

PAPER

Regulatory logic and transposable element dynamics in nematode worm genomes

Janna L. Fierst^{1,2*} and Victoria K. Eggers^{1,2}

¹Biomolecular Sciences Institute and Department of Biological Sciences, Florida International University, 11200 8th Street, 33199, Miami, FL, USA

*Corresponding author. jfierst@fiu.edu

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Genome sequencing has revealed a tremendous diversity of transposable elements (TEs) in eukaryotes but there is little understanding of the evolutionary processes responsible for TE diversity. Non-autonomous TEs have lost the machinery necessary for transposition and rely on closely related autonomous TEs for critical proteins. We studied two mathematical models of TE regulation, one assuming that both autonomous transposons and their non-autonomous relatives operate under the same regulatory logic, competing for transposition resources, and one assuming that autonomous TEs self-attenuate transposition while non-autonomous transposons continually increase, parasitizing their autonomous relatives. We implemented these models in stochastic simulations and studied how TE regulatory relationships influence transposons and populations. We found that only outcrossing populations evolving with Parasitic TE regulation resulted in stable maintenance of TEs. We tested our model predictions in *Caenorhabditis* genomes by annotating TEs in two focal families, autonomous LINEs and their non-autonomous SINE relatives and the DNA transposon *Mutator*. We found broad variation in autonomous - non-autonomous relationships and rapid mutational decay in the sequences that allow non-autonomous TEs to transpose. Together, our results suggest that individual TE families evolve according to disparate regulatory rules that are relevant in the early, acute stages of TE invasion.

Introduction

Transposable elements (TEs) have been described as selfish genomic parasites which are dependent on the eukaryotic cell for maintenance and replication (Doolittle and Sapienza, 1980; Orgel and Crick, 1980; Kidwell and Lisch, 2001), but that view has evolved with evidence that TEs can play critical roles in organismal development (Wang et al., 2022; Chang et al., 2024) and gene regulation (Chuong et al., 2016). The emerging picture is that the evolution of TEs can have complex detrimental and beneficial effects for an organism. TEs can rapidly replicate, mutate and insert throughout the genome (Craig et al., 2002), potentially increasing genome size (Pritham, 2009; Tenaillon et al., 2011; Lehmann et al., 2021). Replication can harm genes and essential functions through disruptive insertions (Finnegan, 1992) but also create pools of novel genetic variation and polymorphism (Fedoroff, 2012). TEs can influence selection and molecular evolution (Castillo et al., 2011) and modulate gene expression levels (Hatanaka et al., 2024). The evolutionary processes generating patterns of TE diversity are not well understood and discovering the factors driving TE dynamics is a central question in biology.

Some TEs are clearly molecular parasites. Autonomous transposons produce the machinery necessary for transposition, including intact recognition sequences, competent proteins and critical enzymes such as transposase, integrase and

transcriptase (Makalowski et al., 2019; Fig. 1). In contrast, non-autonomous transposons have lost their transposition machinery through mutation (Wicker et al., 2007). They rely on closely related autonomous copies for transposition (McClintock, 1950) potentially parasitizing both the genome itself and their transposon relatives (Hartl et al., 1992).

Autonomous and non-autonomous transposons do not share the same interests and likely do not obey the same regulatory logic (Hartl et al., 1992). Empirical evidence suggests the two types of TEs transpose and proliferate differently (Hartl et al., 1992; Robillard et al., 2016). TE insertions can damage critical genes and disrupt genome integrity to the point of harming or killing the host (Finnegan, 1992; Charlesworth et al., 1994). The replication of autonomous transposons is limited by their ability to produce proteins and enzymes such as the transposase necessary for successful insertion. For these reasons, autonomous transposons may have a strong interest in attenuating replication and mutational insertions (Doolittle et al., 1984).

Non-autonomous transposons compete with their autonomous relatives for critical enzymes. Capturing transposition machinery for *trans* regulation (McClintock, 1950) puts them at a substantial disadvantage for the higher affinity resources produced in *cis* by autonomous transposons (Hartl et al., 1992; Fig. 1). For non-autonomous transposons, attenuating activity

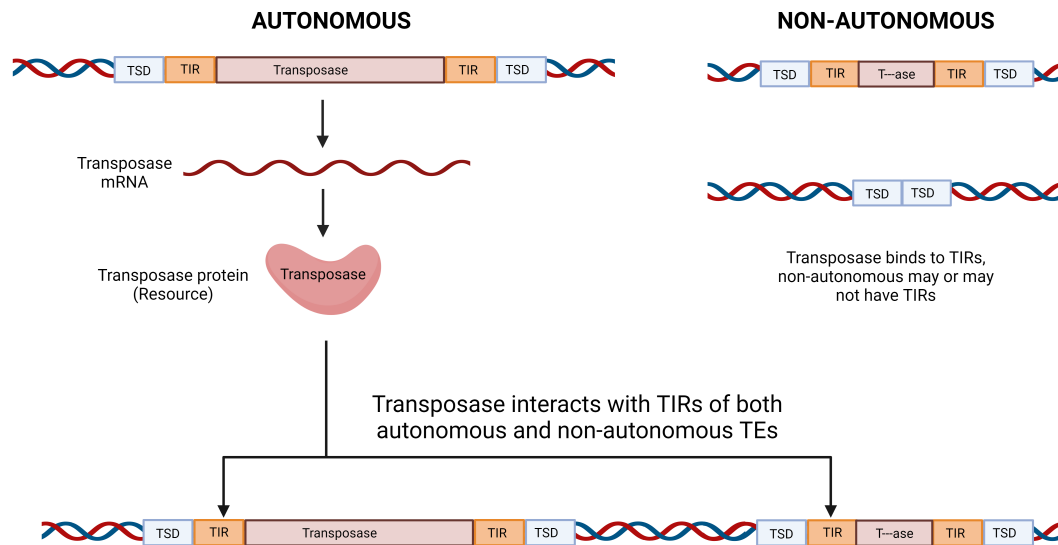


Fig. 1. Autonomous TEs are characterized by competent recognition sequences including target site duplications (TSDs) and terminal inverted repeats (TIRs). They also encode full-length enzymes, here transposase. The transposase protein must be shared between autonomous and non-autonomous relatives and becomes a limiting resource in transposition. Non-autonomous TEs have independent and combined deletions of the sequences necessary to produce competent enzymes and TSD and TIR recognition sequences.

would potentially result in losing the ability to transpose all together.

Here, we ask how transposon regulatory relationships influence evolutionary trajectories by studying two deterministic mathematical models, implementing our models in stochastic simulations, and testing our predictions in nematode genome sequences. Previous models of TE evolution have assumed that autonomous and non-autonomous transposons follow the same regulatory rules (Brookfield, 1996; Le Rouzic and Capy, 2006). In this study we developed two models of TE regulatory logic, one where autonomous elements self-attenuate while non-autonomous elements increase transposition, parasitizing their autonomous relatives (Fig. 2A) and one where both autonomous and non-autonomous elements continually increase transposition rates and compete (Fig. 2B).

TEs are subject to stochastic factors in real populations, some resulting in loss due to drift. To study these processes we implemented our deterministic equations in stochastic, individual-based population genetic simulations and studied the evolution of populations with differing modes of reproduction. Self-fertility and outcrossing influence population genetic factors including the effective population size, recombination rate, genetic variation and ultimately the probability a TE is lost or establishes in the population. Previous models have predicted extreme effects of sexual and asexual reproduction ranging from extinction of the TE in the genome in asexual populations (Hickey, 1982) to population extinction in both sexual (Charlesworth and Langley, 1986) and asexual species (Arkhipova and Meselson, 2005). Critically, none of these early theoretical treatments explain the patterns of TE proliferation and diversity we see in genomes today.

Our deterministic models predicted stable coexistence of both TE types but under population genetic constraints and stochastic conditions only large sexually outcrossing

populations evolving with Parasitic regulatory logic maintained stable numbers of autonomous and non-autonomous TEs. The stochastic model of Competitive TE regulatory logic predicted autonomous TEs outnumber their non-autonomous counterparts while the stochastic model of Parasitic TE regulatory logic predicted the opposite pattern with non-autonomous TEs outnumbering autonomous. We tested these predictions by examining two TE superfamilies containing both autonomous - non-autonomous members, the retrotransposons LINES/SINES and the DNA transposon *Mutator*, annotated in assembled *Caenorhabditis* genome sequences. We found that autonomous LINES far outnumber non-autonomous SINES, supporting a model of Competitive regulation. In contrast, non-autonomous *Mutator* TEs far outnumbered autonomous elements, supporting a model of Parasitic regulation.

Non-autonomous *Mutator* TEs require competent terminal inverted repeat (TIR) sequences for transposase binding. We found that in *Caenorhabditis* these sequences rapidly mutate beyond recognition, suggesting non-autonomous TEs are quickly rendered transposition-incompetent. Our theoretical results demonstrate that regulatory relationships can create intrinsically stable or ephemeral systems of TEs, and our comparative genomic results suggest broad variation in the regulatory logic underlying the evolution of individual TE families. Rapid mutational decay limits the influence of regulatory logic to early stages of invasion, with inactivated sequences lingering in the genome as relics of past TE dynamics.

Methods

The deterministic model

We developed two deterministic models of transposon dynamics, building on previous analytical work (Charlesworth

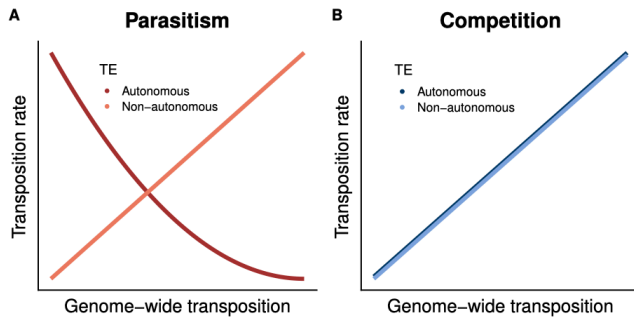


Fig. 2. Under A) a parasitic model of regulatory logic autonomous TEs self-attenuate transposition while their non-autonomous relatives continue to increase transposition rates as genome-wide transposition increases. Under B) a competitive model of regulatory logic both autonomous and non-autonomous transposons increase transposition rates as genome-wide transposition increases. Both autonomous and non-autonomous TEs share the same function but are plotted here with a small displacement for visualization.

and Charlesworth, 1983; Dolgin and Charlesworth, 2006, 2008; Le Rouzic and Cappy, 2006). Both models envision two types of TEs, one fully autonomous and capable of producing all machinery necessary for transposition and one non-autonomous, dependent on its autonomous relative for the sequences and enzymes necessary for transposition. Autonomous element i and non-autonomous element j , respectively, have transposition activities a_i and a_j , roughly the probability of transposition per element. If n_i is the number of autonomous elements and n_j is the number of non-autonomous elements, then n is the total of all TEs in an individual. Each TE carries a small additive fitness cost s and the mean fitness w_n of an individual with n TEs is a linear function, $w_n = 1 + s * n$.

