-N0/2-1)))
  ans <- c(sum(unlist(popin)), rep(NA, Gmax))
  for (gg in 1:Gmax) {
    popin <- reproduct.pop(popin, N=N, gg==1 && !male)
    ans[gg+1] <- sum(unlist(popin))
  }
  ans
}

makeTransparent<-function(someColor, alpha=100)
{
  newColor<-col2rgb(someColor)
  apply(newColor, 2, function(curcoldata){rgb(red=curcoldata[1], green=curcoldata[2],
    blue=curcoldata[3],alpha=alpha, maxColorValue=255)})
}
```

```r
mcsapply <- function (X, FUN, ..., simplify = TRUE, USE.NAMES = TRUE, mc.cores=NA)
{
  if (!require(parallel)) sapply(X, FUN, ..., simplify=simplify, USE.NAMES=USE.NAMES)
  if(is.na(mc.cores)) mc.cores <- detectCores()
    FUN <- match.fun(FUN)
    answer <- mclapply(X = X, FUN = FUN, ..., mc.cores=mc.cores)
    if (USE.NAMES && is.character(X) && is.null(names(answer)))
        names(answer) <- X
    if (!identical(simplify, FALSE) && length(answer))
        simplify2array(answer, higher = (simplify == "array"))
    else answer
}
```

References

[1] Garza D, Medhora M, Koga A, Hartl D (1991) Introduction of the transposable element *mariner* into the germline of *Drosophila melanogaster*. *Genetics* 128(2):303–310.

[2] Bryan GJ, Jacobson JW, Hartl DL (1987) Heritable somatic excision of a drosophila transposon. *Science* 235(4796):1636–1638.

[3] Picot S et al. (2008) The mariner transposable element in natural populations of *Drosophila simulans*. *Heredity* 101(1):53–59.

[4] R Core Team (2014) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).