File S1

Supporting Text: Evaluation of the *Poisson* assumption of the analytical model

We modeled the distribution of TE copy number as *Poisson,* which has been well established in the literature. However, this approximation holds true mainly when the TE population is at near equilibrium and the TE copy number is large. The key part of our model is the spread of a newly invaded TE family, during which the TE population is not at equilibrium and the copy number may be low. To investigate how the deviation from *Poisson* approximation may influence the predictions of our analytical models, we performed full Monte Carlo simulations to evaluate the potential impacts of this assumption.

Monte Carlo Simulations

We used the following Monte Carlo simulation to address this issue. The host population size is 100,000. Each host individual genome is comprised of two parental complements of three chromosomes, each of which has 1,000 potential TE insertion sites and a host locus. Crossover is modeled as *Poisson* process and the crossover rate is set as 0.001 between two potential TE insertion sites, making it averagely one crossover per chromosome per generation. At generation zero, the 0.1% of the population contain on average μ_0 copies of the TE (distribution is *Poisson*). Independently chosen 0.1% of the population have the beneficial allele at the host locus.

A new member of the next generation is simulated through the following steps. Two parents are first chosen and each parent contributes a haploid genome to the offspring (assuming independent assortment and crossing over as described above). Neither TE insertions nor the host locus influence the transmission. Each TE insertion of the offspring then independently undergoes a single replicative transposition with probability equal to the transposition rate *u*, which changes according to the cytotype of the parent (u_0 in hybrid dysgenic cross and u_1 in the other crosses) and the host locus genotype of the offspring (*u(1-d)* in homozygotes of beneficial allele, *u(1-hd)* in heterozygotes of beneficial allele and *u* for the other genotype). If the total number of transposition events in an offspring is above the hybrid dysgenic threshold (*HD*), the offspring's fitness is set to zero and the offspring is not passed to the next generation. If the total number of transposition events in an offspring is below the threshold, its fitness is calculated according to the following equation ($w(n) = e^{-an-bn^2/2}$, where *n* is the total TE copy number and *a* and *b* are 10^{-5} and 10^{-6} respectively). The offspring is transmitted to the next generation with probability equal to it fitness. This process is repeated until 100,000 offspring are generated.

According to the analyses of the analytical modeling, following parameters did not have significant impacts on the dynamics of *I* and were chosen as follows for the simulation: $d = 0.5$, $h = 0.5$, $u_1 = 10^{-4}$ and $\mu_0 = 10$. As discussed in the main text, the spread of newly invaded TE family has almost no impacts on the host gene for cases where u_0 equals 0.1, which is of course less interesting case for our analysis. We thus chose $u_0 = 1$ for our simulation. We did pilot simulations with n_{HD} equals 3, 5, 7, and 10 and found no apparent differences (data not shown) and thus only the case with the greater numbers of simulations, $n_{HD} = 5$, are presented below.

Results of Simulations

Following figures showed the averaged result of 1,000 Monte Carlo simulations for proportion of *P cytotype* (Figure S1), TE copy number (Figure S2) and the frequency of host beneficial allele (Figure S3 and Figure S4), comparing with the prediction of analytical model. The most critical part of our analytical model is from the invasion of the newly invaded TE family to its reaching equilibrium in the population, which takes approximately 100 generations after its first invasion.

Simulations showed the spread and the increase in copy number of the newly invaded TE family is slower than the analytical prediction (Figure S1 and S2). The allele frequency predicted by the analytical model based on the assumed *Poisson* distribution of copy number tends to initially exceed then fall below the simulated host allele frequency (Figure S3). However, the error between analytical approximation and the simulation is always within 2% (Figure S4). Thus, our overall conclusion that the spread of a newly invaded TE family is unlikely to drive the fast evolution of interacting host genes is not sensitive to the naïve assumptions of the analytic model.