TEs insert into the genome according to u_i , the transposition ability of TE i and excise out of the genome according to ν_i . Here, autonomous and non-autonomous transposition abilities are equal and $u_i = u_j$ while $\nu_i = \nu_j$. The mean transposition activity for the genome, \bar{a} , encompasses the transposition of all autonomous and non-autonomous TEs. Under a parasitic model of TE regulation, the autonomous TE i attenuates transposition with increasing genome-wide transposition activity and its transposition rate is given by $u_i * (1 - \bar{a})^2$ (Fig. 2B). In contrast, the non-autonomous TE j does not attenuate as transposition activity increases, transposing instead according to $u_j * \bar{a}$. Under a competitive model of TE regulation both TEs transpose according to $u_i * \bar{a}$, continually amplifying competition for physical transposition resources (Fig. 2A).

Mean transposition is calculated as

$$\bar{a} = \frac{(n_i * a_i + n_j * a_j)}{n} \quad (1)$$

and the rate of change in the population abundance of transposons is determined by

$$S_{\bar{n}} = \frac{\delta(\log w_{\bar{n}})}{\delta(\bar{n})}. \quad (2)$$

At low copy number TE frequencies can be approximated by a Poisson distribution (Le Rouzic and Cappy, 2006) where $P_{n_i=0} = \exp^{-n_i}$ and $P_{n_j=0} = \exp^{-n_j}$. The change in TE copy number

between generations in the parasitic model of regulation is then calculated according to

$$\Delta(n_i) = n_i * (S_{\bar{n}} + (1 - ((1 - P_{n_j=0}) * \bar{a} + P_{n_j=0} * a_i))^2 * u_i - \nu) - \bar{n}_i * m_{ij} + \bar{n}_j * m_{ji} \quad (3)$$

for autonomous TEs and

$$\Delta(n_j) = n_j * (S_{\bar{n}} + ((1 - P_{n_i=0}) * \bar{a} + P_{n_i=0} * a_j) * u_j - \nu) - \bar{n}_j * m_{ji} + \bar{n}_i * m_{ij} \quad (4)$$

for non-autonomous TEs. The terms m_{ij} and m_{ji} represent the potential for elements to mutate from autonomous to non-autonomous and vice versa.

The competitive model of TE regulation evolves according to:

$$\Delta(n_i) = n_i * (S_{\bar{n}} + ((1 - P_{n_j=0}) * \bar{a} + P_{n_j=0} * a_i) * u_i - \nu) - \bar{n}_i * m_{ij} + \bar{n}_j * m_{ji} \quad (5)$$

and

$$\Delta(n_j) = n_j * (S_{\bar{n}} + ((1 - P_{n_i=0}) * \bar{a} + P_{n_i=0} * a_j) * u_j - \nu) - \bar{n}_j * m_{ji} + \bar{n}_i * m_{ij}. \quad (6)$$

The stochastic model

TEs are subject to multiple interacting stochastic factors that are difficult to model with deterministic approaches. We implemented our deterministic mathematical models of parasitic and competitive regulation in forward time simulations using SLiM 4.0.1 (Haller and Messer, 2019, 2023) to study the influence of population genetic factors on TE evolution.

Caenorhabditis nematodes have two primary mating systems. The ancestral reproductive mode is dioecy with males and females. Three species (*C. elegans*, *C. briggsae* and *C. tropicalis*) independently evolved an androdioecious reproductive mode where hermaphrodites produce young by either self-fertilizing or outcrossing with males (Kiontke and Fitch, 2005; Kiontke et al., 2004). Male frequency and outcrossing vary quantitatively from approximately 1 in every 1000 individuals as male in laboratory populations of the *C. elegans* strain N2 to 1 in every 3 individuals as male in populations of the related *C. elegans* strain AB1 (Anderson et al., 2010). Following these features of reproductive mode, our simulated dioecious populations contained males and females. Simulated self-fertile populations had a proportion of individuals reproducing through selfing and a proportion outcrossing. In order to simulate androdioecous populations we constructed 'non-Wright Fisher' populations in SLiM 4.0.1 (Haller and Messer, 2019, 2023) with overlapping generations and variable sizes.

We evolved populations under a number of different evolutionary scenarios including small and large population size (ranging between $N = 100$ and $N = 1000$, respectively), high and low recombination rates (ranging between 1×10^{-2} and 1×10^{-5}) and different levels of self-fertility ranging from

full dioecy with 0% of the population reproducing through self-fertility to 100% of the population reproducing through self-fertility. Each individual had a single diploid chromosome $1x10^5$. Each population was initialized with 4 TEs (2 on each chromosome) with selective effect $s = 0.001$. TEs inserted at a rate $\mu = 0.01$ and excised at $\nu = 0.001$. The activity a_i of the autonomous TE was 1 and the activity a_j of the non-autonomous TE was 0. Simulated populations were evolved for 500,000 generations.

Annotating TEs in genome sequences

We focused on two superfamilies of TEs, one the autonomous retrotransposon Long Interspersed Nuclear Elements (LINEs) and their non-autonomous Short Interspersed Nuclear Elements (SINEs) relatives (Weiner, 2002) and the other autonomous and non-autonomous members of the DNA transposon *Mutator* (Dupeyron et al., 2019; Lisch, 2015; Robertson, 1978, 1986). Although some LINEs have been rendered inactive by mutations (Grimaldi et al., 1984; Ostertag and H H Kazazian, 2001; Beck et al., 2011), members of the family are classified as autonomous and many retain the ability to transpose autonomously. In contrast, SINEs are severely mutated and unable to produce their own transposition machinery. LINEs and SINEs are Class I TEs and require an RNA intermediate to transpose.

In contrast, members of the DNA transposon *Mutator* family are Class II TEs and able to transpose without a molecular intermediate. *Mutator* is a broadly dispersed TE named for its mutagenic potential (Robertson, 1978, 1986). Previous investigations into *Caenorhabditis* genomes revealed large numbers of *Mutator* TEs that are approximately twice as abundant in the genomes of outcrossing species (Adams et al., 2023).

Existing genome assemblies and EDTA annotations were obtained from Adams et al. (2023). Briefly, Extensive *de-novo* TE Annotator (EDTA) v2.0.0 (Ou et al., 2019) annotates TEs by employing several annotation softwares and filtering steps. Long terminal retrotransposons were found by EDTA with LTRharvest (Ellinghaus et al., 2008) and LTR_Finder (Ou and Jiang, 2019) before filtering with LTR_retriever (Ou and Jiang, 2018). Generic Repeat Finder (Shi and Liang, 2019) identified terminal direct repeats (TDRs), miniature inverted transposable elements (MITEs), and terminal inverted repeats (TIRs). TIRs were also found with TIR-Learner (Su et al., 2019). *Helitrons* were found with HelitronScanner (Xiong et al., 2014). Both TIRs and *Helitrons* were investigated for simple sequence repeats using *MUSTv2* (Ge et al., 2017). Identified TEs were used to mask the genome, and RepeatModeler (Smit et al., 2013-2015) then scanned the unmasked portion to find any remaining TEs. False discovery rate was reduced by supplying EDTA with curated Rhabditida repeat libraries and coding sequences (CDS) to filter the TE library. Curated repeats were obtained from RepeatMasker (Smit et al., 2013-2015). CDS files were generated with BRAKER2 (Bruna et al., 2021) or downloaded from either WormBase Parasite (Howe et al., 2017) or NCBI (National Center for Biotechnology Information, 2002). Finally, EDTA classifies TEs with TESorter (Zhang et al., 2019).

We assigned TEs as autonomous and non-autonomous through further analysis of EDTA output files (Ou et al., 2019). For each *Caenorhabditis* genome, EDTA files ending in `*mod.EDTA.TEanno.gff3` were filtered for a specific TE classification, such as *Mutator*. The first, fourth and fifth

columns were cut to make a bed file. We used bedtools getfasta (Quinlan and Hall, 2010) to obtain the TE sequences for each of these specific classification files. Potential genes within individual TE sequences were then identified by searching for open reading frames (ORFs) with a minimum size of 300 base pairs using the function getorf from EMBOSS v.6.6.0 (Rice et al., 2000). These ORFs were supplied to pfam_scan.pl v.1.6 to scan PFAM database v.36.0 for functional annotation (Punta et al., 2012). A web scraping python script was created to search for the word “Transposase” on the webpage descriptions of each unique pfamID found in the TE sequences. For LINEs and SINEs “Transposase” was changed to “Integrase” or “Endonuclease,” for example. TE sequences of appropriate length (>500 nucleotides for *Mutator* and >2000 nucleotides for LINEs) with pfamID descriptions containing these critical proteins were labeled as autonomous while others were labeled as non-autonomous.

For *Mutator*, the DNA binding site for transposase is likely the Terminal Inverted Repeat (TIR) sequences. Thus, parasitism or competition would require non-autonomous elements with TIRs similar enough to be recognized by transposase proteins of autonomous TEs. This was inferred through sequence similarity of TIRs between autonomous and non-autonomous TEs. TIRs were extracted from the last column of EDTA files ending in `*mod.EDTA.intact.gff3`. Levenshtein distance algorithm was used to compute sequence similarity between TIRs and those with a similarity of at least 80% were classified as active non-autonomous TEs.

Results

The deterministic model

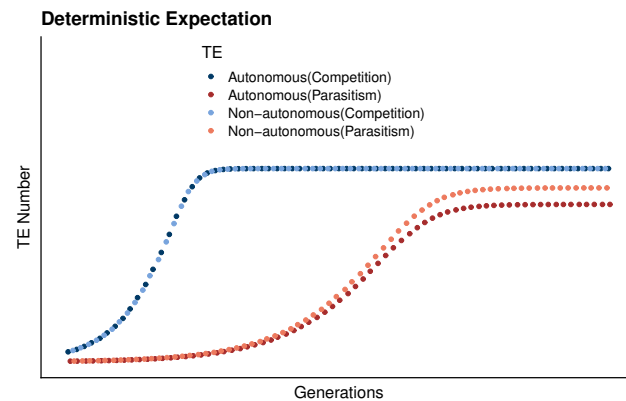


Fig. 3. When autonomous and non-autonomous transposons operate according to (A) the same competitive regulatory logic the deterministic model predicts a small number of transposons will increase until they stabilize in the population. The two TEs have identical numerical trajectories and are plotted on top of one another. (B) When the non-autonomous transposon parasitizes their autonomous relative, the transposons stabilize in the population but it takes more than twice as many generations. At a stable equilibrium the non-autonomous transposon is more abundant than the autonomous transposon. Here, $u_i = 0.01$, $\nu_i = 0.001$ and the selective effect of each TE is $s = 0.001$.

When autonomous and non-autonomous TEs have identical rates of transposition u_i and excision ν_i , the deterministic model with competitive regulation predicts both elements will achieve the same stable copy number in the populations within

1500 generations (Fig. 3). With parasitic regulatory dynamics the two elements achieve stable copy number in the population within 3500 generations but non-autonomous TEs are more abundant than their autonomous relatives.

The stochastic model

We found that the simulated populations followed 4 discrete trajectories according to the population genetic factors they evolved under (Fig. 4). We divided these trajectories into 4 categories to qualitatively describe the impact of population size, recombination rate, reproductive mode and regulatory logic on TE evolution and maintenance. Populations that followed an ‘Early TE Loss’ trajectory would rapidly lose both autonomous and non-autonomous TEs, typically in fewer than 20,000 simulated generations. In other populations both autonomous and non-autonomous TEs were maintained for 20,000-450,000 generations but eventually excised entirely from the population. We labeled populations following this trajectory ‘Eventual TE Loss.’

