The Ecology of the Genome and the Dynamics of the Biological Dark Matter

Alheli Flores-Ferrer, Anne Nguyen, Sylvain Glémin, Jean-Marc Deragon, Olivier Panaud, Sébastien Gourbière

The Ecology of the Genome and the Dynamics of the Biological Dark Matter

Alheli Flores-Ferrer[1], Anne Nguyen[1], Sylvain Glémin[2], Jean-Marc Deragon[1], Olivier Panaud[1], and Sébastien Gourbière[1,3*]

[1]UMR5096 'Laboratoire Génome et Développement des Plantes', Université de Perpignan Via Domitia, Perpignan, France

[2]UMR 6553 'Ecosystèmes, Biodiversité, Evolution', Université de Rennes 1, Rennes, France

[3] Centre for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton BN1 9QG, United Kingdom

alheli.flores-ferrer@univ-perp.fr

annenguyen1984@gmail.com

sylvain.glemin@univ-rennes1.fr

jean-marc.deragon@univ-perp.fr

panaud@univ-perp.fr

gourbiere@univ-perp.fr

Corresponding author: Gourbiere Sébastien, PhD

UMR5096 'Laboratoire Génome et Développement des Plantes', Université de Perpignan Via Domitia, 58 Av. Paul Alduy, 66860 Perpignan, France

gourbiere@univ-perp.fr

Keywords: Genome Ecology, Retrotransposons, Copy Number Variations, Transcriptional and Post-transcriptional Gene Silencing, Bernouilli equation.

Abstract

Transposable elements (TEs) are essential components of the eukaryotic genomes. While mostly deleterious, evidence is mounting that TEs provide the host with beneficial adaptations. How 'selfish' or 'parasitic' DNA persists until it helps species evolution is emerging as a major evolutionary puzzle, especially in asexual taxa where the lack of sex strongly impede the spread of TEs. Since occasional but unchecked TE proliferations would ultimately drive host lineages toward extinction, asexual genomes are typically predicted to be free of TEs, which contrasts with their persistence in asexual taxa. We designed innovative 'Eco-genomic' models that account for both host demography and within-host molecular mechanisms of transposition and silencing to analyze their impact on TE dynamics in asexual genome populations. We unraveled that the spread of TEs can be limited to a stable level by density-dependent purifying selection when TE copies are over-dispersed among lineages and the host demographic turn-over is fast. We also showed that TE silencing can protect host populations in two ways; by preventing TEs with weak effects to accumulate or by favoring the elimination of TEs with large effects. Our predictions may explain TE persistence in known asexual taxa that typically show fast demography and where TE copy number variation between lineages is expected. Such TE persistence in asexual taxa potentially has important implications for their evolvability and the preservation of sexual reproduction.

1. Introduction

How organisms are able to adapt to new environmental conditions is central to evolutionary biology, and to unravel the determinants of such adaptation has increasing socio-economic implications in the context of global changes (1). The scientific knowledge used to address those challenges is rooted in standard genetic and ecological studies. Genetic approaches provide insights into traits adaptive changes by accounting for a description of the underlying sets of genes and their interactions (2). Ecological approaches are essential in identifying the specific forms of frequency- and density-dependent selection that emerge from detailed descriptions of ecological interactions (3). In both cases, the underlying Neo-Darwinian view implies the evolution of genes determining life-history traits and/or reproductive isolation (4).

The revolution in sequencing technologies has revealed that such genes only represent a minor fraction of eukaryotic genomes as non-genes typically account for about two-thirds of our own genome (5) and up to 85% of the maize genome (6). Non-coding sequences were soon referred to as the genomic ‘dark matter’ by analogy with the hypothetical substance that is predicted to account for around five-sixths of the matter of the universe (7). Comparison does not hold far behind such figures as this part of the genome is anything but made of slow moving massive particles that weakly interact with normal matter. Instead, genomic and post-genomic studies have started to shade light on the importance of those non-genic sequences that can potentially dwarf the information of the genes. Not only can non-genes contribute to the regulation of gene expression (8), but they can also be highly mutagenic actors altering genome structure (9), standing genetic variations and species evolvability (10).

These drastic changes in our perception of genomes structure and dynamics are driven by terabytes of data generated by ever-higher throughput genomic, transcriptional and post-transcriptional studies. With such an unprecedented accumulation of information, new opportunities are emerging to better understand the mechanisms underlying micro-evolution (11) and biological innovations (12). As in other fields of biology, theoretical approaches are essential in digging up knowledge from complex datasets (13). This has led to repeated calls to develop an ‘Ecology of the Genome’, an approach that aims at adapting ecological concepts and models to comprehend the interactions between the genic and non-genic entities shaping genomes (14,15).

The broad objective of the present study is to contribute to the emergence of this approach by a theoretical investigation of the dynamics of retro-transposons, i.e. class I transposable elements, thereafter simply referred to as 'TEs' although, *sensu stricto*, the definition 'TEs' also include class II elements. Those TEs are widely spread in eukaryotic species, typically representing 10-50% of their DNA (6) and constituting a substantial amount of the genomic 'dark-matter'. They indeed have a high potential to spread *via* 'copy and paste' mechanisms, which raises a fundamental question at the heart of genome ecology; what is regulating TE proliferation and the correlated increase in genome size by restraining their copy number to an equilibrium level?

Most theoretical answers to this question have been focused on two hypothetical regulatory mechanisms. Modelling studies have shown that i) selection against deleterious effects of TEs can allow for a stable number of copies if such effects show negative synergistic epistasis (16,17), and that ii) regulation of transposition can stabilize the number of TE copies through competitive interactions between copies (18-21). Comparisons with the theory of host-parasite interactions suggest that several features lacking in the above genetic models could have an impact on the predicted TE dynamics. First, host population size is typically considered as a constant in these models, so that they not allow tracking the effect of the spread of TEs on the population size. The corresponding models therefore do not account for the density-dependent feedback between the spread of TEs and the host population dynamics, while it is clearly established in theoretical epidemiology that such feedback plays a key role in regulating 'true' parasite populations (22). Second, genetic models lack a flexible description of the distribution of the number of TE copies per host individual that is considered to be Poisson (23), while experimental (24,25) and theoretical (26,27) studies have shown that over-dispersed TE distribution emerge from self-fertilization or heterogeneities between lineages. The theoretical evidences that aggregation of 'true' parasites in hosts significantly affect the stability of their interactions (28-30) suggest that such over-dispersion of TE distribution is likely to have an impact on TE persistence. Third, most genetic models do not consider TE epigenetic silencing despite its ubiquitous effect on transposition (6,8,9,12) and while similar mechanism of within-host developmental delays, such as dormancy of 'true' parasites, have been shown to significantly change host-parasite dynamics (29,30). While TE copies can also be inactivated by mutations, such as insertions or deletions, we restrained ourselves from considering the distribution of non-transposing remnant copies (as in 31,32), although those can potentially contribute to host adaptation (33) and genome size variations (34).

In this contribution, we developed original 'Eco-genomic' models of host-TE population dynamic based on analogies between TE dynamics and the transmission of 'macro-parasites', such as helminths and parasitic arthropods, in order to provide theoretical insights into the potential effects of the three above determinants that are currently not accounted for in our theoretical understanding of TE population dynamics and evolution.

2. Material and methods

The use of conceptual and modelling analogies between the dynamics of TEs and those of 'true' parasites has been proposed to improve our understanding of the former and its implication for genome structure and evolution (14,15). Here, we take the view that transposition, deletion and the effects of TEs on their hosts are similar to macro-parasite reproduction, death and virulence and we adapt a well-established framework developed for macro-parasites (28-30) to model host-TE interaction dynamics.

2.1. Eco-genomic modelling of TE dynamics

The theory of host-macro-parasites interactions is built on a mathematical framework first proposed by Anderson and May (28) and that split the host population into categories defined with respect to the number i (i=0,...,∞) of parasites carried by individuals at time t (Figure 1A). The host-macro-parasite interaction is then modelled by describing the rates at which individuals enter and leave those categories according to the host demography and parasite transmission and virulence. By considering TE copies as individual parasites, this A&M framework can readily be used to describe the dynamics of a TE family within a population of its host genomes at time t. The following parts describe how we adapted this framework to account for the i) host demography, ii) deleterious impact of TE, iii) distribution of TE copies between individual genomes, and the iv) epigenetic inhibition of their transposition.

2.1.1. Model with no silencing

We first considered the simplifying assumption made by most existing models of TE dynamics that TEs are not silenced. The model shall then predict the numbers of individual genomes carrying i (i=0,...,∞) active elements at time t (Figure 1B) according to the host demographic processes, the transposition and deletion of TEs, and their deleterious effect on host demography.