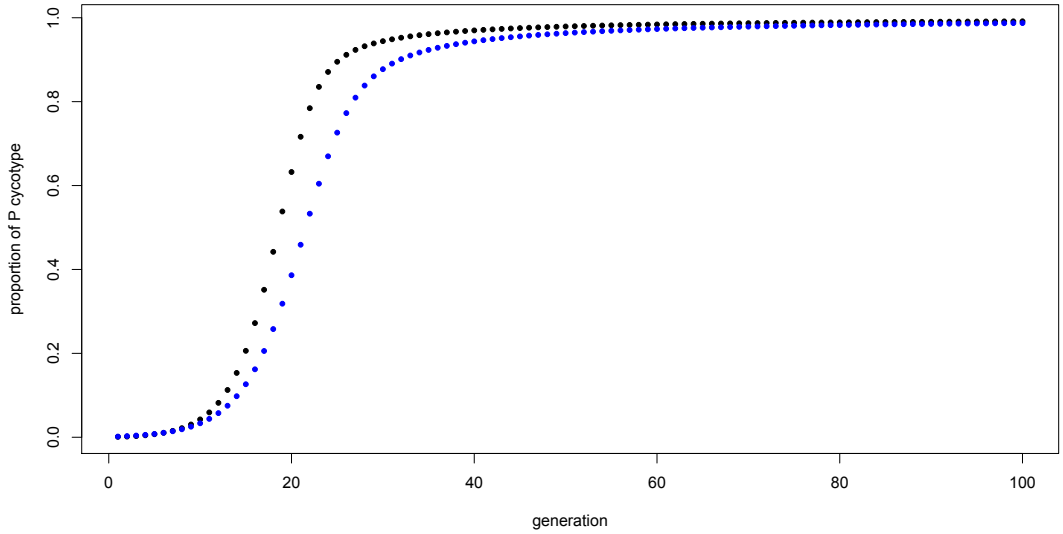

Figure S1 The proportion of *P cytotype* individuals over time. Black and blue dots are the analytical prediction and simulation results respectively.

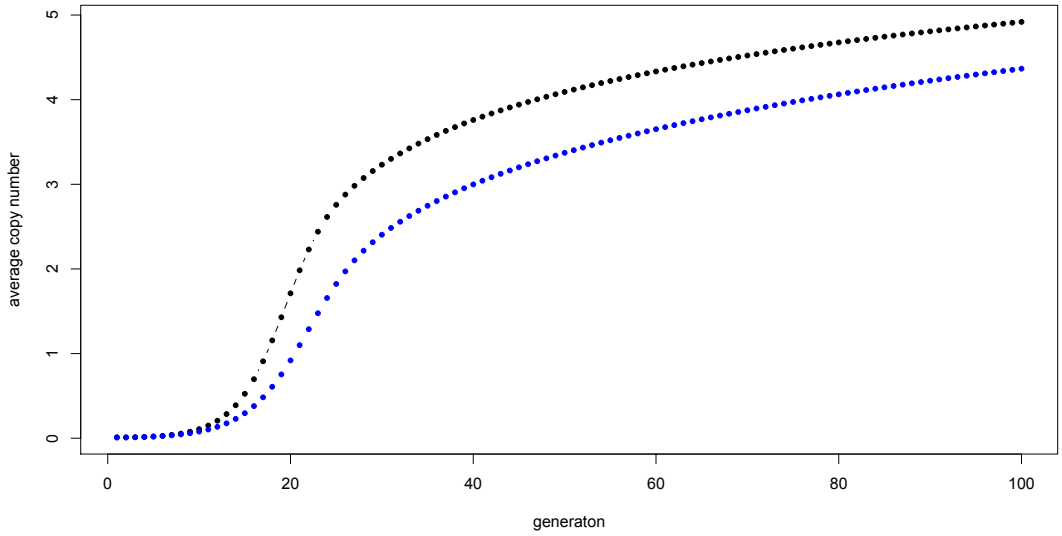

Figure S2 The averaged TE copy number over time. Black and blue dots are the analytical prediction and simulation results respectively.