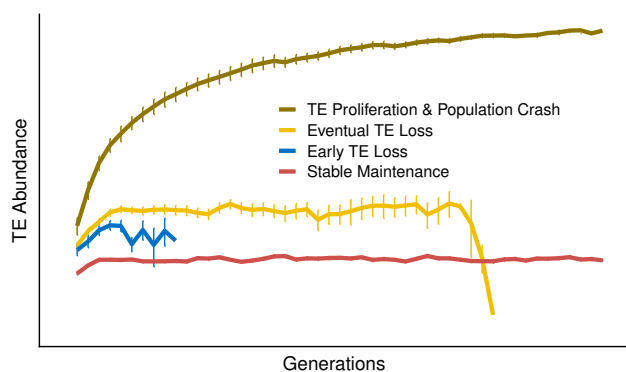


Fig. 4. Across the range of population genetic parameters the simulated populations evolved according to 4 trajectories: 1) Early TE Loss; 2) Eventual TE Loss; 3) TE proliferation, fitness loss and eventual population extinction; or 4) Stable Maintenance

Some populations had TEs that continually proliferated and increased in abundance. These populations lost fitness to the point of eventual population extinction. Importantly, some of the simulated populations had unchecked proliferation of both autonomous and non-autonomous TEs while in other populations the non-autonomous TE was lost early on and the autonomous TE subsequently proliferated to the point of population extinction. We labeled both of these scenarios ‘TE Proliferation & Population Crash.’ Finally, some populations maintained both autonomous and non-autonomous TEs over the entire course of the simulated population evolution and we labeled this trajectory ‘Stable Maintenance.’

We found that three population genetic parameters (population size, reproductive mode and regulatory logic) determined the trajectory of TEs in the simulated populations (Table 1). Small populations frequently had trajectories of Early TE Loss across both reproductive modes and models of TE regulation (Suppl. Fig. 1). Those populations that maintained TEs followed a trajectory of TE Proliferation & Population Crash with TE insertion increasing to the point at which fitness loss was severe and population-wide. Under a Competitive model of TE regulation some simulated

Table 1. TE trajectories are determined by population size, reproductive mode and TE regulation. The trajectories are:

- 1 Early TE Loss
- 2 Eventual TE Loss
- 3 TE Proliferation & Population Crash
- 4 Stable Maintenance

	Competition		Parasitism	
	Selfing	Outcrossing	Selfing	Outcrossing
Population Small	1, 3	1, 3	1, 3	1, 3
Population Large	1, 2, 3	1, 2	1, 3	4

populations decoupled the non-autonomous and autonomous TEs, losing the non-autonomous entirely but maintaining the autonomous TE (Suppl. Fig. 1,2). These populations eventually followed the same TE Proliferation & Population Crash trajectory.

TEs were retained for long periods in some large dioecious populations evolving with Competitive TE regulation (Fig. 5A, B) but gradually lost across the population replicates, following a trajectory of Eventual TE Loss (Fig. 4). TEs were stably maintained in large dioecious populations evolving with Parasitic TE regulation (Fig. 5A, B). Importantly, this was the only set of simulated population parameters that resulted in a trajectory of Stable Maintenance (Table 1; Fig. 3).

Large self-fertile populations evolving with either Competitive or Parasitic TE regulation followed trajectories of Early TE Loss or TE Proliferation % Population Crash (Table 1; Fig. 5,6). Some large self-fertile populations evolving with Competitive TE regulation decoupled the two TE types, excising the non-autonomous TE entirely and maintaining the autonomous TE, but eventually these populations followed the TE Proliferation % Population Crash trajectory (Fig. 6A-D). At low levels of outcrossing, for example 5% of the population outcrossing and 95% selfing, populations evolving under Competitive or Parasitic TE regulation followed the trajectories of fully self-fertile populations (Fig. 6). At 50% outcrossing populations evolving with either Competitive or Parasitic TE regulation followed the trajectories of fully outcrossing populations (Suppl. Fig. 3).

We quantitatively varied population size and found that population of 500 individuals and smaller followed trajectories similar to population sizes of 100 individuals (Suppl. Fig. 1,4). Population sizes of 750 individuals and larger followed trajectories similar to population sizes of 1000 individuals (Suppl. Fig. 5). Importantly, populations evolving under Competitive TE regulation did not achieve the Stable Maintenance trajectory at large population sizes of 1500 individuals. Similarly, we found that varying recombination rate altered the speed with which the simulated populations achieved different trajectories but did not qualitatively change the trajectories determined by population size, reproductive mode and TE regulation (Suppl. Fig. 2). In population genetic models of TEs the rate of recombination heavily influences linkage disequilibrium among suites of TEs (Roze, 2023) depending on the shape of the fitness function (Charlesworth, 1990).

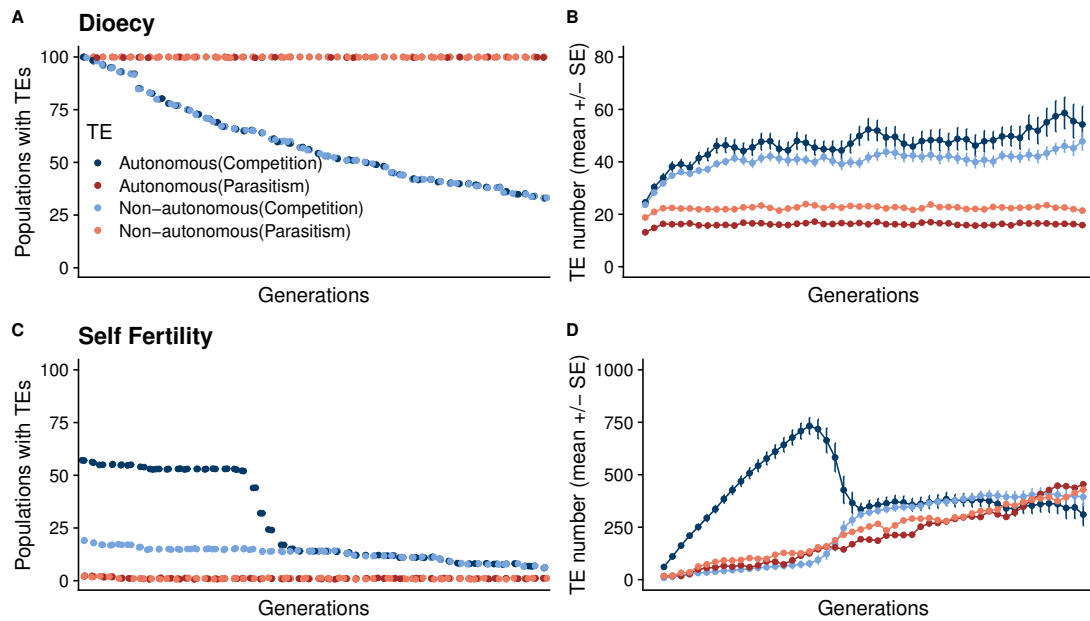


Fig. 5. When the population is large (here, $N=1000$ individuals) and dioecious, populations evolving with (A-B) competitive TE regulation gradually lose TEs. In contrast, populations evolving with (C-D) parasitic TE regulation stably maintain a low number of TEs.

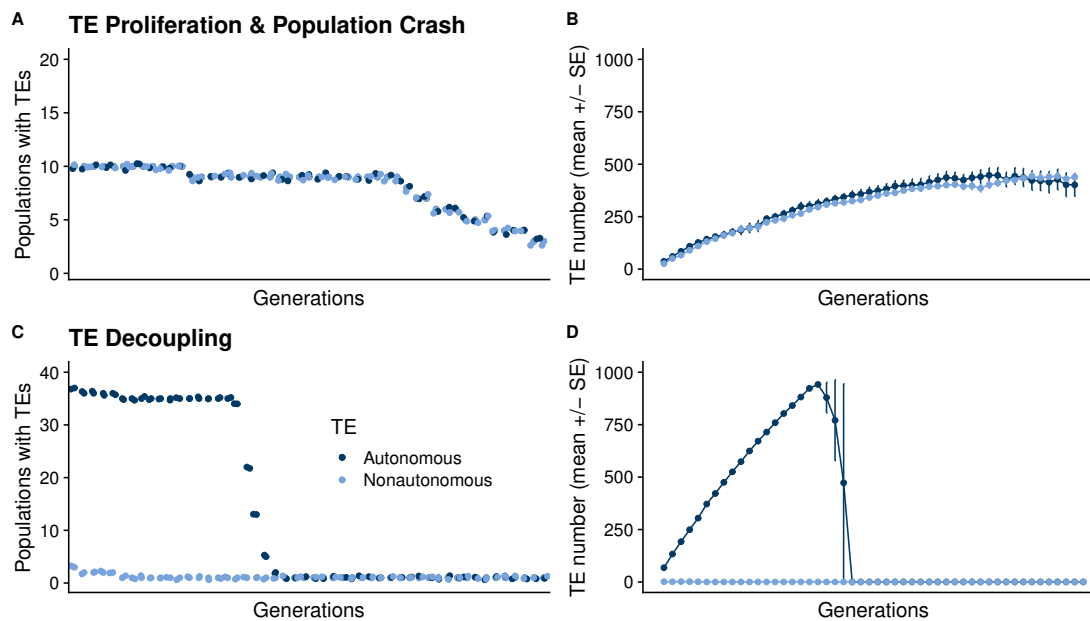


Fig. 6. Large self-fertile populations evolving with competitive TE regulation evolved according to different trajectories. (A-B) Ten of the 100 populations retained both TEs for a period of time, eventually losing both autonomous and non-autonomous TEs. (C-D) Thirty-eight of the populations retained their autonomous TE but lost their non-autonomous TE early in the evolution of the simulated populations. Those populations experienced TE proliferation, fitness loss and eventual population crashes.

Testing predictions from the model

Importantly, the deterministic and stochastic models predict that autonomous and non-autonomous TEs should have differing dynamics in empirical populations. Populations evolving with Competitive TE regulation should have higher abundances of autonomous TEs while populations evolving with Parasitic TE regulation should have higher abundances of non-autonomous TEs. To test these predictions and evaluate

the regulatory logic operating in eukaryotic genomes we studied annotated autonomous and non-autonomous TEs in assembled *Caenorhabditis* genome sequences.