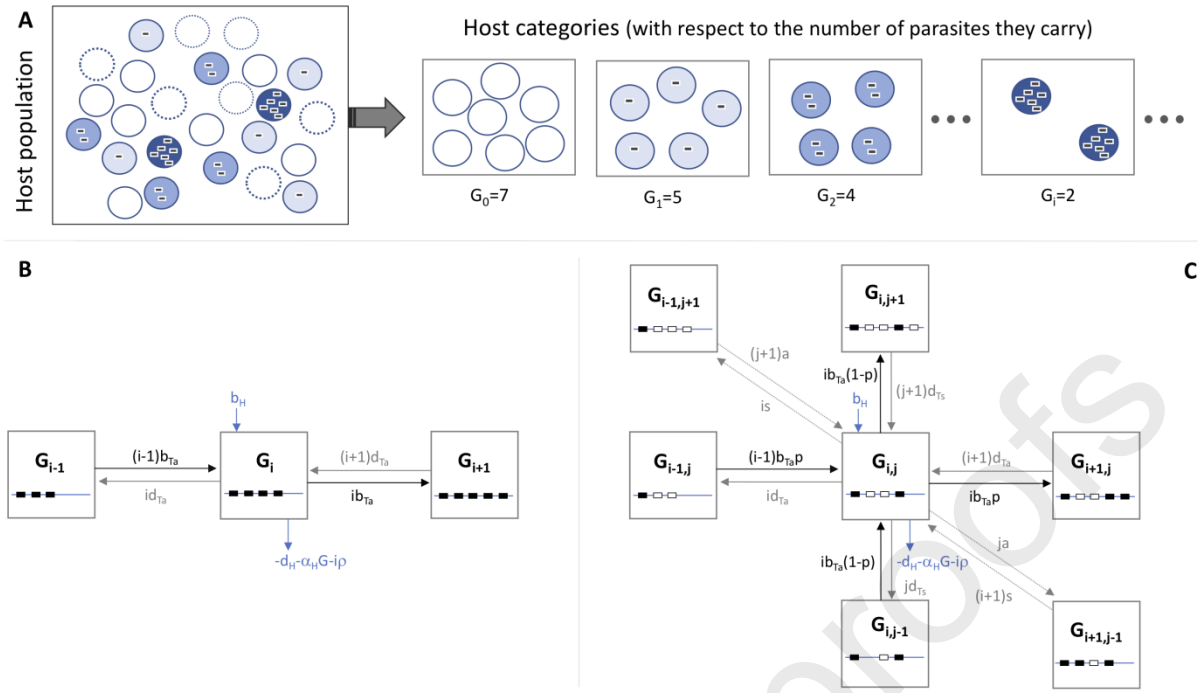

Figure 1. **Eco-Genomic Models of Host and Transposable Elements (TEs) Dynamics.** (A) The Anderson and May (28) modelling of the infection of an host population by macro-parasites. Host individuals are put into categories defined by the number i of parasites they carry, and G_i denotes the number of hosts found in category i. (B) Eco-genomic Model with no silencing. G_i denotes the number of hosts found with i active TEs (black boxes). (C) Eco-genomic Model with silencing. $G_{i,j}$ denotes the number of hosts carrying i active (black boxes) and j silenced (open boxes) TEs. In (A) and (B), the rates at which host individual genomes are gained or lost from categories G_i or $G_{i,j}$ are defined according to the processes and parameters described in the main text. The time dependency of the state variables has been removed for convenience.

Host demography. Each host individual can reproduce and die according to intrinsic per capita host birth (b_H) and death (d_H) rates, with asexual reproduction increasing the number of individuals in the parental category while deaths contribute to decrease it (Figure 1B). When considering only these two basic processes, the host population undergoes exponential growth. Although such a demography would typically occur during the colonization of a new ecological niche, the unbounded growth is sooner or later expected to be down-regulated because of competitive interactions between host individuals. This was modelled by considering a diminishing return term (α_H) increasing host mortality and ultimately leading to a logistic regulation of the host population.

TE transposition-deletion dynamics. Each active TE copy can transpose and be deleted at

constant rates b_{Ta} and d_{Ta}. These processes do not make the number of host genomes to vary, but only change the categories that individuals belong to. Transposition decreases the number of genomes carrying i TE copies and increases the number of genomes bearing i+1 copies, while deletion lower the number of hosts carrying i TE copies and increases the number of hosts bearing i-1 copies (Figure 1B).

Deleterious effects of TEs. The effects of TEs are modelled by defining a per copy increase (φ) of the host death rate and by considering the effects of copies to be additive. The TE-induced effect on host demography is then proportional to the number i of active TE copies (Figure 1B).

The basic Anderson and May's model for active TEs and host genomes. All the above processes summarized in the flow diagram appearing in Figure 1B can be lumped into the following differential equation describing the variations in the number $G_i(t)$ of host genomes carrying i TE copies at time t;

$$\frac{dG_i(t)}{dt} = \left(r_H - \alpha_H \sum_i G_i(t)\right) G_i(t) - i\,\varphi\, G_i(t) + (i-1) b_{Ta} G_{i-1}(t) - i\,(b_{Ta} + d_{Ta}) G_i(t) + (i+1) d_{Ta}\, G_{i+1}(t) \quad (1)$$

where $r_H = b_H - d_H$ and i varies between 0 and infinity.

This infinite system of equations can then be used to obtain the two equations describing the rates of change in the number of host genomes, $G(t)$, and in the number of active TE copies, $T_a(t)$. Since $G(t) = \sum_i G_i(t)$ and $T_a(t) = \sum_i i G_i(t)$, these two equations and the resulting model of host-TE dynamics in the absence of epigenetic silencing stand as (Appendix A):

$$\frac{dG(t)}{dt} = r_H\, G(t) - \alpha_H G^2(t) - \varphi\, T_a(t) \quad (2a)$$

$$\frac{dT_a(t)}{dt} = (r_H - \alpha_H\, G(t) + r_{Ta}) T_a(t) - \varphi\, G(t)\, E(i^2) \quad (2b)$$

where $r_{Ta} = b_{Ta} - d_{Ta}$.

These equations neatly show a first outcome of our modelling. The expectation of the squared number of TEs (i.e. $E(i^2)$) appearing in equation (2b) indeed implies that the TE dynamics depends on between host individual copy number variations. To determine further how TE

dynamics is impacted by such variations we considered two alternatives; a homogeneous distribution, where all individual carry the same number of copies and $E(i^2) = \left(\frac{T_a}{G}\right)^2$, and a heterogeneous distribution, described by a Negative Binomial law whereby $E(i^2) = \frac{T_a}{G} + \frac{k+1}{k}\left(\frac{T_a}{G}\right)^2$. The rational behind the use of a negative binomial is that the TEs distribution is likely to be over-dispersed because of the lower genetic exchanges associated with asexuality and heterogeneities in the transposition rate between host individuals or lineages (see discussion for details). A Poisson law, representing a random distribution of copies, can be recovered from the latter by making k converges to infinity, while finite k values generate over-dispersed distributions with some individual genomes carrying more copies than expected under a random distribution. By assuming the TE distribution to follow one of these distributions at any given time of the host-TE dynamics, the above expressions of $E(i^2)$ can be substituted into equation (2b) to obtain the closed forms of the model (Appendix A) that we analyzed using standard technics of dynamical systems theory (Appendices C and D).

2.1.2. Model with silencing

The above model with no silencing was modified to account for TE epigenetic silencing and therefore predict the numbers of individual genomes carrying i active and j silenced TEs (i,j=0,...,∞) at time t (Figure 1C), according to the processes already modelled in the previous section as well as the epigenetic silencing and (re)activation of TEs.

Silencing and activation of TEs. We accounted for a constant rate of epigenetic silencing applying to all existing active copies (s) and for the silencing of new TE copies that transpose into a silenced area of the genome. The latter was modelled by assuming a new copy to transpose into euchromatine or heterochromatine (where they are silenced) with probability p and 1-p, respectively. We further defined a per copy rate of activation (a) applying to all silenced copies and a specific rate of deletion applying to silenced copies (d_{Ts}). As for transposition and deletion, these processes do not make the number of host genomes to vary, but only change the categories that host individuals belong to (Figure 1C).

The structured Anderson and May's model for active-silenced TEs and host genomes. All the processes represented in Figure 1C can be lumped into the following differential equation

describing the variations in the number $G_{i,j}(t)$ of host genomes carrying i active and j silenced TE copies at time t:

$$\frac{dG_{i,j}(t)}{dt} = \left(r_H - \alpha_H \sum_{i,j} G_{i,j}(t)\right) G_{i,j}(t) - i\,\varphi\, G_{i,j}(t) + (i-1) b_{Ta}\, p\, G_{i-1,j}(t)$$
$$- \left(i\,(b_{Ta} + d_{Ta}) - j\, d_{Ts}\right) G_{i,j}(t) + (i+1) d_{Ta}\, G_{i+1,j}(t)$$
$$+ i\, b_{Ta}(1-p)\, G_{i,j-1}(t) + (j+1) d_{Ts}\, G_{i,j+1}(t)$$
$$+ (i+1)\, s\, G_{i+1,j-1}(t) + (j+1)\, a\, G_{i-1,j+1}(t) - (i\, s + j\, a) G_{i,j}(t) \quad (3)$$

where i and j varying from 0 to infinity.