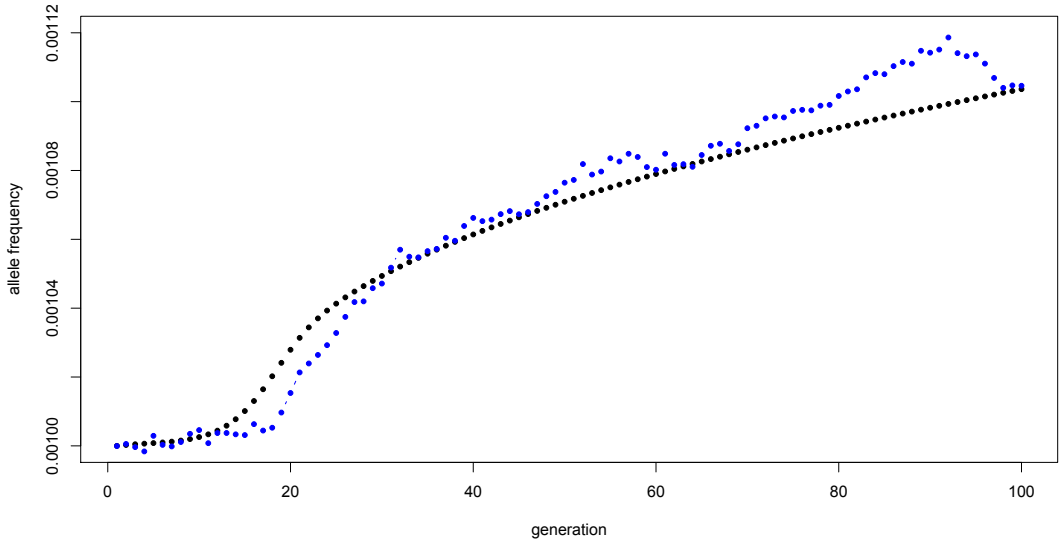

Figure S3 The frequency of host beneficial allele over time. Black and blue dots are the analytical prediction and simulation results respectively.

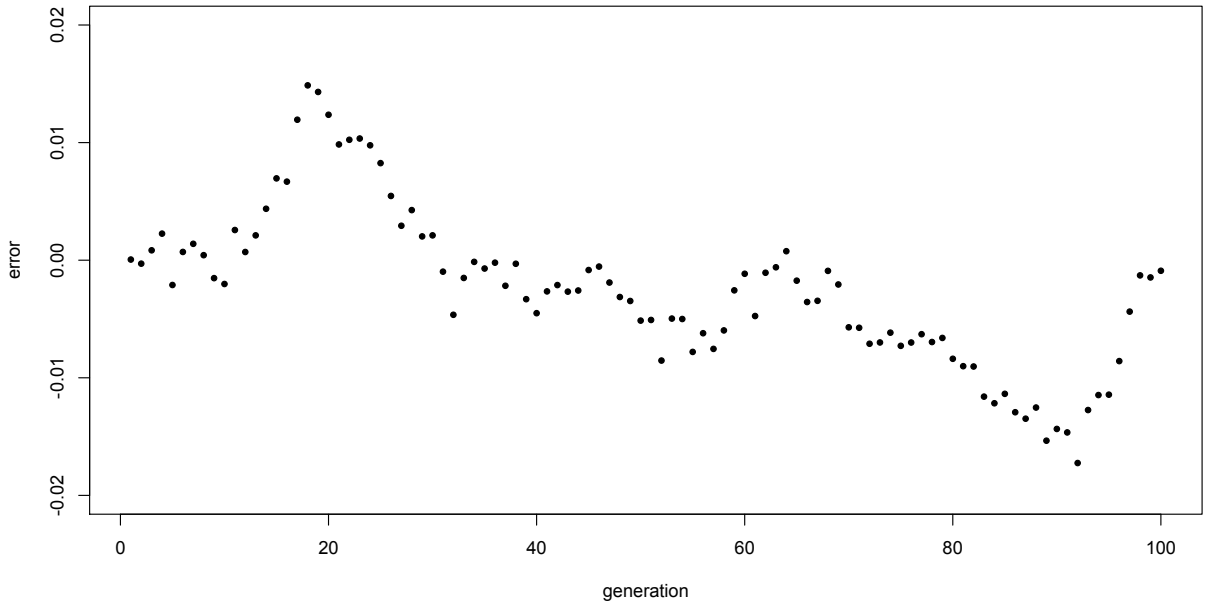

Figure S4 The errors of analytical prediction of the host beneficial allele with respect to simulations.

During the initial invasion phase, the *Poisson* distribution predicts a distribution that has larger mode than the actual simulation (Figure S5A, B). Soon after *the P cytotype* individuals in the population become common (≈ generation 35), the TE copy number distribution is nearly *Poisson* (Figure S5C, D). In addition to the fact that the *Poisson* distribution is a good approximation to the Binomial sampling when the TE copy number is large, the linkage among TE insertions also contributes to the differences between the predictions of analytical model and simulations. In simulations where there is free recombination among TE insertions, the distribution of TE copy number quickly reaches *Poisson* within 15 generations, when the *P cytotype* in the population is still rare (results not shown).