We found that LINE elements were more abundant than SINE elements across both androdioecious and dioecious *Caenorhabditis* species (Fig. 7). The number of annotated LINE elements ranged from 149-2258 in dioecious species and 599-1379 in androdioecious species. In comparison, SINE

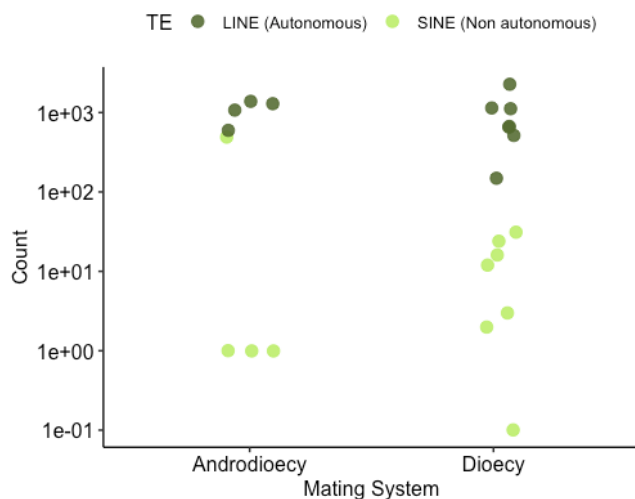


Fig. 7. Autonomous LINE elements are more abundant than mutated non-autonomous SINE elements in both androdioecious and dioecious *Caenorhabditis* genomes.

elements ranged from 0-31 in dioecious species and 1-487 in androdioecious species.

Table 2. Numerical counts of *Mutator*-associated elements in *Caenorhabditis* genome sequences demonstrate uniformly high numbers of *Mutator* elements across the group and uniformly low numbers of intact transposase genes, indicating that virtually all of the annotated *Mutator* elements are either non-autonomous, partial fragments or both. The first column is *Caenorhabditis* species with strain name for species with multiple entries. The second column is total number of annotated *Mutator* elements, the third is MITEs that can be reliably associated to *Mutator* TEs through mutational patterns and the fourth is the number of intact transposase genes associated with the total set of *Mutator* elements. Self-fertile species are the first four rows in the table with species names in bold font.

Species	<i>Mutator</i>	<i>Mutator</i> MITEs	Intact transposase
<i>C. briggsae</i>	18,093	1,854	18
<i>C. elegans</i>	8,796	770	2
<i>C. tropicalis JU1373</i>	12,229	1,917	1
<i>C. tropicalis NIC158</i>	12,163	1,132	1
<i>C. inopinata</i>	53,693	10,138	30
<i>C. latens</i>	28,950	3,414	10
<i>C. nigoni</i>	25,912	3,626	10
<i>C. remanei PX356</i>	35,330	5,695	3
<i>C. remanei PX439</i>	34,880	4,028	13
<i>C. remanei PX506</i>	36,880	4,217	5
<i>C. sinica</i>	26,591	8,355	6

Mutator elements ranged from 8,796 to 36,880 in *Caenorhabditis* (Table 2). The number of non-autonomous MITEs of *Mutator* origin was roughly an order of magnitude lower in number and ranged from 592-3,626 elements. A fully autonomous *Mutator* element requires an intact transposase

sequence. Among our *Mutator* TE set we identified very few intact transposase genes, from a low of 2 in *C. elegans* to a high of 18 in *C. briggsae*. Virtually all of the *Mutator* elements we annotated in *Caenorhabditis* genome sequences were partial fragments, non-autonomous and highly mutated (an example is shown for *C. elegans* in Fig. 8). Both *Mutator* elements and MITEs of *Mutator* origin had a high number of fragments 0-300bp in length. A *Mutator* sequence capable of transposition requires a pair of ~215bp terminal inverted repeats (TIRs) and a ~500bp transposase gene making a competent TE >1000bp in length. The abbreviated lengths and few intact transposase genes indicate the vast majority of *Mutator* elements are non-autonomous.

Table 3. Numerical counts of autonomous *Mutator* elements, non-autonomous *Mutator* elements with intact TIR sequences and non-autonomous *Mutator* elements with TIR sequences with 80% identity show that very few non-autonomous *Mutators* have the ability to transpose. The first column is *Caenorhabditis* species and strain name (self-fertile species are bolded), the second column is total number of intact autonomous *Mutator* elements, the third is non-autonomous *Mutators* with intact, competent TIR sequences and the fourth is non-autonomous *Mutator* elements with TIR sequences with 80% similarity to competent *Mutator* TIRs.

Species	Autonomous <i>Mutators</i>	Non-Aut. Exact TIRs	Non-Aut. 80% TIRs
<i>C. briggsae</i>	8	0	19
<i>C. elegans</i>	2	0	0
<i>C. tropicalis JU1373</i>	1	13	13
<i>C. tropicalis NIC158</i>	1	0	0
<i>C. inopinata</i>	21	49	556
<i>C. latens</i>	7	0	4
<i>C. nigoni</i>	6	0	6
<i>C. remanei PX356</i>	3	0	0
<i>C. remanei PX439</i>	6	0	28
<i>C. remanei PX506</i>	5	3	50
<i>C. sinica</i>	3	0	2

We further investigated autonomous and non-autonomous *Mutator* TEs by identifying those with competent TIRs (Table 3). Despite the high numbers of fragmented *Mutator* sequences in *Caenorhabditis* genomes (Table 2), very few have TIRs with sufficient sequence identity to competently transpose. *C. inopinata*, a close relative of the model *C. elegans* has an active transposable element landscape potentially shaped by disruptive TE insertions into genes critical for the 26G small interfering (si)RNA gene silencing pathway (Kanzaki et al., 2018). The landscape of autonomous and non-autonomous *Mutator* TEs in *C. inopinata* shows that 95% of the non-autonomous *Mutator* TEs originate from just 5 of the 21 autonomous *Mutator* TEs. Phylogenetic reconstruction identified three groups of autonomous sequences (Supp. Fig. 6), potentially indicating three invasion events. One group contained a single autonomous sequence, one contained two sequences and the third group contained the remaining 18 sequences. Analysis of non-autonomous sequences with potentially competent TIRs showed that >98% of these sequences grouped with this large third

group of autonomous sequences in phylogenetic reconstruction (Table 2; Supp. Fig. 7). Together, these patterns suggest most transposed *Mutator* sequences are rapidly mutated and quickly rendered incapable of transposition.

Table 4. Numerical counts of autonomous *Mutator* elements, non-autonomous *Mutator* elements with intact TIR sequences and non-autonomous *Mutator* elements with TIR sequences with 80% identity in *C. inopinata* show that few autonomous *Mutators* produced 94.4% of the non-autonomous elements with recognizable TIRs. The first column is *C. inopinata* autonomous *Mutator* sequence, the second column is non-autonomous *Mutators* with intact, competent TIR sequences and the fourth is non-autonomous *Mutator* elements with TIR sequences with 80% similarity to competent *Mutator* TIRs.

Autonomous Sequence	Non-Aut. Exact TIRs	Non-Aut. 80% TIRs
<i>Sp34_Chr1_3565864_3569901</i>	0	0
<i>Sp34_Chr1_8382683_8383940</i>	0	2
<i>Sp34_Chr2_13584427_13589263</i>	0	0
<i>Sp34_Chr2_17312516_17315633</i>	0	10
<i>Sp34_Chr2_2153458_2156331</i>	1	122
<i>Sp34_Chr3_10008018_10011761</i>	0	0
<i>Sp34_Chr3_12627069_12631530</i>	0	0
<i>Sp34_Chr3_19360477_19362000</i>	0	0
<i>Sp34_Chr3_4378947_4382359</i>	0	0
<i>Sp34_Chr3_6586146_6588574</i>	45	71
<i>Sp34_Chr3_7591352_7592611</i>	0	2
<i>Sp34_Chr3_9754782_9755651</i>	0	0
<i>Sp34_Chr4_12041210_12045282</i>	0	13
<i>Sp34_Chr4_12152380_12154136</i>	1	2
<i>Sp34_Chr4_12855799_12856745</i>	0	0
<i>Sp34_Chr4_1565475_1567272</i>	0	0
<i>Sp34_Chr4_20191332_20193979</i>	0	158
<i>Sp34_Chr4_3934972_3939459</i>	0	105
<i>Sp34_Chr5_23209925_23213952</i>	2	69
<i>Sp34_Chr5_75219_76476</i>	0	0
<i>Sp34_Chr5_144542_146148</i>	0	2

Discussion

The question of how and under what conditions transposable elements might self-regulate has a developed history in evolutionary biology (Charlesworth and Charlesworth, 1983; Charlesworth and Langley, 1986; Langley et al., 1983). Previous theoretical studies have shown stable persistence of transposable elements under narrow sets of conditions and created an apparent paradox with the high transposon content found in most eukaryotic genomes (Kelleher et al., 2020). Here, we have presented deterministic and stochastic models that demonstrate regulatory logic can result in long-term stable

Caenorhabditis elegans

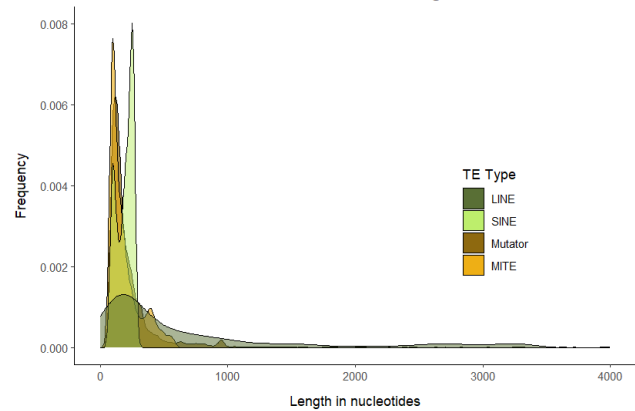


Fig. 8. The majority of LINE, SINE, *Mutator* and *Mutator*-origin MITE sequences are truncated fragments of less than 500bp in length.

persistence of transposable elements. We used our model results to structure investigations of autonomous and non-autonomous elements in *Caenorhabditis* genome sequences and found support for both competitive and parasitic regulatory interactions. Together, our theoretical and comparative genomic results suggest broad variation in regulatory logic across TE families.

Theoretical studies of intrinsic transposable element regulation

Early theoretical studies of transposable elements were fascinated with ideas of self-attenuation and self-regulation (Charlesworth and Charlesworth, 1983; Charlesworth and Langley, 1986; Langley et al., 1983). Without some mechanism of containment, transposable elements could conceivably proliferate until they took over entire genomes or decreased fitness to the point of population extinction. Mapping out the evolutionary processes that could support the broad diversity of TEs found in empirical populations was an important theoretical problem (Kelleher et al., 2020).

Despite extensive theoretical study, the conditions under which transposable elements were predicted to stably persist in populations were fairly narrow. Charlesworth and Charlesworth (1983) found that TEs could be stably maintained with either synergistic interactions for fitness, where the decrease in fitness accumulates faster than linearly with increasing numbers of transposable elements, self-regulation of transposition or both. Self-regulation of transposition was only achieved with unrealistic modeling assumptions including the total absence of selection (Charlesworth and Charlesworth, 1983), free recombination and no linkage (Langley et al., 1983), or highly restrictive rates of recombination (Charlesworth and Langley, 1986). Similarly, empirical studies have found that synergistic epistasis occurs (Lee, 2022) but is not ubiquitous (Maisnier-Patin et al., 2005; de Visser et al., 2011).

More recent theoretical treatments increased the space under which transposable elements may self-regulate by modeling autonomous and non-autonomous elements (Le Rouzic and Capy, 2006; Omole and Czappon, 2024), similar to the models we have developed here. However, these studies still predicted short-term persistence of TEs in the population or required evolution in the absence of selection and fitness effects.

Despite broad empirical patterns of transposable element stability, the predicted space for self-attenuating transposition has remained quite small over 40 years of study. This has led some authors to propose the ‘case closed’ on the potential for transposable elements to engage in any meaningful self-regulation (Kelleher et al., 2020).