From this infinite system of equations, we derived the set of three equations describing the rates of host genomes, $G(t)$, the number of active TE copies, $T_a(t)$, and the number of silenced TE copies, $T_s(t)$ at time t (Appendix A). The resulting model of host-TE dynamics in the absence of silencing stands as:

$$\frac{dG(t)}{dt} = r_H\, G(t) - \alpha_H G^2(t) - \varphi\, T_a(t) \quad (4a)$$

$$\frac{dT_a(t)}{dt} = (r_H - \alpha_H\, G(t) + r_{Ta}) T_a(t) - \varphi\, G(t)\, E(i^2) - s T_a(t) + a T_s(t) \quad (4b)$$

$$\frac{dT_s(t)}{dt} = (r_H - \alpha_H\, G(t) - d_{Ts}) T_s(t) + b_{Ta}(1-p) T_a(t) - \varphi G(t) E(ij) \quad (4c)$$
$$+ s T_a(t) - a\, T_s(t)$$

with $r_{Ta} = b_{Ta}p - d_{Ta}$, which is equivalent to its definition associated with equation (2b) where, in the absence of silencing, p=1.

Further considering that epigenetic silencing and activation occur typically much faster than the birth, death, transposition and deletion processes, we assumed them to be at dynamical equilibrium whenever the other processes occur (Appendix A). The proportion of active TE copies is then $p_a = \left(1 + \frac{s}{a}\right)^{-1}$, which allows re-writing equations 4 to obtain the model of host-TE dynamics in the presence of silencing:

$$\frac{dG(t)}{dt} = r_H\, G(t) - \alpha_H G^2(t) - \varphi\, T_a(t) \quad (5a)$$

$$\frac{dT_a(t)}{dt} = (r_H - \alpha_H\, G(t) + r_{Ta})T_a(t) - \varphi\, G(t)\, E(i^2) \quad (5b)$$

$$\frac{dT_s(t)}{dt} = (r_H - \alpha_H\, G(t) - d_{Ts})T_s(t) + b_{Ta}(1-p)T_a(t) - \frac{s}{a}\varphi G(t) E(i^2) \quad (5c)$$

where $r_T = p_a r_{Ta} - (1-p_a)d_{Ts}$ and $r_{Ta} = b_{Ta}p - d_{Ta}$.

The equation describing the variations of the total number of TEs at time t, referred to as $T(t)$, can then be derived by summation of equations (5b) and (5c):

$$\frac{dT(t)}{dt} = (r_H - \alpha_H\, G(t) + r_T)\, T(t) - \frac{\varphi}{p_a} G(t)\, E(i^2) \quad (5d)$$

where $r_T = p_a r_{Ta} - (1-p_a)d_{Ts}$.

Using the same expression of $E(i^2)$ as in the absence of silencing, the closed forms of the model were obtained under the assumptions of homogeneous and heterogeneous TE distributions (Appendix A), and were analyzed using the standard technics of dynamical systems (Appendices C and D).

2.1.3. A Bernouilli equation for the mean number of TE copies per genome

The models presented above provide equations describing the dynamics of the number of individual genomes and active and silenced TEs. Those equations can further be used to predict the dynamics of the mean number of TE copies per genome, $\overline{T}(t)$ (Appendix B). The equation for the dynamic of $\overline{T}(t)$ then takes on a simple form:

$$\frac{d\overline{T}(t)}{dt} = r_T \overline{T}(t) + \varphi p_a \overline{T}^2(t) - \frac{\varphi}{p_a} E(i^2) \quad (6)$$

where r_T and p_a (as defined above) take on maximal values, i.e. $p_a = 1$ and $r_T = r_{Ta}$, in the absence of epigenetic silencing.

The expectation $E(i^2)$ can again be replaced by its expression assuming a homogeneous or a heterogeneous distribution of TE copies. Whatever be this distribution, equation (6) turns out to be a Bernouilli equation (Appendix B). This equation has a high level of generality as it can be used to predict the dynamics of $\overline{T}(t)$ while considering null or deleterious effect on their host under any assumptions about the host demography, distribution of TEs between individual genomes and the rate at which the silencing and activation occur.

3. Results

We first made predictions for neutral TEs, before to account for the deleterious effects that TEs can have on the host demography. The detailed analyses of all neutral and non-neutral models can be found in Appendix C and are summarized in Appendix D. Their main outcomes are presented below.

3.1. What are the effect of host demography, copy number variation and silencing on neutral TE dynamics?

In the absence of TE effects (φ=0), equations (2b), (5d) and (6) are linear and lead to exponential dynamics. We identified the analytical expressions of the exponential rate of variations to provide clear predictions about the effects of host demography and TE distribution, epigenetic silencing and activation on neutral TE dynamics (Figure 2 and text below).

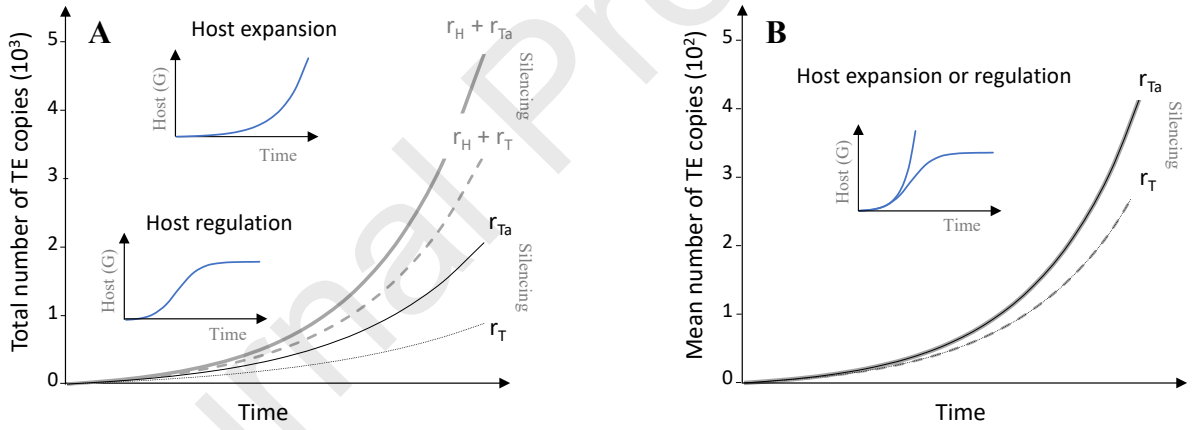

Figure 2. **Dynamics of neutral TE - Host interactions.** Neutral models always predict exponential variations of the *total* or *mean number of TE copies*. (A) The rate of increase of the *total number of TE copies* is higher in host population growing exponentially (bold grey) than in regulated population (thin black) and lower with epigenetic silencing (dotted lines) than without (continuous lines). (B) The rate of increase of the *mean number of TE copies* varies only with the presence or absence of silencing.

Host demography and TE distribution. During the expansion of a host population (α_H=0), the rate of increase of the *total number of TE copies* in the population (T) equals r_H+r_{Ta} (Figure 2A), whatever be the distribution of TE copies (models 1.1.1 in Appendix D). The host demography (r_H=b_H-d_H) and the within-host TE dynamics (r_{Ta}=b_{Ta}-d_{Ta}) then make equal and

additive contributions to the spread of TEs. When the host population dynamics is down-regulated ($\alpha_H \neq 0$), the rate of increase of T is predicted to slow down and converge to r_{Ta} as it is no longer fueled by the host population growth (Figure 2A). This, again, is independent of the TE distribution (models 1.1.2 in Appendix D). Importantly, the above effects of host demography on the dynamic of T during a host expansion vanishes when looking at the *mean number of TE copies* per individual genome ($\overline{T}$), so that the rate of increase of $\overline{T}$ equals r_{Ta} for both exponential and regulated host population (Figure 2B, models 1.1.1 and 1.1.2 in Appendix D). In the absence of TE effects, there is indeed no indirect 'host-mediated' interaction between neutral TE copies. The number of copies then varies at the same rate in every host individual, whatever the number of copies it bears and, accordingly, $\overline{T}$ follows the exact same density-independent transposition-deletion process as the number of neutral TE copies in each host. Obviously, this also explains why the rate of increase of $\overline{T}$ does not depend on the TE distribution (models 1.1.1 and 1.1.2 in Appendix D).