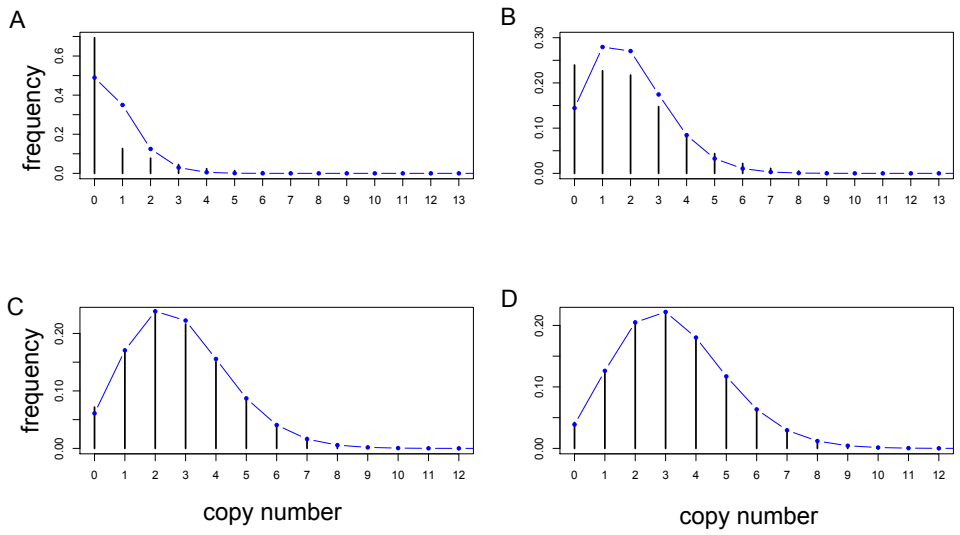

Figure S5 Distribution of TE copy number among host individuals at generation 15 (A), 25 (B), 35 (C) and 45 (D). Black bars are the simulated values while the blue dots are the *Poisson* expectation. The distribution of TE copy number reaches nearly *Poisson* around generation 35, when the *P cycotype* start being common in the population.

Our analytical model initially overestimates the host allele frequency. This is potentially caused by that fact that the *Poisson* approximation has a larger mode than the real copy number distribution. Because the copy number of an individual is generally small and the probability of hybrid dysgenic crosses happening is low, this did not lead to sever deviation between analytical predictions and the actual simulations.

After the *P element* in the population becomes common (around generation 35), our analytical model starts to predict lower host allele frequencies. This could be attributable to the transient linkage disequilibrium among TE insertions in the simulations. In this case, the simulation has a heavier right tail than the expectation from the *Poisson* approximation. The following tables (Table S5, 6, 7) show the proportion of simulations that have more individuals with a particular copy number than the predictions of the *Poisson* approximation. This proportion is universally greater than 50% for individuals with larger copy number, whose offspring are likely to have too many TE transpositions in a single generation. This can lead to stronger selection and a slightly higher host allele frequency change than the analytical model.

Table S5 The proportion of simulations that have more individuals with a particular copy number than the *Poisson* predictions at generation 40. This is the generation when the analytical model starts to underpredict the host allele frequency. The proportion of simulations that have more individuals with a particular copy number than the *Poisson* predictions are shown in the "Proportion" row, with proportion greater than 50% highlighted in blue. Individuals with large *P element* copy number may not present in all simulations and thus the "No Simulations" may not always be 1,000.

Copy Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
No Simulations	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	998	825	361	89	25	3	1
Proportion	1	0.017	0.06	0	0	0.034	0.968	1	1	1	1	0.992	0.92	0.976	1	1	1	1	1

Table S6 The proportion of simulations that have more individuals with a particular copy number than the *Poisson* predictions at generation 50.

Copy Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
No Simulations	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	979	600	215	40	7	3	2
Proportion	1	0.011	0.701	0.493	0.134	0.072	0.271	0.833	0.985	0.979	0.906	0.811	0.692	0.618	1	1	1	1	1	1

Table S7 The proportion of simulations that have more individuals with a particular copy number than the *Poisson* predictions at generation 60.

Copy Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	22
No Simulations	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	999	839	354	100	32	6	1
Proportion	0.969	0.001	0.661	0.854	0.646	0.3	0.233	0.362	0.605	0.652	0.618	0.537	0.517	0.48	0.613	1	1	1	1	1