A major contribution of our results is to demonstrate the theoretical potential for long-term, stable maintenance of transposable elements. Our simulated populations evolving with parasitic regulatory interactions between autonomous and non-autonomous transposable elements showed stable persistence of TEs across a broad range of population genetic factors including varying recombination, mutation and outcrossing rates, and population sizes. Our results suggest that regulatory logic may play an important, determining role in the evolution and stability of transposable elements in eukaryotic genomes.

The maintenance of transposable elements in selfing and asexual populations

Theoretical studies have uniformly predicted little space for transposable element maintenance in asexual or selfing populations (Hickey, 1982; Charlesworth and Charlesworth, 1983; Boutin et al., 2012). Our study largely agrees with these findings and in our models transposable elements were only maintained in partially selfing populations where at least 50% of the individuals reproduced through outcrossing. However, these theoretical results are at odds with empirical studies documenting abundant transposable elements in the genomes of selfing and asexual organisms (Nowell et al., 2021; Fierst et al., 2015; Adams et al., 2023; Bonchev and Willi, 2018; Mattila et al., 2020). How can these two phenomenon-be reconciled?

Importantly, theoretical studies and population models predicting an erosion of transposable elements in asexual and selfing organisms still envision the loss of transposable elements as a process occurring over time. Empirical studies have found mixed evidence for this in experimental evolution studies (Bast et al., 2019; Chen and Zhang, 2021) but found the process dependent on the dynamics of the population, elements and genomic structure (Henault et al., 2024). At the same time, organisms are constantly subject to TE invasions (Panaud, 2016; Peccoud et al., 2017; Guan et al., 2022; Paulat et al., 2023). Transposable element families in asexual and selfing populations may have fundamentally different dynamics from those in sexual populations, with rapid rates of turnover resulting from continual cycles of invasion and loss. An interesting avenue of future research would be phylogenetic comparative analysis of TEs across related sexual and asexual or selfing species, searching for ephemeral dynamics or differing rates across evolutionary time.

The evolution of extrinsic transposable element regulation

We focused on the potential for internal or intrinsic regulatory interactions between and within transposable elements to stabilize TEs in populations. In recent years external or extrinsic TE regulation has been an exciting field of study (Brennecke et al., 2007). Small RNAs, including microRNAs (Oliver et al., 2022), small interfering (si-)RNAs (Wei et al., 2014) and PIWI-interacting or piRNAs (Ozata et al., 2018), target TE sequences for degradation and are an important mechanism of extrinsic regulation. In our study populations

could go to extinction due to TE proliferation. This is likely not a reasonable scenario and in natural populations continually heavy TE burdens may facilitate the evolution of genomic resistance mechanisms, including extrinsic regulation capable of slowing proliferation.

Small RNA regulation systems play vital roles in many organisms but are not a universal mechanism of TE containment and regulation. The systems are absent in some basal eukaryotes (Ryan et al., 2013; Maxwell et al., 2012) and evolutionary reconstruction suggests parallel loss in others (Mondal et al., 2018; Fontenla et al., 2021; Sarkies, 2024; Sarkies et al., 2015). An interesting question for future study is if these organisms have fundamentally different patterns of genomic TEs from those with developed small RNA systems targeting and regulating TEs. Developing genomic data could facilitate comparative studies investigating how the number, diversity and location of transposable element sequences and families differs between organisms with and without small RNA systems of regulation.

Empirical studies of autonomous and non-autonomous transposable elements

Empirical studies of autonomous and non-autonomous TEs have reported interesting dynamics. For example, Robillard et al. (2016) tracked the novel introduction of autonomous and non-autonomous *Mariner* transposons into *Drosophila melanogaster* populations. They found that rapid proliferation of the non-autonomous element led to extinction of both types of elements within 100 generations. In contrast to this rapid extinction, autonomous - non-autonomous partners also appear capable of coexisting and coevolving for long periods (Tanskanen et al., 2007; Wicker et al., 2022). Together, these studies suggest that autonomous - non-autonomous regulatory logic can determine the stability or instability of transposable elements in genomes.

Hatanaka et al. (2024) studied transposons from the *hAT* superfamily that have inserted into the genome of *C. inopinata*, a sister species of *C. elegans*. The *C. elegans* genome lacks *hAT* elements (Woodruff and Teterina, 2020), suggesting a relatively recent insertion event. Active transposable elements in *C. inopinata* have inserted into essential genes in the 26G si-RNA pathway (Kanzaki et al., 2018) and *C. inopinata* and its sister species *C. elegans* appear to have diverged in genomic TE landscape (Woodruff and Teterina, 2020) and TE-associated gene expression (Kawahara et al., 2023). Similar to our comparative genomic results Hatanaka et al. (2024) identified a biased pattern of autonomous - non-autonomous elements (Table 2) with just 4 autonomous *hAT* elements and over 1,000 deletion-related non-autonomous miniature inverted-repeat transposable elements (MITEs). The genomic location of TEs can play important roles in recombination and mutation with differing dynamics and gene density in chromosome arms and centers in *Caenorhabditis* (Rockman and Kruglyak, 2009). The genomic distribution of *hAT*-associated MITEs and other highly abundant non-autonomous elements can influence evolution of the genome and have consequences for multiple dimensions of phenotypic evolution.

The dynamics of transposable element invasion

Our empirical investigation of autonomous and non-autonomous *Mutator* elements in *Caenorhabditis* discovered a ubiquitous pattern of fragmented sequences and mutated TIRs rendering non-autonomous elements incapable of transposition and

essentially ‘silencing’ them in the genome. The course of a transposable element invasion into a genome can be likened to organismal infection with a novel pathogen. The early, initial stages of infection are characterized by dynamic proliferation, here transposition within the genome (Robillard et al., 2016). Within a period of time the infection or invasion is suppressed, here by mutations rendering the sequences incompetent for transposition. After the initial period of dynamic proliferation, TE-associated sequences remain in the genome although inactive and not participating in, or governed by, regulatory logic such as Competitive or Parasitic dynamics. Importantly, these silenced sequences can now create pools of novel variation for regulation (Du et al., 2024) and *de novo* gene creation (Peng and Zhao, 2024). Regulatory logic may play a critical role in the acute stages of transposable element invasion into a genome and lose importance as the invasion shifts into a chronic, silenced state.

Conclusion

Transposable elements constitute large portions of eukaryotic genomes with diverse numbers and types of elements. Despite this, there is little understanding of the evolutionary processes and population genetic factors that produce patterns of TE diversity. We have presented deterministic and stochastic theory studying the role that regulatory logic can play in determining the stability or instability of systems of autonomous and non-autonomous transposons. Our results show that regulatory interactions can interact with population genetic factors to result in highly stable systems of transposable elements. We investigated our predictions in *Caenorhabditis* genome sequences and found that empirical patterns of autonomous - non-autonomous TEs vary broadly. Rapid mutational decay likely renders non-autonomous TEs incapable of transposition and suggests that TE regulatory logic is relevant for early, acute stages of TE invasion. Together, our results demonstrate that regulatory logic can stabilize TEs and that empirical TEs employ diverse regulatory strategies.

Competing interests

No competing interest is declared.

Author contributions statement

Must include all authors, identified by initials, for example: J.L.F. conceived the research(s), V.K.E. and J.L.F. conducted the analyses(s), wrote and reviewed the manuscript.

Acknowledgments

The authors thank the members of the Fierst lab for helpful feedback and critical discussion of the research and manuscript.

Funding

This work was supported by NSF award 2225796 and NIGMS award R35GM147245 to J.L.F.

Data availability

Code and output associated with the deterministic and stochastic models are available at

<https://github.com/jannaferst/TEParasites>.

Bioinformatic scripts and workflows are available at

<https://github.com/vkeggers/>

Caenorhabditis_TE_autonomousANDnonautonomous.

Caenorhabditis genome sequences are available at NCBI under the accessions:

C. remanei PX356 Bioproject PRJNA248909

C. remanei PX439 Bioproject PRJNA248911

C. remanei PX506 Bioproject PRJNA577507

C. latens PX534 Bioproject PRJNA248912

C. briggsae AF16 Bioproject PRJNA20855

C. nigoni JU1422 Bioproject PRJNA384657

C. elegans N2 Bioproject PRJNA158, PRJNA13758

C. tropicalis NIC58, JU1373 Bioproject PRJNA662844

C. inopinata NKZ35 Bioproject PRJDB5687

C. sinica ZZY0401 Bioproject PRJNA194557

References

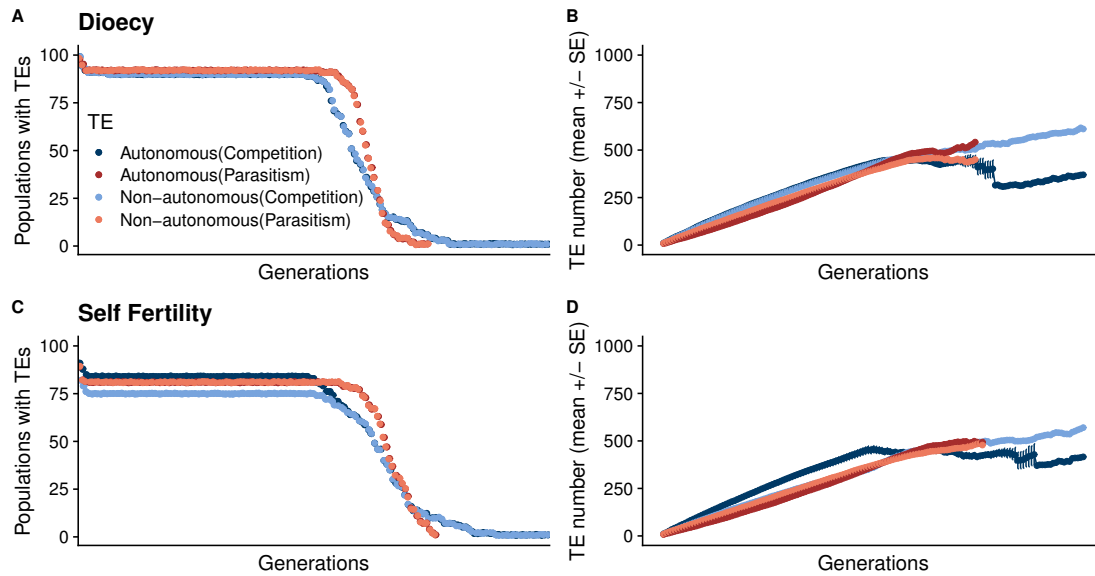
- P. E. Adams, V. K. Eggers, J. D. Millwood, J. M. Sutton, J. Pienaar, and J. L. Fierst. Genome size changes by duplication, divergence and insertion in *Caenorhabditis* worms. *Molecular Biology and Evolution*, 40(3):msad039, 2023.
- J. L. Anderson, L. T. Morran, and P. C. Phillips. Outcrossing and the maintenance of males within *C. elegans* populations. *Journal of Heredity*, 101:S62–74, 2010.
- I. Arkhipova and M. Meselson. Deleterious transposable elements and the extinction of asexuals. *BioEssays*, 27: 76–85, 2005.
- J. Bast, K. S. Jason, D. Schuseil, D. Roze, and T. Schwander. Asexual reproduction reduces transposable element load in experimental yeast populations. *eLife*, 8:e48548, 2019.
- C. R. Beck, J. L. Garcia-Perez, R. M. Badge, and J. V. Moral. LINE-1 elements in structural variation and disease. *Annual Review of Genomics and Human Genetics*, 12:187–215, 2011.
- G. Bonchev and Y. Willi. Accumulation of transposable elements in selfing populations of *Arabidopsis lyrata* supports the ectopic recombination model of transposon evolution. *New Phytologist*, 219:767–778, 2018.
- T. S. Boutin, A. Le Rouzic, and P. Capy. How does selfing affect the dynamics of selfish transposable elements? *Mobile DNA*, 3:5, 2012.
- J. Brennecke, A. A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G. J. Hannon. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, 128(6):1089–1093, 2007.
- J. F. Y. Brookfield. Models of the spread of non-autonomous selfish transposable elements when transposition and fitness are coupled. *Genetical Research*, 67:199–209, 1996.
- T. Bruna, K. J. Hoff, A. Lomsadze, M. Stanke, and M. Borodovsky. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics*, 3(1):lqaa108, 2021.
- D. M. Castillo, J. C. Mell, K. S. Box, and J. P. Blumensiel. Molecular evolution under increasing transposable element burden in *Drosophila*: a speed limit on the evolutionary arms race. *BMC Evolutionary Biology*, 11:258, 2011.