Silencing. The process of epigenetic silencing has no impact on the nature of the growth of the number of neutral TE copies that remains exponential, although it affects the rate of such increase. While the expressions of the rate of increase of T and $\overline{T}$ are similar to those obtained in the absence of silencing (models 1.2 in Appendix D), they now depend on the arithmetic mean (r_T) of the rates of transposition-deletion of active and silenced copies, $r_T = p_a.r_{Ta} - (1-p_a).d_{Ts}$, where p_a is the proportion of active copies. Since silencing and activation typically occur at higher rates than the other processes, the proportion of active TEs is considered to reach the dynamical equilibrium between these two molecular processes; $p_a = \left(1 + \frac{s}{a}\right)^{-1}$ in every host individual. As expected, epigenetic silencing then slows down the spread of TEs as a result of a reduced transposition activity (Figures 2A-B). Potentially, silencing could control the neutral dynamic of $\overline{T}$ by turning its (exponential) increase into decrease. This requires $s>a.\ r_{Ta}/d_{Ts}$, i.e. the rate of epigenetic silencing to exceed the rate of activation of silenced copies by an amount that correspond to their relative contributions to r_T. Again, in the absence of TE effect, the dynamic of the number of TE copies in every individual and its average $\overline{T}$ follow the same transposition-deletion-silencing-activation density-independent process, which explains why the dynamics of $\overline{T}$ is independent of host demography and TE distribution (models 1.2 in Appendix D).

Altogether, these results provide a simple theory of neutral TE dynamics in a population of host individual genomes. The *mean number of TE copies* per genome varies exponentially at a rate given by a meaningful quantity (r_T) that only depends on the within individual genome processes. The host population processes only impact the variation of the *total number of TE copies* during host expansion with the host population growth (r_H) providing a purely additive contribution to those variations.

3.2. When are TE dynamics stabilized by host-mediated purifying selection?

In the presence of TE deleterious effects ($\varphi \neq 0$), equations (2b), (5d) and (6) become non-linear and the host and TE dynamics can feedback on one another, which lead to five different dynamical outcomes (Figure 3). We performed stability analysis to identify the conditions where each of these dynamics occur with respect to the host demography, TE distribution and the absence or presence of epigenetic silencing (Figure 4) in order to predict when purifying selection could allow for TEs to persist in asexual populations.

Host demography and TE distribution. When the distribution of TE copies between individuals is homogeneous (models 2.1.1.1 and 2.1.2.1 in Appendix D), the host demography has no effect on TE dynamics. In exponentially growing ($\alpha_H = 0$) and regulated ($\alpha_H \neq 0$) host populations, the spread of TEs leads to the 'Co-extinction' of both hosts and TEs (Figure 3A).

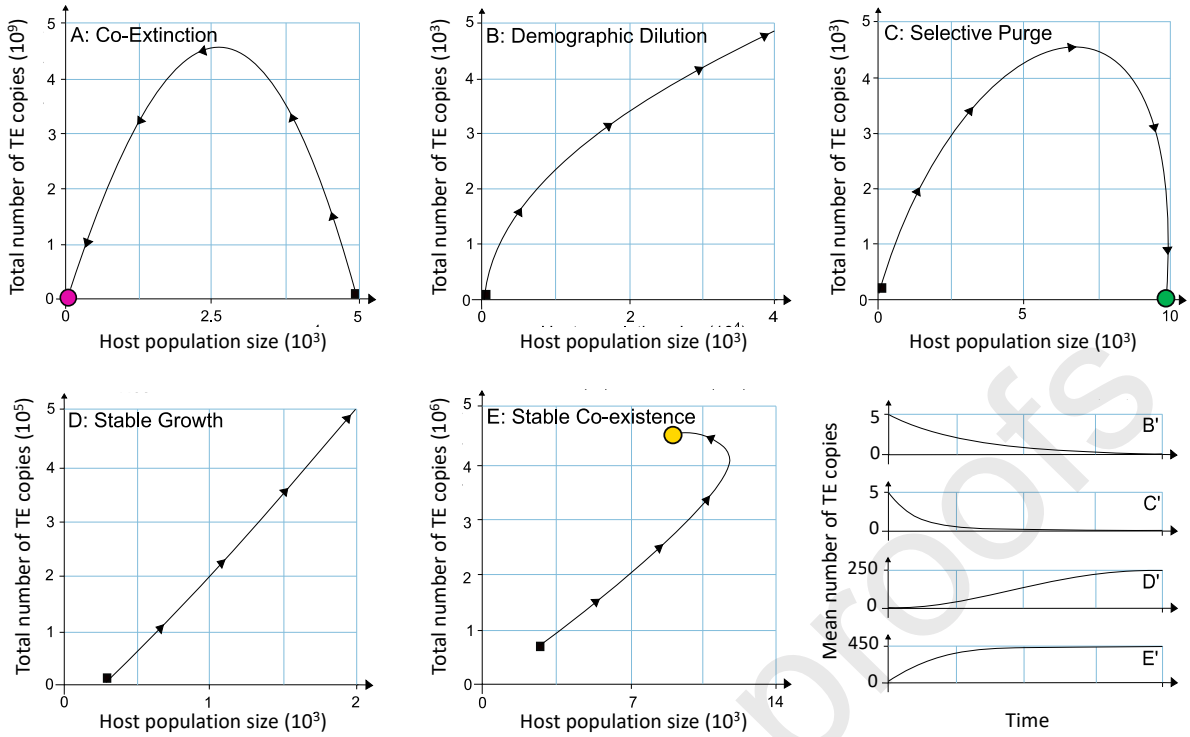

Figure 3. **Dynamics of deleterious TE - Host interactions.** Non-neutral models predict either the 'Co-Extinction' of hosts and TEs (A), a 'Demographic Dilution' as the number of host ultimately grows faster that the *total number of TE copies* (B), a 'Selective Purge' as the host population persist free of TEs (C), a 'Stable Growth' whereby the number of hosts and TE copies ultimately grow at the same rate (D), or a 'Stable coexistence Equilibrium' where the numbers of hosts and TE copies reach a non-trivial equilibrium (E). In cases (B) and (C) *the mean number of TE copies* converges toward 0 (see B', C'), while in cases (D) and (E) it converges to a stable equilibrium (see D', E').

When there are copy number variations between individuals (models 2.1.1.2 and 2.1.2.2 in Appendix D), the dynamical outcomes are much more diverse. During a host expansion (α_H=0) and when copy number variations follows a Poisson distribution ($k\to\infty$), the spread of TEs can also lead to 'Co-Extinction', but only if deleterious effects are weak, i.e. $\varphi < r_{Ta}$ (Figure 4A). Increasing TE effects above that first threshold allows for the host population to persist and for the *total number of copies* in the population (T) to increase (Figures 3B and 4A), although the *mean number of copies* per host $\overline{T}$ converges towards 0 (Figure 3B'). We refer to this dynamic as a 'Demographic Dilution' of TEs as it results from an unbounded increase in the number of individuals with low number of copies overwhelming a similarly unbounded, albeit slower, increase in the number of host individuals carrying more TE copies. This dynamic is associated with host expansion and disappears in regulated host populations (Figure 4B). When TE effects exceed a second threshold, i.e. $\varphi > r_{Ta}+r_H$ (Figure 4A) or $\varphi > r_{Ta}$ (Figure 4B), individuals with

more copies are sufficiently selected against for all TE copies to vanish from the host population, which then persists free of TEs (Figure 3C, C'). We refer to such a dynamic as a 'Selective Purge'.

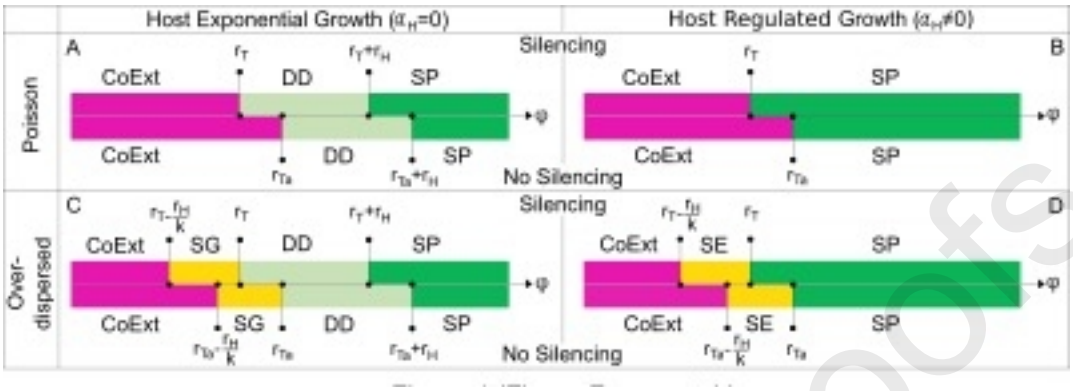

Figure 4. **Conditions for the stabilization of TE - Host dynamics in the presence of TE copy number variations.** The conditions where the dynamics presented in Figure 3 occur are given with respect to φ under the assumptions considered for the host demography, the distribution of TEs, and with or without silencing. CoExt=Co-Extinction, DD=Demographic Dilution, SP=Selective Purge, SG=Stable Growth, SE=Stable coexistence Equilibrium.

Importantly, in all conditions considered above, selection never leads to a stable number of TEs per genome. A stable TE equilibrium only emerged when we considered the copy number variations to correspond to an over-dispersed TE distribution by letting k takes on finite values (Figure 4C, D). To increase TE distribution heterogeneity indeed provides a new type of dynamics whereby the mean number of copies $\overline{T}$ reaches a stable equilibrium (Figure 3D', section 2.1.1.2 in Appendix C), while the total number of hosts and TE copies still grow exponentially in the absence of host demographic regulation (α_H=0, Figures 3D and 4C). Interestingly, such a 'Stable Growth' is only obtained for TE with intermediate effects, i.e. when $r_{Ta} - r_H/k < \varphi < r_{Ta}$ (Figure 4C). Lower values of φ lead to 'Co-Extinction' and larger effects lead to a 'Demographic Dilution' or a 'Selective Purge' in a similar way as described when the copy number variation is assumed to follow a Poisson distribution. In a logistically regulated population (α_H≠0), TE dynamics are similar to what we just described during host expansion, but for two main differences (Figure 4D). First, 'Demographic Dilution' is no longer expected and instead TEs are eliminated since competition between hosts lead to selection against individuals that bear more TE copies. Accordingly, the range of φ-values where a 'Selective Purge' is expected extends to $\varphi > r_{Ta}$. Second, when TE distribution is over-dispersed,

the 'Stable Growth' of TE with intermediate effects is replaced by a 'Stable Equilibrium' where $\overline{T}$ reaches an equilibrium (Figures 3E' and 4D) with the total number of hosts and TEs converging toward stable equilibrium levels (Figure 3E) because of the host population regulation (Appendix C section 2.1.2.2 and next section).

Silencing. The process of epigenetic silencing does not change the types of host-TE dynamics, but it affects the ranges of φ values that lead to each of the dynamics presented in Figure 3A-E. The formal expressions of all thresholds separating those different ranges of φ values (Figure 4A-D) are equivalent to those derived in the absence of silencing, although they now depend on the arithmetic mean of the rates of growth of active and silenced copies (r_T) instead of the rate of active TEs (r_{Ta}). Since r_T is constitutively smaller than r_{Ta}, all thresholds are shifted towards smaller φ values. Accordingly, whatever the assumptions about the host demography and the distribution describing copy number variations, the range of φ values driving hosts and TEs to 'Co-extinction' decreases while the range of φ values allowing for a 'Selective Purge' increases (Figure 4A-D). When the conditions are met for TE to persist, the proportion of active TEs is again considered to reach the dynamical equilibrium between silencing and activation processes, i.e. to be equal to $p_a = \left(1 + \frac{s}{a}\right)^{-1}$.

Altogether the results of this section provide original conclusions about the ability of selection to regulate the proliferation of TE copies within genomes. First, selection can stabilize TE dynamics if copy number variations correspond to an over-dispersed distribution. Second, such a regulation further requires TE effects of intermediate levels since too weak effects lead to TE proliferation and host extinction, while TEs with too strong effects get eliminated. Third, TE silencing can protect host populations in two ways; either by preventing TEs with weak effects to grow in large number and to induce host extinction, or by favoring the elimination of TEs with large effects.

3.3. What are the equilibrium number of TE copies and the '*demographic load*' affecting the host population?

The above results have shown that, in a regulated host population, TEs can i) go extinct with their host population, ii) persist at some stable equilibrium, or iii) be eliminated by purifying selection (Figure 4D). We further aimed to predict the effect of the model demographic and molecular parameters on the actual equilibrium level of the numbers of host individuals and TE copies.

The analyses performed in Appendix C show that the *total number of TE copies* and the host population size at equilibrium are;

$$T^* = \begin{cases} 0 & \text{if } \varphi < r_T - r_H/k \\ \frac{k}{\alpha_H \varphi p_a}(r_H - k(r_T - \varphi))(r_T - \varphi) & \text{if } r_T - r_H/k < \varphi < r_T \\ 0 & \text{if } r_T < \varphi \end{cases} \quad (7)$$

and

$$G^* = \begin{cases} 0 & \text{if } \varphi < r_T - r_H/k \\ \frac{1}{\alpha_H}(r_H - k(r_T - \varphi)) & \text{if } r_T - r_H/k < \varphi < r_T \\ \frac{r_H}{\alpha_H} & \text{if } r_T < \varphi \end{cases} \quad (8)$$

Accordingly, when TEs are able to persist, i.e. if r_T-$r_H/k < \varphi < r_T$, *the mean number of TE copies* per individual at equilibrium is;

$$\overline{T}^* = \frac{k}{\varphi p_a}(r_T - \varphi) \quad (9)$$

This relationship clearly shows that the *mean number of TE copies* per genome decreases with φ from a finite value, $\frac{r_H}{p_a(kr_T - r_H)}$, towards 0 (Figure 5A).

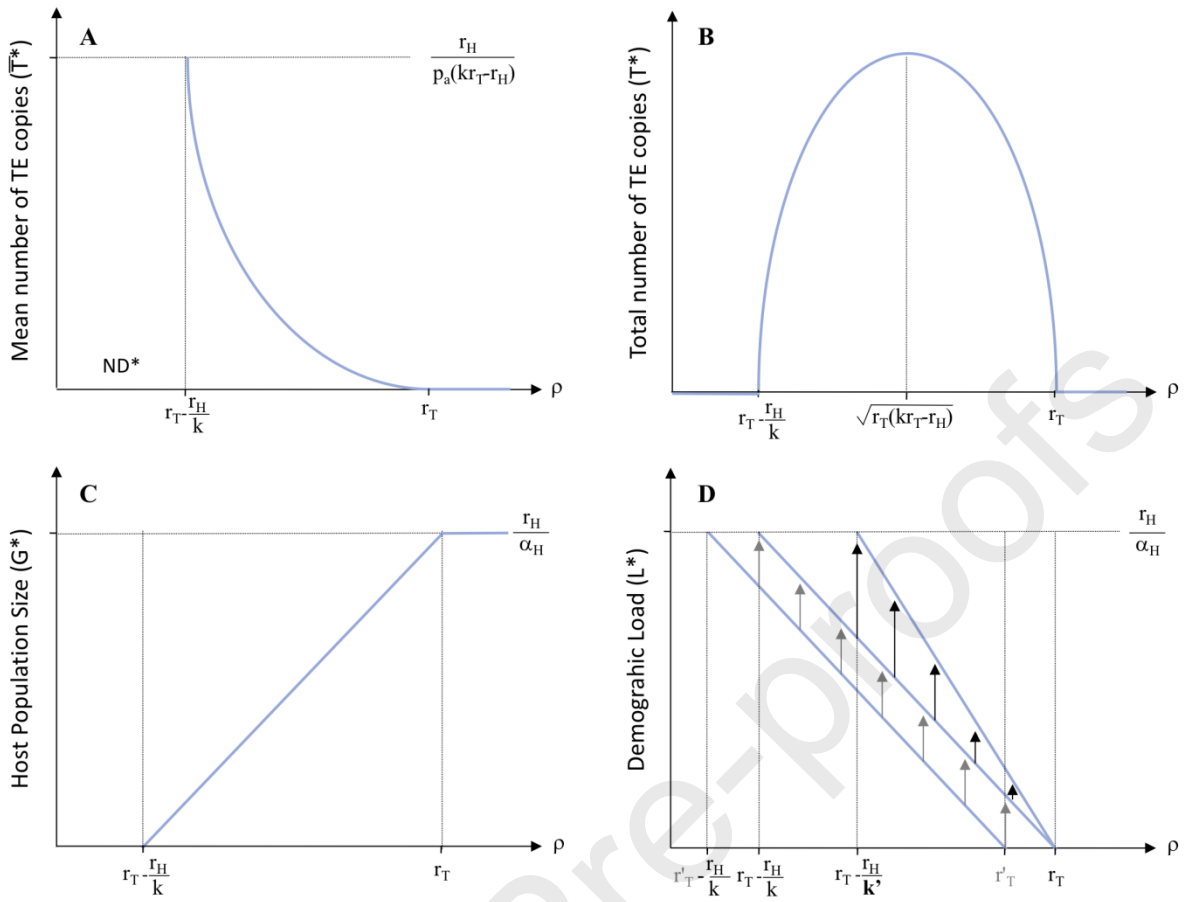

Figure 5. **Coexistence between TEs and their Hosts and the Demographic Load.** Variation of the equilibrium level of the *mean number of TE copies* per genome (A), the *total number of TE copies* (B), the host population size (C) and the Demographic Load (D) with the respect to the effect of TEs on their individual host (φ). In (D), $r_T' < r_T$ and $k' > k$. Grey and black vertical arrows show the increase in L^* due to an increase in r_T (to r_T') and k (to k'), respectively. Results were all obtained with the complete model including both a logistic regulation of the hosts and TE silencing. ND=Not-Defined as both the *total number of TE copies* (T*) and the number of host individual genomes (G*) are 0.