- N.-C. Chang, J. N. Wells, A. Y. Wang, P. Schofield, Y.-C. Huang, V. H. Truong, M. Simoes-Costa, and C. Feschotte. Gag proteins encoded by endogenous retroviruses are required for zebrafish development. *BioRxiv*, 2024.03.25.586437, 2024.
- B. Charlesworth. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genetical Research Cambridge*, 55:199–221, 1990.
- B. Charlesworth and D. Charlesworth. The population dynamics of transposable elements. *Genetical Research*, 42:1–27, 1983.
- B. Charlesworth and C. H. Langley. The evolution of self-regulated transposition of transposable elements. *Genetics*, 112:359–383, 1986.
- B. Charlesworth, P. Sniegowski, and W. Stephan. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*, 371:215–220, 1994.
- P. Chen and J. Zhang. Asexual experimental evolution of yeast does not curtail transposable elements. *Molecular Biology and Evolution*, 38(7):2831–2842, 2021.
- E. B. Chuong, N. C. Elde, and C. Feschotte. Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics*, 18:71–86, 2016.
- N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz, editors. *Mobile DNA II*. ASM Press, Washington, D.C., 2002.
- J. A. G. M. de Visser, T. F. Cooper, and S. F. Elena. The causes of epistasis. *Proceedings Of The Royal Society B-Biological Sciences*, 278(1725):3617–3624, Nov. 2011.
- E. S. Dolgin and B. Charlesworth. The fate of transposable elements in asexual populations. *Genetics*, 174(2):817–27, 2006.
- E. S. Dolgin and B. Charlesworth. The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics*, 178:2169–2177, 2008.
- W. F. Doolittle and C. Sapienza. Selfish genes, the phenotype paradigm and genome evolution. *Nature*, 284:601–603, 1980.
- W. F. Doolittle, T. B. L. Kirkwood, and M. A. H. Dempster. Selfish DNA and self-restraint. *Nature*, 307:501–502, 1984.
- A. Y. Du, J. D. Chobirko, X. Zhuo, C. Feschotte, and T. Wang. Regulatory transposable elements in the encyclopedia of DNA elements. *Nature Communications*, 15:7594, 2024.
- M. Dupeyron, K. S. Singh, C. Bass, and A. Hayward. Evolution of *Mutator* transposable elements across eukaryotic diversity. *Mobile DNA*, 10:12, 2019.
- D. Ellinghaus, S. Kurtz, and U. Willhoeft. LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinformatics*, 9:18, 2008.
- N. V. Fedoroff. Transposable elements, epigenetics, and genome evolution. *Science*, 338:758–767, 2012.
- J. L. Fierst, J. H. Willis, C. G. Thomas, W. Wang, R. M. Reynolds, T. E. Ahernage, A. D. Cutter, and P. C. Phillips. Reproductive mode and the evolution of genome size and structure in *Caenorhabditis* nematodes. *PLoS Genetics*, 11(6):e1005323, 2015.
- D. J. Finnegan. Transposable elements. In D. L. Lindsley and G. Zimm, editors, *The genome of Drosophila melanogaster*, pages 1096–1107. Academic press, New York, 1992.
- S. Fontenla, G. Rinaldi, and J. F. Tort. Lost and found: Piwi and argonaute pathways in flatworms. *Frontiers in Cellular and Infection Microbiology*, 11:653595, 2021.
- R. Ge, G. Mai, R. Zhang, X. Wu, Q. Wu, and F. Zhou. MUSTv2: an improved *de novo* detection program for recently active miniature inverted repeat transposable elements MITEs. *Journal of Integrative Bioinformatics*, 14(3):20170029, 2017.
- G. Grimaldi, J. Skowronski, and M. F. Singer. Defining the beginning and end of KpnI family segments. *EMBO Journal*, 3:1753–59, 1984.
- Z. Guan, S. Shi, M. Diaby, P. Danley, N. Ullah, M. Puzakov, B. Gao, and C. Song. Horizontal transfer of *Buster* transposons across multiple phyla and classes of animals. *Molecular Phylogenetics and Evolution*, 173:107506, 2022.
- B. C. Haller and P. W. Messer. SLiM 3: forward genetic simulations beyond the Wright-Fisher model. *Molecular Biology and Evolution*, 36(3):632–637, 2019.
- B. C. Haller and P. W. Messer. SLiM 4: Multispecies eco-evolutionary modeling. *The American Naturalist*, 201(5), 2023.
- D. L. Hartl, E. R. Lozovskaya, and J. G. Lawrence. Nonautonomous transposable elements in prokaryotes and eukaryotes. *Genetica*, 86(1-3):47–53, 1992.
- R. Hatanaka, K. Tamagawa, N. Haruta, and A. Sugimoto. The impact of differential transposition activities of autonomous and nonautonomous *hAT* transposable elements on genome architecture and gene expression in *Caenorhabditis inopinata*. *Genetics*, 227(2):iyae052, 2024.
- M. Henault, S. Marsit, G. Charron, and C. R. Landry. The genomic landscape of transposable elements in yeast hybrids is shaped by structural variation and genotype-specific modulation of transposition rate. *eLife*, 12:RP89277, 2024.
- D. A. Hickey. Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics*, 101(3-4):519–531, 1982.
- K. L. Howe, B. J. Bolt, M. Shafie, P. Kersey, and M. Berriman. WormBase ParaSite- a comprehensive resource for helminth genomics. *Molecular and Biochemical Parasitology*, 215:2–10, 2017.
- N. Kanzaki, I. J. Tsai, R. Tanaka, V. L. Hunt, D. Liu, K. Tsuyama, Y. Maeda, S. Namai, R. Kumagai, A. Tracey, N. Holroyd, S. R. Doyle, G. C. Woodruff, K. Murase, H. Kitazume, C. Chai, A. Akagi, O. Panda, H. Ke, F. C. Schroeder, J. Wang, M. Berriman, P. W. Sternberg, A. Sugimoto, and T. Kikuchi. Biology and genome of a newly discovered sibling species of *Caenorhabditis elegans*. *Nature Communications*, 9:3216, 2018.
- K. Kawahara, T. Inada, R. Tanaka, M. Dayi, T. Makino, S. Maruyama, T. Kikuchi, A. Sugimoto, and M. Kawata. Differentially expressed genes associated with body size changes and transposable element insertions between *Caenorhabditis elegans* and its sister species, *Caenorhabditis inopinata*. *Genome Biology and Evolution*, 15(4):evad063, 2023.
- E. S. Kelleher, D. A. Barbash, and J. P. Blumenstiel. Taming the turmoil within: new insights on the containment of transposable elements. *Trends in Genetics*, 36(474-489), 2020.
- M. G. Kidwell and D. R. Lisch. Transposable elements, parasitic DNA, and genome evolution. *Evolution*, 55:1–24, 2001.
- K. Kiontke and D. H. Fitch. The phylogenetic relationships of *Caenorhabditis* and other rhabditids. In *WormBook*, pages 1–11. The *C. elegans* Research Community, 11 edition, August 2005.
- K. Kiontke, N. P. Gavin, Y. Raynes, C. Roehrig, F. Piano, and D. H. A. Fitch. *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proceedings of the National Academy of Sciences*, 101(24):