This, however, can be partially deceptive as the *total number of TE copies* is expected to increase with φ until it reaches a maximum at $\varphi = \sqrt{r_T(kr_T - r_H)}$ before to decrease toward 0 (Equation (7), Figure 5B). Such apparent discrepancy is explained by the variations of the number of host genomes (Equation (8), Figure 5C). Interestingly, despite non-linear variations in the total number of TEs, their impact on the host population size is linear. This impact can be measured by a *TE demographic load* defined as the difference between the host population sizes expected in the absence and in the presence of TEs;

$$L^* = \frac{k(r_T - \varphi)}{\alpha_H} \qquad (10)$$

This expression clearly depicts the demographic implications of the purifying selection process; the effect of TEs on the host population size (L^*) decreases when their effect at the individual level (φ) increases (Figure 5D). Host-mediated purifying selection is indeed more efficient against more deleterious TEs, so that such TEs reach lower number of copies per genome, as shown by equation (9), which leads to a smaller reduction of the host population size. This TE demographic load increases with k as larger values of k (such as k' in Figure 5D) lead to less heterogeneous TE distribution, which weaken the efficacy of selection for any given φ value. Meanwhile, any increase in the relative rate of epigenetic silencing and activation (s/a) reduces the impact of TEs on their host population size. Larger s/a ratio indeed lead to a higher proportion of silenced TEs (p_A), which lowers r_T since $r_T = p_a.r_{Ta} - (1-p_a).d_{Ts}$, and in turn reduces L^* for any given φ value.

4. Discussion

An appealing perspective in the 'Ecology of the Genome' approach (14,15) is to understand TE dynamics in a host population whose size depends explicitly on both its ecology and the TE effect on individuals' demography. By adapting a well-established model of host-macro-parasites interactions (28), we developed an innovative 'Eco-genomic' modelling of the spread of retrotransposons in asexual populations. This modelling was intended to provide original insights according to three of its specific features; i) the effect of TEs on the host population size are accounted for, therefore leading to a *density-dependent* purifying selection regime, ii) the copy number variations is allowed to follow an over-dispersed distribution, which, for asexual taxa, appears as a more sensible assumption than the standard Poisson distribution, and iii) the ubiquitous mechanisms of TE epigenetic silencing are considered.

4.1. The implications of copy number variations and density dependent purifying selection on TE dynamics

A main outcome of our modelling is that the average number of TEs per genome can reach a stable equilibrium at which the host population suffers a TE-induced demographic load. Although an equilibrium was previously reported for selfing (26) and asexual (35) populations, this only hold under two strong assumptions; i) the absence of element excision and ii) an infinite host population size. Relaxing the former generates TE free individuals that ineluctably spread to fixation in an infinite population, which purges the population from TEs whatever their impacts on host fitness (35). By contrast, considering finite populations consistently lower the efficiency of purifying selection and lead to an unbounded proliferation of TE copies (35,36). As expected, models relaxing both assumptions predict the existence of thresholds around which small differences in TEs or host features can switch the dynamics from TE elimination to TE accumulation leading to host extinction (26,35,37,38). Those results support the assertion that asexual (or strongly selfing) populations shall not bear TEs, or only for a transitory period (39). However, this contrasts with the presence of TEs in bdelloid rotifers (40) or mites (41) that have been reproducing asexually for million years, parthenogenetic micro-crustaceans (42), asexual parasitoid wasp (43), and in unicellular amoebas (44) or algae (45) that show no or very low levels of sexual reproduction. Several hypotheses have been explored to explain TE persistence. Although these theoretically studies typically consider some level of host sexual reproduction, which favours the spread of TEs, the regulating processes investigated are still relevant for asexual populations. A first mechanism to limit TE expansion is that

549 insertions induce deleterious non-homologous recombination (16,26,46). As the probability of
550 TE-mediated ectopic recombination is expected to increase with the square of copy number, it
551 produces a negative feedback on the TE population growth rate leading to a stable number of
552 copies. While these synergistic effects were included in the two models considering selfing (26)
553 and asexual (35) hosts, they did not allow for an equilibrium to be reached in finite populations.
554 Alternatively, regulating feedbacks could emerge from cis- or trans- acting mechanisms of
555 transposition repression at the host individual level. The rate of transposition of a given copy
556 would then decrease with the number of copies located in its genomic neighbourhood or within
557 the individual genome. These mechanisms of transposition 'immunity' and 'repression' can
558 produce stable copy number in sexual populations (16,18). While such self-regulation of
559 transposition was not included in the two above models considering asexuals or selfers, such
560 populations are expected to be prone to its evolution (19), which could potentially explain TE
561 persistence (41). Finally, horizontal transfers can balance TE elimination through source-sink
562 dynamics (37,47) and beneficial insertions can maintain slightly deleterious donor copies and
563 preserve TEs from extinction in both constant (47,48) and fluctuating (49) environments.
564 Interestingly, none of these regulatory mechanisms were included in our models, and yet a
565 stable equilibrium was reached. This unveils an original mean by which TEs can persist stably
566 in asexuals when indirect host-mediated interactions between copies generate density-
567 dependent demographic regulation and purifying selection. If the demography of a host
568 individual depends on the number of TEs it bears, the TE population growth rate is indeed set
569 to decrease with the square of copy number - even if TEs have purely additive effects on host
570 fitness - as evidenced in equations 5(b-d) and (6). A stable equilibrium can then emerge,
571 provided that the variability in TE copy number between individuals is greater than expected
572 under a Poisson distribution. While measurements of this distribution are lacking, it seems
573 likely to be over-dispersed in asexual and selfing populations for two reasons. First, the
574 variability in the number of TE copies increases with the lower genetic exchanges (26,50) and
575 larger population sizes (23,35) associated with those reproductive modes. Straightforward
576 calculations show that the coefficients of variation of copy number emerging from simulations
577 of TE dynamics in selfing populations (26) are equivalent to low values of k ($\sim 10^{-3}$-10^{-4}). Such
578 values correspond to strongly skewed distributions that enhance the strength of selection and
579 favour the maintenance of TEs by strongly reducing the lower boundary of the range of
580 deleterious effects of TEs where a stable equilibrium is reached, i.e. $r_T - r_H/k$. Second, an over-
581 dispersed TE distribution is expected whenever the transposition rate varies between host
582 individuals or lineages, as it shall be since transposition activity depends upon the local genomic

context (51), and on host (52) and environmental (53) factors. Such variations could be accounted for by a rate of transposition following a gamma distribution, just as substitution rate heterogeneity is modelled in phylogenetic studies (54). Interestingly, the number of TE copies then follow a gamma-Poisson, which effectively is a Negative Binomial distribution, and would thus fit the distribution considered here. Overall, our new modelling approach confirmed the existence of a threshold value whereby TE dynamics switches from elimination to accumulation if the TE distribution among individuals is Poisson. More significantly, it also unravelled that when this distribution is over-dispersed, which seems likely in asexual or selfing populations, a more gradual transition is expected and allows for a stable number of TEs per individual.

4.2. How do molecular mechanisms of epigenetic silencing contribute to the maintenance of TEs?

Theories investigating the success or failure of TE invasion focus on transposition-deletion-selection balances (16-20,23,26,27,35-38,46-49). While a substantial knowledge of the epigenetic mechanisms of TE silencing has accumulated (8,9,55,56), it never was accounted for and, accordingly, there exists no prediction about their effects on TE population dynamics. By accounting for those mechanisms, we showed that the rate of TE increase (r_T) is the arithmetic mean of the rates of transposition-deletion of active copies and the rate of deletion of silenced copies weighted by their relative proportion. This intuitive formulation illustrates that, as expected, silencing reduces the increase of copy number with key implications for the long-term TE dynamics. When the TE distribution is Poisson, silencing does not allow to reach a stable equilibrium, but it changes the ranges of TE effects where their elimination and accumulation are expected. The threshold level of deleterious effects delimiting when those outcomes are expected, i.e. r_T, decreases with the rate of silencing, so that silencing allows for purifying selection to remove elements with lower deleterious effect that would have proliferated otherwise. Although such conclusion may seem counterintuitive, as silenced copies are thought not to have strong deleterious effect on fitness, the trade-off with the lack of transposition actually drives TEs towards elimination. Similar predictions were made when assuming an over-dispersed distribution of copies. Silencing shifts the ranges of deleterious effect where the different dynamical outcomes (TE accumulation, equilibrium and elimination) are expected since the two thresholds values delineating those conditions, namely r_T-r_H/k and r_T, depend upon r_T in the same way. Silencing was then unable to broaden the conditions for a stable number of TE copies to be reached, but it limited the proliferation of elements with lower

effects and facilitates the elimination of elements with larger effects. Interestingly, in slowing down the accumulation of TEs with low effect and in allowing for their persistence in a stable equilibrium, silencing benefits not only the host but TEs themselves as they no longer drive the dynamics of interaction towards co-extinction. Altogether, these results suggest that previous theoretical studies, by not accounting for silencing, are likely to have over-estimated the risk of asexual lineage extinction due to TE accumulation and, accordingly, the associated advantage of maintaining sexual reproduction (35,39). By essence, theoretical predictions are born from modelling assumptions. While we considered a constant per copy probability of silencing (s), the two molecular mechanisms involved in TE epigenetic control, i.e. the small RNA interference (RNAi) and RNA-dependent DNA methylation (RdDm), could make the per copy probability of silencing to vary with the number of copies. The amplification of the pool of 21-22nt small interfering RNAs involved in post-transcriptional silencing (55) and the trans-acting effects of 24nt to guide the methylation of TEs (56) are indeed likely to help better controlling any additional copy. Although no assessments of the strength of such non-linear dynamics of silencing are available, modelling the RNAi pathway on its own (57,58) or with the RdDm pathway (59) has shown that they can regulate TE dynamics at host individual or lineage scales. How far such conclusion would hold in host population models accounting for selection against TE deleterious effect remain an open question. The molecular regulation of TE proliferation within each individual/lineage would indeed concomitantly lowers purifying selection by reducing the variability of the TE distribution between individuals. Investigating the conflict between individual and population scale regulation will require hybrid-models combining molecular and demographic processes, as exemplified in another context (60).

4.3. How deep can we think about TE as macro-parasites, and *vice versa*?

Comparisons are often drawn between TEs and virus as transposition, replication and their silencing involve similar molecular processes (e.g. 61). However, the link between TE number and their effects on host fitness is the same as those between macro-parasites and their impact on host morbidity and mortality (28-30). We thus adapted the modelling framework developed for macro-parasites, rather than SI models used to study infections by other (viral) genomic parasites (22, 37). As it accounts for an explicit description of the host demography and the effects of TE on those processes, our modelling provides a natural environment to investigate the impact of host ecological interactions and network structures that may play an important role in TE dynamics (62). The fate of TEs is indeed predicted while purifying selection emerges from the description of host ecology, which generate frequency- and density-dependent

adaptive dynamics that cannot be anticipated from genetic models assuming a constant population size (3). This background could then help answering topical questions about the role of TEs in host ecology and micro-evolution (9,14,15,53,62). Its second appealing feature is that it relies on compartmental modelling to split the host population into categories. While host individuals were differentiated by their number of active and silenced TE copies, compartments could partition the population into developmental stages, demes or genome structures, such as the division into euchromatine *vs*. heterochromatine, that all influence the rates of transposition and/or silencing. The proposed background thus carries a strong potential to better understand species specific differences in TE content (6). While further adapting the Anderson and May (28) background could provide new insights into the role of TEs in ecology and evolution, this may also feedback on our understanding of host-macro-parasites interactions. Although TEs share important features with macro-parasites, their primarily vertical transmission is very distinctive. Helminths congenital transmission is typically thought to be of secondary importance, although Schistosome ova and larvae have been found in human placenta and trans-placental passages of *Onchocerca volvulus* and *Wuchereria bancrofti* have been reported (63). We may then learn about the role of vertical transmission of macro-parasites from modelling TE dynamics. Already appearing here is that vertical transmission does not lead to asynchronous host and parasite oscillations that typically allow sustaining their interaction (64), but destabilizes the interaction by favouring parasite accumulation, which makes host and parasite co-extinction more likely. Hypothetically, the limited level of vertical transmission in macro-parasites may then be an adaptation to avoid co-extinction dynamics that hosts wouldn't be able to prevent through molecular mechanisms of silencing targeting such parasites, which is consistent with the higher level of congenital transmission observed for viruses whose replication can be controlled by the same molecular pathways as TEs.

We hope this study will contribute to the development of the 'Ecology of the genome' that is emerging as a key interdisciplinary field in today's biology in order to strengthen the analyses of massively accumulated genomic and post-genomic data. It exemplifies how incorporating biological knowledge from different fields into generic mathematical models helps circulating concepts across disciplinary boundaries.

682 # Acknowledgements
683

684 This work has benefited from a PhD fellowship to AFF (CONACYT, Person Number 239540).
685 This study is set within the framework of the "Laboratoires d'Excellences (LABEX)" TULIP
686 (ANR-10-LABX-41). The funders played no role in the study design, data collection and
687 analysis, decision to publish, or preparation of the manuscript.

References

1. Carroll SP, Jorgensen PS, Kinnison MT, Bergstrom CT, Denison RF, Gluckman P, Smith TB, Strauss SY, Tabashnik TB. 2014. Applying evolutionary biology to address global challenges. *Science*. **346**(6207). (DOI :1245993-1-1245993–10).
2. Mauricio R. 2005. Genetics of Adaptation. Dordrecht: Springer Netherlands. 209p. https://www.springer.com/gb/book/9781402034763
3. Dieckmann U, Doebeli M, Metz JAJ, Tautz D. Adaptive Speciation. 2004. Cambridge: Univ. Press. 460 p. (DOI : 10.1017/CBO9781139342179)
4. Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol Evol*. **23**(1):38–44. (DOI : 10.1016/j.tree.2007.09.008)
5. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. 2011. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genetics.***7**(12):e1002384. (DOI : 10.1371/journal.pgen.1002384)
6. Tenaillon MI, Hollister JD, Gaut BS. 2010. A Triptych of the evolution of plant transposable elements. *Trends in Plant Science.* **15**(8): 471-478. (DOI : 10.1016/j.tplants.2010.05.003)
7. Rubin VC, Ford WKJ, Thonnard N. 1980. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophys. J.* **238**, 471–787. (DOI : 10.1086/158003)
8. Elbarbary RA, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. *Science* **351**, aac7247–aac7247. (DOI : 10.1126/science.aac7247)
9. Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8**, 272–285. (DOI : 10.1038/nrg2072)
10. Lynch, VJ. GENETICS. A copy-and-paste gene regulatory network. 2016. *Science* **351**, 1029–1030. (DOI : 10.1126/science.aaf2977)

11.van't Hof, AE, Campagne, P. 2016. The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **534**, 102–105. (DOI: 10.1038/nature17951)

12. Lynch VJ, Leclerc RD, May G, Wagner GP. 2011. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* **43**, 1154–1159.

13. Cohen JE. 2004. Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better. *PLoS Biol.* **2**, e439.

14. Brookfield JFY. 2005. THE ECOLOGY OF THE GENOME — mobile DNA elements and their hosts. *Nat. Rev. Genet.* **6**, 128–136.

15. Venner S, Feschotte C, Biémont C. 2009. Dynamics of transposable elements: towards a community ecology of the genome. *Trends Genet.* **25**, 317–323.

16. Charlesworth B, Charlesworth D. 1983. THE POPULATION DYNAMICS OF TRANSPOSABLE ELEMENTS. *Genet. Res.* **42**, 1–27. (DOI : 10.1017/S0016672300021455)

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

18. Langley CH, Brookfield JFY, Kaplan N. 1983. Transposable elements in mendelian populations. I. A theory. *Genetics* **104**, 457–471.

19. Charlesworth B, Langley CH. 1986. THE EVOLUTION OF SELF-REGULATED TRANSPOSITION OF TRANSPOSABLE ELEMENTS. *Genetics* **112**, 359–383.

20. Le Rouzic AL, Capy P. 2006. Population Genetics Models of Competition Between Transposable Element Subfamilies. *Genetics* **174**, 785–793.

21. Basten CJ, Moody ME. 1991. A branching-process model for the evolution of transposable elements incorporating selection. *Journal of Mathematical Biology* **29**(8):743-761.

737 22. Diekmann O, Heesterbeek, HJ. 2000. *Mathematical Epidemiology of Infectious Diseases:*
738 *Model Buildign, Analysis and Interpretation*. Wiley, Chichster.
739 23. Brookfield JFY, Badge RM. 1997. Population genetics models of transposable elements.
740 *Genetica* **100**, 281–294.
741 24. Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH. 2000. Genome evolution of
742 wild barley (Hordeum spontaneum) by BARE-1 retrotransposon dynamics in response to
743 sharp microclimatic divergence. *Proc. Natl. Acad. Sci.* **97**, 6603–6607. (DOI :
744 10.1073/pnas.110587497)
745 25. Vielle-Calzada J-P, Martínez de la Vega O, Hernández-Guzmán G, Ibarra-Laclette E,
746 Alvarez-Mejía C, Vega-Arreguín JC, Jiménez-Moraila B, Fernández-Cortés A, Corona-
747 Armenta G, Herrera-Estrella L, *et al.* 2009. The Palomero genome suggests metal effects on
748 domestication. *Science* **326**, 1078. (DOI : 10.1126/science.1178437)
749 26. Wright SI, Schoen DJ. 1999. Transposon dynamics and the breeding system. *Genetica* **107**,
750 139–148 .
751 27. Iranzo J, Cuesta JA, Manrubia S, Katsnelson MI, Koonin EV. *2017.* Disentangling the
752 effects of selection and loss bias on gene dynamics. *PNAS*. **11**;114(28):E5616–24. (DOI :
753 10.1073/pnas.1704925114)
754 28. Anderson RM, May RM. 1978. REGULATION AND STABILITY OF HOST-PARASITE
755 POPULATION INTERACTIONS: I. REGULATORY PROCESSES. *J. Anim. Ecol.* **47**,
756 219. (DOI : 10.2307/3933)
757 29. Dobson AP, Hudson PJ. 1992. Regulation and Stability of a Free-Living Host-Parasite
758 System: Trichostrongylus tenuis in Red Grouse. II. Population Models. *J. Anim. Ecol.* **61**,
759 487–498. (DOI : 10.2307/5339)