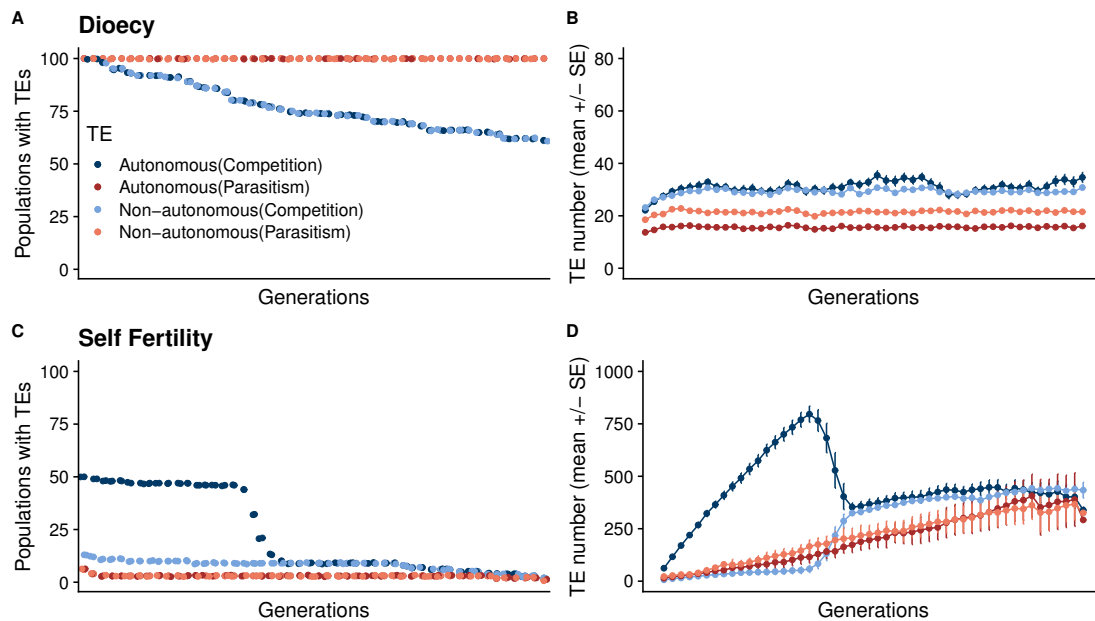
- 9003–9008, June 2004.
- C. H. Langley, J. F. Brookfield, and N. Kaplan. Transposable elements in mendelian populations. i. a theory. *Genetics*, 104(3):457–471, 1983.
- A. Le Rouzic and P. Capy. Population genetics models of competition between transposable element subfamilies. *Genetics*, 1976(2):785–793, 2006.
- Y. C. G. Lee. Synergistic epistasis of the deleterious effects of transposable elements. *Genetics*, 220(2):iyab211, 2022.
- R. Lehmann, A. Kovarik, K. Ocalewicz, L. Kirtiklis, A. Zuccolo, J. NTegner, J. Wanzenbock, L. Bernatchez, D. K. Lamatsch, and R. Symonova. DNA transposon expansion is associated with genome size increase in mudminnows. *Genome Biology and Evolution*, 13(10):evab228, 2021.
- D. Lisch. Mutator and MULE transposons. *Microbiology Spectrum*, 3(2):MDNA3–0032, 2015.
- S. Maisnier-Patin, J. R. Roth, A. Fredriksson, T. Nystrom, O. G. Berg, and D. I. Andersson. Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nature Genetics*, 37(12):1376–9, 2005.
- W. Makalowski, V. Gotea, A. Pande, and I. Makalowska. Transposable elements: classification, identification, and their use as a tool for comparative genomics. *Methods in Molecular Biology*, 1910:177–207, 2019.
- T. M. Mattila, B. Laenen, and T. Slotte. Population genomics of transitions to selfing in brassicaceae model systems. In J. Y. Dutheil, editor, *Statistical Population Genomics*, chapter 11, page 269. Humana Press, 2020.
- E. K. Maxwell, J. F. Ryan, C. E. Schnitzler, W. E. Browne, and A. D. Baxevanis. MicroRNAs and essential components of the microRNA processing machinery are not encoded in the genome of the ctenophore *Mnemiopsis leidyi*. *BMC Genomics*, 13(1):714, 2012.
- B. McClintock. The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences*, 36:344–355, 1950.
- M. Mondal, P. Klimov, and A. S. Flynt. Rewired RNAi-mediated genome surveillance in house dust mites. *PLoS Genetics*, 14(1):e1007183, 2018.
- National Center for Biotechnology Information. The reference sequence (refseq) project. Bethesda (MD), 2002.
- R. M. Nowell, C. G. Wilson, P. Almeida, P. H. Schiffer, D. Fontaneto, L. Becks, F. Rodriguez, I. R. Arkhipova, and T. G. Barraclough. Evolutionary dynamics of transposable elements in bdelloid rotifers. *eLife*, 10:e63194, 2021.
- C. Oliver, M. L. Annacondia, Z. Wang, P. E. Jullien, R. K. Slotkin, C. Kohler, and G. Martinez. The miRNome function transitions from regulating developmental genes to transposable elements during pollen maturation. *The Plant Cell*, 34(2):784–801, 2022.
- A. D. Omole and P. Czuppon. Maintenance of long-term transposable element activity in genomes through regulation by nonautonomous elements. *BioRxiv*, page doi.org/10.1101/2024.07.13.603364, 2024.
- L. E. Orgel and F. H. Crick. Selfish DNA: The ultimate parasite. *Nature*, 284:604–607, 1980.
- E. M. Ostertag and J. H. H. Kazazian. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genome Research*, 11:2059–65, 2001.
- S. Ou and N. Jiang. LTR retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiology*, 180(4):1803–1815, 2018.
- S. Ou and N. Jiang. LTR FINDER parallel: parallelization of LTR FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mobile DNA*, 10(1):48, 2019.
- S. Ou, W. Su, Y. Liao, K. CHougule, J. R. A. Agda, A. J. Hellinga, C. Santiago, B. Lugo, T. A. Elliott, D. Ware, T. Peterson, N. Jiang, C. N. Hirsch, and M. B. Hufford. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology*, 20(275), 2019.
- D. M. Ozata, I. Gainetdinov, A. Zoch, D. O’Carroll, and P. D. Zamore. PIWI-interacting RNAs: small RNAs with big functions. *Nature Reviews Genetics*, 20:89–108, 2018.
- O. Panaud. Horizontal transfers of transposable elements in eukaryotes: The flying genes. *Comptes Rendus Biologies*, 339(7-8):296–299, 2016.
- N. S. Paulat, J. M. Storer, D. D. Moreno-Santillan, A. B. OSmanski, K. A. M. Sullivan, J. R. Grimshaw, J. Korstian, M. Halsey, C. J. Garcia, C. Crookshanks, J. Roberts, A. F. A. Smit, R. Hubley, J. Rosen, E. C. Teeling, S. C. Vernes, E. Myers, M. Pippel, T. Brown, M. Hiller, Z. Consortium, D. Rojas, L. M. Davalos, K. Lindblad-Toh, E. K. Karlsson, and D. A. Ray. Chiropterans are a hotspot for horizontal transfer of DNA transposons in mammalia. *Molecular Biology and Evolution*, 40(5):msad092, 2023.
- J. Peccoud, V. Loiseau, R. Cordaux, and C. Gilbert. Massive horizontal transfer of transposable elements in insects. *Proceedings of the National Academy of Sciences of the United States of America*, 114(18):4721–4726, 2017.
- J. Peng and L. Zhao. The origin and structural evolution of de novo genes in *Drosophila*. *Nature Communications*, 15:810, 2024.
- E. J. Pritham. Transposable elements and factors influencing their success in eukaryotes. *Journal of Heredity*, 100(5):648–655, 2009.
- M. Punta, P. C. Coggill, R. Y. Eberhardt, J. Mistry, J. Tate, C. Bournsnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E. L. L. Sonnhammer, S. R. Eddy, A. Bateman, and R. D. Finn. The PFam protein families database. *Nucleic Acids Research*, 40(D1):D290–D301, 2012.
- A. R. Quinlan and I. M. Hall. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6):841–842, 2010.
- P. Rice, I. Longden, and A. Bleasby. EMBOSS the european molecular biology open software suite. *Trends in Genetics*, 16(6):276–277, 2000.
- D. S. Robertson. Characterization of a mutator system in maize. *Mutational Research*, 51:21–8, 1978.
- D. S. Robertson. Genetic studies on the loss of mu mutator activity in maize. *Genetics*, 113:765–773, 1986.
- E. Robillard, A. Le Rouzic, Z. Zhang, P. Capy, and A. Hua-Van. Experimental evolution reveals hyperparasitic interactions among transposable elements. *Proceedings of the National Academy of Sciences of the United States of America*, 113(51):14763–14768, 2016.
- M. V. Rockman and L. Kruglyak. Recombinational landscape and population genomics of caenorhabditis elegans. *PLoS Genetics*, 5(3):e1000419, Jan 2009. doi: 10.1371/journal.pgen.1000419.
- D. Roze. Causes and consequences of linkage disequilibrium among transposable elements within eukaryotic genomes. *Genetics*, 224(2):iyad058, 2023.
- J. F. Ryan, K. Pang, C. E. Schnitzler, A. D. Nguyen, R. T. Moreland, D. K. Simmons, B. J. Koch, W. R. Francis, P. Havlak, N. C. S. Program, S. A. Smith, N. H. Putnam, S. H. D. Haddock, C. W. Dunn, T. G. Wolfsberg,

- J. C. Mullikin, M. Q. Martindale, and A. D. Baxevanis. The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science*, 342(6164):1242592, 2013.
- P. Sarkies. The curious case of the disappearing piRNAs. *WIREs RNA*, 15:e1849, 2024.
- P. Sarkies, M. E. Selkirk, J. T. Jones, V. Blok, T. Boothby, B. Goldstein, B. Hanelt, A. Ardila-Garcia, N. M. Fast, P. M. Schiffer, C. Kraus, M. J. Taylor, G. Koutsovoulos, M. L. Blaxter, and E. A. Miska. Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. *PLoS Biology*, 13(2):e1002061, 2015.
- J. Shi and C. Liang. Generic repeat finder: A high-sensitivity tool for genome-wide de novo repeat detection. *Plant Physiology*, 180(4):1803–1815, 2019.
- A. F. A. Smit, R. Hubley, and P. Green. Repeatmasker open-4.0. 2013-2015.
- W. Su, X. Gu, and T. Peterson. TIR-Learner, a new ensemble method for TIR transposable element annotation, provides evidence for abundant new transposable elements in the maize genome. *Molecular Plant*, 12:447–60, 2019.
- J. A. Tanskanen, F. Sabot, C. Vicient, and A. H. Schulman. Life without GAG: the BARE-2 retrotransposon as a parasite's parasite. *Gene*, 390:166, 2007.
- M. I. Tenailon, M. B. Hufford, B. S. Gaut, and J. Ross-Ibarra. Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxurians*. *Genome Biology and Evolution*, 3:219–229, 2011.
- L. Wang, L. Tracy, W. Su, F. Yang, Y. Feng, N. Silverman, and Z. Z. Zhao Zhang. Retrotransposon activation during *Drosophila* metamorphosis conditions adult antiviral responses. *Nature Genetics*, 54(12):1933–1945, 2022.
- L. Wei, L. Gu, X. Song, X. Ciu, Z. Lu, M. Zhou, L. Wang, F. Hu, J. Zhai, B. C. Meyers, and X. Cao. Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice. *Proceedings of the National Academy of Sciences of the United States of America*, 111(10):3877–3882, 2014.
- A. M. Weiner. SINEs and LINEs: the art of biting the hand that feeds you. *Current Opinion in Cell Biology*, 14(3):343–50, 2002.
- T. Wicker, F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhou, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, and A. H. Schulman. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, 8:973–982, 2007.
- T. Wicker, C. Stritt, A. G. Sotiropoulos, M. Poretti, C. Pozniak, S. Walkowiak, H. Gundlach, and N. Stein. Transposable element populations shed light on the evolutionary history of wheat and the complex co-evolution of autonomous and non-autonomous retrotransposons. *Advanced Genetics*, 3(1):2100022, 2022.
- G. C. Woodruff and A. A. Teterina. Degradation of the repetitive genomic landscape in a close relative of *Caenorhabditis elegans*. *Molecular Biology and Evolution*, 37(9):2549–2567, 2020.
- W. Xiong, L. He, J. Lai, H. K. Dooner, and C. Du. HelitronScanner uncovers a large overlooked cache of helitron transposons in many plant genomes. *Proceedings Of The National Academy Of Sciences*, 111:10263–8, 2014.
- R. Zhang, Z. Wang, S. Ou, and G. Li. TESorter: lineage-level classification of transposable elements using conserved protein domains. *BioRxiv*, 800177, 2019.