760 30. Gaba S, Gourbiere S. 2008. To delay once or twice: the effect of hypobiosis and free-living
761 stages on the stability of host-parasite interactions. *J. R. Soc. Interface* **5**, 919–928. (DOI :
762 10.1098/rsif.2007.1282)
763 31. Sheinman M., Ramisch A., Massip F., Arndt PF. 2016. Evolutionary dynamics of selfish
764 DNA explains the abundance distribution of genomic subsequences. *Scientific reports*
765 **6**:30851.
766 32. Banuelos M., Sindi S. 2018. Modeling transposable element dynamics with fragmentation
767 equations. *Mathematical Biosciences* **302**: 46-66.
768 33. Casacuberta E., Gonzales J. 2013. The impact of transposable elements in environmental
769 Adaptation. Molecular Ecology. *Molecular Ecology* **22**:1503–1517.
770 34. Fischer S, Bernard S, Beslon G., Knibbe C. 2014. A Model for Genome Size Evolution.
771 *Bull Math Biol* **76**:2249–2291.
772 35. Dolgin ES, Charlesworth B. 2006. The Fate of Transposable Elements in Asexual
773 Populations. *Genetics* **174**, 817–827. (DOI : 10.1534/genetics.106.060434)
774 36. Docking TR, Saadé FE, Elliott MC, Schoen DJ. 2006. Retrotransposon Sequence Variation
775 in Four Asexual Plant Species. *J. Mol. Evol.* **62**, 375–387. (DOI : 10.1007/s00239-004-0350-
776 y)
777 37. Rankin DJ, Bichsel M, Wagner A. 2010. Mobile DNA can drive lineage extinction in
778 prokaryotic populations: Mobile DNA can drive lineage extinction in prokaryotic
779 populations. *J. Evol. Biol.* **23**, 2422–2431. (DOI : 10.1111/j.1420-9101.2010.02106.x)
780 38. Boutin TS, Le Rouzic A, Capy P. 2012. How does selfing affect the dynamics of selfish
781 transposable elements? *Mob. DNA* **3**, 5.(DOI : 10.1186/1759-8753-3-5)
782 39. Arkhipova I, Meselson M. 2005. Deleterious transposable elements and the extinction of
783 asexuals. *BioEssays* **27**, 76–85. (DOI : 10.1002/bies.20159)

784 40. Rodriguez F, Kenefick A. Arkhipova I. 2017. LTR-Retrotransposons from Bdelloid
785 Rotifers Capture Additional ORFs Shared between Highly Diverse Retroelement Types.
786 *Viruses* **9**(4), 78. (DOI : 10.3390/v9040078)
787 41. Bast J. Schaefer I, Schwander T, Maraun M, Scheu S, Kraaijeveld K. 2016. No
788 Accumulation of Transposable Elements in Asexual Arthropods. *Mol. Biol. Evol.* **33**, 697–
789 706. (DOI : 10.1093/molbev/msv261)
790 42. Rho M, Schaack S, Gao X, Kim S, Lynch M, Tang H. 2010. LTR retroelements in the
791 genome of *Daphnia pulex*. *BMC Genomics* **11**, 425 (DOI : 10.1186/1471-2164-11-425).
792 43. Kraaijeveld AR, Layen SJ, Futerman P H, Godfray HCJ. 2012. Lack of Phenotypic and
793 Evolutionary Cross-Resistance against Parasitoids and Pathogens in *Drosophila*
794 *melanogaster*. *PloS One* **7**, e53002. (DOI : 10.1371/journal.pone.0053002)
795 44. Pritham EJ. 2009. Transposable elements and factors influencing their success in
796 eukaryotes. *J. Hered.* **100**, 648–655 (DOI : 10.1093/jhered/esp065).
797 45. Blanc-Mathieu R, Krasovec M, Hebrard M, Yau S. 2017. Population genomics of
798 picophytoplankton unveils novel chromosome hypervariability. *Sci. Adv.* **3**, e1700239.
799 (DOI : 10.1126/sciadv.1700239)
800 46. Langley CH, Montgomery E, Hudson R, Kaplan N, Charlesworth B. 1988. On the role of
801 unequal exchange in the containment of transposable element copy number. *Genet. Res.* **52**,
802 223–235.
803 47. Bichsel M, Barbour, AD, Wagner A. 2010. The early phase of a bacterial insertion sequence
804 infection. *Theor. Popul. Biol.* **78**, 278–288. (DOI : 10.1016/j.tpb.2010.08.003)
805 48. Startek M, Le Rouzic A, Capy P, Grzebelus D, Gambin A. 2013. Genomic parasites or
806 symbionts? Modeling the effects of environmental pressure on transposition activity in
807 asexual populations. *Theor. Popul. Biol.* **90**, 145–151. (DOI : 10.1016/j.tpb.2013.07.004)

808 49. McGraw JE, Brookfield JFY. 2006. The interaction between mobile DNAs and their hosts
809 in a fluctuating environment. *J. Theor. Biol.* **243**, 13–23. (DOI : 10.1016/j.jtbi.2006.05.029)
810 50. Burt A, Trivers R. 1998. Genetic conflicts in genomic imprinting. *Proc. Biol. Sci.* **265**,
811 2393–2397. (DOI : 10.1098/rspb.1998.0589)
812 51. Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger T. 2009. Selective epigenetic
813 control of retrotransposition in *Arabidopsis*. *Nature* **461**, 427–430. (DOI :
814 10.1038/nature08328)
815 52. Badge RM, Brookfield JFY. 1997. The Role of Host Factors in the Population Dynamics
816 of Selfish Transposable Elements. *J. Theor. Biol.* **187**, 261–271.(DOI :
817 10.1006/jtbi.1997.0432)
818 53. Bonchev G, Parisod C. 2013. Transposable elements and microevolutionary changes in
819 natural populations. *Mol. Ecol. Resour.* **13**, 765–775. (DOI : 10.1111/1755-0998.12133)
820 54. Yang Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. *Trends*
821 *Ecol. Evol.* **11**, 367–372. (DOI : 10.1016/0169-5347(96)10041-0)
822 55. Fultz D, Choudury SG, Slotkin RK. 2015. Silencing of active transposable elements in
823 plants. *Curr. Opin. Plant Biol.* **27**, 67–76. (DOI : 10.1016/j.pbi.2015.05.027)
824 56. Slotkin, RK, Freeling M, Lisch D. 2005. Heritable transposon silencing initiated by a
825 naturally occurring transposon inverted duplication. *Nat. Genet.* **37**, 641–644. (DOI :
826 10.1038/ng1576)
827 57. Groenenboom MAC, Marée AFM, Hogeweg P. 2005. The RNA Silencing Pathway: The
828 Bits and Pieces That Matter. *PLoS Comput. Biol.* **1**. (DOI : 10.1371/journal.pcbi.0010021)
829 58. Rodrigo G, Carrera J, Jaramillo A, Santiago E. 2011. Optimal viral strategies for bypassing
830 RNA silencing. *J. R. Soc. Interface* **8**, 257–268. (DOI : 10.1098/rsif.2010.0264).

831 59. Roessler K, Bousios A, Meca E, Gaut BS. 2018. Modeling Interactions between
832 Transposable Elements and the Plant Epigenetic Response: A Surprising Reliance on
833 Element Retention. *Genome Biol. Evol.* **10**, 803–815. (DOI : 10.1093/gbe/evy043)
834 60. Jin H, Lei J. 2016. A hybrid model of molecular regulation and population dynamics for
835 yeast autophagy. *J. Theor. Biol.* **402**, 45–53. (DOI : 10.1016/j.jtbi.2016.04.019)
836 61. Bergstrom CT, McKittrick E, Antia R. 2003. Mathematical models of RNA silencing:
837 Unidirectional amplification limits accidental self-directed reactions. *Proc. Natl. Acad. Sci.*
838 **100**, 11511–11516. (DOI : 10.1073/pnas.1931639100)
839 62. Venner S, Miele V, Terzian C, Biémont C, Daubin V, Feschotte C, Pontier D. 2017
840 Ecological networks to unravel the routes to horizontal transposon transfers. *PLOS Biol.* **15**,
841 e2001536. (DOI : 10.1371/journal.pbio.2001536)
842 63. Carlier Y, Truyens C, Deloron P, Peyron F. 2012. Congenital parasitic infections: a review.
843 *Acta Trop.*121(2):55–70. (DOI : 10.1016/j.actatropica.2011.10.018)
844 64. Rosà R, Pugliese A. 2002. Aggregation, Stability, and Oscillations in Different Models for
845 Host-Macroparasite Interactions. *Theor. Popul. Biol.* **61**, 319–334. (DOI :
846 10.1006/tpbi.2002.1575)