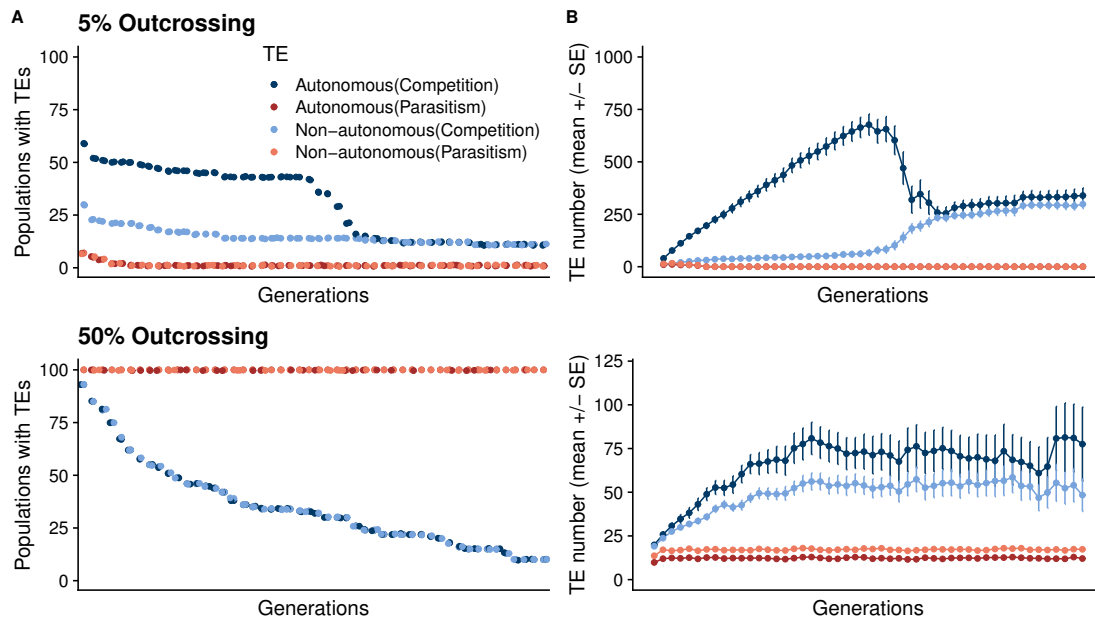
Supplementary Figures



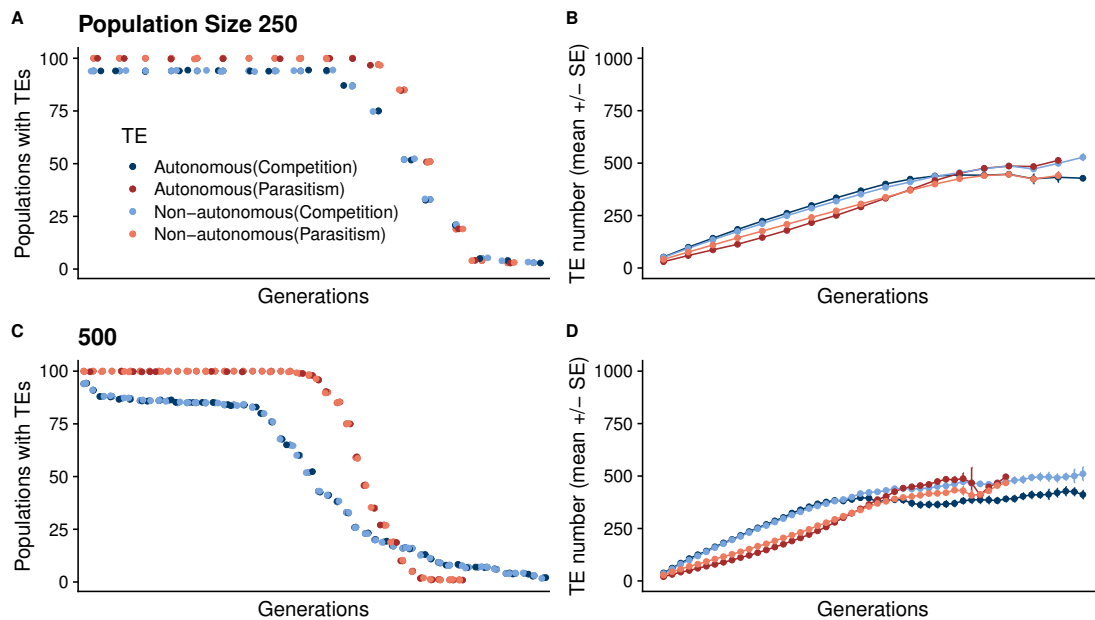
Supplementary Figure 1. In small (here, $N=100$ individuals) dioecious populations both (A) Competitive and Parasitic TE regulation resulted in one of two trajectories, either early TE loss or TE proliferation to the point of extreme fitness loss and population crash. The populations were evolved for 500,000 generations and all simulations started with 100 populations. (B) The mean number of TEs (\pm SE) were calculated for populations with TEs present. (C) In small self-fertile populations (here, $N=100$ individuals) both Competitive and Parasitic TE regulation resulted in similar trajectories of early TE loss or TE proliferation, fitness loss and population crash. (D) TEs could proliferate until the population crashed



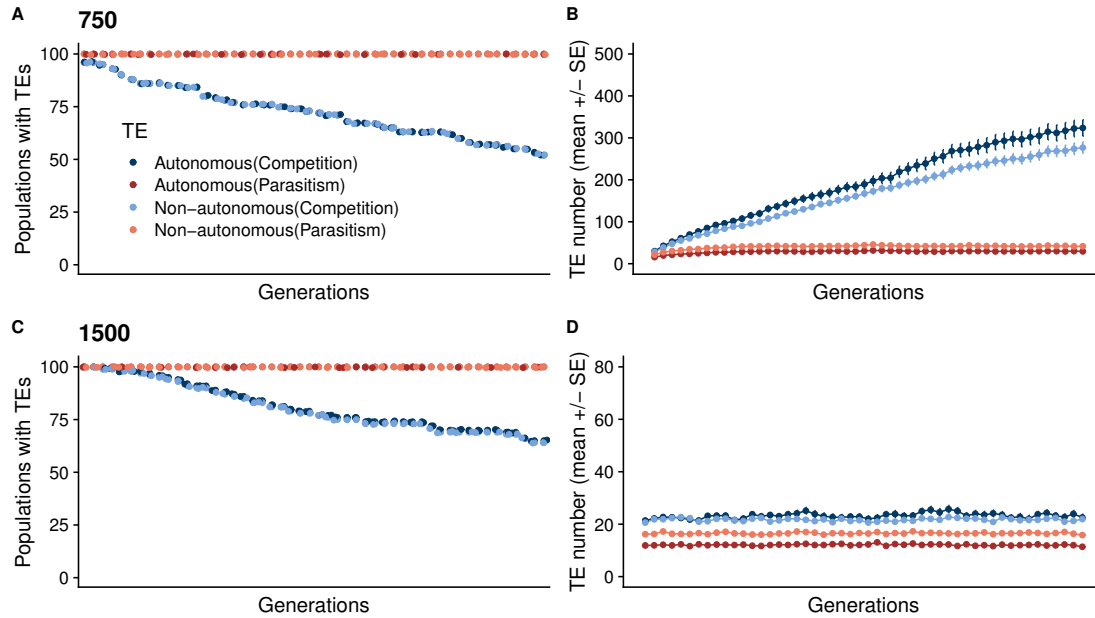
Supplementary Figure 2. Increasing the recombination rate to $1e-2$ did not alter the qualitative trajectory of TEs in the simulated populations. At large size (here, $N=1000$) dioecious populations (A) evolving with Parasitic TE regulation achieved Stable Maintenance while populations evolving with Competitive TE regulation did not. (B) For those populations that did retain TEs, the mean abundance was lower when compared with populations evolving with lower recombination rates. (C) Self-fertile populations showed the same trajectories of TE proliferation and loss when evolving with high recombination rates including (D) decoupling of the autonomous and non-autonomous TEs.



Supplementary Figure 3. Populations evolving with both Competitive and Parasitic TE regulation follow the trajectories of fully self-fertile populations at A,B) 5% outcrossing and fully outcrossing populations at C,D) 50% outcrossing.

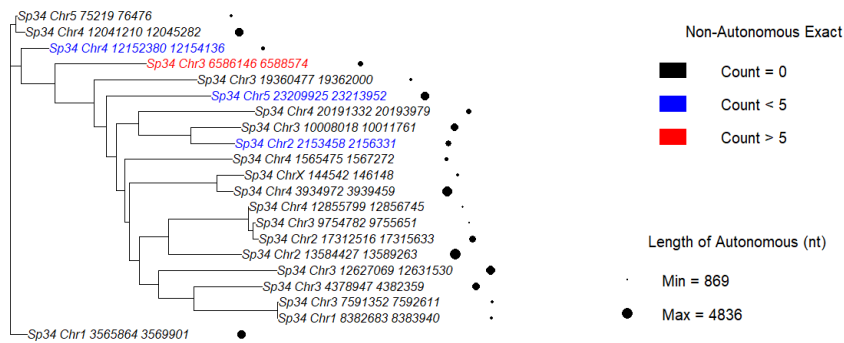


Supplementary Figure 4. Small populations of A,B) 250 and C,D) 500 individuals show similar patterns to populations evolving with 100 individuals (SFigure 1). TEs follow either the 'Early TE Loss' or 'TE Proliferation & Population Crash' trajectories.



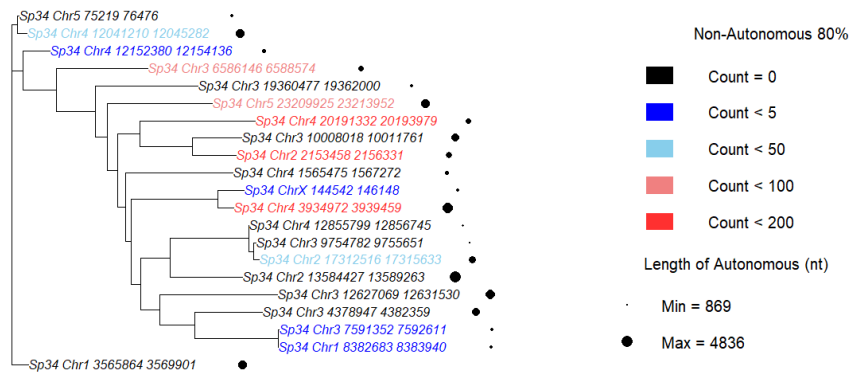
Supplementary Figure 5. Populations evolving with larger population sizes show similar TE patterns to populations evolving with 1000 individuals (Fig. 5) Parasitic TE regulation achieve the stable maintenance trajectory at population sizes of A-B) 750 and C-D) 1500 individuals. With Competitive regulation TEs are maintained for longer periods but lost according to the ‘Eventual TE Loss’ trajectory. In populations of D) 1500 individuals TEs are maintained at lower mean levels in populations evolving with 750 individuals.

Autonomous *C.inopinata* with Corresponding Exact Non-Autonomous TIRs



Supplementary Figure 6. Phylogenetic reconstruction of autonomous *Mutator* sequences in *C. inopinata* (also known as *Caenorhabditis* Species 34) shows the transposons form 3 clades, potentially reflecting 3 separate *Mutator* invasions. One clade contains a single sequence, one contains two sequences and the third clade contains the remaining 18 sequences. Of these, 17 have no corresponding non-autonomous elements with intact TIRs, three have <5 corresponding non-autonomous elements with intact TIRs and the remaining one autonomous element has >5 corresponding non-autonomous elements.

Autonomous *C.inopinata* with Corresponding 80% Non-Autonomous TIRs



Supplementary Figure 7. Phylogenetic reconstruction of autonomous *Mutator* sequences in *C. inopinata* shows that 7 autonomous sequences each retain >80% TIR identity with >50 non-autonomous sequences. The remaining 14 autonomous sequences retain >80% identity with 5 or fewer non-autonomous *Mutator* elements.